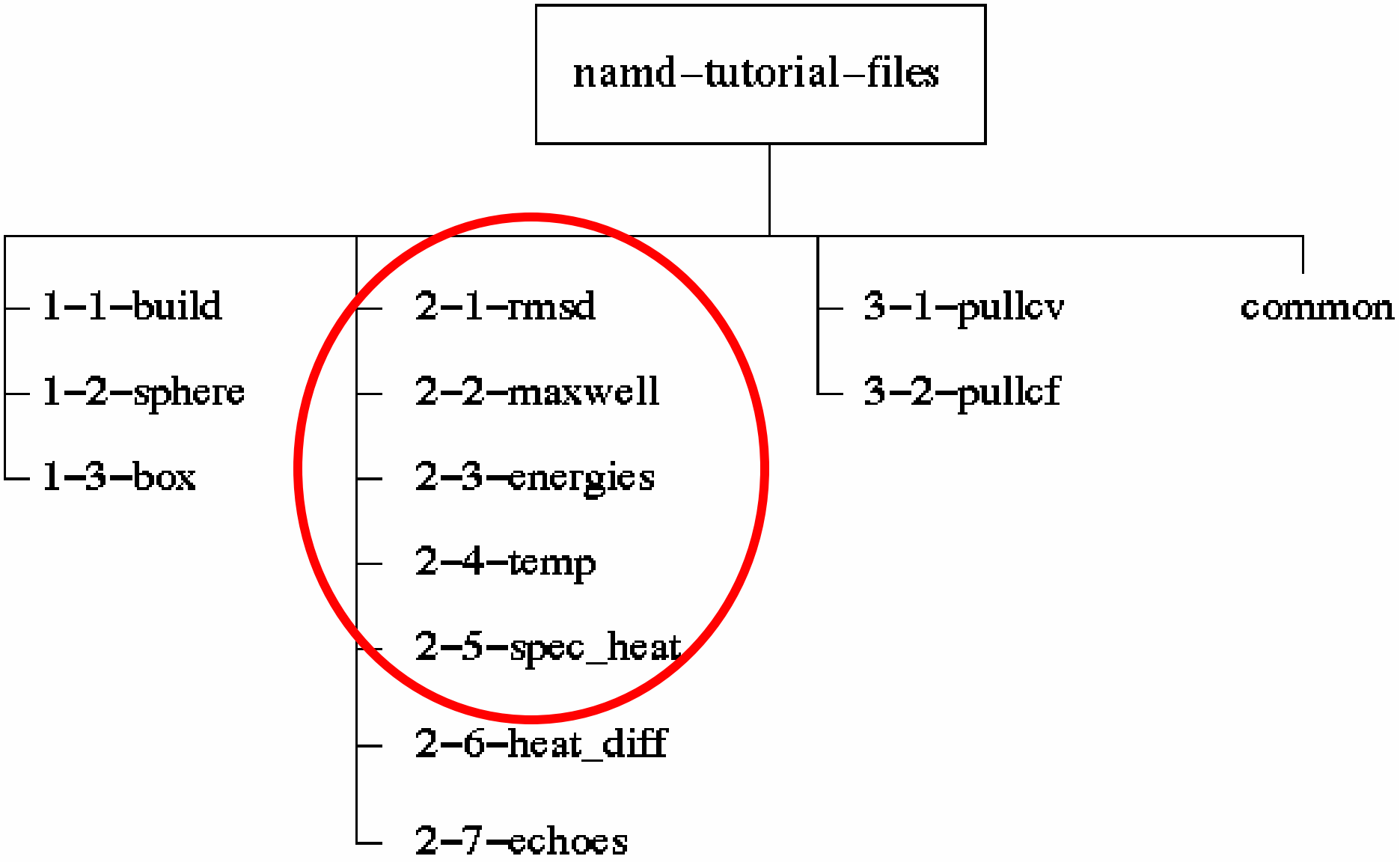
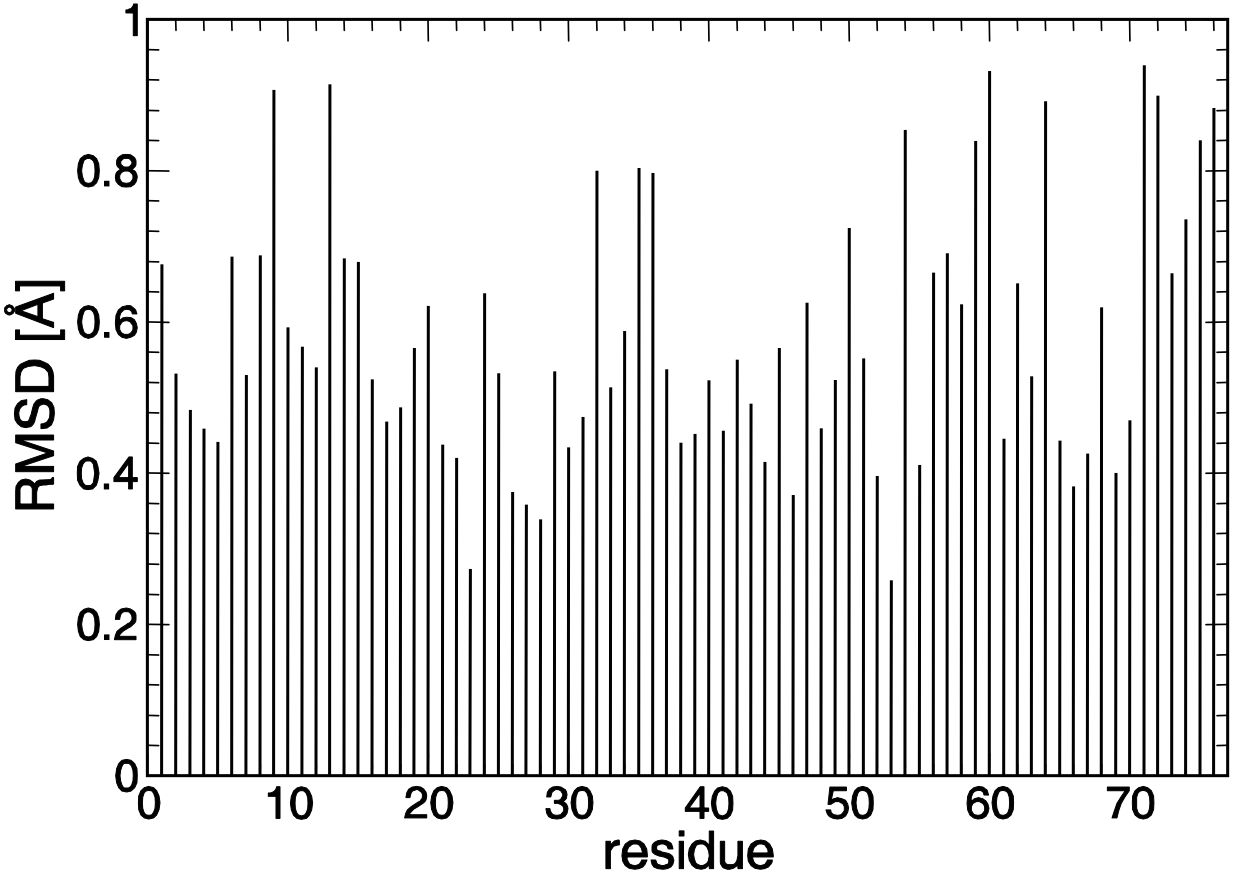
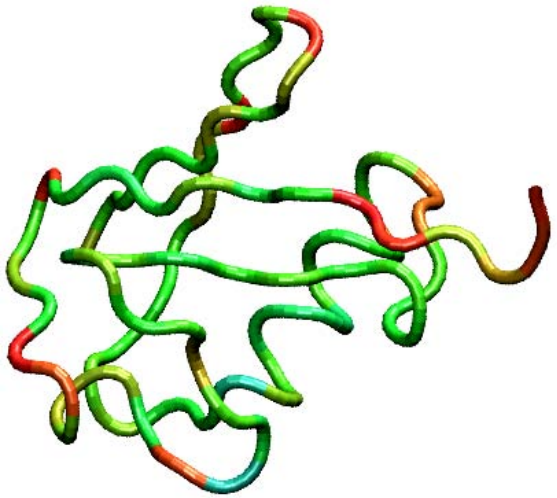
Organization of NAMD Tutorial Files



2.1.1. RMSD for individual residues

Objective: Find the average RMSD over time of each residue in the protein using VMD. Display the protein with the residues colored according to this value.

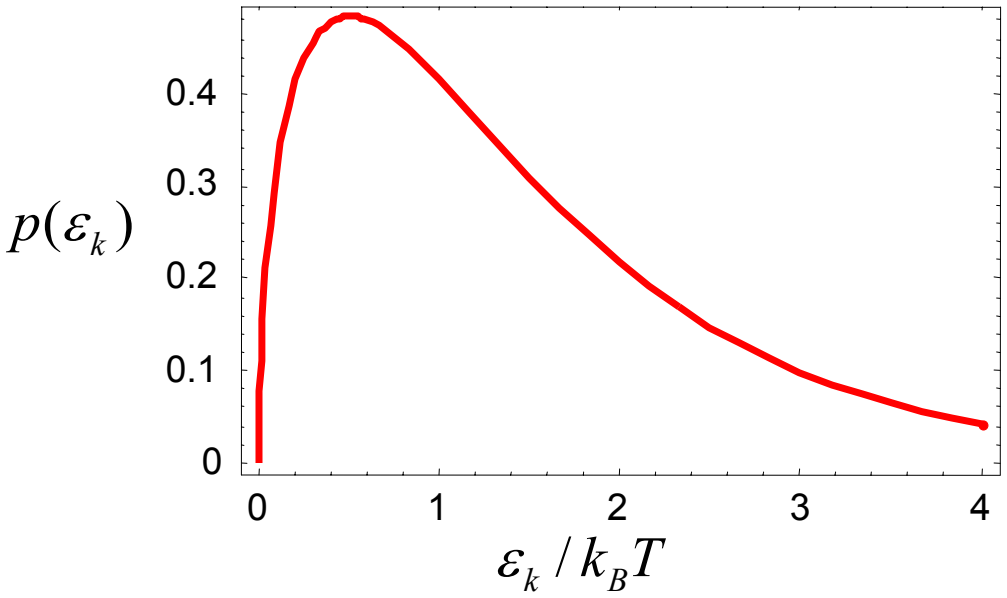




2.1.2 Maxwell-Boltzmann Distribution

Objective: Confirm that the kinetic energy distribution of the atoms in a system corresponds to the Maxwell distribution for a given temperature.

$$p(\varepsilon_k) = \frac{2}{\sqrt{\pi}} (k_B T)^{-3/2} \sqrt{\varepsilon_k} \exp\left(-\frac{\varepsilon_k}{k_B T} \right)$$



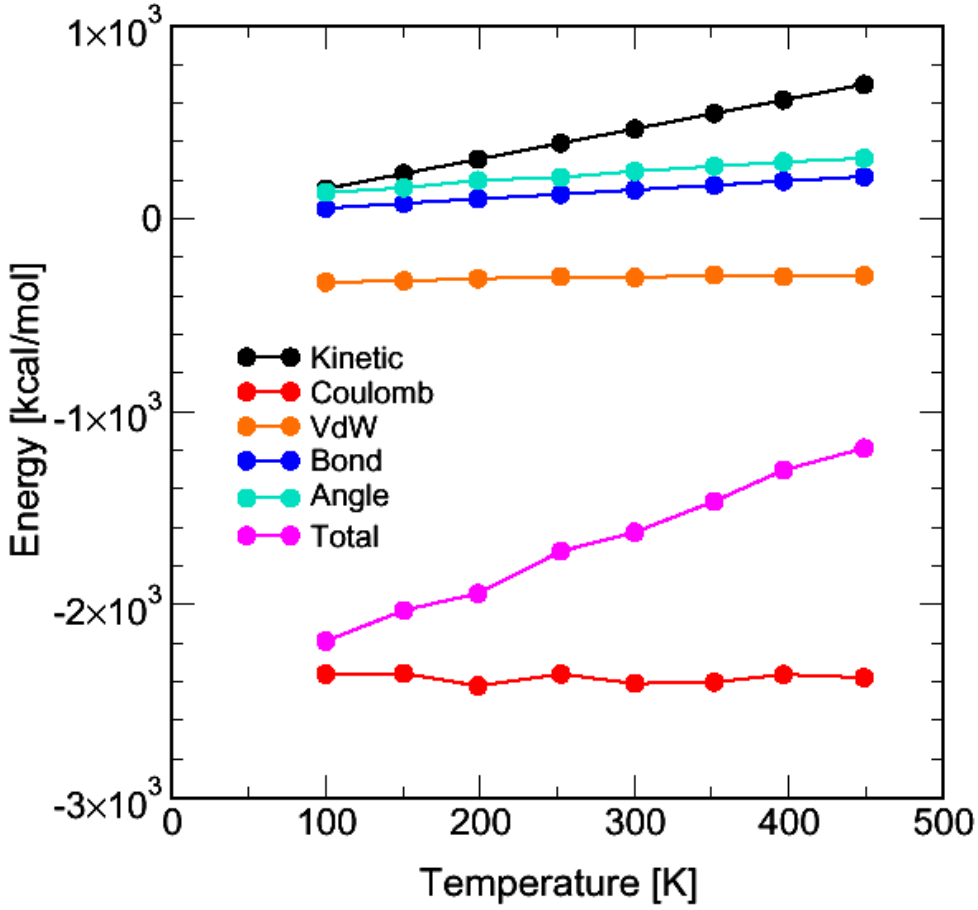
normalization condition:

$$\int_0^\infty d\varepsilon \, p(\varepsilon) = 1$$

2.1.3 Energies

Objective: Plot the various energies (kinetic and the different internal energies) as a function of temperature.

sample:



2.1.4 Temperature Fluctuations

Temperature time series: $T(t) = \frac{2}{3k_B}\langle K(t)\rangle = \frac{1}{3Nk_B}\sum_{i=1}^{N} m_i v_i^2(t)$

$$= \sum_{i=1}^{N} X_i(t)\,, \qquad X_i(t) \equiv \frac{m_i v_i^2(t)}{3Nk_B} = \frac{2\varepsilon_i}{3Nk_B}$$

According to the central limit theorem:

$$\langle T\rangle = N\langle X\rangle = \frac{2}{3k_B}\left\langle \frac{m_i v_i^2}{2}\right\rangle = \frac{2}{3k_B}\frac{3}{2}k_B T_0 = T_0$$

thermodynamic temperature

$$\sigma_0^2 = \langle X^2\rangle - \langle X\rangle^2 = \frac{2T_0^2}{3N^2} \Rightarrow \sigma^2 = \sigma_0^2 / N = \frac{2T_0^2}{3N}$$

$$\boxed{p(x) = \left(\frac{4\pi T_0^2}{3N}\right)^{-1/2} \exp\left[-\frac{3\left(T - T_0\right)^2}{4T_0^2}\right]}$$

Analysis of MD Data

1. Structural properties
2. Equilibrium properties
3. **Non-equilibrium properties**

Can be studied via both **equilibrium** and/or **non-equilibrium** MD simulations

Time Correlation Functions

$$C_{AB}(t-t') = \underbrace{\langle A(t)\,B(t')\rangle = \langle A(t-t')\,B(0)\rangle}_{\text{since } \rho_{eq} \text{ is } t \text{ independent !}}$$

$A \neq B$ cross-
$A = B$ auto-
} correlation function

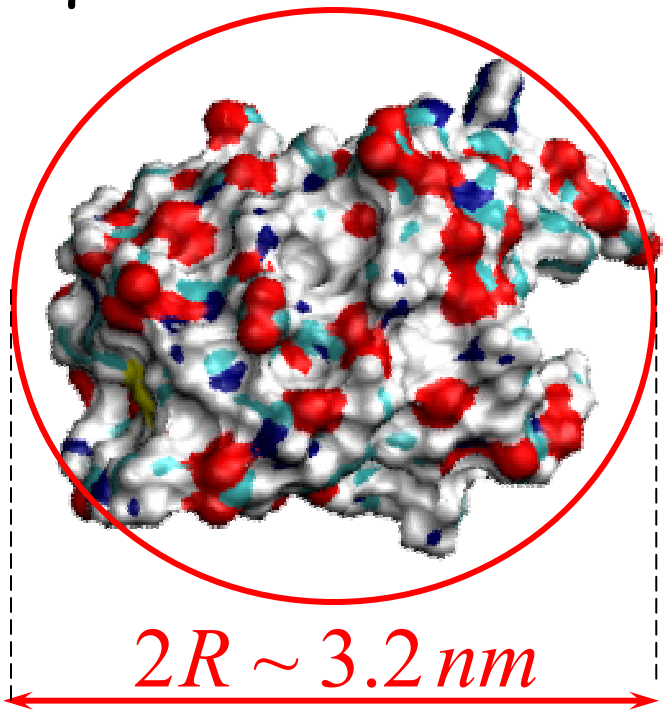
Correlation time: $\tau_c = \int_0^\infty dt\, C_{AA}(t)/C_{AA}(0)$

Estimates how long the "memory" of the system lasts

In many cases (but not always): $C(t) = C(0)\exp(-t/\tau_c)$

Free Diffusion (Brownian Motion) of Proteins

- in living organisms proteins exist and function in a viscous environment, subject to stochastic (random) thermal forces
- the motion of a globular protein in a viscous aqueous solution is diffusive

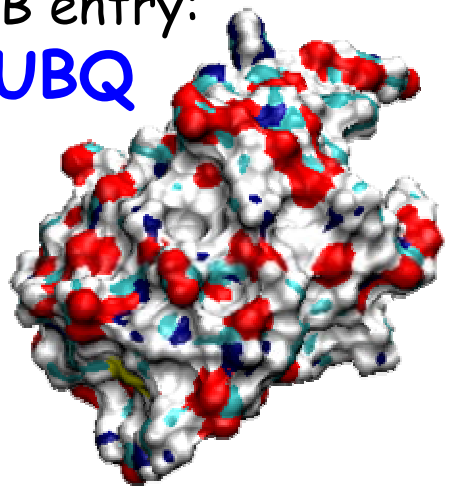


- e.g., *ubiquitin* can be modeled as a spherical particle of radius R~1.6nm and mass $M=6.4kDa=1.1\times10^{-23}$ kg

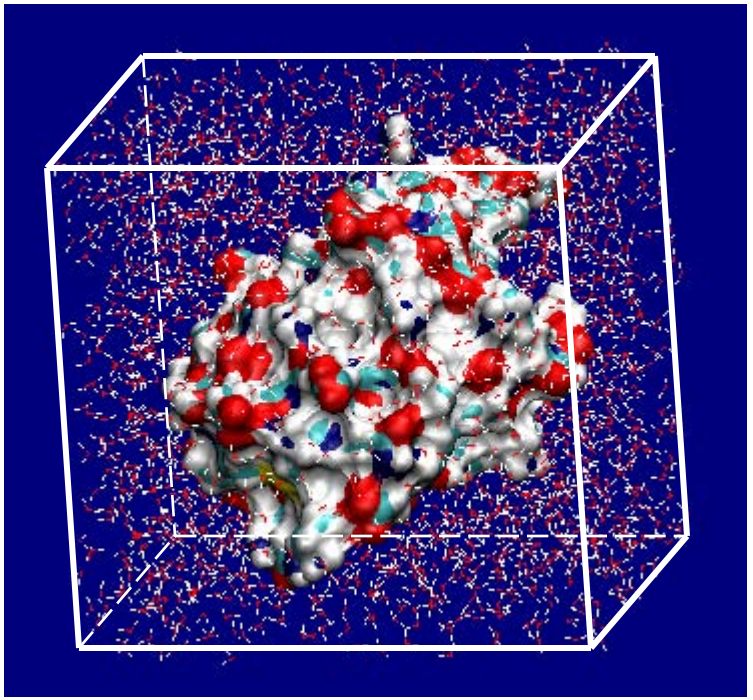
Diffusion can be Studied by MD Simulations!

ubiquitin in water

PDB entry:
1UBQ



solvate



total # of atoms: 7051 = 1231 (protein) + 5820 (water)

simulation conditions: NpT ensemble (T=310K, p=1atm), periodic BC, full electrostatics, time-step 2fs (SHAKE)

simulation output: Cartesian coordinates and velocities of all atoms saved at every other time-step (10,000 frames = 40 ps) in separate DCD files

Goal: calculate D and τ

by fitting the theoretically calculated center of mass (COM) velocity autocorrelation function to the one obtained from the simulation

- **theory:**

$$C_{vv}(t) = \langle v(t)\, v(0) \rangle = \langle v_0^2 \rangle e^{-t/\tau}$$

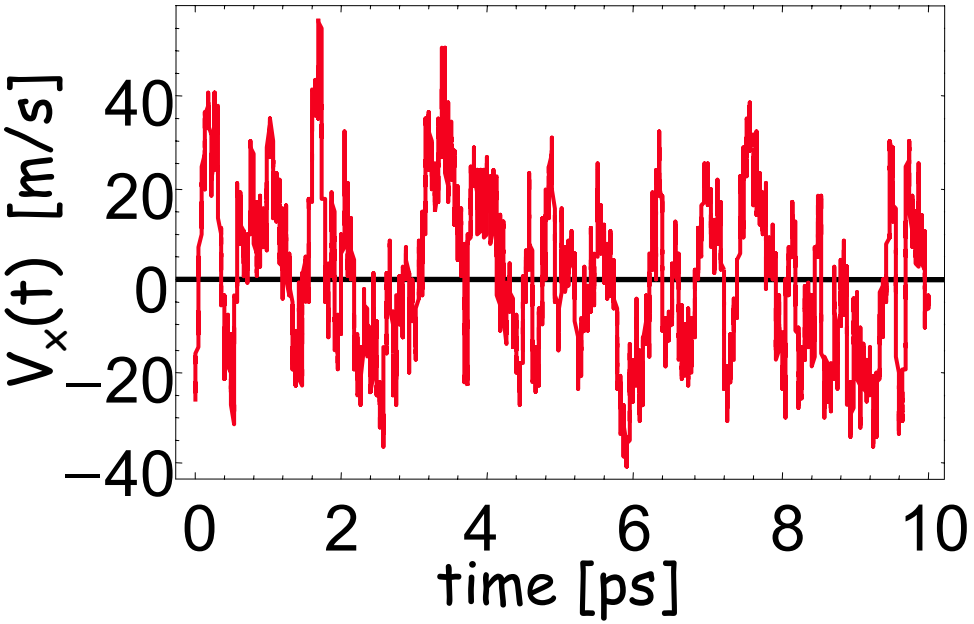
$$\langle v_0^2 \rangle = \frac{k_B T}{M} = \frac{D}{\tau}$$ (equipartition theorem)

- **simulation:** consider only the x-component $(v_x \to v)$
  replace ensemble average by time average

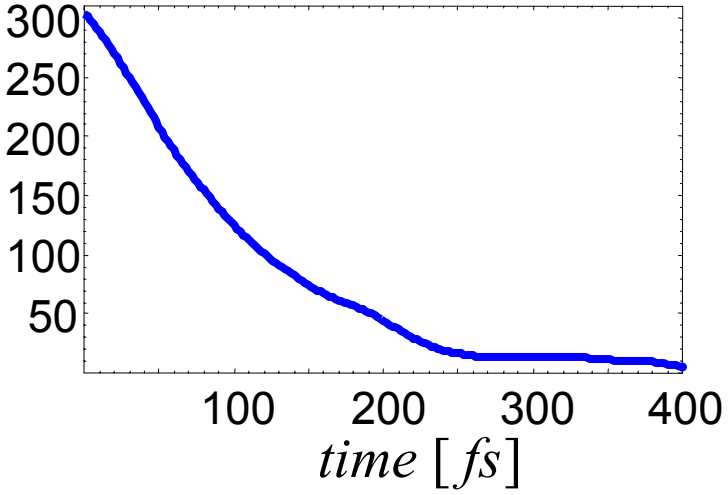
$$C_{vv}(t) \approx C_i = \frac{1}{N-i} \sum_{n=1}^{N-i} v_{n+i}\, v_n$$

$t \equiv t_i = i\Delta t$, $v_n = v(t_n)$, $N = \#$ of frames in vel.DCD

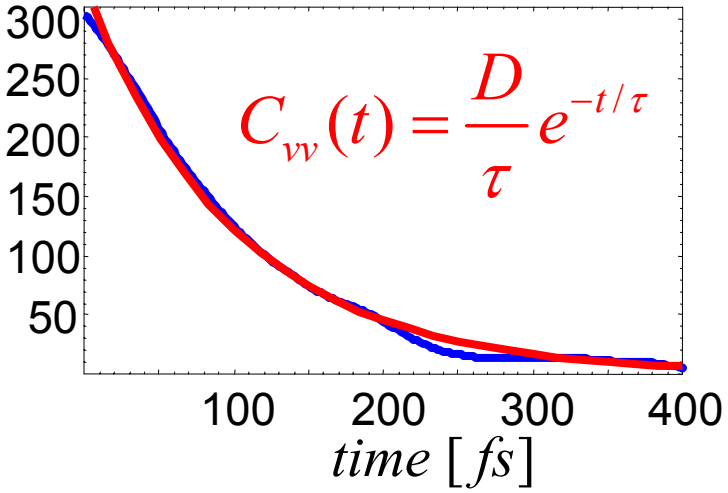
Velocity Autocorrelation Function



$C_{vv}(t)$



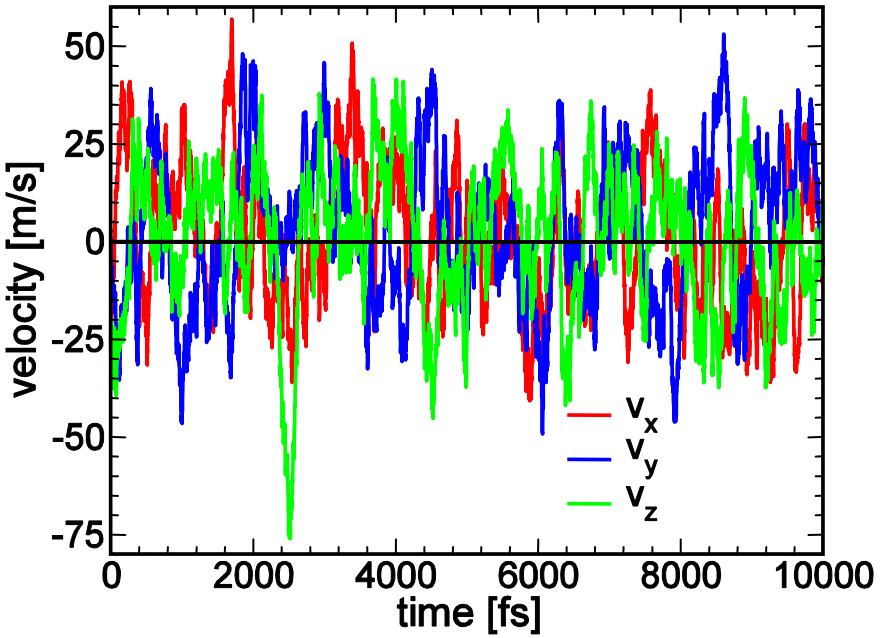
Fit

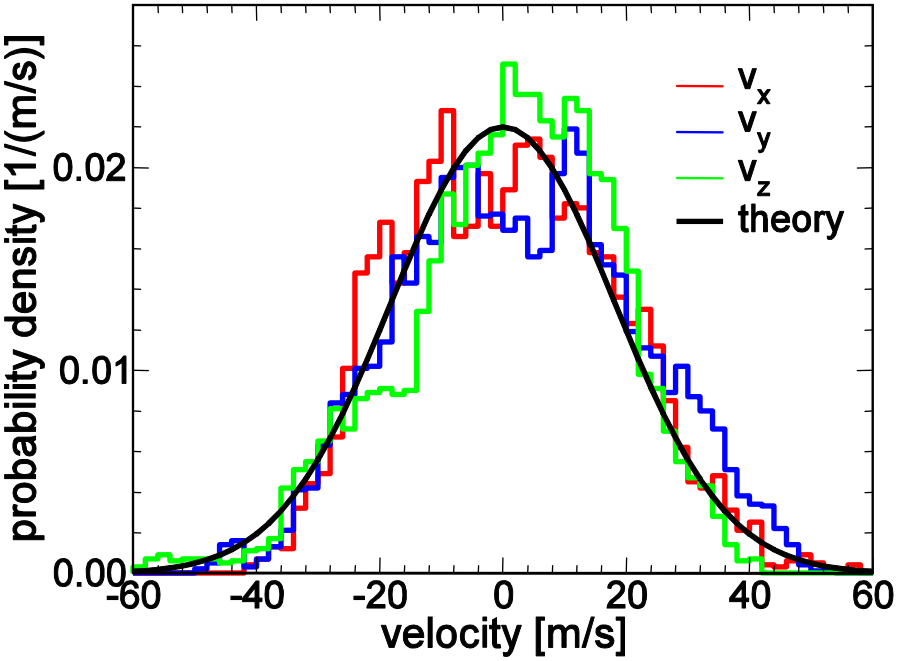


$$\tau \approx 0.1\ ps$$

$$D = k_B T/\gamma = \langle v_x^2 \rangle \tau \approx 3.3 \times 10^{-11}\ m^2 s^{-1}$$

Probability distribution of $v_{x,y,z}$



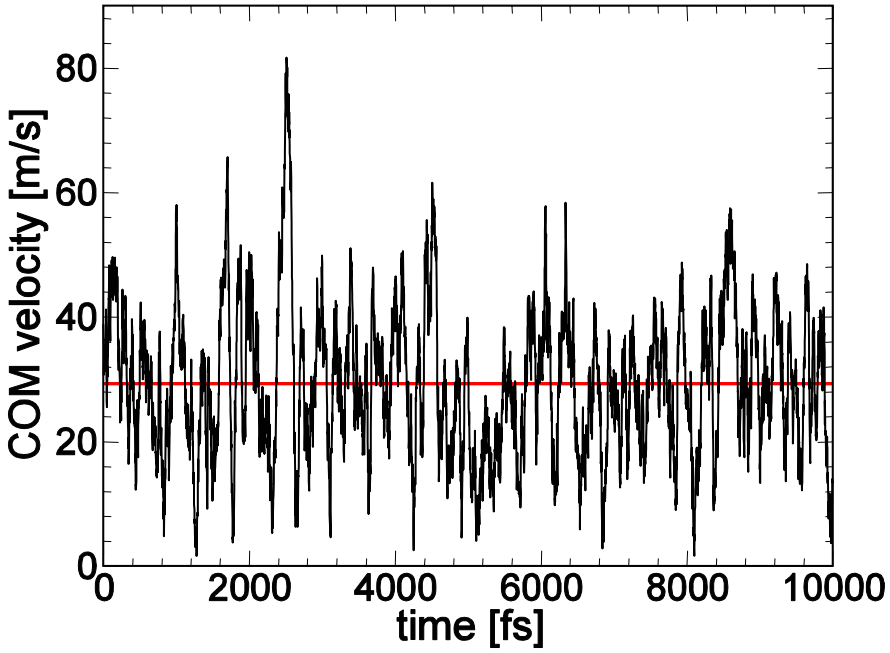


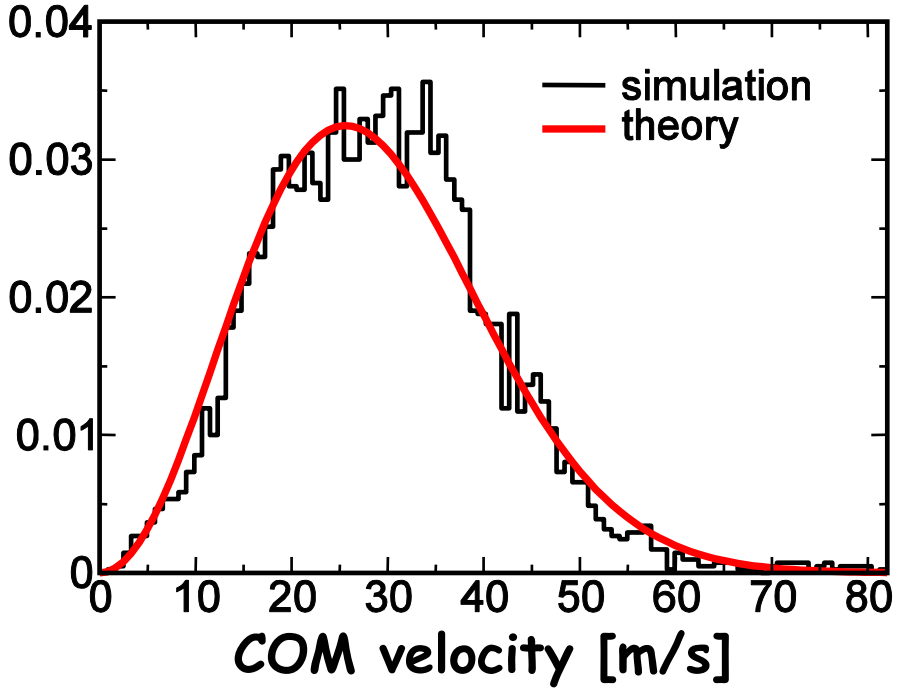
$$p(v) = \left(2\pi\left\langle v^2\right\rangle\right)^{-1/2} \exp\left(-v^2/2\left\langle v^2\right\rangle\right)$$

$$= \sqrt{\tau/2\pi D}\,\exp\left(-\tau v^2/2D\right)$$

with $v \equiv v_{x,y,z}$

Maxwell distribution of v_{COM}





$$P(v)dv = p(v_x)p(v_y)p(v_z)dv_x dv_y dv_z$$

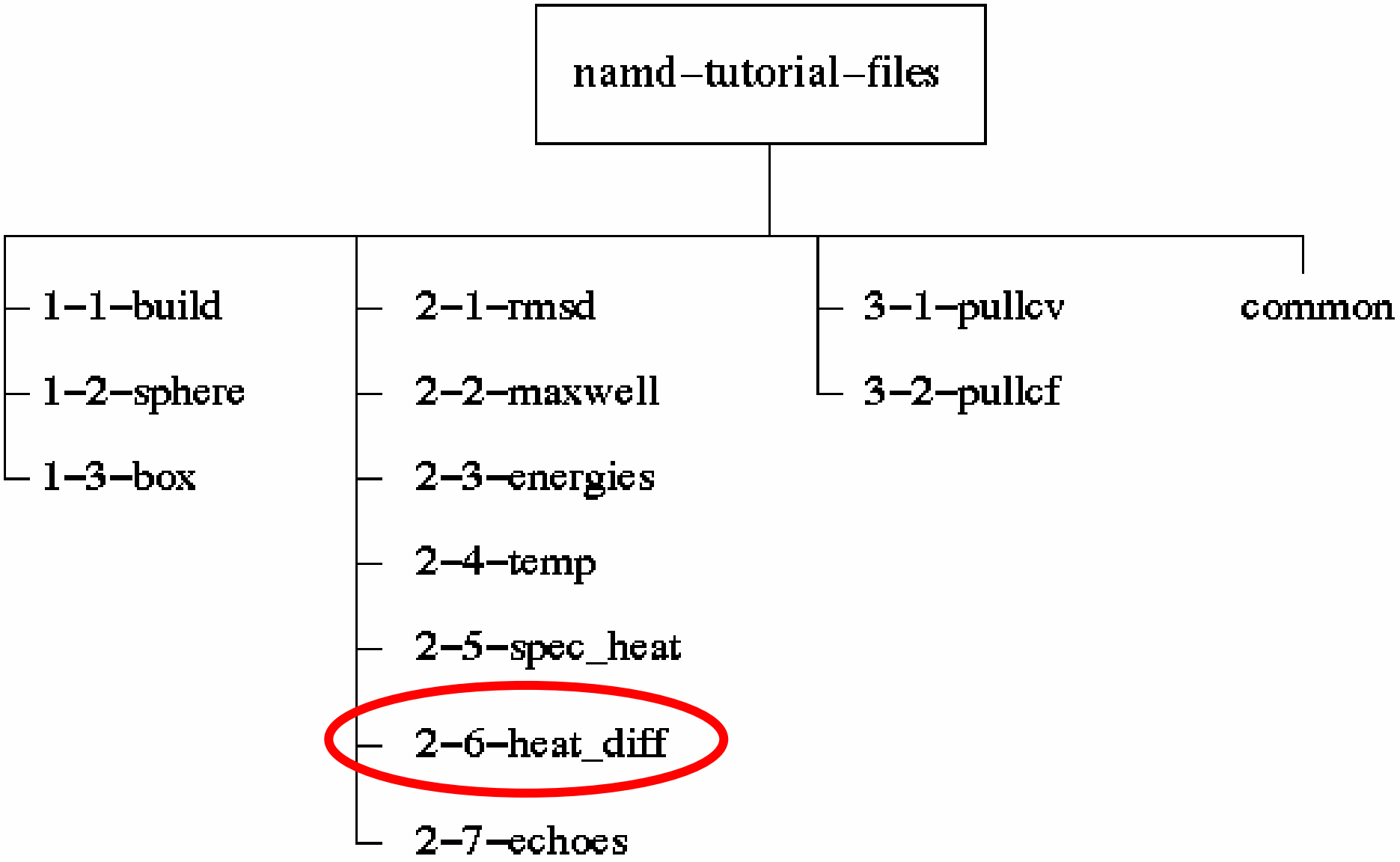
$$= \left(\tau/2\pi D\right)^{-3/2} \exp\left(-\tau v^2/2D\right) 4\pi v^2 dv$$

$$= \sqrt{\frac{2}{\pi}} \left(\frac{M}{k_B T}\right)^{-3/2} v^2 \exp\left(-\frac{Mv^2}{2k_B T}\right) dv$$

NAMD Tutorial (Part 2)

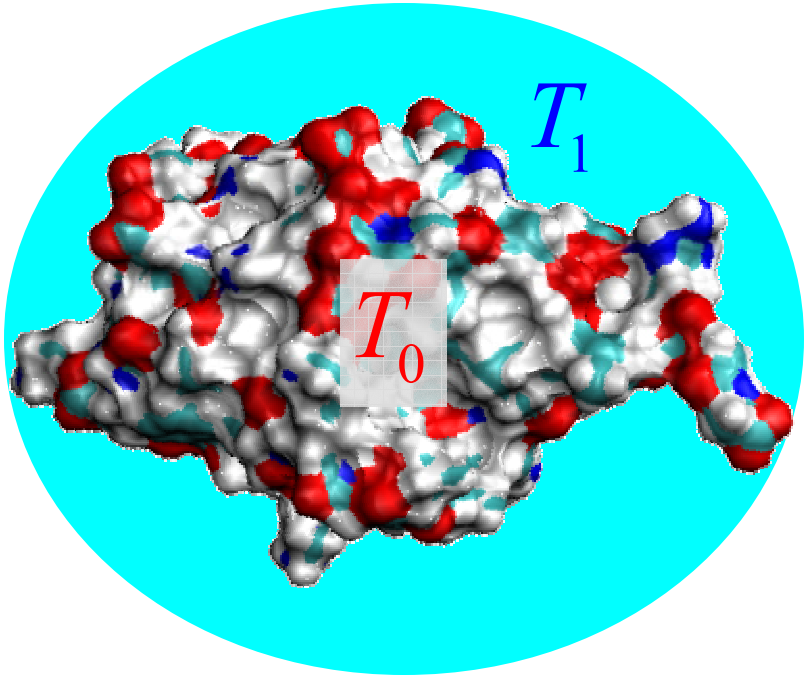
- **2 Analysis**
  - 2.1 Equilibrium
    - 2.1.1 RMSD for individual residues
    - 2.1.2 Maxwell-Boltzmann Distribution
    - 2.1.3 Energies
    - 2.1.4 Temperature distribution
    - 2.1.5 Specific Heat
  - **2.2 Non-equilibrium properties of protein**
    - **2.2.1 Heat Diffusion**
    - 2.2.2 Temperature echoes

Organization of NAMD Tutorial Files



Simulated Cooling of Ubiquitin

- Proteins function in a narrow (physiological) temperature range. What happens to them when the temperature of their surrounding changes significantly (temperature gradient) ?
- Can the heating/cooling process of a protein be simulated by molecular dynamics ? If yes, then how?



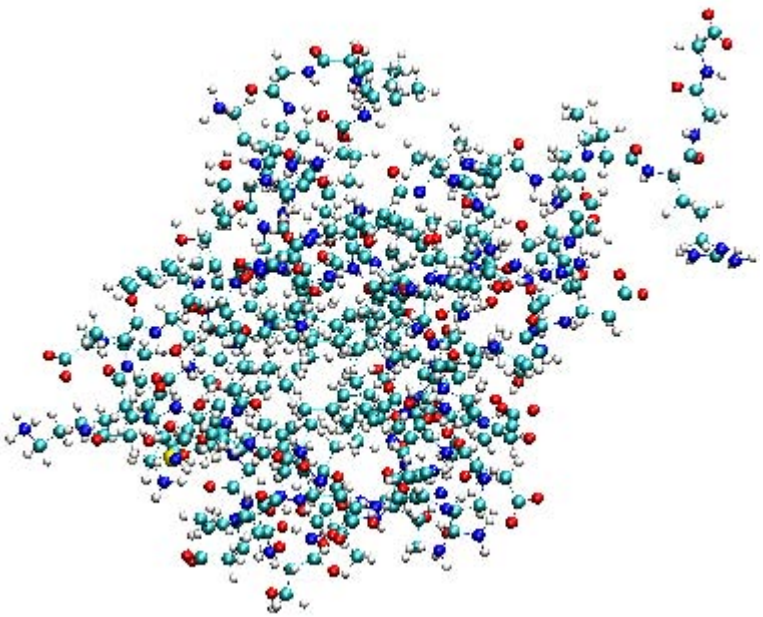
- What can we learn from the simulated cooling/heating of a protein ?

Nonequilibrium (Transport) Properties

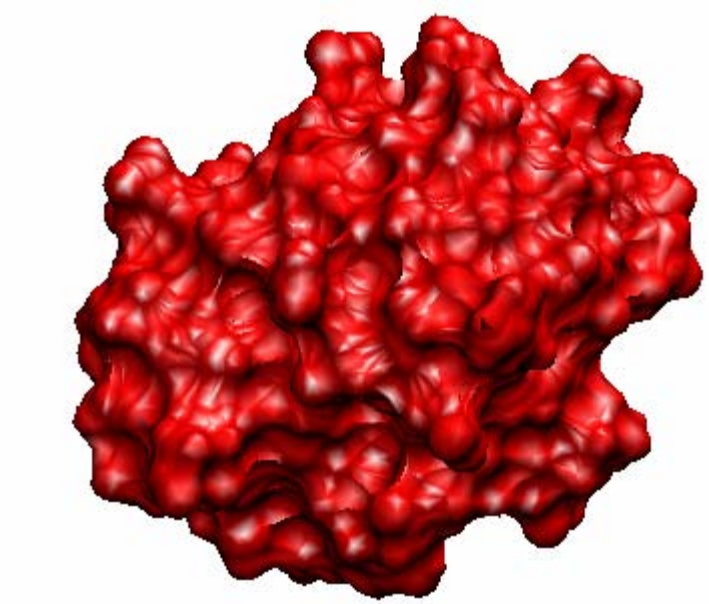
- macromolecular properties of proteins, which are related to their biological functions, often can be probed by studying the response of the system to an external perturbation, such as *thermal gradient*
- "small" perturbations are described by linear response theory (LRT), which relates transport (nonequilibrium) to thermodynamic (equilibrium) properties
- on a "mesoscopic" scale a globular protein can be regarded as a continuous medium $\Rightarrow$ within LRT, the local temperature distribution $T(r,t)$ in the protein is governed by the heat diffusion (conduction) equation

$$\frac{\partial T(\mathbf{r},t)}{\partial t} = D\nabla^2 T(\mathbf{r},t)$$

Atomic vs Mesoscopic



- each atom is treated individually
- length scale ~ 0.1 Å
- time scale ~ 1 fs



- one partitions the protein in small volume elements and average over the contained atoms
- length scale ≥ 10 Å = 1nm
- time scale ≥ 1 ps

Heat Conduction Equation

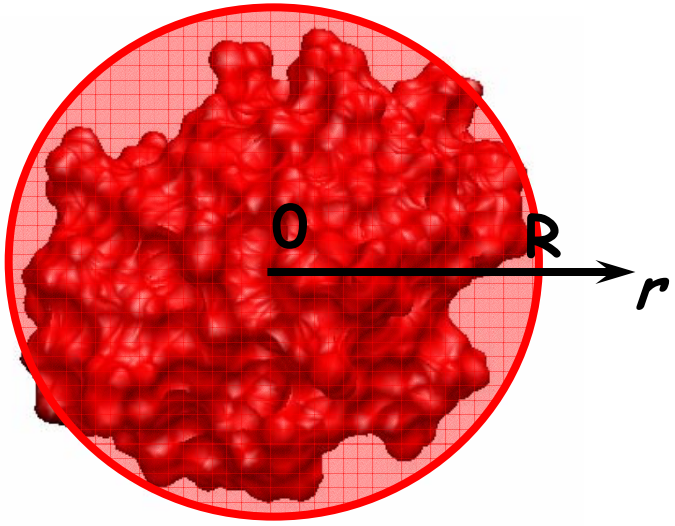
$$\frac{\partial T(\mathbf{r},t)}{\partial t} = D\nabla^2 T(\mathbf{r},t)$$

mass density

thermal diffusion coefficient → $D = K/\rho c$

specific heat

thermal conductivity



- approximate the protein with a homogeneous sphere of radius R~20 Å
- calculate *T(r,t)* assuming initial and boundary conditions:

$$T(r,0) = T_0 \quad for \quad r < R$$

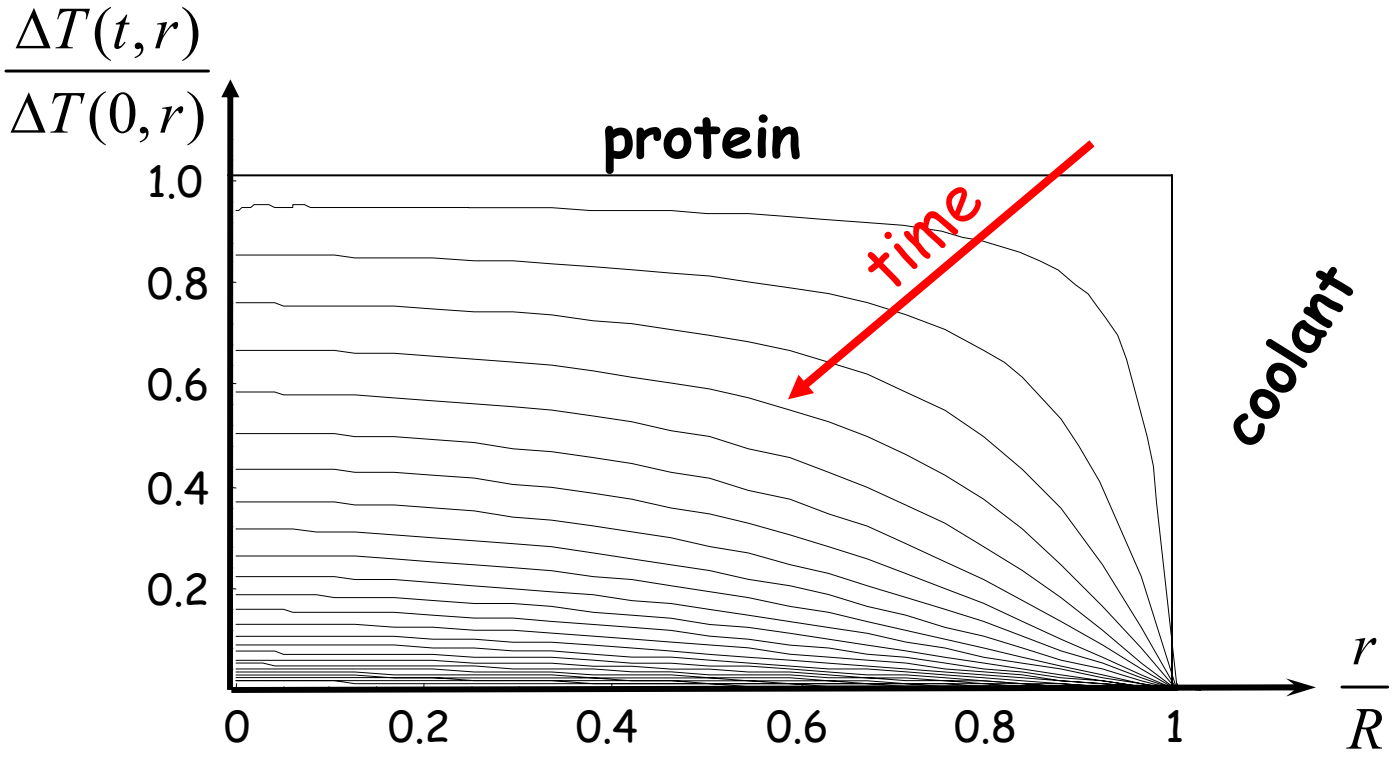
$$T(R,t) = T_{bath}$$

Solution of the Heat Equation

$$\langle \Delta T(t) \rangle = \langle \Delta T(0) \rangle \times \frac{6}{\pi^2} \sum_{n=1}^{\infty} \frac{1}{n^2} \exp\left(-n^2 \pi^2 t / \tau_0\right)$$

where $\langle \Delta T(t) \rangle \equiv \langle T(t) \rangle - T_{bath}$, $\tau_0 = R^2/D$

averaged over the entire protein!



How to simulate cooling ?

- In laboratory, the protein is immersed in a coolant and the temperature decreases from the surface to the center
- Cooling methods in MD simulations:

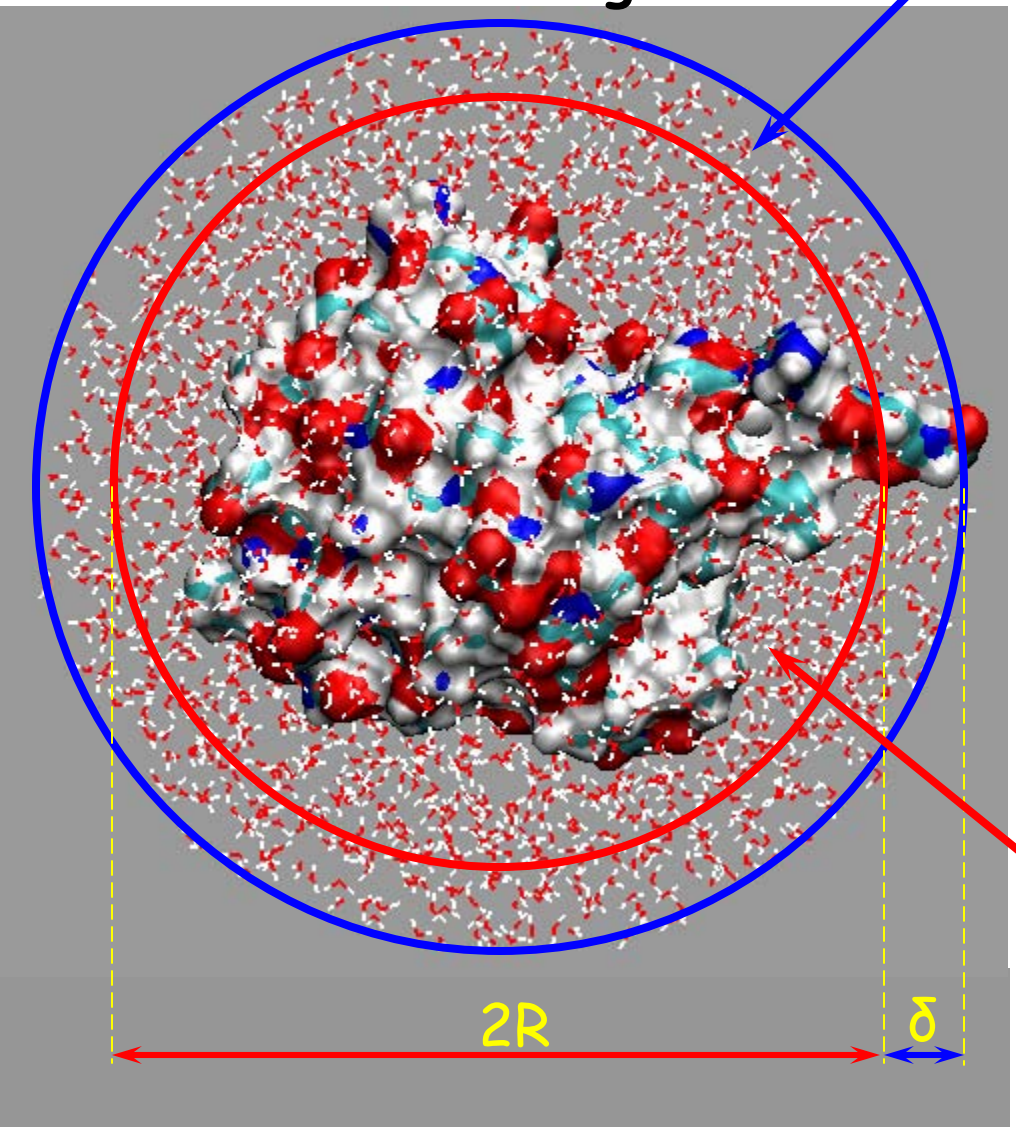
  1. **Stochastic boundary method**
  2. Velocity rescaling (rapid cooling, biased velocity autocorrelation)

$$\langle T(t) \rangle_{sim} = \frac{\sum_{i=1}^{N_d} m_i v_i^2}{N_d k_B} \quad \Rightarrow \quad v_i' = v_i \sqrt{\frac{T_{new}}{T_{old}}}$$

  3. Random reassignment of atomic velocities according to Maxwell's distribution for desired temperature (velocity autocorrelation completely lost)

Stochastic Boundary Method

Heat transfer through mechanical coupling between atoms in the two regions



coolant layer of atoms

motion of atoms is subject to stochastic Langevin dynamics

$$m\ddot{\boldsymbol{r}} = \boldsymbol{F}_{FF} + \boldsymbol{F}_H + \boldsymbol{F}_f + \boldsymbol{F}_L$$

$\boldsymbol{F}_{FF} \rightarrow$ force field

$\boldsymbol{F}_H \rightarrow$ harmonic restrain

$\boldsymbol{F}_f \rightarrow$ friction

$\boldsymbol{F}_L \rightarrow$ Langevin force

atoms in the inner region follow Newtonian dynamics

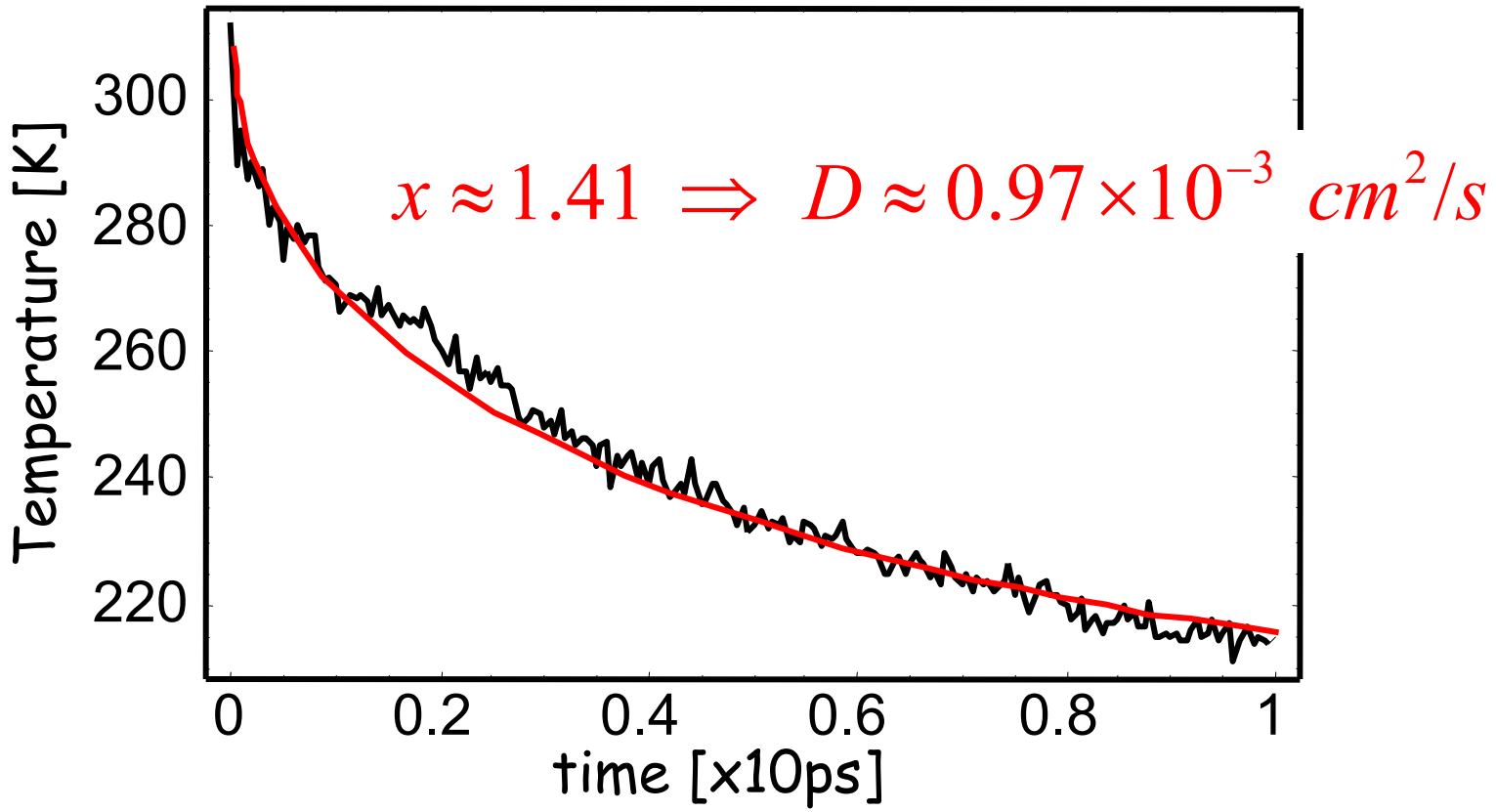
$$m\ddot{\boldsymbol{r}} = \boldsymbol{F}_{FF}$$

Determine D by Fitting the Data

$$\frac{\langle \Delta T(t) \rangle}{\langle \Delta T(0) \rangle} = \frac{6}{\pi^2} \sum_{n=1}^{\infty} \frac{1}{n^2} \exp\left(-n^2 x\, t/t_0\right)$$

(use the first 100 terms in the series !)

$$t_0 = 10\ ps, \quad x = \textit{fitting parameter}, \quad D = \left(\frac{R}{\pi}\right)^2 \frac{x}{t_0}$$



Thermal Conductivity of UBQ

$$K = D\rho c$$

$$C_V = \langle \delta E^2 \rangle / k_B T^2 = \left(\langle E^2 \rangle - \langle E \rangle^2 \right) / k_B T^2$$

$$D \approx 0.97 \times 10^{-3} \ cm^2/s$$

$$\rho \approx 1 \times 10^3 \ kg/m^3$$

NAMD Tutorial (Part 2)

- **2 Analysis**
  - 2.1 Equilibrium
    - 2.1.1 RMSD for individual residues
    - 2.1.2 Maxwell-Boltzmann Distribution
    - 2.1.3 Energies
    - 2.1.4 Temperature distribution
    - 2.1.5 Specific Heat
  - 2.2 Non-equilibrium properties of protein
    - 2.2.1 Heat Diffusion
    - **2.2.2 Temperature echoes**