

**Presenter: Alexey Savelyev**

**Topic:  
Molecular Dynamics II**

**Copy of Lecture at:**

<https://demeler.uleth.ca/biophysics/archive/Savelyev>

## Last Lecture... (#1)

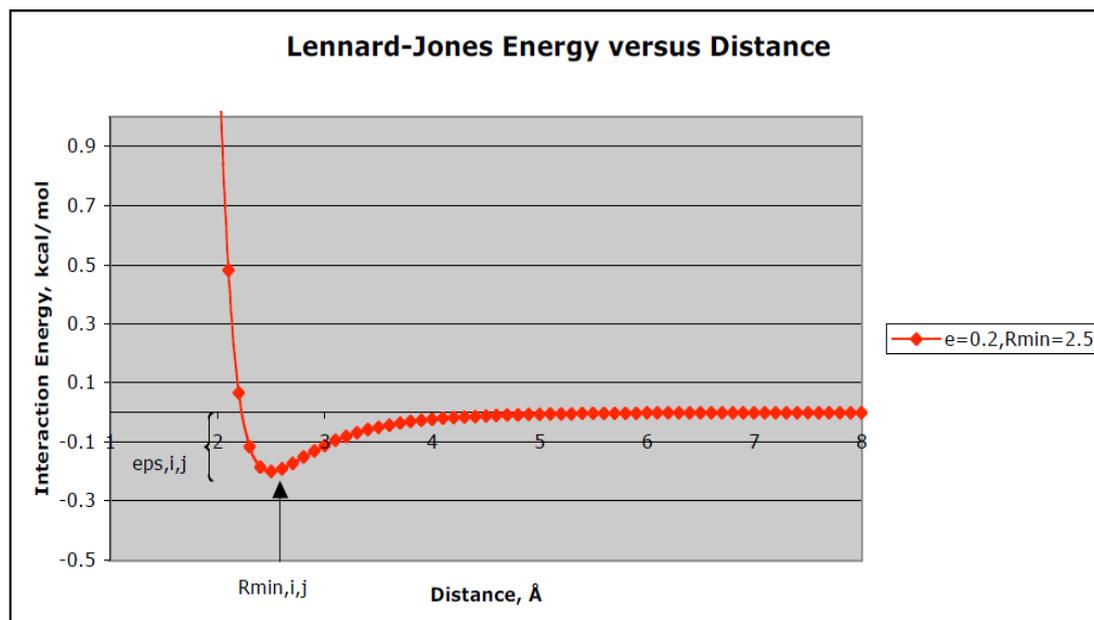
### • Basic Concepts of the Classical MD Simulations:

- › Properties of the “particle” (no directions, fixed charge, connected by springs)
- › MD simulation engine (Newton’s eqns.)
- › Integration algorithm (Verlet-type)
- › MD flowchart (structure initialization, minimization, equilibration etc)
- › **Periodic Boundary Conditions**
- › **Molecular Force Field (interaction potentials)**

### Non-bonded interactions:

- **Lennard-Jones** potential

$$\epsilon_{ij} \left[ \left( \frac{R_{\min,ij}}{r_{ij}} \right)^{12} - 2 \left( \frac{R_{\min,ij}}{r_{ij}} \right)^6 \right]$$



# Classical MD Simulation: Molecular FF (cont. Lecture #1)

- **Non-bonded** (*inter*-molecular) interaction potentials in all-atom models
  - **Lennard-Jones** potential

VdW interactions are a subset of electrostatic interactions (do not confuse with Coulomb term !) involving permanent or induced dipoles.

These include:

- **Repulsive**  $(1/r)^{12}$ , approx. for Pauli exclusion principle
- **Attractive**  $(1/r)^6$  term:
  - › Permanent dipole – dipole interactions (Keesom force)
  - › Dipole-induced – dipole interactions (Debye force)
  - › Induced dipole – Induced dipole interactions (London dispersion forces)

Hydrogen bonds are NOT treated as vdW (LJ) forces: special treatment !!!

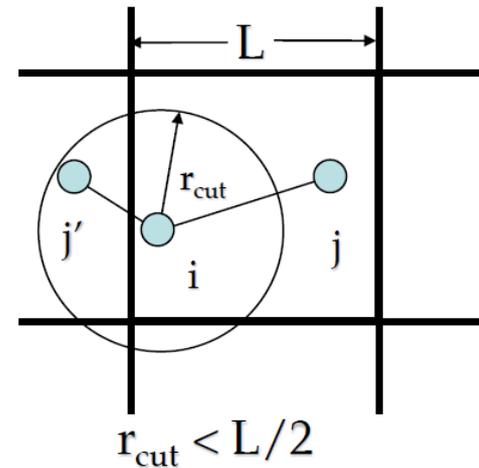
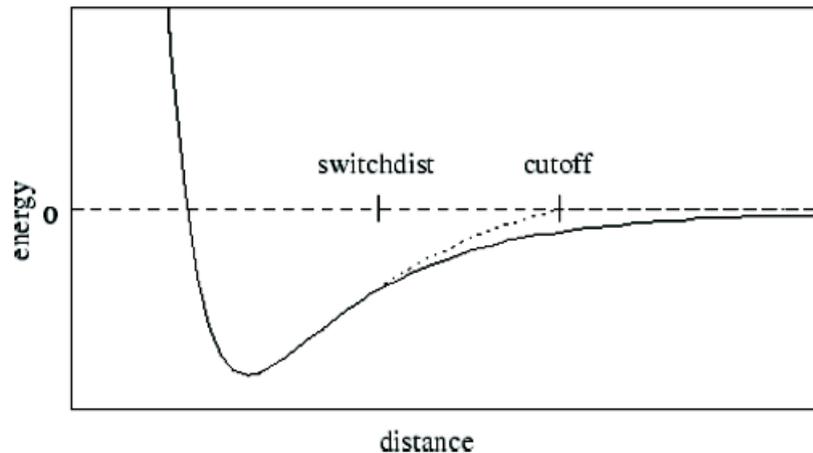
---

**Van der Waals forces** are weak compared to covalent bonds and electrostatic interactions, but play a fundamental role chemistry, biology, nanotechnology, surface science.

# Classical MD Simulation: Molecular FF (cont. Lecture #1)

- **Non-bonded** (*inter*-molecular) interaction potentials in all-atom models

- **Periodic Boundary Conditions**



- $j$  and  $j'$  are identical particles
- To calculate  $U_{ij}$  we have to use

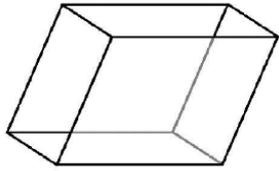
$ij'$  NOT  $ij$

- **vdW interactions are spatially truncated** using cutoff distance. The simulation cell has to be larger than 2 x cutoff.
- **Electrostatic interactions** can be computed fully with minimal additional cost using **Particle Mesh Ewald (PME) method (1970)**.
  - The direct summation of interaction energy between point particles is replaced by **two summations: a direct sum of the short-ranged potential in real space and a summation in Fourier space of the long-ranged part**. Both summations converge quickly, so they may be truncated with little loss of accuracy and great improvement in required computational time. The method uses FFT which requires that the density field be evaluated on a discrete lattice in space (mesh).

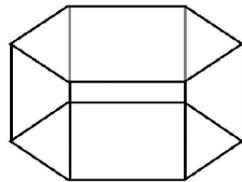
# Classical MD Simulation: Molecular FF (cont. Lecture #1)

- **3D Periodic Boundary Conditions** (other than *Rectangular/Cubic*)

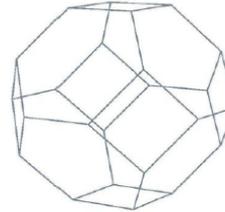
Most of MD simulations !!!



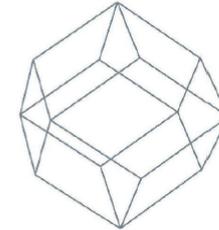
*Triclinic*



*Hexagonal prism*



*Truncated octahedron*

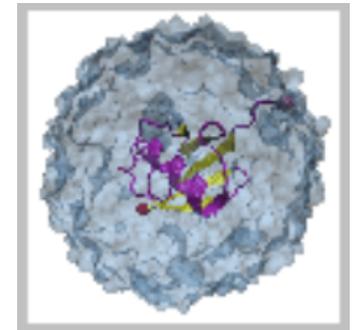


*Rhombic dodecahedron*

- **Stochastic Boundary Conditions (non-Periodic)**

- **Spherical BC**

- The region of interest is solvated in a water sphere at 1atm. The water molecules are submitted to an additional force field that restrain them in the sphere; *resemblance to bulk water ??*

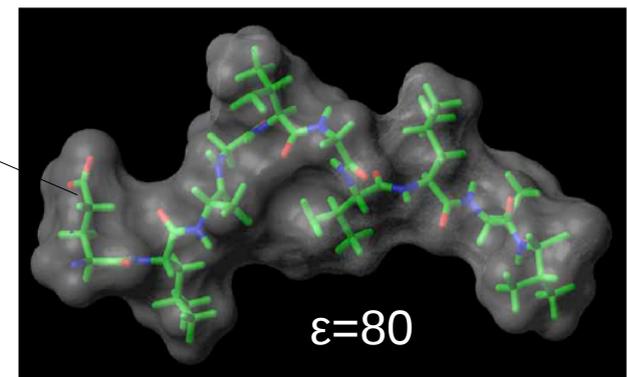


- **None**

- **Implicit solvent simulations**

- Poisson-Boltzmann equation for electr. pot.  $\varphi(\mathbf{r})$ , knowing biomolecular charge density  $\rho(\mathbf{r})$   
*very time consuming !!!*
- Generalized Born Equation, SASA, EEF1 etc..

$\epsilon=1$



$\epsilon=80$

# Classical MD Simulation: Molecular FF (cont. Lecture #1)

- Summary

$$\begin{aligned}
 U(\vec{R}) = & \underbrace{\sum_{\text{bonds}} k_i^{\text{bond}} (r_i - r_0)^2}_{U_{\text{bond}}} + \underbrace{\sum_{\text{angles}} k_i^{\text{angle}} (\theta_i - \theta_0)^2}_{U_{\text{angle}}} + \\
 & \underbrace{\sum_{\text{dihedrals}} k_i^{\text{dihe}} [1 + \cos(n_i \phi_i + \delta_i)]}_{U_{\text{dihedral}}} + \underbrace{\sum_i \sum_{j \neq i} 4 \epsilon_{ij} \left[ \left( \frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left( \frac{\sigma_{ij}}{r_{ij}} \right)^6 \right]}_{U_{\text{nonbond}}} + \sum_i \sum_{j \neq i} \frac{q_i q_j}{\epsilon r_{ij}}
 \end{aligned}$$

PDB file → geometry (red box)  
 Topology PSF file →  $n_i$ ,  $\phi_i$ ,  $\delta_i$ ,  $q_i$ ,  $q_j$   
 Parameter file → parameters (blue box)

## Force-field development is a great challenge:

- >  $K_{\{\text{bond,angle,torsion}\}}$ , equilibrium values, multiplicities, phases,  $\epsilon$ ,  $\sigma$ , partial charges

Combination of empirical (MM) and QM calculations, fit range of condensed phase data; force field has to be tested to check if reproduces correctly the structural, dynamical and thermodynamical properties of molecules that have been well-characterized experimentally.

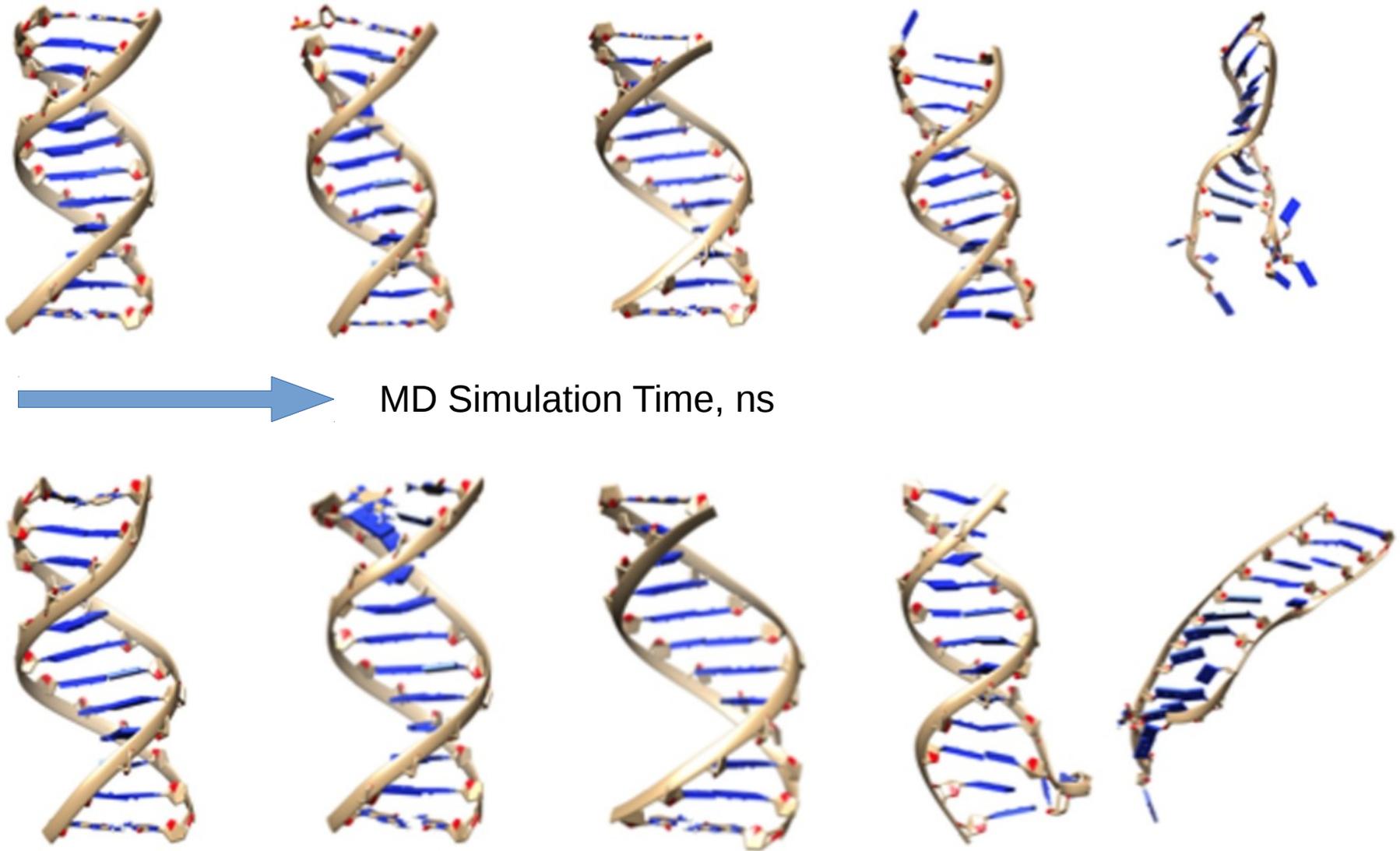
# Molecular Force Field Parametrization (cont. Lecture #1)

$$E = \sum_{\text{bonds}} K_b (r - r_0)^2 + \sum_{\text{angles}} K_\theta (\theta - \theta_0)^2 + \sum_{\text{dihedrals}} K_\phi [1 + \cos(n\phi - \delta)] + \sum_{\text{impropers}} K_\psi (\psi - \psi_0)^2 + \sum_{i>j} 4\epsilon \left[ \left( \frac{\sigma}{r} \right)^{12} - \left( \frac{\sigma}{r} \right)^6 \right] + \sum_{i>j} \frac{q_i q_j}{4\pi\epsilon_0 \epsilon r}$$

Type of data	Type of system	Phase	Type of properties	Force field parameters
structural data (exp.)	small molecules	crystalline solid phase	molecular geometry: bond lengths, angles	$r_0, \theta_0, \psi_0$
spectroscopic data (exp.)	small molecules	gas phase	intra-molecular vibrations: force constants	$K_b, K_\theta, K_\psi$
quantum-chemical calculations : energy profiles (theor.)	small molecules	gas phase	torsional-angle rotational profiles	$K_\phi, \delta, n$
quantum-chemical calculations : electron densities (theor.)	small molecules	gas phase	atom charges	charges $q_i$ (initial)
thermodynamic data (exp.)	molecules in solution, mixtures	condensed phase	heat of vaporisation, density, free energy of solvation	v. d. Waals : $\sigma_i, \epsilon_i$ charges $q_i$ (final)
dielectric data (exp.)	small molecules	condensed phase	dielectric permittivity, relaxation	charges $q_i$
transport data (exp.)	small molecules	condensed phase	transport coefficients: diffusion, viscosity	v. d. Waals : $\sigma_i, \epsilon_i$ charges $q_i$

## Example: Bad Molecular Force Field (cont. Lecture #1)

- Model **parametrization** is a great challenge !!!!



~1.5 years to stabilize double-stranded DNA

# MD Simulations: Connection to Statistical Mechanics

- **Newtonian dynamics**

$$m \frac{d^2 \mathbf{r}_i}{dt^2} = \mathbf{F}_i(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N), \quad i = 1, 2, \dots, N$$

$$U = \sum_{i=1}^N \sum_{j>i}^N u(r_{ij}) \quad \mathbf{F}_i = \sum_{j \neq i}^N \mathbf{f}_{ij}, \quad \mathbf{f}_{ij} = -\frac{du(r_{ij})}{dr_{ij}} \cdot \frac{\mathbf{r}_{ij}}{r_{ij}}$$

$$\mathbf{r}_{i=1, \dots, N}(t), \quad \mathbf{v}_{i=1, \dots, N}(t); \quad t = [t_b - t_e]$$

- Time step:

$$\Delta t \sim [0.5 - 2] \text{fs}$$

for most all-atom simulations

- **Pluses:**

- First-principle approach
- Easy to integrate
- Deterministic (no stochasticity, reproducible)

- **Minuses:**

- Abstract closed system (energy conserved)

$$E = E_{kin}(\{\mathbf{v}_i\}) + U(\{\mathbf{r}_i\})$$

- **Does not correspond to any realistic experiment !!!**

- Corresponds to microcanonical (N,V,E) ensemble in statistical mechanics

# MD Simulations: Connection to Statistical Mechanics

- **Microcanonical Ensemble**

$$\mathcal{H}(\mathbf{p}^N, \mathbf{r}^N) = \mathcal{K}(\mathbf{p}^N) + \mathcal{U}(\mathbf{r}^N)$$

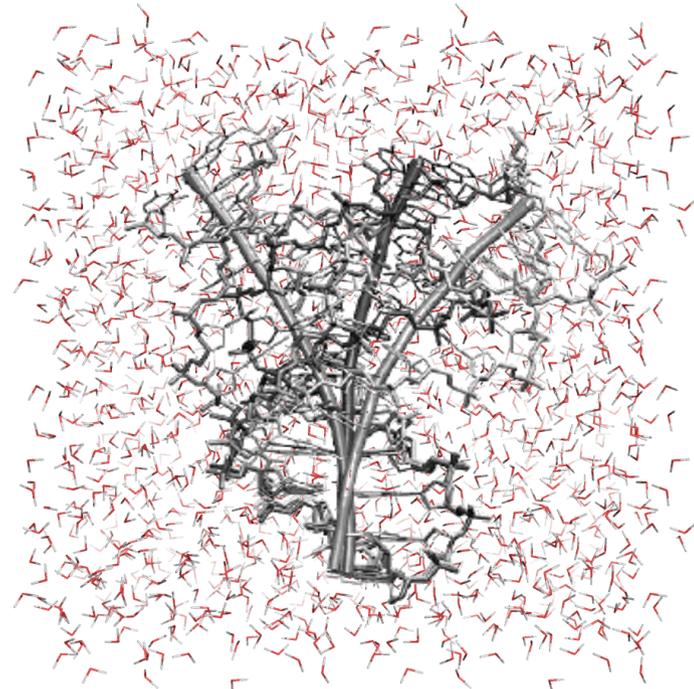
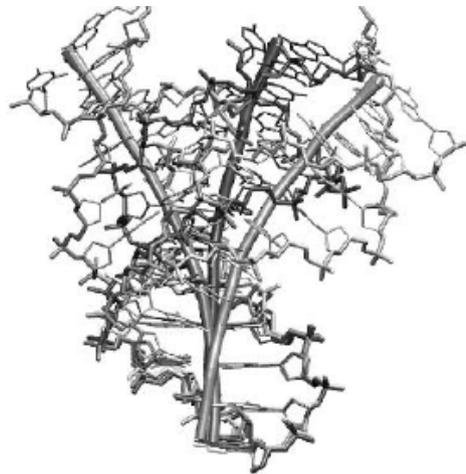
$$\Omega(N, V, E) \equiv E_0 C_{\{N\}} \int d\mathbf{p}^N d\mathbf{r}^N \delta(\mathcal{H}(\mathbf{r}^N, \mathbf{p}^N) - E)$$

$$C_{\{N\}} = \frac{1}{h^{3N} (N_A! N_B! \dots)}$$

$$S(N, V, E) = k_B \ln \Omega(N, V, E)$$

► Microstates:  $U(\{\mathbf{r}\})$ ,  $U(\{\alpha, \beta, \gamma, \delta, \epsilon, \zeta\})$

**E=constant**



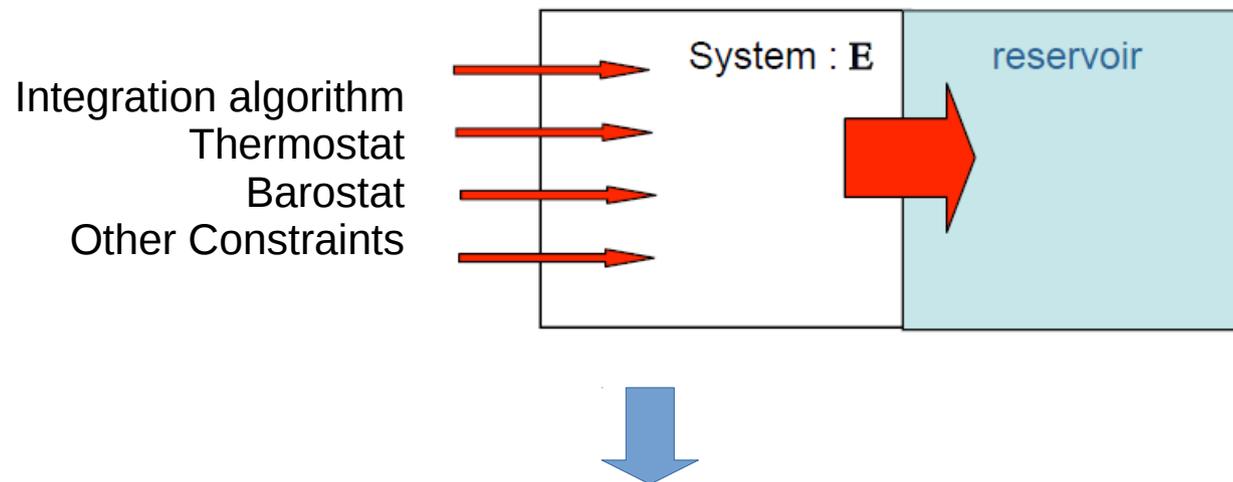
$$P_E = \frac{1}{\Omega(N, V, E)}$$

# MD Simulations: Connection to Statistical Mechanics

- **Microcanonical Ensemble: Newtonian Dynamics in Practice**

Constant energy dynamics (N,V,E) is rarely used in MD for 2 reasons:

- In real experiment it's way more easier to operate with other thermodynamic parameters: **T, P**
- **Inaccuracies** of the MD algorithm lead to **heating up** the system:
  - **We can couple the system to a heat reservoir to absorb the excess heat**



- **Canonical Ensemble:**

# MD Simulations: Connection to Statistical Mechanics

- Canonical Ensemble (N,V,T)

$$1/T = \left( \frac{dS(E)}{dE} \right)_{V,N} \quad \beta \equiv \frac{1}{k_B T}$$

$$Z(N, V, T) \equiv \int d\mathbf{p}^N d\mathbf{r}^N \exp[-\beta \mathcal{H}(\mathbf{r}^N, \mathbf{p}^N)]$$

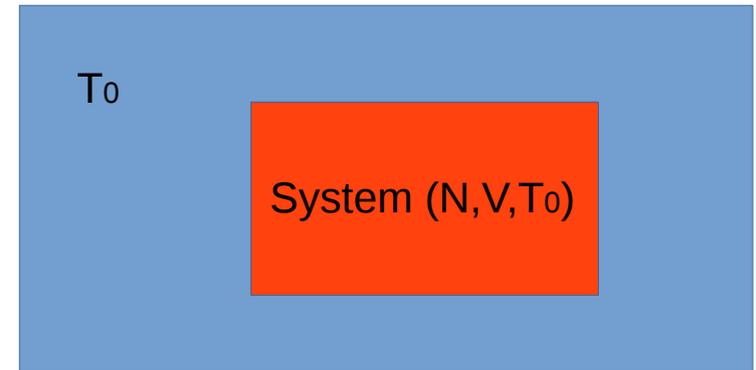
$$\mathcal{H}(\mathbf{p}^N, \mathbf{r}^N) = \mathcal{K}(\mathbf{p}^N) + \mathcal{U}(\mathbf{r}^N)$$

$$\text{Pr}_{NVT}(\mathbf{p}^N) d\mathbf{p}^N = \left( \frac{\beta}{2\pi m} \right)^{3/2} \exp\left[-\frac{\beta p^2}{2m}\right] d\mathbf{p}^N$$

$$\text{Pr}_{NVT}(\mathbf{r}^N) d\mathbf{r}^N = \frac{\exp[-\beta \mathcal{U}(\mathbf{r}^N)] d\mathbf{r}^N}{Z(N, V, T)}$$

$$F = -k_B T \ln Z(N, V, T)$$

$$\Delta F = -k_B T \ln \left( \frac{Z_2}{Z_1} \right) \longleftarrow \text{Can be measured from MD!!!}$$

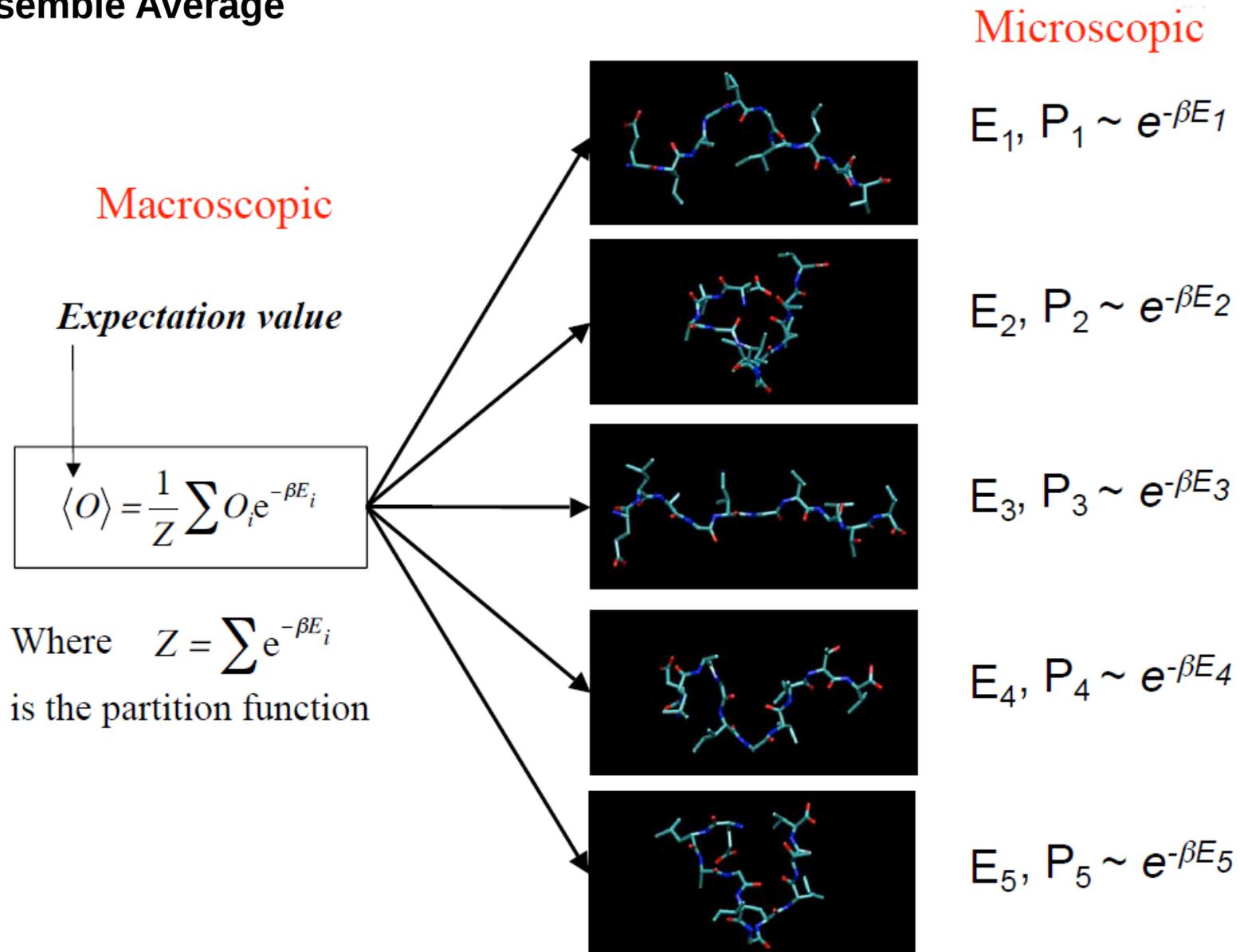


The partition function is a very complex function to compute, because it represents a measure of the whole space accessible to the system

- Maxwell-Boltzmann distribution for system velocities for canonical microstate
- Boltzmann probability distribution for canonical microstate
- Free Energy (Hemholtz)

# MD Simulations: Connection to Statistical Mechanics

- Ensemble Average



# MD Simulations: Connection to Statistical Mechanics

- **Canonical Ensemble (N,P,T)**

$$Z(N, P, T) \equiv \int dV \exp[-\beta PV] \int d\mathbf{r}^N \exp[-\beta \mathcal{U}(\mathbf{r}^N)]$$

Allow volume to fluctuate

$$\text{Pr}_{NPT}(\mathbf{r}^N) d\mathbf{r}^N = \frac{\exp[-\beta PV] \exp[-\beta \mathcal{U}(\mathbf{r}^N)] d\mathbf{r}^N}{Z(N, P, T)}$$

$$G(N, P, T) = -k_B T \ln Q(N, P, T)$$

› Gibbs Free Energy

---

$$A + B \rightleftharpoons AB, \quad K_a = \frac{[AB]}{[A][B]}, \quad K_d = \frac{1}{K_a}$$

› Binding Constants

$$K_d = \exp[\Delta G / RT]$$

# MD Simulations: Connection to Statistical Mechanics

- **Ergodic hypothesis:** marriage of statistical mechanics & computational physics (MD)

$$\langle A \rangle = \frac{\int_{\text{phase space}} A(\mathbf{r}_N) e^{-\beta E(\mathbf{r}_N)} d\mathbf{r}_N}{\int_{\text{phase space}} e^{-\beta E(\mathbf{r}_N)} d\mathbf{r}_N}$$

Ensemble Average



=

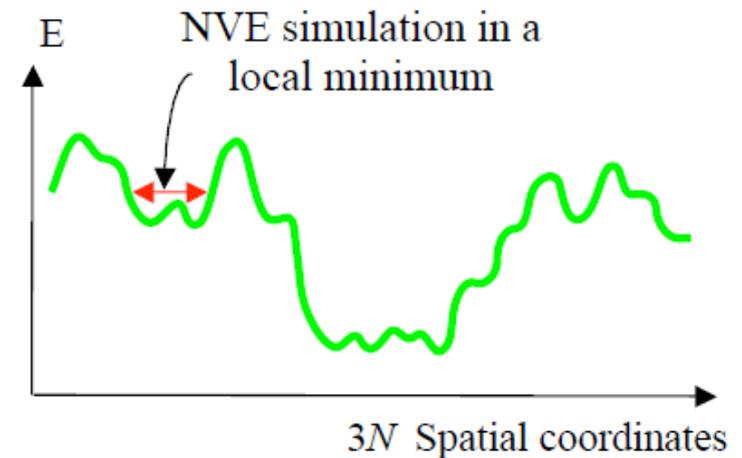
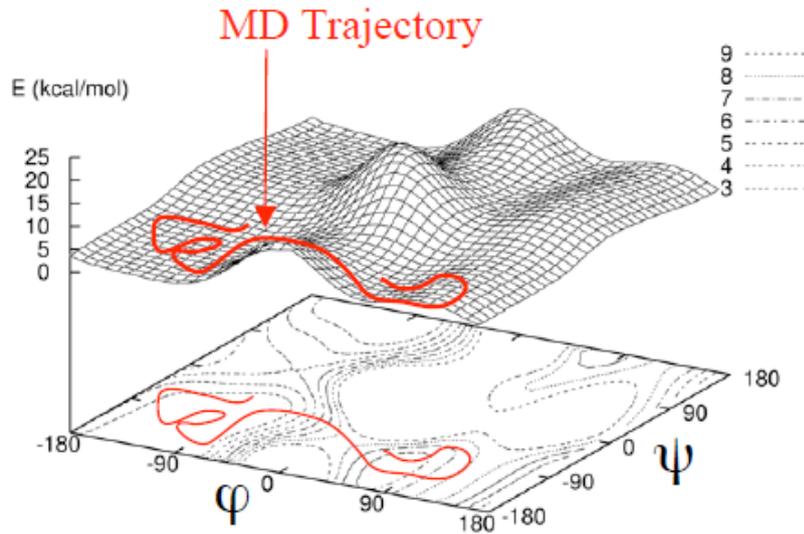
$$\langle A \rangle^\tau = \lim_{\tau \rightarrow \infty} \frac{1}{\tau} \int_0^\tau A(t) dt$$
$$\langle A \rangle = \frac{1}{M} \sum_{m=1}^M A(\mathbf{r}_N(t_m))$$

Time Average

- *The microstates sampled by MD are usually a small subset of the entire ensemble*
- *Phase space of the system has to be continuous, all (sub)-minima should be accessible on the time scale of MD simulation and be sampled with correct probabilities*

# MD Simulations: Connection to Statistical Mechanics

- **Ergodic hypothesis:** intuitive view

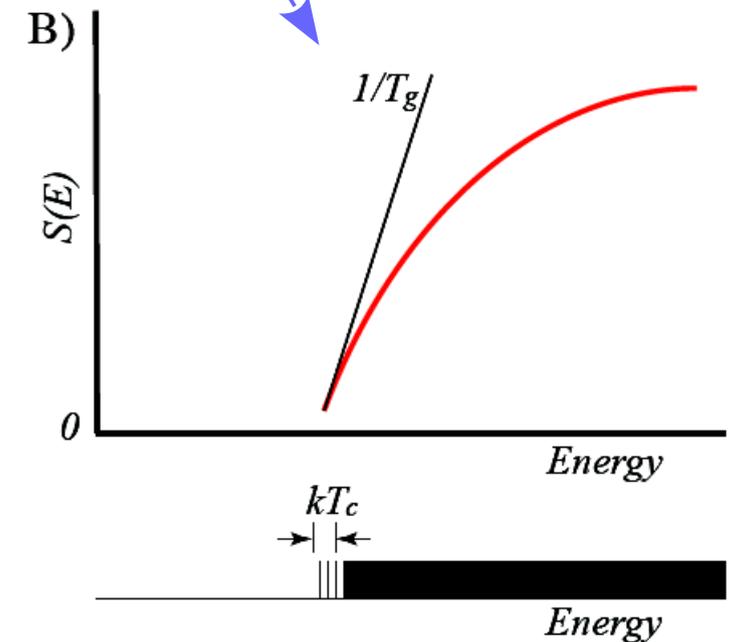
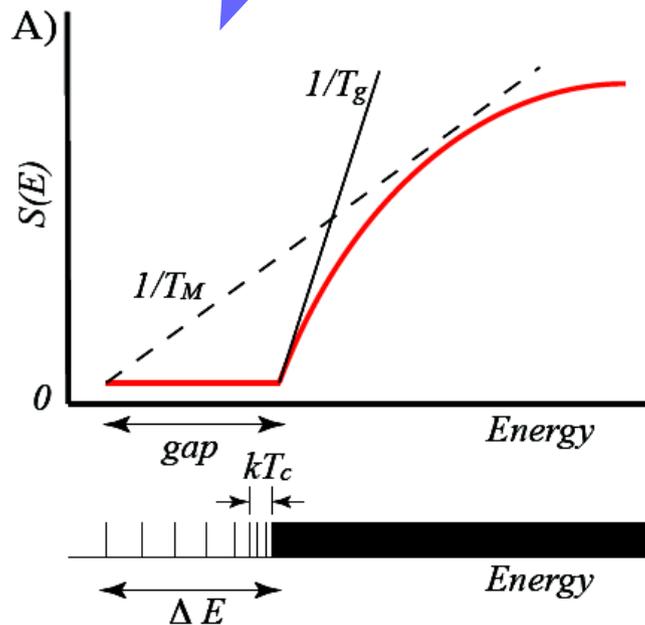
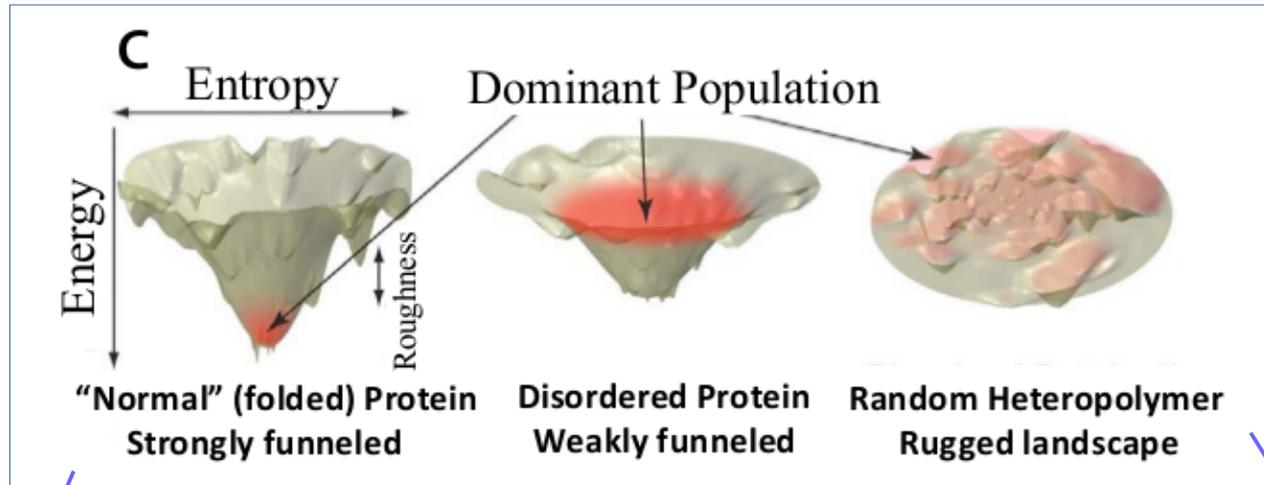


$$\langle O \rangle_{ensemble} = \frac{1}{Z} \int O(r, p) e^{-\beta E(r, p)} dr dp \quad ? \quad = \quad \frac{1}{\tau} \int_0^{\tau} O(t) dt = \langle O \rangle_{time}$$

- **Non-ergodic** systems have parts of the phase space that are inaccessible, i.e. they are not explored during timescale of interest.
- System examples:
  - Glasses
  - *Intrinsically disordered proteins*

# MD Simulations: Connection to Statistical Mechanics

- 2D representations of the protein's energy landscape



# MD Simulations: Connection to Statistical Mechanics

- **“Normal” proteins** fold via “all-or-none” mechanism
  - 200-residue protein has  $\sim 2^{100}$  conformations; it would take  $\sim 10^{10}$  years to fold (Levinthal’s paradox) – NATURE comes to rescue!!!

- **IDPs** are structurally diverse:

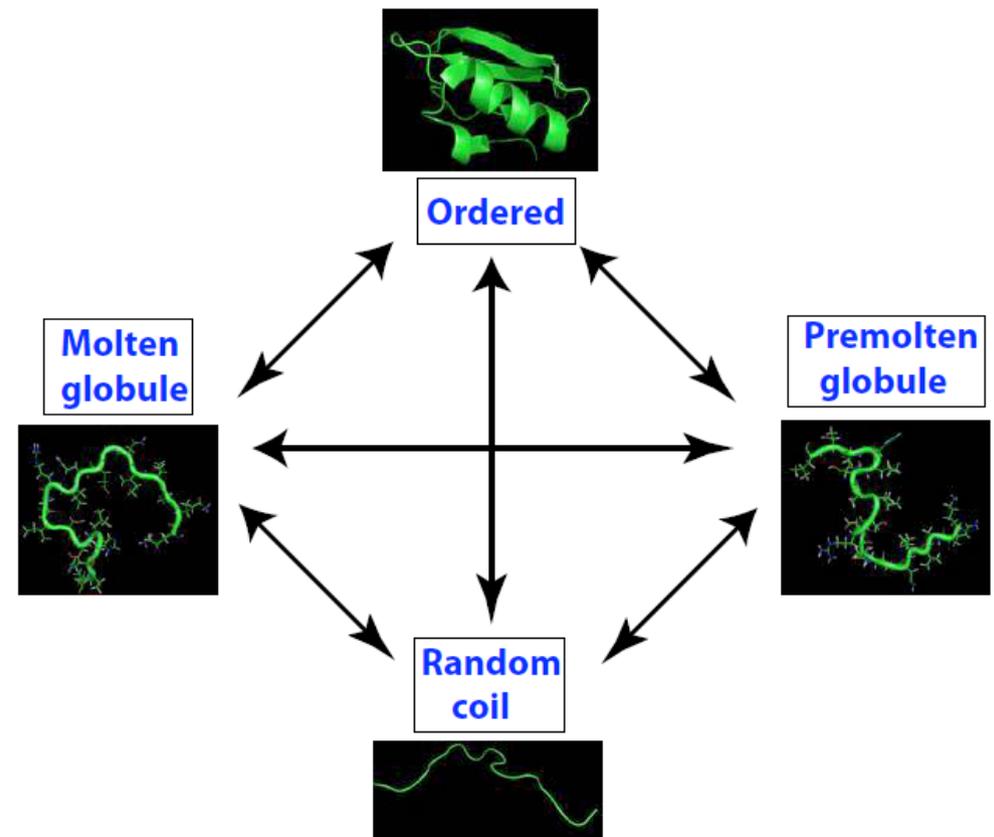
- Existence of stable intermediates
- Many folding pathways
- Tremendously hindered kinetics



Hard to study  
experimentally &  
theoretically

- **~30% of eukaryotic proteins are IDPs**
  - Can function in “normal” way
  - Can regulate many transcriptional and signal transduction processes
  - Can act as highly controlled enzymes
  - Implicated in neurodegenerative diseases (Alzheimer’s, Parkinson’s, Huntington’s )

## Quartet model for IDP functioning



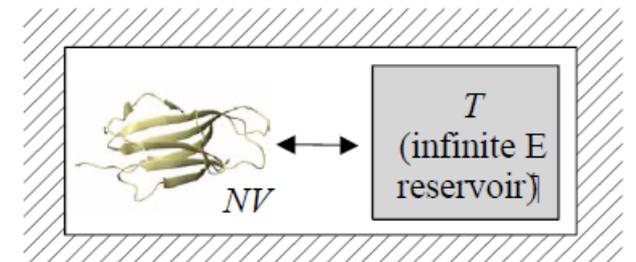
# MD Simulations: Connection to Statistical Mechanics

- **Thermodynamic Ensembles: Summary**

- **Microcanonical ensemble (constant  $N, V, E$ )**  
sampling is obtained by simple integration of the Newtonian dynamics:
  - Verlet-like algorithms



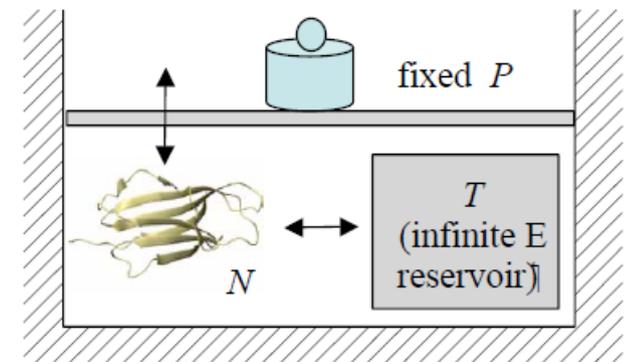
- **Canonical ensemble: constant  $N, V, T$**   
sampling is obtained by using thermostats:
  - Andersen thermostat
  - Berendsen thermostat
  - Nose-Hoover thermostat
  - Langevin Dynamics



Non-Hamiltonian dynamics

Most often used in MD

- **Canonical ensemble: constant  $N, P, T$** 
  - In addition to the thermostat, the volume of the system is allowed to fluctuate, and is regulated by barostat algorithms



# MD Simulations: Connection to Statistical Mechanics

- **Non-Hamiltonian** Dynamics to Generate **Canonical Ensembles**

- › Based on positions and velocities we have to calculate statistical quantiles such as **temperature**.

$$\left\langle \frac{1}{2} m v_i^2 \right\rangle = \frac{1}{2} k_B T$$

Equipartition theorem (kinetic theory of gases)

$$\sum_{i=1}^N \frac{m_i}{N_f k_B} v_i^2 = T(t)$$

Microscopic definition of temperature

Fluctuations on the order of  $1/\sqrt{N_f}$   $N_f=3N-3$

- › **Velocity scaling algorithms**

$$v^{new} = v^{old} \sqrt{\frac{T_{desired}}{T(t)}}$$

Temperature relaxation algorithms:

- Stochastic coupling to hypothetical heat bath
- Stochastic dynamics (Langevin)

# MD Simulations: Connection to Statistical Mechanics

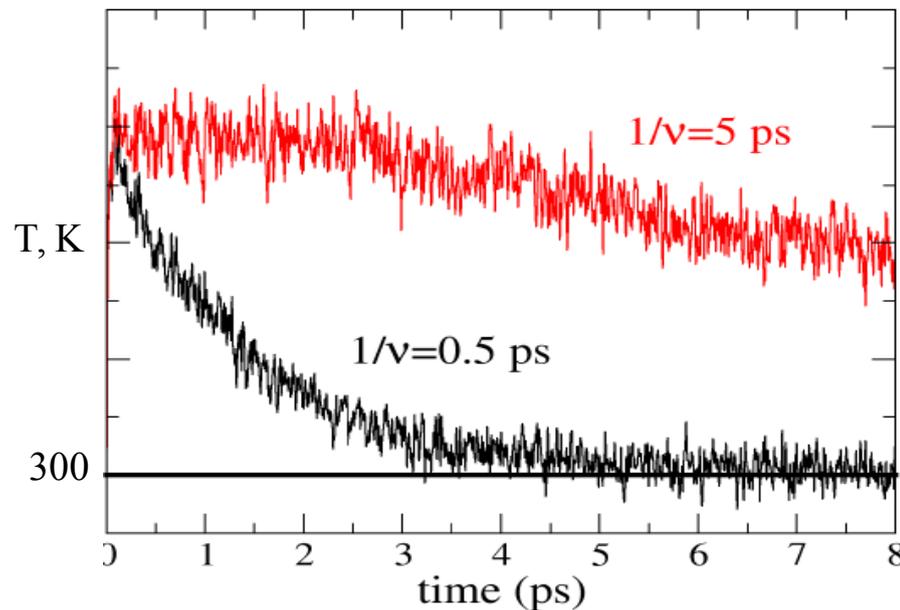
- Thermostatting Algorithms

**Berendsen algorithm:** To maintain the temperature, the system is coupled to an external heat bath with fixed temperature  $T_0$ .

Velocities are scaled at each step, such that the rate of change of temperature is proportional to the difference in temperature:

$$\frac{dT(t)}{dt} = \frac{1}{\tau} (T_0 - T(t))$$

$\tau$  being the rise time (strength of the coupling to a hypothetical heat bath).



Straight forward but not accurate

# MD Simulations: Connection to Statistical Mechanics

- Thermostatting Algorithms

**Andersen algorithm:** Coupling of the system with a bath is performed through a **stochastic process**. Modifies velocities of particles by the presence of instantaneous forces.

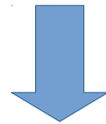
- ❑ Between these stochastic "collisions", the system evolves with the usual Newtonian dynamics.
- ❑ Coupling strength is controlled by a collision frequency denoted by  $\nu$ .
- ❑ Stochastic collisions are assumed to be totally uncorrelated. This leads to a Poissonian distribution of collisions  $P(\nu, t)$ :

$$P(\nu, t) = e^{-\nu t}$$

Produces a canonical distribution

$$v^{new} = v^{old} \sqrt{\frac{T_{desired}}{T(t)}}$$

But... it becomes a combination of MD and MC... !!!



What about dynamical properties ?

$$D = \lim_{t \rightarrow \infty} \frac{1}{6Nt} \left\langle \sum_{i=1}^N [\mathbf{r}_i(t) - \mathbf{r}_i(0)]^2 \right\rangle$$

$$C_v(t) = \frac{\sum_{i=1}^N \langle \mathbf{v}_i(t) \cdot \mathbf{v}_i(0) \rangle}{\sum_{i=1}^N (\mathbf{v}_i(0))^2}$$

Velocities and, thus, positions are brutalized by such rescaling approaches

# MD Simulations: Connection to Statistical Mechanics

- Thermostatting Algorithms

**Nosé-Hoover thermostat:** a clever way of using extended Lagrangian formalism (lecture 1)

Phase space extended by two extra variables :  $(\underbrace{r, p}_{\text{physical variables}}, \eta, p_\eta)$

$$\begin{aligned} \dot{r}_i &= \frac{p_i}{m_i} && \leftarrow \text{Newton} \\ \dot{p}_i &= -\frac{\partial \Phi}{\partial r_i}(r, t) - \frac{p_\eta}{Q} p_i && \leftarrow \text{friction term} \\ \dot{\eta} &= \frac{p_\eta}{Q} \\ \dot{p}_\eta &= \sum_i \frac{p_i^2}{m_i} - N_{df} k_B T . && \leftarrow \text{temperature regulation} \end{aligned}$$

- One can demonstrate that the **canonical distribution is reproduced** for the physical variables

$$Z(N, V, T) = \int dr dp e^{-\beta H(r, p)}$$

- Conserved quantity :

$$H'(\Gamma, t) = \sum_{i=1}^N \frac{p_i^2}{2m_i} + \Phi(r) + \frac{p_\eta^2}{2Q} + k_B T \bar{\eta}$$

Very accurate, deterministic.

# MD Simulations: Connection to Statistical Mechanics

- Thermostatting (and Sampling) by Means of Langevin Dynamics

$$m_i \ddot{\mathbf{r}}_i(t) = \underbrace{-\frac{\partial \mathcal{U}(\mathbf{r}^N(t))}{\partial \mathbf{r}_i}}_{\text{Newton's equation}} - \underbrace{\int_0^t d\tau \dot{\mathbf{r}}_i(\tau) \gamma_i(t-\tau)}_{\text{Friction force with memory kernel}} + \underbrace{\xi_i(t)}_{\text{Random force}}$$

## Langevin equation

$$m_i \ddot{\mathbf{r}}_i(t) = -\frac{\partial \mathcal{U}(\mathbf{r}^N(t))}{\partial \mathbf{r}_i} - \gamma_i \dot{\mathbf{r}}_i(t) + \xi_i(t)$$

Friction coeff.  $\gamma$ :

*too small (weak coupling), too big (no acceleration), constant/different for all atoms*

“White” noise

$$\langle \xi_i(t) \rangle = 0$$

$$\langle \xi_i(0) \xi_j(t) \rangle = \underline{2m_i k_B \gamma_i T} \delta(t) \delta_{ij}$$

Fluctuation-dissipation theorem

**Overdamped Langevin: Brownian Dynamics**

$$\dot{\mathbf{r}}_i(t) = \gamma_i^{-1} m_i^{-1} [\mathbf{F}_i(t) + \xi_i(t)]$$



**Brownian Motion**

$$\dot{\mathbf{r}}_i(t) = \gamma_i^{-1} m_i^{-1} \xi_i(t)$$

# MD Simulations: Connection to Statistical Mechanics

- Thermostatting (and Sampling) by Means of Langevin Dynamics

$$m_i \ddot{\mathbf{r}}_i(t) = -\frac{\partial \mathcal{U}(\mathbf{r}^N(t))}{\partial \mathbf{r}_i} - \gamma_i \dot{\mathbf{r}}_i(t) + \xi_i(t)$$

Identical algorithm is used for both **thermostatting** (*explicit* solvent simulations) and actual **dynamics in frictionous environment** (e.g. *implicit* solvent simulations)

Example: Na<sup>+</sup> ion diffusion

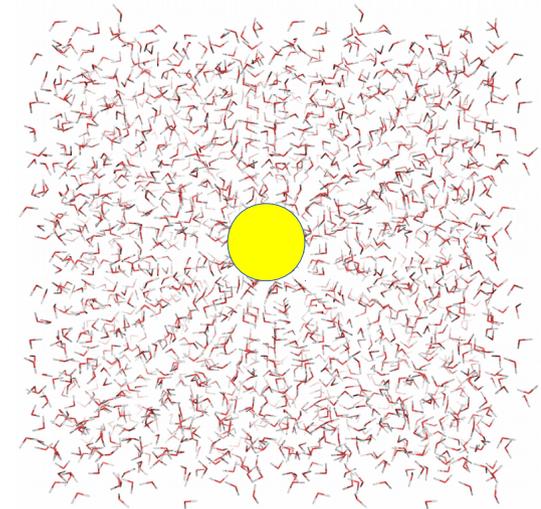
*implicit* solvent



experimental

$$D(\text{Na}^+) = 133 \text{ \AA}^2/\text{ns}$$

*explicit* solvent



damping\_coeff.  $\sim [1 - 5] \text{ ps}^{-1}$

$$\text{NAMM: } D(\text{Na}^+) = 134 \text{ \AA}^2/\text{ns}$$

damping\_coeff.  $= k_B T / (m D) = 81 \text{ ps}^{-1}$

$$\text{NAMM: } D(\text{Na}^+) = 136 \text{ \AA}^2/\text{ns}$$

# MD Simulations: Connection to Statistical Mechanics

- **Brownian Motion**

$$\dot{\mathbf{r}}_i(t) = \gamma_i^{-1} m_i^{-1} \boldsymbol{\xi}_i(t)$$

It can be demonstrated (Fokker-Plank formalism) that density of Brownian particles obeys diffusion equation:

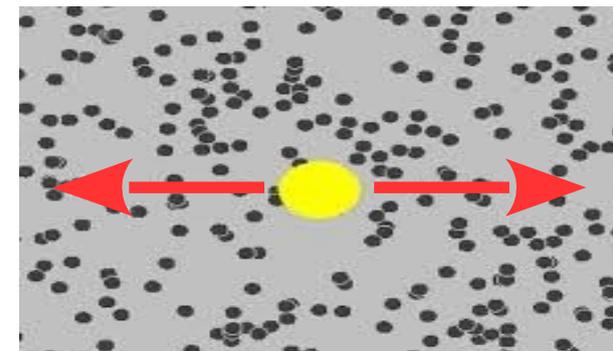
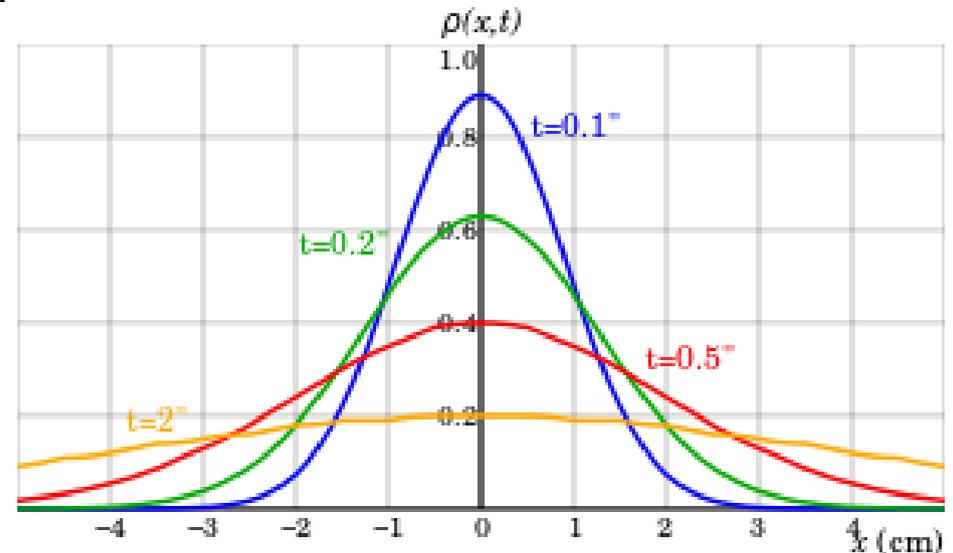
$$\frac{\partial \rho}{\partial t} = D \cdot \frac{\partial^2 \rho}{\partial x^2} \quad \triangleright \text{Diffusion equation}$$

$$D = k_B T / \gamma \quad \triangleright \text{Einstein relation}$$

$$D = \frac{k_B T}{6\pi \eta r} \quad \triangleright \text{Stokes-Einstein relation}$$

$$\rho(x, t) = \frac{N}{\sqrt{4\pi D t}} e^{-\frac{x^2}{4Dt}}$$

$$\langle x(t) \rangle = 0, \quad \langle x(t)^2 \rangle = 2 D t$$

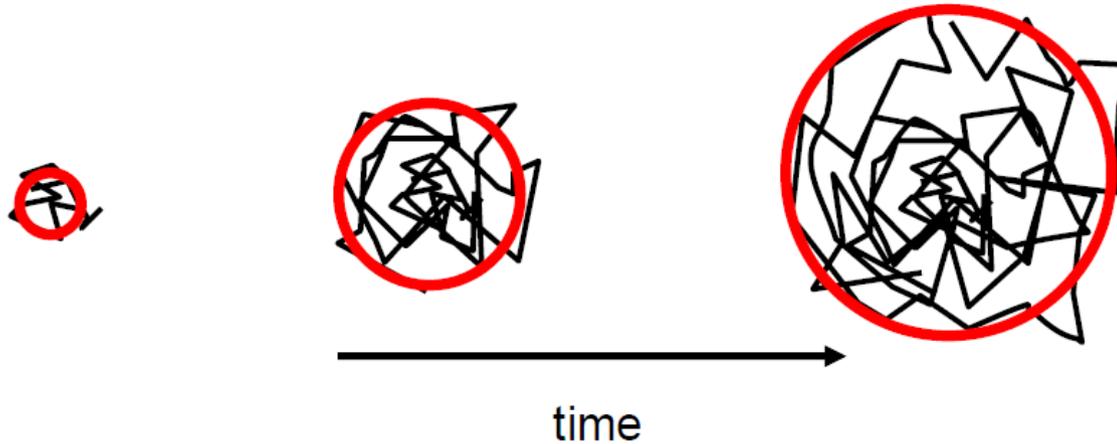


x



# MD Simulations: Connection to Statistical Mechanics

- Diffusion Coefficient from MD simulations

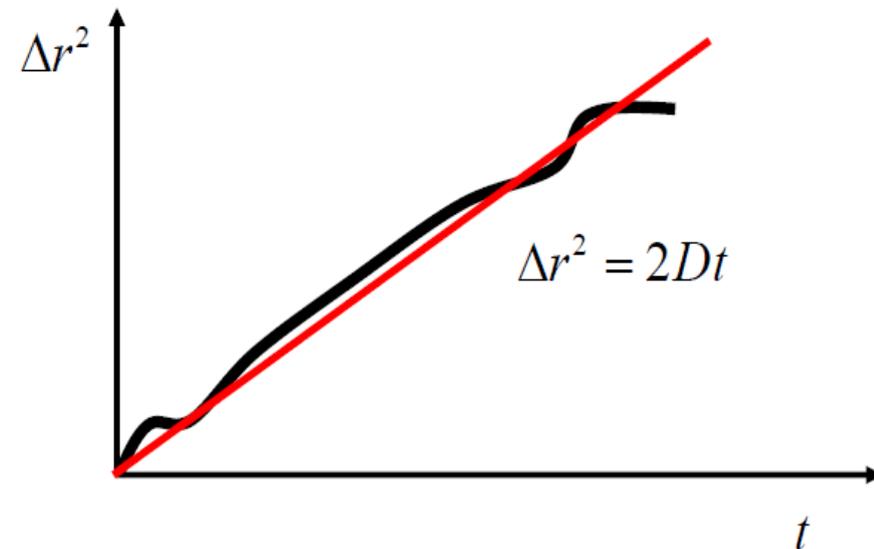


$$D = \frac{1}{2} \frac{1}{d} \frac{\Delta r^2}{\Delta t}$$

Factor 1/2 = no directionality in (forth/back)

Factor  $d = 1, 2, \text{ or } 3$  due to 1D, 2D, 3D (dimensionality)

$$\langle \Delta r^2(t) \rangle = \frac{1}{N} \sum_i \underbrace{(r_i(t))}_{\text{Position of atom } i \text{ at time } t} - \underbrace{r_i(t=0)}_{\text{Position of atom } i \text{ at time } t=0}$$



# MD Simulations: Connection to Statistical Mechanics

- Diffusion Coefficient from MD simulations: Periodic Boundary Effects

*J. Phys. Chem. B* **2004**, *108*, 15873–15879

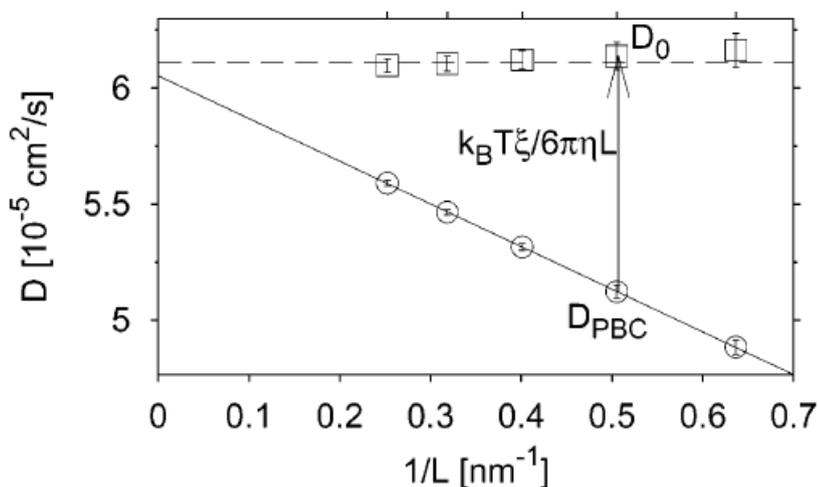
## System-Size Dependence of Diffusion Coefficients and Viscosities from Molecular Dynamics Simulations with Periodic Boundary Conditions

In-Chul Yeh and Gerhard Hummer\*

*Laboratory of Chemical Physics, Building 5, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, Maryland 20892-0520*

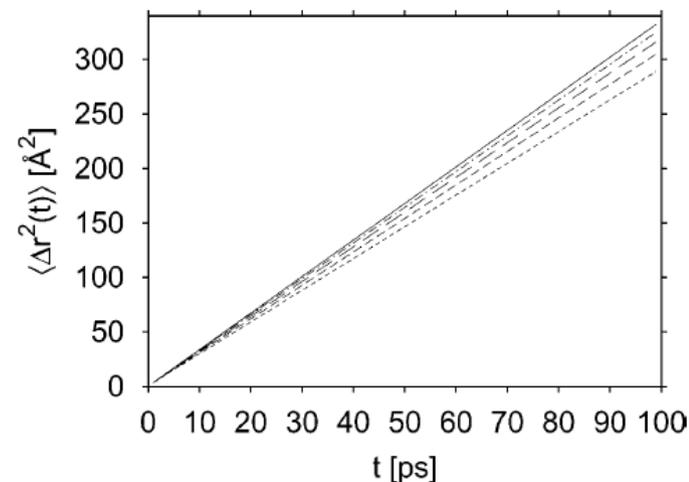
*Received: May 26, 2004; In Final Form: July 23, 2004*

$$D_{\text{PBC}} = D_0 - \frac{k_B T \xi}{6\pi\eta L}$$



←  
correction

$$D_{\text{PBC}} = \lim_{t \rightarrow \infty} \frac{\partial}{\partial t} \frac{\langle |\mathbf{r}(t) - \mathbf{r}(0)|^2 \rangle}{6}$$

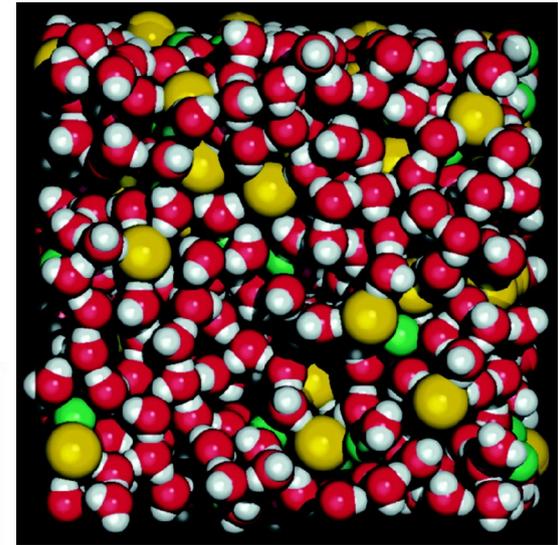
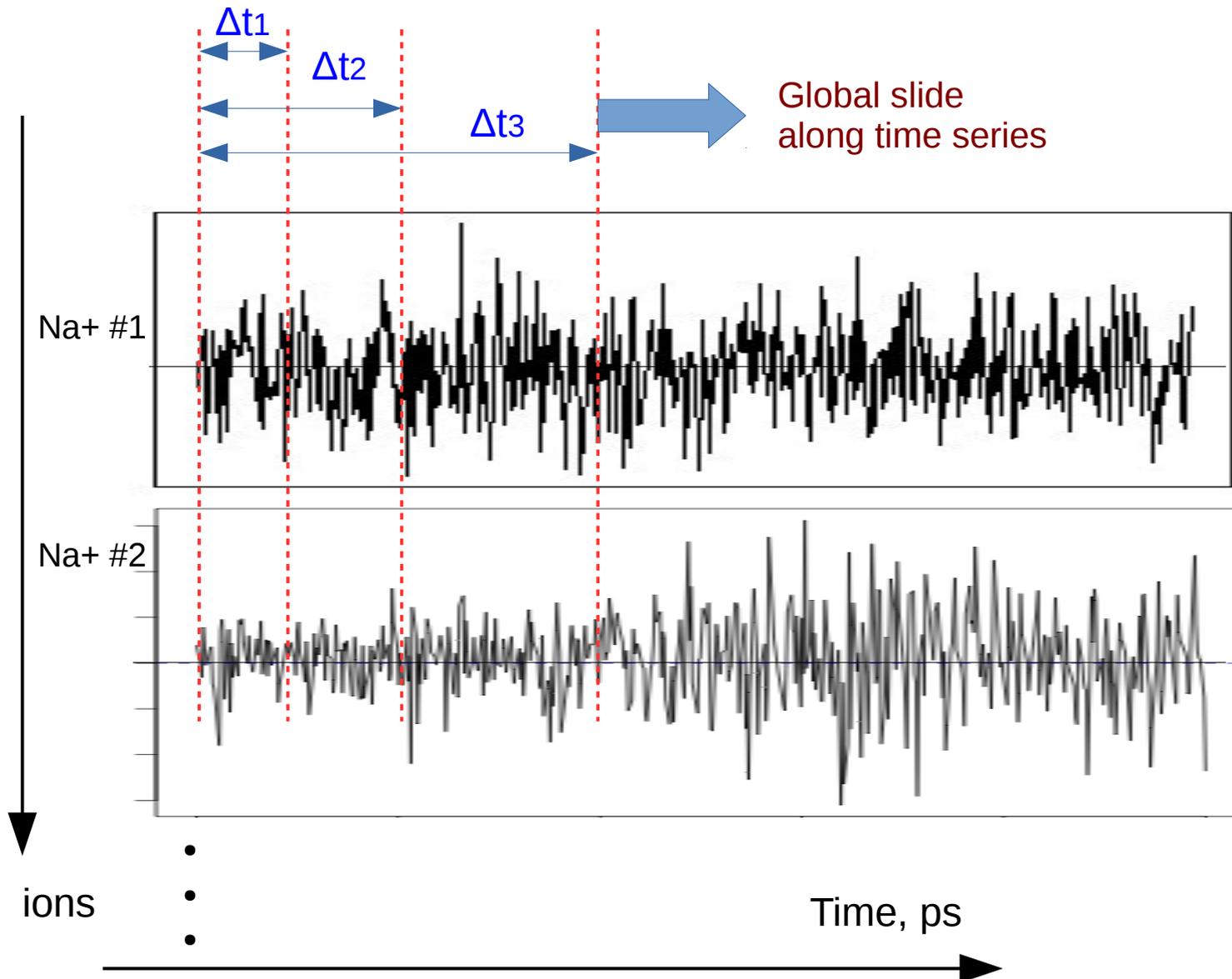


**Figure 1.** Mean-square displacement of TIP3P water as a function of time  $t$ . The lines are for  $N = 128, 256, 512, 1024,$  and  $2048$  water molecules (bottom to top).

# MD Simulations

- Time series and **time correlation** functions

NaCl at  $c \sim 150$  mM in water;  $\#(\text{NaCl}) = 15$



**Double averaging:**

- Over time interval
- Over molecules

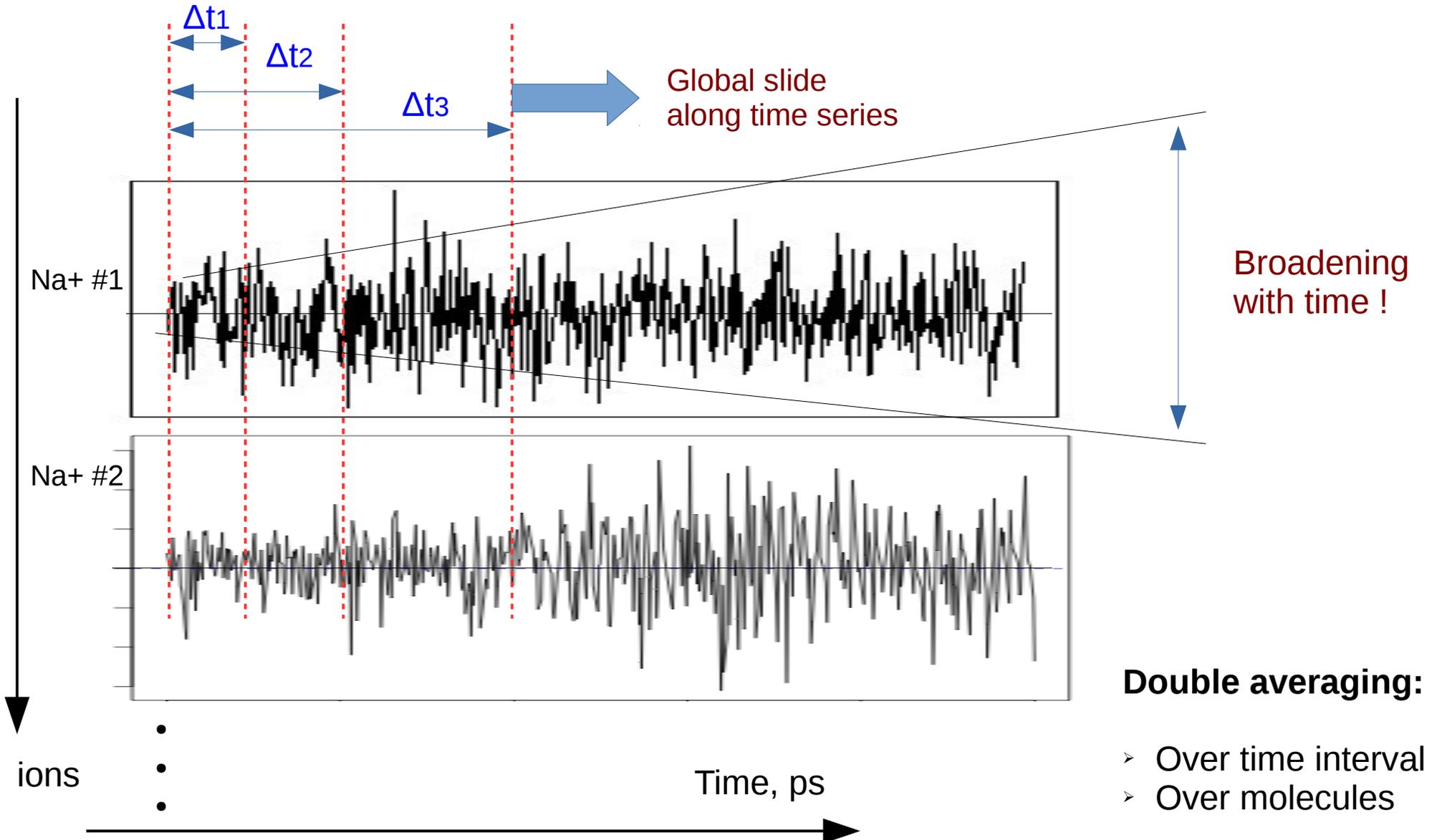
**Q:**

- Is these time series for particle's displacement (diffusion) is realistic ?

# MD Simulations

- Time series and **time correlation** functions

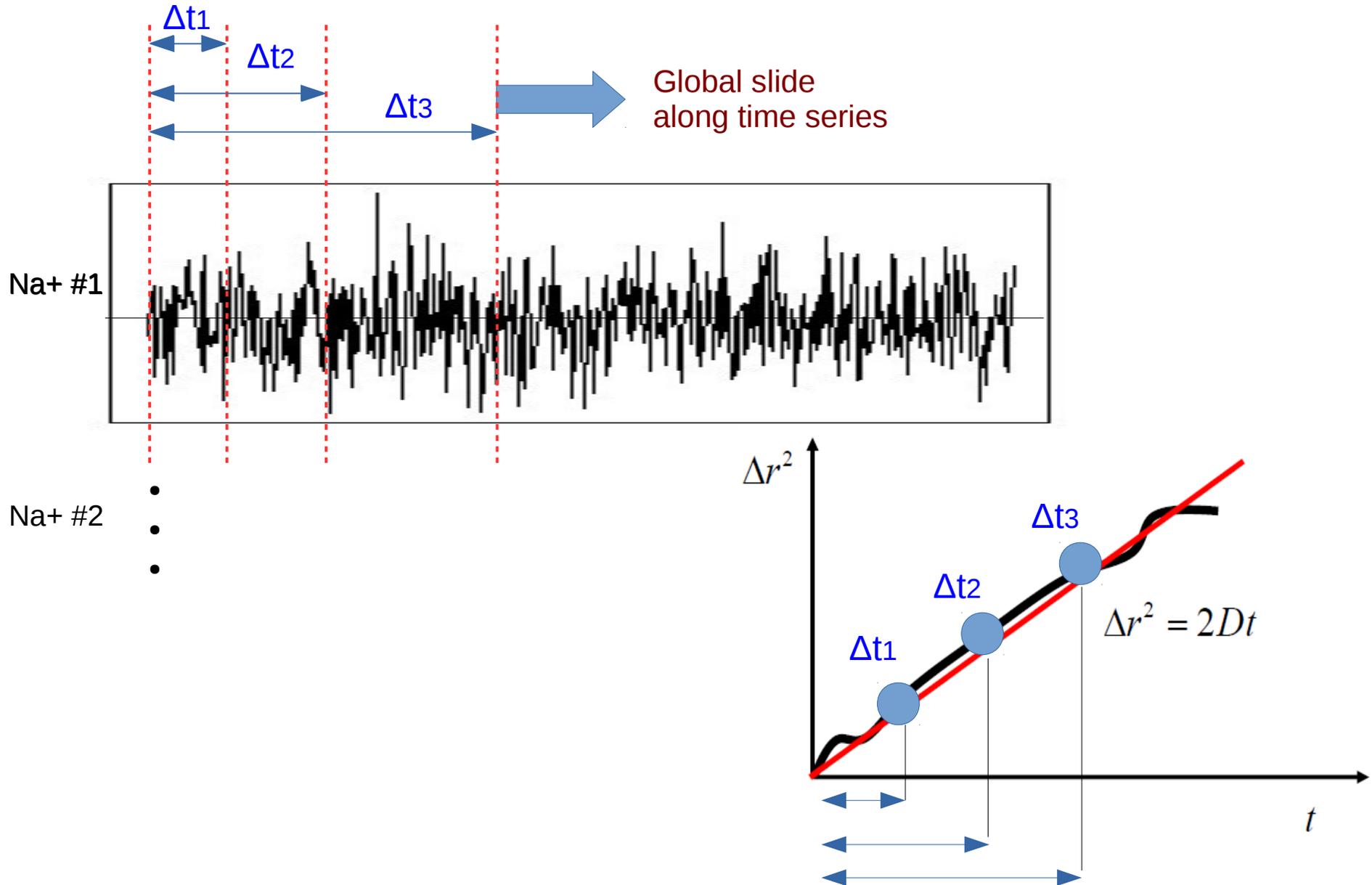
NaCl at  $c \sim 150$  mM in water;  $\#(\text{NaCl}) = 15$



# MD Simulations

- Time series and **time correlation** functions

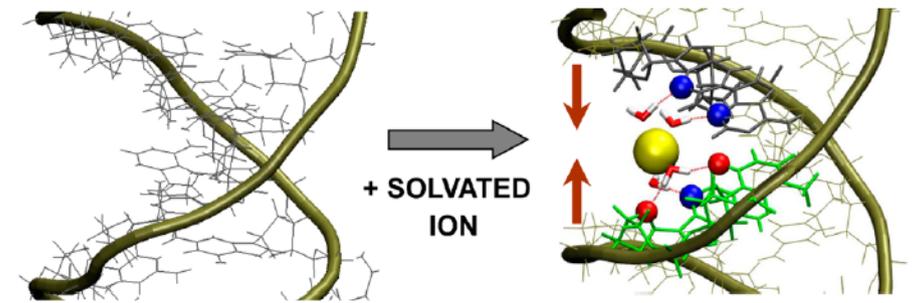
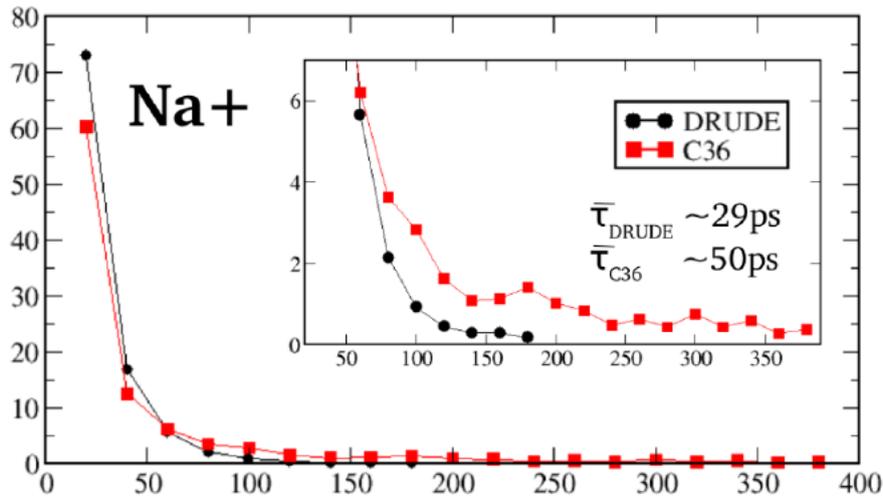
NaCl at  $c \sim 150$  mM in water;  $\#(\text{NaCl}) = 15$



# MD Simulations

- Time series and **time correlation** functions: other examples

- Na<sup>+</sup> ion residence time in the DNA minor groove



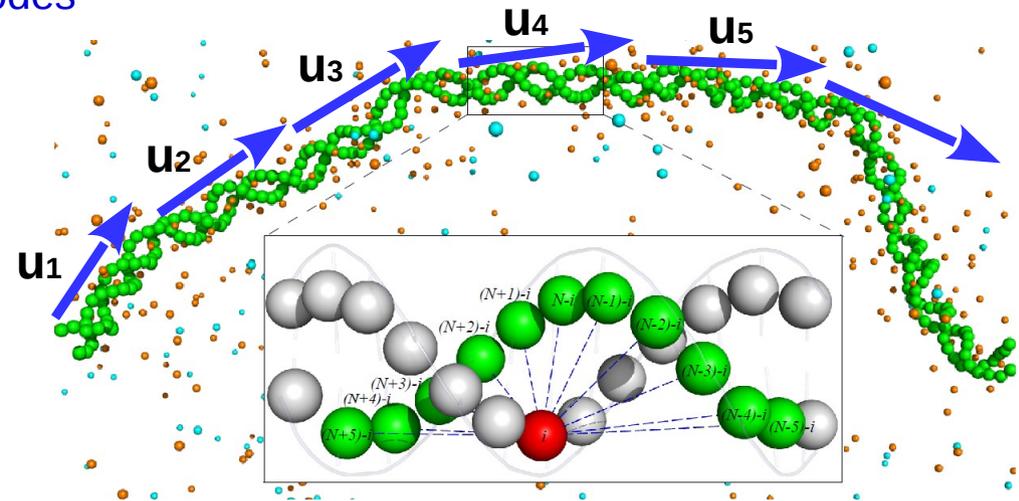
J. Chem. Theory Comput. 2015, 11, 4473–4485

- DNA persistence length and correlation modes

$$\langle \mathbf{u}_1 \mathbf{u}_{i+1} \rangle = \exp(-i \cdot \bar{a} / l_p)$$

$$C^i(t) = \langle \cos[\Delta\alpha^i(t)] \rangle, \quad \Delta\alpha^i \equiv \alpha^i(0) - \alpha^i(t)$$

Longevity of MD simulation  $\sim 500\tau_{\text{max}}$



PNAS, 2010 107 (47) 20340-20345

# MD Simulations

- Time series and **spacial correlation** functions

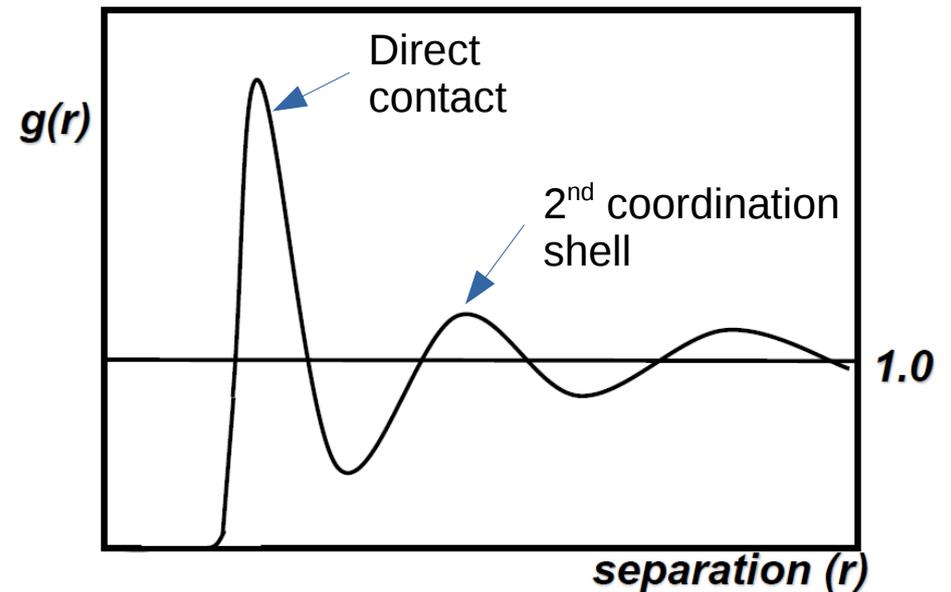
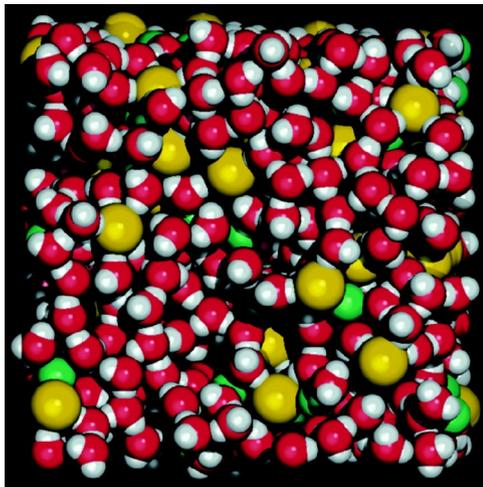
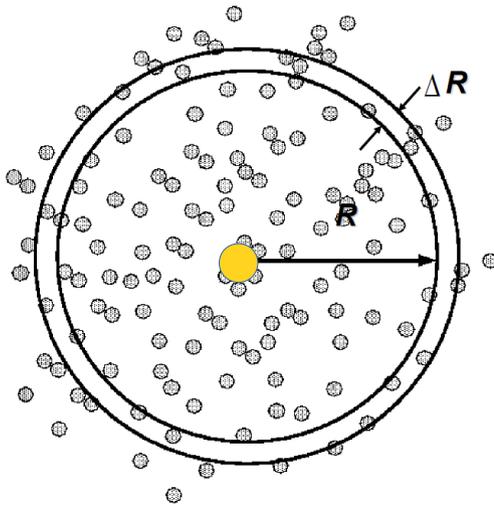
Radial Distribution Function: simple case of spherical symmetry

$$\rho g(\mathbf{r}) = \frac{1}{N} \left\langle \sum_{i=1}^N \sum_{j \neq i}^N \delta(\mathbf{r} - \mathbf{r}_{ij}) \right\rangle$$

$$g(r) = \frac{V}{N} \cdot \frac{\overline{N(r)}}{J(r)}$$

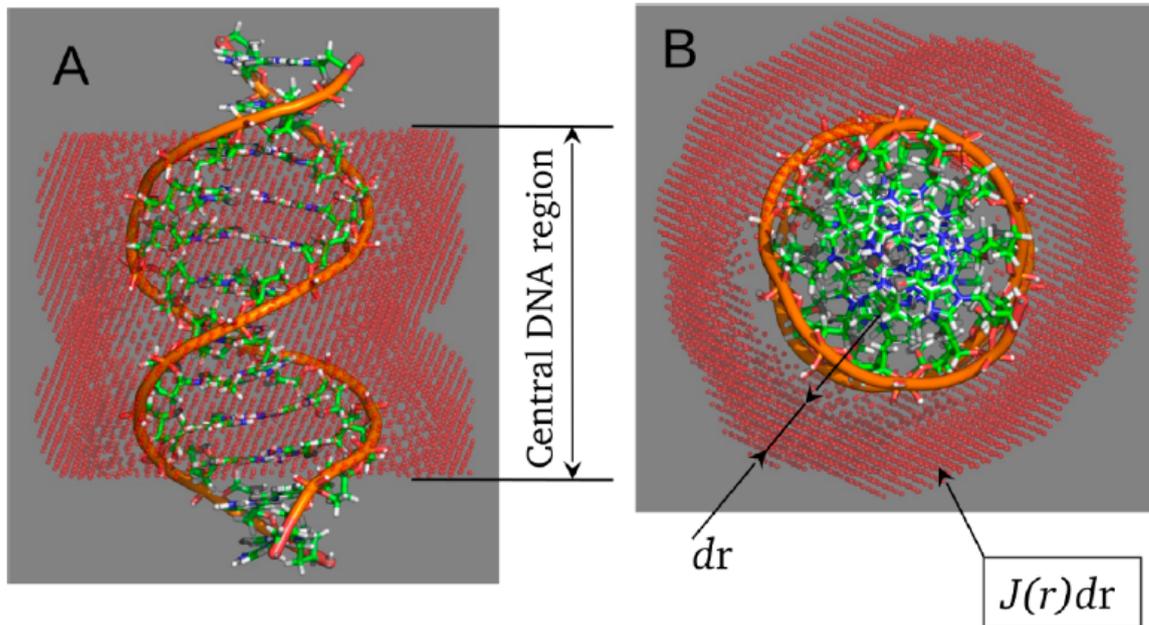
$$n(r) = \rho \int_0^r g(r) J(r) dr$$

$$J(r) = 4\pi r^2$$



# MD Simulations

- Time series and **spacial correlation** functions  
Radial Distribution Function: asymmetrical case



$$g(r) = \frac{V}{N} \cdot \frac{\overline{N(r)}}{J(r)}$$

$J(r)$  Needs to be  
computed  
numerically

- Competitive binding of ions/small molecules to the solute
- Binding to specific areas of biomolecule (DNA minor groove)
- Biomolecular partial specific volume (study changes in hydration layer caused by changes in environment, e.g. salt concentration)

# MD Simulations

- **Partial specific volume (PSV):**
- PSV is intimately related to statistical-mechanical formulation of excess volume caused by insertion of the solute into the solvent:

$$\Delta V = \int_{|\vec{r}| < \lambda} d\vec{r} \rho(\vec{r}) \left( \frac{1}{\rho(\vec{r})} - \frac{1}{\rho_0} \right)$$

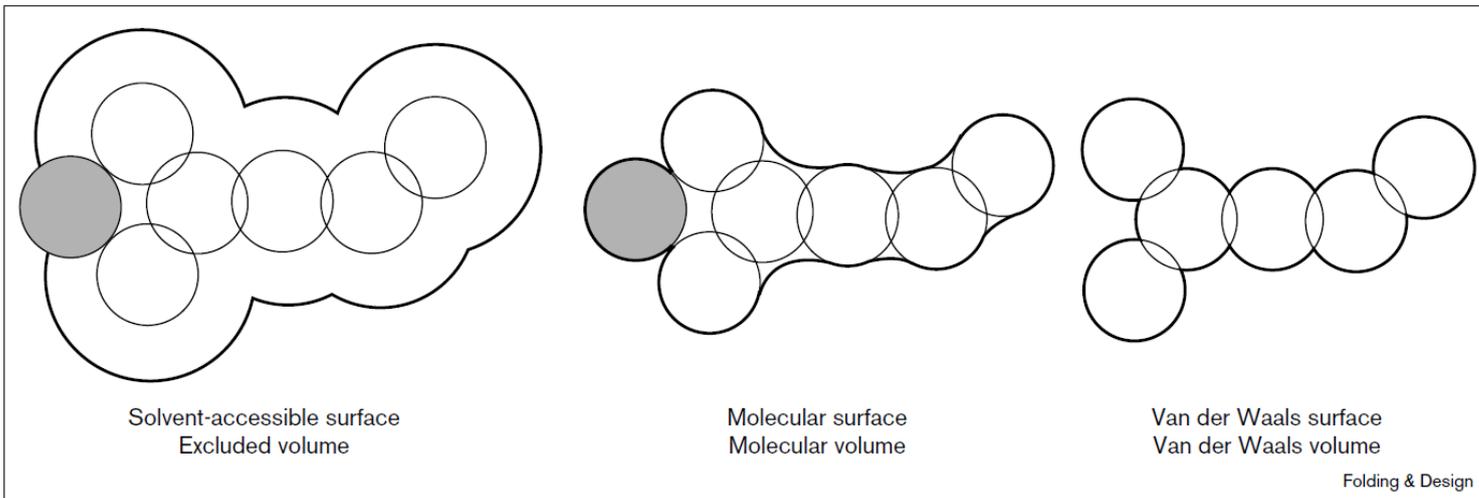


$$\bar{v}_2 = v_2 + \delta_1(v_1 - v_1^0)$$

$$= - \int_{|\vec{r}| < \lambda} d\vec{r} (g_{uv}(\vec{r}) - 1) \stackrel{\text{def}}{=} \Delta V(\lambda)$$

- $v_2$  - "intrinsic" solute volume
- $\delta_1$  - # of waters in the hydration layer
- $v_1$  - PSV of the water in hydration layer
- $v_0$  - PSV of the water in the bulk

## Surface area & intrinsic volume definitions

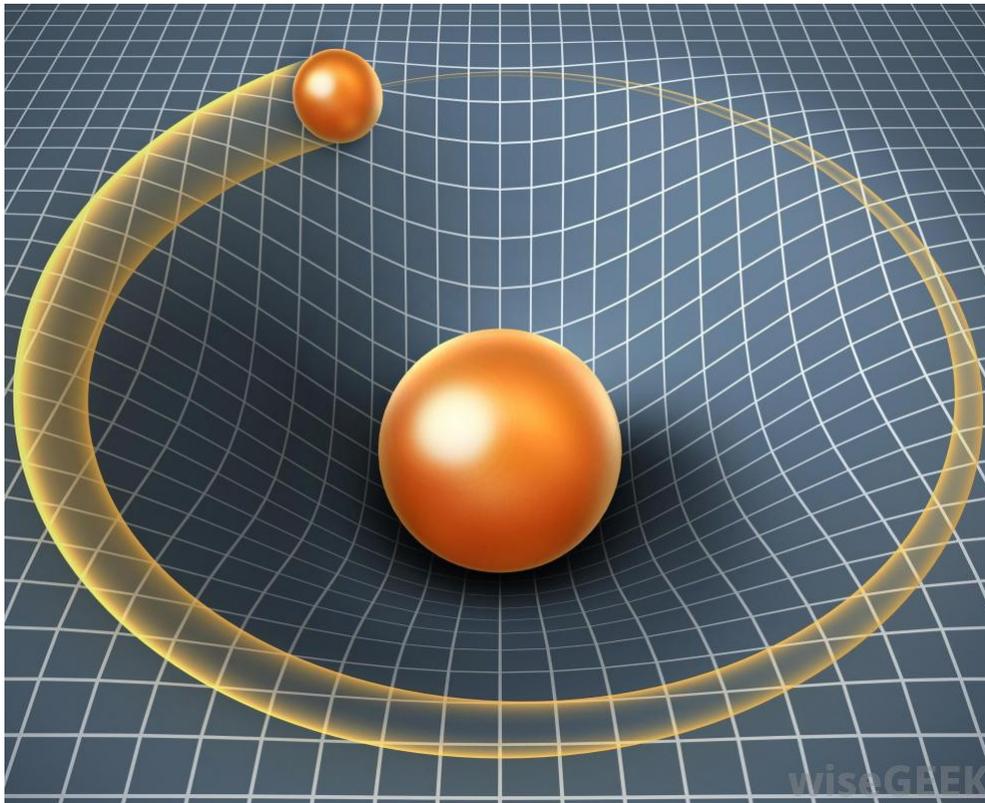


MD can address all these issues and provide closest correspondence with the theory!!

# MD Simulations

- **Partial specific volume (PSV): intuitive picture**

Analogy with disturbance of time-mater by gravitation



- › Time-matter → solvent
- › “Black hole” → biomolecule
- › Moon → ions/other solvent molecules

PSV may be thought of as a quantitative measure of perturbation of the bulk solvent by biological molecule

# MD Simulations

- Partial specific volume (PSV):

$$\Delta V = \int_{|\vec{r}| < \lambda} d\vec{r} \rho(\vec{r}) \left( \frac{1}{\rho(\vec{r})} - \frac{1}{\rho_0} \right)$$
$$= - \int_{|\vec{r}| < \lambda} d\vec{r} (g_{uv}(\vec{r}) - 1) \stackrel{\text{def}}{=} \Delta V(\lambda)$$



$$\bar{v}_2 = v_2 + \delta_1(v_1 - v_1^0)$$

- $v_2$  - “intrinsic” solute volume
- $\delta_1$  - # of waters in the hydration layer
- $v_1$  - PSV of the water in hydration layer
- $v_0$  - PSV of the water in the bulk

