Exercise 05

Gromacs Tools for the Analysis of the MD Data

Deadline: Please hand in your protocol in pdf format by Thursday, 15^{th} June, 10.15 a.m. to **saleksic@zedat.fu-berlin.de**. The proctol should contain GROMACS commands, Python code, plots, and comments if necessary.

5.1 Root Mean Square Fluctuations (10)

Provide a GROMACS comand that calculate RMSF values for the C-α atoms of the each residue in Langerin sequence from the given trajectory. (1)
How many residues are in Langerin monomer? (1)
Visualize the RMSF plot via Python script **(file rmsf.xvg)**. List the residues that have the RMSF values greater then 2 $\mathrm{\AA}$ (name them by using one letter code). Include RMSF plot into your report. Script should contain the code for plotting, as well as for getting residues with the RMSF values greater then 2 $\mathrm{\AA}$. (8)

5.2 Root Mean Square Deviation (10)

Provide a GROMACS comand that calculate RMSD values for the all atoms in Langerin over the whole trajectory by taking into account every 10th frame. (2)
Visualize the RMSD plot via a Python script **(file rmsd.xvg)**. Include RMSD plot into your report and discuss the range of RMSD values observed in plot (report RMSD values in Å). (8)

5.3 Backbone Flexibility (40)

Provide a GROMACS comand that extracts the backbone torsion angle distributions Are values for GLY-198 and PRO-325 avaiable in the output file? Run the given bash script on the **rama.xvg** file **Extract_Backbone_dihedrals.sh** to obtain ascii files suitable for importing into Python. (5)
Write a Python script which produces Ramachandran plots for all the residues in the Langerin sequence. Pseudo code for the computing Ramachandran plots looks as following:

- Load ϕ-ψ trajectory of the respective residue
- Convert ϕ-ψ trajectory into the 2D histogram (hint: numpy.histogram2d) with 360 bins in range [-180,180],
- Modify 2D histogram, by taking logarithm of the non-zero entries
- Plot histogram as a heatmap (hint: imshow of matplotlib) with colorbar with ticks set in range [0,6], axis labels (check how to add greek letters as axis labels), axis thicks in range [-180,0,180], and save as a *.png figure with resolution of 96 dpi. Zero entries of the 2D histogram should be visualized in white. Close figure at each iteration. (25)

Condiseder residues GLY-248, ALA-258, GLY-262, PRO-283, and ALA-289, and discuss how many populations (minima) are in the respective Ramachandran plots ? Name visited regions of the Ramachandran plain and include plots in your report. (10)

5.4 Sidechain Flexibility (20)

Provide a GROMACS comand that extracts the sidechain torsion angle distributions **g_chi** command by considering the first 350 ns of the simulation. (4)
Are there output files for all 20 canonical amino acids? (1)
Write a Python script which plots χ-1, χ-2, χ-3, χ-4 angle distributions as normalized histograms (distribution probability should be normalized to 1). Histogram bins should be in the range [-180,180], and bin size should be 10°. Figure features such as axis labels, axis thicks, and title should be set accordingly. Output files of choice are called chi*LYS60.xvg. (15)

5.5 Computing Atom Distances (20)

Provide a GROMACS comand that extracts the distance between C-α atoms of M260, and G290 with **g_mindist** command. Modify given index.ndx file so that numbers correspond to the numbers of the respective C-α atoms in the langerin.pdb file. (5)
Plot distance distribution as a histogram. How many maxima can you observe in the histogram? (8)
Match the given files 1.pdb 2.pdb 3.pdb with the respective peaks of the histogram (hint measure the same distance in VMD, the value is given in Å. Visualize all three strucure in VMD by setting the view style as New Cartoon, highligth M260, and G290 in Licorice view style, and comment the loop mobility around M260. (7)

5.6 Files

Files can be downloaded under:

`https://www.dropbox.com/sh/or6h8q2zvfljfwe/AAAzIgASxX-UE8h3U_GPt_Kza?dl=0`