**LAB 5: DYNAMICS SIMULATION WALKTHROUGH**

This lab will serve as basic walkthrough of a dynamics simulation. The objectives for this lab are as follows:

1. Build a system, including building a water box, neutralization, and configuration of the ligand.
2. Heat a system in preparation for a dynamics simulation.
3. Perform a short dynamics simulation
4. Analyze the simulations with simple analysis tools in VMD.

This lab will draw upon what you have learned in previous labs regarding energy minimization, atom selections, and constraints.

*Luciferase*

The molecular dynamics simulation you will be performing is an analysis of the effect of an electric field on the protein configuration of the Luciferase enzyme. Luciferase is an enzyme that produces bioluminescence in fireflies through converting luciferin and ATP into oxyluciferin and AMP, releasing a photon of light as oxyluciferin returns to ground state.

The luciferase that we will be analyzing is found on PDB.org under the identifier 2D1S. Its crystal structure contains a luciferyl-AMP intermediate within its active site.

Dr. Brian Mazzeo at BYU has conducted several experiments where he analyzed the effect of an electric field on luciferase fluorescence. His experiments are currently inconclusive, but dynamics simulations could certainly offer further insight. The dynamics simulations you perform in this lab should demonstrate how the protein and ligand react to a strong electric field.

**I. Building the System**

A crucial step in performing dynamics simulations is to properly set up your system. This includes *solvation* (surrounding the protein in generated water molecules), *neutralization* (adding ions to give the system a net charge of 0), and performing energy minimizations. The complexity of this process depends on the complexity of the dynamics simulation you wish to perform.

*Prepare the Protein and Ligand*

The combine\_protein\_lig.inp script inside the directory 1\_put\_ligand\_in\_protein is a simple input script that combines the protein and the ligand in one coordinate file and performs some minimizations on the system.

Open the input script and answer the following question:

*Question 1: Describe this script. What are its inputs? What is being manipulated and why? What is the output?*

Make sure the submission script’s input and output files are designated correctly and execute the submit script. The output is a CRD and PSF that contain the protein with the ligand in the active site. Feel free to view the output in VMD.

*Solvation*

This is a very simplified solvation script. The 100x100 angstrom water box has already been generated by [www.charmm-gui.org](http://www.charmm-gui.org). You can generate your own water box in CHARMM, but using the CHARMM-GUI can speed things up a bit. Make sure the solvation script refers to the psf and crd you generated previously. Open solvator.inp and take a look at the script. You’ll notice a deletion command that is written as follows:

delete atom sort -

select .byres. (resn TIP3 .AND. type oh2 .and. -

((.not. (resn TIP3 .OR. hydrogen)) .around. 2.8)) end

*Question 2: What is this deletion doing? Why is this important?*

Execute the submission script to solvate the molecule.

*Question 3: How many water molecules did you generate?*

*Neutralization*

Dr. Mazzeo performed his experiments in a solution with a .2M concentration of ammonium sulfate. For the purpose of our experiment, we’ll be using potassium chloride. Using www.charmm-gui.og, it was determined that 108 KCl molecules will result in a .2M concentration. This means 108 K+ and 108 Cl-. However, the protein carries a net charge of +1, so a total of 109 Cl- ions is necessary to neutralize the system.

Run the neutralize.inp script. This script streams 3 additional subscripts: watervars.str, ioncoor.str, and watercolor.str. These three scripts identify 217 water molecules to delete through randomly generated numbers, making room for the ions to be added. The water ions are then deleted, and the ions replace the water ions former positions, with 2 exceptions: 2 chloride ions are added within the protein to take the place of chloride ions that were found in the original pdb script.

**II. Minimize and Heat your System**

The system must be heated to room temperature prior to simulation. Open the folder 4\_minimize\_heat in WinSCP and open the submit file batch.submit. This file submits your script to the Fulton supercomputer. The line **#SBATCH -N1 -n16 --mem-per-cpu=2G -t10:00:00 -C 'm7'** designates several important settings for your submission. **–N1** refers to the number of nodes that you’ll run your script on. Generally speaking, the more nodes you request the faster your script will run. **–n16** refers to the number of processors requested per node, with a maximum of 16. **–mem-per-cpu** will usually be held at a constant **2G**. **–t10:00:00** designates the amount of time you anticipate your script to run. **-C 'm7'** designates that we’ll be running our scripts on the m7 supercomputer. Change the number of requested nodes to 16, and the requested time to 8 hours.

Open the input file minimize\_heat.inp. Scroll down to the “! CRYSTAL” portion of the script. This command generates an infinite crystal of our system. This will allow the system to remain confined to a cubic structure, and the outer edges of the cube will factor in the Van der Waals of the molecules that are within 14 angstroms of the outer edge.

*Question 4: What do you think “crystal defi cubi 100.0 100.0 100.0 90.0 90.0 90.0” is designating?*

View the “! CONSTRAINTS” portion of the script. Three constraints will be applied in this script. We’ll initially apply a harmonic constraint on the protein, to ease the impact of the electric field. PULL EFIELD 1E9 XDIR 1.0 SELE ALL END creates our 1E9-volt electric field constraint in the X-direction, which is applied to the whole system. Finally, we must apply a center-of-mass constraint to the protein, which will prevent the charged luciferase enzyme from migrating through the water box as a result of the electric field.

View the “! MINIMIZATION” portion of the script.

*Question 5: How many minimization steps and what kind of minimizations are being performed?*

View the “! HEAT SYSTEM: HEAT TO 300”segment of the script. You can see that two files are opened for writing: 0.rst and 0.dcd. After these files are opened, our heating script follows:

dynamics cpt leap verl strt nstep 40000 time 0.0015 -

iunrea -1 iunwri 31 iuncrd 32 -

isvfrq 2000 nsavc 200 nsavv 0 -

inbfrq -1 nprint 250 iprfrq 0 ntrfrq 100 -

pcon pgam 25 pmass 500 pref 1.0 surface tension 0.0 -

ihtfrq 2 teminc 0.06 -

imgfrq 50 ixtfrq 1000 cutim 14 -

iasors 0 iasvel 1 iscvel 1 iseed 1 -

firstt 0.0 finalt 300.0

There are several values we need to understand when running this heating script:

**dynamics cpt leap verl strt** – this begins our dynamics simulation as a constant pressure and temperature (cpt) leap verlet simulation.

**nstep 40000** – this designates the number of steps to be performed.

**time 0.0015** – this designates the length in time (in picoseconds) of a single step, or how long a simulation is allowed to move before the velocities of the system are recalculated and adjusted. .0015 picoseconds, or 1.5 femtoseconds.

*Question 6: How many picoseconds of simulation would be carried out with the current script?*

**nsavc 200** – this determines the number of steps before the coordinates are saved to the .dcd file

**iunwri 31** – this will write our RESTART file to the open unit 31. Restart files are essential for restarting our simulations from where our previous simulation finished off.

**iuncrd 32** – this will write our .dcd file to the open unit 32.

**pcon** – this command and its subsequent values designate that the simulation will run under constant pressure.

**firstt 0.0 finalt 300.0 –** this will heat our script from 0.0 K to 300.0 K.

Now run the script. Submit your script to the supercomputer in the UNIX console by typing

sbatch batch.submit

You can view the progress of your script in several ways by typing in the command line:

watch squeue –u (your fsl username)

This script will take about 7 hours to complete. The finished product has been provided for you.

**III. Perform Dynamics Simulations**

Now we are going to perform our dynamics simulation. Open the folder 5\_simulate, and the input file 1E9\_sim. You’ll notice that this file is very similar to the heating script, without the minimizations. The differences are primarily in the dynamics simulation.

Analyze the differences in the dynamics settings:

dynamics cpt leap verl restart nstep 400000 time 0.00125 -

iunrea 30 iunwri 31 iuncrd 32 -

isvfrq 2000 nsavc 800 -

inbfrq -1 nprint 2000 iprfrq 2000 ntrfrq 100 -

PCONst pmass 500.0 pgamma 0.0 tbath 300.0 PREFerence 1.0 -

hoov tmass 1000 tbath 300 tcoup 5.0 tref 300 -

imgfrq 50 ixtfrq 1000 cutim 14 -

ichecw 0 iscvel 1 -

finalt 300.0

*Question 7: What are the differences between this simulation and the heating script? Ask your TA what these differences mean, or look them up on* [*www.charmm.org*](http://www.charmm.org)*.*

Now prepare the submission script. You’ll notice that there are various folders within 5\_simulate. This is necessary to organize the output files of our simulation. To perform a long simulation we have to pause the simulation and restart it from where we left off. This is to decrease the amount of consecutive wall time on the supercomputer, and to protect us from having to restart the entire simulation should an error occur.

If you were performing this simulation yourself, your next script would take about 3 days to complete. Luckily, we have the finished product already available to you.

**IV. ANALYSIS**

The analyses that we are going to perform on our system are the Root Mean Square Deviation (RMSD), and a graph of the distance between two residues over time. These are simple analyses, and we’re going to be using VMD to perform them.

*RMSD Analysis*

In WinSCP, transfer the 0.psf, 0.crd, and 1.dcd files onto your local computer. When complete, open VMD on your computer. Go to File > New Molecule. Load 0.crd, be sure to specify that the file contains CHARMM coordinates in the drop down menu. Load 0.psf and 1.dcd.

Once all the frames are loaded, go to Extensions > Analysis > RMSD Visualizer Tool. In the new window, type “all segid PROA” in the atom selection dialogue box. Next, click “ALIGN” and then “RMSD”. Select “Plot Result”

*Question 8: Describe the data that you see. What causes the changes in RMSD?*

*Distance between Atoms over Time*

This is another easy analysis that graphs the distance between atoms over time. Create a bond in VMD between any atom in the ligand and an atom near it in the surrounding luciferase pocket. Then go to Graphics > Labels, select your bond, click the “Graph” tab, and Click the “Graph” Button.

*Question 9: Describe the data that you see, and propose a possible conclusion.*