TUTORIAL 4

BB 101

MODULE: *PHYSICAL BIOLOGY*

Ambarish Kunwar

Lab No. 204
Department of Biosciences and Bioengineering
IIT Bombay

akunwar@iitb.ac.in

http://www.bio.iitb.ac.in/~akunwar/

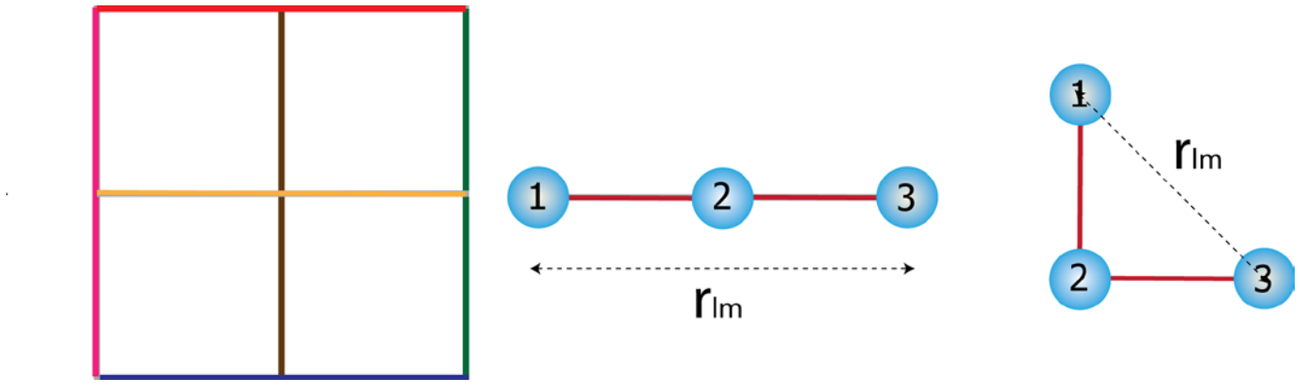
1. Explain why a freely jointed chain left to itself tries to coil itself. What is it trying to maximize by doing that?

2. A hypothetical polymer chain of 100 segments of length b = 3 Å has the root-mean-square (RMS) end-to-end distance of 100 Å. Does this chain behave as an ideal freely-jointed chain?

3. DNA is stiff because it is a double helix. At room temperature, the persistence length of DNA molecule was found to be 500 Å. Calculate the bending stiffness of DNA in kcal mol-1 ? (Hint: Formula discussed in the lecture is for one molecule)

4. Prove the relation $G = -k_B T ln\, Z$ using relations G=<U>-TS. Given <U>=$\sum_i u_i p_i$, S=-$k_B \sum_i p_i ln p_i$ and Z =$\sum_i e^{-\beta u_i}$ where $\beta = \frac{1}{k_B T}$ and $p_i = \frac{1}{Z} e^{-\beta U_i}$

5. Imagine a protein made of three identical/indistinguishable connected positive charges. The length of the bond between two neighboring charges of protein is $1\,nm$. This three-charge protein is lying on a 3×3 square lattice in 2D (or a 2D grid connecting 9 lattice sites) as shown below. Color of the grid line denote the spatial inhomogeneity such that all possible conformations/microstates become unique and are not related by rotational symmetry



The Coulomb energy of the protein, in a conformation/microstate i is given by the typical formula for energy,

$$U_i = \sum_{l=1}^{2} \sum_{m=l+1}^{3} \frac{A}{r_{lm}}$$

Where r_{lm} is the distance between charges l and m. Assume $A = 1\ k_B T\ nm$. Note that the charges can only lie on the sites of the lattice and the bonds on the edges.

(a) What is the energy of the protein in the conformation/microstate when all the three charges on a straight line?

(b) What is the energy of the protein in the conformation/microstate that is bent (non-straight; when one bond is making 90^o angle with the other one)?

(c) How many straight conformations are possible on this square lattice?

(d) How many bent conformations are possible on this square lattice?

(e) What is the probability that you will find the protein in a straight structural state or straight macrostate?

(f) What is the probability that you will find the protein in a bent structural state/macrostate?

6. During evolution, some genes get mutated and the resulting proteins get altered. In biology, it is very useful (and often important) to find out the DNA sequence that is “conserved" during evolution. Entropy can be a simple measure of this conservation (or the lack of it) during evolution. Let us imagine you got 10 DNA sequences (say, from 10 different organism). Each of these sequences have 3 bases as shown below.

AAT

AGT

ATA

ACG

ATT

AGT

ACT

AAC

ATT

AGT

(i) Calculate the entropy (disorder) at each position (column) using following relation

$$S = -k_B \sum_{i=1}^{M} p_i \ln p_i$$

where M is the number of different letters in each position (column) and $p_i = n_i$/N, where n_i is the number of letters of type i in the column, and N is the total number of letters in that position (column).

(ii) Calculating entropy for each position (column)? Find out which position is more “conserved" over evolution and which position is least conserved over evolution

Notes: Those highly conserved positions are likely to have some crucial role in the function/folding of the protein. This also tells you how to use information theory {theory used for communication by electrical engineers} to understand information content in biological sequences.