

# Masters Practical 2012

offered by

Dr. A. Schafferhans, Dr. E. Kloppmann, Dr. M. Offmann

# **Molecular Dynamics Analysis**

presented by

Fanny Gatzmann and Susann Vorberg



Happy Birthday Julia !

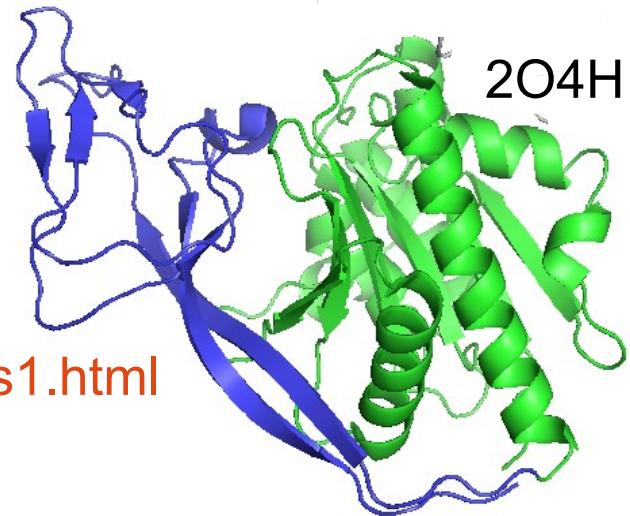


Happy Birthday Julia !



# Overview

- Gromacs Output
- MD Analysis – based on tutorial:  
<http://md.chem.rug.nl/~mdcourse/analysis1.html>
- Using analysis methods provided by Gromacs:
  - g\_energy, gmxcheck, g\_mindist, g\_rmsf, g\_rms,...
- Very useful:  
[http://www.gromacs.org/Documentation/Gromacs\\_Utilities](http://www.gromacs.org/Documentation/Gromacs_Utilities)



# Gromacs Results

_md.mdp	Simulation parameter input file: (time step, type of simulation, electrostatics, van der Waals,...)
_md.tpr	portable binary run input file: (generated by grompp, executed by mdrun) starting structure (coordinates and velocities), molecular topology and all simulation parameters
_md.edr	Binary portable energy file: energy terms that are saved in a simulation
_md.trr	full-precision trajectory file: coordinate, velocity, and force information
_md.xtc	compressed version of the trajectory: only coordinate, time, and box vector information

# standard checks: gmxcheck

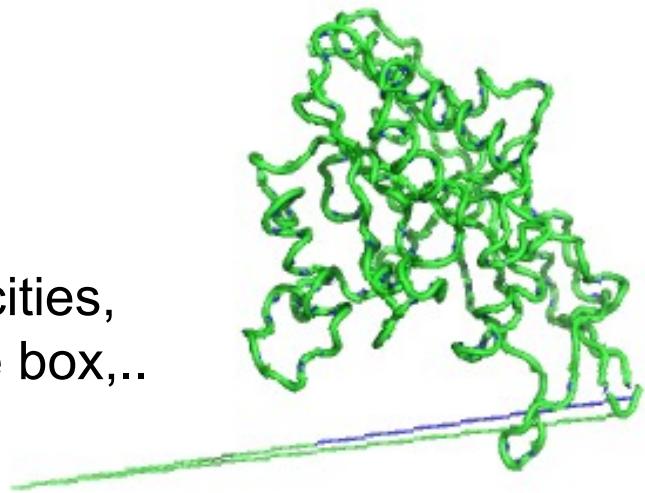
- gmxcheck: *Getting the Right Output Means no Artefacts in Calculating Stuff*

```
gmxcheck -f .xtc
```

→ simulation finished properly?

```
gmxcheck -c .tpr
```

→ presence of coordinates, velocities,  
close contacts, atoms outside the box,..



- Check log file (\_md.log): lot of statistics

	(Mnbfs)	(GFlops)	(ns/day)	(hour/ns)
Performance:	698.741	54.668	44.634	0.538

trjconv – the Swiss army knife tool

```
trjconv -s topol.tpr -f traj.xtc  
-o prot.pdb -pbc nojump -dt 10
```

- Convert trajectories to pdb file
- extract 1000 frames
- remove the jumps over the boundaries

# Visualisation in Pymol



# Quality assurance

- tests for the convergence of thermodynamic parameters  
→ equilibrium reached?
  - Temperature
  - Pressure
  - potential and the kinetic energy ...
- No convergence of values:  
simulation has not yet reached thermal equilibrium  
→ extend simulation
- extracted from \_md.edr file with g\_energy

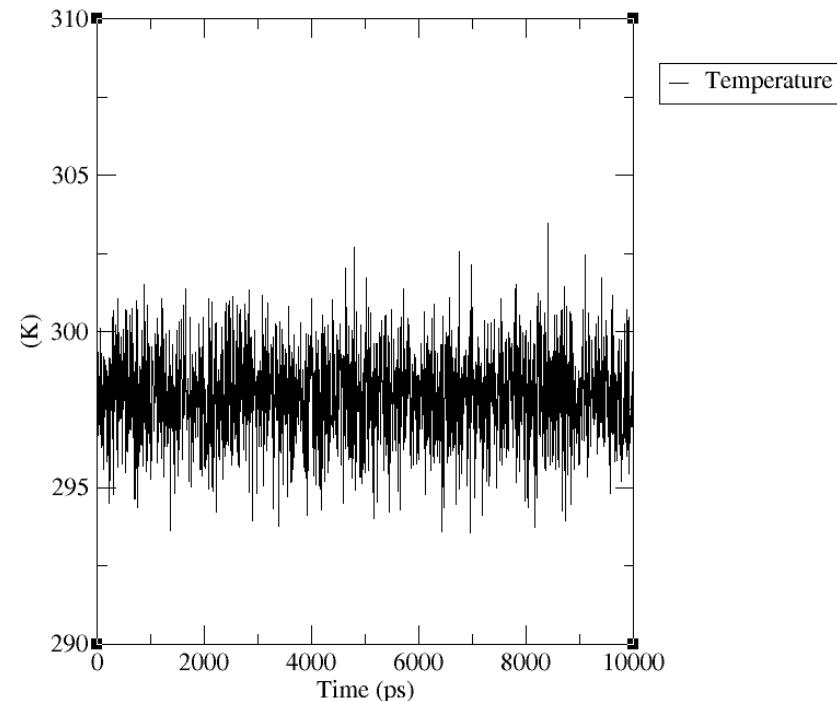
# thermal equilibrium: g\_energy

```
echo 12 0 | g_energy -f _md.edr  
-o temperature.xvg
```

Energy	Average	Err.Est.	RMSD	Tot-Drift
Temperature	297.914	1.38693	0.0072	-0.00042403 (K)

Gromacs Energies

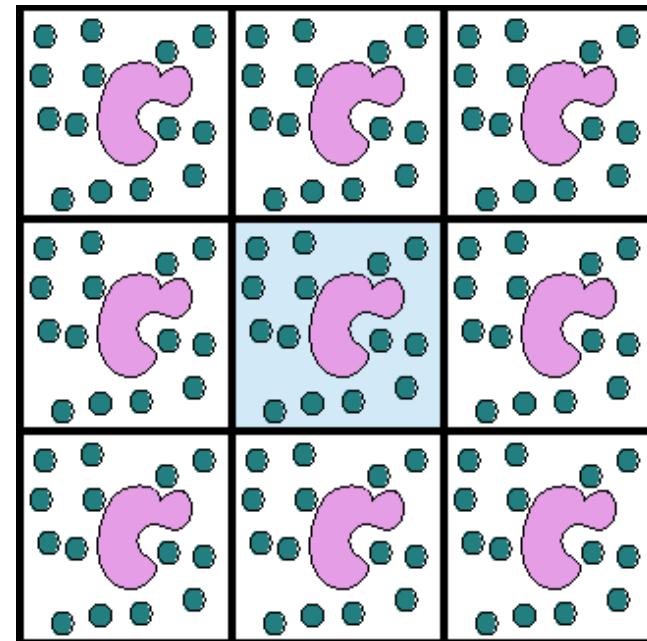
- reference temperature from \_mdmdp: 298K
- Visualize .xvg files
  - xmgrace file.xvg



# Periodic images

- Using periodic boundary conditions:
  - forces between atoms calculated across periodic boundaries
- Check for direct interactions between periodic boxes:
  - ~ unphysical self-interactions
- minimal distance should be  $\geq 2\text{nm}$

```
g_mindist -f traj.xtc  
-s topol.tpr  
-od minperdist.xvg -pi
```



Karplus M, Kuriyan J., PNAS 2005 102 (19) 6679-6685

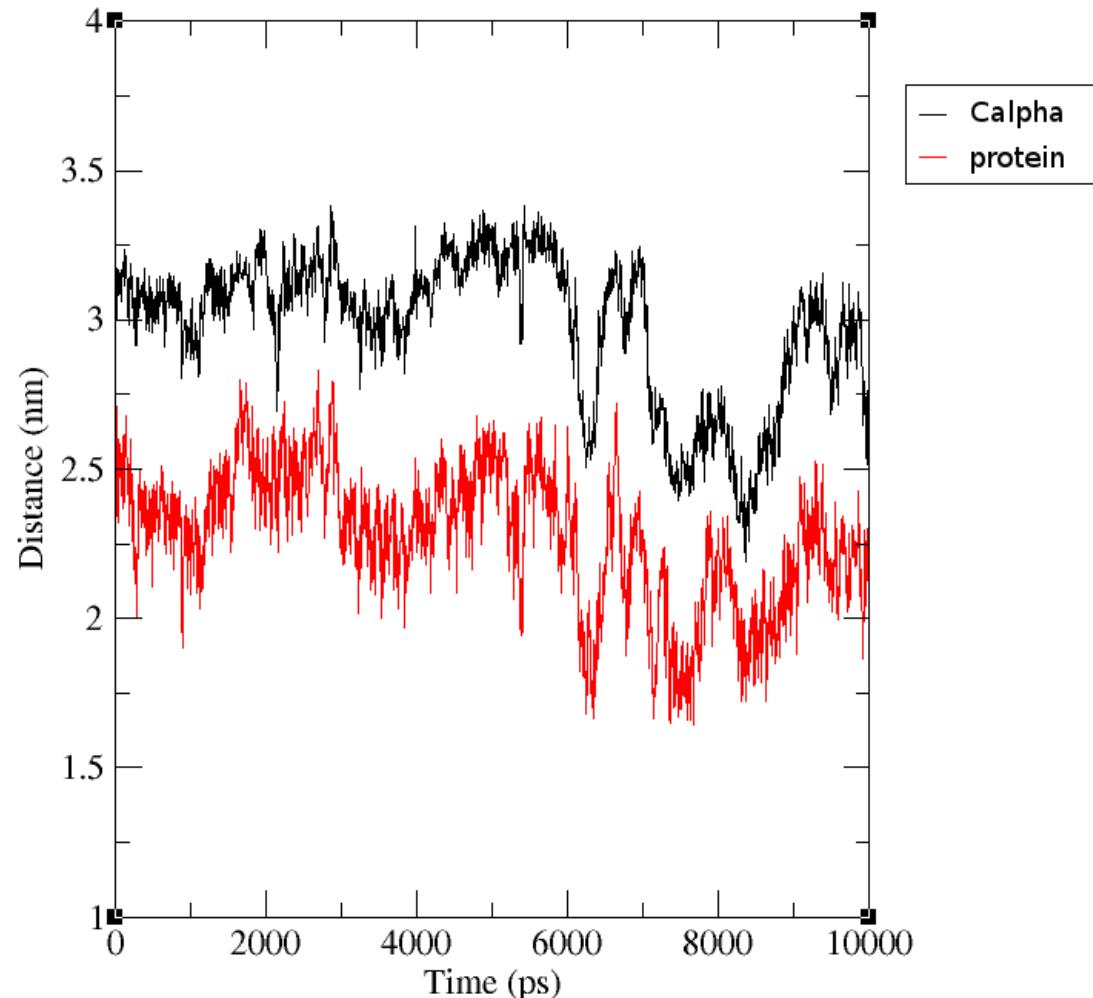
pi: Calculate minimum distance with periodic images

# Periodic images II

The shortest periodic distance is 1.6456 (nm) at time 7675 (ps), between atoms 15 and 4507

Minimum distance to periodic image

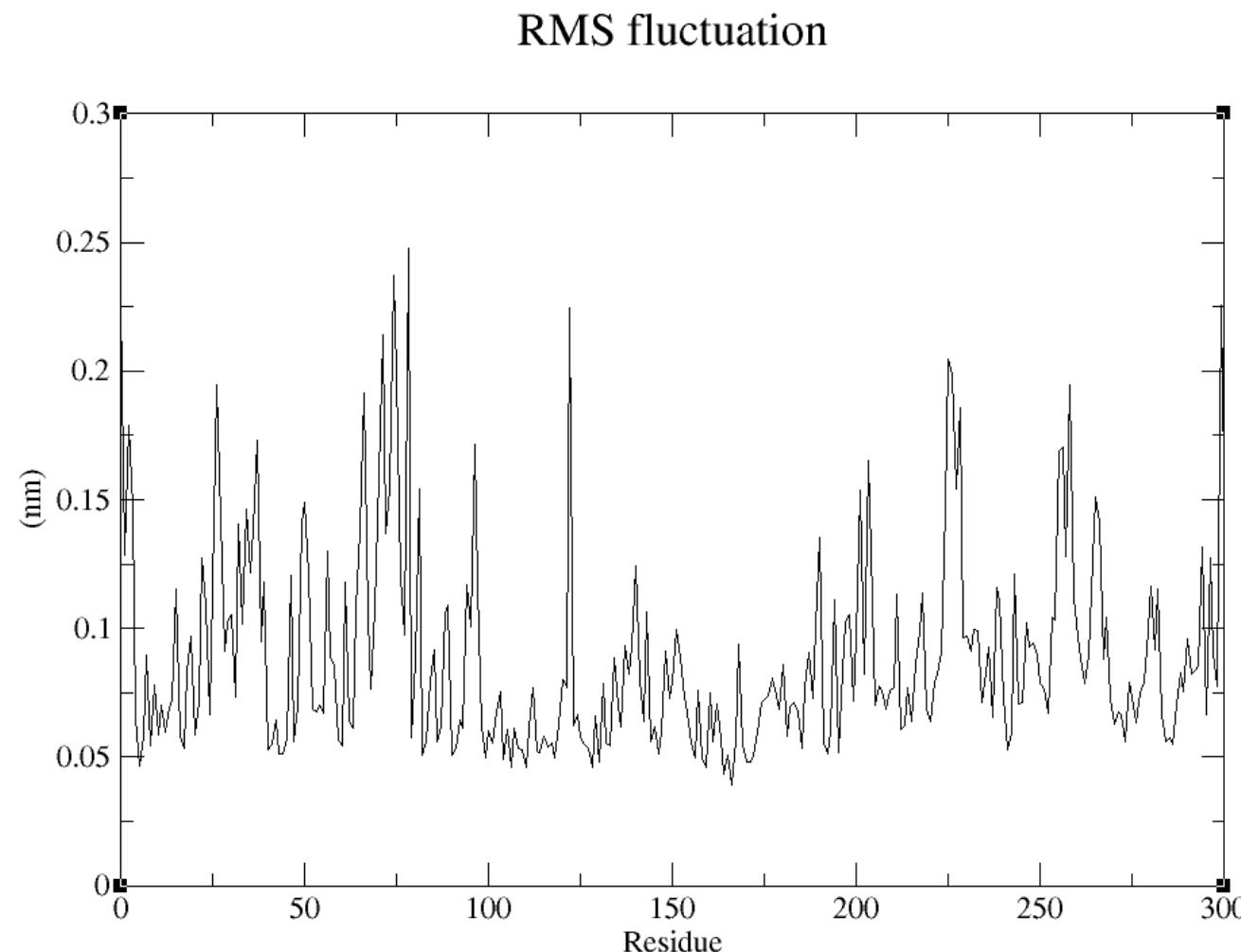
and maximum internal distance



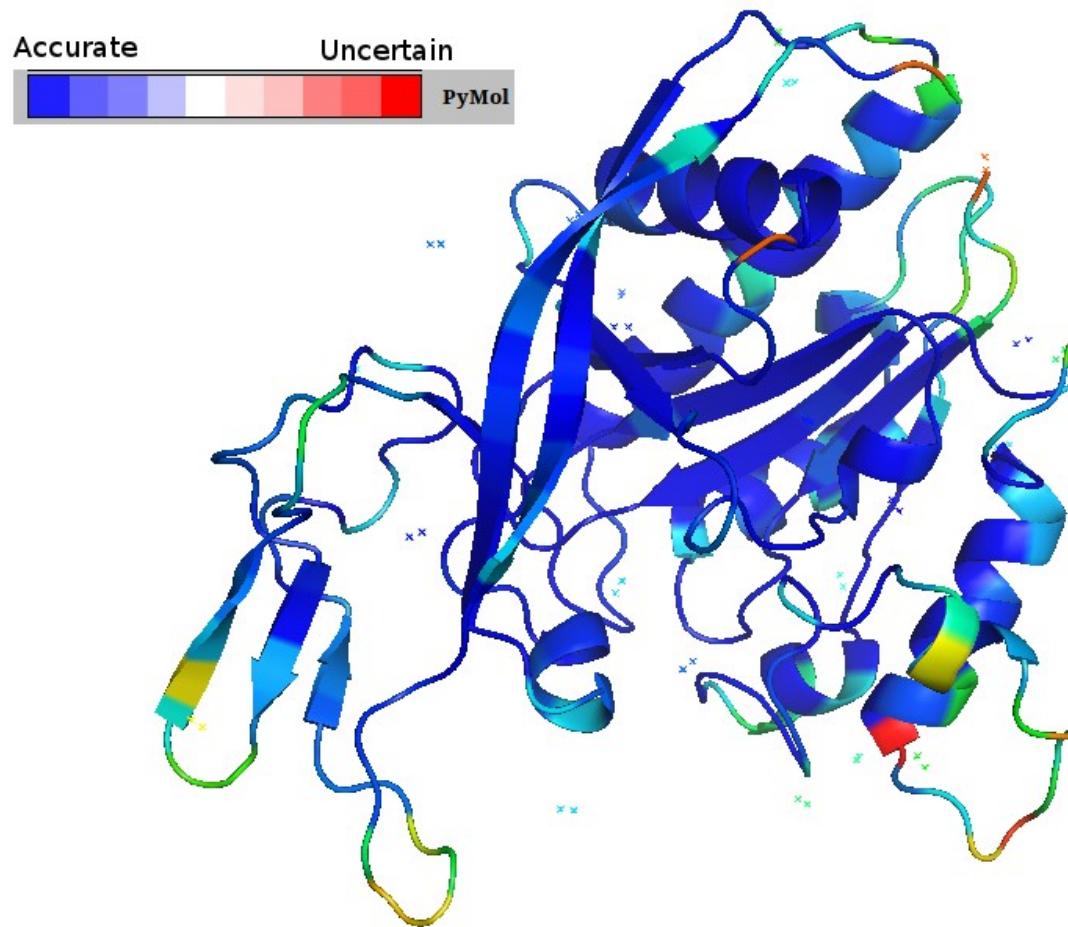
# RMSF: g\_rmsf

- = Root mean square fluctuations
- = fluctuation of an atom about its average position
- ~ crystallographic b-factors → flexibility
- `g_rmsf -f traj.xtc -s topol.tpr  
-o rmsf-per-residue.xvg -ox average.pdb  
-oq bfactors.pdb -res`
- per residue RMSD
- calculate b-factors

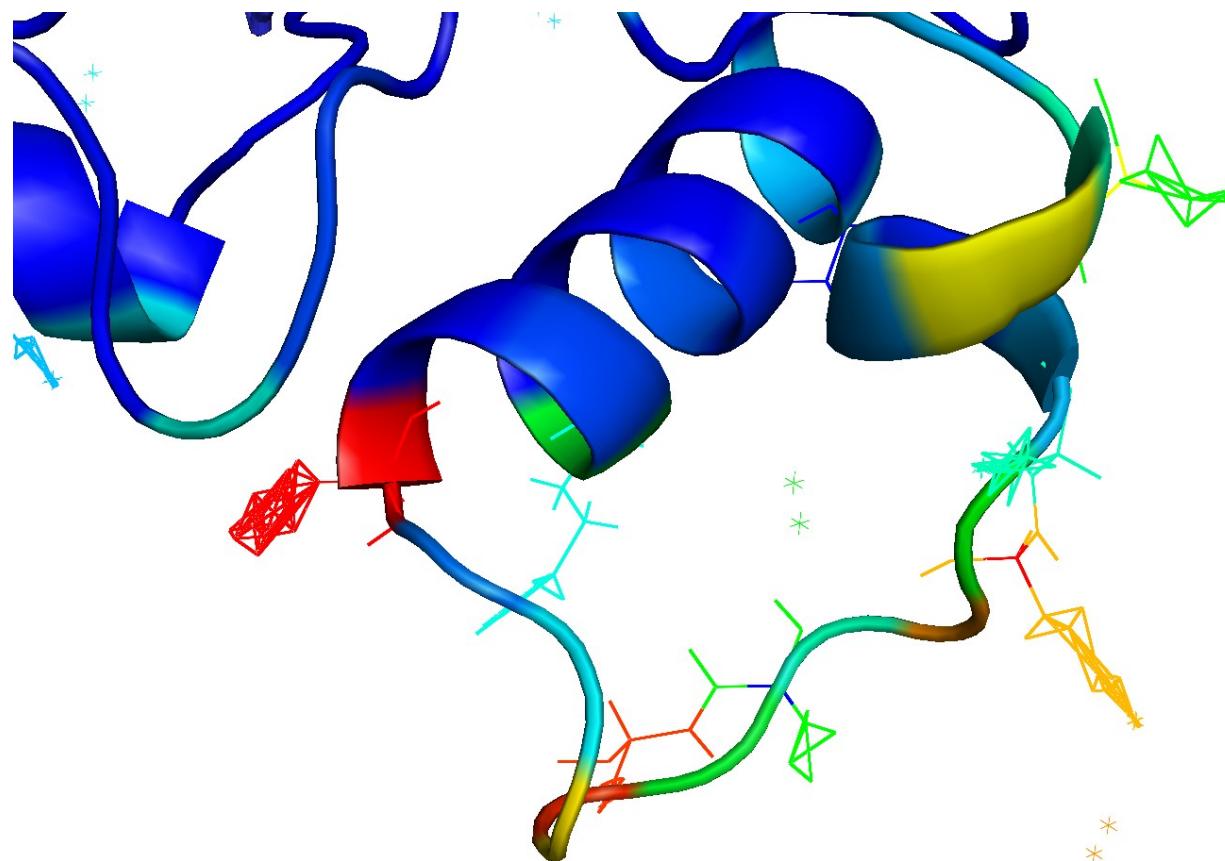
# RMSF – per residue



# RMSF - bfactors



# RMSF - average



Average pdb is an unphysical structure:  
effect of averaging over conformations.

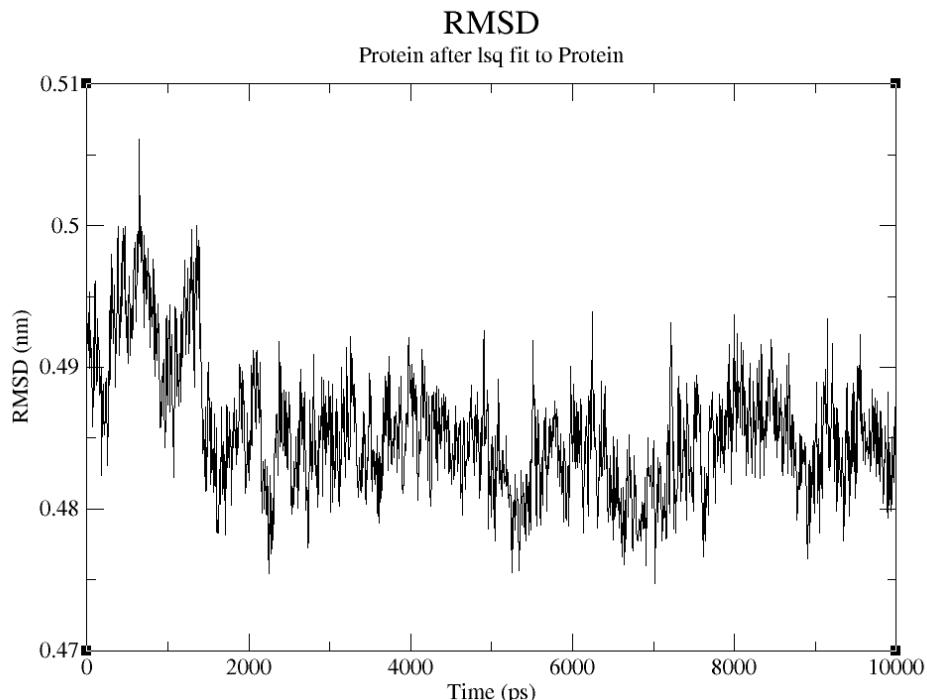
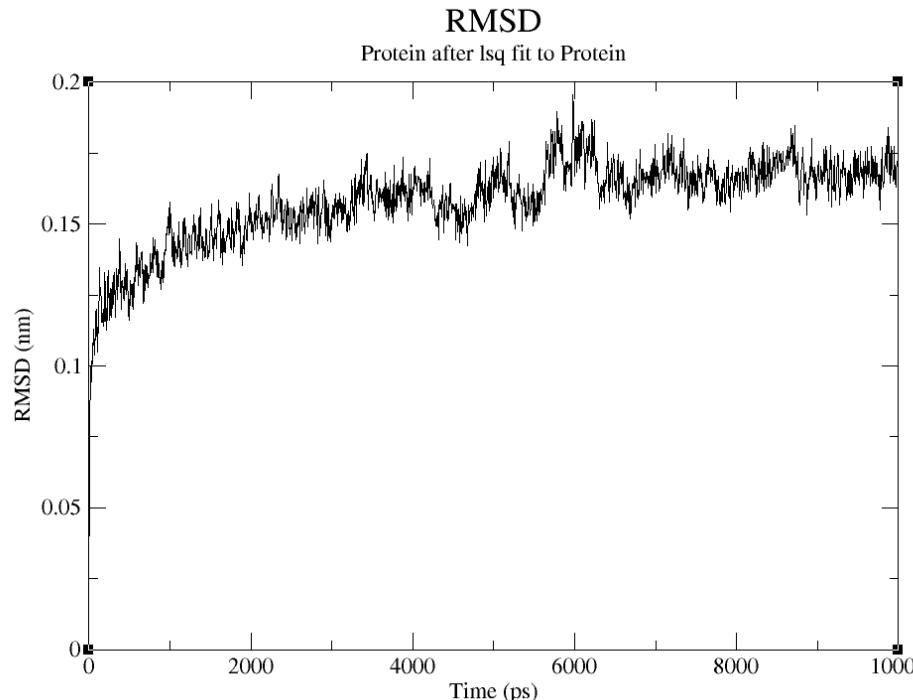
- g\_rmsf also calculates average structure
- Calculate RMSD with respect to the average structure  
→ indicator of convergence of the structure towards an equilibrium state
- RMSD to start structure:  

```
g_rms -f traj_nojump.xtc -s topol.tpr
-o rmsd-backbone-vs-start.xvg
```
- RMSD to average structure:  

```
g_rms -f protein.xtc -s average.pdb -o
rmsd-backbone-vs-average.xvg
```

# Convergence of RMSD

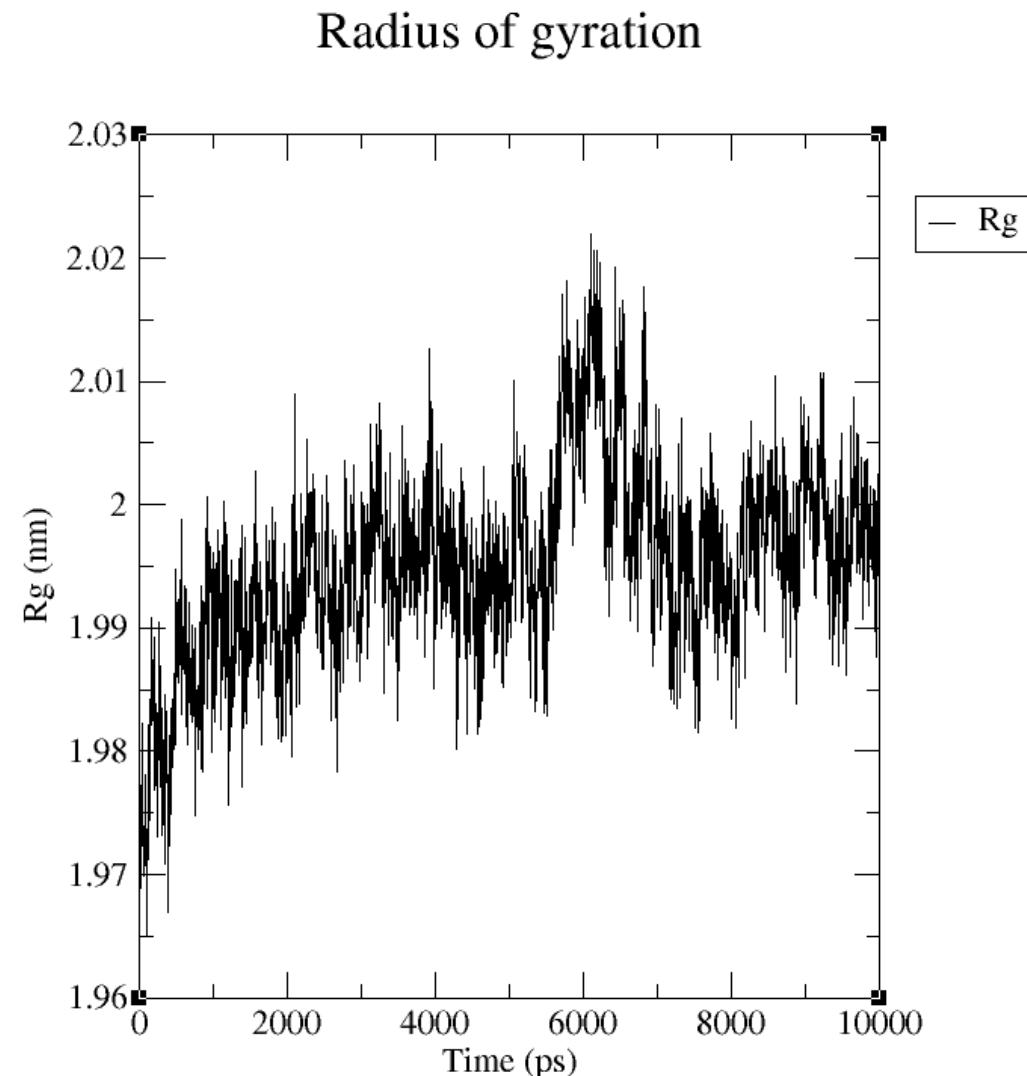
- RMSD all atoms to starting structure
- RMSD all atoms to average structure



# Radius of gyration

- indication of the shape of the molecule at each time
- Should also converge

```
g_gyrate  
-f traj.xtc  
-s topol.tpr  
-o radius-of-  
gyration.xvg
```



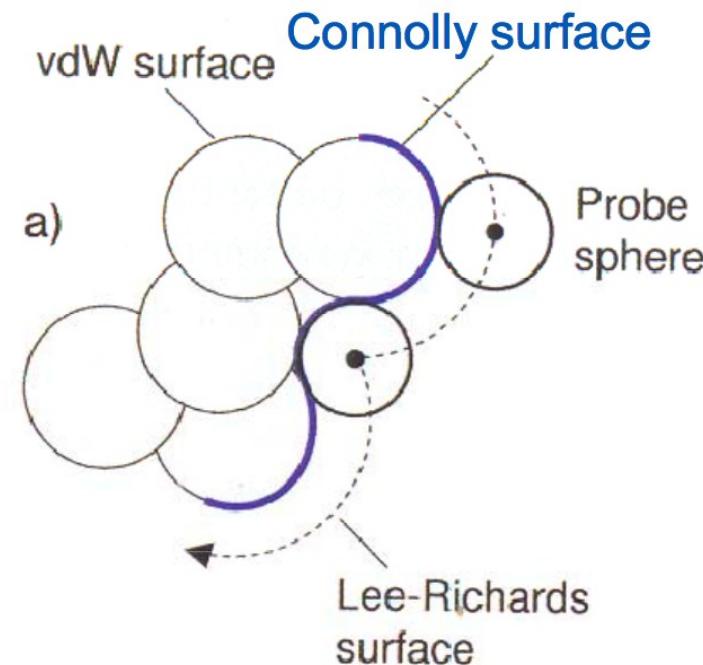
# Solvent Accessible Surface: g\_sas

- can compute: hydrophilic, hydrophobic or total SAS
- can use SAS for estimate of free solvation energy

**Call:** g\_sas

```
-f traj.xtc  
-s topol.tpr  
  
-o area.xvg
```

**(Lee-Richards = SAS)**

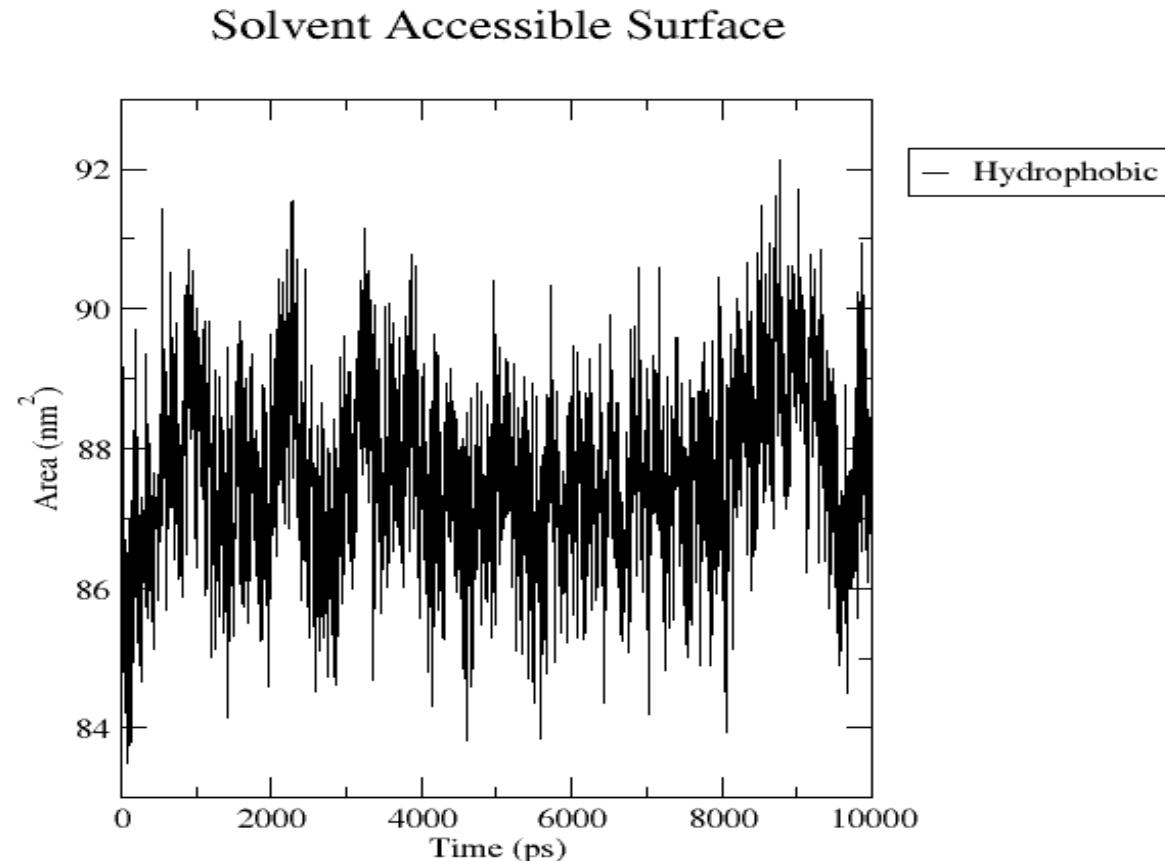


# Solvent Accessible Surface: g\_sas

- can compute: hydrophilic, hydrophobic or total SAS
- can use SAS for estimate of free solvation energy

**Call:** g\_sas

```
-f traj.xtc  
-s topol.tpr  
-o area.xvg
```



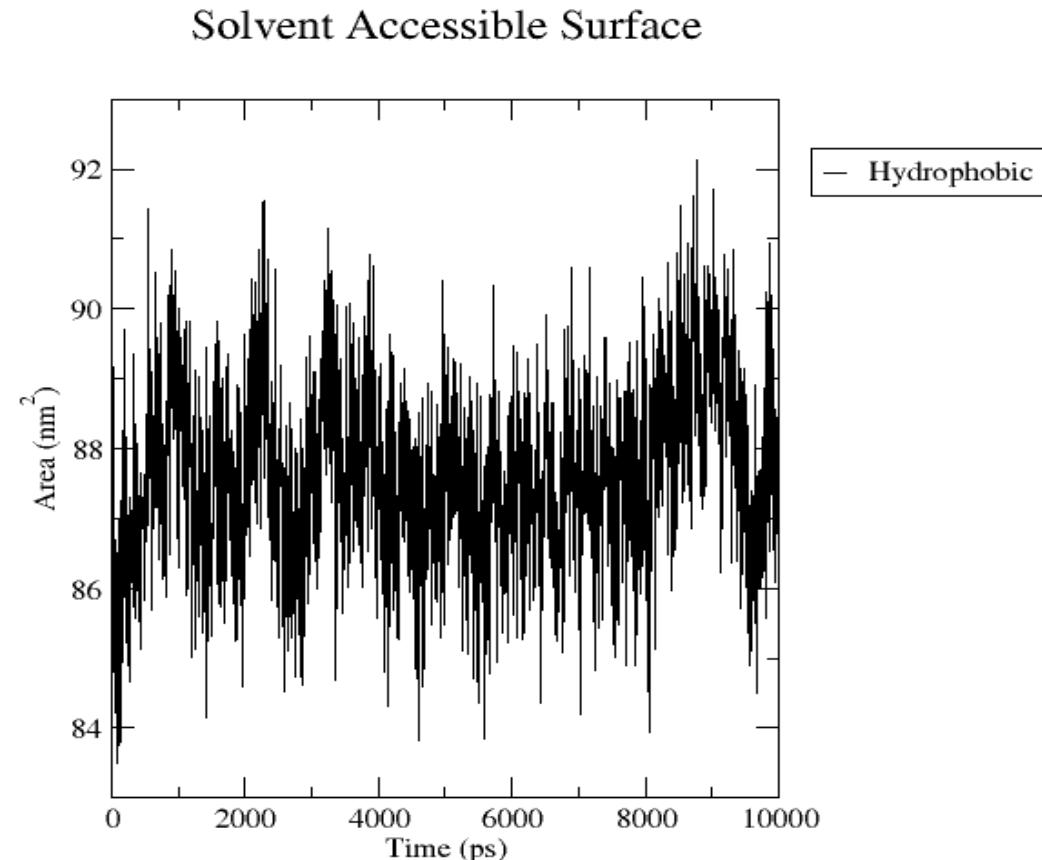
# Solvent Accessible Surface: g\_sas

- can compute: hydrophilic, hydrophobic or total SAS
- can use SAS for estimate of free solvation energy

**Call:** g\_sas

```
-f traj.xtc  
-s topol.tpr  
  
-o area.xvg
```

- **Additional options:**
  - Connolly surface (SES)
  - Avrg and  $\sigma$  of the area
  - Total volume & density

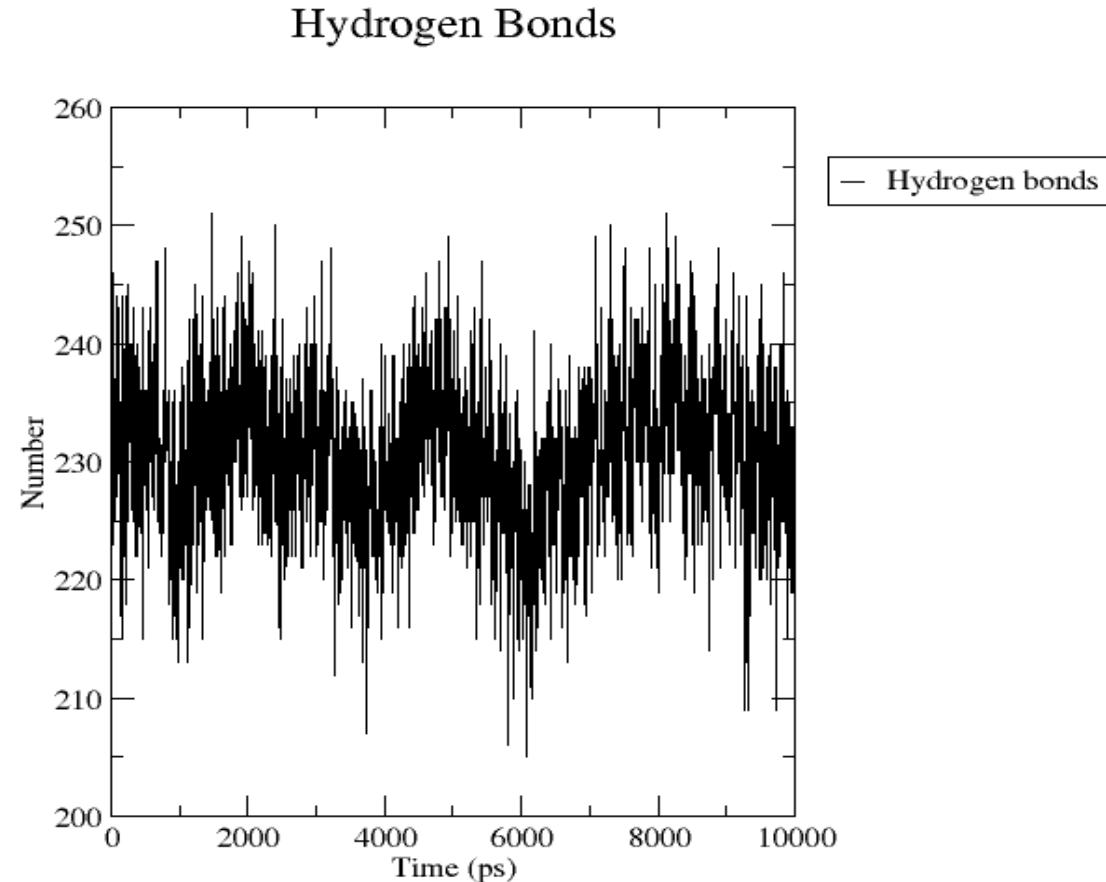


# Hydrogen bonds: g\_hbond

- Analyse hbonds btw all possible donors and acceptors
- Angle + distance criterion to define an hbond

- Call: **g\_hbond**

```
-f traj.xtc  
-s topol.tpr  
  
-num hbonds.xvg
```



# Hydrogen bonds: g\_hbond

- Analyse hbonds btw all possible donors and acceptors
- Angle + distance criterion to define an hbond

- **Call:** `g_hbond`

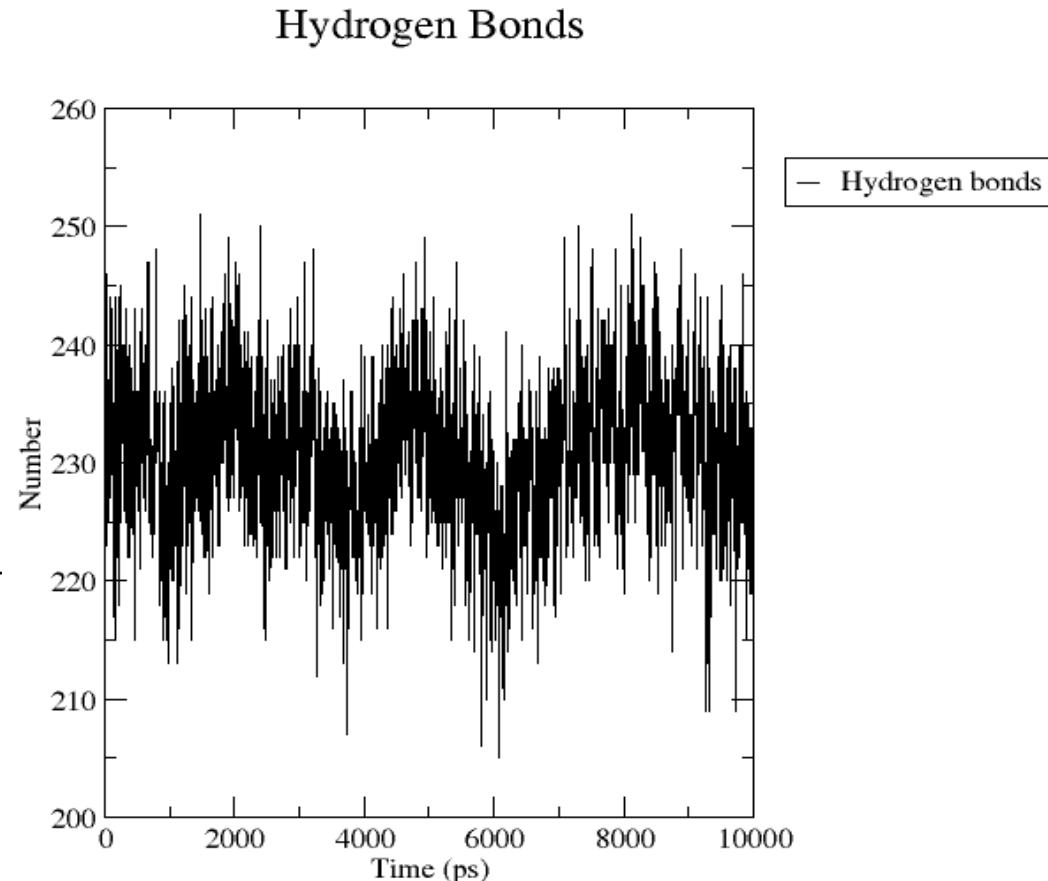
- f traj.xtc

- s topol.tpr

- num hbonds.xvg

- **Additional options:**

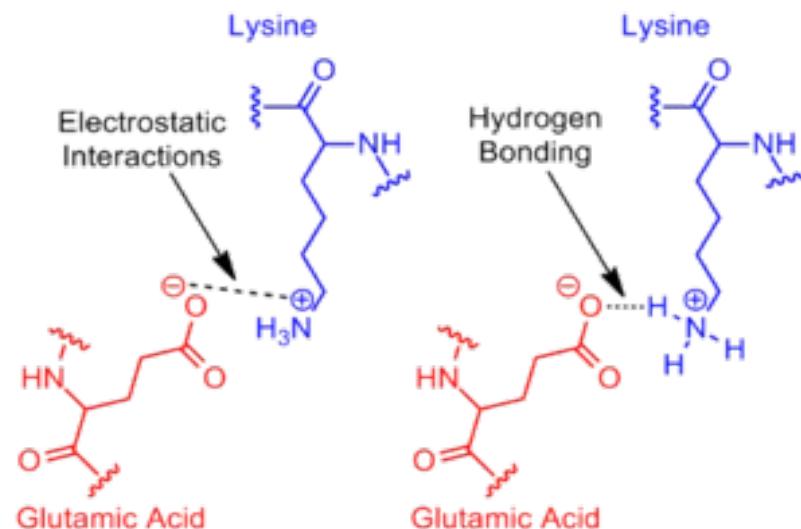
- Distance distribution
- Angle distribution
- #hbonds btw given res
- Lifetime of an hbond
- etc



# Saltbridges: g\_saltbr

- Stabilise structure and are therefore interesting for further analysis
- Plots distance between oppositely charged residues as a function of time

- Call: **g\_saltbr**
  - f traj.xtc
  - s topol.tp

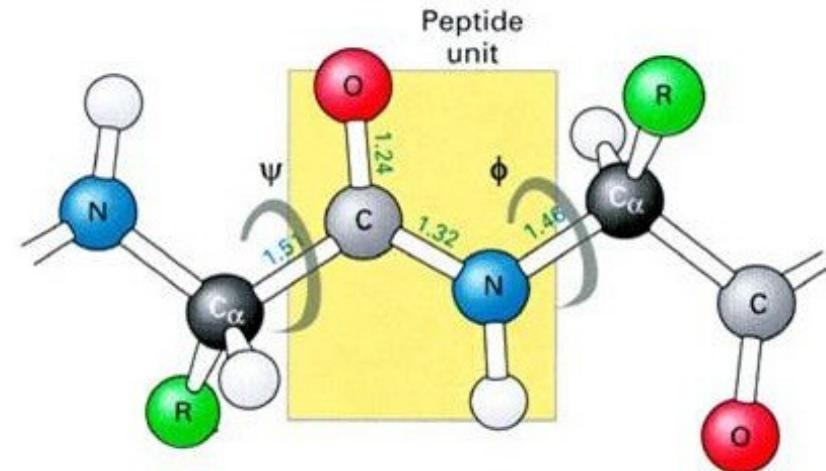


wikipedia.org

# Ramachandran plots: g\_rama

- calculates  $\phi$ - $\psi$  angle combinations for each residue as a function of time:

- Foreach residue:
  - Foreach frame:
    - Calc  $\phi$  and  $\psi$  angle
    - Write to file



<https://wiki.cmpb.ru/nl/images/c/ce/Prncpsi.jpg>

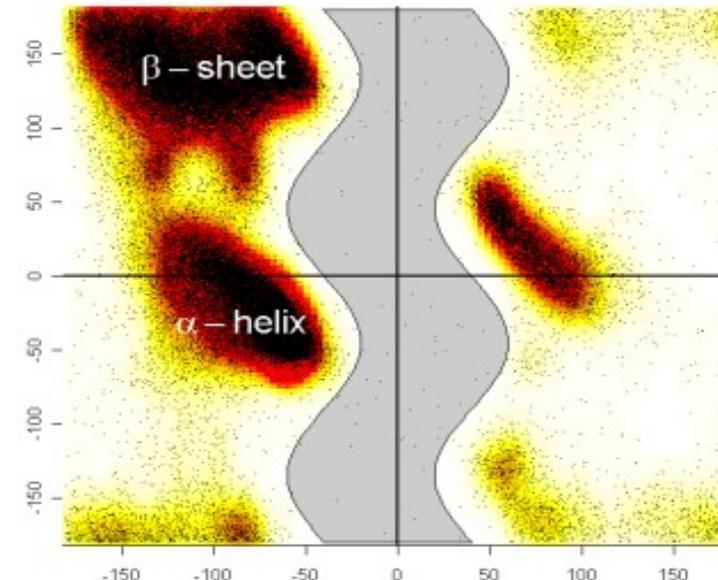
- Call: **g\_rama**
  - f traj.xtc
  - s topol.tp
  - o ramachandran.xvg

# Ramachandran plots: g\_rama

- calculates  $\varphi$ - $\psi$  angle combinations for each residue as a function of time:

- For each residue:

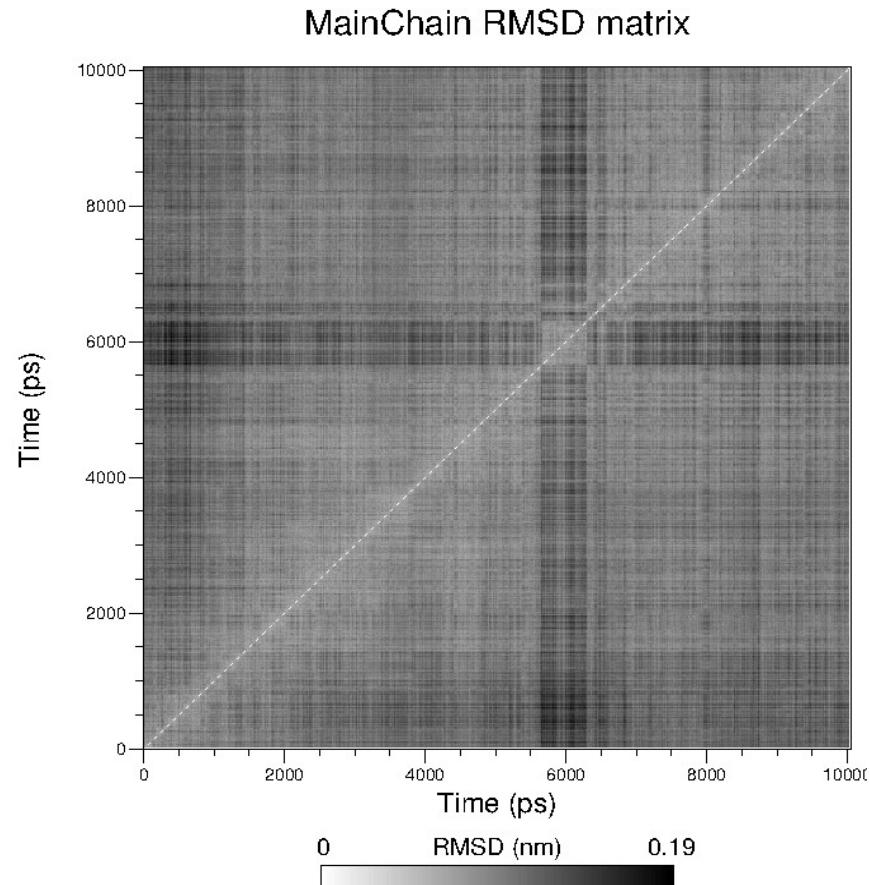
- For each frame:
  - Calc  $\varphi$  and  $\psi$  angle
  - Write to file



- extract angles for one residue for each frame (grep) to compute a ramachandran plot for that residue
- extract angles for one residue type for each frame to compute a ramachandran plot for that amino acid

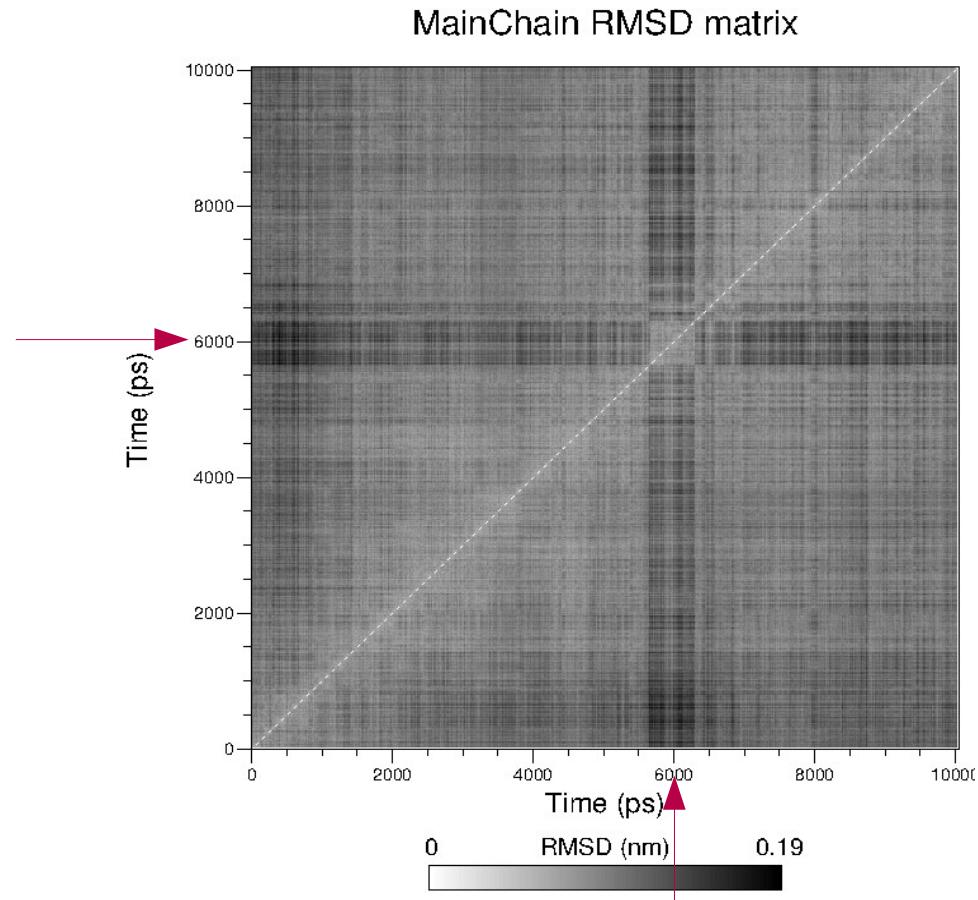
# RMSD for grouping structures: g\_rms

- Pairwise comparison of structures from trajectories:
- Plot frame against frame in a matrix
- Comparison measure:  
RMSD, rho, or rhosc
- Call: **g\_rms**
  - f traj.xtc
  - s topol.tpr
  - o rmsd.xpm (matrix)



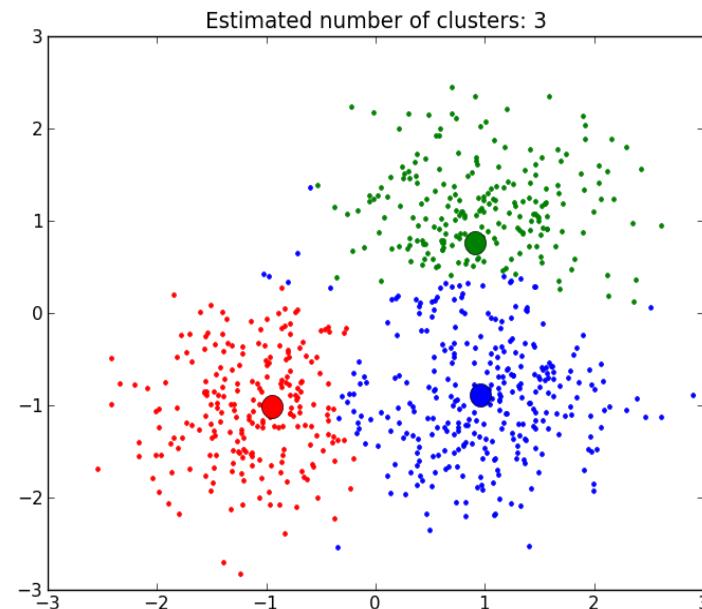
# RMSD for grouping structures: g\_rms

- Pairwise comparison of structures from trajectories:
- Plot frame against frame in a matrix
- Comparison measure:  
RMSD, rho, or rhosc
- Good for detecting  
transitions  
→ here: frame ~6000



# Cluster structures: g\_cluster

- Cluster either from .xtc or .xpm file
- Available clustering algorithms:
  - Single linkage
  - Gromos
  - MC
  - More
- **Call:** `g_rms`
  - f traj.xtc OR
  - dm matrix.xpm
  - o cluster.xpm

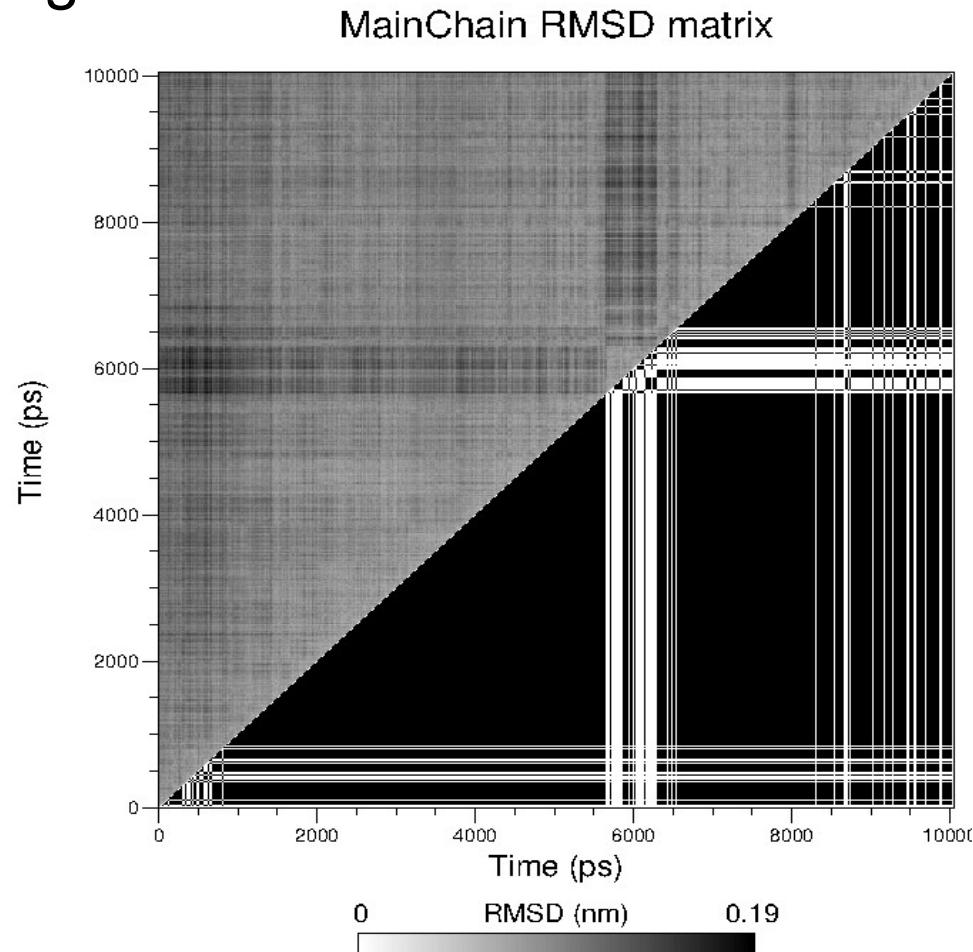


# Cluster structures: g\_cluster

- Cluster either from .xtc or .xpm file
- Offers several clustering algorithms

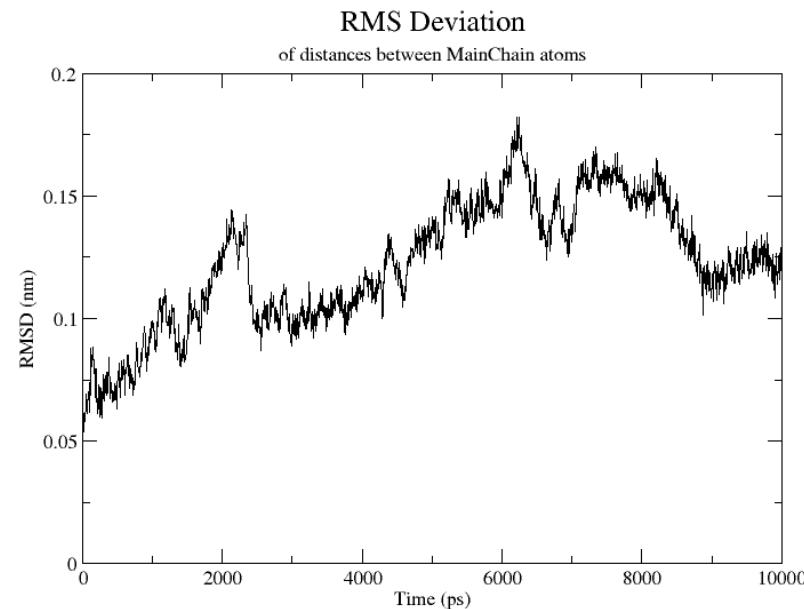
## .xpm matrix:

- upper half: pairwise RMSD values
- lower half: whether two structures are in same cluster



# Internal RMSD: g\_rmsdist

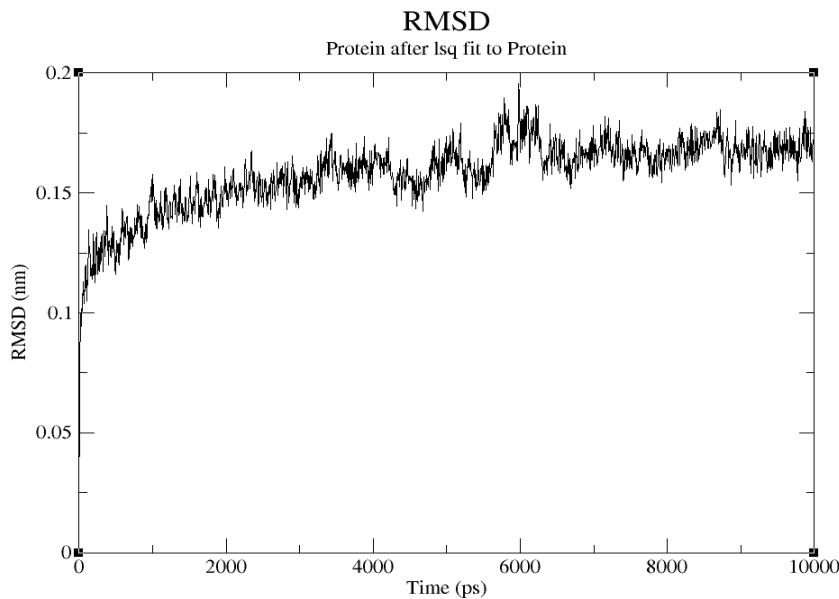
- Drawback of classical RMSD: involves least squares fitting → use distance based RMSD
- Call: **g\_rmsdist**
  - f traj.xtc
  - s topol.tpr
  - o dist-rmsd.xvg



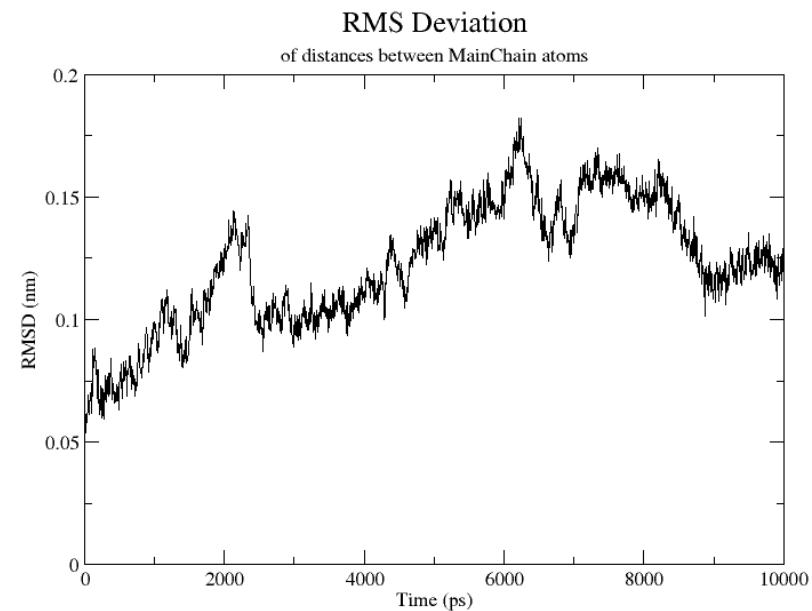
Internal RMSD

# Internal RMSD: g\_rmsdist

- Drawback of classical RMSD: involves least squares fitting → use distance based RMSD



RMSD



Internal  
RMSD

# Thanks for listening!

# Masters Practical 2012

offered by

Dr. A. Schafferhans, Dr. E. Kloppmann, Dr. M. Offmann

## **Molecular Dynamics Analysis**

presented by

Fanny Gatzmann and Susann Vorberg



Happy Birthday Julia !



Happy Birthday Julia !



## Overview



- Gromacs Output 204H
- MD Analysis – based on tutorial:  
<http://md.chem.rug.nl/~mdcourse/analysis1.html>
- Using analysis methods provided by Gromacs:  
– g\_energy, gmxcheck, g\_mindist, g\_rmsf, g\_rms,...
- Very useful:  
[http://www.gromacs.org/Documentation/Gromacs\\_Utils](http://www.gromacs.org/Documentation/Gromacs_Utils)

# Gromacs Results



_mdmdp	Simulation parameter input file: (time step, type of simulation, electrostatics, van der Waals,...)
_md.tpr	portable binary run input file: (generated by grompp, executed by mdrun) starting structure (coordinates and velocities), molecular topology and all simulation parameters
_md.edr	Binary portable energy file: energy terms that are saved in a simulation
_md.trr	full-precision trajectory file: coordinate, velocity, and force information
_md.xtc	compressed version of the trajectory: only coordinate, time, and box vector information

## standard checks: gmxcheck



- gmxcheck: *Getting the Right Output Means no Artefacts in Calculating Stuff*

```
gmxcheck -f .xtc
```

→ simulation finished properly?

```
gmxcheck -c .tpr
```

→ presence of coordinates, velocities,  
close contacts, atoms outside the box,,,



- Check log file (\_md.log): lot of statistics

	(Mnbfs)	(GFlops)	(ns/day)	(hour/ns)
Performance:	698.741	54.668	44.634	0.538

09.07.12

Molecular Dynamics Analysis

5

- f

Will show errors

How long simulation run

-c check tpr file

# Visualisation in Pymol: trjconv



trjconv – the Swiss army knife tool

```
trjconv -s topol.tpr -f traj.xtc  
-o prot.pdb -pbc nojump -dt 10
```

- Convert trajectories to pdb file
- extract 1000 frames
- remove the jumps over the boundaries

# Visualisation in Pymol



09.07.12

Molecular Dynamics Analysis

7

## Quality assurance



- tests for the convergence of thermodynamic parameters  
→ equilibrium reached?
  - Temperature
  - Pressure
  - potential and the kinetic energy ...
- No convergence of values:  
simulation has not yet reached thermal equilibrium  
→ extend simulation
- extracted from \_md.edr file with g\_energy

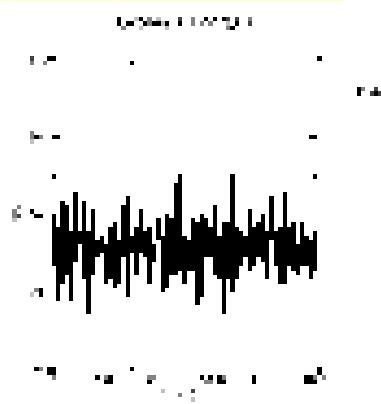
## thermal equilibrium: g\_energy



```
echo 12 0 | g_energy -f _md.edr  
-o temperature.xvg
```

Energy	Average	Err.Est.	RMSD	Tot-Drift
Temperature	297.914	1.38693	0.0072	-0.00042403 (K)

- reference temperature from  
\_mdmdp: 298K
- Visualize .xvg files
  - xmGrace file.xvg



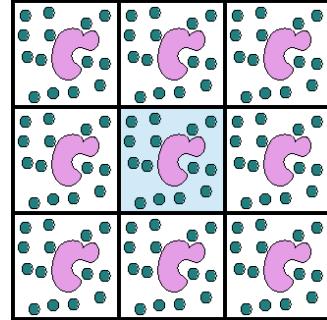
# Periodic images



- Using periodic boundary conditions:
  - forces between atoms calculated across periodic boundaries
- Check for direct interactions between periodic boxes:
  - ~ unphysical self-interactions
- minimal distance should be  $\geq 2\text{nm}$

```
g_mindist -f traj.xtc  
-s topol.tpr  
-od minperdist.xvg -pi
```

**pi:** Calculate minimum distance with periodic images

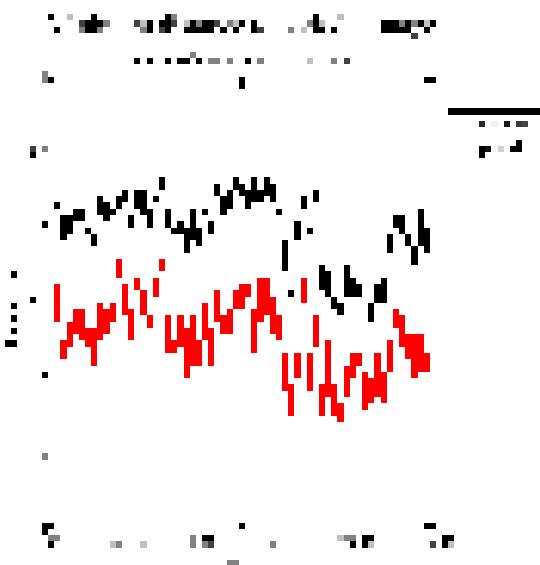


Karplus M, Kuriyan J., PNAS 2005 102 (19) 6679-6685

## Periodic images II



The shortest periodic distance is 1.6456 (nm) at time 7675 (ps), between atoms 15 and 4507

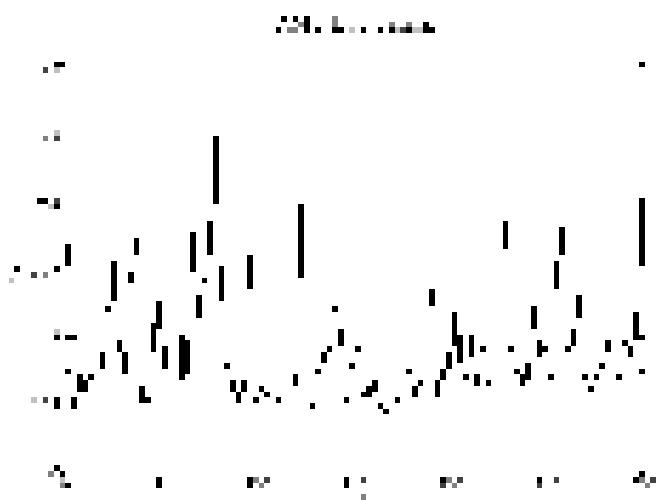


## RMSF: g\_rmsf



- = Root mean square fluctuations
  - = fluctuation of an atom about its average position
  - ~ crystallographic b-factors → flexibility
- 
- `g_rmsf -f traj.xtc -s topol.tpr  
-o rmsf-per-residue.xvg -ox average.pdb  
-oq bfactors.pdb -res`
  - per residue RMSD
  - calculate b-factors

## RMSF – per residue

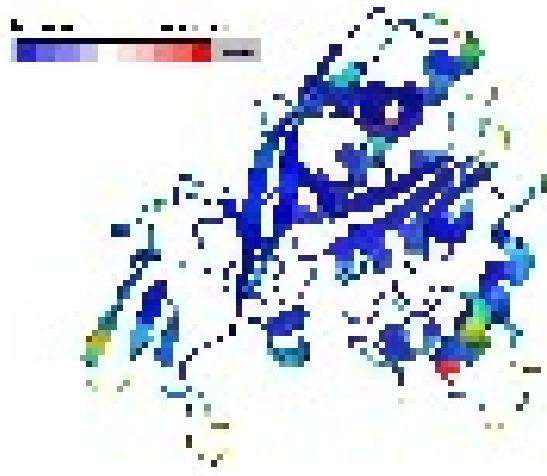


09.07.12

Molecular Dynamics Analysis

13

## RMSF - bfactors



09.07.12

Molecular Dynamics Analysis

14

## RMSF - average



Average pdb is an unphysical structure:  
effect of averaging over conformations.

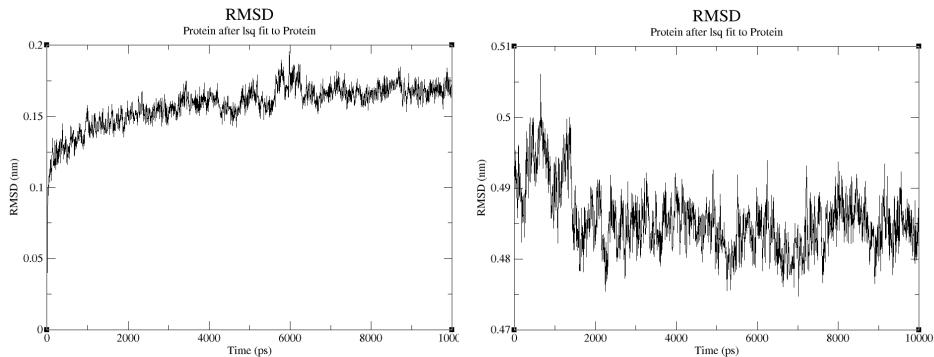
## RMSD: g\_rms



- g\_rmsf also calculates average structure
- Calculate RMSD with respect to the average structure
  - indicator of convergence of the structure towards an equilibrium state
- RMSD to start structure:  
`g_rms -f traj_nojump.xtc -s topol.tpr  
-o rmsd-backbone-vs-start.xvg`
- RMSD to average structure:  
`g_rms -f protein.xtc -s average.pdb -o  
rmsd-backbone-vs-average.xvg`

## Convergence of RMSD

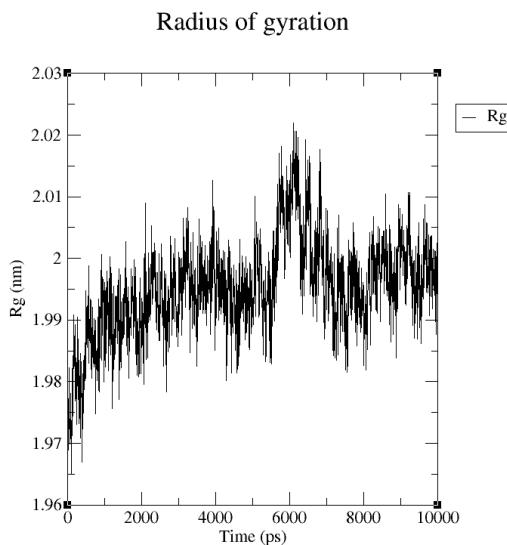
- RMSD all atoms to starting structure
- RMSD all atoms to average structure



## Radius of gyration

- indication of the shape of the molecule at each time
- Should also converge

```
g_gyrate  
-f traj.xtc  
-s topol.tpr  
-o radius-of-  
gyration.xvg
```



## Solvent Accessible Surface: g\_sas

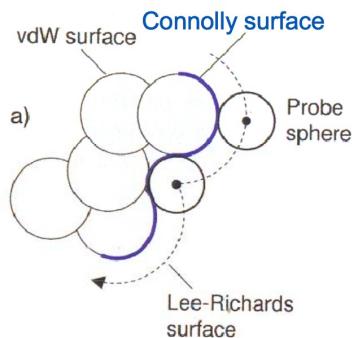


- can compute: hydrophilic, hydrophobic or total SAS
- can use SAS for estimate of free solvation energy

Call: **g\_sas**

```
-f traj.xtc  
-s topol.tpr  
-o area.xvg
```

**(Lee-Richards = SAS)**



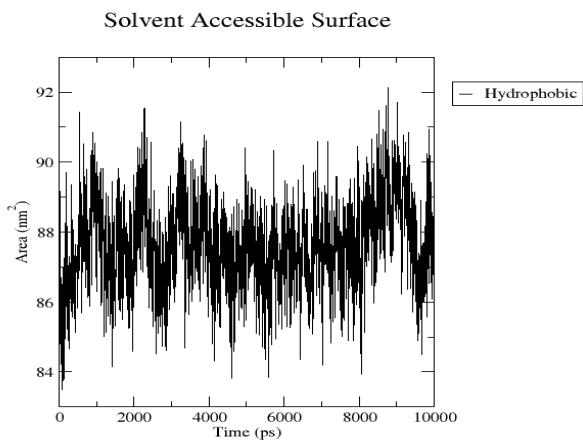
## Solvent Accessible Surface: g\_sas



- can compute: hydrophilic, hydrophobic or total SAS
- can use SAS for estimate of free solvation energy

Call: **g\_sas**

```
-f traj.xtc  
-s topol.tpr  
  
-o area.xvg
```



## Solvent Accessible Surface: g\_sas



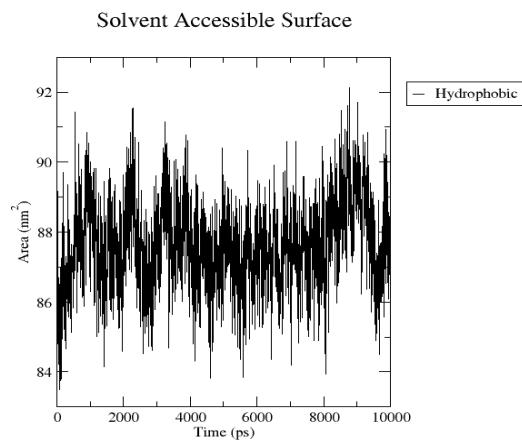
- can compute: hydrophilic, hydrophobic or total SAS
- can use SAS for estimate of free solvation energy

Call: **g\_sas**

```
-f traj.xtc  
-s topol.tpr  
  
-o area.xvg
```

- **Additional options:**

- Connolly surface (SES)
- Avrg and  $\sigma$  of the area
- Total volume & density

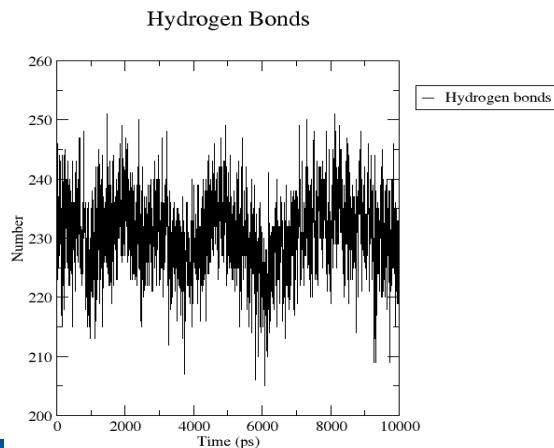


## Hydrogen bonds: g\_hbond



- Analyse hbonds btw all possible donors and acceptors
- Angle + distance criterion to define an hbond

```
• Call: g_hbond  
  -f traj.xtc  
  -s topol.tpr  
  
  -num hbonds.xvg
```



## Hydrogen bonds: g\_hbond

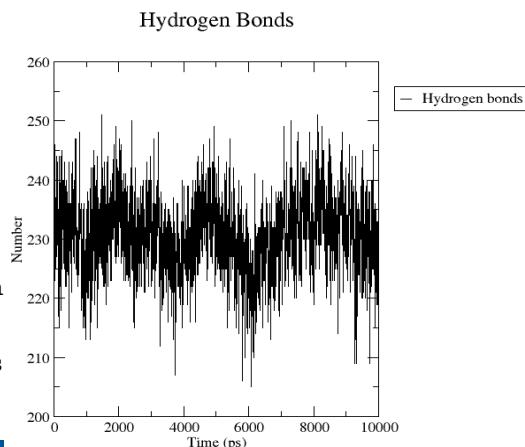
- Analyse hbonds btw all possible donors and acceptors
- Angle + distance criterion to define an hbond

- **Call:** g\_hbond

```
-f traj.xtc  
-s topol.tpr  
  
-num hbonds.xvg
```

- **Additional options:**

- Distance distribution
- Angle distribution
- #hbonds btw given res
- Lifetime of an hbond
- etc

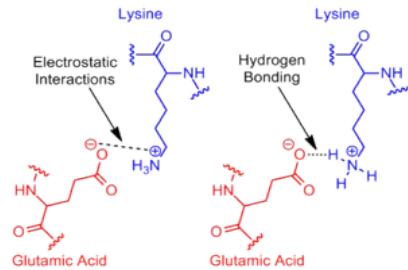


## Saltbridges: g\_saltbr



- Stabilise structure and are therefore interesting for further analysis
- Plots distance between oppositely charged residues as a function of time

- **Call:** g\_saltbr
  - f traj.xtc
  - s topol.tp



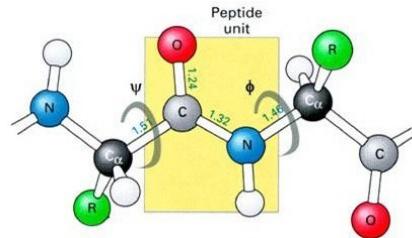
wikipedia.org

## Ramachandran plots: g\_rama



- calculates  $\phi$ - $\psi$  angle combinations for each residue as a function of time:

- For each residue:
  - For each frame:
    - Calc  $\phi$  and  $\psi$  angle
    - Write to file
- Call: **g\_rama**
  - f traj.xtc
  - s topol.tp
  - o ramachandran.xvg



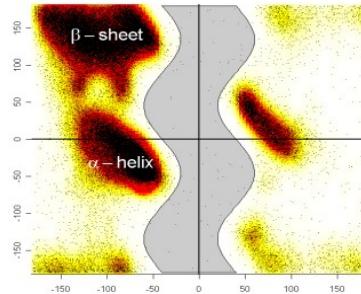
<https://www.cmpb.ru.nir/images/3/00/rmipsi.jpg>

## Ramachandran plots: g\_rama



- calculates  $\phi$ - $\psi$  angle combinations for each residue as a function of time:

- Foreach residue:
  - Foreach frame:
    - Calc  $\phi$  and  $\psi$  angle
    - Write to file

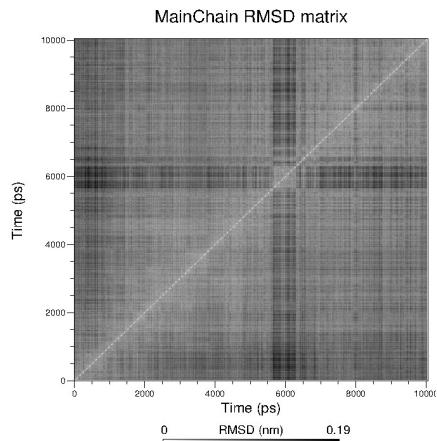


→ extract angles for one residue for each frame (grep) to compute a ramachandran plot for that residue

→ extract angles for one residue type for each frame to compute a ramachandran plot for that amino acid

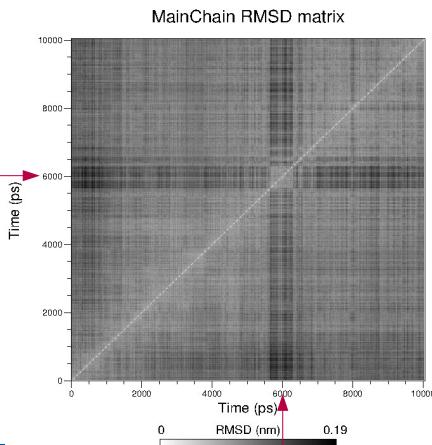
## RMSD for grouping structures: g\_rms

- Pairwise comparison of structures from trajectories:
- Plot frame against frame in a matrix
- Comparison measure:  
RMSD, rho, or rhosc
- Call: **g\_rms**
  - f traj.xtc
  - s topol.tpr
  - o rmsd.xpm (matrix)



## RMSD for grouping structures: g\_rms

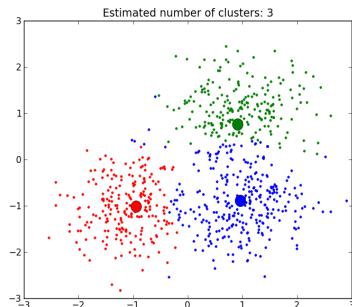
- Pairwise comparison of structures from trajectories:
- Plot frame against frame in a matrix
- Comparison measure:  
RMSD, rho, or rhosc
- Good for detecting transitions  
→ here: frame ~6000



## Cluster structures: g\_cluster



- Cluster either from .xtc or .xpm file
- Available clustering algorithms:
  - Single linkage
  - Gromos
  - MC
  - More
- Call: **g\_rms**
  - f traj.xtc OR
  - dm matrix.xpm
  - o cluster.xpm



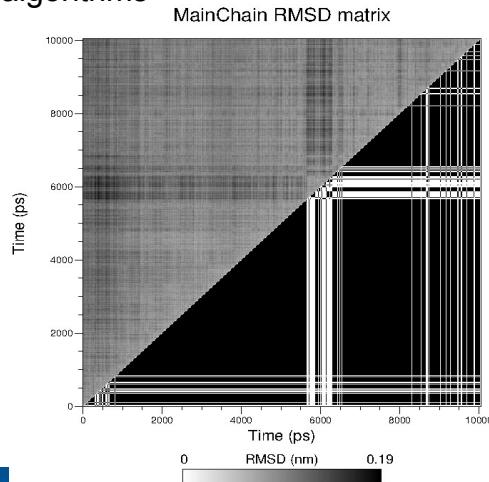
## Cluster structures: g\_cluster



- Cluster either from .xtc or .xpm file
- Offers several clustering algorithms

### .xpm matrix:

- upper half: pairwise RMSD values
- lower half: whether two structures are in same cluster



09.07.12

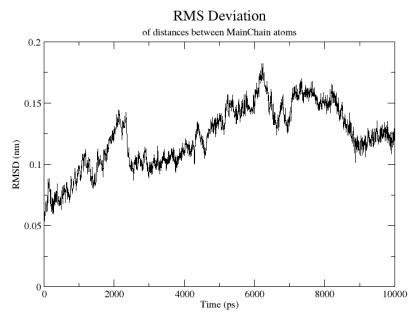
Molecular Dynamics Analysis

30

## Internal RMSD: g\_rmsdist



- Drawback of classical RMSD: involves least squares fitting → use distance based RMSD
- **Call: g\_rmsdist**  
-f traj.xtc  
-s topol.tpr  
-o dist-rmsd.xvg

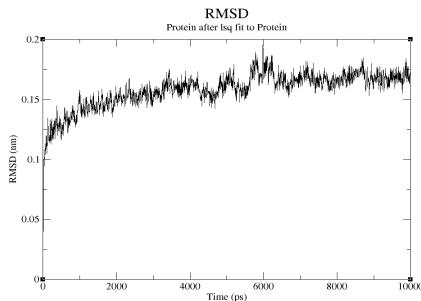


Internal RMSD

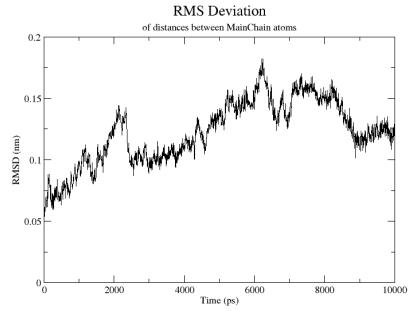
## Internal RMSD: g\_rmsdist



- Drawback of classical RMSD: involves least squares fitting → use distance based RMSD



RMSD



Internal  
RMSD

# Thanks for listening!