The Jyväskylä Summer School 2017

NANO1. Demo 08-08-2017

In silico titration of Asp-Ala-His peptide

Let's condider a short peptide of sequence Aspartic acid, Alanine, Histidine (Asp-Ala-His). This is one of the peptides you worked with during yesterday's demo. This peptide contains two titratable groups, Aspartic acid and Histidine, which can be protonated or deprotonated. The N-terminus and C-terminus of the peptide are capped with N-methyl and Acetyl groups, respectively, and are, therefore, neutral, and cannot change protonation state. To model the protonation of this peptide at different pH, we will use today molecular dynamics simulations at constant pH.

Molecular dynamics at constant pH

Molecular dynamics (MD) simulations are used to model the behaviour of a system as a function of time. Your system consist here of a short peptide in water. Note that water is modelled here explicitely, i.e, in addition to the peptide there are also thousands of water molecules. We provide the program the coordinates of each atom of the system (asp-ala-his-water.pdb), a topology file (topol.top) to model the interactions between each atom in the system, and a file (md.mdp) with the parameters of the simulation (lenght, temperature, pressure, etc.). The molecular dynamics simulation program calculates the interactions between the atoms in the system at every time step in the simulation (0.002 picoseconds), and re-computes the coordinates at the following time step, by applying Newton's equations of motion. The final result is a trajectory of your system, where the positions of the peptide and water atoms are stored at each time frame. In constant pH MD, also the protonation state of the titratable groups in the peptide can change during the simulation, and it is stored in a similar way, at each time frame.

Running the constant pH MD

We will run the constant pH MD at the computer science center (CSC) in Espoo. For this you need to login at CSC via a terminal connection, in the same way as you did yesterday. Open the WinSPC program (under All programs > WinSCP) and type under the host name

taito.csc.fi

and your CSC username and password.

Then from the Menu *Commands* choose *Open in PUTTY*. You are now connected to the taito server at CSC via a terminal.

The server uses UNIX based commands. For a quick reference guide you can download
https://research.csc.fi/documents/48467/85840/CSC+Quick+Reference.pdf

Now type in the terminal the following commands to move to the working directory and copy the files you need

cd $WRKDIR
cp /appl/courses/JY_NANO1_summer_school_2017/GROMACS/cphmd_practical.tar.gzip .

Then

gunzip cphmd_practical.tar.gzip
tar -xvf cphmd_practical.tar
cd cphmd_practical/asp-ala-his
ls -l

You will find a directory called *run*. Move to the *run* directory. Choose a pH value at which you want to model your peptide, e.g. pH 4, and type this value in the ph.dat file. You can edit this file using the command

nano ph.dat

To start your job type

sbatch –reservation=training job

You have now submitted your job to the queue. You can check the status of your job with the command

squeue -u username

and list the files of your directory using

ls -l

To have a look at the system you are simulating, transfer the asp-ala-his-water.pdb file from the directory *asp-ala-his* in taito.csc.fi to your computer, and look at it using the program Rasmol (on your computer under All programs > RasWin). Use the Menu *File* and then *Open*

In the rasmol command window type

select 1-5
zoom selected

to center the peptide. You can change the representation of the peptide from the *Display* Menu of Rasmol.

Analyzing the constant pH MD

It will take few hours to complete your job. For this reason, in the directory *asp-ala-his/RESULTS*, you find completed jobs, at different pH values (directories *pH_X*), which you can analyze in the mean time. Choose a pH directory and list the files. In each pH directory, the molecular dynamics simulation has produced a trajectory file (traj_comp.xtc) and an energy file (ener.edr). You can check the manual of the GROMACS MD program to find how to look at these files (check chapters 8.2 and 8.3 to start with)

ftp://ftp.gromacs.org/pub/manual/manual-5.0.7.pdf

For this practical, we have used a special version of GROMACS which models the protonation state of titratable sites[Donnini *et al.*, 2011, Donnini *et al.*, 2016]. The protonation state is written as a function of simulation time (in picoseconds) in the file *l_dynamics_groups_X.dat*. The *X* is the number of the titratable group, in the order in which it appears in the file *lambda_groups.dat*. Transfer this file to your computer and plot it using the program Gunplot (or any other program you are familiar with). If you are using Gnuplot, from the Menu *ChDir* move to the directory where the file is, and type

plot "l_dynamics_groups_X.dat"

and to quit the program

quit

The X-axis is the time (in picoseconds) and the Y-axis is the protonation state of the titratable site. At a value of 0, the titratable group is protonated, and at a value of 1 it is deprotonated (note that this is the opposite of what you had in the practical yesterday!). At a value of 0.5 there is a half proton on the protonatable gruop.

Compare the *l_dynamics_groups_X.dat* at different pH values. Going from low pH to high pH the time spent at values of 1 increases. Calculate the average protonation state using

../../../calc_average l_dynamics_groups_X.dat

Collect the averages at different pH values and plot them as a function of pH (you can do this also on paper). You have now a titration curve for the titratable group. The pH at which the average protonation is 0.5 is the pKa of the group. What pKa do you find? If you sum the averages of the two titratable sites you can plot the titration curve for the whole peptide.
Compare the pKa value you obtained for the Aspartic acid and Histidine with the respective reference pKa values (refpKa) in the file pka_data.dat. The reference pKa is the pKa of the aminoacid alone, in water. It is often the case that the interactions of the aminoacid with the rest of the peptide shift the pKa value from its reference value. Discuss with your group what could be the reason for this shift.

The transition rate between the protonated and deprotonated states is a parameter of the simulation. Because values of protonation between 0 and 1 are not physical, we want to have fast transitions between the protonated (0) and deprotonated states (1). To achieve this, we define a barrier betweent the two states. Here the barrier height was set to 5 kJ/mol. To check the effect of increasing the barrier from 5 to 8 kJ/mol, compare the *l_dynamics_groups_X.dat* files at pH 4 in the directories *pH_4* and *pH_4_barrier8*. What is different in the two trajectories?

When the protonation state of a group changes, its charge changes. Aspartic acid is neutral when the protonation has a value of 0 and is charged -1 when the protonation has a value of 1. Histidine is +1 when protonation is 0, and neutral when protonation is 1. Calculate which is the average distance between the two groups at different pH values. Move to a *pH_X* directory and type

../calc_mindist

which executes a script to perform a GROMACS command to calculate the distance between the carboxyl oxygens of the Aspartic acid and the ND nitrogen of Histidine as a function of time (*mindist.xvg*). Note that here we deprotonate only one of the two nitrogens of Histidine, the ND nitrogen. You should know after yesterday's demo that the proton affinity of ND is lower than that of NE, therefore ND has a higher probability of being deprotonated than NE. If you are interested in using a more accurate model for Histidine check reference [Donnini *et al.*, 2011], or ask for a tutorial.
In the file *mindist.xvg*, the distance is plotted as a funtion of time. Collect the average distances at different pH values and plot them as a function of pH. How does the curve look like?
If you have time you can also try to see if there is a correlation between the distance and the protonation state of the titratable groups at a certain pH.

In silico titration of the Asp-Ala-Asp peptide

We considere here a second peptide of sequence Aspartic acid, Alanine, Aspartic acid. Also this peptide is familiar to you from yesterday. Repeat what you have done for the first peptide.

What pKa values do you obtain for the two Aspartic Acids? Is there a difference with respect to the reference pKa value of Aspartic acid? If yes, is this difference similar to the difference between calculated and reference pKa of Aspartic acid that you obtained for the Asp-Ala-His peptide?

How does the plot of the average distance between the two Aspartic acid as a function of pH looks like? Is it different from the average distance plot you obtained for the Asp-Ala-His peptide?

Discuss your results with your group. Did you get different results? What is your explanation for the results you obtained?

Compare the pKa's you obtained for the two peptides yesterday and today. Discuss with your group about

why are they simular or dissimilar, and about main differences between the methods you used during these two demos.

Towards a more accurate model of protonation equilibria in constant pH MD

We already mentioned that in our model only the protonation of one nitrogen of Histidine was modelled. In the same way also in Aspartic acid, only deptotonation of one oxygen was modelled. There are more complex models of these aminoacids which can be used to obatin a more realistic descritption of the peptide.

In our model the peptide was solvated in water. However, typically the solution contains also salt, which can screen interactions between charged groups, and therefore have an effect on the proton affinities. During our simulation the net charge of the system changed as the protonation state of the peptide changes. In reality, solutions are neutral. There are some methods to assure neutrality during constant pH MD simulations. If you want to know more about these check reference[Donnini *et al.*, 2016], or ask for a tutorial.

Can you think of something else which you can do to get more accurate pKa estimates of the simulated peptides?

References

[Donnini *et al.*, 2011] Donnini, S., Tegeler, F., Groenhof, G., & Grubmüller, H. (2011). *J. Chem. Theor. Comp.* **7**, 1962–1978.

[Donnini *et al.*, 2016] Donnini, S., Ullmann, R. T., Groenhof, G., & Grubmüller, H. (2016). *J. Chem. Theor. Comp.* **12**, 1040–1051.