International Spring School Statistical Thermodynamics, Santiago de Chile
Monday, November 27, 2017
Lecture 14

Spatial boundary conditions and geometric constraints in molecular simulation

Prof. Dr. Wilfred F. van Gunsteren

ETH Zürich, Switzerland

Molecular simulation: algorithms, boundary conditions and constraints

1. Generating configurational ensembles

   a. Integrating equations of motion

      1. Newton: molecular dynamics simulation (MD)

      2. Langevin: stochastic dynamics simulation (SD)

   b. Monte Carlo type approach

      1. Metropolis Monte Carlo simulation (MC)

      2. Replica exchange technique (RE)

2. Boundary conditions

     1. Spatial boundary conditions

     2. Thermodynamic boundary conditions

       - temperature
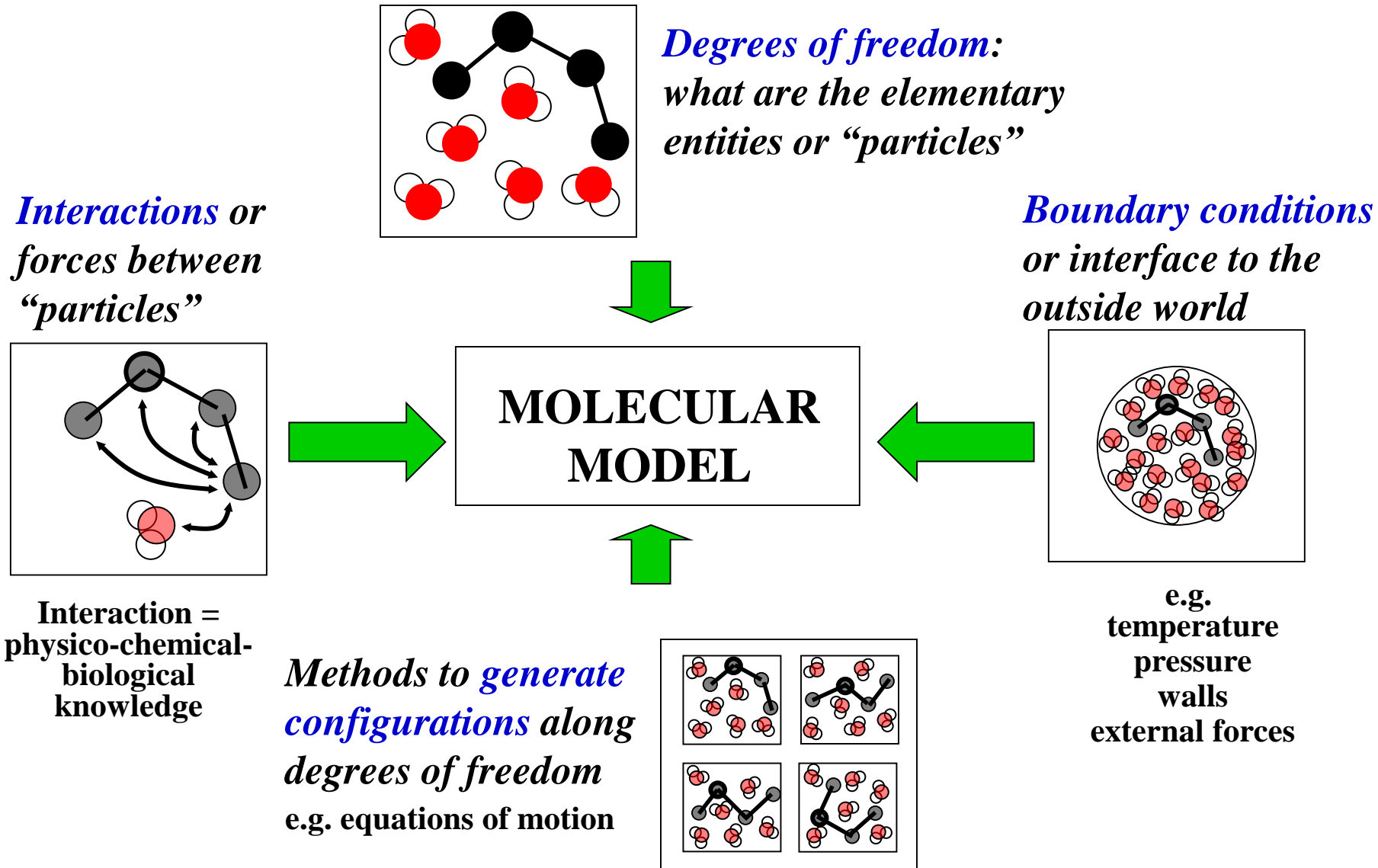
       - pressure

       - other

3. Constraints

     1. Distance constraints
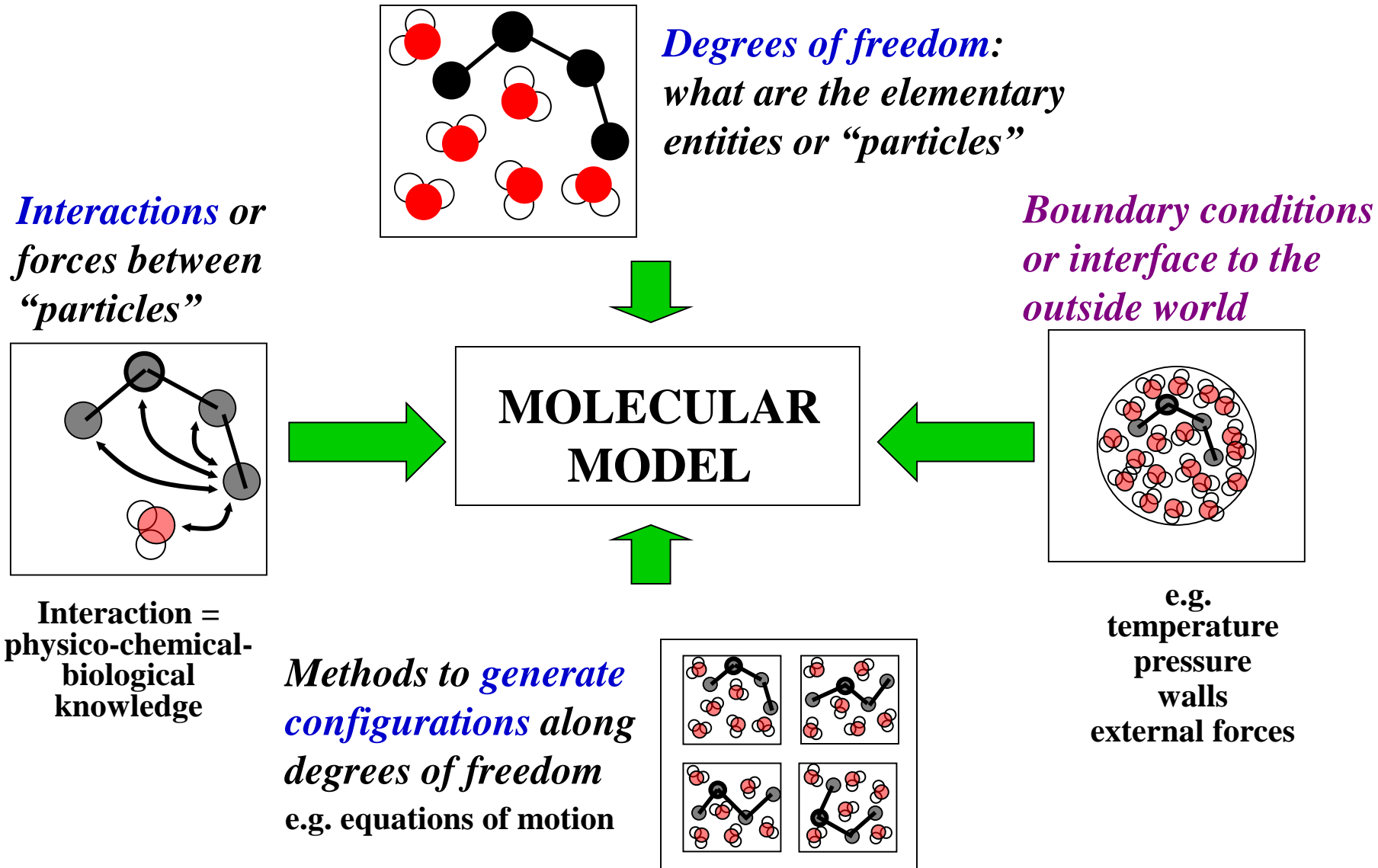
     2. Other constraints

4. MD algorithm

Definition of a model for a computer simulation

For any system *four choices* have to be made

Degrees of freedom:
what are the elementary
entities or "particles"

Interactions or
forces between
"particles"

Boundary conditions
or interface to the
outside world

**MOLECULAR
MODEL**

**Interaction =
physico-chemical-
biological
knowledge**

*Methods to generate
configurations* along
degrees of freedom
e.g. equations of motion

**e.g.
temperature
pressure
walls
external forces**

Definition of a model for a computer simulation

For any system *four choices* have to be made

Degrees of freedom: *what are the elementary entities or "particles"*

Interactions or forces between "particles"

Boundary conditions or interface to the outside world

MOLECULAR MODEL

Interaction = physico-chemical-biological knowledge

Methods to generate configurations along degrees of freedom
e.g. equations of motion

e.g. temperature pressure walls external forces

Molecular simulation: algorithms, boundary conditions and constraints

1. Generating configurational ensembles

    a. Integrating equations of motion

        1. Newton: molecular dynamics simulation (MD)

        2. Langevin: stochastic dynamics simulation (SD)

    b. Monte Carlo type approach

        1. Metropolis Monte Carlo simulation (MC)

        2. Replica exchange technique (RE)

2. Boundary conditions

    1. Spatial boundary conditions

    2. Thermodynamic boundary conditions

        - temperature

        - pressure

        - other

3. Constraints

    1. Distance constraints

    2. Other constraints

4. MD algorithm

TREATMENT OF SPATIAL BOUNDARIES

Number of atoms N in system $<<$ $N_{avogadro}$ = 10^{23}

1. No wall: in vacuo

distortive effects:

- **surface tension will reduce surface area**

    -> makes molecules more spherical:

            DNA, insulin: deformed

            myoglobin, cytochrome: less deformed

    partial remedy: use solvent to make system spherical

                     use solvation area force
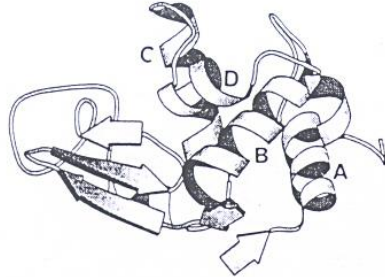
- **dielectric permittivity of vacuum = 1**

    charge – charge interactions in vacuum are larger than in polar

    solvent : $\varepsilon > 1$; water: $\varepsilon = 80$

    partial remedy: reduce charges or ε
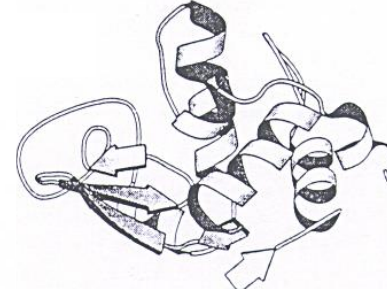
DISTORTIVE EFFECT OF VACUUM BOUNDARY CONDITIONS
e.g. hen-egg-white lysozyme (protein, 129 residues)
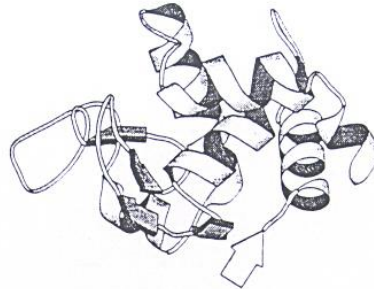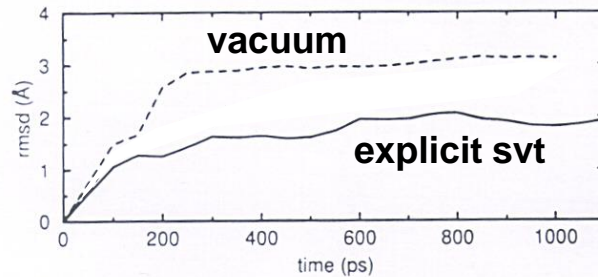

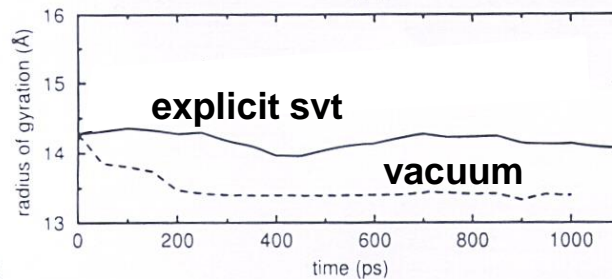**X-ray structure**   **after 1ns, vacuum**   **after 1ns, explicit svt (PBC)**
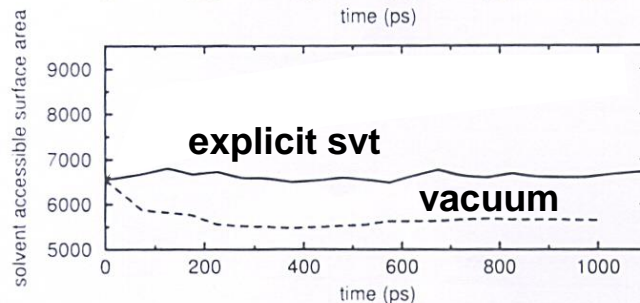

deviation from X-ray structure

$\Rightarrow$ **larger deviations from experiment**

radius of gyration

$\Rightarrow$ **too compact shape**

solvent-accessible surface area

$\Rightarrow$ **more spherical shape, reduced surface area**

$\Rightarrow$ **+ too stable salt bridges**
    + too stable H-bonds
    + slower dynamics

Simulation in vacuum

- **Simulating a molecule in vacuum:**

dialanine
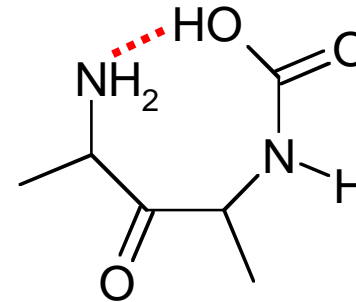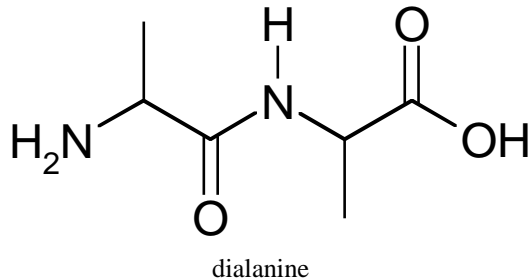
- **Tends to get stuck in a single conformation**
  – Surface tension leads to globular molecules
  – No collisions and interactions with solvent molecules
  – No dielectric screening:

    electrostatic interactions are not reduced by solvent

$$\mathcal{V}^{el} = \frac{1}{4\pi\varepsilon_0\varepsilon}\frac{q_i q_j}{r_{ij}}$$

80 for water

Adaptation of a force field to a vacuum boundary condition

GROMOS force field: not vacuum
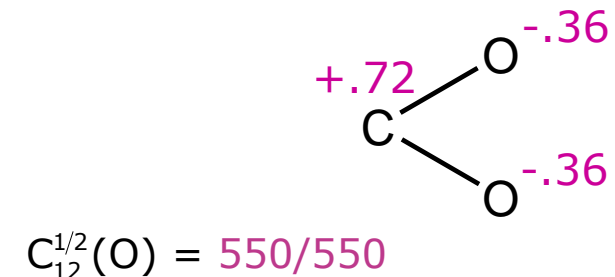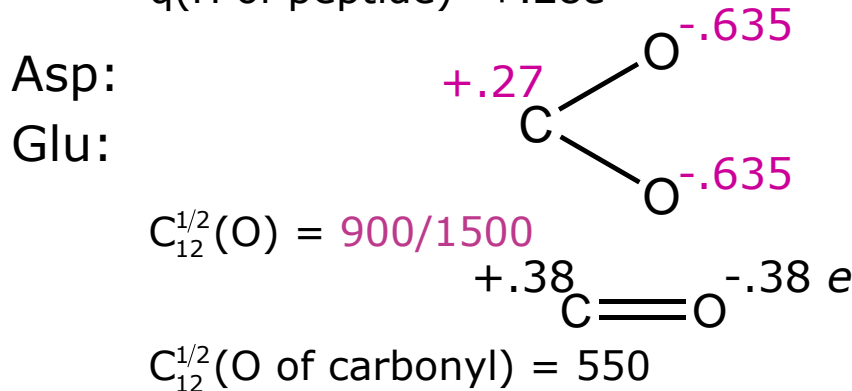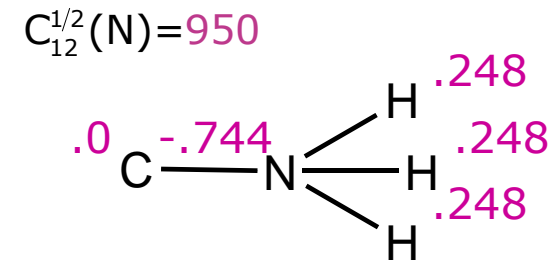
What to do when using it for *in vacuo* simulations?

solution parameters (37C4):
GROMOS: 43A1 (45A4, 53A6)

vacuo parameters(37D4):
43B1 (45B4, 53B6)

neutralise charged groups retaining H-bond capacity

Lys: $C_{12}^{1/2}(N)=1500\left[kcal/mol\,Å^{12}\right]^{1/2}$

$C_{12}^{1/2}(N)=950$

$C_{12}^{1/2}(N \text{ of peptide})= 950$

$q(H \text{ of peptide})=+.28e$

Asp:
Glu:

$C_{12}^{1/2}(O) = 900/1500$

$C_{12}^{1/2}(O) = 550/550$

$C_{12}^{1/2}(O \text{ of carbonyl}) = 550$
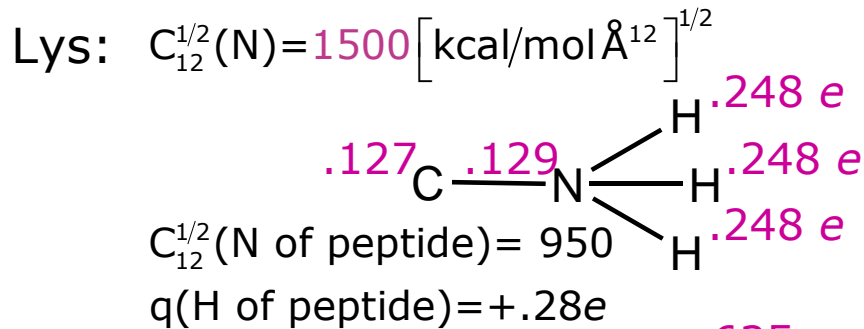
Arg, N-terminus, C-terminus: likewise

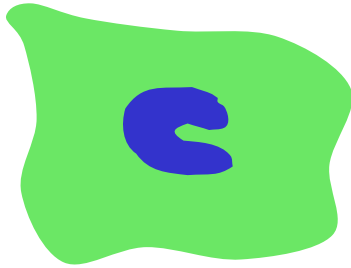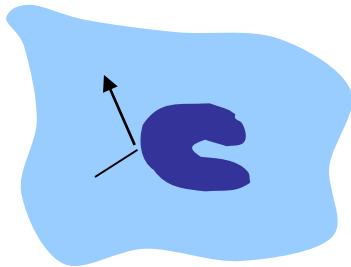# Implicit solvation models I

Vacuum

- Surface effects (surface tension)
- No dielectric screening

Implicit solvation

- Surrounding is a continuum with dielectric constant ε > 1
- Different ε within and outside the solute: solving Poisson-Boltzmann equation
- Nonpolar interaction proportional to solvent accessible surface area

Stochastic dynamics

- Change of the equations of motion
- Additional frictional coefficient represents viscosity of solvent
- Random (stochastic) force represents collision with solvent

$$\mathbf{F}(t) = -\frac{\partial \mathcal{V}}{\partial \mathbf{r}} - \zeta \mathbf{p}(t) + \mathbf{F}_{stoch}(t)$$

Implicit solvent models II

- **Take the solvent into account in a mean-field manner:**
    - average dielectric screening and solvation captured
    - no structure in the solvent
    - no explicit interactions (hydrogen bonds) with the solute
    - still distortive boundary effects

- **In general:**
    - short range interactions are really bad
    - long range interactions may be reasonably described
    - depends strongly on the parameters of the model
        - surface tension of an apolar solute:
          solvent-accessible-surface-area (SASA) proportionality term
   - choice of dielectric permittivity ε in the protein

Advice: Do avoid implicit solvation models

*J.R. Allison et al., J. Phys. Chem. B 115 (**2011**) 4547 - 4557*

DISTORTIVE EFFECT OF VACUUM BOUNDARY CONDITIONS

2. extended wall region: buffer region

restrict motion of atoms in layer adjacent to the vacuum to
inhibit distortion

- (harmonic) position restraining
- mean force from outside (vacuo)
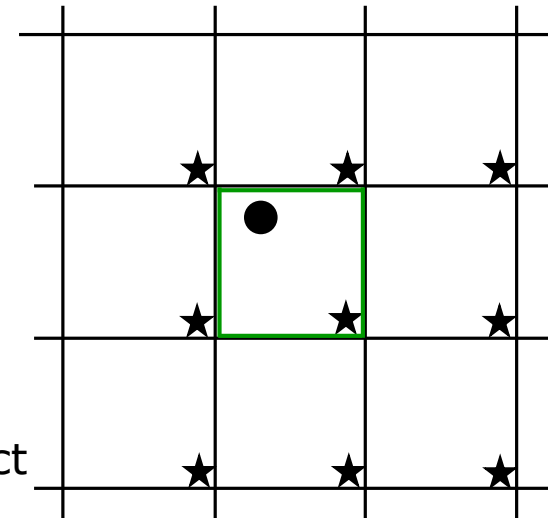- stochastic force from outside

vacuum

MD

restrict motion
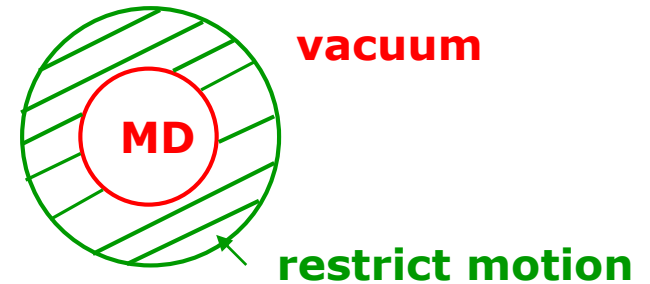
3. periodic: infinite periodic system

- cube, rectangle
- dodecahedron
- truncated octahedron

**more
spherical**

distortive effect: anisotropy -> cube-corner effect

PERIODIC BOUNDARY CONDITIONS: avoiding surface effects

- **The simulated system (solute + solvent) consists of particles within a reference computational box of space-filling shape (e.g. cube)**

- **At each simulation step, particles exiting the box through one face are translated so that they reenter the box through the opposing face**

- **This procedure mimics a system consisting of an infinite lattice of periodic copies of the reference box ($\rightarrow$no interface to vacuum !)**

- **Only the coordinates of particles in the central box are actually stored in the computer**

reference box
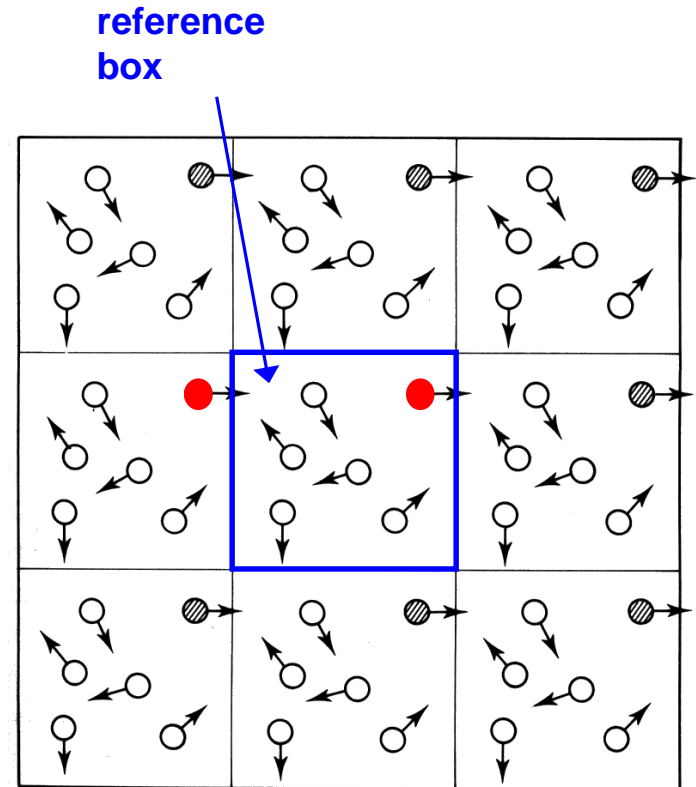
Periodic boundary conditions
Space filling box shapes

cubic periodic
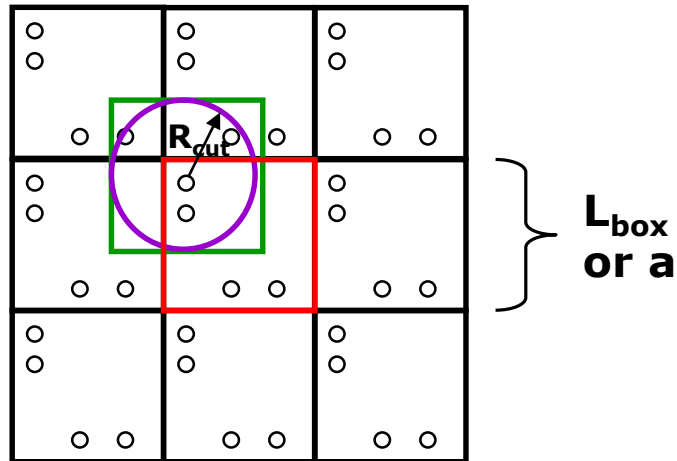
truncated octahedron periodic

R_{cut}

L_{box} or a
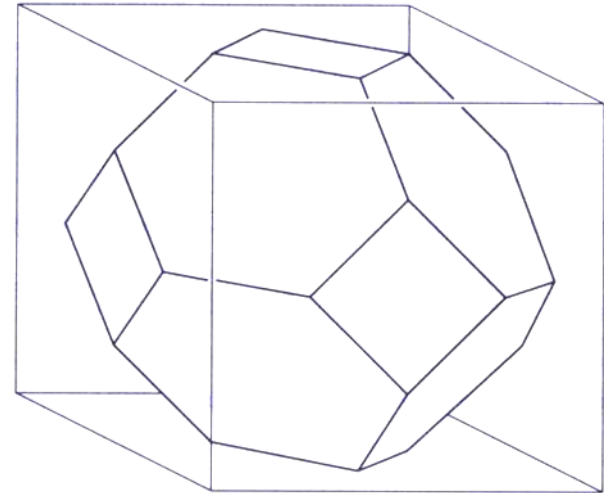
Preferred over extended wall boundary conditions

Spatial periodic boundary conditions:

Periodic boundary conditions:
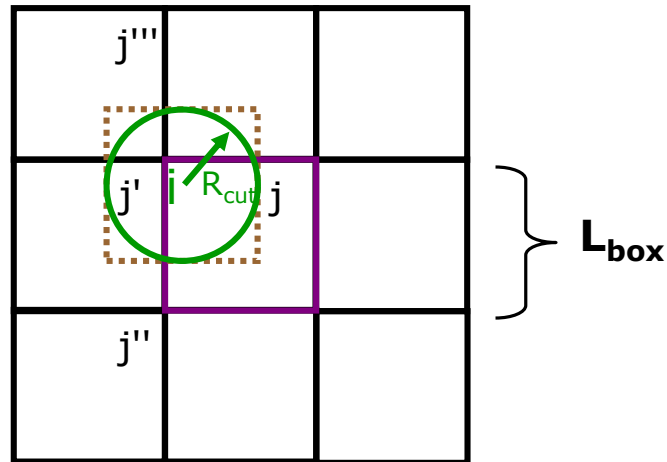
Calculation of force on particle i due to particle j or *one* of its periodic images j`, j``, j```

Condition: $2\ R_{cut} < L_{box}$

In order to avoid interaction between periodic images

For a solute of radius R_{sol} we have

Additional condition: $2 R_{sol} + R_{cut} < L_{box}$

$$L_{box}$$
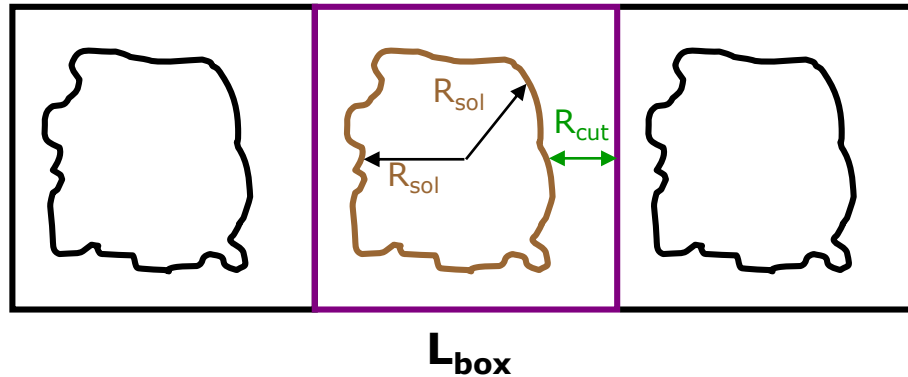
Upon violation of this condition:

Aggregation of periodic images

Molecular simulation: algorithms, boundary conditions and constraints

1. Generating configurational ensembles

    a. Integrating equations of motion

        1. Newton: molecular dynamics simulation (MD)

        2. Langevin: stochastic dynamics simulation (SD)

    b. Monte Carlo type approach

        1. Metropolis Monte Carlo simulation (MC)

        2. Replica exchange technique (RE)

2. Boundary conditions

    1. Spatial boundary conditions

    2. Thermodynamic boundary conditions

        - temperature

        - pressure

        - other

3. Constraints

    1. Distance constraints

    2. Other constraints

4. MD algorithm

Constraints in molecular simulation

1. **Why constraints, consequences**

2. **Techniques to impose constraints**

3. **Enhanced efficiency**

4. **Metric tensor effects**

5. **Physical effects**

Constrained ensembles

Why use constraints?

- elimination of highest frequency motions

    → often uninteresting

    → maybe unphysical

    → limit time step Δt

- biasing of the sampling

    → towards phase space region of interest

Consequence:

- modifies ensemble averages when using non-Cartesian coordinates:

    → mass-metric tensor effect

Methods to impose constraints:

- Use of generalized (non-Cartesian) coordinates

- Use of Cartesian coordinates and Lagrange multipliers:

  Non-linear constraint equations

- Analytical solution (SETTLE, M-SHAKE) *J. Comput. Chem. 22 (**2001**) 501*

- Matrix methods (LINCS)

- Iterative methods (SHAKE) *J. Comput. Phys. 23 (**1977**) 327*

Constrained versus unconstrained ensemble averages:

- General case

- Tri-atomic molecule

- Bond and bond-angle constraints

Constraints in MD simulation

Goal:

save computer time,

approximation of quantum degrees of freedom

Integration time step: $\Delta t << \nu_{max}^{-1}$ (of system)

Four conditions for application of constraints:

1. **Frequencies** of frozen degrees of freedom **>> rest** of frequencies
2. **Frozen degrees of freedom weakly coupled** to remaining ones
3. **Metric tensor effects** play a **minor role**
4. **Property of interest independent** of degrees of freedom that are to be frozen

Examples:

- Bond-lengths   (no constraints: time step Δt = 0.0005 ps)
  - to H-atoms (**factor 2 speed-up**)   (Δt = 0.001 ps)
  - to all atoms (**factor 3 speed-up**)   (Δt = 0.002 ps)
- Bond-lengths + bond-angles (**artefacts, entropy loss**)

Methods:

A. Generalized coordinates: rigid molecules

    (flexible molecules: **impractical**)

B. Cartesian coordinates: flexible, holonomic

1. Matrix method + variations   matrix inversion
2. SHAKE + variations   iterative, **efficient**

Nomenclature:

Holonomic	$\sigma_k(\vec{r}_1...\vec{r}_n, t) = 0$	Rheonomous: σ_k functions of t
Non-holonomic	> 0	Scleronomous: independent of t

Constraints in molecular simulation

1. **Why constraints, consequences**

2. **Techniques to impose constraints**

3. **Enhanced efficiency**

4. **Metric tensor effects**

5. **Physical effects**

Classical dynamical equations of motion

Use of generalized versus Cartesian coordinates

Cartesian Coordinates: x_i $i=1,2,...,N_{df}$

Equations of motion (Newton):

$$m_i \frac{d^2 x_i}{dt^2} = - \frac{\partial V(x_1, x_2, ..., x_{N_{df}})}{\partial x_i} \qquad i=1,2,...,N_{df}$$

Generalized coordinates: θ_i $i=1,2,...,N_{df}$

Example: torsional angles θ_i

Polymer: *bonds*

          bond-angles   constrained

Equations of motion (Lagrange):

$$\sum_{j=1}^{N_{df}} a_{ij} \frac{d^2\theta_j}{dt^2} = -\frac{\partial V(\theta_1, \theta_2, \ldots, \theta_{N_{df}})}{\partial \theta_i} -$$

$$\sum_{j=1}^{N_{df}} b_{ij} \left(\frac{d\theta_j}{dt} \right)^2 - \sum_{j=1}^{N_{df}} \sum_{k=1}^{N_{df}} c_{ijk} \left(\frac{d\theta_j}{dt} \right) \left(\frac{d\theta_k}{dt} \right)$$

$$i = 1, 2, \ldots, N_{df}$$

Coefficients a_{ij}, b_{ij}, c_{ijk} depend on

- atomic masses

- topology of polymer

Note: *very inefficient* (compared to Cartesian coordinates)

- **additional summations (over j and k) $\Rightarrow$ coupled equations**

- **additional non-linear terms (velocities)**

*U. Stocker et al., Mol. Sim. 29 (**2003**) 123-138*

Application of constraints using Lagrange multipliers

Distance constraints:

$$\sigma_k(\vec{r}) \equiv r^2_{k_1 k_2} - d^2_{k_1 k_2} = 0 \qquad\qquad k=1,2,\ldots,N_c$$

Lagrangian multipliers:

(Add a zero term to the potential energy in Newton's equations)

Lagrange multipliers depending on t

$$m_i \frac{d^2 \vec{r}_i(t)}{dt^2} = -\frac{\partial}{\partial \vec{r}_i}\left[V(\vec{r}) + \sum_{k=1}^{N_c} l_k(t)\sigma_k(\vec{r}) \right]$$

$$\hookrightarrow = 0$$

$$= \vec{f}_i^{\,uc}(t) + \vec{f}_i^{\,c}(t)$$

$$\hookrightarrow = -2\sum_{k=1}^{N_c} l_k(t)\left[\delta_{ik_1} - \delta_{ik_2} \right]\vec{r}_{k_1 k_2}(t)$$

Unconstrained forces

Constraint forces

Leap-frog time-integration scheme:

$$\vec{r}_i^{uc}(t + \Delta t) = \vec{r}_i(t) + \vec{v}_i(t + \Delta t / 2)\Delta t + m_i^{-1}\vec{f}_i^{uc}(t)(\Delta t)^2$$

$$\vec{r}_i(t + \Delta t) = \vec{r}_i^{uc}(t + \Delta t) + m_i^{-1}\vec{f}_i^{c}(t)(\Delta t)^2$$

$\downarrow$ must satisfy constraint: $\sigma_k = 0$

$$\left[\vec{r}_{k_1}^{uc}(t + \Delta t) + m_{k_1}^{-1}\vec{f}_{k_1}^{c}(t)(\Delta t)^2 - \vec{r}_{k_2}^{uc}(t + \Delta t) - m_{k_2}^{-1}\vec{f}_{k_2}^{c}(t)(\Delta t)^2\right]^2 - d_{k_1 k_2}^2 = 0$$

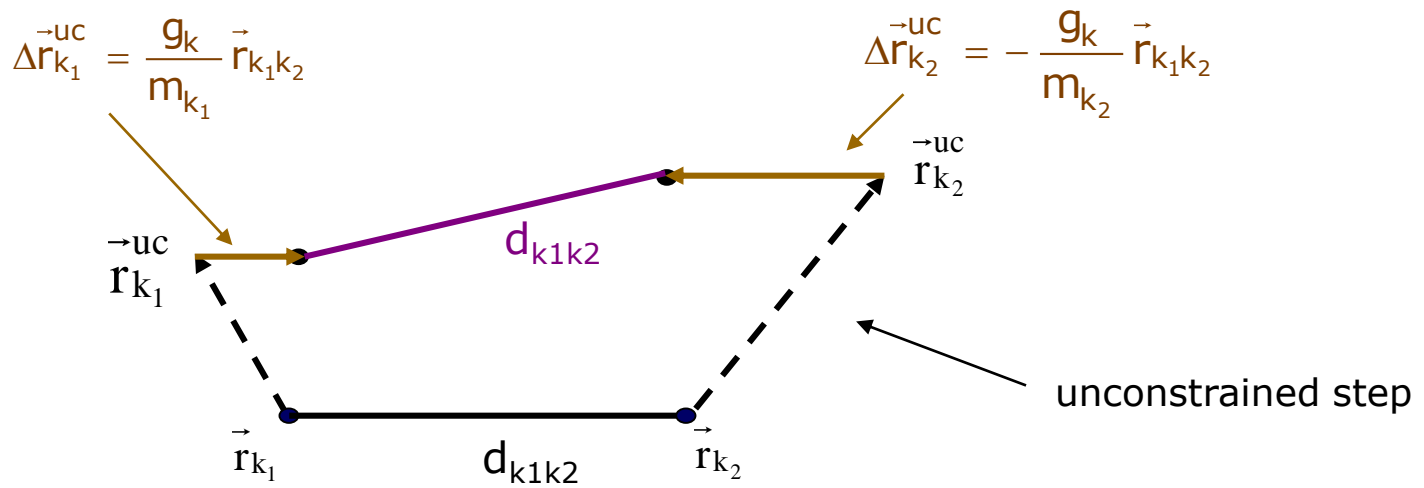$\downarrow$ quadratic equations in $l_k(t)$

$$\left[\vec{r}_{k_1 k_2}^{uc}(t + \Delta t) - 2l_k(t)\left[m_{k_1}^{-1} + m_{k_2}^{-1}\right]\vec{r}_{k_1 k_2}(t)(\Delta t)^2\right]^2 - d_{k_1 k_2}^2 = 0 \qquad k=1,2,...,N_c$$

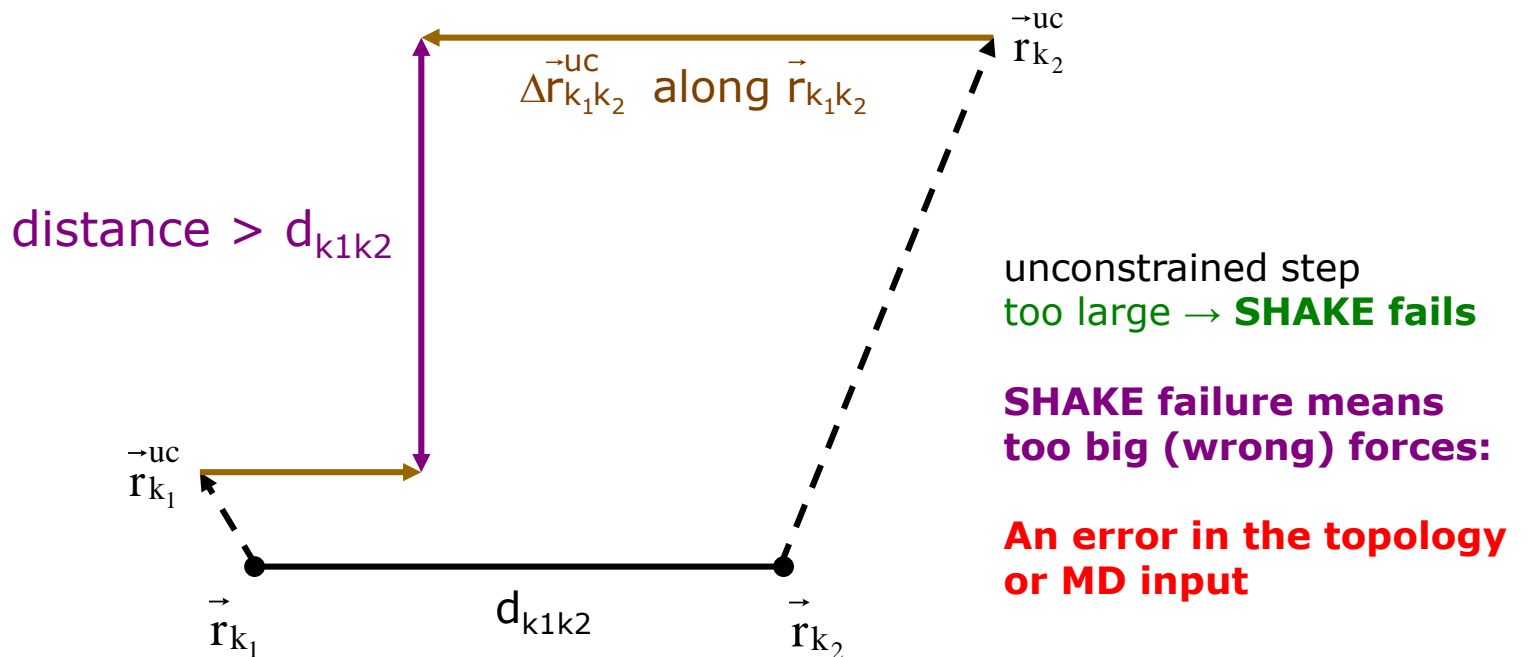$$\downarrow \quad \text{solve :} \begin{cases} \text{analytical} \\ \text{matrix} \\ \text{iterative (SHAKE)} \end{cases}$$

Position corrections due to constraints (SHAKE):

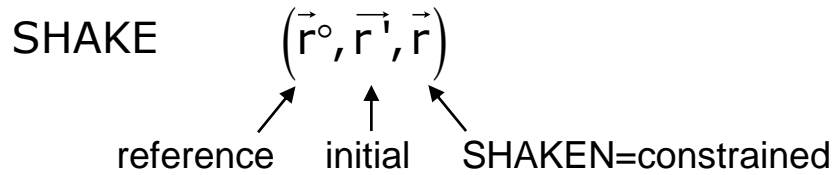$$\Delta\vec{r}_{k_1}^{uc}(t + \Delta t) = -2(\Delta t)^2 m_{k_1}^{-1} l_k(t)\vec{r}_{k_1 k_2}(t)$$

$$= m_{k_1}^{-1}g_k(t)\vec{r}_{k_1 k_2}(t)$$

$$\Delta \vec{r}_{k_1}^{uc} = \frac{g_k}{m_{k_1}} \vec{r}_{k_1 k_2}$$

$$\Delta \vec{r}_{k_2}^{uc} = - \frac{g_k}{m_{k_2}} \vec{r}_{k_1 k_2}$$

$\vec{r}_{k_2}^{uc}$

$\vec{r}_{k_1}^{uc}$

d_{k1k2}

unconstrained step

$\vec{r}_{k_1}$ d_{k1k2} $\vec{r}_{k_2}$

Positional corrections induced by SHAKE

$\vec{r}_{k_2}^{uc}$

$\Delta \vec{r}_{k_1 k_2}^{uc}$ along $\vec{r}_{k_1 k_2}$

distance > d_{k1k2}

unconstrained step
too large → **SHAKE fails**

**SHAKE failure means
too big (wrong) forces:**

**An error in the topology
or MD input**

$\vec{r}_{k_1}^{uc}$

$\vec{r}_{k_1}$ d_{k1k2} $\vec{r}_{k_2}$

Constrained positions:

SHAKE $\left(\vec{r}^{\circ}, \vec{r}^{\,i}, \vec{r} \right)$

reference   initial   SHAKEN=constrained

Constrained velocities:

Leap-frog: $\quad \vec{r}_i(t + \Delta t) = \vec{r}_i(t) + \vec{v}_i(t + \Delta t / 2)\Delta t$

Invert: $\qquad \vec{v}_i(t) = \dfrac{\left\{ \vec{r}_i(t + \Delta t) - \vec{r}_i(t) \right\}}{\Delta t}$

Where $\left. \begin{array}{c} \vec{r}_i(t + \Delta t) \\ \vec{r}_i(t) \end{array} \right\}$ have been shaken $\longrightarrow$ satisfy the constraints
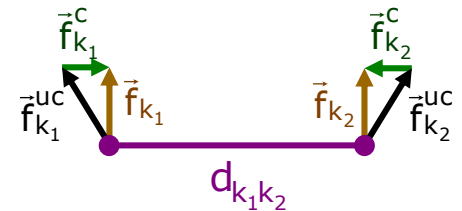
Constrained forces:

Likewise form: $\quad \vec{r}_i(t + \Delta t) = \vec{r}_i^{\,uc}(t + \Delta t) + m_i^{-1} \vec{f}_i^{\,c}(t)(\Delta t)^2$

So **constrain_t_** force:
$$\vec{f}_i^{\,c}(t) = \frac{\left\{ \vec{r}_i(t + \Delta t) - \vec{r}_i^{\,uc}(t + \Delta t) \right\} m_i}{(\Delta t)^2}$$

And **constrain_ed_** force:

$\quad \vec{f}_i(t) = \vec{f}_i^{\,uc}(t) + \vec{f}_i^{\,c}(t)$

   = component along the constraint is equal to

      that at the other end of the constraint

Constraints in molecular simulation

1. **Why constraints, consequences**

2. **Techniques to impose constraints**

3. **Enhanced efficiency**

4. **Metric tensor effects**

5. **Physical effects**

COMPARISON OF MD ALGORITHMS

Complex system: trypsin inhibitor (BPTI) ≈ 500 atoms

Effect of constraints: bond forces: harmonic   ⇒ Gear better
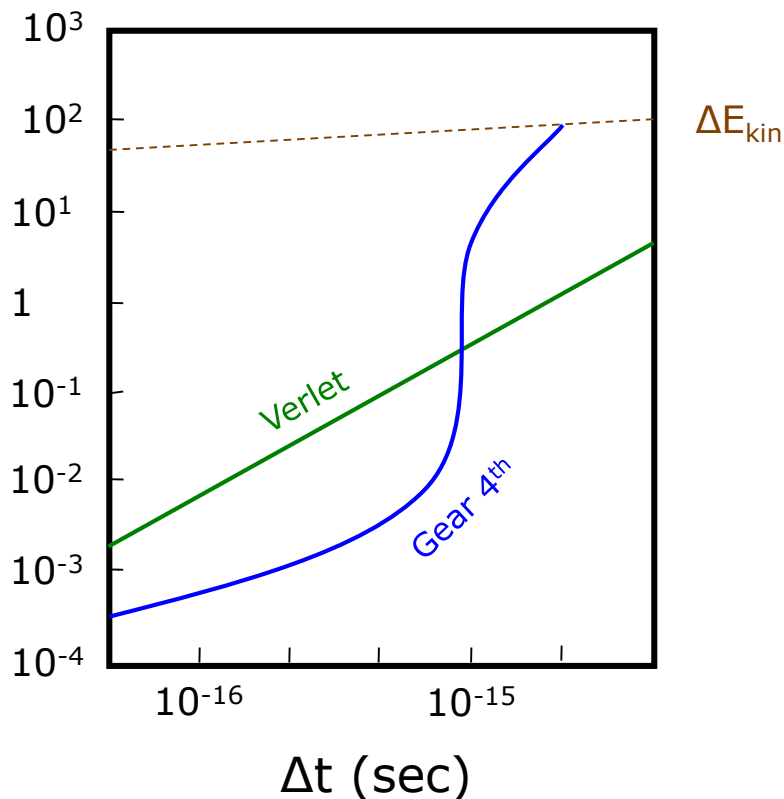
                             other forces: noisy   ⇒ Verlet better
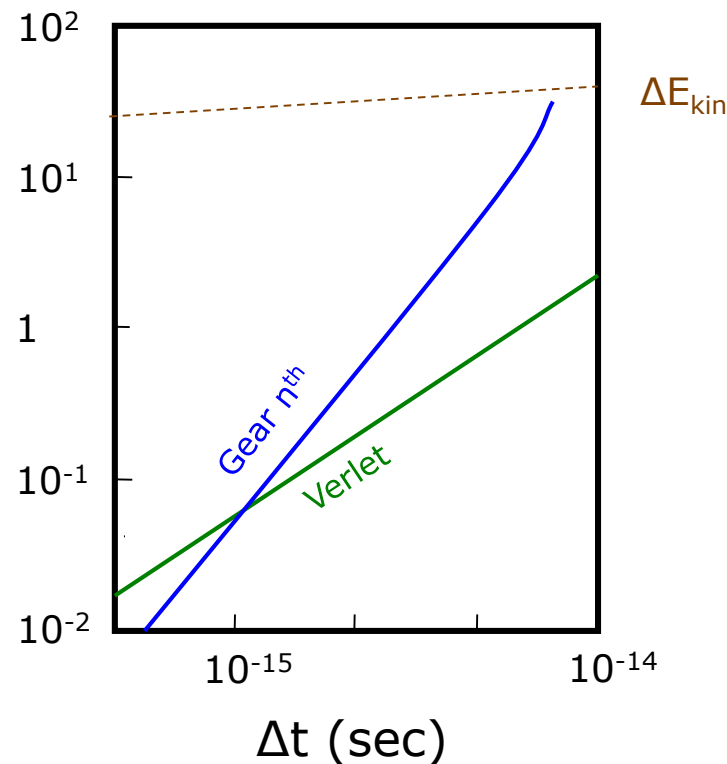
Fluctuation total energy, ΔE_{tot} (kJ/mol)

with harmonic bond forces

with bond constraints

*van Gunsteren & Berendsen, Mol. Phys. 34 (**1977**) 1311 - 1327*

Constraints in molecular simulation

1. **Why constraints, consequences**

2. **Techniques to impose constraints**

3. **Enhanced efficiency**

4. **Metric tensor effects**

5. **Physical effects**

6. **Work done by constraint forces**

7. **Use of constraints to calculate a free energy profile**

METRIC TENSOR EFFECT

Definitions:

N atoms   3N   Cartesian coordinates $\vec{x}$

conjugated momenta $\vec{p}_x = \partial K \big/ \partial \dot{\vec{x}}$

Hamiltonian $H(\vec{p}_x, \vec{x}) = K(\vec{p}_x) + V(\vec{x})$

Equilibrium average of observable A: $<A> = \dfrac{\int A e^{-H(\vec{p}_x, \vec{x})/k_B T} d\vec{p}_x d\vec{x}}{\int e^{-H(\vec{p}_x, \vec{x})/k_B T} d\vec{p}_x d\vec{x}}$

Transform to generalised coordinates (q,p)

     K_h hard coordinates $\vec{q}^\beta$ (e.g. bond-lengths)
     $3N - K_h$ soft coordinates $\vec{q}^\alpha$ (e.g. rest)
     conjugated momenta $\vec{p}^\beta, \vec{p}^\alpha$

Equilibrium average of A: $<A> = \dfrac{\int A e^{-H(\vec{p}, \vec{q})/k_B T} d\vec{p} d\vec{q}}{\int e^{-H(\vec{p}, \vec{q})/k_B T} d\vec{p} d\vec{q}}$

Volume element dpdq is invariant under canonical transformation

Consider two cases

1. Flexible model:

Hard variables $\vec{q}^{\beta}$ are harmonic with infinitely high force constant

$$V(\vec{q}) = V_{soft}(\vec{q}_{\alpha}, \vec{q}_{\beta} \approx \vec{b}) + V_{hard}(\vec{q}_{\beta})$$

$$V_{hard}(\vec{q}_{\beta}) = \tfrac{1}{2}(\vec{q}_{\beta} - \vec{b})^{\tau} \underline{F}(\vec{q}_{\beta} - \vec{b})$$

$\underline{F}$ matrix of harmonic force constants

Integrate over hard variables: $\left. \begin{array}{c} \vec{q}^{\beta}\vec{p}^{\beta} \\ \\ \text{and over: } \quad \vec{p}^{\alpha} \end{array} \right\} < A > = \int ...d\vec{q}^{\alpha}$

$$< A > = \frac{\int A \left| \underline{G} \right|^{1/2} e^{-V(\vec{q}^{\alpha})/k_{B}T} d\vec{q}^{\alpha}}{\int \left| \underline{G} \right|^{1/2} e^{-V(\vec{q}^{\alpha})/k_{B}T} d\vec{q}^{\alpha}}$$

2. Rigid model:

Remove hard variables $\vec{q}^{\beta}$ and $\vec{p}^{\beta}$ from Hamiltonian

Integrate over: $\left. \vec{p}^{\alpha} \right\} < A > = \int ...d\vec{q}^{\alpha}$

$$< A > = \frac{\int A \left| \underline{G}^{\alpha} \right|^{1/2} e^{-V(\vec{q}^{\alpha})/k_{B}T} d\vec{q}^{\alpha}}{\int \left| \underline{G}^{\alpha} \right|^{1/2} e^{-V(\vec{q}^{\alpha})/k_{B}T} d\vec{q}^{\alpha}}$$

Flexible model

Determine conjugate momenta p:

Kinetic energy: $K = \dfrac{1}{2}\sum_{k=1}^{3N} m_k \dot{x}_k^2 = \dfrac{1}{2}\dot{\vec{q}}^{\,\tau}\underline{G}\dot{\vec{q}}$ with $G_{ij} = \sum_{k=1}^{3N} m_k \dfrac{\partial x_k}{\partial q_i}\dfrac{\partial x_k}{\partial q_j}$

Conjugate momenta   (mass-metric tensor)

$\vec{p} = \partial K \Big/ \partial \dot{\vec{q}} = \underline{G}\dot{\vec{q}} \Rightarrow K = \tfrac{1}{2}\vec{p}^{\,\tau}\underline{G}^{-1}\vec{p}$

Integration over $\vec{p}$

$$\int e^{-\frac{1}{2}\vec{p}^{\,\tau}\underline{G}^{-1}\vec{p}/k_BT}\, d\vec{p} = (2\pi k_BT)^{3N/2}\left|\underline{G}\right|^{1/2}$$

Integration over $\vec{q}^{\,\beta}$

$$\int e^{-\frac{1}{2}(\vec{q}^{\,\beta}-\vec{b})^{\tau}\underline{F}(\vec{q}^{\,\beta}-\vec{b})/k_BT}\, d\vec{q}^{\,\beta} = (2\pi k_BT)^{3N/2}\left|\underline{F}\right|^{-1/2}$$

Equilibrium average of A

$$<A> = \frac{\int A\left|\underline{G}\right|^{1/2}\left|\underline{F}\right|^{-1/2} e^{-V(\vec{q}^{\,\alpha})/k_BT}\, d\vec{q}^{\,\alpha}}{\int \left|\underline{G}\right|^{1/2}\left|\underline{F}\right|^{-1/2} e^{-V(\vec{q}^{\,\alpha})/k_BT}\, d\vec{q}^{\,\alpha}}$$

Good assumption: $\left|\underline{F}\right|^{-1/2}$ independent of $\vec{q}^{\,\alpha}$, so it drops out

Rigid model

Kinetic energy:
Conjugate momenta

$$K^\alpha = \dot{\vec{q}}^{\alpha^\tau} \underline{G}^\alpha \dot{\vec{q}}^\alpha \qquad \text{with } G_{ij}^\alpha = \sum_{k=1}^{3N} m_k \frac{\partial x_k}{\partial q_i^\alpha} \frac{\partial x_k}{\partial q_j^\alpha}$$

$$\vec{p}^\alpha = \partial K^\alpha \Big/ \partial \dot{\vec{q}}^\alpha = \underline{G}^\alpha \dot{\vec{q}}^\alpha \Rightarrow K^\alpha = \tfrac{1}{2}\vec{p}^{\alpha^\tau} \underline{G}^{\alpha^{-1}} \vec{p}^\alpha$$

Integration over $\vec{p}^\alpha$

$$\int e^{-\frac{1}{2}\vec{p}^{\alpha^\tau} \underline{G}^{\alpha^{-1}} \vec{p}^\alpha / k_B T} d\vec{p}^\alpha = (2\pi k_B T)^{(3N-K_h)/2} \left| \underline{G}^\alpha \right|^{1/2}$$

Equilibrium average of A

$$< A > = \frac{\int A \left| \underline{G}^\alpha \right|^{1/2} e^{-V(\vec{q}^\alpha)/k_B T} d\vec{q}^\alpha}{\int \left| \underline{G}^\alpha \right|^{1/2} e^{-V(\vec{q}^\alpha)/k_B T} d\vec{q}^\alpha}$$

Calculation of mass-metric tensors

$$\left. \begin{array}{c} \left| \underline{G} \right| \\ \left| \underline{G}^\alpha \right| \end{array} \right\} \text{ too difficult, but } H_{ij}^\beta = \sum_{k=1}^{3N} m_k^{-1} \frac{\partial q_i^\beta}{\partial x_k} \frac{\partial q_j^\beta}{\partial x_k} \text{ can be evaluated}$$

Fixman's theorem: $\left| \underline{G}^\alpha \right| \Big/ \left| \underline{G} \right| = \left| \underline{H}^\beta \right|$

Correction potential energy term $V' = \tfrac{1}{2} k_B T \ln \left| \underline{H}^\beta \right|$
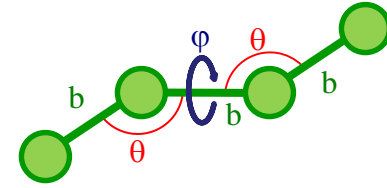
= free energy due to entropy $\quad -\tfrac{1}{2} k_B \ln \left| \underline{H}^\beta \right|$

Example of mass-metric tensor effect

n-butane

Hard degrees of freedom (5):   bonds b

angles θ   } β

Soft degree of freedom (1):   torsion φ } α

$$V'(\varphi) = \tfrac{1}{2} k_B T \ln \left| \underline{H}^\beta \right| \qquad \left| \underline{H}^\beta \right| = \sum_{n=0}^{4} c_n f^n(\varphi)$$

$$f(\varphi) = \left[\tfrac{1}{2} - \tfrac{1}{2} \cos\theta \right]^{1/2} \left[\cos\theta - (1 + \cos\theta) \cos\varphi \right]$$
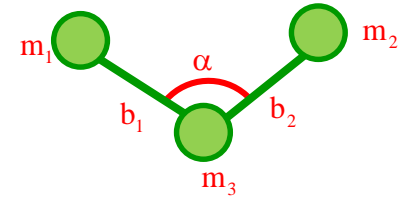
c_n = complicated functions of θ

crosses: φ-distribution (from SD)
    without V' term
    (not flat)

squares: φ-distribution (from SD)
    with V' term
    (flat)
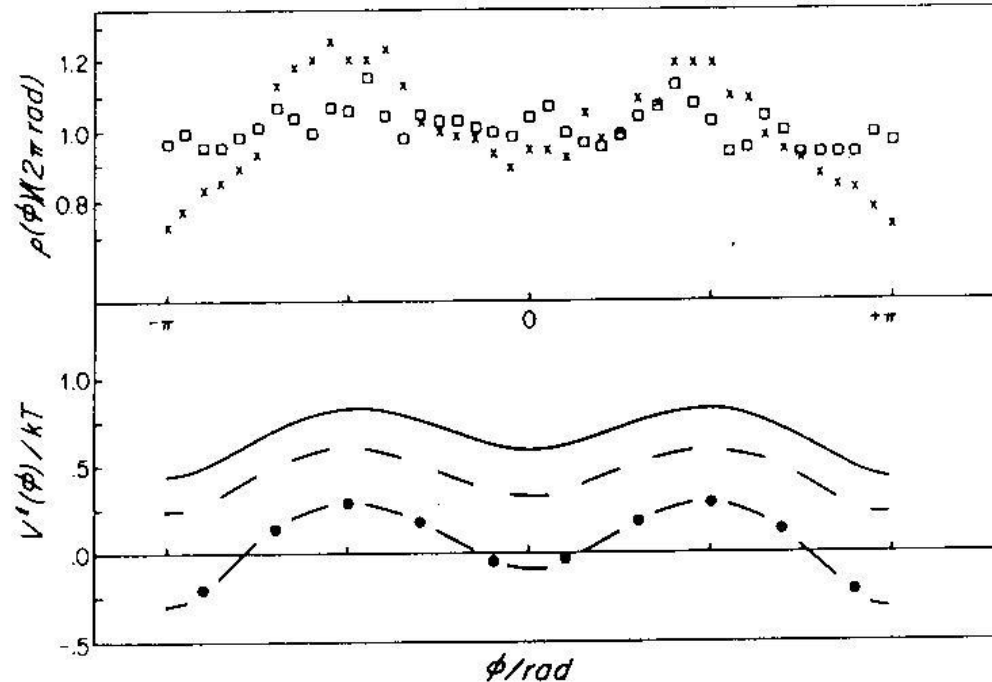
Example of mass-metric tensor calculation

Triatomic molecules

Hard degrees of freedom b_1, b_2
Soft degree of freedom α

$$q_1^\beta = b_1 = \sqrt{x_{13}^2 + y_{13}^2 + z_{13}^2} = r_{13}$$

$$q_2^\beta = b_2 = \sqrt{x_{23}^2 + y_{23}^2 + z_{23}^2} = r_{23}$$

H-matrix

$$q_3^\alpha = \cos\alpha = \frac{r_{13}^2 + r_{23}^2 - r_{12}^2}{2r_{13}r_{23}}$$

$$H_{ij}^\beta = \sum_{k=1}^N m_k^{-1}\left\{\frac{\partial q_i^\beta}{\partial x_k}\frac{\partial q_j^\beta}{\partial x_k} + \frac{\partial q_i^\beta}{\partial y_k}\frac{\partial q_j^\beta}{\partial y_k} + \frac{\partial q_i^\beta}{\partial z_k}\frac{\partial q_j^\beta}{\partial z_k}\right\}$$

$$\underline{H}^\beta = \begin{pmatrix} m_1^{-1} + m_3^{-1} & m_3^{-1}\cos\alpha \\ m_3^{-1}\cos\alpha & m_2^{-1} + m_3^{-1} \end{pmatrix} = \frac{1}{m_3}\begin{pmatrix} \frac{m_1+m_3}{m_1} & \cos\alpha \\ \cos\alpha & \frac{m_2+m_3}{m_2} \end{pmatrix}$$

$$\left|\underline{H}^\beta\right| = (m_1^{-1} + m_3^{-1})(m_2^{-1} + m_3^{-1}) - m_3^{-2}\cos^2\alpha = \frac{\left|\underline{G}^\alpha\right|}{\left|\underline{G}\right|} = \frac{\text{rigid}}{\text{flexible}}$$

$$V' = \tfrac{1}{2}k_B T\ln\left\{(m_1^{-1} + m_3^{-1})(m_2^{-1} + m_3^{-1}) - m_3^{-2}\cos^2\alpha\right\}$$

For restricted variation of angle α: metric tensor has almost no effect
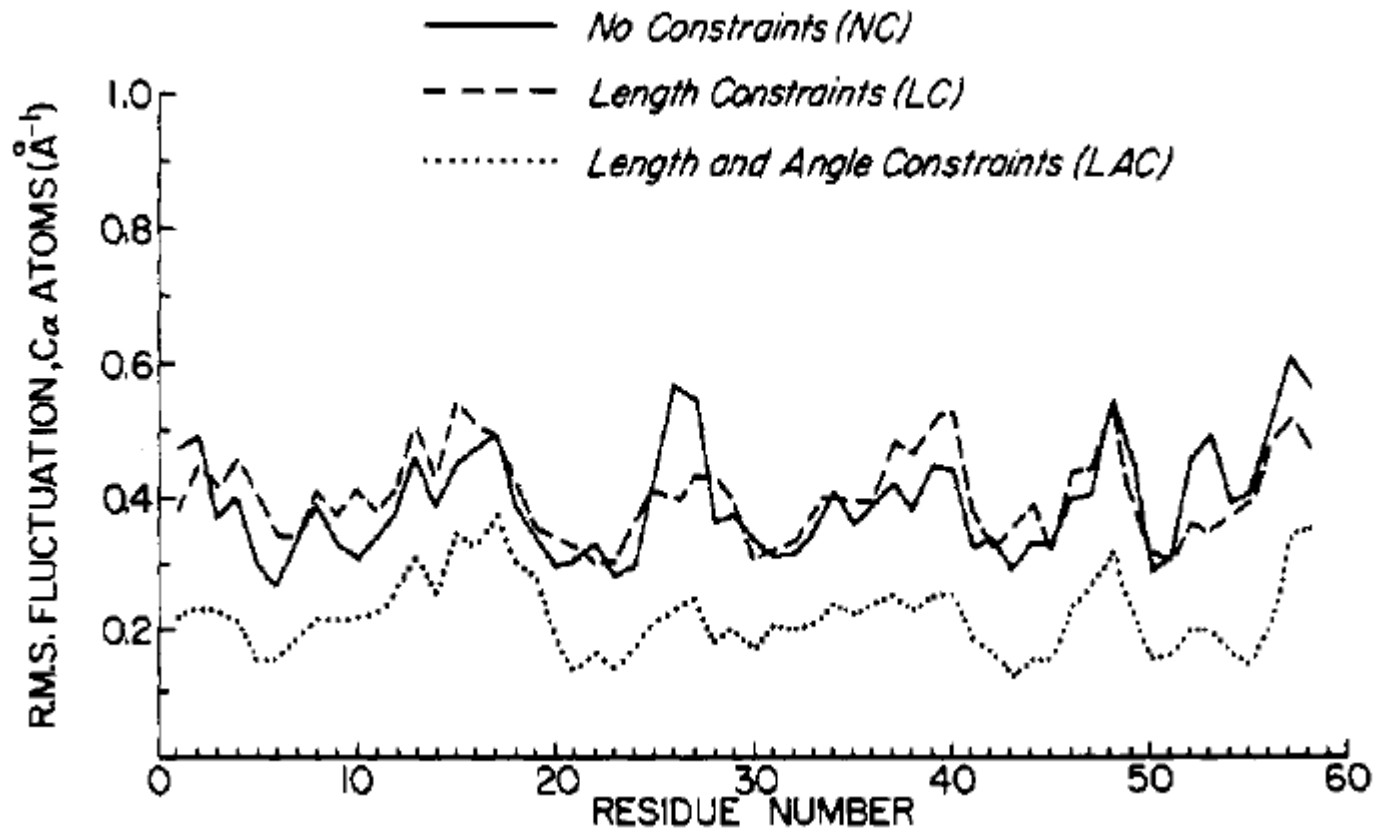
Constraints in molecular simulation

1. **Why constraints, consequences**

2. **Techniques to impose constraints**

3. **Enhanced efficiency**

4. **Metric tensor effects**

5. **Physical effects**

Physical effect of application of constraints

Atom-positional fluctuations

Protein: Bovine pancreatic trypsin inhibitor (BPTI) MD simulations

Bond-length constraints do **not** affect atomic motions
Bond-angle constraints do **considerably restrict** the atomic motions (entropy loss)

Physical effect of application of constraints
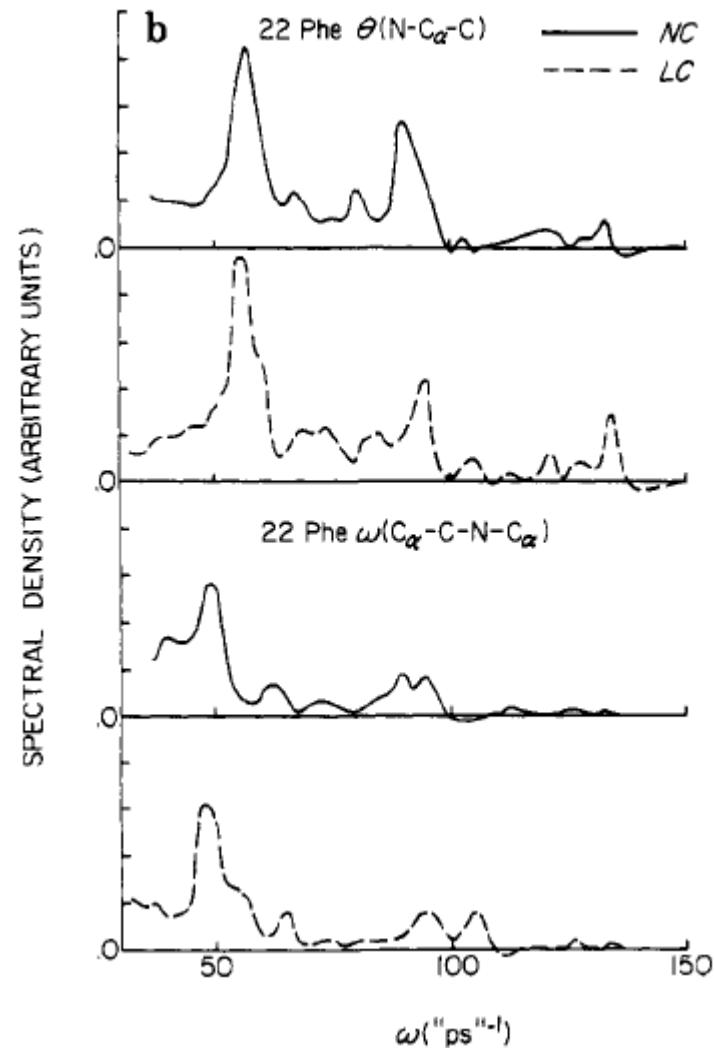
Angular motion

Protein: BPTI MD simulations

NC = no constraints
LC = bond-length constraints

Bond-length constraints **do not significantly affect** **the motions of** **the bond angles or** **the torsional angles**

On the relative merits of flexible versus rigid models for use in computer simulation of molecular liquids

*Ilario Tironi et al., Chem. Phys. Lett. 250 (**1996**) 19-24*

Rigid   **Flexible**

Arguments:

I Principle: approximation of quantum mechanics

  No internal vibrations: Classical vibrations:

  $\hbar\omega$ $1000 - 4000 \text{cm}^{-1}$ harmonic distribution

  $k_B T$ 200cm^{-1}

II Technical: computational efficiency, elegance

A. Large Δt (4x)   Multiple time step methods (since '90s)

B. Less different frequencies   High↔ low frequency modes

    fast relaxation   slow energy exchange

                                    may lead to different temperatures

                                    in high versus low frequency modes

                    Remedy: a) Ewald/exact forces

                                b) Separate T-scaling (Rot/Trans/Vib)

III Practical: liquid properties better?

    Compare: Model fitted   ⎤ structural   ⎤
            - to same data   ⎥ thermodynamic ⎥ PROPERTIES
            - in same manner   ⎦ electric   ⎦

Molecular simulation: algorithms, boundary conditions and constraints

1. Generating configurational ensembles

   a. Integrating equations of motion

     1. Newton: molecular dynamics simulation (MD)

     2. Langevin: stochastic dynamics simulation (SD)

   b. Monte Carlo type approach

     1. Metropolis Monte Carlo simulation (MC)

     2. Replica exchange technique (RE)

2. Boundary conditions

    1. Spatial boundary conditions

    2. Thermodynamic boundary conditions
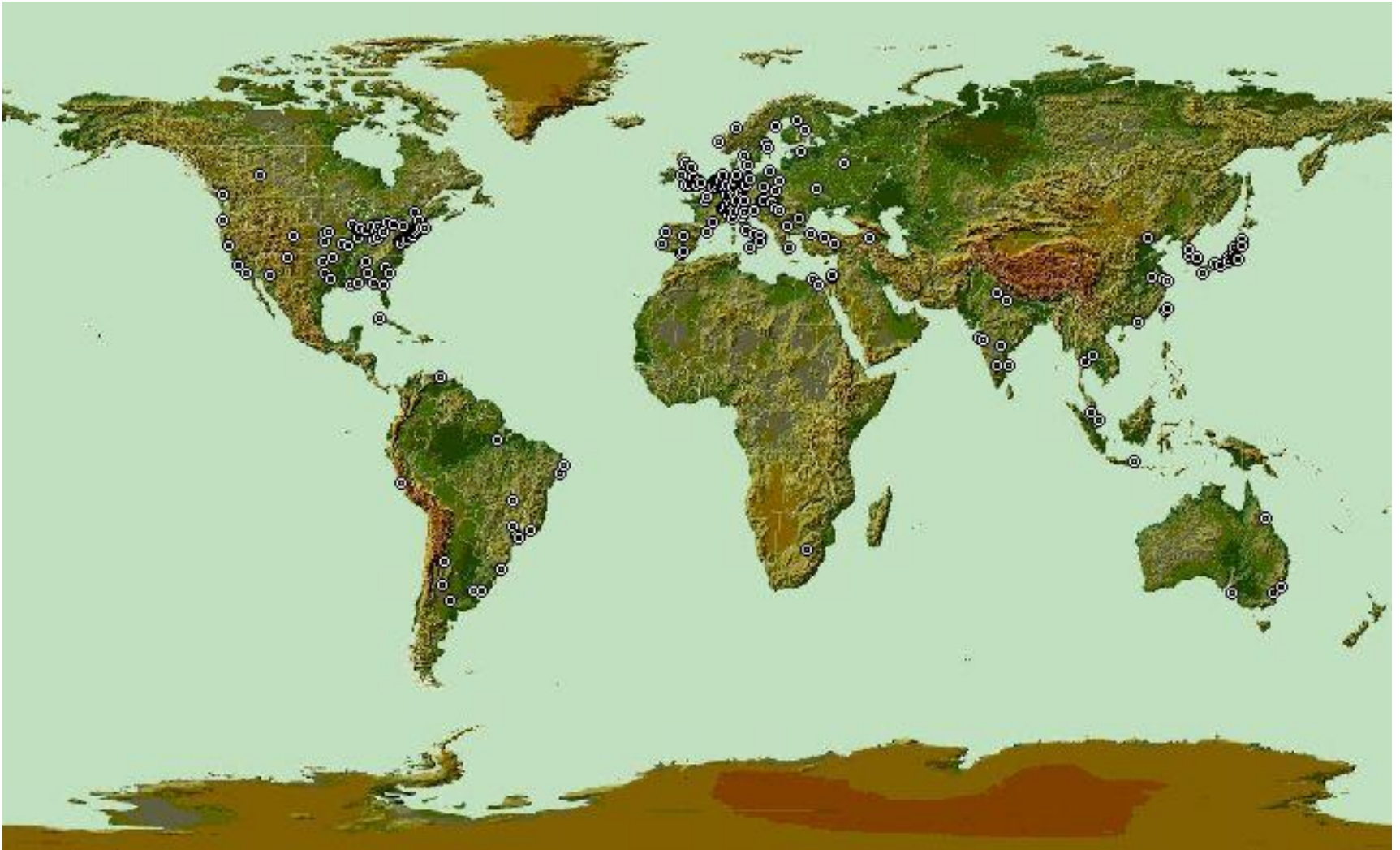
      - temperature

      - pressure

      - other

3. Constraints

    1. Distance constraints

    2. Other constraints

4. MD algorithm

Spatial distribution of licences
GROMOS biomolecular simulation software

GROMOS = Groningen Molecular Simulation + GROMOS Force Field

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