20.110/5.60/2.772 Fall 2005

Homework # 5

Due Friday October 21

1.) **ELVIS is everywhere**

a.) Given 20 naturally occurring amino acids, what is the probability that the amino acid sequence ELVIS occurs in a stretch of a protein sequence?

b.) What is the probability if the order of the amino acids did not matter, i.e., VLSEI, etc.?

2.) **Levinthal's Paradox**:

Cyrus Levinthal, a former MIT professor, claimed that proteins can not go from unfolded to folded state by sampling of all possible states. He claimed that it would take way too long for a protein to fold if they were to sample all possible conformations (proteins fold on the order of ~ms to ~μs). Assume that a protein has 100 amino acids and each amino acid has two degrees of freedom (angles of rotation called ϕ and ψ), also assume that it takes a femtosecond to sample each conformation. Using simple statistics show that Prof. Levinthal was indeed right and that it would take a very long time (much longer than our lifetimes!) if the protein was to sample all possible conformations. If proteins were to fold by sampling all possible states, calculate the change in entropy upon folding, assume that the unfolded state has all the possible states and the folded state is unique (i.e there is only one folded state). (Hint : Use Boltzmann's definition of entropy).

3.) **The statistical thermodynamics of a cooperative system.** (Dill and Bromberg)

Perhaps the simplest statistical mechanical system having 'cooperativity 'is the three-level system in the following table.

Energies	$2\varepsilon_0$	ε_0	0
Degeneracies	γ	1	1

a) Write an expression for the partition function q as a function of energy ε, degeneracy γ ,and temperature T.

b) Write an expression for the average energy $< \varepsilon >$ versus T .

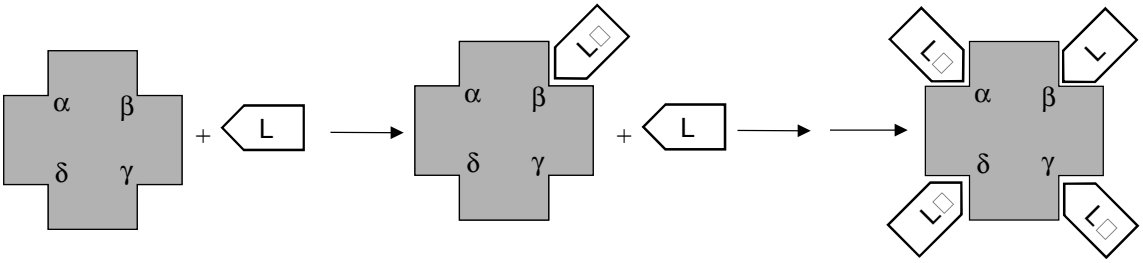
c) For ε 0 $/kT$ =1 and γ =1, compute the populations, or probabilities, p_1^*, p_2^*, p_3^* of the three levels.

d) Now if ε_0=2 kcal mol^{-1} and γ =1000, find the temperature T_0 at which $p_1 = p_3$.

e) Under condition (d), compute p_1^*, p_2^*, p_3^* at temperature T_0.

4.) **Multiple ligand binding sites on a protein.**

The protein below has four distinguishable binding sites (α, β, γ, δ) for the ligand L. We would like to find its equilibrium binding population. For now assume that the association and dissociation constants are equal.



a) Calculate **Ω** and the entropy (in units of k) for the situation in which

i) 0 ligands are bound ($N_L = 0$)

ii) 1 ligand is bound ($N_L = 1$)

iii) 2 ligands are bound ($N_L = 2$)

iv) 3 ligands are bound ($N_L = 3$)

v) 4 ligands are bound ($N_L = 4$)

b) Which states have the **highest** entropy?

c) Which states have the **lowest** entropy?

e) Let's say that the binding constants are not equal, i.e., ligand L has a higher probability of being bound:

$p_{bound} = 0.75$

$p_{unbound} = 0.25$

Plot the probability distribution for all of the states.

5.) **Entropy depends on distinguishability** (Dill and Bromberg)

Given a system of molecules at $T = 300$ K, $q = 1 \times 10^{30}$,and $\Delta U = 3740$ J mol^{-1} ,

(a)What is the molar entropy if the molecules are distinguishable?

(b)What is the molar entropy if the molecules are indistinguishable?

6.) **The pressure reflects how energy levels change with volume.** (Dill and Bromberg)

If energy levels $\varepsilon_i(V)$ depend on the volume of a system, show that the pressure is the average

$$p = -N\left\langle \frac{\partial \varepsilon}{\partial V} \right\rangle$$

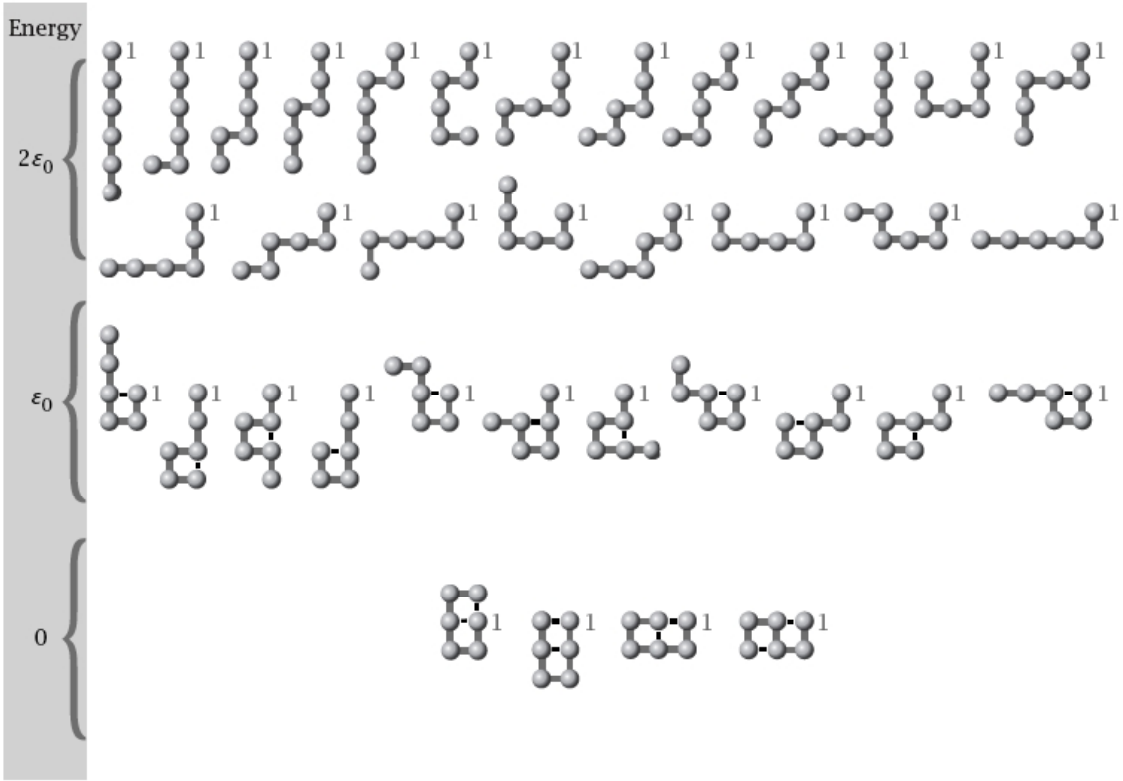
7) **Populations of energy levels**

a)Given a two state system, where the energy level of the lower state is at 600 cal/mol and the higher level energy is 1800 cal/mol, calculate the partition function, average energy and entropy at T = 300 K and T = 500 K.

b) Now imagine that you have a five state system, with probabilities p_1= 0.85, p_2= 0.1, p_3=0.04, p_4=0.009 and p_5 = 0.001 at T =300 K. What are the energies of the third and the fifth energy level relative to the ground state, relative to the first excited state?

8) **Denaturation occurs over a temperature range**.

An important prediction we can make with simple statistical mechanical models is the dependence of protein folding on temperature. Let's show this for a variation on the polypeptide model we used in class. The model is sketched below: we will consider a 6-bead chain model of the polypeptide, and we shall define the energies of the system differently than in lecture, as shown below:



(Dill and Bromberg, Molecular Driving Forces)

The energy increment of the unfolded states, ε_o, is positive. Answer the following questions for this model:

a. What is the degeneracy of each energy state?
b. What is the partition function for the single peptide chain?
c. Write an expression for p(E = 0) and p(E = 2ε), the probability of observing the polypeptide in the folded and unfolded states, respectively, as a function of temperature. Plot these two quantities vs. T for the case where ε = 2.06E10-20 J, and show that every conformation of the peptide becomes equally likely as the temperature approaches ∞.

Additional Problems (no credit)

1.) **Uniqueness of DNA sequences in a gene.**

```
mRNA 5'...GTAGCCGTGAATCGATGGTCCAAATACCG...5'
complement 3'CGGCACT.... 5'
             |----l=?---|
```

Antisense gene regulation is a technique in which short DNA strands are hybridized to mRNA, blocking translation into a protein. The DNA strand needs to hybridize to a specific site on the mRNA. You have a mRNA strand that is 1×10^9 bases long. You would like to choose a site for the DNA sequence to hybridize to that is unique. What is the minimum length of the sequence such that it could be unique (i.e., it is possible that it does not occur anywhere else in the genome)?

2.) SAB 16.1

3.) SAB 16.2