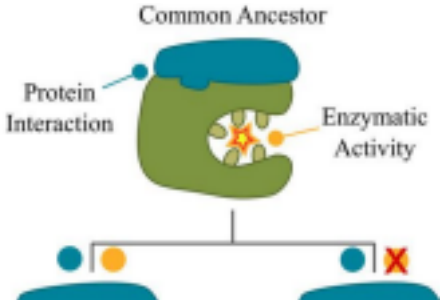
# Molecular Dynamics Simulation of Viral Lysozymes (5KI1 & 206L) Using Gromacs

*Flavio Scipione*
*Andrea Belli Contarini*
*Federico Loreto*
*Krakowski Patryk*
*Vasileios Chatzitolios  Matteo Cacioppo*

Computational Biophysics 2023-2024

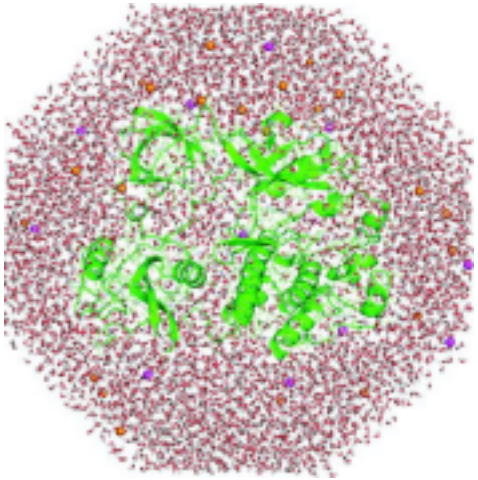
## Introduction & Objectives

- Pseudoenzymes
  - Catalytically-deficient variants
  - Same folding structure



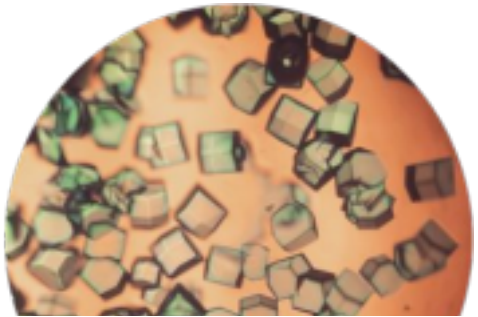
- Molecular Dynamics
    - How are dynamics affected by point mutations? ○ What happens near mutated residues?

# Introduction & Objectives



- Pseudoenzymes
    - Catalytically-deficient variants
    - Same folding structure
- Molecular Dynamics
    - How are dynamics affected by point mutations? ○ What happens near mutated residues?

# Lysozyme

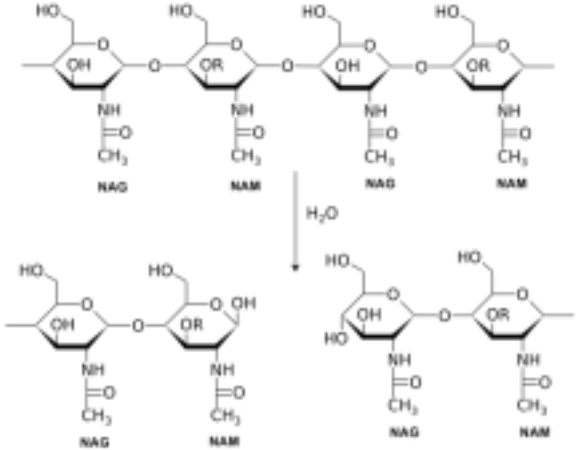


- Antimicrobial enzyme found in animal secretions ○ Tears, saliva, mucus and milk
- Acts by degrading peptidoglycans in the bacterial

walls ● Main active sites in Glu35 and Asp52

● Non-enzymatic function

# Lysozyme




● Antimicrobial enzyme found in animal secretions

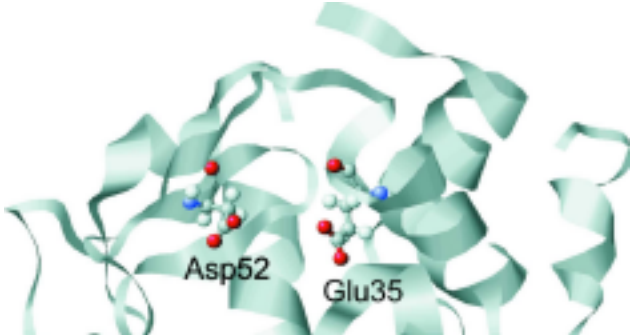
● Acts by degrading peptidoglycan in the bacterial walls ○ Binds peptidoglycan in the cleft between its two domains ○ Hydrolysis of the glycosidic bonds

● Main active sites in Glu35 and Asp52

● Non-enzymatic function

# Lysozyme

● Antimicrobial enzyme found in animal secretions ● Acts by degrading peptidoglycan in the bacterial walls ● Main active sites in Glu35

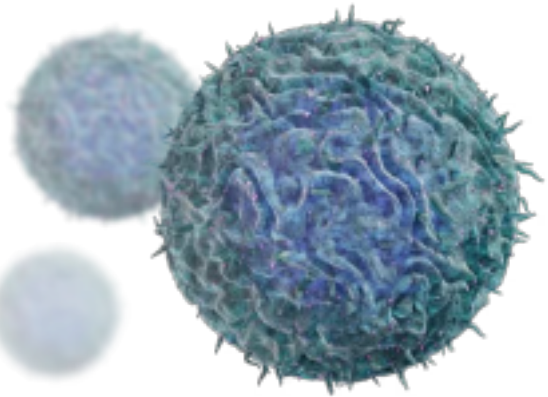


and Asp52

- Substitution leads to malfunction

- Non-enzymatic function

# Lysozyme

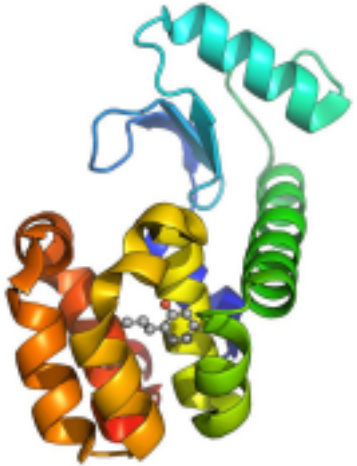


- Antimicrobial enzyme found in animal secretions
- Acts by degrading peptidoglycan in the bacterial walls
- Main active sites in Glu35 and Asp52
- Non-enzymatic function
  - Asp52 → Ser52 keeps some antimicrobial properties
  - Possible interactions with T-cell receptors

# T4 Lysozyme

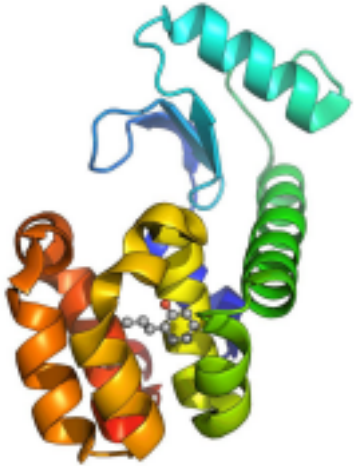
- Lysozyme variant expressed in Tequatrovirus
  - Used to infect bacteria
  - Main active sites in Glu11 and Asp20
- 206L Phage T4 Variant
- 5KI1 Pseudo T4 Mutant

# T4 Lysozyme

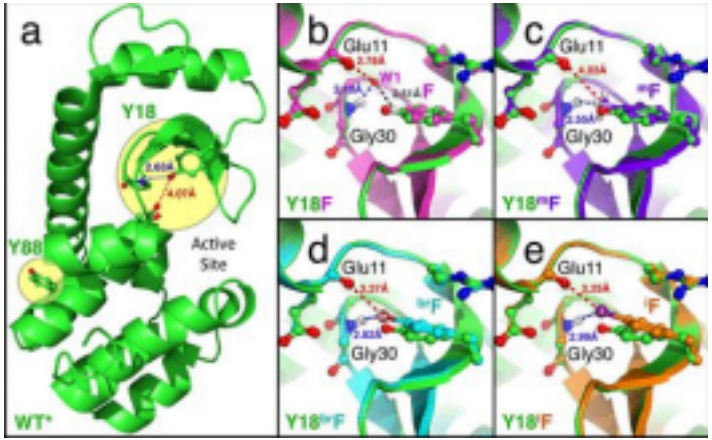



- Lysozyme variant expressed in Tequatrovirus
- 206L Phage T4 Variant
  - Expressed in E. Coli
  - Displays standard function
- 5KI1 Pseudo T4 Mutant

# T4 Lysozyme

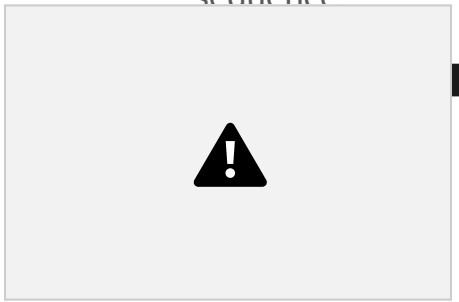
- Lysozyme variant expressed in Tequatrovirus
- 206L Phage T4 Variant
- 5KI1 Pseudo T4 Mutant
  - Ser42 → Ala42
  - Tyr18 → Phe18
  - The lost connection is bridged by water
  - Structure remains intact




# The viral lysozymes studied


- **206L & 5KI1:**

  Both are from the same organism (bacteriophage Tequatrovirus T4), expressed in Escherichia coli. However, they differ in specific name, due to specific mutations made to amino acids in the sequence



## n tools: Gromacs & Computers' processors

GROMACS is one of the fastest and most popular software packages available designed for simulations of proteins, lipids and nucleic acids.

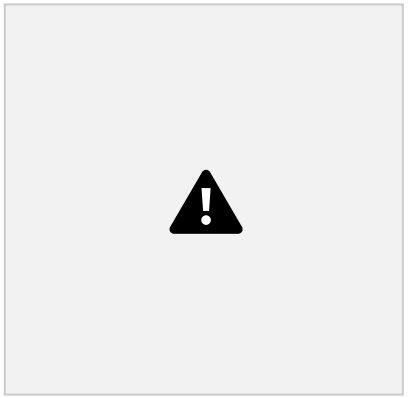


**Computing power of computers :**

# MD Simulation Process

1. Topology Preparation & Examination 2. Defining the Unit Cell & Adding Solvent 3. Ions Addition
4. Energy Minimization
5. Equilibration
6. Production MD
7. Data Analysis

# 1. We clean our proteins



Topology Preparation & Examination



[redacted]

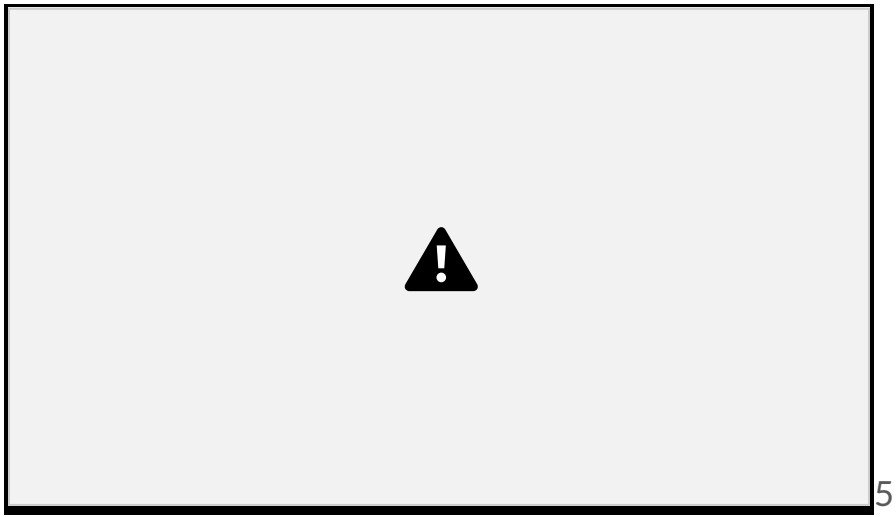
[redacted] - Best performance for liquid simulations - LMP2 (Local Møller–Plesset perturbation)

- Better RMSD (Root Mean Squared Deviation)

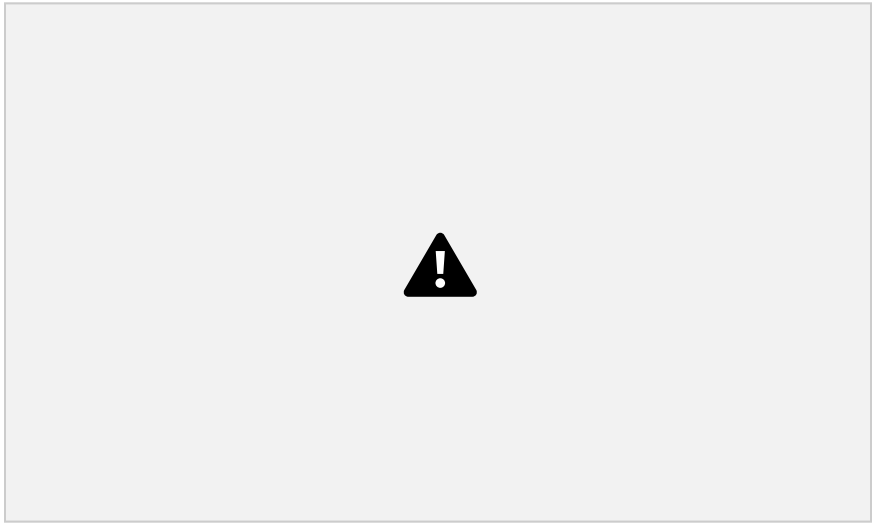
## Defining the Unit Cell & Adding Solvent

- [redacted]
- [redacted]
- [redacted]

[redacted]



5KI1



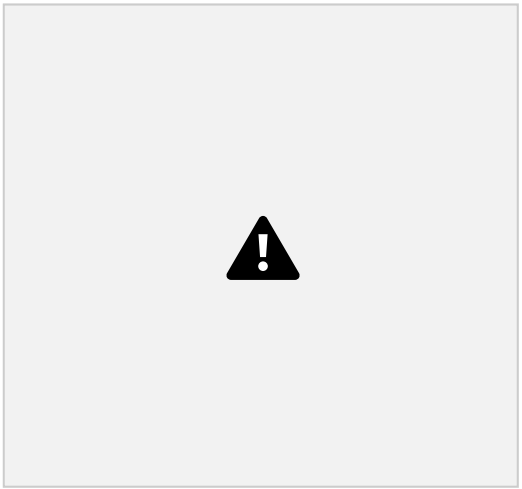
206L

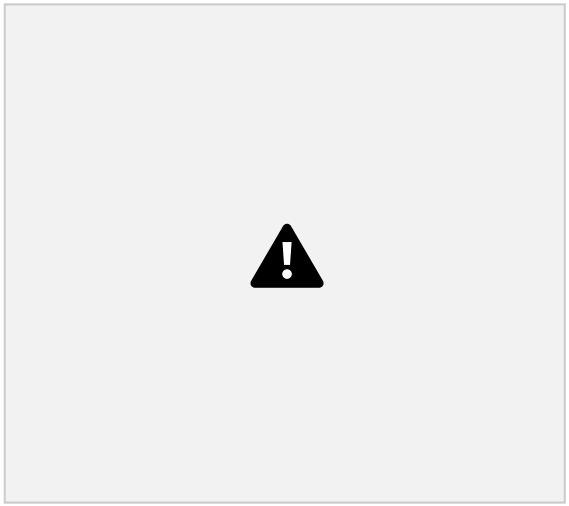
## Energy Minimization

http://www.mdtutorials.com/gmx/lysozyme/Files/minim.mdp









## minimization

**Parametri:**
**emtol = 1000.0** (F_max=1000 KJ/Mol/nm) **emstep = 0.01** (step size)
**nsteps = 50000**

factors to evaluate to determine if EM was successful:

- $E_{pot}$ should be negative on the order of $10^5$-$10^6$
- the minimum is reached before the maximum number of steps

**Equilibration reached through a steepest descent**

## Equilibration

**NVT ensemble**

**Equilibration reached through a modified Berendsen thermostat:**

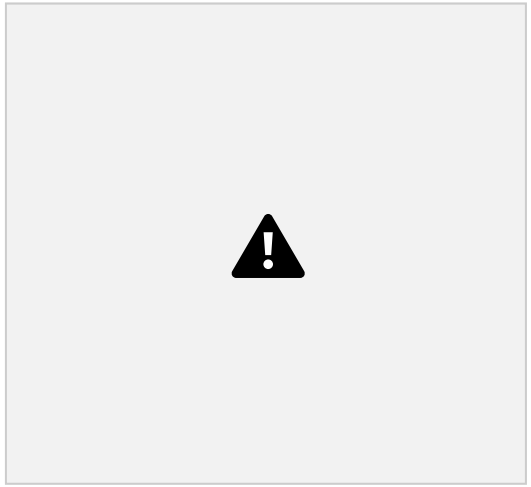
$\tau\tau$= 0.1 *ps*,
$T_{bath}$=*300 K*

$T_{206L}$=300.001±1.508 K
$T_{5KI1}$=299.992±1.480 K

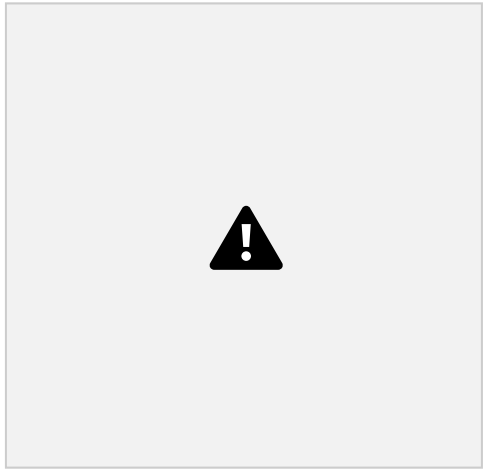
in both cases, temperature of the system quickly reaches the target value (300 K), and remains stable for 100 ps.

## Equilibration

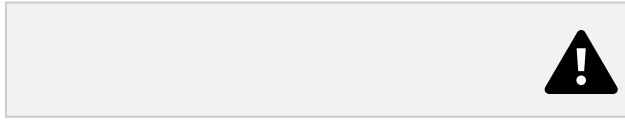
NPT ensemble



$\rho\rho$=1015.93 ±2.54 kg/m$^3$



$\rho\rho$=1016.34 ±2.54 kg/m$^3$



Equilibration reached through a **Parrinello-Rahman barostat**

$P_{5KI1}$=1.011±127.961 bar

Pressure widely fluctuate, but statistically in can't be distinguished by the target value (see http://www.mdtutorials.com/gmx/lysozyme





$\tau_P$= 2 *ps*, **k**=44.5*10-5 $bar^{-1}$,$P_{bath}$=1 bar

$P_{206L}$=0.834±124.955 bar

The densities are compatible with the predicted densities of the system /07_equil2.html)

## Production MD

**We are now ready to release the position restraints and run production MD for data collection**



**Using**



**NPT ensemble:**

- Berendsen thermostat: $\tau_T$= 0.1ps
- Parrinello Rahman barostat: $\tau_P$= 2 ps,
  **K**=44.5*10-5 $bar^{-1}$

- Time step: 2 fs (good compromise for computability)
- $3*10^7$ total steps

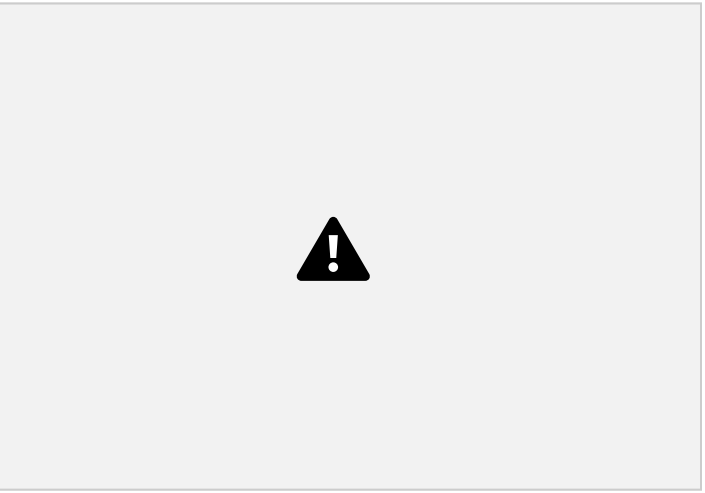
http://www.mdtutorials.com/gmx/lysozyme/Files/md.m

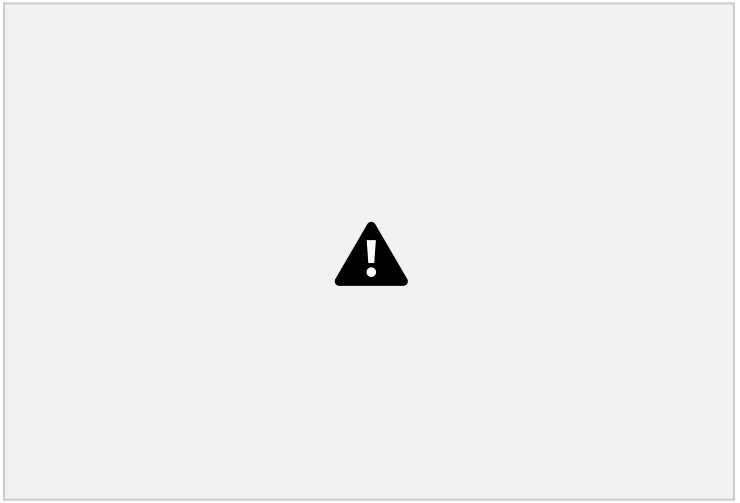
dp

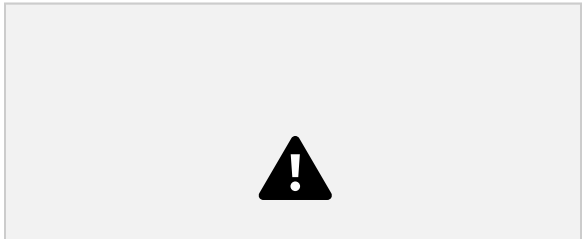
*x(t)* → position at step *t*
*v(t+1/2)* → *velocity at step (t+1/2)* *a(t)* → acceleration
*dt* → time-step

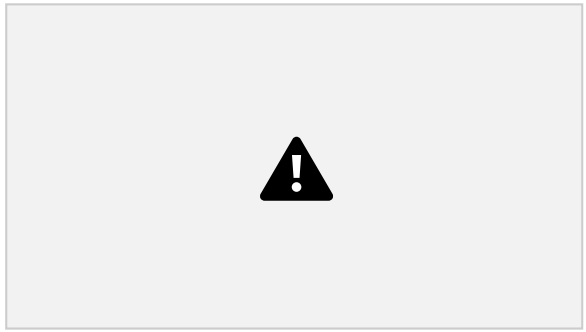
## Production MD

## MD 206L MD 5KI1









Energy: 206L Energy: 5KI1

energy fluctuation: ~ 0.2% energy fluctuation: ~ 0.2%

# Comparison between the two lysozymes: alignment



Changes:

- number *18*:

  tyrosine phenylalanine

*makes hydrogen-bonding with 30 when not mutated, otherwise $H_2O$ bridges.*

- number *42*:

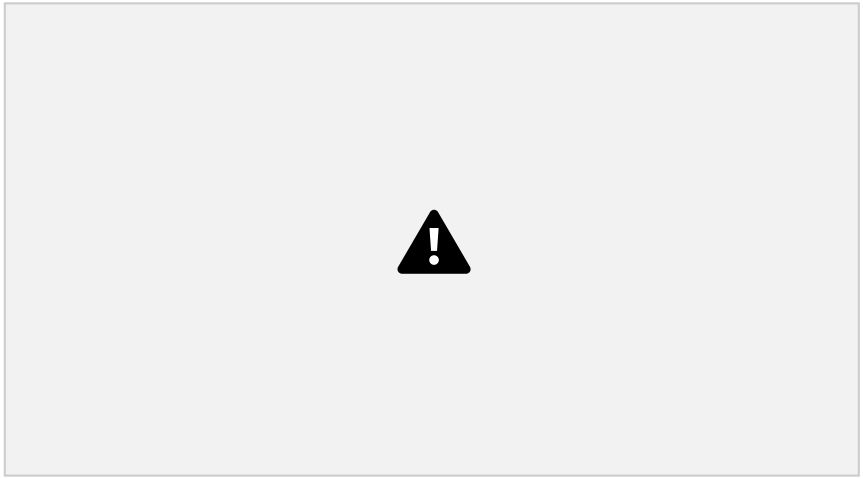
  serine alanine

*makes hydrogen-bonding with: 39, 17 & 45.*
*Check everything on link.*

Sequences alignment, using *BLAST*:
https://www.ncbi.nlm.nih.gov/

## Data Analysis

# Comparison between the two lysozymes: RMSDs



Possible causes in fluctuations difference:

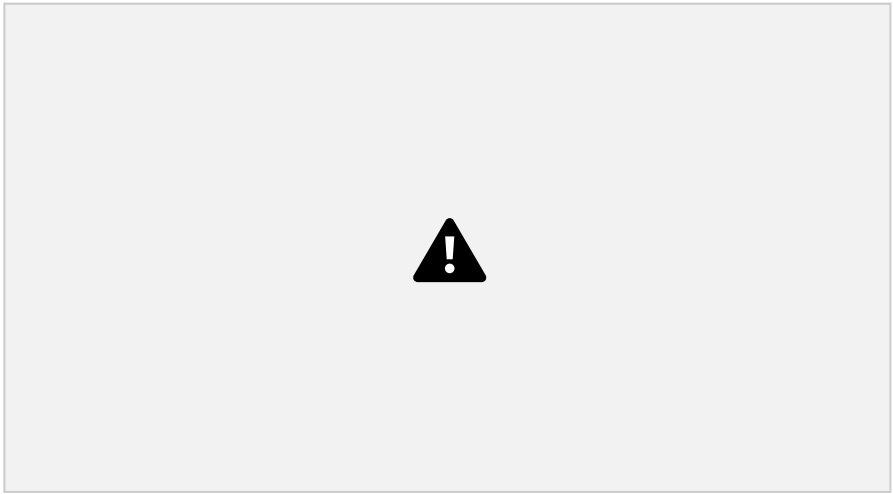
● Intrinsic conformation: 5KI1's greater intrinsic flexibility. ● Molecular interactions: more dynamic or unstable interactions with its surroundings.

● More frequent binding events or conformational changes.

## Data Analysis

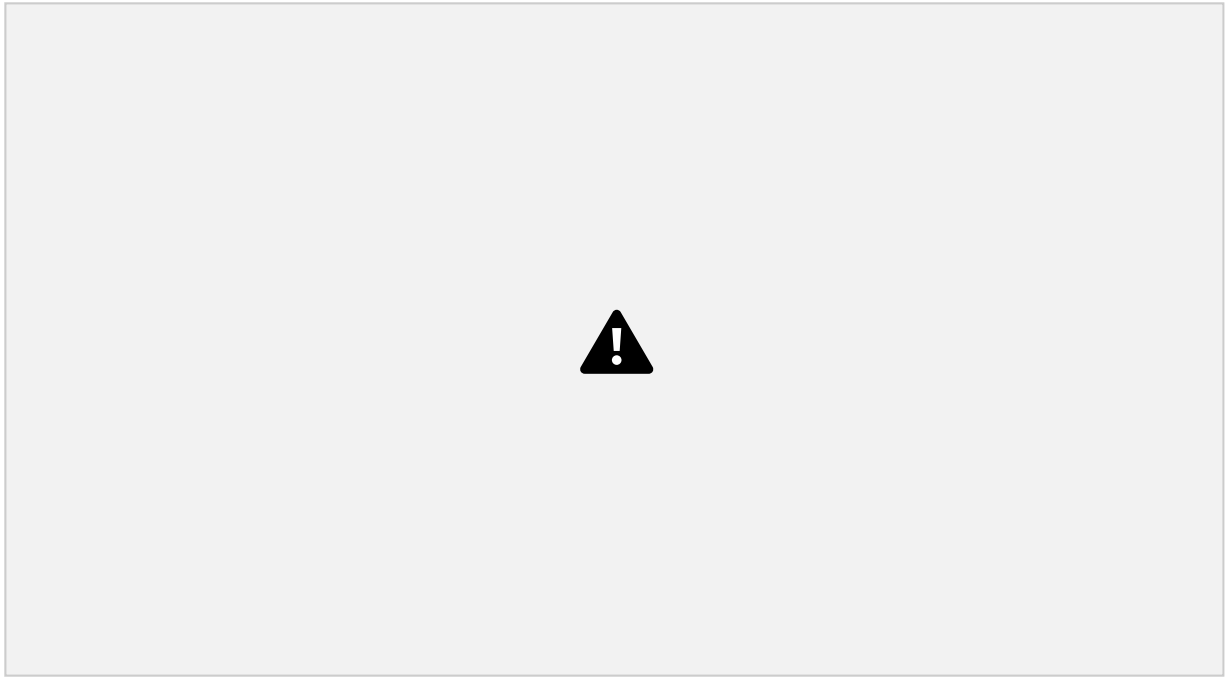
# Comparison between the two lysozymes: RMSDs

Specific plots of the RMSD for the protein and for the backbone



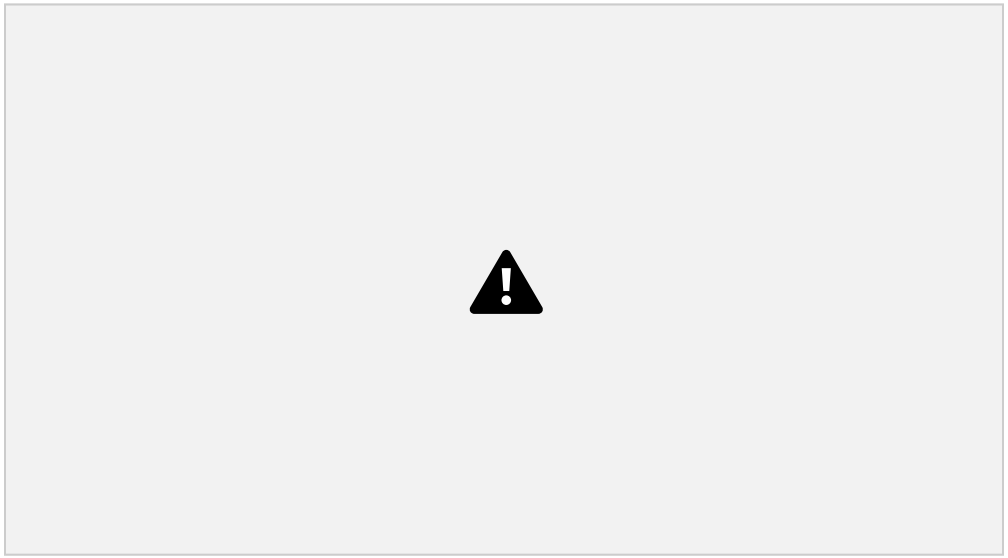
Data Analysis

# Comparison between the two lysozymes: RMSDs



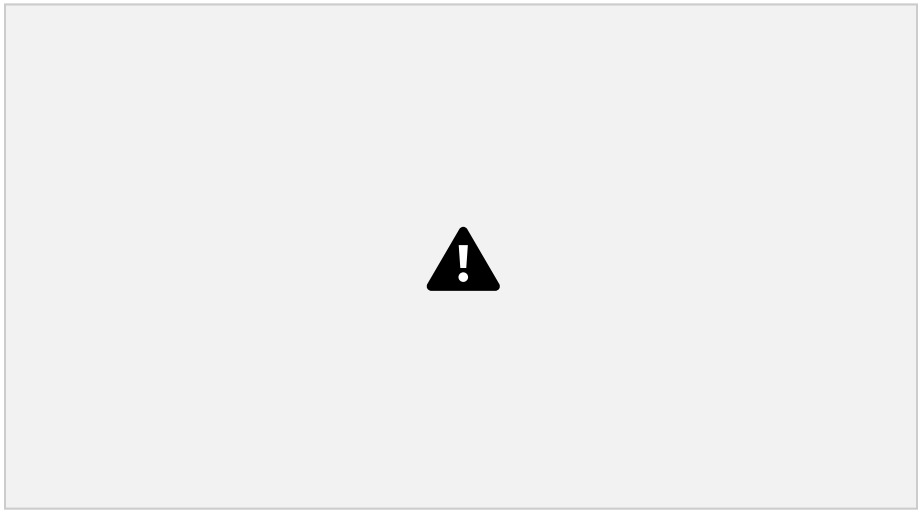
## Data Analysis

# Comparison between the two lysozymes: RMSFs



# Comparison between the two lysozymes: RMSFs

Radius of gyration



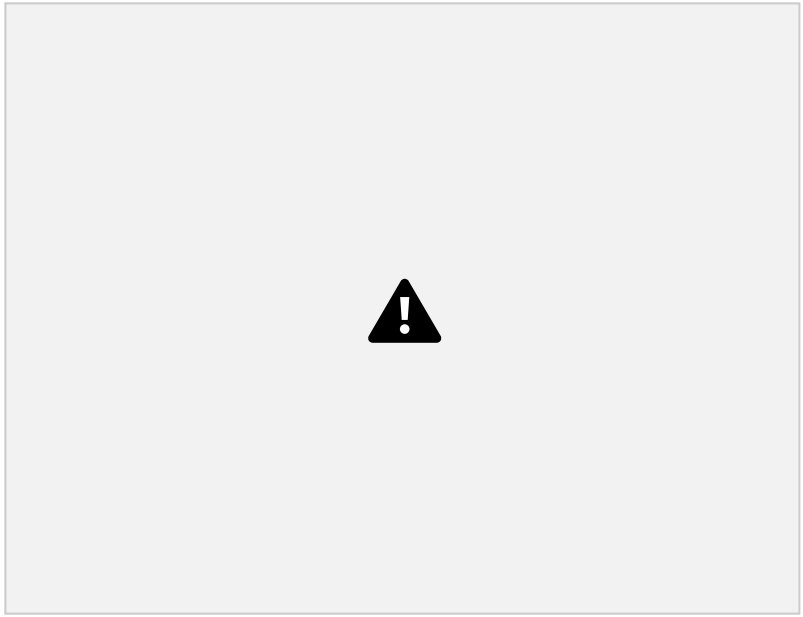
comparison:

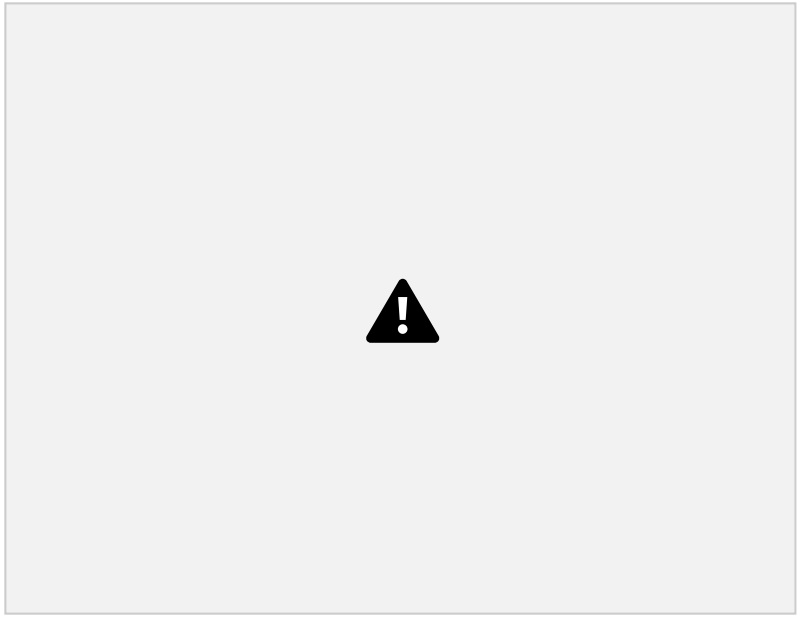
Possible causes in fluctuations difference

- **Molecular structure**: Differences in the 3D structures could influence the dynamics.
- **Binding characteristics**: for example a.a #18 cannot make hydrogen bonds with #11 and #30.
- **Residue-residue interactions**: interactions between residues (i.e. H-B, ionic bonds, hydrophobic bonds) may influence structural stability, affecting the RMSF.

## Data Analysis

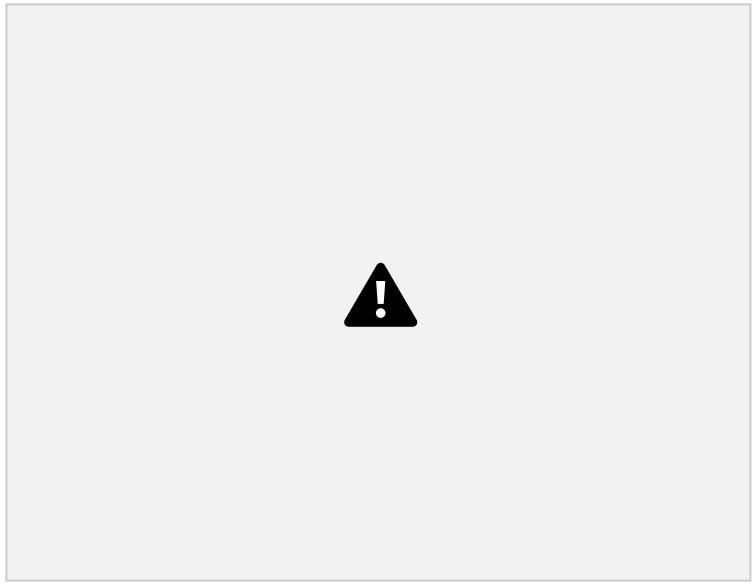
# Comparison between the two lysozymes: R18 & R30

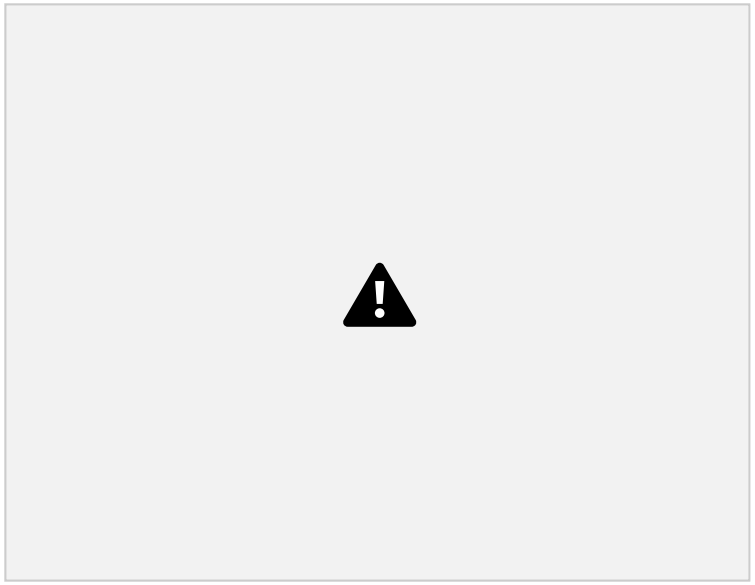




Data Analysis

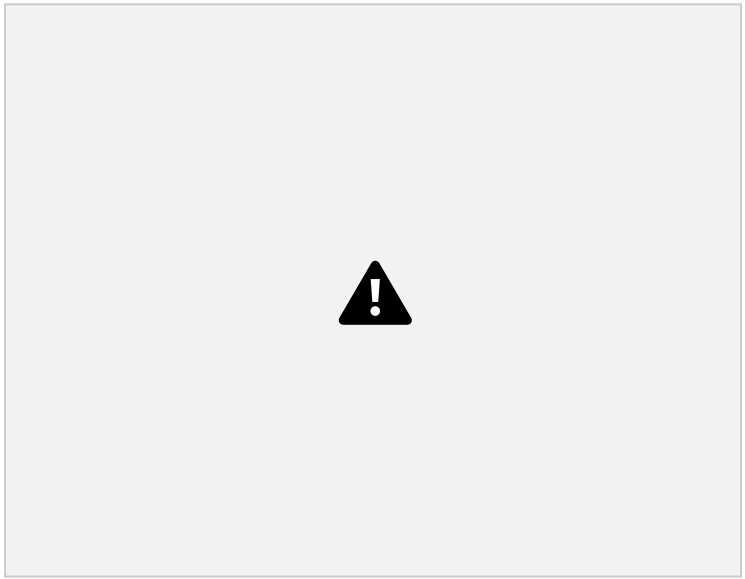
# Comparison between the two lysozymes: R18 & R11

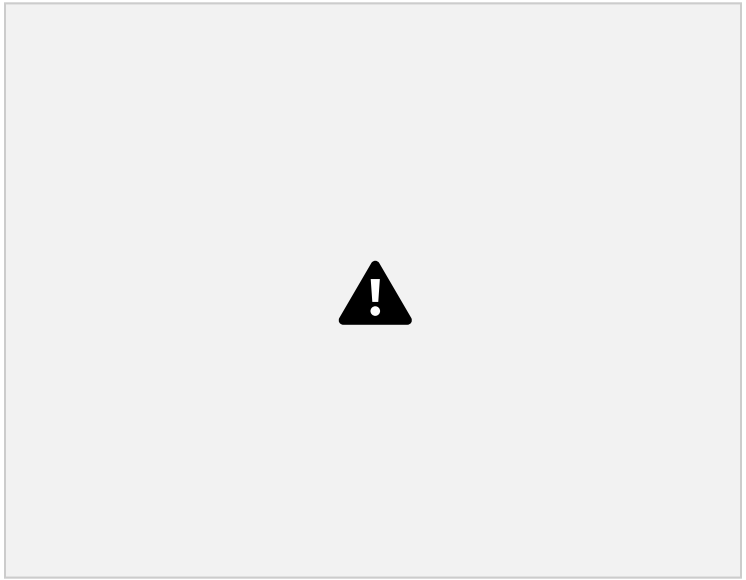




Data Analysis

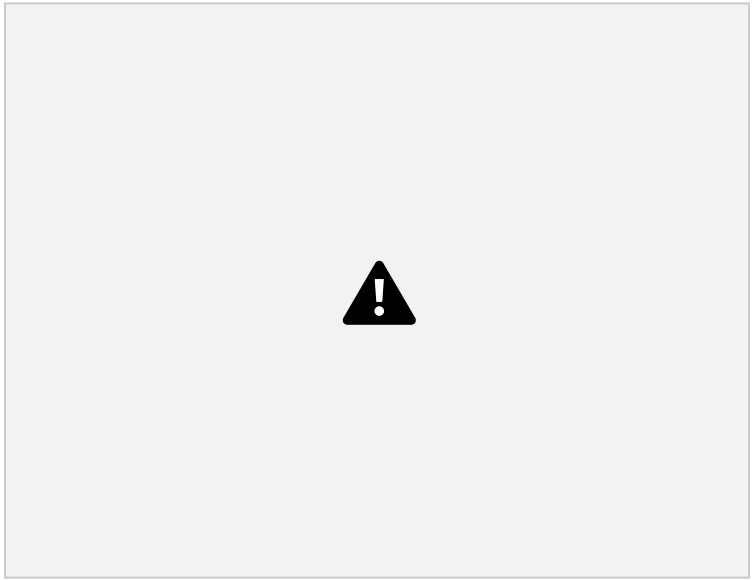
# Comparison between the two lysozymes: R42 & R17

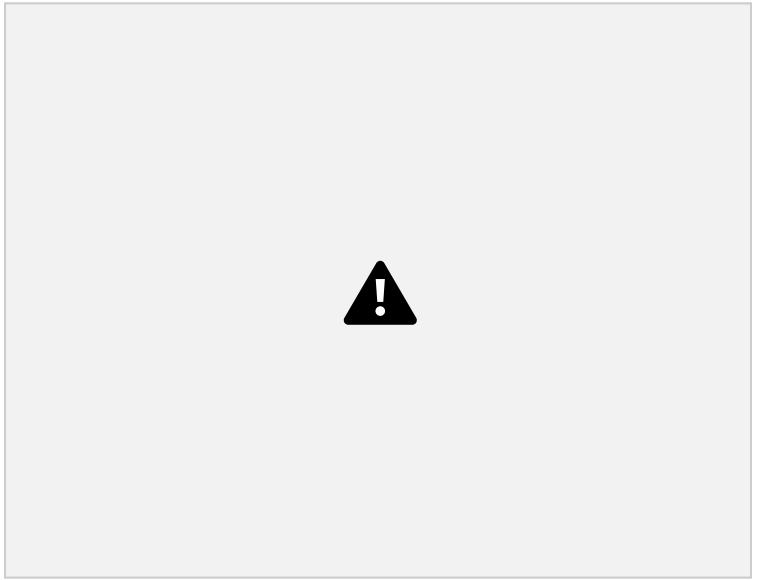




Data Analysis

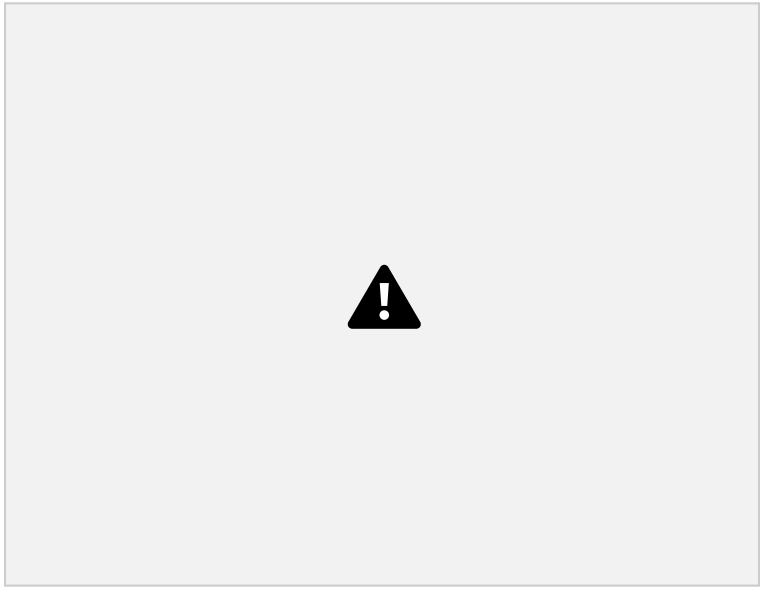
# Comparison between the two lysozymes: R42 & R39

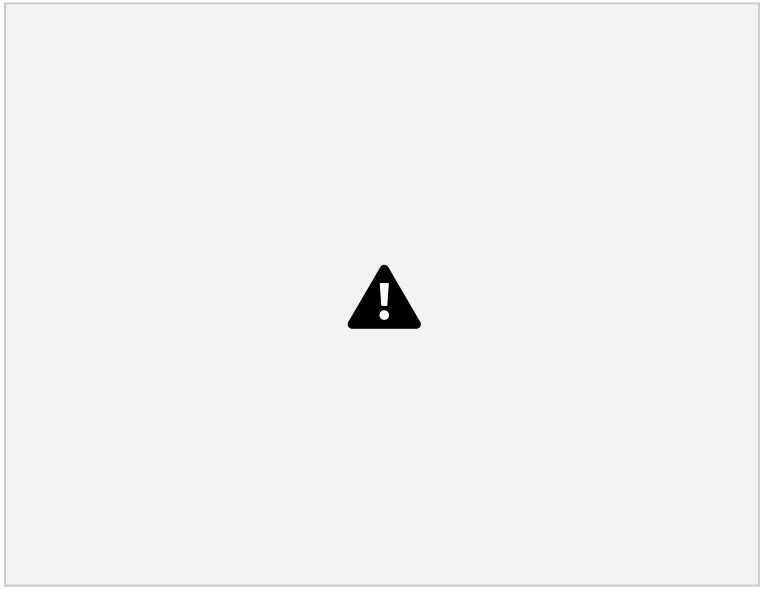




Data Analysis

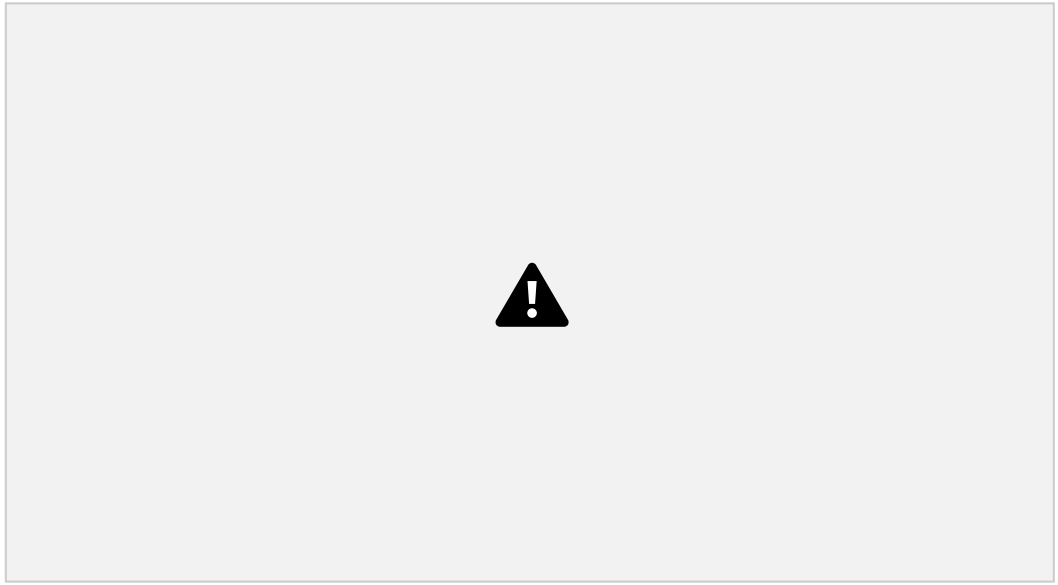
# Comparison between the two lysozymes: R42 & R45





Data Analysis

# Results- Cross Correlation Dynamical Matrices (CCDM)



206L 5KI1



# Conclusions

● The overall structure does not change → in the mutation in site #18 $H_2O$ is able to bridge; however the overall activity is reduced, compared to the mutated version, as anticipated in the article. ● Using a rhombic dodecahedron box is computationally less expensive (≃30% less water). ● Additional softwares to Gromacs (i.e. *Blast*, *Biosig*, …) are very useful for a deeper understanding of the molecules behavior.