PHYS3511-Biological Physics Fall 2018, Assignment #8

Exercise 1) Now suppose you purify a DNA sample of an organism, consisting of double-stranded DNA molecules each of length 2000000 basepairs (bp), and suspended in a salt solution. Using a light scattering microscope, with resolution of $0.2\mu m$, each molecule appears to be a blob of diameter $1.2\mu m$.

a) Using the appropriate equation estimate the length of one basepair. Compare your results with the known size of ~0.34 nm per basepair.

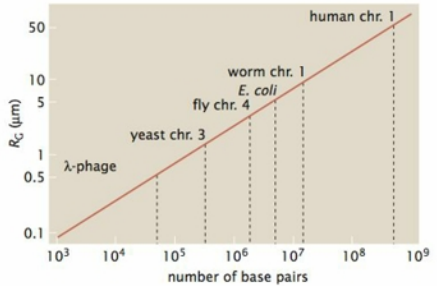
Using equation 8.4 $\langle R^2\rangle = Na^2$, with a being the Kuhn length, and N is the number of Kuhn length in the DNA/polymer. On page 319, in the example "Estimate: End-to-End Probability for the E. Coli genome", it is stated that "**An open DNA chain can be modeled as a random walk…. Since the Kuhn length for bare DNA is roughly 300 bp".** If we assume this then we have $a = 300L_{bp}$, where L_p is the average length of one basepair (bp), and $N = \frac{N_{bp}}{300}$, where $N_{bp}$is the number of bp of the DNA. This gives $\langle R^2\rangle = Na^2 = 300N_{bp}L_{bp} \rightarrow \sqrt{\langle R^2\rangle} = \sqrt{300N_{bp}}L_{bp}$. Since the diameter is $1.2\mu m$, we have $\sqrt{\langle R^2\rangle} = \sqrt{300N_{bp}}L_{bp} = \frac{1.2\mu m}{2} \rightarrow L_{bp} = \frac{0.6\mu m}{\sqrt{300\times 2000000}} = 2.5\times 10^{-5}\mu m$, which is $L_{bp} = 0.025\, nm$. On the other hand if we assume that $a = L_{bp}$, then $\langle R^2\rangle = N_{bp}L_{bp}^2$, and we have $L_{bp} = \frac{0.6\mu m}{\sqrt{2000000}} = 4.24\times 10^{-4}\mu m = .424nm$, which is only a bit larger than the known size of ~0.34 nm per basepair.

b) Consider the DNA of a particular strain of E. Coli, which is a single-stranded DNA of length 5386 bp. Would this DNA be visible using the same light microscope as part a)? If your answer is no, what is the minimum resolution of a microscope that can resolve this DNA blob?

Let's assume for the purpose of this question that the equation $\langle R^2\rangle = N_{bp}L_{bp}^2$, with $L_{bp} = 0.34nm$. For a system of $N_{bp} = 5386$, $\langle R^2\rangle = N_{bp}L_{bp}^2 \rightarrow size{\sim}2\times\sqrt{\langle R^2\rangle} = 2\sqrt{N_{bp}}L_{bp} = 2\sqrt{5386}\times 0.34nm = 49nm$ or $0.049\mu m$, which means that it will not be visible to a microscope of resolution $1.2\mu m$.

Important point: I copied this question from another assignment, without knowing that there are a few errors. For example, E. Coli DNA has $5.4\times 10^6 bp$, and ~5400 genes. The question erroneously states that it has about 5386 bp.

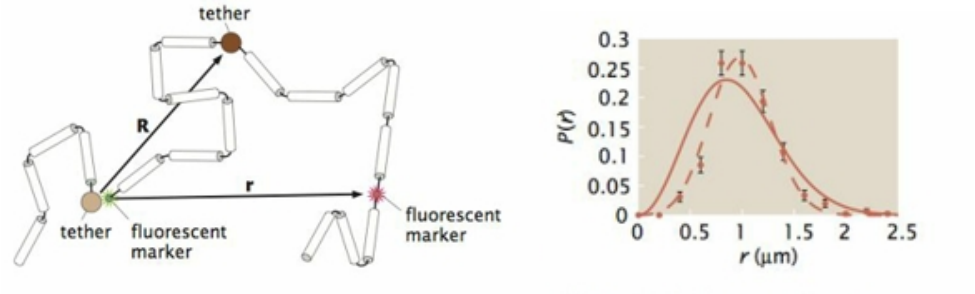
More accurate information can be gleaned from looking at the Figure 8.6



For example, in the graph E. Coli has about $5.4\times 10^6 bp$. If we use the equation, $\langle R^2\rangle = Na^2$, with a = 300 bp$\times\frac{1}{3}nm = 100nm$, and the number of Kuhn length, $N = \frac{N_{bp}}{300} = \frac{5.4\times 10^6}{300}$, or $N = 18000$, $size{\sim}2\times\sqrt{\langle R^2\rangle} = 2\times\sqrt{N}a = 2\times\sqrt{18000}\times 100nm = 2.7\times 10^4 nm = 26\mu m$, which is still larger than $\sim 10\mu m$ on the graph. But the plot use radius of gyration from equation 8.32 and 8.33

$size \sim \sqrt{\langle R_G^2 \rangle} = \sqrt{\frac{L\xi_p}{3}}$, with the persistent length $\xi_p = \frac{a}{2} = 50nm$, and $L = N_{bp}L_{bp}$, $L_{bp} = 0.34nm$. For E.Coli, $size \sim \sqrt{\langle R_G^2 \rangle} = \sqrt{\frac{5.4\times10^6 bp \times (0.34nm/bp) \times 50nm}{3}} = 5531nm \sim 5.5\mu m$, which matches the data of Figure 8.6. **Use this method for the final exam.******

Exercise 2) Consider equation 8.36 for a free-polymer distribution, and 8.37 for a tethered polymer. The result is shown in the plot of figure 8.12 (see below), for distance 100 kb between the two fluorescent tags, and $N/a^2 = 0.5\mu m^2$, $R \sim 0.9\mu m$, usually the Kuhn is taken to be about a = 300 bp, with the length of a base pair being $\sim \frac{1}{3}nm = 0.34nm$. In the equation N is the Kuhn segments between the two markers, and $N/$ is the Kuhn segments between the second tether and the marker.



$$P(r) = \left(\frac{3}{2\pi Na^2}\right)^{3/2} 4\pi r^2 e^{-3r^2/2Na^2}, \qquad (8.36)$$

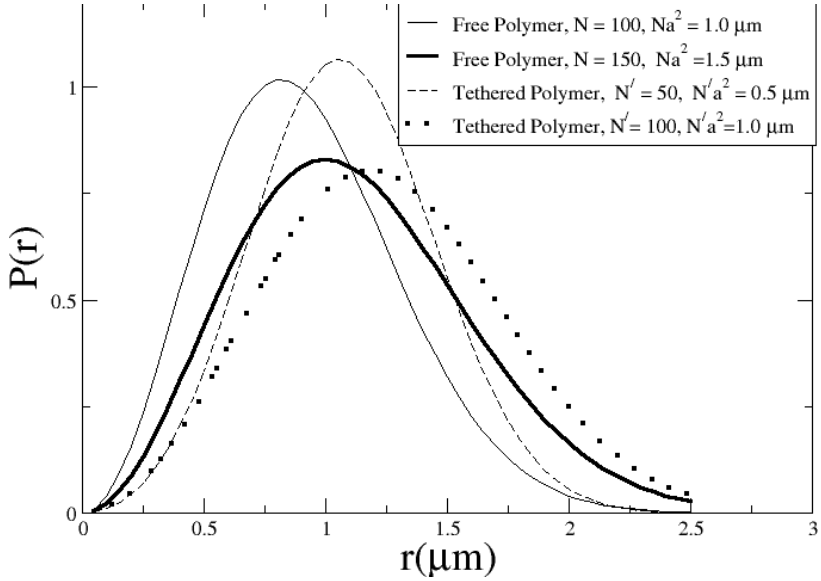
for the free-polymer case, and

$$P(r) = \left(\frac{3}{2\pi N'a^2}\right)^{1/2} \frac{r}{R}\left(e^{-3(r-R)^2/2N'a^2} - e^{-3(r+R)^2/2N'a^2}\right) \qquad (8.37)$$

A) Reproduce the plot above
If we assume that the original plot use $N/a^2 = 0.5\mu m^2$, with a = 100 nm, then $N/a^2 = 0.5\mu m^2 \rightarrow N/ = \frac{0.5\times10^6 nm^2}{(100nm)^2} = 50$ **Kuhn segments**. If we assume as in figure caption of 8.12 that the DNA has 100 kb = $10^5 bp$, then we should have $N = \frac{N_{bp}}{300} = 333$ Kuhn segments. However, if you read page 326, it is clear that figure 8.12 uses the value $Na^2 = 1\mu m^2$, which gives N = 100 Kuhn segments. If we use this number for plotting equation 8.36 and 8.37, we obtain the result shown below (on next page): thin solid line for free polymer with $Na^2 = 1\mu m^2$; dashed line for tethered polymer with $N/a^2 = 0.5\mu m^2$. Note that the positions of the maxima are the same as in figure 8.12, but the probability values P(r) are higher. This will be explained below.

B) Double the number of segments $N/$, and plot the result and compare with A).

If we doubled $N/ \rightarrow N/ = 100$, then $N/a^2 = 1.0\mu m^2$. Looking at the picture of the polymer we can also say that this means $N \rightarrow N = 150$, and $Na^2 = 1.5\mu m^2$. The plot is shown on the next page.



****The most important point to note is that the plots of part B, shift the maxima to the right, and broaden the plots.

****Important Point:** You may note that the probability of some plot has value greater than 1, P(r) > 1. This seems illogical since probability should not be > 1. But please note that the probability:

$$P(r) = \left(\frac{3}{2\pi N a^2}\right)^{3/2} 4\pi r^2 e^{-\frac{3r^2}{2Na^2}},$$

and

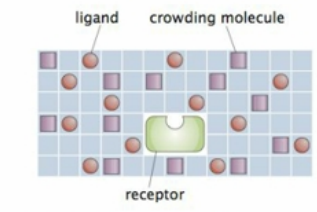
$$P(r) = \left(\frac{3}{2\pi N a^2}\right)^{1/2} \frac{r}{R}\left(e^{-\frac{3(r-R)^2}{2N/a^2}} + e^{-\frac{3(r+R)^2}{2N/a^2}}\right)$$

has unit of inverse length (m^{-1}). This means that the probability must be multiply by an arbitrary length $\Delta r \rightarrow P(r)\Delta r$. If we use $\Delta r = 0.25\mu m$, then we the probability will have the same values as in Figure 8.12 (on previous page).

****ADVICE** verify in the above that P(r) has unit of inverse length.

Exercise 3)

A) Using Stirling's approximation derive equation 14.3, where the values of L and C are not negligible. Reproduce the plot of 14.7, but using one lattice volume being $V_{box} = 1nm^3$, with P_{bound} vs. $[L]$ in M.



From class notes: We begin by calculating the partition function $Z = \sum_E g(E)e^{-\beta E}$, $g(E)$is the multiplicity (number of microstates) of states with energy E.

Definition:

- Ω is the number of lattice
- L is the number of ligands
- C is the number of crowding molecules

The multiplicity is a function of Ω, L and C

$$g(\Omega, L, C) = \frac{\Omega!}{L!\, C!\, (\Omega - L - C)!}$$

$$P_{bound} = \frac{1}{1 + \frac{g(\Omega, L, C)}{g(\Omega, L-1, C)} e^{\beta(\varepsilon_b - \varepsilon_L^{sol})}}$$

Assume Ω, L and C are large $\frac{g(\Omega,L,C)}{g(\Omega,L-1,C)} = \frac{(L-1)!(\Omega-L-C+1)!}{L!(\Omega-L-C)!} = \frac{\Omega-L-C}{L}$

$$P_{bound} = \frac{1}{1 + \frac{\Omega - L - C}{L} e^{\beta\Delta\varepsilon}}$$

$$\Delta\varepsilon = \varepsilon_b^{sol} - \varepsilon_L^{sol}$$

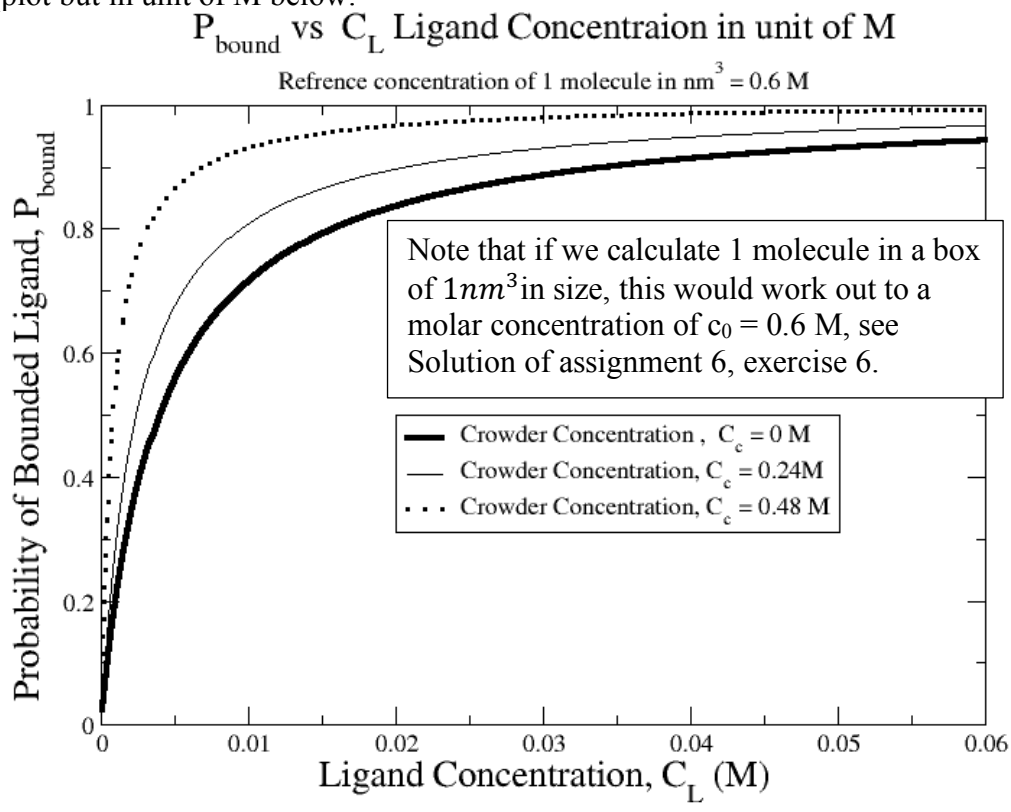
Defining the total volume as ΩV_{box}, with V_{box} being the arbitrary box size.

$$P_{bound} = \frac{1}{1 + \frac{\frac{1}{V_{box}} - \frac{L}{\Omega V_{box}} - \frac{C}{\Omega V_{box}}}{\frac{L}{\Omega V_{box}}} e^{\beta\Delta\varepsilon}}$$

In this case we define the concentration: reference, $c_0 = \frac{1}{V_{box}}$; Ligand concentration, $c_L = \frac{L}{\Omega V_{box}}$; crowders concentration $c_c = \frac{C}{\Omega V_{box}}$, and

$$P_{bound} = \frac{1}{1 + \frac{c_0 - c_L - c_c}{c_L} e^{\beta\Delta\varepsilon}}.$$

As usual we use $c_0 = 0.6M$, and from figure 14.7, $\frac{\Delta\varepsilon}{k_B T} = -5.0$, we reproduce the plot but in unit of M below:



B) Consider the opposite of Figure 14.9, where a crowding molecule occupy **four lattices**, while a ligand occupies **one lattice**. Derive an equation similar to 14.4.

$$P_{bound} = \frac{1}{1 + \frac{g(\Omega, L, C)}{g(\Omega, L-1, C)} e^{\beta(\varepsilon_b - \varepsilon_L^{sol})}}$$

$$g(\Omega, L, C) = \frac{\Omega!}{C!\,(\Omega - C)!} \times \frac{(r\Omega - rC)!}{L!\,(r\Omega - rC - L)!}$$

$$\frac{g(\Omega, L, C)}{g(\Omega, L-1, C)} = \frac{(L-1)!}{L!} \frac{(r\Omega - rC - L + 1)!}{(r\Omega - rC - L)!} = \frac{r\Omega - rC - L + 1}{L}$$

For large Ω, C, L

$$P_{bound} = \frac{1}{1 + \frac{r\Omega - rC - L}{L} e^{\beta\Delta\varepsilon}}, \Delta\varepsilon = \varepsilon_b^{sol} - \varepsilon_L^{sol}$$

Using the same definition of concentration as before, and with r = 4,

$$P_{bound} = \frac{1}{1 + \frac{4c_0 - c_L - 4c_c}{c_L} e^{\beta\Delta\varepsilon}}$$

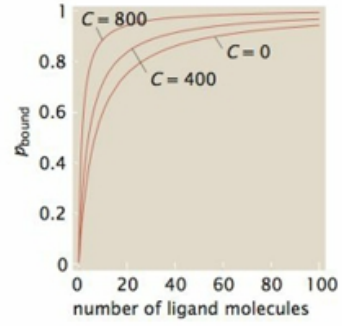


Figure 14.7: Probability of a protein binding site being occupied by a ligand for a number of different concentrations of the crowding molecules. The reaction volume is $\Omega = 1000$ and $\Delta\varepsilon_L = -5\ k_BT$.

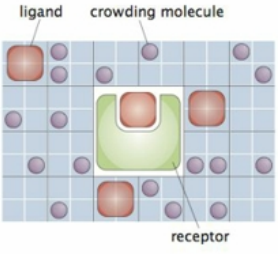


Figure 14.9: Lattice model for large ligands. This lattice model describes binding in the presence of crowding agents where the size of the crowder is different from that of the ligands. This is represented by using different-size boxes for the crowder and the ligand.

****Note**: Conclusion for exercise 3
It is clear in part A) that a crowded environment induce cooperativity, i.e. increase the probability of ligand binding. But do the **large crowding molecules** of part B) increase binding cooperativity? To answer note that $\frac{c_0 - c_L - c_c}{c_L}$ of A) is smaller than $\frac{4c_0 - c_L - 4c_c}{c_L}$ of B), for the same values of c_L and c_c. This means that P_{bound} of part B) is smaller. Can you see why? Ask yourself whether larger crowder are **more** or **less** effective at **inducing cooperativity in binding**?

$$p_{\text{bound}} = \frac{1}{1 + \frac{\Omega - L - C}{L} e^{\beta\Delta\varepsilon_L}}, \qquad (14.3)$$

$$p_{\text{bound}} = \frac{1}{1 + \frac{Z_{\text{sol}}(L, C)}{Z_{\text{sol}}(L-1, C)} e^{\beta\Delta\varepsilon_L}} = \frac{1}{1 + \frac{\Omega}{L}(1 - \phi_C)^r e^{\beta\Delta\varepsilon_L}}$$

Exercise 4) Chapter 14, Problem 14.1
Chapter 14 of the textbook argued that the mean spacing between molecules in an *in vitro* biochemical experiment is roughly 100 nm at μM concentration, while in the cell the spacing are a factor of 10 smaller. Justify these statements with simple estimates. The biochemical "**standard state**" is often taken as 1M. Work out the mean spacing at this concentration.

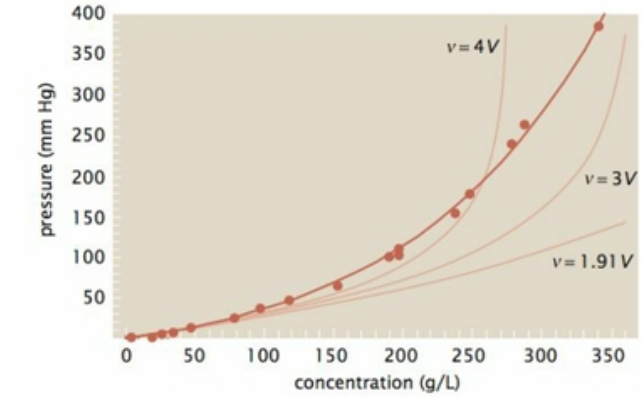
For *in vitro* $1\mu M = 1 \times 10^{-6} M$, with 1 M $= \frac{1mol}{L} = \frac{1mol \times 6.023 \times 10^{23} mol^{-1}}{1L \times 10^{-3} m^3 \cdot L^{-1}} = 6.023 \times 10^{26} m^{-3}$, so that $1\mu M = 6.023 \times 10^{20} m^{-3}$. The mean spacing is found by taking the inverse cube root $\left(\frac{1}{6.023 \times 10^{20} m^{-3}}\right)^{1/3} = 1.18 \times 10^{-7} m$, which is about 118 nm.

For *in vivo* (living cell), with 1 M $= \frac{1mol}{L} = \frac{1mol \times 6.023 \times 10^{23} mol^{-1}}{1L \times 10^{-3} m^3 \cdot L^{-1}} = 6.023 \times 10^{26} m^{-3}$.
The mean spacing is found by taking the inverse cube root $\left(\frac{1}{6.023 \times 10^{26} m^{-3}}\right)^{1/3} = 1.18 \times 10^{-8} m$, which is about 11.8 nm, which is ten time lower than the *in vitro* value.

Exercise 5) Chapter 14, Problem 14.3
Consider equation 14.11 on the partial pressure of non-dilute solute

$$p = k_B T[H](1 + x + 0.625x^2 + 0.287x^3 + 0.110x^4)$$

In this case $[H]$ is the number concentration of hemoglobin proteins. The solid line in the plot below show that equation 14.11 is able to quantified the pressure vs concentration of hemoglobin in solution. This exercise asks students to verify this claim, by substituting one or more data points into the equation.



Detail:
1) mass of 1 Hemoglobin, M_H = 64 000 Da
2) with V = volume of sphere of diameter 5.8 nm = $1 \times 10^{-25} m^3$
3) For concentration 200 g/L $\rightarrow x = V[H] \sim 0.748$, $P \sim 17000 Pa$

Note: 1 mm Hg $\sim 133 Pa \rightarrow P \sim 17000 Pa \div 133 mm \frac{Hg}{Pa} = 127\ mm\ Hg$, which more or less matches the value on the graph above.