GROMACS Tutorial: Atomistic Simulation of a Protein in Water Using CHARMM36m

This tutorial presents a complete workflow for preparing and simulating a protein in aqueous solution using GROMACS 2024.1 with CHARMM36m. The commands follow a stepwise logic, with each step explained in detail. Simulations are GPU-accelerated and suitable for production-level calculations.

1. System Preparation

1.1 Generate the Protein Topology

```
echo '1' | gmx pdb2gmx -f 4IU3_ok.pdb -o complex.gro -water tip3p -ss yes -ignh
```

This command generates a GROMACS-compatible structure and topology.

- -f 4IU3_ok.pdb specifies the input PDB file.
- -o complex.gro writes the output coordinates in GROMACS .gro format.
- -water tip3p selects the water model. CHARMM36m is parametrized with a modified TIP3P model, and this selection ensures consistency.
- -ss yes adds secondary structure information from DSSP if available, which is important for some force fields.
- -ignh ignores existing hydrogen atoms and rebuilds them according to the force field.

You will be prompted to select a force field. Choose CHARMM36m. This will produce:

- complex.gro: coordinate file with hydrogen atoms.
- topol.top: system topology file.
- posre.itp: position restraint definitions, active during equilibration.

1.2 Define the Simulation Box

```
gmx editconf -f complex.gro -o newbox.gro -bt dodecahedron -c -d 1.0
```

Defines the shape and dimensions of the unit cell.

- -bt dodecahedron uses a rhombic dodecahedral box, reducing the number of solvent molecules needed.

- -c centers the protein in the box.
- -d 1.0 sets a 1.0 nm minimum distance between the protein and box edges, which prevents artifacts due to periodic boundary conditions.

1.3 Solvate the Protein

```
gmx solvate -cp newbox.gro -cs spc216.gro -p topol.top -o solv.gro
```

This fills the box with water molecules.

- -cp newbox.gro is the input coordinate file from the previous step.
- -cs spc216.gro provides the water box. Despite its name, it is used for TIP3P because CHARMM uses the modified TIP3P geometry.
- -p topol.top updates the topology file with the number of added water molecules.
- -o solv.gro outputs the solvated system.

1.4 Add Ions to Neutralize and Set Ionic Strength

Generate a pre-ionization input file:

```
gmx grompp -f ions.mdp -c solv.gro -p topol.top -o ions.tpr -maxwarn
1
```

- ions.mdp should contain basic settings for energy minimization (no dynamics are run here).
- This prepares a .tpr file that defines the system geometry and topology.

Replace solvent molecules with ions:

```
echo '15' | gmx genion -s ions.tpr -o solv_ions.gro -p topol.top
-pname NA -nname CL -neutral -conc 0.150
```

- -pname NA and -nname CL define ion names.
- -neutral ensures the system has no net charge.
- -conc 0.150 adds NaCl to 150 mM to approximate physiological conditions.
- echo '15' selects the solvent group, typically group 15, which is the water molecules to be replaced.

This outputs the updated solvated and neutralized system as solv_ions.gro.

2. Energy Minimization

```
gmx grompp -f minim.mdp -c solv_ions.gro -p topol.top -o em.tpr -r
solv_ions.gro -maxwarn 2

gmx mdrun -v -deffnm em -ntmpi 2 -ntomp 8 -pin on
```

Purpose: relax bad contacts between atoms introduced during solvation and ion placement.

- minim.mdp should use steepest descent and define a force tolerance (emtol, typically 1000.0 kJ/mol/nm).
- -r sets the reference positions for any restraints.
- mdrun performs the minimization, outputting energy (em.edr), structure (em.gro), and a log (em.log).

Energy minimization ensures the system starts in a physically reasonable configuration.

---

3. Equilibration

3.1 Temperature Equilibration (NVT)

Equilibration is divided into four short NVT stages. Each uses output from the previous step as input and maintains position restraints on the heavy atoms of the protein.

Example for NVT1:

```
gmx grompp -f nvt1.mdp -c em.gro -r em.gro -p topol.top -o nvt1.tpr
-maxwarn 2

gmx mdrun -v -deffnm nvt1 -gpu_id 0 -ntmpi 2 -ntomp 8 -pin on -nb gpu
-bonded gpu -pme gpu -npme 1 -dlb auto
```

Repeat this sequence with nvt2.mdp, nvt3.mdp, and nvt4.mdp, chaining the input and output files accordingly.

Each .mdp file should gradually raise the temperature (e.g., 50 K increments up to 300 K), and apply a thermostat such as velocity rescaling with a 0.1 ps coupling time.

3.2 Pressure Equilibration (NPT)

```
gmx grompp -f npt.mdp -c nvt4.gro -r nvt4.gro -t nvt4.cpt -p
topol.top -o npt.tpr -maxwarn 2

gmx mdrun -v -deffnm npt -gpu_id 0 -ntmpi 2 -ntomp 8 -pin on -nb gpu
-bonded gpu -pme gpu -npme 1 -dlb auto
```

- The system is now equilibrated at constant temperature and pressure (NPT ensemble).
- Use the Parrinello-Rahman barostat with ref_p = 1.0 and tau_p = 2.0.
- Maintain position restraints.

---

4. Production Simulation

```
gmx grompp -f md.mdp -c npt.gro -t npt.cpt -p topol.top -o
md_0_10.tpr -maxwarn 2

gmx mdrun -deffnm md_0_10 -gpu_id 0 -ntmpi 2 -ntomp 8 -pin on -nb gpu
-bonded gpu -pme gpu -npme 1 -dlb auto
```

This performs an unrestrained simulation of 10 ns.

- Use a time step of 2 fs.
- Use constraints on all bonds (LINCS).
- Save frames every 10 ps.

Simulations can be extended using gmx convert-tpr and restarted with mdrun -cpi.