Tutorial: MD simulations of proteins in NAMD + VMD

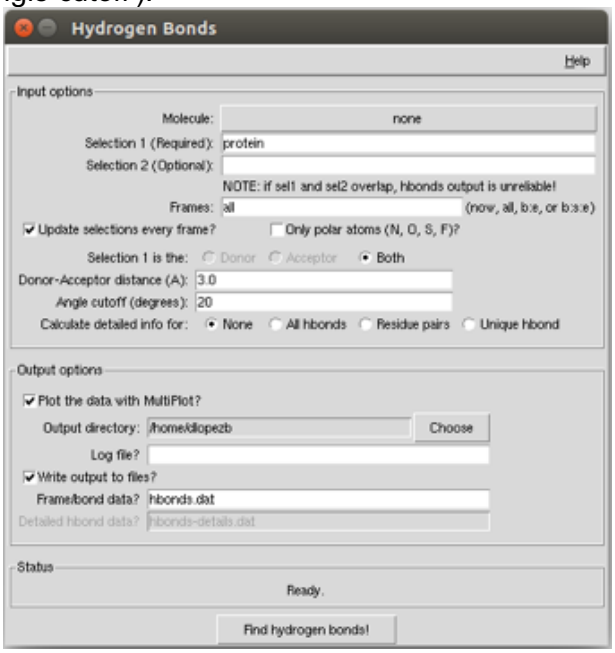
How to set up a MD simulation in NAMD using VMD

1. Load the .pdb file of your protein in VMD
2. Create the .psf file using the CHARMM force field
    a. Go to Extension > Modeling > Automatic PSF builder.
    b. Add files containing the force field parameters, if needed.
        i. Step 1: Load all the FF files
        ii. Step 2: Click on "Guess chains"
        iii. Step 3: Click on "Create chains"
    c. Close the window
3. Calculate the dimensions of the protein to determine the dimension of the water box size. For that, you will need to open the TK Console and type the following commands:
    a. *set set [atomselect top all]* ;#these selects all the atoms in the protein
    b. *measure minmax $sel* ;#this will determine the minimum and maximum dimensions of your protein. It will return two points {xmin_protein, ymin_protein, zmin_protein} and {xmax_protein, ymax_protein, zmax_protein}
    c. With those two points, you can calculate the location and dimensions of the water box. Normally, you should make your water box extend at least the cutoff distance for long range interactions in each dimension. So as an example, your water box dimensions could be:
        i. xmin_water_box = xmin_protein - cutoff
        ii. ymin_water_box = ymin_protein - cutoff
        iii. zmin_water_box = zmin_protein - cutoff
        iv. xmax_water_box = xmax_protein + cutoff
        v. ymax_water_box = ymax_protein + cutoff
        vi. zmax_water_box = zmax_protein + cutoff
4. Solvate your protein, using the {xmin_water_box, ymin_water_box, zmin_water_box} and {xmax_water_box, ymax_water_box, zmax_water_box} dimensions. To solvate, just go to Extensions > Modeling > Add Solvation Box and provide the box dimensions you just calculated.
5. You can create an ionized version of your water box if you need (e.g., to ensure charge neutrality in your system) by going to Extensions > Modeling > Add ions
    a. If you want your system to be just charge neutral, click on "Only neutralize the system with NaCl"
    b. If you want to add a certain salt concentration, click on "Neutralize and set NaCl concentration to XXXX mol/L"
    c. Click on Autoionize
6. Your final system will be automatically stored as ionized.psf and ionized.pdb.
7. You can check the final dimensions of the system, which should correspond to those calculated in step 3c.
    a. *set set [atomselect top all]* ; # select all the atoms in your system (protein+water)
    b. *measure minmax $sel* ;#this will give you the coordinates of the water box {xmin_water_box, ymin_ water_box, zmin_ water_box} and {xmax_ water_box, ymax_ water_box, zmax_ water_box}. You will use those to calculate Δx, Δy and Δz of your simulation box. You will also need to calculate the central point of your box:
        i. central point x: xmin_box+ Δx/2
        ii. central point y: ymin_box+ Δy/2
        iii. central point z: zmin_box+ Δz/2

8. Add that information to the NAMD input file.
9. Adjust the parameters you want to have for the simulation (timestep, cutoff, temperature, pressure, minimization timesteps, run timesteps, output frequency, etc), and run NAMD.

How to analyze the output of MD simulations in NAMD using VMD

1. Copy the file sec_struct.tcl from this repository into the folder where all the dcd files are contained
2. If you have multiple .dcd files from your trajectory, combine them into a single file (e.g., merged.dcd). You can do that with several VMD scripts, like catdcd.
3. Open the .psf/.pdb files of your starting system in VMD
    - *mol load psf system.psf*
    - *mol load pdb system.pdb*
4. Load the file containing all the dcd trajectories: right-click on the protein name in VMD Main, and select “Load data into Molecule”.
5. Extract a .dcd trajectory file just from the protein that will be used for analysis
    - *set sel [atomselect top protein]*
    - *$sel writepsf protein.psf*
    - *$sel writepdb protein.pdb*
    - *animate write dcd protein.dcd sel $sel*
6. Open only the .psf/.pdb protein files in VMD
7. Right-click on the protein name in VMD Main, and select “Load data into Molecule” to load protein.dcd
8. Calculate the secondary structure using the TK Console to output the file sec_struct_ext.txt that will contain the secondary structure of each frame from your protein.dcd trajectory
    - *source sec_struct.tcl*
9. Calculate the number of intraprotein hydrogen bonds (Extensions > Analysis > Hydrogen bonds). Select your desired geometric criteria ('Donor-Acceptor distance' and 'Angle cutoff').



10. Calculate the number of solvent accessible surface area (SASA) and the radius of gyration of your protein using the TK Console to output the sasa_rgyr.txt file that will contain those parameters for each frame of your .dcd file.
    - *source sasa_rgyr.tcl*

11. Calculate the number of number of molecules in the hydration layer from your protein using the TK Console.
    - *source hydration_shell.tcl*
12. You can also obtain other parameters from your protein, such as the root mean square displacement of the atomic positions (using the RMSD Trajectory Tool in VMD), or the energy (using the NAMD Energy tool). All these tools allow to export .txt files containing the values of those parameters on each frame of your .dcd trajectory.
13. At the end of this process, you will several .txt files containing the information extracted from the trajectory of your protein after MD simulations. Import them into an Excel document and analyze the trends for each system.

Acknowledgements: this project has received funding from the European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie grant agreement SUPERB 892369. More information can be found in *dlopezbarreiro.github.io*