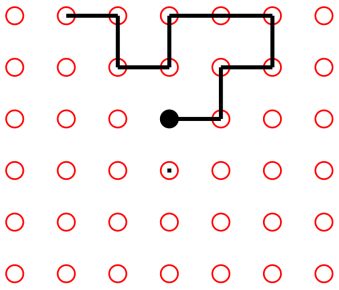
CHEM-AD315/Spring 15

Homework Assignment 2

Due @ 11:59pm Thursday 12th March by email serdal@nyu.edu or to ERB 114

This sign means python or any kind of scientific computing is needed to get the answer. Please provide your answer with your python code.

1. To study the statistical mechanics of protein folding scientists have been using lattice models for the last three decades. In this simplest representation the conformational space is restricted to discrete lattice points in 2 or 3 dimensions that allows reducing the complexity of the problem drastically. Here each amino acid is represented by one bead that is connected to the adjacent neighbors with a bond. A protein conformation is generated using Self Avoiding Walk (SAW) algorithm. In this algorithm we start from an arbitrary point on the lattice and walk one step in one of the four available directions and put our second amino acid. From the second amino acid we repeat this for the third one and so on, with a condition that no two amino acid will occupy the same lattice point.



In this model the amino acids (so the beads) are of two kinds: Hydrophobic (H) residues, and Polar (P) residues with interaction energies tabulated below.

Figure 1. Self-Avoiding walk of a protein chain in 2D lattice

The distance dependence of the interactions is given in the equation below:

$$V_{ij} = \begin{cases} \varepsilon & d_{ij} = 1 \\ 0 & \text{otherwise} \end{cases}$$

ε(kcal/mol)	H	P
H	-1	0
P	0	0.

Table1. Interactions of different kind of amino acid pairs

a. All possible configurations for a 10-residue mini-protein are generated computationally by Self Avoiding Walk algorithm and is given in file N10.txt. Each line of the file represents one conformation with values x1,y1, x2,y2, …, x10,y10. Assume that you have a protein sequence *HPHPPHPPHH.* Find the minimum energy state, report its energy and plot the structure. (0.5pts)

b. Find all conformations with the same energy. How many states do you observe? Plot a set of representative conformations for each state. Plot degeneracy as a function of energy. What do you see? (0.5pts)

c. Find the population of each state as a function of temperature. Plot them with different colors for the range $100K < T < 500K$ with 1K intervals. What happens to the protein in high temperatures? (0.5pts)

d. Overall size and shape of a protein is measured by Small Angle X-Ray scattering experiments. SAXS measurements can be used to estimate the Radius of gyration. Radius of gyration can be directly calculated from the positions of beads as $R_g^{\,2} = \frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}(r_i - r_j)^2$ where N is the number of beads in our chain. Estimate the average value of the Radius of gyration for the solution of 10^{23} proteins and plot R_g^2 as a function of temperature for $100K<T<500K$. (0.5pts)

e. Assume that you have 10^{23} proteins in your solution. Calculate internal energy per protein as a function of temperature. Use $100K < T < 500K$ as your temperature range and $\partial T = 1K$ your temperature interval. To calculate the internal energy use the expectation value formula $\langle u \rangle = \sum_{i}^{N_{conf}} E_i p_i$ where N_{conf} is the total number of conformations. (0.5pts)

f. Calculate the partition function as a function of temperature and calculate internal energy from $\langle u \rangle = k_B T^2 \frac{\partial \ln(Q)}{\partial T}$ by using finite difference method $\frac{\partial \ln(Q)}{\partial T} \cong \frac{\ln Q(T_{i+1}) - \ln Q(T_i)}{T_{i+1} - T_i}$. Use the same temperature interval as **e** and compare your result by plotting both methods. (0.5pts)

g. The stability of proteins can be studied by varying the temperature in the experiments. Here we slowly increase the temperature until the folded protein goes to a conformational change and unfold, forming an extended

structure. The temperature at which the folded and unfolded populations are equal in solution is called the folding temperature. To find folding temperature of our proteins we will use partition function. Folding temperature is the maxima in the heat capacity versus temperature plot. Calculate the heat capacity as a function of temperature using $C_V = N\left(\frac{\partial\langle u\rangle}{\partial T}\right)$. Again to obtain the derivative use the finite difference method. You have already calculated internal energy as a function of temperature in 1f. Compare your protein with HPPHPPHPHH and HHHHHPPPPPP. These sequences have the same ratio of H/P. Do you see a change in their stability. Order them according to their stability. (0.5pts)

h. Can you design a sequence that is more stable that the ones here keeping the H/P ratio fixed? (Bonus 0.5pts) 

2. Suppose that we have N molecules of Ar gas in a container of Volume V and temperature T. The gas is ideal thus no interaction between individual atoms. Hence the Hamiltonian is only Kinetic energy and it is $H = \frac{1}{2}mv^2$.

a. Calculate the velocity distribution in the container. Find the average of the velocity. (Hint use $\int_{-\infty}^{\infty} e^{-ax^2}\,dx = \sqrt{\frac{\pi}{a}}$ and $p(v_x)=p(v_y)=p(v_z)$). (0.5pt)

b. Calculate $\langle v_x^{\,2}\rangle$ using $p(v_x)$ from a and derive that average kinetic energy is $k_B T/2$. (Hint use $\int_{-\infty}^{\infty} x^2 e^{-ax^2}\,dx = (1/2a)\sqrt{\pi/a}$) (0.5 pts).

3. Sarah is an NMR specialist who is trying to combine experimental data and theoretical calculations. One day she came up with a very interesting observation: The dominant molecular conformation in her sample is not the lowest energy state according to the theoretical calculations. She is puzzled with this result. How can that be happening? The minimum energy state should have a higher Boltzmann factor. Can you help her to understand her findings? What do you think the reason is? What should she do to understand the reason behid it further? (0.5pts)