Running MD simulation by NAMD

Adam Pirnia

July 2, 2024

1 Background

```
BASH COMMANDS CHEAT SHEET

SYSTEM INFO
uname -a : Display kernel version
lsb_release -a : Display OS version

FILE AND DIRECTORY COMMANDS
pwd : Print working directory
ls -la : List files (detailed)
cd /path/to/dir : Change directory
mkdir dir_name : Create directory
rmdir dir_name : Remove directory
touch file_name : Create empty file
cp source dest : Copy file
cp -r source_dir dest : Copy directory
mv old_name new_name : Rename or move file
rm file_name : Remove file
rm -r dir_name : Remove directory and contents
cat file_name : View file content
head -n 5 file_name : View first 5 lines of the file contents
tail -n 5 file_name : View last 5 lines of the file contents


FILE PERMISSIONS
chmod +x file : Give execution permissions for a file


NETWORK COMMANDS
ssh user@host : SSH to remote host # example: ssh asuID@sol.asu.edu
scp source dest : Secure copy between hosts

FILE SEARCHING
grep -r "pattern" path_directory : Recursive search for text in all files in path_directory

TEXT EDITORS
nano file_name : Open file in Nano
vim file_name : Open file in Vim
```

2 Softwares

NAMD, Nanoscale Molecular Dynamics, is a computer software for molecular dynamics simulation.

VMD, Visual Molecular Dynamics is a molecular modelling and visualization computer program. VMD is developed mainly as a tool to view and analyze the results of molecular dynamics simulations.

TCL programming language, is a high-level, general-purpose, interpreted, dynamic programming language. It was designed with the goal of being very simple but powerful. VMD already has a built in TCL compiler (interpreter).

** Installation, depending on the OS, is always included in the README document.**

3 Ingredients

PDB file (Protein Data Bank), the primary structure. It contains names and coordinates of every single atom in protein. Also, what residue and chain they belong to.

PSF file (Protein Structure File), other structural info such as what bonds, angles, dihedrals, and improper angles exist in the protein. Additional info about the atoms, such as partial charge, and the type of the atom regarding their chemical situation (what atoms and groups they are bonded to) are included.

Force field topology file contains general structural information such as what bonds and angles exist in a particular residue. It is needed to generate a PSF file.

Force field parameter file is a mathematical expression of the potential which atoms in the system experience. CHARMM, X-PLOR, AMBER, and GROMACS are four types of force fields, and NAMD is able to use all of them. The parameter file defines bond strengths, equilibrium lengths, etc. For example, a CHARMM forcefield parameter file contains all of the numerical constants needed to evaluate forces and energies, given a PSF structure file and atomic coordinates.

Configuration file (also called a config file, .conf file, or .namd file) is given to NAMD on the command line and specifies virtually everything about the simulation to be done.

4 Actual work

4.1 PSF generation

```
cd ~/namd_tutorial_files/1-1-build/example-output
vmd -dispdev text -e ubq_pgn.tcl
```

```
package require psfgen
# loading the PSF generator package (included in the VMD software)

topology top_all27_prot_lipid.inp
# the topology file is needed to generate PSF

pdbalias residue HIS HSE
# change the residue name of the HIS to HSE, this should be customized based on the redox state

pdbalias atom ILE CD1 CD
# change the atom name of carbon delta in ILE residue for compatibility purpose

segment U {pdb ubqp.pdb}
# laods ubqp.pdb (should be in the same directory) and changes the segment name to U
# segment U also adds H to the protein as X-ray crystallography cannot resolve them

coordpdb ubqp.pdb U
# read coordinates from PDB file, matching segment, residue and atom names

guesscoord
# guesses coordinates of atoms for which they were not explicitly set

writepdb ubq.pdb
# write pdb file with all new info

writepsf ubq.psf
# write psf file

exit
```

Figure 1: Comments describing each line of the TCL code.

4.2 Solvating the protein

```
cd ~/namd_tutorial_files/1-1-build/example-output
vmd -dispdev text -e solvate.tcl
```

```
package require solvate
# loading the solvate package (included in the VMD software)

solvate ubq.psf ubq.pdb -t 5 -o ubq_wb
# loading the psf and pdb files. The -t option creates the water box dimensions such that
# there is a layer of water 5Å in each direction from the atom with the largest
# coordinate in that direction. The -o option creates the output files ubq_wb.pdb
# and ubq_wb.psf for ubiquitin with the water box.

exit
```

4.3 Simulation cell

```
cd ~/namd_tutorial_files/1-1-build/example-output
vmd -dispdev text -e sim_cell.tcl
```

```tcl
mol load pdb ubq_wb.pdb

set all [atomselect top all]
# selecting all the atoms in the pdb file

set cn [measure center $all]
# measuring the geometrical center of the system

set vect [vecsub {0.0 0.0 0.0} $cn]
# making a vector from the geometrical center to the origin

$all moveby $vect
# moving all atoms along the vector to recenter to the origin

set mm [measure minmax $all]
# measuring the maximum and minimum x, y, and z among all coordinates
#-------------------------------------------------------#
# split the mm variable to extract the coordinates
set vec1 [split [lindex $mm 0]]
set vec2 [split [lindex $mm 1]]

# extract individual coordinates
set x1 [lindex $vec1 0]
set y1 [lindex $vec1 1]
set z1 [lindex $vec1 2]
set x2 [lindex $vec2 0]
set y2 [lindex $vec2 1]
set z2 [lindex $vec2 2]

# calculate the simulation cell edges
set newX [expr {$x2 - $x1}]
set newY [expr {$y2 - $y1}]
set newZ [expr {$z2 - $z1}]
#-------------------------------------------------------#

puts $newX
puts $newY
puts $newZ
# print the simulation cell edges

exit
```

Executing the last commands above, the output last three lines will be three nubmers that are going to be the length of the edges of the whole system. These are needed for telling the NAMD about an imaginary box in which NAMD will put the system to perform the simulation.

4.4 Running NAMD / Configuration file

Open the PDF file named conf_file.pdf to see the full explanation on the configuration file.

Before running NAMD, modify the configuration file as needed. For example, specifying the temperature, and cimulation cell (using the output of the previous section (sim_cell.tcl)).

The below commands will take you to the corresponding directory and call NAMD software and do the simulation based on the configurations set in the ubq_wb_eq.conf file

```
cd ~/namd_tutorial_files/1-3-box
path_to_namd/namd2 ubq_wb_eq.conf > output.log
```