Connor Aghili

5/30/16

Questions for MD and NAMD tutorial video

1. In your own words, when are molecular dynamics simulations useful? What type of information can we hope to learn from them?

Molecular dynamics (MD) is useful for comparing known properties to experimental data (NMR spectroscopy, X-ray crystallography etc.), describe motion of molecules and used to refine molecular structures. MD is helpful for interpreting the results of experiments with many moving parts (macromolecules). MD can also be used to describe interactions on the atomic level which cannot be observed directly. However, MD is limited to the force fields themselves, computing power and the parameter sets (such as neglecting contributing factors or simplifying factors like Van der Waals forces with Lennard – Jones potentials).

Information we hope to gain from MD is refined models on pervious experimental data, atomic interactions which cannot be observed, understanding of macromolecular motion, molecular processes, etc.

2. What are the four types of interatomic interactions that we can model with a potential / energy function?

$$U(\vec{R}) = \underbrace{\sum_{bonds} k_i^{bond}(r_i - r_0)^2}_{U_{bond}} + \underbrace{\sum_{angles} k_i^{angle}(\theta_i - \theta_0)^2}_{U_{angle}} +$$

$$\underbrace{\sum_{dihedrals} k_i^{dihe}[1 + \cos(n_i\phi_i + \delta_i)]}_{U_{dihedral}} +$$

$$\underbrace{\sum_i \sum_{j \neq i} 4\epsilon_{ij}\left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right] + \sum_i \sum_{j \neq i} \frac{q_i q_j}{\epsilon r_{ij}}}_{U_{nonbond}}$$

Figure 1: Potential Energy Function

Bonding: oscillations about the equilibrium bond length

Angle: oscillations of 3 atoms about an equilibrium bond angle

Dihedral: torsional rotation of 4 atoms about a central bond

Non-bonding: coulombic interactions (electrostatics) and van der Waals interactions (Lenard-Jones)

Bonds are harmonic potentials (springs), angles are also harmonic, and dihedrals are periodic since they rotate freely. Some models have another term called the improper term which may be added to enforce planarity. Other models add more terms to get more accurate results of energetics of proteins such as the CMAP term. Some models also need a quantum mechanical treatment.

3. What are the timescales for bond-stretching, elastic vibration, and rotation of side chains?

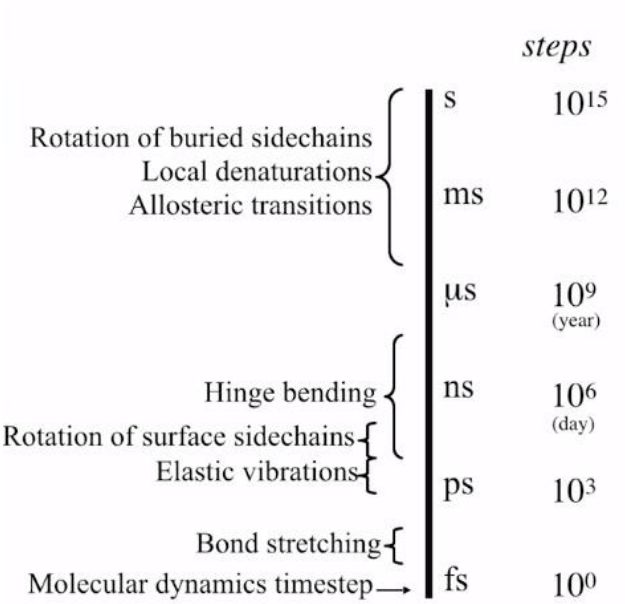

Figure 2: Time Steps; These values are based on atomistic modeling

The time step depends on the process that needs to be simulated. MD timesteps are on a femtosecond scale (fs). It takes about 10 frames or femtoseconds to describe molecular motion properly. Bond stretching is between 10 -10^3 fs. Elastic vibrations are about 10^3 fs and rotation of surface sidechain takes about 10^5 fs. For coarse-grain the timestep can increase to 20fs or even microseconds. Time steps are crucial, if the integration is not fine enough the system can crash.

4. What are the steps you need to perform a typical molecular dynamics simulation?

The Protein structure file (psf) stores topology and information about the system except coordinates. The coordinates/ positions come from the pdb file. The system reads a pdb file and generates a psf file using topology files. This gives the information one needs where the molecule is concerned. Then another file defines the rules and the force fields involved. A configuration file is used to specify the timestep, temperature, steps of minimization, etc. These files set everything up to get ready for calculations.

One needs to run a minimization step in order to bring the system down to a temperature of 0. This step removes a lot of the large forces that may tamper with the calculations. Pdb files are usually predefined as minimized however if there are modifications one will need to minimize it before heating. After heating one equilibrates the system with user defined settings. One then simulates the equilibrated system to analyze the dynamics.

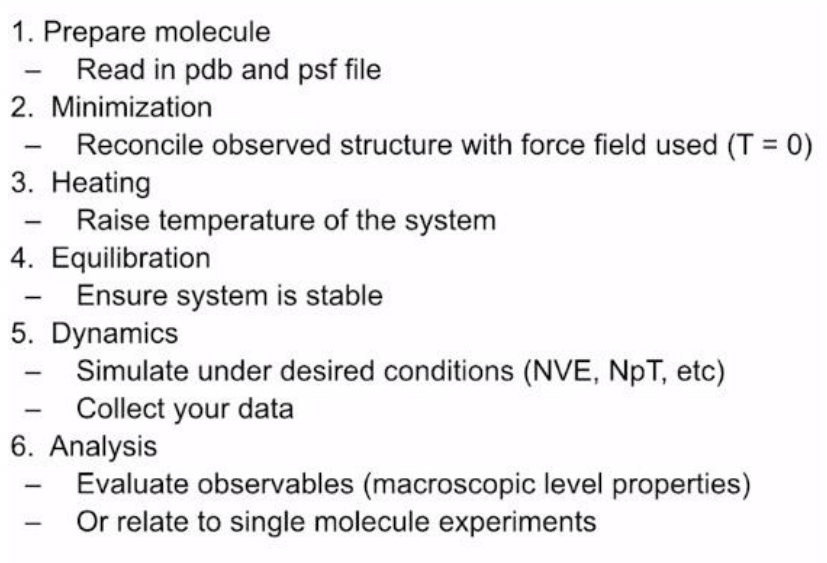

Figure 3: Steps that need to be taken using the CHARMM language

5. Why do we need to take into account periodic boundary conditions in MD simulations and how do we correct artificial periodic boundary effects?

We need to take into account periodic boundary conditions because many molecular systems are in membrane or water. We need the right density of water around the boundaries. So, to correct these boundaries the simulation is essentially copied around itself. An example of this is simulating periodic boundary conditions in a rectangular box with molecules in it. One will generate exact copies of the box around the original box. The copies follow the central unit's design exactly. Since these units are touching each other you have the correct density of water. Things may move from one of the boxes in to the central unit while items move out of the central unit. There are other arrangements besides boxes such as hexagonal, spherical, etc.

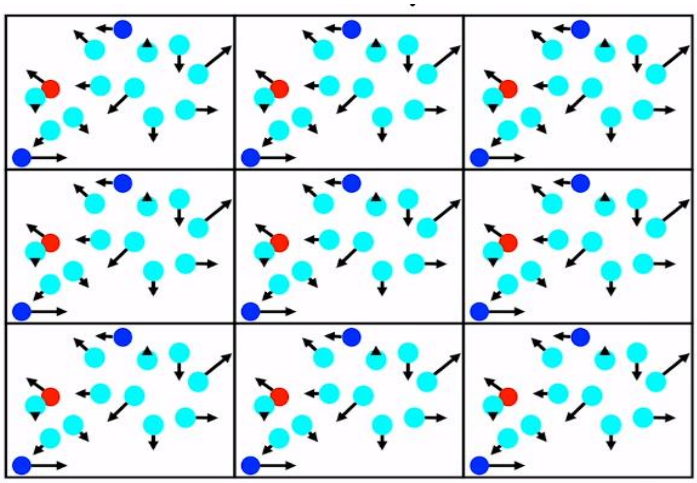

Figure 4: The box example

6. What are the different parts of a NAMD input file?

NAMD has great parallel performance (on many cores) compared to CHARMM. It also can use tcl commands. The different parts of the configuration file are: the structure file (.psf), coordinate file (.pdf), parameter file for constants, the variables such as temperature, and tcl variable commands (define loops and conditions). The simulation in the lecture discusses generating random velocities from given temperatures. Using the same example one can also input files from a previous run using a binary file for coordinates and a binary file for velocities.

7. What is FFTK and why is it useful?

ffTK stands for Force Field Toolkit. It is a set of tools that aid in the development of force field parameters which includes charges, bonds, angles and dihedrals. ffTK assumes that all quantum mechanical (QM) data is generated using Gaussian.

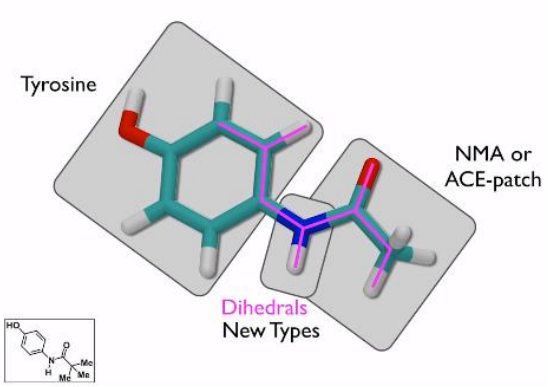

Figure 5: Acetaminophen

This program can fragment the data into parts for example Acetaminophen, it breaks it into Tyrosine, "a new tape" and NMA. You need to assign parameters to include the new type.

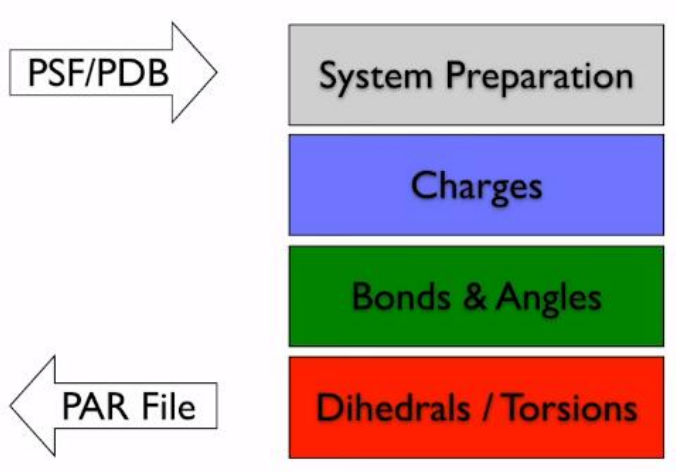

Figure 6: General Steps involved in ffTK

In general, you put together a molecule. You get the charges, bonds, angels, dihedrals and then put everything into the parameter file. However, it is much more complicated than that ffTK has various steps involved. The first step is to prepare the system and find the missing parameters. QM calculations are done during this step for geometric optimization. These calculations then generate input files. A Gaussian file is then created where the distance and the energies are calculated. The output files form all the pervious steps are then extracted and used for charge optimization. A hessian calculation is then done to calculate bond and angle parameters. Then more QM calculations are done, where you rotate your dihedrals and calculate their energies. Finally, everything is then put together to fit the model.

The toolkit is very useful. It can be used in VMD while your molecule is loaded. This system is useful because it generates default values, reads all the log files that you generated, easy to use interface and then optimizes it for you. This program makes calculating force fields significantly easier.

8. Why is coarse-grained modeling a useful approach? What type of information can we hope to learn from such simulations?

Coase-grained modeling is useful because it groups atoms together if you are not interested in the atomistic approach. You can use these groups or beads to represent the system. The number of particles are greatly reduced and the vibration of the particles are at a much lower frequency therefore getting a larger timestep. Essentially these models can reach longer time and or length scales. The main thing with membranes is that you need to have a distinction between a hydrophobic area and a polar area. The distinction, shape and size are the only things you need to reproduce these structures which can be done easily with coarse grain modeling. These models are computationally less demanding and help us learn and understand the interplay between driving forces.

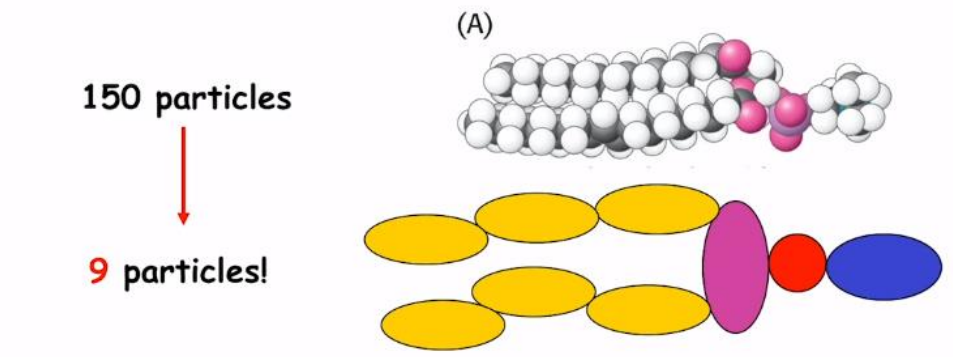

Figure 7: coarse graining of a lipoid

An application of this type of modeling is the MARTINI model. The MARTINI model has shown how to combine structural accuracy and thermodynamic reproducibility in a systematic way.

Another way to apply coarse graining is ultra – coarse graining[2] which keeps the system constant but coarse grains the parts which are irrelevant during the simulation. This provides precise results and lowers computational cost.

REFRENCES:

[1]Tajkhorshid, Emad. (2015, September 15). *Introduction to VMD and NAMD - Emad Tajkhorshid.* Retrieved from https://www.youtube.com/watch?v=VdfeUSB3VZA

[2] *Dama, J. F., Sinitskiy, A. V., Mccullagh, M., Weare, J., Roux, B., Dinner, A. R., & Voth, G. A. (2013). The Theory of Ultra-Coarse-Graining. 1. General Principles. J. Chem. Theory Comput. Journal of Chemical Theory and Computation, 9(5), 2466-2480. doi:10.1021/ct4000444*