

The pipeline/Procedure for:
SAXS/WAXS calculations using all-atom MD simulations

Table of Contents

<i>The pipeline/Procedure for:.....</i>	<i>- 1 -</i>
<i>SAXS/WAXS calculations using all-atom MD simulations.....</i>	<i>- 1 -</i>
1. INSTALLATION DETAILS.....	- 3 -
I. Installing program.....	- 3 -
2. DESCRIPTION OF INPUT FILES.....	- 3 -
3. HOW TO RUN STEPS.....	- 3 -
4. OUTCOME OF PROGRAM AND HOW TO ANALYZE IT.....	- 8 -
I. Plot distance distribution function $P(r)$	- 8 -
II. Plot SAXS/WAXS profile $I(q)$	- 8 -
5. ERROR REPORTS AND FIXES IF DONE.....	- 8 -
1) Error 1: one-line description.....	- 8 -
2) Error 2: one-line description.....	- 8 -
3) Error 3: one-line description.....	- 8 -

1. INSTALLATION DETAILS

I. Installing program

Step 1: Install the custom version of GROMACS from this link:

<https://gitlab.com/cbjh/gromacs-saxs/-/wikis/Clone-and-run>

Step 2: Ask IT for help in environment configuration (GPU/CUDA, etc.)

Step 3: Confirm the path of the software for future use:

e.g. `/home/user111/saxs/gromacs-saxs/build/dist/bin/GMXRC`

Step 4: Installation done

2. DESCRIPTION OF INPUT FILES

Buffer	Solution	Description
<code>solvent-topol.top</code>	<code>saxs-topol.top</code>	Original topologies produced by GROMACS (<i>copy topol.top >>> solvent/saxs-topol.top</i>)
<code>solvent.pdb/.gro/.tpr</code>	<code>solution.pdb/.gro/.tpr</code>	Structure reference file (product structure)
<code>solvent.xtc</code>	<code>solution.xtc</code>	Selected frames
	<code>index.ndx</code>	Index file

3. HOW TO RUN STEPS

Generating Cromer-Mann Parameters

Step 1: Make two new folders named as “Solution” and “Buffer”

Step 2: Copy the required files (`saxs-topol.top`, `solution.pdb/.gro`, `solution.xtc`, `index.ndx`, etc.) to your “Solution” folder.

- For S/WAXS computing, you actually don't need too many frames (the more frames you take, the longer time it will run.). If that's the case, select the frames by:

```
> gmx trjconv -f solution.xtc -s solution.tpr -dt XXX -center (y/n) -o solution_n.xtc
```
- If your initial box dimension is not big enough, centering the trajectory of “solute” may help: use `-center` flag

Step 3: Copy the required files to your ‘Buffer’ folder.

- Note that you need a simulation which includes only the ions and water but mimics the ion strength of “Solution” system. For example, you already have simulation of RNA molecule in 200mM KCl [Solution], you need to run a short (~20ns) MD simulation of pure 200mM KCl [Buffer].

Step 4: Go to “Solution” folder, create a dummy tpr file:

- First of all, enable the SAXS-version GROMACS engine:


```
module purge
module load all gcc cuda/10.0
source /home/user111/saxs/gromacs-saxs/build/dist/bin/GMXRC
```
- Make an empty mdp file and generate the dummy.tpr file as well as an index file


```
> touch empty.mdp
> gmx trjconv -f empty.mdp -c solution.gro -p saxs-topol.top
-
o dummy.tpr
> gmx make_ndx -f solution.gro -o index.ndx
```

Step 5: Generate a Cromer-Mann parameters file for your molecule of interest

- Prepare Cromer-Mann parameters file named:


```
“scatter_<name-of-molecule>.itp”
> gmx genscatt -s dummy.tpr -vsites
> Select the molecule group of interest
```
- Check the resulting file and make sure that no line reads: CROMER_MANN_UNKNOWN.
- Copy explicit parameters file cromer_mann_defs.itp to “Solution” folder.
The file can be found in GROMACS installation folder:


```
> cp
./../gromacs-saxs/build/dist/share/gromacs/top/cromer-mann-defs.itp .
```
- Replace each CROMER_MANN_<element> macro definition with explicit parameters from “cromer_mann_defs.itp”
Example:


```
find CROMER_MANN_O in “cromer_mann_defs.itp”
#define CROMER_MANN_O .....

> sed -i -e 's/CROMER_MANN_O/...../g'
scatter_<name-of-molecule>.itp
```
- Include this generated scattering topology file in your saxs-topol.top and make sure it's added before the lines of tip3p.itp (take tip3p water as an example) and ions.itp:


```
; Include Position restraint file
#ifdef POSRES
#include "posre.md.itp"
#endif

; Include scattering topology ;
#include "scatter_RNA.itp"

; Include water topology
#include "../amber14sb.ff/tip3p.itp"
```

- Backup the force field folder:

```
> cp -r amber14sb.ff amber14sb_scatt.ff
```
- Go to the `amber14sb_scatt.ff` folder, the scattering parameters of ions and water molecules should also be added to `ions.itp` and `tip3p.itp`, respectively.

o `tip3p.itp`:

Check file “`cromer_mann_defs.itp`”, find water parameters:

```
#define CROMER_MANN_Owat .....
#define CROMER_MANN_Hwat .....
```

At the bottom, add these cromer_mann parameters:

```
#ifdef SCATTER
[ scattering_params ]
; i ft a1-a4, bi-b4, c
1 1 1.74364 2.56338 3.42821 0.95684 0.33326 5.70713 13.32240 33.18740 0.26793 ;Owat
2 1 0.07852 0.26770 0.16044 0.01140 3.20156 10.91840 28.25590 69.45850 0.00195 ;Hwat
3 1 0.07852 0.26770 0.16044 0.01140 3.20156 10.91840 28.25590 69.45850 0.00195
#endif
```

o `ions.itp`:

Check file “`cromer_mann_defs.itp`”, find ion parameters: (e.g. KCl)

```
#define CROMER_MANN_K_plus .....
#define CROMER_MANN_Cl_minus .....
```

Add these cromer_mann parameters to `[moleculetype]` K and Cl, respectively.

```
[ moleculetype ]
; Name
K

atoms
18 1 18.29150 7.20840 6.53370 2.33860 0.00660 1.17170 19.54240 60.44860 -16.37800

#ifdef SCATTER
[ scattering_params ] ; Cl_minus
; i ft a1-a4, bi-b4, c
1 1 18.29150 7.20840 6.53370 2.33860 0.00660 1.17170 19.54240 60.44860 -16.37800
#endif
```

- Now `amber14sb_scatt.ff` is ready to for use, edit the forcefiled information in `saxs-topol.top` file:

```
sed -i -e 's/amber14sb.ff/amber14sb_scatt.ff/g'
saxs-topol.top
```

Building the envelope

Step 1: In the “Solution” folder, create a structure reference file (gro/pdb) and a trajectory file (xtc) that includes ONLY the molecule of interest (e.g. RNA)

```
> gmx trjconv -f solution.xtc -n index.ndx -s dummy.tpr -o
solution_<molecule-name>.xtc
> gmx trjconv -f solution.xtc -n index.ndx -s dummy.tpr -dump
<time-frame> -o solution_<molecule-name>.pdb
```

Step 2: Generate the envelope files. The flag `-d` controls the size of the envelope as the distance from the surface of the molecule (typically 1.0 nm).

```
> gmx genenv -d 1.0 -f solution_<molecule-name>.xtc -s
solution_<molecule-name>.pdb
> The generated files envelope-ref.gro & envelope.dat are what
you
need.
> Note the information of pbc atom showed on the screen:
```

e.g.

```
##### G O O D P B C A T O M #####
N.B.: Solute atom number 427 is near the center of the bounding sphere -
it would make a good waxes-pbc atom (distance = 0.153606)
```

```
Global atom number = 427 (name N3, residue A-13)
#####
```

Creating the tpr files

Step 1: In the “Solution” folder, write an mdp file for scattering calculation: `rerun.mdp`

- The relevant differences to a usual .mdp file are highlighted below:

```
; File 'mdout.mdp' was generated
; By user: wh1165 (3032577)
; On host: login-0-2.local
; At date: Thu Oct 8 17:49:17 2020
;
; Created by:
; :-) GROMACS - gmX grompp, 2018.8-dev-20200318-a093ee8-unknown (-:
[.....]

; VARIOUS PREPROCESSING OPTIONS
; Preprocessor information: use cpp syntax.
; e.g.: -I/home/joe/doe -I/home/mary/roe
include =
; e.g.: -DPOSRES -DFLEXIBLE (note these variable names are case sensitive)
define = -DSCATTER

; RUN CONTROL PARAMETERS
integrator = md
; Start time and timestep in ps
tinit = 0
dt = 0.002
nsteps = 4000000
[.....]

; Scattering coupling stuff: xray and/or neutron (multiple neutron possible)
scatt-coupl = xray
; Selection of solute, solvent and fit group
waxs-solute = <name of solute> (e.g. RNA)
waxs-solvent = <name of solvent> (e.g. Water and ions)
waxs-rotfit =
; WAXS pbc atom list near center of solute (uses global indices, 0 = number-wise center, -1 = used atomic
distances)
waxs-pbcatom = <enter atom from step 2 of building the envelope> (e.g. 427)
; Coupling time constant (ps). (0=no averaging; -1=non-weighted average)
waxs-tau = -1
; Gradual switch target SWAXS curve from current to experimental curve within this time (ps)
waxs-t-target = 0
; WAXS coupling potential type (linear or log)
waxs-potential = linear
; WAXS coupling weights (uniform, exp, exp+calc, exp+solvdens, exp+calc+solvdens)
waxs-weights = uniform
; Extra overall force constant (one per scatt-coupl type)
waxs-fc = 1
; Frequency of spectrum calculation (steps)
waxs-nstcalc = 1
; Number of pure solvent structure factors to use (-1 = take from xtc)
waxs-nfrsolvent = 1000
; Frequency of output (steps)
waxs-nstlog = 1
; Nr of q values, starting and ending q (one per scatt-coupl type)
waxs-nq = 100 ; <nq>
waxs-startq = 0 ; <The range, nm-1>
waxs-endq = 10 ; <The range, nm-1>
; Nr of point for spherical quadrature (0 = auto)
waxs-nsphere = 1500
; Fit I(exp) on the fly to I(calc): no / scale-and-offset /scale
waxs-lexp-fit = no
; Electron density of solvent (0 = use xtc, 334 = experiment)
waxs-solvdens = 334
; Relative uncertainty of solvent density (e.g., 0.01 means 1% uncertainty)
waxs-solvdens-uncert = 0.005
; Bayesian sampling of solvent density uncertainty (no/yes)
waxs-solvdens-uncert-bayesian = no
; Buffer subtraction reduced by solute volume (no/yes)
waxs-correct-buffer = no
; Deuterium concentration (0 - 1), for each neutron group
waxs-deuter-conc =
; Scale I(q=0) to target pattern while coupling
waxs-scale-i0 = no
; Warn if solvation layer is thinner than:
waxs-warnlay = 0.3
; Anisotropic pattern: no, yes, cos2alpha
waxs-anisotropic = no
; Energy of X-ray beam
waxs-energy = 12
; WAXS ensemble type (none, bayesian-one-refined, maxent-ensemble)
waxs-ensemble-type = none
; Ensemble refinement - number of states (will read files intentisty_stateXX.dat)
waxs-ensemble-nstates = 0
; Ensemble refinement - initial weights and force constant
waxs-ensemble-init-w =
```

```
waxs-ensemble-fc = 0
```

```
[.....]
; Electric fields
; Format for electric-field-x, etc. is: four real variables:
; amplitude (V/nm), frequency omega (1/ps), time for the pulse peak (ps),
; and sigma (ps) width of the pulse. Omega = 0 means static field,
; sigma = 0 means no pulse, leaving the field to be a cosine function.
electric-field-x = 0 0 0 0
electric-field-y = 0 0 0 0
electric-field-z = 0 0 0 0
```

- Create the final .tpr file for the “Solution” system:

```
> gmx grompp -f rerun.mdp -p saxs-topol.top -c solution.gro
-n
index.ndx -o saxs.tpr
```

Step 2: Go to “Buffer” folder:

- Copy amber14sb_scatt.ff from “Solution” folder
- Copy rerun.mdp from “Solution” folder, but renamed to solvent.mdp. Erase the waxs_pbcatom and the solute group name. Keep the solvent group name.

```
; Selection of solute, solvent and fit group
waxs-solute = <name of solute> (e.g. RNA)
waxs-solvent = <name of solvent> (e.g. Water_and_ions)
waxs-rotfit =
; WAXS pbc atom list near center of solute (uses global indices, 0 = number-wise center, -1 = used
atomic distances)
waxs-pbcatom = <enter atom from step 2 of building the envelope> (e.g. 427)
```

- Similarly, don’t forget to edit the forcefiled information in solvent-topol.top file:
sed -i -e 's/amber14sb.ff/amber14sb_scatt.ff/g'
solvent-topol.top
- Create the final .tpr file for the “Buffer” system:
> gmx grompp -f solvent.mdp -p solvent-topol.top -c
solvent.gro -n -o solvent.tpr

Running the scattering

Step 1: All files are ready now. Make a slurm script job.sh:

```
#!/bin/bash
#SBATCH -n 1
#SBATCH -c 12
#SBATCH --time=48:00:00
#SBATCH -p nvidia

# load modules
module purge
module load all
module load gcc cuda/10.0

# import gmx-saxs
source /...../saxs/gromacs-saxs/build/dist/bin/GMXRC

# define the following environment variables
export GMX_WAXS_FIT_REFFILE=solution/envelope-ref.gro
export GMX_ENVELOPE_FILE=solution/envelope.dat

# run
srun gmx mdrun -s solution/saxs.tpr -rerun solution/solution.xtc -sw buffer/solvent.tpr -fw buffer/solvent.xtc
```

Step 2: Submit the job

Step 3: Several output files will be obtained, the most important of which are waxs.log (contains radius of gyration from Guinier fit together with other statistics) and waxs_final.xvg (the actual intensity as a function of q).

4. OUTCOME OF PROGRAM AND HOW TO ANALYZE IT

I. Plot distance distribution function $P(r)$

Obtain the pair distance distribution function using [GNOM](#):

```
> datgnom4 waxes_final.xvg -r <rg from Guinier approx> -o  
  pairwise_dist.dat  
> p 'pairwise_dist.dat' u 1:2 w 1
```

II. Plot SAXS/WAXS profile $I(q)$

```
> p 'waxes_final.xvg' u 1:2 w 1
```

5. ERROR REPORTS AND FIXES IF DONE

1) Error 1: one-line description

What: Segmentation fault when preparing the files

How to fix: Make sure that you used the modified GROMACS to generate all files.

2) Error 2: one-line description

What: Envelope do not fit into solvent/buffer box.

How to fix: Simulate a larger solvent box.

3) Error 3: one-line description

What: The envelope does not fit into solute/solution box.

How to fix: Use trjconv to center your trajectory; Use a smaller envelope; Rerun the “Solution” system with larger box.