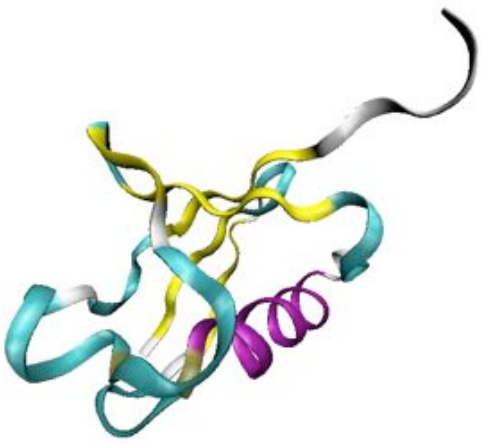
Statistical Mechanics of Proteins



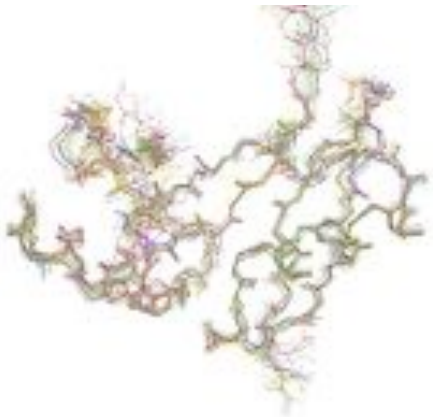
Ubiquitin

Equilibrium Properties of Proteins

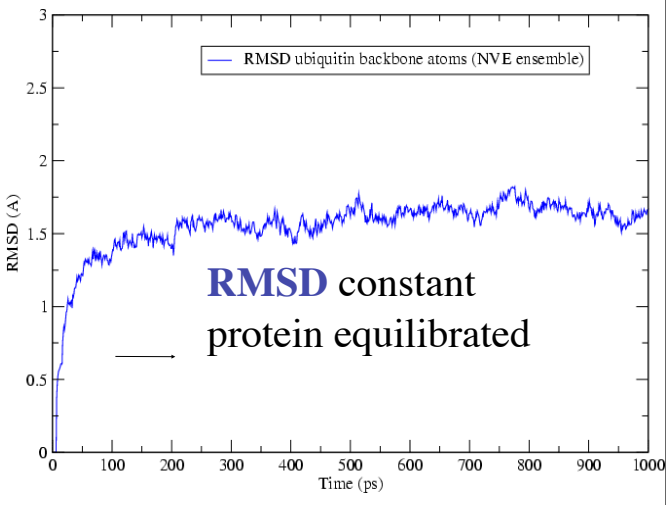
Ubiquitin

Root Mean Squared Deviation: measure for equilibration and protein flexibility

$$RMSD_\alpha = \sqrt{\frac{\sum_{j=1}^{N_t} \sum_{\alpha=1}^{N_\alpha} (\vec{r}_\alpha(t_j) - \langle \vec{r}_\alpha \rangle)^2}{N_\alpha}}$$



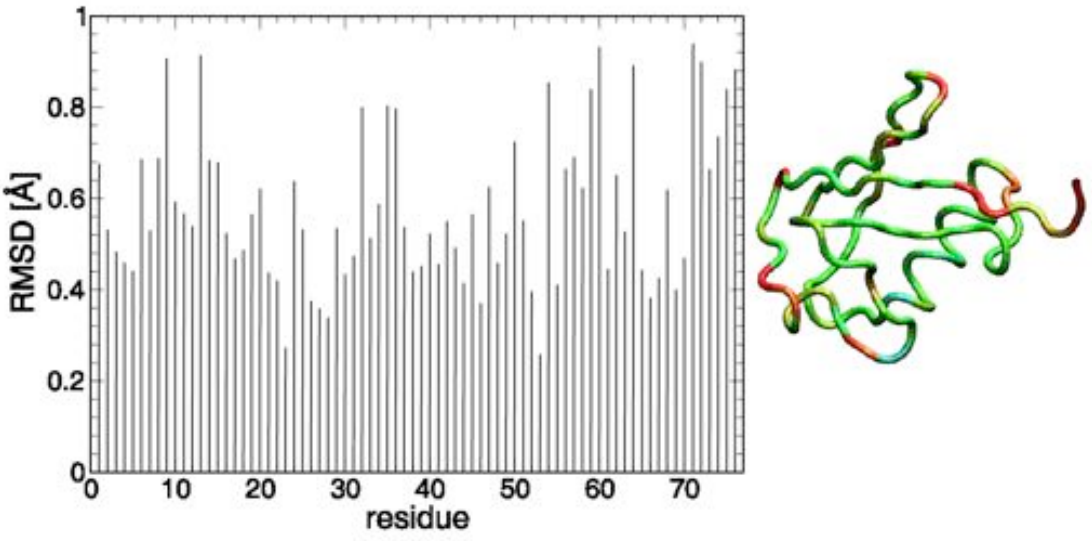
NMR structures
aligned together to see flexibility



RMSD constant
protein equilibrated

Thermal Motion of Ubiquitin from MD

RMSD values per residue



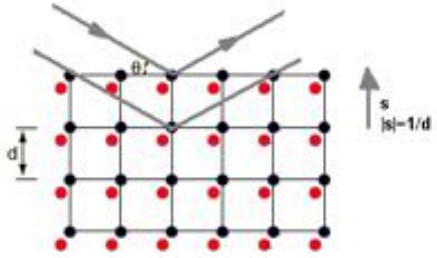
Protein sequence exhibits characteristic permanent flexibility!

Thermal Motion of Ubiquitin from MD

Temperature Dependence of Crystal Diffraction (Debye-Waller factor)

Bragg's law

$$2\,d\,\sin\theta \;=\; \lambda$$



due to contributin from each copy *j*

$$f_j \exp[-i\vec{s}\cdot\vec{r}_j]$$ structure factor f_j

But the atom carries out thermal vibrations around equilibrium position $\vec{x}_j$

$$\vec{r}_j(t) \;=\; \vec{x}_j \;+\; \vec{u}_j(t)$$

Accordingly:

$$\langle f_j \exp[-i\vec{s}\cdot\vec{r}_j]\rangle \;=\; f_j \exp[-i\vec{s}\cdot\vec{x}_j]\,\langle \exp[-i\vec{s}\cdot\vec{u}_j]\rangle$$

Thermal Motion of Ubiquitin from MD

Temperature Dependence of Crystal Diffraction (Debye-Waller factor)

One can expand:

$$\langle \exp[-i\vec{s}\cdot\vec{u}_j]\rangle \;=\; 1 \;-\; i\underbrace{\langle \vec{s}\cdot\vec{u}_j\rangle}_{=0} \;-\; \tfrac{1}{2}\langle(\vec{s}\cdot\vec{u}_j)^2\rangle \;+\ldots$$

Spatial average for harmonic oscillator: $\langle(\vec{s}\cdot\vec{u}_j)^2\rangle \;=\; \tfrac{1}{3}s^2\langle u_j^2\rangle$

One can carry out the expansion further and show all further terms vanish

$$\langle \exp[-i\vec{s}\cdot\vec{u}_j]\rangle \;=\; \exp\left[-\tfrac{1}{6}s^2\,\langle u_j^2\rangle\right]$$

Using for the thermal amplitude of the harmonic oscillator

$$\tfrac{1}{2}m\omega^2 u_j^2 \;=\; \tfrac{3}{2}k_BT$$

thermal "weakening" of struct. factor

one obtains

$$\langle f_j \exp\left[-i\vec{s}\cdot\vec{r}_j\right]\rangle \;=\; f_j\overbrace{\exp[-s^2k_BT/2m\omega^2]}^{\text{Debye-Waller factor}}\exp[-i\vec{s}\cdot\vec{x}_j]$$

Equilibrium Properties of Proteins

Energies: kinetic and potential

?

temperature dependence

$$\langle\sum_j \frac{1}{2}m_j v_j^2\rangle$$

Kinetic energy (quadratic)

$$U(\vec{R}) \;=\; \underbrace{\sum_{bonds} k_i^{bond}(r_i-r_0)^2}_{U_{bond}} + \underbrace{\sum_{angles} k_i^{angle}(\theta_i-\theta_0)^2}_{U_{angle}} +$$

$$\underbrace{\sum_{dihedrals} k_i^{dihe}[1+\cos(n_i\phi_i+\delta_i)]}_{U_{dihedral}} +$$

$$\underbrace{\sum_i\sum_{j\neq i} 4\epsilon_{ij}\left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12}-\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right] + \sum_i\sum_{j\neq i}\frac{q_iq_j}{\epsilon r_{ij}}}_{U_{nonbond}}$$

Potentialc energy (not all quadratic)

Equilibrium Properties of Proteins

Energies: kinetic and potential

?

temperature dependence

$$\langle \sum_j \frac{1}{2} m_j v_j^2 \rangle = \frac{3}{2} N k_B T$$

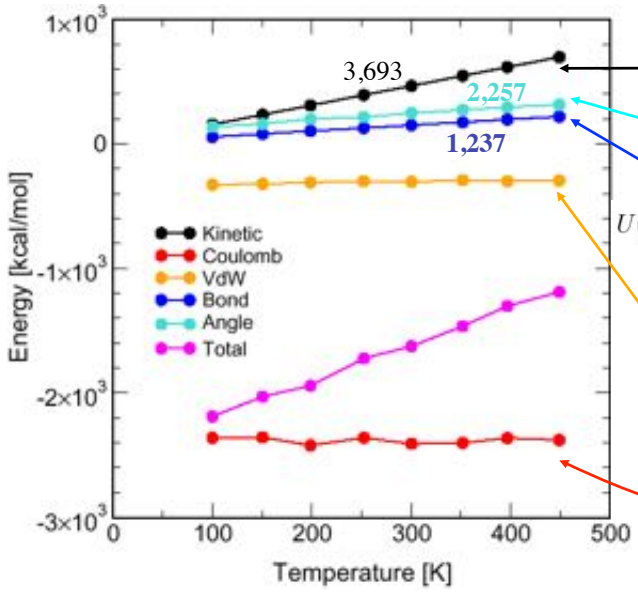
Kinetic energy (quadratic)

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

Potentialc energy (not all quadratic)

Equilibrium Properties of Proteins

Energies: kinetic and potential



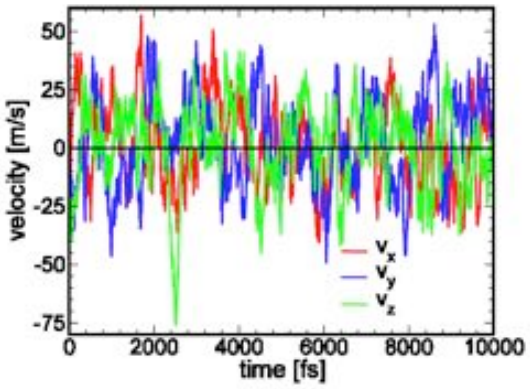
$$\langle \sum_j \frac{1}{2} m_j v_j^2 \rangle = \frac{3}{2} N k_B T$$

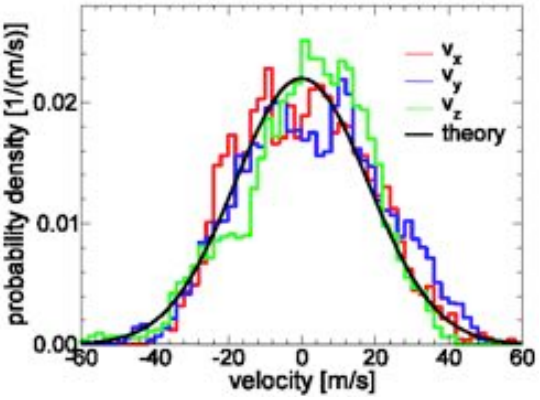
Kinetic energy (quadratic)

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

Potentialc energy (not all quadratic)

Maxwell Distribution of Atomic Velocities

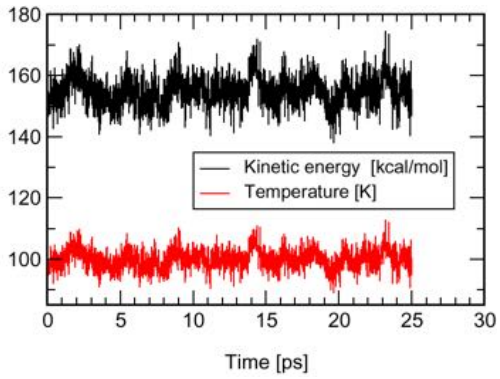


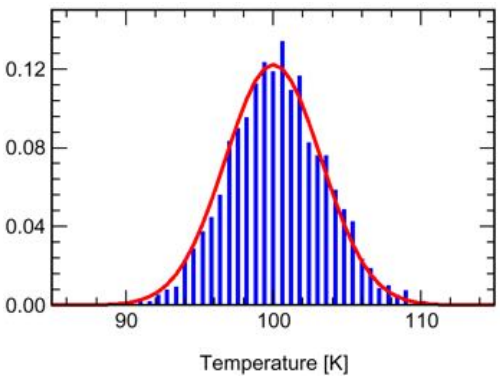


$$p(v_\sigma) = \sqrt{\frac{m}{2\pi k_B T}} \exp\left[-\frac{m v_\sigma^2}{2 k_B T}\right]$$

$$\sigma = x,\ y,\ z$$

Analysis of E_{kin}, T (free dynamics)





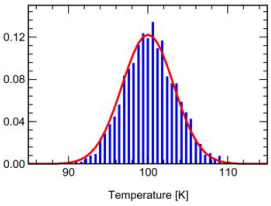
Definition of Temperature

$$\langle \sum_j \frac{1}{2} m_j v_j^2 \rangle = \frac{3}{2} N k_B T$$

$$T = \frac{2}{3N k_B} \langle \sum_j \frac{1}{2} m_j v_j^2 \rangle$$

The atomic velocities of a protein establish a thermometer, but is it accurate? Apparently not! Why!

The atomic velocity thermometer is inaccurate due to the finite size of a protein!



Temperatur Fluctuations

Maxwell distribution

$$dP(v_n) = c\,\exp(-m\,v_n^2/2k_BT)\,dv_n \tag{7}$$

Individual kinetic energy $\epsilon_n = mv_n^2/2$

$$dP(\epsilon_n) = (\pi T_0\epsilon_n)^{-1/2}\exp(-\epsilon_n/k_BT_0)\,d\epsilon_n \tag{8}$$

One can derive

$$\langle\epsilon_n\rangle = T_0/2 \tag{9}$$

$$\langle\epsilon_n^2\rangle = 3\,T_0^2/4 \tag{10}$$

$$\langle\epsilon_n^2\rangle - \langle\epsilon_n\rangle^2 = T_0^2/2 \tag{11}$$

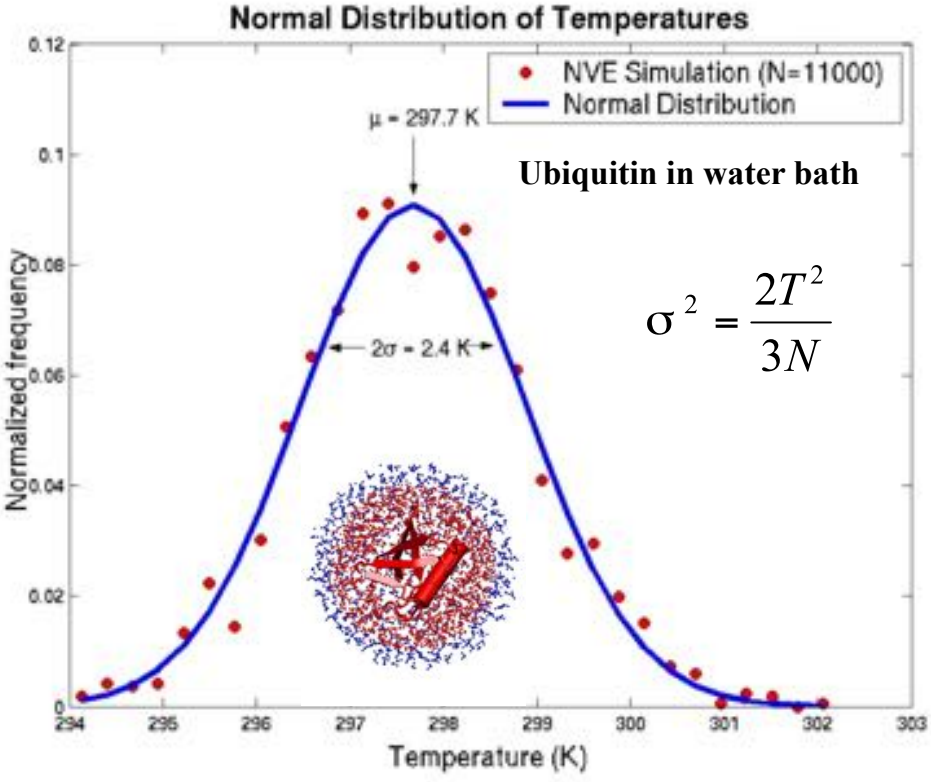
The distribution of the total kinetic energy $E_{kin} = \sum_j \frac{1}{2}m_jv_j^2$, according to the central limit theorem, is approximately Gaussian

$$P(E_{kin}) = c\,\exp\left(\frac{-(E_{kin}-\langle E_{kin}\rangle)^2}{2\left(\frac{3Nk_B^2T_0^2}{2}\right)}\right) \tag{12}$$

The distribution function for the temperature ($T = 2E_{kin}/3k_B$) fluctuations $\Delta T = T - T_0$ is then

$$P(\Delta T) = c\,\exp[-(\Delta T)^2/2\sigma^2]\,, \qquad \sigma^2 = 2\,T^2/3N \tag{13}$$

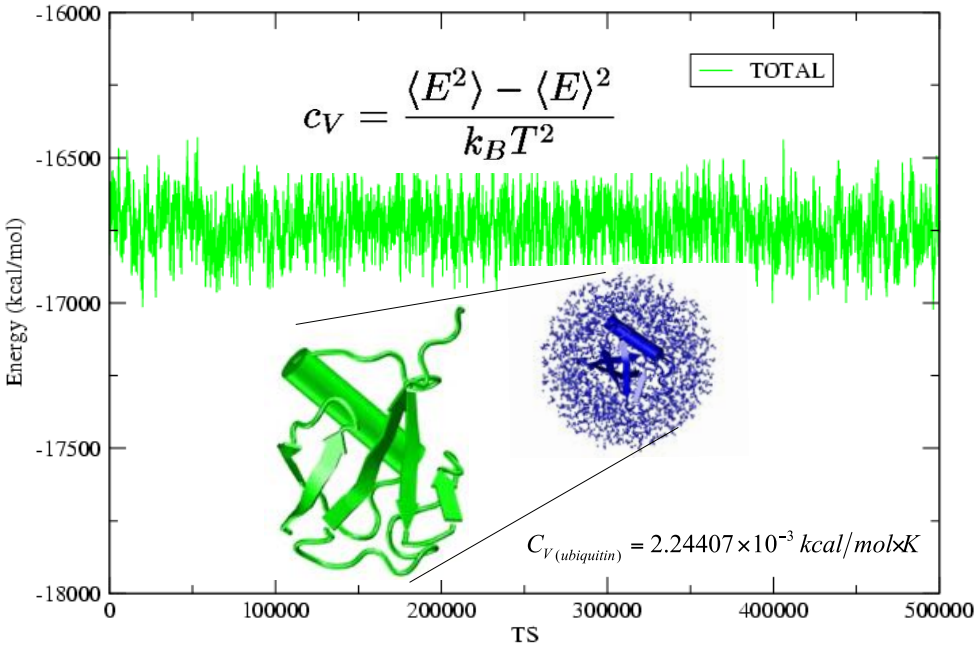
For $T_0 = 100$K and $N = 557$, this gives $\sigma = 3.6$.



$$\sigma^2 = \frac{2T^2}{3N}$$

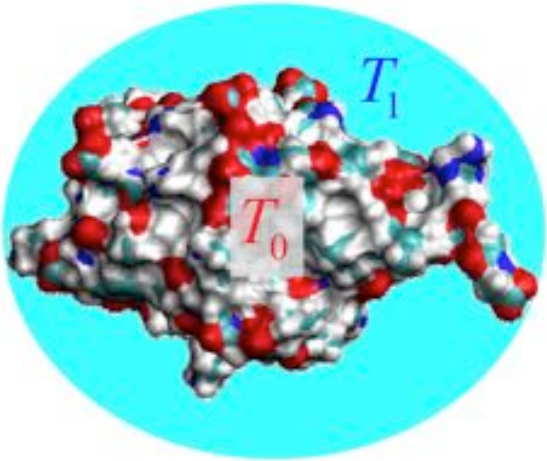
Specific Heat of a Protein

Total energy of ubiquitin (NVE ensemble)



Simulated Cooling of Ubiquitin

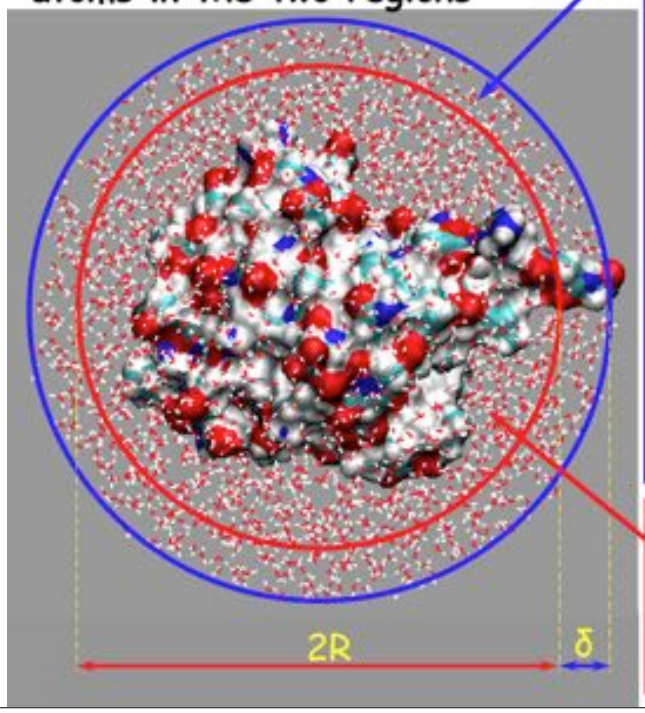
- Proteins function in a narrow (physiological) temperature range. What happens to them when the temperature of their surrounding changes significantly (temperature gradient) ?
- Can the heating/cooling process of a protein be simulated by molecular dynamics ? If yes, then how?



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

How to simulate cooling ?

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

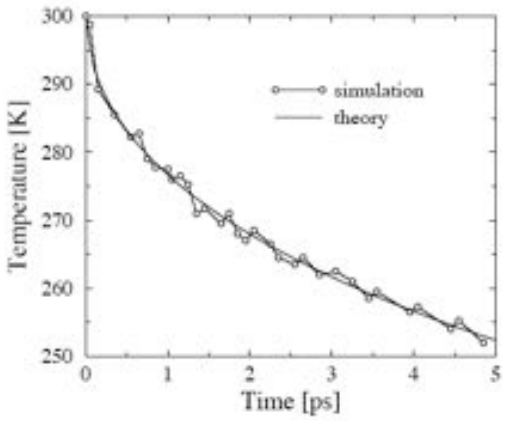
atoms in the inner region follow Newtonian dynamics

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

Simulated Cooling - Result

t	$\langle T_{sim} \rangle$	t	$\langle T_{sim} \rangle$	t	$\langle T_{sim} \rangle$	t	$\langle T_{sim} \rangle$
0.05	298.75	1.05	276.00	1.95	267.00	3.25	261.00
0.15	289.25	1.15	276.50	2.05	268.50	3.45	258.50
0.35	285.50	1.25	275.25	2.25	266.50	3.55	259.50
0.55	282.25	1.35	271.00	2.35	264.50	3.95	256.50
0.65	282.75	1.45	271.75	2.55	263.50	4.05	257.25
0.75	279.00	1.65	269.50	2.65	264.50	4.45	254.00
0.85	277.75	1.75	271.00	2.85	262.00	4.55	255.25
1.00	277.50	1.85	268.00	3.05	262.50	4.85	252.00

Table 1: Mean temperature $\langle T_{sim} \rangle$ [K] of the protein as a function of time t [ps].



Result from simulation that cools a protein, initially at T = 300 K, from the edge down to 200 K. Only first 5 ps are shown in which short time interval T drops to 250 K.

Heat Conduction Equation

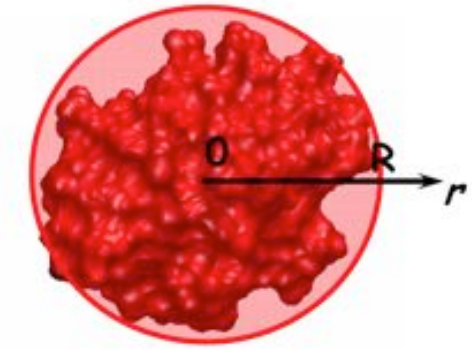
$$\frac{\partial T(\mathbf{r},t)}{\partial t} = D\nabla^2 T(\mathbf{r},t)$$

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:

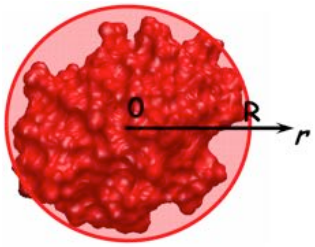
$$T(r,0) = T_0 \quad for \; r < R$$

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

Solution of the Heat Equation

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

$$D = K/\rho c\,,$$



Initial condition

$$T(\mathbf{r},0) = \langle T_{sim}\rangle(0) \qquad \text{for } r < R\,,$$

Boundary condition

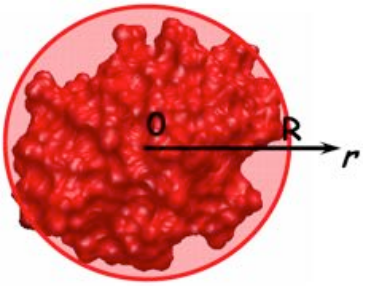
$$T(R,t) = T_{bath}\,.$$

Solution of the Heat Equation

Spherical symmetry

$$\frac{\partial T(r,t)}{\partial t} = D\,\frac{1}{r}\partial_r^2 r\,T(r,t)$$

T_{bath}



We assume

$$T(r,t) = T_{bath} + \sum_{n=1}^{\infty} a_n e^{\lambda_n t} u_n(r)$$

difference from bath

Here u_n are the eigenfunctions of the spherical diffusion operator

$$L \equiv \frac{D}{r}\frac{d^2}{dr^2}r$$

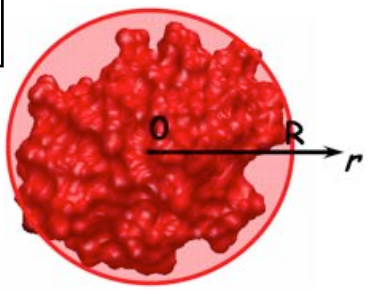
$$\frac{D}{r}\frac{d^2}{dr^2}r u_n(r) = \lambda_n\, u_n(r) \quad , \quad u_n(0) = \text{finite}\,, \quad u_n(R) = 0$$

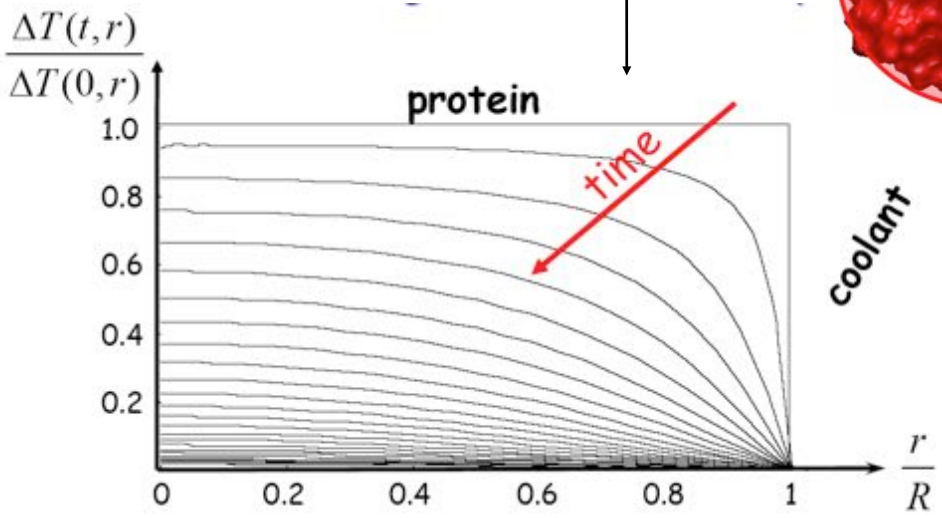
Solution of the Heat Equation

$$T(r,t) = T_{bath} + \boxed{\sum_{n=1}^{\infty} a_n \exp\left[-\left(\frac{n\pi}{R}\right)^2 D\,t\right]\frac{\sin\left(n\pi r/R\right)}{r}}$$

$$a_m = \frac{2R}{m\pi}\,\Delta T\,(-1)^{m+1}$$

T_{bath}

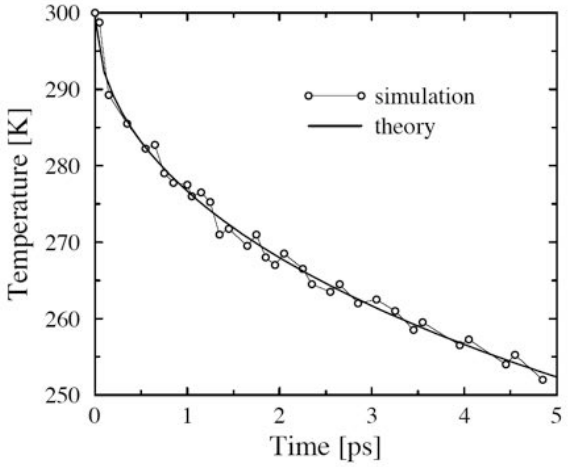


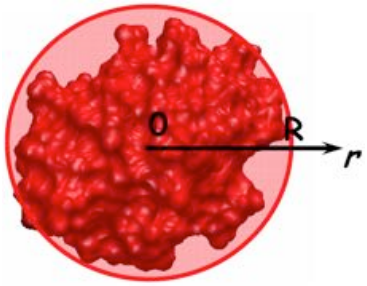


Solution of the Heat Equation

Temperature averaged over volume

$$\begin{aligned}
\langle T \rangle(t) &= \left(\frac{4\pi R^3}{3}\right)^{-1} \int d^3\mathbf{r}\, T(\mathbf{r},t) = \frac{3}{R^3}\int_0^R r^2 dr\, T(r,t) \\
&= T_{bath} + \sum_{n=1}^{\infty} a_n \exp\left[-\left(\frac{n\pi}{R}\right)^2 D\,t\right] \frac{3}{R^3}\int_0^R r dr \sin\left(\frac{n\pi r}{R}\right) \\
&= T_{bath} + 6\frac{\Delta T}{\pi^2}\sum_{n=1}^{\infty}\frac{1}{n^2}\exp\left[-\left(\frac{n\pi}{R}\right)^2 D\,t\right]
\end{aligned}$$





$D \approx 0.38 \times 10^{-3}\mathrm{cm}^2\mathrm{s}^{-1}$

water $1.4 \times 10^{-3}\mathrm{cm}^2\mathrm{s}^{-1}$