Molecular Dynamics of Proteins
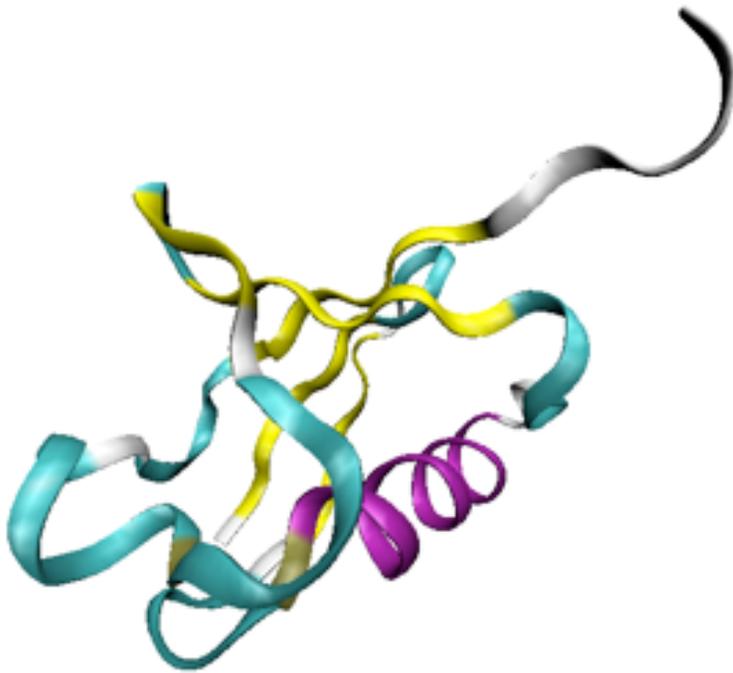
Ubiquitin
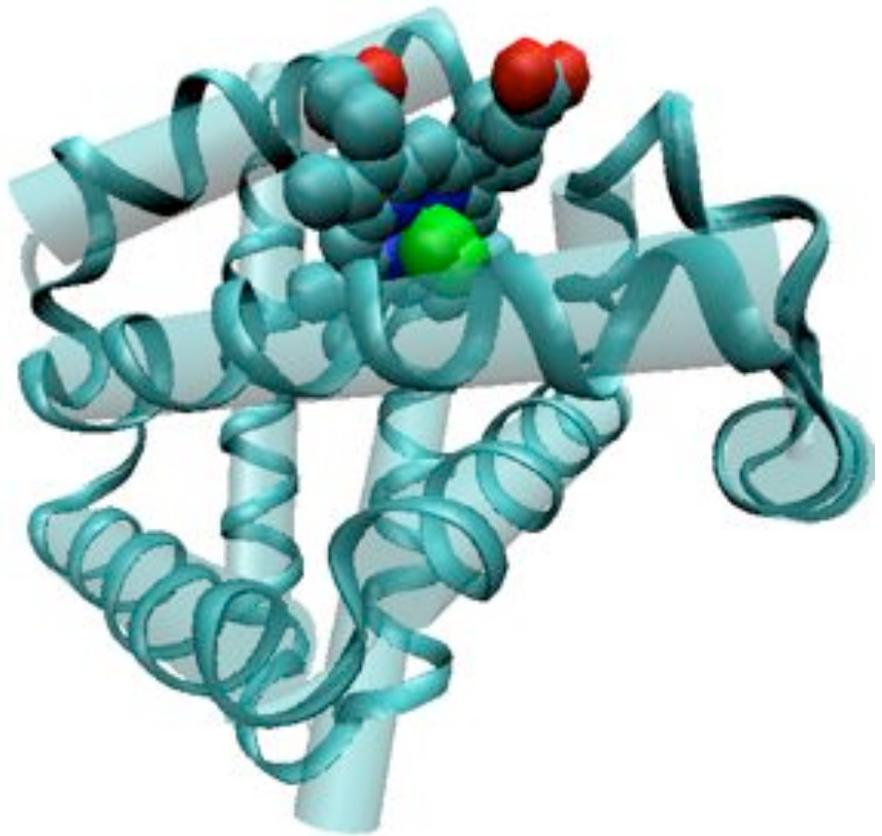
**Myoglobin**

Equilibrium Properties of Proteins

Ubiquitin

Root Mean Squared Deviation: measure for equilibration and protein flexibility

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

— RMSD ubiquitin backbone atoms (NVE ensemble)

RMSD constant
$\longrightarrow$ protein equilibrated

Protein sequence exhibits characteristic permanent flexibility!
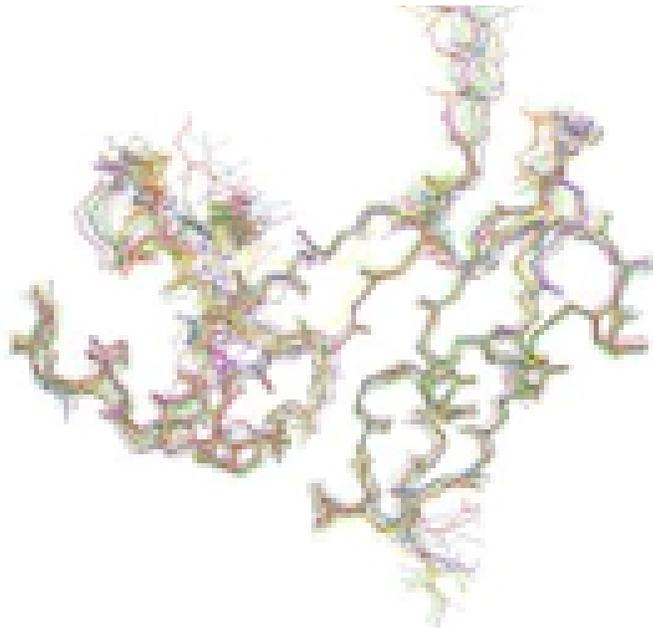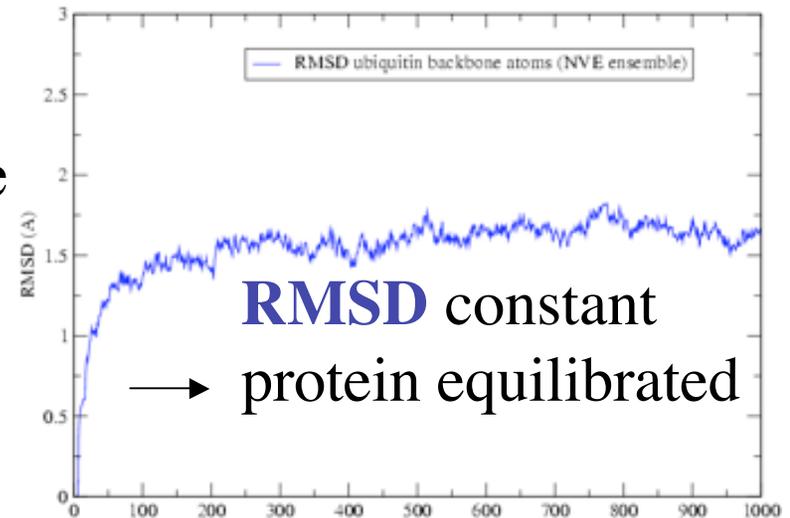

NMR structures
aligned together to see flexibility
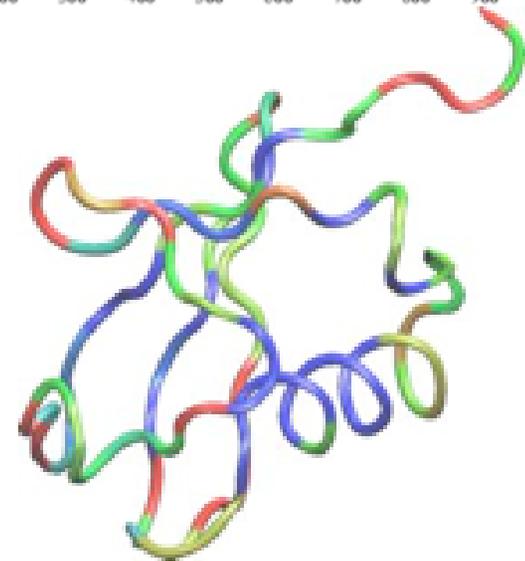
MD simulation
The color represents mobility of the protein through simulation (red = more flexible)

Thermal Motion of Ubiquitin from MD

RMSD values per residue
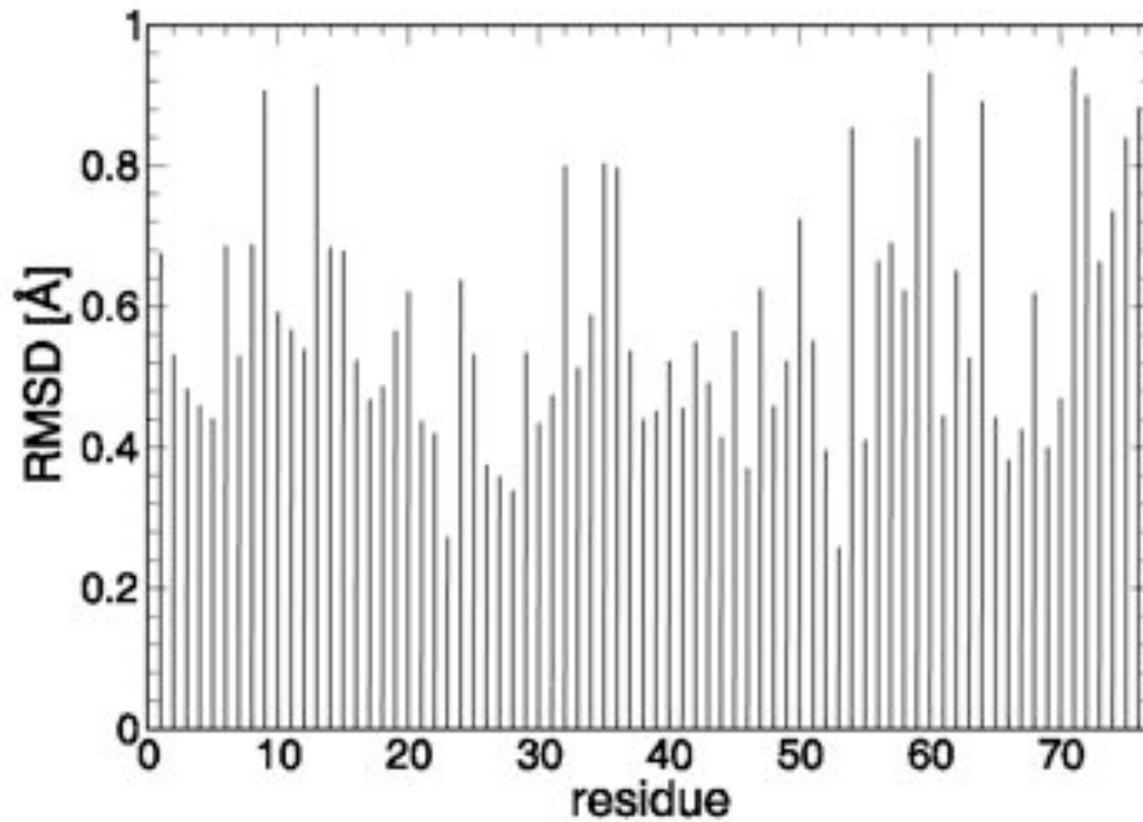
Thermal Motion of Ubiquitin from MD

Temperature Dependence of Crystal Diffraction (Debye-Waller factor)

Bragg's law

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

structure factor

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

The diffraction signal is the sum of the structure factors of all atoms in the crystal.
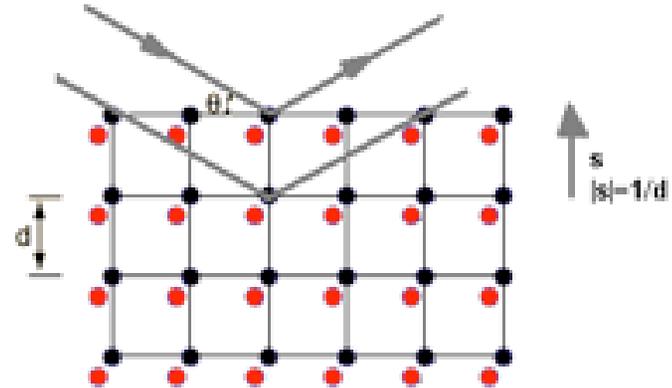
But the atom carries out thermal vibrations around equilibrium position $\vec{x}_j$

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

Accordingly:

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

Thermal Motion of Ubiquitin from MD

Temperature Dependence of Crystal Diffraction (Debye-Waller factor)

One can expand:

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

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

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

Using for the thermal amplitude of the harmonic oscillator

$$\tfrac{1}{2}m\omega^2 u_j^2 = \tfrac{3}{2}k_B T$$

one ob
tains

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

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)

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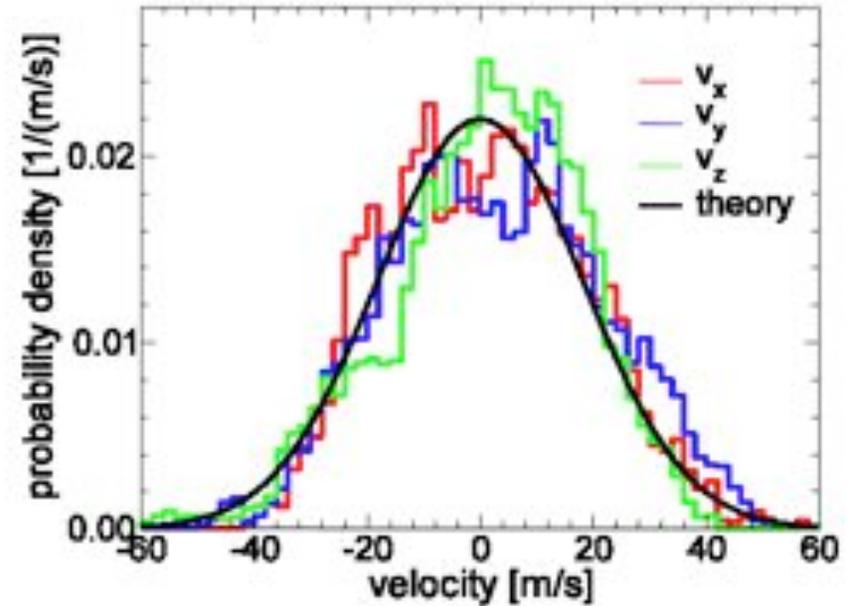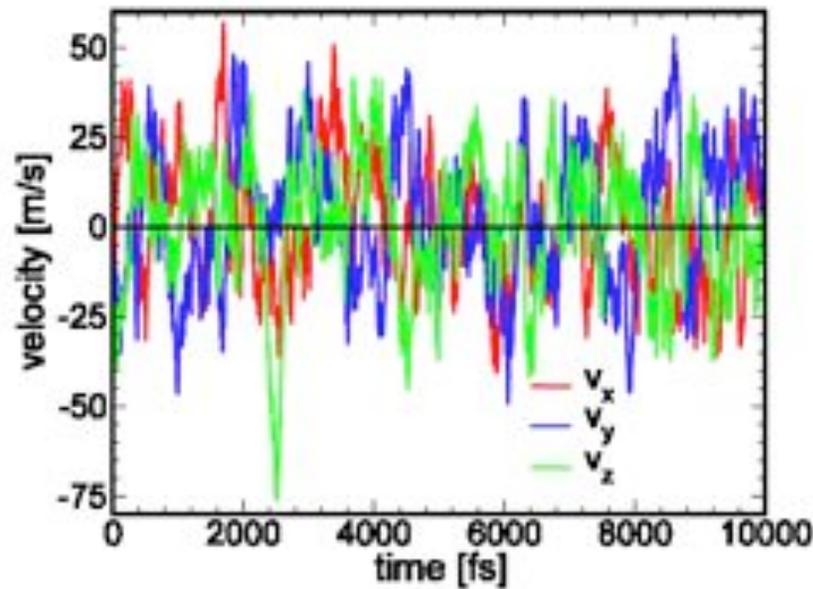
$$p(v_\sigma) = \sqrt{\frac{m}{2\pi k_B T}} \exp\left[-\frac{m v_\sigma^2}{2 k_B T}\right]$$

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

Mean Kinetic Energy
Exercise in Statistics

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

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

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

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

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

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

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

# Maxwell Kinetic EnergyDistribution
Second Exercise in Statistics

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

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

(factor 2 from restriction of integration to positive values)

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

$$\tilde{p}(\epsilon_k) =$$

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

Distribution of velocities at T=297.8 K
- MD Simulation
- Maxwell Distribution

Number of atoms

Energy [kcal/mol]
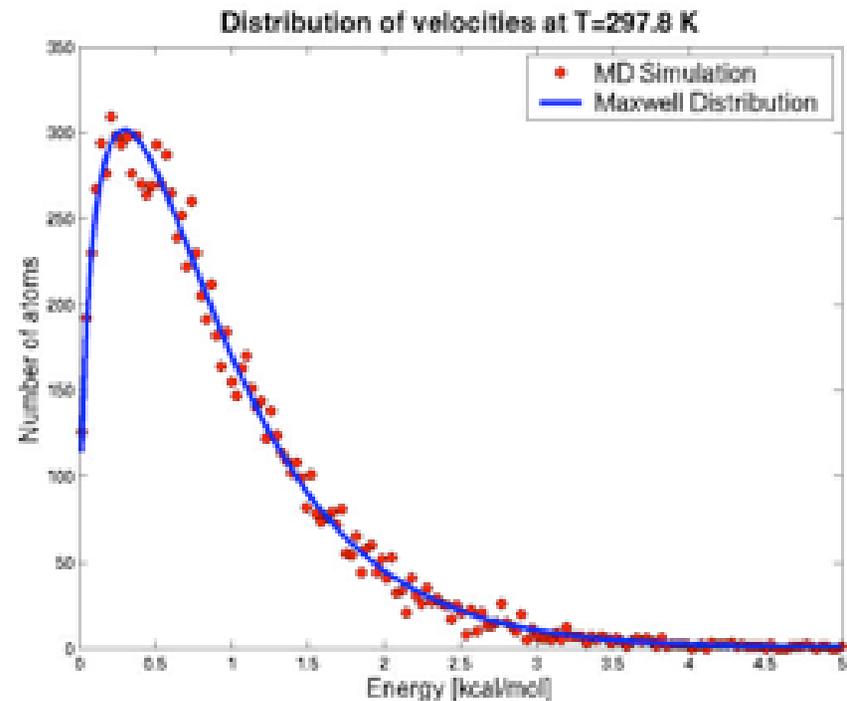
Analysis of E_{kin}, T (free dynamics)
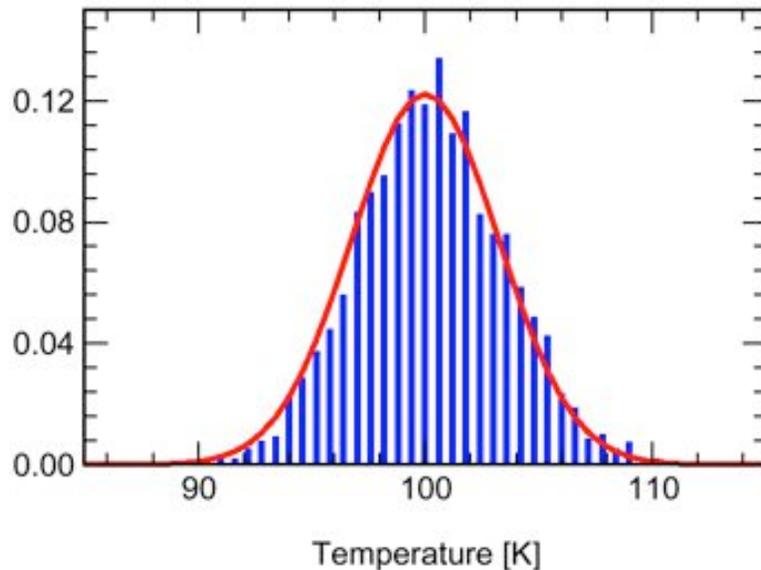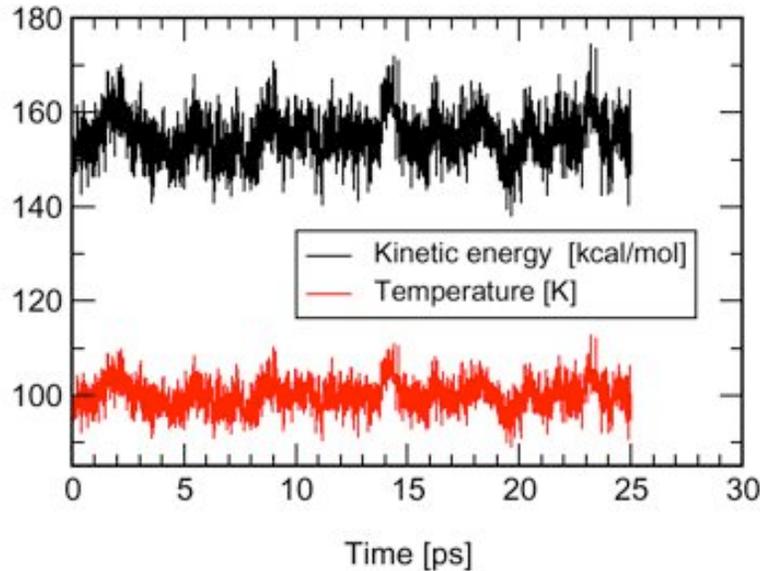
Definition of Temperature

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

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

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

Temperatur Fluctuations

Maxwell distribution

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

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

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

One can derive

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

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

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

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

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

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

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

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

The atomic velocity thermometer is inaccurate due to the finite size of a protein!
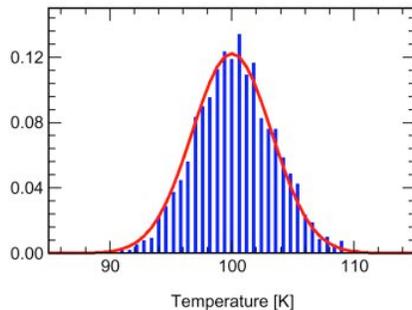
Temperature [K]

Normal Distribution of Temperatures

- NVE Simulation (N=11000)
- Normal Distribution

$\mu = 297.7$ K

$2\sigma = 2.4$ K

Normalized frequency

Temperature (K)

Another example, ubiquitin in water bath

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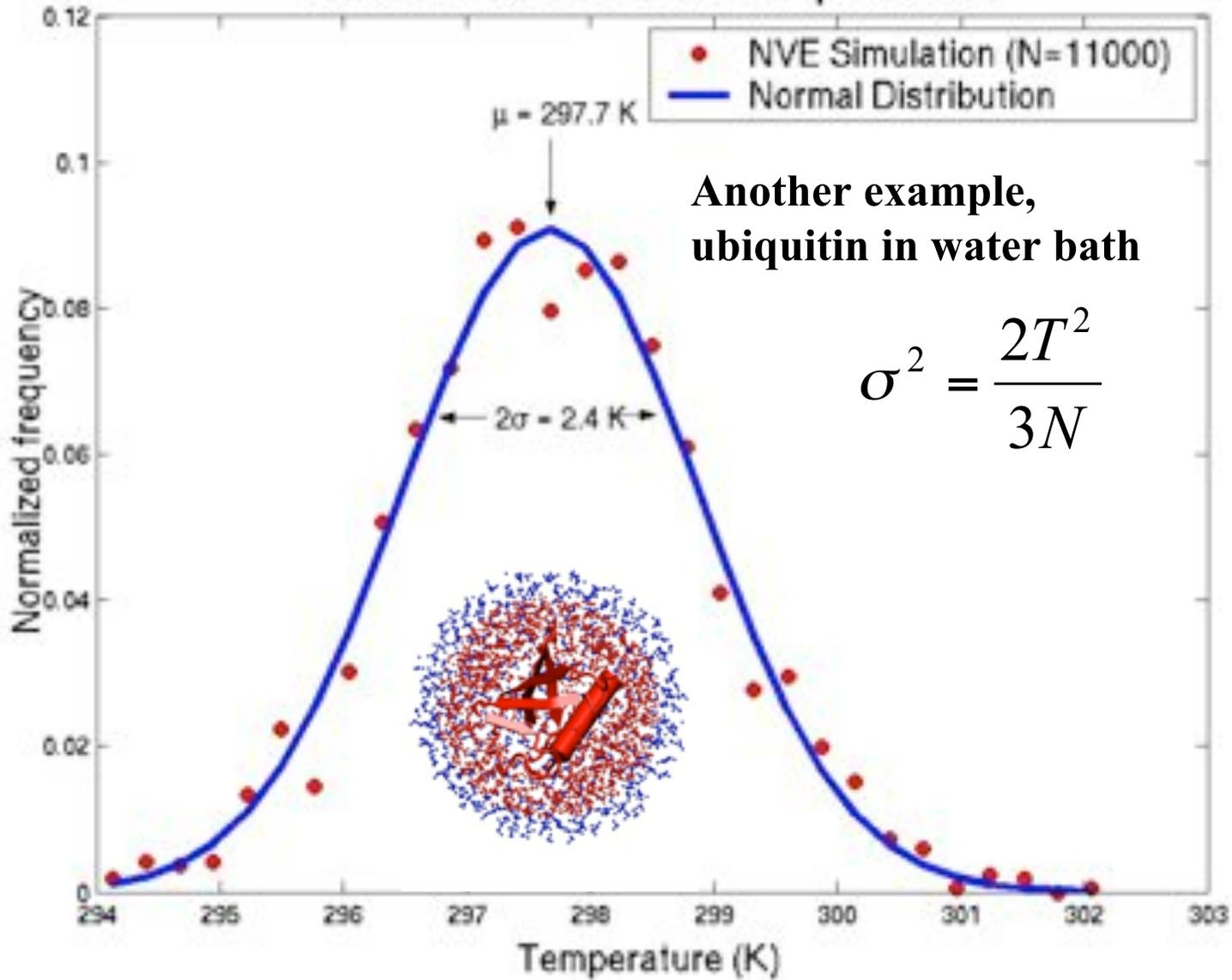
Show BPTI trajectory

Dihedral Angle

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

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

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

Cα

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

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

dihedral angle Φ (degree)

time (ps)
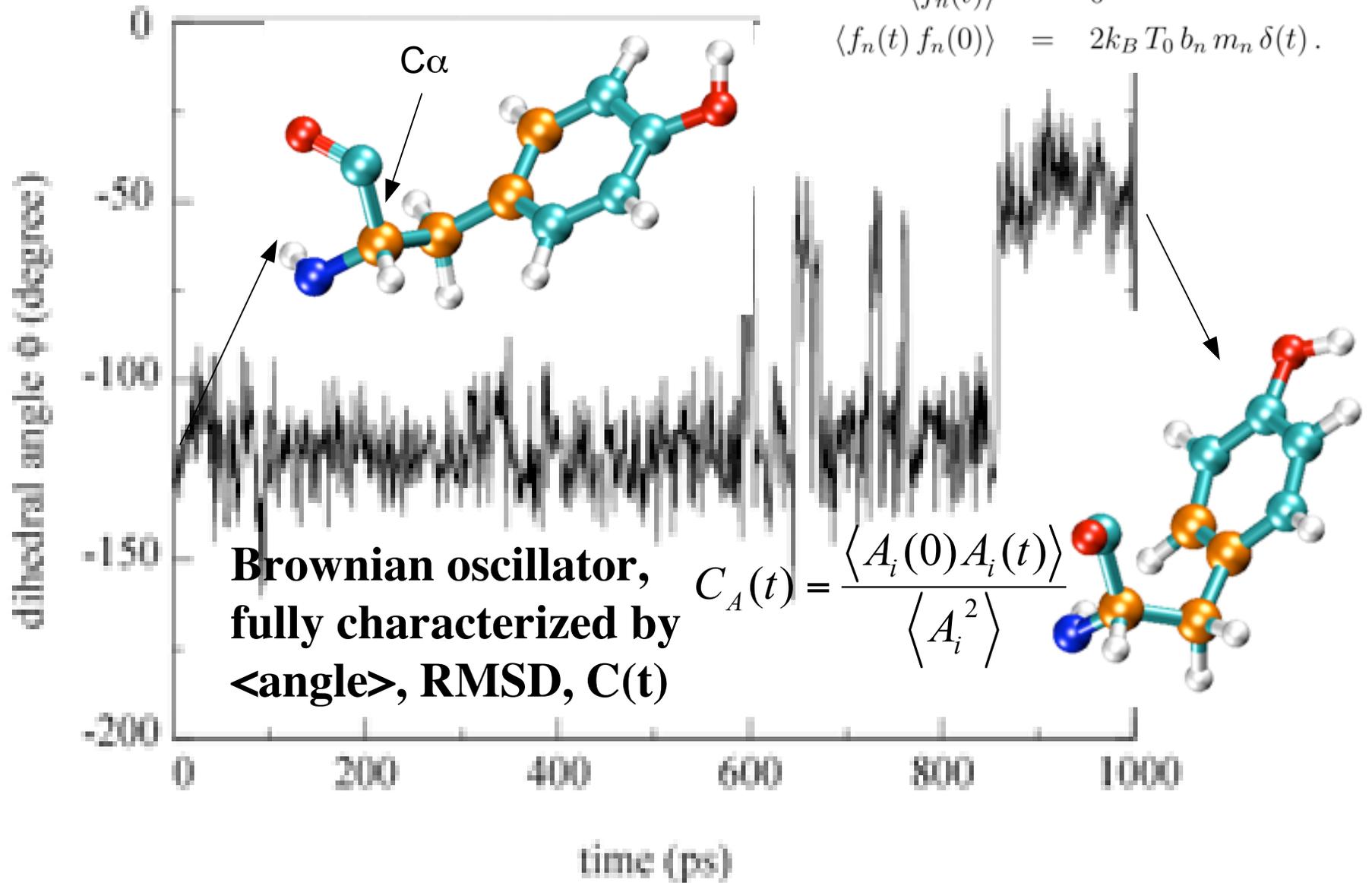
Specific Heat of a Protein

Total energy of ubiquitin (NVE ensemble)

$$c_V = \frac{\langle E^2 \rangle - \langle E \rangle^2}{k_B T^2}$$

$$C_{V\,(ubiquitin)} = 2.24407 \times 10^{-3} \; kcal/mol \cdot K$$
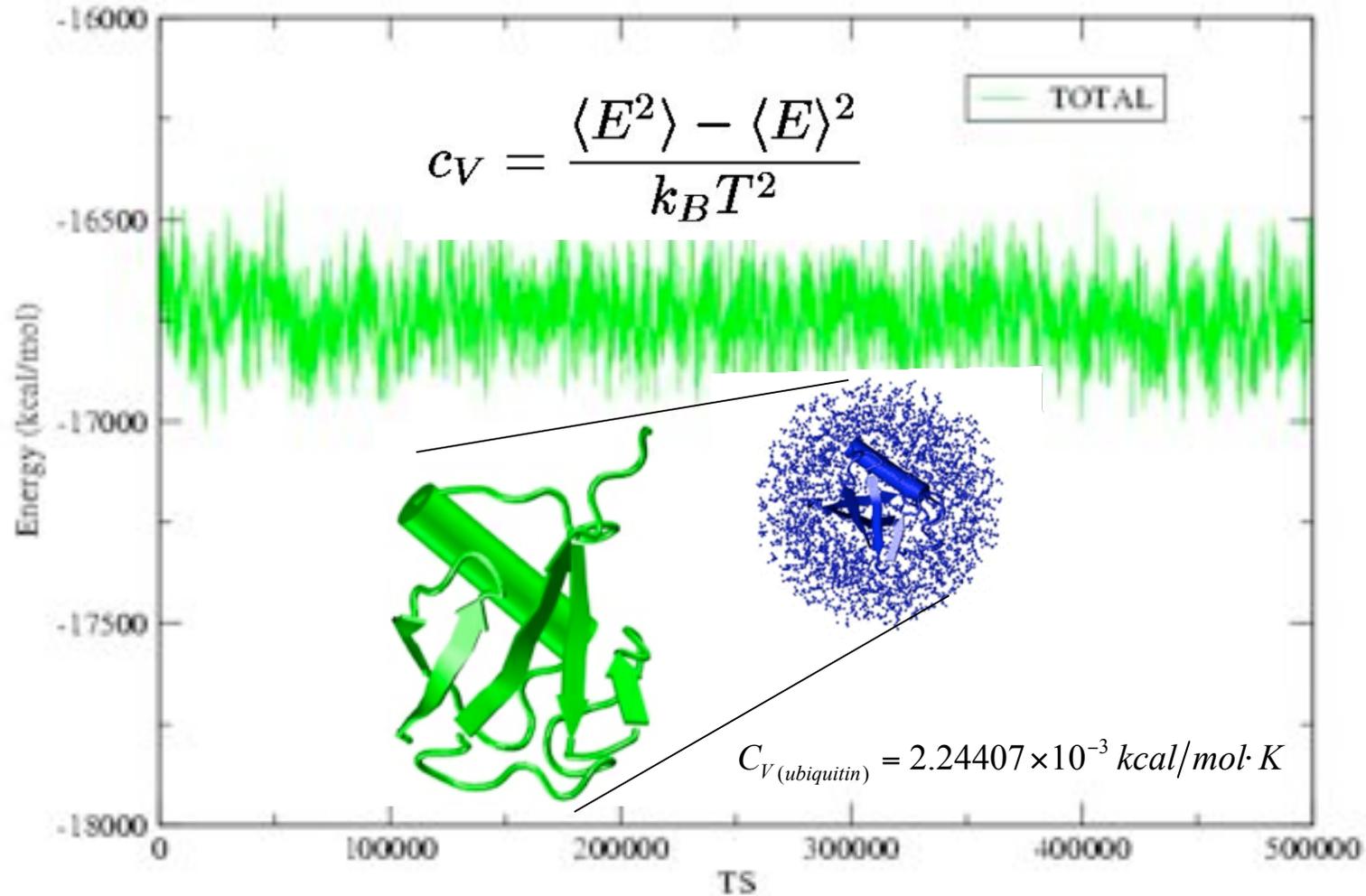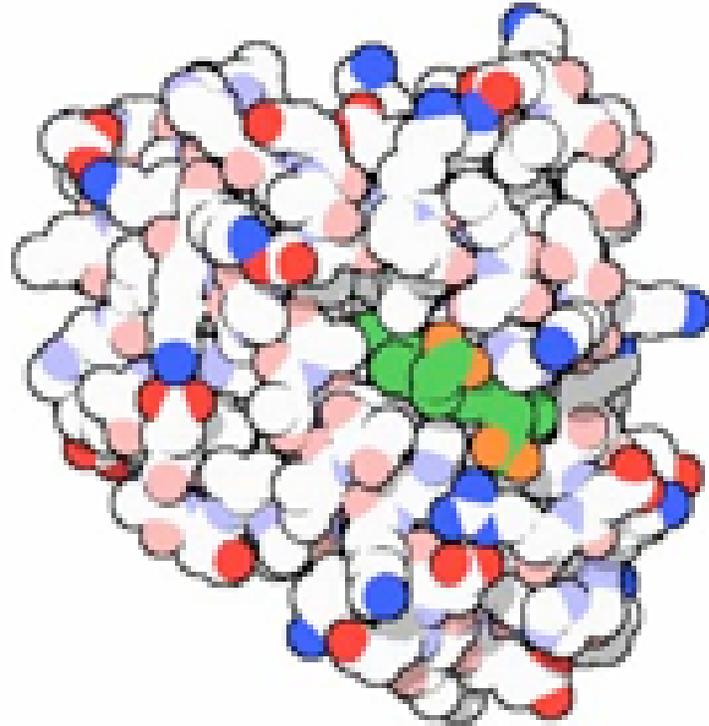
Myoglobin

Myoglobin is a small, bright red protein. It is very common in muscle cells, and gives meat much of its red color. Its job is to store oxygen, for use when muscles are hard at work. If you look at John Kendrew's PDB file, you will notice that the myoglobin that he used was taken from sperm whale muscles. As you can imagine, marine whales and dolphins have a great need for myoglobin, so that they can store extra oxygen for use in their deep dives undersea.

PDB Molecule of the Month: Myoglobin

Oxygen Bound to Myoglobin

This structure of myoglobin, with the accession code **1mbo**, shows the location of oxygen. The iron atom at the center of the heme group holds the oxygen molecule tightly. Compare the two pictures. The first shows only a set of thin tubes to represent the protein chain, and the oxygen is easily seen. But when all of the atoms in the protein are shown in the second picture, the oxygen disappears, buried inside the protein.So how does the oxygen get in and out, if it is totally surrounded by protein? In reality, myoglobin (and all other proteins) are constantly in motion, performing small flexing and breathing motions. Temporary openings constantly appear and disappear, allowing oxygen in and out. The structure in the PDB is merely one snapshot of the protein, caught when it is in a tightly-closed form. Looking at the static structure held in the PDB, we must imagine the dynamic structure that actually exists in nature.The two pictures above were created with RASMOL. You can create similar pictures by accessing the PDB file **1mbo**, and then clicking on "View Structure." Try switching between the two types of pictures shown above, to prove to yourself that the oxygen is buried in this structure!
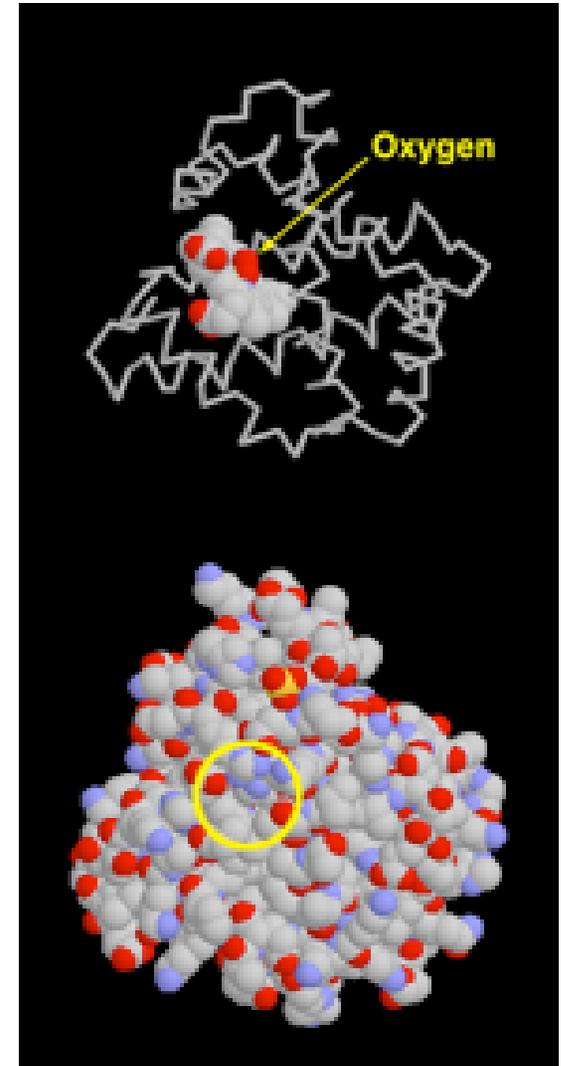
PDB Molecule of the Month: Myoglobin

Myogobin, the first protein with known structure
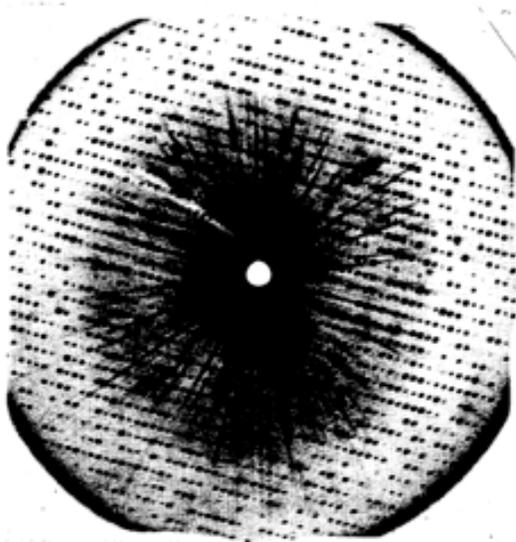
Diffraction pattern observed

**John Cowdery Kendrew
Nobel Prize in Chemistry
Jointly with Max Perutz**


Struture model at 6 A resolution

Higher resolution
Model:

1) Construct
   electron density
   map

2) Build model
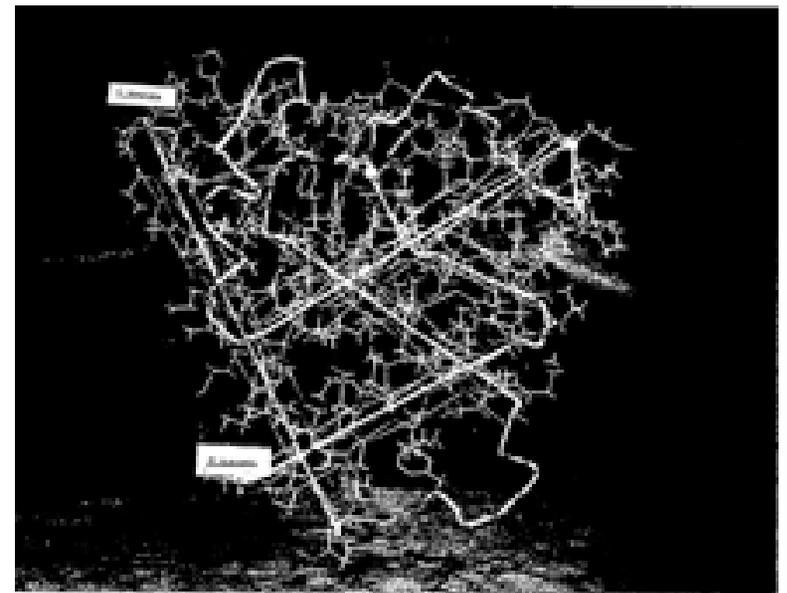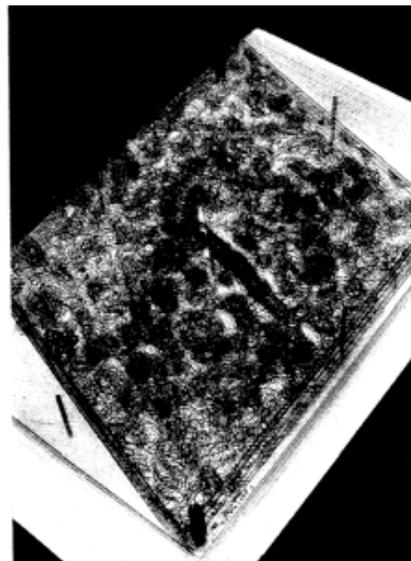
Myoglobin with heme group
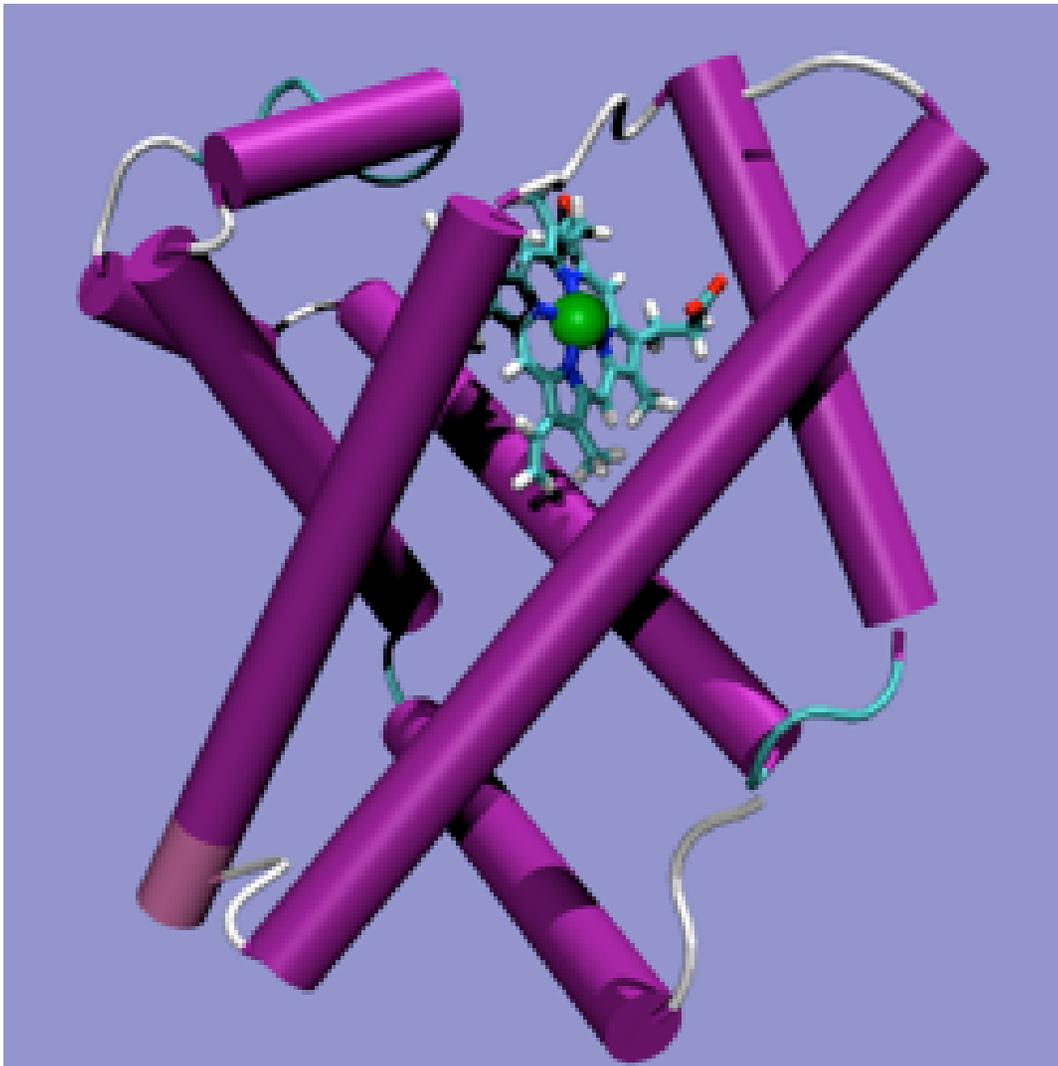
- Myoglobin from PDB structure 1A6M

- X-ray crystal structure at 1.00 A resolution.

- Steps seen in RMSD are due primarily to tilting of the helix to the upper right of the heme in the picture…

Myoglobin Dynamics to Probe Motion of Fe
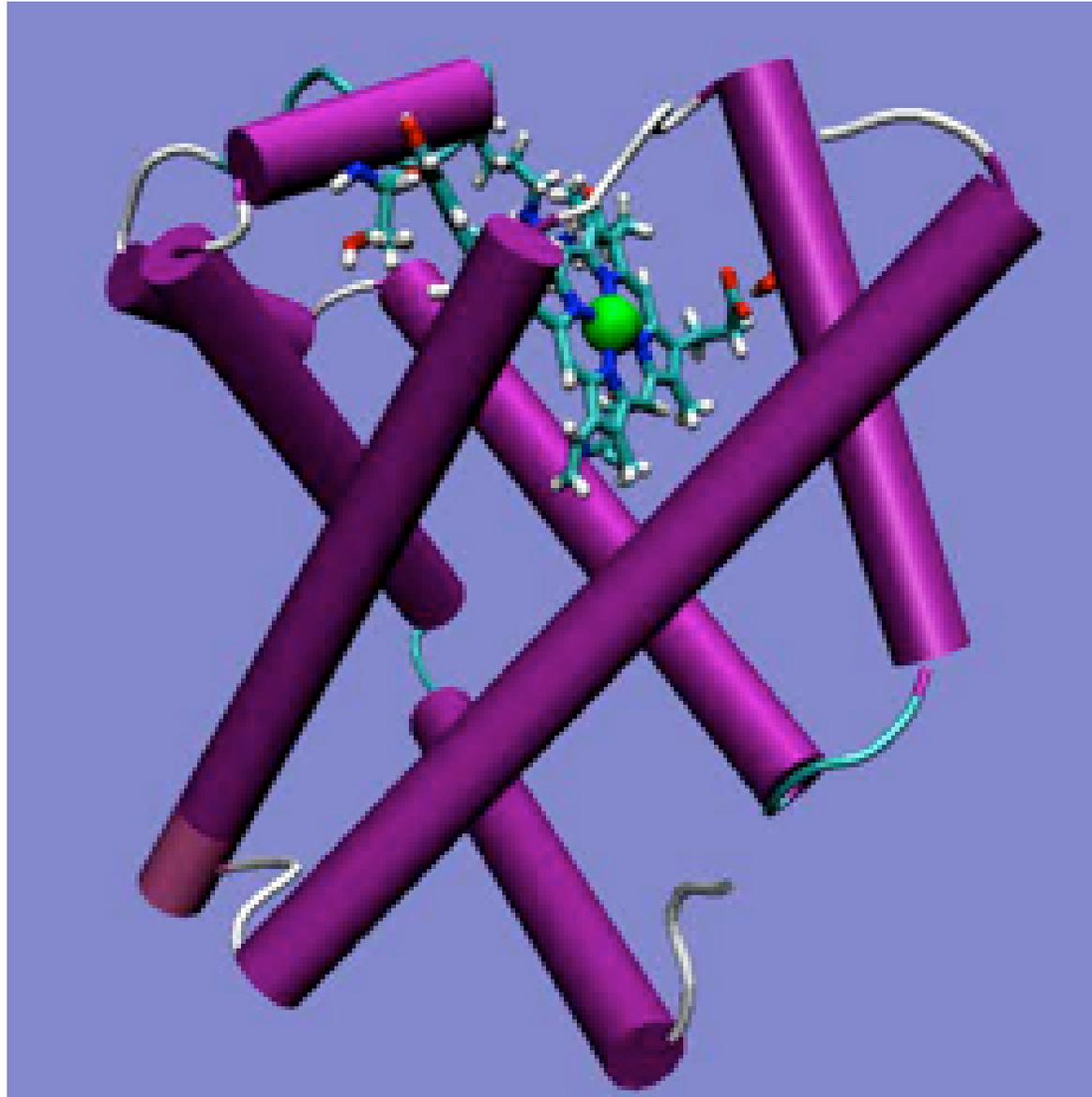
Setup and Equilibration

- Remove oxygen liganded to Fe
- Minimize 1000 steps, fixing the C_α atoms.
- Heat for 5 ps with Langevin dynamics at 300 K, fixed C_α atoms.
- Simulate in NVT ensemble for 19 ns, saving coordinates every ps.


short

long
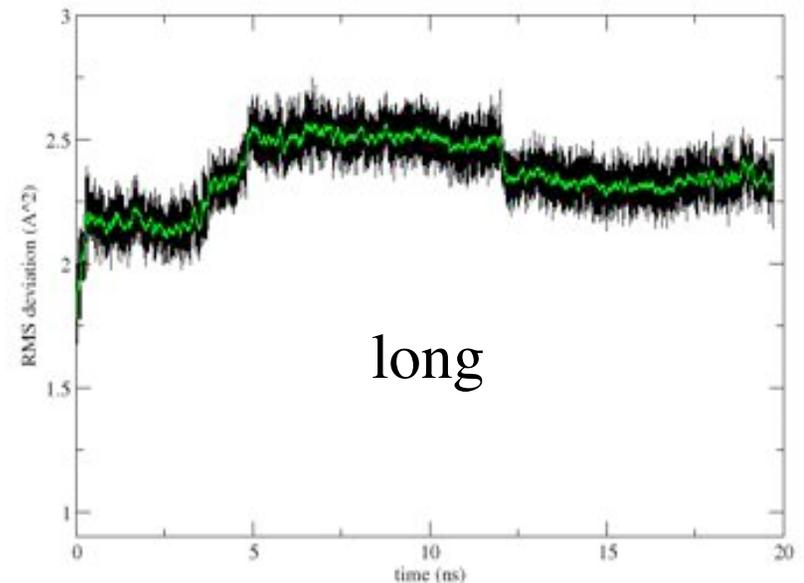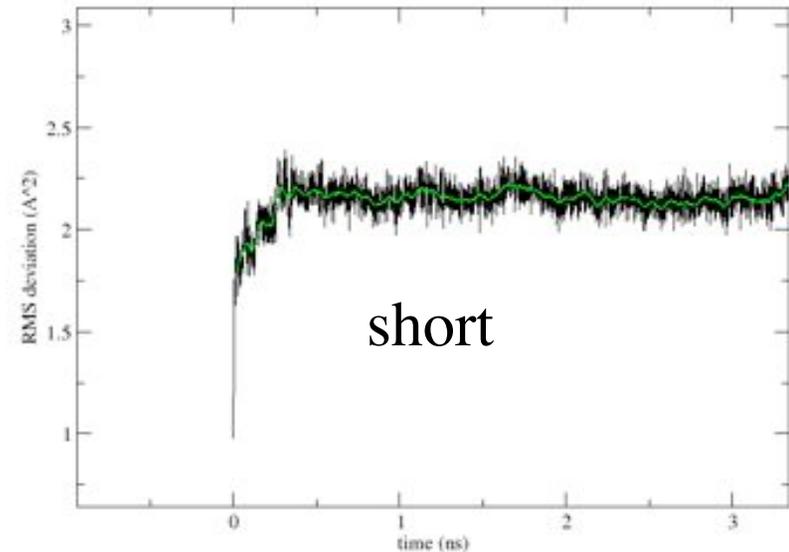
Obtain "f" from position distribution

- Best fit "by eye" is $\sigma = 0.528$ Å.

- However: standard deviation gives $\sigma = 0.36$ ($f = kT/\sigma^2$) = 319 pN/A; this is what we use below.

Fit: $\sigma = 0.528$ Å
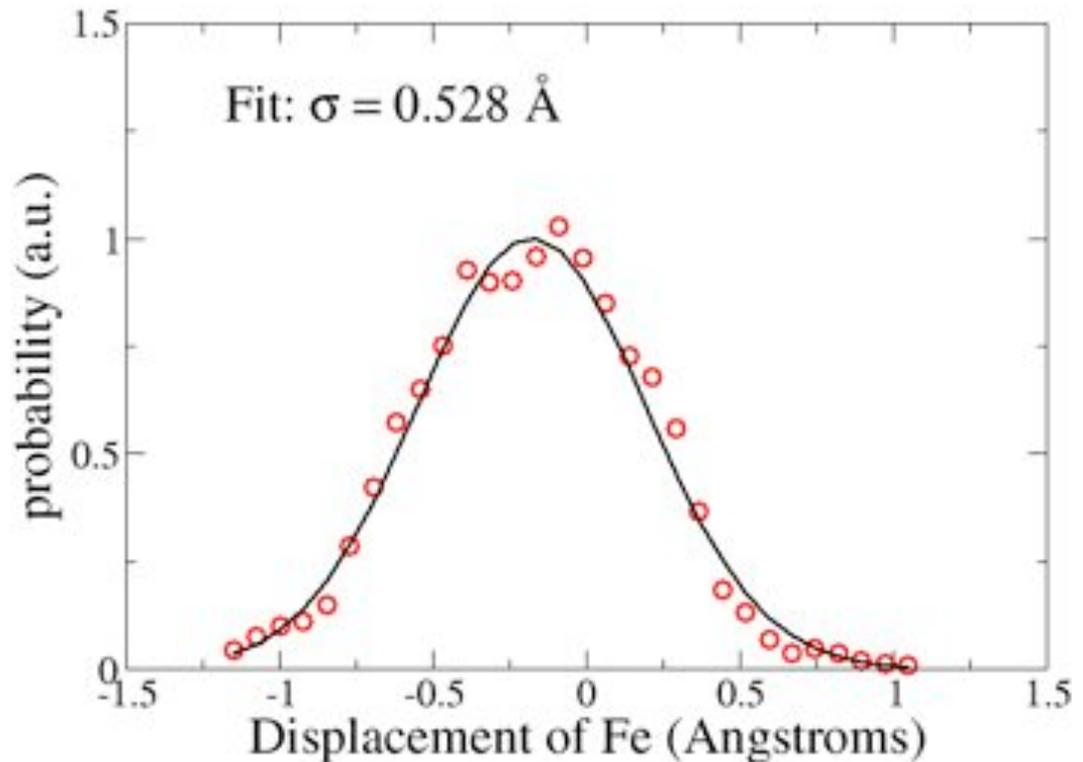
Obtain diffusion coefficient from position autocorrelation function

Once we know the restoring force, the diffusion coefficient can be obtained from the position autocorrelation function:

$$\langle x(t)x(0)\rangle = \frac{k_B T}{2f} \exp[-Dft/k_B T]$$

D f/kT = .0321
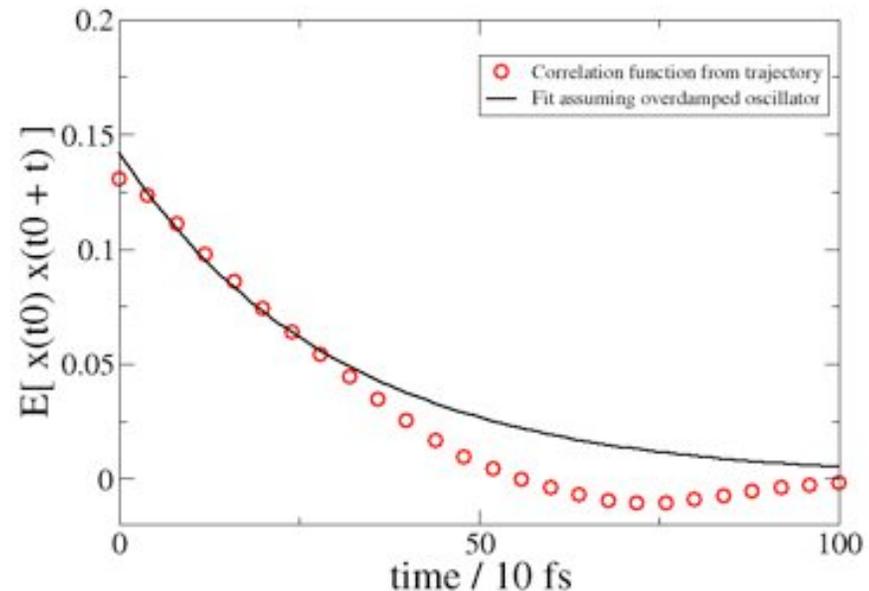
D = 0.0042 A^2/10 fs
   = 0.42 A^2 /ps.

Compare: D$_{water}$ = 0.24 A^2 /ps

Position autocorrelation: underdamped case

The Langevin equation governing underdamped motion is

$$\ddot{x} + b\dot{x} + \omega^2 x = \eta(t)$$

The position correlation function is given by

$$\langle x(t)x(0) \rangle = \langle x(0)^2 \rangle e^{-bt/2} \left[\cos \Omega(t) + \frac{b}{2\Omega} \sin \Omega(t) \right]$$

Using $\quad \Omega = \sqrt{\omega^2 + b^2/4}$

$$m = f/\omega^2 \quad \text{(from F = ma)}$$
$$\gamma = mb$$
$$D = k_B T/\gamma = \frac{k_B T}{f} \frac{\omega^2}{b}$$

we can solve for D.

Diffusion coefficient from underdamped oscillator
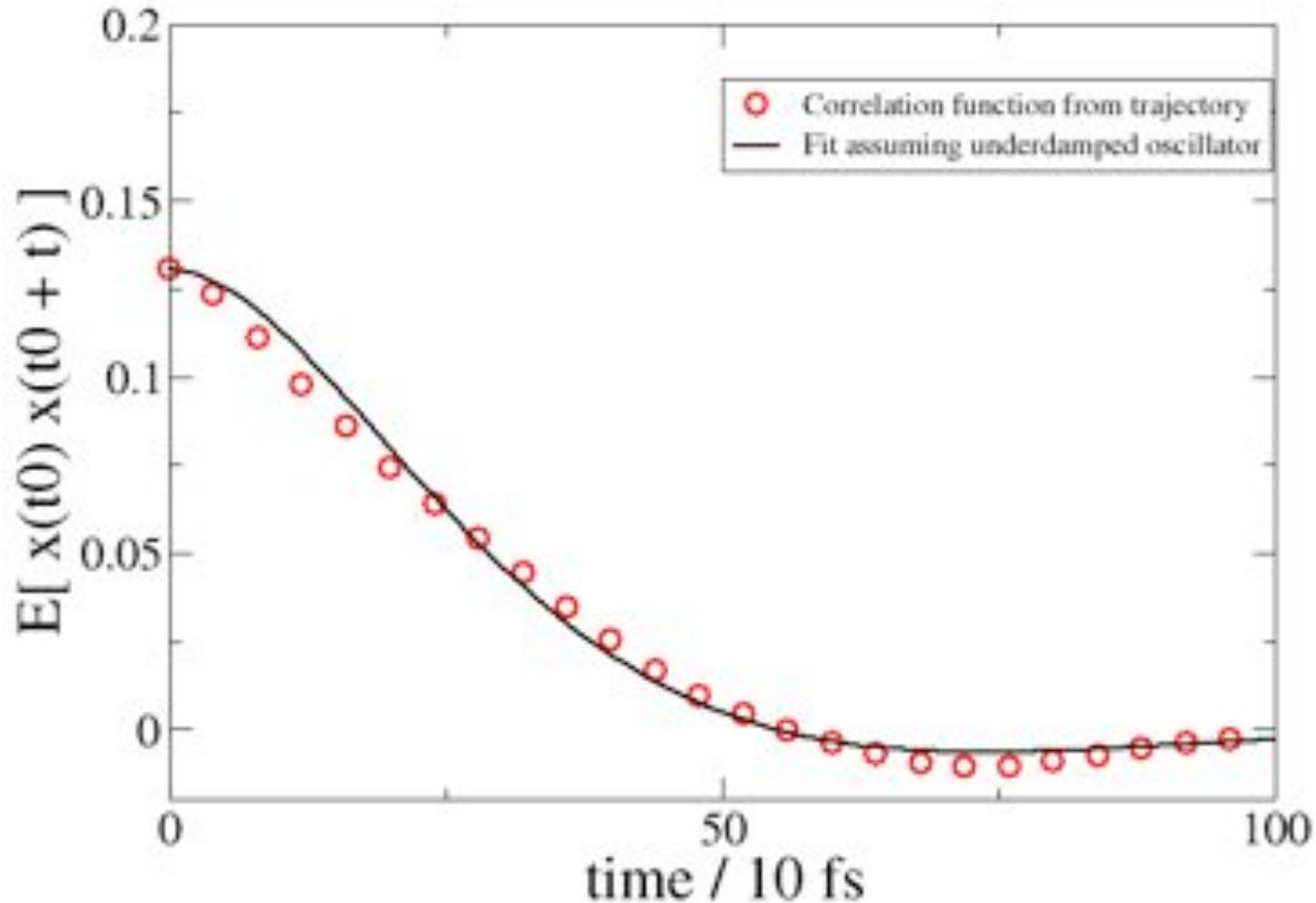
Fitting parameters: $\Omega = 0.0426$; $b = 0.0811$; $\omega^2 = \Omega^2 + b^2/4 = 34.59/ps^2$. $D = 0.556 \, A^2/ps$.

Mossbauer line shape function

The lineshape $I(\omega)$ we are trying to calculate is given by

$$I(\omega) = \frac{\sigma_0 \Gamma}{4} \int_{-\infty}^{\infty} dt\, e^{-i\omega t - \frac{1}{2}\Gamma|t|} G(\mathbf{k}, t) \qquad (1)$$

where $G(\mathbf{k}, t)$ is given by

$$G(\mathbf{k}, t) = \int d\mathbf{r} \int d\mathbf{r}_o e^{i\mathbf{k}\cdot(\mathbf{r}-\mathbf{r}_o)} p(\mathbf{r}, t|\mathbf{r}_o, 0) p_0(\mathbf{r}_o)$$

$$(2)$$

Notice that $G(\mathbf{k}, t)$ is just the autocorrelation function of $\exp(-i\mathbf{k}\cdot\mathbf{r})$.

Moessbauer Line Shape Function - Sampled and Matched to Analytical Formula

I(w) using D and f from position correlation

Positions sampled every 10 fs for 100 ps

$$\Sigma(\omega) = \int_{-\infty}^{+\infty} dt \, \langle \exp[i\vec{k}\cdot\vec{r}(t)] \, \exp[-i\vec{k}\cdot\vec{r}(0)]\rangle \, \exp[-i\omega t - \frac{1}{2}\Gamma|t|]$$

$$I(\omega) = \frac{1}{2}\sigma_o \Gamma \, \Re[\Sigma(\omega)]$$