Lab III: MD II & Analysis – Linux & OS X

In this lab you will:

1) Add water, ions, and set up and minimize a solvated protein system
2) Learn how to create fixed atom files and systematically minimize a new structure.
3) Set up NVT and NPT molecular dynamics simulations including PBC/PME.
4) Analyze output files from MD runs, print and plot thermodynamic and structural variables.

**Working directory: $HOME/Desktop/ccp-sas_2015/exercises/lab3**

NOTE: in this lab, the ">" symbol means that you are in an active terminal window and what follows this symbol is what you type. Make sure you know what directory you are in (> pwd) and what files are in your current directory (> ls). The symbol "vmd>" means you type what follows inside the VMD console (not a Linux terminal). One final note, in problem 5 (d&e) you will run a plotting program called Gnuplot. Once started, it presents the user with a prompt "gnuplot>". Therefore, you type what it past this prompt in the active Gnuplot program.

## 1. Add water and ions to solvate the protein hydrolase.

In Lab II yesterday you built several water, protein, and complex systems. You minimized one (or more structures) in vacuum. In this problem, you will build a more applicable model system for the protein hydrolase ("protein_II") from Lab II. You will use VMD GUI tools to add a solvation box, then you will use another VMD GUI to add ions, and you will type some TCL commands in the VMD console to measure the resulting periodic box size. You will need these numbers for subsequent problems.

Directories:
solvate_and_add_ions/

Files: you will generate these!

Procedure:

**a)** Change to the solvate_and_add_ions/ directory and the copy the complete directory for "protein_II" that you generated yesterday to this location.

**i.** **>cp –Rp $HOME/Desktop/ccp-sas_2015/exercises/lab2/psfgen/protein_II .**

**ii.** NOTE the trailing period at the end of the command above.

**iii.** NOTE2: in the copy command (cp) the –Rp tells the copy program to copy all of the files in the directory and that you maintain ownership of the file. The trailing period means "put it HERE".

**b)** Open *VMD* and load the new_hydrolase.psf and new_hydrolase.pdb files.

**c)** Using the pull-down menu, choose the Extensions -> Modeling -> Add Solvation Box item.

**d)** Uncheck the 'Waterbox Only' box (if it is checked).

**e)** Check the Input PSF: and Input PDB: fields to make sure that they are pointing at the full path and filenames for the "new_hydrolase" system.

**f)** In the Output field, name the output file "solvated_proteinII".

**g)** Make sure that the 'Use Molecular Dimensions' box is selected (checked).

**h)** Enter the value "12.0" for each of the Box Padding for all min and max values of x, y, and z.

**i)** Hit the 'Solvate' button at the bottom.

**j)** The resulting files (solvated_proteinII.psf and solvated_proteinII.pdb) may be created in different directories depending on the system and the place you invoked *VMD*. Move them to your current location if there were not generated in-space. For example, if they were generated on your home directory:

i. **> mv ~/solvated_proteinII* .**

ii. NOTE the trailing period at the end of the command above.

iii. NOTE2: The "~/" is a Linux short-cut to your home directory.

iv. NOTE3: The "*" symbol merely tells the "mv" program to move all files with that start with ionized_solvated_proteinII and have any additional characters following (including NO additional characters). The "*" is called a wild-card. The trailing period means "put it HERE".

**k)** Now, using the pull-down menu, choose the Extensions -> Modeling -> Add Ions box item.

**l)** Input the newly created solvated_proteinII (.psf & .pdb) files (Browse if needed).

**m)** Enter "ionized_solvated_proteinII" in the Output prefix: field.

**n)** Enter "ION" in the Segment name of placed ions: field.

**o)** Make sure that the 'Only neutralize with NaCl' box is checked.

**p)** Look at Figure 1 below to check your input values. Then, click the Autoionize button at the bottom.

**q)** Just like solvated_proteinII, the resulting files (ionized_solvated_proteinII.psf and ions_solvated_proteinII.pdb) may be created in different directories depending on the system and the place you invoked *VMD*. Move them to your current location if there were not generated in-place. For example, if they were generated on your home directory:

.
> **mv ~/ionized_solvated_proteinII* .**

**r)** Using the "2" button (or the Mouse -> Label -> Bonds from the VMD GUI), measure the x, y, z dimensions of the system. Write down your values.

**s)** Enter the commands listed below into the *VMD console* to use TCL commands to measure the dimensions of your system. Write down your values. Type the text AFTER the "vmd > " prompt below.

```
vmd > set everyone [atomselect top all]
vmd > measure minmax $everyone
```

**t)** Compare your measured values in step (q) above to those obtained in step (r).

**u)** Visualize your final system using VMD. Use VMD "Representations" to depict the protein as "New Cartoon", the water molecules by "Points" and the ions by VDW spheres.

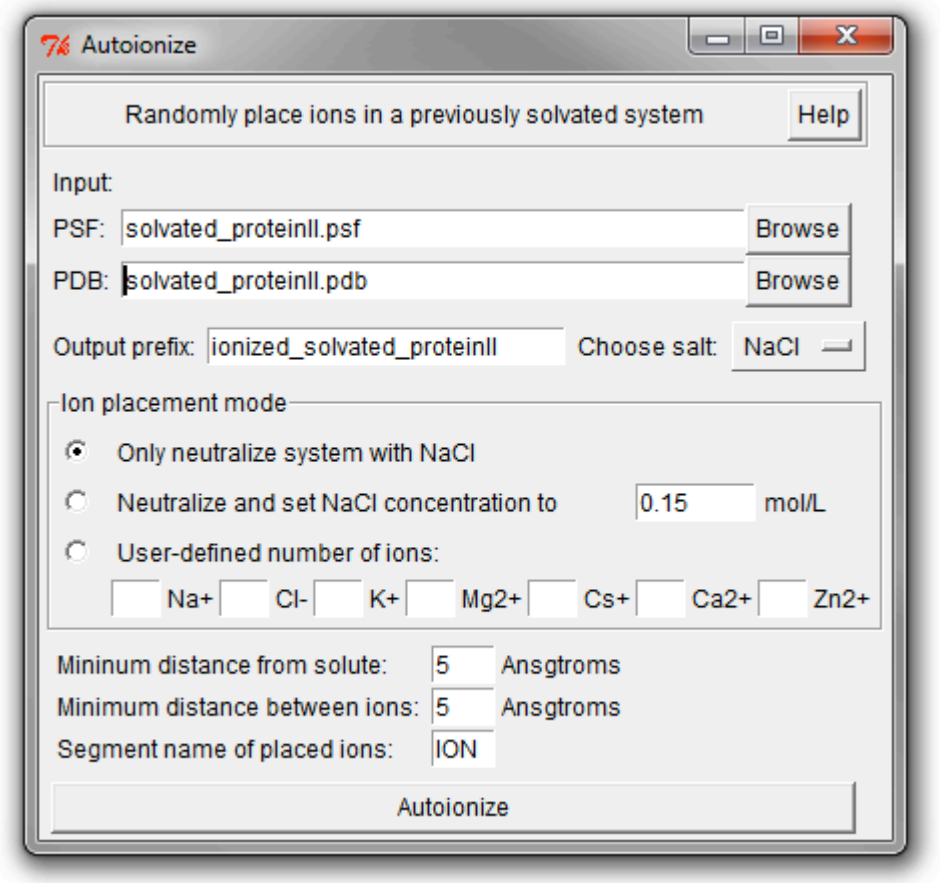


Figure 1. Autoionize input GUI.

**2. Create fixed atom files to be used to systematically minimize the solvated hydrolase system.**

In the lecture, we went through the process of generating fixed atom files so that we could keep parts of our system rigid while we relax everything else. In this problem you will generate ONE of the same fixed atom files. The others will be provided to you due to time limitations. Of course, it would be beneficial if you can create all three fixed atom files yourself. If you wish to do this, refer to the slides in the talk for the correct TCL commands to do this. You will use these fixed atom PDB files in the remainder of this lab.

Directories:

namd/

minimization_II/

output/

Files: there are files in both the minimization_II and minimization_II/output directories.

minimization_II/

smin0

minimization_II/output/

fixed_backbone_atoms.pdb

fixed_secondary_structure.pdb

Procedure:

*a)* Change to the namd/minimization_II/ directory.

*b)* Copy the final files you created in (1) to the namd/minimization_II/output directory.

**i.** > **`cp ../../solvate_and_add_ions/ionized*  output/`**

**ii.** NOTE the "../../" at the beginning of the statement above means go up to directory levels (of course one set of "../" would tell the copy command to go up a single level.

**iii.** NOTE2: the "*" is the wild-card and the final "output/" tells to put the files into the local directory named "output/". Note that we don't have a trailing period here since we are not putting the files "HERE".

*c)* If VMD is still open, close the program and *restart VMD* by clicking the icon on the desktop.

*d)* Load the ionized_solvated_proteinII.psf file and then load the coordinates into the same molecule by loading the ionized_solvated_proteinII.pdb file.

*e)* Create a new fixed atom filed named: fixed_protein_atoms.pdb using the commands listed in the figure below *into the VMD console.*

With the **ionized_solvated_proteinII.pdb** file open in VMD . . . at the console type:

```
vmd > set everyone [atomselect top all]
atomselect32

vmd > $everyone set beta 0.00

vmd > set sel1 [atomselect top "protein"]
atomselect33

vmd > $sel1 set beta 1.00

vmd > $everyone writepdb fixed_protein_atoms.pdb

Info) Opened coordinate file fixed_protein_atoms.pdb for writing.
Info) Finished with coordinate file fixed_protein_atoms.pdb.
```

Figure 2. *NAMD TCL commands to create the fixed_protein_atoms.pdb file. Enter these commands into your VMD console (not the Linux terminal).*

Examine the resulting output fixed_protein_atoms.pdb file. Confirm that the correct atoms have either 1.00 or 0.00 in the beta column.

*f)* Make sure that the newly created fixed_protein_atoms.pdb file is in the namd/minimization_II/output directory (move it there if it is created somewhere else such as "~/" or "namd/minimization/" etc.)

*g)* List the minimization_II/output directory and inspect the other fixed_atom files that are there (fixed_backbone_atoms.pdb and fixed_secondary_structure.pdb). Make sure that you understand what these files are doing (if needed, open them up as new molecules and color using the BETA value).

*h)* If you have time, you can try to create your own versions of fixed_backbone_atoms.pdb and fixed_secondary_structure.pdb files. The slides in the MD-II lecture cover have the exact commands to execute.

**3. Set up and run a series of restrained minimization runs of the ionized_solvated_proteinII system using the fixed atom files created in step (2).**

In the lecture we showed how you sequentially relax a newly built system in order to retain the original biological structure. In this problem, you will use the fixed atom files (one or more of which you generated in (2) above) in a series of minimization runs to relax your newly built system. You will write (and edit) a set of input files for the minimization runs and run the jobs using NAMD. The final file from this sequence of minimization runs (smin3.coor) will be used in problem 4 below.

Directories:

namd/
minimization_II/
output/

Files: there are files in both the minimization_II and minimization_II/output directories.

minimization_II/
smin0

minimization_II/output/
fixed_backbone_atoms.pdb
fixed_secondary_structure.pdb
fixed_protein_atoms.pdb (created in step (2e) above)
ionized_solvated_proteinII.psf (copied here in step (2b) above)
ionized_solvated_proteinII.pdb (copied here in step (2b) above)

Procedure:

a) Change to the namd/minimization_II/ directory.
b) Inspect the NAMD input file named smin0.
   i. Inspect the cellbasisvector lines, how do they compare to the dimensions of the system as measured in step 2?
c) Start the minimization by typing:
   a. > **namd2 smin0 >& smin0.out &**
   b. ALTERNATIVELY, since we have dual-core CPUs, you could run this job in parallel using the command:
      i. > **charmrun +p 2 $(PATH_TO_NAMD)/namd2 smin0 > smin0.out &**
      or
      **namd2 +p 2 smin0 > smin0.out &**
      See lab2 for an explanation of the two commands.
      ii. DON'T DO IT BOTH WAYS AT THE SAME TIME. Choose one. In either case the output is saved to a file called "smin0.out".
d) Observe the progress of the minimization by typing
   a. > **tail -90f smin0.out**
e) When the run is finished, hit "control c" to exit out of the "tail -90f …" process.
f) Copy the smin0 file and to a new file called smin1
   a. > **cp smin0 smin1**

g) Edit the smin1 file to use the output coordinates from the last run and change the fixed atom file to use the fixed_backbone_atoms.pdb file. Also change the name of the output file (see lecture slides).
h) Minimize the system using similar commands as used above (c-e).
i) Copy the smin1 file and to a new file called smin2
    **a. > cp smin1 smin2**
j) Edit the smin2 file to use the output coordinates from the last run and change the fixed atom file to use the fixed_secondary_structure.pdb file. Also change the name of the output file (see lecture slides).
k) Minimize the system using similar commands as used above (c-e). You need to change the actual names of the input and output files accordingly!!!
l) Copy the smin2 file and to a new file called smin3
    **a. > cp smin2 smin3**
m) Edit the smin3 file to use the output coordinates from the last run and remove the fixed atom lines. This final minimization run will let all atoms move. Also change the name of the output file (see lecture slides). Make sure that the output file for this run is smin3 (you will need smin3.coor for the problems described below).
n) Minimize the system using similar commands as used above (c-e).
o) Examine the final structure(s) with VMD. NOTE: the final coordinate file you should visualize is in a file called smin3.coor in the output/ directory. Since we save the coordinates as text files (i.e. not using the binary output option) NAMD writes these files with the ".coor" suffix, but it is really a PDB file. When you open this file with VMD it will think that the input file type is NAMDBIN. You have to use the pull down menu to tell VMD that it is a PDB file. Alternatively, you can merely make a copy of the smin3.coor file so that it has the ".pdb" suffix that VMD will associate with a PDB formatted file.

## 4. NAMD NVT and NPT molecular dynamics.

In this problem you will take your minimized hydrolase system and start a molecular dynamics trajectory in the NVT ensemble and then continue the same run in the NPT ensemble. NVT merely means that the number of particles (atoms), volume, and temperature will be held fixed, while NPT means that the number of particles (atoms), pressure and temperature will be held fixed. We will perform these simulations at 1 atmosphere at 300 K.

The NPT MD run is set up to run for 1 million time-steps (i.e. $10^6$ femtoseconds = 1 nanoseconds). Your laptop will not be able to finish this run in time for our analysis lab tomorrow, so while you can let this job continue for the duration of the course, we'll use completed NVT & NPT runs in the next problem (i.e. we'll give you the finished files). You should probably cancel this run before you start other assignments in the days ahead as they are computationally demanding as well.

Directories:
    namd/
        dynamics/
                output/

Files: there are files in the dynamics directory. You will move files created in steps (2) and (3) above to the dynamics/output directory.

dynamics/
        dyn0
        dyn1

dynamics/output/

**v)** Change to the namd/dynamics directory.

**w)** Copy the files ionized_solvated_proteinII.psf and smin3.coor into the output/ directory. You created these files in part 1o and 3n above.

**i.** > **cp ../minimization_II/output/ionized_solvated_proteinII.psf output/**

**ii.** > **cp ../minimization_II/output/smin3.coor output/**

**x)** Inspect the NAMD NVT dynamics input file (dyn0). Write down the name of the output files that will be created.

**y)** Start the dynamics run using your input file.

**i.** > **namd2 dyn0 >& dyn0.out &**

**ii.** ALTERNATIVELY, since we have dual-core CPUs, you could run this job in parallel using the command:

**charmrun +p 2 $(PATH_TO_NAMD)/namd2 dyn0 > dyn0.out &**

**or**

namd2 +p 2 dyn0 > dyn0.out &

See lab2 for an explanation of the two commands.

DON'T DO IT BOTH WAYS AT THE SAME TIME. Choose one. In either case the output is saved to a file called "dyn0.out".

**iii.** THIS RUN WILL TAKE SOME TIME, so proceed to section 5 below before completing the rest of this section.

**z)** Examine the output file and resulting final structure using VMD.

**aa)** Inspect the NAMD NPT dynamics input file (dyn1). Write down the name of the output files that will be created.

**bb)** The dyn1 script will take a really long time to run (at least a day). You can start the run to convince yourself that it works using the similar commands to thonse you used for dyn0 (pressing Ctrl and C at the same time will quit the run). However, a better way than tying up you laptop for a day is to run it using SASSIE-web.

**cc)**

**i.** Login to your SASSIE-web account.

**ii.** From the main menu select 'Simulate'

**iii.** Click the button to enter the 'Energy Minimization' module (note: this doesn't only perform energy minimization but allows you to specify a NAMD input file).

**iv.**

**v.** Enter the appropriate files from the current and minimization_II directories so that your screen looks like the following Figure.

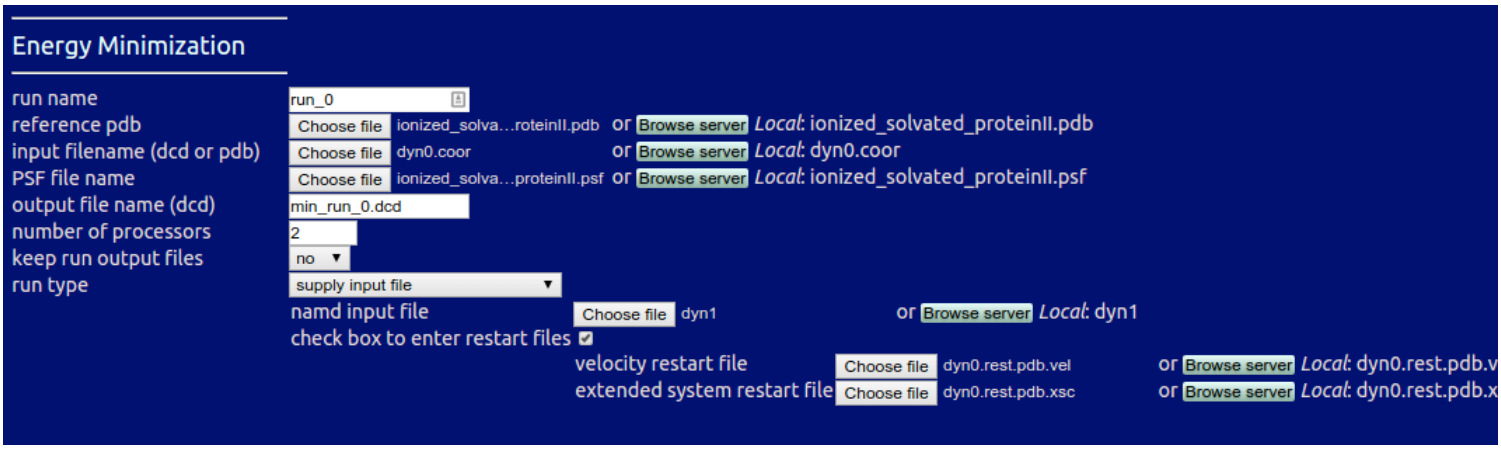


Figure 3: Input for the SASSIE-web NPT molecular dynamics run

**vi.** Click 'Submit'

**vii.** This run will take a long time but your laptop remains free. In the section 5 we will analyse data from a similar run that has been run previously. You can reconnect to this job at any point from the Job Management screen.

**dd)** Examine the output file (dyn0.out) and resulting final structure from the NVT run (dyn0.coor) using VMD.

## 5. Analysis of MD runs.

In this problem, you will analyze the output from a completed 1 nanosecond trajectory of the hydrolase system you set up in this lab. This lab will introduce you to the way you can "source" TCL scripts (code, commands, etc.) in the VMD console to perform simple analyses.

Directories:

namd/

finished_dynamics/

output/

Files: there are files in both the finished_dynamics/ and finished_dyanmics/output/ directories.

finished_dynamics/

dyn0
dyn1
dyn0.out
dyn1.out
dyn1_no_wrap
dyn1_no_wrap.out
rmsd.tcl
rmsd_no_correction.tcl
residue_rmsd.tcl
namdstats.tcl

finished_dynamics/output/

dyn0.coor
dyn0.dcd
dyn0.rest.pdb.coor
dyn0.rest.pdb.vel

dyn0.rest.pdb.xsc
dyn0.vel
dyn0.xsc
dyn0_nw.coor
dyn0_nw.dcd
dyn0_nw.rest.pdb.coor
dyn0_nw.rest.pdb.vel
dyn0_nw.rest.pdb.xsc
dyn0_nw.vel
dyn0_nw.xsc
dyn1.coor
dyn1.dcd
dyn1.rest.pdb.coor
dyn1.rest.pdb.vel
dyn1.rest.pdb.xsc
dyn1.vel
dyn1.xsc
dyn1.xst
dyn1_nw.coor
dyn1_nw.dcd
dyn1_nw.rest.pdb.coor
dyn1_nw.rest.pdb.vel
dyn1.rest.pdb.xsc
dyn1_nw.vel
dyn1_nw.xsc
dyn1_nw.xst
ionized_solvated_proteinII.psf
smin3.coor

Some notes: Essentially there are four sets of files in the finished_dynamics and finished_dynamics/output directories.

NVT, 10 ps with wrapped coordinates → dyn0
NVT, 10 ps without wrapped coordinates → dyn0_nw

NPT, 1 ns with wrapped coordinates → dyn1
NPT, 1 ns without wrapped coordinates → dyn1_nw

a) Change to the namd/finished_dynamics directory.
b) First, let's look at the averages by analyzing the output file from the 1 ns NPT run (with wrapped coordinates).
    - Open *VMD* and change the working directory to the namd/finished_dynamics directory by typing

vmd > **`cd $HOME/Desktop/ccp-sas/exercises/lab3/namd/finished_dynamics`**

    - At the vmd prompt type: vmd > **`source namdstats.tcl`**
    - At the vmd prompt type: vmd > **`data_avg dyn1.out`**

c) Now, let's create files that contain only some of the thermodynamics variables. Your laptop is probably too slow to run more than one of these … pick one of the following. You will have to alter your plot command in section (d) below if you chose something other than TEMP.
    - vmd > **`data_time TEMP dyn1.out`**
    - vmd> **`data_time VDW dyn1.out`**
    - vmd> **`data_time  ELECT dyn1.out`**

d) View these files by using gnuplot. In a Linux terminal that is in the namd/finished_dynamics/ directory , type
   - > **gnuplot**
   - Then within the gnuplot command window, you can type, for example,
   - gnuplot> **plot "TEMP.dat" w l**
   - NOTE that the gnuplot program has a "gnuplot>" prompt!

*e)* Now, let's take a look at the root-mean square deviation of the protein structure over the 1 ns trajectory.
   - *Open a new VMD session* and change the working directory to the namd/finished_dynamics directory by typing

vmd > **cd $HOME/Desktop/ccp-sas_2015/exercises/lab3/namd/finished_dynamics**

   - Load the output/ionized_solvated_proteinII.psf and output/dyn1.dcd into a new molecule
   - At the vmd prompt type:
     i. vmd> **source rmsd_no_correction.tcl**
     ii. vmd> **source rmsd.tcl**
   - This will create two new files with rmsd data: rmsd_no_correction.dat and rmsd.dat.
   - View the results using Gnuplot in a similar manner as was done in section (d) above.

f) Close VMD and reopen, then load output/ionized_solvated_proteinII.psf and output/dyn1_nw.dcd into a new molecule

g) See Figure 4 below for the commands that you should type into the vmd console (not a Linux terminal).

h) To color the protein with RMSD values for each individual amino acid follow averaged over the entire 1 ns trajectory we are going to use a script supplied by VMD to calculate the RMSD for each alpha carbon atom. Then you will use the VMD GUI to color the molecule through the "User" selection. The TCL script keeps the results in a special place (called "User") so that the calculated values can be shown on the molecule in question. This is similar to what we have done with the BETA field, but by using the "User" capability we do not have to create a special PDB file with BETA values that correspond to the RMSD of each particular amino acid. You access the "User" data by merely selecting the "User" option in the "Coloring Method" pull down menu of the "Representation" pop-up menu in VMD. The higher version of VMD may list "User" option under "Trajectory" in the "Coloring Method" pull down menu.

i) Since the RMSD calculation is for the protein part only, you may want to exclude the solvent in VMD representation. In "Graphical Representations", type "protein" in the "Selected Atoms" box.

j) VMD has an odd default coloring scheme (counter intuitive). If you want to color the molecule so that "blue" corresponds to "cold" or "rigid" portions of the molecule and RED corresponds to "hot" or "flexible" portions, then you can reconfigure the color scale settings
   - In the VMD GUI, use the pull-down menu to get the color scale settings by clicking on Graphics → Colors, then choose the "Color Scale" tab. Then change the default setting from RWB to BWR.

k) Repeat the analysis for (b) through (f) using the 1 ns NPT run (without wrapped coordinates). If time allows, repeat (b) through (f) for the 10 ps NVT run(s).

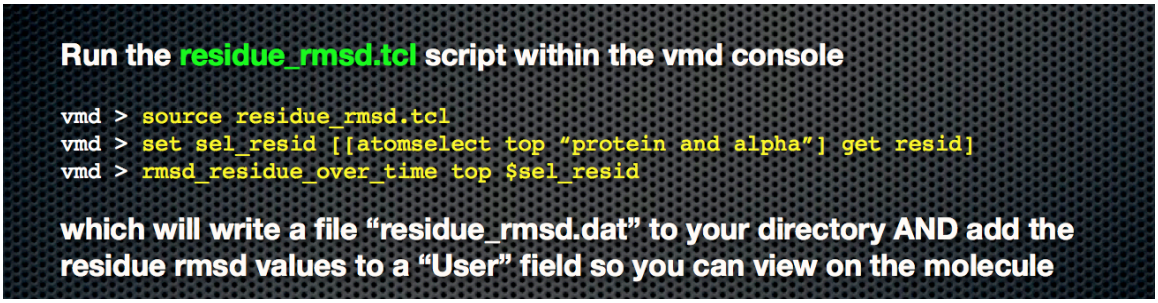


Figure 4. *VMD TCL commands to load the residue_rmsd.tcl script, select a subset of atoms and then run the rmsd_residue_over_time function on your subset of atoms to calculate the RMSD over the entire trajectory.*

Lab III: MD II & Analysis - Windows

In this lab you will:

1) Add water, ions, and set up and minimize a solvated protein system
2) Learn how to create fixed atom files and systematically minimize a new structure.
3) Set up NVT and NPT molecular dynamics simulations including PBC/PME.
4) Analyze output files from MD runs, print and plot thermodynamic and structural variables.

**Working directory: C:\Users\yourname\Desktop\ccp-sas_2015\exercises\lab3**

NOTE: in this lab, the ">" symbol means that you are in an active command prompt window and what follows this symbol is what you type. Make sure you know what directory you are in (it is displayed before the ">" prompt) and what files are in your current directory (> dir). The symbol "vmd>" means you type what follows inside the VMD console (not the Command Prompt). One final note, in problem 5 (d&e) you will run a plotting program called Gnuplot. Once started, it presents the user with a prompt "gnuplot>". Therefore, you type what it past this prompt in the active Gnuplot program.

## 1. Add water and ions to solvate the protein hydrolase.

In Lab II yesterday you built several water, protein, and complex systems. You minimized one (or more structures) in vacuum. In this problem, you will build a more applicable model system for the protein hydrolase ("protein_II") from Lab II. You will use VMD GUI tools to add a solvation box, then you will use another VMD GUI to add ions, and you will type some TCL commands in the VMD console to measure the resulting periodic box size. You will need these numbers for subsequent problems.

Directories:
solvate_and_add_ions\

Files: you will generate these!

Procedure:

a) Change to the solvate_and_add_ions directory and the copy the complete directory for "protein_II" that you generated yesterday to this location.

**i.**
```
>robocopy C:\Users\yourname\Desktop\CCP-SAS_2015ccp-
sas_2015\exercises\lab2\psfgen\protein_II protein_II
/S
```

***ii.*** NOTE: The /S flag to robocopy copies subdirectories. If your version of robocopy does not correctly copy the expected files replace the /S flag with /s /i.

b) *Open VMD* and change to the working directory ($HOME\embo_labs\exercises\lab3\solvate_and_add_ions\)
c) Load the new_hydrolase.psf and new_hydrolase.pdb files.
d) Using the pull-down menu, choose the Extensions -> Modeling -> Add Solvation Box item.
e) Uncheck the 'Waterbox Only' box.
f) Check the Input PSF: and Input PDB: fields to make sure that they are pointing at the full path and filenames for the "new_hydrolase" system.
g) In the Output field, name the output file "solvated_proteinII".
h) Make sure that the 'Use Molecular Dimensions' box is selected (checked).
i) Enter the value "12.0" for each of the Box Padding for all min and max values of x, y, and z.
j) Hit the 'Solvate' button at the bottom.

k) The resulting files (solvated_proteinII.psf and solvated_proteinII.pdb) will be created in the VMD working directory. If you placed them somewhere other than Desktop\ccp-sas_2015\exercises\lab3\solvate_and_add_ions\ then move them to your current location:

   **i.** `move <location>\solvated_proteinII* .`

   **ii.** <location> should be replaced with the full path to the location in which the files were saved (e.g. C:\\Users\yourname\Desktop\exercises\stuff).

   **iii.** NOTE the trailing period at the end of the command above.

   **iv.** NOTE: The "*" symbol merely tells the "move" program to move all files that start with ionized_solvated_proteinII and have any additional characters following (including NO additional characters). The "*" is called a wild-card. The trailing period means "put it HERE".

l) Now, using the pull-down menu, choose the Extensions -> Modeling -> Add Ions box item.
m) Input the newly created solvated_proteinII (.psf & .pdb) files (Browse if needed).
n) Enter "ionized_solvated_proteinII" in the Output prefix: field.
o) Enter "ION" in the Segment name of placed ions: field.
p) Make sure that the 'Only neutralize with NaCl' box is checked.
q) Look at Figure 1 below to check your input values. Then, click the Autoionize button at the bottom.
r) The resulting files (ionized_solvated_proteinII.psf and ions_solvated_proteinII.pdb) will again be created in the VMD working directory. If necessary move them to the Desktop\ccp-sas_2015\exercises\lab3\solvate_and_add_ions\ directory with a similar command to before (replacing solvated_proteinII with ionized_solvated_proteinII).

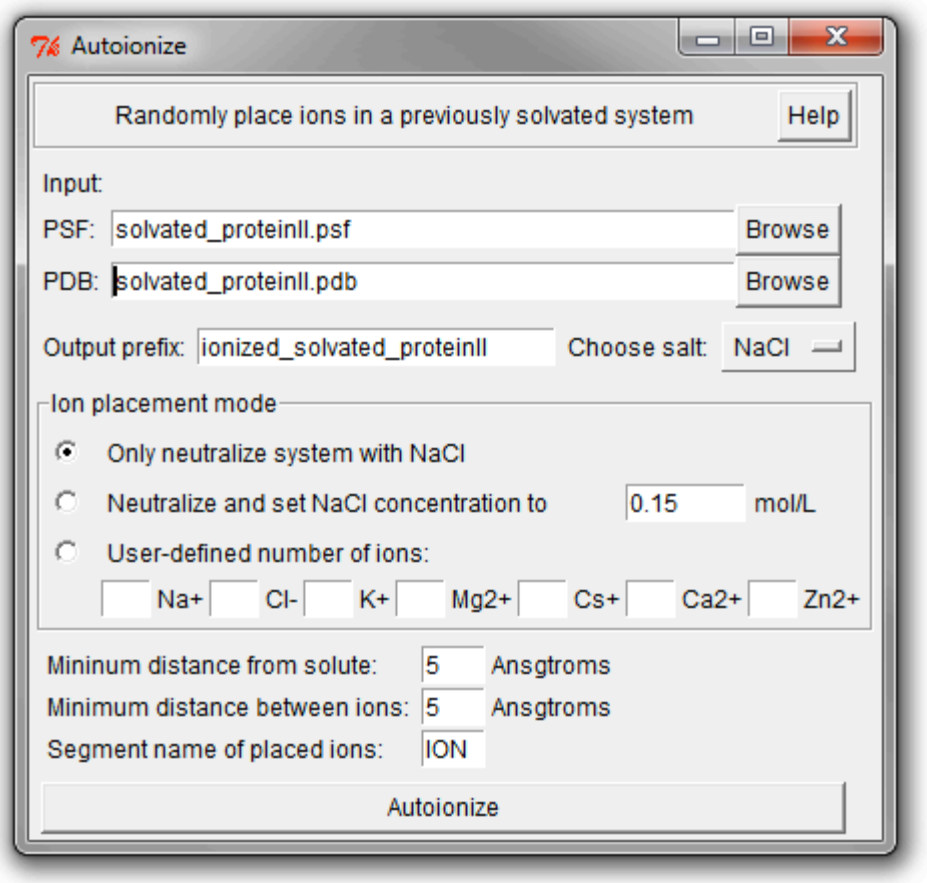


Figure 1: Autoionize input GUI.

## 2. Create fixed atom files to be used to systematically minimize the solvated hydrolase system.

In the lecture, we went through the process of generating fixed atom files so that we could keep parts of our system rigid while we relax everything else. In this problem you will generate ONE of the same fixed atom files. The others will be provided to you due to time limitations. Of course, it would be beneficial if you can create all three fixed atom files yourself. If you wish to do this, refer to the slides in the talk for the correct TCL commands to do this. You will use these fixed atom PDB files in the remainder of this lab.

Directories:

namd\
minimization_II\
output\

Files: there are files in both the minimization_II and minimization_II\output directories.

minimization_II\
smin0

minimization_II\output\
fixed_backbone_atoms.pdb
fixed_secondary_structure.pdb

Procedure:

*a)* Change to the namd\minimization_II\ directory.
*b)* Copy the final files you created in (1) to the namd\minimization_II\output directory.
   - **i.** `> copy ..\..\solvate_and_add_ions\ionized*  output\`
   - **ii.** NOTE the "..\..\" at the beginning of the statement above means go up to directory levels (of course one set of "..\" would tell the copy command to go up a single level.
   - **iii.** NOTE2: the "*" is the wild-card and the final "output\" tells to put the files into the local directory named "output\".
*c)* If VMD is still open, close the program and *restart VMD* by clicking the icon on the desktop.
*d)* Load the ionized_solvated_proteinII.psf file and then load the coordinates into the same molecule by loading the ionized_solvated_proteinII.pdb file.
*e)* Create a new fixed atom filed named: fixed_protein_atoms.pdb using the commands listed in the figure below *into the VMD console.*

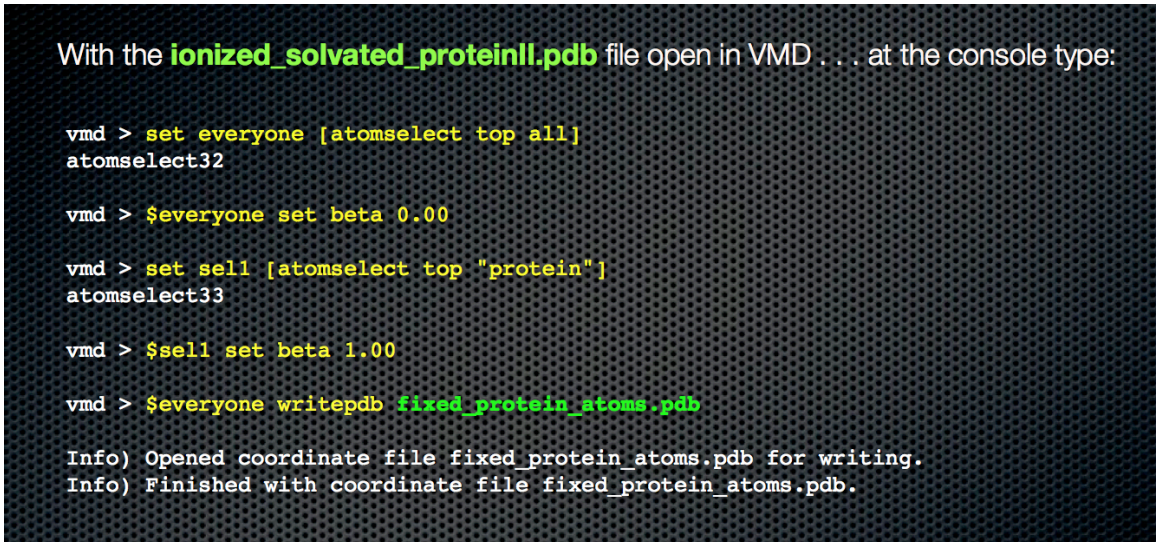


Figure 2. *NAMD TCL commands to create the fixed_protein_atoms.pdb file. Enter these commands into your VMD console (not the Linux terminal).*

Examine the resulting output fixed_protein_atoms.pdb file. Confirm that the correct atoms have either 1.00 or 0.00 in the beta column.

*f)* Make sure that the newly created fixed_protein_atoms.pdb file is in the namd\minimization_II\output directory (move it there if it is created somewhere else, such as "namd\minimization\").

*g)* List the minimization_II\output directory and inspect the other fixed_atom files that are there (fixed_backbone_atoms.pdb and fixed_secondary_structure.pdb). Make sure that you understand what these files are doing (if needed, open them up as new molecules and color using the BETA value).

*h)* If you have time, you can try to create your own versions of fixed_backbone_atoms.pdb and fixed_secondary_structure.pdb files. The slides in the MD-II lecture cover have the exact commands to execute.

**3. Set up and run a series of restrained minimization runs of the ionized_solvated_proteinII system using the fixed atom files created in step (2).**

In the lecture we showed how you sequentially relax a newly built system in order to retain the original biological structure. In this problem, you will use the fixed atom files (one or more of which you generated in (2) above) in a series of minimization runs to relax your newly built system. You will write (and edit) a set of input files for the minimization runs and run the jobs using NAMD. The final file from this sequence of minimization runs (smin3.coor) will be used in problem 4 below.

Directories:

namd\
minimization_II\
output\

Files: there are files in both the minimization_II and minimization_II\output directories.

minimization_II\
smin0

minimization_II\output\

fixed_backbone_atoms.pdb
fixed_secondary_structure.pdb
fixed_protein_atoms.pdb (created in step (2e) above)
ionized_solvated_proteinII.psf (copied here in step (2b) above)
ionized_solvated_proteinII.pdb (copied here in step (2b) above)

Procedure:

a) Change to the namd\minimization_II\ directory.
b) Inspect the NAMD input file named smin0.
    i. Inspect the cellbasisvector lines, how do they compare to the dimensions of the system as measured in step 2?
c) Start the minimization by typing:
    **a.** `> START "" namd2 smin0 >& smin0.out`
    **b.** ALTERNATIVELY, since we have dual-core CPUs, you could run this job in parallel using the command:
        **i.** `> START "" namd2 +p 2 min0 > min0.out`
        **ii.** DON'T DO IT BOTH WAYS AT THE SAME TIME. Choose one. In either case the output is saved to a file called "smin0.out".
d) Observe the progress of the minimization by typing :
   `> tail -90 smin0.out`
e) epeat the command as the run progresses
f) Copy the smin0 file and to a new file called smin1
    **a.** `> copy smin0 smin1`
g) Edit the smin1 file to use the output coordinates from the last run and change the fixed atom file to use the fixed_backbone_atoms.pdb file. Also change the name of the output file (see lecture slides).
h) Minimize the system using similar commands as used above (c-e).
i) Copy the smin1 file and to a new file called smin2
    **a.** `> copy smin1 smin2`
j) Edit the smin2 file to use the output coordinates from the last run and change the fixed atom file to use the fixed_secondary_structure.pdb file. Also change the name of the output file (see lecture slides).
k) Minimize the system using similar commands as used above (c-e). You need to change the actual names of the input and output files accordingly!!!
l) Copy the smin2 file and to a new file called smin3
    **a.** `> copy smin2 smin3`
m) Edit the smin3 file to use the output coordinates from the last run and remove the fixed atom lines. This final minimization run will let all atoms move. Also change the name of the output file (see lecture slides). Make sure that the output file for this run is smin3 (you will need smin3.coor for the problems described below).
n) Minimize the system using similar commands as used above (c-e).
o) Examine the final structure(s) with VMD. NOTE: the final coordinate file you should visualize is in a file called smin3.coor in the output/ directory. Since we save the coordinates as text files (i.e. not using the binary output option) NAMD writes these files with the ".coor" suffix, but it is really a PDB file. When you open this file with VMD it will think that the input file type is NAMDBIN. You have to use the pull down menu to tell VMD that it is a PDB file. Alternatively, you can merely make a copy of the smin3.coor file so that it has the ".pdb" suffix that VMD will associate with a PDB formatted file.

## 4. NAMD NVT and NPT molecular dynamics.

In this problem you will take your minimized hydrolase system and start a molecular dynamics trajectory in the NVT ensemble and then continue the same run in the NPT ensemble. NVT merely means that the number of particles (atoms), volume, and temperature will be held fixed, while NPT means that the number of particles (atoms), pressure and temperature will be held fixed. We will perform these simulations at 1 atmosphere at 300 K.

The NPT MD run is set up to run for 1 million time-steps (i.e. $10^6$ femtoseconds = 1 nanoseconds). Your laptop will not be able to finish this run in time for our analysis lab tomorrow, so while you can let this job continue for the duration of the course, we'll use completed NVT & NPT runs in the next problem (i.e. we'll give you the finished files). You should probably cancel this run before you start other assignments in the days ahead as they are computationally demanding as well.

Directories:

namd\
dynamics\
output\

Files: there are files in the dynamics directory. You will move files created in steps (2) and (3) above to the dynamics\output directory.

dynamics\
dyn0
dyn1

dynamics\output\

s) Change to the namd\dynamics directory.
t) Copy the files ionized_solvated_proteinII.psf and smin3.coor into the output\directory. You created these files in part 1o and 3n above.

**i.** `> cp ../minimization_II/output/ionized_solvated_proteinII.psf output/`

**ii.** `> cp ../minimization_II/output/smin3.coor output/`

u) Inspect the NAMD NVT dynamics input file (dyn0). Write down the name of the output files that will be created.
v) Start the dynamics run using your input file.

**i.** `> START "" namd2 dyn0 >& dyn0.out`

**ii.** ALTERNATIVELY, you could run this job in parallel using the command:

**i.** `> START "" namd2 +p 2 dyn0 > dyn0.out`

**ii.** DON'T DO IT BOTH WAYS AT THE SAME TIME. Choose one. In either case the output is saved to a file called "dyn0.out".

**iii.** THIS RUN WILL TAKE SOME TIME, so proceed to section 5 below before completing the rest of this section.

w) Examine the output file and resulting final structure using VMD.
x) Inspect the NAMD NPT dynamics input file (dyn1). Write down the name of the output files that will be created.
y) The dyn1 script will take a really long time to run (at least a day). You can start the run to convince yourself that it works using the similar commands to thonse you used for dyn0 (pressing Ctrl and C at the same time will quit the run). However, a better way than tying up you laptop for a day is to run it using SASSIE-web.

z) The dyn1 script will take a really long time to run (at least a day). You can start the run to convince yourself that it works using the similar commands to thonse you used for dyn0 (pressing Ctrl and C at the same time will quit the run). However, a better way than tying up you laptop for a day is to run it using SASSIE-web.

**i.** Login to your SASSIE-web account.

**ii.** From the main menu select 'Simulate'

**iii.** Click the button to enter the 'Energy Minimization' module (note: this doesn't only perform energy minimization but allows you to specify a NAMD input file).

**iv.** Select 'supply input file' from the 'run type' listbox and check the 'checkbox to enter restart files' checkbox.

**v.** Enter the appropriate files from the current and minimization_II directories as input such that your input page looks like that shown in Figure 3.

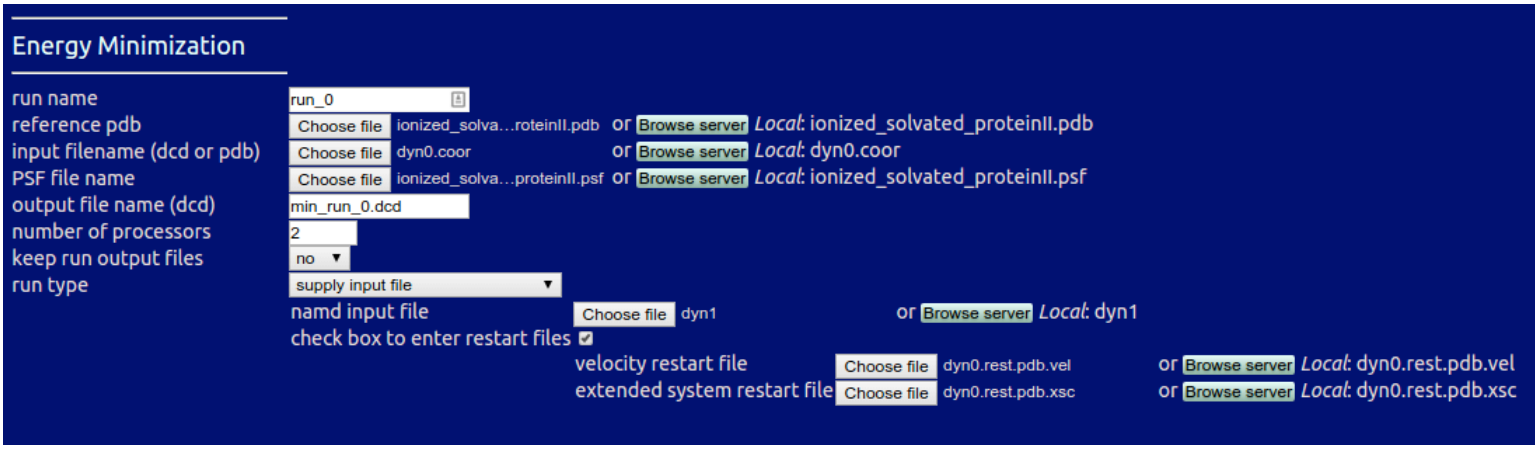


Figure 3: Input for the SASSIE-web NPT molecular dynamics run

**vi.** Click 'Submit'

This run will take a long time but your laptop remains free. In the section 5 we will analyse data from a similar run that has been run previously. You can reconnect to this job at any point from the Job Management screen.

aa) Examine the output file (dyn0.out) and resulting final structure from the NVT run (dyn0.coor) using VMD.

## 5. Analysis of MD runs.

In this problem, you will analyze the output from a completed 1 nanosecond trajectory of the hydrolase system you set up in this lab. This lab will introduce you to the way you can "source" TCL scripts (code, commands, etc.) in the VMD console to perform simple analyses.

Directories:

namd\
finished_dynamics\
output\

Files: there are files in both the finished_dynamics\ and finished_dyanmics\output\ directories.

finished_dynamics\
dyn0
dyn1
dyn0.out

dyn1.out
dyn1_no_wrap
dyn1_no_wrap.out
rmsd.tcl
rmsd_no_correction.tcl
residue_rmsd.tcl
namdstats.tcl

finished_dynamics\output\
dyn0.coor
dyn0.dcd
dyn0.rest.pdb.coor
dyn0.rest.pdb.vel
dyn0.rest.pdb.xsc
dyn0.vel
dyn0.xsc
dyn0_nw.coor
dyn0_nw.dcd
dyn0_nw.rest.pdb.coor
dyn0_nw.rest.pdb.vel
dyn0_nw.rest.pdb.xsc
dyn0_nw.vel
dyn0_nw.xsc
dyn1.coor
dyn1.dcd
dyn1.rest.pdb.coor
dyn1.rest.pdb.vel
dyn1.rest.pdb.xsc
dyn1.vel
dyn1.xsc
dyn1.xst
dyn1_nw.coor
dyn1_nw.dcd
dyn1_nw.rest.pdb.coor
dyn1_nw.rest.pdb.vel
dyn1.rest.pdb.xsc
dyn1_nw.vel
dyn1_nw.xsc
dyn1_nw.xst
ionized_solvated_proteinII.psf
smin3.coor

Some notes: Essentially there are four sets of files in the finished_dynamics and finished_dynamics\output directories.

NVT, 10 ps with wrapped coordinates → dyn0
NVT, 10 ps without wrapped coordinates → dyn0_nw

NPT, 1 ns with wrapped coordinates → dyn1
NPT, 1 ns without wrapped coordinates → dyn1_nw

a) Change to the namd\finished_dynamics directory.
*b)* First, let's look at the averages by analyzing the output file from the 1 ns NPT run (with wrapped coordinates).
- *Open VMD* and change the working directory to the namd\finished_dynamics directory by typing something like (depending on the location from which you start):

vmd> **cd CCP-SAS\exercises\lab3\namd\finished_dynamics**

- At the vmd prompt type: vmd > **source namdstats.tcl**
- At the vmd prompt type: vmd > **data_avg dyn1.out**

c) Now, let's create files that contain only some of the thermodynamics variables. Your laptop is probably too slow to run more than one of these … pick one of the following. You will have to alter your plot command in section (d) below if you chose something other than TEMP.

- vmd> **data_time TEMP dyn1.out**
- vmd> **data_time VDW dyn1.out**
- vmd> **data_time ELECT dyn1.out**

d) View these files by using gnuplot. In a Linux terminal that is in the namd\finished_dynamics\ directory , type

- > **gnuplot**
- Then within the gnuplot command window, you can type, for example,
- gnuplot> **plot "TEMP.dat" w l**
- NOTE that the gnuplot program has a "gnuplot>" prompt!

*e)* Now, let's take a look at the root-mean square deviation of the protein structure over the 1 ns trajectory.

- *Open a new VMD session* and change the working directory to the namd\finished_dynamics directory by typing something like (depending on the location from which you start):

vmd > **cd ccp-sas_2015\exercises\lab3\namd\finished_dynamics**

- Load the output\ionized_solvated_proteinII.psf and output/dyn1.dcd into a new molecule
- At the vmd prompt type:
  - **i.** vmd> **source rmsd_no_correction.tcl**
  - **ii.** vmd> **source rmsd.tcl**
- This will create two new files with rmsd data: rmsd_no_correction.dat and rmsd.dat.
- View the results using Gnuplot in a similar manner as was done in section (d) above.

f) Close VMD and reopen, then load output\ionized_solvated_proteinII.psf and output/dyn1_nw.dcd into a new molecule

g) To color the protein with RMSD values for each individual amino acid follow averaged over the entire 1 ns trajectory we are going to use a script supplied by VMD to calculate the RMSD for each alpha carbon atom. Then you will use the VMD GUI to color the molecule through the "User" selection. The TCL script keeps the results in a special place (called "User") so that the calculated values can be shown on the molecule in question. This is similar to what we have done with the BETA field, but by using the "User" capability we do not have to create a special PDB file with BETA values that correspond to the RMSD of each particular amino acid. You access the "User" data by merely selecting the "User" option in the "Coloring Method" pull down menu of the "Representation" pop-up menu in VMD.

h) See Figure 4 below for the commands that you should type into the vmd console (not the Command Prompt).

i) Repeat the analysis for (b) through (f) using the 1 ns NPT run (without wrapped coordinates). If time allows, repeat (b) through (f) for the 10 ps NVT run(s).

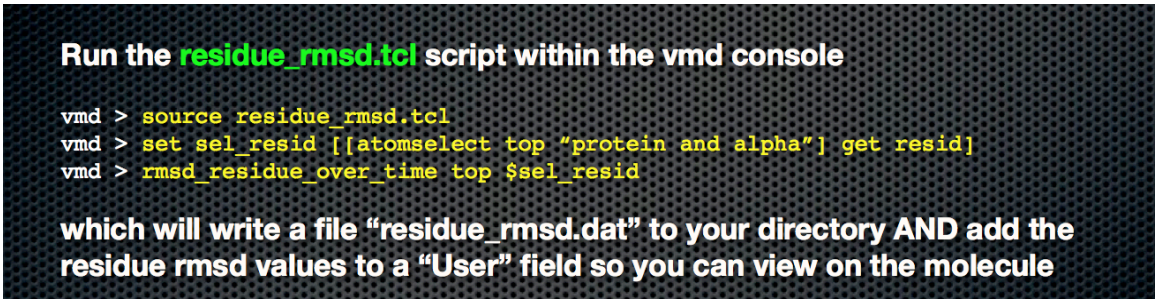


Figure 4. *VMD TCL commands to load the residue_rmsd.tcl script, select a subset of atoms and then run the rmsd_residue_over_time function on your subset of atoms to calculate the RMSD over the entire trajectory.*

j) VMD has an odd default coloring scheme (counter intuitive). If you want to color the molecule so that "blue" corresponds to "cold" or "rigid" portions of the molecule and RED corresponds to "hot" or "flexible" portions, then you can reconfigure the color scale settings

- In the VMD GUI, use the pull-down menu to get the color scale settings by clicking on Graphics → Colors, then choose the "Color Scale" tab. Then change the default setting from RWB to BWR.