

Generate custom molecule parameters

We often want to run simulations with small molecules or chemically-modified macromolecules. To do so, we need to generate parameters for these ligands. The following protocol generates CHARMM36 forcefield parameters for use in the GROMACS simulation package.

Under the hood, this protocol builds parameters assuming that the molecule of interest is similar to molecules already in the forcefield. This works fairly well, particularly for molecules that have moieties found in standard biomolecules. You should be fine using this approach to model standard post-translational modifications, small molecules built out of CHNOPS atoms, and molecules like glycolipids that are made up of standard molecules. Once you get outside of these use cases, however, BEWARE. (I would not use this protocol to build an iron-sulfur cluster!)

There are three steps between deciding, "I want to simulate non-standard molecule X" and having the input directory used to run an MD simulation: 1) Generate CHARMM parameters, 2) Convert those parameters to GROMACS format, and then 3) Set up a simulation using those parameters. This protocol goes through each of those steps.

The "ligand-param/" directory has example outputs generated using this protocol.

If you use this protocol, *PLEASE CITE THE FOLLOWING*:

S. Jo, T. Kim, V.G. Iyer, and W. Im (2008) CHARMM-GUI: A Web-based Graphical User Interface for CHARMM. [J. Comput. Chem. 29:1859-1865](#)

J. Lee, X. Cheng, J.M. Swails, M.S. Yeom, P.K. Eastman, J.A. Lemkul, S. Wei, J. Buckner, J.C. Jeong, Y. Qi, S. Jo, V.S. Pande, D.A. Case, C.L. Brooks III, A.D. MacKerell Jr, J.B. Klauda, and W. Im (2016) CHARMM-GUI Input Generator for NAMD, GROMACS, AMBER, OpenMM, and CHARMM/OpenMM Simulations using the CHARMM36 Additive Force Field. [J. Chem. Theory Comput. 12:405-413](#)

S. Jo, X. Cheng, S.M. Islam, L. Huang, H. Rui, A. Zhu, H.S. Lee, Y. Qi, W. Han, K. Vanommeslaeghe, A.D. MacKerell, Jr., B. Roux, and W. Im (2014) CHARMM-GUI PDB Manipulator for Advanced Modeling and Simulations of Proteins Containing Non-standard Residues. [Adv. Protein Chem. Struct. Biol. 96:235-265](#)

I. Generate CHARMM parameters for a custom ligand

Protocol

1. Create a PDB file that has only your ligand of interest (no protein, solvent, or other molecules).
 - You can do this from PyMOL by loading a structure that has the ligand of interest docked in the correct location, going to "File->Export molecule...", and then selecting only the ligand as output. When you get to the final save dialog, make sure to select "PDB", not "PDBx/mmCIF."
 - Alternatively, you can manually edit a PDB file with a protein and ligand in a text editor, deleting any non-ligand atoms.
 - If you are simulating LPS and want to generate a specific LPS variant, see the "Generate an LPS molecule" section at the end of this document.
 - To get structures for other ligands, check out pubchem.ncbi.nlm.nih.gov. You can download 3D structures of many small molecules. You may need to convert small molecules to the PDB format from some other format (e.g. SDF, MOL2, etc.). I recommend OpenBabel (openbabel.org) for this purpose.
2. Go to charmm-gui.org and log in.
3. Click "Input generator" on the left-hand navigation bar.
4. Click "PDB Reader & Manipulator" on the left-hand navigation bar. This will bring up a page to upload your file.
5. Upload your ligand PDB file using the box at the bottom of the page.
6. Click "Next Step: Select Model/Chain" on the bottom right. A new page should appear that has your ligand as a chain with a checkbox next to it.
7. Click "Next Step: Manipulate PDB" on the bottom right. A new page will appear that has a bunch of options for how to treat your ligand.
8. Set whatever options that are appropriate for the chemistry of your ligand. (In my experience, the software is pretty intelligent at selecting how to treat molecules, but you should check carefully and make sure you understand the modeling decisions the software is making).
9. Click "Next Step: Generate PDB" on the bottom right. This will bring up a new page allowing you to download your results.
10. Click the red button titled "download.tgz" on the top right. This will download a compressed folder with your newly generated model to your computer.

11. On your computer, open the downloaded file. This will create a folder that looks something like "charmm-gui-1285168560", where the number is random and unique to the job you just ran.

Output

A directory called charmm-gui-BIG-NUMBER with your model. The .PDB file has the coordinates, while the .PSF file the topology and parameters. The .STR files have specific constraints and additional forcefield parameters to treat your ligand.

II. Convert CHARMM parameters to GROMACS format

Protocol

1. Go to charmm-gui.org and log in.
2. Click "Input generator" on the left-hand navigation bar.
3. Go to "Force Field Converter"
4. Set up the conversion:
 - Upload the PSF and Coord files from the CHARMM parameter generation step (usually "step1_pdbreader.psf" and "step1_pdbreader.crd").
 - Click "Upload Additional Topology and Parameter Files" and use the interface to upload any ".str" files that are in the output directory from your topology generation step. (The number and names of these files vary depending on the ligand.) You do NOT need to upload any .str files inside the "toppar" sub-directory.
 - Un-click "Setup PBC".
5. Click "Next step: Check System Components" on the bottom right. This will bring up a new page allowing you to select conversion options.
6. Set up the conversion:
 - Under "Forcefield options", select "CHARMM36m". No other boxes should be checked.
 - Under "Input Generation Options" select "GROMACS". No other boxes should be checked.
 - Under "Equilibration Options", select "NPT" at a temperature 303.15 K.

7. Click "Next step: Generate Inputs" on the bottom right. This will bring up a new page allowing you to download your results.
8. Click the red button titled "download.tgz" on the top right. This will download a compressed folder with your newly generated model to your computer.
9. On your computer, open the downloaded file. This will create a folder that looks something like "charmm-gui-1285168560", where the number is random and unique to the job you just ran.
10. Make a new directory in a convenient location. Copy "charmm-gui-BIG-NUMBER/gromacs" into that directory. Then, copy "charmm-gui-BIG-NUMBER/gromacs/step3_input.pdb" to "model-to-pose.pdb" in your new directory.

Output

A directory with a folder named "gromacs" (holding the topology and parameters in GRO-MACS format) and a file named "model-to-pose.pdb" (holding the structural model of the ligand).

III. Set up a simulation with custom parameters

Protocol

1. Align your protein structure to the ligand structure in DIRECTORY/model-to-pose.pdb, where "DIRECTORY" is the directory generated in the last step. This can be done in a variety of ways; the following describes how to do so using PyMOL.
 - Find a PDB file containing both your protein AND your something similar to your ligand of interest. For example, we might want to simulate testosterone bound to the estrogen receptor. There are multiple structures of the estrogen receptor bound to estradiol, but none to testosterone. To simulate testosterone bound, we could generate a model for testosterone using the steps above, then place the protein and testosterone into the correct relative orientation by overlying the estradiol and testosterone.
 - Load your PDB file with protein and ligand into PyMOL.
 - In the same PyMOL session, load the PDB file for your ligand ("DIRECTORY/model-to-pose.pdb").

- On the PyMOL command line, run the command "align SELECTION, model-to-pose", where SELECTION selects the ligand already in the PDB file. (For the example above, we'd use SELECTION="resname EST", selecting the estradiol.) This will align the ligands, dragging your protein along for the ride. At this point, your protein structure is now oriented correctly relative to the model.
 - You might want to remove clashes between atoms in the protein and ligand. The software will also remove clashes later, but if something is egregious, it's worth taking care of now. A simple way to do this is via the PyMOL "Mutagenesis Wizard," which allows you to sample side chain rotamers.
 - When you are happy with the model, save out the protein structure, but *not* the ligand. This can be done by going to "File -> Export molecule..." and setting "Selection" to "polymer.protein". When you get to the final save dialog, make sure to select "PDB", not "PDBx/mmCIF."
 - *If you have to change the conformation of the model-to-pose, you can use PyMOL's editing mode. **You must then save out the altered model-to-pose exactly as follows.*** Go to "File -> Export molecule..." and use a "Selection" that only grabs your altered model. Then, make sure that "Original atom order" and "Retain atom ids" are both checked in the export dialog. (These are not selected by default). When you get to the final save dialog, make sure to select "PDB", not "PDBx/mmCIF." Then use this new PDB file you just saved out instead of "model-to-pose.pdb" in the instructions below.
2. At this point, you have a PDB file with the protein (we'll pretend it's called PROTEIN_PDB.pdb), a ligand in the correct orientation relative to the pocket (DIRECTORY/model-to-pose.pdb), and a topology for that ligand in GROMACS format (DIRECTORY/gromacs). To set up the MD run, use the following command.

```
setup_md.py output --standard PROTEIN_PDB.pdb --custom DIRECTORY/model-to-pose.pdb --params DIRECTORY/gromacs/
```

3. Copy the "output/final" folder to the cluster.

Output

The folder "output/final" has everything necessary to start a simulation. It can be uploaded to the cluster and used to launch an MD simulation with the command `sbatch run_md.srun`.

Appendix: note on methodology

Instead of generating parameters for a single ligand, you can also upload a PDB with the protein and any ligand(s) to charmm-gui. Section I, steps #6, #7, and #8, will help you select appropriate chains and chemistry for the whole system. The directory you download at the end of Section I will then have parameters for all atoms in the combined system. You can follow the steps in Section II, unchanged, to convert this to GROMACS format. Finally, in Section III, you can run the following command. (Note that we only specify custom, not standard, as the protein is in the custom pose file).

```
setup_md.py output --custom DIRECTORY/model-to-pose.pdb --params
DIRECTORY/gromacs/
```

I generally do not use this whole-system protocol for two reasons. First, in my experience, the more complicated the system, the more likely it is that charmm-gui spits out some incomprehensible error. It is (usually) simpler to just run ligands through charmm-gui individually. Second, the ligand-alone protocol yields a directory with parameters that you can use for future simulations. This is important, as the algorithm used to generate the parameters on the server could change. If you want to compare a simulation you run today to one you run in a year, it's best to use identical parameters. The ligand-alone protocol thus provides greater consistency and control.

Appendix: generate an LPS molecule

If you use this protocol to generate LPS molecules, please *CITE THE FOLLOWING* in addition to the papers noted at the top of the protocol.

J. Lee, D.S. Patel, J. Stähle, S-J. Park, N.R. Kern, S. Kim, J. Lee, X. Cheng, M.A. Valvano, O. Holst, Y. Knirel, Y. Qi, S. Jo, J.B. Klauda, G. Widmalm, and W. Im (2019) CHARMM-GUI Membrane Builder for Complex Biological Membrane Simulations with Glycolipids and Lipoglycans. *J. Chem. Theory Comput.* 15:775-786

Protocol

1. Go to charmm-gui.org and log in.
2. Click "Input generator" on the left-hand navigation bar.
3. Go to "LPS modeller". This will bring up a page with a million options for your LPS molecule.
4. Choose the options you want for your LPS molecule.

5. Click "Next Step: Generate PDB/PSF" on the bottom right. This will bring up a new page allowing you to download your results.
6. Click the red button titled "download.tgz" on the top right. This will download a compressed folder with your newly generated model to your computer.
7. On your computer, open the downloaded file. This will create a folder that looks something like "charmm-gui-1285168560", where the number is random and unique to the job you just ran.
8. The outputs can then be used as inputs for conversion to a GROMACS parameters following the instructions given in Section II. The only difference will be the names of the .PSF and .CRD files: with three small differences: "lps_modeler.psf" instead of "step1_pdbreader.psf"; "lps_modeler.crd" instead of "step1_pdbreader.crd".

Output

A directory called charmm-gui-BIG-NUMBER with your LPS model. The .PDB file has the coordinates, while the .PSF file the topology and parameters. The .STR files have specific constraints and additional forcefield parameters to treat your ligand.