**Announcements**

Office hours this week:
4-5 pm (right after section) on Tuesday the 8th

Wednesday from 7-8:30 pm on Zoom

or by appointment.

Section next week:
In-person on November 15th at 3 pm in M217.

Paper of the week Not new papers, but two classic papers for modeling biopolymers:

- *Stretching DNA* Macromolecules 1995, 28, 26, 8759–8770—section IIA derives the classic interpolated force-extension relationship for worm-like chains.
- *Stiff Chains and Filaments Under Tension* Macromolecules 1995, 28, 20, 7016–7018—considers a worm-like chain that can both bend and undergo elastic deformation and derives a useful force-extension relationship. This version of the worm-like chain model is particularly useful for ssDNA and RNA under high forces.

**Problem 1: The mean squared extension of a worm-like chain**

In this problem, we will compute the mean-squared end-to-end extension of a worm-like chain polymer. This is a similar (if slightly harder) calculation to problem 3 on your homework.
**a)** We can obtain the mean-squared extension by summing over all possible paths $p$ of the polymer:

$$\langle \vec{r}^2 \rangle = \sum_p \vec{r_p}^2 P_p, \tag{1}$$

where $r_p$ is extension of a path and $P_p$ is the corresponding probability. We can rewrite this expression to more explicitly show how $\vec{r_w}^2$ is defined:

$$\langle \vec{r}^2 \rangle = \sum_p \left[ \int_0^L \vec{t}(s) ds \right] \left[ \int_0^L \vec{t}(s) ds \right] P_p. \tag{2}$$

Now for the sneaky part: rearrange this expression so that it is a double integral over the expected tangent-tangent correlation of the polymer. Then, replace this expectation value with the functional form we found in class.

**b)** Evaluate this double integral. You may find it helpful to split the inner integral into two terms.

**c)** Evaluate your expression from (b) in two limits: $L \ll \ell_p$ and $L \gg \ell_p$. Comment on the results.

**Problem 2: DNA looping probabilities and energies**

DNA looping is crucial to gene expression. The Lac repressor is a tetrameric protein that binds to DNA and downregulates the expression of proteins that are used in lactose metabolism.

There are a total of three specific binding sites on the DNA (O1, O2, O3) for the Lac repressor, and full repression requires that both O1 and O2 are bound, or that both O1 and O3 are bound. The looped region can be either 401 or 92 base pairs in length.[1]

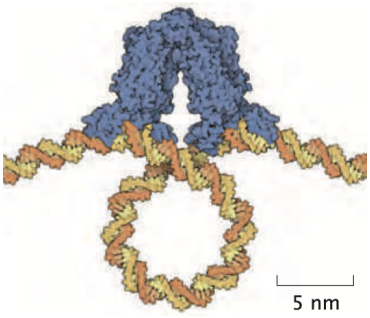


**Figure 8.19:** Model for DNA loop formation by the Lac repressor. The interface between the protein and the DNA was determined by X-ray crystallography (PDB 1lbh, 1efa), but the overall position and shape of the DNA in the loop is an artist's rendition. (Courtesy of D. Goodsell.)

(Image from PBOC)

**a)** The probability density function for the displacement $\mathbf{r}_{AB}$ between two points $A$ and $B$ on a polymer in the freely jointed chain model in 3D is

$$P(\mathbf{r}_{AB}) = \left(\frac{3}{2\pi N b^2}\right)^{3/2} \exp\left(-\frac{3|\mathbf{r}_{AB}|^2}{2Nb^2}\right) \tag{3}$$

for a free polymer in 3D. $N$ is the number of links, and $b$ is the length of each link.

We want to compute the probability $p_{\text{loop}}$, that the polymer happens to form a loop connecting the two ends. We will say that a loop is formed when they come within a critical distance $\rho$.

Set up a three-dimensional integral (in spherical coordinates!) for $p_{\text{loop}}$ using the probability distribution given.

**b)** Simplify the integrand assuming that the critical distance $\rho$ is much smaller than the typical polymer extent $Nb^{1/2}$. This should let you evaluate the integral in one line. How does the looping probability depend on $N$?

**c)** Using $b = 2\ell_p$, estimate the looping probability for DNA at the length of the lambda phage genome ($\approx 50,000$ base pairs); you will have to convert the number of base pairs into a length[2]. Use $\ell_p$ = 50 nm for dsDNA. Propose a reasonable value for $\rho$.

**d)** Now let's look at the energetic cost of forming a loop. In class we derived that

$$E_{\text{loop}} = \frac{\alpha\pi}{R} \tag{4}$$

where $R$ is the radius of the loop and $\alpha = \ell_p k_B T$ is the bending modulus. Compute the energy it takes to loop a strand of DNA consisting of $N_{bp}$ base pairs.

Note that the 92 base pair loop is quite difficult to form — 92 base pairs is even smaller than the persistence length of DNA (!!)

**e)** Compute the entropy loss of forming a loop $\Delta S_{\text{loop}}$ as well as the energetic cost of loop formation $\Delta E_{\text{loop}}$, using the results from parts **b)** and **c)**.

What is the free energy $\Delta G_{\text{loop}}$ for loop formation as a function of $N_{bp}$? Where is it minimized?

(This problem is partly from Physical Biology of the cell and partly from Prof. Brian Camley.)

[1] By the way, how the protein finds its appropriate binding site is a question we explored in section 6! https://en.wikipedia.org/wiki/Lac_repressor#Search_kinetics_of_DNA_binding

[2] http://bionumbers.org is a great resource for looking up parameters!