Problem Set 4, due Nov. 8 (Monday, but I may delay until Nov. 10)
I will talk about some of these problems in Tutorial Tuesday.
Reading for Lectures 22—24: PKT Chapter 8

Receptor Cooperativity: Models for Hemoglobin binding of O_2 (PKT, Ch. 7)

A. Dimeric Hemoglobin (model):

The purple blobs are proteins which are aggregated into a dimer.
The clear circles represent binding sites where a ligand may attach.
In the simplest form the binding sites are independent and each one may be described by a σ variable. There are now a total of 4 microstates.
The energetics of binding is described by
$(E-\mu N)(\sigma_1,\sigma_2)=(\varepsilon-\mu)\sigma_1+(\varepsilon-\mu)\sigma_2$.
The probability $P(\sigma_1,\sigma_2)=\frac{1}{\Xi}e^{-\beta[(\varepsilon-\mu)\sigma_1+(\varepsilon-\mu)\sigma_2]}$, with normalisation

$\Xi=\sum_{\sigma_1,\sigma_2=0,1}e^{-\beta[(\varepsilon-\mu)\sigma_1+(\varepsilon-\mu)\sigma_2]}=\left(1+e^{-\beta(\varepsilon-\mu)}\right)^2$, since the two binding sites are independent.

The calculation of physical quantities reduces (in this language) to calculating averages of the "sigma" variables. Thus, the probability of binding at each of the two sites: $P_{B1}=\langle\sigma_1\rangle, P_{B2}=\langle\sigma_2\rangle$ with

$$P_B=P_{B1}=P_{B2}=\langle\sigma_1\rangle=\frac{1}{\Xi}\left(e^{-\beta(\varepsilon_B-\mu)}\right)\left(1+e^{-\beta(\varepsilon_B-\mu)}\right)=\frac{\left(e^{-\beta(\varepsilon_B-\mu)}\right)}{\left(1+e^{-\beta(\varepsilon_B-\mu)}\right)},$$

and the mean number of ligands bound:

$$\langle N\rangle=\langle\sigma_1+\sigma_2\rangle=\frac{1}{\Xi}\sum_{\sigma_1,\sigma_2=0,1}(\sigma_1+\sigma_2)e^{-\beta[(\varepsilon-\mu)\sigma_1+(\varepsilon-\mu)\sigma_2]}=2P_B,$$

which is not surprising, since the two site are independent.
This model gets more interesting when the binding on each site reinforces the other.
Suppose there is an *extra* energy (lowering) J (<0), over and above 2ε when both sites are occupied.
This might come about, for example, by a mechanism whereby a local deformation occurs when one ligand binds which lowers the energy of the second binding site.
To take an extreme example, suppose that ε is large and positive, so that single-ligand binding is very improbable, but J is negative and gives an overall-attractive binding energy of (2ε+J) for two ligands simultaneously. (see Lecture 12 and the Hill Equation).
Then the microstate energies become
$(E-\mu N)(\sigma_1,\sigma_2)=(\varepsilon-\mu)\sigma_1+(\varepsilon-\mu)\sigma_2+J\sigma_1\sigma_2$.
The coupling term $J\sigma_1\sigma_2$ means that the two variables σ_1 and σ_2 are no longer independent. Thus, the Grand partition function no longer factors into two parts. This is what causes correlations between the two variables and, thus, leads to cooperativity.

Calculate Grand partition function: $\Xi=1+2e^{-\beta(\varepsilon-\mu)}+e^{-\beta(2\varepsilon+J-2\mu)}$.
Calculate the mean number of bound ligands:

$$\langle N\rangle=\frac{1}{\Xi}\sum_{\sigma_1,\sigma_2=0,1}(\sigma_1+\sigma_2)e^{-\beta[(\varepsilon-\mu)\sigma_1+(\varepsilon-\mu)\sigma_2+J\sigma_1\sigma_2]}=\frac{\left[0+2e^{-\beta(\varepsilon-\mu)}+2e^{-\beta(2\varepsilon+J-2\mu)}\right]}{\left[1+2e^{-\beta(\varepsilon-\mu)}+e^{-\beta(2\varepsilon+J-2\mu)}\right]}$$

Can also write down similar expressions for P_n, the probability of binding n ligands.
All this could have been done by any one of the other 3 methods, too. This is just a little easier.
Of course, we can always write this back in terms of the ligand density n_L by substituting for the (ligand)

chemical potential, $e^{\frac{\mu}{k_BT}} = n_L\lambda_{th}^3, L e^{\frac{\varepsilon_s}{k_BT}}$: $\langle N\rangle = \dfrac{\left[2e^{-\beta(\varepsilon-\varepsilon_s)}n_L\lambda_{th}^3 + 2e^{-\beta(2(\varepsilon-\varepsilon_s)+J)}\left(n_L\lambda_{th}^3\right)^2\right]}{\left[1+2e^{-\beta(\varepsilon-\varepsilon_s)}n_L\lambda_{th}^3 + e^{-\beta(2(\varepsilon-\varepsilon_s)+J)}\left(n_L\lambda_{th}^3\right)^2\right]}$. **22.2**

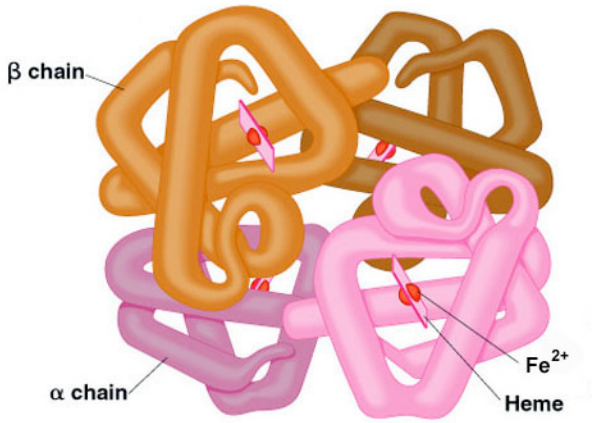
At first sight, this may look rather different from what we had before. To make one more connection, imagine $(\varepsilon-\varepsilon_s)$ is large and positive, so the single-occupancy terms vanish above and below, but $2(\varepsilon-\varepsilon_s)+J<0$, so the double occupancy state is preferred. Then, things simplify to

$\langle N\rangle = \dfrac{2n_L^2}{\left[\dfrac{e^{\beta(2(\varepsilon-\varepsilon_s)+J)}}{\left(\lambda_{th}^3\right)^2} + n_L^2\right]}$, which is just the Hill Equation of order 2. (Lect. 12.2)

B. Hemoglobin:

Hemoglobin is even more complicated than this. It is a symmetrical tetramer with 4 oxygen binding sites and exhibits a highly cooperative binding curve. As discussed in PKT (Ch. 7), some quite involved models of the type described here have been used to model the binding.

from: http://themedicalbiochemistrypage.org/hemoglobin-myoglobin.html



There are now four σ variables, each with two states 0 and 1 for a total of 16 microstates. The "energies" entering the ensemble are written:

$(E-\mu N)\left(\{\sigma_k\}_{k=1}^4\right) = \sum_{k=1}^4(\varepsilon-\mu)\sigma_k$ uncorrelated model, no cooperativity

$(E-\mu N)\left(\{\sigma_k\}_{k=1}^4\right) = \sum_{k=1}^4(\varepsilon-\mu)\sigma_k + J\sum_{\langle k\neq l\rangle}\sigma_k\sigma_l$ Pauling model (extension of above to 4 sites)

$(E-\mu N)\left(\{\sigma_k\}_{k=1}^4\right) = \sum_{k=1}^4(\varepsilon-\mu)\sigma_k + J\sum_{\langle k,l\rangle}\sigma_k\sigma_l + K\sum_{\langle k,l,m\rangle}\sigma_k\sigma_l\sigma_m + L\sum_{\langle k,l,m,n\rangle}\sigma_k\sigma_l\sigma_m\sigma_n$ Adair model

Note how the different binding sites are treated symmetrically.

The Grand partition function always involves a sum over the 16 microstates.

For the uncorrelated model $\Xi = \left(1+e^{-\beta(\varepsilon-\mu)}\right)^4$; however, for Pauling and Adair you just have to sum over the 16 microstates.

It is straightforward to calculate $\langle N\rangle$ or $P_n, n=0,1,2,3,4$.

Two-state models and 2^N-state models are common in biophysics.

They can be used to fit (often quite well) data for rather complex systems.

Relating the couplings, J, K, L, to, e.g., the cooperative distortions of the proteins is harder.

What are polymers?
Covalently linked linear chains of molecules A-B-C-D-...
When monomers, A, B, C, D,..., are all the same, they are called *homopolymers*; otherwise, *heteropolymers*. In heteropolymers, ordering can be orderly (e.g., A-B-A-B-…, a sequence of dimers) or disordered/random.
Major classes of biopolymers include:
DNA, RNA (monomers are the 4(5) nucleotides) "random."
proteins (monomers are the 20 amino acids) "random."
structural polymers (actin, spectrin, etc.) typically structural elements, ordered.
All these function biologically in aqueous solution.
The number N of linked monomers—the degree of polymerization—can range from small ("oligamers"), to order 300 for typical proteins, up to 10^6—10^9 for DNA.

A key feature of polymers is that they have many configurations/conformations ("microstates"). Details depend on the bonding. But, roughly speaking, we distinguish three classes:

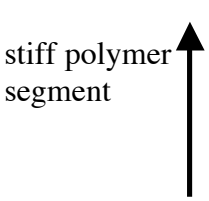
1. "Rod-like" polymers, which resist bending over relevant length-scales and are used as structural elements.
2. "Random-coils" which behave more or less like random walks (more below).
3. "Compact" structures, like enzymes, which are wrapped up/folded into specific functional shapes.

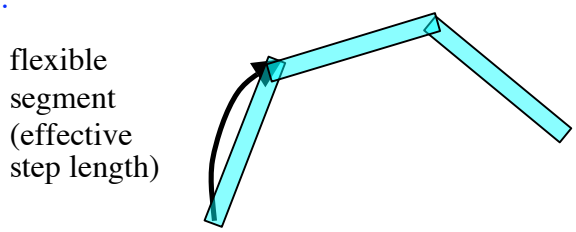
These classes are not fixed but transform or blend into one another as a function of length-scales and temperature T.

For a freely-jointed polymer, as in the random-walk models, all configurations {n} are equally likely (equal a priori probability); however, if these configurations have different energies E_n, then the probability of each configuration goes as $e^{-\frac{E_n}{k_B T}}$ (canonical ensemble).

The number of configurations of a long polymer (N units) is exponentially large: Take our cubic-lattice example: Each successive step can take z=2d directions or z=(2d-1) if you exclude immediate reversals. Thus, $W(N) = z^N = e^{N \ln z}$. This soon becomes a *very* large number.

The central theme of polymer configurational statistics is the trade-off between energy and entropy: At low-enough temperatures the polymer will select its lowest energy state (or one of its relatively few low-energy states). But, for a long polymer (high-enough N) or large-enough T, the very large numbers of random-coil configurations will normally win.
All polymers have a kind of local "stiffness" that prevents them from bending locally. This is due partly to bonding structure and partly to steric interference. In the random-walk model we have talked about, this is incorporated into the typical step-length a.
But, no polymers are completely stiff. If you build a long-enough polymer, it can still behave like a random coil at a large enough length scale.





Persistence length ξ_p: The contour length over which the polymer becomes effectively flexible.
Stiffness is controlled by temperature T, since it costs energy to bend polymer over a certain length scale