Statistical Mechanics of Proteins

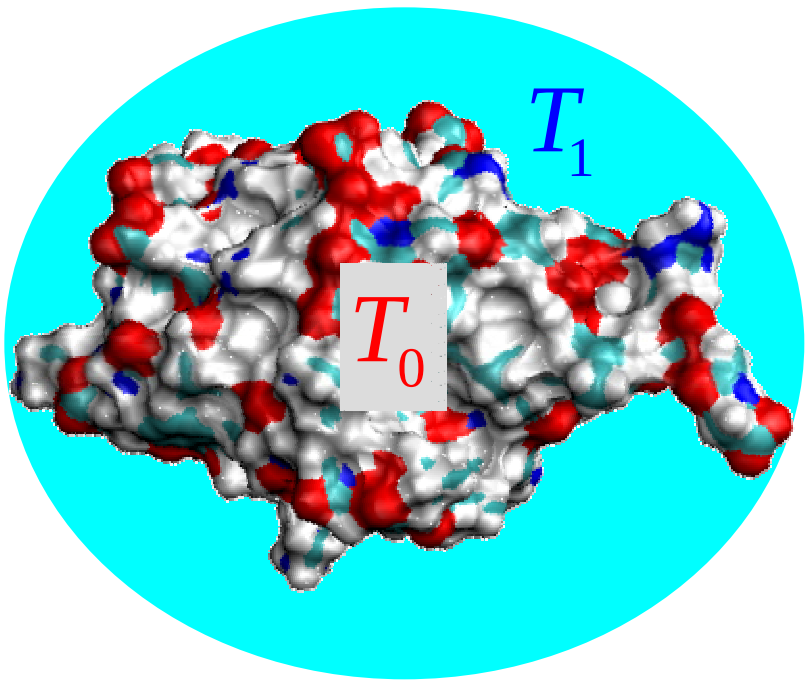
Ioan Kosztin

Department of Physics & Astronomy
University of Missouri - Columbia

- Equilibrium and non-equilibrium properties of proteins
  - Free diffusion of proteins
- Coherent motion in proteins: temperature echoes
- **Simulated cooling of proteins**

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?



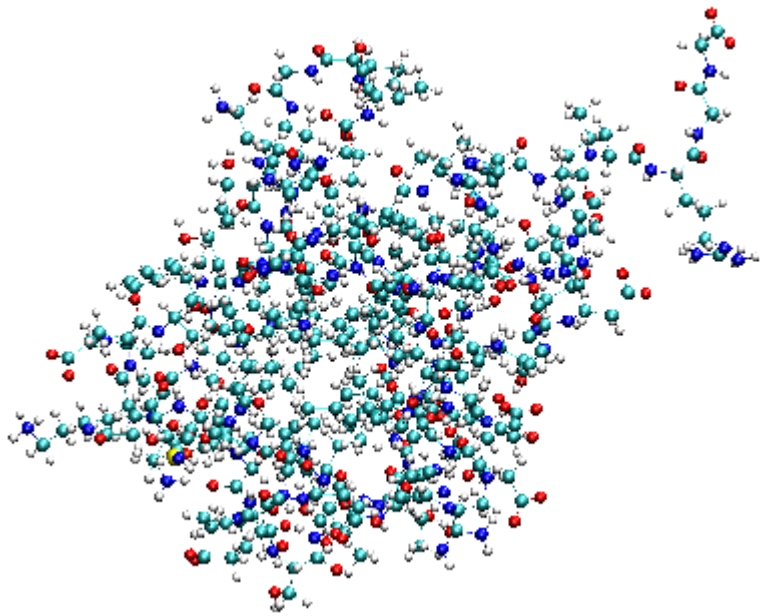
- What can we learn from the simulated cooling/heating of a protein ?

Nonequilibrium (Transport) Properties

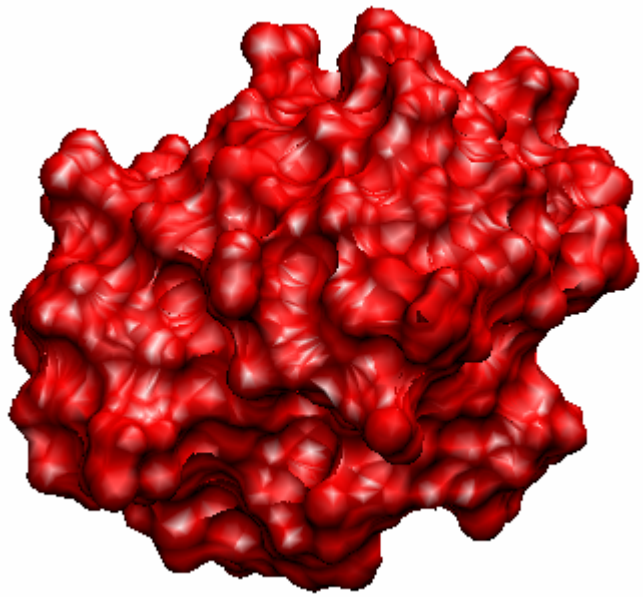
- macromolecular properties of proteins, which are related to their biological functions, often can be probed by studying the response of the system to an external perturbation, such as *thermal gradient*
- "small" perturbations are described by linear response theory (LRT), which relates transport (nonequilibrium) 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

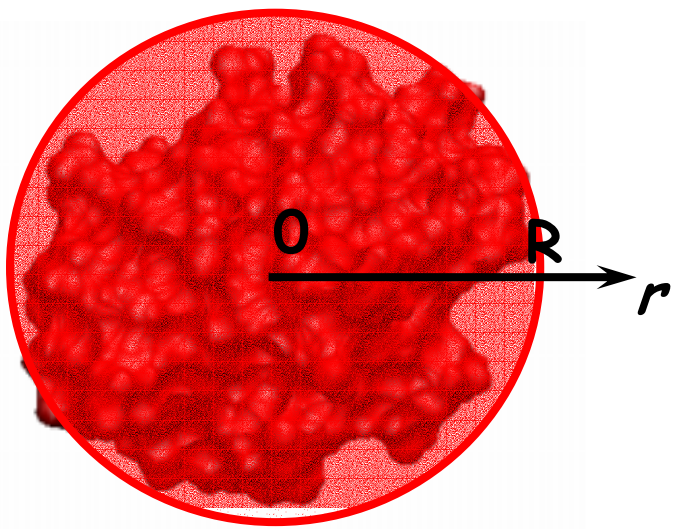
Heat Conduction Equation

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

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:

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

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

Thermal diffusion coefficient D=?

D is a phenomenological transport coefficient which needs to be calculated either from a microscopic (atomistic) theory, or derived from (**computer**) **experiment**

"Back of the envelope" estimate:

$$\frac{\partial T(\mathbf{r},t)}{\partial t} = D\nabla^2 T(\mathbf{r},t) \Rightarrow \frac{\Delta T}{\tau_0} \sim D\frac{\Delta T}{R^2} \Rightarrow D \sim R^2/\tau_0$$

$$R \sim 10\,\text{Å} = 10^{-9}\ m, \quad \tau_0 \sim 10\ ps = 10^{-11}\ s \quad ???$$

$$\Rightarrow D \sim R^2/\tau_0 \sim 10^{-7}\ m^2/s = 10^{-3}\ cm^2/s$$

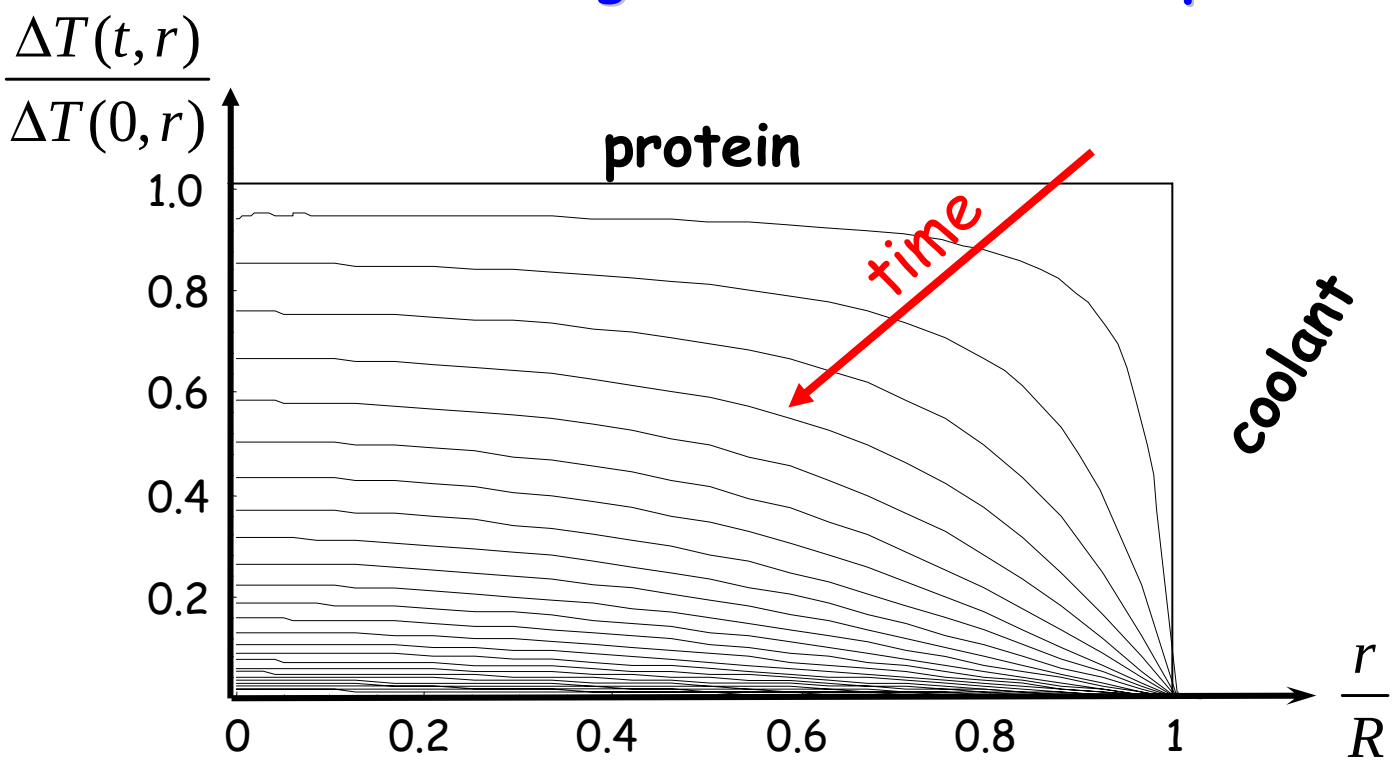
From MD simulation !

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)

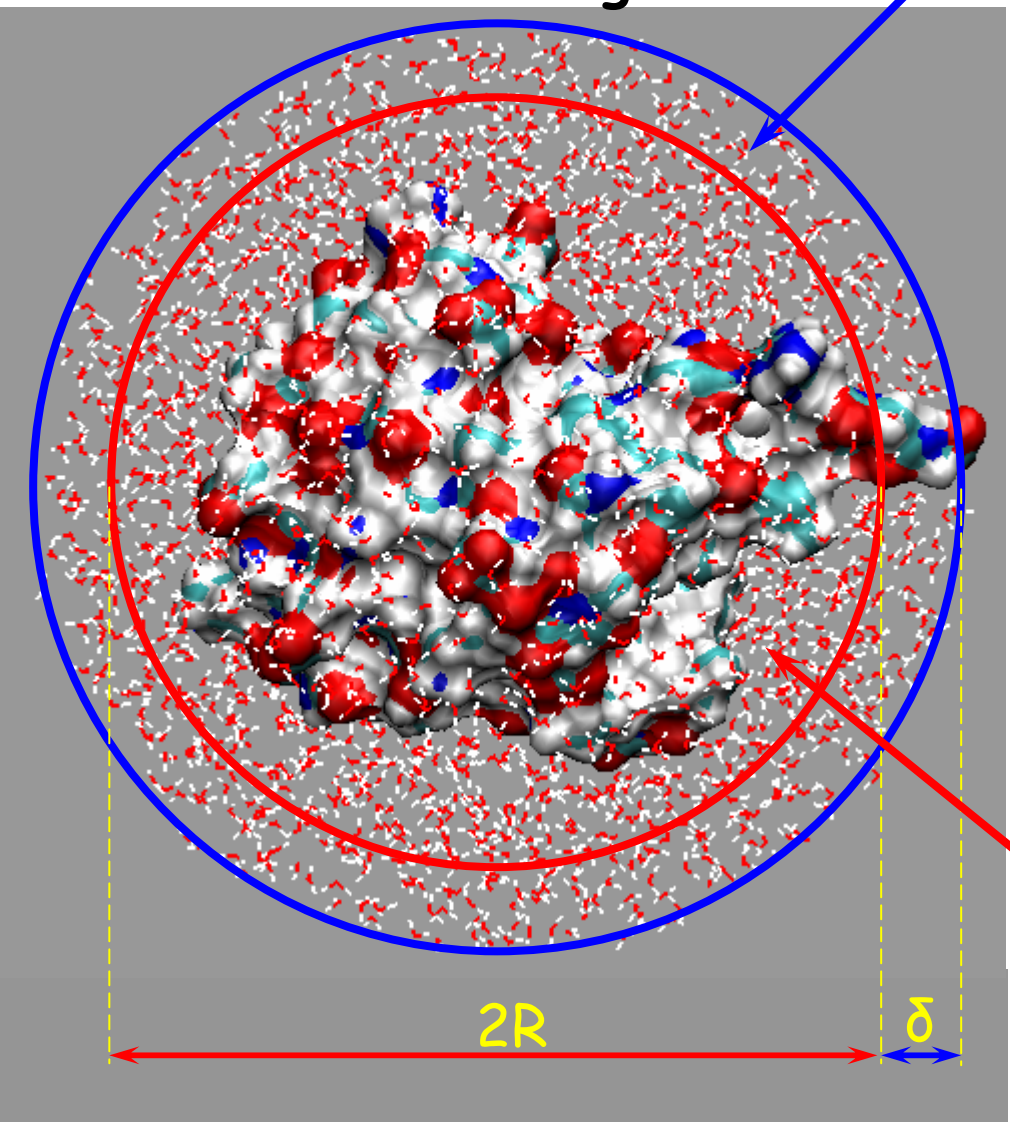
$$\langle T(t) \rangle_{sim} = \frac{\sum_{i=1}^{N_d} m_i v_i^2}{N_d k_B} \Rightarrow v_i' = v_i \sqrt{\frac{T_{new}}{T_{old}}}$$

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

NAMD User Guide: Temperature Control

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

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

2-6-heat_diff: Simulated Cooling of UBQ

Start from a pre-equilibrated system of UBQ in a water sphere of radius 26Å
mol load psf ubq_ws.psf namdbin ubq_ws_eq.restart.coor

Create the a coolant layer of atoms of width 4Å

Select all atoms in the system:
set selALL [atomselect top all]

Find the center of the system:
set center [measure center $selALL weight mass]

Find X, Y and Z coondinates of the system's center:
foreach {xmass ymass zmass} $center { break }

2-6-heat_diff: Simulated Cooling of UBQ

Select atoms in the outer layer:
set shellSel [atomselect top "not (sqr(x-$xmass) + sqr(y-$ymass) + sqr(z-$zmass) <= sqr(22)) "]

Set beta parameters of the atoms in this selection to 1.00:
$shellSel set beta 1.00

Select the entire system again:
set selALL [atomselect top all]

Create the pdb file that marks the atoms in the outer layer by 1.00 in the beta column:
$selALL writepdb ubq_shell.pdb

NAMD configuration file: ubq_cooling.conf

```
# Spherical boundary conditions
# Note: Do not set other bondary conditions and PME if spherical
# boundaries are used
if {1} {
sphericalBC          on
sphericalBCcenter    30.30817, 28.80499, 15.35399
sphericalBCr1        26.0
sphericalBCk1        10
sphericalBCexp1      2
}

# this is to constrain atoms
if {1} {
constraints                    On
consref                        ubq_shell.pdb
consexp                        2
conskfile                      ubq_shell.pdb
conskcol                       B
}
```

NAMD configuration file: ubq_cooling.conf

```
# this is to cool a water layer
if {1} {
tCouple                 on
tCoupleTemp             200
tCoupleFile             ubq_shell.pdb
tCoupleCol              B
}
```

RUN THE SIMULATION FOR 10 ps

(5000 steps; timestep = 2 ps)

The (kinetic) temperature *T(t)* is extracted from the simulation log (output) file; it can be plotted directly with

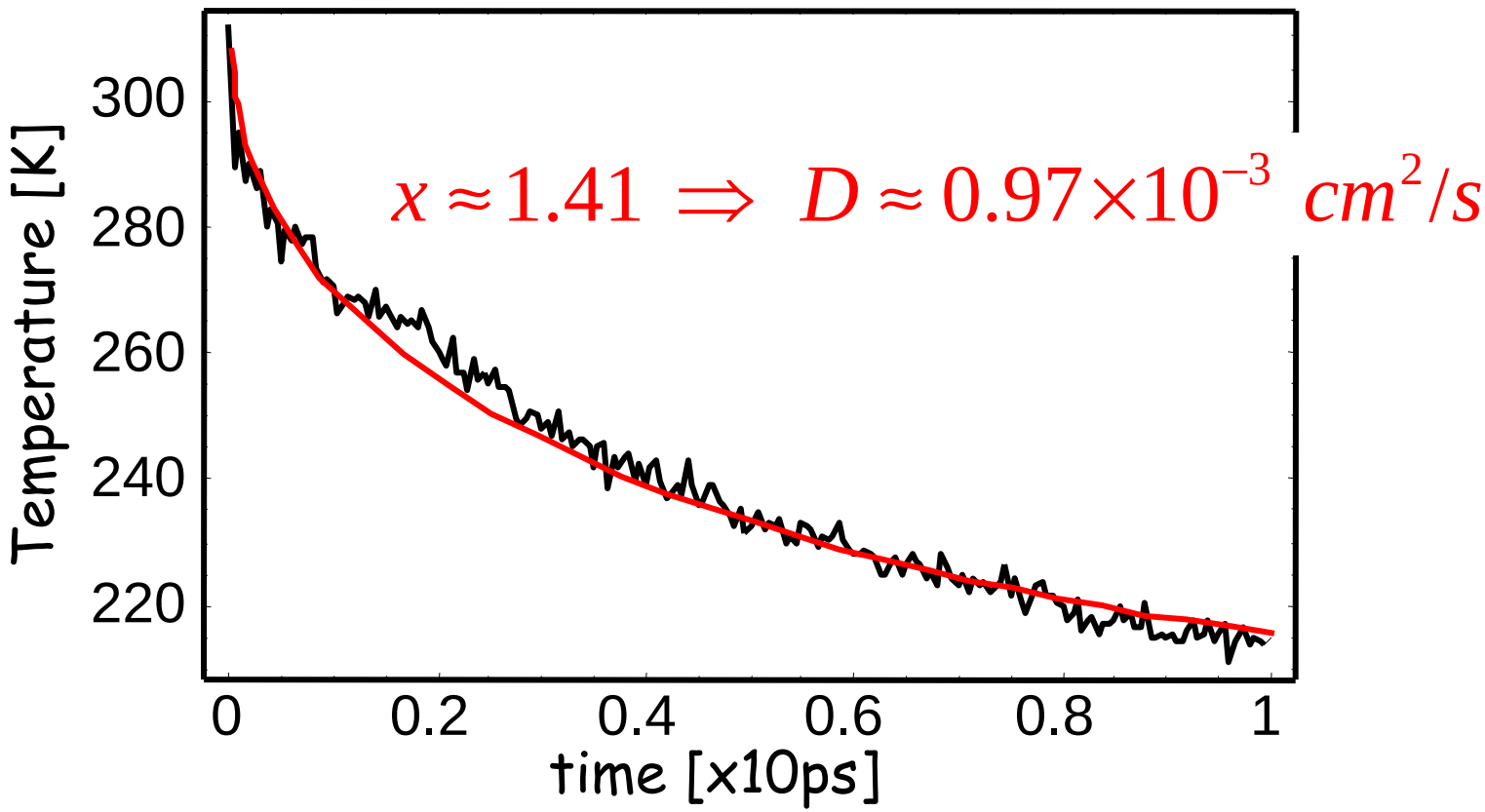
namdplot TEMP ubq_cooling.log

Is this procedure of getting T(t) correct ?

Determine D by Fitting the Data

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

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



$$x \approx 1.41 \Rightarrow D \approx 0.97 \times 10^{-3}\ cm^2/s$$

Thermal Conductivity of UBQ

$$K = D\rho c$$

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

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

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