Laboratory Instruction for Analyzing NAMD Output

Introduction

The focus of this lab is to learn how to analyze the output of the NAMD (Molecular Dynamic software) simulation. The output file generated by NAMD contains a wealth of information, such as potential energy, electrostatic force, van der Waals force, and more. The goal of this lab is to learn how to extract desired data from the NAMD output file.

In the first step, we will run NAMD to generate the output file. Then we will use shell scripts to extract data from the output file.

Overview of Key Files in NAMD Simulations

PDB (Protein Data Bank) Files

- Purpose: PDB files store the 3D structures of biological molecules such as proteins and nucleic acids. They contain atomic coordinates and may include information about the molecule's secondary structure, connectivity, and other relevant data.

- Format: Text-based with specific columns for atom names, residue names, chain identifiers, and XYZ coordinates.

```
ATOM      1  N   MET A   1    38.428  13.847  22.521  1.00 51.51      N
ATOM      2  CA  MET A   1    37.788  12.536  22.881  1.00 21.18            C
```

PSF (Protein Structure File)

- Purpose: PSF files describe the structural topology of a molecular system, including information about atoms, bonds, angles, dihedrals, and impropers. They are essential for defining the molecular connectivity.

- Format: Includes sections for atoms, bonds, angles, dihedrals, and other structural elements.

```
PSF
        1 !NTITLE
REMARKS Created by psfgen
        5 !NATOM
        1 ALA  1   N   NH3  -0.300000  0
        2 ALA  1   HT1 H       0.330000        0
        3 ALA  1   HT2 H       0.330000        0
```

Configuration Files in NAMD

- Purpose: Define the parameters for a simulation, including input files, simulation parameters, output settings, and more.

- Format: Plain text with commands to specify parameters and settings.

```
structure       mystructure.psf
coordinates     mycoordinates.pdb
```

```
parameters      myparameters.prm
temperature     300
outputName      myoutput
```

Topology Files

- Purpose: Define the molecular topology, including residues, atoms, bonds, angles, dihedrals, and other parameters necessary for building the PSF file.

- Format: Text-based following a specific format required by the simulation software.

```
RESI ALA
GROUP
        ATOM N   NH3  -0.3000
        ATOM HT1 H   0.3300
        ATOM HT2 H   0.3300
        ATOM HT3 H   0.3300
```

Parameter Files

- Purpose: Contain the force field parameters required for simulations, such as bond lengths, angles, dihedral angles, and non-bonded interaction parameters.

- Format: Text-based with sections for different types of parameters.

```
BOND
NH3  H          1.04   450.0
CA   CB   1.53   300.0

ANGLE
H  NH3  H  109.5   35.0
CA  CB  CG 109.5   70.0
```

Running NAMD

1. Navigate to the directory where your configuration file is located:

cd minimization/namdConf

2. Run the NAMD simulation:

/Users/ChemLab/Desktop/NAMD_2.14_MacOSX-x86_64-multicore/namd2 1ubq_min.300.restart.conf > 1ubq_min.300.restart.conf.out

Extracting Energy Data from NAMD Output

1. Extract energy data using `grep` and `awk`:

grep "ENERGY: " 1ubq_min.300.restart.conf.out
grep "ENERGY: " 1ubq_min.300.restart.conf.out | awk '{printf "%f\n",$7*0.0433634}' >> potentialE.dat

2. Automate data extraction with a shell script:

Create a script named `getEnergy.sh`:

```
#####################################################################

#!/bin/bash

nss=1
nsf=10
#---------------------------------------------------------------------
echo "..... START ANALYZING ....."
#---------------------------------------------------------------------

for (( i=${nss}; i<=${nsf}; i++ ))
do
        echo "*** ANALYZING ${i}ns ***"
        # Column number 7 should be multiplied by 0.0433634 to be converted to eV.
        grep "ENERGY:  " ${i}ns.out | awk '{printf "%f\n",$7*0.0433634}' >> potentialE.dat

        # Uncomment the following line to extract force data if needed
        # grep "ELECT_FORCE: " ${i}ns.dat | awk '{printf "%f %f
%f\n",$10*0.0433634,$11*0.0433634,$12*0.0433634}' >> force.eVPeA.dat.tmp
done

#---------------------------------------------------------------------
echo "..... DONE ANALYZING ALL DATA ....."
#---------------------------------------------------------------------

exit

#####################################################################
```

3. Make the script executable:

```
chmod +x getEnergy.sh
```

4. Run the script:

```
./getEnergy.sh
```