Computational Biology Lab

Assignment 3 - BE21B037

Q1.a) What is RMSD? Why is it useful to measure RMSD?

RMSD stands for Root Mean Squared Deviation. It is useful to measure RMSD as it provides us the average distance between atoms and will give an idea of similarity between two sets of atomic coordinates.

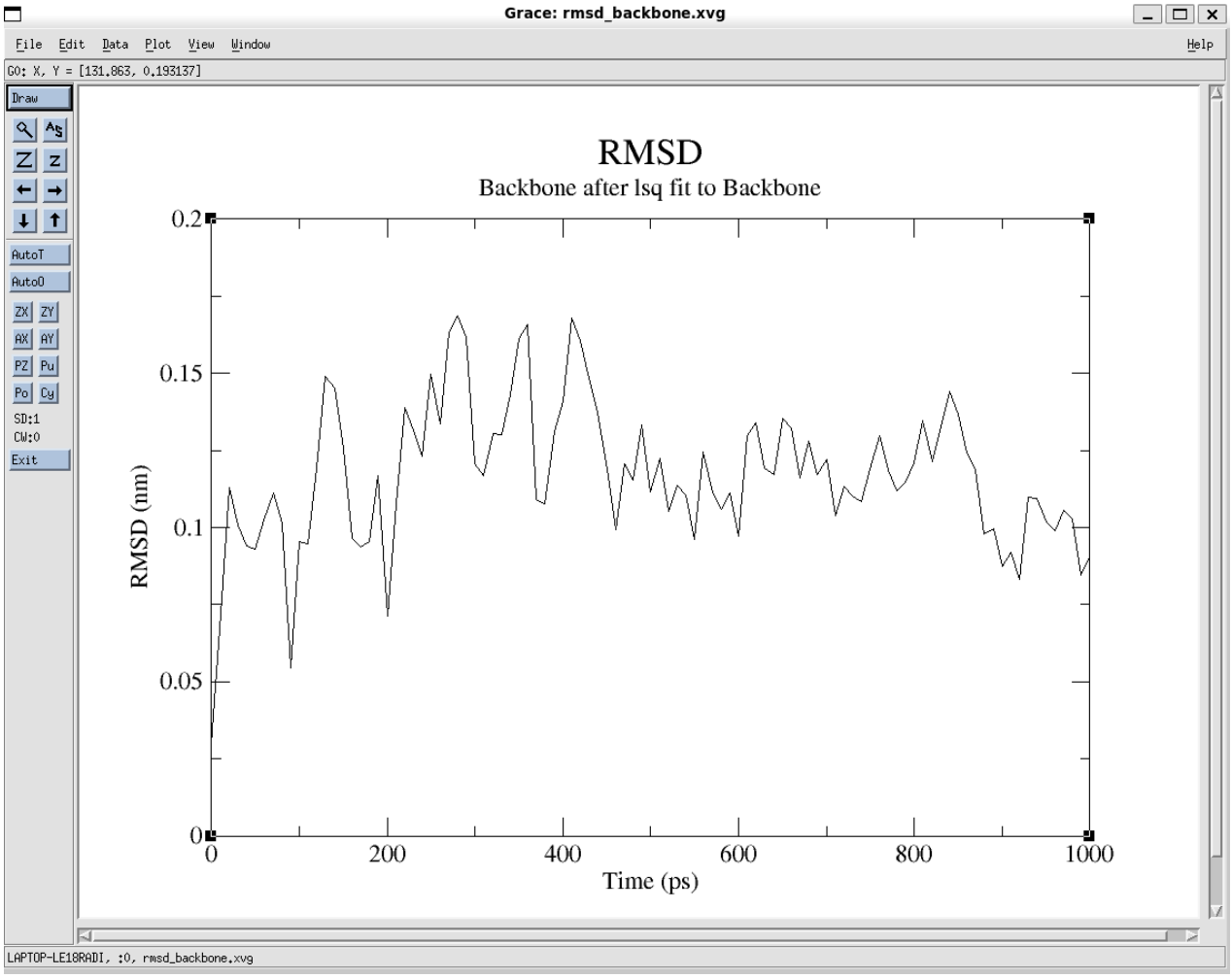
RMSD will provide us a sense of how much the molecule varies with respect to some set state.(ex:the minimum energy state)

b) Calculate the RMSD of the protein backbone after the production run with reference to the backbone of the crystal structure (em.tpr). Provide a plot in grace.

The code run to find the RMSD of the Protein backbone after the production run with reference to the backbone of the crystal structure was

gmx rms -s em.tpr -f md_noPBC.xtc -o rmsd_backbone.xvg

The output plot was plotted using xmgrace



Q2. a) What is Rg? If the Rg of a protein increases with time, what does it indicate?

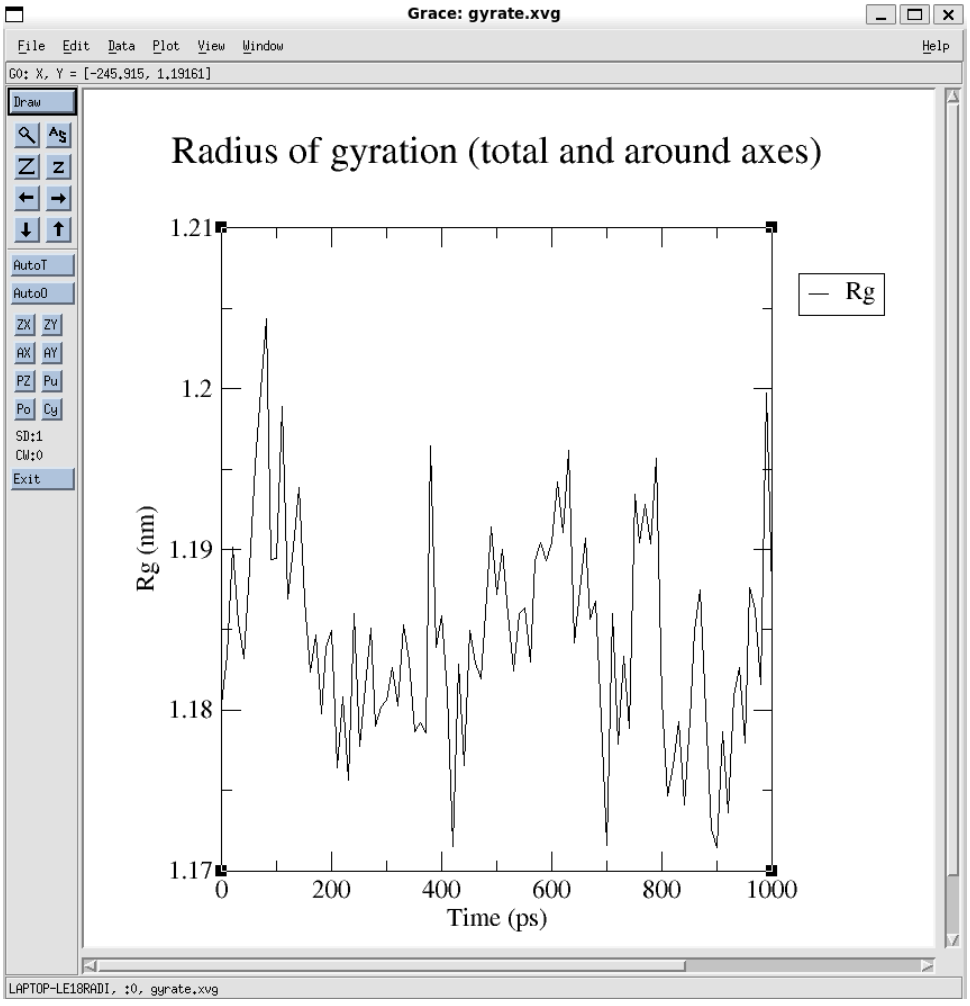
Rg stands for Radius of Gyration and is defined as the RMSD between each atom of the protein and the center of mass of the protein. It measures compactness of the protein, if the Rg of a protein increases with time then it means the protein is unfolding/opening up.

b) What is the Rg of the protein at the end of production MD? Plot Rg vs time.

The Rg value at the end of the production run is

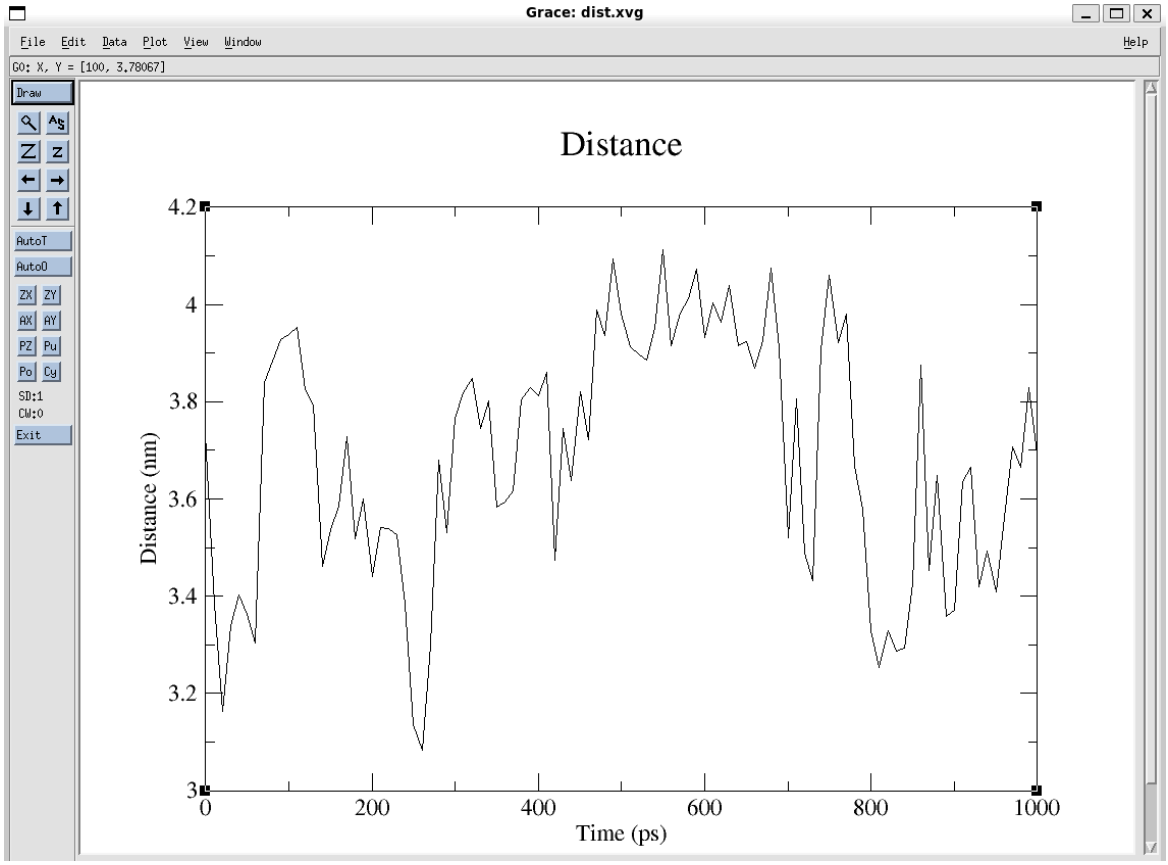
No.	Rg	Rg(x)	Rg(y)	Rg(z)
1000	1.18744	0.983947	0.912508	1.00955

The plot of Rg with respect to time is plotted using xmgrace



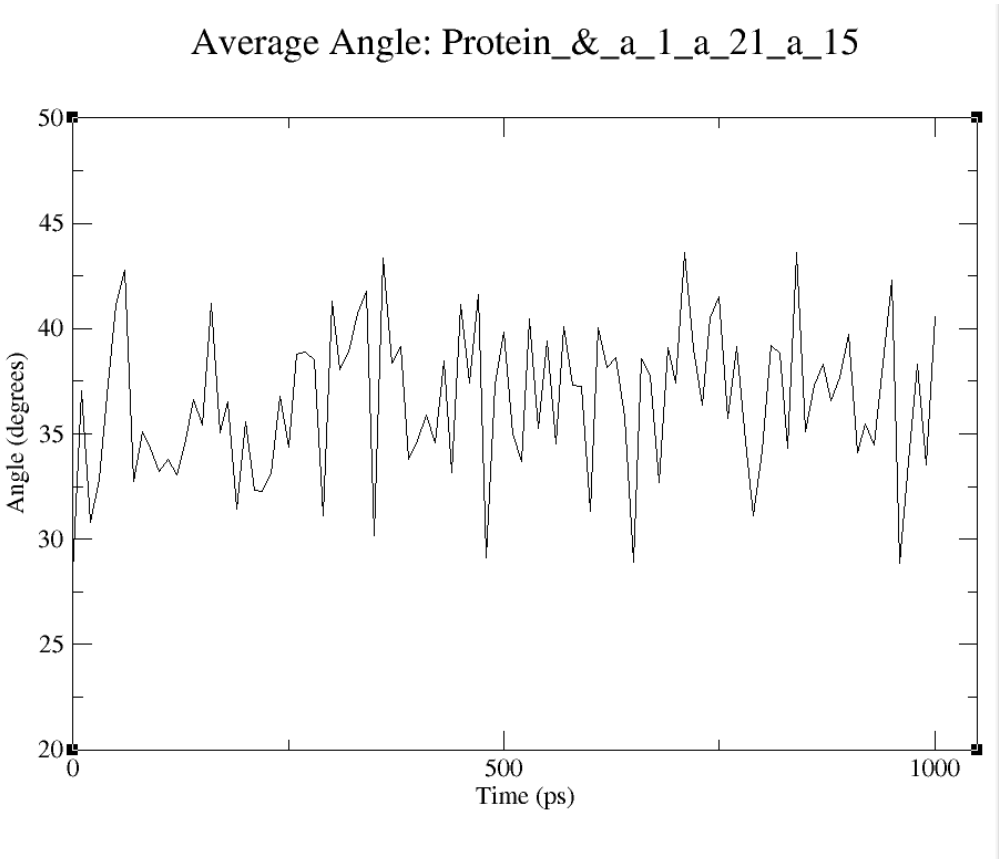
Q3. a) How does the distance between the first and last atoms of your protein vary over the course of the simulation? Provide a plot for the same.

The distance between the first and last atom varies around the 3.6nm range, having the largest distance at 500-600 ps timesteps, and reaching the closest at the 260 time step with a distance of 3.084nm.



b) Calculate the Angle between any three atoms of different residues of your protein that vary over the course of the simulation. Provide a plot for the same.

We calculate the average angle between atoms **1, 21 and 15.**

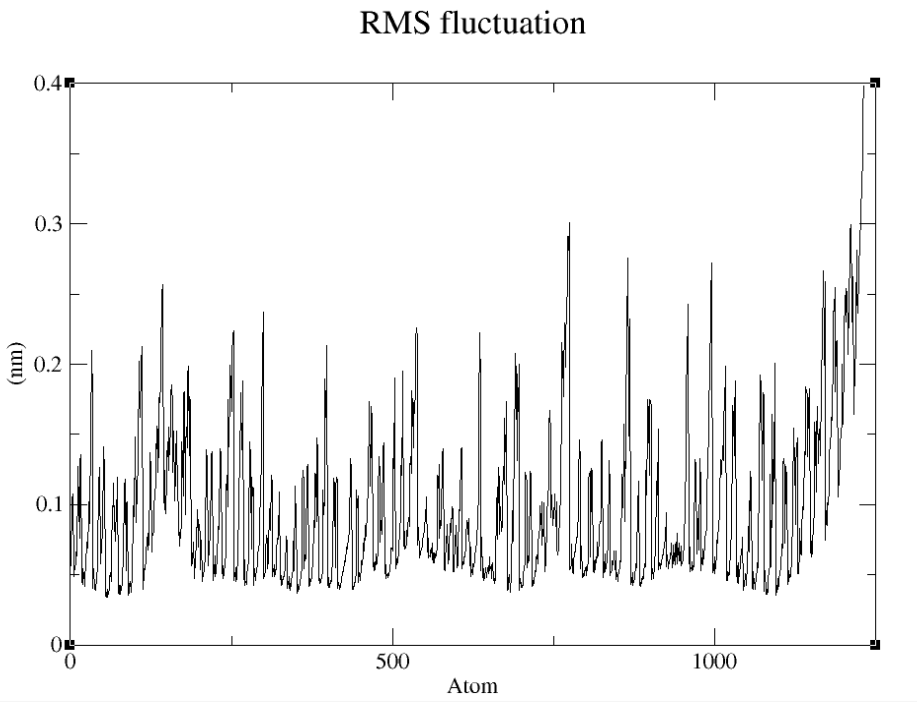


Q4. a) Expand RMSF. Why is it useful to measure RMSF?

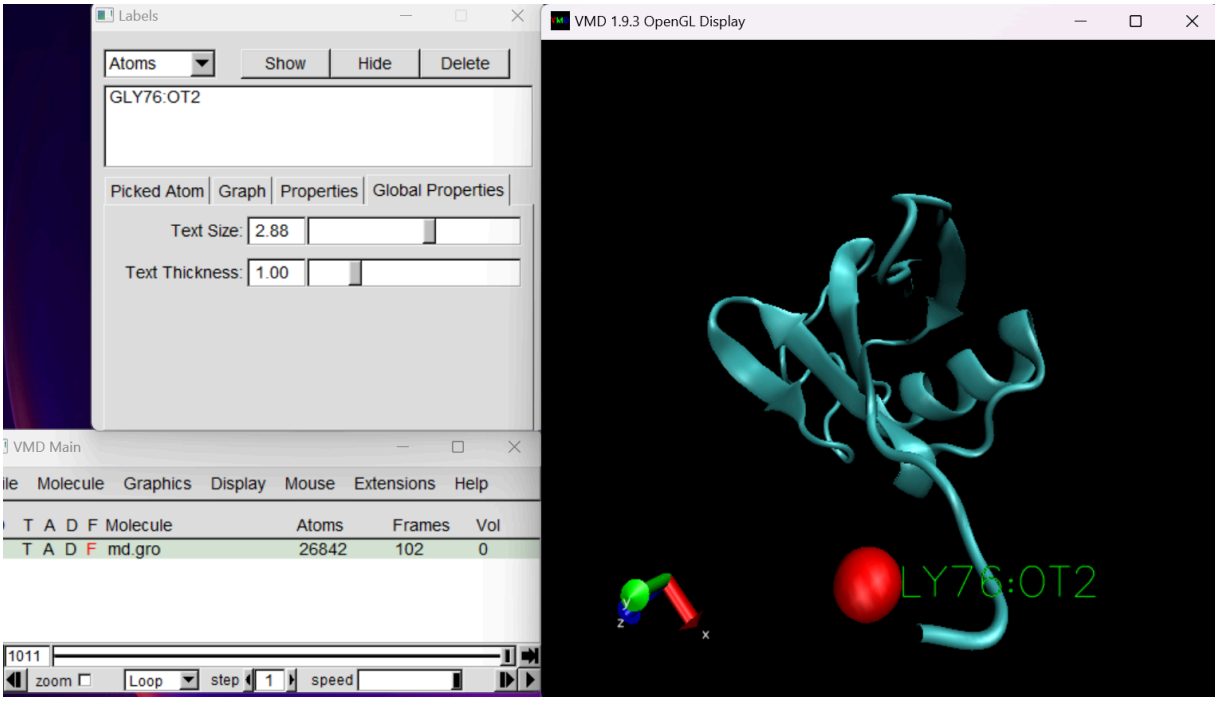
RMSF = Root Mean Squared Fluctuations. This analyzes the extent of fluctuation of individual atoms over the course of the simulation. We can identify regions of the protein structure that fluctuate the most and least during the entire simulation.

b) Calculate the RMSF of the protein after the production run and provide the plot for the same. Which atom fluctuates the most in your simulation? Name the residue this atom belongs to.

The RMSF fluctuations per atom is plotted using xmgrace and provided below



The last atom (Serial 1231) is the one that fluctuates the most. We analyze this atom in VMD.



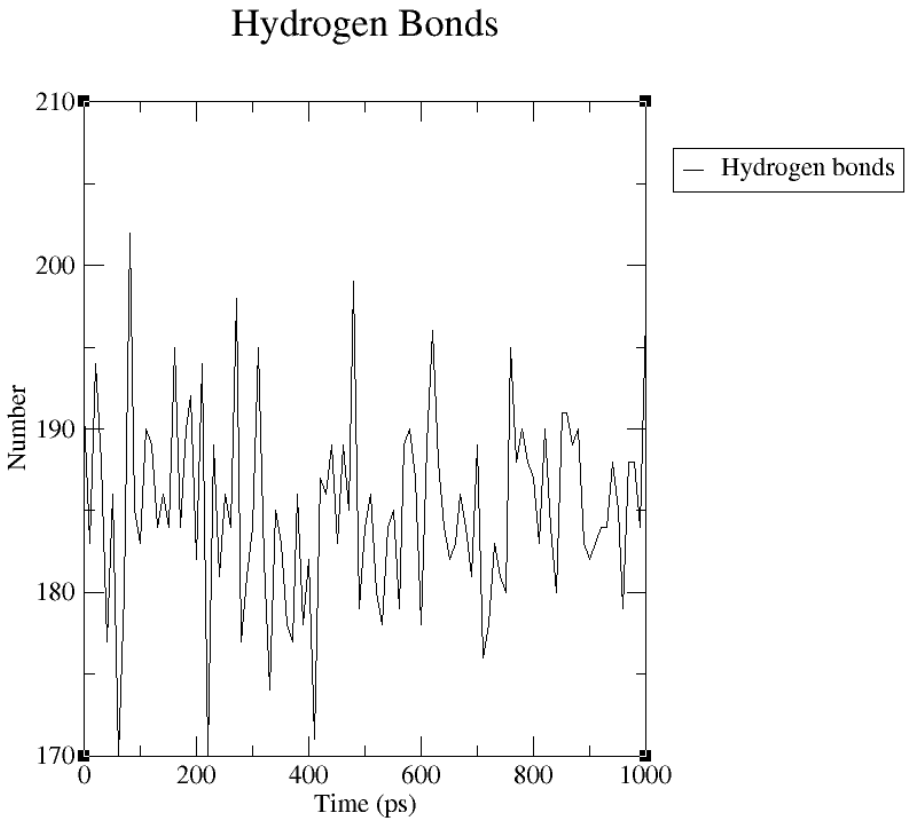
The residue this atom belongs to is GLY76

Q5. Report the average number of hydrogen bonds per time frame and provide plots for the number of hydrogen bonds versus time for the following groups of molecules:

a) Protein-water

```
Select a group: 1
Selected 1: 'Protein'
Select a group: 12
Selected 12: 'Water'
Checking for overlap in atoms between Protein and Water
Calculating hydrogen bonds between Protein (1231 atoms) and Water (25611 atoms)
Found 8649 donors and 8760 acceptors
Reading frame       0 time    0.000
Will do grid-search on 15x15x15 grid, rcut=0.34999999
Frame loop parallelized with OpenMP using 16 threads.
Last frame        100 time 1000.000
Average number of hbonds per timeframe 185.129 out of 3.78826e+07 possible

GROMACS reminds you: "Life would be tragic if it weren't funny." (Stephen Hawking)
```

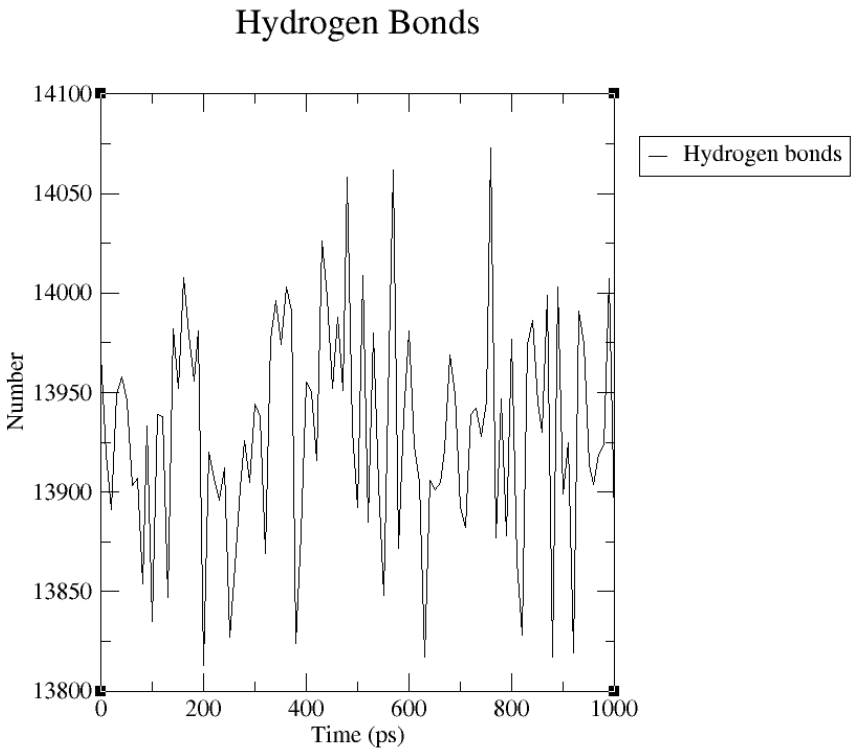


The average number of hbonds per timeframe is 185.129 for protein-water interactions

b) Water-water

```
Select a group: 12
Selected 12: 'Water'
Select a group: 12
Selected 12: 'Water'
Calculating hydrogen bonds in Water (25611 atoms)
Found 8537 donors and 8537 acceptors
Reading frame       0 time    0.000
Will do grid-search on 15x15x15 grid, rcut=0.34999999
Frame loop parallelized with OpenMP using 16 threads.
Last frame        100 time 1000.000
Average number of hbonds per timeframe 13930.327 out of 3.64402e+07 possible

GROMACS reminds you: "I think everybody should like everybody." (Andy Warhol)
```



The average number of hbond per timeframe interactions for protein-water interactions is 13930.327