Introduction to Molecular Dynamics on GROMACS

In *silico* Methodologies for Biochemist

2020

pierre.leprovost@oulu.fi

Outline

Molecular Dynamics (MD)

GROMACS 101

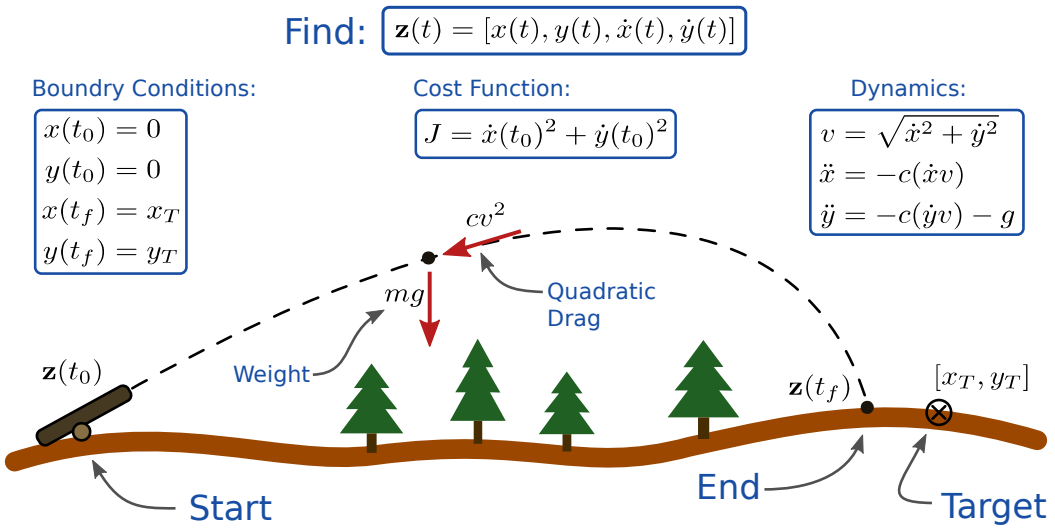
Exercices

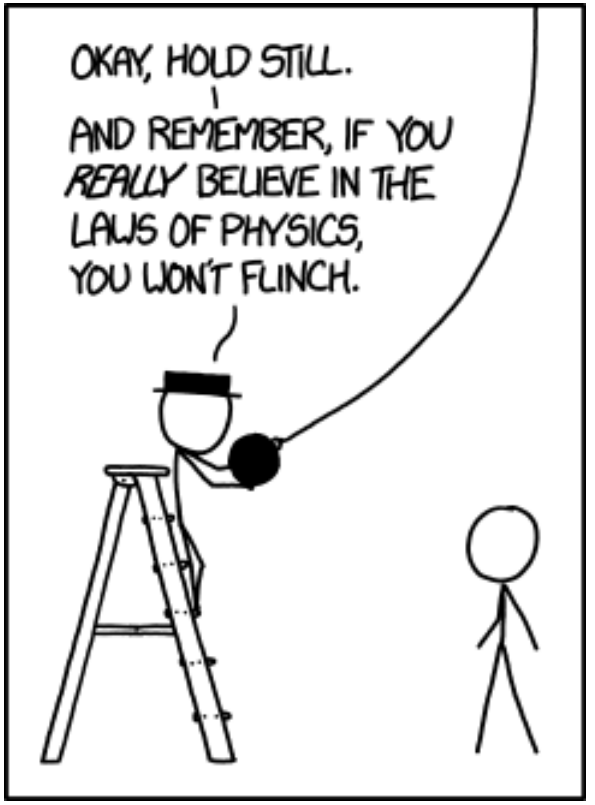
- 01 Water Box
- 02 Small Peptide
- 03 Coarse Grain MD with MARTINI

Motion

- Newtonian Mechanics
  - Second Law

    F=m**a**





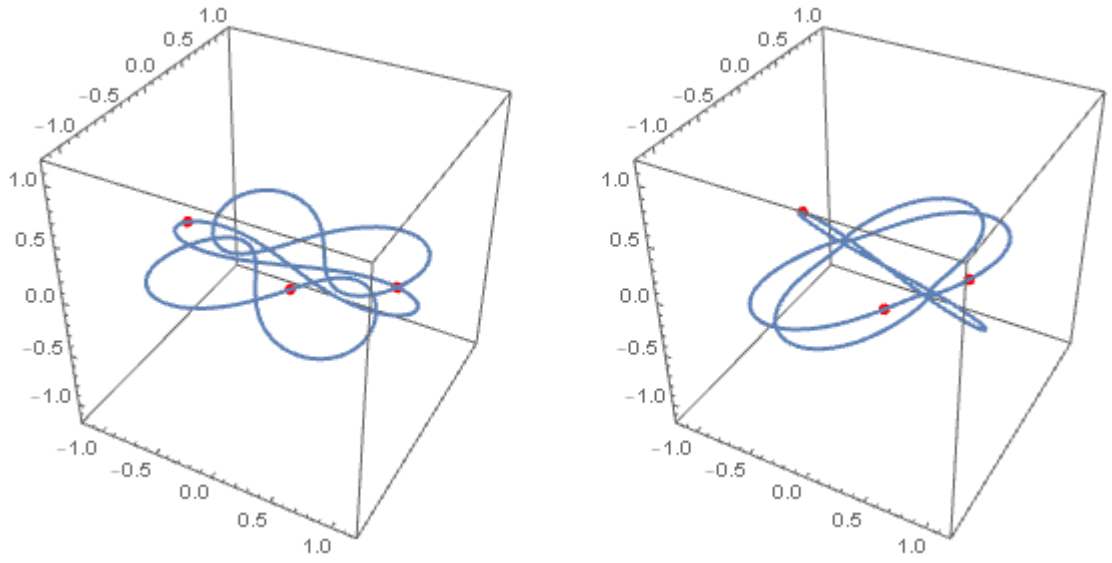
Molecular Dynamics

Motion

- The three body problem

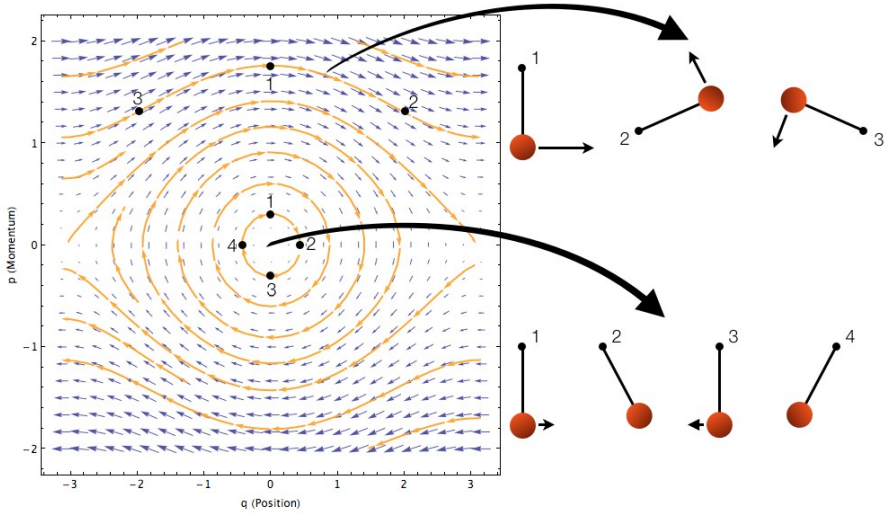
  Above 3 body analytical solution do not exist

  (3 three body problem has analytical solution in particular cases)



Solution ?

- Numerical Integration
  - Sample of the phase space of n-body systems
    - State of the system described by position and momentum



The idea:

- Iterative
- Calculate solutions for time intervals

Molecular Dynamics

Numerical Integration

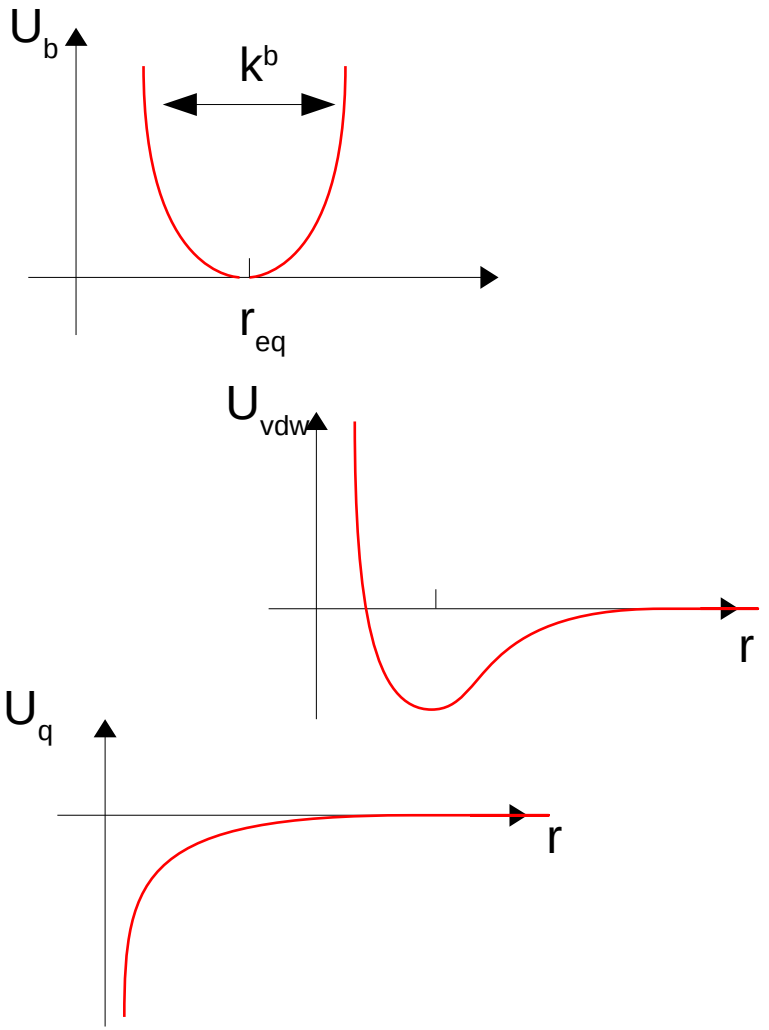
- A short algorithm


```
> Get initial positions
> Compute initial force
> Compute initial velocities
> For n in number of steps:
> Do
    Compute Force
    Update Position
    Update Velocity
> Done
```

Force Fields

Describe the interaction between bodies

- Example: Atoms
  - Bonded
    - Bond
    - Angle
    - Dihedral
  - Non-bonded
    - Electrostatics
    - Van der Waals



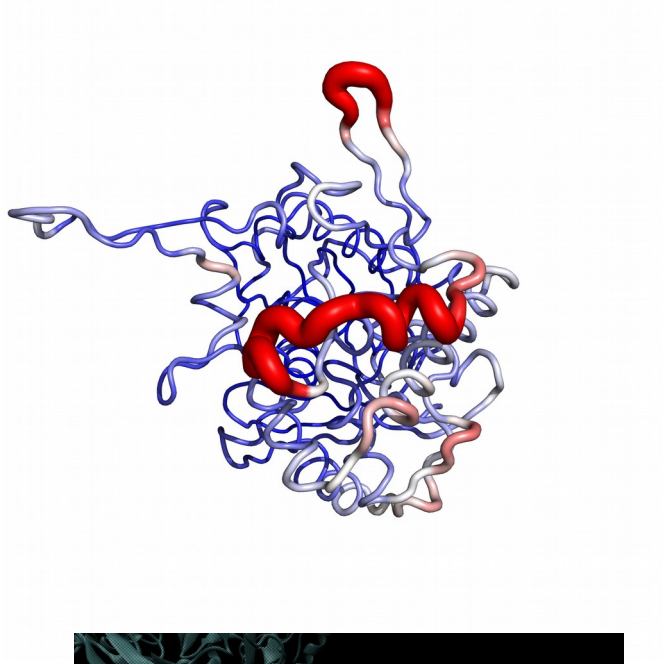
Molecular Dynamics

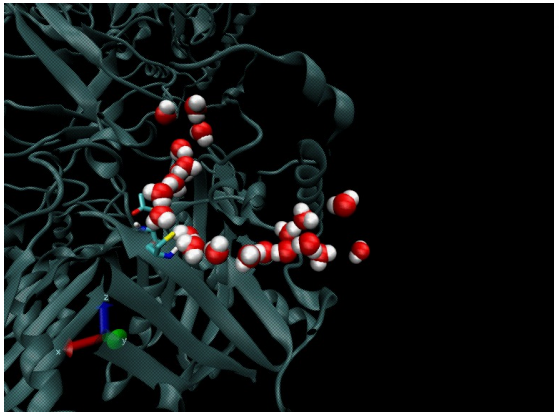
Key Properties

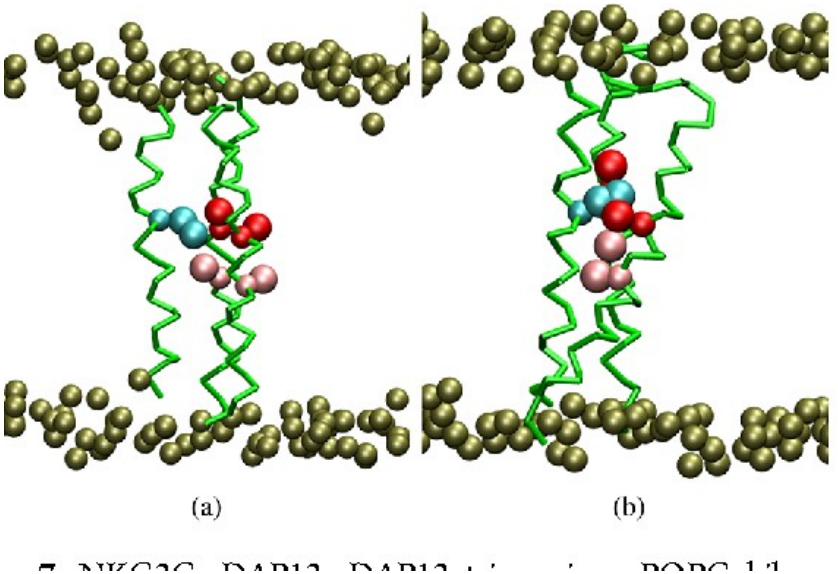
- Atoms keep moving
  - They can't get perfectly trapped in low energy minimum.
  - With time atoms sample a statistical distribution.
  - But there is not enough time sample all configurations.
- Total energy should be conserved
  - Numerical integration comes with error that tend to increase the total energy.
  - There are methods to ensure that.

Applications

- Structural Changes and Mechanism





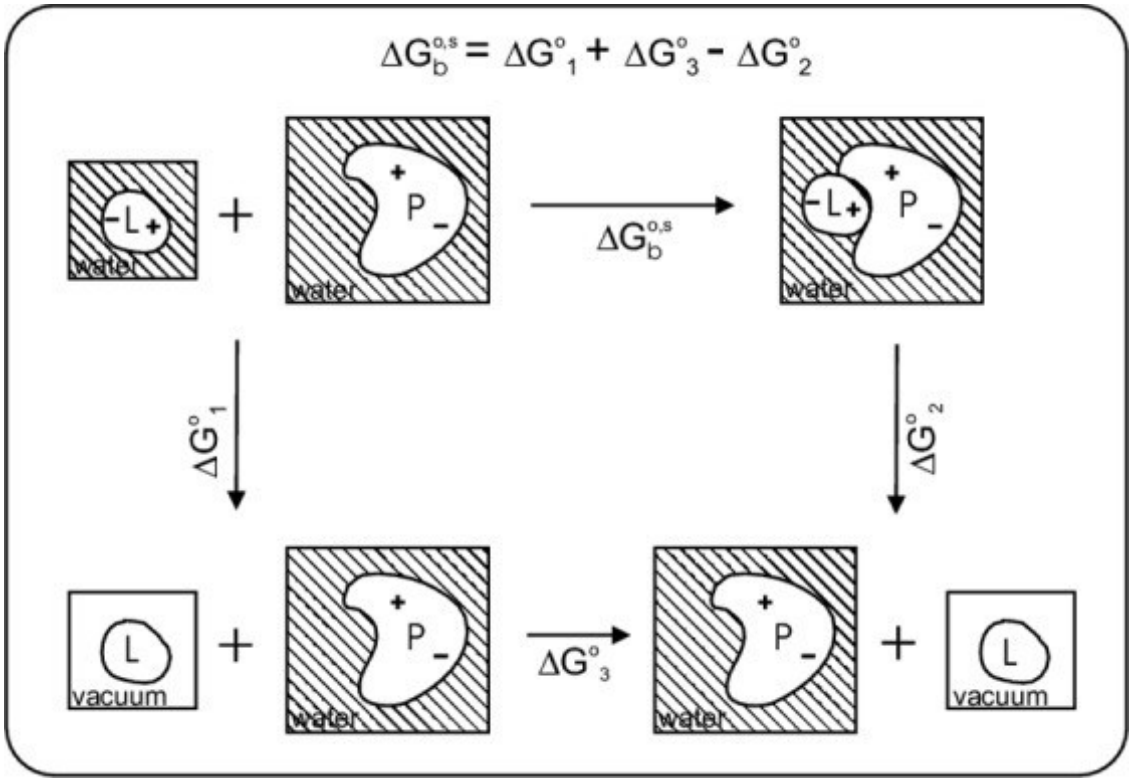


Sharma, Juffer 2013

Molecular Dynamics

Applications

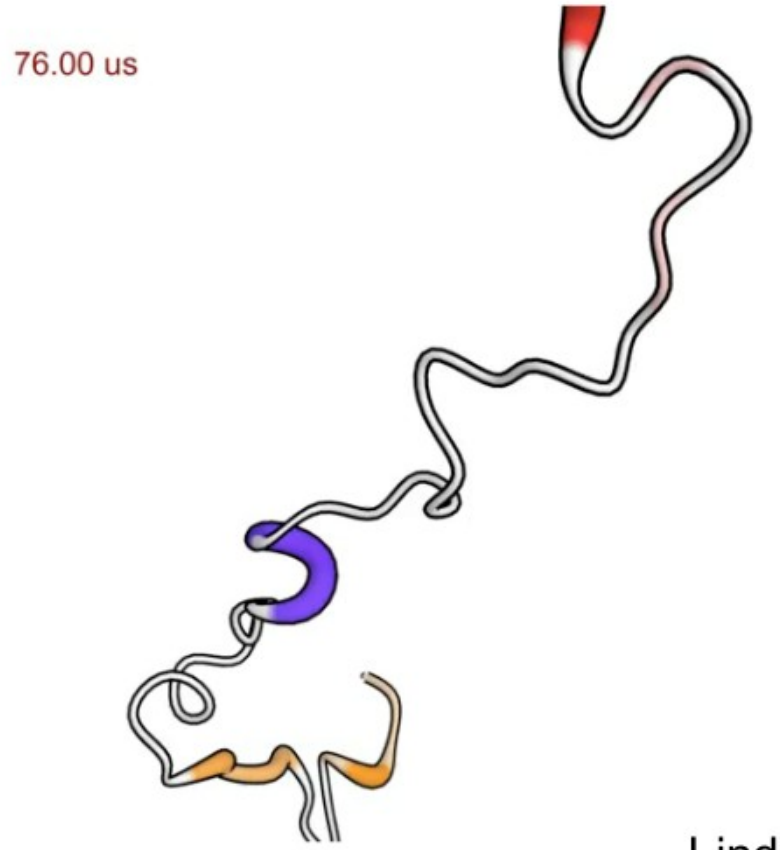
- Ligand binding and affinities



Donnini, Juffer 2003

Applications

- And many more
  - Folding



Lindorff-Larsen et al., *Science* 2011

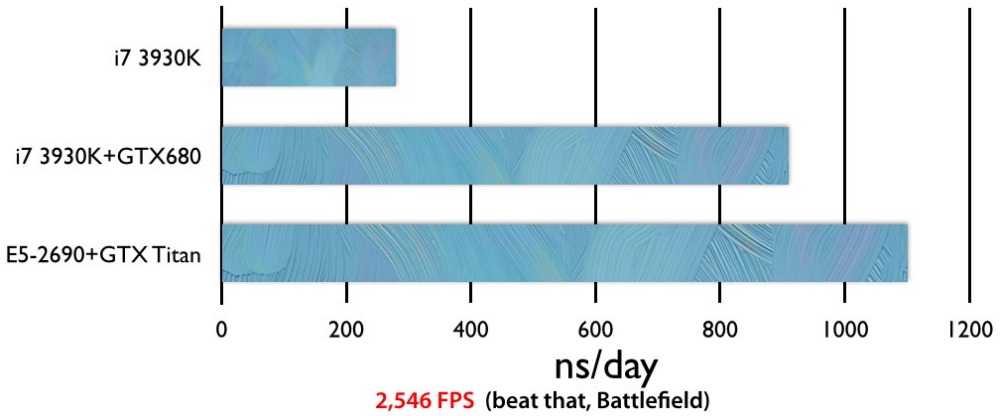
What is GROMACS ?

- GROMACS is a versatile package to perform molecular dynamics i.e. simulate the Newtonian equations of motion for system with hundred to millions of particles
- Aimed at biochemical system but can also use for non-organic system (carbon nano-tubes or galaxies)

Why GROMACS ?

GROMACS

- High Performance Computing
- Make good use of GPU
- User Friendly
- Well Documented
- Active Community
- Tons of Tools
- Free



The Force Field Jungle

- All Atoms
  - Amber
  - CHARMM
  - OPLS
- United Atoms
  - GROMOS
- Coarse Grained
  - MARTINI
  - GO

Using GROMACS

- Prefix

  All GROMACS commands are preceded by gmx

- If you are lost GROMACS is here to help

```
gerard@plop:~$ gmx help
```

- Running a command

```
gerard@plop:~$ gmx <command> -<field> <input>
gerard@plop:~$ gmx help solvate
```

01 - Water Box

- Preparation

Exercises

```
gerard@plop:~$ mkdir md_exercises
gerard@plop:~$ cd md_exercises
gerard@plop:~$ mkdir 01_water_box
```

01 – Water Box

Using `gmx solvate` generate of water of 5 by 7 by 10

Exercises

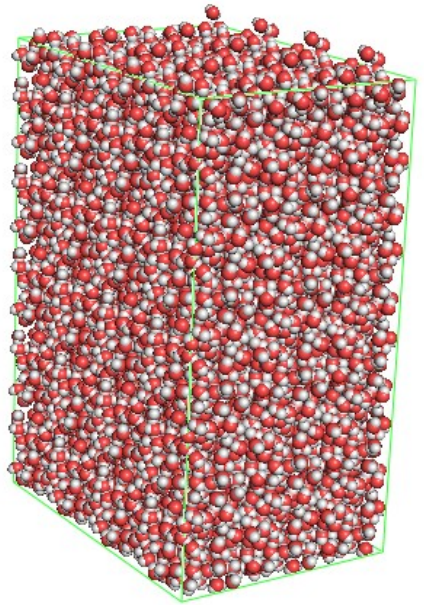
01 – Water Box

Exercises

Using `gmx solvate` generate of water of 5 by 7 by 10

```
gmx solvate -cs -box 5 7 10
```

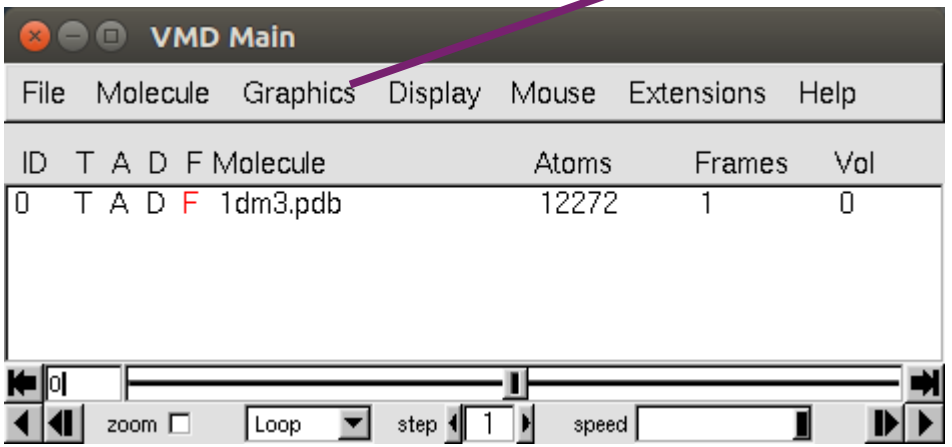
```
vmd out.gro
```

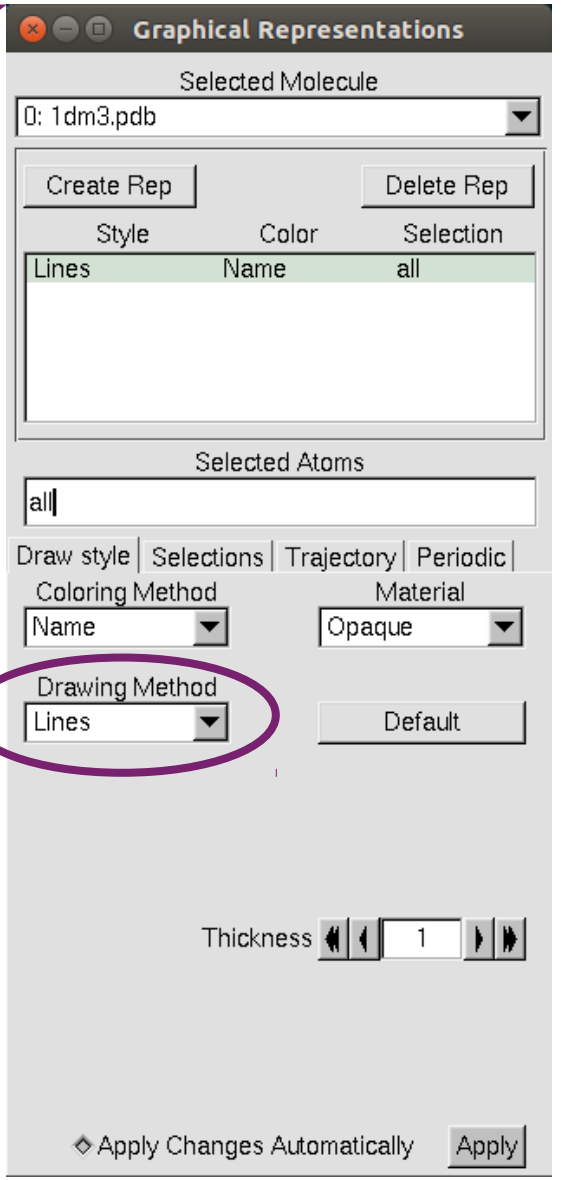


01 – Water Box

- VMD

Exercises



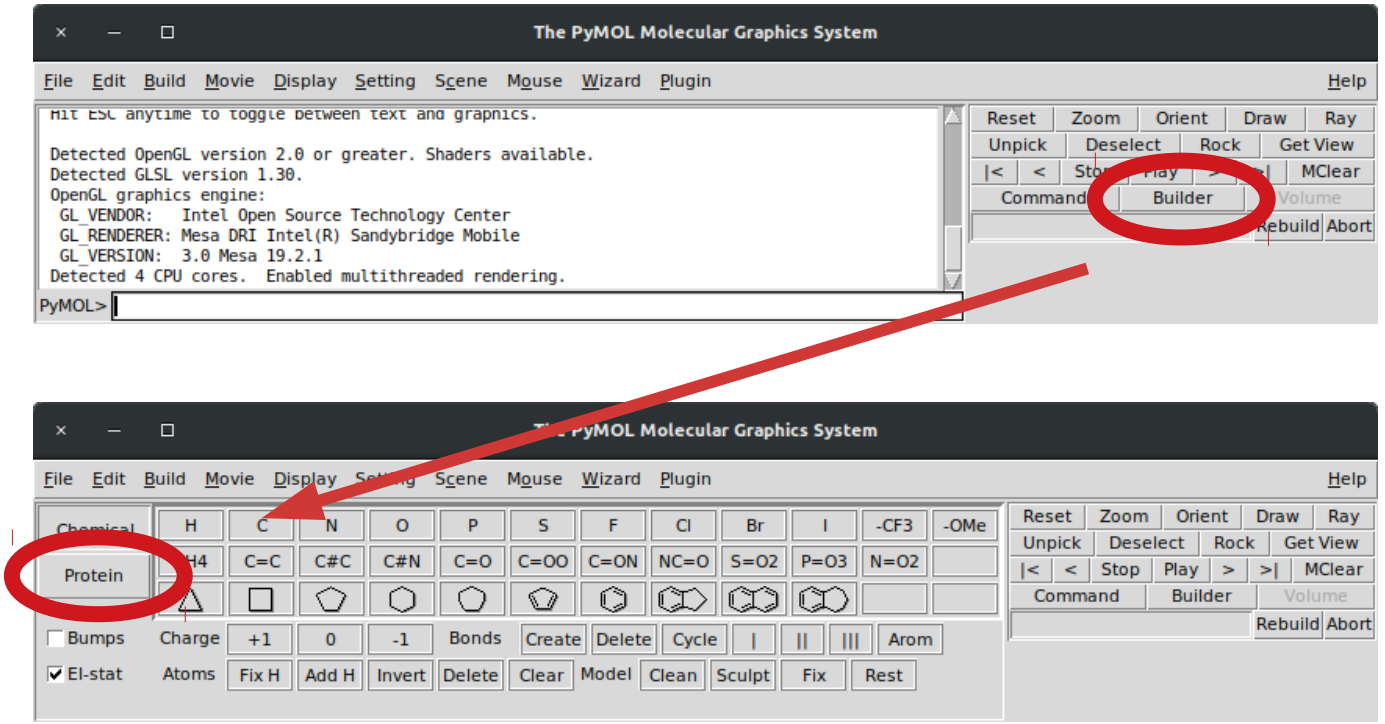


02 – Small Peptide

- Builder Tool on PyMOL

Build a small chain of Ala-His-Ala-Ser-Ala

Exercises



02 – Small Peptide

- Prepare the peptides

  Protonation states with Propka:

  http://nbcr-222.ucsd.edu/pdb2pqr_2.0.0/

  Propka is a Heuristic Program for computing pKa of residues in protein. Thus allow us to choose meaningful protonation states.

Exercises

What are the pKa of the Histidine ?

02 – Small Peptide

- Prepare the peptides

  Protonation states with Propka:

  http://nbcr-222.ucsd.edu/pdb2pqr_2.0.0/

  Propka is a Heuristic Program for computing pKa of residues in protein. Thus allow us to choose meaningful protonation states.

Exercises

What are the pKa of the Histidine ?

```
Open the file peptide.propka
Around the standard pKa 6.5, neutral histidine
```

02 – Small Peptide

- Generate a topology

  gmx pdb2gmx is your friend … and your enemy

```
gerard@plop:~$ gmx pdb2gmx -f peptide.pdb
gerard@plop:~$ gmx pdb2gmx -f peptide.pdb -ignh
```

  Choose AMBER99 and TIP3P

Exercises

What is the total charge of the system ?

02 – Small Peptide

- Generate a topology

  gmx pdb2gmx is your friend … and your enemy

```
gerard@plop:~$ gmx pdb2gmx -f peptide.pdb
gerard@plop:~$ gmx pdb2gmx -f peptide.pdb -ignh
```

  Choose AMBER99 and TIP3P

Exercises

What is the total charge of the system ?

```
The program is prompting a lot of info it is in it.
0.0
```

02 – Small Peptide

Exercises

- Add a box

```
gmx editconf -f conf.gro -d 1.0 -bt triclinic -o boxed.gro
```

- Solvate

```
gmx solvate -cp boxed.gro -cs -o solvated.gro -p topol.top -o ions.tpr
```

- Visualize the peptide in its box

  Check the protonation of the Histidine

02 – Small Peptide

- Neutralize and add ion pressure

  To replicate biological condition.

```
gmx grompp -f mdps/em.mdp -c solvated.gro -p topol.top -o ions.tpr
gmx genion -s ions.tpr -o neutralized.gro -p topol.top -neutral -conc 0.1
```

Exercises

What type are the ions ? How many are there ?
Visualize them using PyMOL or VMD.

PyMOL cannot read .gro files.
If you want to use PyMOL you must convert it to .pdb
Hint: use gmx editconf

02 – Small Peptide

- Minimization

  Optimization of the structure of the protein and the solvent to minimize the potential energy of the system. This is NOT MD yet.

Exercises

```
gmx grompp -f mdps/em.mdp -c solvated.gro -p topol.top -o ions.tpr
gmx genion -s ions.tpr -o neutralized.gro -p topol.top -neutral -conc 0.1
```

If your run converged, what is the final value of Fmax ?

Open the log file and scroll to the end

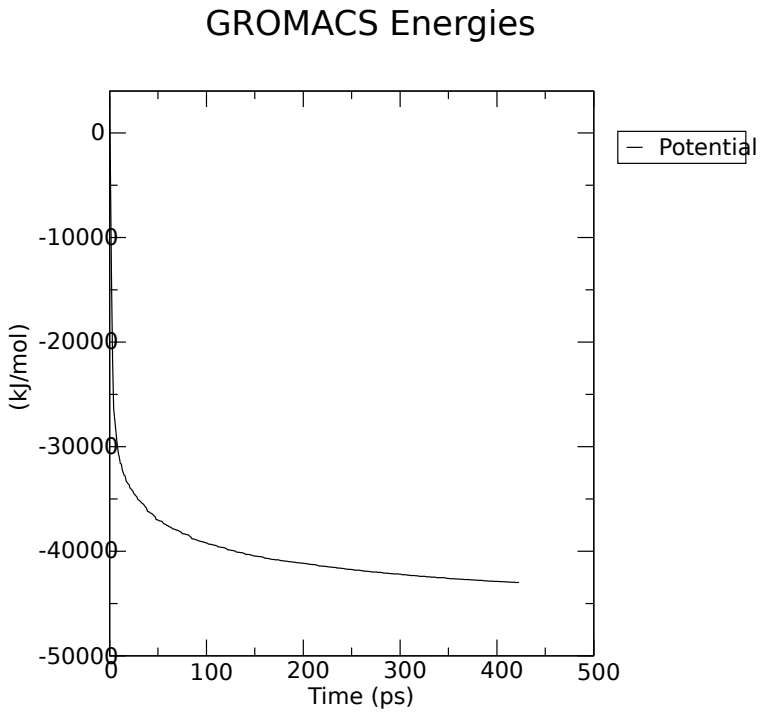
02 – Small Peptide

- Minimization

  Grace is a Data visualization program

```
gmx energy -f em.edr -o potential.xvg
xmgrace potential.xvg
```

Exercises



02 – Small Peptide

- Temperature Coupling

  "Weak Coupling"

  Berendsen, Andersen, Nosé-Hoover, Velocity Re-scaling

  Velocity Rescaling: Heat flow in and out of the system by scaling the velocity of each particles plus a stochastic terms ensuring correct kinetic energy

Exercises

02 – Small Peptide

- Equilibration NVT

  N: Constant number of particles

  V: Constant Volume

  T: Constant temperature

  Bring and keep the system at the desired temperature.

  Make use of Temperature Coupling Methods.

Exercises

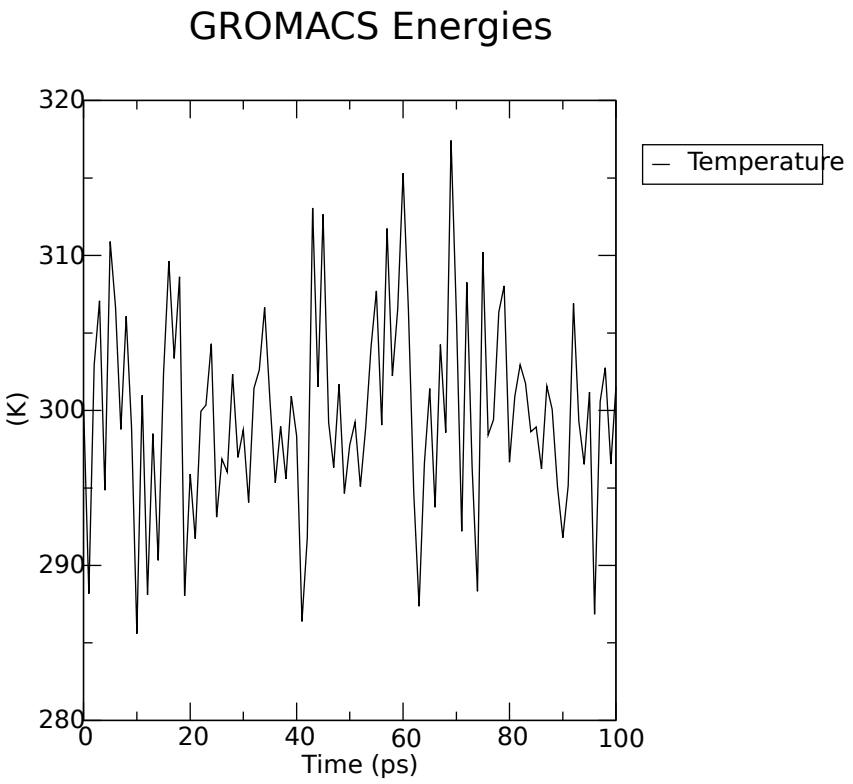
```
gmx grompp -f mdps/nvt.mdp -c em.gro -r em.gro …
-p topol.top -o nvt.tpr
gmx mdrun -deffnm nvt -v
```

02 – Small Peptide

- Equilibration NVT

```
gmx energy -f nvt.tpr -o temperature.xvg
xmgrace temperature.xvg
```

Exercises



02 – Small Peptide

- Pressure Coupling

  Pressure baths:

  Berendsen (again), Parinello-Rahman, MTTK, Surface tension …

  Either scales the volume of the box, or consider the box as a particle to “follow” the motion of the system.

Exercises

02 – Small Peptide

- Equilibration NPT

  N: Constant number of particle

  P: Constant Pressure

  T: Constant temperature

  Ideally closet ensemble to laboratory conditions

  Has both temperature and pressure coupling applied with positions restrain.

Exercises

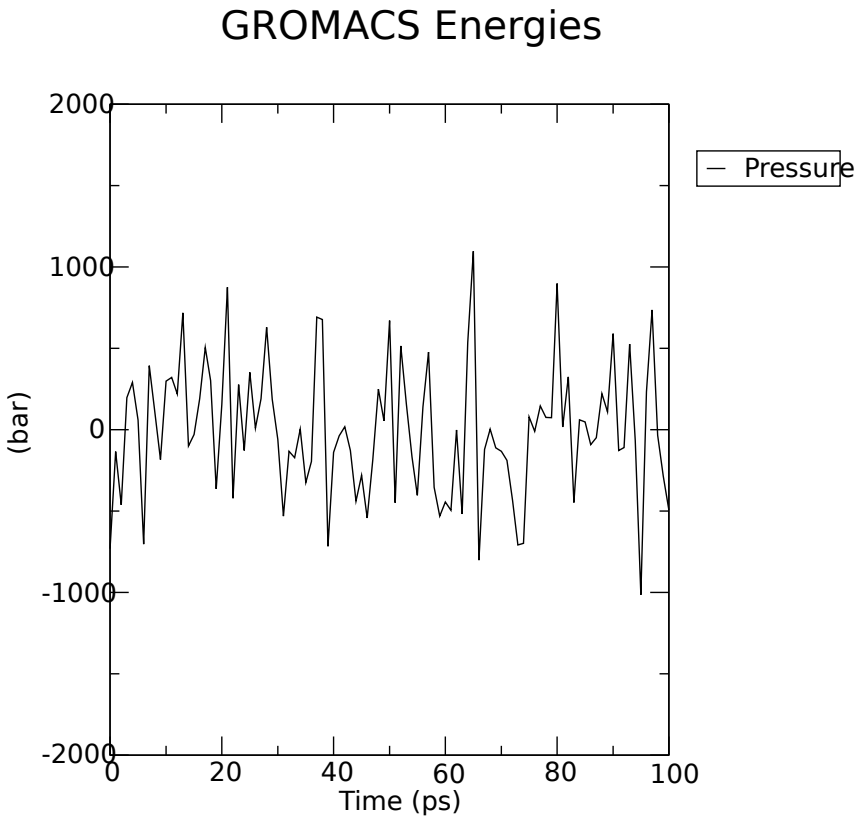
```
gmx grompp -f mdps/npt.mdp -c nvt.gro -r npt.gro …
-p topol.top -o nvt.tpr -t nvt.cpt
gmx mdrun -deffnm nvt -v
```

02 – Small Peptide

- Equilibration NVT

```
gmx energy -f npt.tpr -o pressure.xvg
xmgrace pressure.xvg
```

Exercises



02 – Small Peptide

- Production in NPT

  Same thing as in NPT equilibration but without restrain.

Exercises

```
gmx grompp -f mdps/prod.mdp -c npt.gro -r npt.gro …
-p topol.top -o prod.tpr -t npt.cpt
gmx mdrun -deffnm prod -v
```

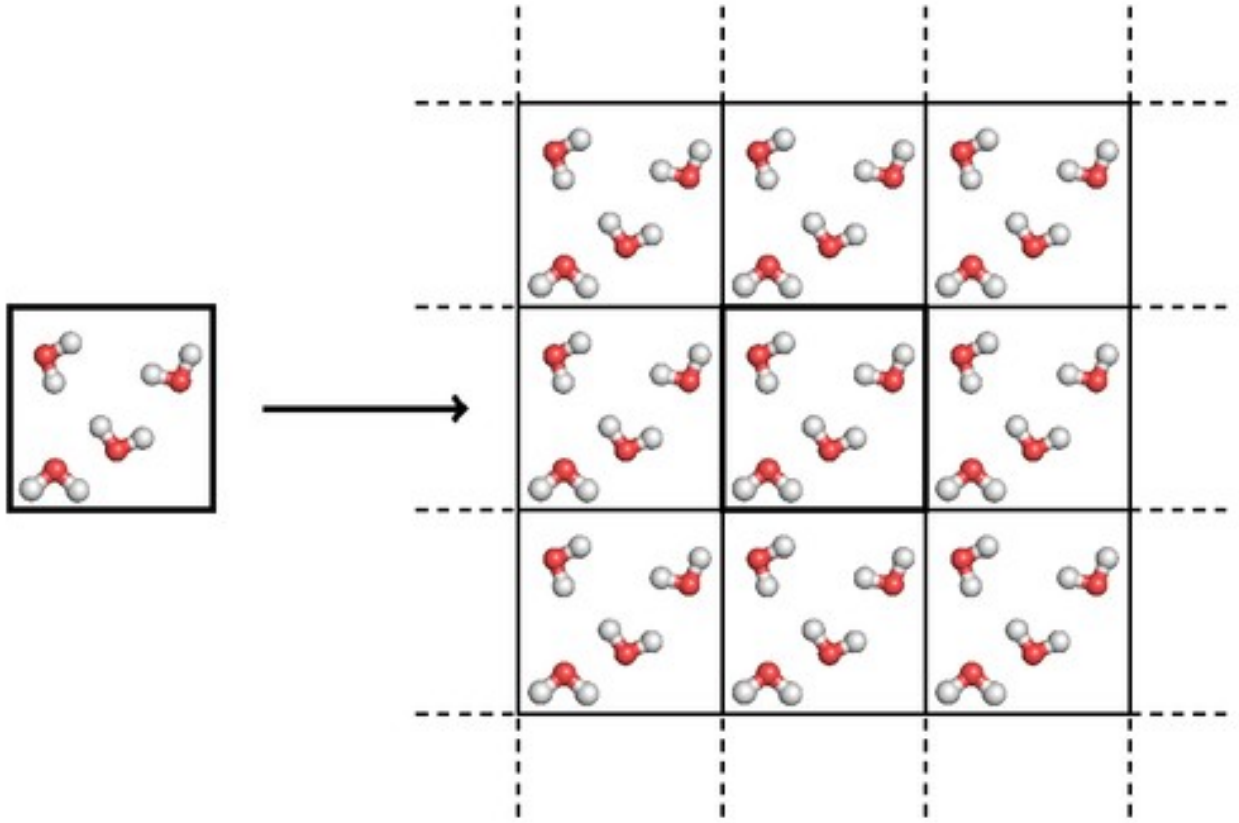
Depending on your computer your may have time for a coffee … or two …

How many ns can you run a day ?

02 – Small Peptide

- Correct for Periodic Boundaries

Exercises



http://www.texample.net/media/tikz/examples/PNG/periodic-boundaries-conditions.png

02 – Small Peptide

- Root Mean Square Deviation

  The first tool to check if you simulation did exploded (after the crash).

  Select the C-Alpha atoms

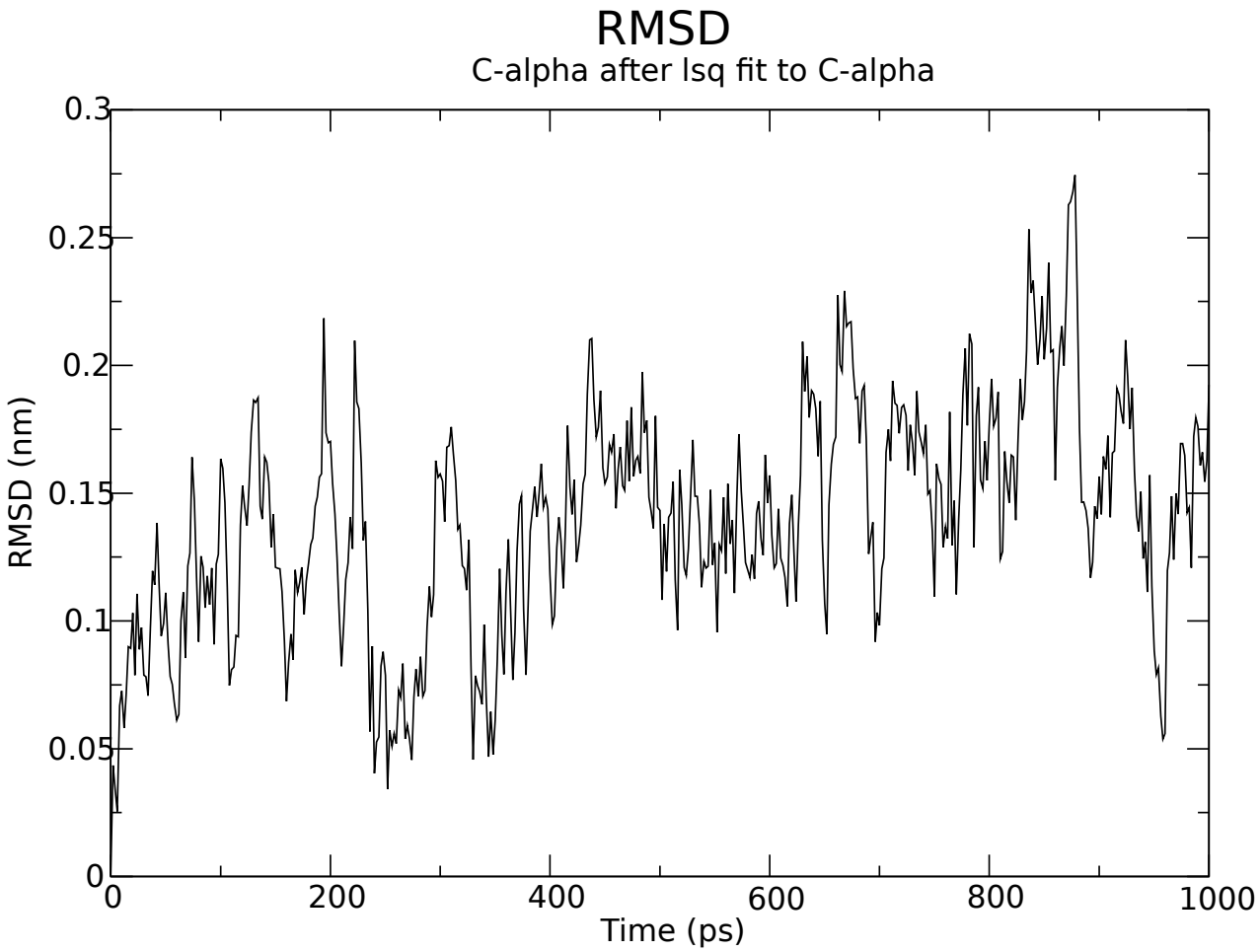
Exercises

```
gmx rms -f prod.xtc -prod.tpr -o rmsd.xvg
xmgrace rmsd.xvg
```

What is RMSD at the end of the production?

02 – Small Peptide

- Root Mean Square Deviation



Exercises

02 – Small Peptide

- Root Mean Square Fluctuation

  Proportional to Beta factors from X-Ray crystallography

  Usefull to estimate the degree of flexibility of the different region of the protein.

  Select the Backbone atoms

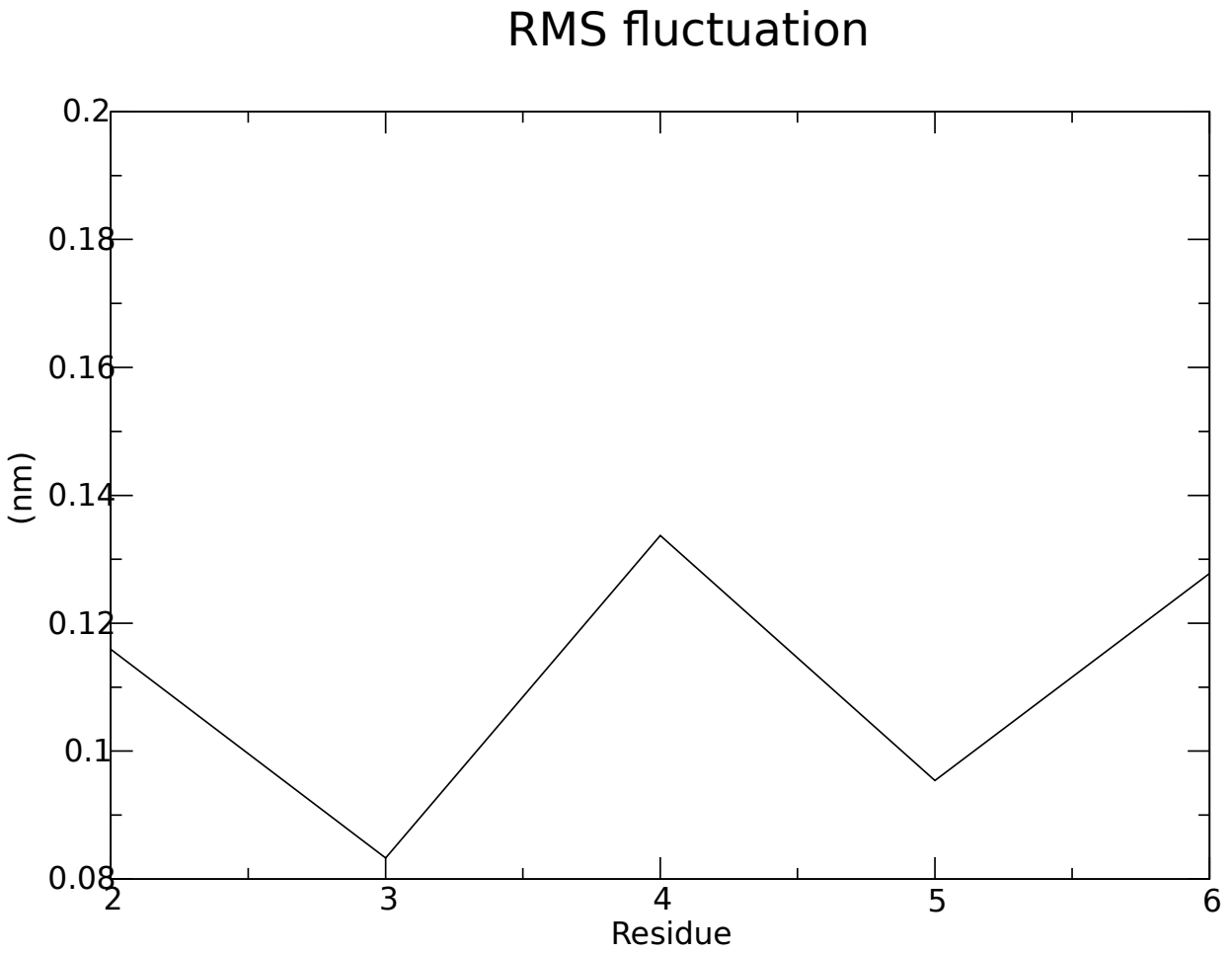
Exercises

```
gmx rmsf -f prod.xtc -prod.tpr -o rmsf.xvg -res
xmgrace rmsf.xvg
```

02 – Small Peptide

- Root Mean Square Fluctuation

Exercises



02 – Small Peptide

- Analysis using VMD

Exercises

```
vmd nojump.gro nojump.xtc
```

03 – Protein G

Exercices

- Coarse Graining ?

  Study membrane and protein structure over longer timescale

- Get the structure from RCSB

  Protein G B1 Domain PDB id: 3MP9

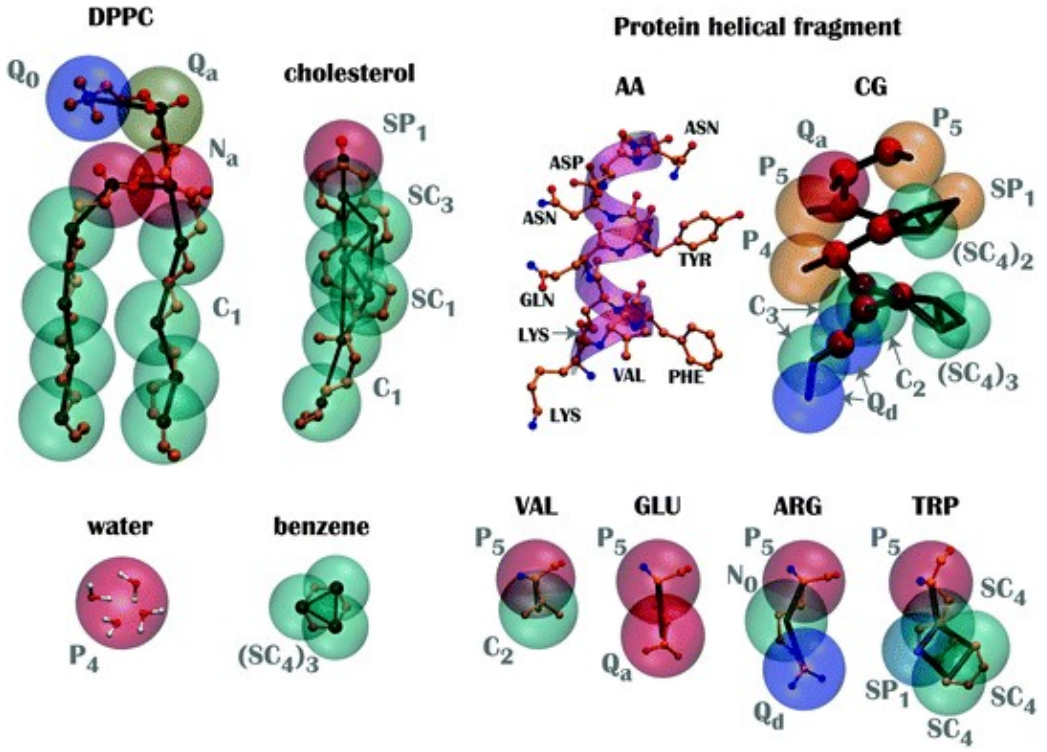
- Preparation

```
gerard@plop:~$ mkdir 03_protein_g
gerard@plop:~$ cd 03_protein_g
gerard@plop:~$ mv ~/Downloads/3mp9.pdb .
gerard@plop:~$ pymol 3mp9.pdb
```

03 – Protein G

- Preparation MARTINI

http://cgmartini.nl/



Exercices

Open the pbd file in PyMOL. Are there any ligand? Are they relevant?

03 – Protein G

- Preparation MARTINI

http://cgmartini.nl/

Exercices

Open the pbd file in PyMOL. Are there any ligand? Are they relevant?

No, we need to remove FMT and the waters

```
sed '/HOH/d;/FMT/d' 3mp9.pdb > 3mp9_ready.pdb
pymol 3mp9_ready.pdb
*remove the chain B and save the pdb*
```

03 – Protein G

- The script

  Open the script do-cg.sh using your favorite text editor

Exercices

03 – Protein G

- Running the Script

```
bash do_cg.sh 3mp9_ready martini
```

And probably go for coffee

Exercices

How many ns can you run a day ?

03 – Protein G

- Correct for Periodic boundaries

  But keep the whole system in the output

  Select system

Exercices

```
gmx trjconv -f prod.xtc -s prod.tpr …
-o viz.xtc -pbc nojump
gmx trjconv -f viz.xtc -s prod.tpr …
-o viz.gro -dump 0
```

03 – Protein G

- RMSD

  Groups of atoms are not properly recognzed by GROMACS

  Select type BB

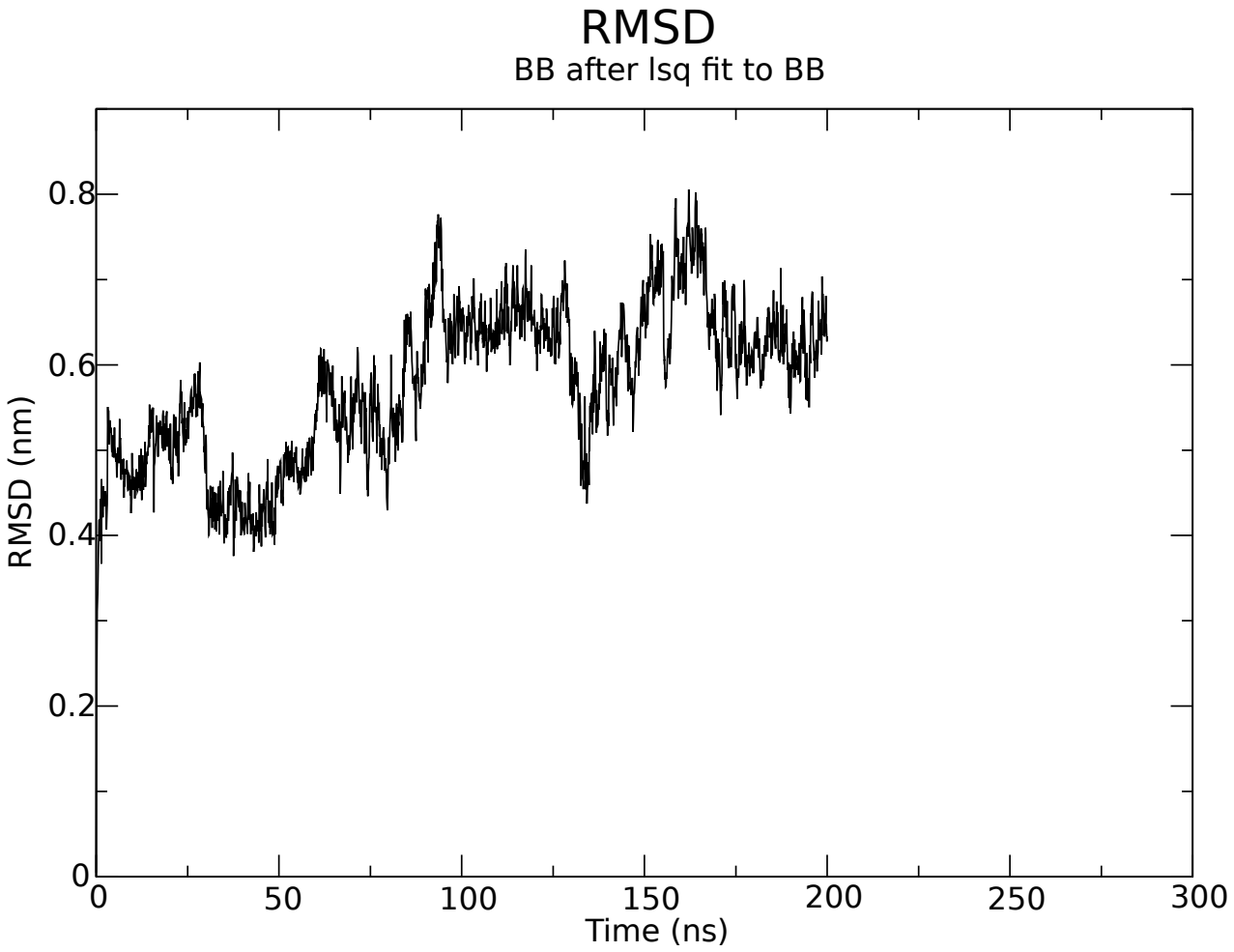
Exercices

```
gmx make_ndx -f prod.tpr -n index.ndx
gmx rms -f viz.xtc -prod.tpr -o rmsd.xvg …
-n index.ndx -tu ns
xmgrace rmsd.xvg
```

03 – Protein G

- RMSD

Exercices

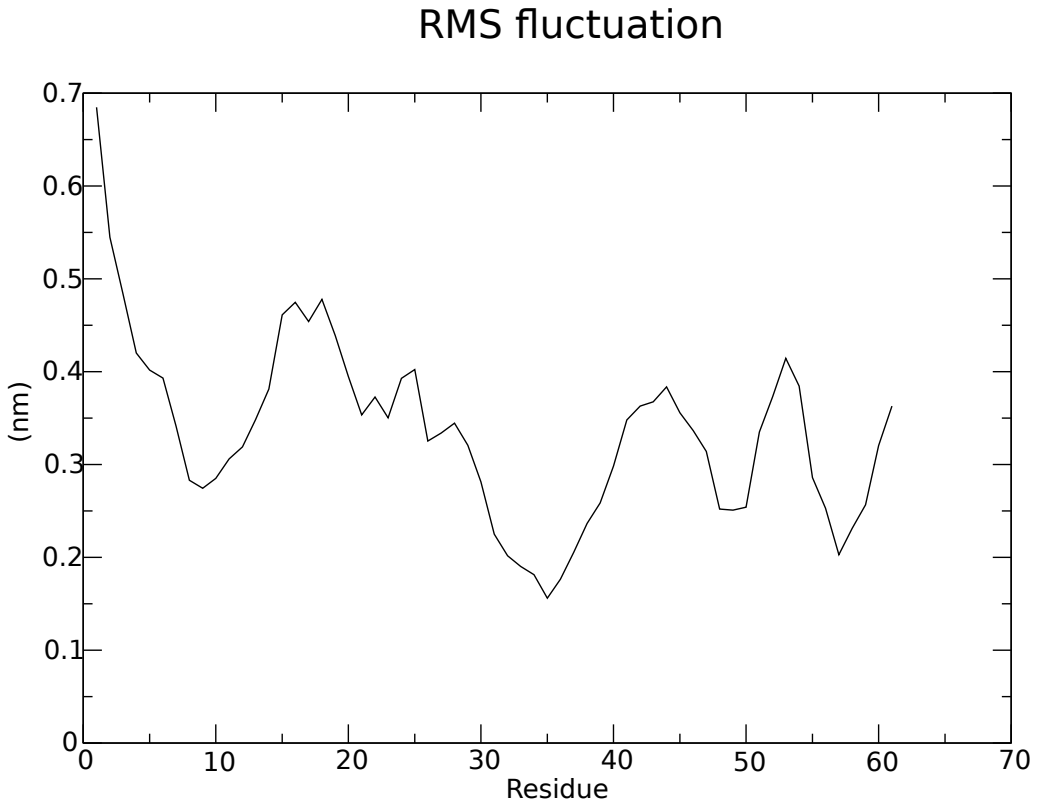


03 – Protein G

- RMSF

```
gmx rmsf -f viz.xtc -prod.tpr -o rmsf.xvg …
-n index.ndx -res
xmgrace rmsf.xvg
```

Exercices



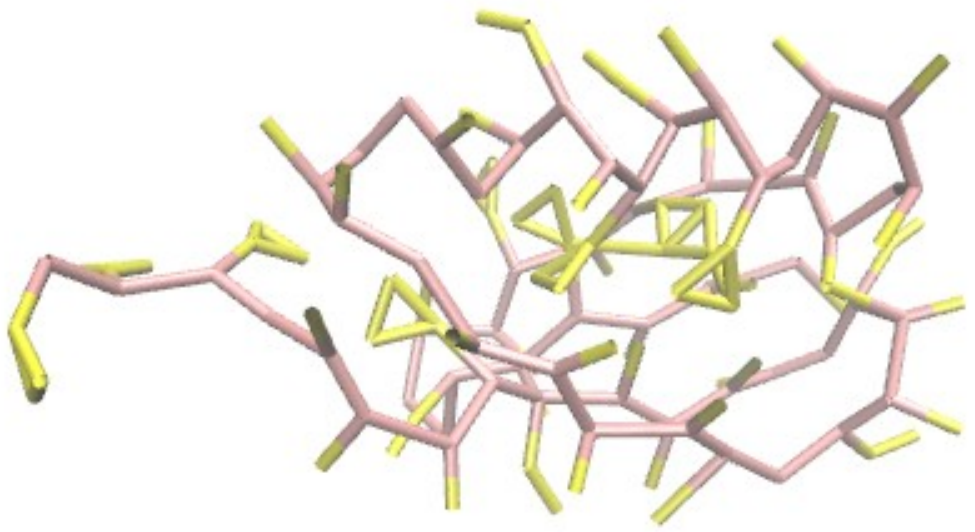
03 – Protein G

- Visualization under VMD

  Using dynamic bonds

```
vmd viz.gro viz.xtc
```

Exercices



Conclusion

- Computational methods are well established
- Yet they are not oracles
- I hope you had a (not traumatizing) introduction to MD

Conclusion (bis)

- I have demonstrated the superiority of computing work over experimental work.
- You must now reconsider you choice of career
- All hail to the Biocomputing group