Basics of membrane-protein interactions using molecular dynamics with NAMD, VMD, and CHARMM-GUI

Michael J. Hallock and Taras V. Pogorelov

School of Chemical Sciences, University of Illinois at Urbana-Champaign

August 17, 2016, Edition 1.1

This tutorial is designed to introduce the preparation, modeling, and analysis of membrane-protein system with molecular dynamics (MD) using *NAMD*, VMD, and *CHARMM-GUI.* The system of interest will be the outer membrane phospholipase A from *E. coli* (OmpLA, PDB ID 1QD5, Snijder HJ, …, Dijkstra BW, Nature, 401, 717, 1999). In a recent experimental study, the Fleming Lab (Johns Hopkins) used spontaneous folding of OmpLA into the membrane to design a hydrophobicity scale (Moon PC and Fleming KG, PNAS 108, 10174, 2011). Note: NAMD and VMD are developed at the NIH Center for Macromolecular Modeling hosted at the University of Illinois at Urbana-Champaign (http://www.ks.uiuc.edu). CHARMM-GUI is developed by the Im Lab at the University of Kansas (http://www.charmm-gui.org)

Estimated time to complete this tutorial is 1.5 hr.

Outline: we will use the SCS computer cluster, triton, to model the systems.

Step 1: prepare membrane-OmpLA structure using CHARMM-GUI web-server;

Step 2: download the system to local desktop computer;

Step 3: submit the system to the local computer cluster;

Step 4: when job is completed, analyze the results.

Software/Hardware

Triton: *NAMD2*. Personal computer: Terminal (Mac OSX) or xserver (e.g. *Xming* for a Windows machine), and CHARMM-GUI web server.
This tutorial can be completed on iMacs in the SCS VizLab (151 Noyes Lab) or with personal computers with Windows or Mac OS. Small modifications might be needed when used on other computers.

1. Prepare membrane-OmpLA structure

Open the CHARMM-GUI website (http://www.charmm-gui.org/) and go to Input Generator->Membrane Builder -> Bilayer Builder.

1.1 CHARMM-GUI: load, analyze, and position PDB

● Protein/Membrane System

Download PDB File: 1qd5 Download Source: OPM

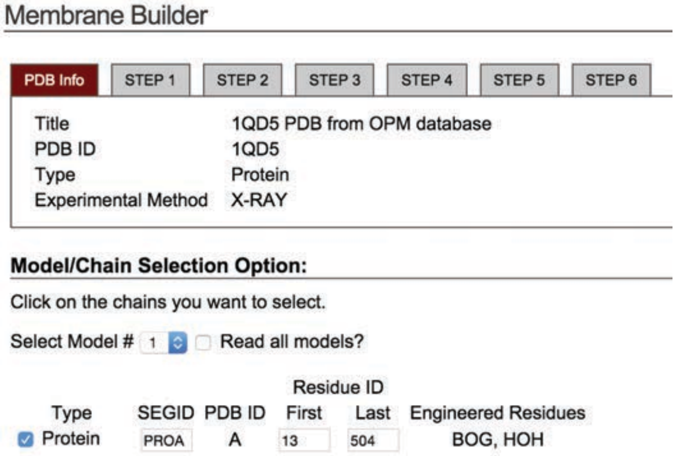
Upload PDB File: No file chosen

PDB Format: ○ RCSB ○ CHARMM

○ Membrane Only System

Figure 1 Protein-membrane window

Enter the PDB ID, keep OPM as the selection and proceed to the next step (Fig.1).



Title	1QD5 PDB from OPM database
PDB ID	1QD5
Type	Protein
Experimental Method	X-RAY

Model/Chain Selection Option:

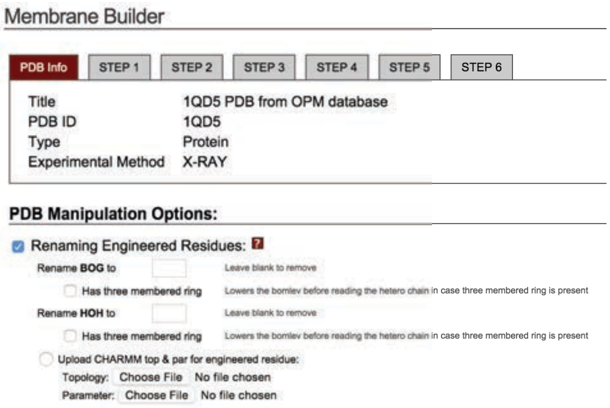
Click on the chains you want to select.

Select Model # 1 ☐ Read all models?

Type	SEGID	PDB ID	Residue ID First	Residue ID Last	Engineered Residues
☑ Protein	PROA	A	13	504	BOG, HOH

Figure 2 Select models available in the PDB

Note that CHARMM-GUI automatically located the protein in the OPM database (http://opm.phar.umich.edu/), detected the residue IDs and engineered residues. OPM (Orientations of Proteins in Membranes) is a depository of precalculated membrane protein positions within a membrane and provides a good starting point. Here one can change the name of the segment to be used if needed. Leave the default selections intact for now. CHARMM-GUI analyzes of the provided PDB and offers possible modifications. Proceeed to the next step: PDB Manipulation Optioins (Fig. 3)



Title	1QD5 PDB from OPM database
PDB ID	1QD5
Type	Protein
Experimental Method	X-RAY

PDB Manipulation Options:

☑ Renaming Engineered Residues:

Rename **BOG** to ____ Leave blank to remove

☐ Has three membered ring Lowers the bomlev before reading the hetero chain in case three membered ring is present

Rename **HOH** to ____ Leave blank to remove

☐ Has three membered ring Lowers the bomlev before reading the hetero chain in case three membered ring is present

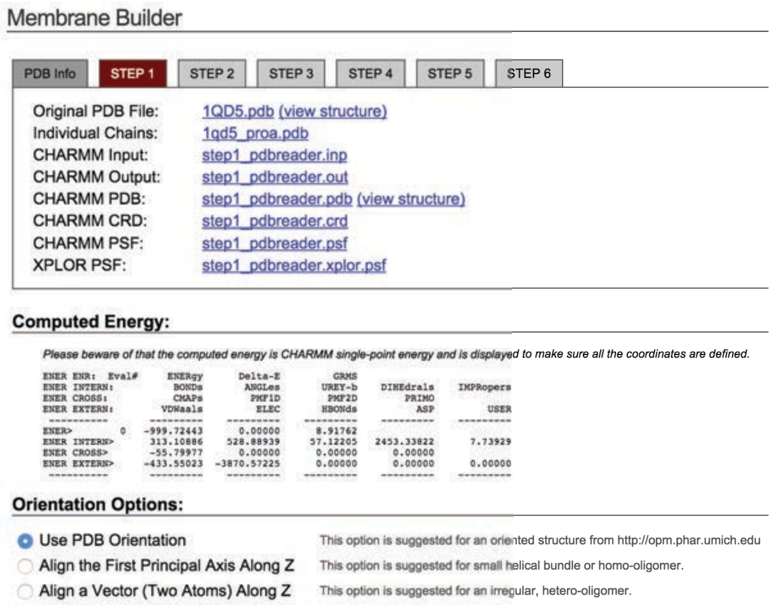
○ Upload CHARMM top & par for engineered residue:

Topology: No file chosen

Parameter: No file chosen

Figure 3 PDB manipulations

CHARMM-GUI detected non-protein components (BOG and HOH). Leave cells blank (Fig. 3) so octyl glucoside (n-octyl-β-D-glucoside – BOG) and oxygen atoms of water (HOH) will be removed. Proceed to the next step.



Original PDB File:	1QD5.pdb (view structure)
Individual Chains:	1qd5_proa.pdb
CHARMM Input:	step1_pdbreader.inp
CHARMM Output:	step1_pdbreader.out
CHARMM PDB:	step1_pdbreader.pdb (view structure)
CHARMM CRD:	step1_pdbreader.crd
CHARMM PSF:	step1_pdbreader.psf
XPLOR PSF:	step1_pdbreader.xplor.psf

Computed Energy:

Please beware of that the computed energy is CHARMM single-point energy and is displayed to make sure all the coordinates are defined.

```
ENER ENR:  Eval#     ENERgy      Delta-E         GRMS
ENER INTERN:          BONDs       ANGLes       UREY-b    DIHEdrals    IMPRopers
ENER CROSS:           CMAPs        PMF1D        PMF2D        PRIMO
ENER EXTERN:        VDWaals         ELEC       HBONds          ASP         USER
 ----------       ---------    ---------    ---------    ---------    ---------
ENER>        0   -999.72443      0.00000      8.91762
ENER INTERN>      313.10886    528.88939     57.12205   2453.33822      7.73929
ENER CROSS>       -55.79977      0.00000      0.00000      0.00000
ENER EXTERN>     -433.55023  -3870.57225      0.00000      0.00000      0.00000
 ----------       ---------    ---------    ---------    ---------    ---------
```

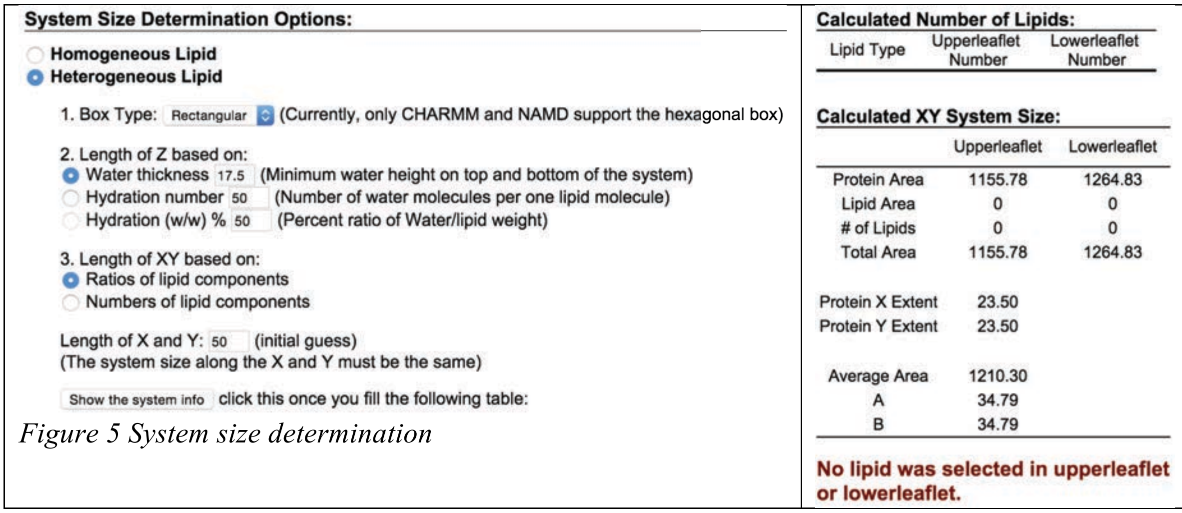
Orientation Options:

● Use PDB Orientation — This option is suggested for an oriented structure from http://opm.phar.umich.edu
○ Align the First Principal Axis Along Z — This option is suggested for small helical bundle or homo-oligomer.
○ Align a Vector (Two Atoms) Along Z — This option is suggested for an irregular, hetero-oligomer.

Figure 4 PDB orientation

CHARMM-GUI provides choice for protein orientation in the membrane. Keep the PDB orientaion as a choice (Fig. 4) – this assumes the the Z-axis is aligned with the principal axis of inertia of the protein. Note other options. Proceed to the next step.

1.2 CHARMM-GUI: select lipids and water thickness. Cross sectional density will be displayed (not shown) as well as system size options (Fig. 5). Enter 50 Å as the initial estimate for the size of the cell (Fig. 5, left) and click the "Show the system info" button. The results of the estimate will be shown in the table (Fig. 5, right). We can keep the defult value of the water thickness at 17.5 Å.



System Size Determination Options:

○ Homogeneous Lipid
● Heterogeneous Lipid

1. Box Type: Rectangular (Currently, only CHARMM and NAMD support the hexagonal box)

2. Length of Z based on:
● Water thickness 17.5 (Minimum water height on top and bottom of the system)
○ Hydration number 50 (Number of water molecules per one lipid molecule)
○ Hydration (w/w) % 50 (Percent ratio of Water/lipid weight)

3. Length of XY based on:
● Ratios of lipid components
○ Numbers of lipid components

Length of X and Y: 50 (initial guess)
(The system size along the X and Y must be the same)

Show the system info click this once you fill the following table:

Figure 5 System size determination

Calculated Number of Lipids:

Lipid Type	Upperleaflet Number	Lowerleaflet Number

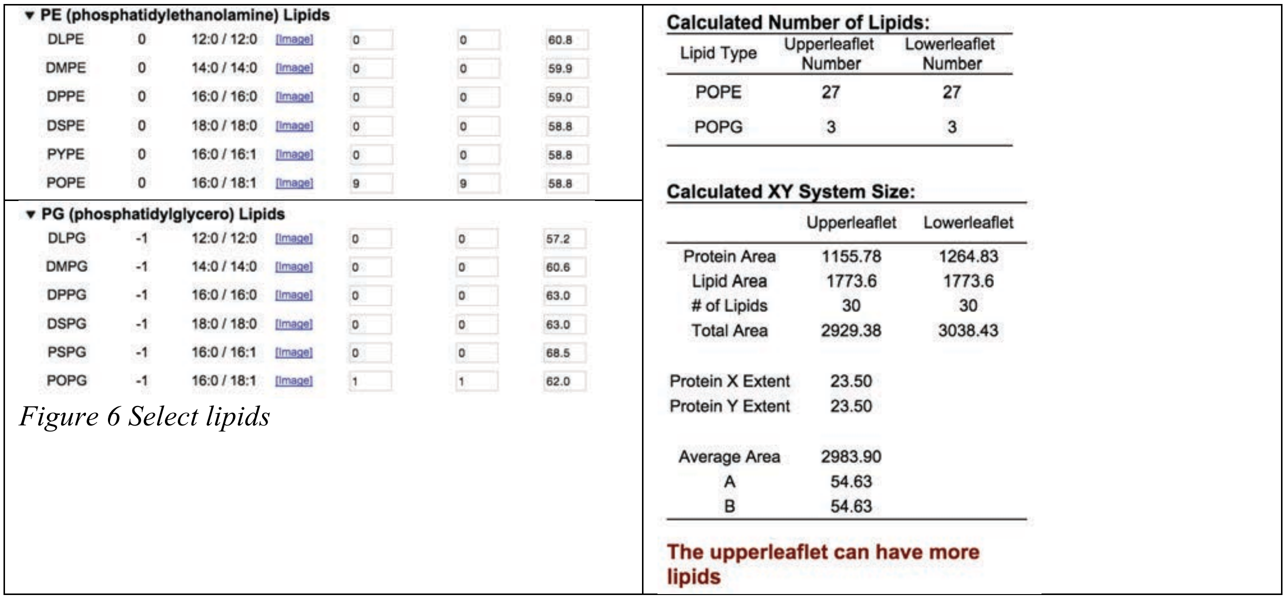
Calculated XY System Size:

	Upperleaflet	Lowerleaflet
Protein Area	1155.78	1264.83
Lipid Area	0	0
# of Lipids	0	0
Total Area	1155.78	1264.83
Protein X Extent	23.50	
Protein Y Extent	23.50	
Average Area	1210.30	
A	34.79	
B	34.79	

No lipid was selected in upperleaflet or lowerleaflet.

Select lipids under the clickable triangles (Fig. 6): Build a membrane consisting of phosphatidylethanolamine (POPE) 90% and phosphatidylglycerol (POPG) 10 %, which will be our model for the outer membrane of *E.coli*. Enter “90” in the upper and lower leaflet boxes for POPE, and “10” in the upper and lower leaflet boxes for POPG. As a reminder, PE is a zwitterionic (formally neutral) while PG is charged lipid. Hover over the links “Image” in the row of the chosen lipids and lipid structure will be shown. Click the “Show the system info” button to update.

Remark: for *E.coli* outer membrane composition see e.g., Lugtenberg EJJ & Peters R, BBA, 441, 38–47, 1976. **Click again the “Show the system info” button.** Observe that summary now has list of lipids (Fig. 6, right). Proceed to the next step.



▾ PE (phosphatidylethanolamine) Lipids

DLPE	0	12:0 / 12:0	[Image]	0	0	60.8
DMPE	0	14:0 / 14:0	[Image]	0	0	59.9
DPPE	0	16:0 / 16:0	[Image]	0	0	59.0
DSPE	0	18:0 / 18:0	[Image]	0	0	58.8
PYPE	0	16:0 / 16:1	[Image]	0	0	58.8
POPE	0	16:0 / 18:1	[Image]	9	9	58.8

▾ PG (phosphatidylglycero) Lipids

DLPG	-1	12:0 / 12:0	[Image]	0	0	57.2
DMPG	-1	14:0 / 14:0	[Image]	0	0	60.6
DPPG	-1	16:0 / 16:0	[Image]	0	0	63.0
DSPG	-1	18:0 / 18:0	[Image]	0	0	63.0
PSPG	-1	16:0 / 16:1	[Image]	0	0	68.5
POPG	-1	16:0 / 18:1	[Image]	1	1	62.0

Figure 6 Select lipids

Calculated Number of Lipids:

Lipid Type	Upperleaflet Number	Lowerleaflet Number
POPE	27	27
POPG	3	3

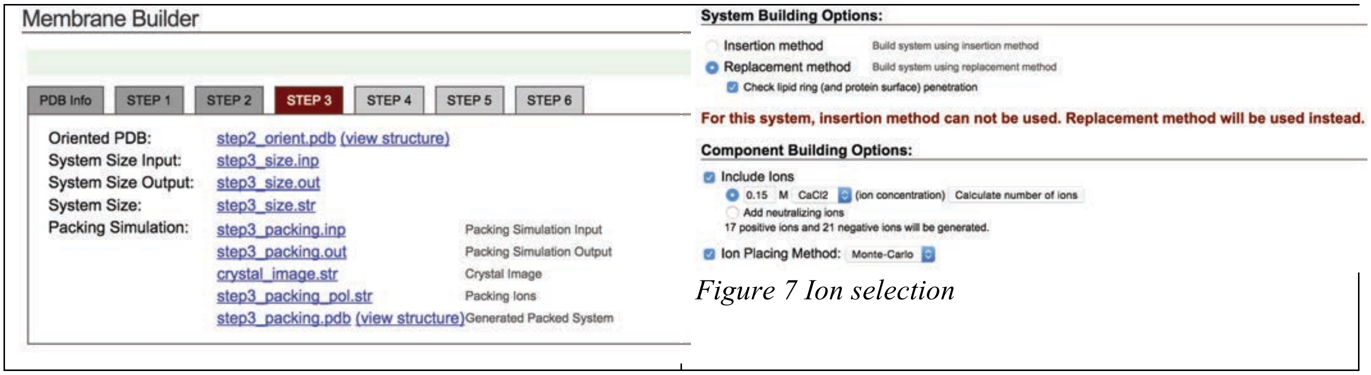
Calculated XY System Size:

	Upperleaflet	Lowerleaflet
Protein Area	1155.78	1264.83
Lipid Area	1773.6	1773.6
# of Lipids	30	30
Total Area	2929.38	3038.43
Protein X Extent	23.50	
Protein Y Extent	23.50	
Average Area	2983.90	
A	54.63	
B	54.63	

The upperleaflet can have more lipids

1.3 CHARMM-GUI: select ions

The system size summary will be displayed (not shown) as well as two choices for protein insertion into the membrane (Fig. 7, right). We will use the default Replacement method. As for the ions, we will use calcium ions for system neutralization and additional concentration. Select $CaCl_2$ as the salt of choice (Fig. 7, right). Proceed to the next step.



Membrane Builder

Oriented PDB:	step2_orient.pdb (view structure)	
System Size Input:	step3_size.inp	
System Size Output:	step3_size.out	
System Size:	step3_size.str	
Packing Simulation:	step3_packing.inp	Packing Simulation Input
	step3_packing.out	Packing Simulation Output
	crystal_image.str	Crystal Image
	step3_packing_pol.str	Packing Ions
	step3_packing.pdb (view structure)	Generated Packed System

System Building Options:

○ Insertion method — Build system using insertion method
● Replacement method — Build system using replacement method
☑ Check lipid ring (and protein surface) penetration

For this system, insertion method can not be used. Replacement method will be used instead.

Component Building Options:

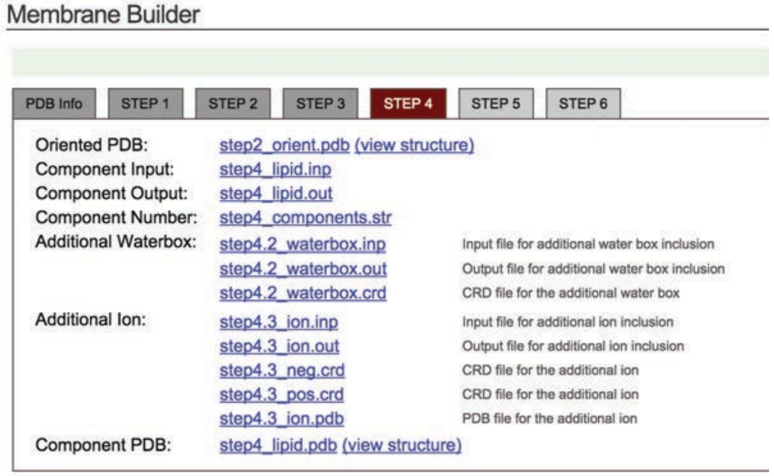
☑ Include Ions
● 0.15 M CaCl2 (ion concentration) Calculate number of ions
○ Add neutralizing ions
17 positive ions and 21 negative ions will be generated.

☑ Ion Placing Method: Monte-Carlo

Figure 7 Ion selection

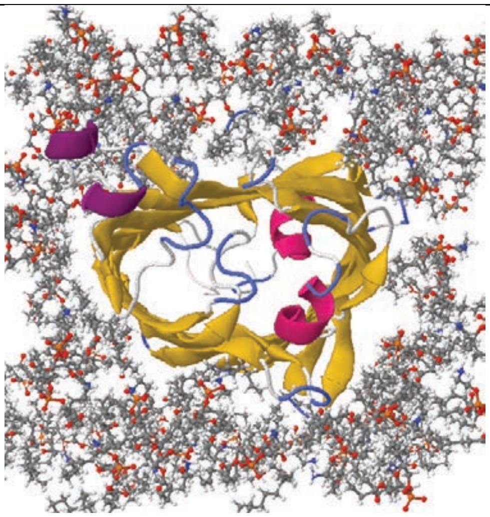
1.4 CHARMM-GUI: system preparation

CHARMM-GUI reports on possible lipid penetrations of rings (not shown). This could be a problem that will not be resolved by minimization and the system has to be regenerated. Proceed to the next step to combine membrane components. This may take few minutes. If this step is not working the instructor will provide pre-calculated files.



Oriented PDB:	step2_orient.pdb (view structure)	
Component Input:	step4_lipid.inp	
Component Output:	step4_lipid.out	
Component Number:	step4_components.str	
Additional Waterbox:	step4.2_waterbox.inp	Input file for additional water box inclusion
	step4.2_waterbox.out	Output file for additional water box inclusion
	step4.2_waterbox.crd	CRD file for the additional water box
Additional Ion:	step4.3_ion.inp	Input file for additional ion inclusion
	step4.3_ion.out	Output file for additional ion inclusion
	step4.3_neg.crd	CRD file for the additional ion
	step4.3_pos.crd	CRD file for the additional ion
	step4.3_ion.pdb	PDB file for the additional ion
Component PDB:	step4_lipid.pdb (view structure)	

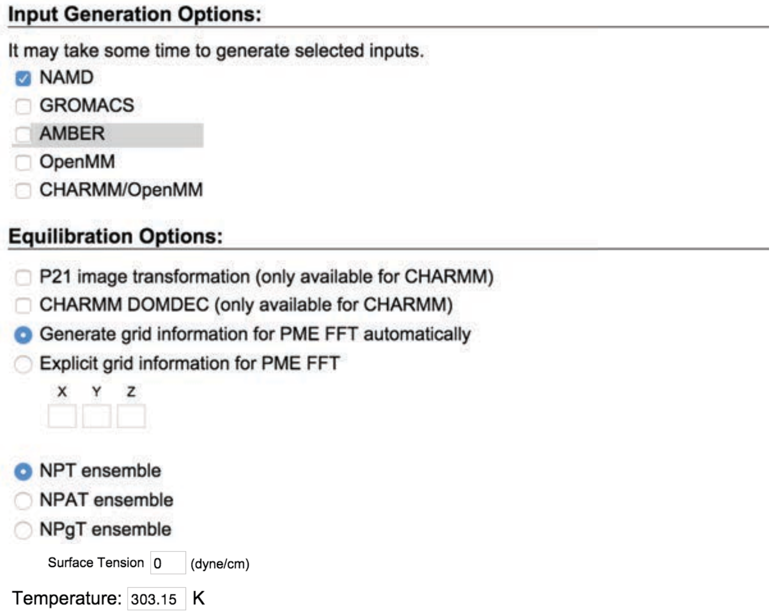
Figure 8 System assembly



CHARMM-GUI will display the list of prepared files (Fig.8, left). Click the “view structure” link next to the Component PDB file link. A pop-up window will be displayed where one can visually inspect the structure (Fig. 8, right). Proceed to the next step to complete the assembly.

1.5 CHARMM-GUI: generate input files

The list of files prepared so far, as well as the system size, will be displayed (not shown). Select only NAMD for the Input Generation Options as shown on Fig. 9. Proceed to the next step to complete the preparation.



Input Generation Options:

It may take some time to generate selected inputs.

☑ NAMD
☐ GROMACS
☐ AMBER
☐ OpenMM
☐ CHARMM/OpenMM

Equilibration Options:

☐ P21 image transformation (only available for CHARMM)
☐ CHARMM DOMDEC (only available for CHARMM)
● Generate grid information for PME FFT automatically
○ Explicit grid information for PME FFT

X Y Z

● NPT ensemble
○ NPAT ensemble
○ NPgT ensemble

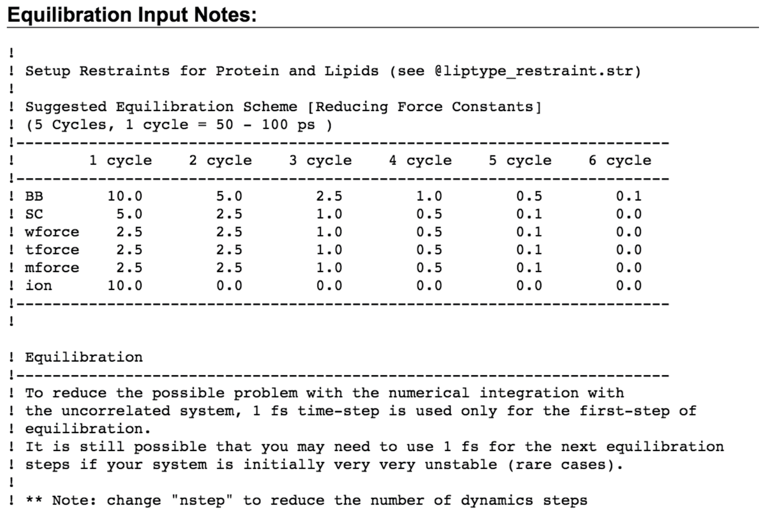
Surface Tension 0 (dyne/cm)

Temperature: 303.15 K

Figure 9 Input files options.

1.6 CHARMM-GUI: input file download.

The updated list of all generated files will be displayed again (not shown). The coordinate file is named step5_assembly.pdb, while the structure/topology file is step5_assembly.psf. Click the "download .tgz" in the right top corner. CHARMM-GUI will also display suggested steps for the simulation (Fig. 10). To expedite the process for the purpose of this tutorial the instructor will provide a single input file.



Equilibration Input Notes:

```
!
! Setup Restraints for Protein and Lipids (see @liptype_restraint.str)
!
! Suggested Equilibration Scheme [Reducing Force Constants]
! (5 Cycles, 1 cycle = 50 - 100 ps )
!-------------------------------------------------------------------------
!        1 cycle    2 cycle    3 cycle    4 cycle    5 cycle    6 cycle
!-------------------------------------------------------------------------
! BB       10.0        5.0        2.5        1.0        0.5        0.1
! SC        5.0        2.5        1.0        0.5        0.1        0.0
! wforce    2.5        2.5        1.0        0.5        0.1        0.0
! tforce    2.5        2.5        1.0        0.5        0.1        0.0
! mforce    2.5        2.5        1.0        0.5        0.1        0.0
! ion      10.0        0.0        0.0        0.0        0.0        0.0
!-------------------------------------------------------------------------
!

! Equilibration
!-------------------------------------------------------------------------
! To reduce the possible problem with the numerical integration with
! the uncorrelated system, 1 fs time-step is used only for the first-step of
! equilibration.
! It is still possible that you may need to use 1 fs for the next equilibration
! steps if your system is initially very very unstable (rare cases).
!
! ** Note: change "nstep" to reduce the number of dynamics steps
```

Figure 10 Suggested MD steps

2. Run simulation on cluster

We will run a short 200 ps simulation (1000 steps minimization followed by MD) using a single NAMD input file. Please note we are using a small number of lipids and a short simulation time for purposes of the tutorial only.

2.1 Prepare directory for submission

Copy* the template directory from provided template to your directory:

```
> cp -r /mnt/people/pogorelo/teaching/biophyscamp/namd_run_ompla_Clean .
```

You will need to uncompress the CHARMM-GUI directory (`tar -xzf *tgz`), change to the directory, and copy the coordinate and structure files step5_assembly.pdb/step5_assembly.psf to the template directory provided.

```
> cp charmm-gui/step5_assembly.p* ./namd_run_ompla_Clean
```

2.2 Submit job to queue

Hint: To preserve an intact copy of the starting directory for future use make a copy before you will submit a job:

```
> cp -r namd_run_ompla_Clean/ namd_run_ompla.lipid.1
```

Now change to the working directory `>cd namd_run_ompla.lipid.1` and load NAMD module:

```
> module load  namd2/multicore-2.11
```

Finally submit the job:

```
> submit-namd2 -n 32 -o namd_run_ompla.0.lipid.1.log namd_run_ompla.0.conf
```

2.3 Check queue and your job status.

Check the job status:

```
> qstat -u mylogin
```

To check how fast your job will progress run the following command that will search for all mentions of the word “Bench” in the .log file:

```
> grep Bench *log
```

The last entry will be the estimate for the speed of the NAMD simulation on the amount of cores you have chosen (32, if the previous command was used and in this case it is likely the speed is on the order of 0.11 day/ns or 1 hr/ns or ~26 ns/day). The cpu core vs atom count is an important consideration for selecting number of cores. Often most efficient NAMD runs are in the range of 300-1,000 atoms/core. Check the size of your system using Linux word count command. Should be ~24,000, hence you are running NAMD with 750 atoms/core:

```
> wc *pdb
```

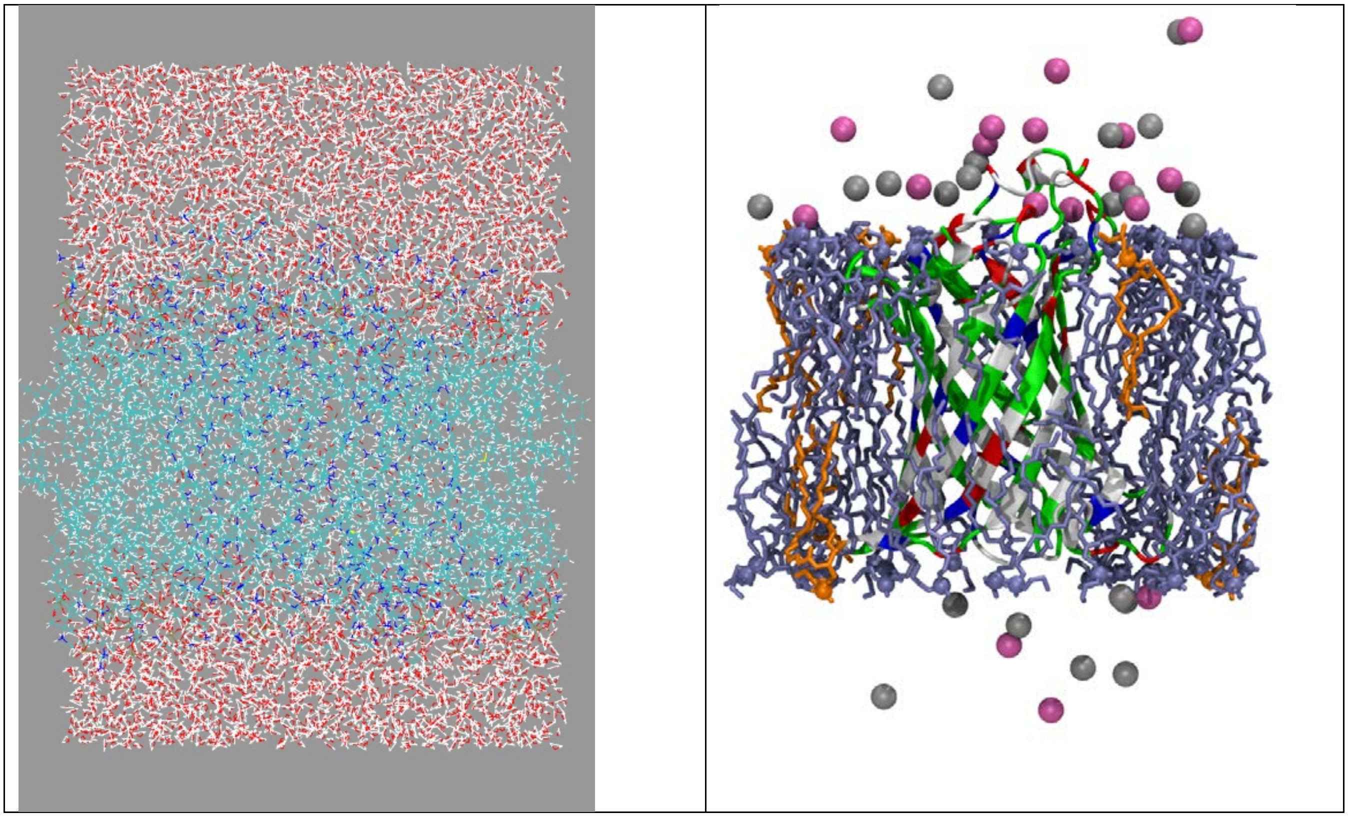
*Note: at the SCS the file system is shared between the VizLab computers and SCS cluster. Thus one does not need to copy files to the cluster. When using other clusters please refer to the Linux tutorial (http://computing.scs.illinois.edu/tutorials/) for a reminder on how to copy files to remote machines.

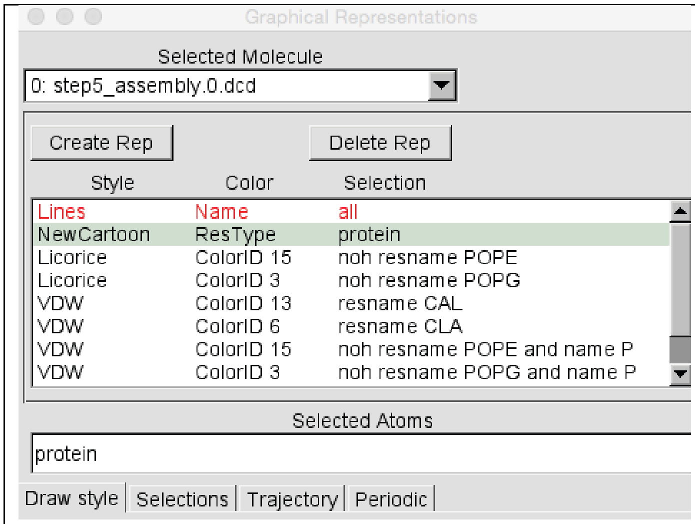
3. Analyses of the simulation with VMD

When your job is completed, open VMD (NAMD's companion program produced by the NIH Center at the Beckman Institute, http://www.ks.uiuc.edu/Research/vmd/). Load first the trajectory file (.dcd) and second structure/topology file (.psf). Please note a detailed introduction to VMD is beyond the scope of this tutorial. To learn more please see the VMD Tutorial produced by the makers of the program (http://www.ks.uiuc.edu/Training/Tutorials/).

3.1 Visual inspection of the system: residues, lipids, and ions

When you load the system you will see the default representation (Fig. 11, left). Lets now introduce other representation that will demonstrate properties of the system (Fig. 11, right). Open from the Main window: Graphics -> Representations window and create additional seven representations (see Fig. 11 lower left). Please note the components of the representations and keywords.





Selected Molecule: 0: step5_assembly.0.dcd

Style	Color	Selection
Lines	Name	all
NewCartoon	ResType	protein
Licorice	ColorID 15	noh resname POPE
Licorice	ColorID 3	noh resname POPG
VDW	ColorID 13	resname CAL
VDW	ColorID 6	resname CLA
VDW	ColorID 15	noh resname POPE and name P
VDW	ColorID 3	noh resname POPG and name P

Selected Atoms: protein

Figure 11 VMD analysis

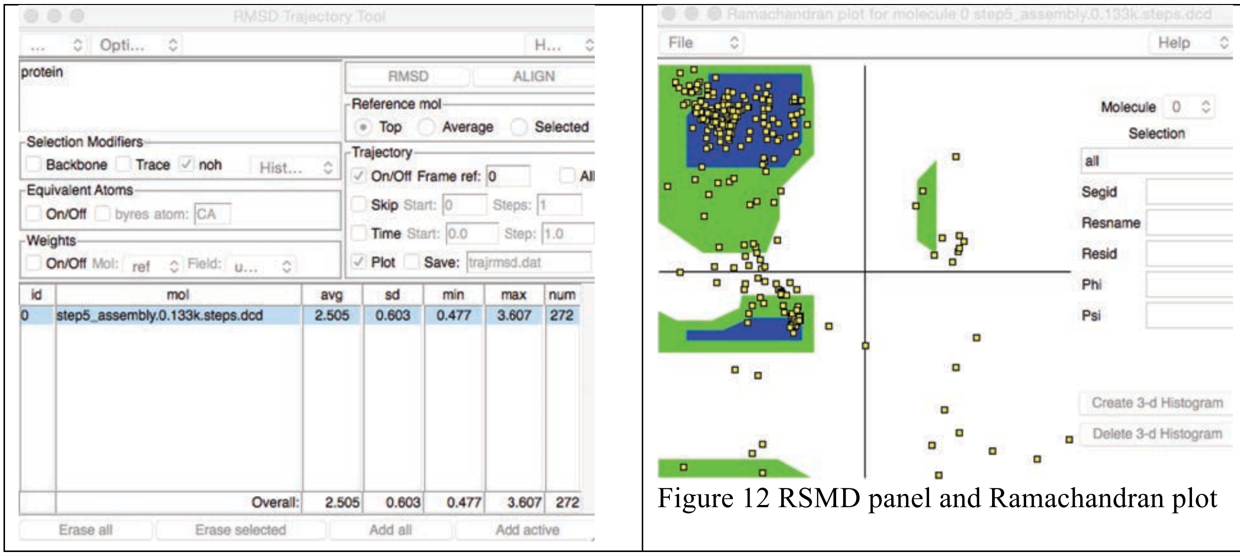
Components of the representations

1) Style
2) Color
3) Selection

In particular, Color:ResType color residues by the type (blue – positively charged, red – negatively charged, green – polar, and white – non-polar).

Questions/comments to discuss:

1) The OMP database placed membrane proteins using a particular energy function. What do you think is one of the main components of this function? Hint: the internal part of the membrane is hydrophobic.
2) Observe the locations of the charged residues (shown in red and blue): some are on the loops above the membrane and some are in the core of the membrane. Solvent accessible surface areas of the side chains can be measured with VMD (measure sasa). For a quick check, make a protein representation in VDW style and observe whether charged residues inside of the membrane are exposed.
3) Play the trajectory and observe the fluidity of the membrane and how ions are starting to interact with lipid. Although to capture lipid-ion interactions more accurately, we need a much longer simulation.
4) To quantitatively characterize the protein one can calculate: RMSD (root mean square deviation) and the Ramachandran plot (see Fig. 12). Try them from the Main window Extensions menu (RMSD Trajectory Tool and Ramachandran Plot).
5) The Moon and Fleming study (see reference to PNAS 2011 on page 1) used OmpLA point mutants at the position 210. Display it in VMD (resid 210) and comment on why this particular location was chosen.



protein

☐ Backbone ☐ Trace ☑ noh

☐ On/Off ☐ byres atom: CA

☐ On/Off Mol: ref Field: u...

Reference mol: ● Top ○ Average ○ Selected

Trajectory: ☑ On/Off Frame ref: 0 ☐ All

☐ Skip Start: 0 Steps: 1

☐ Time Start: 0.0 Step: 1.0

☑ Plot ☐ Save: trajrmsd.dat

id	mol	avg	sd	min	max	num
0	step5_assembly.0.133k.steps.dcd	2.505	0.603	0.477	3.607	272
	Overall:	2.505	0.603	0.477	3.607	272

Molecule 0

Selection: all

Figure 12 RSMD panel and Ramachandran plot

4. Summary

This tutorial covered basics of using *NAMD/VMD/CHARMM-GUI*: building membrane-protein systems, performing MD simulations on computer clusters and basic analysis on desktop computers.

5. Contact

If you found errors/typos or have suggestions or comments on material in this tutorial please contact us at the SCS Computer Center (pogorelo@illinois.edu and mhallock@illinois.edu). We are looking forward to hearing from you. http://computing.scs.illinois.edu