Protocol SIRAH FF in GROMACS

Hybrid solvation : plugging SIRAH solvent to your atomistic system

This tutorial is based on the official SIRAH tutorial for hybrid solvation by Matias Machado. It shows how to apply the hybrid solvation approach of SIRAH Force Field (FF) to speed up the simulation of an atomistic solute. The example system contains a hPSP protein surrounded by a shell of fine-grained (FG) waters, which are embedded in coarse-grained (CG) molecules, called WT4, to represent bulk water.

The main references for this tutorial are:
- Darré et al. *WAT4 ?* [JCTC, **2010**, 6:3793]
- Darré et al. *All-atoms/CG solvation* [JCTC, **2012**, 8:3880]
- Gonzalez et al. *Transferable All-atoms/CG solvation* [J Phys Chem B, **2013**, 117:14438],
- Machado et al. *SIRAH Tools* [Bioinformatics, **2017**, 32:1568].

Ionic strength is better represented by including **FG ions to the FG shell** in addition of CG ions in the CG region, as demonstrated in previous references.

Required Software: GROMACS version >= 4.5.5
Prior knowledge: How to perform a standard atomistic molecular dynamic simulation with GROMACS.

0. Retrieve the starting files

Download the archive with all the necessary stuff and extract it :

```
curl -L "https://github.com/eddydu44/SIRAH-FF-Hybrid-
Solvation/raw/main/Protocol.Hybrid.Solvation.FILES.tar.gz" | tar xz
```

You will get a folder containing the force field definition of SIRAH inside the sirah.ff folder (version 2.2_20-07), the necessary MDP files to perform the simulation, the topology files, and the protein already in the GROMACS format. The protein was prepared using the command pdb2gmx over the protein structure of chain A, retrieved from rscb.org with the PDB code 6HYJ, without the ligand. Then, the ligand was added using ACPYPE server and the GAFF2 Force Field.
SIRAH FF was called inside the topol.top file using these lines :

```
#include "./sirah.ff/hybsol_comb2.itp"
#include "./sirah.ff/solv.itp"
#include "amber99sb-ildn.ff/ions.itp"
```

Be vigilant on the position of these lines because it can cause conflicts with the AMBER FF.

1. Box creation

```
gmx editconf -f 0.hPSP-pSER.gro -o 1.boxed.gro -c -d 2.0 -bt cubic
```

2. Solvate the system with a 1 nm FG shell

```
gmx solvate -cp 1.boxed.gro -cs spc216.gro -o 2.water.gro -shell 1 -p topol.top
```

The command returns one essential information on success, the density of solvent molecules in the box :

```
Output configuration contains 13384 atoms in 3522 residues
Volume                 :      1021.42 (nm^3)
Density                :       139.669 (g/l)
Number of solvent molecules:   3299
```

3. Add FG ions for a concentration of 0.15 mol/L

Prepare a full precision system topology :

```
gmx grompp -f ions.mdp -c 2.water.gro -p topol.top -o 2.water.tpr
```

Now we need to correct the concentration of ions wanted as we are not in a full box of water. Using earlier density, we **can apply a proportional coefficient** as we know that a full box of water would have a density of around 1000g/L (if you want to be more precise, you can run the command of point 2. without the shell option and get the density).

Consequently, the wanted concentration would be of : $C = 0.139 * 0.15 = \mathbf{0.02\ mol/L}$

Run the following command to add ions into the FG shell, select the SOL molecules to be substituted with.

```
gmx genion -s 2.water.tpr -o 3.ions.gro -p topol.top -pname NA -nname CL -neutral -conc 0.02
```

4. Add CG water

Same method as point 2., but this time we don't use the shell option to fill the rest of the box with the *WT*4 CG representation of water.

```
gmx solvate -cp 3.ions.gro -cs ./sirah.ff/wt4tip3p.gro -o 4.CG.gro -p topol.top
```

5. Add CG ions

Prepare a full precision system topology, correct number of molecules in topol.top if needed :

```
gmx grompp -f ./sirah.ff/tutorial/2/CPU/em_HYBSOL.mdp -p topol.top -c 4.CG.gro -o 4.CG.tpr -maxwarn 2
```

We need to calculate how many CG ions must be added to reach a concentration of 0.15 mol/L. Retrieve the number of WT4 molecules by looking at the topology file "topol.top", or by looking at command output at point 4.

The available ionic species in SIRAH force field are Na+ (NaW), K+ (KW) and Cl- (ClW). One ion pair (e.g., NaW-ClW) each 34 WT4 molecules renders a salt concentration of ~0.15M.

For 2273 *WT*4 molecules we need 67 pairs of CG ions. Add them similarly to point 3. :

```
gmx genion -s 4.CG.tpr -o 5.CG.ions.gro -np 67 -pname NaW -nn 67 -nname ClW -p topol.top
```

6. MD Simulation

Create an index and add the following groups for the MD simulations :
Protein_SEP_CA, Water_and_ions, WT4_NaW_ClW

```
gmx make_ndx -o index.ndx -f 5.CG.ions.gro
```

Then you are ready to run Energy Minimization, Equilibrations (NVT then NPT) and Production (200ns) using the following series of commands :

```
#Optimisation
gmx grompp -f em.mdp -c 5.CG.ions.gro -o 6.EM.tpr
gmx mdrun -deffnm 6.EM

#Equilibration nvt_2
gmx grompp -f nvt_2.mdp -c 6.EM.gro -p topol.top -o 7.nvt_2.tpr -n index.ndx
gmx mdrun -deffnm 7.nvt_2

#Equilibration nvt_3
gmx grompp -f nvt_3.mdp -t 7.nvt_2.cpt -p topol.top -o 8.nvt_3.tpr -n index.ndx
gmx mdrun -deffnm 8.nvt_3

#Equilibration nvt_4
gmx grompp -f nvt_4.mdp -t 8.nvt_3.cpt -p topol.top -o 9.nvt_4.tpr -n index.ndx
gmx mdrun -deffnm 9.nvt_4

#Equilibration npt_5
gmx grompp -f npt_5.mdp -t 9.nvt_4.cpt -p topol.top -o 10.npt_5.tpr -n index.ndx
gmx mdrun -deffnm 10.npt_5

#Production npt_6
gmx grompp -f npt_6.mdp -t 10.npt_5.cpt -p topol.top -o 11.npt_6.tpr -n index.ndx
gmx mdrun -deffnm 11.npt_6

#Production npt_9
gmx grompp -f npt_9.mdp -t 11.npt_6.cpt -p topol.top  -o 12.npt_9.tpr -n index.ndx
gmx mdrun -deffnm 12.npt_9
```

You can get a look at each MDP file to get the parameters used at each step.