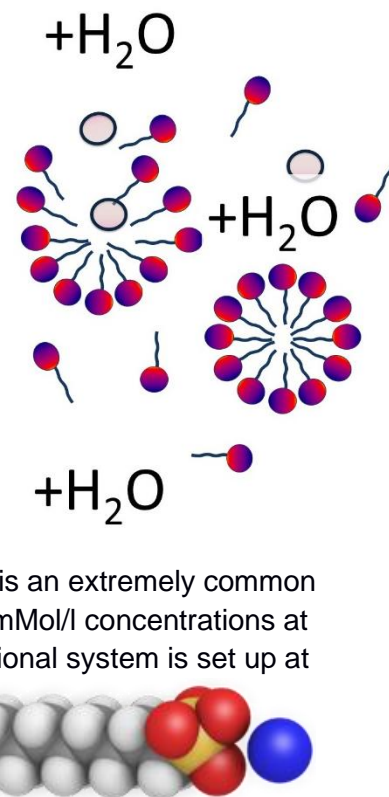


CHEM-E4115 Molecular Dynamics Exercise 3: Surfactants in water and at water-air interface: molecular dynamics by Gromacs

In the exercise, you study the behavior of surfactants in water and at water-air interface via molecular simulations. Surfactants are molecules that lower the surface tension between two liquids or between a liquid and a solid. They are important, for example, as detergents, solubilization agents, dispersants, emulsifiers, and processing agents. Surfactants and their aggregates also play a key role in basic biology for example in cell membranes and trafficking. See for example, Wikipedia for further information about surfactants and their solutions.

This particular exercise examines sodium dodecyl sulfate (SDS, $\text{CH}_3(\text{CH}_2)_{11}\text{SO}_4\text{Na}$) solvated in water. The setup is so that it enables examining the surfactants both in water and at water-air interface. SDS is an extremely common anionic surfactant, which in bulk water solution forms micelles above 8 mMol/l concentrations at temperatures higher than approximately room temperature. Our simulational system is set up at $T=323\text{K}$ and the concentration is significantly above micellization concentration.



The aims of the exercise are to learn and practice

- Constructing, initializing and running a real aqueous environment molecular simulation
- Doing basic structural analysis of an aqueous environment molecular simulation
- Critical assessment of simulations and data produced by them

Exercise 3: Set up this system, make a hypothesis what you expect the simulation outcome to be, and get at least one production run running.

In exercise 4, you can continue with the results and analysis, comparison of the findings with your hypothesis and meaning of the differences. There will be another, Monte Carlo task for exercise 4 but the analysis of this one cannot be completed during Ex 3.

- **Gromacs module is loaded in Mylly2 by *module load gromacs/2023*** You need to give this command to be able to run Gromacs commands in Mylly2
- Gromacs preprocessing (=preparing the files for simulation) can be done in Mylly2 head node.
- **Energy minimization, the relaxation runs and the production run etc.** should be submitted to the queue system by *jsub -np 4 gromacs mdrun [the remainder of the command]* or for the longer runs *jsub -np 12 gromacs mdrun [the remainder of the command]*.
- Note that the *gmx* of tutorial is replaced by *gromacs* in Mylly2 submit commands.

The MyCourses folder for this exercise contains

- The coordinates of 1 SDS molecule (Na⁺ counterion not included)
- A topology file for the SDS molecule (please note that for SDS it is not possible to generate the topology in the same way as for proteins). This SDS topology has been generated by the LigParGen server at <http://zarbi.chem.yale.edu/ligpargen/index.html> . The partial charges have been modified from those of the server to add up to -1 unit charge.
- mdp files for adding the ions, running energy minimization and a 5ns NVT simulation
- **The files are also available in Mylly2 directory /home/maria/sds** and can be copied to your current directory by `cp /home/maria/sds/* ./`[enter]
- Files to get exercise points
 - 1) the final topology file (top-file) and 2) either a VMD snapshot or a gro file of 40/80 SDS molecules

Preparing and setting running a simulation system with a water-air interface with surfactants in the water

1. Construct a simulation box that contains N solvated SDS molecules in random placement and their counterions.
 - a. To get a simulation that runs <1h with 12 processors in Mylly2, put 40 SDS molecules in a 5nm x 5 nm x 5nm simulation box. To get a simulation that runs ~4h with 12 processors, use a system size of 80 SDS molecules in a 6nm x 6nm x 6nm box. Instructions below are for 40 SDS, if you use a larger system, change the numbers.
 - b. The steps to do this are: Insert the 40 SDS, solvate the SDS molecules by water, add 40 Na ions (counterions of the surfactants) to make the system charge neutral by adding ions and finally run energy minimization on the system.
 - c. Tips for doing all this in Mylly2
 - i. Help on any Gromacs command “`gmx help [command]`”
 - ii. Inserting molecules “`gmx insert-molecules ...`”
 - iii. Solvating molecules “`gmx solvate ...`”
 - iv. Adding ions (as in exercise 1, the tutorial): First “`gmx grompp ...`” then “`gmx genion ...`”
 - v. Energy minimization (as in exercise 1): First “`gmx grompp ...`” then “`jsub -np 12 gromacs mdrun ...`”
 - vi. vmd is good for checking whether each step has succeeded. Vmd shows the simulation box if you provide the command “`pbx box`” in the vmd terminal window.
 - vii. **If the software complains about the number of atoms not matching, check that the number of SDS molecules in your .top file matches with the coordinate file**

2. Make the aqueous solution of the SDS molecules and the Na⁺ ions into a system that has an air interface (actually 2 interfaces because of the periodic boundary conditions).
 - a. Tip: Add 8 nm vacuum (air) in the Z direction by using “gmx editconf ...” and energy minimize the resulting simulation configuration. Check by VMD that you actually have a system that has surfactants and ions in water surrounded by a vacuum area.
 - i. The energy minimization may complain about water molecules that “cannot be settled”. These are water molecules at high-energy configurations at the interface. Typically, the algorithm manages to achieve convergence but if yours does not, return to the step where the SDS were solvated.
3. Start an NVT simulation using the provided mdp file on this system. The run will take ~2h so ideally leave it running in Mylly2 and continue with the analysis in the next exercise.
4. Make a hypothesis what you expect the simulation outcome to be.
 - a. What is your hypothesis about what will happen in this surfactant system? What do you expect will happen in a real surfactant solution? What do you expect will happen in this simulation? Why? Do these differ? Again, why?
5. The analysis can be started already when the simulation is running (1-2ns of trajectory is sufficient to get sensible data).

Results and analysis

What is your hypothesis about what will happen in this surfactant system? What do you expect will happen in a real surfactant solution? What do you expect will happen in this simulation? Why? Do these differ? Again, why?

Analyse the simulation for

1. What is the location of water, the SDS molecules and the NA ions in the simulation box? What are your observations and what is the significance of your observations?
 - a. Tips for doing this
 - i. Periodic boundary conditions cause the molecules to split across the boundary. To get a trajectory file without split molecules, run *gmx trjconv -s [tpr file] -f [trajectory file] -o [output trajectory file name] -pbc mol -ur compact*
 - ii. Vmd (visual analysis)
 - iii. “gmx density …” and xmgrace
 - iv. xmgrace [data set 1] [data set 2] [data set 3] plots three different data sets in the same graph
2. Analyse the NA ion distribution with respect to SDS head group positions and the solvation shell of the NA ions in this model.
 - a. Tip “gmx make_ndx” with commands “a OA*”, “a OW”, “a HW*” to create an index file with the sulfonate oxygens, water oxygens and water hydrogens as groups with respect you can calculate distributions. For simplicity, all 4 sulfonate oxygens are in the same group although more detailed info would be obtained by differentiating between them.
 - b. Where are the NA counter ions with respect to locations of the SO₄ head groups of SDS (the O atoms)? What is the significance of your observations regarding the ions?
 - i. Tip “gmx rdf …” with the above created index file.
 - c. Analyse a bit the NA ion solvation in this model. How thick or large is the NA solvation shell? Does the data capture multiple solvation shells?
 1. Tip “gmx rdf …”. Gromacs also contains dipole orientation analysis tools
3. Conclusions / discussion
 - a. What are your observations regarding the surfactant system? Do these match or differ from your hypotheses? Why?
 - b. Which ones of these result from the simulation setup or model used and which ones are “real” findings? Why?
 - c. What is the significance or the conclusions?
 - d. Could the setup be improved for better description of the system?
 - e. What kind of properties could be mapped via simulations better than experimentally?
 - f. Calculate the SDS concentration in the simulation. How does this value compare with your favorite SDS buffer or some concentration of SDS you have run into in the lab? What is the significance?