Tutorial 2. SIRAH forcefield in GROMACS

Hybrid solvation: Plugging SIRAH solvent to your atomistic system

By Matias Machado
Mail any comment or suggestion to *spantano@pasteur.edu.uy*

This tutorial shows how to apply the hybrid solvation approach of SIRAH forcefield to speed up the simulation of an atomistic solute. The system contains a DNA fiber surrounded by a shell of atomistic waters, which are embedded in coarse-grained (CG) molecules, called WT4, to represent bulk water. The general procedure is extensible to any other solute. The hybrid solvation methodology is well tested for SPC, SPC/E, TIP3P atomistic water models. The main references for this tutorial are: *Another Coarse Grain Model for Aqueous Solvation: WAT Four?* Darré L, Machado MR, Dans PD, Herrera FE, Pantano S. JCTC. 2010, 6:3793. *Mixing Atomistic and Coarse-Grain Solvation Models for MD Simulations: Let WT4 Handle the Bulk.* Darré L, Tek A, Baaden M, Pantano S. JCTC. 2012, 8:3880. We strongly advise you to read these articles before starting the tutorial.

Software required
GROMACS 4.5.5 or later version properly installed in your computer. A molecular visualization program, we recommend VMD (freely available at: *http://www.ks.uiuc.edu/Research/vmd/*)

Previous knowledge required
How to perform a standard atomistic molecular dynamic simulation with GROMACS.

Hands on
0) Download the file *sirah.gmx.tgz* from *www.sirahff.com* and uncompress it into your work directory:

```
tar -xzvf sirah.gmx.tgz
```

You will get a folder *sirah.ff/* containing the forcefield definition, a set of useful tools in *sirah.ff/tools/* and the required material to perform the tutorial in *sirah.ff/tutorial/2/*.

Make a new folder for this tutorial:

```
mkdir tutorial2; cd tutorial2
```

1) Convert the PDB file of the solute *dna.pdb* into Gromacs format:

```
pdb2gmx -f ../sirah.ff/tutorial/2/dna.pdb -o dna.gro
```

Choose AMBER99SB and then TIP3P as water model.

2) For Gromacs to recognize SIRAH, edit your topology file *topol.top* adding the following lines **after** the forcefield definition:

Topology before editing	Topology after editing

`; Include forcefield parameters`
`#include "amber99sb.ff/forcefield.itp"`
	`; Include forcefield parameters`
	`#include "amber99sb.ff/forcefield.itp"`
	
	`; SIRAH forcefield`
	`#include "../sirah.ff/hybsol_comb2.itp"`
	`#include "../sirah.ff/solv.itp"`

Notice: *hybsol_comb2.itp* is a parameter file, while *solv.itp* links to the topologies of WT4 and SIRAH ions. The choice of the right parameter file depends on the chosen atomistic forcefield (see Table 1). The plug-in works smoothly with the implemented forcefields of the GROMACS distribution. When using a customized forcefield (e.g. for lipids) choose the parameter file according to the combination rule and check that the atom type you are using for the atomistic water match the DEFAULT in the SIRAH parameter file, which is OW. If they don't match then rename OW accordingly to your definition.

3) Solvate the system

Define the simulation box of the system:

```
editconf -f dna.gro -o dna_box.gro -bt octahedron -d 2
```

Then add an atomistic water shell of 1 nm around the solute:

```
genbox -cp dna_box.gro -cs spc216.gro -o dna_shell.gro -shell 1
```

Finally add the CG solvent:

```
genbox -cp dna_shell.gro -cs ../sirah.ff/wt4tip3p.gro -o dna_solv.gro
```

Update *[molecules]* section from *topol.top* to include the number of SOL and WT4 molecules:

Hint! You can use the following command line to get these numbers

```
grep -c OW dna_solv.gro; grep -c WP1 dna_solv.gro
```

Topology before update	Topology after update

`[ molecules ]`
`; Compound  #mols`
`DNA_chain_A  1`
`DNA_chain_B  1`
	`[ molecules ]`
	`; Compound  #mols`
	`DNA_chain_A  1`
	`DNA_chain_B  1`
	`SOL  3578`
	`WT4  2663`

Add counterions:

```
grompp\
 -f ../sirah.ff/tutorial/2/em_HYBSOL.mdp\
 -p topol.top\
 -c dna_solv.gro\
 -o dna_solv.tpr

genion -s dna_solv.tpr -o dna_ion.gro -np 38 -pname NaW
```

When prompted, choose to substitute NaW ions for WT4 molecules.

Notice: The available ionic species in SIRAH forcefield 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. Be aware that SIRAH ions remain within the CG phase. So, if the presence of atomistic electrolytes in close contact with the solute is important to describe the physics of the system you will have to add them.

Update *[molecules]* section in *topol.top* to include the 38 NaW ions and the correct number of WT4.

Before running the simulation it may be a good idea to visualize your molecular system. CG molecules are not recognized by molecular visualizers and will not display correctly. To fix this problem you may generate a PSF file of the system using the script *top2psf.pl*:

```
../sirah.ff/tools/top2psf.pl -i topol.top -o dna.psf
```

Notice: This is the basic usage of the script *top2psf.pl*, you can learn other capabilities from its help:

```
../sirah.ff/tools/top2psf.pl -h
```

Use VMD to check how the hybrid system looks like:

```
vmd dna.psf dna_ion.gro -e ../sirah.ff/tools/sirahradii.tcl
```

Notice: VMD assigns default radius to unknown atom types, the script *sirahradii.tcl* sets the right ones.

4) Run the simulation

You can find typical input files for energy minimization, equilibration and production runs in *em_HYBSOL.mdp*, *eq_HYBSOL.mdp* and *md_HYBSOL.mdp*, respectively. Please check carefully the input flags therein.

Create an index files adding a group for WT4 and NaW:

```
echo -e "r WT4 | r NaW\nq\n" | make_ndx -f dna_ion.gro -o dna.ndx
```

Make a new folder for the run:

```
mkdir run; cd run
```

Energy Minimization:

```
grompp\
 -f ../../sirah.ff/tutorial/2/em_HYBSOL.mdp\
 -p ../topol.top\
 -n ../dna.ndx\
 -c ../dna_ion.gro\
 -o dna_em.tpr

mdrun -deffnm dna_em &
```

Equilibration (NPT):

```
grompp\
 -f ../../sirah.ff/tutorial/2/eq_HYBSOL.mdp\
 -p ../topol.top\
 -n ../dna.ndx\
 -c dna_em.gro\
 -o dna_eq.tpr

mdrun -deffnm dna_eq &
```

Production (10ns):

```
grompp\
 -f ../../sirah.ff/tutorial/2/md_HYBSOL.mdp\
 -p ../topol.top\
 -n ../dna.ndx\
 -c dna_eq.gro\
 -o dna_md.tpr

mdrun -deffnm dna_md &
```

That's it! Now you can check the simulation using VMD:

```
vmd ../dna.psf dna_md.xtc
```

Table 1. SIRAH parameter file to include in the topology according to the chosen atomistic forcefield.

		Atomistic forcefield				
Parameter file	Combination rule	GMX	GROMOS	AMBER	CHARMM	OPLS
hybsol_comb1.itp	1	X	X			
hybsol_comb2.itp	2			X	X	
hybsol_comb3.itp	3					X