10.34 – Fall 2006
Homework #10
Due Date: Friday, December 1st, 2006 – 9 AM

*** Note: Please read the entire problem set before starting, there is important information throughout, even at the very end. For this problem, you do NOT need to have the Matlab code generate all of the results for part A-E by running it once. However, it should be able to take a temperature and number of points and generate all of the desired plots for that set of inputs.*

The most popular way to experimentally test a proposed geometrical structure for a large molecule (such as a protein) is by X-ray crystallography. However, some proteins are hard to crystallize; for these proteins, proposed geometrical structures can be tested using nuclear magnetic resonance (NMR). NMR measures the through-space magnetic coupling between two atoms which are not directly bonded to each other; this magnetic coupling is proportional to $<1/R^6>$, where R is the distance between the two atoms. The symbol < > means the Boltzmann average over all the possible molecular geometries; in the classical limit and neglecting some minor complications due to the integral over the kinetic energy we can write:

$$\left\langle \frac{1}{R^6} \right\rangle = \frac{1}{Q} \iiint \frac{1}{\left[R\left(x_1, x_2, ..., x_N\right) \right]^6} \cdot \exp\left[\frac{-V\left(x_1, x_2, ..., x_N\right)}{k_B T} \right] d^3x_1 \, d^3x_2 ... d^3x_N$$

where ***Q*** is the classical partition function:

$$Q = \iiint \exp\left[\frac{-V\left(x_1, x_2, ..., x_N\right)}{k_B T} \right] d^3x_1 \, d^3x_2 ... d^3x_N$$

This high-dimensional integral can be computed for a proposed structure using Monte Carlo techniques. Of course for a molecule with a large number of atoms this can be quite challenging. Here we instead ask you to compute this integral for a small molecule.

Note that it is very easy to figure out the equilibrium geometry from this analytical expression for V (note V = 0 at the equilibrium geometry). We suggest you use Metropolis's method, and start your Monte Carlo steps from the equilibrium geometry.

Write a set of Matlab functions which use Monte Carlo integration to compute $<1/R_{HH}^6>$ at a given Temperature, where R_{HH} is the distance between the two H atoms in HOOH.

A. Determine the equilibrium structure at 0 K of HOOH by minimizing the potential energy of HOOH. Plot the structure of the molecule in 3-D using the `plot3` command in Matlab. State the 0 K equilibrium values for $<1/R_{HH}^6>$ and $<R_{HH}>$ in Angstroms. Do you expect the value of $<R_{HH}>$ to be different for T = 300 K, why?

B. Use your code to solve for the value of $\langle 1/R_{HH}^6 \rangle$ at 300 K. Report the value obtained for $\langle 1/R_{HH}^6 \rangle$, the value of $\langle R_{HH} \rangle$, and the number of Monte Carlo steps attempted and accepted.

C. Plot the 3-D location of all of the MC points obtained in the above simulation using the `plot3` command again. It may also be instructive to plot the equilibrium structure underneath the MC points. This can be done with something like (obviously, this syntax will need to be modified to your problem)**:**

```
plot3(equil,'-','linewidth',4.0);
hold on;
plot3(MC_points,'.');
hold off;
```

D. Repeat this to generate plots for temperatures of 600 K, 1000 K, and 5000 K. Generate these plots with a minimum of 10000 MC steps. Generate a histogram of the $\langle R_{HH} \rangle$ values for each temperature, using the same x-axis scale for all figures. Also create a histogram (50 bins) showing the distribution of potential energies that the molecule adopts for each temperature (you don't need to use the same x-axis scale). Find the bin with the largest frequency and compare this energy value with the value of $k_B T$.

E. Generate a plot for each T showing the evolution of the $\langle 1/R_{HH}^6 \rangle$ as the number of MC points increases (ideally this curve will converge to the actual value of as N $\rightarrow \infty$).

F. For your answer in part B, give your best guess at the uncertainty in your predicted value of $\langle 1/R_{HH}^6 \rangle$, and explain how you derived it.

Assume this is the expression for the potential energy of HOOH:

$$V = V_{OH}\left(R_{O_1H_1}\right) + V_{OH}\left(R_{O_2H_2}\right) + \frac{1}{2}k_{OO}\cdot\left(R_{OO} - L_0\right)^2 + \frac{1}{2}k_\theta \cdot\left[\left(\theta_{HOO} - \theta_0\right)^2 + \left(\theta_{OOH} - \theta_0\right)^2\right] + V_\phi$$

where:

$$V_{OH}\left(r\right) = D_{OH}\cdot\left(1 - \exp\left[-\alpha\cdot\left(r - L_H\right)\right]\right)^2$$

$$V_\phi = 80\,\tfrac{kJ}{mole}\cdot\sin\left(\tfrac{1}{2}\theta_{HOO}\right)\cdot\sin\left(\tfrac{1}{2}\theta_{OOH}\right)\cdot\left[\cos\left(\phi\right) - \cos\left(\phi_0\right)\right]$$

$D_{OH} = 360\ \frac{kJ}{mole}$ $\quad L_H = 1.05\ \text{Å}$ $\quad \alpha = 1.5\ \text{Å}^{-1}$ $\quad k_{OO} = 300\ \frac{J}{m^2}$

$L_0 = 1.6\ \text{Å}$ $\quad k_\theta = 10^{-6}\ \frac{pJ}{radian^2}$ $\quad \theta_0 = 1.8\ radians$ $\quad \phi_0 = 1.7\ radians$

The R's are Cartesian distances between the atoms. θ_{HOO} is the angle defined by H_1-O_1-O_2 and θ_{OOH} is the angle defined by O_1-O_2-H_2 (you can compute these using law of cosines). An expression for the dihedral angle ϕ is given below.

Hint: *Molecule fixed axes*
Molecular potentials V do not depend on the position of the molecule in space, nor on its orientation, but only on the relative position of the atoms. Hence one can usually cut 6 degrees of freedom (corresponding to the position of the molecule and its angular orientation (Euler angles)) out of molecular problems. In this particular problem, we suggest using molecule-fixed axes where the position of atom O_1 sets the origin, atom O_2 lies on the x axis, and atom H_1 lies in the x-y plane. Then one can remove these 6 degrees of freedom from the problem: (x_{O1},y_{O1},z_{O1},y_{O2},z_{O2},z_{H1}). (You can set them all equal to zero). When you remove the orientational degrees of freedom you pick up some Jacobian volume elements; including these the new expression for the integral (again approximating away some minor terms related to rotational kinetic energy) is:

$$\left\langle \frac{1}{R_{HH}^6} \right\rangle = \frac{1}{Q_{red}} \iiint \left(\frac{x_{O_2}^2 \cdot \left| y_{H_1} \right|}{\left[R_{HH} \right]^6} \cdot \exp\left[\frac{-V}{k_B T} \right] \right) dx_{O2}\, dx_{H1}\, dy_{H1}\, dx_{H2}\, dy_{H2}\, dz_{H2}$$

$$Q_{red} = \iiint \left(x_{O_2}^2 \cdot \left| y_{H_1} \right| \cdot \exp\left[\frac{-V}{k_B T} \right] \right) dx_{O2}\, dx_{H1}\, dy_{H1}\, dx_{H2}\, dy_{H2}\, dz_{H2}$$

In this molecule-fixed axis system, the expression for the dihedral angle is:

$$\cos(\phi) = \frac{y_{H_1} y_{H_2}}{\left| y_{H_1} \right| \cdot \sqrt{y_{H_2}^2 + z_{H_2}^2}}$$

One other point of interest how to move around this multi-dimensional space, and the step sizes to take. When you are moving atoms around in 3-D Cartesian space (not using internal coordinates), you are not very well constrained along the normal modes of the atom, i.e. you have a small chance of making a large jump from one low energy position to another. For example, consider a water molecule with the O centered at (0,0), one H1 at (0,y1), and H2 at (x2,y2). A reasonable structure would have H1 at (0,1) and H2 at (0.95,-0.3). However, an equally plausible structure would have H2 at (-0.95,-0.3). So if one was doing a MC simulation, with a random step of < 2 in x and y, then you could imagine that you may "hop" from one node to the other, since they are energetically equivalent. However, this is rare because all 3 variables must step the appropriate amounts: y1 step must be small, y2 step must be small, and x2 must be about -1.9. The odds of all of these conditions happening during the same step are low, so many MC steps may be needed to "hop" between nodes. You will also have to reject a large number of steps because many steps would lead to very large energy values.

This leaves two options for Cartesian simulations:

1. Take small steps which will be accepted more often, and will give reliable information about the areas of similar energy that are not separated by significant barriers. This will not allow you to probe multiple low energy wells separated by large barriers.

2. Take large steps that will be rejected most of the time (thus not give you very good statistics), but will allow for the possibility of finding other low-energy features that are far away or separated by large barriers. This can be useful when trying to minimize a function with many local minima.

A better method for molecules would be to take steps in internal coordinates (bond lengths, bond angles, and dihedral angles), so that you can better probe the feasible positions of the atoms because realistic step constraints can be more easily defined than with Cartesian coordinates. **It is NOT necessary to solve this problem using internal coordinates. We suggest taking small random steps with a maximum displacement in each direction of 0.10 Å. If you use another method to probe phase space, please state it explicitly.**