WORKFLOW DESCRITPTION FOR ATOMISTIC MOLECULAR DYNAMICS SIMULATION OF BIOMOLECULES WITH GROMACS 5.1

VI-SEEM application MULTIDRUG

Initial files

It is convenient to store all parameter files in one 'master' folder – the MDP files, the force field (FF) folder, the initial structure *.pdb file, etc. To obtain parameters for the biomolecule of interest (and check if all the information in the FF is in order), the following command should be executed:

```
gmx pdb2gmx -f molecule.pdb -o molecule.gro
```

When no error massages appear, the FF and the type of water molecules are chosen from an interactive menu. Apart from the initial configuration of the biomolecule, the program will generate a molecular topology (.top → to be renamed to .itp) and restraints (*_restraints.itp) file for it. A system topology file (.top) may be adopted from a previous simulation. A sample one (membrane.top) is enclosed to this workflow description.

General remark on file types: The *.mdp files contain commands, keywords, and parameters for the calculations, while *.top files represent the description of the model system contents and call the particular force field used to compute energies and some files (*.itp) containing the topology and some force field parameters for the biomolecules. Some sample files are enclosed to this workflow. Please consult the manual of Gromacs for further details.

Sometimes, index files are needed. They put user-specified atom selections in one group. An index file is created with:

```
gmx make_ndx -f molecule_water_ions.gro -o molecule.ndx
```

In the interactive menu, the desired selection is specified and (optional) all the unnecessary groups are deleted.

It is recommended that periodic boundary conditions are applied in the three directions throughout the simulations. This is specified in a .mdp file.

System preparation

The next step in a bio-simulation would be to recreate as closely as possible physiological conditions. First, the molecule is put in a simulation box with the desired size and shape with the command:

```
gmx editconf –f molecule.gro -box 5 -c -o molecule_box.gro
```

This example creates a cubic box with edge length of 5 nm and puts the molecule in the geometrical center of the box.

The following command is used to solvate the molecule, i.e., to fill the box with water molecules:

```
gmx solvate -cp molecule_box.gro –cs –o molecule_water.gro
```

The number of the added solvent molecules should be included in the system topology file.

To neutralize the molecule charge and add ions for a certain ionic strength the following commands should be executed:

```
gmx grompp -f any_mdp_file.mdp -c molecule_water.gro -p molecule.top -o molecule_for_ions.tpr
```

```
gmx genion -s molecule_for_ions.tpr -np XX -nn YY -o molecule_water_ions.gro
```

The option –np requires the number of positive ions and the one –nn – the number of negative ions. The program replaces solvent molecules with ions, so the corresponding changes in the system topology file after this step.

Minimization (environment)

The first stage of the MD simulation is energy minimization of the structure of the environment – position restraints on the biomolecule are imposed. The following command is used to create a .tpr file:

```
gmx grompp -f min_env.mdp -c molecule_water_ions.gro -p molecule.top -o molecule_min_env.tpr
```

Use the *run_mdrun.pbs* script to run it. The script is suitable for the VI_SEEM facility Avitohol.

Check (1a)

See if the stage ended at the bottom of the log file – the maximum force should have converged below the criterion. Check the evolution of the potential energy during the procedure, by plotting the output from:

```
gmx energy –f  molecule_min_env.edr –o molecule_min_env_ener.xvg
```

In the interactive menu, the property to be check is picked up (in this case – "Potential"). It has to drop down continuously up to a constant value.

It is beneficial also to check the trajectory in a visualization program (for example, VMD). The following trajectory processing is required before visualization:

```
gmx trjconv –f  molecule_min_env.trr –s olecule_min_env.tpr –pbc whole –o molecule_min_env_whole.trr
```

A similar command may be used to visualize the trajectories after each subsequent stage.

Minimization (whole system)

The next step is to minimize the energy of the whole system. To create a .tpr file:

```
gmx grompp -f min_whole.mdp -c molecule_water_ions.gro -p molecule.top -o molecule_min_whole.tpr
```

Use the *run_mdrun.pbs* script (after changing it accordingly) to run the calculation.

Check (1b)

Once again, follow the potential energy evolution. Plot the output from:

```
gmx energy –f molecule_min_whole.edr –o molecule_min_whole_ener.xvg
```

It should fluctuate around a constant value.

Heating

Then, the system should be heated to the desired temperature (310 K is the temperature of the human body – appropriate for simulations of biomolecules). A .tpr file is created from the minimized structure and run with *run_mdrun.pbs* (after changing it accordingly):

```
gmx grompp -f heat.mdp -c molecule_min_whole.gro -p molecule.top -o molecule_heat.tpr
```

Check (2)

The evolution of the total energy of the system and the change in temperature (total and that of separate groups) should be checked after this stage:

```
gmx energy –f  molecule_heat.edr –o molecule_heat_ener_temp.xvg
```

The values of both thermodynamic properties should rise steadily with time, until the desired temperature is reached – then they should fluctuate around constant values.

Relaxation

After heating, constant pressure or surface tension (files provided as examples) maintenance should be introduced to the system and a relaxation stage is needed to reach equilibrium. To create a .tpr file:

```
gmx grompp -f relax.mdp -c your_molecule_heat.gro -p your_molecule.top -o your_molecule_relax.tpr
```

Use the *run_mdrun.pbs* script (after changing it accordingly) to run the calculation.

Check (3)

There is no predefined relaxation time, it all depends on the evolution of some thermodynamic and structural parameters. The evolution of total energy, temperature, pressure and density, should be checked by plotting the output of:

```
gmx energy -f  molecule_relax.edr -o molecule_relax_stat.xvg
```

They should fluctuate around their predefined constant values. The root-mean-square deviation (RMSD) of the coordinates with respect to the initial structure provides information about the structural dynamics of the biomolecule. To calculate it, the trajectory should first be processed as follows:

```
gmx trjconv -f  molecule_relax.trr -s molecule_relax.tpr -pbc whole
-o molecule_relax_whole.trr
```

```
gmx trjconv -f  molecule_relax_whole.trr -s molecule_relax.tpr -
pbc nojump -o molecule_relax_whole_nojump.trr
```

For the RMSD calculation, it is recommended to use the minimization .tpr to track the change from the initial structure:

```
gmx rms -f molecule_relax_whole_nojump.trr -s
molecule_min_whole.tpr -o molecule_relax_whole_nojump_RMSD.xvg
```

When the RMSD also fluctuates around constant value(s) (one or more substructures), it may be assumed that the system is in equilibrium.

Production stage

With a properly equilibrated system, the calculation may proceed with the final production stage. To create a tpr for the desired trajectory duration:

```
gmx grompp -f prod.mdp -c molecule_relax.gro -p molecule.top -o
molecule_prod.tpr
```

Run the simulation with *run_mdrun.pbs* (after changing it accordingly).

Trajectory

At the previous step a trajectory was obtained, which can be prepared for further check, analysis, and visualization as follows:

```
gmx trjconv -f  molecule_prod.trr -s molecule_prod.tpr -pbc whole
-o molecule_prod_whole.trr
```

```
gmx trjconv –f  molecule_prod_whole.trr –s molecule_prod.tpr –pbc nojump –o molecule_prod_whole_nojump.trr
```

Analysis

First, the mandatory checks as in the previous stage of the simulation (total energy, temperature, pressure, density, RMSD) should be performed:

```
gmx energy –f  molecule_prod.edr –o molecule_prod_stat.xvg
```

```
gmx rms -f molecule_prod_whole_nojump.trr -s molecule_min_whole.tpr -o molecule_prod_whole_nojump_RMSD.xvg
```

If the fluctuations of those parameters are constant, other analyses specific for the investigated problem may be performed. Please consult the manual of Gromacs for various options.

Data

Now, you have obtained data, which you can rationalize with the help of your knowledge and intuition :-).