Final Project.

Due Friday, March 11, 2016

Questions:

Q.1. Choose **one** paper from Papers #1 to #7 from the *Readings* section in the course website and read it. Discuss the main scientific question addressed in the paper, the approach, and the conclusions.

Problems:

P.1. θ-temperature of polymers. Here you will estimate the θ-temperature of our simple polymer model by determining the crossover of $\langle R_g^2 \rangle / N$ vs $\langle T \rangle$ for two different lengths of the polymer. For both, make sure you compile in **three** dimensions (file 'defs.h' uncomment line "#define NDIM 3") and build the program by typing: % make 1.

A. N=30
   1. Before running, copy the initial collapsed-coil configuration to the 'inputs' directory by typing (**important**: you may have to create the 'inputs' directory AND note that the destination filename is different from the source filename):
      ```
      % cp positions_N30.xyz  inputs/positions.xyz
      ```
   2. Run the code by typing:
      ```
      % ./run_me input_N30.txt
      ```
      editing the input file and using for the input **T** values from **4.0** to **12.0** in step increments of **1.0**. Wait until conclusion of the run and calculate the average temperature $\langle T \rangle$ (third column of 'thermo.dat') and the mean square radius of gyration $\langle R_g^2 \rangle$ (from file 'rg2.dat') for each case. Plot $\langle R_g^2 \rangle / N$ vs $\langle T \rangle$.

B. N=40
   1. Again, before running, copy the other initial collapsed-coil configuration to the 'inputs' directory by typing:
      ```
      % cp positions_N40.xyz  inputs/positions.xyz
      ```
   2. Run the code by typing:
      ```
      % ./run_me input_N40.txt
      ```
      editing the input file and using for the input **T** values from **4.0** to **8.0** in step increments of **0.5**. As above, wait until conclusion of the run and calculate the average temperature $\langle T \rangle$ and the mean square radius of gyration $\langle R_g^2 \rangle$ for each case. Plot $\langle R_g^2 \rangle / N$ vs $\langle T \rangle$ in the *same graph* as above.

C. From your plot of $\langle R_g^2 \rangle / N$ vs $\langle T \rangle$ for N=30 and N=40 estimate the θ-temperature for this system. Discuss the meaning of the θ-temperature and explain why is it valid to estimate it by examining the crossover of both graphs.

P.2. All-atom protein simulation. In this exercise, you will gain a taste of a "real-world" application of molecular dynamics to simulate a protein in water using the GROMACS simulation package. Although this exercise looks long, it is rather straightforward and simple if directions are followed one after the other. The simulation uses a real PDB file, the OPLS force field, and explicit water molecules using the TIP3P force field. Follow the steps below to complete this exercise and generate your response that will consist of answers and images.

1. Download a PDB file for the lyzozyme (7LYZ) protein from www.rcsb.org/pdb. You can find it by typing the 7LYZ code in the *search* box. Choose the "text" version of the file from the

rhs download choices.

2. Download also the files *ions.itp* and *md.mdp* from the course website and make sure everything is in the same directory.
3. Visualize everything with 'vmd' and include a screen (or window) capture in your work

   % vmd 7LYZ.pdb

4. Edit this PDB file and delete all of the header lines at the beginning of the file up to the lines starting with "ATOM". Save as *7LYZ-clean.pdb* .
5. Convert this PDB file to a GROMACS conformation file *conf.gro* using 'pdb2gmx' by typing

   % pdb2gmx -f 7LYZ-clean.pdb -water tip3p

   Choose the OPLS-AA force field. Note that you are already choosing the TIP3P water model. Also note that upon completion of this command, you will have three extra files: *conf.gro, topol.top,* and *posre.itp.* From the output, report on the number of dihedrals, impropers, angles, pairs, and bonds. Also report on the total mass and the total charge.
6. Visualize the molecule again with

   % vmd conf.gro

   Type in the 'vmd' console window to see the periodic box

   > draw pbcbox

   Capture the molecule and box (screenshot will do) in an image to be included in your work. Discuss the contents and outlook of the image.
7. Type

   % editconf -f conf.gro -c -o conf-b.gro

   Repeat step 6 using now *conf-b.gro*, including a new image capture and a short discussion of its contents and outlook.
8. Make space for the water in the box of about 10Å between the protein and the edge of the box, by typing

   % editconf -f conf.gro -c -d 1.0 -o conf-b.gro -bt cubic

9. Solvate the protein by typing

   % genbox -cp conf-b.gro -cs -p -o conf-solvated.gro

10. Visualize and screen capture the solvated protein now in file *conf-solvated.gro.*
11. Pre-process files for the next step by typing

    % grompp -f md.mdp -c conf-solvated.gro

    Examine the output from this command and copy/paste in your work the part of the output that indicated the global charge of the system.
12. Neutralize your system by typing and selecting the SOL group

    % genion -s -o conf-neutral.gro -p -nn 8 -nname Cl

    Explain what did this command do, how did it do it, and what atom it use to do it. You may examine the output to complete your answer.

13. For the production run, make a sub-directory and copy *topol.top, conf-neutral.gro, ions.itp* and *md.mdp* to that directory. Change to that sub-directory and prepare for a production run by typing:

    % grompp -f md.mdp -c conf-neutral.gro -r conf-neutral.gro -p topol.top -o pep_md.tpr

14. Run the production run by typing

    % mdrun -v -s pep_md.tpr -o pep_md.trr -c conf-neutral.gro -g md.log -e md.edr

    Copy/paste your computer's "performance" found at the end of the production run output.

15. Visualize and play the trajectory by typing

    % vmd traj.xtc conf-neutral.gro

    Answer the following: Describe the motion. Does the protein unfold or otherwise change its shape? Is the trajectory sufficient for calculating thermodynamics properties? Explain in very simple terms what should be done or what should be modified in the run to improve on this simulation.