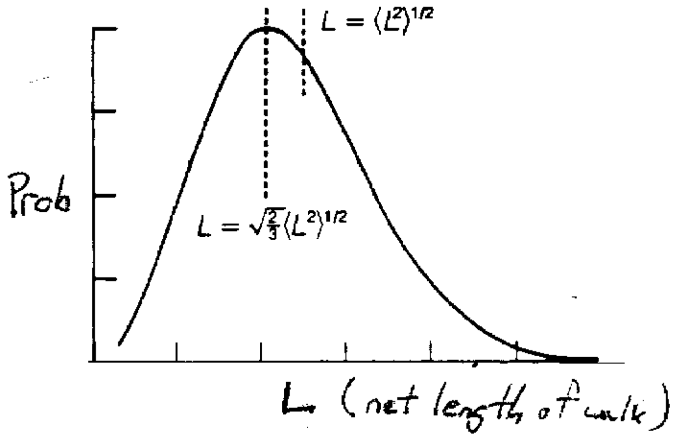
Neil Clarke
Topics in Biophysical Chemistry

Random walks: polymers and diffusion

Double-stranded DNA is a long polymer that (unlike proteins) does not adopt a well-ordered three-dimensional shape

"Random walks" are important in understanding the polymer properties of DNA; they also provide the basis for understanding diffusion.

Imagine starting at some origin and taking a large number of steps (N) of length l_e. Suppose also that the direction of each new step is completely random. The end result of the walk is that you have probably moved away from the origin, but you almost certainly are not at a distance $N*l_e$ because that would require each step to be in the same direction as the previous step. The probability for the end of the random walk being at various distances from the origin can be expressed in a fairly simple way. The calculated probability distribution is shown below for a three-dimensional random walk (The units are arbitrary).



Although it may seem like a peculiar measure, a convenient shorthand for this distribution is the root-mean-squared (rms) distance from the origin $\langle L^2\rangle^{1/2}$. The mean square distance $\langle L^2\rangle$ is obtained by squaring all of the possible distances (L) from the origin, weighting the contribution of each distance on the basis of its probability, and then taking the mean of the squared distances. The square root of this number is the rms distance.

The reason this number is convenient and revealing is that it can be shown that:

$$\text{rms distance} = \left\langle L^2 \right\rangle^{\frac{1}{2}} = l_e N^{\frac{1}{2}}.$$

Thus, there is a simple relationship between the "average" distance traveled and the length of each step, which makes a lot of sense. What may not have been so obvious to start with is that the "average" distance travelled goes up only with the square root of the number of steps. This has important implications for the efficiency of diffusion as a transport process.

Note that there is also a simple relationship between the rms distance traveled and the most probable distance traveled.

$$\text{most probable distance} = \left(\frac{2}{3}\right)^{\frac{1}{2}} l_e N^{\frac{1}{2}}$$

Polymers (e.g, DNA) as random walks

The path taken by very long polymers like DNA can (under appropriate conditions) approximate random walks. The structure of such a random walk polymer is called a random coil. However, its clear that we can't consider each basepair of DNA to be a step in a random walk because a random walk requires that each step be in a random direction. In DNA, each basepair can differ from the direction of the previous basepair only slightly. The trick here is to redefine what we mean by a step. If we define the step as that length of the polymer which is sufficiently long that the subsequent stretch of polymer of that length is randomly oriented, then we can consider a polymer to be like a random coil.

The length (l_e) of this effectively random step is called the "statistical segment length" (l_e). We have already used l_e to denote the length of a random step so we need not change that. The number of statistical segment lengths in the polymer (that is, the number of effectively random steps) is denoted N_e. For an imaginary polymer in which each monomer is completely independent of the others, N_e equals the number of monomers. For a real polymer $N_e=(C/l_e)$where C is the total length of the polymer (or "contour" length). Nothing has changed except that we've defined the length in such a way that the random walk derivation is appropriate.

$$\text{rms distance} = l_e N_e^{1/2}$$

The statistical segment length is a measure of the stiffness of a polymer. For DNA, it is about 350 basepairs.

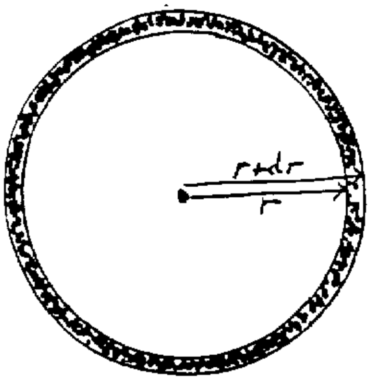
Even using the notion of a statistical segment length, polymers display perfect random coil characteristics only under special conditions. The most important difference between a polymer and an ideal random walk is the excluded volume effect. Unlike the path taken in a random walk by a single point, the path taken through space by a long polymer creates areas of space which cannot be simulataneously occupied by the polymer further down the chain. The chain is sterically prevented from getting as close to the origin as the random walk model would otherwise predict. Thus, the root-mean-squared distance from one end of the polymer (origin) to the other increases with the number of steps faster than would be predicted by the simple random walk model. The corrected form of the equation for rms distance that approximates the excluded volume effect is

$$\text{rms distance} = l_e \left(N_e^{(1+\varepsilon)} \right)^{1/2}$$

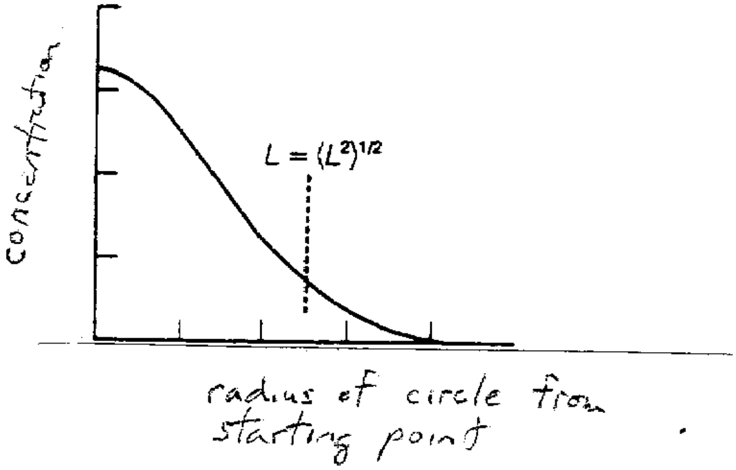
e is a variable that depends on the properties of the polymer and the solvent.

Diffusion as a random walk

Molecules in solution move in a way that can be modeled as a random walk. Imagine a point-source for some dye in solution, such that at time t=0 all of the dye is concentrated in a single point. What is the concentration of dye at different distances from the origin after some time t? As the dye diffuses from the origin, each molecule has a probability of being at some distance from the origin as described by the distribution in the figure above that we've already discussed. Since we are talking about a very large number of molecules, each distributed according to this probability function, we can think about the distribution as representing the number of molecules at each distance. Is this also the distribution of the concentration of molecules as a function of distance, then? Not quite. We have to remember that the volume of the shell that lies between two spheres, one of radius r and the second of radius r+dr, becomes larger the greater the radius.



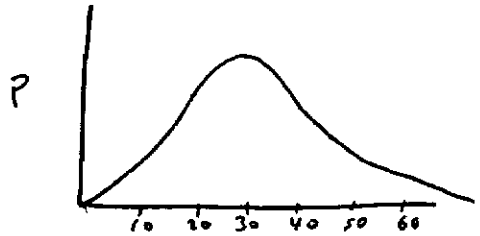
Dividing the relative number of molecules at some distance between r and r+dr by the volume of the shell gives the relative concentration. The figure below shows the concentration distribution for the same random walk illustrated in the previous figure. Note that at the distance described by the rms distance value, the concentration is about 1/4 that at the origin, and that it falls off quickly from there.



Problems for random walks and diffusion

(1) Suppose a small molecule has a diffusion coefficient of 10^{-6} cm^2/sec for diffusion through membranes. How long would you expect it to take for the molecule to get from outside the cell to inside if the membrane is 50Å thick? (An angstom (Å) is 10^{-10} meters or 10^{-8} cm). Now suppose the diffusion constant for a protein in the cytoplasm of a cell is also 10^{-6} cm^2/sec. How long would it take to diffuse from one end of the cell to the other if the cell is 50 microns long (a micron (um) is 10^{-6} meters)?. Finally, some nerve cells can be 10000 times longer (or more) than the cell considered above. How long would it take the protein to diffuse the length of such a nerve cell? ?

(2) The figure below shows the probability distribution for a random walk of 1000 steps of length 1 meter. Sketch the distribution for a walk of 4000 steps.



(3) Show that, for polymers of the same length but of differing stiffness, that the root-mean-square end-to-end distance varies with the square root of the statistical segment length.

(4) The E. coli chromosome is $4x10^9$ base pairs in length. For the sake of this problem we'll assume the chromosome is linear (its actually circular). Suppose further that the the DNA adopts a random coil conformation. The statistical segment length of DNA is about 350 basepairs. The length of DNA is about 3.4Å per basepair. What is the root-mean-square end-to-end distance for bacterial DNA? E. coli is about 2 microns in the longest dimension. What are the implications of your calculation?

(5) Suppose there is a high concentration of some molecule on the outside of the cell and a low concentration inside. A pore in the cell membrane is opened that allows free diffusion of this molecule into and out of the cell. How, if at all, would you expect the following to affect the rate at which the concentration in the cell increases: (a) higher temperature, (b) higher viscosity, (c) conformational change from sphere-like to rod-like, (d) time