THE UNIVERSITY of EDINBURGH

Classical Molecular Dynamics Simulations

Antonia Mey — antonia.mey@ed.ac.uk

Slides adapted from Dr Fernanda Duarte

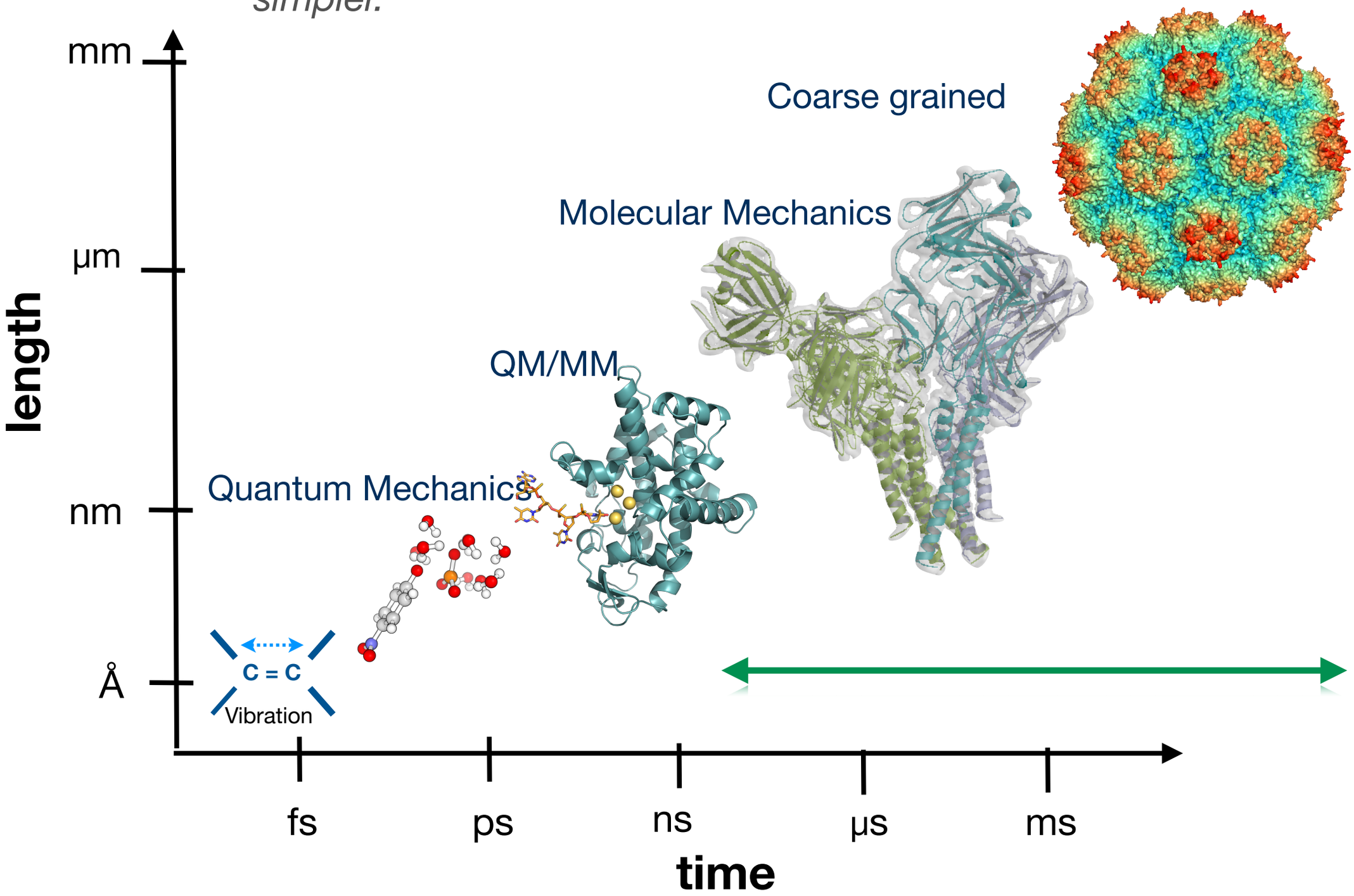
School of Chemistry

University of Edinburgh, UK

December 2018 – SUPA

Computational Modelling

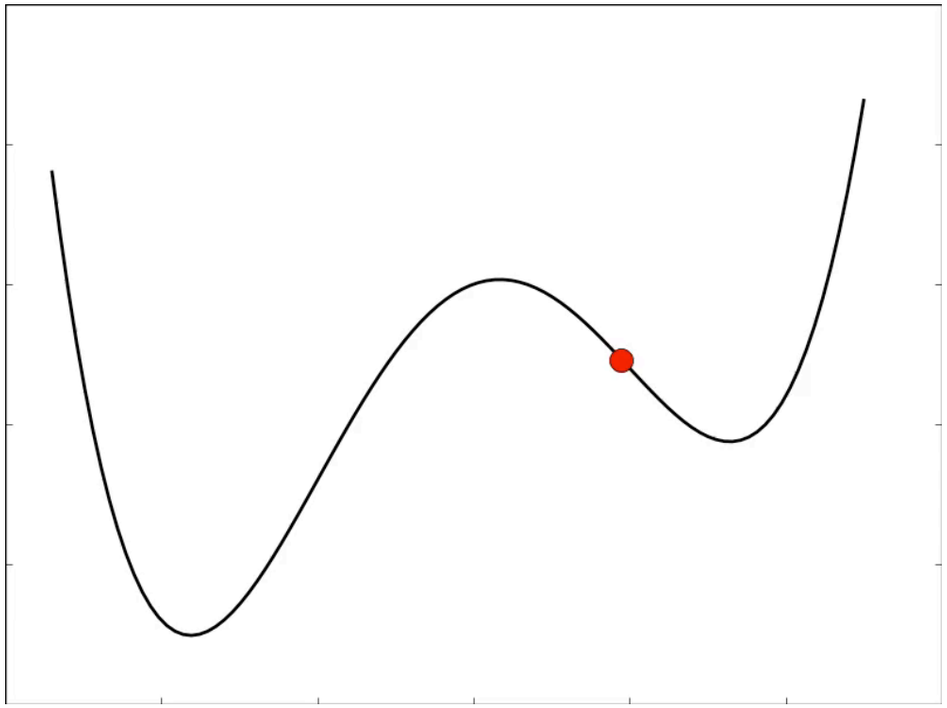
"Everything should be made as simple as possible, but not simpler."

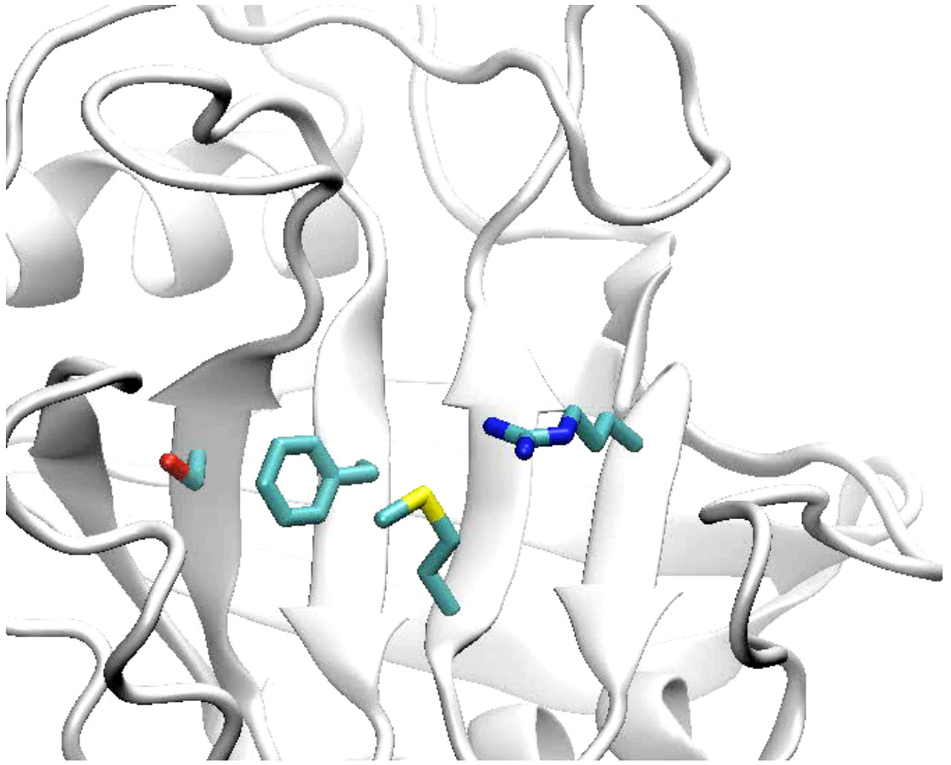


https://quoteinvestigator.com/2011/05/13/einstein-simple/

Overview

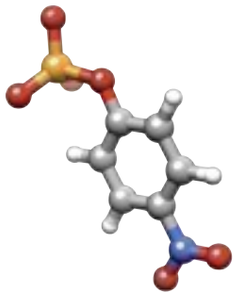
- What is - and why run - a Classical Simulation?
- Force fields/Atom Types/Solvent Models
- Molecular Dynamics–Framework and Steps Involved





Computational Modelling

"Everything should be made as simple as possible, but not simpler."



Quantum Mechanics

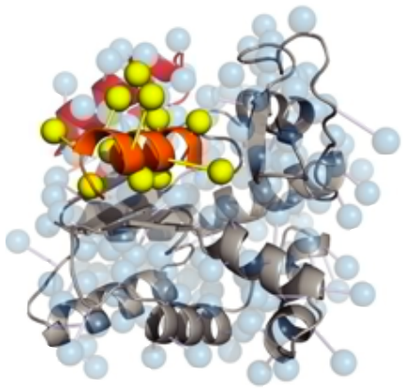
Atomistic

Electronic Structure

(Schrödinger equation)

Accurate ⇒Expensive

(DFT N^{3}- CCSD(T) N^{7})



Molecular Dynamics

Monte Carlo

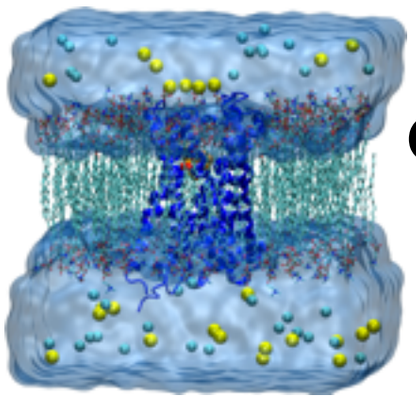
Atomistic

Empirical Forces

(Newton's equations/ Random move)

LESS accurate ⇒Faster

(N *log*N/N^{2})



Coarse grained

Not atomistic

Reduction of number of degrees of freedom

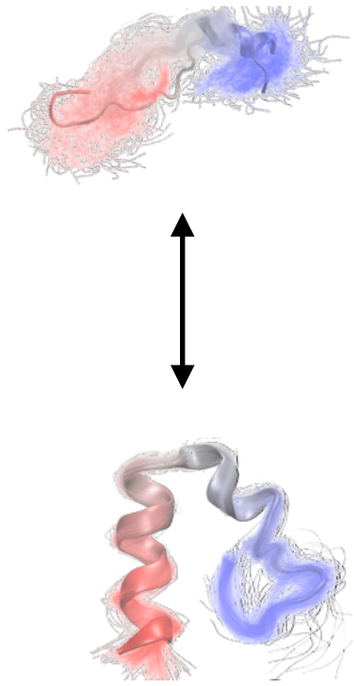
Reaction and diffusion

Can be even faster

Computational Modelling

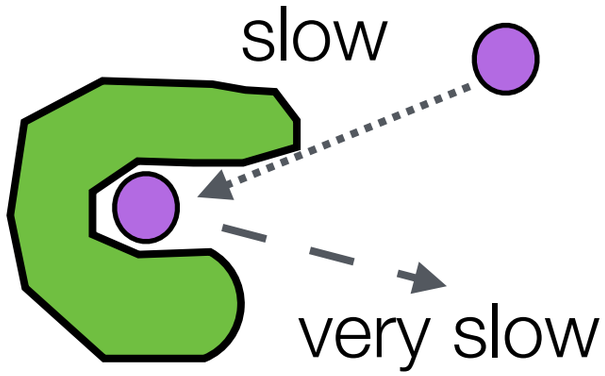
What is the scientific question we want to address?
How much detail do we need to describe the phenomenon?
What are the computational resources?

Large or small conformational dynamics



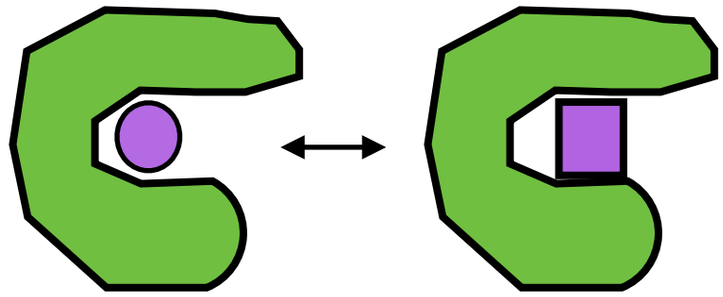
Standard MD

Binding thermodynamics/ dynamics



Enhanced sampling

Reaction dynamics



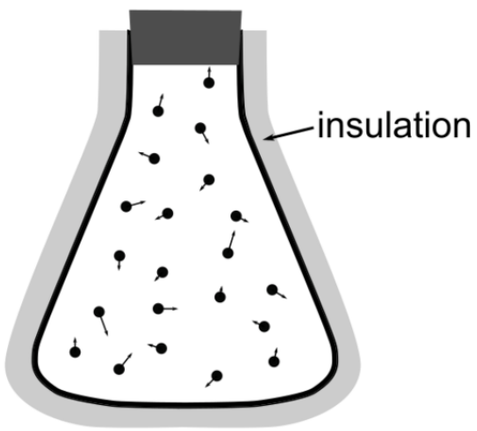
QM/MM

Nature Commun. 2011, 2, 41

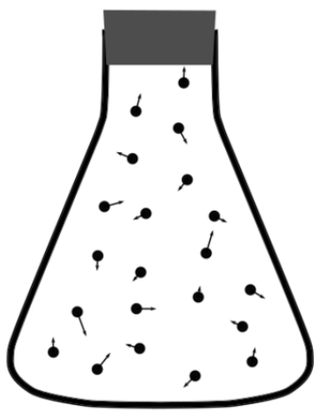
Computational Modelling

Experiments measure **ensemble averages**, of macroscopic properties

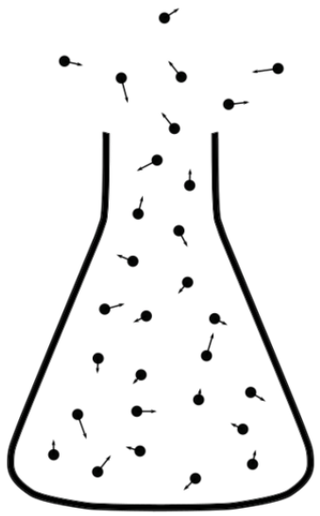
The knowledge of a **single** structure, even of a global energy minimum is **NOT sufficient.**



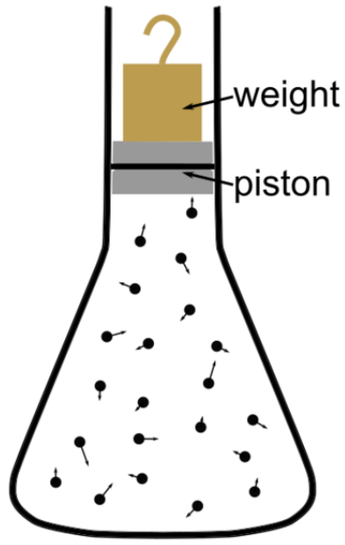
Microcanonical
(NVE)



Canonical
(NVT)



Grand canonical
(μVT)



Isobaric-Isothermal
(NPT)

https://commons.wikimedia.org/wiki/File%3AStatistical_Ensembles.png

Computational Modelling

Statistical mechanics to the rescue

Macroscopic Observables

Equilibrium Properties

Binding Constant of an Inhibitor

Potential Energy of a System

Structure of a Liquid

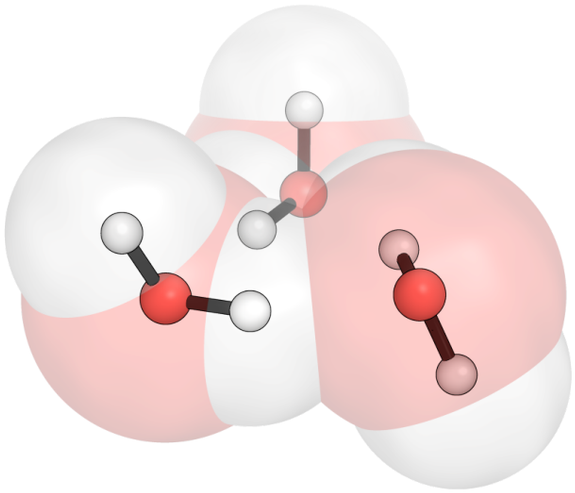
Non-Equilibrium Properties

Viscosity of a Liquid

Diffusion Processes in Membranes

Reaction Kinetics

Computational Modelling



Statistical Mechanics

Microscopic Information on Atomic Positions and Velocities

Force Fields

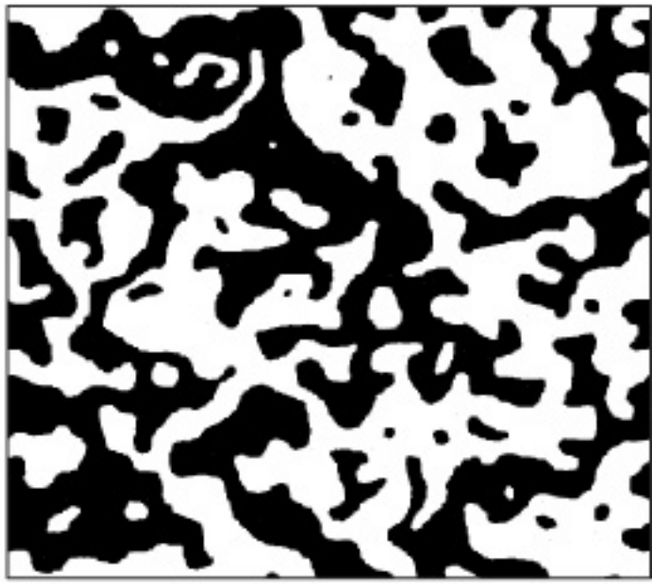
Macroscopic Observables: pressure, energy, heat capacity etc.

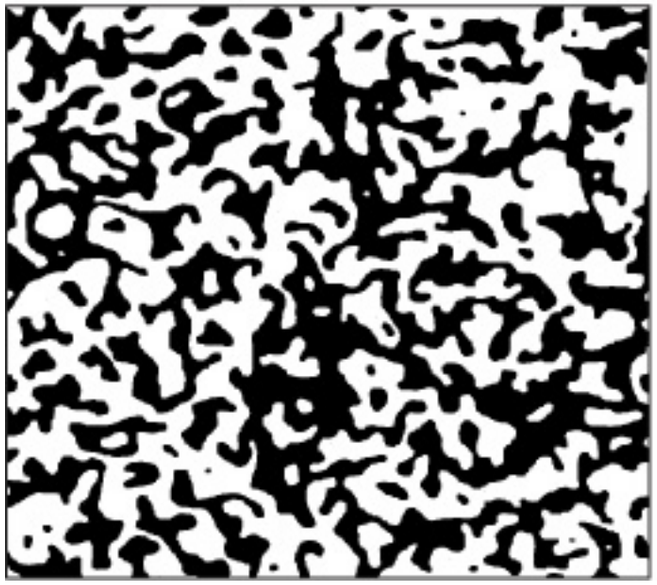
Generation of a Representative Equilibrium Ensemble

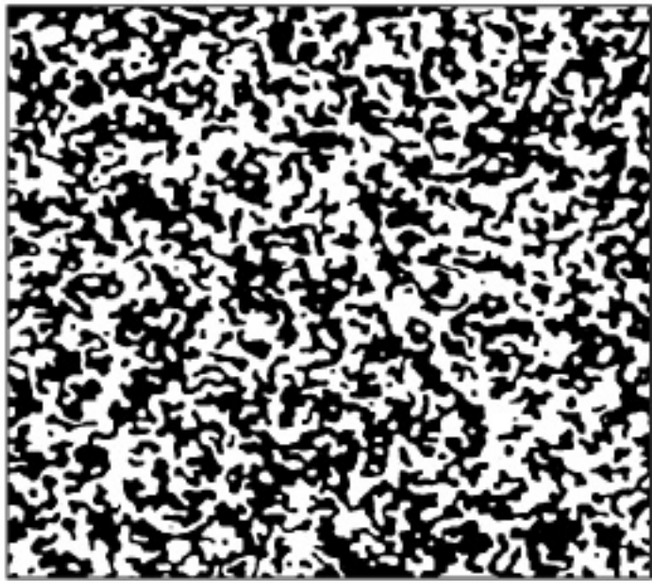
Metropolis Monte Carlo (MC)	**Molecular Dynamics (MD):**
Random Walk algorithm	Newton's Equation
No true analogue of time	Simulation times from ps to ms
Faster, MC compare energies no forces.	Require more computer time to achieve same level of convergence
Thermodynamics and structural properties	Thermodynamics, structural and dynamic (equilibrium and non-equilibrium) properties

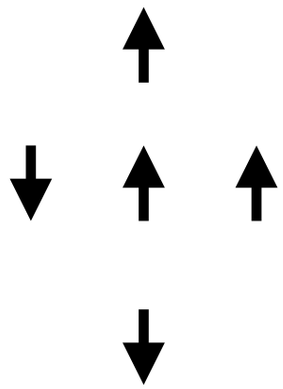
Jorgensen W.L. and Tirado-Rives J. J*. Phys. Chem.* (**1996**) 100, 14508

Monte Carlo — Ising model









What is this E?

$$P(\text{accept}) = \min(1, \exp(-\beta[E_f(x) - E_i(x)]))$$

F Caccioli et al. J. Stat. Mech. (2008) L07001

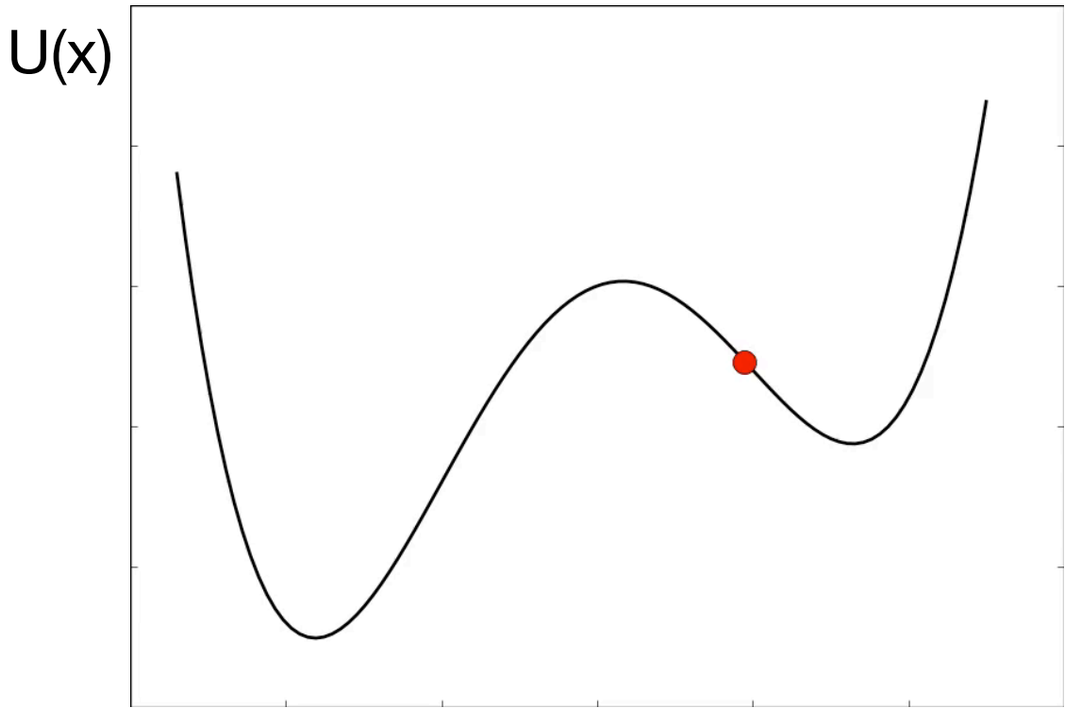
Langevin dynamics

Langevin Dynamics (LD)

Stochastic differential equation in which two force terms have been added to Newton's second law to approximate the effects of neglected degrees of freedom.

$$\vec{F}_i - \gamma_i \vec{v_i} + \vec{R}_i(t) = m_i \vec{a}_i$$

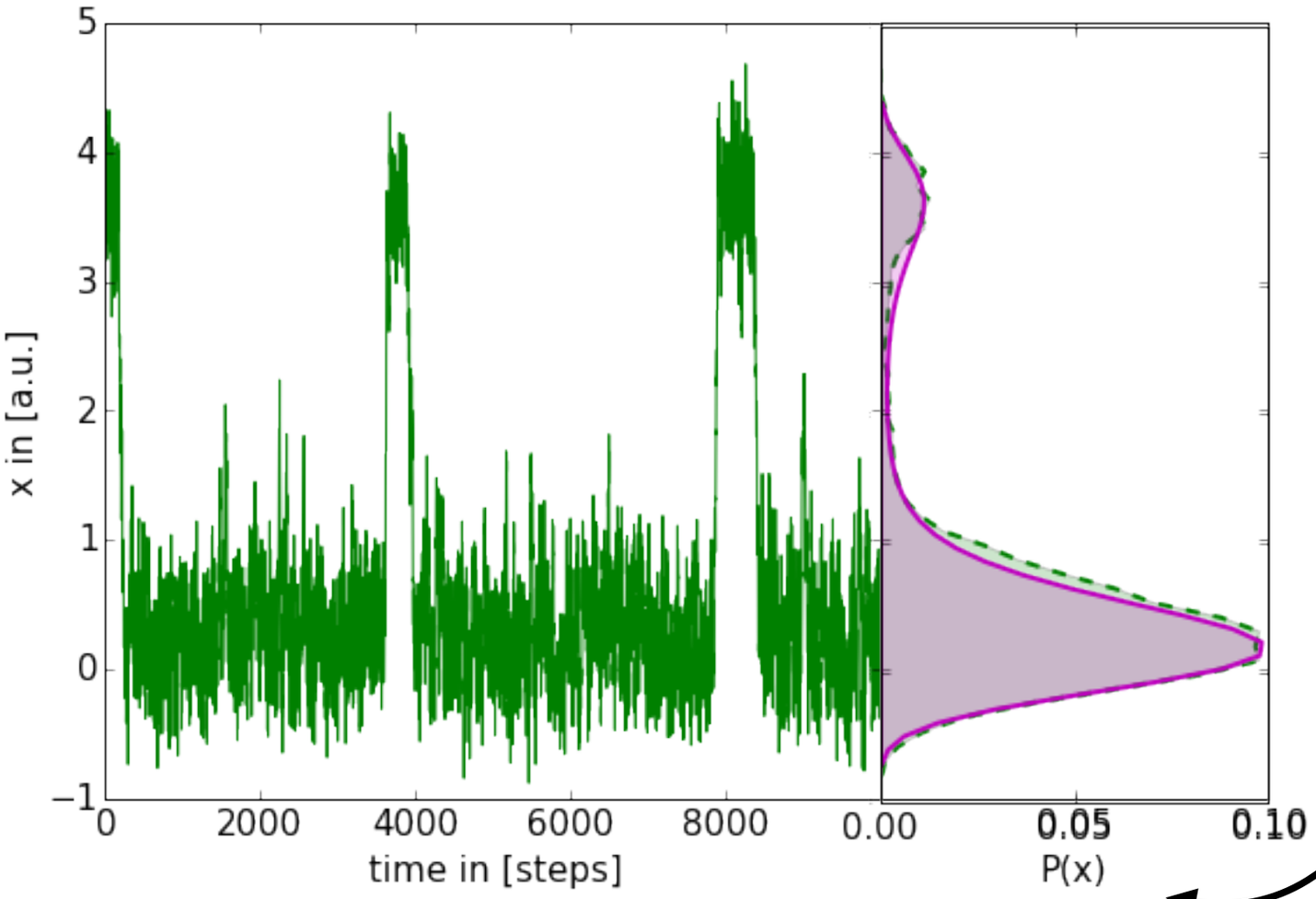
Frictional force **Random force**



Jorgensen W.L. and Tirado-Rives J. J. *Phys. Chem.* (**1996**) 100, 14508

A closer look at LD?

Boltzmann tells us, why it is easier to sample this potential at higher temperature



random force will sample both wells

Boltzmann factor

$$P(x) = \frac{\exp(-U(x)/kT)}{\int_{-\infty}^{+\infty} \exp(-U(x)/kT)}$$

Partition function: Z

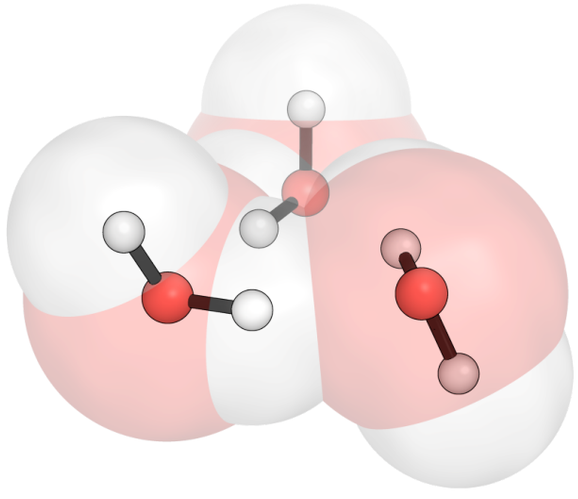
Free Energy:

F = -ln Z

Possible to compute free energy differences between configurations, or thermodynamic states.

Molecular force fields

Example of a potential energy (force field) function:



- It contains all the parameters for this expression plus masses
- Parameters defined in terms of atom type (e.g. sp^3, sp^2 and sp atoms).
- In most cases pair-additive, non-polarisable.

$$U = \sum_{\text{bonds}} K(\mathbf{r} - \mathbf{r_{eq}})^2$$

$$+ \sum_{\text{angles}} K_\theta(\theta - \theta_{eq})^2$$

$$+ \sum_{\text{dihedral}} \frac{V_n}{2}[1 + \cos(n\phi - \gamma)]$$

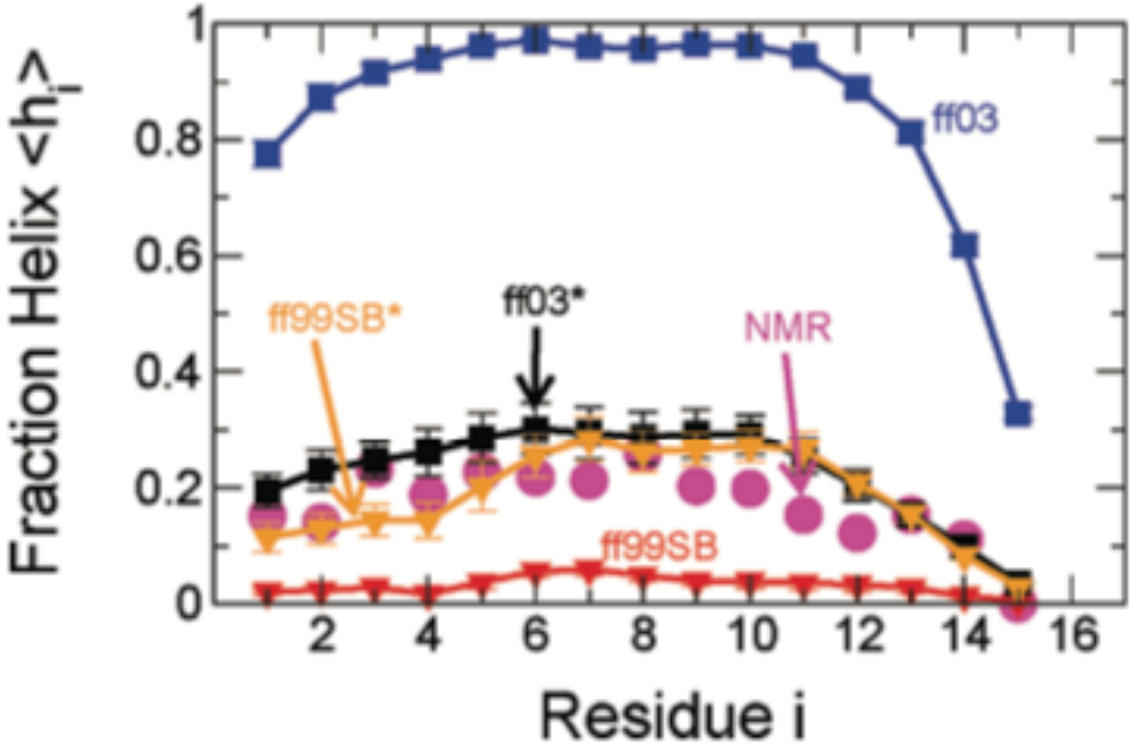
$$+ \sum_{\text{non-bonded}} [\frac{A_{ij}}{R_{ij}^{12}} - \frac{B_{ij}}{R_{ij}^{6}} + \frac{q_i q_j}{\epsilon R_{ij}}]$$

Molecular force fields — so many of them

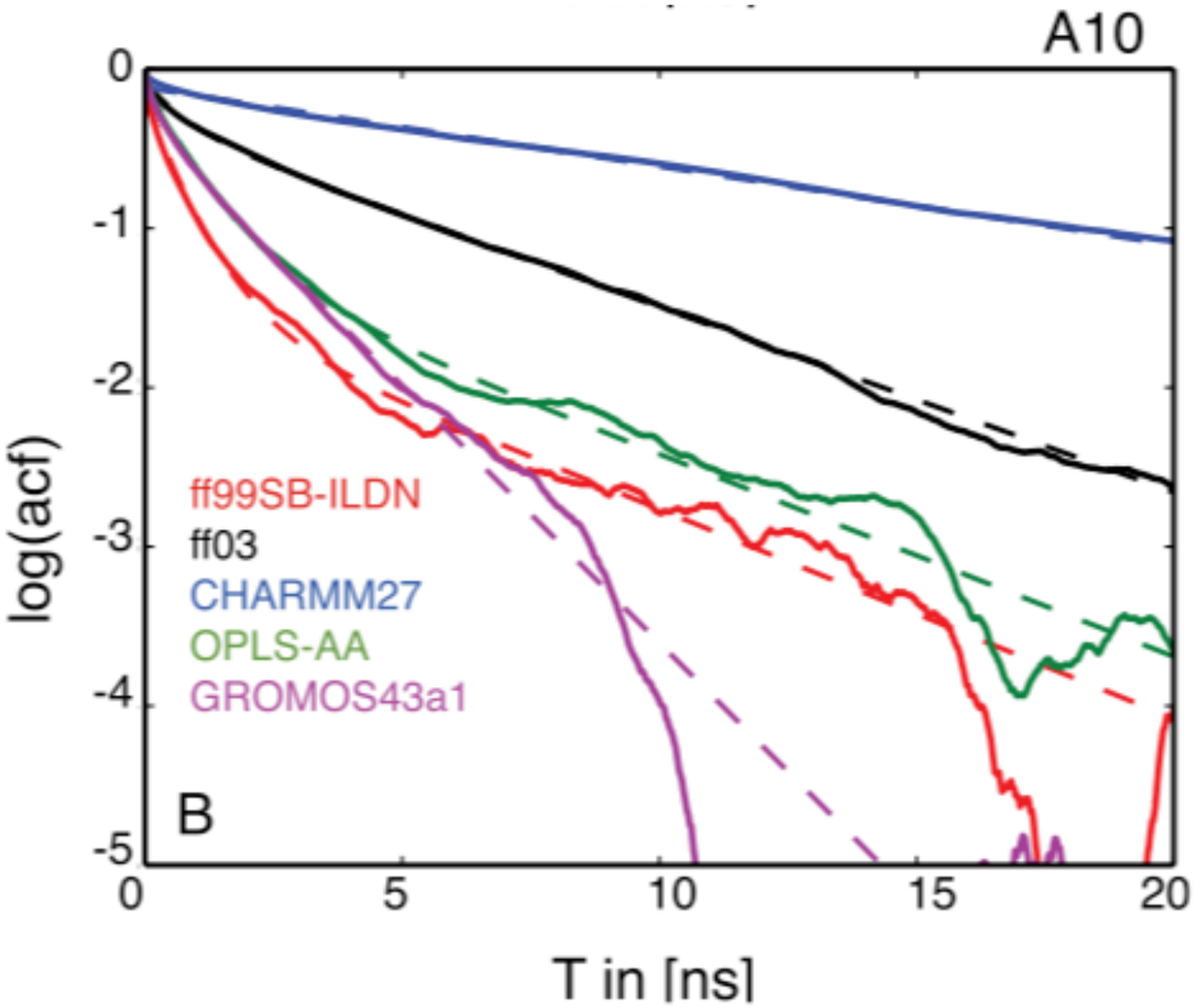
- They can be modified, different implementations depending on the softwa
- Pay attention to conversion rules!!

Amber	DNA, proteins and lipids. Generalized Amber Force Field (GAFF) covers most organic space.
OPLS	Organic molecules in the liquid phase. Available as all-atom (AA) or united-atom (UA) form.
CHARMM	DNA, proteins, lipids, sugars.
GROMOS	General purpose. Organic and biochemical space.
Dreiding	General purpose. Includes some parameters for metals and main group elements.
MM2/MM3	General purpose. MM2 used for hydrocarbons. MM3 includes most of organic space plus some other main group and metal atom types.
Water	Rigid and flexible models available. TIP3P widely used in biological simulation.

Molecular force fields — which one to use?



Best et al. J. Phys. Chem. B 2009, **113**, 9004–9015

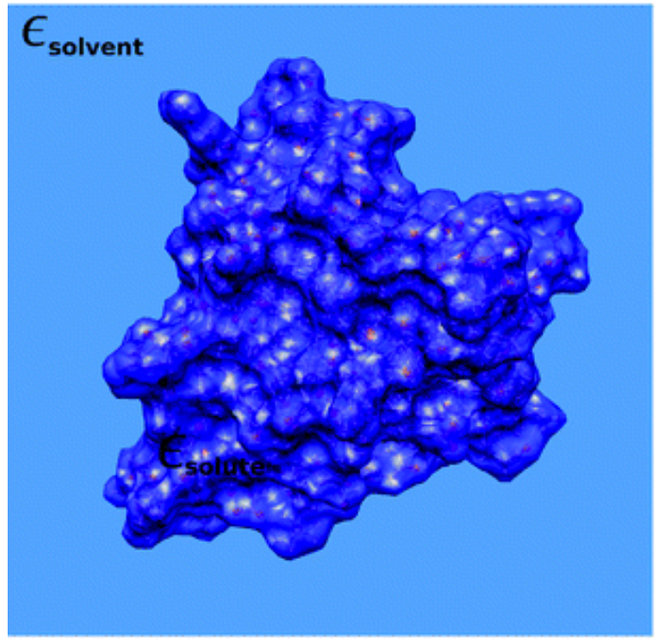


Vitalini, Mey et al, J. Chem. Phys. **142**, 084101 (2015)

What about water?

Implicit Solvent Model

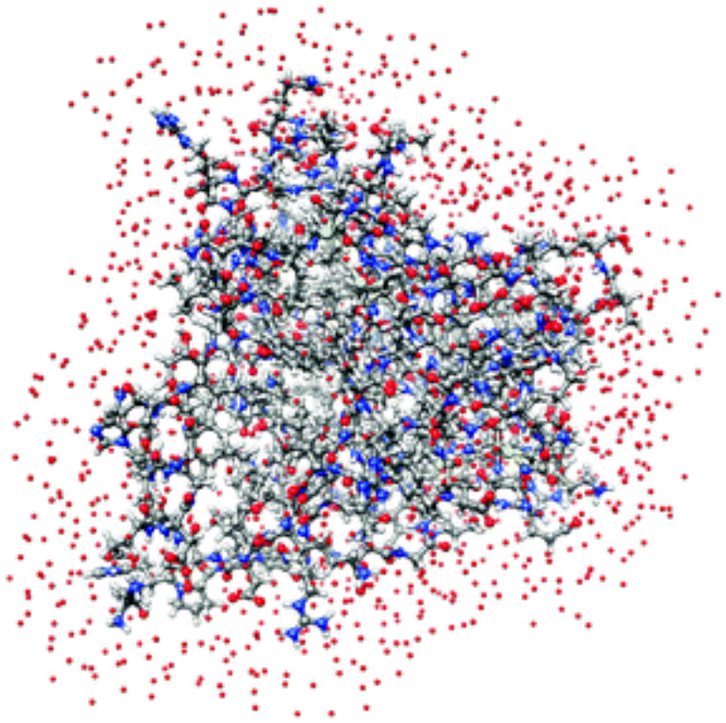
- Macromolecule interacts only with itself.
- Solvent effects accounted for in dielectric constant.



Generalised Born (GB): A more advanced implicit solvent model. GBSA is a GB model augmented with the hydrophobic solvent accessible surface area (SA) term.

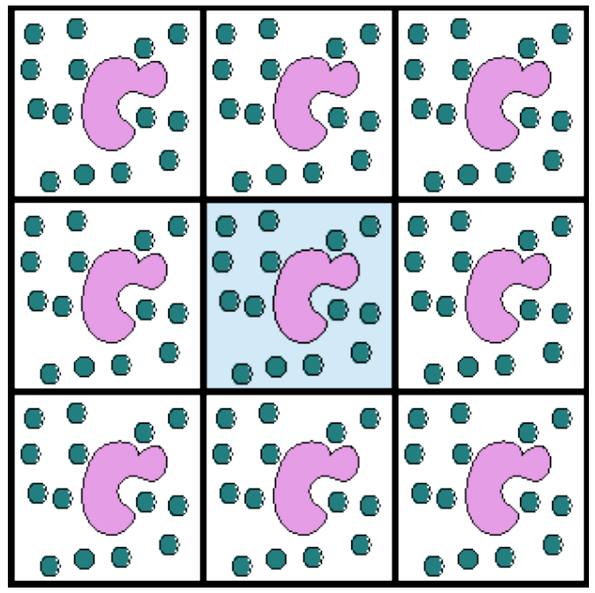
Explicit Solvent Model

- Macromolecule is surrounded by solvent molecules (e.g. water).
- Specific non-bonded interactions are calculated.



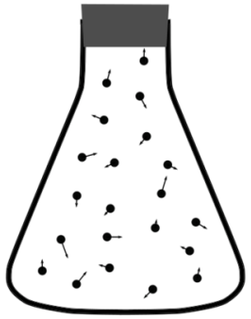
Shugar et al. *Mol. BioSyst.*, **2011**, 7, 2923-2949

Periodic boundary conditions



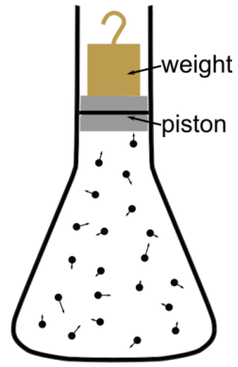
- Boundary Conditions are unnatural however, we cannot simulate infinite systems.
- We use PBC to avoid real phase boundaries.
- **Minimum-Image Convention:** cut-off radius for the Lennard-Jones (Coulomb) interactions cannot exceed half the box size.
- Truly bad things can happen in charged systems (accumulation of charges at the cut-off boundary, wrong energies). **Particle mesh Ewald algorithm.**

Barostat/Thermostat



Canonical
(NVT)

Requires a thermostat to make sure the temperature of the system stays the same.



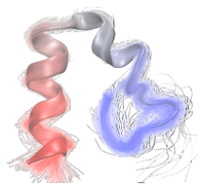
Isobaric-Isothermal
(NPT)

Requires a thermostat and barostat to make sure the temperature and pressure of the system stays the same.

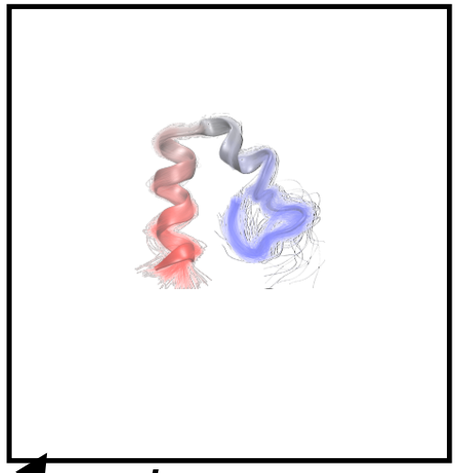
Careful! While Berendsen T and P control are simple to implement and use, they can steadily drive the system state far from equilibrium toward equilibrated state.

Good choice! **Nosé-Hoover thermostat with Rahman-Parrinello barostat**

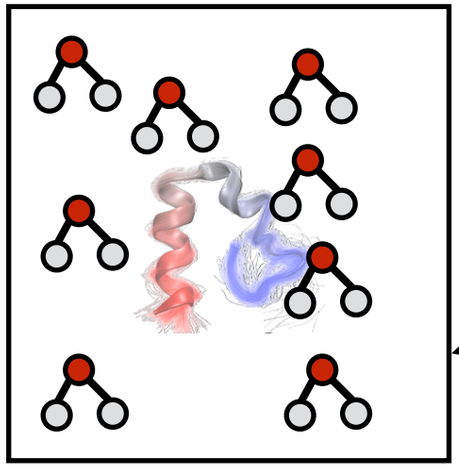
Putting everything together



protein crystal



box+ protein



box+ protein+ water > 100k particles

or implicit solvent

Decide on: Thermostat, Barostat and integrator

Each atom has a position and velocity and is integrated according to Newtonian dynamics, with a timestep of **2 fs.**

— Reaching interesting timescales means millions of MD steps

Timestep and integrator

Depends on the system that you are interested in

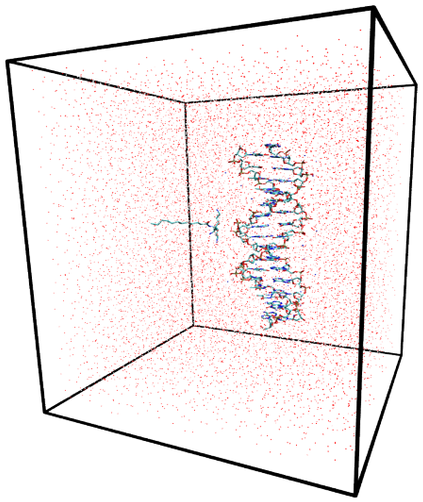
- Too small - only a small region of phase space will be covered
- Too large - instabilities may arise due to atom overlaps

For flexible molecules the time step should be ~ 1/10th the time of the fastest motion (often C-H vibrations).

System	Types of motion present	Suggested time step (s)
Atoms	Translation	10^{-14}
Rigid molecules	Translation and rotation	5×10^{-15}
Flexible molecules, rigid bonds	Translations, rotation, torsion	2×10^{-15}
Flexible molecules, flexible bonds	Translation, rotation, torsion, vibration	10^{-15} or 5×10^{-16}

Simple example of MD workflow

System prep

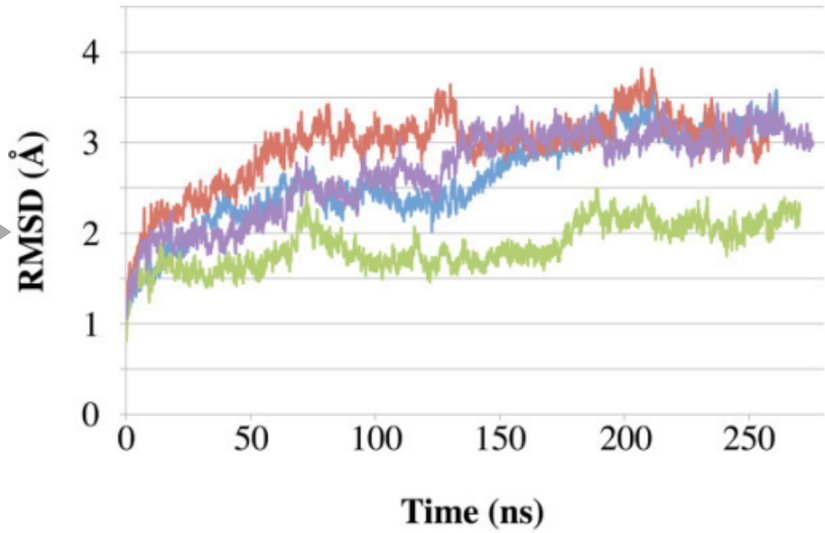


Simulation

The IBM Blue Gene/P supercomputer

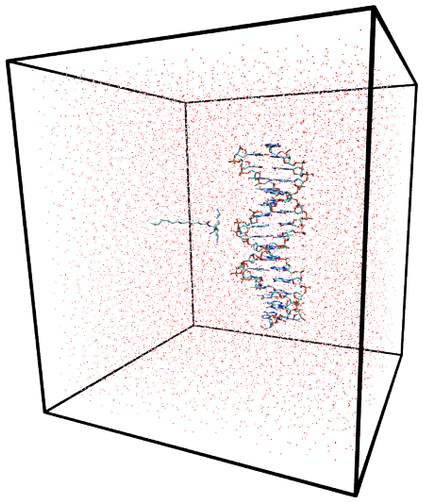
Cloud computing

Analysis



Simple example of MD workflow

System prep



Simulation

The IBM Blue Gene/P supercomputer

What will actually happen

More detailed actual workflow

Generate Initial Coordinates

Assign Force Field Parameters

Minimisation

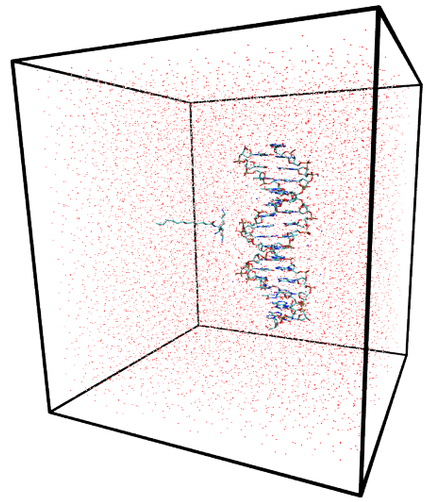
Assign Initial Velocities
Heat the system

Rescaling

Equilibration Dynamics

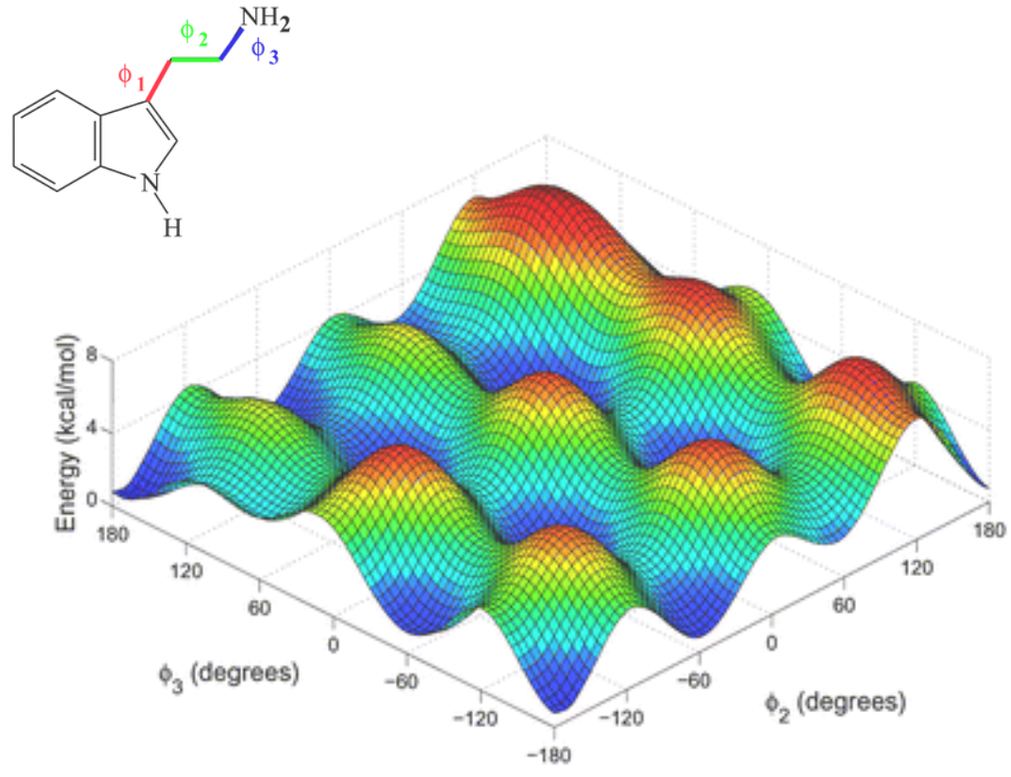
Production

Analysis
T/P RMSD,
PCA, MSM, MI



Minimisation

The potential energy function of a molecular system is a very complex landscape.



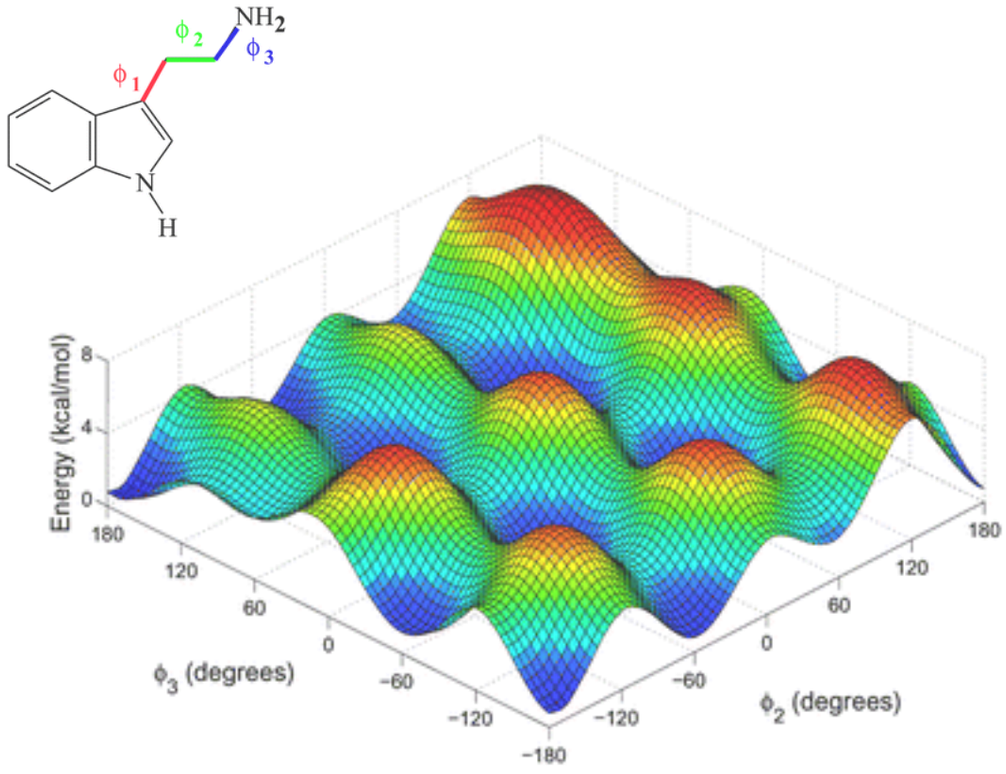
It has a deepest point, **global minimum**, and a very large number of **local minima.**

The **gradient** ($-\nabla V(\mathbf{r^N})$) *i*s zero & the **Hessian matrix** ($H(\mathbf{r^N})$) has non-negative eigenvalues.

In between local minima there are saddle points: $H(\mathbf{r^N})$ has **only one** negative eigenvalue.

Sturdy Y.K and Clary D.C. *Phys. Chem. Chem. Phys.*, **2007**, 9, 2065

Minimisation



The potential energy function of a molecular system is a very complex landscape.

NO minimisation method can guarantee the determination of the global minimum …..

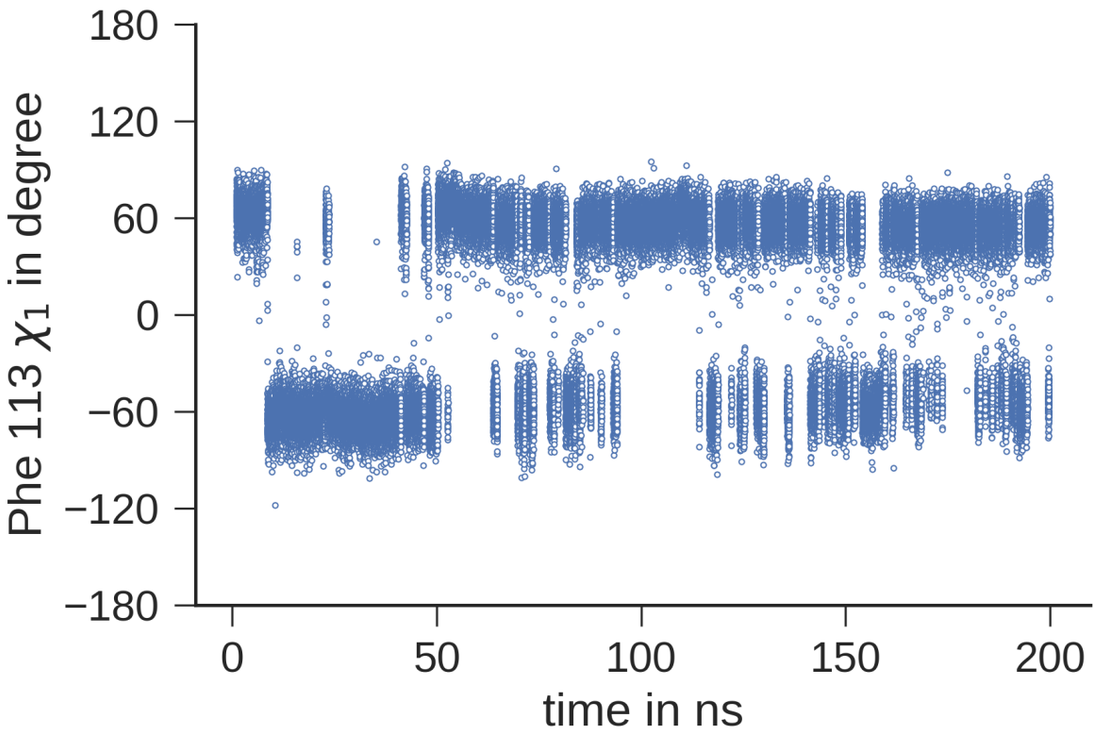
Finding this nearest local minimum is all these methods can do for you!

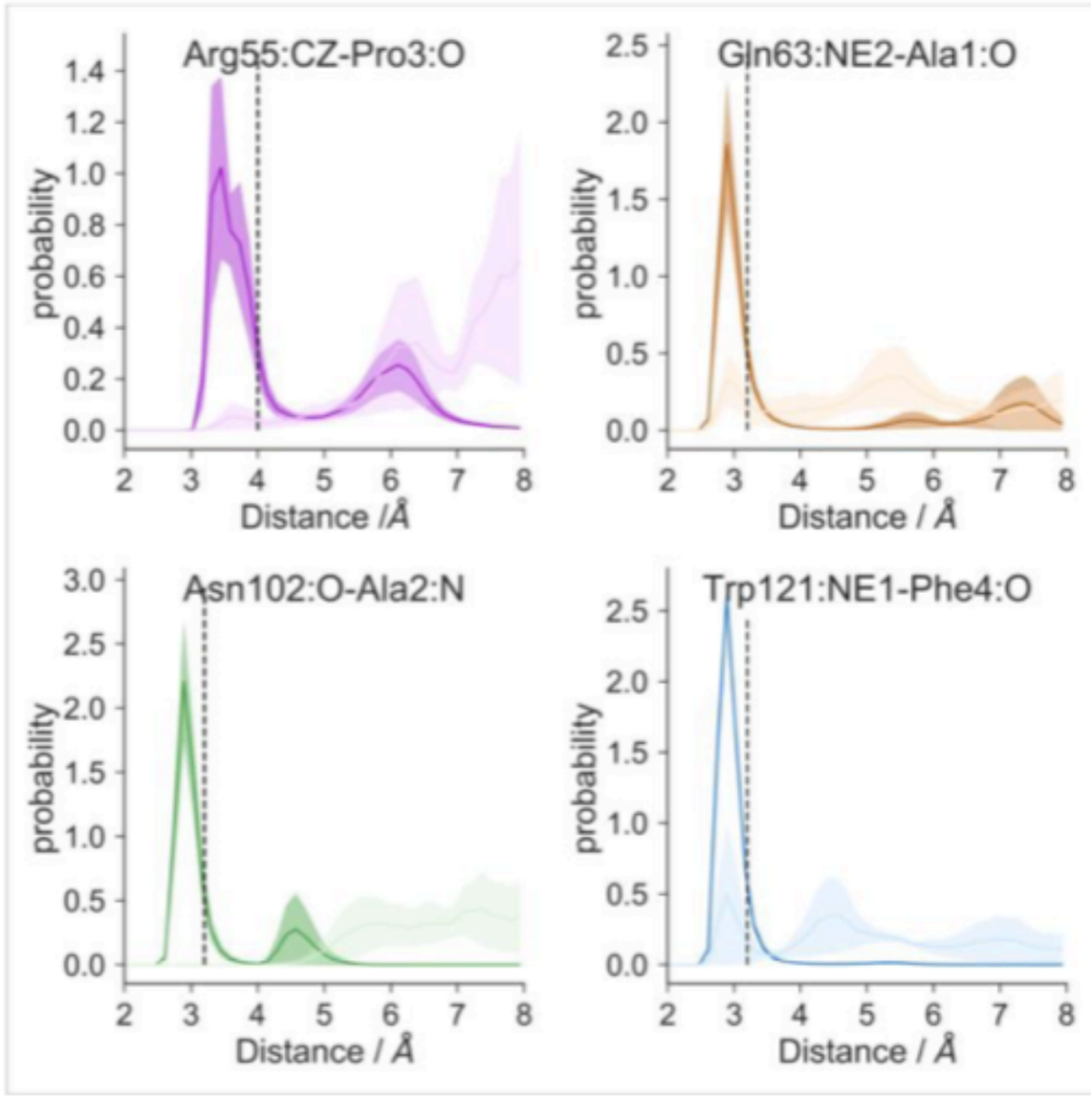
Methods

- Ones that require only function evaluations
- Ones that use derivative information. **Conjugate Gradient/ Steepest Descent**
- Ones that use second derivative information as well (QM approaches)

Sturdy Y.K and Clary D.C. *Phys. Chem. Chem. Phys.*, **2007**, 9, 2065

Analysis





Enhanced analysis techniques:

PCA

Markov State models

Mutual information

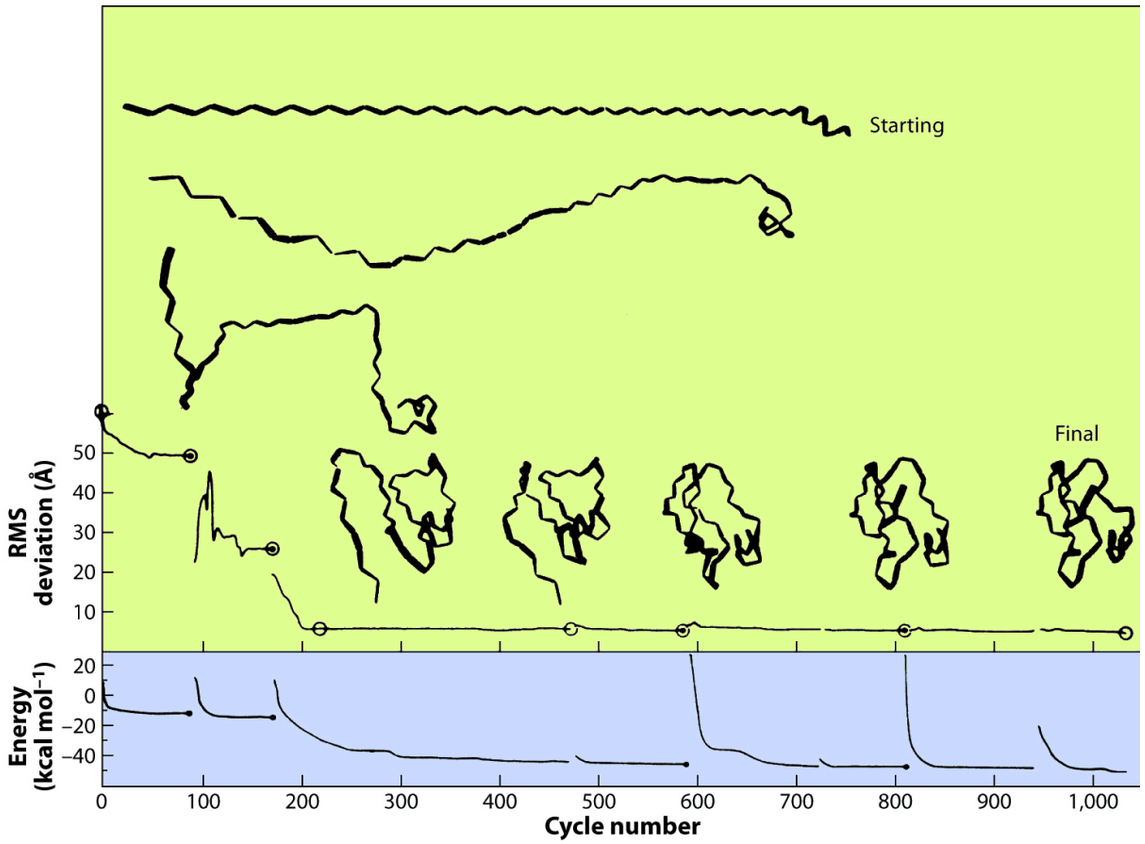
- > Also enhanced simulation techniques

Wapeesittipan et al https://doi.org/10.1101/224329

Applications

Computer Simulations of Protein Folding

First MD simulation of a biological folding process (BPTI) (Levitt & Warshel, 1975)



The Nobel Prize in Chemistry 2013

Martin Karplus Michael Levitt Arieh Warshel

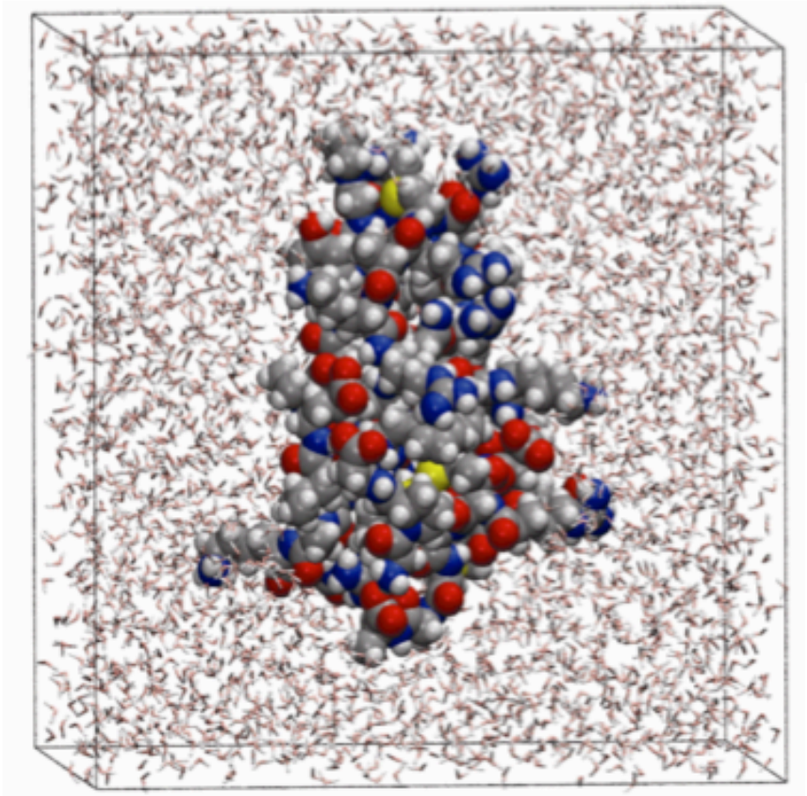
https://youtu.be/_hMa6G0ZoPQ?t=40

Levitt, M; A Warshel *Nature* **1975** 253, 694.
A. Warshel Nature ***1976*** 260, 679.

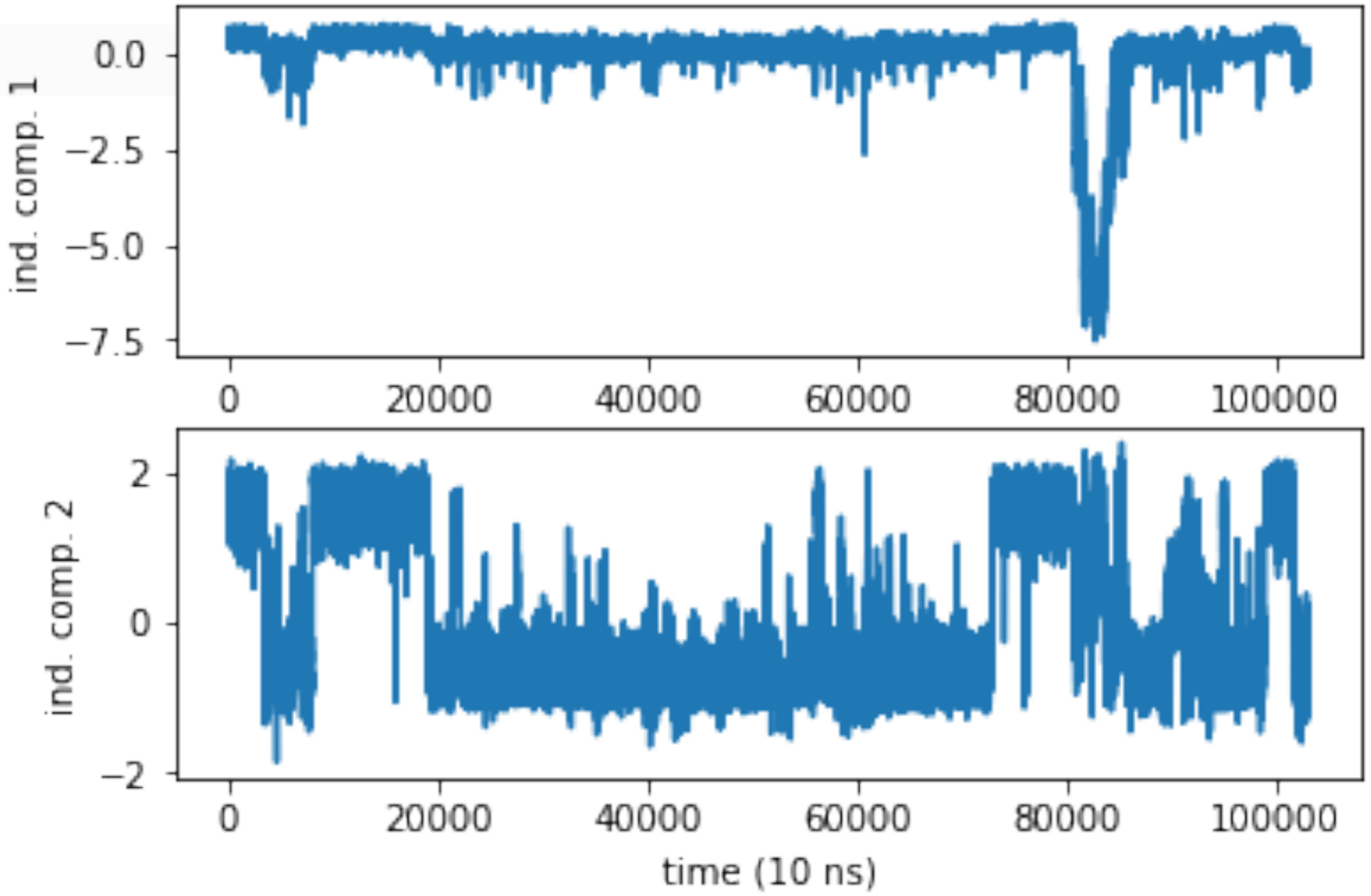
Applications

Computer Simulations of Protein Folding

BPTI revisited — 1st millisecond simulation using Anton



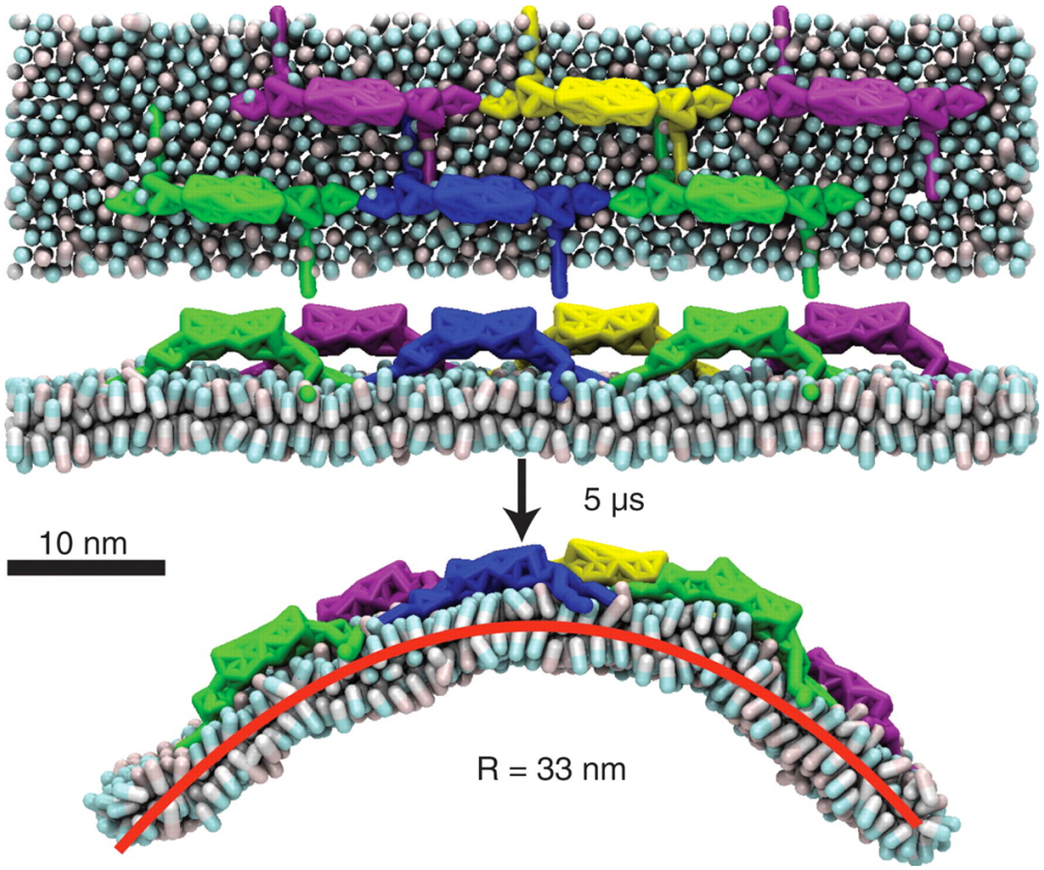
(a)



D.E. Shaw et al.
pyEMMA project: pyemma.org

Applications

Four-Scale Description of Membrane Sculpting by BAR Domains



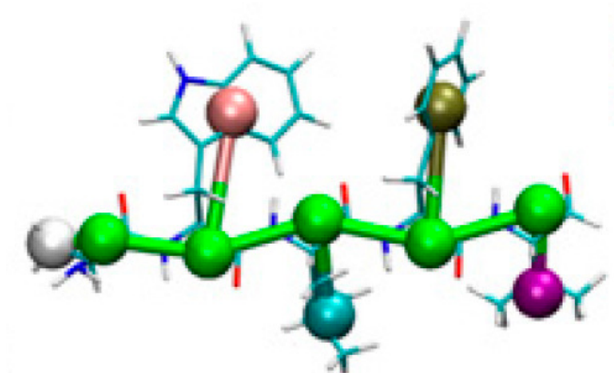
Computation was done using four levels of description

a) All-atom MD,

b) Residue-based CG (RBCG) MD ~10 atoms per CG bead

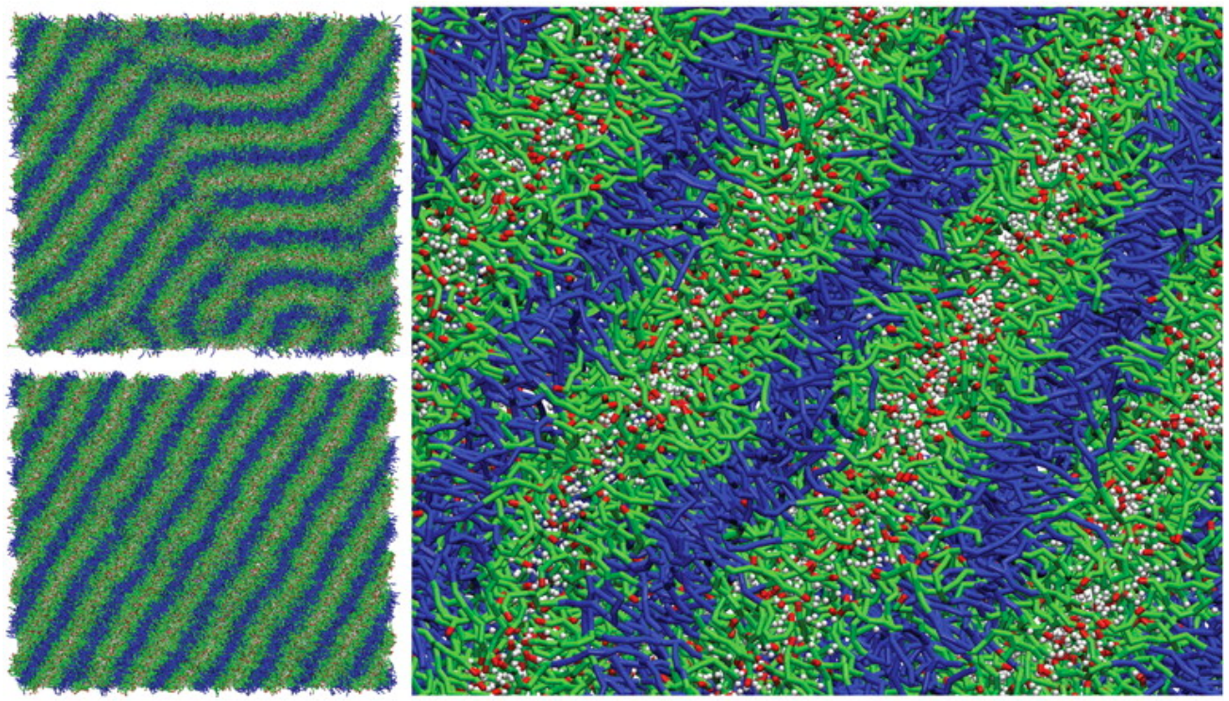
c) Shape-based CG (SBCG) MD ~150 atoms per CG bead

d) Continuum elastic membrane model. NAMD Software



A. Arkhipov, Y. Yin, K. Schulten *Biophys J* **2008**, 95, 2806

Applications



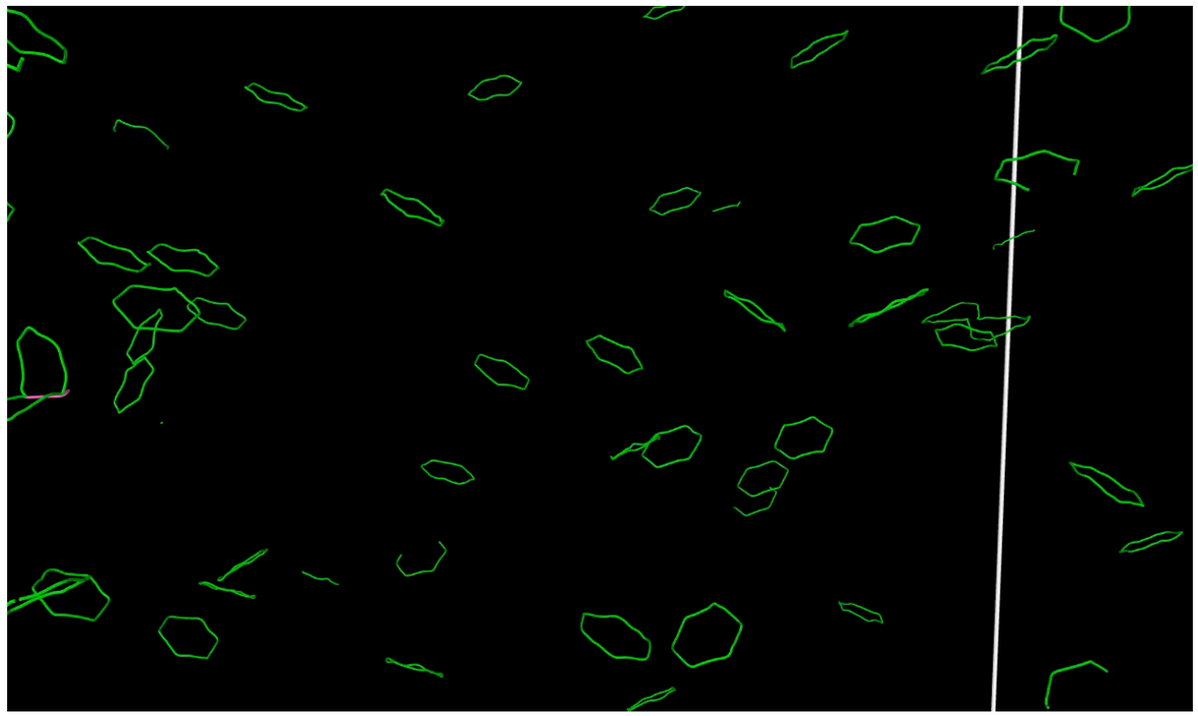
Large-Scale Molecular Dynamics Simulations of Self-Assembling Systems

Science, **2008,** 321, 798-800

Software: LAMMPS

snapshots of a solution with 80% surfactant (C12E6) and 20% water

800,000 Coarse grained particles.



Probing the limits of metal plasticity with molecular dynamics simulations

Nature, **2017**, 550, 492

Software: LAMMPS

Fragments of tantalum single crystals embedded into an infinite crystal under PBC (about 268million atoms)

The crystal was compressed either at a constant or variable rate

Useful Reading

- GROMACS manual www.gromacs.org
- Gromacs Tutorial: http://www.bevanlab.biochem.vt.edu/Pages/Personal/justin/gmx-tutorials/
- OpenMM Tutorial: http://openmm.org/tutorials/index.html
- LAMMPS Tutoriales: http://lammps.sandia.gov/tutorials.html
- Charmm Tutorial - http://www.ch.embnet.org/MD_tutorial/
- Allen, M.P., Tildesley, D.J. “Computer Simulation of Liquids”, 1987.
- Frenkel, D., Smit B. “Understanding Molecular Simulation”, Academic Press, 1996.
- Keffer, D.J., “A Working Person’s Guide to Molecular Dynamics”, http://utkstair.org/clausius/docs/che548/pdf/md_sim.pdf.

Useful Tools

Classical simulation codes

- Gromacs www.gromacs.org
- OpenMM http://openmm.org/
- Amber – http://ambermd.org/
- NAMD http://www.ks.uiuc.edu/Research/namd/
- CHARMM - http://www.charmm.org
- DL_POLY – http://www.stfc.ac.uk/cse/25526.aspx
- LAMMPS – http://lammps.sandia.gov
- BioSimSpace — biosimspace.org (Workflow design tool)

Visualisation software

- VMD – http://www.ks.uiuc.edu/Research/vmd/
- Pymol - http://www.pymol.org
- Chimera https://www.cgl.ucsf.edu/chimera/