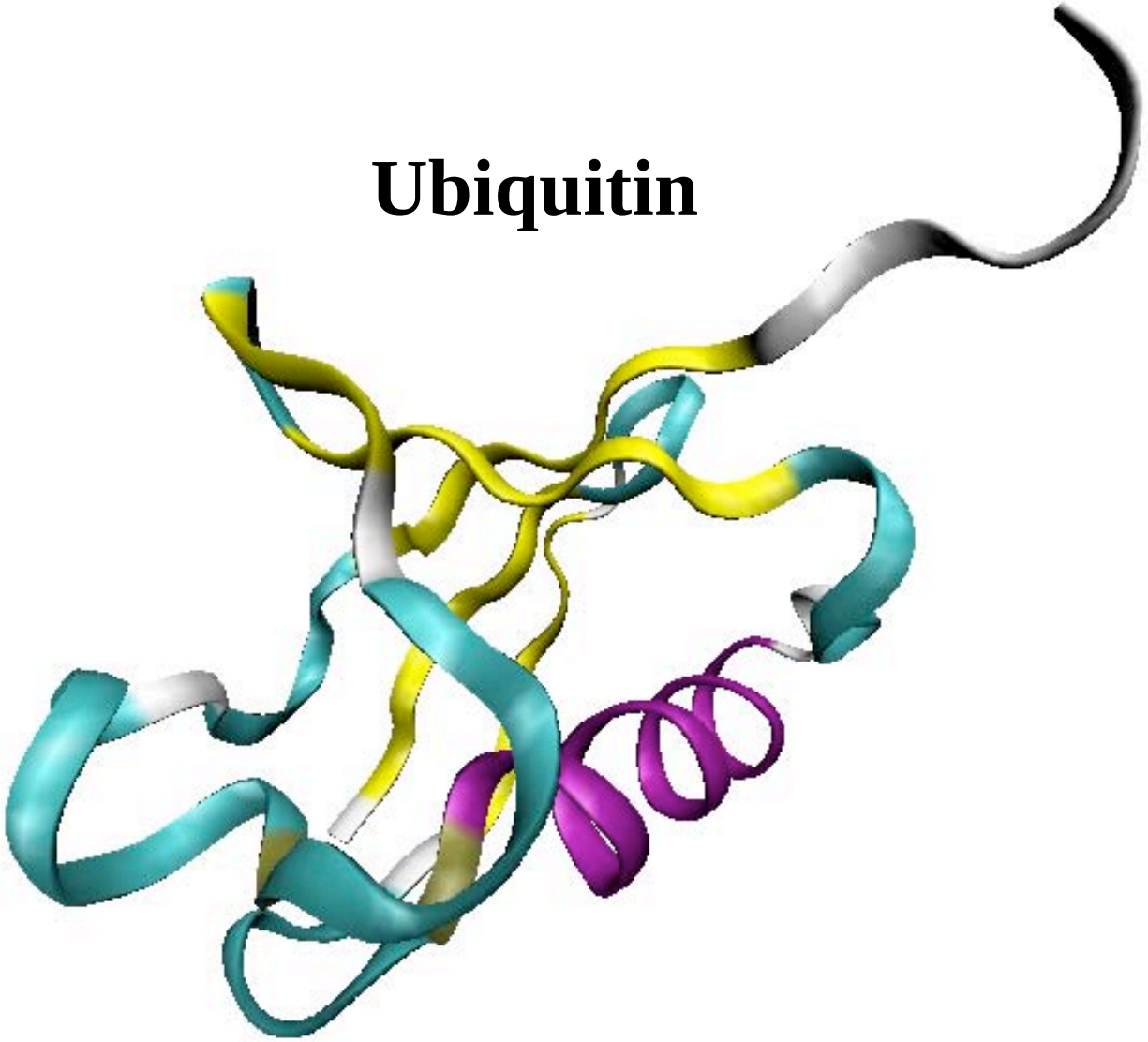
Molecular Dynamics of Proteins

Ubiquitin

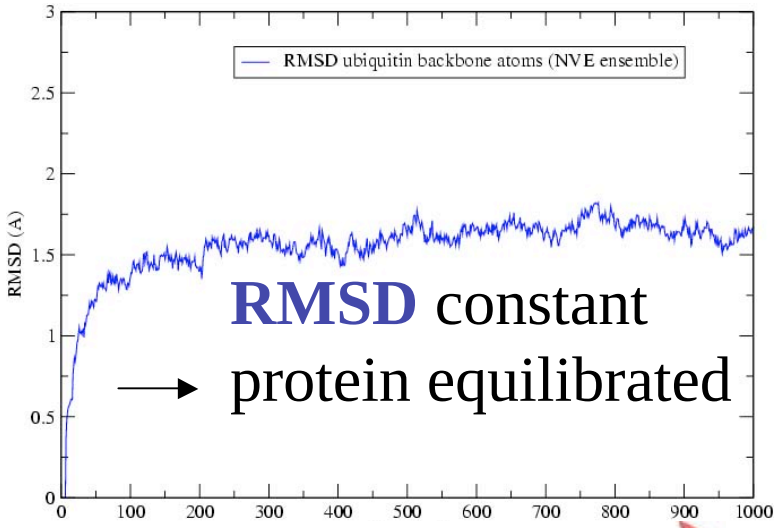


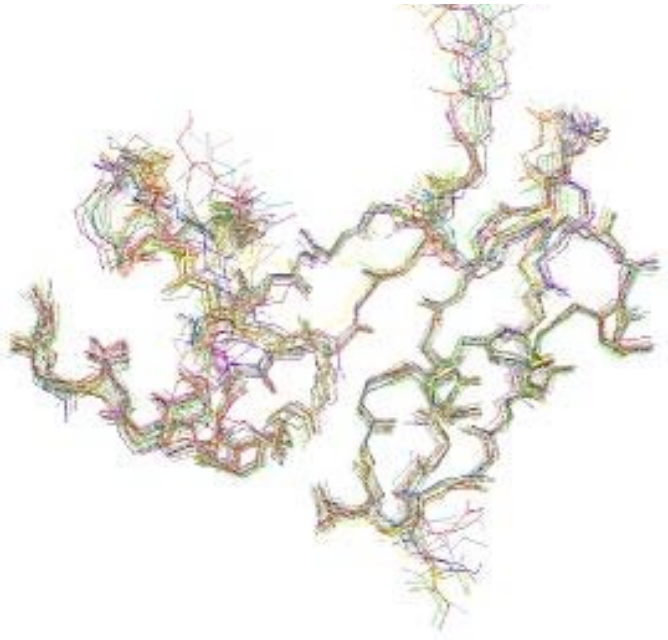
Equilibrium Properties of Proteins

Ubiquitin

Root Mean Squared Deviation: measure for equilibration and protein flexibility

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

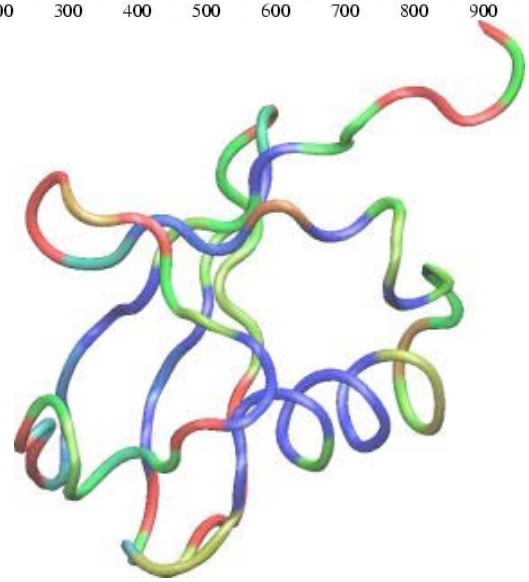




NMR structures

aligned together to see flexibility

Protein sequence exhibits characteristic permanent flexibility!

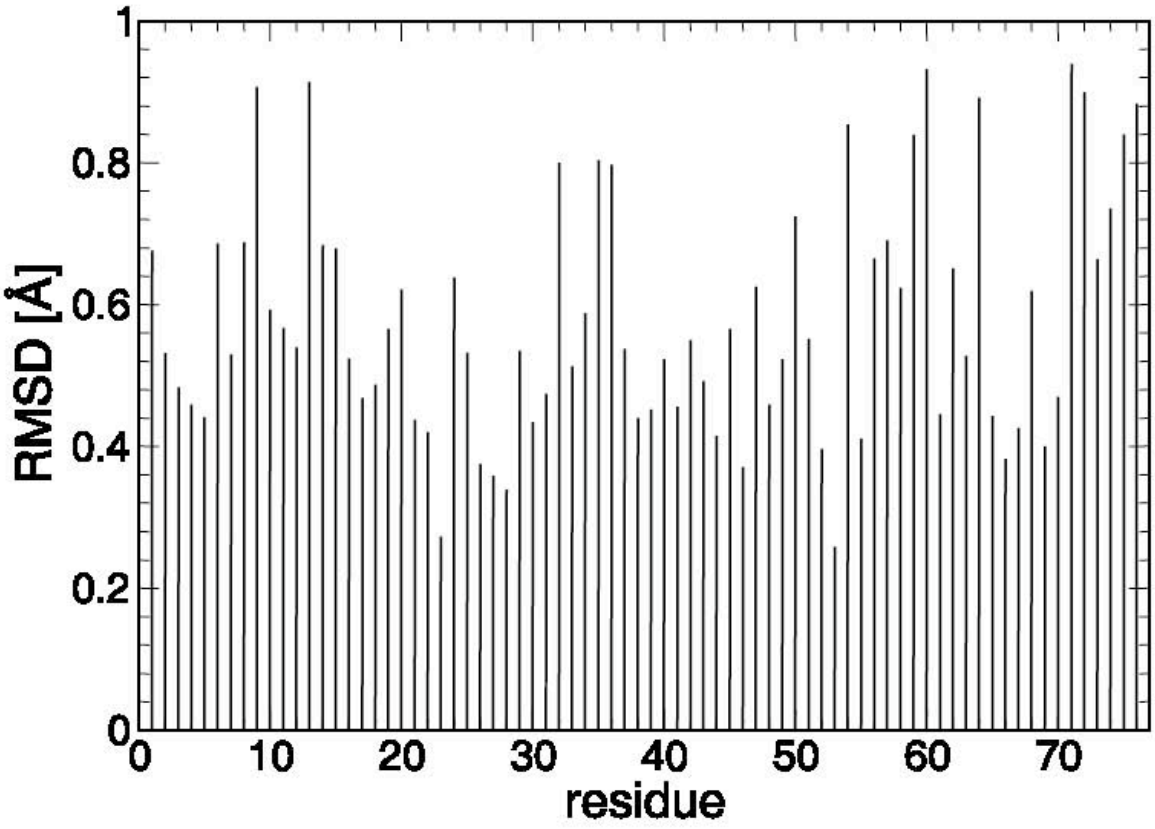


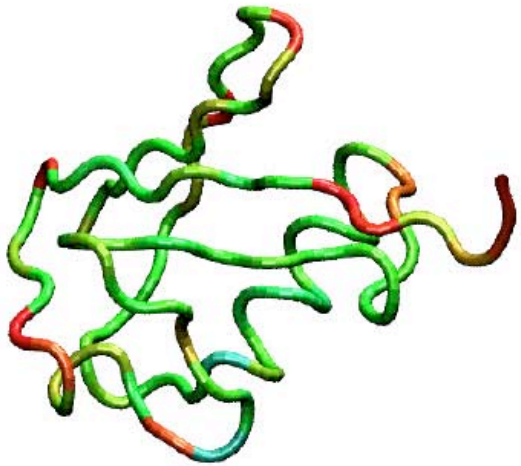
MD simulation

The color represents mobility of the protein through simulation (red = more flexible)

Thermal Motion of Ubiquitin from MD

RMSD values per residue



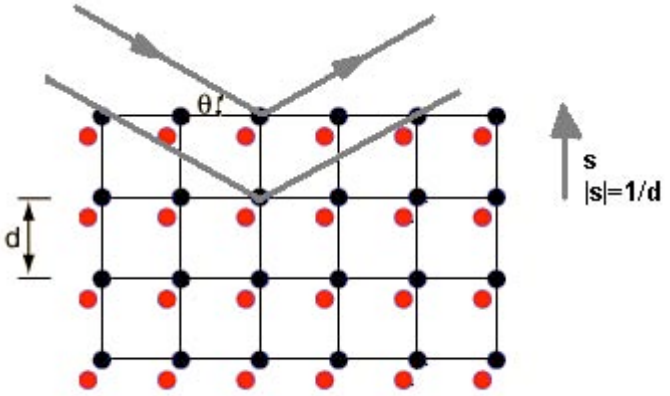


Thermal Motion of Ubiquitin from MD

Temperature Dependence of Crystal Diffraction (Debye-Waller factor)

Bragg's law

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



structure factor

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

The diffraction signal is the sum of the structure factors of all atoms in the crystal.

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:

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

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

One can carry out the expansion further and show

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

Using for the thermal amplitude of the harmonic oscillator

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

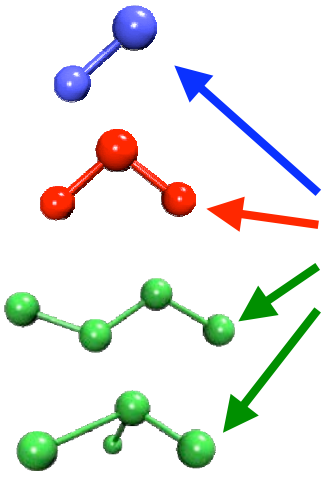
one obtains

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

Equilibrium Properties of Proteins

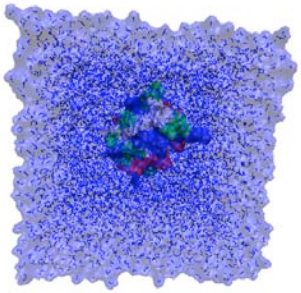
Energies: kinetic and potential

?



from ubiquitin.psf

26190	!NATOM:	atoms
17867	!NBOND:	bonds
10572	!NTHETA:	angles
3293	!NPHI:	dihedrals
204	!NIMPHI:	impropers



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

Kinetic energy (quadratic)

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

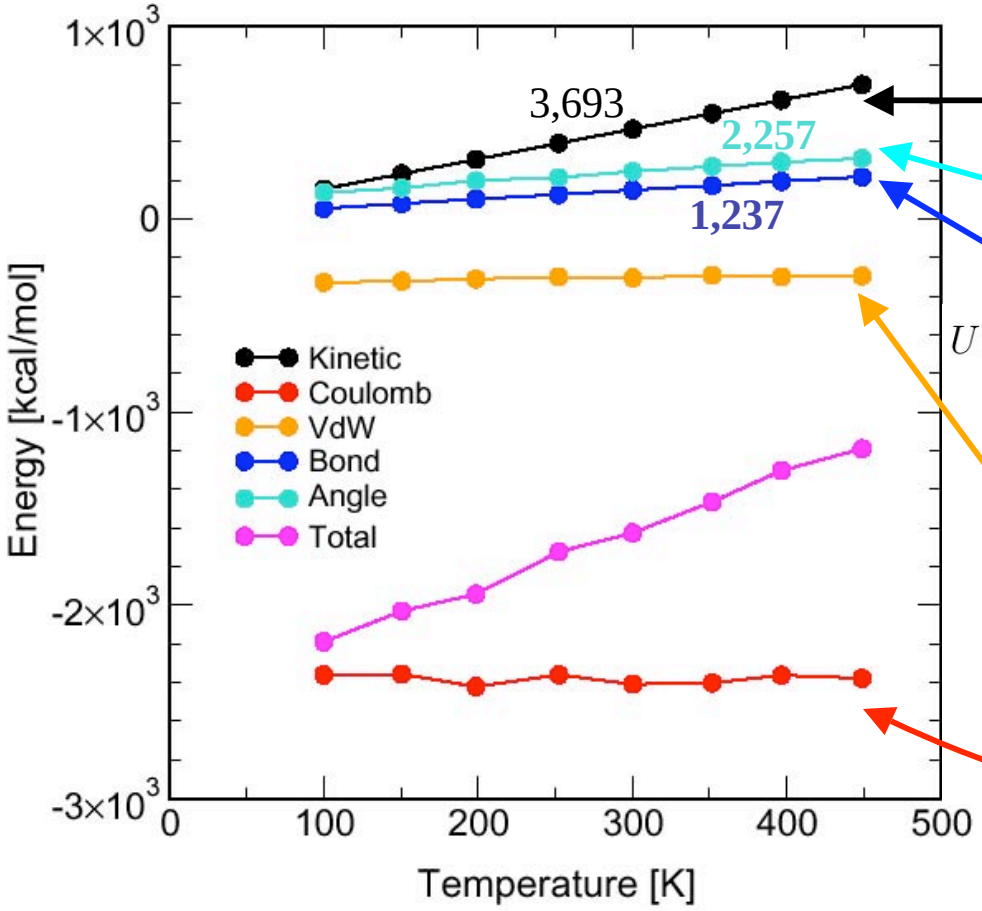
$$\underbrace{\sum_{dihedrals} k_i^{dihe}[1 + \cos(n_i \phi_i + \delta_i)]}_{U_{dihedral}} +$$

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

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)

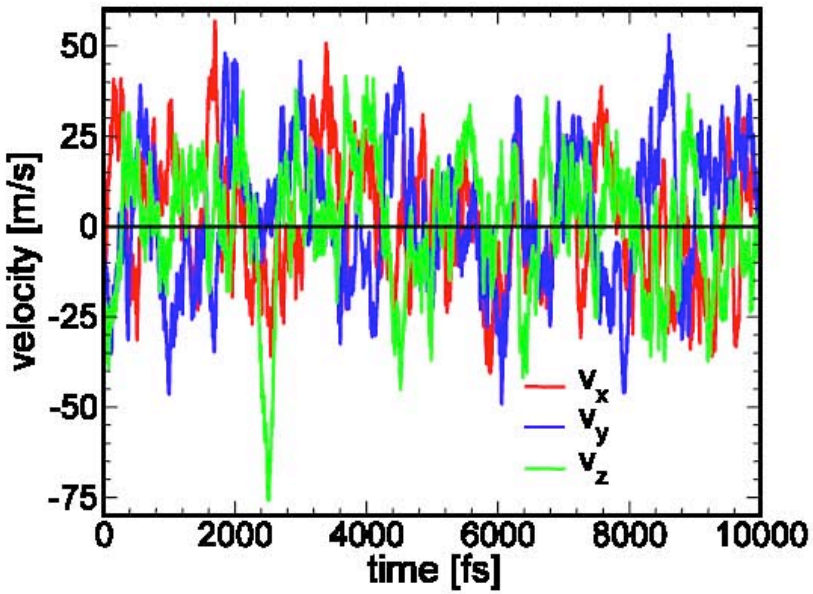
$$U(\vec{R}) = \underbrace{\sum_{bonds} k_i^{bond}(r_i - r_0)^2}_{U_{bond}} + \underbrace{\sum_{angles} k_i^{angle}(\theta_i - \theta_0)^2}_{U_{angle}} +$$

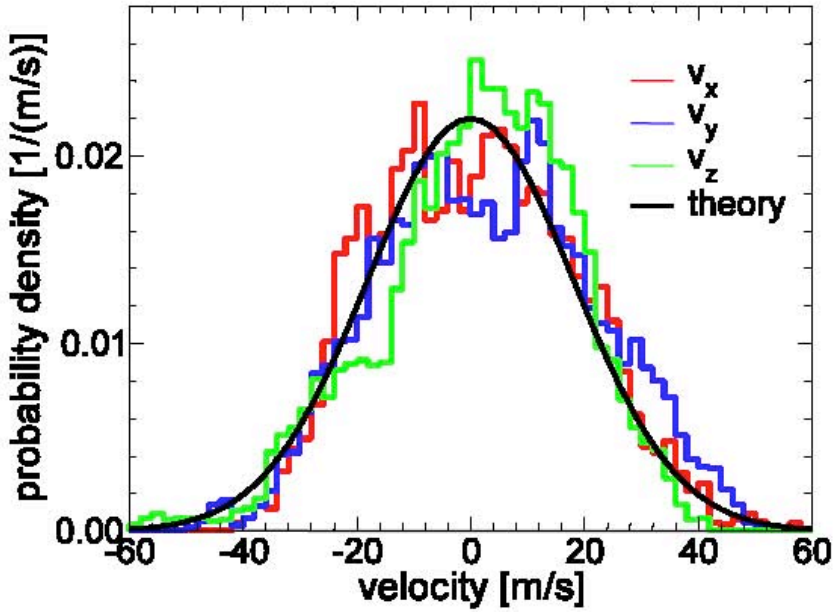
$$\underbrace{\sum_{dihedrals} k_i^{dihe}[1 + \cos(n_i \phi_i + \delta_i)]}_{U_{dihedral}} +$$

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

Potentialc energy (not all quadratic)

Maxwell Distribution of Atomic Velocities





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

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

Mean Kinetic Energy

Exercise in Statistics

$$\left\langle \tfrac{1}{2}mv^2 \right\rangle = \int_{-\infty}^{\infty} dv \left(\tfrac{1}{2}mv^2\right) p(v)$$

$$= \sqrt{\frac{m}{2\pi k_B T}} \int_{-\infty}^{\infty} dv \left(\tfrac{1}{2}mv^2\right) \exp\left[-\frac{mv^2}{2k_B T}\right]$$

$$= k_B T \sqrt{\frac{1}{\pi}} \int_{-\infty}^{\infty} \sqrt{\frac{m}{2k_B T}}\, dv \left(\frac{mv^2}{2k_B T}\right) \exp\left[-\frac{mv^2}{2k_B T}\right]$$

$$= k_B T \sqrt{\frac{1}{\pi}} \int_{-\infty}^{\infty} dy\, y^2 \exp -y^2]$$

Use formula below: $\left\langle \tfrac{1}{2}mv^2 \right\rangle = \tfrac{1}{2}k_B T$

$$\int_0^{\infty} dy\, y^m \exp[-y^2] = \tfrac{1}{2}\Gamma\left(\tfrac{m+1}{2}\right)$$

$$\Gamma(x+1) = x\,\Gamma(x)\,, \quad \Gamma\left(\tfrac{1}{2}\right) = \sqrt{\tfrac{1}{2}}$$

Maxwell Kinetic EnergyDistribution

Second Exercise in Statistics

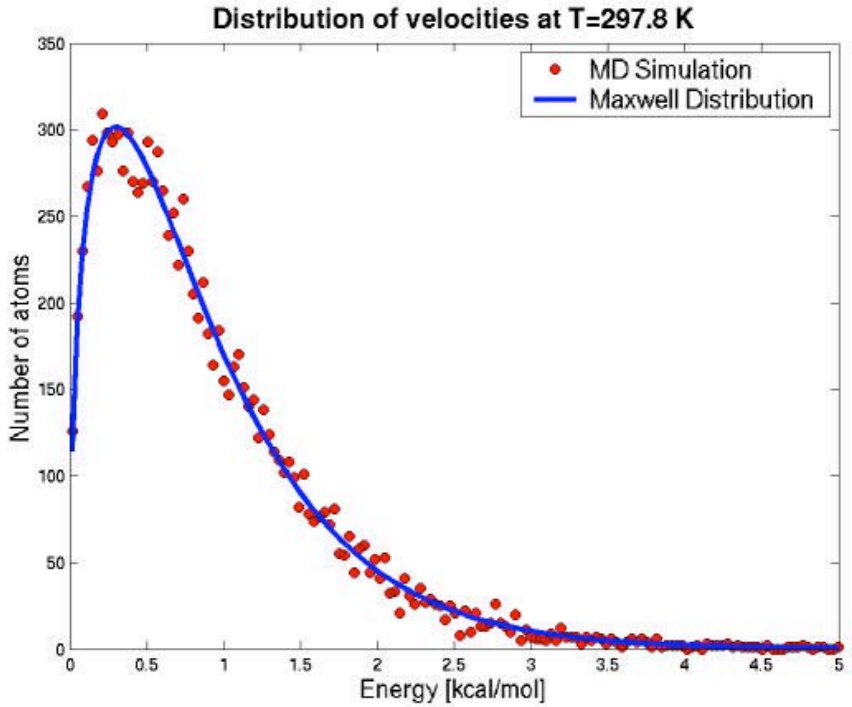
One-dimensional kinetic energy: $\epsilon_k = \frac{1}{2} m v_\sigma^2$

$$\tilde{p}(\epsilon_k) = p(v_\sigma)\frac{dv_\sigma}{d\epsilon_k} \rightarrow \tilde{p}(\epsilon_k) = \sqrt{1/\pi k_B T}\sqrt{1/\epsilon_k}\,\exp[-\epsilon_k/k_B T]$$

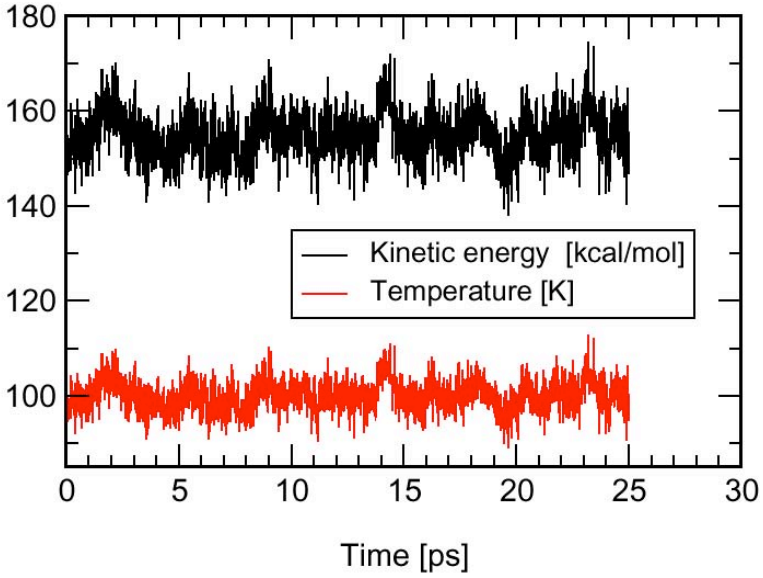
(factor 2 from restriction of integration to positive values)

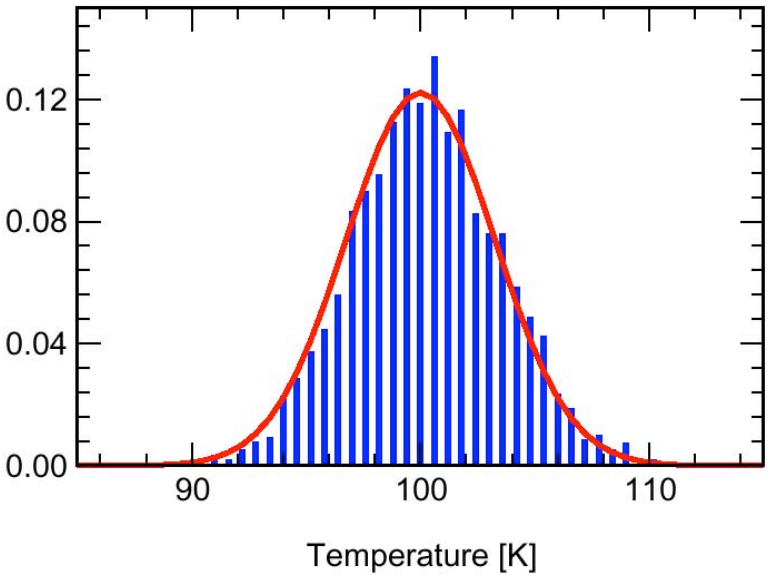
For the total kinetic energy (in three dimensions) holds then

$$\tilde{p}(\epsilon_k) = \frac{4}{\sqrt{\pi}}(k_B T)^{-\frac{3}{2}}\sqrt{\epsilon_k}\,\exp[-\epsilon_k/k_B T]$$



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





Definition of Temperature

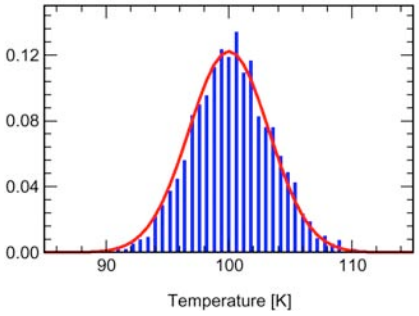
$$\langle \sum_j \frac{1}{2} m_j v_j^2 \rangle = \frac{3}{2} N k_B T$$

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

The atomic velocities of a protein establish a thermometer, but is it accurate?

Temperatur Fluctuations

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



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

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

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

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

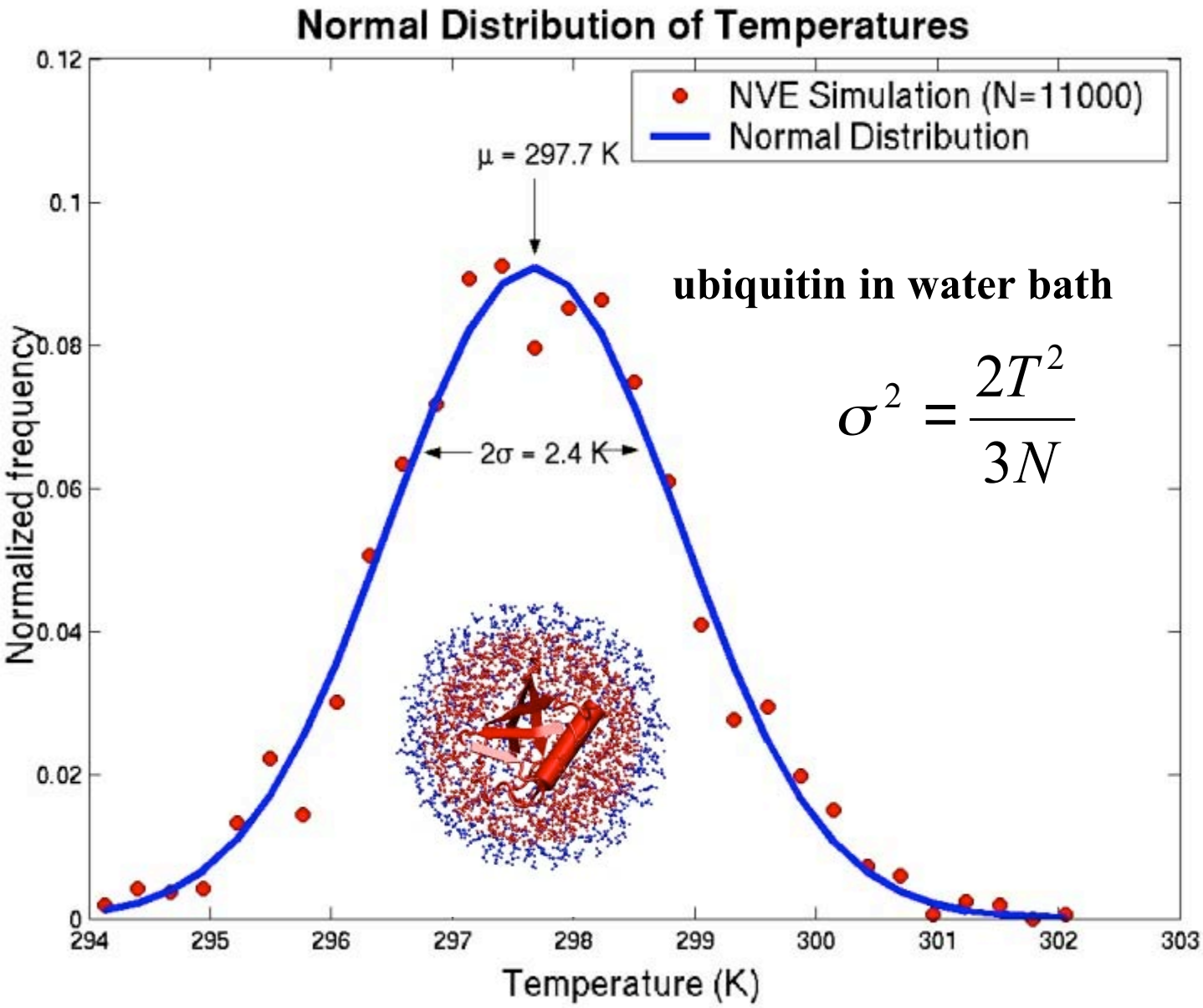
The distribution of the total kinetic energy $E_{kin} = \sum_j \frac{1}{2}m_jv_j^2$, according to the central limit theorem, is approximately Gaussian

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

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 $\sigma = 3.6$.

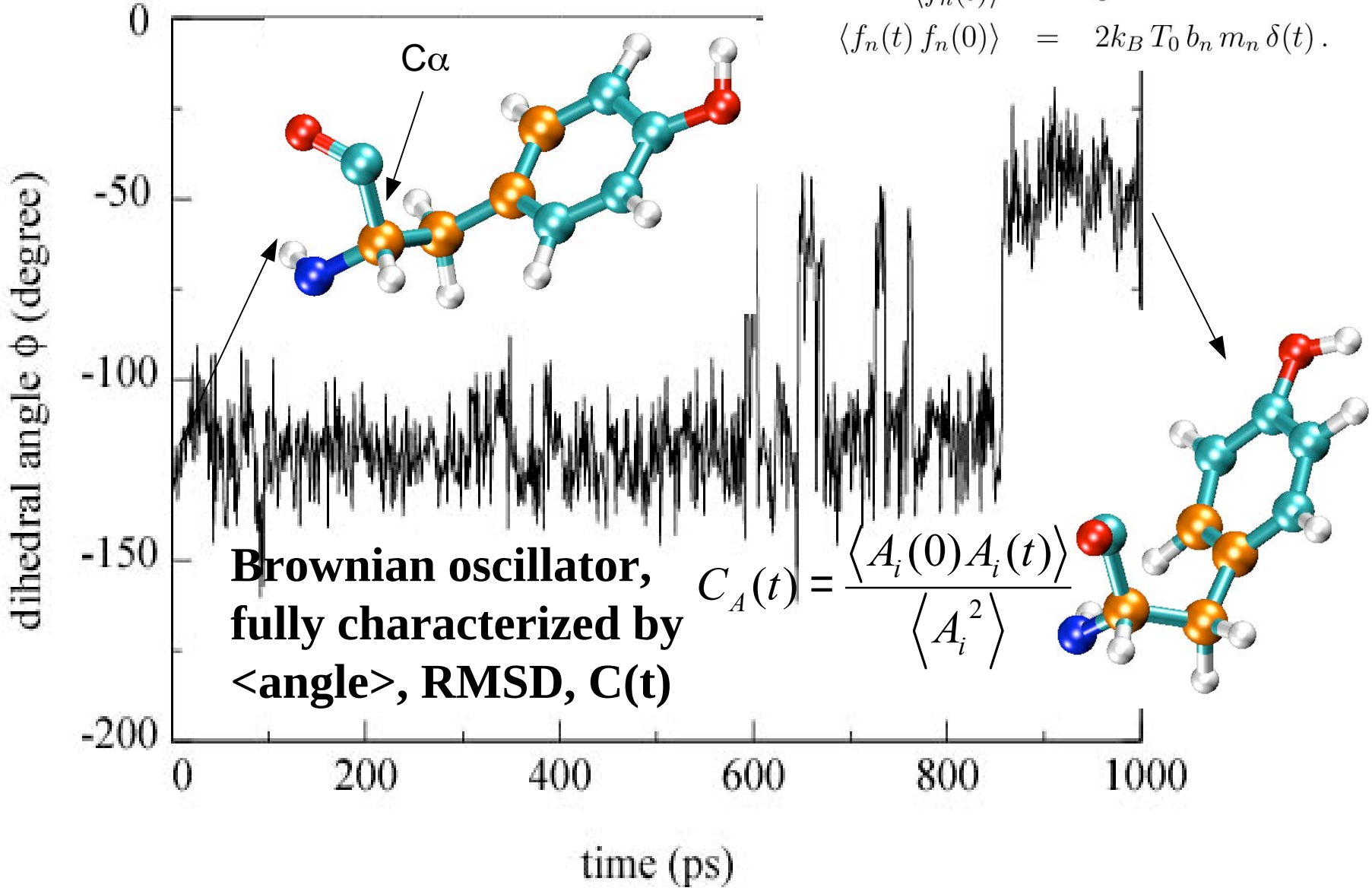

Normal Distribution of Temperatures
NVE Simulation (N=11000)
Normal Distribution
μ = 297.7 K
2σ = 2.4 K
ubiquitin in water bath
$\sigma^2 = \frac{2T^2}{3N}$
Normalized frequency
Temperature (K)
0.12
0.1
0.08
0.06
0.04
0.02
0
294
295
296
297
298
299
300
301
302
303

$$m_n \frac{d^2 x_n}{dt^2}(t) = \nabla_{x_n} V - m_n b_n \frac{dx_n}{dt}(t) + f_n(t)$$

Dihedral Angle

Langevin dynamics in strong friction limit

$$\langle f_n(t) \rangle = 0$$
$$\langle f_n(t)\, f_n(0) \rangle = 2 k_B T_0\, b_n\, m_n\, \delta(t)\,.$$



Brownian oscillator, fully characterized by <angle>, RMSD, C(t)

$$C_A(t) = \frac{\langle A_i(0) A_i(t) \rangle}{\langle A_i^2 \rangle}$$

Specific Heat of a Protein