Molecular rare event simulations

Peter Bolhuis

van 't Hoff institute for Molecular Sciences
University of Amsterdam, The Netherlands

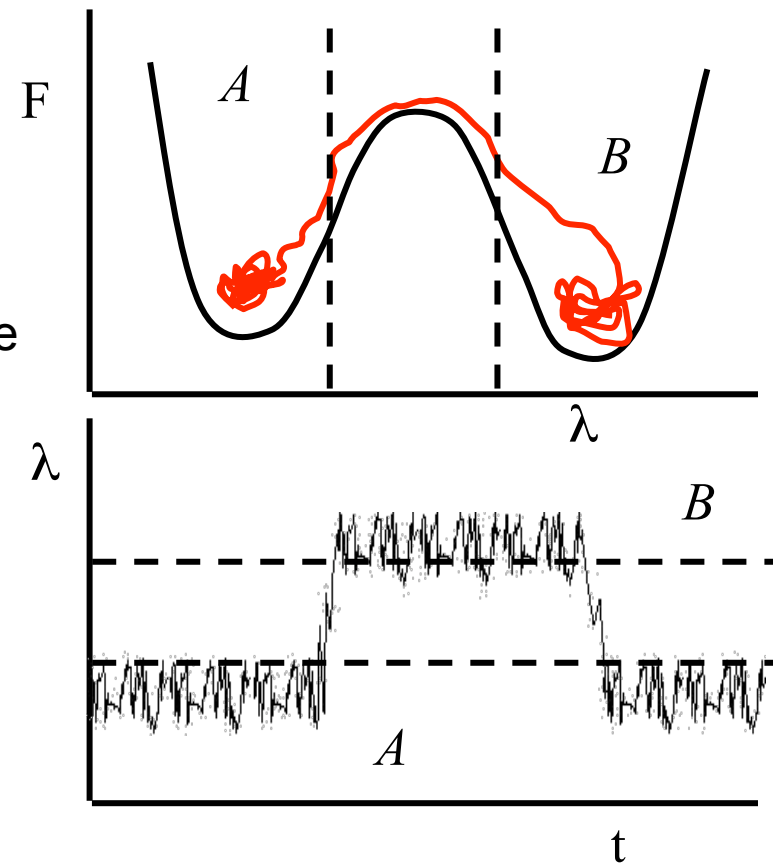# *Rare events*

Interesting transitions in complex systems

- – solution chemistry
- – protein folding
- – enzymatic reactions
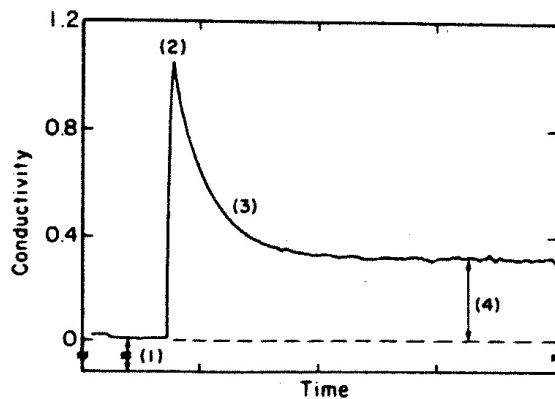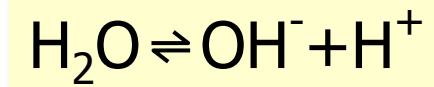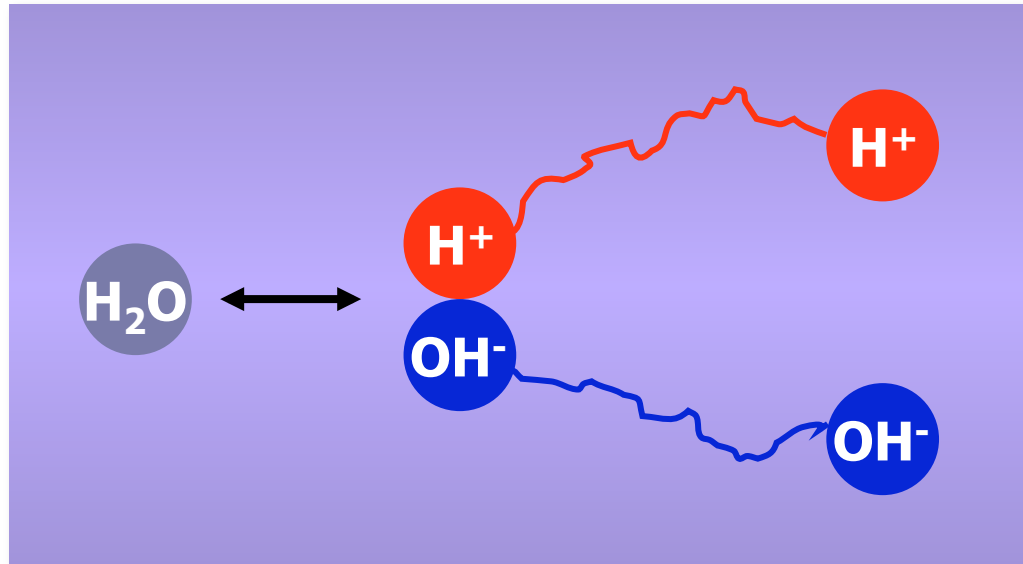- – complex surface reactions
- – nucleation of phases

These reactions happen on a long time scale compared to the molecular timescale (eg solvent motion)

dominated by collective, rare events

Example: Autoionization in liquid water

$$H_2O \rightleftharpoons OH^- + H^+$$

Average life time > 10 h

M. Eigen and L. De Maeyer, *Z. Elektrochemie* **59**, 987 (1955)
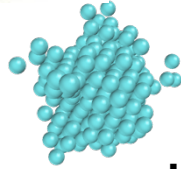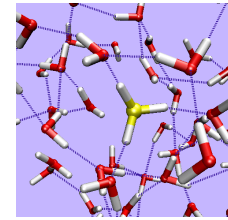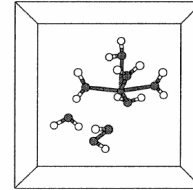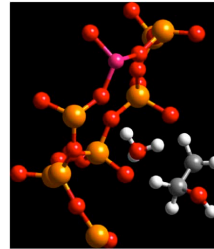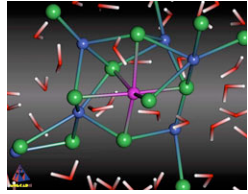
crystallisation

catalysis

solution reactions
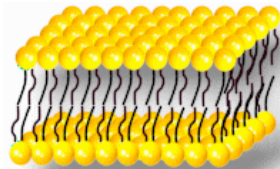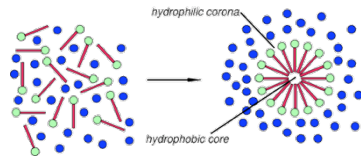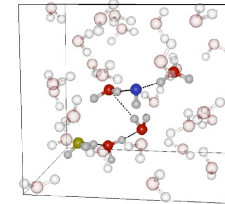
complex fluids

RARE EVENTS

enzyme reactions
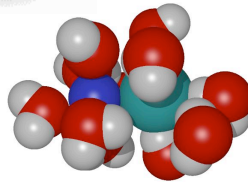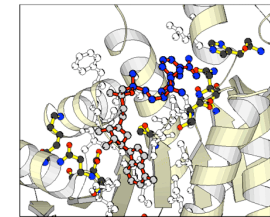
solvent effects

folding & binding

isomerization

DNA-Polymerase-beta

fingers

palm

thumb

8-kD
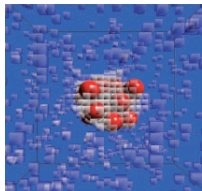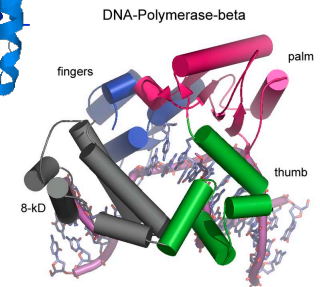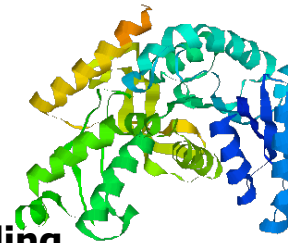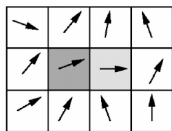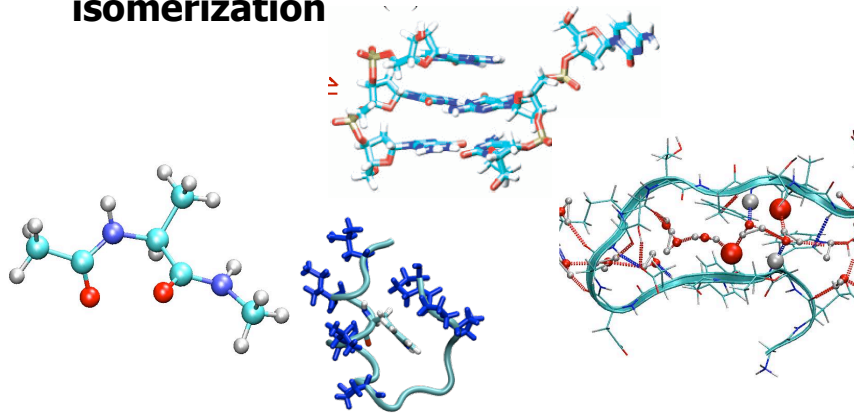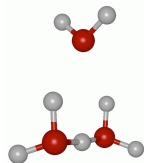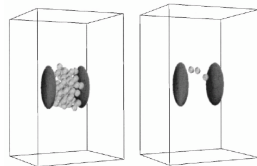
Smooth vs rough energy landscapes
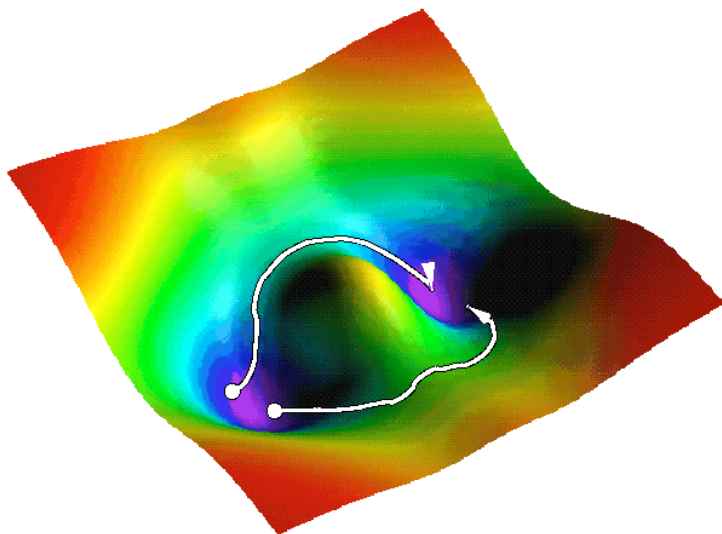
- # saddle points limited
- reaction coordinate known
- pathways can be enumerated
- Arrhenius (TST) applies

- saddle points uncountable
- reaction coordinates unknown
- entropy important, many pathways
- Arrhenius still applies????

How do we explore?
Usually by free energy surface

Dellago logo™
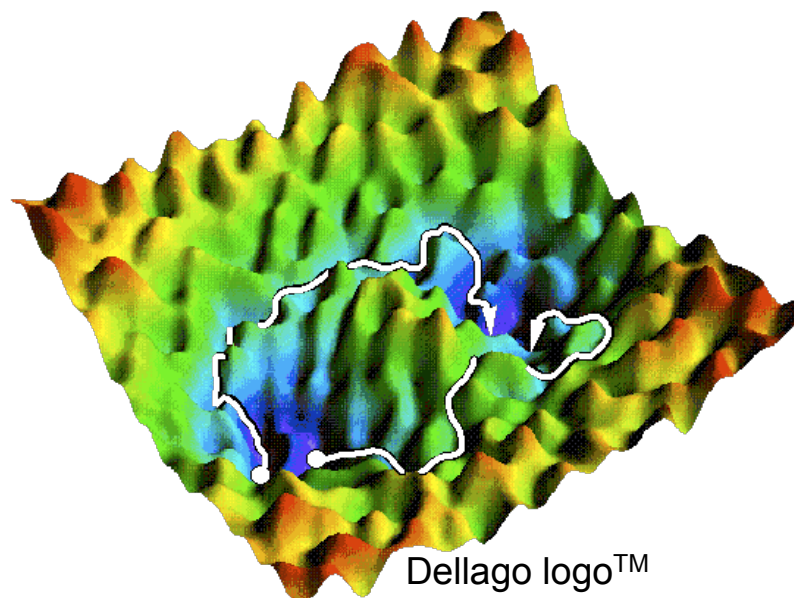
Free energy for protein folding
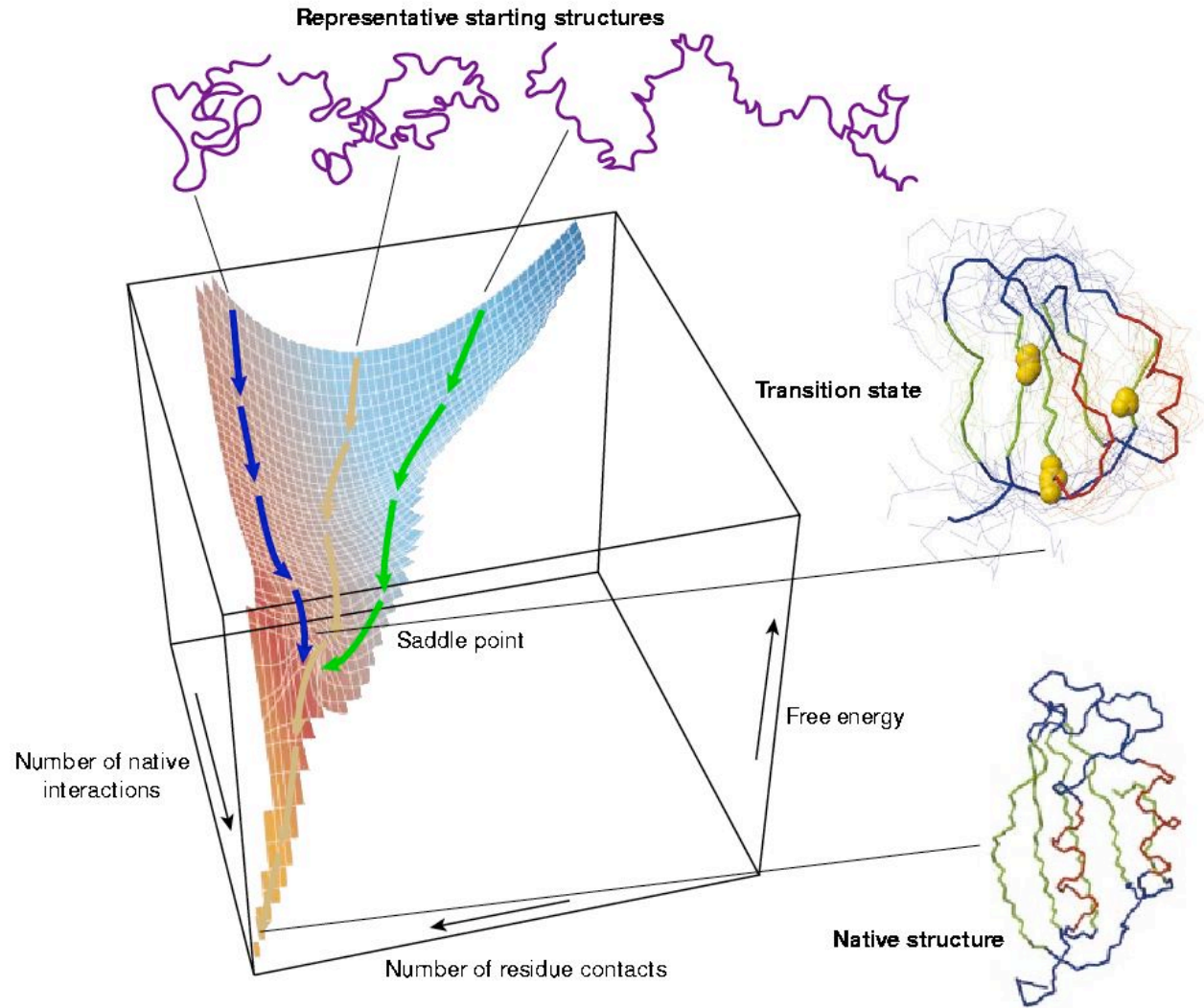
Taken from Dobson, Nature, 2003

Questions

- How do we get free energy? (thermodyamics)

- How do we compute rate constants? (kinetics)

- How do find transition states? (mechanism)

- What is the reaction coordinate of a reaction/process?

- How do we know we have the correct reaction coordinate?

Outline

- Lecture 1 :Introduction molecular simulation
  - Monte Carlo
  - Molecular dynamics
  - Application

- Lecture 2 : Rare events simulations
  - Transition State theory
  - Computing free energies
  - Bennett-Chandler approach

- Lecture 3: Advanced trajectory based methods
  - Transition Path Sampling
  - Analyzing reaction coordinates

Han sur Lesse

Molecular simulation of complex systems

◆ Molecular dynamics: solve equations of motion
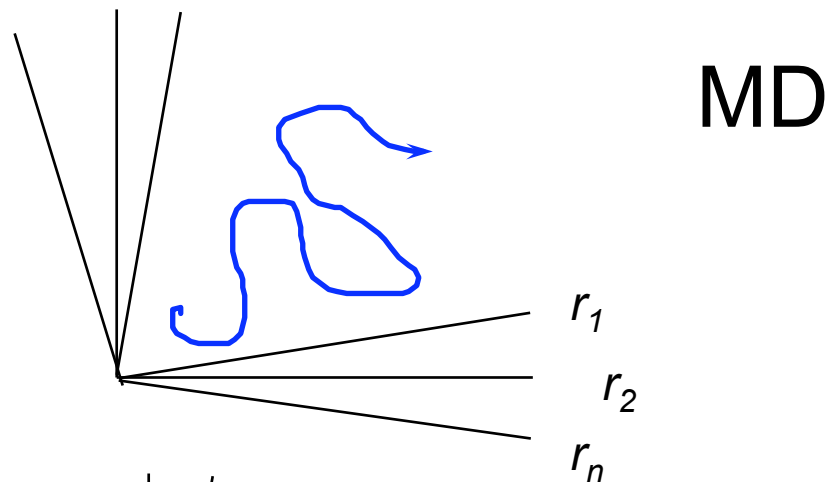
◆ Monte Carlo: importance sampling

MD

r_1

r_2

r_n

MC

r_1

r_2

r_n
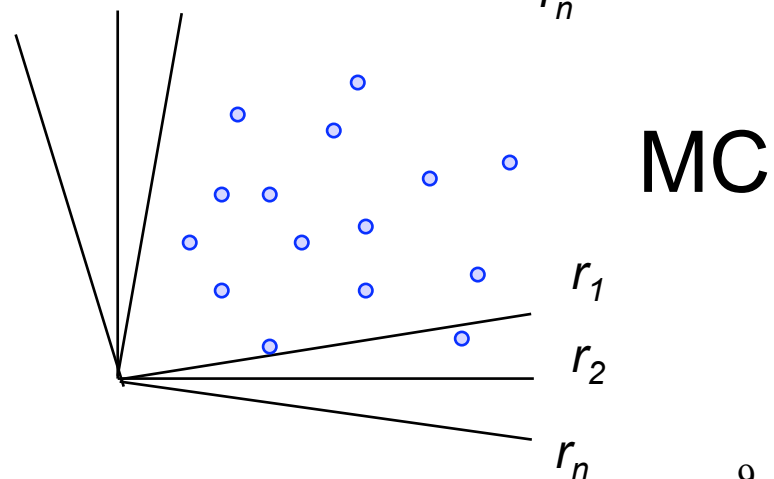
Monte Carlo

Aim : compute thermal averages (to obtain thermodynamics)

$$\langle A \rangle = \frac{\sum_i \exp(-\beta \epsilon_i) A_i}{\sum_i \exp(-\beta \epsilon_i)}$$

Where i labels all eigenstates of the system, and

$$A_i = \langle i | A | i \rangle$$

Phase space integral

In the classical limit we replace the sum over quantum states by an integral over phase space

$$\langle A \rangle = \frac{\int dp^N\, dr^N\, A(p^N, r^N) \exp(-\beta \mathcal{H}(p^N, r^N))}{\int dp^N\, dr^N \exp(-\beta \mathcal{H}(p^N, r^N))}$$

r is position vector and p is momentum vector.

N is number of particles

H is the Hamiltonian of the system

and $\beta = 1/k_B T$

In replacing the sum by an integral, we have attributed a "volume" h^{3N} to every quantum state

Hamiltonian is sum of potential and kinetic energy

$$H(p^N, r^N) = U(r^N) + \sum_i \frac{p_i^2}{2m_i}$$

If A does not depend on momenta the integration of momenta can be done analytically (Gaussian integral)

$$\langle A \rangle = \frac{\int dr^N A(r^N) \exp(-\beta \mathcal{U}(r^N))}{\int dr^N \exp(-\beta \mathcal{U}(r^N))}$$

Problems with this integral

- We cannot compute the sum over all quantum states because there are so many

- And we cannot compute the classical integral either (except the integration over momenta).

- Consider "normal" numerical integration of 100 particles, 3 dimensions, 10 points in every direction.

- Requires 10^{300} points for a very poor estimate of the integral…

Daan Frenkel's analogy

A similar but much less
serious problem:

Measure the depth of the
Nile by quadrature.

Not very efficient…

NILE →

© D. Frenkel

The solution

A better strategy:

measure depth while walking around in the Nile

Importance sampling

© D. Frenkel

Importance sampling

We wish to perform a **random walk** in configuration space, such that the number of times that each point is visited, is proportional to its Boltzmann weight

$$\mathcal{N}(r^N) = c \exp[-\beta U(r^N)]$$

Then

$$\langle A \rangle \approx \frac{1}{L} \sum_{1=1}^{L} A(r_i^N)$$

how to achieve such a random walk?

Balance

Whatever our rule is for moving from one point to another, it should not destroy the equilibrium distribution.

That is: in equilibrium we must have balance

$$\pi(o \rightarrow n)$$

$$\pi(n \rightarrow o)$$
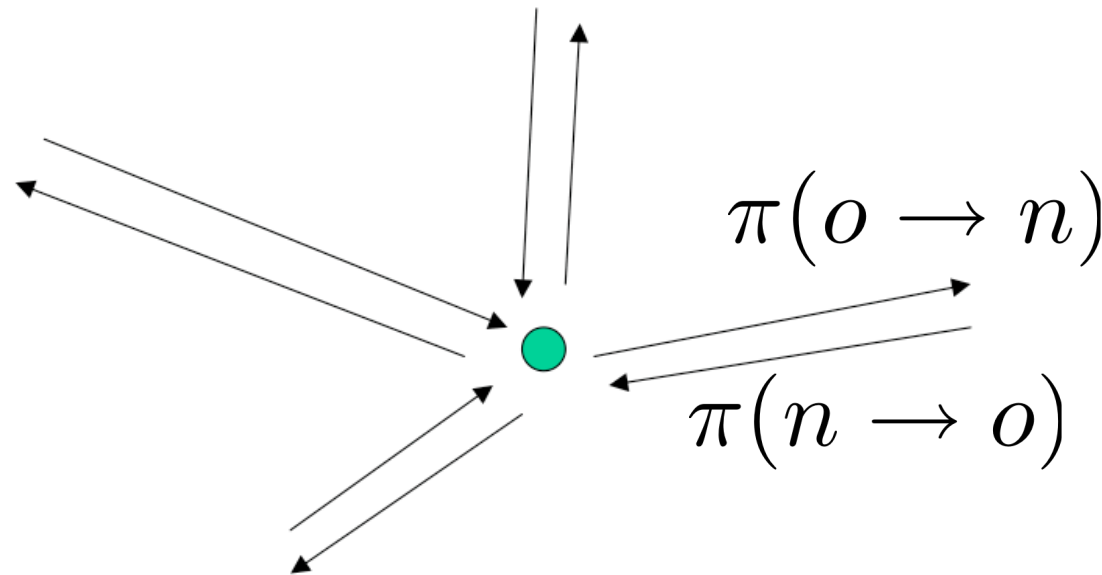
Detailed balance

denoting **o** old configuration and **n** the new state reachable from **o** this
balance condition is

$$\mathcal{N}(o) \sum_n \pi(o \to n) = \sum_n \mathcal{N}(n)\pi(n \to o)$$

A stronger condition is

$$\mathcal{N}(o)\pi(o \to n) = \mathcal{N}(n)\pi(n \to o)$$

for each pair of states o, n

Detailed balance (automatically implies balance)

Importance Sampling Random Walk

A move starting from one point consist of generating a **trial move** and accept or reject such a move.

Transition probabilities are a product of the **generation probability** and the **acceptance probability**

$$\pi(o \rightarrow n) = \alpha(o \rightarrow n) \times \mathrm{acc}(o \rightarrow n)$$

Detailed balance implies

$$\mathcal{N}(o)\alpha(o \rightarrow n) \times \mathrm{acc}(o \rightarrow n) =$$
$$\mathcal{N}(n)\alpha(n \rightarrow o) \times \mathrm{acc}(n \rightarrow o)$$

Metropolis algortihm

Generation probabilities are often chosen symmetric.

$$\alpha(o \rightarrow n) = \alpha(n \rightarrow o)$$

Therefore

$$\mathcal{N}(o) \times \text{acc}(o \rightarrow n) = \mathcal{N}(n) \times \text{acc}(n \rightarrow o)$$

$$\frac{\text{acc}(o \rightarrow n)}{\text{acc}(n \rightarrow o)} = \frac{\mathcal{N}(n)}{\mathcal{N}(o)} = e^{-\beta[U(n)-U(o)]}$$

the choice of Metropolis, Rosenbluth, Rosenbluth,Teller and Teller (1953)

$$\text{acc}(o \rightarrow n) = \min\left(1, e^{-\beta[U(n)-U(o)]}\right)$$

Importance Sampling Random Walk

A move starting from one point consist of generating a **trial move** and **accept or reject** such a move.

$$\mathrm{acc}\big(o \rightarrow n\big) = \min\left(1, e^{-\beta[U(n)-U(o)]}\right)$$

Metropolis, Rosenbluth, Rosenbluth,Teller and Teller (1953)

- try to change energy state
- compute $\Delta E = E_{new} - E_{old}$
- accept new state if ran < exp($\Delta E/kT$)
- reject otherwise
- sample the state of the system
- repeat
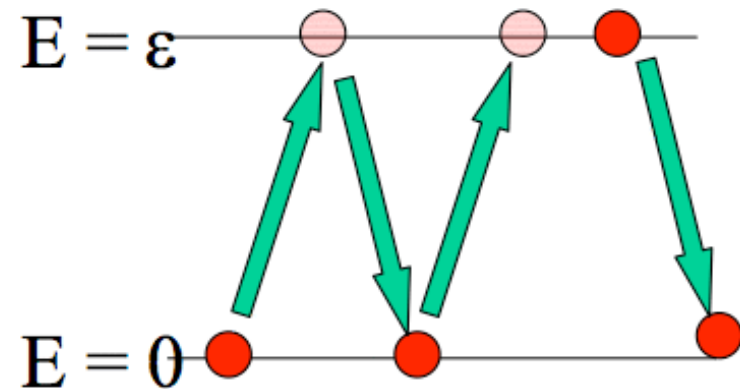
Parallel tempering/Replica Exchange

Ergodicity problems can occur, esspecially in glassy systems: biomolecules, moleuclar glasses, gels, etc.

The solution: go to high temperature

High barriers in energy landscape: difficult to sample

Barriers effectively low: easy to sample

low T

phase space

high T
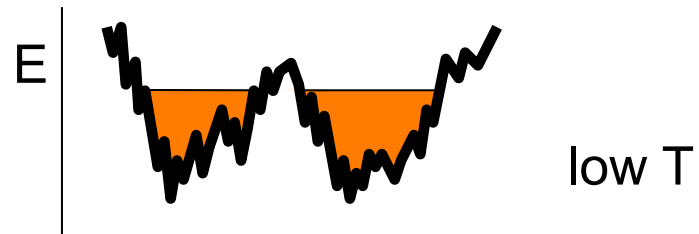
phase space
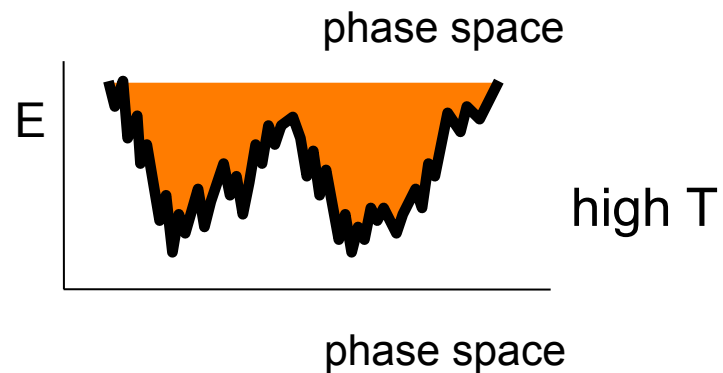
Parallel tempering/Replica Exchange

Simulate two systems simultaneously

system 1 temperature T_1	system 2 temperature T_2

$$e^{-\beta_1 U_1(r^N)} \qquad\qquad e^{-\beta_2 U_2(r^N)}$$

total Boltzmann weight:

$$e^{-\beta_1 U_1(r^N)} e^{-\beta_2 U_2(r^N)}$$

Swap move

- Allow two systems to swap

system 2 temperature T_1

system 1 temperature T_2

$$e^{-\beta_1 U_2(r^N)}$$

$$e^{-\beta_2 U_1(r^N)}$$

total Boltzmann weight:

$$e^{-\beta_1 U_2(r^N)} e^{-\beta_2 U_1(r^N)}$$

$$\mathrm{acc}(1 \leftrightarrow 2) = \min\left(1, e^{(\beta_2 - \beta_1)[U_2(r^N) - U_1(r^N)]}\right)$$

Acceptance rule

The ratio of the new boltzmann factor over the old one is

$$\frac{\mathcal{N}(n)}{\mathcal{N}(o)} = e^{(\beta_2 - \beta_1)[U_2(r^N) - U_1(r^N)]}$$

the swap acceptance ratio is

$$\mathrm{acc}(1 \leftrightarrow 2) = \min\left(1, e^{(\beta_2 - \beta_1)[U_2(r^N) - U_1(r^N)]}\right)$$

More replicas

Consider M replica's in the NVT ensemble at a different temperature.
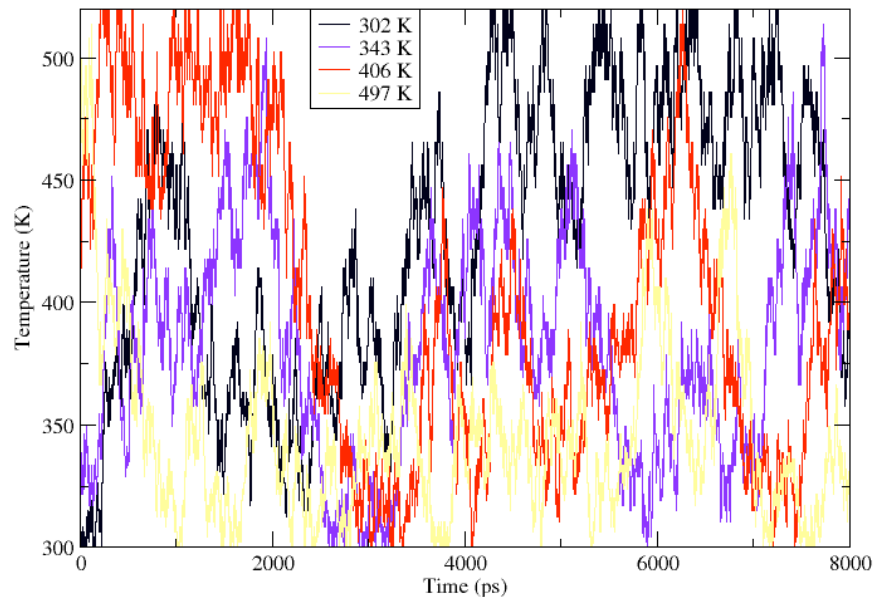
A swap between two systems of different temperatures (T_i, T_j) is accepted if their potential energies are near.

other parameters can be used: Hamiltonian exchange

Molecular dynamics

Is based on Newton's equations.

$$F_i = m_i a_i = m_i \frac{d^2 x_i(t)}{dt^2}$$

for i=1 .. N particles

the force F is given by the gradient of the potential

$$F_i = -\frac{\partial V(r^N)}{\partial r_i}$$

given the potential, one can integrate the trajectory *x(t)* of the whole system as a function of time.

Numerical integration

This is an N-body problem, which can only be solved numerically (except in very special cases)

$$x(t + \Delta t) = x(t) + \dot{x}(t)\Delta t + \frac{1}{2}\ddot{x}(t)\Delta t^2 + \frac{1}{6}\dddot{x}(t)\Delta t^3 + \ldots$$

at least, in principle..

Naïve implementation: truncation of Taylor expansion

$$x(t + \Delta t) = x(t) + \dot{x}(t)\Delta t + \frac{1}{2}\ddot{x}(t)\Delta t^2$$

Wrong!

The naive "forward Euler" algorithm

-           is not time reversible
-           does not conserve volume in phase space
-           suffers from energy drift

Better approach: "Verlet" algorithm

Verlet algorithm

compute position in next and previous time steps

$$x(t + \Delta t) = x(t) + \dot{x}(t)\Delta t + \frac{1}{2}\ddot{x}(t)\Delta t^2 + \frac{1}{6}\dddot{x}(t)\Delta t^3 + \frac{1}{24}\ddddot{x}(t)\Delta t^4 ...$$

$$x(t - \Delta t) = x(t) - \dot{x}(t)\Delta t + \frac{1}{2}\ddot{x}(t)\Delta t^2 - \frac{1}{6}\dddot{x}(t)\Delta t^3 + \frac{1}{24}\ddddot{x}(t)\Delta t^4 ...$$

+

$$x(t + \Delta t) + x(t - \Delta t) = 2x(t) + \ddot{x}(t)\Delta t^2 + \mathcal{O}(\Delta t^4)...$$

or

$$x(t + \Delta t) = 2x(t) - x(t - \Delta t) + \ddot{x}(t)\Delta t^2$$ **Verlet**

Verlet algorithm

- – is time reversible

- – does conserve volume in phase space

- – (is "symplectic")

- – does not suffer from energy drift
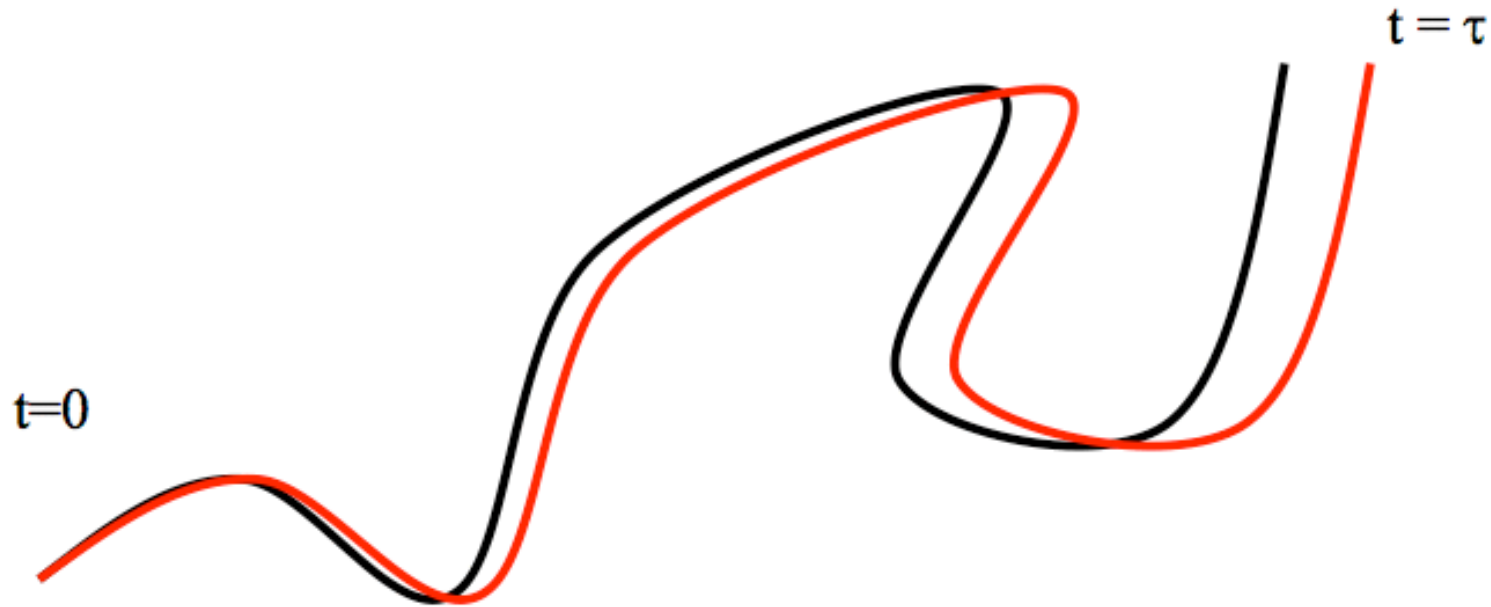
...but is it a good algorithm?

i.e. does it predict the time evolution of the system correctly???

Molecular chaos

Dynamics of "well-behaved" classical many-body system is chaotic.

Consequence: Trajectories that differ very slightly in their initial conditions diverge exponentially ("Lyapunov instability")

$t = \tau$

$t=0$

Lyapunov instability

The Lyapunov disaster in action...
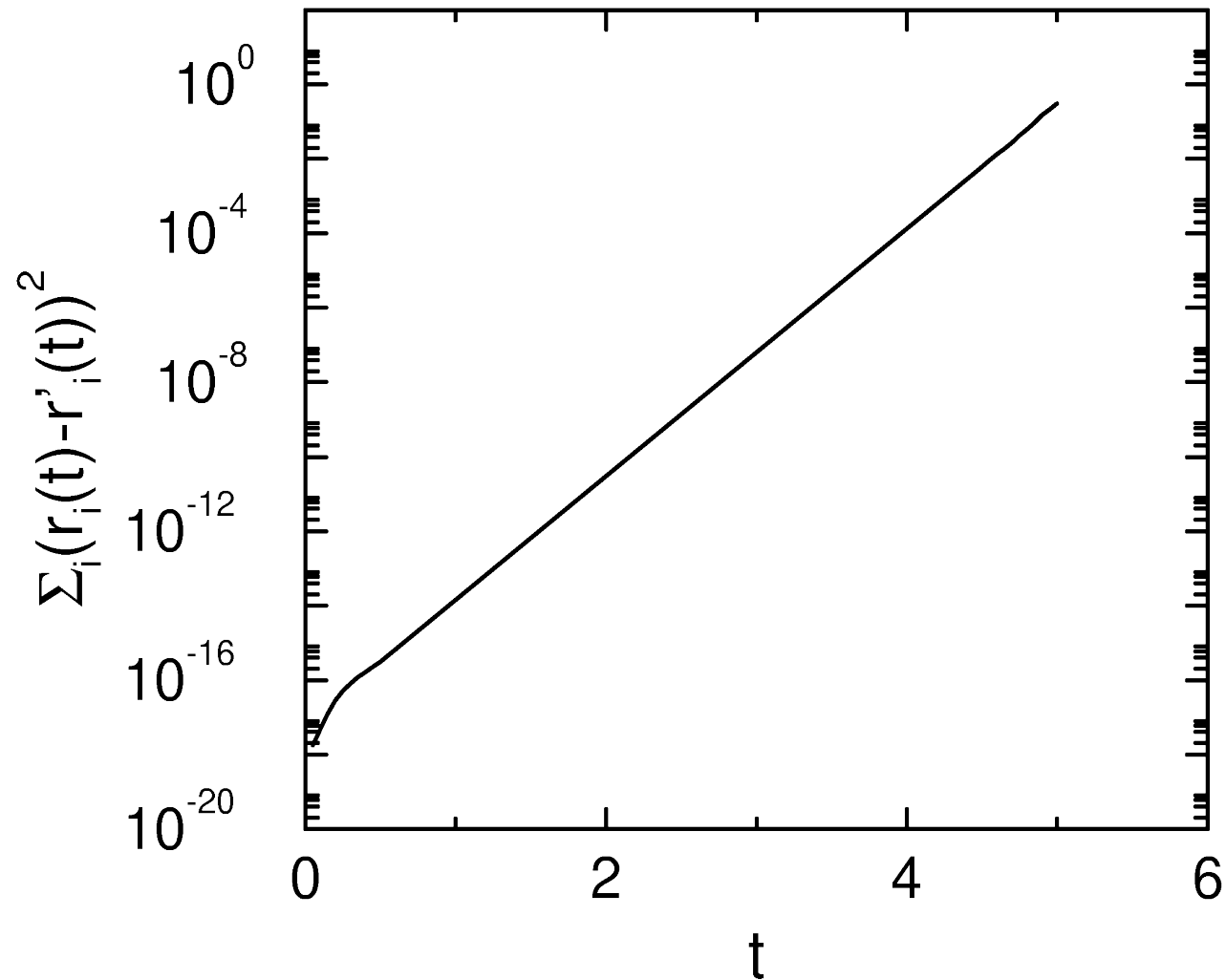
Any small error in the numerical integration of the equations of moti on, will blow up exponentially....

always...

...and for any algorithm!!

SO:

Why should anyone believe Molecular Dynamics simulations ???

Answers:

1. In fact, one should not…

2. Good MD algorithms (e.g. Verlet) can also be considered as good Monte Carlo algorithms –they therefore yield reliable STATIC properties ("Hybrid Monte Carlo")

3. What is the point of simulating dynamics, if we cannot trust the resulting time-evolution???

4. All is well (probably), because of...
   The Shadow Theorem.

Shadow theorem

- For any realistic many-body system, the shadow theorem is merely a hypothesis.

- It basically states that Good algorithms generate numerical trajectories that are "close to" a REAL trajectory of the many-body system.

- Question: Does the Verlet algorithm indeed generate "shadow" trajectories?

- Take a different look at the problem.
  - Do not discretize NEWTON's equation of motion…
  - ...but discretize the ACTION

Lagrangian Classical mechanics

- Newton:

$$F_i = m_i \frac{d^2 x_i(t)}{dt^2}$$

- Lagrange:
    - Consider a system that is at a point r_0 at time t=0 and at point r_t at time t=t, then the system follows a trajectory r(t) such that:

$$S \equiv \int_0^t dt' \mathcal{L}(r(t'))$$

is an extremum. The Lagrangianis L defined as:

$$\mathcal{L}(r(t)) = K - U(r)$$

kinetic energy

Langrangian

For example, if we use cartesian coordinates:

$$\mathcal{L}(r(t)) = \sum_{i=1}^{N} \frac{1}{2} m_i \dot{r}_i^2 - U(r_1, r_2, \ldots r_N)$$

What does this mean?

Consider the "true" path R(t), with R(0)=r_0 and R(t)=r_t.
Now, consider a path close to the true path:

$$r(t') = R(t') + \delta r(t')$$

Then the action S is an extremum if

$$\frac{\partial S}{\partial r(t')} = 0 \ \text{ for all } t$$

what does this mean?

Discretized action

$$S_{cont} = \int_{t_0}^{t_1} dt \mathcal{L}(t)$$

$$S_{disc} = \Delta t \sum_{i=0}^{i_{max}} \mathcal{L}(t_i) \qquad\qquad \mathcal{L}(t_i) = K(t_i) - U(t_i)$$

For a one dimensional system this becomes

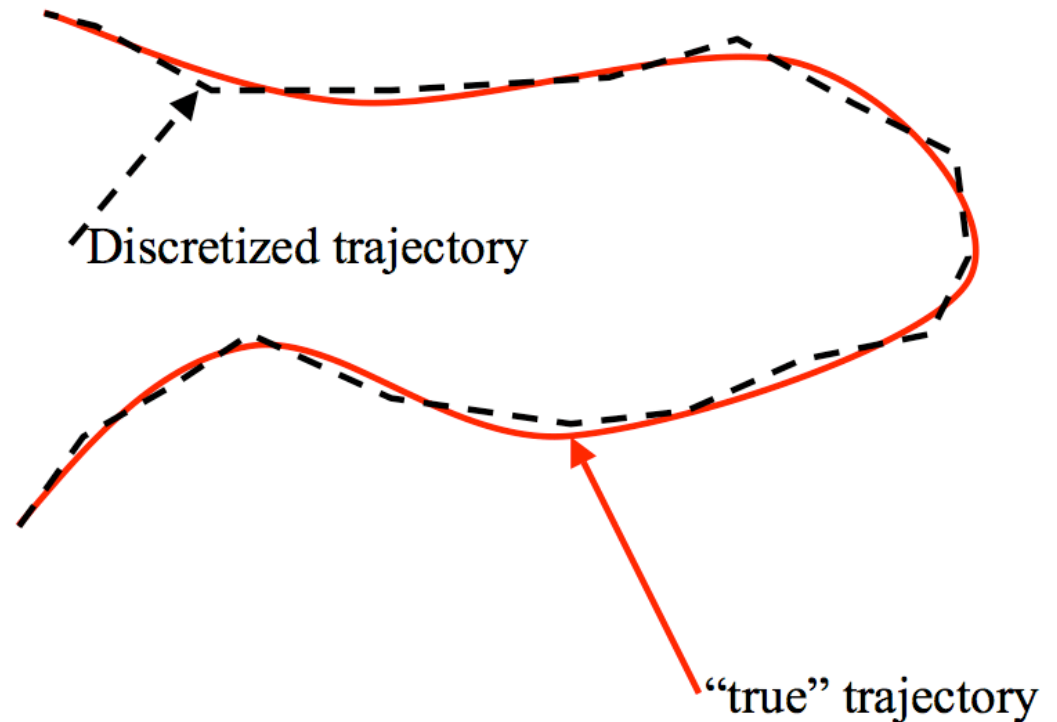$$\mathcal{L}(t_i)\Delta t = \frac{1}{2} m \Delta t \frac{(x_{i+1} - x_i)^2}{\Delta t^2} - U(x_i)\Delta t$$

$$S_{disc} = \sum_{i=1}^{i_{max}} \left[\frac{m(x_{i+1} - x_i)^2}{2\Delta t} - U(x_i)\Delta t \right]$$

Minimize the action

Now do the standard thing: Find the extremum for small variations in the path, i.e. for small variations in all x_i.

$$\frac{\partial S_{disc}}{\partial x_i} = 0 \ \text{ for all } i$$

This will generate a discretized trajectory that starts at time t_0 at X, and ends at time t at X_t.

Discretized trajectory

"true" trajectory

Minimizing the action

$$\frac{\partial S_{disc}}{\partial x_i} = \frac{\partial}{\partial x_i} \sum_{i=1}^{i_{max}} \left[\frac{m(x_{i+1} - x_i)^2}{2\Delta t} - U(x_i)\Delta t \right]$$

$$\frac{\partial S_{disc}}{\partial x_i} = \frac{-m(x_{i+1} - x_i) + m(x_i - x_{i-1})}{\Delta t} - \Delta t \frac{\partial U(x_i)}{\partial x_i}$$

$$0 = \frac{m}{\Delta t} \left(2x_i - x_{i+1} - x_{i-1} - \frac{\Delta t^2}{m} \frac{\partial U(x_i)}{\partial x_i} \right)$$

$$0 = 2x_i - x_{i+1} - x_{i-1} - \frac{\Delta t^2}{m} \frac{\partial U(x_i)}{\partial x_i}$$

$$x_{i+1} = 2x_i - x_{i-1} + \frac{\Delta t^2}{m} F(x_i)$$

- which is the Verlet algorithm!

- The Verlet algorithm generates a trajectory that satisfies the boundary conditions of a REAL trajectory –both at the beginning and at the endpoint.

- Hence, if we are interested in statistical information about the dynamics (e.g. time-correlation functions, transport coefficients, power spectra...)

  ...then a "good" MD algorithm (e.g. Verlet) is fine.

Velocity Verlet

Downside regular verlet algorithm: velocity is not known.

Velocity verlet (Andersen 1983):

$$\mathbf{r}(t + \Delta t) \approx \mathbf{r}(t) + \mathbf{v}(t)\Delta t + \frac{\Delta t^2}{2m}\mathbf{f}(t)$$

$$\mathbf{v}(t + \Delta t) \approx \mathbf{v}(t) + \frac{\Delta t}{2m}\left[\mathbf{f}(t + \Delta t) + \mathbf{f}(t)\right]$$

(Is based on Trotter decomposition of Liouville operator formulation, also basis of Multiple time steps).

Lagrangian approach

Lagrangian is sum of two terms

$$\mathcal{L}(\dot{r}, r) = K(\dot{r}) - U(r) = \frac{m\dot{r}^2}{2} + U(r)$$

$$\frac{\partial \mathcal{L}}{\partial \dot{r}} = \frac{\partial K}{\partial \dot{r}} = p \qquad\qquad \frac{\partial \mathcal{L}}{\partial r} = -\frac{\partial U}{\partial r} = F$$

$$p = \frac{\partial \mathcal{L}(\dot{r}, r)}{\partial \dot{r}} \qquad\qquad \dot{p} = \frac{\partial \mathcal{L}(\dot{r}, r)}{\partial r}$$

Newton : F=ma

Hamiltonian approach

The Hamiltonian is defined as

$$H(p, r) = p\dot{r} - \mathcal{L}(\dot{r}, r)$$

$$H(p^N, r^N) = U(r^N) + \sum_i \frac{p_i^2}{2m_i}$$

Hamilton's equations are then

$$\dot{r} = \frac{\partial \mathcal{H}(r, p)}{\partial p} = \frac{p}{m}$$

$$\dot{p} = -\frac{\partial \mathcal{H}(r, p)}{\partial r} = -\frac{\partial U(r^N)}{\partial r}$$

Integrating equations of motion (by Verlet) conserves the Hamiltonian

Conservation of Hamiltonian

$$dH(p,r) = \frac{\partial H}{\partial p}\,dp + \frac{\partial H}{\partial r}\,dr$$

$$\frac{\partial H}{\partial p} = \dot{r} \qquad \frac{\partial H}{\partial r} = -\dot{p}$$

$$\frac{dH(p,r)}{dt} = \frac{\partial H}{\partial p}\,\dot{p} + \frac{\partial H}{\partial r}\,\dot{r} = \dot{r}\dot{p} - \dot{p}\dot{r} = 0$$

So a solution to the Hamiltonians equation conserves the TOTAL energy

$$E = K + U$$

MD generates NVE ensemble

In general the MC phase space density is

$$\rho(x) = e^{-\beta \mathcal{H}(x)} / Z \qquad\qquad Z = \int e^{-\beta \mathcal{H}(x)} dx$$

with $x = \{p^N, r^N\}$

Integrating over momenta gives

$$Z = \frac{1}{N! \Lambda^{3N}} \int e^{-\beta \mathcal{U}(r)} dr$$

N! comes from indistinguishability of particles.

But MD conserves Hamiltonian H= E = constant (and constant total P).

$$\rho(x) = \delta[E - \mathcal{H}(x)] / g(E) \qquad g(E) = \int dx \, \delta[E - \mathcal{H}(x)]$$

with instantaneous temperature $k_B T = \sum_{i=1}^{N} \frac{m v_i^2}{N_f}$

Thermostat: From NVE to NVT

Introduce thermostat in MD trajectory:

- stochastic thermostats
  - **Andersen**
  - Langevin
  - Bussi (2007)

- deterministic thermostat
  - **Nose-Hoover**

All of these alter the velocities such that the trajectory samples the canonical NVT ensemble, and the partition function becomes

$$Z = \frac{1}{N!\Lambda^{3N}} \int e^{-\beta \mathcal{U}(r)} dr$$

These thermostats differ in how they achieve this

Andersen Thermostat

- Every particle has a fixed probability to collide with the Andersen demon

- After collision the particle is give a new velocity

$$P(v) = \left(\frac{\beta}{2\pi m}\right)^{3/2} \exp\left[-\beta m v^2 / 2\right]$$
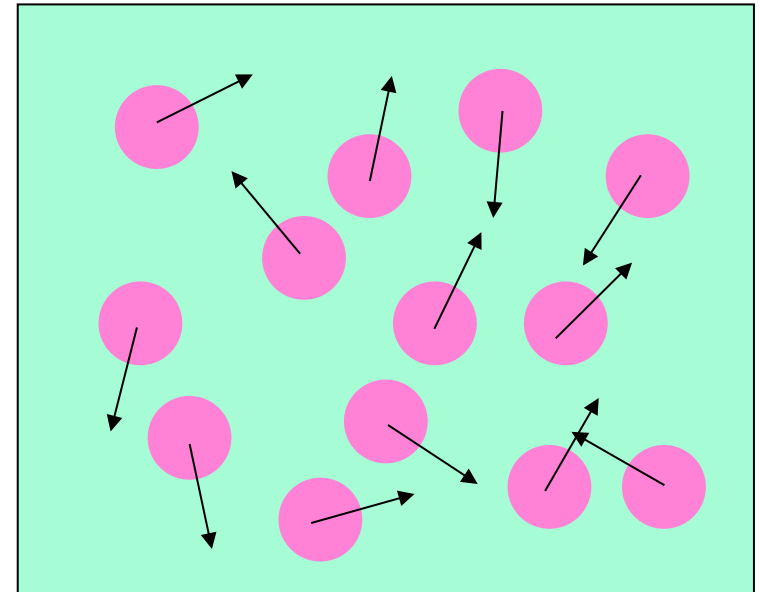
- The probabilities to collide are uncorrelated (Poisson distribution)

$$P(t;v) = v \exp\left[-vt\right]$$

Nose Hoover thermostat

goal: compute MD trajectory sampling NVT ensemble.

Take kinetic energy out of the system and put it back in via a 'piston'.

piston can be seen as additional variable *s* storing kinetic energy

Approach: extended Lagrangian

extended variable

$$L_{\text{Nose}} = \sum_{i=1}^{N} \frac{1}{2} m s^2 \dot{r}_i^2 - U\left(r^N\right) + \frac{1}{2} Q \dot{s}^2 - \frac{g}{\beta} \ln s$$

effective mass

constant to be set

$$p_i = \frac{\partial L}{\partial \dot{r}_i} = m s^2 \dot{r}_i \qquad p_s = \frac{\partial L}{\partial \dot{s}} = Q \dot{s}$$

$$H_{\text{Nose}} = \sum_{i=1}^{N} \frac{p_i^2}{2 m s^2} + \frac{p_s}{2Q} + U\left(r^N\right) + \frac{g}{\beta} \ln s$$
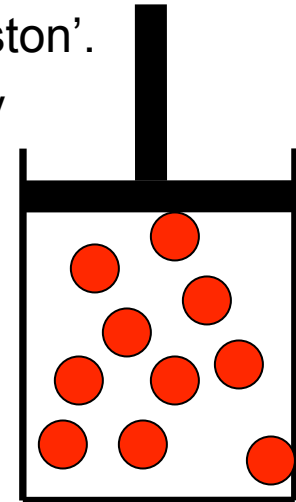
Nose-Hoover Thermostat

$$H_{\text{Nose}} = \sum_{i=1}^{N} \frac{p_i^2}{2ms^2} + \frac{p_s}{2Q} + U\left(r^N\right) + \frac{g}{\beta}\ln s$$

now define

$$\mathcal{H}(p', r) = \sum_{i=1}^{N} \frac{p_i'^2}{2m_i} + U(r^N) \qquad\qquad p' = p/s$$

then it is possible to show that the partition function Z_{nose} is

$$Z_{nose} \propto \frac{1}{N!} \int dp'^N \, dr^N \exp\left[-\beta\frac{3N+1}{g}\mathcal{H}(p', r)\right]$$

for g=3N+1 the system samples the canonical distribution if p' is interpreted as the real momentum

Nose-Hoover Thermostat

$$H_{\text{Nose}} = \sum_{i=1}^{N} \frac{p_i^2}{2ms^2} + \frac{p_s}{2Q} + U\left(r^N\right) + \frac{g}{\beta} \ln s$$

equations of motion follow from Hamilton's equations.

$$\frac{dr_i}{dt} = \frac{\partial H_{\text{Nose}}}{\partial p_i} = \frac{p_i}{ms^2} \qquad \frac{dp_i}{dt} = -\frac{\partial H_{\text{Nose}}}{\partial r_i} = -\frac{\partial U\left(r^N\right)}{\partial r_i}$$

$$\frac{ds}{dt} = \frac{\partial H_{\text{Nose}}}{\partial p_s} = \frac{p_s}{Q} \qquad \frac{dp_s}{dt} = -\frac{\partial H_{\text{Nose}}}{\partial s} = \frac{1}{s}\left(\sum \frac{p_i^2}{ms^2} - \frac{g}{\beta}\right)$$

Effect of mass Q

Lennard-Jones fluid
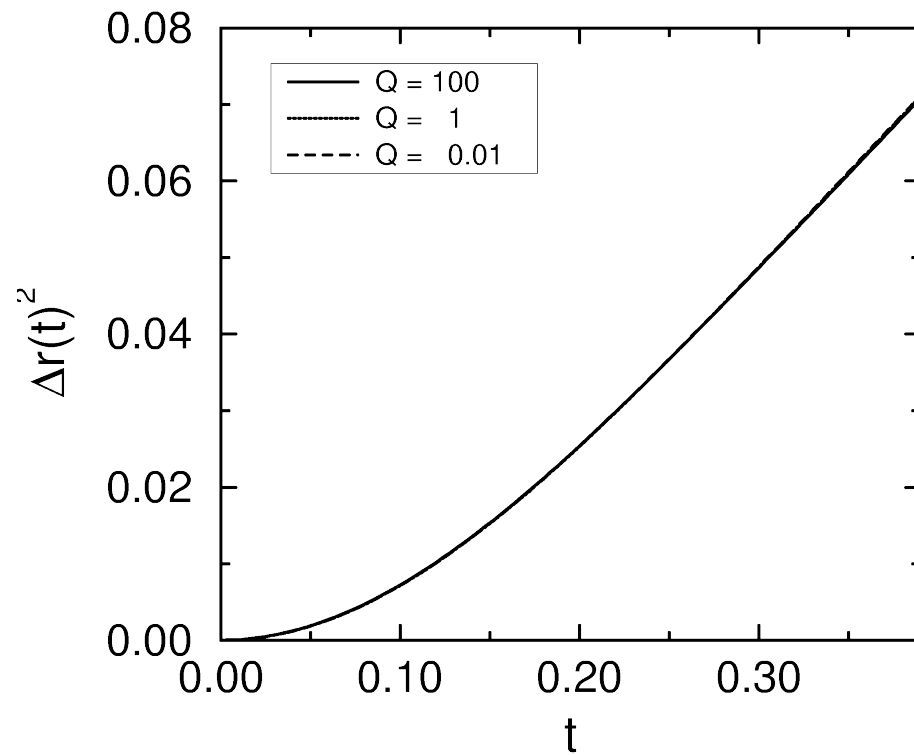
mean square displacement
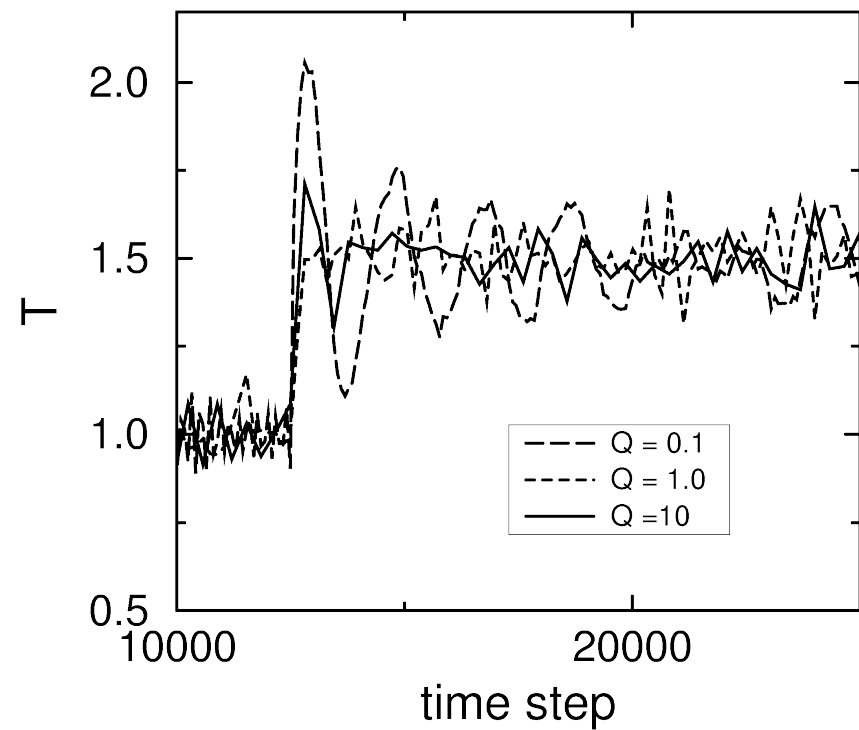
temperature relaxation

Ergodicity theorem

time averages over a NVT MD trajectory

$$\bar{A} = \frac{1}{\mathcal{T}} \int_0^{\mathcal{T}} A(t)dt$$

ensemble average

$$\langle A \rangle = \frac{\int dr^N A(r^N) \exp(-\beta \mathcal{U}(r^N))}{\int dr^N \exp(-\beta \mathcal{U}(r^N))}$$

Ergodicity theorem states that for an 'ergodic system'

$$\bar{A} = \langle A \rangle$$

Practical MD

- Initialization
  - Total momentum should be zero (no external forces)
  - Temperature rescaling to desired temperature
  - Particles/atoms/molecules start on a lattice/ or random positions

- Force calculations
  - Periodic boundary conditions
  - Straightforward force: Order N^2 algorithm:
  - neighbor lists, linked cell: Order N:
  - Electrostatics: Ewald summation $O(N^{1.5})$ or PME: $O(N\log N)$

- Integrating the equations of motion
  - Verlet or velocity verlet
  - Thermostat
  - optionally : multiple time steps

Periodic boundaries
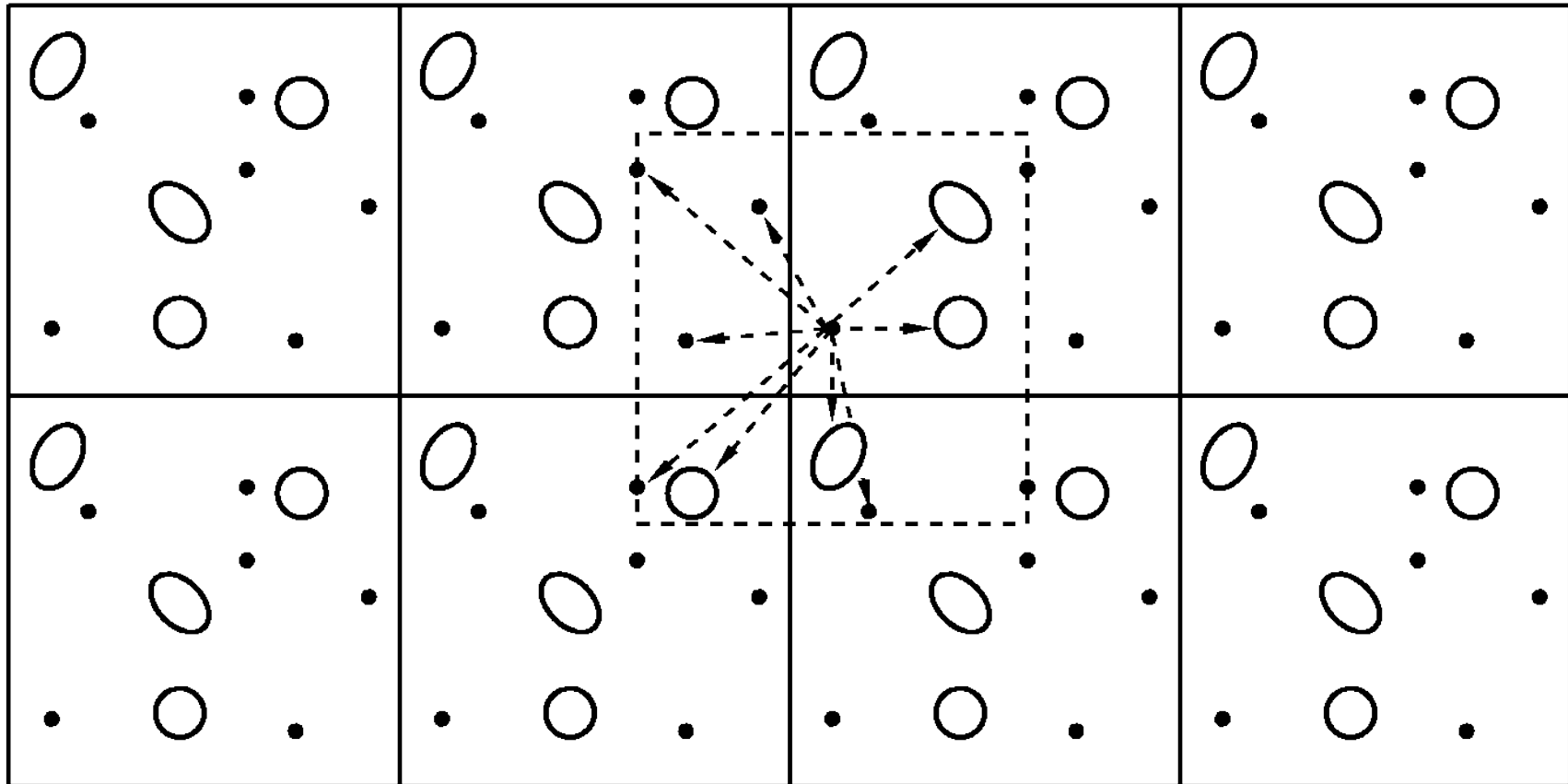
Saving CPU

- Cell list

Verlet List

r_C

r_V

r_C

i

i
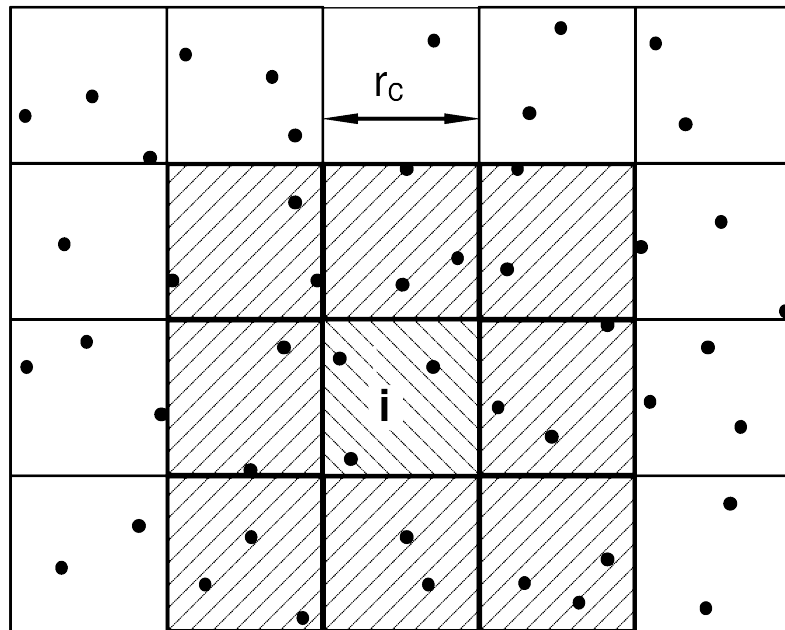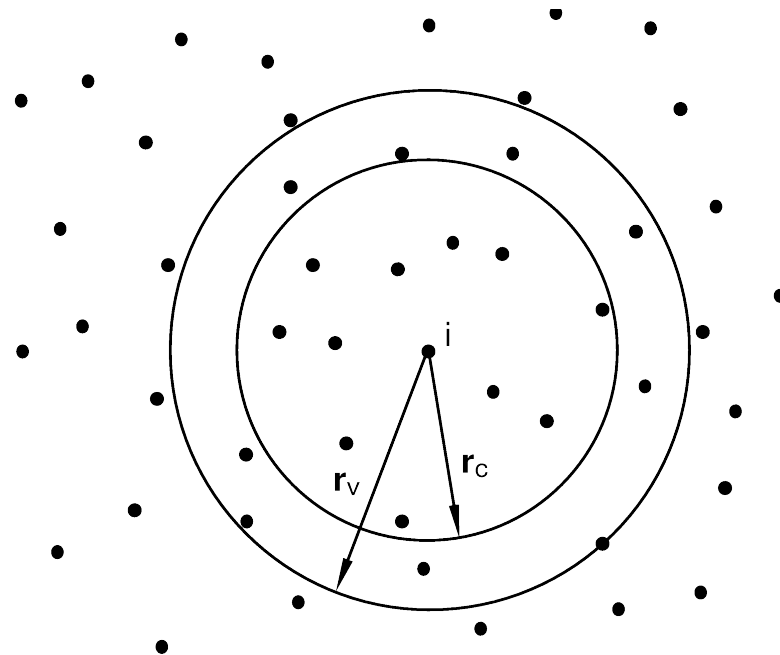
All-atom force fields for biomolecules

- Potential energy for protein

$$V(\mathbf{r}) = \sum_{bonds} k_r (r - r_{eq})^2 + \sum_{angles} k_\theta (\theta - \theta_{eq})^2 + \sum_{dihedrals} \frac{1}{2} v_n (1 + \cos(n\phi - \phi_0))$$

$$+ \sum_{i<j} \left(\frac{a_{ij}}{r_{ij}^{12}} - \frac{b_{ij}}{r_{ij}^6} + \frac{q_i q_j}{\varepsilon r_{ij}} \right)$$
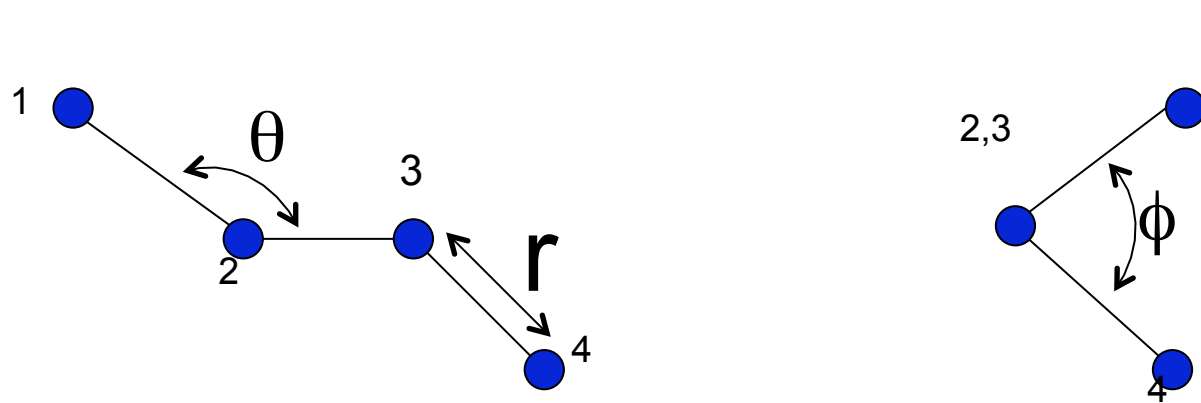
vdW interactions only between non-bonded |i-j|>4

Currently available empirical force fields

- CHARMm   (MacKerrel et 96)
- AMBER   (Cornell et al. 95)
- GROMOS   (Berendsen et al 87)
- OPLS-AA   (Jorgensen et al 95)
- ENCAD   (Levitt et al 83)

- Subtle differences in improper torsions, scale factors 1-4 bonds, united atom rep.
- Partial charges based on empirical fits to small molecular systems
- Amber & Charmm also include ab-initio calculations
- Not clear which FF is best : top 4 mostly used

- Water models also included in description
  - TIP3P, TIP4P
  - SPC/E
- Current limit: 10^5 atoms, 1000 ns

Photoactive Yellow Protein
bacterial blue-light sensor

N-terminal domain   PAS domain
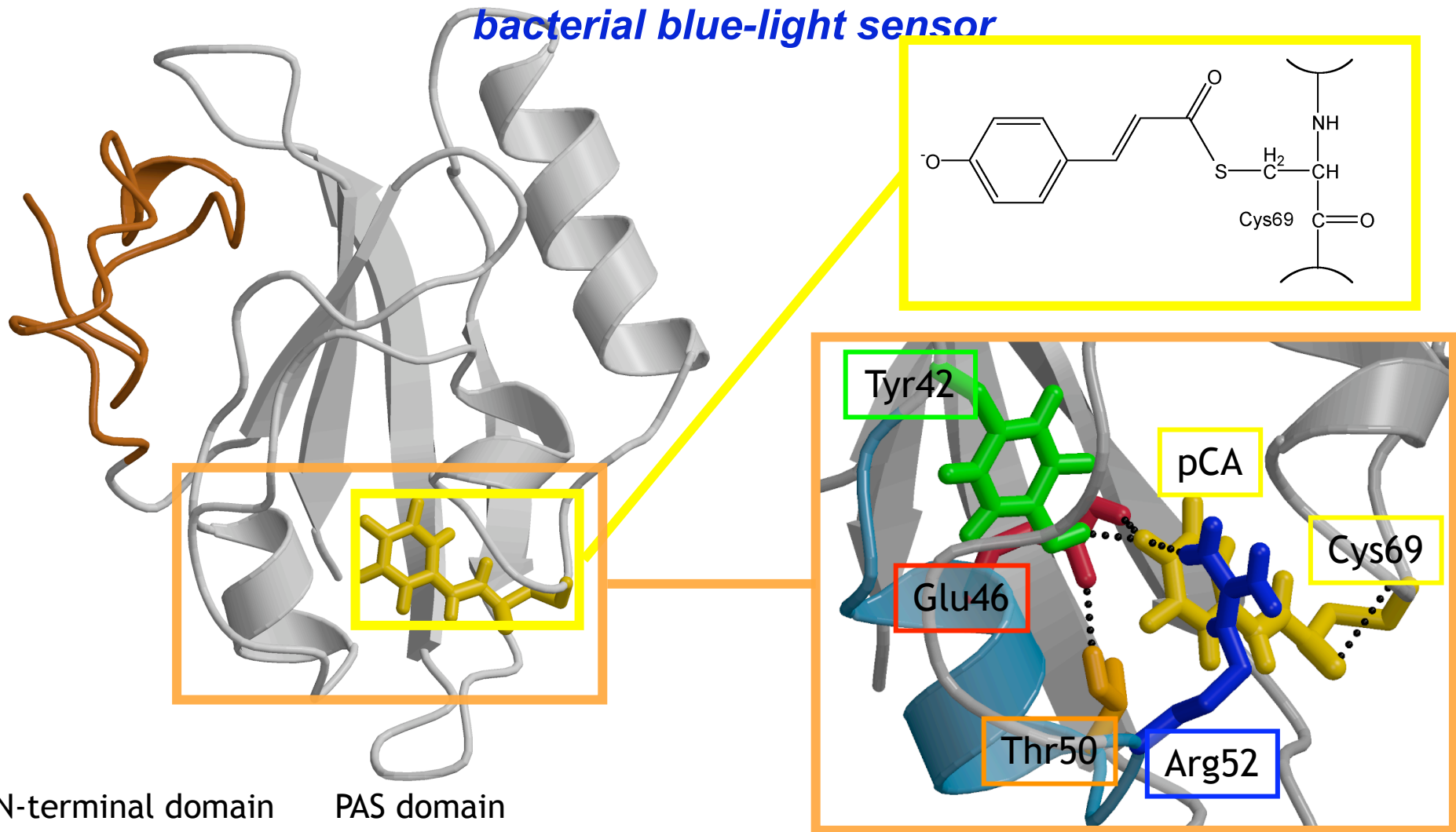
Absorption of a blue-light photon triggers the photo cycle

J. Vreede et al. Biophys. J. 2005

pG

pG*

0.7s

0.1 ms

1 ps

deprot

trans

pB

pB

0.2 ns

pR

0.3 s

10 μs

pB*

pB
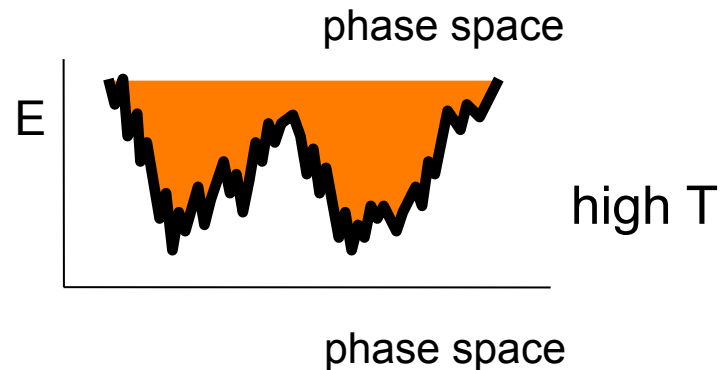
hv

pR'

1 ms

0.3 μs

molecular simulation
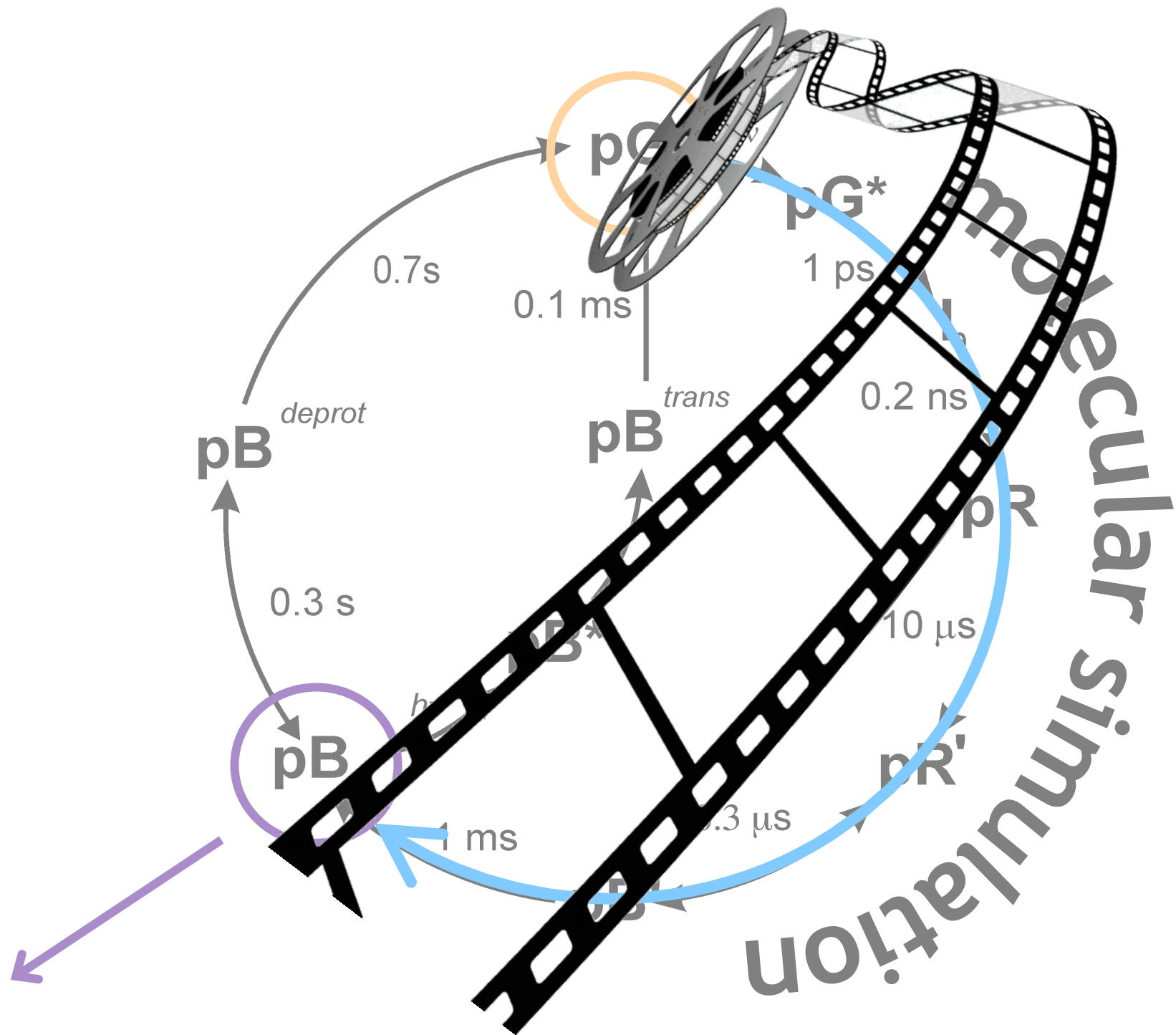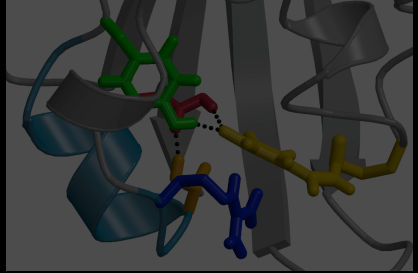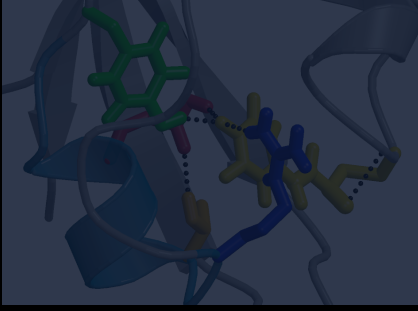
n
s

μ
s

m
s

Formation of signaling state

S—Cys69

$^-$O

O

Glu46 O OH

$^-$O

S—Cys69

O
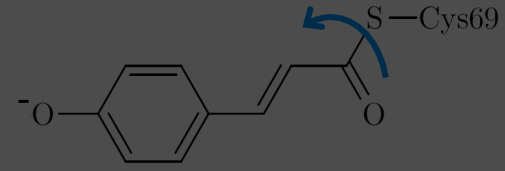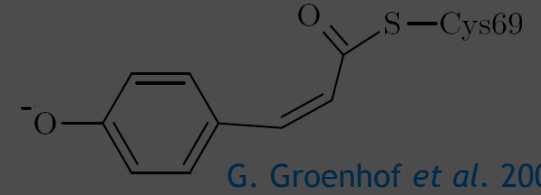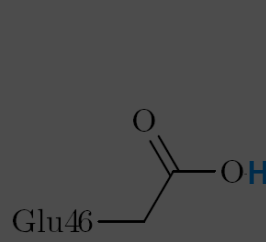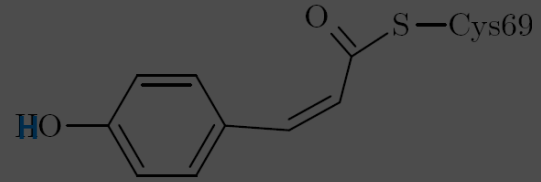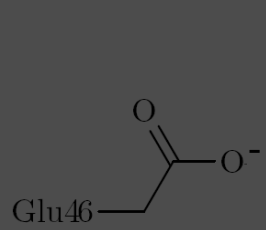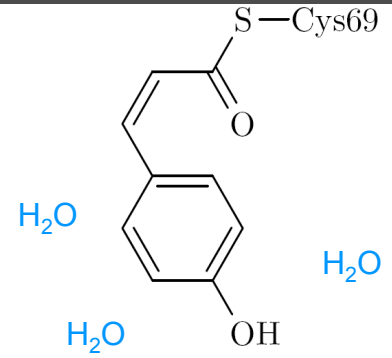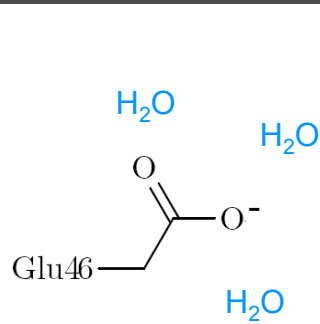
G. Groenhof *et al.* 2005, 2008

Glu46 O O$^-$

HO

O

S—Cys69

S—Cys69

H₂O

H₂O

Glu46 O O$^-$

H₂O

H₂O

H₂O

O

H₂O

H₂O

OH

Replica Exchange works for MD

High barriers in energy
landscape: difficult to sample

E | low T

phase space

Barriers effectively low: easy to
sample

E | high T
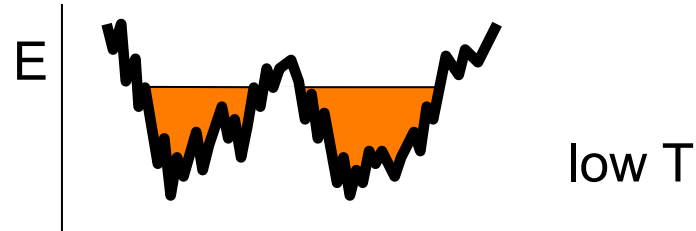
phase space

Consider M replica's in the NVT ensemble at a different temperature.
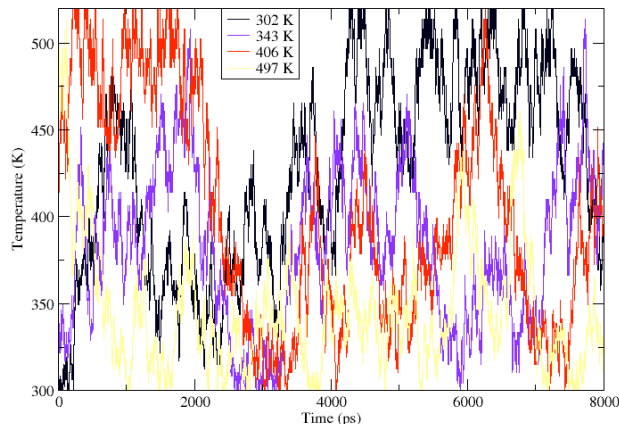
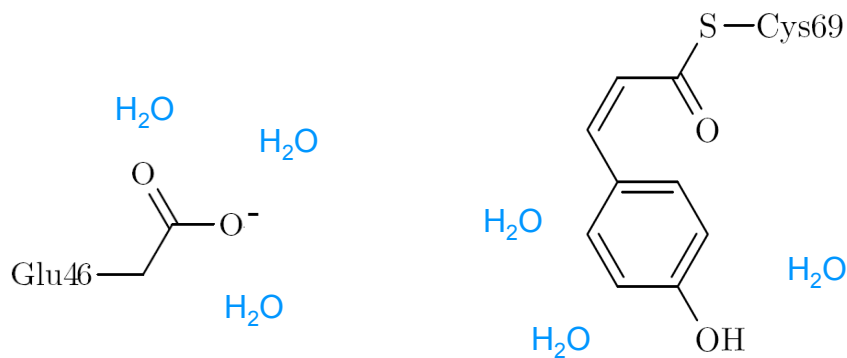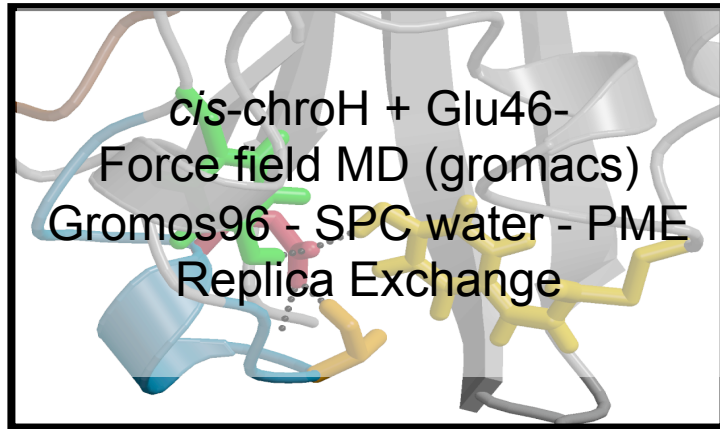$$\frac{i \rightarrow j}{j \rightarrow i} = \exp-\left[(\beta_i - \beta_j)(U(j) - U(i))\right]$$

A swap between two systems of different temperatures (T_i, T_j) is accepted if their potential energies overlap.
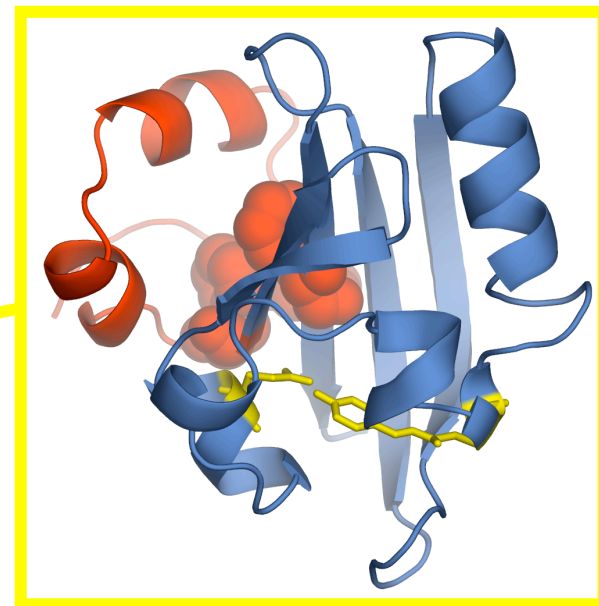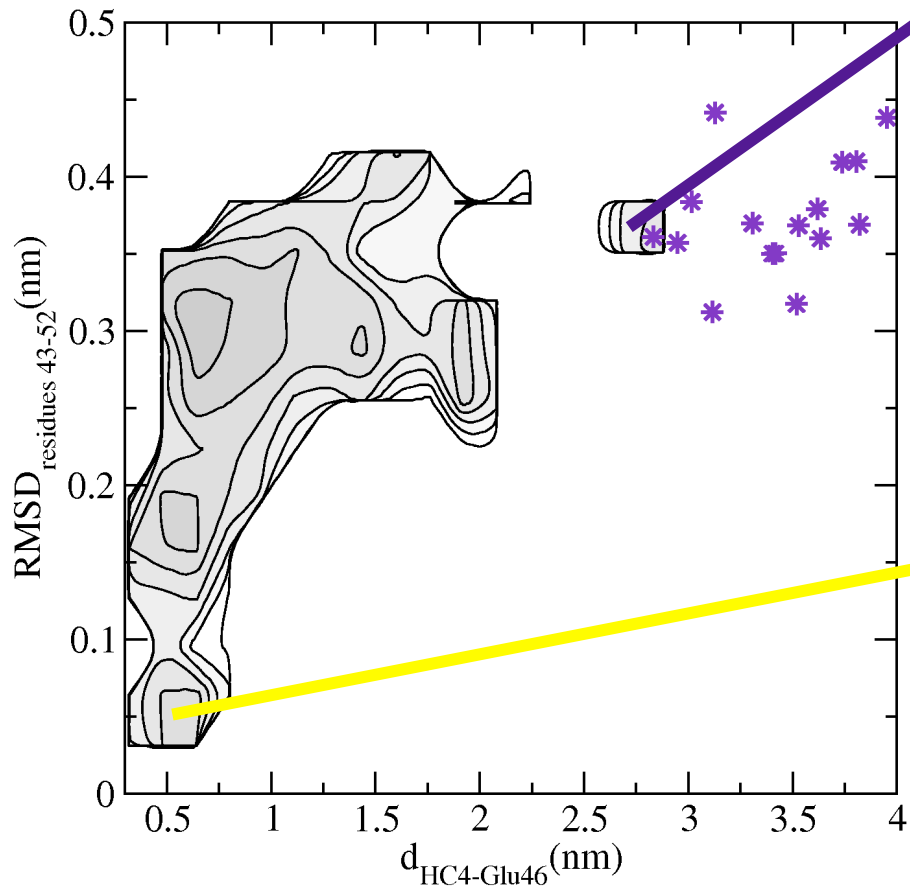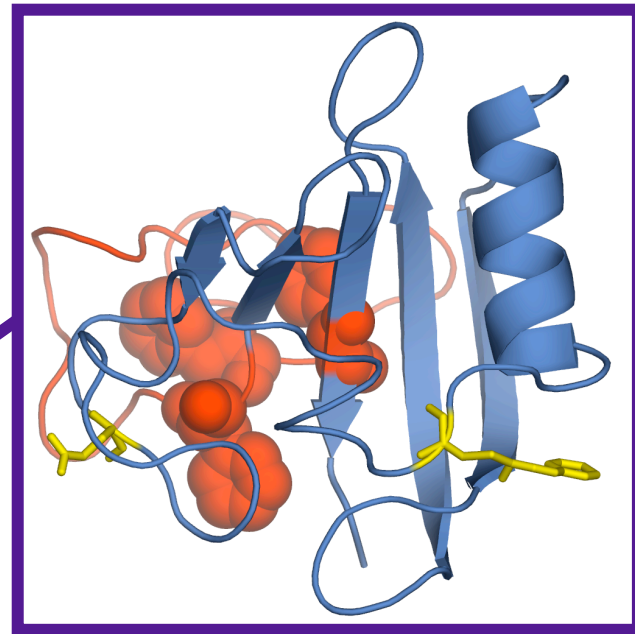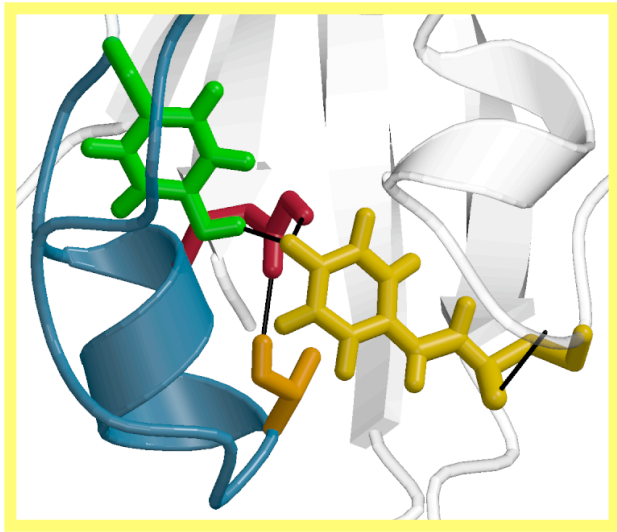
Advantage: better sampling phase space

Partial unfolding

- Loss of α-helical structure
- Exposure of hydrophobic groups
- Increased flexibility in parts of the protein backbone

Molecular simulation

cis-chroH + Glu46-
Force field MD (gromacs)
Gromos96 - SPC water - PME
Replica Exchange

$d_{HC4-Glu46}$(nm)

RMSD$_{residues\ 43-52}$(nm)
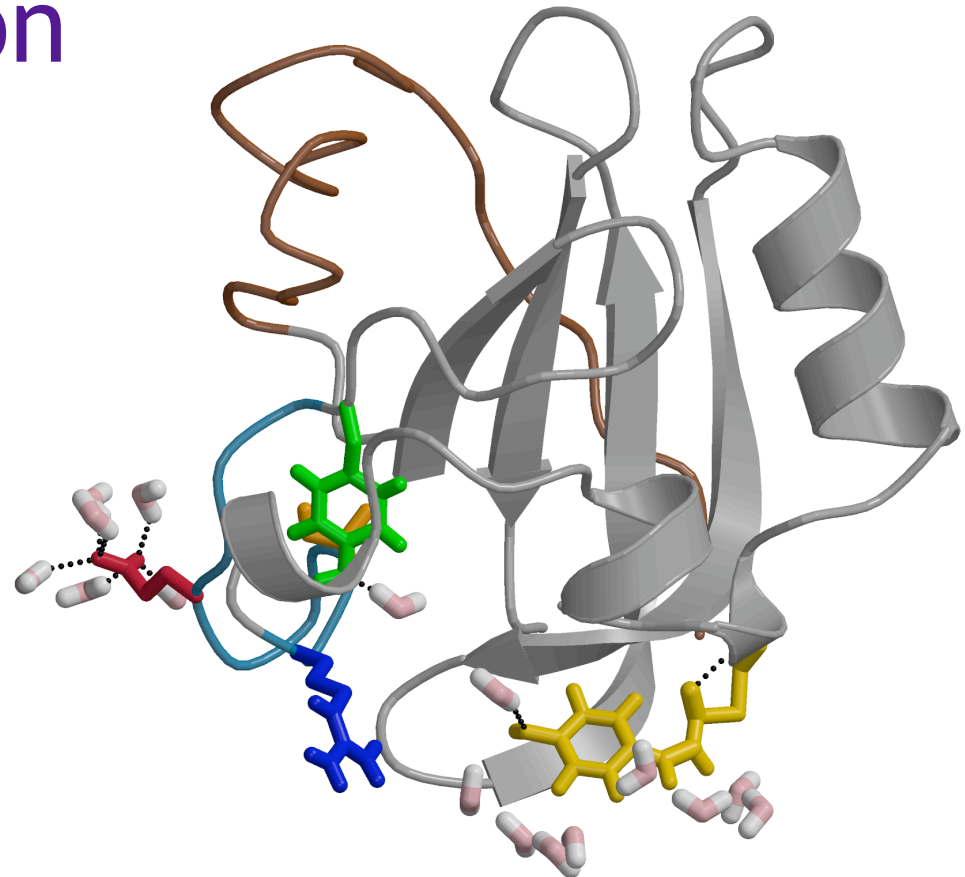
Exp: Bernard et al. Structure 2005

A REMD trajectory