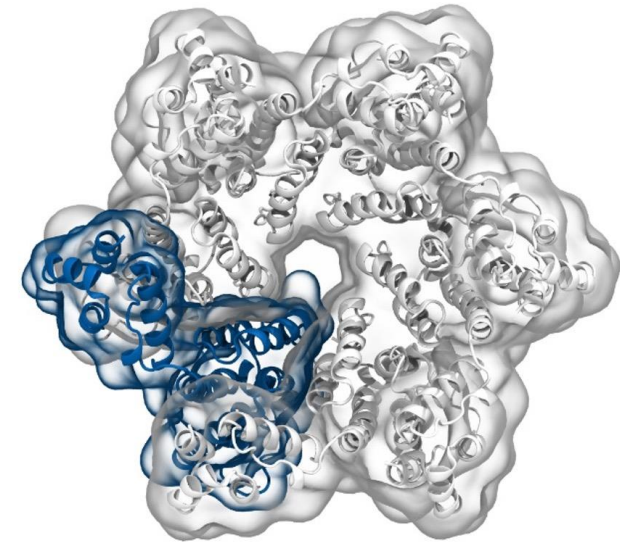
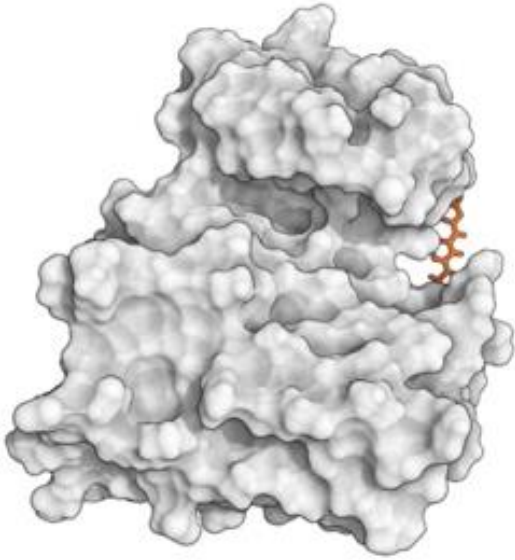


Simulation of Biomolecules

Basic Simulation Analysis



Dr Matteo Degiacomi
Durham University

matteo.t.degiacomini@durham.ac.uk

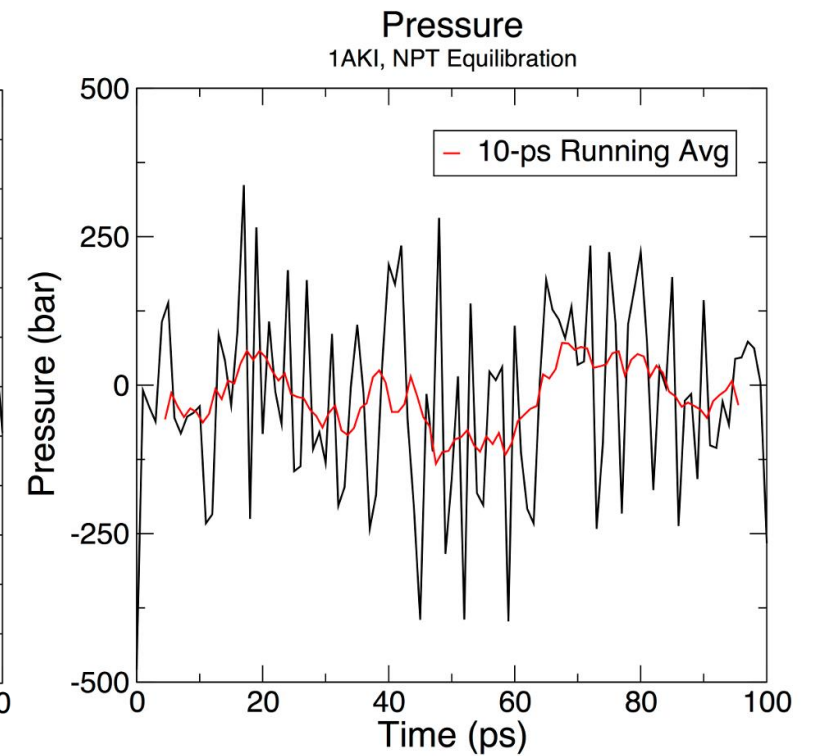
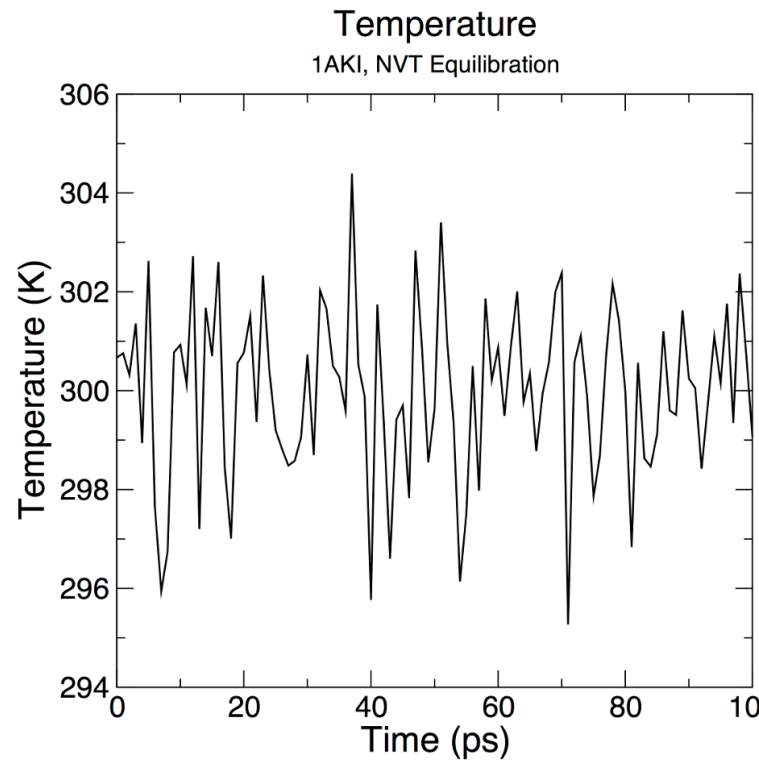
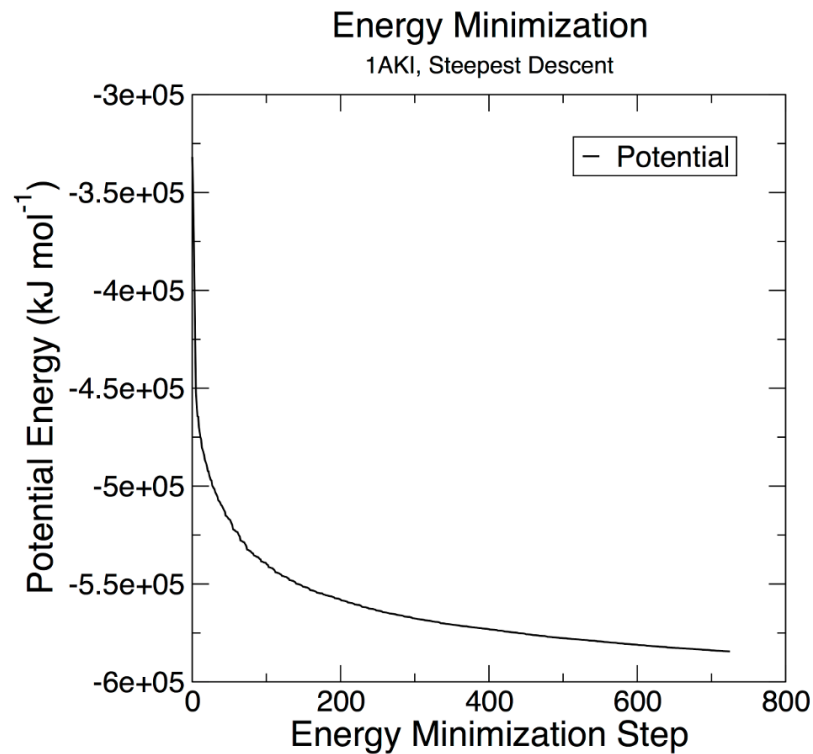
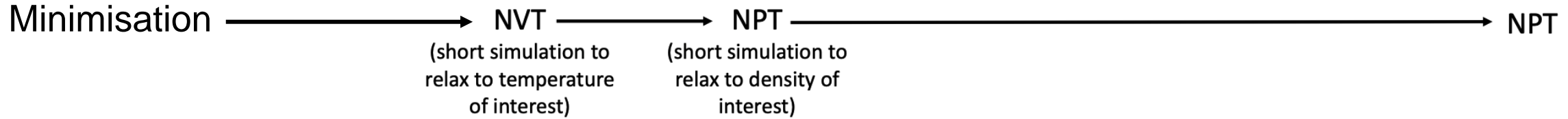
Dr Antonia Mey
University of Edinburgh

antonia.mey@ed.ac.uk

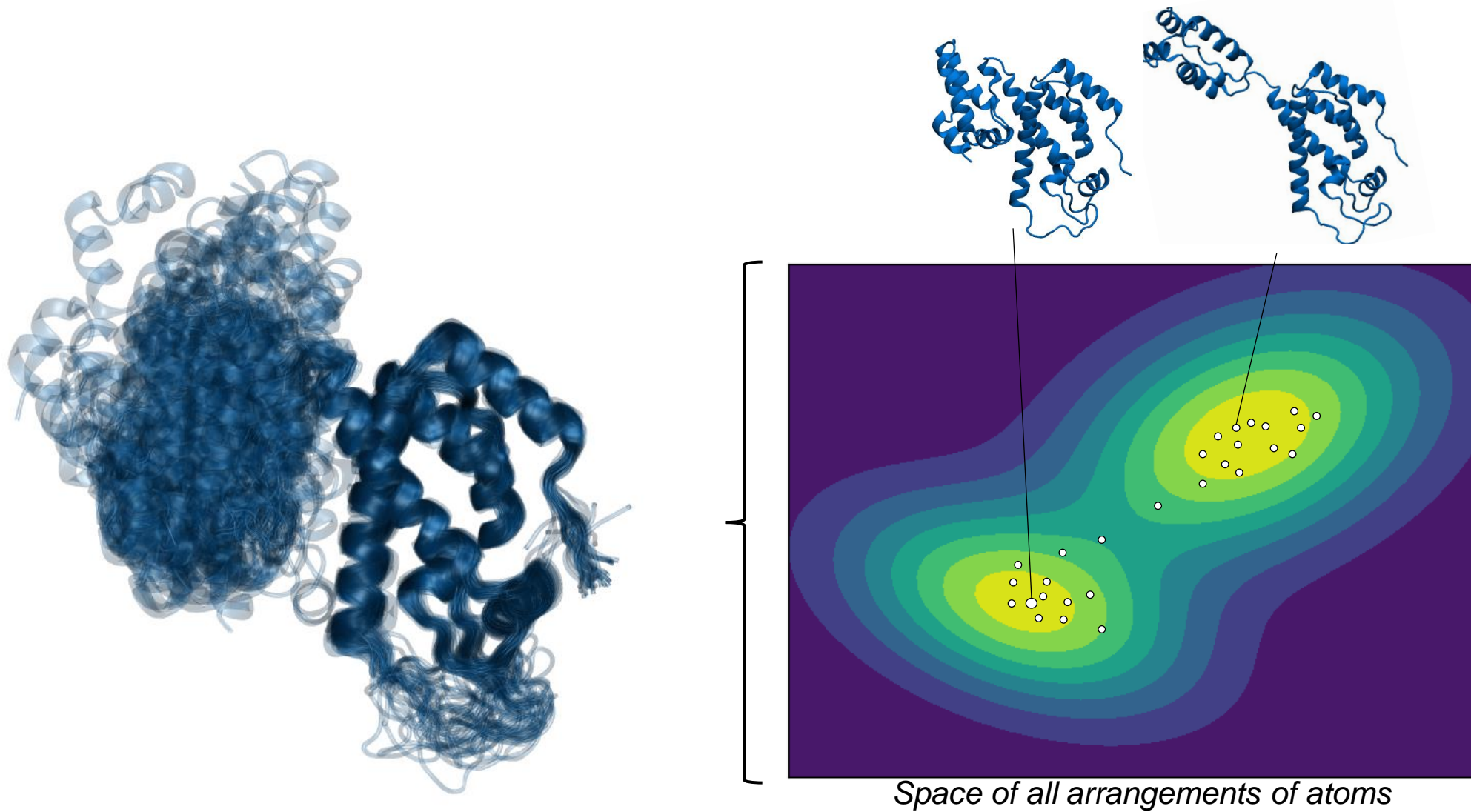
Volume and pressure equilibration

Steps until production:

Production ensemble



Sampling the conformational space

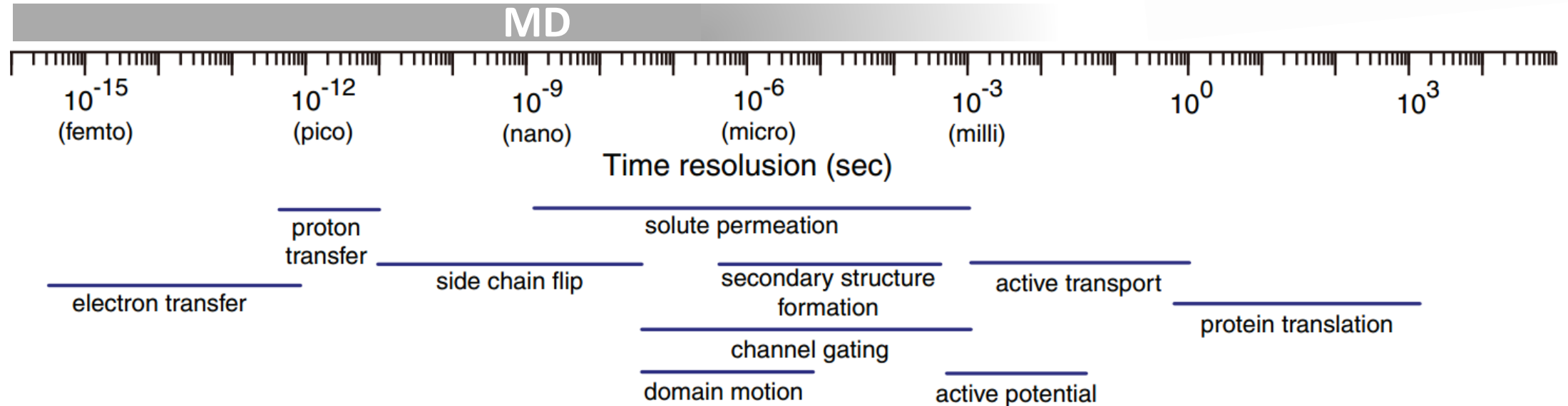
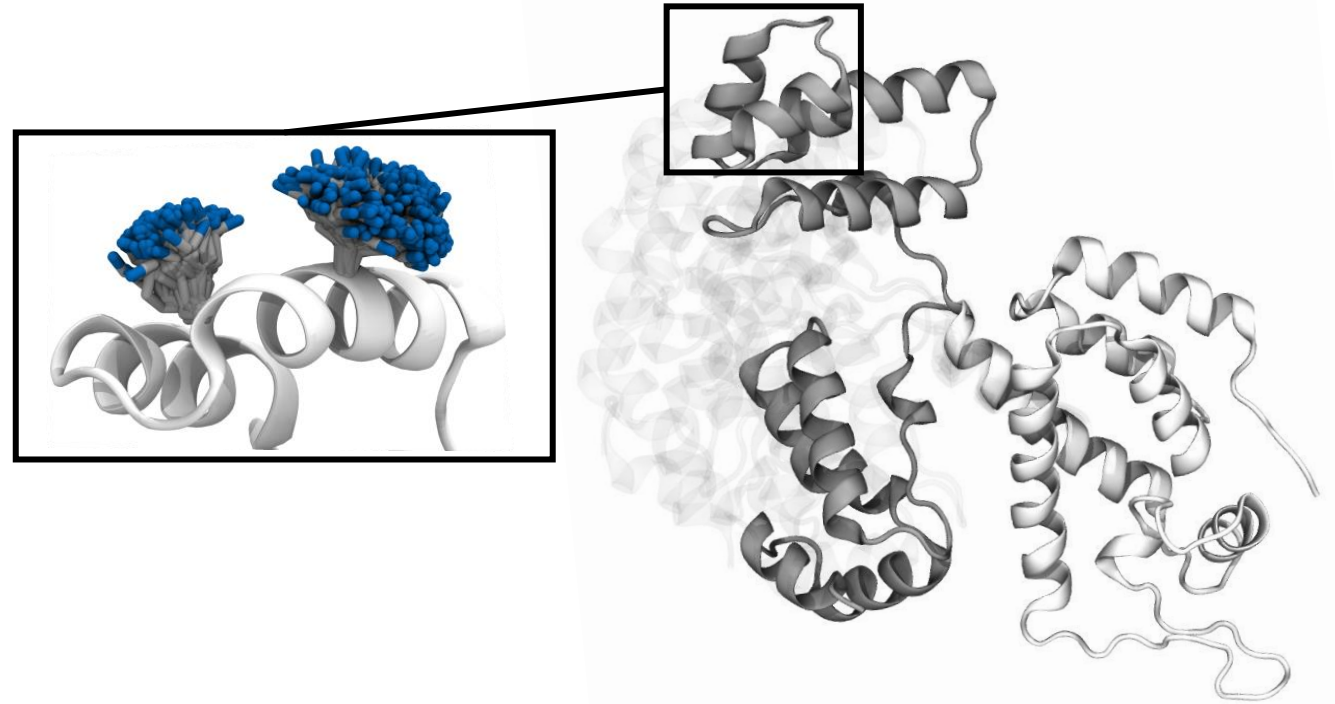


Probability of sampling a conformation is inversely proportional to its energy: $p_i \propto e^{-\epsilon_i/kT}$

Timescales in biology

Different regions, different timescales:

- Side chains faster than backbone
- Loops faster than helices and sheets
- backbone faster than side chains
- Protein surface faster than core

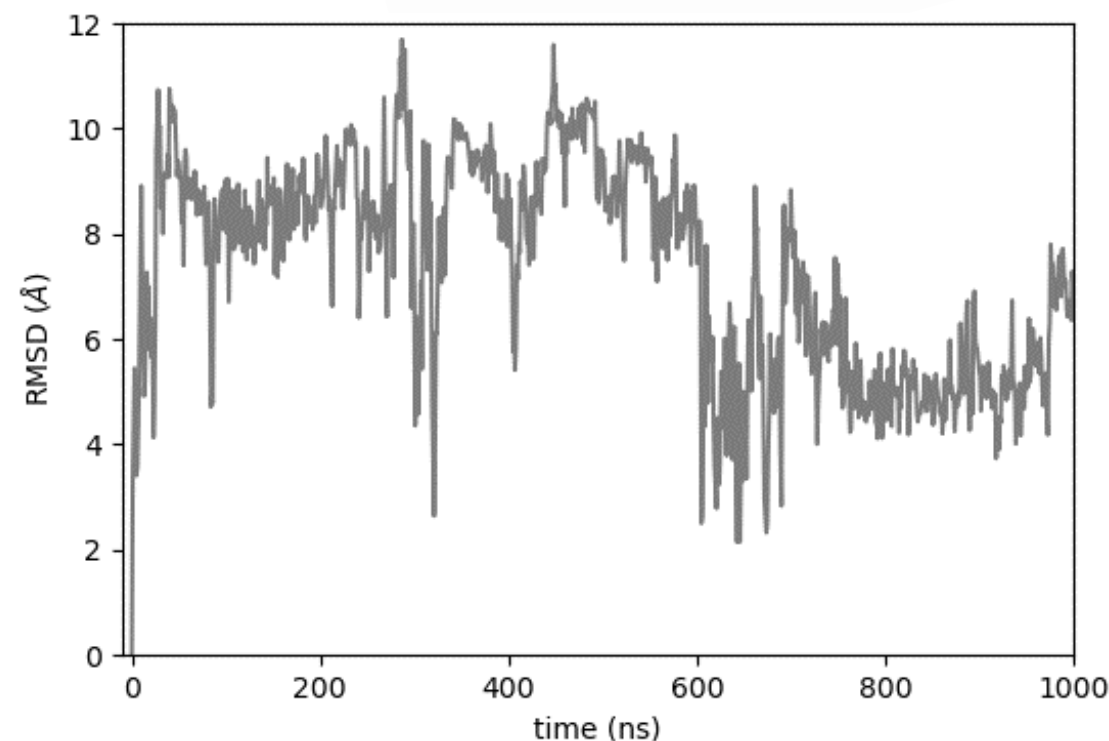
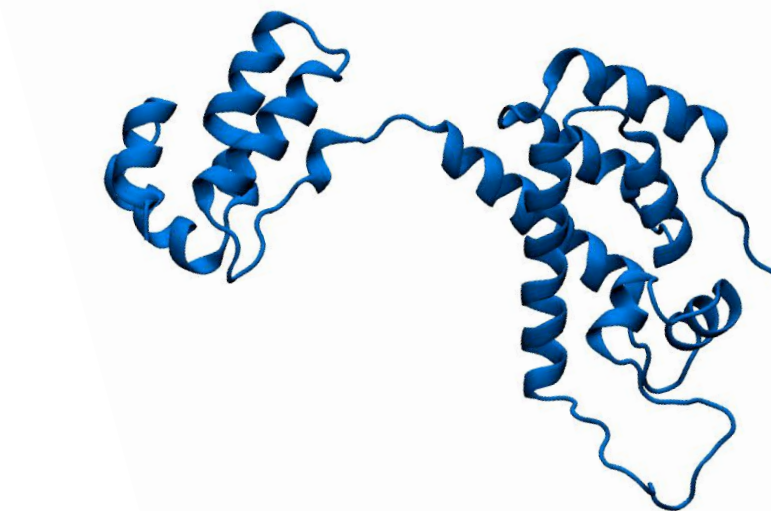


Root Mean Square Deviation (RMSD)

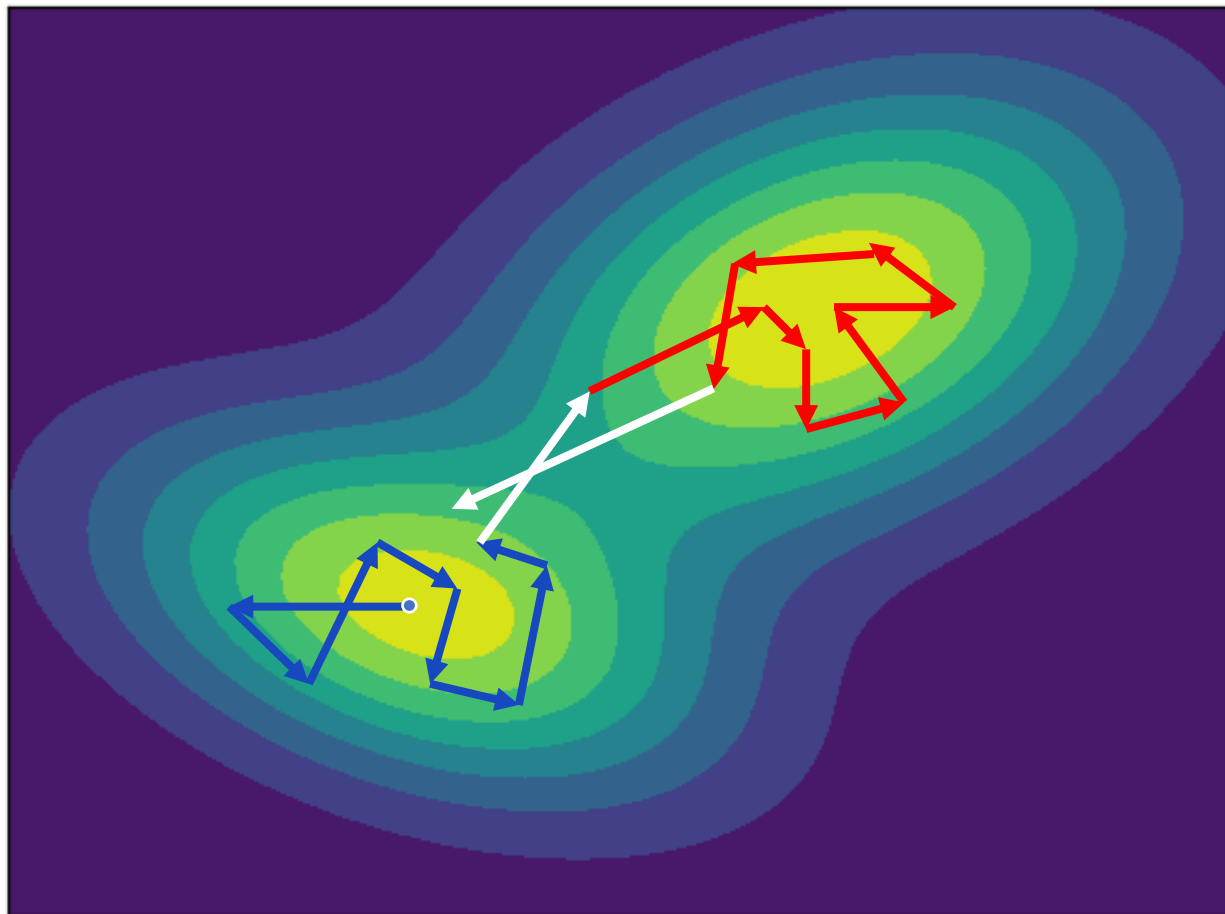
Given a system with N atoms, and a reference arrangement x_0 :

$$RMSD = \sqrt{\frac{1}{N} \sum_{i=0}^N (X_i - x_0)^2}$$

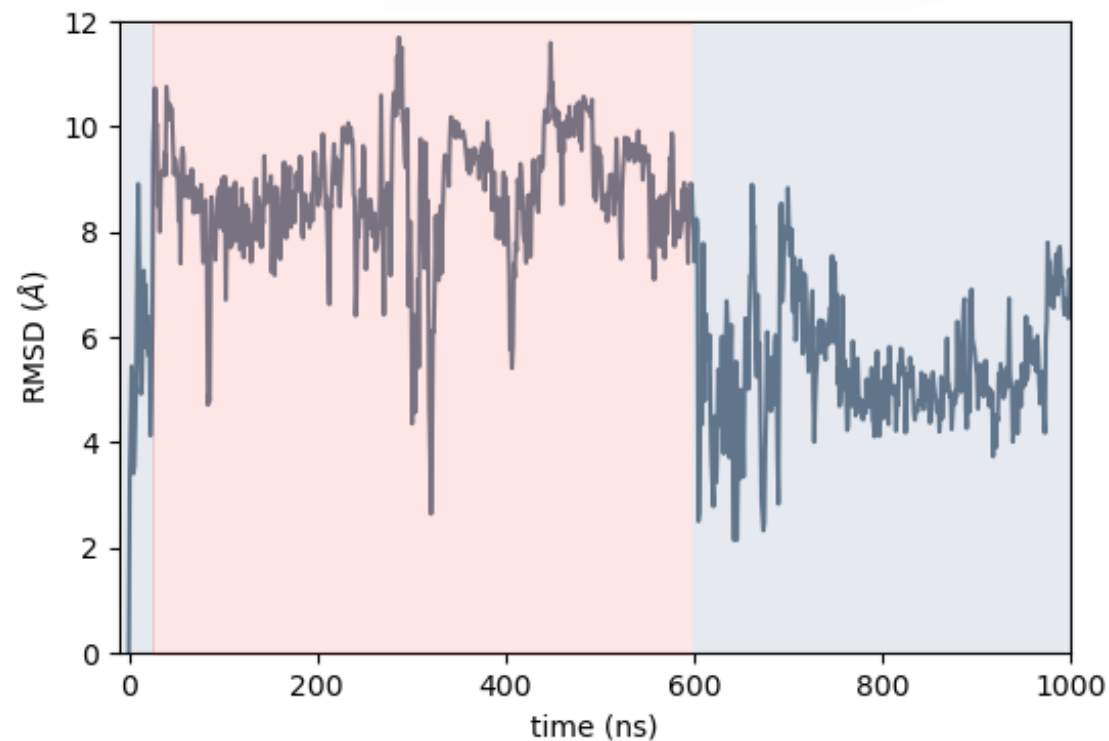
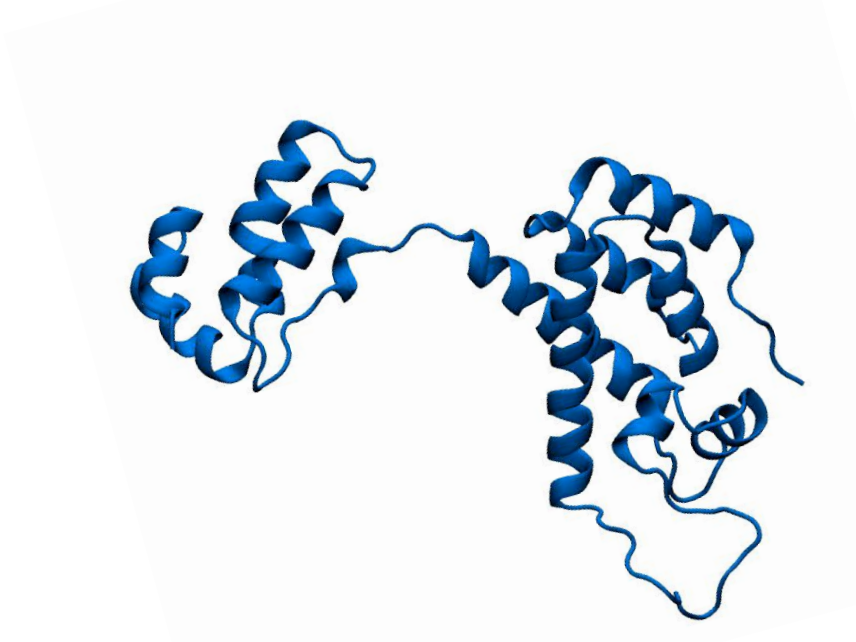
In MD, x_0 is often the first conformation in the simulation.



Convergence?

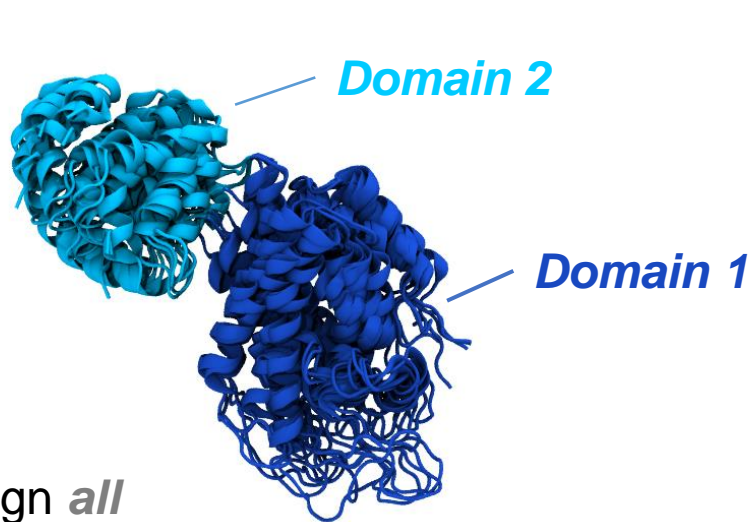


Refrain from using RMSD as a single indicator of simulation convergence.

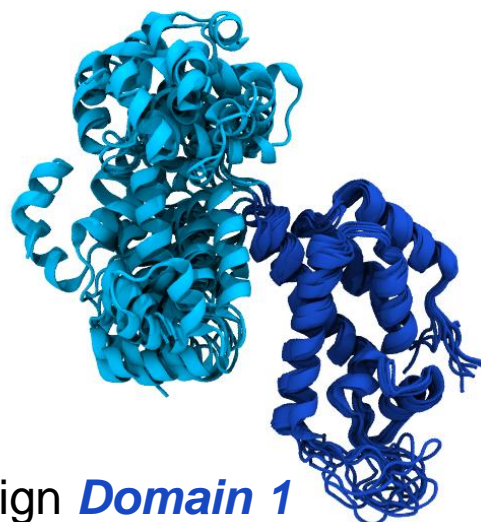
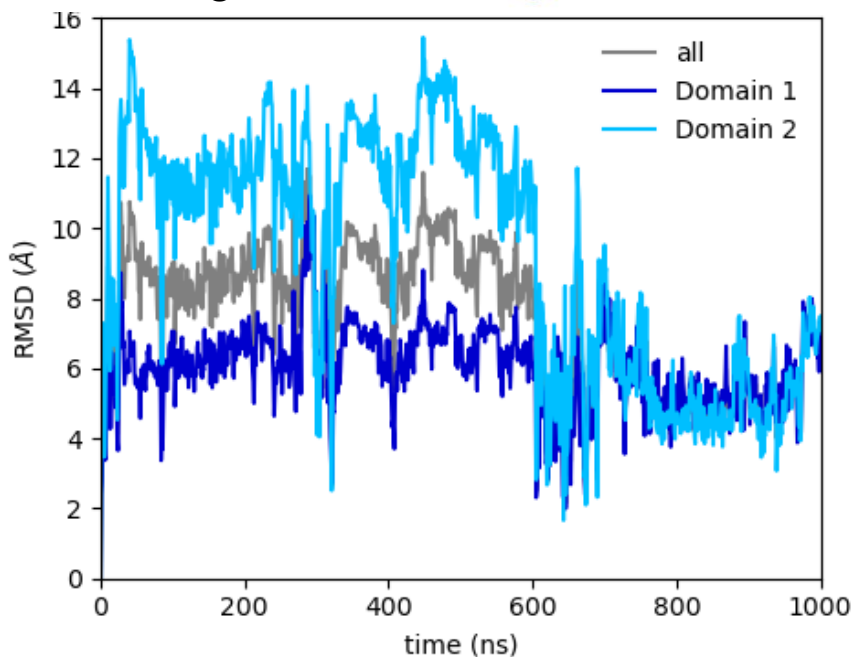


HIV-1 capsomer simulation from: *Degiacomi & Dal Peraro, Structure, 2013*

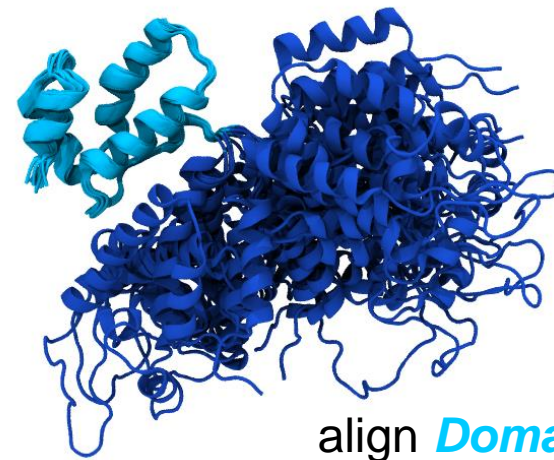
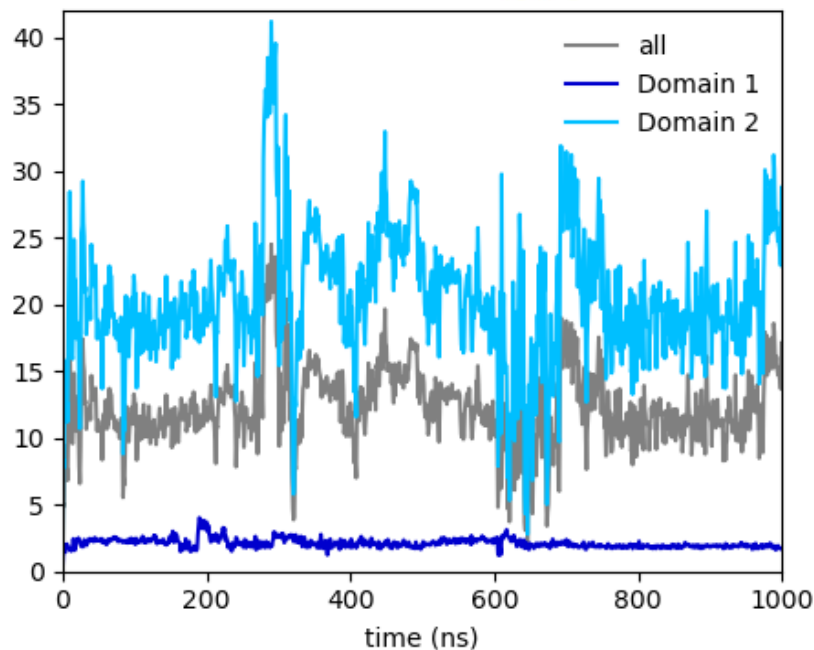
RMSE is alignment- and selection-dependent



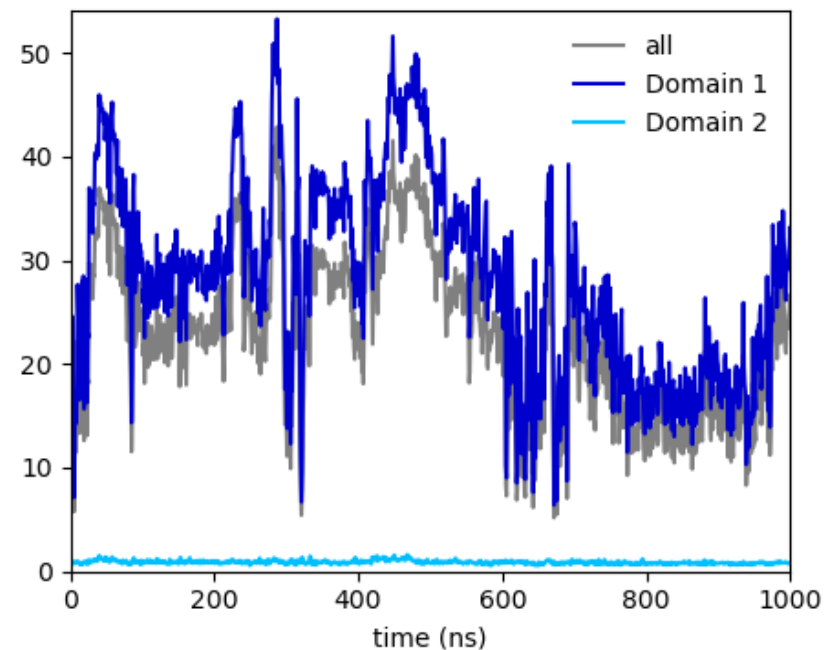
align *all*



align *Domain 1*

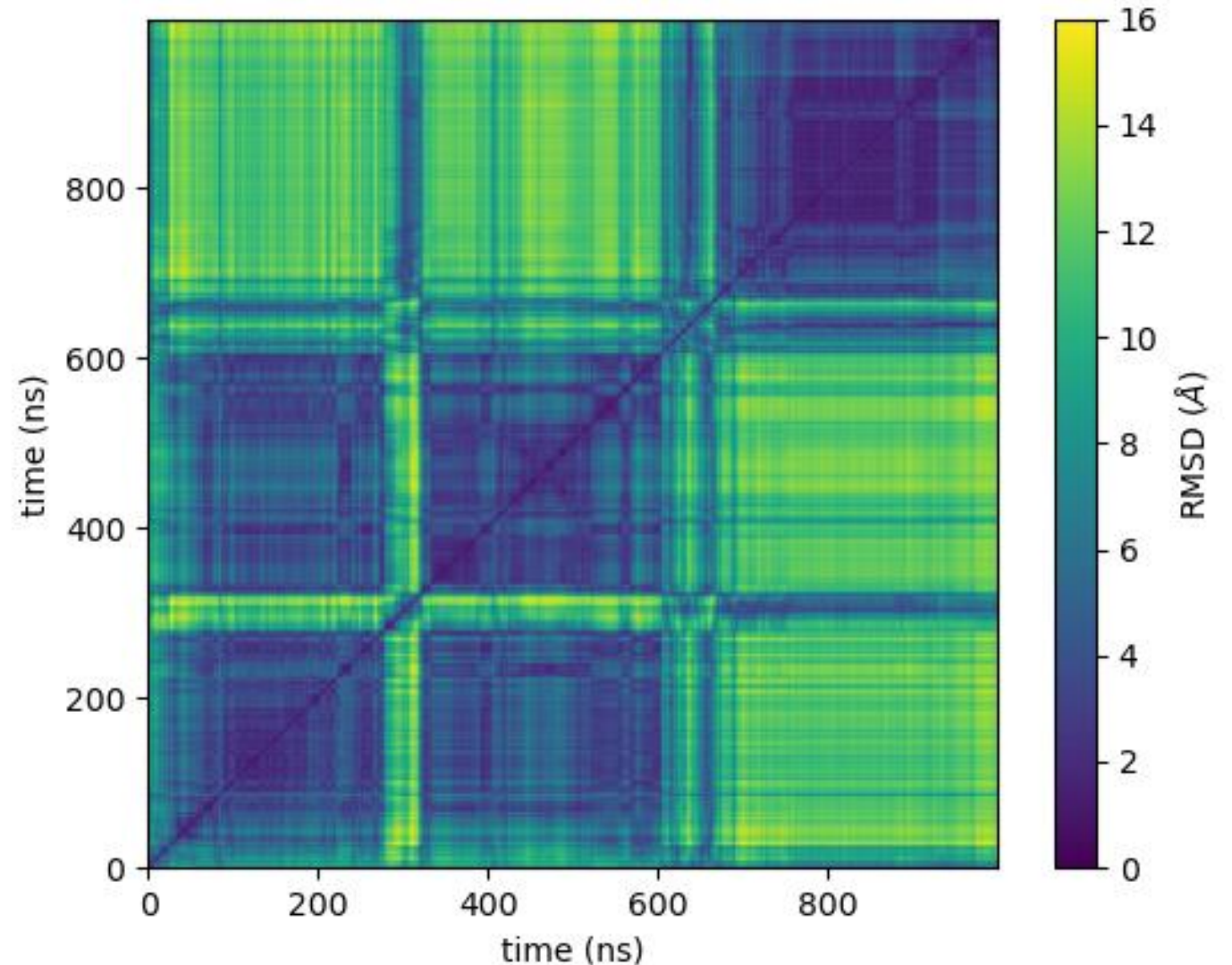


align *Domain 2*



Pairwise RMSD

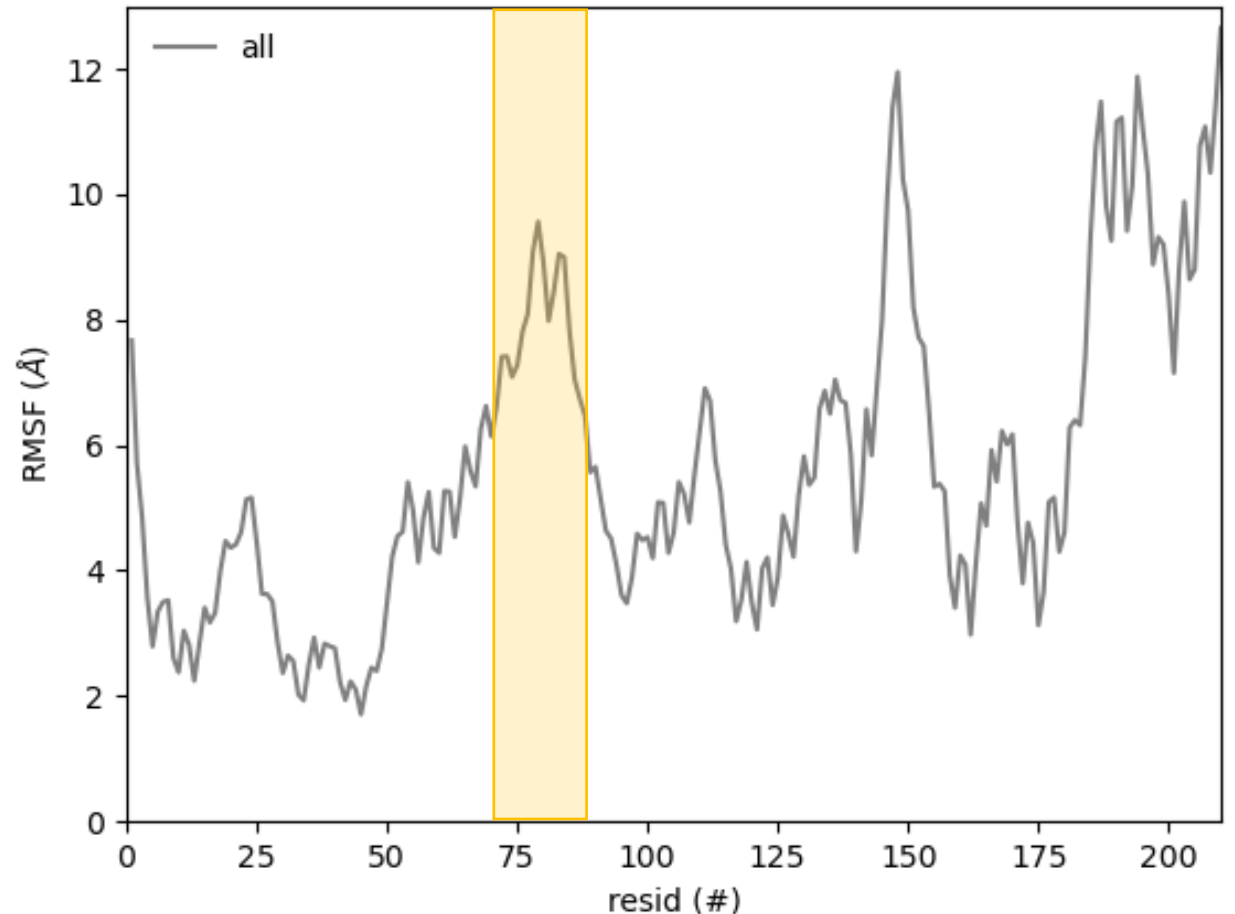
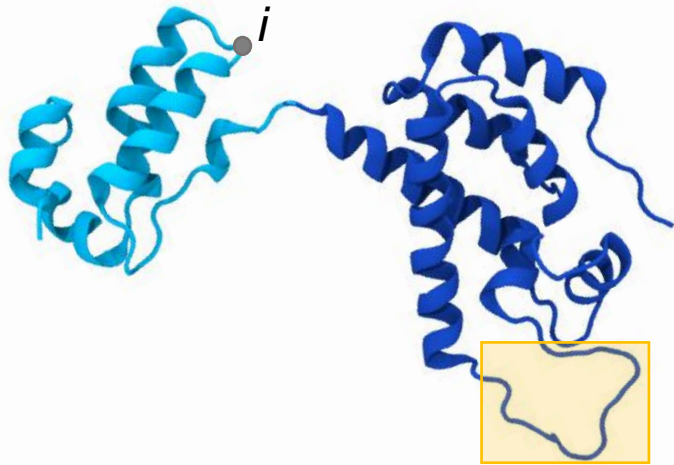
- Two structures with same RMSD from a reference are not forcefully similar to each other.
- Pairwise RMSD helps seeing if protein re-visits conformations throughout the simulation.



Root Mean Square Fluctuation (RMSF)

The RMSF σ_i of atom i calculates how much it fluctuates around its mean position $\langle X_i \rangle$.

$$\sigma_i = \sqrt{\langle (X_i - \langle X_i \rangle)^2 \rangle}$$

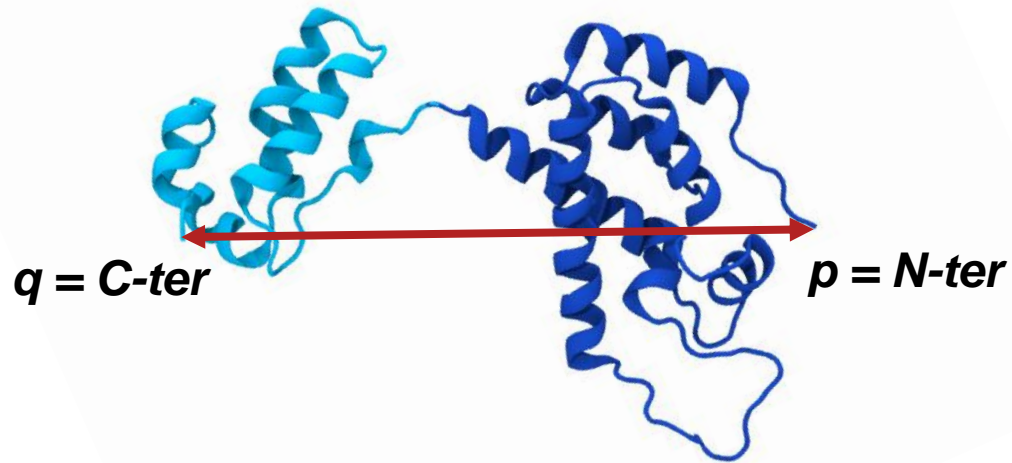


Typically done on C_α atoms.
Result depends on alignment!

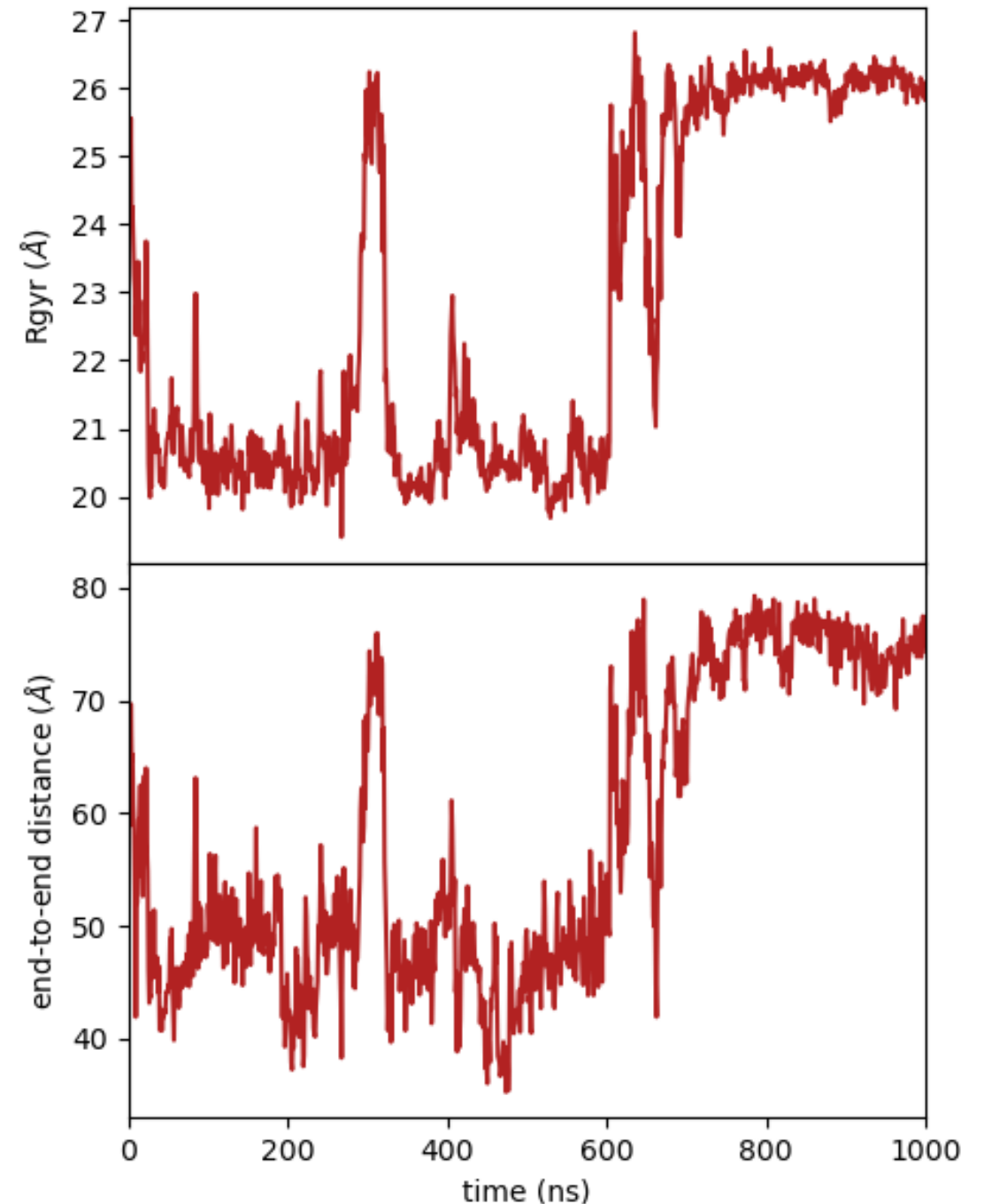
end-to-end distance and Radius of Gyration (Rg)

$$d(p, q) = \|p - q\|$$

$$R_g = \sqrt{\frac{1}{N} \sum |r_k - r_{mean}|^2}$$



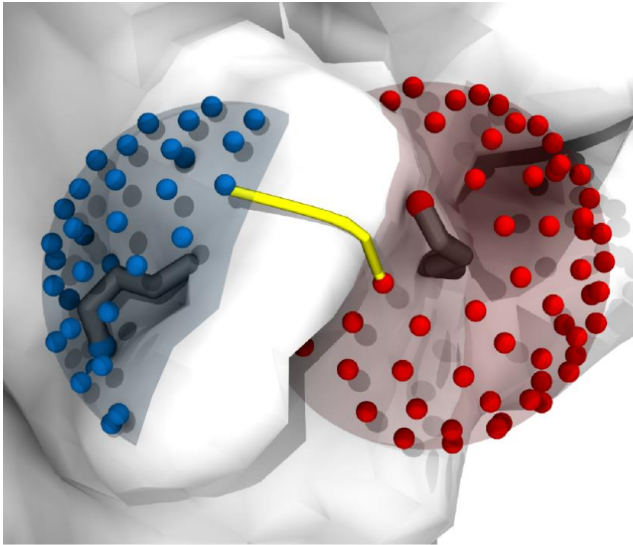
Internal properties do *not* depend on alignment



Some quantities are not *directly* measurable

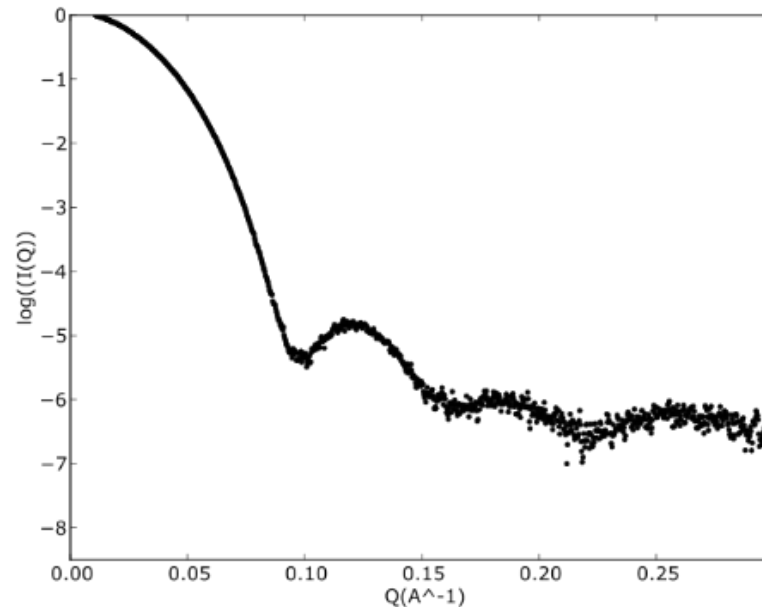
Submit MD conformers to external software simulating experimental data, e.g.:

Chemical cross-linking



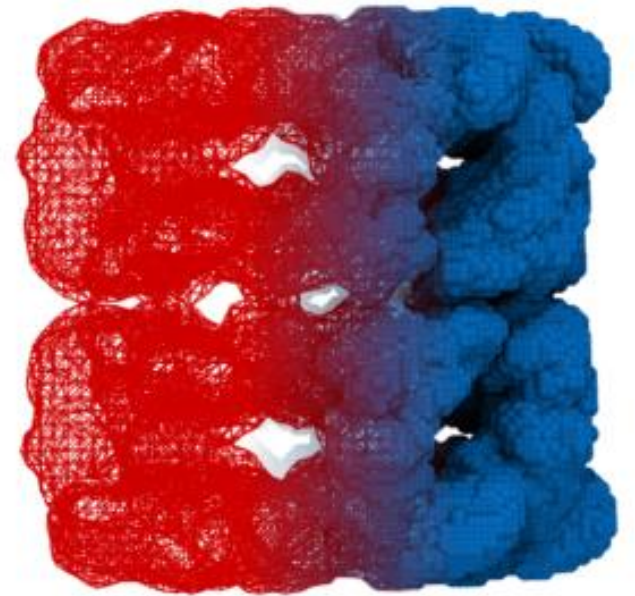
DynamXL, Xwalk, ...

SAXS



CRY SOL, FoXS, ...

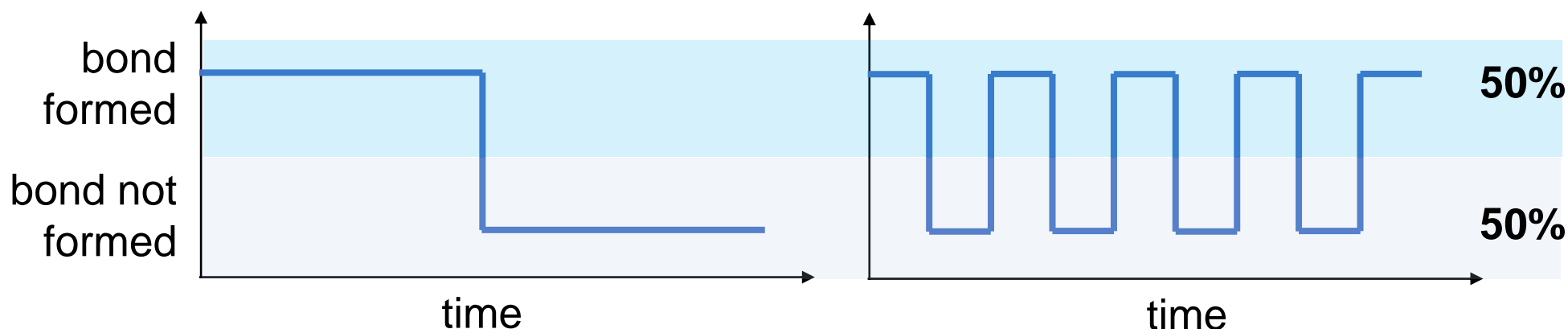
Collision cross-section



IMPACT, MobCal,
CollisionCode, ...

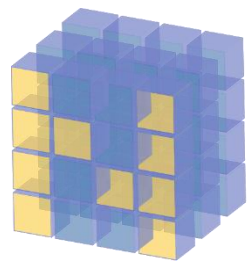
Warning: time averaging may hide processes

Thought experiment: typically hydrogen bond is considered established if donor-acceptor distance $< 2.5 \text{ \AA}$, and donor-acceptor-hydrogen angle $< 20^\circ$.



Reporting % time a bond is established in simulation can be misleading!

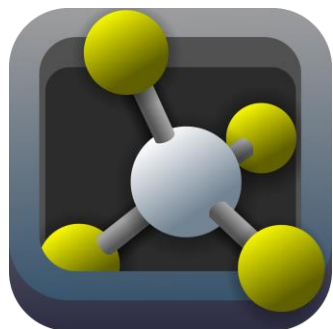
Large ecosystem of software and packages for data analysis



NumPy



SciPy



Open-Source Cheminformatics
and Machine Learning



VMD

Visual Molecular Dynamics



MDTRAJ

And many more...

MDAnalysis



The **Universe** contains everything about a MD system

- *Static information*: atoms and their connectivities
- *Dynamic information*: The trajectory

Data accessible via a hierarchy of containers

