CHEM 436 / CHEM 630: Molecular Modelling of Proteins
TUTORIAL #4a: Molecular dynamics: Setup

INTRODUCTION

In this tutorial, we will use the CHARMM-GUI server (http://charmm-gui.org) to convert the all-atom PDB structure created in Tutorial #2 into PDB and PSF files compatible with the CHARMM force field and to build a simulation system describing the protein in solution.

Do not use the CHARMM-GUI server blindly! At each step, take the time to read the instructions to make sure you understand what is being done.

REQUIRED PRE-LAB READING

None

PRE-LAB REPORT

Compile the following information for your protein model:

- In preparation for STEP 1, decide which type of N-terminus and C-terminus you will be using and briefly explain why.
- List all His residues that need to be protonated on the ε-nitrogen instead of the δ-nitrogen (if any).
- List all Asp/Glu residues that should be protonated (if any) and all His residues that should be doubly protonated (if any).
- List all pairs of Cys residues forming a disulfide bond (if you have any).
- List all residues coordinating a zinc atom (if you have any).
- List all buried water molecules you have decided to keep, and what protein residues they are likely forming a hydrogen-bond with.

READING

Chapter 7 of Leach (“Molecular Dynamics Simulation Methods”): Sections 7.1, 7.4, and 7.9.

REFERENCE MATERIAL

PDB format:
http://www.ks.uiuc.edu/Training/Tutorials/namd/namd-tutorial-unix-html/node22.html

PSF format:
http://www.ks.uiuc.edu/Training/Tutorials/namd/namd-tutorial-unix-html/node23.html

CHARMM topology file format:
http://www.ks.uiuc.edu/Training/Tutorials/namd/namd-tutorial-unix-html/node24.html

CHARMM parameter file format:
http://www.ks.uiuc.edu/Training/Tutorials/namd/namd-tutorial-unix-html/node25.html

NAMD configuration files:
http://www.ks.uiuc.edu/Training/Tutorials/namd/namd-tutorial-unix-html/node26.html

NAMD standard output:
http://www.ks.uiuc.edu/Training/Tutorials/namd/namd-tutorial-unix-html/node27.html

PROCEDURE

STEP 1: Read the PDB file of your protein model

Create a directory "`tutorial4`" and put a copy of the PDB file of your final, all-hydrogen model (obtained from STEP 14 of Tutorial #2c). Give that file a short name with no space of punctuation characters (example: "`protein.pdb`").

From the "Front Page" of the CHARMM-GUI server (http://charmm-gui.org), go to the "Input Generator" section and select the "Quick MD Simulator" module.

Upload your PDB file. For the PDB format, check "RCSB".

Proceed to the next step ("Select Model/Chain").

Model/Chain Selection Option:

Make sure CHARMM-GUI has detected all segments in your PDB file: The protein (from the first to the last residue), the heteroatoms (all residues), and the water molecules. Keep the PDB file open in a text editor by the side to confirm that all residues were properly read.

Select the "Protein" chain and all "Water" chains, but be aware that not all "Hetero" chains present in your model can be selected—because not all of them can be described by the CHARMM force field. Ask the instructor about which of your hetero chains should be included and which should not.

Proceed to the next step ("Manipulate PDB").

PDB Manipulation Options:

Terminal Group Patching:

Decide whether the first and last amino acids of your protein chain should be "patched" with a standard charged N-terminus group ("NTER", containing an ammonium $–NH_3^+$ group) and a standard charged C-terminus group ("CTER", containing a carboxylate $–COO^–$ group), or with a neutral acetylated N-terminus ("ACE") and a neutral N-methylamide C-terminus ("CT3"). This decision depends on whether your sequence corresponds to a truncated protein chain or not.

Preserve hydrogen coordinates:

Leave this box unchecked unless you are sure you have correctly included all hydrogen atoms. (Keep in mind that no hydrogen atoms were added to hetero ligands during the homology modeling process.)

Mutation:

Check this option if you have histidine residues that should be "HSE" instead of the default "HSD"—based on what you have observed in Tutorial #2. Select those "HIS" residues using their SEGID and RESID values and mutate them to "HSE". (Use the "Add Mutation" button if needed.)

Protonation:

Check this option if you have Asp/Glu residues that should be protonated (i.e., neutral) or His residues that should be doubly protonated (i.e. positively charged)—based on what you have observed in Tutorial #2. Select the residues that should be protonated using their Residue name and ID. (Use the "Add Protonation" button if needed.)

(Be careful with the "Mutation" and "Protonation" steps! Double check everything!)

Disulfide Bonds:

Check this option if your protein structure has disulfide bonds, and add bonds as needed.

Symmetry Operation Options:

(No need to pick any option.)

Proceed to the next step (“Generate PDB”).

STEP 2: Solvate your protein model

Waterbox Size Options:

Use a rectangular water box, with a 10 Å edge distance. (This will create a layer of solvent at least 20 Å thick between each image of the protein.)

Add Ions:

Include 0.15 M potassium chloride. Place the ions using the Monte Carlo method.

Proceed to next step (“Solvate Molecule”).

System Size:

Write down the parameters that have been chosen by CHARMM-GUI.

Periodic Boundary Conditions Options:

Generate grid information for PME FFT automatically.

Proceed to the next step (“Setup Periodic Boundary Conditions”).

Input Generation Options:

Ask only for the NAMD inputs (uncheck all other options).

Equilibration Input Generation Options:

Choose the NVT Ensemble.

Dynamics Input Generation Options:

Choose the NPT Ensemble and set the temperature to 298.15 K.

Proceed to the next step (“Generate Equilibration and Dynamics Input”).

Download the files prepared by CHARMM-GUI by clicking on the button called “download .tgz”. Save the file “`charmm-gui.tgz`” in your “`tutorial4`” directory. (You might want to keep a copy of that file somewhere else.)

STEP 3: Inspect the CHARMM-GUI input files

Decompress the archive using the following command from the terminal:

```
$ tar -z -xvf charmm-gui.tgz
```

This will create a directory called “`charmm-gui`” in which all the files needed for the simulation can be found.

Files “`step1_*`” to “`step4_*`” correspond to what the CHARMM-GUI server has done (using the software CHARMM) to read the PDB file provided, to solvate the system, to set up periodic boundary conditions, and to perform a quick equilibration.

File “`step5.1_production.inp`” could be run using CHARMM. On the other hand, files “`namd/step4_equilibration.inp`” and “`namd/step5_production.inp`” (in the “`namd`” subdirectory) can be run using NAMD.

Load the PSF and PDB files of the system in VMD:

```
$ vmd step3_pbcsetup.xplor.ext.psf step3_pbcsetup.pdb
```

and examine the structure built by CHARMM-GUI.

If CHARMM-GUI read any of your hetero residues, check that their positions and structures are correct. You may want to check once again that the His/Glu/Asp residues are correctly protonated and that the buried water molecules are at the right place.

✦ How many atoms does your system contain? What fraction of this number is water? How many K^+ and Cl^- ions were generated?

✦ Using either VMD or PyMOL, create a image of the system showing the protein as cartoon, the ions as spheres, and the water molecules as lines. If any metal atoms are present, render them as spheres and render their side chain ligands as sticks. If any disulfide bonds are present, render them as sticks (see page 3 of Tutorial #1c). Join this image to your report.

You may want to discuss your system with the instructor before proceeding to STEP 4.

STEP 4: If you have zinc atoms in your system, add extra coordination bonds

The standard non-bonded CHARMM force field tends to yield poor coordination geometries for zinc. To maintain the tetrahedral coordination, you need to define additional bonds between each zinc atom and the four atoms coordinating it (either histidine nitrogen atoms or water oxygen atoms).

• Edit files "`namd/step4_equilibration.inp`" and "`namd/step5_production.inp`" to add the following two lines after the "`rigidBonds`" statements:

```
extraBonds on;
extraBondsFile extrabonds.txt;
```

• Create a text file "`namd/extrabonds.txt`" (in the "`namd`" subdirectory) containing as many "bond" entries as you have coordination bonds involving zinc atoms:

```
bond [zincAtomIndex] [atomIndex1] [k] [ref]
bond [zincAtomIndex] [atomIndex2] [k] [ref]
bond [zincAtomIndex] [atomIndex3] [k] [ref]
bond [zincAtomIndex] [atomIndex4] [k] [ref]
```

(Read the "extraBonds" section of the NAMD User Guide for more information.)

• To find the value of "`[zincAtomIndex]`", load the PDB file "`step3_pbcsetup.pdb`" in VMD and find the index of the zinc atom by typing "0" and clicking on it (as explained in Step 11 of Tutorial #3b). The index number will be printed in the VMD terminal window. Find the indices of the four coordinating atoms the same way.

• Use a value of 200 for "`[k]`" and a value of 2.00 for "`[ref]`".

STEP 5: Prepare the molecular dynamics simulation

Edit file "`namd/step4_equilibration.inp`" as following:

• Replace all instances of "500" by "5" (for "restartFreq" and "reassignFreq") and all instances of "1000" by "10" (for "dcdFreq", "xstFreq", and "outputTiming").

• Set "outputEnergies" to 1 (instead of 125).

• Set "stepspercycle" to 10 (instead of 20).

• Set "minimize" to 100 (instead of 10000) and "run" to 50 (instead of 25000).

Run a benchmark equilibration simulation in the background:

```
$ namd2 step4_equilibration.inp > step4_equilibration.log &
```

✦ This calculation performs 100 minimization steps and 50 molecular dynamics steps. Examine the log file and report the average “CPU time” per step and average “wall time” per molecular dynamics step. Estimate how much time (“CPU” and “wall”) would be required for a 1.0 ns simulation (500,000 steps instead of 50) if the simulation were performed on that computer.

Edit file “`namd/step5_production.inp`” as following:

- Replace all instances of “500” by “5” (for “restartFreq”, “xstFreq”, “outputTiming”) and all instances of “1000” by “10” (for “dcdFreq”).
- Set “outputEnergies” to 1 (instead of 125).
- Set “stepspercycle” to 10 (instead of 20).
- Set “numsteps” and “run” to 50 (instead of 500000).

Once the equilibration simulation is done, run a benchmark production simulation in the background:

```
$ namd2 step5_production.inp > step5_production.log &
```

Load the PSF and the trajectory in VMD and make sure the simulations run smoothly:

```
$ vmd ../step3_pbcsetup.xplor.ext.psf step5_production.dcd
```

Pay attention to buried water molecules, the cofactors, and the metal atoms bound to the protein. Inform the instructor of any unusual change in coordination, keeping in mind that Zn^{2+} tends to be tightly tetracoordinated by soft ligands (nitrogen- or sulfur-containing) and that Ca^{2+} tends to be loosely hexacoordinated by hard ligands (water, carboxylates, carbonyls, etc.).

STEP 6: Submit the simulation

In file “`namd/step4_equilibration.inp`”, change the values of “restartFreq”, “reassignFreq”, “dcdFreq”, “xstFreq”, “outputTiming”, “outputEnergies”, “minimize”, and “run” back to their original values.

In file “`namd/step5_production.inp`”, change the values of “restartFreq”, “xstFreq”, “outputTiming”, “dcdFreq”, “outputEnergies”, “numsteps”, and “run” back to their original values.

Note:
An easier way to revert to the old version of the files is to decompress the original CHARMM-GUI archive again, as described in STEP 3. Note that this will overwrite all changes you have made to the input files! If you have zinc atoms in your system, make sure you re-introduce the “extraBonds” and “extraBondsFile” lines as described in STEP 4.

Create a new archive using the following command from the terminal:

```
$ cd ../..
$ tar -z -cvf charmm-gui_????.tgz charmm-gui
```

Replace ???? with your last name.

Call the instructor when your “.tgz” file is ready to be transferred to the cluster.