International Spring School Statistical Thermodynamics, Santiago de Chile
Tuesday, November 28, 2017
Lecture 18

Methods for searching and sampling configurational space

Prof. Dr. Wilfred F. van Gunsteren

ETH Zürich, Switzerland

Searching and sampling configuration space

A. Types of methods for searching configuration space

A. Systematic or exhaustive search

B. Heuristic search

1. Non-step methods (e.g. Distance Geometry)
2. Step methods: change of a complete configuration (e.g. MC, MD, SD)
3. Step methods: build-up of a configuration (e.g. CBMC)

B. Types of search enhancement techniques

1. Deformation or smoothening of the potential energy surface

- Soft-core non-bonded interaction
- Local-elevation search
- Coarse graining of the molecular model

2. Scaling of system parameters

- Temperature annealing
- Tight coupling to a heat bath
- Mass scaling
- Mean-field approaches

3. Multi-copy searching and sampling

- Replica-exchange and multi-canonical algorithms
- Cooperative search: SWARM MD

Methods for searching configuration space for configurations r^N with low $V(r^N)$=energy

I. Molecular coordinates as variables:

A. Systematic or exhaustive search

$\vec{r}^{\,N}$ cartesian

φ^N angles

- Scan complete space (SS)

  small molecules

B. Heuristic search

- Generate tiny set of representative conformers:

1. Non-step methods

Distance geometry algorithms (DG)

Distribution? Solvent?

2. Step methods: change of a complete configuration

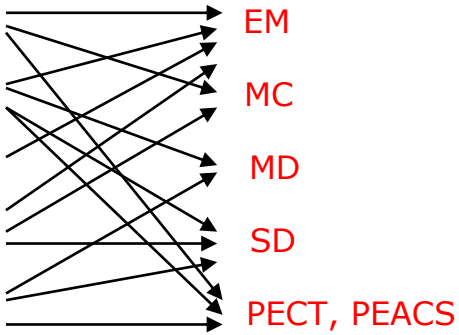
a. Energy: V(x)

b. Gradient: $-\frac{\partial V}{\partial x}$

c. 2nd Derivative: $\frac{\partial^2 V}{\partial x \partial x'}$

d. Random:

e. Memory:



3. **Step methods: build-up of a configuration**

configurational bias MC

combinatorial chain growing

Techniques to enhance the searching and sampling power of simulation methods

1. **Deformation or smoothening of the potential energy surface**
   a. omission of high-resolution structure factor data in structure refinement based on X-ray diffraction data
   b. gradual introduction of longer-range distance restraints in variable target structure refinement based on NMR NOE data
   c. softening of the hard core of atoms in the non-bonded interaction (soft-core atoms)
   d. reduction of the ruggedness of the energy surface through a diffusion-equation type of scaling
   e. avoiding the repeated sampling of an energy well through local potential energy elevation or conformational flooding
   f. softening of geometric restraints derived from experimental (NMR, X-ray) data through time-averaging of these
   g. circumvention of energy barriers through an extension of the dimensionality of the Cartesian space (4D-MD)
   h. freezing of high-frequency degrees of freedom through the use of constraints
   i. coarse-graining the model by reduction of the number of interaction sites

2. **Scaling of system parameters**
   a. temperature annealing
   b. tight coupling to heat bath
   c. mass scaling
   d. mean-field approaches

3. **Multi-copy searching and sampling**
   a. genetic algorithms
   b. replica-exchange and multi-canonical algorithms
   c. cooperative search: SWARM

*Review: J. Comput. Chem. 29 (**2007**) 157-166*

Methods for searching configuration space for configurations r^N with low $V(r^N)$=energy

I. Molecular coordinates as variables:

A. Systematic or exhaustive search

$\vec{r}^{\,N}$ cartesian

φ^N angles

- Scan complete space (SS)

  small molecules

B. Heuristic search

- Generate tiny set of representative conformers:

1. Non-step methods

Distance geometry algorithms (DG)

Distribution? Solvent?

2. Step methods: change of a complete configuration

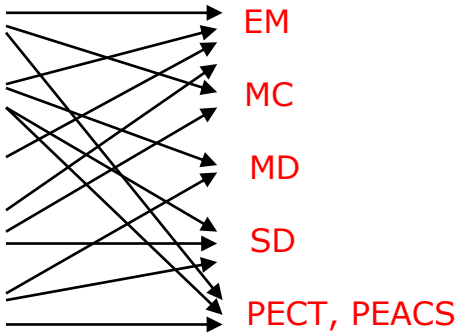
a. Energy: V(x)

b. Gradient: $-\frac{\partial V}{\partial x}$

c. 2nd Derivative: $\frac{\partial^2 V}{\partial x \partial x'}$

d. Random:

e. Memory:



3. **Step methods: build-up of a configuration**

configurational bias MC

combinatorial chain growing

Systematic search of conformational space

1. **Scan complete or significant part of space**
2. **Exclude subspaces based on:**
   - physical/chemical knowledge
   - solutions obtained so far

Exponential growth of computing effort as function of number of degrees of freedom

only for small molecules

Methods for searching configuration space for configurations r^N with low $V(r^N)$=energy

I. Molecular coordinates as variables:

A. Systematic or exhaustive search

$\vec{r}^{\,N}$ cartesian

φ^N angles

- Scan complete space (SS)

  small molecules

B. Heuristic search

- Generate tiny set of representative conformers:

1. Non-step methods

Distance geometry algorithms (DG)

Distribution? Solvent?

2. Step methods: change of a complete configuration

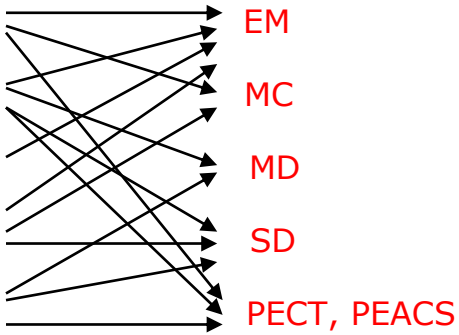
a. Energy: V(x)

b. Gradient: $-\frac{\partial V}{\partial x}$

c. 2nd Derivative: $\frac{\partial^2 V}{\partial x \partial x'}$

d. Random:

e. Memory:



3. **Step methods: build-up of a configuration**

configurational bias MC (CBMC)

combinatorial chain growing

Generating spatial structures or searching conformational space

Step methods: change of a complete configuration

1. Energy minimization:

Only local minimum is found, no escape from it

$$\Delta x = -\frac{\partial V_{pot}}{\partial x} \cdot constant$$

2. Monte Carlo:

Good for liquids of small molecules, not for folded long chain molecules

$$\Delta x = random$$

$$\text{acceptance probability} = e^{-\left[V_{pot}(x+\Delta x)-V_{pot}(x)\right]/kT}$$

3. Molecular Dynamics (Newton):

Surmounts energy barriers ~ $k_B T$

$$m\frac{d^2x}{dt^2} = -\frac{\partial V_{pot}}{\partial x} \begin{cases} v_{new} = v_{old} - \frac{1}{m}\frac{\partial V_{pot}}{\partial x}\Delta t \\ x_{new} = x_{old} + v_{new}\Delta t \end{cases}$$

4. Stochastic dynamics (Langevin):

MD + randomisation

$$m\frac{d^2x}{dt^2} = -\frac{\partial V_{pot}}{\partial x} + \text{random force} - m\gamma\frac{dx}{dt}$$

friction

5. Modified molecular dynamics (**PEACS**):

Enhances barrier crossing

Potential Energy Annealing Conformational Search

MD plus $$\frac{dV_{pot}(t)}{dt} = \frac{1}{\tau_v}\left[V_{reference} - V_{pot}(t)\right]$$

slowly lowered

*R.C. van Schaik et al., J. Comp.-Aided Mol. Des. 6 (**1992**) 97-112*

Techniques to enhance the searching and sampling power of simulation methods

1. **Deformation or smoothening of the potential energy surface**
   a. omission of high-resolution structure factor data in structure refinement based on X-ray diffraction data
   b. gradual introduction of longer-range distance restraints in variable target structure refinement based on NMR NOE data
   c. softening of the hard core of atoms in the non-bonded interaction (soft-core atoms)
   d. reduction of the ruggedness of the energy surface through a diffusion-equation type of scaling
   e. avoiding the repeated sampling of an energy well through local potential energy elevation or conformational flooding
   f. softening of geometric restraints derived from experimental (NMR, X-ray) data through time-averaging of these
   g. circumvention of energy barriers through an extension of the dimensionality of the Cartesian space (4D-MD)
   h. freezing of high-frequency degrees of freedom through the use of constraints
   i. coarse-graining the model by reduction of the number of interaction sites

2. **Scaling of system parameters**
   a. temperature annealing
   b. tight coupling to heat bath
   c. mass scaling
   d. mean-field approaches

3. **Multi-copy searching and sampling**
   a. genetic algorithms
   b. replica-exchange and multi-canonical algorithms
   c. cooperative search: SWARM

*Review: J. Comput. Chem. 29 (**2007**) 157-166*

Use of soft-core non-bonded interactions

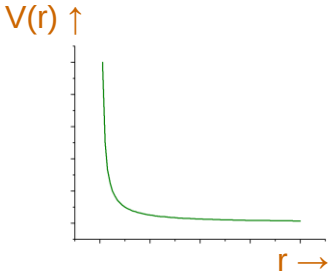
Thomas Beutler et al. Chem. Phys. Letters 222 (1994) 529-539

Use of non-physical potential energy terms

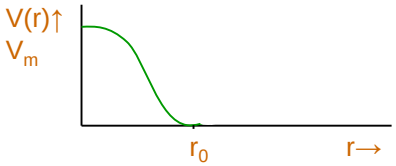
Physical non-bonded term:

van der Waals $\frac{C_{12}}{r_{ij}^{12}}$

Coulomb $\frac{q_i q_j}{r_{ij}}$



Non-physical softer non-bonded term that allows atoms to pass through each other:



Conditions: $V(0) = V_m \quad V'(0) = 0$

$V(r_0) = 0 \quad V'(r_0) = 0$

1. V(r) is a function of r:

$$f(r) = a + br + cr^2 + dr^3$$

$$V(r) = V_m \left[1 - 3\left(\frac{r}{r_0}\right)^2 + 2\left(\frac{r}{r_0}\right)^3 \right]$$

$$f'(r) = -6V_m \frac{r}{r_0^2}\left[1 - \left(\frac{r}{r_0}\right) \right]$$

2. V(r) is a function of r²:

$$g(r) = a + br^2 + cr^4$$

$$V(r) = V_m\left[1 - \left(\frac{r}{r_o}\right)^2\right]^2$$

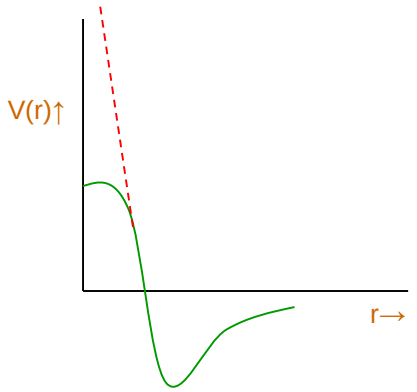
$$g'(r) = -4V_m \frac{r}{r_0^2}\left[1 - \left(\frac{r}{r_0}\right)^2\right]$$

3. V(r) is general van der Waals plus Coulomb plus reaction-field form:

$$V(r) = \frac{C_{12}}{\left[\alpha + r_{ij}^6\right]^2} - \frac{C_6}{\left[\alpha + r_{ij}^6\right]} + \frac{q_i q_j}{4\pi\varepsilon_0\varepsilon_r}\left(\frac{1}{\left[\alpha_c + r_{ij}^2\right]^{1/2}} - \frac{0.5C_{rf}r^2}{\left[\alpha_c + R_{rf}^2\right]^{1/2}} - \frac{1 - 0.5C_{rf}}{R_{rf}}\right)$$

$$\lim_{\alpha \to 0} = \text{standard form}$$

$$\alpha \neq 0: \quad V(0) = \text{finite} \qquad V'(0) = 0$$



Techniques to enhance the searching and sampling power of simulation methods

1. **Deformation or smoothening of the potential energy surface**
   a. omission of high-resolution structure factor data in structure refinement based on X-ray diffraction data
   b. gradual introduction of longer-range distance restraints in variable target structure refinement based on NMR NOE data
   c. softening of the hard core of atoms in the non-bonded interaction (soft-core atoms)
   d. reduction of the ruggedness of the energy surface through a diffusion-equation type of scaling
   e. avoiding the repeated sampling of an energy well through local potential energy elevation or conformational flooding
   f. softening of geometric restraints derived from experimental (NMR, X-ray) data through time-averaging of these
   g. circumvention of energy barriers through an extension of the dimensionality of the Cartesian space (4D-MD)
   h. freezing of high-frequency degrees of freedom through the use of constraints
   i. coarse-graining the model by reduction of the number of interaction sites

2. **Scaling of system parameters**
   a. temperature annealing
   b. tight coupling to heat bath
   c. mass scaling
   d. mean-field approaches

3. **Multi-copy searching and sampling**
   a. genetic algorithms
   b. replica-exchange and multi-canonical algorithms
   c. cooperative search: SWARM

*Review: J. Comput. Chem. 29 (**2007**) 157-166*

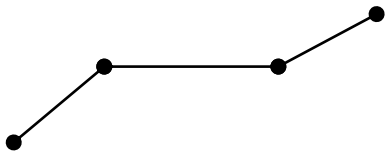
Methods to search conformational space

Idea: Include information obtained so far during the simulation into the search scheme: **memory function**

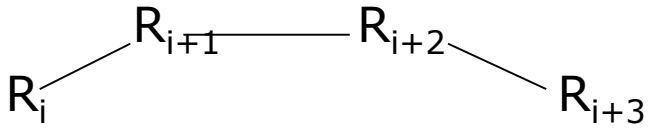
A. Characterize molecular conformations using:

- cartesian coordinates *too many*

- *torsional angles* φ, ψ, χ



- *dihedral angles* spanning residues:



*Review searching: M. Christen & W.F. van Gunsteren J. Comput. Chem. 29 (**2007**) 157 - 166*

B. Penalize the visited conformations by changing the energy function V as function of time

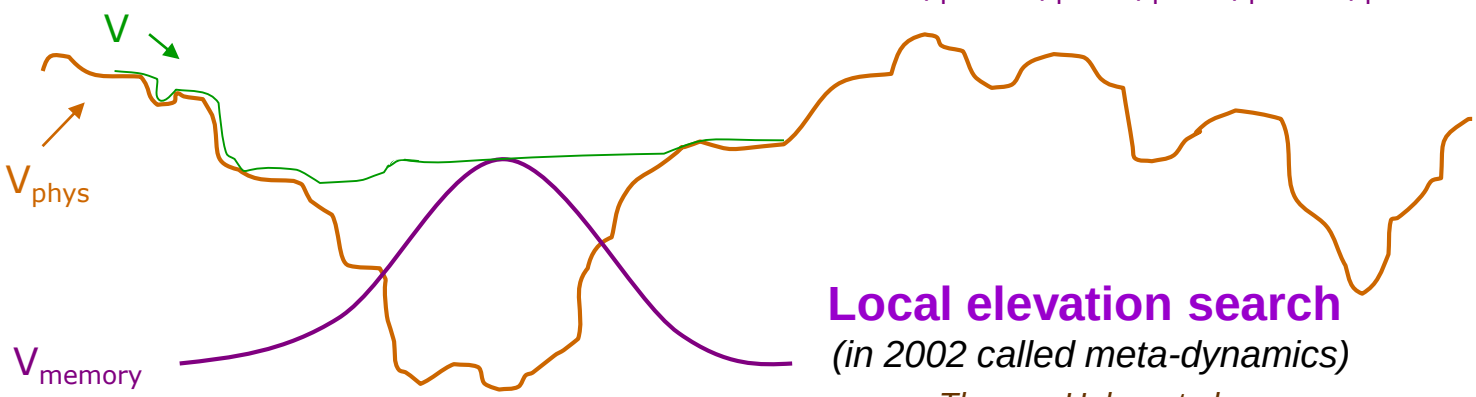
$$V\left(\{\varphi_i\}\right) = V_{phys}\left(\{\vec{r}_i\}\right) + V_{memory}\left(\{\varphi_i\}\right)$$

- potential energy term that pushes molecule out of the current conformation $\{\varphi_i^0\}$

$$V_{memory} = c \cdot \text{Number}\left(\varphi_i \text{ at } \varphi_i^0\right) \cdot e^{-\sum_i \left(\varphi_i - \varphi_i^0\right)^2 / 2\omega^2}$$

$$\downarrow$$

of conformations for which $\varphi_i^0 - \Delta\varphi_i < \varphi_i < \varphi_i^0 + \Delta\varphi_i$



Local elevation search

(in 2002 called meta-dynamics)

Thomas Huber et al.

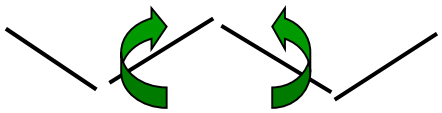
*J. Comp. Aided Mol. Design 8 (**1994**) 695*

Implementation

1. Use torsion angles, ϕ_i
2. Each conformation $\phi_1, \phi_2, \phi_3, \ldots, \phi_n = \phi^n$
3. Discretise to M parts $\rightarrow$ M^n grid points $\phi^n{}_0$
4. Gaussian function at grid points: $V_{mem}\left(\phi^n\right) = k_{mem} N_{\phi_0^n} e^{-\frac{\left(\phi^n - \phi_0^n\right)^2}{2w^2}}$
5. $V_{total} = V_{phys} + V_{mem}$

A toy application

Pentane (two torsional angles)



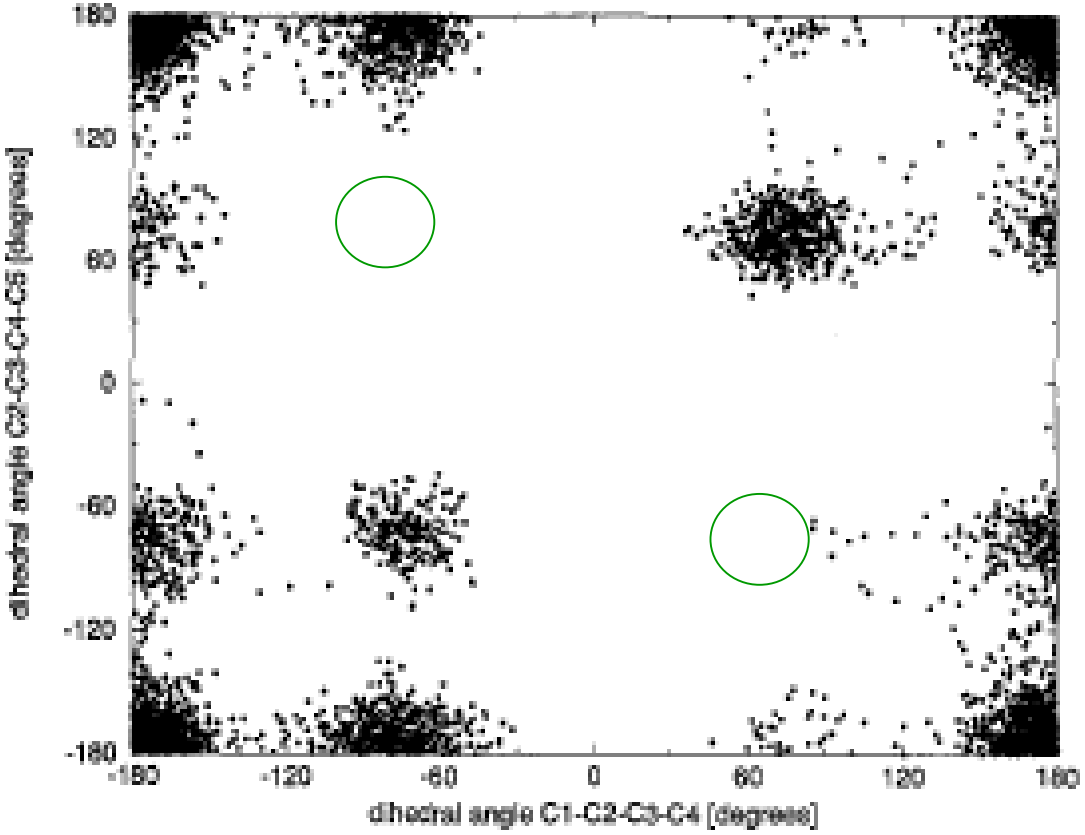
Complete space can be mapped out

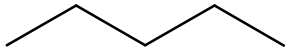
Test case: pentane

Thomas Huber et al.
J. Comp. Aided Mol. Design 8 (1994) 695

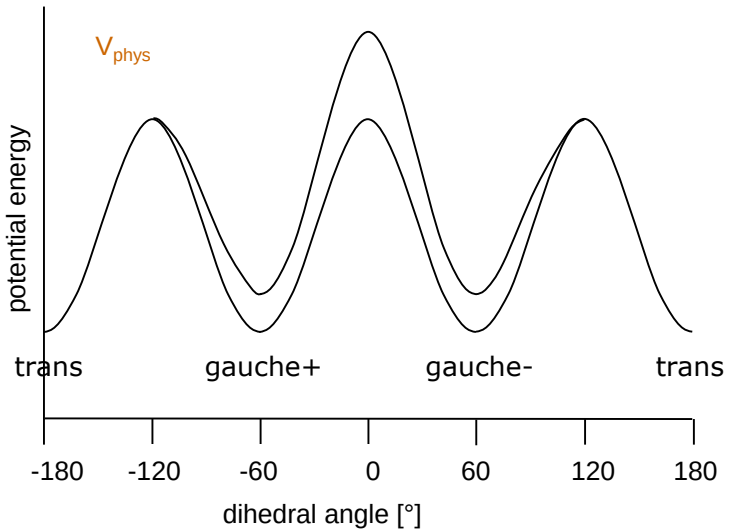
2 dihedral angles (3 minima each) → 9 low V_{phys} conformers

free SD-simulation (united atoms)
simulation time 100 ps, T=300 K, GROMOS force field





trans - trans *lowest*



cis - cis | gauche+ gauche+ | gauche+ gauche-

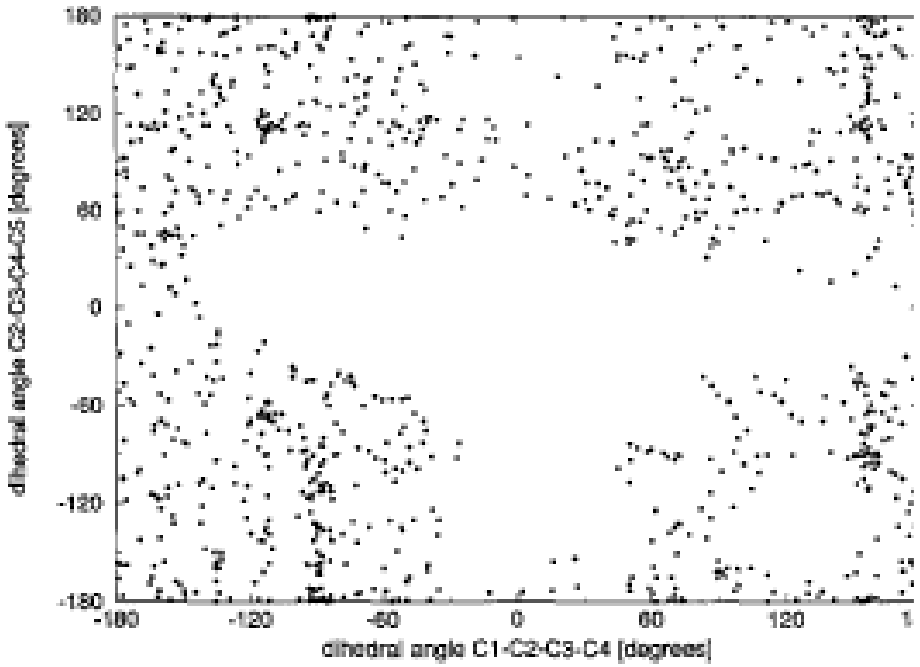
highest *low* *higher*

Higher-energy conformers are not (yet) sampled in 100 ps *normal* MD(SD) simulation

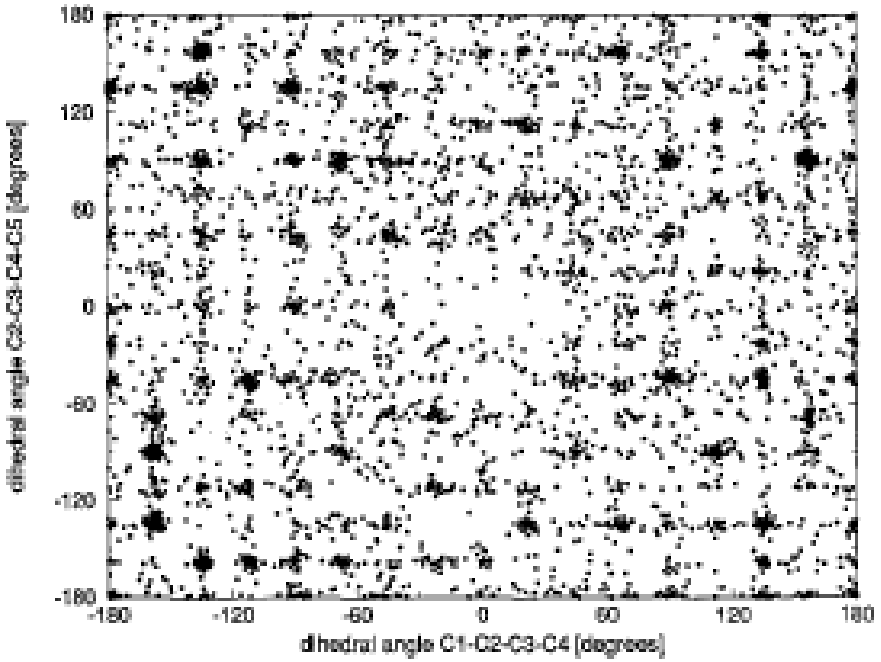
Local elevation search: pentane

Local-elevation simulation of pentane (united atoms) T=300 K,
Gaussian local-elevation function with k=5kJ/mol per MD step

simulation time **20ps**



simulation time **100ps**



Higher-energy conformations *are* sampled in 20 ps *local-elevation MD simulation*

Almost all conformations are sampled in 100 ps *LE*-MD simulation

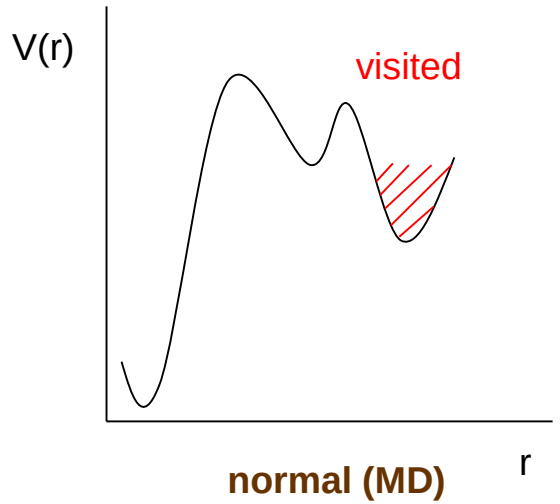
The local elevation simulation method

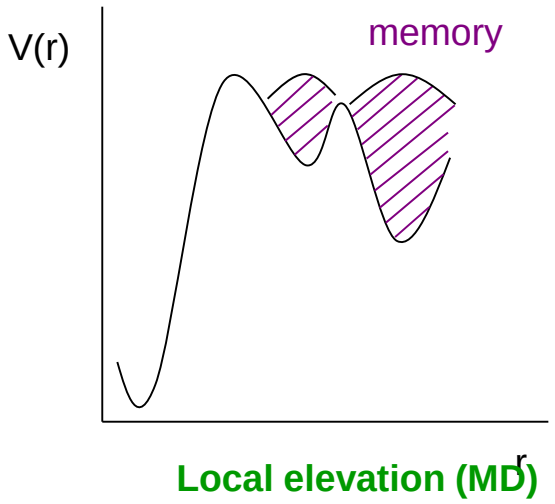
Normal simulation: relevant properties

- Many conformers
  few visited
- Compact representation should be possible

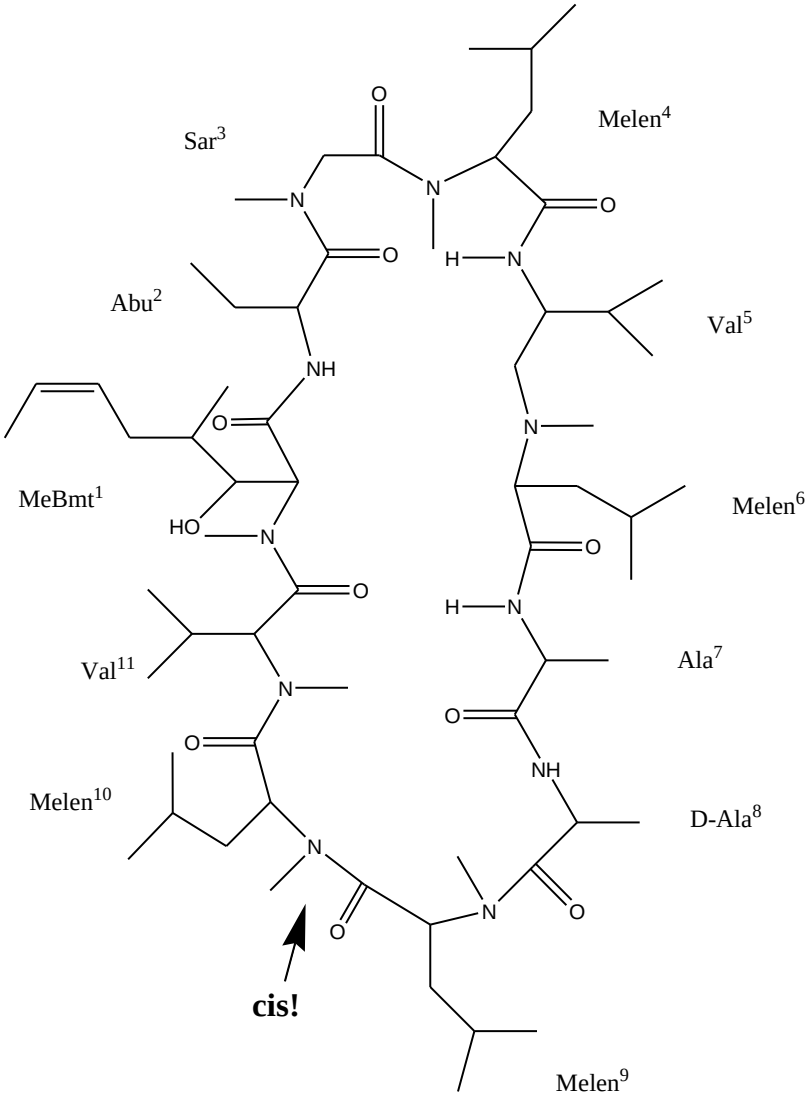
Local-elevation simulation:

- run simulation
- store visited conformations (using compact representation)
- push system away when old conformation is seen



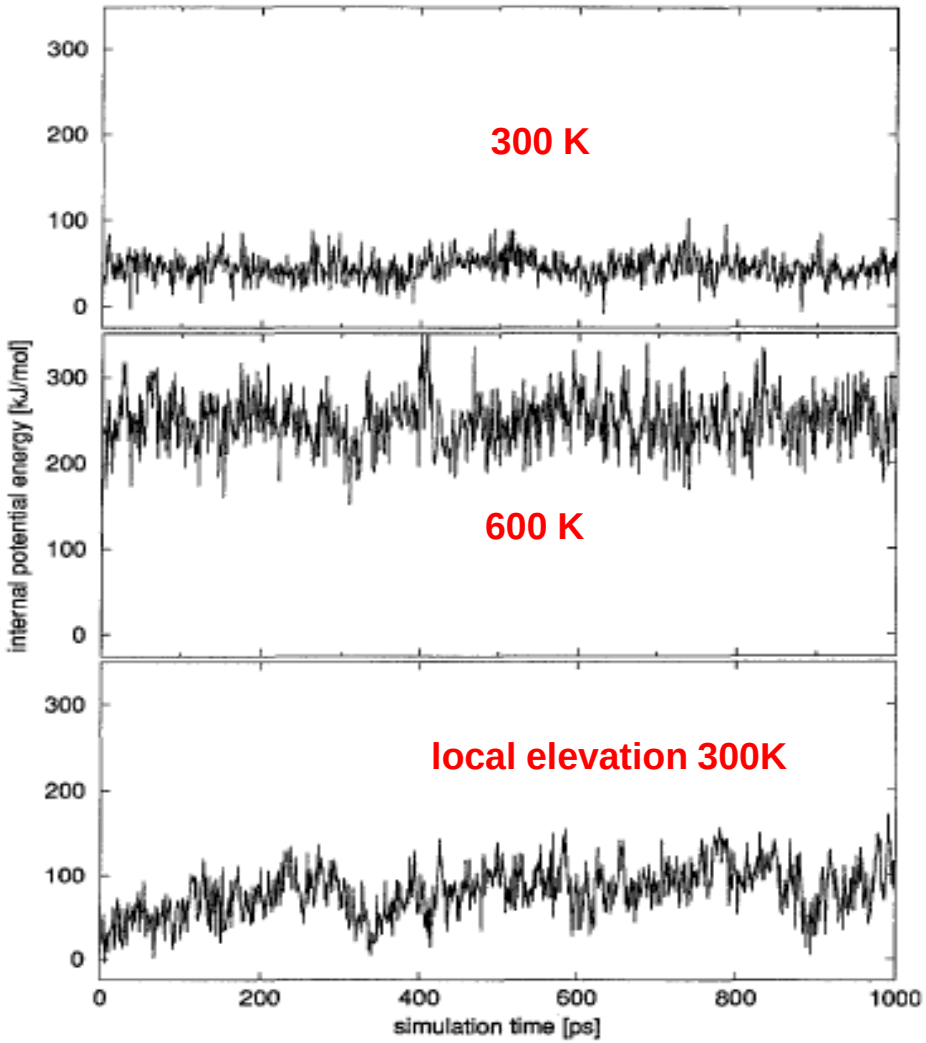


Cyclosporin A



- Amid bond (fixed to trans) ω-dihedral
- central bond of φ-dihedral
- central bond of ψ-dihedral

Cyclosporin A: potential energy



MD

MD

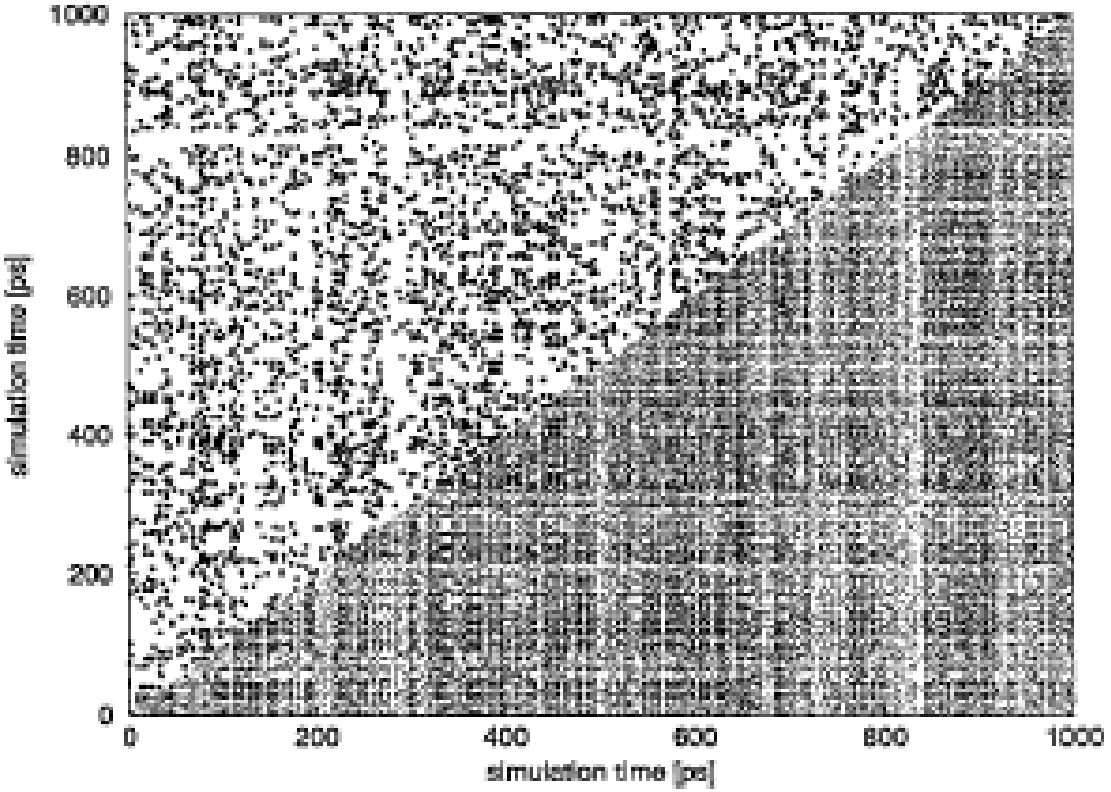
LE-MD

Cyclosporin A: similarity of conformations

Criterion: $\Delta\varphi_i$ ≤ 30° (upper)
≤ 45° (lower) } for each of the 11 φ-angles

164 = average number of visits of same conformer

SD simulation at 300 K

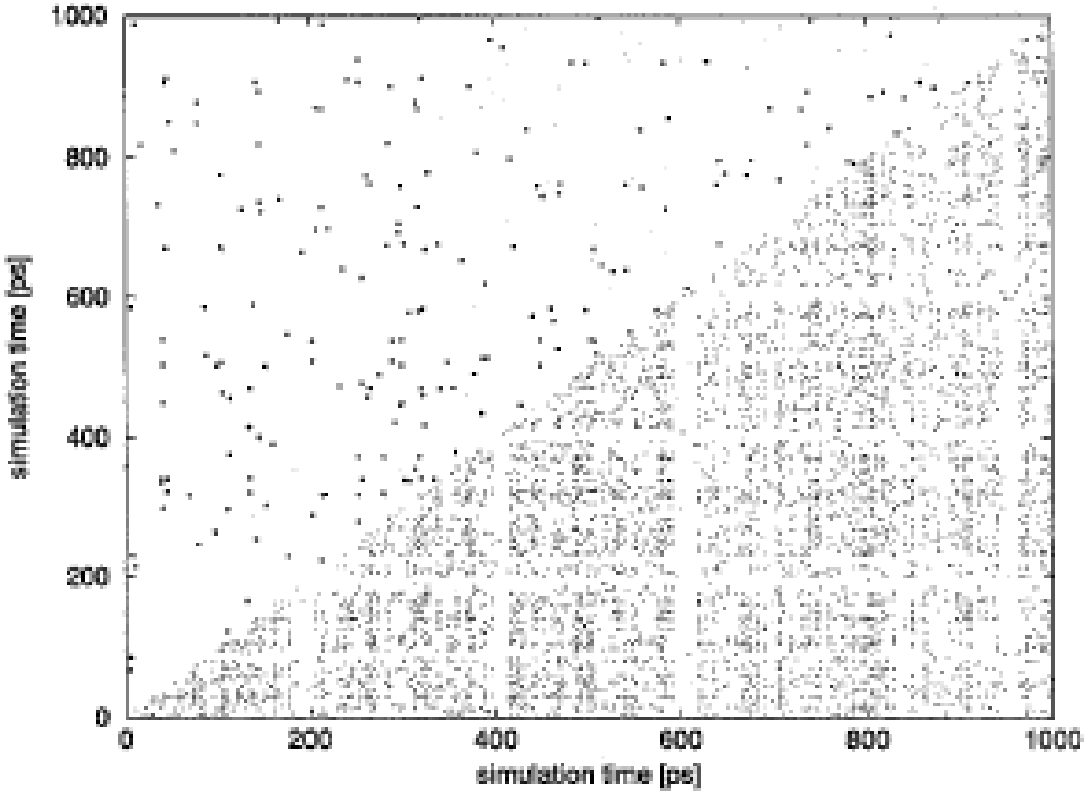


Cyclosporin A: similarity of conformations

Criterion: $\Delta\varphi_i$ ≤ 30° (upper)
≤ 45° (lower)
} for each of the 11 φ-angles

26 visits on average

SD simulation at *600* K

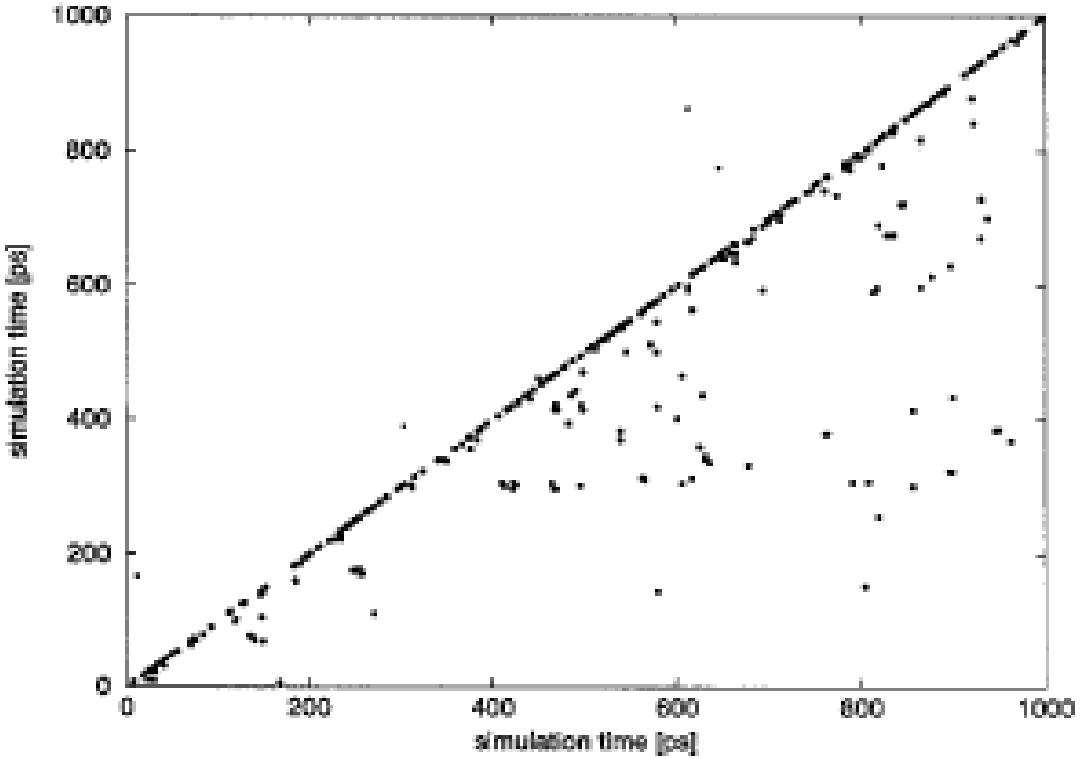


Cyclosporin A: similarity of conformations

Criterion: $\Delta\varphi_i$ ≤ *45°* (upper)
≤ *60°* (lower) } for each of the 11 φ-angles

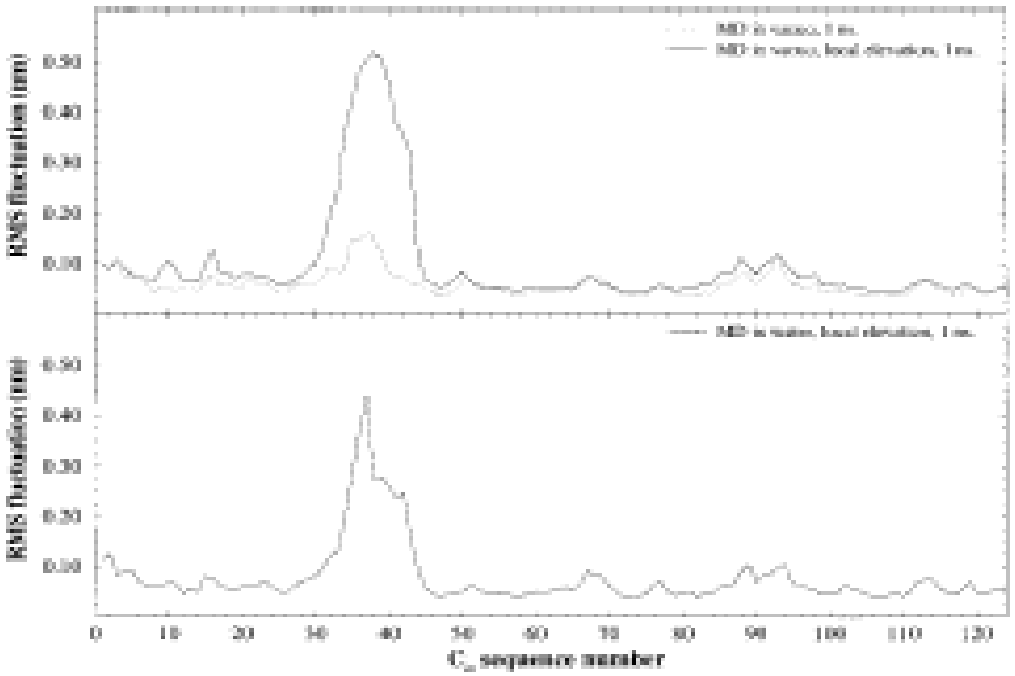
1.6 visit on average

Local elevation simulation at *300* K



Ribonuclease A: RMS fluctuations in the loop region

Loop search



in vacuo

dashed: MD
solid: LE-MD

in solution

solid: LE-MD

Figure 2. Root mean square fluctuations of atomic positions in ribonuclease A. Upper graph: simulations in vacuo with and without local elevation search. The simulation using local elevation search produces larger positional fluctuations than without, which is indicative of the larger conformational space searched. Lower graph: simulation in solvent with local elevation search. The fluctuations are smaller than those encountered in the vacuum simulations.

Local-elevation MD searches a much larger conformational space

Ribonuclease A: loop conformations

Standard MD **Local-elevation MD**

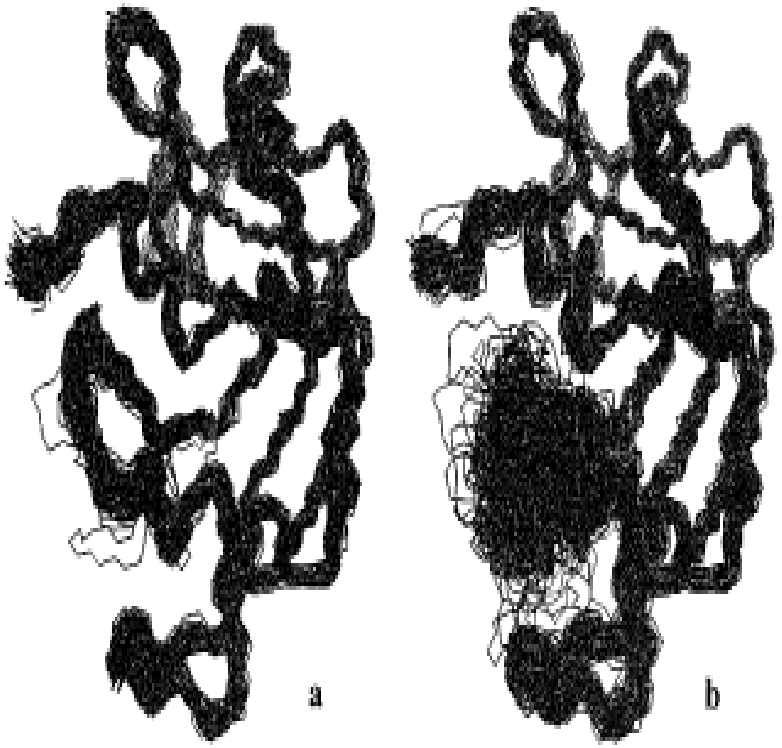


Figure 3. Superposition of 200 conformations of ribonuclease A taken at 5 ps intervals (total of 1 ns simulation time) in vacuo: (a) simulation without local elevation; (b) with local elevation search in the loop consisting of residues 33–43. The larger conformational space searched is apparent.

Local-elevation MD searches a much larger conformational space

*Scott et al., J. Phys. Chem. A103 (**1999**) 3596-3607*

Techniques to enhance the searching and sampling power of simulation methods

1. **Deformation or smoothening of the potential energy surface**
   a. omission of high-resolution structure factor data in structure refinement based on X-ray diffraction data
   b. gradual introduction of longer-range distance restraints in variable target structure refinement based on NMR NOE data
   c. softening of the hard core of atoms in the non-bonded interaction (soft-core atoms)
   d. reduction of the ruggedness of the energy surface through a diffusion-equation type of scaling
   e. avoiding the repeated sampling of an energy well through local potential energy elevation or conformational flooding
   f. softening of geometric restraints derived from experimental (NMR, X-ray) data through time-averaging of these
   g. circumvention of energy barriers through an extension of the dimensionality of the Cartesian space (4D-MD)
   h. freezing of high-frequency degrees of freedom through the use of constraints
   i. coarse-graining the model by reduction of the number of interaction sites

2. **Scaling of system parameters**
   a. temperature annealing
   b. tight coupling to heat bath
   c. mass scaling
   d. mean-field approaches

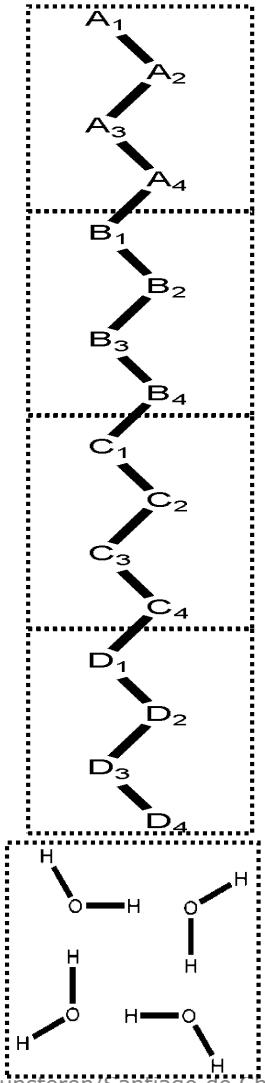
3. **Multi-copy searching and sampling**
   a. genetic algorithms
   b. replica-exchange and multi-canonical algorithms
   c. cooperative search: SWARM

*Review: J. Comput. Chem. 29 (**2007**) 157-166*

Coarse-grained versus fine-grained models

liquid alkanes: hexadecane

AL (λ=0)
All-atom model
(non-hydrogen)
16 (CH_2 or CH_3) atoms



MAP
"mapped"
all-atom
configurations

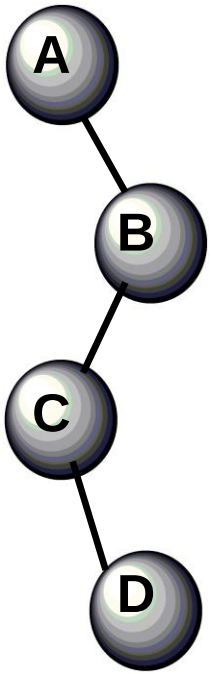
Centre of mass
$A_1 - A_4$

Centre of mass
$B_1 - B_4$

Centre of mass
$C_1 - C_4$

Centre of mass
$D_1 - D_4$

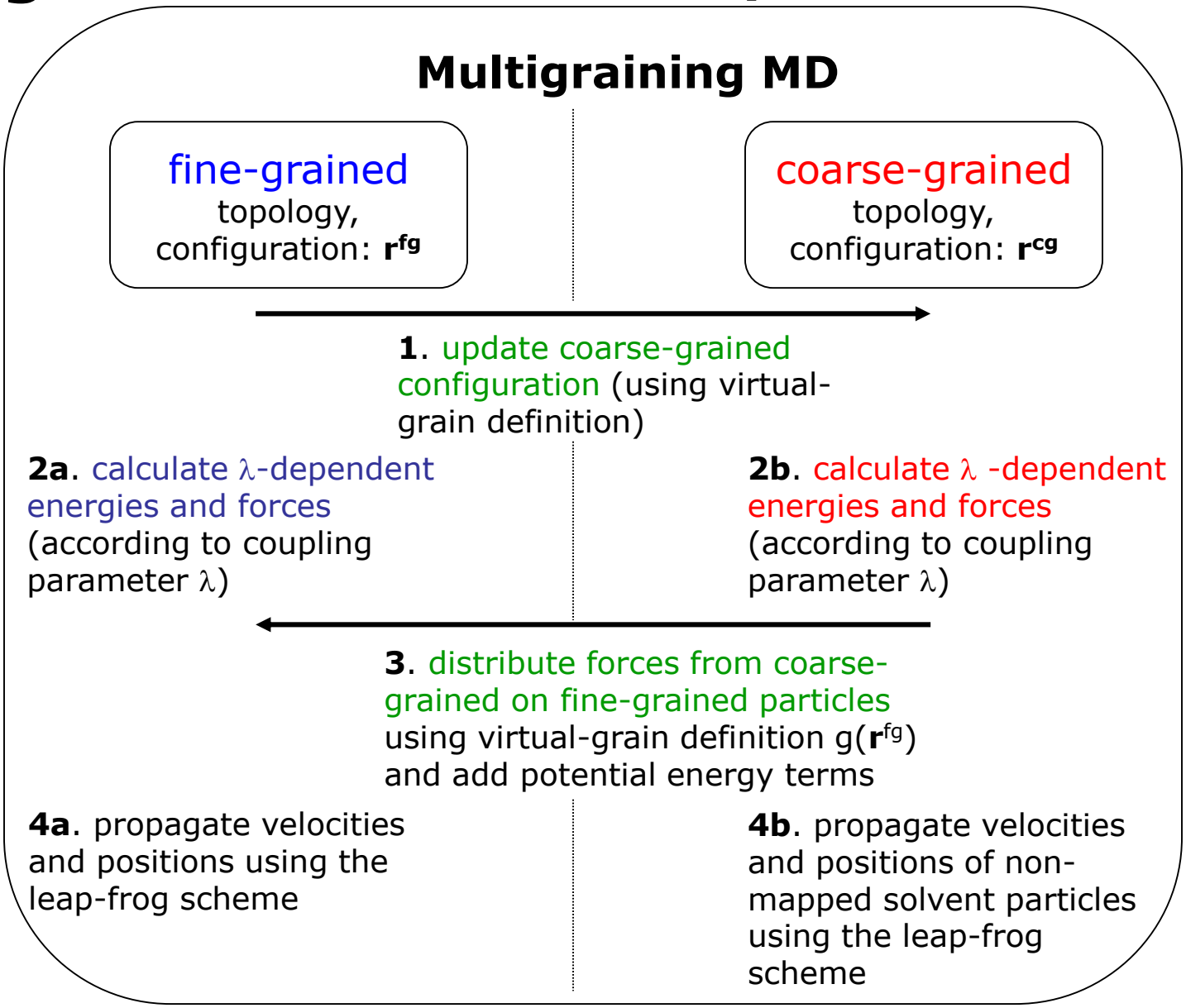
CG (λ=1)
Coarse-grained model
4 atoms



Compare: - structural characteristics
- energetic / entropic characteristics

*M. Christen & WFvG, J. Chem. Phys. 124 (**2006**) 154106*

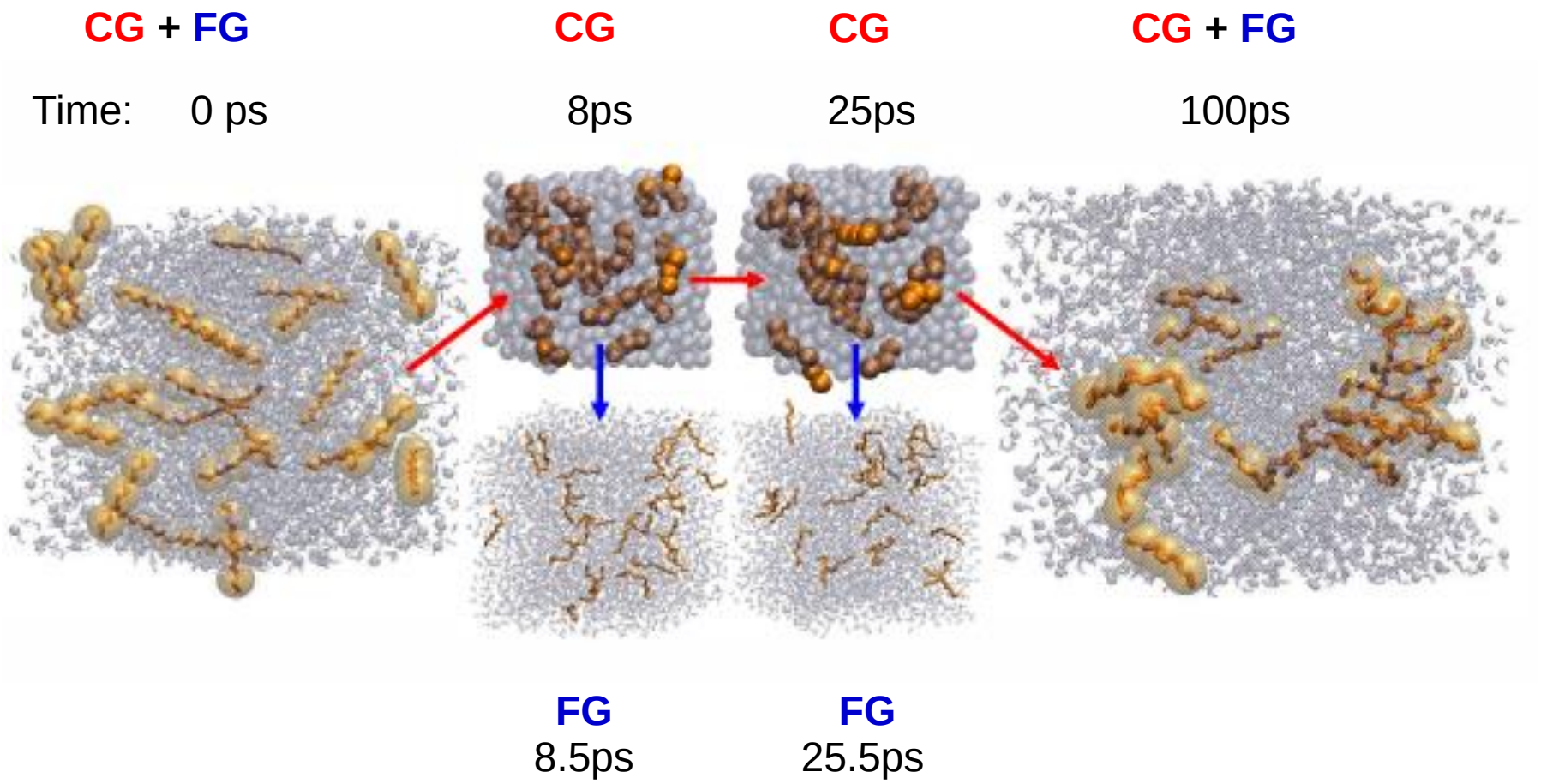
Algorithm for mixed FG/CG simulation



*M. Christen & WFvG, J. Chem. Phys. 124 (**2006**) 154106*

Multi-grained simulation of 25 hexadecanes in water

*M. Christen & W.F. van Gunsteren, J. Chem. Phys, **124** (2006) 154106*



CG level simulation with occasional switching to FG level enhances exploration of FG conformational space

Techniques to enhance the searching and sampling power of simulation methods

1. **Deformation or smoothening of the potential energy surface**
   a. omission of high-resolution structure factor data in structure refinement based on X-ray diffraction data
   b. gradual introduction of longer-range distance restraints in variable target structure refinement based on NMR NOE data
   c. softening of the hard core of atoms in the non-bonded interaction (soft-core atoms)
   d. reduction of the ruggedness of the energy surface through a diffusion-equation type of scaling
   e. avoiding the repeated sampling of an energy well through local potential energy elevation or conformational flooding
   f. softening of geometric restraints derived from experimental (NMR, X-ray) data through time-averaging of these
   g. circumvention of energy barriers through an extension of the dimensionality of the Cartesian space (4D-MD)
   h. freezing of high-frequency degrees of freedom through the use of constraints
   i. coarse-graining the model by reduction of the number of interaction sites

2. **Scaling of system parameters**
   a. temperature annealing
   b. tight coupling to heat bath
   c. mass scaling
   d. mean-field approaches

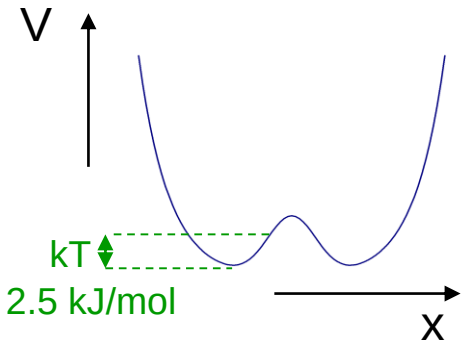
3. **Multi-copy searching and sampling**
   a. genetic algorithms
   b. replica-exchange and multi-canonical algorithms
   c. cooperative search: SWARM

*Review: J. Comput. Chem. 29 (**2007**) 157-166*

Search enhancement techniques

Scale system parameters

1. Search at high temperature, T annealing



e^{-V/k_BT} higher T, lower V $\Rightarrow$

in principle: equivalent

in practice: T global

V local

T 10 × higher $\xrightarrow{k_BT=mv^2}$ v 3 × higher $\longrightarrow$ Δt 3 × smaller, CPU 3 × more

More barrier crossings:

1. Wanted O.K.
2. Unwanted e.g. chirality →
   → V_{chir} larger →
   → more rigidity →
   → less barrier crossing

annealing: lower T infinitely slow

Scaling of system parameters

2. Tight coupling to a temperature bath:

velocities are maintained going up hill →

→ **more barrier crossings ?**

3. Scaling of atomic masses:

Equilibrium properties are independent of masses

increase some masses → more inertia →

→ **more barrier crossings ?**

4. Enhanced sampling via a mean-field approach:

Example: *Huber et al., Biopolymers **39** (1996) 103-114*

Optimization methods for conformational sampling using a Boltzmann-weighted mean-field approach

Techniques to enhance the searching and sampling power of simulation methods

1. **Deformation or smoothening of the potential energy surface**
   a. omission of high-resolution structure factor data in structure refinement based on X-ray diffraction data
   b. gradual introduction of longer-range distance restraints in variable target structure refinement based on NMR NOE data
   c. softening of the hard core of atoms in the non-bonded interaction (soft-core atoms)
   d. reduction of the ruggedness of the energy surface through a diffusion-equation type of scaling
   e. avoiding the repeated sampling of an energy well through local potential energy elevation or conformational flooding
   f. softening of geometric restraints derived from experimental (NMR, X-ray) data through time-averaging of these
   g. circumvention of energy barriers through an extension of the dimensionality of the Cartesian space (4D-MD)
   h. freezing of high-frequency degrees of freedom through the use of constraints
   i. coarse-graining the model by reduction of the number of interaction sites

2. **Scaling of system parameters**
   a. temperature annealing
   b. tight coupling to heat bath
   c. mass scaling
   d. mean-field approaches

3. **Multi-copy searching and sampling**
   a. genetic algorithms
   b. replica-exchange and multi-canonical algorithms
   c. cooperative search: SWARM

*Review: J. Comput. Chem. 29 (**2007**) 157-166*

Metropolis Monte Carlo method

Algorithm:

DO large number of steps n=1,2,….

1. Make a step $\Delta\vec{r}$ in configuration space: $\vec{r}^{\,N}_{n+1} = \vec{r}^{\,N}_{n} + \Delta\vec{r}$

2. Calculate the change in potential energy ΔE: $E_n = V\left(\{\vec{r}^{\,N}_{n}\}\right)$

$$\Delta E = E_{n+1} - E_n$$

3. IF ($\Delta E \leq 0$) THEN

   Accept new configuration $\vec{r}^{\,N}_{n+1}$

   ELSE IF ($e^{-\Delta E/k_B T}$ > random number ε(0,1)) THEN

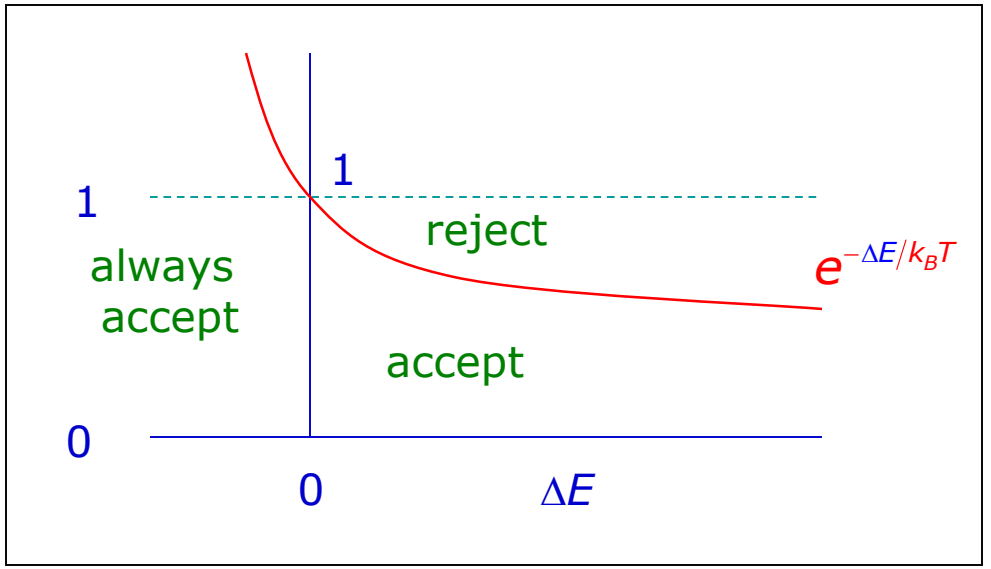
   Accept $\vec{r}^{\,N}_{n+1}$

   ELSE

   Take $\vec{r}^{\,N}_{n}$ again.

   END IF

END DO



Metropolis Monte Carlo method

Rationale: consider the probability of **transiton** $t_{x\to y}$ from state **x** to state **y**

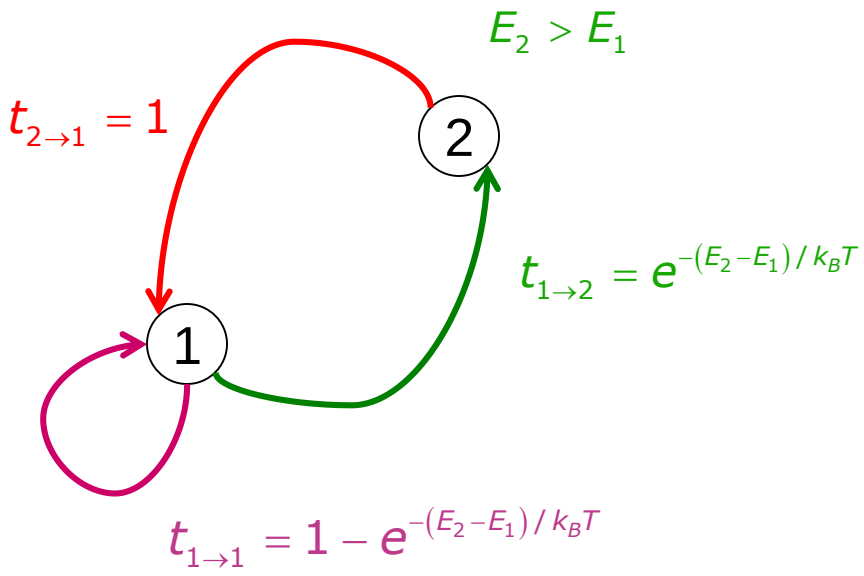
Equilibrium:

$$P_2 \cdot t_{2\to 1} = P_1 \cdot t_{1\to 2}$$

$$P_2 \cdot 1 = P_1 \cdot e^{-(E_2-E_1)/k_BT}$$

or

$$\frac{P_2}{P_1} = \frac{e^{-E_2/k_BT}}{e^{-E_1/k_BT}}$$



Each configuration occurs with **Boltzmann** probability.

Note:

- **Assumption**: *detailed balance between states*
- MC steps must cover Cartesian space **uniformly**: *dxdydz*
- Using polar coordinates, sample $\propto r^2 \sin\theta dr d\theta d\varphi$: ***non-uniform*** *sampling*
- **Unphysical steps** are possible: exchange of particles

Replica-exchange simulation

Idea: - **A number of replicas of a system that do *not* interact with each other**, is simulated simultaneously.

- The **replicas** are
  a) at different thermodynamic state points
     e.g. different temperatures (T_s), or
  b) characterised by different Hamiltonians,
     e.g. interatomic interactions, $H(p^N, r^N; \lambda_s)$

- **From time to time, after time period τ_{ex}, a Monte Carlo exchange of configurations r_m and r_n between two replicas m and n is attempted** with exchange probability (assumption: *detailed balance*)

$$p(m \leftrightarrow n) = min(1, exp(-\Delta_{mn})) = \begin{cases} 1 & for\ \Delta_{mn} \le 0 \\ exp(-\Delta_{mn}) & for\ \Delta_{mn} > 0 \end{cases}$$

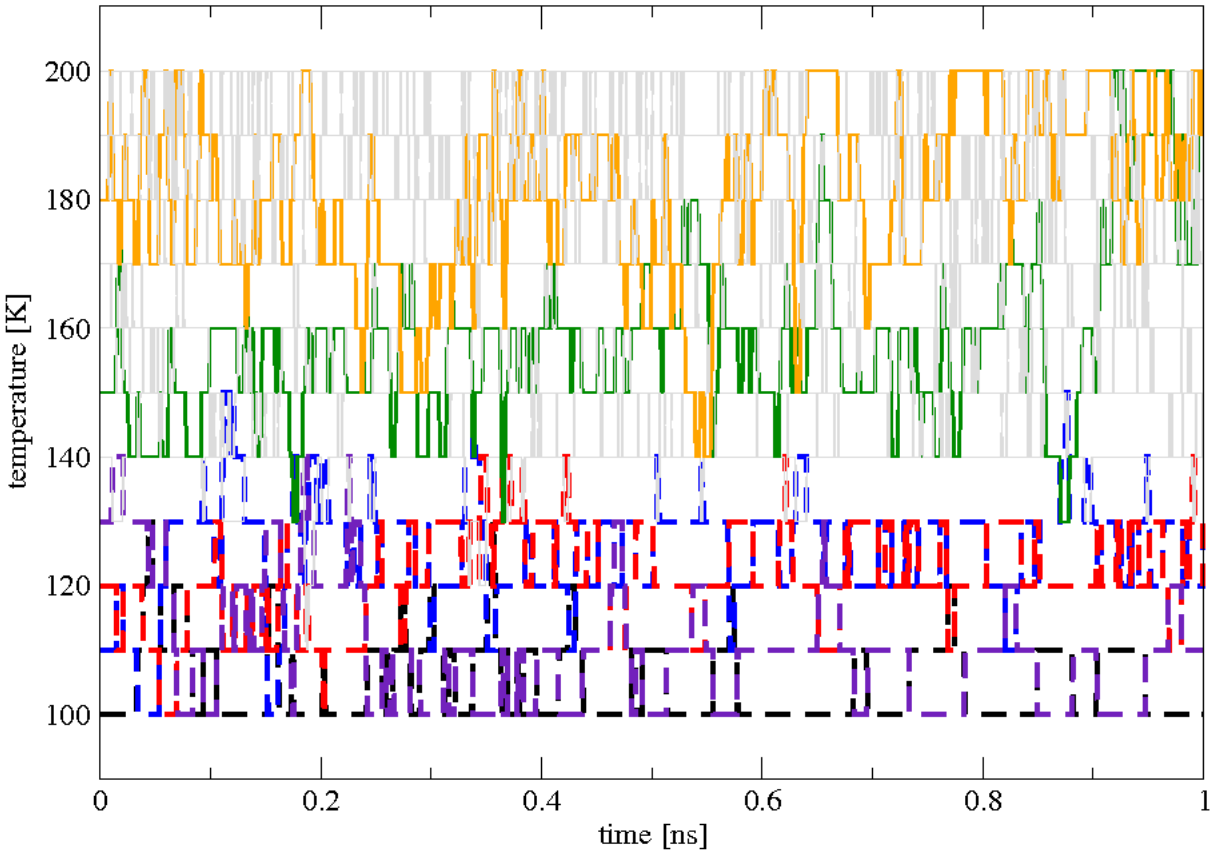
with (T-REMD) $$\Delta_{mn} = \frac{\rho_{T_m}(\vec{p}_n^N, \vec{r}_n^N)\rho_{T_n}(\vec{p}_m^N, \vec{r}_m^N)}{\rho_{T_m}(\vec{p}_m^N, \vec{r}_m^N)\rho_{T_n}(\vec{p}_n^N, \vec{r}_n^N)}$$

with (H-REMD) $$\Delta_{mn} = \frac{\rho_{H_m}(\vec{p}_n^N, \vec{r}_n^N)\rho_{H_n}(\vec{p}_m^N, \vec{r}_m^N)}{\rho_{H_m}(\vec{p}_m^N, \vec{r}_m^N)\rho_{H_n}(\vec{p}_n^N, \vec{r}_n^N)}$$

and probability density $$\rho(\vec{p}^N, \vec{r}^N) \propto exp(-H(\vec{p}^N, \vec{r}^N)/k_B T)$$

Temperature replica-exchange SD-simulation of 11 replicas (ΔT=10K) of a box with 512 n-butanes

exchange frequency: $\tau_{ex} = 1ps^{-1}$

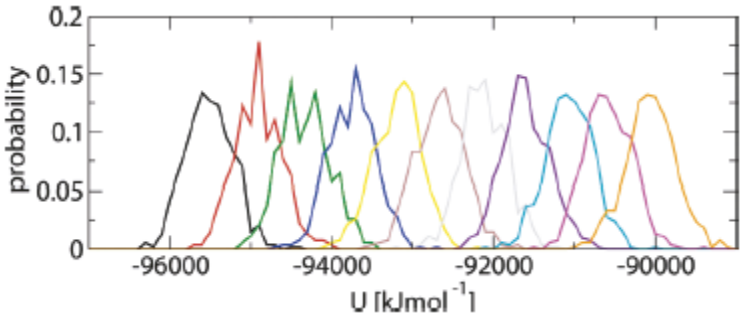


Not all replicas are exchanging ⟶ **ΔT too large**

Temperature replica-exchange MD simulation of 11 replicas (ΔT=5K) of a box with Ca^{2+} and SO_4^{2-} ions in water (2002 molecules)

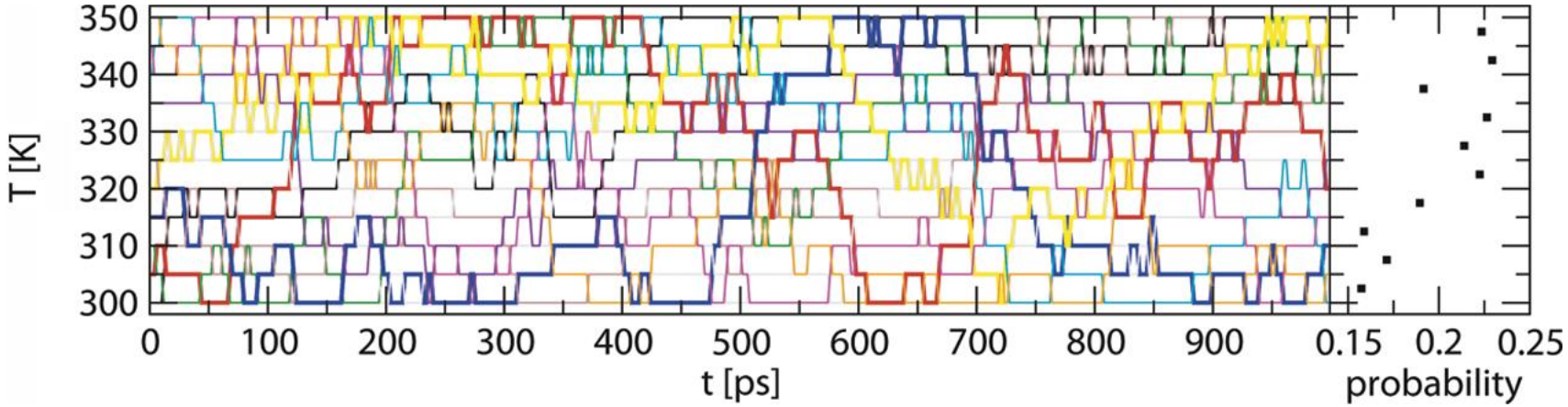
exchange frequency: $\tau_{ex} = 2ps^{-1}$

1. Are the potential energy distributions overlapping ? *YES*



T	D (Ca^{2+})	D (SO_4^{2-})
K	*nm/ns*	*nm/ns*
300	**1.06**	**1.23**
325	1.17	1.31
350	**1.63**	**1.84**

2. Are the replicas exchaging sufficiently often ? *YES*



3. Is the ion diffusion enhanced by more than a factor $N_{replicas} = 11$?

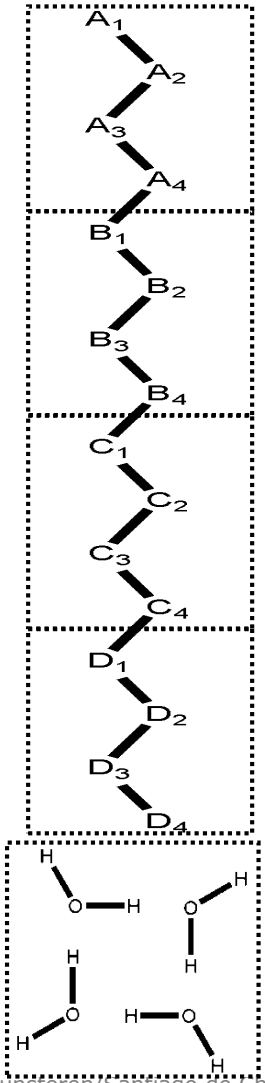
NO: T-replica-exchange MD is not more efficient than standard MD in this case

*A.P.E. Kunz & WFvG, J. Phys. Chem B 115 (**2011**) 2931-2936*

Coarse-grained versus fine-grained models

liquid alkanes: hexadecane

AL (λ=0)
All-atom model
(non-hydrogen)
16 (CH_2 or CH_3) atoms



MAP
"mapped"
all-atom
configurations

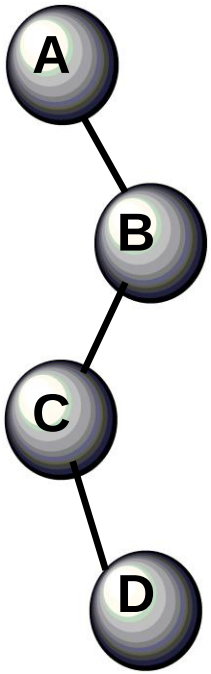
Centre of mass
$A_1 - A_4$

Centre of mass
$B_1 - B_4$

Centre of mass
$C_1 - C_4$

Centre of mass
$D_1 - D_4$

CG (λ=1)
Coarse-grained model
4 atoms

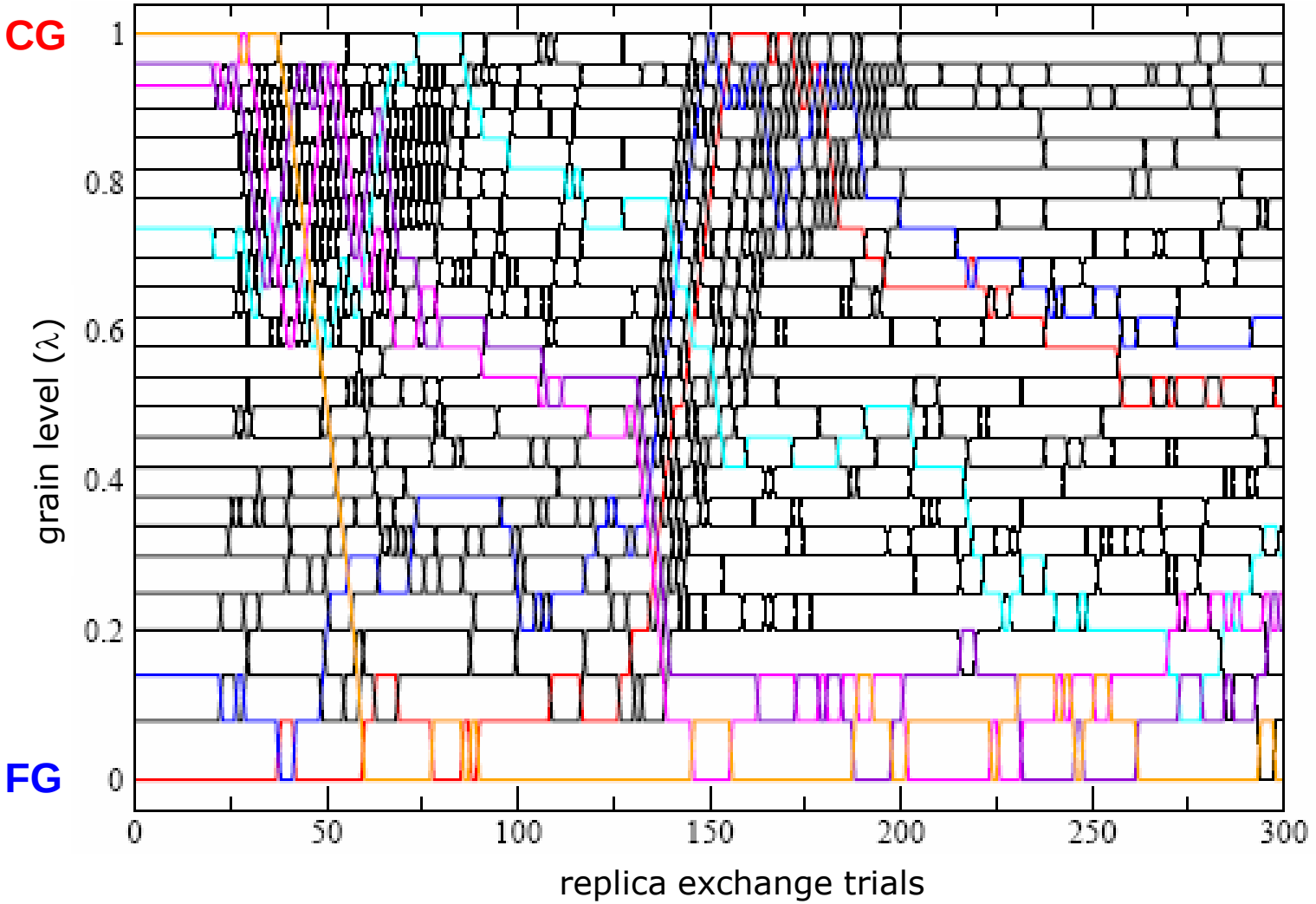


Compare: - structural characteristics
- energetic / entropic characteristics

*M. Christen & WFvG, J. Chem. Phys. 124 (**2006**) 154106*

Multi-grained simulation of liquid octane

grain level of the 24 replicas during 300 replica exchange steps

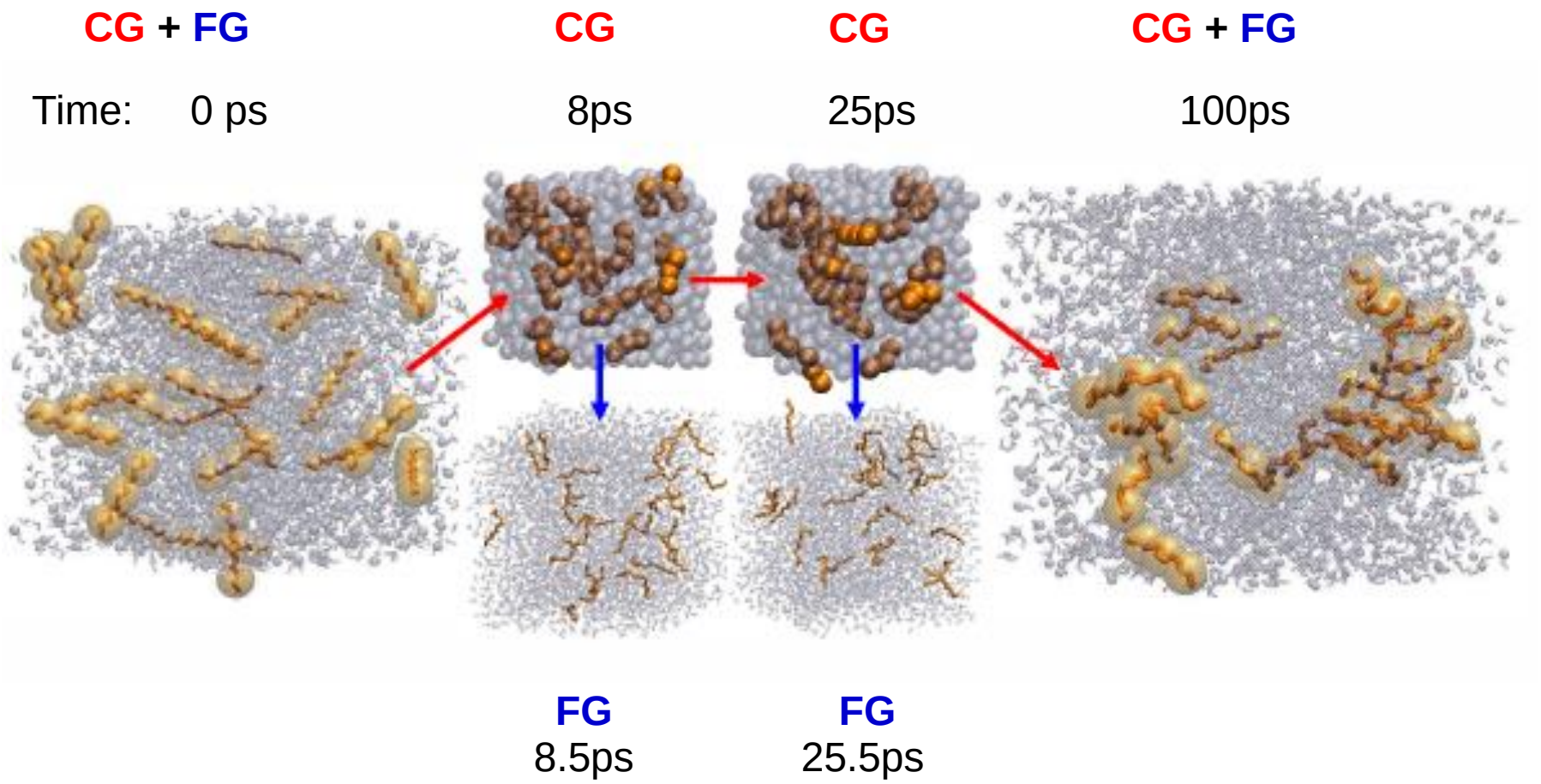


FG/CG replica-exchange simulation enhances sampling

*M. Christen & WFvG, J. Chem. Phys. 124 (**2006**) 154106*

Multi-grained simulation of 25 hexadecanes in water

*M. Christen & W.F. van Gunsteren, J. Chem. Phys, **124** (2006) 154106*



CG level simulation with occasional switching to FG level enhances exploration of FG conformational space

Techniques to enhance the searching and sampling power of simulation methods

1. **Deformation or smoothening of the potential energy surface**
   a. omission of high-resolution structure factor data in structure refinement based on X-ray diffraction data
   b. gradual introduction of longer-range distance restraints in variable target structure refinement based on NMR NOE data
   c. softening of the hard core of atoms in the non-bonded interaction (soft-core atoms)
   d. reduction of the ruggedness of the energy surface through a diffusion-equation type of scaling
   e. avoiding the repeated sampling of an energy well through local potential energy elevation or conformational flooding
   f. softening of geometric restraints derived from experimental (NMR, X-ray) data through time-averaging of these
   g. circumvention of energy barriers through an extension of the dimensionality of the Cartesian space (4D-MD)
   h. freezing of high-frequency degrees of freedom through the use of constraints
   i. coarse-graining the model by reduction of the number of interaction sites

2. **Scaling of system parameters**
   a. temperature annealing
   b. tight coupling to heat bath
   c. mass scaling
   d. mean-field approaches

3. **Multi-copy searching and sampling**
   a. genetic algorithms
   b. replica-exchange and multi-canonical algorithms
   c. cooperative search: SWARM

*Review: J. Comput. Chem. 29 (**2007**) 157-166*

Multi-copy search techniques: the SWARM method

Idea: combine a swarm of molecules with molecular trajectories into a cooperative system that searches conformational space (like a swarm of insects)

Implementation:

each molecule is, in addition to the physical forces,

subject to (artificial) forces that drive the trajectory of each molecule toward an average of the trajectories of the swarm of molecules

*Huber and van Gunsteren: J.Phys.Chem. A102 (**1998**) 5937-5943*

SWARM-MD: Searching configurational space by cooperative molecular dynamics

Methods for searching configuration space for configurations r^N with low $V(r^N)$=energy

I. Molecular coordinates as variables:

A. Systematic or exhaustive search

$\vec{r}^{\,N}$ cartesian

φ^N angles

- Scan complete space (SS)

small molecules

B. Heuristic search

- Generate tiny set of representative conformers:

1. Non-step methods

Distance geometry algorithms (DG)

Distribution? Solvent?

2. Step methods: change of a complete configuration

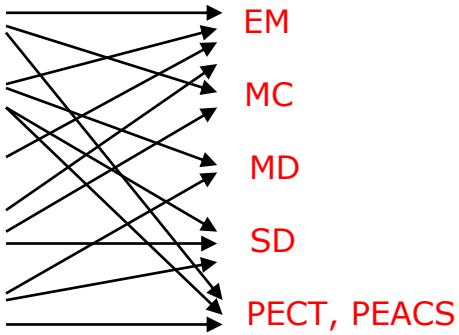
a. Energy: V(x)

b. Gradient: $-\frac{\partial V}{\partial x}$

c. 2nd Derivative: $\frac{\partial^2 V}{\partial x \partial x'}$

d. Random:

e. Memory:



3. **Step methods: build-up of a configuration**

configurational bias MC

combinatorial chain growing

Techniques to enhance the searching and sampling power of simulation methods

1. **Deformation or smoothening of the potential energy surface**
   a. omission of high-resolution structure factor data in structure refinement based on X-ray diffraction data
   b. gradual introduction of longer-range distance restraints in variable target structure refinement based on NMR NOE data
   c. softening of the hard core of atoms in the non-bonded interaction (soft-core atoms)
   d. reduction of the ruggedness of the energy surface through a diffusion-equation type of scaling
   e. avoiding the repeated sampling of an energy well through local potential energy elevation or conformational flooding
   f. softening of geometric restraints derived from experimental (NMR, X-ray) data through time-averaging of these
   g. circumvention of energy barriers through an extension of the dimensionality of the Cartesian space (4D-MD)
   h. freezing of high-frequency degrees of freedom through the use of constraints
   i. coarse-graining the model by reduction of the number of interaction sites

2. **Scaling of system parameters**
   a. temperature annealing
   b. tight coupling to heat bath
   c. mass scaling
   d. mean-field approaches

3. **Multi-copy searching and sampling**
   a. genetic algorithms
   b. replica-exchange and multi-canonical algorithms
   c. cooperative search: SWARM

*Review: J. Comput. Chem. 29 (**2007**) 157-166*

Biased sampling

Problem: If $V(\vec{r}^{\,N})$ is large for given regions in configurational space, the sampling in MC, MD, SD simulations will be poor.

Remedy: Add a biasing potential energy term $V^{BIAS}(\vec{r}^{\,N})$ to the (physical) Hamiltonian $H(\vec{r}^{\,N}, \vec{p}^{\,N})$, possibly based on experimental data, that focuses the sampling on a given region of configurational space.

Ensemble average of a quantity $Q(\vec{r}^{\,N})$:

$$< Q >= \frac{\iint Q(\vec{r}^{\,N}) e^{-H(\vec{r}^{\,N},\vec{p}^{\,N})/k_B T} d\vec{p}^{\,N} d\vec{r}^{\,N}}{\iint e^{-H(\vec{r}^{\,N},\vec{p}^{\,N})/k_B T} d\vec{p}^{\,N} d\vec{r}^{\,N}}$$

$$= \frac{\iint Q e^{+V^{BIAS}/k_B T} e^{-\left(H+V^{BIAS}\right)/k_B T} d\vec{r}^{\,N} d\vec{p}^{\,N}}{\iint e^{-\left(H+V^{BIAS}\right)/k_B T} d\vec{r}^{\,N} d\vec{p}^{\,N}} \bullet \frac{\iint e^{-\left(H+V^{BIAS}\right)/k_B T} d\vec{r}^{\,N} d\vec{p}^{\,N}}{\iint e^{+V^{BIAS}/k_B T} e^{-\left(H+V^{BIAS}\right)/k_B T} d\vec{r}^{\,N} d\vec{p}^{\,N}}$$

$$= \frac{< Q e^{+V^{BIAS}/k_B T} >_{BIAS}}{< e^{+V^{BIAS}/k_B T} >_{BIAS}}$$

***Unbiased* ensemble average of Q can be obtained from *two biased* ensemble averages.**

Searching and sampling configuration space

A. Types of methods for searching configuration space

A. Systematic or exhaustive search

B. Heuristic search

1. Non-step methods (e.g. Distance Geometry)
2. Step methods: change of a complete configuration (e.g. MC, MD, SD)
3. Step methods: build-up of a configuration (e.g. CBMC)

B. Types of search enhancement techniques

1. Deformation or smoothening of the potential energy surface

- Soft-core non-bonded interaction
- Local-elevation search
- Coarse graining of the molecular model

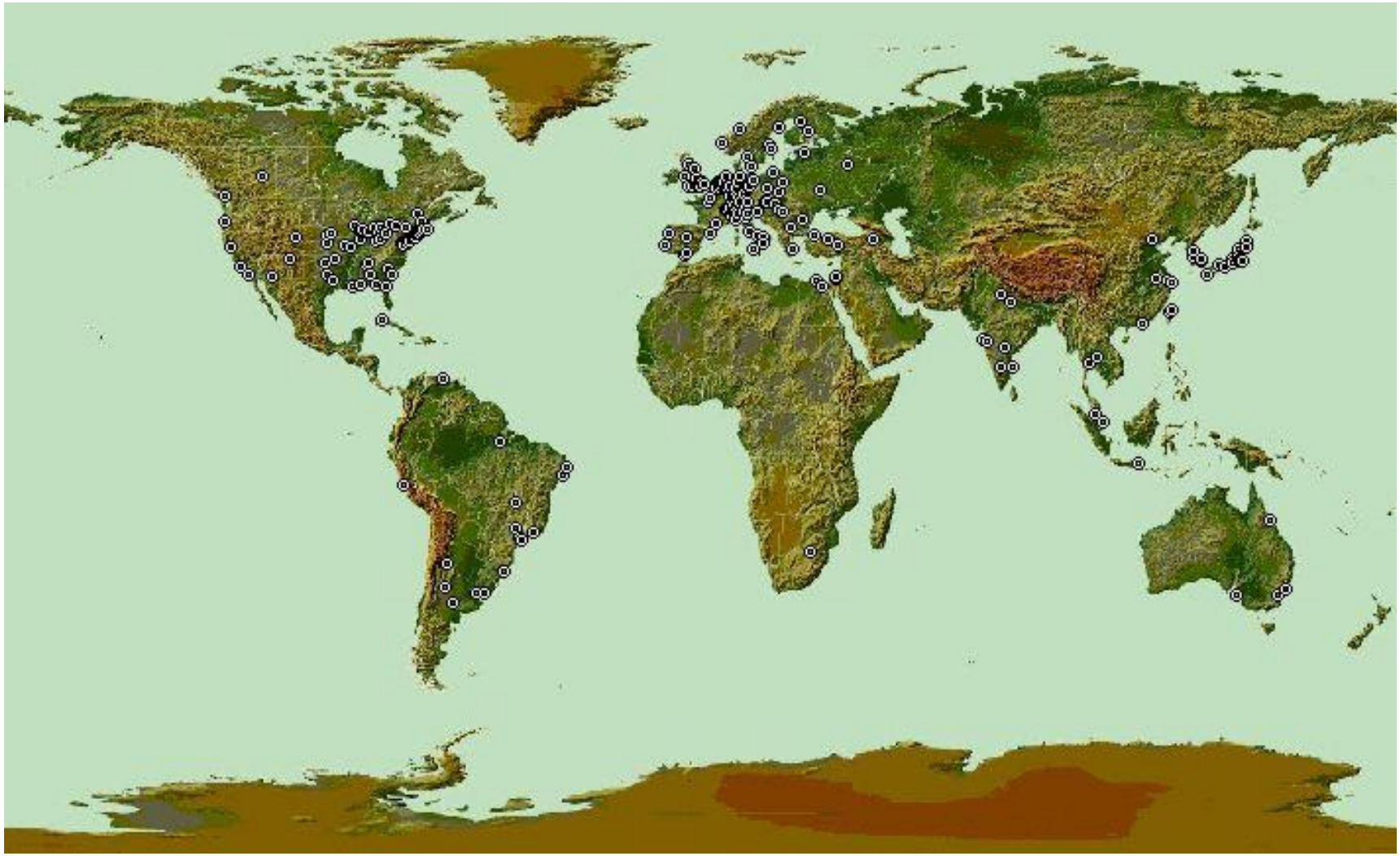
2. Scaling of system parameters

- Temperature annealing
- Tight coupling to a heat bath
- Mass scaling
- Mean-field approaches

3. Multi-copy searching and sampling

- Replica-exchange and multi-canonical algorithms
- Cooperative search: SWARM MD

Spatial distribution of licences
GROMOS biomolecular simulation software



GROMOS = Groningen Molecular Simulation + GROMOS Force Field

Generally available: http://www.gromos.net