Molecular Dynamics Simulation by using NAMD-VMD and Gromacs

Jitendra Sasumana[1] and Vikas Kaushik[1*]

[1]Department of Bioinformatics, School of Bio Engineering and Bio Sciences, Lovely Professional University, Punjab, India

Abstract: Molecular Dynamics Simulation is a form of computer simulation, where atoms and molecules are allowed to interact for a period of time under some laws of physics. Calculates the time dependent behaviour of a molecular system. Simulations have provided detailed information on the fluctuations and conformational changes of proteins and nucleic acids. Used to investigate the structure, dynamics and thermodynamics of biological molecules and their complexes. One of the principal tools in the theoretical study of biological molecules.

Introduction: Molecular Dynamics simulation is the important tool in the theoretical study of biomolecules. This procedure technique calculates the time dependent behavior of a molecular system. MD simulations have provided careful data on the fluctuations and conformational changes of proteins and nucleic acids [1]. We carry out computer simulations within the hope of understanding the properties of assemblies of molecules in terms of their structure and therefore the microscopic interactions between them. This is a complement to standard experiments, enabling us to find out something new, something that can't be acknowledged in other ways [4].

The two main families of simulation technique are molecular dynamics (MD) and Monte Carlo (MC); additionally, there's an entire range of hybrid techniques which combine features from both. During this study, we shall consider MD. The apparent advantage of MD over MC is that it gives a route to dynamical properties of the system: transport coefficients, time-dependent responses to perturbations, rheological properties and spectra [2]. Computer reenactments go about as an extension between minuscule length and time scales and the plainly visible universe of the research facility: we give a theory at the connections among atoms, and get 'accurate' expectations of mass properties. The forecasts are 'definite' as in they can be made as precise as we like, subject to the confinements forced by our PC spending plan. Simultaneously, the shrouded detail behind mass estimations can be uncovered. A model is the connection between the dissemination coefficient furthermore, speed autocorrelation work (the previous simple to quantify tentatively, the last mentioned a lot harder) [3].

At last we might need to make direct correlations with trial estimations made on explicit materials, in which case a decent model of atomic collaborations is basic. The point of supposed abdominal muscle initio atomic elements is to lessen the measure of fitting what's more, mystery right now a base [3]. Then again, we might be intrigued in marvels of a fairly nonexclusive nature, or we may basically need to segregate between great and awful hypotheses. With regards to points of this sort, it isn't important to have a flawlessly practical sub-atomic model; one that contains the fundamental material science might be very reasonable [4].

To perform a molecular dynamics simulation we use certain software’s like VMD and NAMD, GROMACS, and Desmond. By these we solvation of biomolecules, simulating the molecules to calculate the minimum energy for which we do energy minimization and we analyze the trajectory, rmsd values and the energies of the molecule after the simulation.

Significance of MD Simulation

- MD simulations permit the study of complex, dynamic processes that occur in biological systems. These include, for example;
    - Protein Stability
    - Conformational changes
    - Protein folding
    - Molecular recognition: proteins, DNA, membranes, complexes
    - Ion transport in biological systems
    - Drug Design
    - Refinement of the structure determined through X-ray Crystallography and NMR Spectroscopy.

Performing MD Simulation:

- To run a simulation several things are needed like:
    - A file containing the coordinates for all atoms
    - Information on the interactions (bond angles, charges, Vander Waals)
    - Parameters to control the simulation
    - Processors to run the simulation

The **PDB file** contains the coordinates for all atoms and is the input structure file for MD Simulation. The interactions are listed in the **topology** (**.top**) file and the input parameters are put into a **.mdp** file.

Methodology: VMD and NAMD are the main molecular dynamics simulation software packages, which will work together to know about the structural information of biomolecules. By doing these molecular simulations can give the knowledge to researchers regarding the roles and functions of various biomolecules in life science research [5]. Molecular dynamics simulation has a great significance for biological scientists to study the physical basis of the structure and the function of proteins of bio-molecules since the internal motions of individual atoms play an essential role in the functional mechanism of the proteins in living organisms [6].

NAMD is a molecular dynamics software designed for high performance simulation of large bimolecular system. This software is designed to run on a large number of processors of computer clusters or it can be a large number of cores in graphics processing unit hardware [5]. Usually NAMD works together with molecular graphics software VMD; it provides easy accessible tools, which explore the structural information of bio-molecules. In a glimpse, VMD is visualization software for simulation of a molecule, displaying and analyzing large bio-molecule with the help of 3D graphics and built-in script. VMD and NAMD are available free of cost with source codes and can be supported on different OS platforms like such as Linux, Windows and Mac OS [8].

Download and Install VMD Software in C Drive of your computer. Download the NAMD tutorial files and Save it in VMD folder along with vmd.exe file and all other files related to VMD software [8].

After installation open VMD, go to Extensions tab and open Tk console.

Through Tkconsole enter into NAMD tutorial files and then 1-1-build folder. Now open File tab in VMD and go to new molecule and browse the protein of your own (In this tutorial I am using 5eui protein) and load it in VMD [9].

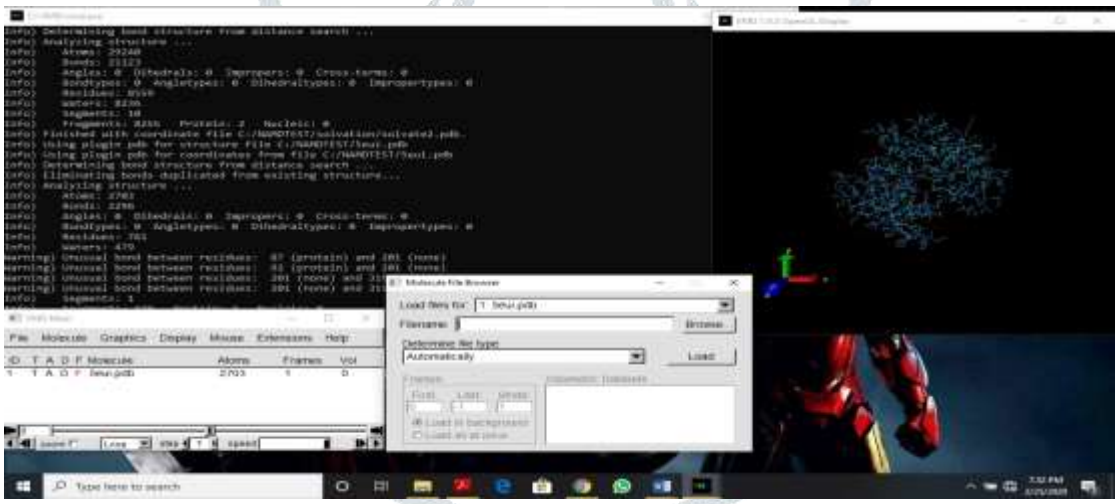


Fig 1: Browse the protein and load it in VMD.

To run the simulation we need to have the coordinates file, .psf extension file, parameter file and configuration file.

Now in Tk console give the command

Set 5eui [atomselect top protein]

$5euiwritepdbeuip.pdb

This command will extract all the coordinates of your protein molecule and will save it in output in pdb format.

Now delete the protein molecule from VMD which have been loaded first and load newly generated coordinates file in VMD. Now go to Extensions tab → Modeling → Automatic PSF Builder and Convert the coordinated file which is in euip.pdb to eui.psf file.

Now the solvation of the protein will be in two ways, which are:

1. Non-periodic conditions by creating water sphere.
2. Periodic conditions by creating water box.

Non-periodic conditions by creating water sphere:

Through Tk console go to the NAMD tutorial files and then 1-1-build folder and use these commands.

sourcewat_sphere.tcl

There will be two output files will be generated which are eui_ws.pdb and eui_ws.psf.

Through VMD → File → New molecule → Browse the eui_ws.psf and load it. We will not able to see it in visualizer but we will be able to see it in window.

Through VMD → File → New molecule → Browse the eui_ws.pdb and load it. We will be able to see it in visualizer as well as in window.

Delete the eui_ws.psf file and eui_ws.pdb file from the VMD window.

Fig 2: Non-Periodic condition after creating Water sphere.

Periodic conditions by creating water box:

Through Tk console go to the NAMD tutorial files and then 1-1-build folder and use these commands.

package require solvate

solvate eui.psf eui.pdb -t 5 -o eui_wb

There will be two output files will be generated which are eui_wb.pdb and eui_wb.psf.

Through VMD → File → New molecule → Browse the eui_wb.psf and load it. We will not able to see it in visualizer but we will not able to see it in window.

Through VMD → File → New molecule → Browse the eui_wb.pdb and load it. We will be able to see it in visualizer as well as in window.

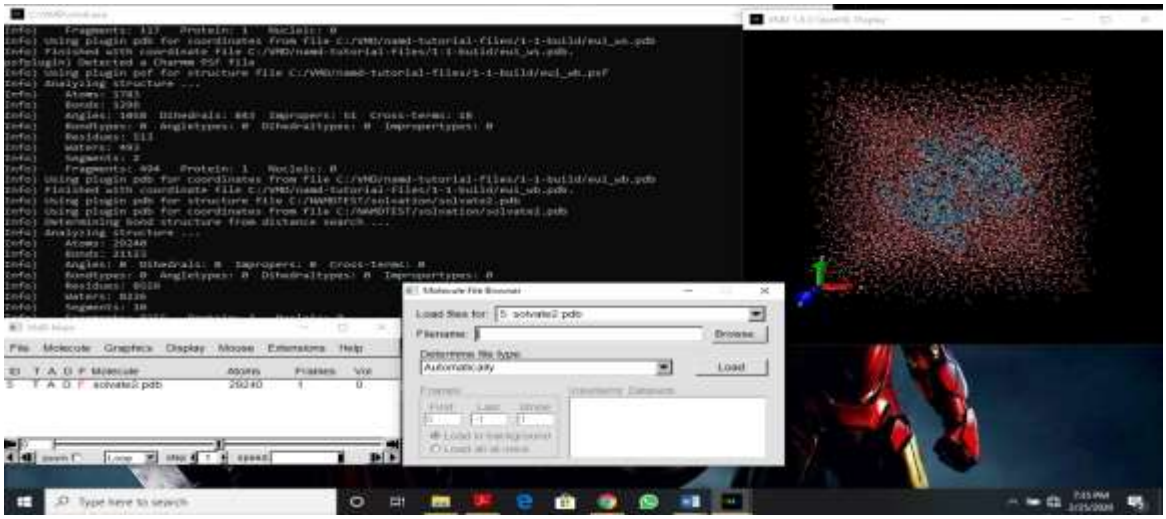


Fig 3: Periodic condition after creating Water Box.

Through Tk console in 1-1-build folder use commands

Set everyone [atomselect top all]

Measureminmax $everyone

It helps to measuring the minimum and maximum values of x,y and z coordinates of the entire water system.

X min: -41.3409 Y min: -52.1500 Z min: -48.3680

X max:28.9680 Y max: 18.1560 Z max: 12.9379

Quit VMD.

Copy all the 6 files euip.pdb, eui.psf, eui_ws.pdb,eui_ws.psf, eui_wb.pdb and eui_wb.psf into common folder along with some parameter files to run the simulation.

Simulation in non-periodic conditions for which we have created water sphere:

Go to NAMD tutorial files, open the 1-2-sphere folder, and analyze the configuration file parameters. Minimization is generally performed after setting all the atomic velocities to zero.

Download the NAMD Software in C Drive of your computer [9].

As we discussed earlier that NAMD will work together with VMD, so we need to set up NAMD. Right click on my computer → Properties → Advance system settings → Environmental variables → Edit Path and give the complete Path of NAMD.

Now use command prompt to enter VMD → NAMD tutorial files → 1-2-sphere folder. Now give the command

Namd2 eui_ws_eq.conf> eui_ws_eq.log &

Several output files will be generated.

Simulation in periodic conditions for which we have created water box:

Now use command prompt to enter VMD → NAMD tutorial files → 1-3-box folder. Now give the command

Namd2 eui_wb_eq.conf> eui_wb_eq.log &

Several output files will be generated.

After finishing, now analyzing the Molecular Dynamics simulation result by considering 3 files

- Log file
- RMSD File
- Configuration of protein molecules

Now use command prompt to enter 1-2-sphere folder in NAMD tutorial files open eui_ws_eq.log text document and analyze it.

Open VMD → Analyze → NAMD Plot

Open file → select NAMD log file → enter 1-2-sphere → eui_ws_eq.log text document-click on TEMP checkbox as well.

Open file → plot selected data

Open VMD → NAMD tutorial files → common folder →eui_ws_psf→ load

Open VMD → NAMD tutorial files → 1-2-sphere folder--- eui_ws_eq.dcd→ load

Now we can see all 10 frames uploaded in VMD.

Through Tk console in 1-2-sphere folder Use commands

sourcermsd.tcl

This command will generate a file rmsd.dat, we can open that in excel sheet and plot a graph.

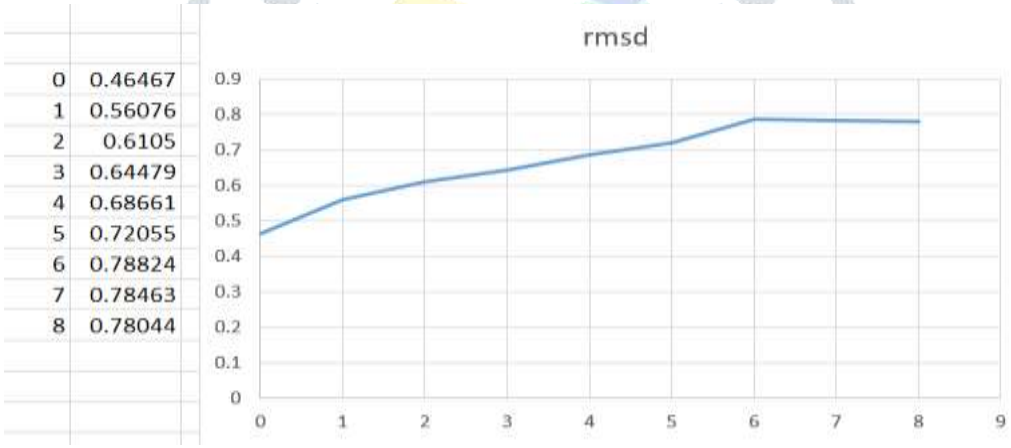


Fig 4: RMSD Values after plotting a graph in excel sheet.

Comparing the protein confirmation before and after simulation in periodic condition:

Open VMD → NAMD tutorial files → common folder→eui_wb_psf--- load

Open VMD → NAMD tutorial files → 1-3-box folder →eui_wb_eq.dcd--- load

Open VMD → Graphics → Representation →can visualize confirmations in different ways.

Selected atom → Protein

Coloring Method → Name

Drawing Method → Lines

Click on eui_wb_psf indicated in window and then

VMD file → save coordinates, selected atom → protein, file type .pdb and save it.

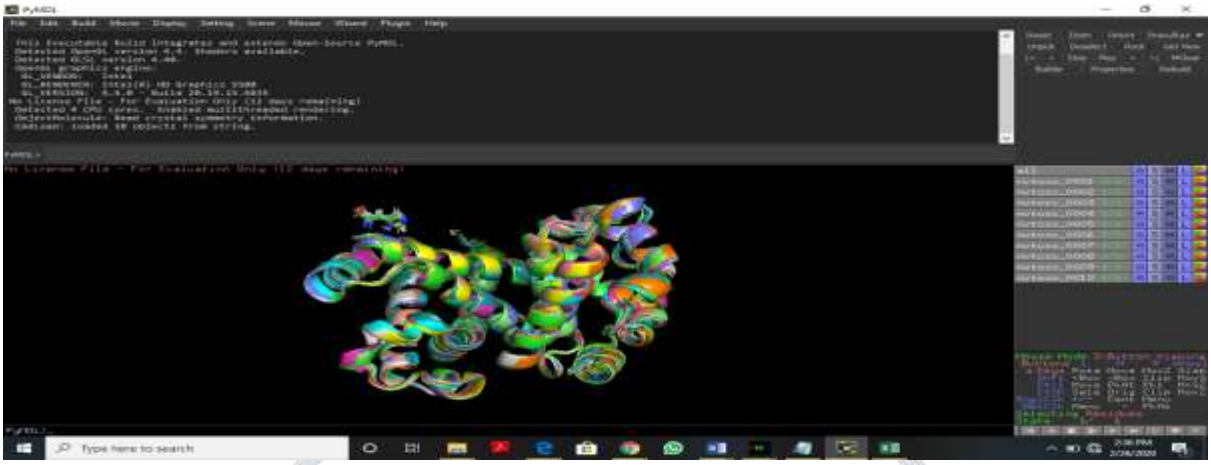


Fig 5: Visualizing all coordinates of 10 frames in Pymol.

Comparing the protein confirmation before and after simulation in non-periodic condition:

Open VMD → NAMD tutorial files → common folder→eui_ws_psf→ load

Open VMD → NAMD tutorial files → 1-2-sphere folder →eui_ws_eq.dcd→ load

Open VMD → Graphics → Representation →can visualize confirmations in different ways.

Selected atom → Protein

Coloring Method → Name

Drawing Method → Lines

Click on eui_wb_psf indicated in window and then

VMD file → save coordinates, selected atom → protein, file type .pdb and save it.

Fig 6: Visualizing all coordinates of 10 frames in PyMol.

There are some file required Molecular dynamics simulation through VMD and NAMD. Those are

- Water_sphere.tcl
- Configuration file
- Rmsd.tcl
- Parameters file to run simulation

Water_sphere.tcl is used solvating the protein in water sphere and water box.

Configuration file is to run the simulation in water sphere and water box.

Rmsd.tcl file is to generate the rmsd values after the simulation

Script for the Water_sphere.tcl:

```
### Script to immerse capsid protein in a sphere of water just large enough
### to cover it

set molname 12y

mol new ${molname}.psf
mol addfile ${molname}.pdb

### Determine the center of mass of the molecule and store the coordinates
set cen [measure center [atomselect top all] weight mass]
set x1 [lindex $cen 0]
set y1 [lindex $cen 1]
set z1 [lindex $cen 2]
set max 0

### Determine the distance of the farthest atom from the center of mass
foreach atom [[atomselect top all] get index] {
  set pos [lindex [[atomselect top "index $atom"] get {x y z}] 0]
  set x2 [lindex $pos 0]
  set y2 [lindex $pos 1]
  set z2 [lindex $pos 2]
  set dist [expr pow(($x2-$x1)*($x2-$x1) + ($y2-$y1)*($y2-$y1) + ($z2-$z1)*($z2-$z1),0.5)]
  if {$dist > $max} {set max $dist}
  }

mol delete top

### Solvate the molecule in a water box with enough padding (15 A).
### One could alternatively align the molecule such that the vector
### from the center of mass to the farthest atom is aligned with an axis,
### and then use no padding
```

```
package require solvate
solvate ${molname}.psf ${molname}.pdb -t 15 -o del_water

resetpsf
package require psfgen
mol new del_water.psf
mol addfile del_water.pdb
readpsf del_water.psf
coordpdb del_water.pdb

### Determine which water molecules need to be deleted and use a for loop
### to delete them
set wat [atomselect top "same residue as {water and ((x-$x1)*(x-$x1) + (y-$y1)*(y-$y1) + (z-$z1)*(z-$z1))<($max*$max)}"]
set del [atomselect top "water and not same residue as {water and ((x-$x1)*(x-$x1) + (y-$y1)*(y-$y1) + (z-$z1)*(z-$z1))<($max*$max)}"]
set seg [$del get segid]
set res [$del get resid]
set name [$del get name]
for {set i 0} {$i < [llength $seg]} {incr i} {
  delatom [lindex $seg $i] [lindex $res $i] [lindex $name $i]
  }
writepsf ${molname}_ws.psf
writepdb ${molname}_ws.pdb

mol delete top

mol new ${molname}_ws.psf
mol addfile ${molname}_ws.pdb
puts "CENTER OF MASS OF SPHERE IS: [measure center [atomselect top all] weight mass]"
puts "RADIUS OF SPHERE IS: $max"
mol delete top
```

Configuration File NAMD Simulation in Water sphere and Water Box:

```
structure          ../common/eui_ws.psf
coordinates        ../common/eui_ws.pdb

set temperature    310  ;# target temperature used several times below

# starting from scratch
temperature         $temperature    ;# initialize velocities randomly

# continuing a run
#set inputname      myinput          ;# only need to edit this in one place!
#binCoordinates     $inputname.coor  ;# coordinates from last run (binary)
#binVelocities      $inputname.vel   ;# velocities from last run (binary)
#extendedSystem    $inputname.xsc   ;# cell dimensions from last run

outputName         eui_ws_eq    ;# base name for output from this run

restartfreq         500     ;# 500 steps = every 1ps
dcdfreq             500
xstFreq             500
outputEnergies      100     ;# 100 steps = every 0.2 ps
outputTiming       1000     ;# shows time per step and time to completion

# Force-Field Parameters
paraTypeCharmm      on
parameters          ../common/par_all36_lipid_prot_carb.prm

# These are specified by CHARMM
exclude             scaled1-4
1-4scaling          1.0
switching           on

# You have some freedom choosing the cutoff
cutoff              12. ;# may use smaller, maybe 10., with PME
switchdist          10. ;# cutoff - 2.

# Promise that atom won't move more than 2A in a cycle
pairlistdist        14. ;# cutoff + 2.

# Integrator Parameters
timestep            2.0  ;# 2fs/step
rigidBonds          all  ;# needed for 2fs steps
nonbondedFreq       1    ;# nonbonded forces every step
fullElectFrequency  2    ;# PME only every other step

# Constant Temperature Control
langevin            on             ;# langevin dynamics
langevinDamping     1.             ;# damping coefficient of 1/ps
langevinTemp        $temperature   ;# random noise at this level
langevinHydrogen    no             ;# don't couple bath to hydrogens

# Periodic Boundary conditions
cellBasisVector1 70.30900001525879 0 0
cellBasisVector2 0 70.30600166320801 0
cellBasisVector3 0 0 61.305999755859375
cellOrigin -6.241332530975342 -16.991395950317383 -17.680776596069336


wrapWater           on               ;# wrap water to central cell
wrapAll             on               ;# wrap other molecules too
wrapNearest         off              ;# use for non-rectangular cells

langevinPiston        on
langevinPistonTarget  1.01325       ;# pressure in bar -> 1 atm
langevinPistonPeriod  100.          ;# oscillation period around 100 fs
langevinPistonDecay   50.           ;# oscillation decay time of 50 fs
langevinPistonTemp    $temperature ;# coupled to heat bath

minimize            500           ;# lower potential energy for 1000 steps
#reinitvels          $temperature  ;# since minimization zeros velocities
#run 50000 ;# 100ps
```

Script for RMSD Calculation:

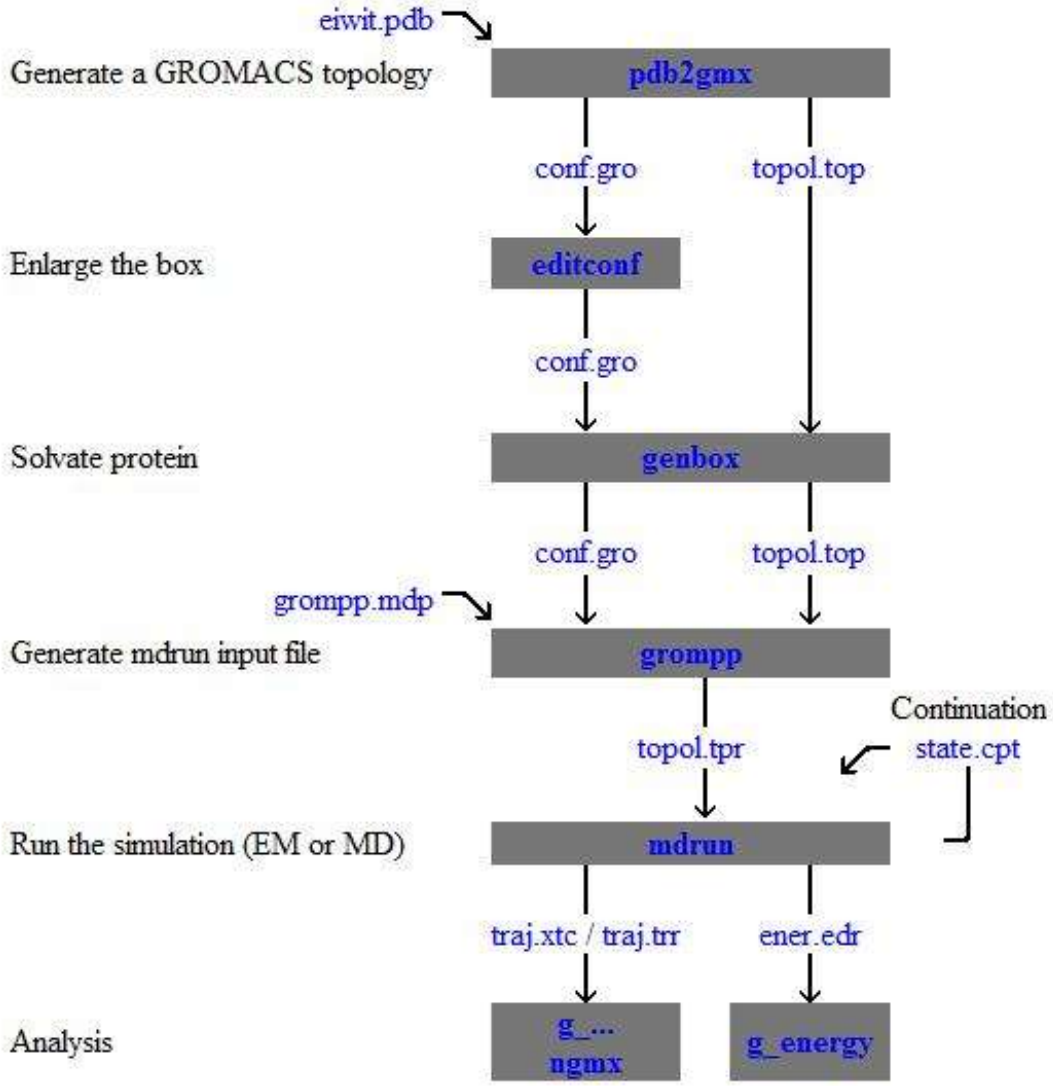
```
set outfile [open rmsd.dat w];
set nf [molinfo top get numframes]
set frame0 [atomselect top "protein and backbone and noh" frame 0]
set sel [atomselect top "protein and backbone and noh"]
# rmsd calculation loop
for {set i 1 } {$i < $nf } { incr i } {
    $sel frame $i
    $sel move [measure fit $sel $frame0]
    puts $outfile "[measure rmsd $sel $frame0]"
}
close $outfile
```

GROMACS:

Gromacs is a software used for Molecular dynamics simulation, which means simulating the Newtonian equations of motion for systems with hundred to millions of particles. GROMACS is usually designed for biochemical particles like proteins, lipids and nucleic acids which have a many complicated bonded and non-bonded interactions. GROMACS dominate simulations by calculating the non-bonded interactions very fast. We can also use it for doing research on non-biological systems like polymers [10].

- GROMACS (Groningen Machine for Chemical Simulations) is a molecular dynamics simulation package originally developed in the University of Groningen.
- GROMACS is an engine to perform molecular dynamics simulations and energy minimization.
- A high performance Molecular Dynamics Program.
- It is reasonably well optimized for low memory usage.
- GROMACS is the fastest program for molecular simulations.
- The support of different force fields make GROMACS very flexible.
- It is a high performance research tool which can be run on almost every platform.
- This program is free software; it can be redistributed and/or modified under the terms of the GNU General Public License as published by the Free Software Foundation.

Flow chart of Molecular dynamics simulation by GROMACS:



First we have downloaded the structure through pdb and then visualized it by using visualization software such as VMD, Chimera or Pymol. Now we have to delete the water. To delete the water molecules (residue "HOH") use this command [10].

grep –v HOH input.pdb > output.pdb

Now the water molecules are deleted and we will have only the necessary atoms present in the PDB file, containing only protein atoms and we are ready to input the first module of GROMACS, pdb2gmx. The pdb2gmx is generating three files:

- Topology file
- A position restraint file.
- A post- processed structure file.

The topology file, which is topol.top, will have the information to describe the molecule in a simulation. The information is all about non-bonded parameters and bonded parameters [10].

Running Simulation using GROMACS

```
pdb2gmx -f 1AKI_clean.pdb -o 1AKI_processed.gro -water spce
editconf -f 1AKI_processed.gro -o 1AKI_newbox.gro -c -d 1.0 -bt cubic
solvate -cp 1AKI_newbox.gro -cs spc216.gro -o 1AKI_solv.gro -p topol.top
grompp -f ions.mdp -c 1AKI_solv.gro -p topol.top -o ions.tpr
genion -s ions.tpr -o 1AKI_solv_ions.gro -p topol.top -pname NA -nname CL -
neutral
grompp -f minim.mdp -c 1AKI_solv_ions.gro -p topol.top -o em.tpr
mdrun -v -deffnm em
energy -f em.edr -o potential.xvg
grompp -f nvt.mdp -c em.gro -r em.gro -p topol.top -o nvt.tpr
energy -f nvt.edr -o temperature.xvg
grompp -f npt.mdp -c nvt.gro -r nvt.gro -t nvt.cpt -p topol.top -o npt.tpr
energy -f npt.edr -o pressure.xvg
energy -f npt.edr -o density.xvg
grompp -f md.mdp -c npt.gro -t npt.cpt -p topol.top -o md_0_1.tpr
mdrun -deffnm md_0_1
trjconv -s md_0_1.tpr -f md_0_1.xtc -o md_0_1_noPBC.xtc -pbc mol -center
rms -s md_0_1.tpr -f md_0_1_noPBC.xtc -o rmsd.xvg -tu ns
rms -s em.tpr -f md_0_1_noPBC.xtc -o rmsd_xtal.xvg -tu ns
```

Now execute the **mdrun** and the trajectory, energies and confirmations of the output file. We can observe several graphs of energies, nvt equilibration, npt equilibration and trajectory.

Conclusion:

GROMACS is preferred over other Molecular Dynamics Simulation Software packages because it is a versatile package to perform molecular dynamics. It is user friendly, with topologies, parameter files, and error messages written in clear text format. It provides extremely high performance compared to all other programs. There is a lot of consistency checking. There is no scripting language. Comes with a large selection of flexible tools for trajectory analysis

References:

1.

2. Allen, M. P. (2004). Introduction to molecular dynamics simulation. *Computational soft matter: from synthetic polymers to proteins*, *23*, 1-28.
3. Feller, S. E., Zhang, Y., Pastor, R. W., & Brooks, B. R. (1995). Constant pressure molecular dynamics simulation: the Langevin piston method. *The Journal of chemical physics*, *103*(11), 4613-4621.

4. Brooks, B. R., Brooks III, C. L., Mackerell Jr, A. D., Nilsson, L., Petrella, R. J., Roux, B., ... & Caflisch, A. (2009). CHARMM: the biomolecular simulation program. *Journal of computational chemistry*, *30*(10), 1545-1614.
5. Wong, A. K., & Goscinski, A. M. (2012). A VMD plugin for NAMD simulations on Amazon EC2. *Procedia Computer Science*, *9*, 136-145.
6. Wong, A. K. L., & Goscinski, A. M. (2012). The design and implementation of the VMD plugin for NAMD simulations on the Amazon cloud. *International Journal of Cloud Computing and Services Science*, *1*(4), 155.

8. Murail, S. (2017). Simulation of ligand binding to membrane proteins. In *Membrane Protein Structure and Function Characterization* (pp. 359-381). Humana Press, New York, NY.
9. Gumbart, J., Roux, B., & Chipot, C. (2011). Protein: ligand standard binding free energies: A tutorial for alchemical and geometrical transformations. *University of Illinois at Urbana-Champaign, USA*.
10. Makarewicz, T., & Kaźmierkiewicz, R. (2013). Molecular dynamics simulation by GROMACS using GUI plugin for Pymol.