APh161: Physical Biology of the Cell
Homework 4
Due Date: 5 pm, Friday February 12, 2010

Comments from RP to Class:

This fourth homework explores several problems I have done in class in much greater detail.

1. Free energy scale of entropy of chemical interaction in k_BT units.

The goal of this problem is to get you to examine entropy, free energy and chemical potentials from a few different angles. In addition, the results of part (b) will permit you to do the problem on the critical micelle concentration later in the homework. NOTE: in this problem I am a little imprecise with my free energy definitions, using the Gibbs free energy G in the first part and the Helmholtz free energy F in the second part.

(a) *Entropy per molecule due to configurations.* Imagine the chemical reaction $A + B \rightleftharpoons AB$. The goal in this part of the problem is to estimate the free energy scale due to entropy when a single reaction occurs changing the number of molecules from $N_A,\ N_B$ and N_{AB} to $N_A - 1,\ N_B - 1$ and $N_{AB} + 1$. Estimate the free energy difference by using $\Delta G = -T(S_f - S_i)$ in conjunction with the Boltzmann expression for the entropy $S = k_B \ln W$. As a result, we have

$$\Delta G = -k_B T \ln \frac{W_f}{W_i}, \tag{1}$$

where W_f and W_i are the multiplicities for the system in the final and initial states, respectively. Hence, to obtain our estimate, all you need to do is construct the ratio W_f/W_i. Use the lattice models favored in class in order to write this result. Make sure to explain how you convert from the lattice model which involves the number of ligands and the number of lattice sites to a representation in terms of concentrations. Once you have a formula for this free energy, make a *numerical* estimate of its value by using some "typical" concentrations for reactants and products.

(b) *Entropy per molecule including the distribution of kinetic energy as part of the entropy.* In this second part of the problem, we do something a

little more sophisticated for the same reaction by considering the contribution to the entropy coming from the ways of distributing the kinetic energy. That is, all of the molecules in our solution (or gas) share a certain amount of total kinetic energy. But there are many different ways that they can do this sharing and this provides a contribution to the entropy. This result will prove useful to us in the next problem when we work out the critical micelle concentration.

In this case, we will use the expression for the relation between the Helmholtz free energy and the partition function, namely, $F = -k_B T \ln Z$. To compute this, we use the fact that the total partition function is given by

$$Z_{tot} = \frac{Z_1^N}{N!}, \tag{2}$$

where Z_1 is the partition function of a single-molecule. This partition function can be computed in turn as

$$Z_1 = \frac{1}{h^3} \int d^3\mathbf{q} \int d^3\mathbf{p} e^{-\beta p^2/2m}, \tag{3}$$

where h is Planck's constant, $\mathbf{q}$ is the generalized coordinate and $\mathbf{p}$ is the momentum for the molecule of interest. The first integral is just an integral over all the configurations of the system and simply yields V, the volume of the box in which the molecules of interest are jiggling around. Use the identify $\int_\infty^\infty dp e^{-\alpha p^2} = \sqrt{\pi/\alpha}$. and show that

$$Z_1 = \frac{V}{h^3}(2\pi m k_B T)^{\frac{3}{2}} \tag{4}$$

and

$$Z_{tot} = \frac{V^N}{N! h^{3N}}(2\pi m k_B T)^{\frac{3N}{2}}. \tag{5}$$

Use the fact that we can write the chemical potential as

$$\mu = F(N) - F(N-1) \tag{6}$$

and by using the fact that $F = -k_B T \ln Z_{tot}$, show that

$$\mu = -k_B T \ln \frac{Z_{tot}(N)}{Z_{tot}(N-1)}. \tag{7}$$

Now, demonstrate that this implies

$$\mu = k_B T \ln c\lambda_{th}^3, \tag{8}$$

where c is the concentration given by $c = N/V$ and λ_{th} is the thermal de Broglie wavelength and is given by

$$\lambda_{th} = \frac{h}{\sqrt{2\pi m k_B T}}. \tag{9}$$

With this result in hand, once again make an *estimate* of the free energy scale associated with a single reaction by using the fact that

$$\Delta G = \mu_{AB} - \mu_A - \mu_B. \tag{10}$$

Make sure to explain the conceptual underpinnings of your estimate as well as the numerical values you invoked. Like in part (a), you will have to think about the "typical" concentrations in a chemical reaction and also, make some decision about the masses of your reactants and products. The simplest idea might be $m_A = m_B = m_{AB}/2$.

2. Energy and Entropy Competition in Micelle Formation.

In this problem we examine some of the interesting forces that arise in the crowded environs of the cellular interior in a way that will also permit you to practice using the chemical potential you derived in part (b) of the previous problem in an important and intuitive biological setting that you already know a lot about - namely, the properties of lipids in solution. In class I gave a quick impression of the hydrophobic effect as an idea that is invoked often with great explanatory power. In this problem, you will estimate the magnitude of the interfacial energy that is assigned to having certain chemical groups in contact with water. This will give us an idea of how much free energy is gained when different molecules come into contact and sequester these hydrophobic structural elements. The essential argument is that the water molecules that surround the hydrophobic region of a molecule are deprived of some of their entropy because they can adopt fewer hydrogen bonding configurations. In particular, the water molecules are thought to form cages known as clathrate structures such as are shown in the accompanying figure.

(a) Estimate the entropy lost for each water molecule by appealing to the schematic of the tetrahedron shown in fig. 2. The basic idea is that if we think of the O of the water molecule as being situated at the center of the tetrahedron then the two H atoms can be associated with any two adjacent vertices (or, there are a total of six configurations). However, when in the presence of the hydrophobic molecule, one of the faces of the tetrahedron can be thought of as facing that hydrophobic molecule and hence all

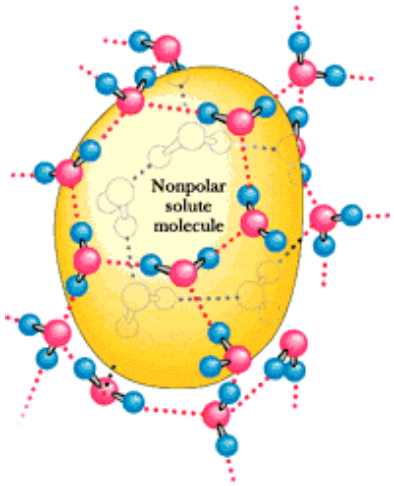

Figure 1: Schematic of the clathrate structure adopted by water molecules surrounding a hydrophobic molecule.

configurations (three of the edges) facing that molecule are unavailable for hydrogen bonding. How many configurations are available now? Compute the entropy change of a single water molecule as a result of this configurational inhibition using the celebrated Boltzmann equation already invoked in problem 1, namely,

$$S = k_B \ln W, \tag{11}$$

where S is the entropy and W is the number of microscopic configurations available to the system in the given macroscopic state of interest. NOTE: Section 5.5.1 of PBoC essentially does this calculation from start to finish.

(b) Next, we need to estimate how many water molecules neighbor a given hydrophobic molecule. Consider the case of methane and ethane and approximate them as "spheres" and estimate the radius of sphere that represents the hydrophobic surface area they present. Next, estimate how many water molecules neighbor these molecules and hence the total free energy difference because of the lost entropy using the results you obtained in the previous part of the problem. Convert your result into an interfacial energy and use units both of J/m^2 and cal/mol $\mathring{A}^2$. Compare the result to the rule of thumb I quoted in passing in class which is 25 cal/mol $\mathring{A}^2$.

(c) Next, use the same argument as part (b) and work out a formula for the

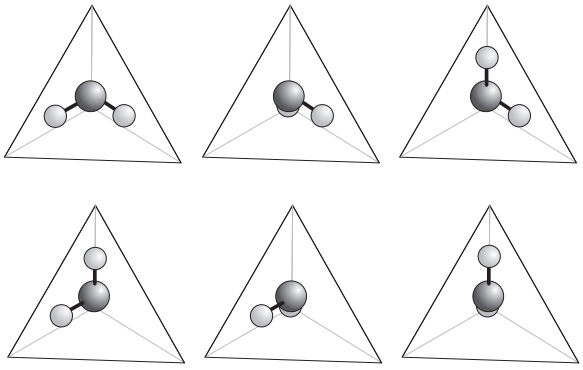

Figure 2: Schematic of the arrangements available to a water molecule when in a complete network of other water molecules.

hydrophobic cost for an individual lipid in solution as a function of both the length and the number of tails. This may sound vague. I am trying to get you to think about how to construct a "toy model" of this problem. In particular, consider the lipid tails as cylinders of length L and radius r and use the formula for the area of a cylinder to approximate the hydrophobic area presented by the tails. So really, all you have to do is figure out some clean idea for the length of lipids as a function of the number of carbons in the tails. Avanti Polar Lipids (a company that is a big provider of lipids) has a website with properties of lipids, but really, I am just looking for a sensible approximation and an explanation for the approximation rather than a particular "right" answer.

(d) Now you will use what we learned above to construct a "toy model" of lipids in solution and in aggregate form and then use this model to derive the critical micelle concentration for lipid molecules. I mentioned this idea in class. The point is that as you increase the concentration of lipids in solution, there will be a concentration at which the free energy competition favors the creation of lipid spheres with tails pointing inwards as opposed to having free lipids wandering around in solution. There are many different measurable parameters such as the surface tension, the osmotic pressure and turbidity that permit a measurement of the onset of micellization.

The concept of your calculation is that we will construct a chemical potential both for individual lipids and for micelles and the critical micelle concentration will be that concentration at which these two chemical poten-

tials are equal. For ease, let's take the entropy of the micelles to be zero. That is, let's ignore the fact that they are free to jiggle around in solution or to rotate and hence pretend that they have no entropy. In addition, let's assume that the only contribution to the energetic part of the free energy comes from the hydrophobic effect associated with the lipid tails. What this means is that in the micellar state, we will also assert that the *energetic* contribution to the micelle free energy (and chemical potential) is zero since in those structures the lipid tails are sequestered. Further, when computing the entropy of the lipid molecules, treat them as "point" particles. What I mean by this is that you will merely compute their conformational entropy and will ignore the fact that they can rotate, that their chains can wiggle around. As a result, the chemical potential for the isolated lipids will have two contributions: i) the energy contribution you worked out in part (c) of this problem and the entropy like that you worked out in problem 1(b). Note that the entropy you worked out will depend upon the mass of the lipids through the thermal de Broglie wavelength (this is the same as depending upon the length).

Work out the critical micelle concentration by equating the chemical potentials for isolated lipids with that for micelles and do so for both single-tailed and double-tailed lipid molecules. Look up examples of each type of lipid and use the length of such tails in $\mathring{A}$ as part of your estimate. Also, make a plot of the critical micelle concentration as a function of the chain lengths, for chains varying in length between 12 and 20 carbons per chain. Experimental data for the critical micelle concentration is shown in fig. 3. Make sure to explain the qualitative trends you calculate and that are shown in the experimental data.

3. Chemotaxis and Receptor Binding.

As described in class, bacterial chemotaxis is claimed to be the best studied signal transduction problem in biology. In this problem, we work through some of the statements and results in a few of the classic papers I presented in class. We develop a feeling for the numbers by examining direct quotations from the experimental papers that have really driven the field recently as well as a commentary on this work by Dennis Bray. Begin by reading both of these papers which are attached on the website.

In their 2002 paper in PNAS entitled "Receptor sensitivity in bacterial chemotaxis", Sourjik and Berg say: "The changes in receptor occupancy encountered by bacteria swimming in spatial gradients (e.g., near the mouth

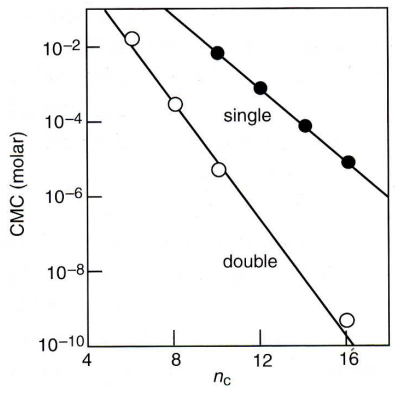

Figure 3: Experimental data for the critical micelle concentration. (taken from David Boal, **Mechanics of the Cell**). "Single" and "double" refer to the number of tails.

of a capillary tube in the capillary assay) are very small. For example, in the tracking experiments, cells about 0.6 mm from the tip of a capillary tube containing 1 mM aspartate moved in a gradient of steepness 0.02 μM/μm at a mean concentration of about 8 μM. A 10-μm run straight up such a gradient would change the concentration from 8 to 8.2 μM, i.e., by 2.5 %. Assuming K_d values for aspartate of 7.1 μM and 62 mM (see above), this step gives a fractional change in receptor occupancy of about 0.003". **RP to class:** the two K_d values correspond to the fact that two of the different chemotactic receptors (Tar and Tsr) will bind aspartate, but with quite different affinities. For your estimates, only consider the smaller K_d since the larger one will be irrelevant at the concentrations of interest here.

Your job is to carry out calculations that exploit the numbers given above and using what you know about the definitions of concentration, the size of *E. coli* cells and about the meaning of K_d and simple binding curves (i.e. $p_{bound} = (L/K_d)/(1 + (L/K_d))$). First, use the steady-state diffusion equation for a spherically symmetric source to estimate the concentration at 0.6 mm from the pipette. The idea is to solve the 3D diffusion equation in spherical coordinates, given that the concentration at the source (i.e. the pipette) is 1 mM and that the concentration in the far field is zero. (NOTE: we work this out in chap. 13 of PBoC in a different context, but the ideas

are all the same.) Do you agree with them about the concentration being 8 μM at a distance of 0.6 mm? Next, examine the statement about the consequences of a 10-μm run and also about the fractional change in occupancy. Do you agree with their numbers? Do you agree with the qualitative thrust of their statements?

In his commentary on the paper of Sourjik and Berg, Dennis Bray says: "The mystery can be expressed in a different way. Estimates of the binding affinity of aspartate to the membrane receptor of wild-type *E. coli* typically give a dissociation constant in the range 15 μM. A bacterium responding to a change in occupancy of 0.1% is therefore sensing concentrations of aspartate of a few nanomolar. And yet we know from decades of observations that the same bacterium is also capable of responding to gradients of aspartate that extend up to 1 mM. Somehow, *E. coli* is able to sense aspartate over a range of at least 5 orders of magnitude in concentration by using just one molecular species of receptor!"

Your job is to actually do the estimate/calculation that supports the claim made by Bray. In particular, examine a 0.1% change in occupancy and see what that means about the change in concentration given that the K_d has the value claimed. Also, if there is a change of concentration of order a few nanomolar, how many fewer molecules are there in a box of size 1 μm^3 due to such a concentration difference at the front and back of a cell?

4. Concentrations by Dilution.

In this problem we consider the concentration of mRNA or proteins as a function of time in dividing cells. This exercise provides some of the conceptual tools we will need to write down rate equations describing gene expression. In particular, the point of this problem is to work out the concentration of mRNA or protein given that we start with a single parental cell that has N copies of this mRNA or protein (in the experiments of Golding *et al.* they watch the mRNA dilution effect while in the experiments of Rosenfeld *et al.* this is a fluorescently-labeled transcription factor). In the Rosenfeld experiment, at some point while the culture is growing, the production of the protein is stopped by providing a chemical in the medium and then the number of copies per cell is reduced as a result of dilution as the cells divide.

(a) For this part of the problem, let's focus on the protein dilution effect. Work out a differential equation for the change in protein concentration as

a function of the time that has elapsed since production of the protein was stopped. Solve the equation and relate the decay constant to the cell cycle time. Note that here we are only interested in the dilution that results from the original N copies of the protein being partitioned into an ever-larger number of daughter cells, not in the dilution that occurs as each individual cell lengthens in preparation for the next round of division. Note also that in this part we're interested in a continuous model—you'll look at the discrete version in part (b). HINT: there are two ways to approach this problem. You can consider the change in the concentration as a function of the change in the number of cells into which the original N proteins are partitioned. Or you can note that for a bacterium like *E. coli*, it is a reasonable assumption to imagine that the cell diameter is unchanged and that the size is controlled by the cell length, such that the change in volume with time is simply the change in length with time times a constant prefactor; then consider the change in protein concentration as a function of the change in the total volume into which the original N proteins are diluted.

(b) We can repeat a calculation like that given above using a discrete language in which the number of proteins per cell is a discrete integer. Imagine that before cell division, the number of copies of a given transcription factor in the cell is N. In particular, for every cell doubling, the number of proteins is reduced by a factor of 2. Using such a picture, write a formula for the average number of proteins per cell as a function of the number of cell divisions and relate this result to that obtained in part (a). Furthermore, by using the fact that $2 = \exp(\ln 2)$, reconcile the discrete and continuous pictures precisely.

(c) Interestingly, the model used in part (b) opens the door to one of the most important themes in physics, namely, that of fluctuations. In particular, as the cells divide from one generation to the next, each daughter does not really get $N/2$ copies of the protein since the dilution effect is a stochastic process. Rather the partitioning of the N proteins into daughter cells during division follows the binomial distribution. Analyzing these fluctuations can actually lead to a quantification of the number of copies of a protein in a cell. In this part of the problem, work out the expected fluctuations after each division by noting that the fluctuations can be written as $\sqrt{< (N_1 - N_2)^2 >}$, where N_1 and N_2 are the number of proteins that end up in daughter cells 1 and 2 respectively. Show that $\sqrt{< (N_1 - N_2)^2 >} = N$ (hint: you'll need to use the binomial theorem.)

Next, look at the Rosenfeld paper and explain how measuring fluores-

cence variations can be used to calibrate the exact number of copies of the fluorescent protein in a cell. Assume that the fluorescence intensity in each cell can be written as $I = \alpha N$, where α is some calibration factor and N the number of proteins. Make a plot of $\sqrt{< (I_1 - I_2)^2 >}$ versus I_{tot} and explain how to get the calibration factor α from this plot.

(d) Now we are going to repeat the Rosenfeld experiment numerically in order to *fit* the calibration factor. Consider a fluorescent protein such that the calibration factor between the intensity and the number of fluorophores is 50. Generate intensity data by choosing $N_1 + N_2 = 10, 50, 100, 1000$ and 5000 and for each case, "partition" the proteins from the mother cell to the two daughters 100 times (i.e. as if you are looking at 100 mother cells divide for each choice of the protein copy number). Then, make a plot of the resulting $\sqrt{< (I_1 - I_2)^2 >}$ vs I_{tot} just as we did analytically in the previous problem. What I mean is that you need to make a plot of all of your simulation results. Then, do a fit to your "data" and see how well you recover the calibration factor that you actually put in by hand. Plot the fit on the same graph as all of the "data".