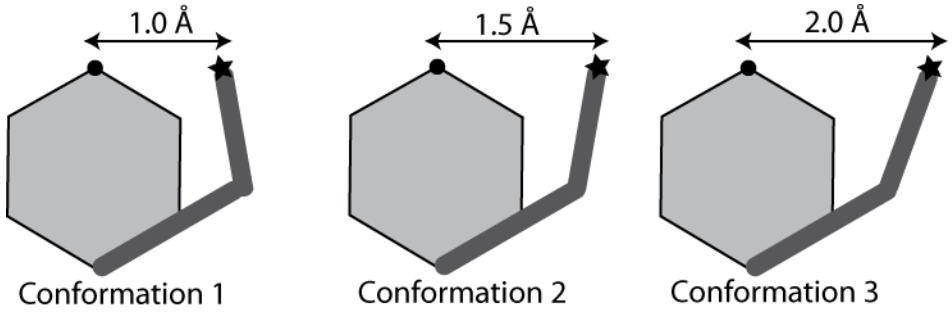
BB 101: Physical Biology

TUTORIAL 3

1. A peptide loop on a protein molecule was probed using fluorescence spectroscopy to measure the separation l from a point on the loop to a point on the protein as shown in figure below. After experiment you decide to model the loop as having following three different conformations:

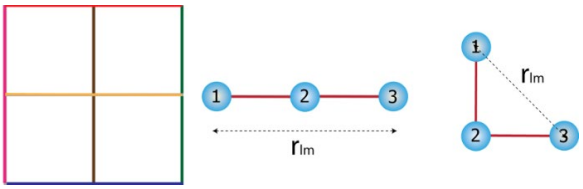


(1) In first conformation, the loop sticks to the side of the loop with separation l=1.0 Å; and you define this as the ground state, with energy ϵ=0 pN nm
(2) In second conformation, the loop is more distant from the protein with separation l=1.5 Å; you define this as first excited state with energy ϵ=4.14 pN nm
(3) In third conformation, the loop is far away from the protein with separation l=2.0 Å; and you define this as second excited state with energy ϵ=8.28 pN nm

Using above model, calculate following at T=300 K:

(a) Partition function for the loop **(b)** Average separation i.e. $\langle l \rangle$ **(c)** Average energy of the loop $\langle \epsilon \rangle$

2. 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° 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?

3. Show that expression for entropy $S = k_B ln W$ is equivalent to $S = -k \sum_i p_i \, ln p_i$ if k_B is replaced by k for a single macro-state assuming that all microstates have identical energies.

4. 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.