Why does the random coil work so well?
To be followed by discussions of DNA packing and protein folding.
Problem Set 4, officially due Nov. 8 but will be accepted Wednesday, Nov. 10.
Reading for Lectures 22—24: PKT Chapter 8

Polymeric microsates ("configurations") n weighted by canonical factor $e^{-\beta E_n}$.
What interactions contribute to E_n?
"Short-range" interactions along the axis of the chain, including covalent bond geometry and flexibility, steric interactions, and bond-angle restrictions.
"Long-range" interactions, i.e., interactions with other monomers that are "distant" in terms of labeling along the chain but get close in space because the chain loops back on itself, e.g., self-avoidance and specific interactions between monomeric units (charges, hydrogen bonding, salt bridges, etc.).

We will now look back at the "random-coil" polymer and see how it is modified by interactions:
Recall from Lect. 10:
For random +/-a steps in 1D:
$\langle X^2 \rangle = a^2 \langle M^2 \rangle = Na^2$ with full distribution of end-to-end length,

$$P_N(X) = \frac{1}{\sqrt{2\pi \langle X^2 \rangle}} e^{-\frac{X^2}{2\langle X^2 \rangle}}.$$

For random steps along cubic axes in 3D:
$\langle \vec{R}^2 \rangle = Na^2$ with full distribution,

$$P_N(\vec{R}) = \left(\frac{3}{2\pi \langle R^2 \rangle} \right)^{3/2} e^{-\frac{3R^2}{2\langle R^2 \rangle}}.$$

Q: How general are these results?
A: In the absence of long-range interactions, they are quite general, except that the step-length a may change and even become temperature dependent. Distribution remains Gaussian.

Short-Range interactions only: Random-Coil models
Consider a general model of N links (N+1 monomers) $\vec{a}_n = a_n \hat{t}_n$, with $\hat{t}_n \cdot \hat{t}_n = 1$, with variable length and direction.

The end-to-end displacement is $\vec{R} = \sum_{n=1}^{N} \vec{a}_n$.

$$\langle \vec{R} \cdot \vec{R} \rangle = \langle R^2 \rangle = \langle (\vec{a}_1 + \vec{a}_2 + \vec{a}_3 + \cdots) \cdot (\vec{a}_1 + \vec{a}_2 + \vec{a}_3 + \cdots) \rangle = \sum_{n=1}^{N} \langle a_n^2 \rangle + 2 \sum_{m<n}^{N} \langle \vec{a}_m \cdot \vec{a}_n \rangle,$$

where the average is over all configurations.
If the links are independent, then $\langle \vec{a}_m \cdot \vec{a}_n \rangle = \langle \vec{a}_m \rangle \cdot \langle \vec{a}_n \rangle$, so, as long as each one averages to zero, we
have $\langle R^2 \rangle = \sum_{n=1}^{N} \langle a_n^2 \rangle$. Note that this argument does not assume that the links $\vec{a}_n$ are along cubic axes.
Example 1: The freely-jointed chain:
Suppose the step-length is fixed at a but the links are freely jointed (and if we ignore self-aviodance/steric effects), then each link is distributed freely over a sphere, so $\langle \vec{a}_m \cdot \vec{a}_n \rangle = \langle \vec{a}_m \rangle \cdot \langle \vec{a}_n \rangle = 0$
and $\langle R^2 \rangle = Na^2$, exactly as for the random-walk on the cubic lattice.
I haven't shown you that the distribution of end-to-end length is as given above. But, it is.

I am now going to introduce you to a series of models for which $\langle R^2 \rangle = N \times (\text{microscopic length})^2$, just as for the freely jointed chain/random walk. In writing these results it is traditional to quote $\langle R^2 \rangle$ in the form $\boxed{\langle R^2 \rangle = \ell_K L}$, where L is the "contour length" of the polymer following the backbone and ℓ_K is called the "Kuhn length."

For the freely jointed chain: L=Na and ℓ_K=a.

The significance of the "Kuhn length" is that it is the segment length into which you have to break up the chain, so that (on this length scale) it looks like a freely jointed chain: Thus, for a polymer of contour length L, the number of "Kuhn segments" would be $N_K = \dfrac{L}{\ell_K}$. If these were treated as a freely jointed chain, we would get $\langle R^2 \rangle = \ell_K^2 N_K = \ell_K L$.

More examples:
As long as the direction of the steps is random, the step length does not have to be fixed:

2. The links might have independently some random length distribution $P(a)$ with $\langle a \rangle = \int_0^\infty da\, a P(a)$ and $\langle a^2 \rangle = \int_0^\infty da\, a^2 P(a)$.

Then, we calculate: $\langle R^2 \rangle = \sum_{n=1}^{N} \langle a_n^2 \rangle = N \langle a^2 \rangle = \dfrac{\langle a^2 \rangle}{\langle a \rangle}(\langle a \rangle N) = \ell_K L$ with $\ell_K = \dfrac{\langle a^2 \rangle}{\langle a \rangle}$.

3. The links might be alternating in length as in an A-B--A-B—A-B--… heteropolymer. Then,

$$\langle \vec{R}^2 \rangle = \frac{N}{2}(a^2 + b^2) = \left(\frac{N}{2}(a+b)\right) \cdot \frac{(a^2 + b^2)}{(a+b)} = \ell_K L \text{ with } \ell_K = \frac{(a^2 + b^2)}{(a+b)}.$$

But, because of short-range interactions, the successive links are *NOT* generally independent of one another for real polymers.
The reason for this is that—contrary to the assumptions of the random-walk-type models--the direction of the nth link is limited by/dependent on the configuration of the (n-1)th link. It's just chemical bonding and steric hinderance. So, it is not yet entirely clear, why the random-coil results should carry over to real polymers.
Thus, it is still far from clear why a random-walk result should continue to hold.
To get a feel for why this might still carry over, I am going to consider three examples, one a lattice model (not very realistic chemically) and the other a (semi-realistic) model of alkane chains, and finally an important continuum model.

4. Another cubic-lattice model: Suppose that at each step we forbid the immediate reversal, so there are now z=5 possibilities instead of z=6 (3D) that we would have in a random lattice walk.
This reflects steric hinderance and/or local (but not global) self-avoidance (i.e., hard-core repulsion).

In this situation, it is easy to see that $\langle \vec{a}_n \cdot \vec{a}_{n+1} \rangle = a^2\left(\frac{1}{5}\right)$, $\langle \vec{a}_n \cdot \vec{a}_{n+2} \rangle = a^2\left(\frac{1}{25}\right)$, $\langle \vec{a}_n \cdot \vec{a}_{n+3} \rangle = a^2\left(\frac{1}{125}\right)$,

etc.
Here's the rigorous argument:
The straight configuration always contributes and gives the $(1/5)^n$ contribution. Transverse bonds are irrelevant, since they always contribute "0" to the dot product. Thus, the average of the dot product is just $(1/5)^n$ times the net excess of forward bonds at the nth step. But, except for the straight configuration, every configuration includes a least one transverse bond. Any trajectory with a forward nth bond can be paired with a trajectory with a backwards one simply by reflecting in the transverse

plane containing that bond. And, similarly for any trajectory with a backwards nth bond. Thus, the only uncompensated nth bond is from the straight configuration (and it goes forward).

Thus,

$$\langle R^2 \rangle = \langle \vec{R} \cdot \vec{R} \rangle = \sum_{n=1}^{N} \langle a_n^2 \rangle + 2 \sum_{n=1}^{N-1} \langle \vec{a}_n \cdot \vec{a}_{n+1} \rangle + 2 \sum_{n=1}^{N-2} \langle \vec{a}_n \cdot \vec{a}_{n+2} \rangle + \ldots + 2 \langle \vec{a}_1 \cdot \vec{a}_N \rangle$$

$$= a^2 \left[N + 2(N-1)\left(\frac{1}{5}\right) + 2(N-2)\left(\frac{1}{5}\right)^2 + \ldots \right] = a^2 \left[N\left(1 + \frac{2}{5}\left(1 + \frac{1}{5} + \frac{1}{5^2} + \cdots\right)\right) + O(1) \right] = \frac{3}{2} N a^2 + a^2 O(1),$$

which for large N is just the same as a random walk with $\langle R^2 \rangle = \frac{3}{2} a \cdot (Na) = \frac{3}{2} aL$, so $\ell_K = \frac{3}{2} a$.

Note for future reference the geometric (i.e., exponential) decrease of the directional correlations!
Again, I have not shown you that the whole distribution is Gaussian. But, it is. Upshot: at long distance you can't distinguish this from a random walk with the step-length b_K.

5. Alkane model

This model was introduced in tutorial yesterday. Chemistry of bonding fixes (quite tightly) the angle between $\vec{a}_n$ and $\vec{a}_{n+1}$; however, there is free rotation about the direction of $\vec{a}_n$, so the "head" of $\vec{a}_{n+1}$ lies on the red circle. Clearly, $\langle \vec{a}_n \cdot \vec{a}_{n+1} \rangle = \cos\theta$.

Q: What about $\langle \vec{a}_n \cdot \vec{a}_{n+2} \rangle$?

A: $\langle \vec{a}_n \cdot \vec{a}_{n+2} \rangle = (\cos\theta)^2$.

The argument is made by pairing the opposite points on the red circle. They average to the centre. The projection back to the direction of $\vec{a}_n$ gives $(\cos\theta)^2$.

The same logic can be applied iteratively.

Upshot: $\langle \vec{a}_n \cdot \vec{a}_{n+k} \rangle = (\cos\theta)^k$.

So,

$$\langle R^2 \rangle = N a^2 + 2a^2 \left[(N-1)\langle \vec{a}_n \cdot \vec{a}_{n+1} \rangle + (N-2)\langle \vec{a}_n \cdot \vec{a}_{n+2} \rangle + (N-3)\langle \vec{a}_n \cdot \vec{a}_{n+3} \rangle + \cdots \right]$$

$$= a^2 \left[N + 2(N-1)(\cos\theta) + 2(N-2)(\cos\theta)^2 + \ldots \right] = a^2 \left[N\left(1 + 2\cos\theta\left(1 + \cos\theta + \cos^2\theta + \cdots\right)\right) + O(1) \right]$$

$$= N a^2 \left[1 + \frac{2\cos\theta}{1 - \cos\theta} \right] + a^2 O(1) = N a^2 \left[\frac{1 + \cos\theta}{1 - \cos\theta} \right] + \text{small}.$$

Thus, finally, $\langle R^2 \rangle = b_\ell L$ with $\ell_K = a\left(\frac{1 + \cos\theta}{1 - \cos\theta}\right)$.

Note: the Kuhn length diverges as $\theta \to 0$, illustrating increased "persistence" of the forward direction.