Hackathon Problem: Lysozyme and Water Variations

Reminder, run these calculations in your /gscratch!

In-class, we did a MD simulation with GROMACS of a lysozyme protein and water. For the hackathon, you will be running a similar system. Instead of hen egg-white lysozyme, you will now be running echidna milk lysozyme (1JUG). I have provided the cleaned .pdb file of the lysozyme for you with the crystal water removed (1JUG_clean.pdb).

You will be running identical calculations from the in-class lecture, so you can use the same input files and similar commands. The difference for this problem, though, is that you are asked to change parameters related to the molecular dynamics lecture. This includes the force fields, the timestep for the simulations, and the type of integrator. After running each of these calculations, plot the specified data and analyze for differences. **Your deliverable for this project is a final document that includes the plots, your analysis and conclusions, and a brief literature review of each of the force fields below. Find the original publications, cite them, and discuss their advantages and disadvantages.**

You will need to utilize the GROMACS documentation. While this is a variation on the example problem done in-class, you will have to go beyond it to understand how to change your input files.

Each of these calculations are pretty quick. However, there are many calculations to do and a lot of data to analyze, so try to keep yourself organized so that you don't miss any data. I recommend making a directory for each simulation so that you don't overwrite any data.

1. Force Fields. Test each of the force field combinations below and compare the results.
    - CHARMM27 (protein) and SPC/E (water)
    - CHARMM27 (protein) and TIP4P (water)
    - OPLS-AA (protein) and TIP4P (water)
    - OPLS-AA (protein) and SPC/E (water)
    - AMBERGS (protein) and TIP4P (water)
    - AMBERGS (protein) and SPC/E (water)

Make plots to analyze the **potential energy from the energy minimization, the temperature from the NVT equilibration, the density from the NPT equilibration, the radius of gyration of the protein from the production run, and the mean squared displacement of water from the production run.** Compare the results for each force field. Discuss the effect that changing the force field had on each property.

2. Timesteps. Select at least one of the force field combinations you have above and vary the timestep. Make this change in npt.mdp, nvt.mdp, and md_prod.mdp. The timestep in minim.mdp can remain 0.01 fs.
    - 0.002 ps
    - 0.001ps
    - 0.0005 ps

- 0.0002 ps

Keep the length of your simulation the same for each of these tests. This means that you'll have to change the number of steps that you take in order to keep your total simulation length the same (100 ps for NVT and NPT, 1000 ps for the production run).

Make plots to analyze the **temperature from the NVT equilibration, the density from the NPT equilibration, and the radius of gyration of the protein from the production run.** Compare the results. Discuss how the change in timestep effects the oscillations over the simulation length.

3. Integrators. Again, select at least one of the force field combinations and your best performing timestep to test a variation of integrators with. When I say "best timestep", consider the accuracy of the results *and* the computational cost. Again, change the integrator only for nvt.mdp, npt.mdp, and md_prod.mdp.
    - md-vv (velocity Verlet)
    - md-vv-avek (velocity Verlet with half step kinetic energy averages)
        i. For this run, in addition to changing your integrator to md-vv-avek, include the following two lines in your .mdp files:
        nsttcouple = 1
        nstpcouple = 1
    - md (leap-frog)

Make plots of the **temperature and kinetic energy from the NVT equilibration, the density from the NPT equilibration, and the temperature, kinetic energy, and radius of gyration for the protein from the production run**. Read the description of the differences between the integrators in the GROMACS documentation. Discuss the effect the different integrators have on your results, particularly the temperature and the kinetic energy.

Assemble your plots and discussion in a document of your preferred format, i.e., Word, LaTeX, etc.

Use the GROMACS documentation. It has very useful information on the force fields, the integrators, and what keywords can be used with each command. If you feel up to it, play around with additional keywords and see if there are other properties/plots that you'd like to make.

HINT: To calculate the diffusion coefficient of water, search the documentation for "mean square displacement" to find the GROMACS command.
https://manual.gromacs.org/current/index.html