The Jyväskylä Summer School 2017

NANO1. Demo 07-08-2017

Calculating protonation probabilities with MEAD and a Monte Carlo titration programme

In this section you will learn how to use the programme package MEAD, and the Monte Carlo titration programme mcti for calculating protonation probabilities.

Asp-Ala-Asp peptide

Let's condider at first a short peptide of sequence Aspartic acid, Alanine, Aspartic acid (Asp-Ala-Asp). This peptide contains two titratable Aspartic acids. 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 calculate the titration curves for the two Aspartic acids we will use the programme multiflex of MEAD and the programme mcti.

We will use these programs at the computer science center (CSC) in Espoo. For this you need to login at CSC via a terminal connection. Open the WinSPC program (under All programs > WinSCP) and type under the host name

taito.csc.fi

and your CSC username and password (which you will receive at the practical).

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/MEAD/mead_practical.tar.gzip .

Then

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

To have a look at the structure of the peptide, transfer the asp-ala-asp.pqr file 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. You can also plot distances between two atoms (which you select by clicking on them) using the command

set picking distance

Now, follow the list of commands given in master.sh, but run the different scripts one by one. To know more about the input files and the programs you are goint to use, you can look for the README file in the directories
/appl/courses/JY_NANO1_summer_school_2017/MEAD/programmes/mead-2.2.8a
and
/appl/courses/JY_NANO1_summer_school_2017/MEAD/programmes/mcti

Have a look into every script you run: try to understand what these scripts do, and how they do it. Do not bother too much about the perl scripts (*.pl). Also do not bother about the "used only once"-warnings issued by *collect_curves.pl*. When you have completed this section, please remove the rather large **.potat* files.

Once you have completed the commands in master.sh, you should have the file *curves.out*, which contains the titration curves of the two Aspartic acids of the peptide. Copy the file
cp curves.out curves_ASP-2.out
and edit the new file, e.g. using the program nano (or if you prefer you can copy the file to your computer and use Word or any program you are familiar with)
nano curves_ASP-2.out
Delete all lines in *curves_ASP-2.out* except the pH and protonation (prot.) columns of the first Aspartic acid (which is number 2, according the the numbering in the asp-ala-asp.pqr file). Do the same for the other Aspartic acid (ASP-4). Now, if you have not done that yet, move *curves_ASP-2.out* and *curves_ASP-2.out* to you computer, and plot the titration curves using the Gnuplot program (or any other program you are familiar with). If you are using Gnuplot, open the program, and from the Menu *ChDir* move to the directory where the file is. Then, type

plot “curve_ASP-2.out”,“curve_ASP-4.out”

or to plot with lines

plot “curve_ASP-2.out” with lines,“curve_ASP-4.out” with lines

and to quit the program

quit

In the plot, the X-axis is the pH, the Y-axis is the fraction of protonated acid (1 means all acid is protonated, 0 means all acid is deprotonated). You will see that the fraction of protonated acid decreases with increasing pH. The pKa corresponds to the pH at which the fraction of the protonated acid is 0.5 (pK(1/2)). Aspartic acid has a charge of -1 when deprotonated. Find out what is the charge of the asp-ala-asp peptide at pH 4.

The pKa’s of the two Aspartic acids are written in the file *pkhalf.out*. Are the two pKa’s the same? Discuss with your group why.

Compare the pKa values you obtained to the pKa of an Aspartic acid in water (the reference pKa), which is written on the first line of the file *ASP.st*. Are there differences? Discuss with your group why.

In the file *curves.out* you find also the Hill coefficient of each titratable group (“Slope of Hill plot”). Find out what is the Hill coefficient, and what it means when it is different from one.

Asp-Ala-His peptide

Let’s know consider a different peptide, which contains a Histidine aminoacid. Move to a different directory

cd $WRKDIR/mead_practical/asp-ala-his_eps_and_del

Have a look at the peptide (asp-ala-his.pqr) with Rasmol, and try to find where are the ND and NE nitrogens of Histidine. Histine has a +1 charge when it is protonated (both ND and NE are protonated). When Histidine is neutral, the proton is on the ND or on the NE nitrogen. These nitrogens are not equivalent, like the two oxygens of the carboxyl group of Aspartic acid. That is why there are two different files (*HISdel.st* and *HISeps.st*), which define the charges of Histidine in the protonated and deprotonated states. Also, you can read from the first line of the **.st*-files, that the reference pKa's differ, and NE has a higher affinity for the proton than ND. When calculating protonation probabilities one has to consider both possible protonation sites of Histidine. That is why the number of input and output files are larger than for the previous peptide.

Similarly to the previous peptide, run the commands in the *master.sh* file. Notice that now, the multiflex program is called more than once from the script *run_mol_multimead.sh*.

Look at the titration curve of Aspartic acid. What pKa value do you obtain? Compare the titration curves and pKa values of Aspartic acid obtained for the asp-ala-asp and asp-ala-his peptides. Do they differ? Discuss with your group what can be the reason for that.

Try to understand what the different columns in the Histidine sections of *curves.out* represent. To plot all the columns together with Gnuplot (after you have copied the curves of Histidine to a sepatate file, e.g. *curves-HIS-4.out*) use the command

plot "curves_HIS-4.out" using 1:2, "curves_HIS-4.out" using 1:3, "curves_HIS-4.out" using 1:4

What is the predominant protonation and/or tautomer state of Histidine at pH 7? Discuss with your group what could be the structural reason for that.

Look for the Hill coefficients of the Aspartic acid and Histidine in the file *curves.out*. What values they have? Discuss with your group why are they different from the Hill coefficients of the Aspartic acids in the asp-ala-asp peptide?

Cytochrome c

Now you should do a titration calculation on horse heart cytochrome c. Move to the directory

cd $WRKDIR/mead_practical/cytc

Have a look at the protein structure. Pay attention to the heme group and how it is coordinated by the protein.

Run a multiflex calculation as before, using the scripts listed in *master.sh*. From the list of **.st*-files provided, you can see which residues are in principle considered titratable. The titratable residues are listes in the **sites.del* and **sites.eps* files (the residue numbering corresponds to the numbering in the *cytc.pqr* file). Notice that the N-terminus and one of the Histidines are not titratable. Which one is the non titratable Histidine? Why is that?

The Monte Carlo titration calculation is run by the script *runmcti.sh*. Compare this script with the one of the previous calculation. What is different? Remember to remove the **.potat*-files when you are done.

Plot the titration curves of sites with pKa between pH 5 and 9. In addition, plot the titration curves of the heme propionates. What is peculariar about them? Find out what is the charge of the protein at pH 7.

Which are the titratable sites which show the largest deviation with respect to their reference pKa? And which are the residues with the smallest Hill coefficient? Discuss with your group what could be the structural reason for that.

For this practical we have used only one conformation of the protein (the one described by the coordinates

in the **.pqr* file). If you have read carefully the README for MEAD you have seen that it is possible to use more conformations of a protein. If you are done with the practical above, you can try to set up a calculation with more conformations for the peptides (ask from where to get them), or ask for a more advanced tutorial on this topic.