Biomolecular modeling IIII

Marcus Elstner and Tomáš Kubař

2017, December 13

Coarse-grained models

United-atom force fields

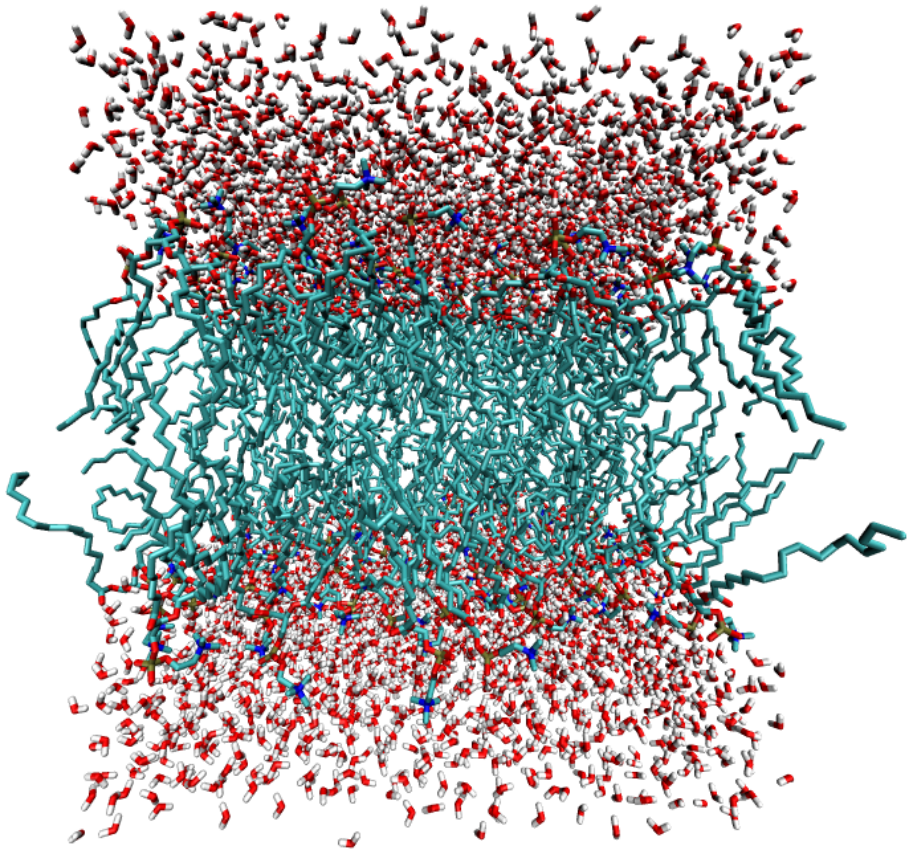
Early biomolecular force fields (e.g. Weiner 1984)
– united-atom approach
– hydrogen atoms considered as condensed to the heavy atom
– mass and charge represent such a group of atoms as a whole
– number of atoms reduced considerably relative to all-atom FF
– popular in the 1990's

This approach works very well for non-polar C–H bonds,
so a methyl group constituting of one united atom works good.

A substitution of a polar O–H group by a single particle
would be very crude (without any correction terms in FF)
$\rightarrow$ only non-polar hydrogens are usually condensed with heavy

United-atom force fields

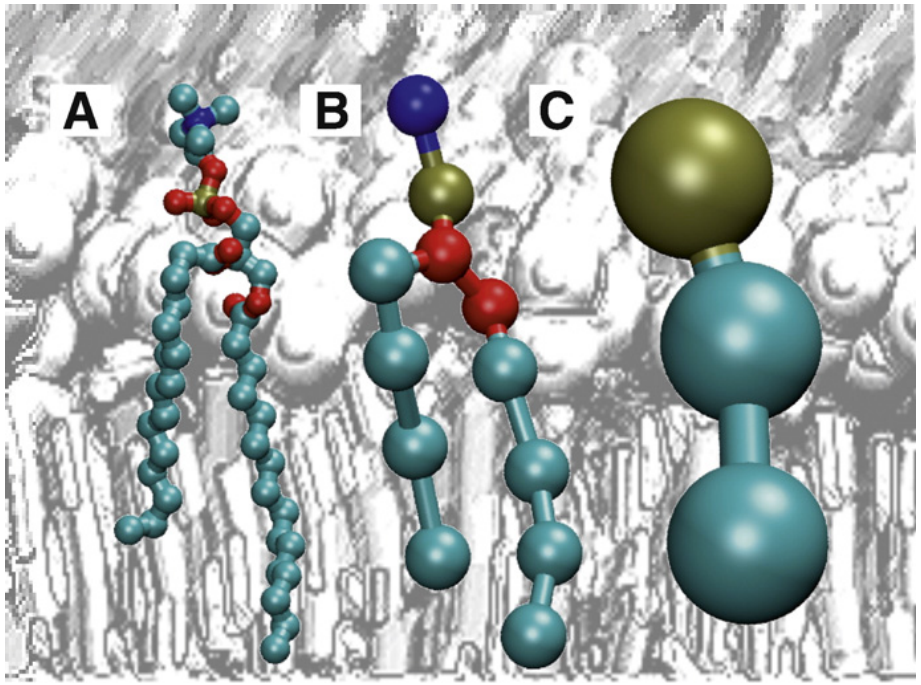
– still used e.g. to describe lipids, where each CH_2 is a united atom



– simulation of a DOPC bilayer in water – Berger FF for the lipid

from the website of Rainer Böckmann

United-atom and coarse-grained force fields



(A) united-atom, (B) specific and (C) generic coarse-grained
from Marrink et al., Biochim. Biophys. Acta 2009

Coarse-grained models

Coarse graining – an advanced and sophisticated approach
to reduce the computational expense of simulations

The same idea – reduction of the number of particles
Considered are particles composed of several atoms – beads
The number of inter-particle interactions decreases,
reducing the computational expense largely.

The necessary parameters of the force field are often obtained
by fitting to all-atom force fields.

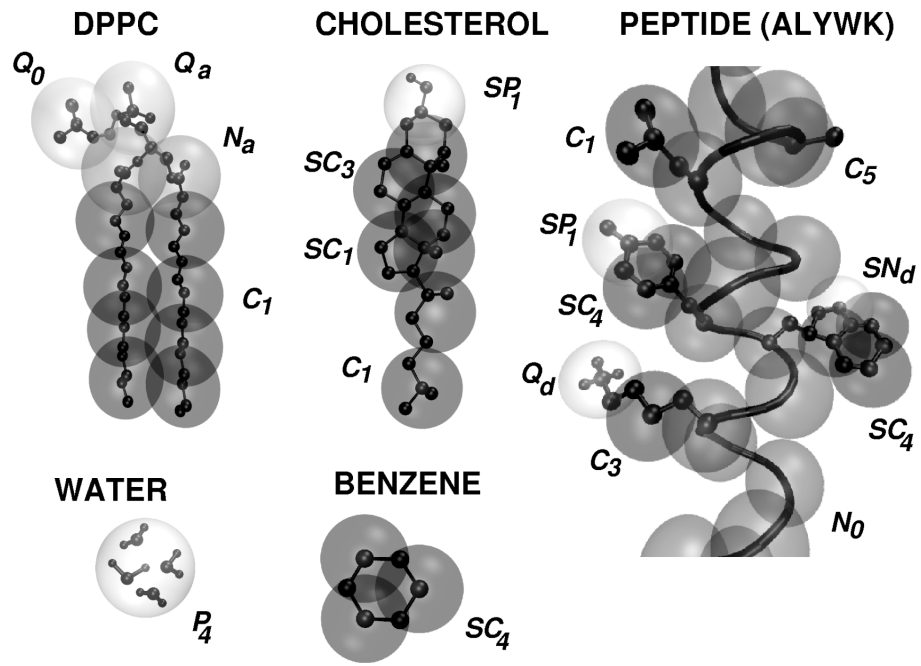
Coarse-grained models

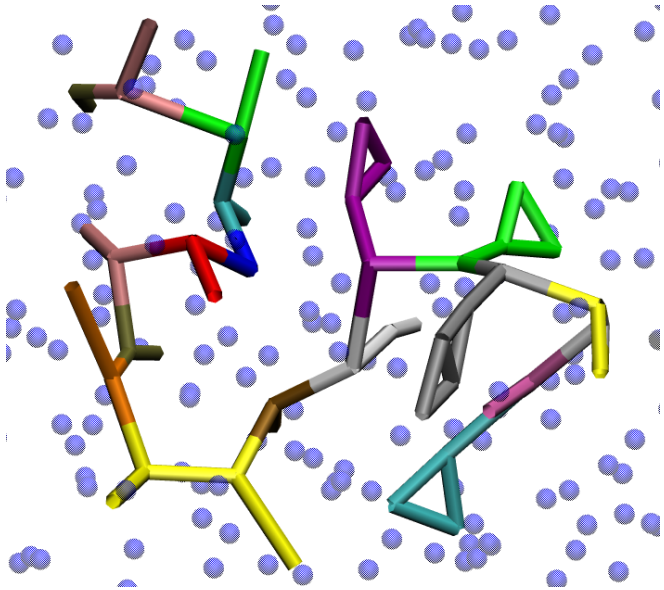
Every bead usually represents several atoms,
and a molecule is composed of several beads.
For the solvent, there is e.g. a 'water bead'
composed of four H_2O molecules.

Note that some of the transferability of all-atom FF is lost
– e.g. secondary structure of proteins is fixed with Martini FF
Also, hydrogen bonding cannot be described with beads!
solution – compensation with Lennard-Jones contributions

Such CG force fields are particularly useful for simulations
of large-scale conformational transitions, which involve
exceedingly large molecular systems, excessive time scales,
or both.

Martini force field





left – mapping of beads onto molecular fragments with Martini FF

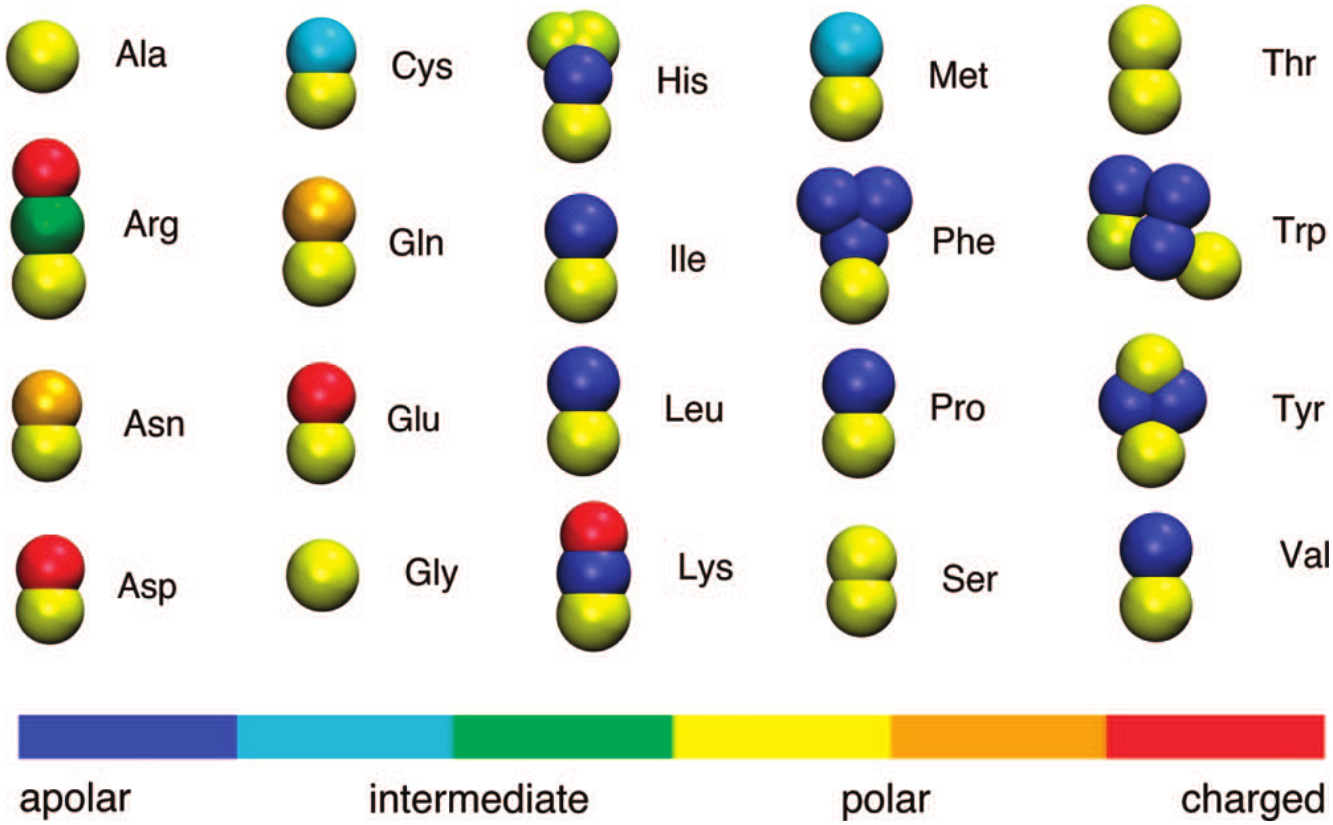
– 3 to 4 heavy atoms compose one bead ('4-to-1 mapping')

– mass of beads – 72 u (= 4 H_2O), or 45 u in ring structures

right – a solvated peptide with Martini

from the Martini website

Martini force field



The CG force field Martini – amino acids

from Monticelli et al., J. Chem. Theory Comput. 2008

Acceleration of the simulation

Why does a coarse-grained simulation run faster?

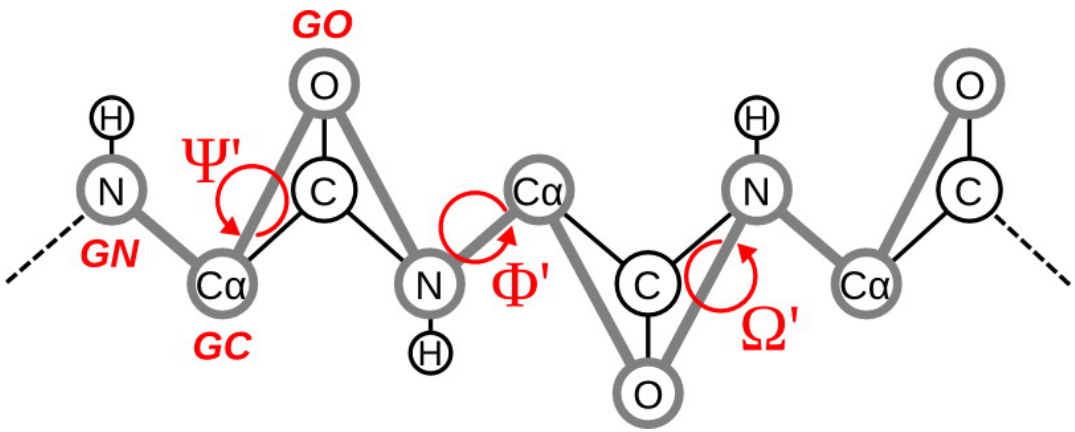
- smaller number of particles $\rightarrow$ fewer interactions
- long integration time step due to large masses of beads
  – 25 fs with Martini (i.e. 100 fs effectively, see below)
- FF often constructed for use with faster simulation algorithms
  – e.g. cut-off for electrostatics with Martini
- smaller number of DoF $\rightarrow$ smoother free energy surfaces
  $\rightarrow$ fewer barriers $\rightarrow$ acceleration of all processes
  (by a factor of 3 to 8 for Martini, but not uniformly!
  – factor of 4 for acceleration of diffusion in water)

"... length and time scales that are 2 to 3 orders of magnitude larger compared to atomistic simulations, providing a bridge between the atomistic and the mesoscopic scale."

Coarse-grained models

SIRAH force field

- somewhat less coarse-grained, closer to united-atom
- representation of backbone dihedral angles retained

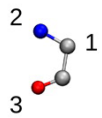
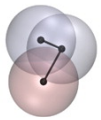
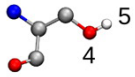
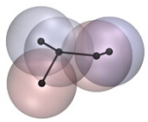


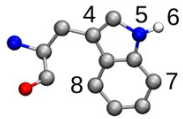
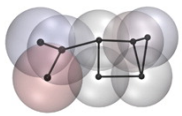
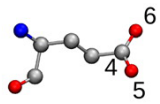
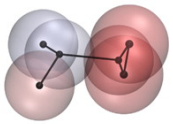
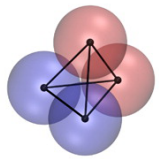
from Pantano et al., J. Chem. Theory. Comput. 2015

Coarse-grained models

SIRAH force field

- less coarse-grained → possibly improved transferability
- explicit solvent, long-range electrostatics (no cut-off)

FG	CG	SIRAH name	q (e)	σ (nm)	ε (kJ/mol)
G 		1: GC	0,10	0,40	0,55
		2: GN	0,125	0,40	0,55
		3: GO	-0,225	0,40	0,55
S 		4: BOG	-0,20	0,41	0,35
		5: BPG	0,20	0,40	0,01
Na^+ and 6 water molecules		1: NaW	1,00	0,58	0,55

FG	CG	SIRAH name	q (e)	σ (nm)	ε (kJ/mol)
W 		4: BCG	0	0,35	1,70
		5: BNE	-0,10	0,35	0,10
		6: BPE	0,10	0,35	0,01
		7: BCZ	0	0,35	1,70
		8: BCE	0	0,35	1,70
E 		4: BCD	-0,30	0,40	0,35
		5: BOE1	-0,35	0,45	0,55
		6: BOE2	-0,35	0,45	0,55
WT4 11 water molecules		1: WN1	-0,41	0,42	0,55
		2: WN2	-0,41	0,42	0,55
		3: WP1	0,41	0,42	0,55
		4: WP2	0,41	0,42	0,55

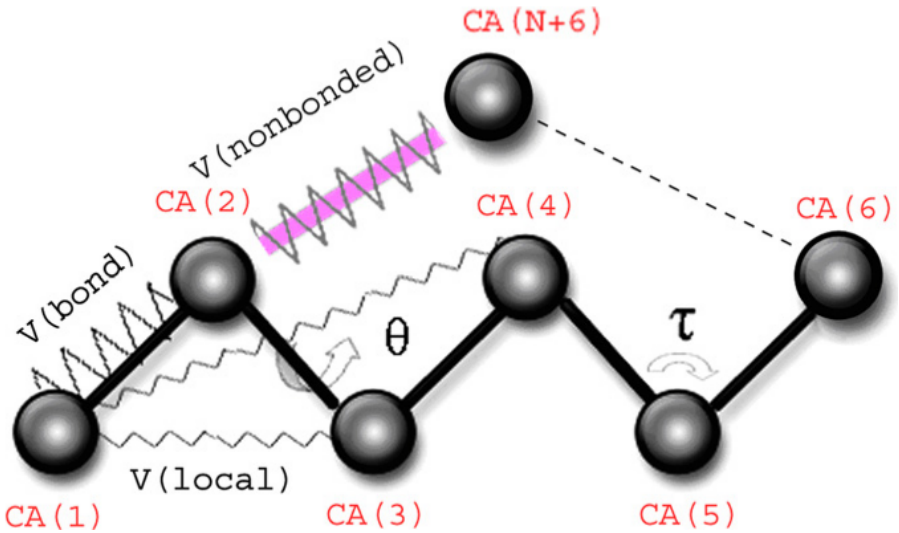
- illustration – different compromises may be made

from Pantano et al., J. Chem. Theory. Comput. 2015

Coarse-grained models

Yet another example – Vamm force field for proteins:

- every amino acid – a single bead at Cα
- more coarse than Martini



from Korkut & Hendrickson 2009

Enhanced sampling

How to save time, and time is money

Problem

with normal nanosecond length MD simulations:

It is difficult to overcome barriers to conformational transitions,
and only conformations in the neighborhood of the initial structure
may be sampled,
even if some other (different) conformations are more relevant,
i.e. have lower free energy

Special techniques are required to solve this problem.

Note – do not be afraid of Arrhenius

How often does something happen (in a simulation)?

$k = A \times \exp\left[-E_A/kT\right]$, let us have $A = 1 \times 10^9$ s^{-1}

E_A	k	$1/k$
kcal/mol	$1/$s	μs
1	0.19×10^9	0.005
3	6.7×10^6	0.15
5	0.24×10^6	4.2
7	8.6×10^3	120

So, if the process has to overcome a barrier of 5 kcal/mol,
we will have to simulate for 4 μs to see it happen once on average.

Replica-exchange molecular dynamics

REMD (or parallel tempering) – method to accelerate the sampling of configuration space, which can be applied
even if the configurations of interest are separated by high barriers.

Several (identical) replicas of the molecular system are simulated at the same time, with different temperatures.

The coordinates+velocities of the replicas may be switched (exchanged) between two temperatures.

Probability of replica exchange

The probability of the replica exchange between T_1 and T_2 is determined in (regular) time intervals from the instantaneous potential energies U_1 and U_2 in the corresponding simulations as

$$\mathcal{P}(1 \leftrightarrow 2) = \begin{cases} 1 & \text{if } U_2 < U_1, \\ \exp\left[\left(\frac{1}{k_B T_1} - \frac{1}{k_B T_2}\right) \cdot (U_1 - U_2)\right] & \text{otherwise.} \end{cases}$$

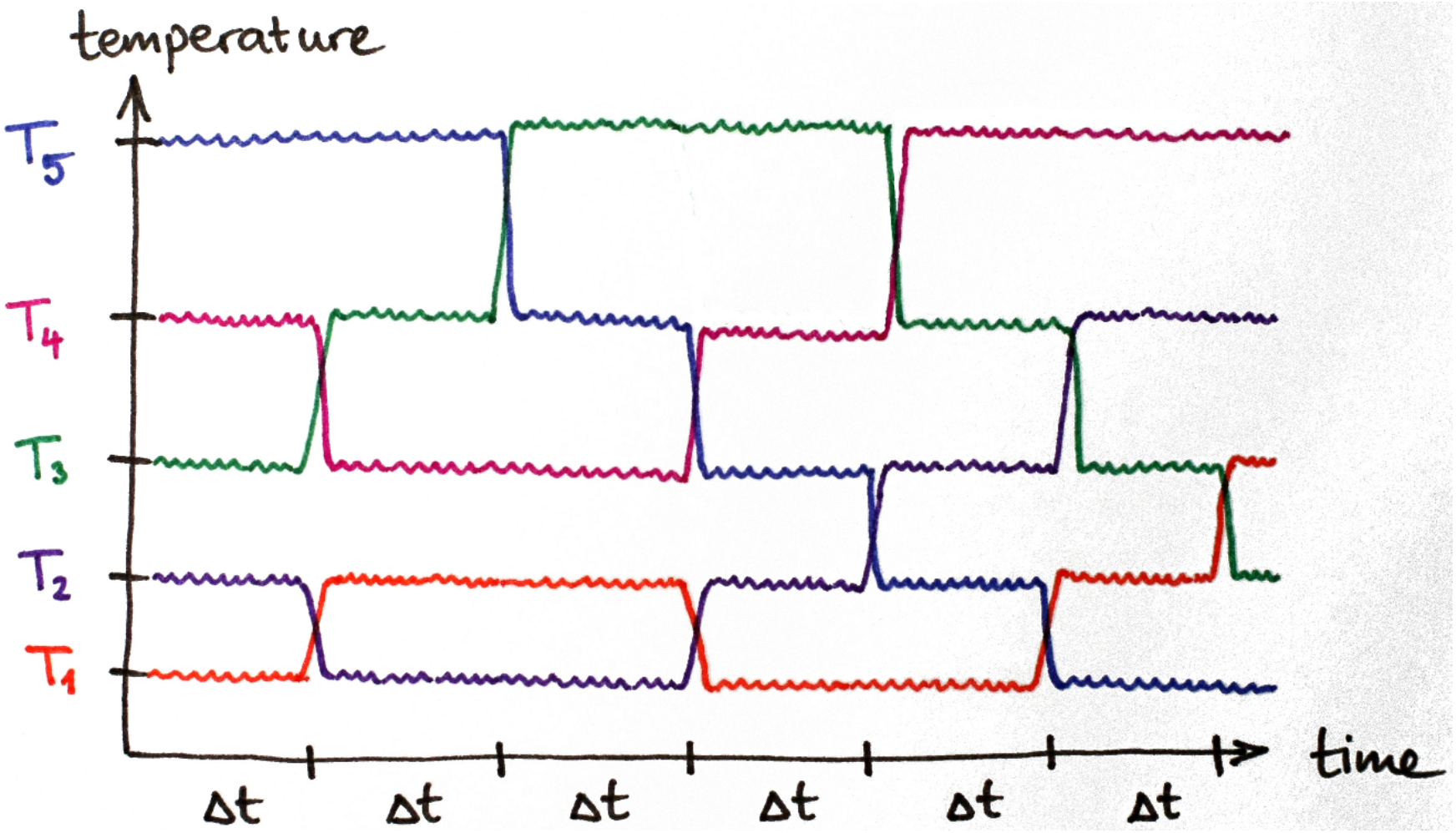
Then, if $\mathcal{P}(1 \leftrightarrow 2)$ is larger than a random number from $(0, 1)$, the replicas in simulations at T_1 and T_2 are exchanged.

Setup of the simulation of replicas

Simulated one replica at the temperature of interest ($T_1 = 300$ K) and several other replicas at higher temp. ($T_1 < T_2 < T_3 < \ldots$).

After (say) 1000 MD steps, attempt exchanges $1 \leftrightarrow 2$, $3 \leftrightarrow 4$ etc., and after next 1000 steps do the same for $2 \leftrightarrow 3$, $4 \leftrightarrow 5$ etc.
so only try to exchange replicas at "neighboring" temperatures

Setup of the simulation of replicas



Advantages of REMD

- faster sampling and more frequent crossing of energy barriers
- correct sampling at (not only) the temperature of interest
- increased computational cost (multiple simulations)
  pays off with largely accelerated sampling
- simulations running at different temperatures are independent
  except at attempted exchanges $\rightarrow$ easy parallelization
- first application – protein folding (Sugita & Okamoto, Chem. Phys. Lett. 1999)
- disadvantages:
  – large number of atoms $\rightarrow$ low $\mathcal{P} \rightarrow$ low efficiency
  – high temperature $\rightarrow$ sensitive biostructures may not survive

Hamiltonian replica exchange – HREX

– way to apply the replica-exchange idea and avoid those issues
– also called 'Replica exchange with solute tempering' (REST)

$$P = \exp\left[-\frac{U}{kT}\right] = \exp\left[-\beta U\right]$$

- note: $\frac{1}{2}U$ would be the same as $2T$
- force field energy U is combined from many individual terms
  – let us scale selected terms (not all of them!)
  – is not possible for temperature scaling (a single T)
  – 'heating' of a (small) part of the system
  – typically, a group of atoms – a ligand, or several AAs...

Hamiltonian replica exchange – HREX

Simulations 1 and 2 are performed
with different force fields U_1 and U_2

How to calculate the probability of exchange?
(q_1 and q_2 – coordinates of atoms in simulations 1 and 2)

$$\Delta = \frac{U_1(q_2) - U_1(q_1) - U_2(q_1) + U_2(q_2)}{kT}$$

$$\mathcal{P}(1 \leftrightarrow 2) = \begin{cases} 1 & \text{if } \Delta \leq 0, \\ \exp\left[-\Delta\right] & \text{otherwise.} \end{cases}$$

Then, if $\mathcal{P}(1 \leftrightarrow 2)$ is larger than a random number from $(0, 1)$, the replicas in simulations with U_1 and U_2 are exchanged.

HREX – a good variant

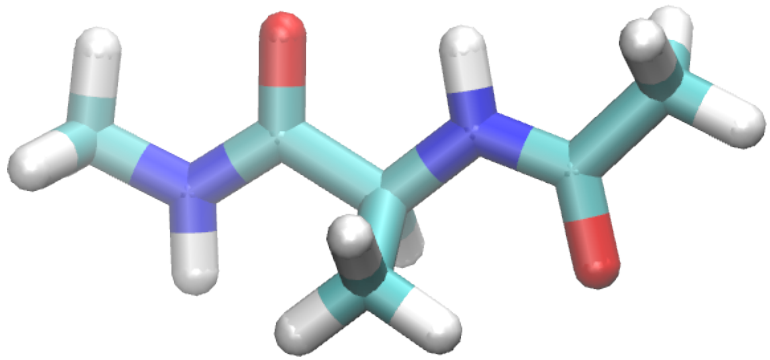
- divide the system into two parts:
- hot – a relatively small number of atoms
  – scale down the parameters of the force field
  $\rightarrow$ increase the effective temperature
- cold – all of the rest

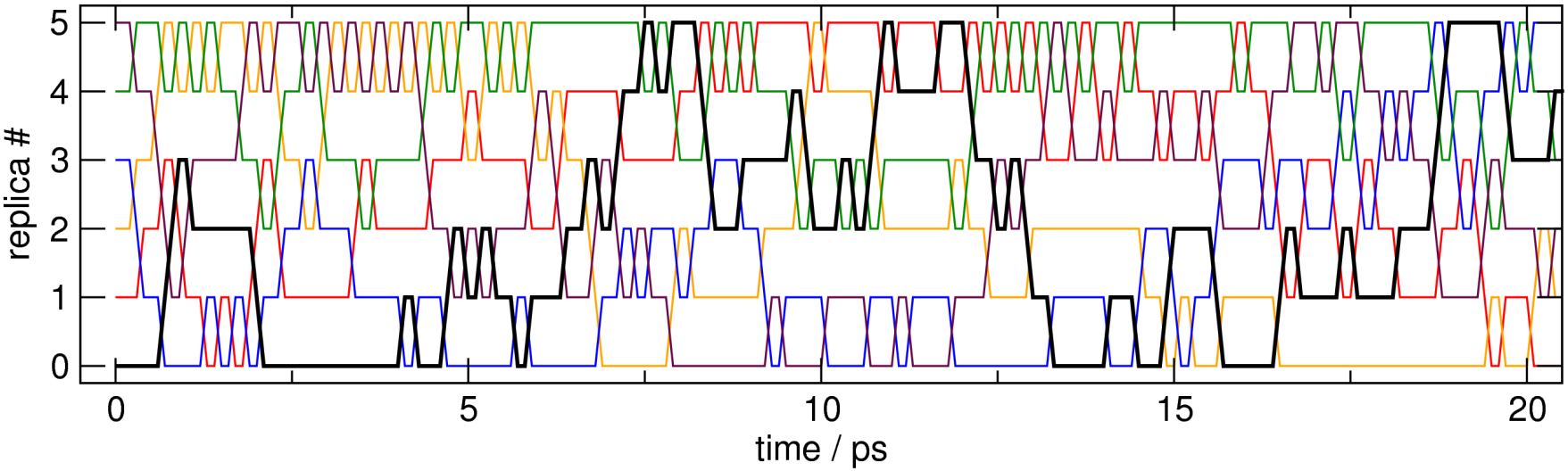
Meaning of temperature

- kinetic energy $\leftarrow$ velocities
  – does not change, is the same in hot and cold (300 K)
  – simulation settings need not be adjusted (time step!)
  – unlike in parallel tempering
- factor affecting the population of states
  – we play with this

HREX – example – solute tempering

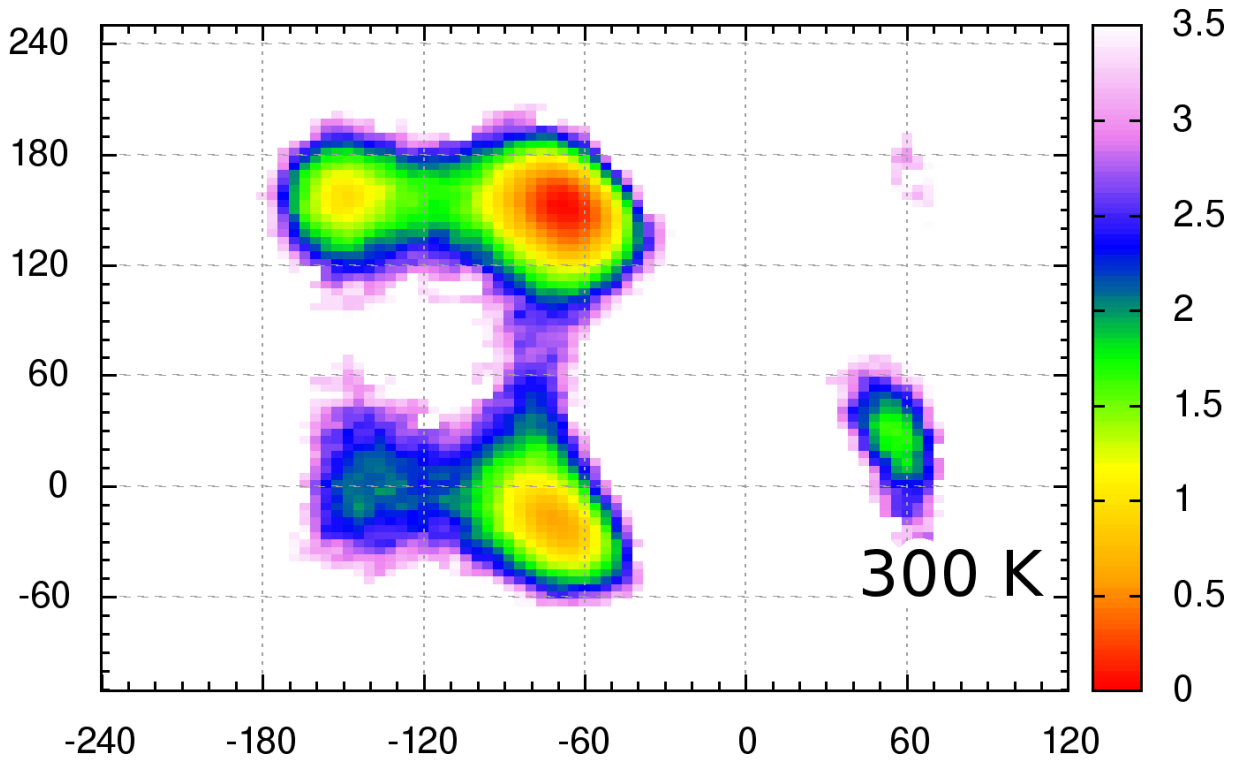
- alanine dipeptide – 22 atoms,
  1 pair of $\varphi - \psi$
- 5 replicas, $\lambda = 1 \ldots 0.18$
  i.e. $T_m = 300 \ldots 1700$ K
- exchange every 0.1 ps,
  observed $\overline{\mathcal{P}}$ =0.25–0.50





HREX – example

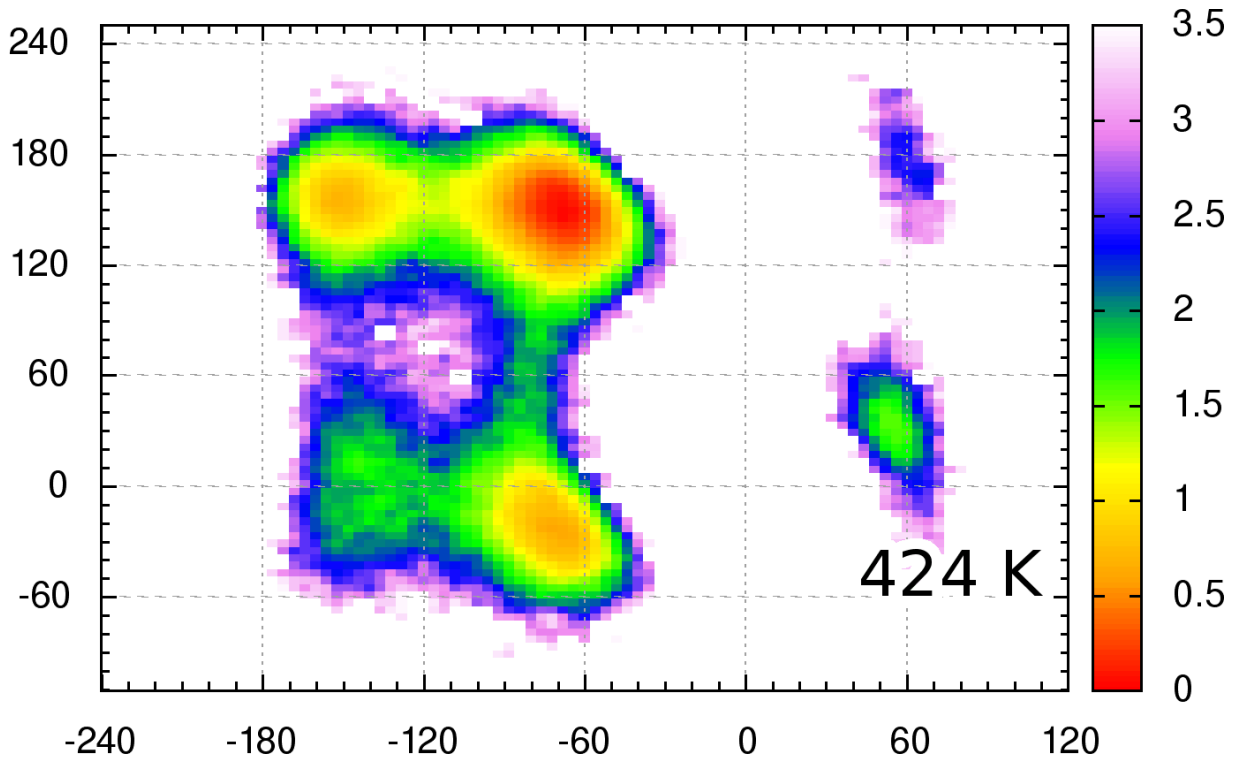
Solute tempering – dialanine – replica #0



$\varphi - \psi$ in degrees, ΔF in kcal/mol

HREX – example

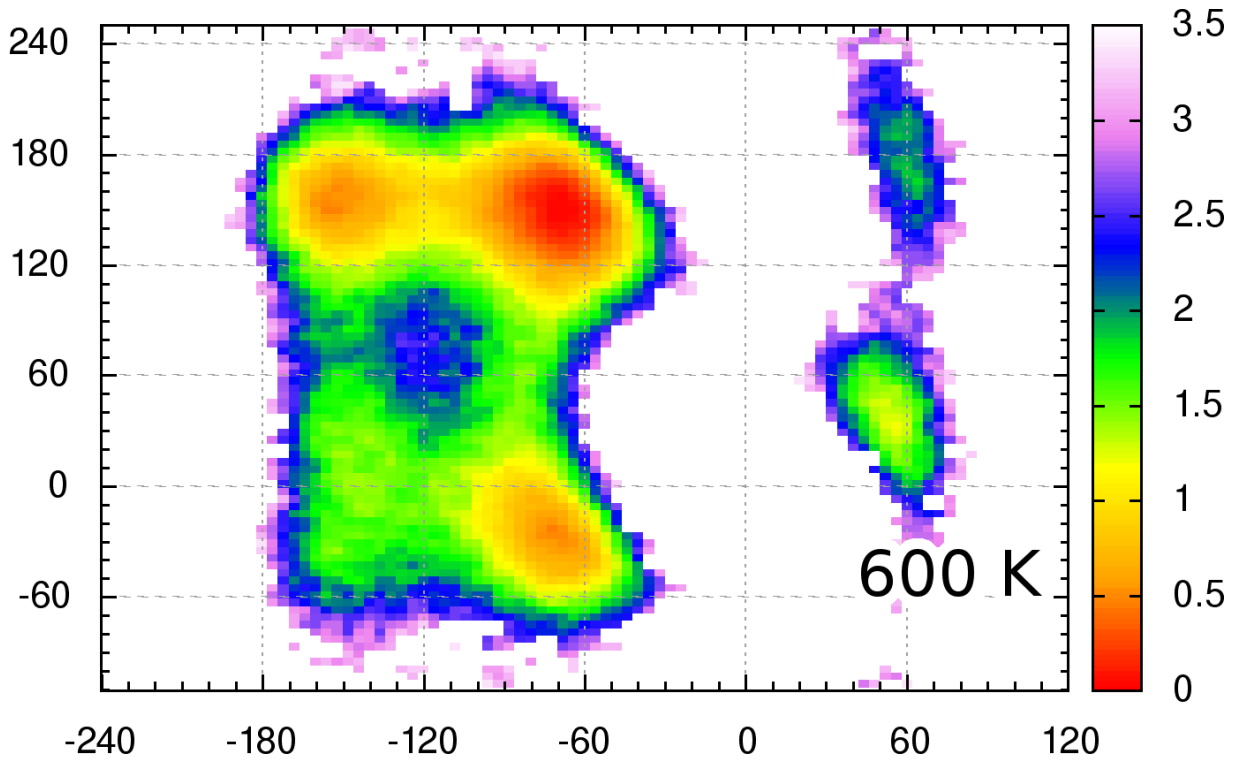
Solute tempering – dialanine – replica #1



$\varphi - \psi$ in degrees, ΔF in kcal/mol

HREX – example

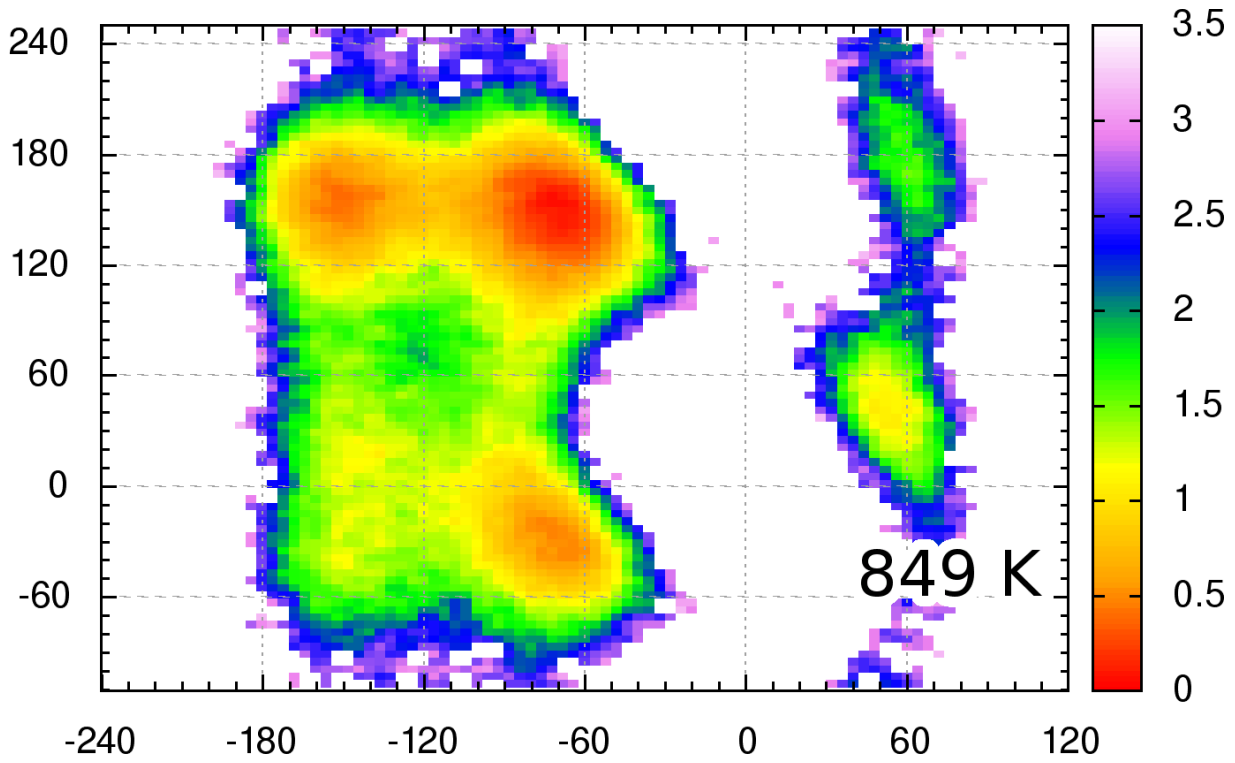
Solute tempering – dialanine – replica #2



$\varphi - \psi$ in degrees, ΔF in kcal/mol

HREX – example

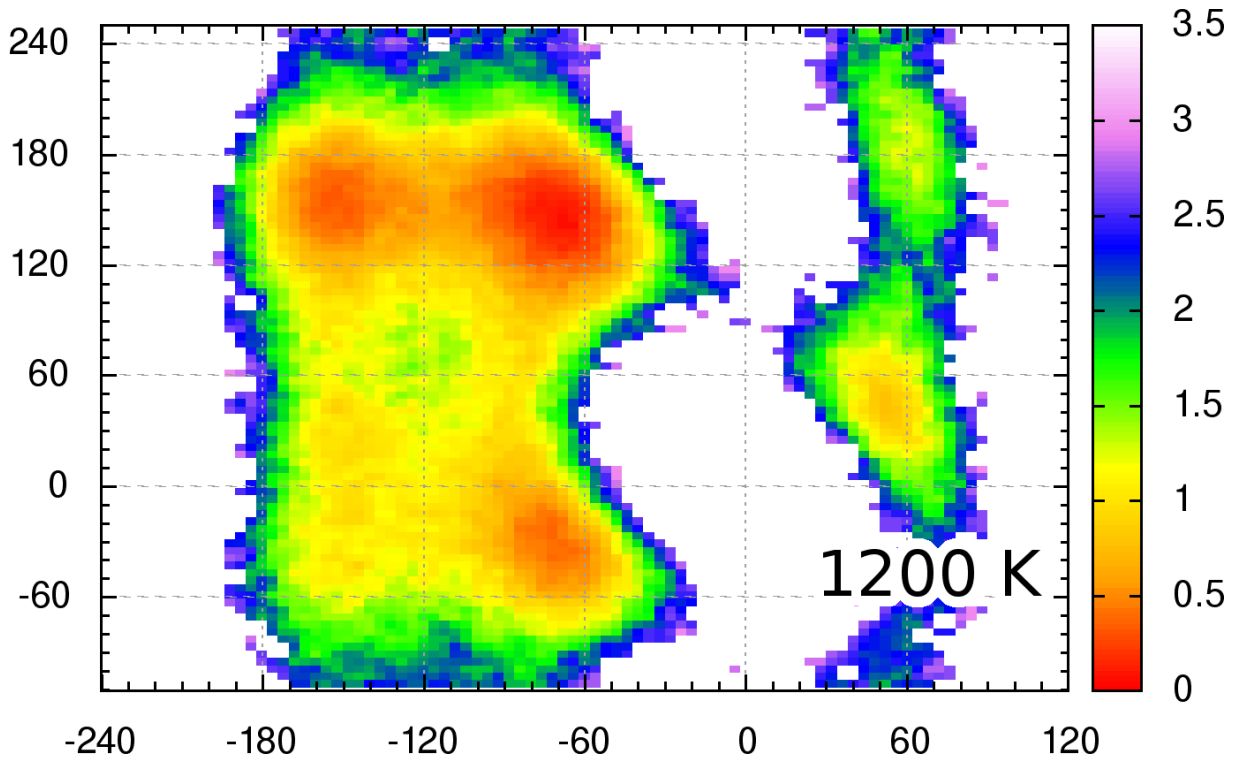
Solute tempering – dialanine – replica #3



$\varphi - \psi$ in degrees, ΔF in kcal/mol

HREX – example

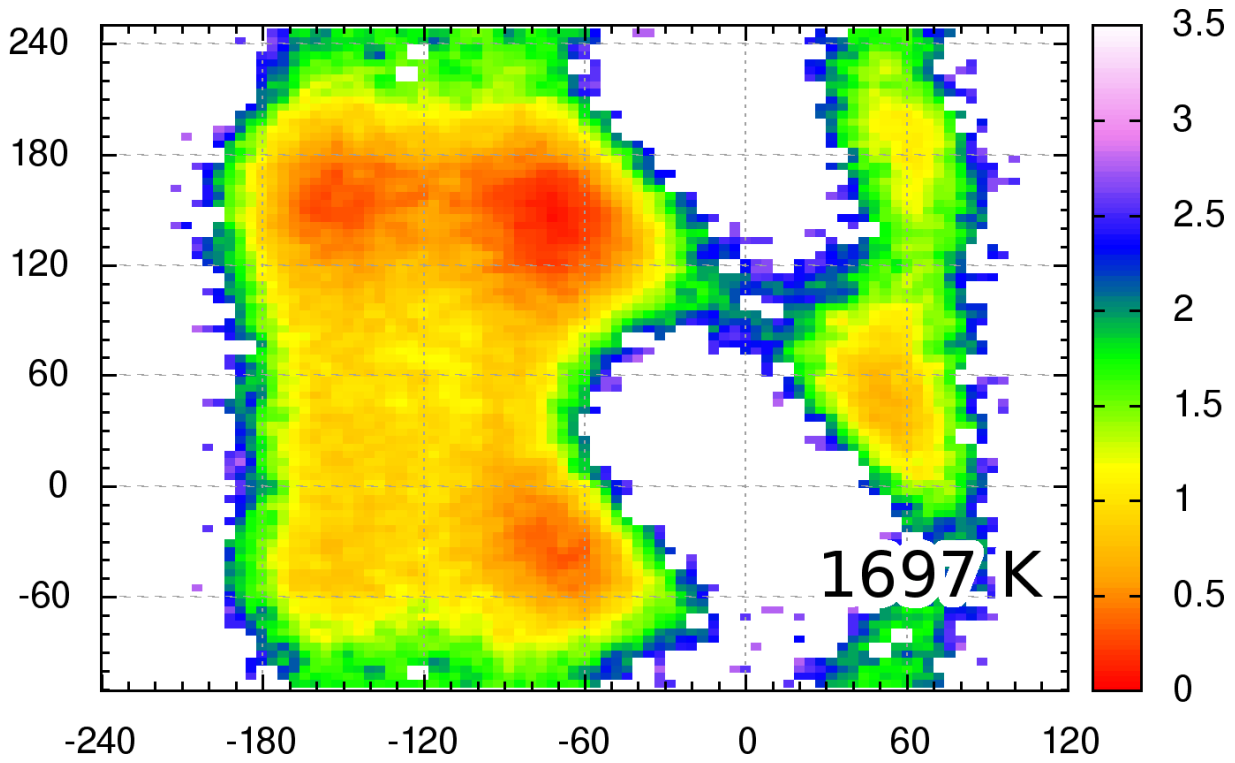
Solute tempering – dialanine – replica #4



$\varphi - \psi$ in degrees, ΔF in kcal/mol

HREX – example

Solute tempering – dialanine – replica #5



$\varphi - \psi$ in degrees, ΔF in kcal/mol

Methods using biasing potentials

Other approaches use a different idea:

It is easy to introduce an additional contribution
to the potential energy of the molecule

Example – the extra potential may force the molecule
over an energy barrier, to explore other conformations

It is 'unrealistic' – we do not simulate a real molecule
but this bias may be removed by a right post-processing

Note: use of NMR-based distance restrains in MD simulations
$\rightarrow$ 'NMR-refined' structure of the molecule (e.g. PDB ID 1AC9)

Metadynamics

– aimed at reconstructing the multidimensional free energy of complex systems (Laio & Parrinello 2002)

– based on an artificial dynamics (metadynamics) performed in the space of a few collective variables S (e.g. normal modes)

– at regular time intervals during the simulation,
an additional biasing energy function is added to the force field
    – a Gaussian that is centered on the current structure

using quotations by Alessandro Laio

Metadynamics – how it works

a new Gaussian is added at every time interval t_G,
and the biasing potential at time t is given by

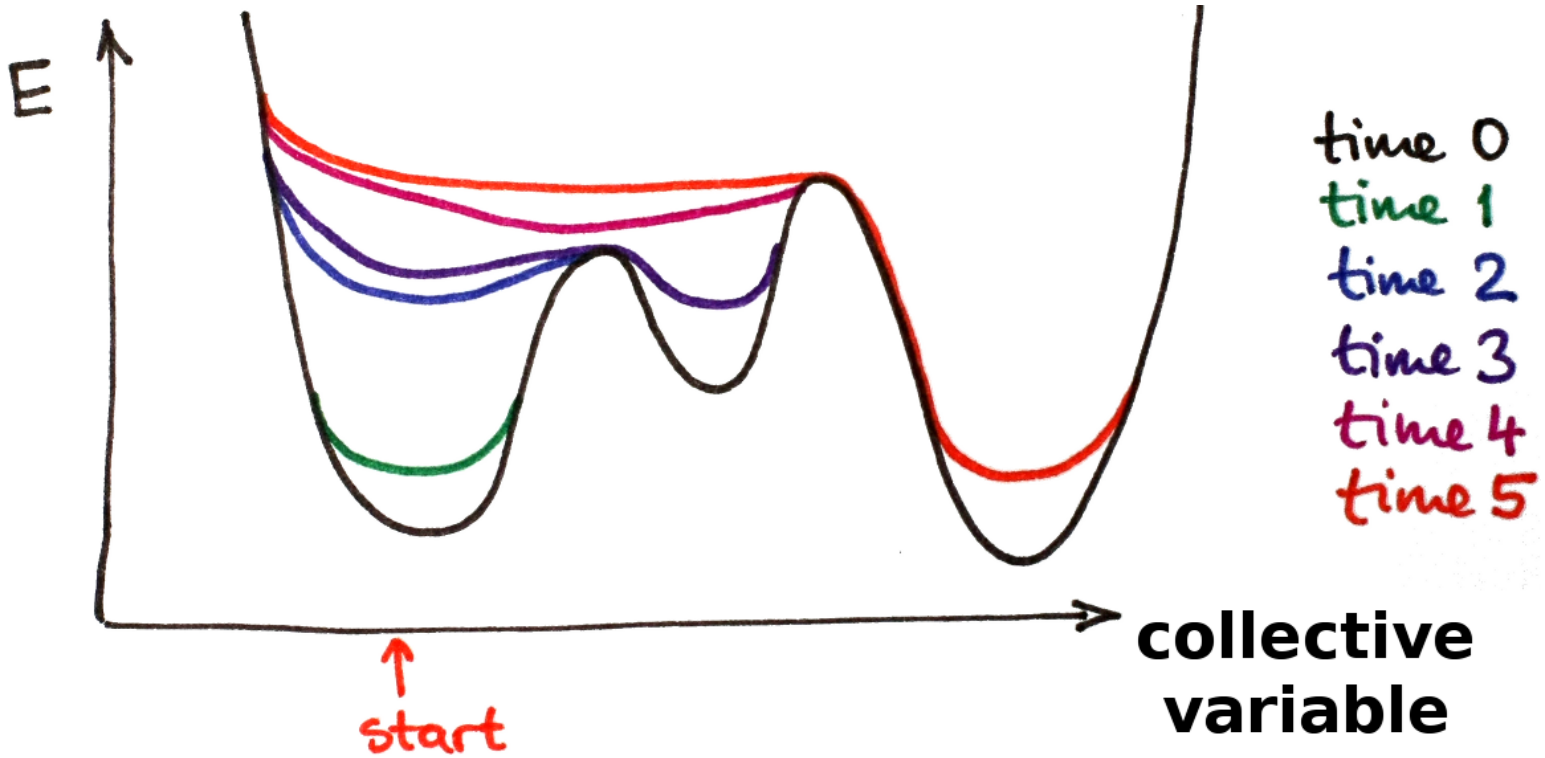
$$V_G(S(x), t) = \sum_{t'=t_G, 2t_G, 3t_G, \ldots} w \cdot \exp\left[-\frac{(S(x) - s_{t'})^2}{2 \cdot \delta s^2}\right]$$

w and δs – height and width of the Gaussians
$s_t = S(x(t))$ – value of the collective variable at time t

In the course of the simulation, this potential is filling the minima on the free energy surface that the system is traveling through

So, the MD has a memory via the biasing potential

Metadynamics – what it looks like



https://www.youtube.com/watch?v=IzEBpQ0c8TA
https://www.youtube.com/watch?v=iu2GtQAyoj0

Properties of metadynamics

Metadynamics – to explore new reaction pathways,
accelerate rare events,
and also to estimate the free energies efficiently.

- The system escapes a local free energy minimum through the lowest free-energy saddle point.
- The dynamics continues, and all of the free-energy profile is filled with the biasing Gaussians.
- At the end, the sum of the Gaussians provides the negative of the free energy.

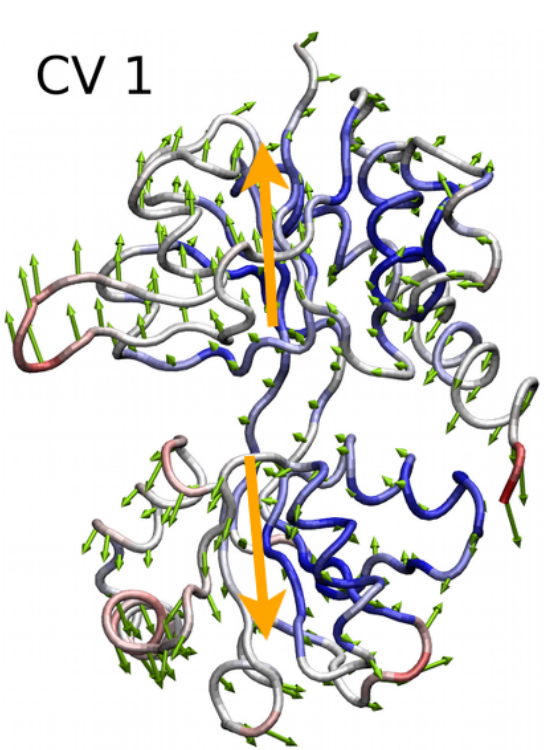
Properties of metadynamics

Crucial task – prior to simulation:
identify the collective variables of interest
that are difficult to sample because of high barriers

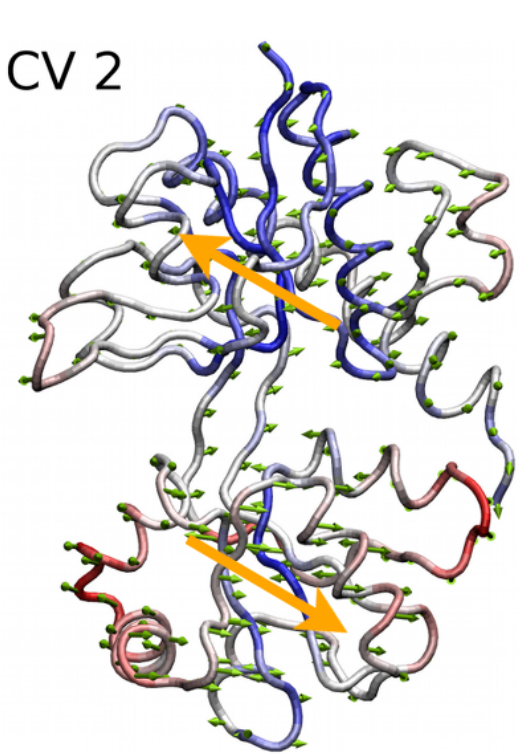
These variables $S(x)$ are functions of the coordinates of the system;
practical applications – up to 3 such variables,
and the choice depend on the process being studied.

Typical choices – principal modes of motion obtained with PCA
Still, the choice of S may be difficult

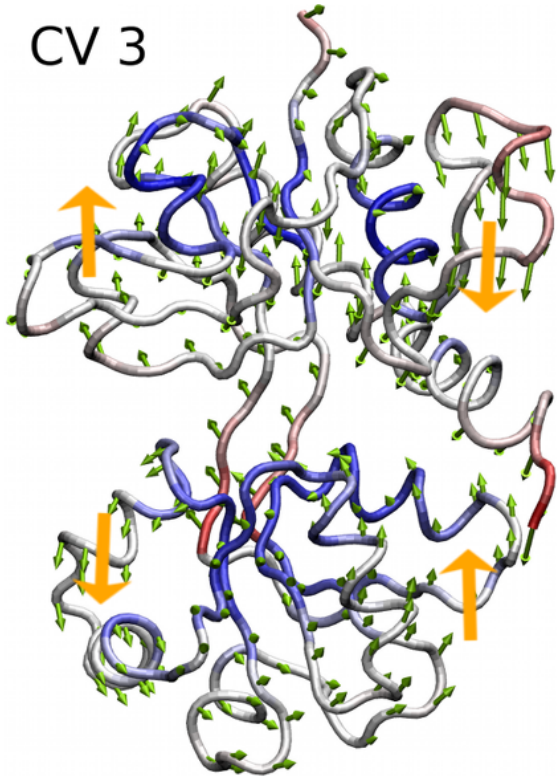
Example – opening of a protein binding pocket



clamshell



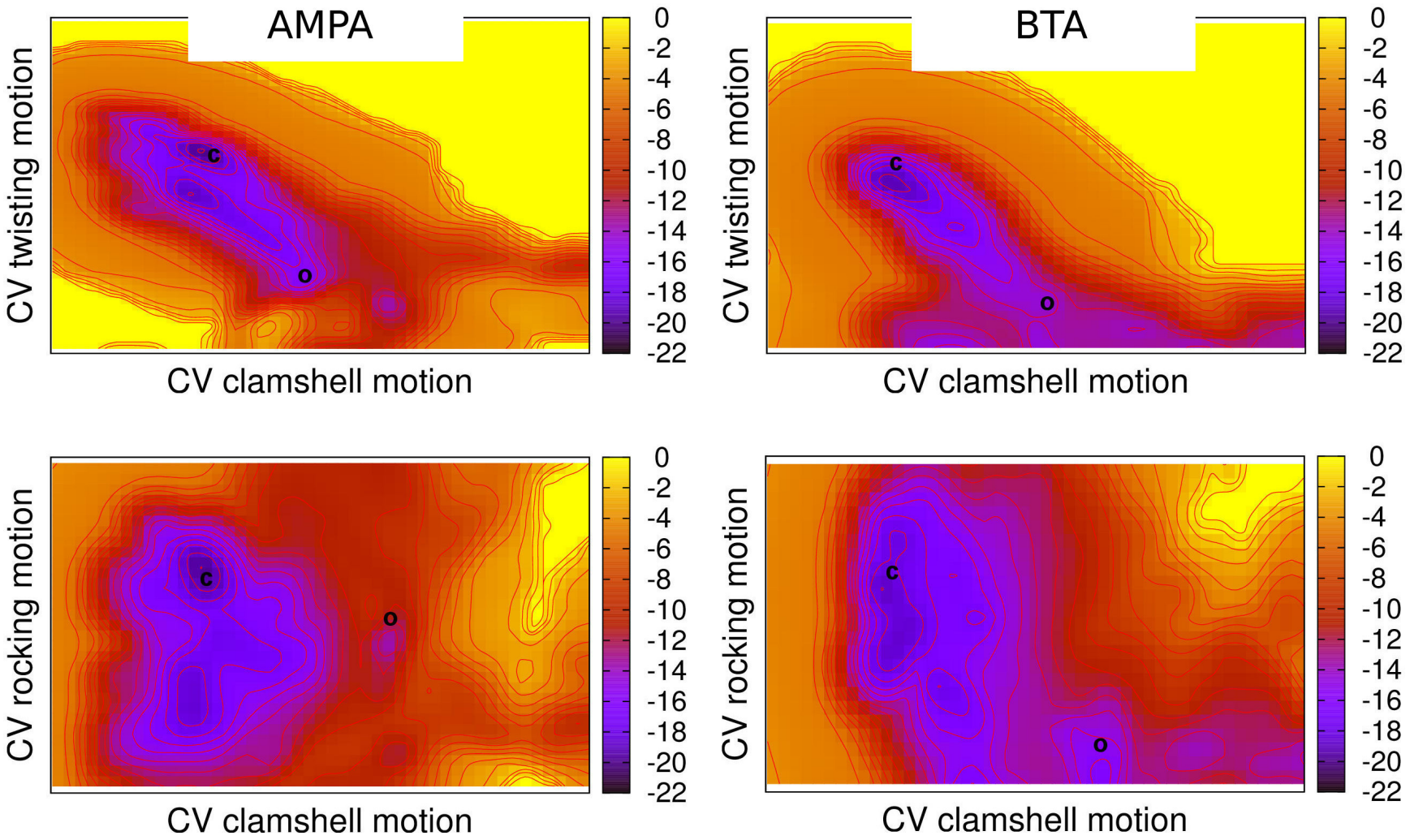
twisting



rocking

courtesy Tino Wolter

Example – opening of a protein binding pocket



courtesy Tino Wolter

Enhanced sampling methods – comparison

Biasing potential methods – metadynamics, umbrella sampling

- required: a priori choice of reaction coordinate(s) to be biased
- problem – success depends on that choice, possibly non-trivial

REMD with parallel tempering

- + no such required, can be used rather blindly
- – all of the system heated $\rightarrow$ may destroy something
- – no knowledge of the system may be embedded
- – poor efficiency for big systems: $\overline{\mathcal{P}(1 \leftrightarrow 2)} \approx \exp\left[-2\varepsilon^2 N\right]$
  $\rightarrow$ critical problem

Enhanced sampling methods – comparison

Hamiltonian replica exchange (HREX)

- in intermediate position
  between metadynamics/US and REMD-PT
- simpler to use than metadynamics/US
  – results depend not so strongly on the choices to be made
- efficiency does not depend on the overall system size

Free energy simulations

Motivation

free energies – Helmholtz F or Gibbs G
– determine whether processes (reactions) run spontaneously or not
– are extremely important and difficult to calculate

convergence in MD simulations – especially desperate for F and G:

$$F = k_B T \cdot \ln \left\langle \exp \left[\frac{E}{k_B T} \right] \right\rangle$$

problem – the large energy values enter an exponential, so
if high-energy structures are undersampled, then F / G are wrong

$\rightarrow$ calculation of free energies impossible from free MD simulation,
special methods needed!

Tackling the issue

two fundamental approaches:
free energy perturbation and thermodynamic integration

several computational tricks for particular types of reactions:
alchemical simulations or umbrella sampling

important: not necessary to find the absolute value of free energy;
for a chemical reaction, we only need
the free energy difference (ΔF, ΔG) of reactant and product

"reaction" – not necessarily chemical bonds created or broken
– ligand binding a protein
– passage of a molecule/ion through membrane
– protein folding
...

Thermodynamic integration

Free energy as function of reaction coordinate λ: $F = F(\lambda)$,
with $\lambda = 0$ for reactant, $\lambda = 1$ for product

$$\Delta F = F(1) - F(0) = \int_0^1 \frac{\partial F(\lambda)}{\partial \lambda} \mathrm{d}\lambda$$

Free energy is a state function
$\rightarrow$ the result is independent of the chosen path $0 \rightarrow 1$
$\rightarrow$ reaction coordinate may be even an unphysical process
– change of chemical identity of atoms (alchemical simulations)

Principle of TI – the derivative of total MM energy E is evaluated in the simulation directly, and then easily averaged

$$\Delta F = \int_0^1 \left\langle \frac{\partial E_\lambda}{\partial \lambda} \right\rangle_\lambda \mathrm{d}\lambda$$

How to do it practically

- perform a MD simulation for each chosen value of λ:
  – usually, equidistant values in the interval (0,1) are taken:
  0, 0.05, . . . , 0.95 and 1
- each of these simulations runs with a different parameter set
  – interpolation of parameters between reactant and product
- each of these simulations produces a value of $\left\langle \frac{\partial E}{\partial \lambda} \right\rangle_\lambda$
  – we obtain the derivative of F in discrete points for $\lambda \in (0, 1)$
- this function is integrated numerically,
  – the result is the desired free energy difference ΔF

Computational alchemy

TI looks complicated, but it is rather straightforward,
– common simulation programs run TI conveniently

Computational alchemy
– change of chemical identities of atoms or functional groups

Using a parameter λ, the force-field parameters of state 0 are changed to those of state 1 gradually:

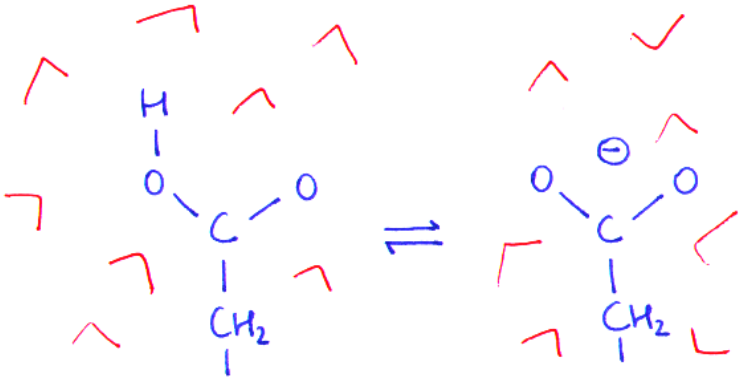
$$E_\lambda = (1 - \lambda) \cdot E_0 + \lambda \cdot E_1$$

Advantages of TI

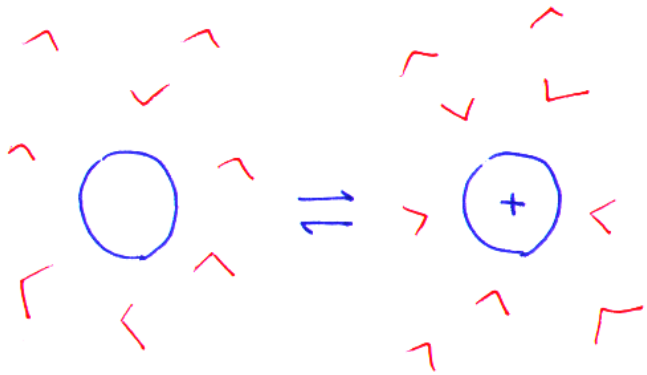
- evaluate the derivative of energies,
  no need to sample for the (large) total energies first
- it is not important what happens outside of the region
  where the reaction takes place (no contrib. to $E_1 - E_0$)
- the ensemble of structures that have to be sampled thoroughly
  is much smaller, and shorter simulation length is required

Examples

deprotonation of amino acid



ionization of a molecule



Example

Free energy of hydration of rare gas (neon)

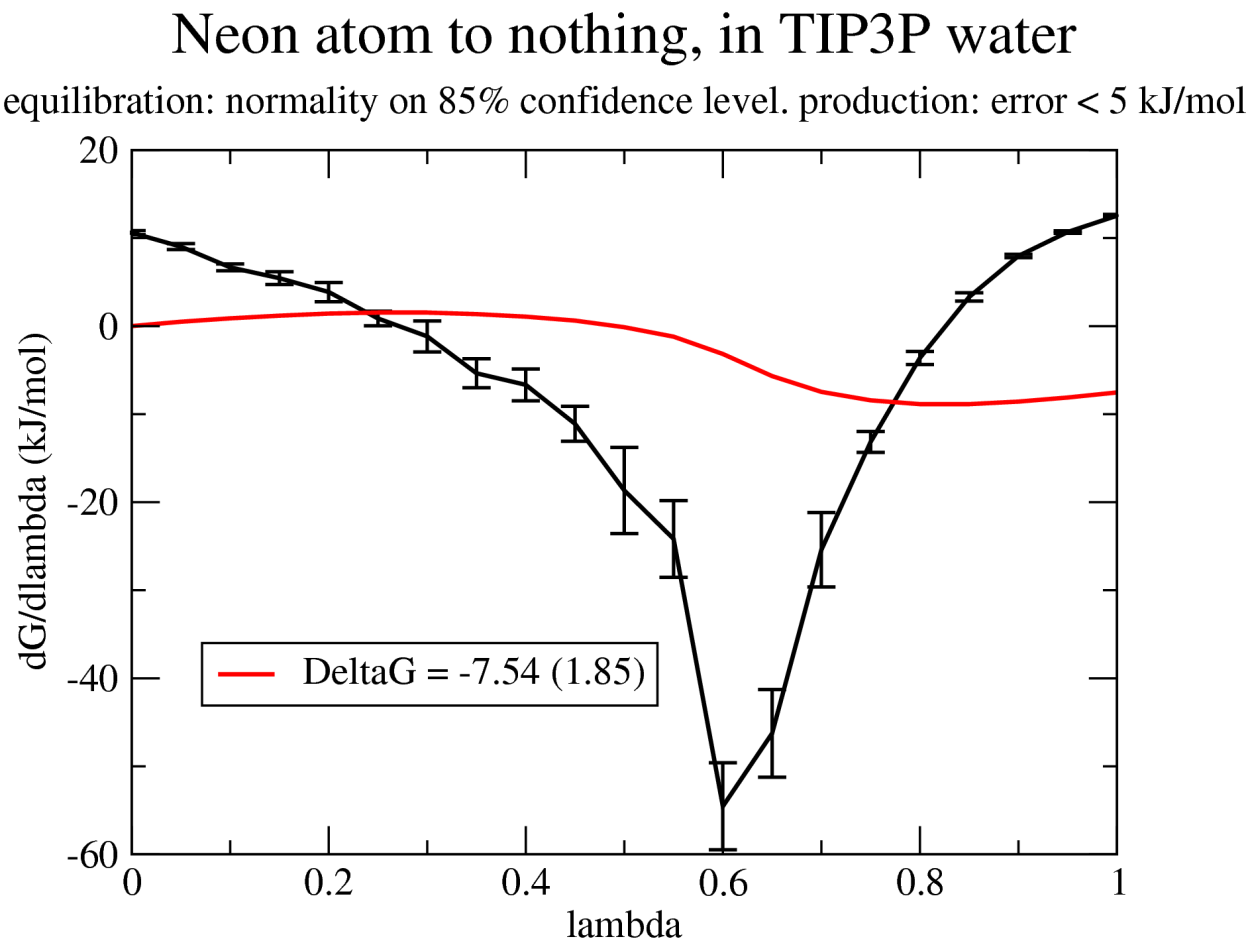
van der Waals parameters of the neon are gradually switched off by means of λ, so that the atom is effectively disappearing

The derivative of total energy with respect to λ is evaluated for 21 values of λ ranging from 0 to 1.

Then, TI gives the Gibbs energy difference of two states:

- a neon atom in water
- no neon atom in water $\equiv$
  $\equiv$ a neon atom outside of the solution, in vacuo

Example


Neon atom to nothing, in TIP3P water
equilibration: normality on 85% confidence level. production: error < 5 kJ/mol
dG/dlambda (kJ/mol)
lambda
DeltaG = -7.54 (1.85)
20
0
-20
-40
-60
0
0.2
0.4
0.6
0.8
1

Differences of differences

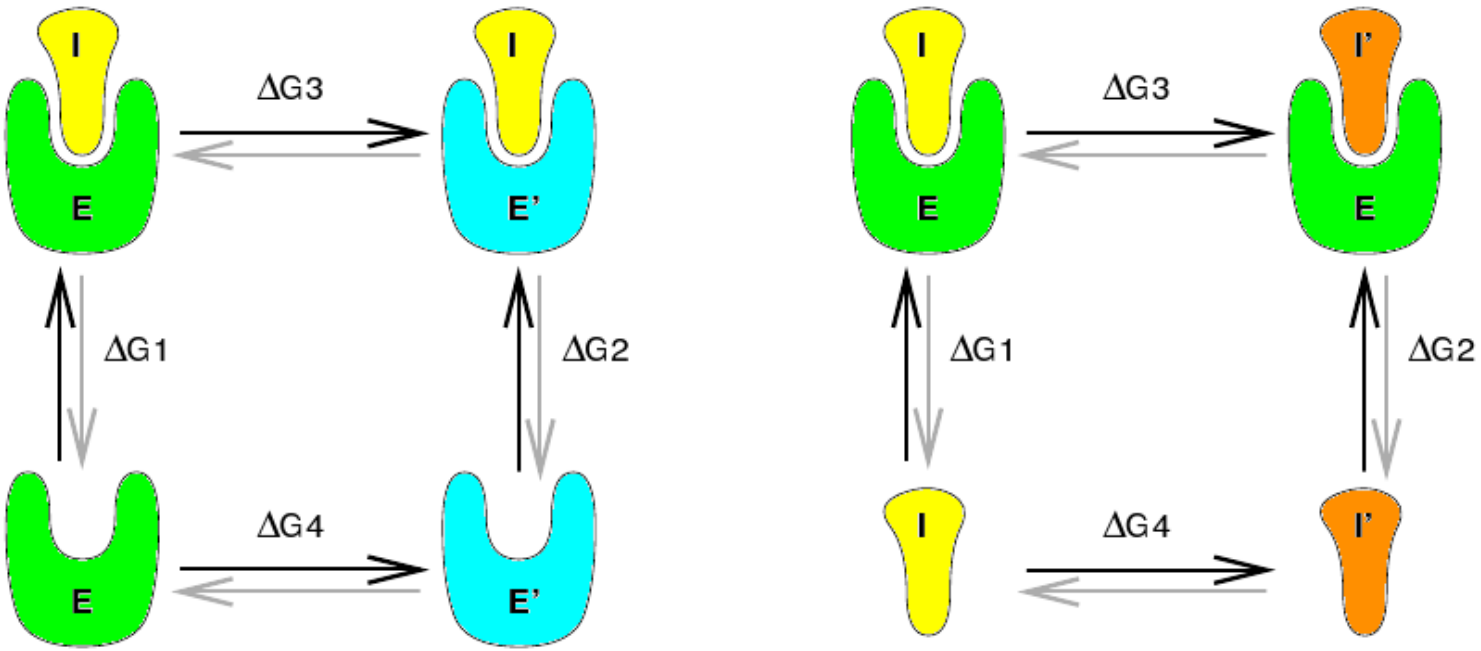
Often – we are interested not in the absolute free energies
and not even in the reaction free energies,
rather, in the difference (Δ) of reaction free energies (ΔF)
of two similar reactions:

$$\Delta\Delta F \text{ or } \Delta\Delta G$$

Reaction free energy difference

Example left: binding of an inhibitor molecule I to an enzyme E, difference of binding free energies to similar enzymes E and E′:

$$\begin{aligned} E + I &\rightleftharpoons EI & \Delta G_1 \\ E' + I &\rightleftharpoons E'I & \Delta G_2 \end{aligned}$$



Reaction free energy difference

The simulation of the ligand binding process itself – very difficult
(possibly large structural changes in the enzyme upon binding)

Solution of the problem – do not simulate the reaction of binding,
rather, the alchemical transmutation of enzyme E to E′.

E and E′ are very similar, so this may be easy to do.
(example: mutation of a single AA, e.g. leucine to valine)

Then, the structure of complexes EI and E′I may be similar as well,
and the simulation may provide converged free energy.

Reaction free energy difference

Free energy is a state function $\rightarrow$ the sum of free energies around a thermodynamic cycle vanishes:
(e.g. clockwise in figure left):

$$\Delta G_1 + \Delta G_3 - \Delta G_2 - \Delta G_4 = 0$$

The difference of binding free energies equals
the difference of free energies calculated in alchemical simulations:

$$\Delta\Delta G = \Delta G_1 - \Delta G_2 = \Delta G_3 - \Delta G_4$$

Geometric reaction coordinate

Sometimes, we need to know how the free energy changes along a geometric reaction coordinate q

The free energy is then a function of q

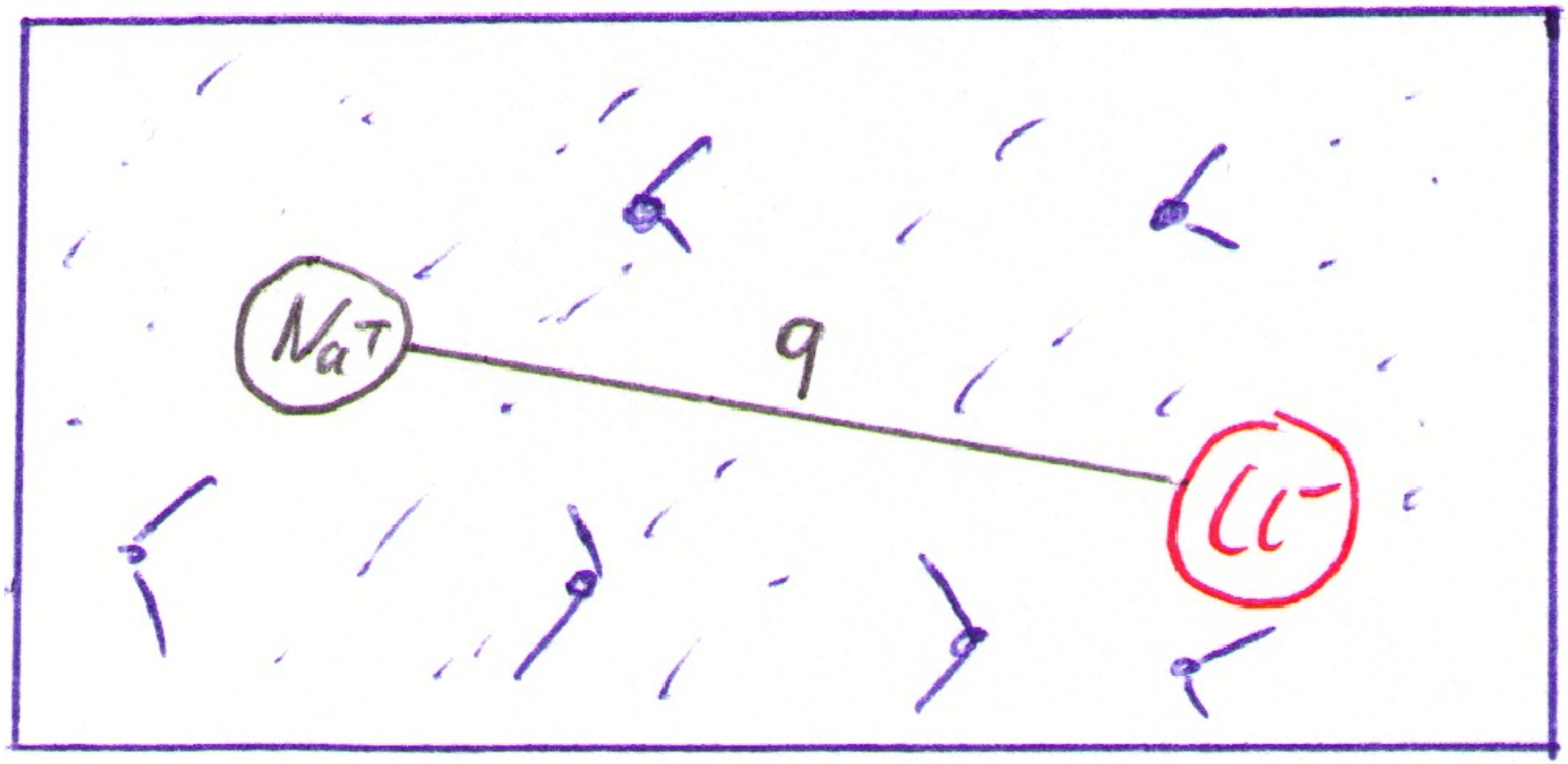
Such a function $F(q)$ is called the potential of mean force.

Examples:

- distance between two particles in a dissociating complex
- the dihedral angle when dealing with conformational changes
- the position of a proton for a reaction of proton transfer

Example

free energy of formation of an ion pair in solution:



Working principle

The problem:
If a high barrier has to be crossed to come from A to B,
a free MD simulation may not reach the product B,
or at least the barrier region is described poorly

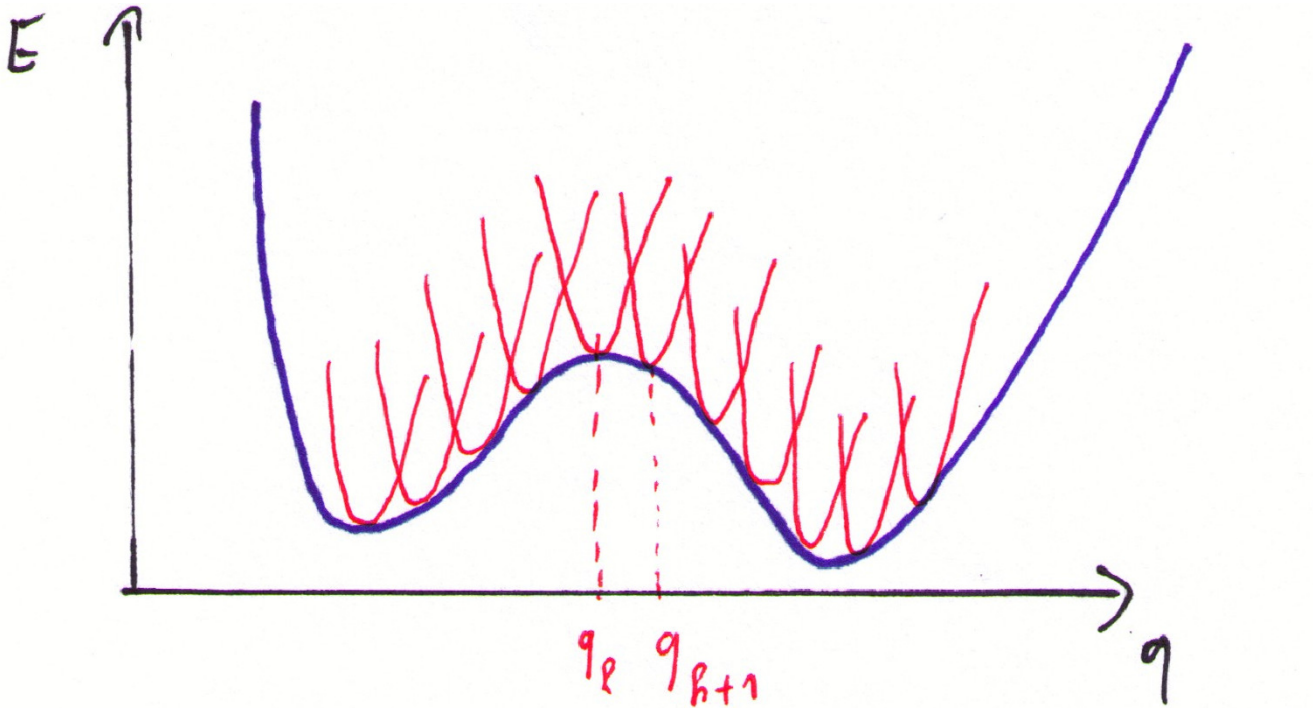
The solution:
Apply an additional potential, also called biasing potential
to restrain the system to values of reaction coordinate
that would otherwise remain possibly undersampled

This is the principle of the umbrella sampling.

The additional potential will become a part of the force field,
and it shall depend only on the reaction coordinate: $V = V(q)$

Practical umbrella sampling

We can use this scheme efficiently, by way of moving a biasing harmonic potential along the reaction coordinate:



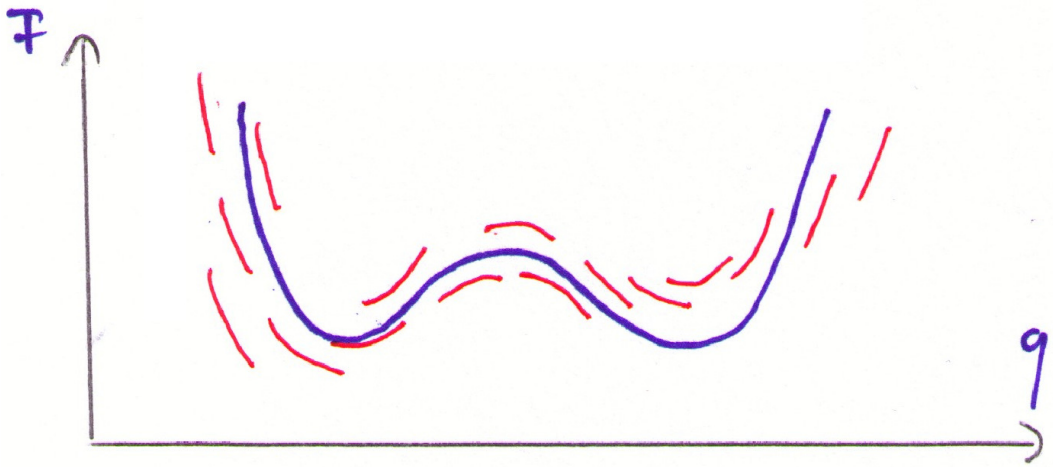
Potentials of mean force

- We perform k simulations with biasing potentials V_k, which are designed to cover the interesting range of q
- From each of the k simulations, we extract the biased probability $P^*(q)$ of finding the system at the certain value of q – this differs from the real, unbiased probability $P(q)$
- We correct for bias to obtain the unbiased free energy $F(q)$

$$F_k(q) = -k_B T \ln P^*(q) - V_k(q) + K_k$$

where K_k is a constant shift that has to be determined

Potentials of mean force



- the curves $F_k(q)$ for simulations k and $k+1$ differ by a constant shift, difference of K
- main task – match the pieces to provide a continuous curve
- one way – fit the values K_k to obtain a resulting $F(q)$ curve that is as continuous and smooth as possible
- requirement – sufficient 'overlap' of the pieces F_k and F_{k+1}
- WHAM method – included in modern simulation programs

Potentials of mean force