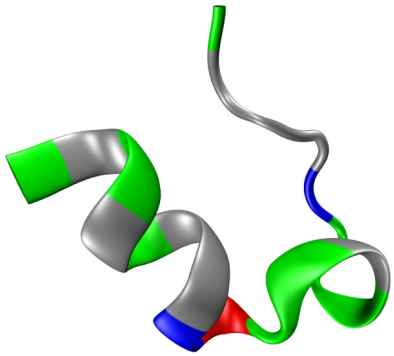
Exercise “Analysis of MD simulations”

In this hands-on session you will analyze a **100-ns MD simulation of the 20-residue trp-cage mini-protein (PDB 1L2Y)**. The MD simulation was initiated from the NMR structure:



The simulation details were as following:

- 1 ns equilibration
- 100-ns production run at 298 K and 1 atm (NPT ensemble)
- Nose-Hoover thermostat, Parrinello-Rahman barostat
- periodic boundary conditions with PME for electrostatics calculation
- 1 Cl^- ion for neutralization

The trajectory file is called “*trpc.xtc*”. Snapshots were saved *every 10 ps*. The energy file is called “*trpc.edr*”. The analysis consists of following steps:

1) Visualize the trajectory using VMD and show a picture of the protein.

2) Energy and temperature analysis : Plot

- the total energy, potential energy and kinetic energy of the system;
- the LJ and Coulomb (short-range) energy of the system, water and protein
- the temperature of the protein and the water.

Comment the plots (each question in **2-4 sentences**):

- Why is the kinetic energy positive and the potential energy negative?
- What is the largest contribution to the potential energy?
- Why are there differences in the temperature of the protein and the water?

3) RMSD, RMSF : Plot

- the RMSD for the backbone (fitting also applied to the backbone only)
- the RMSF for the C_α atoms

4) Radius of gyration :
Plot the radius of gyration (R_g).

5) Minimum distance between termini:
Plot the minimum distance between the N- and the C-terminus (d_{NC}).
Compare RMSD, RMSF, R_g and d_{NC} with each other? (Hint: Describe the above plots and explain the difference and similarities between them in **4-8 sentences**)

6) Hydrogen bonds: Plot the number of H-bonds between

- protein and protein
- protein and water

Shortly describe on the H-bond result in **2-3 sentences**.

7) DSSP analysis: Plot

- the time evolution of the secondary structure
- number of secondary stucture elements (helix, β-sheet and bridge plus turn) - c_{SS}

Describe the secondary structure transitions in the peptide in **2-4 sentences**.

8) Principal component analysis (PCA)
Perform the principal component analysis and plot the eigenvalue spectrum.
Describe, compare and explain the motion for the 1st and 50th eigenvalue in **2-4 sentences**.

9) Free energy surface (FES): Generate two different FESs using

- PCA *
- two order parameters of your choice, e.g., RMSD and R_g , RMSD and d_{NC} , or RMSD and c_{SS} etc.

Choose one of the two FESs and extract representative structures (maximal 2-3) for the free energy minima.

10) Clustering:
Cluster the sampled conformations (using the Daura algorithm) with a cut-off of 2 Å (option *-cutoff*, unit is in nm!), fitting and RMSD calculation applied to the C_α atoms and the output should be the whole protein.
Show the structures of the cluster centers for the 3 most populated clusters.
Compare the structures from 9) and 10) based on the similarity in **2-4 sentences**.

* **Help:**

PCA:

~$ gmx covar -f trpc.xtc -s trpc.tpr

~$ gmx anaeig -v eigenvec.trr -f input.xtc -noxvgr -s input.tpr -first 1 -last 2 -2d 2dproj.xvg

<remove all lines starting with “@” from 2dproj.xvg>

~$ python generateFES.py 2dproj.xvg -6 6 -6 6 100 100 300 fes-pca.out