Statistical Mechanics of Proteins

Ubiquitin

Show molecular dynamics trajectory in VMD

Equilibrium Properties of Proteins

Ubiquitin

Root Mean Squared Deviation: measure for equilibration and protein flexibility

$$RMSD_{\alpha} = \sqrt{\frac{1}{N} \sum_{i=1}^{N} \sum_{\alpha=1}^{N} (\rho(\alpha, t_i) - \langle \rho \rangle)^2}$$

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

---

**Temperature Dependence of Crystal Diffraction (Debye-Waller factor)**

Bragg’s law

\[ 2d \sin \theta = \lambda \]

structure factor

\[ f_j \exp[-is \cdot \vec{r}_j] \]

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:

\[ \langle f_j \exp[-is \cdot \vec{r}_j] \rangle = f_j \exp[-is \cdot \vec{x}_j] \langle \exp[-is \cdot \vec{u}_j] \rangle \]
Thermal Motion of Ubiquitin from MD

Temperature Dependence of Crystal Diffraction (Debye-Waller factor)

One can expand:
\[ \langle \exp[-i \vec{s} \cdot \vec{u}_j] \rangle = 1 - i \sum_{\vec{u}_j} \frac{(\vec{s} \cdot \vec{u}_j) - \frac{1}{2} (\vec{s} \cdot \vec{u}_j)^2}{\frac{(\vec{s} \cdot \vec{u}_j)^2}{3}} + \ldots \]

Spatial average:
\[ \langle (\vec{s} \cdot \vec{u}_j)^2 \rangle = \frac{1}{3} s^2 \langle u_j^2 \rangle \]

One can carry out the expansion further and show
\[ \langle \exp[-i \vec{s} \cdot \vec{u}_j] \rangle = \exp \left[ -\frac{1}{6} s^2 \langle u_j^2 \rangle \right] \]

Using for the thermal amplitude of the harmonic oscillator
\[ \frac{1}{2} m \omega^2 u_j^2 = \frac{3}{2} k_B T \]

one obtains
\[ \langle f_j \exp[-i \vec{s} \cdot \vec{r}_j] \rangle = f_j \exp[-s^2 k_B T / 2 m \omega^2] \exp[-i \vec{s} \cdot \vec{x}_j] \]

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)

Potential energy (not all quadratic)
Equilibrium Properties of Proteins

Energies: kinetic and potential

\[ \langle \sum_j \frac{1}{2} m_j \dot{v}_j^2 \rangle = \frac{3}{2} N k_B T \]

Kinetic energy (quadratic)

\[ U(\vec{r}) = \sum \frac{k_i^b (r_i - r_0)^2}{2} + \sum \frac{k_i^{\alpha\beta} (\alpha_i - \alpha_0)^2}{2} + \sum \frac{k_i^a [1 + \cos (n_i \phi_i + \delta_i)]}{2} + \sum \frac{9 \sigma_i^2}{2} \]

Potential energy (not all quadratic)

Maxwell Distribution of Atomic Velocities

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

\[ \sigma = x, y, z \]
Analysis of $E_{\text{kin}}$, $T$ (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

Maxwell distribution

$$dP(\epsilon_a) = e \exp(-\epsilon_a/2k_B T) d\epsilon_a$$

Individual kinetic energy $\epsilon_a = m v_a^2/2$

$$dP(\epsilon_a) = (\pi m k_B T)^{1/4} \exp(-\epsilon_a/k_B T_0) d\epsilon_a$$

One can derive

$$\langle \epsilon_a \rangle = T_0/2$$

$$\langle \epsilon_a^2 \rangle = 3 T_0^2/4$$

$$\langle \epsilon_a^4 \rangle - \langle \epsilon_a^2 \rangle^2 = T_0^2/2$$

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

$$P(E_{\text{kin}}) = e \exp\left(\frac{-\left(E_{\text{kin}} - \langle E_{\text{kin}} \rangle\right)^2}{2 \left(\frac{\sqrt{\sigma^2}}{\sqrt{2}}\right)}\right)$$

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

$$P(\Delta T) = e \exp(-\Delta T^2/2\sigma^2)$$

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!
**Normal Distribution of Temperatures**

- \( \mu = 297.7 \text{ K} \)
- \( \sigma = 2.4 \text{ K} \)

Another example, ubiquitin in water bath

\[
\sigma^2 = \frac{2T^2}{3N}
\]

**Specific Heat of a Protein**

Total energy of ubiquitin (NVE ensemble)

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

\( c_V \text{ ubiquitin} = 2.24407 \times 10^{-7} \text{ kcal/molK} \)
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

\[ m \ddot{r} = F_{FF} + F_H + F_f + F_L \]

- \( F_{FF} \rightarrow \) force field
- \( F_H \rightarrow \) harmonic restrain
- \( F_f \rightarrow \) friction
- \( F_L \rightarrow \) Langevin force

Atoms in the inner region follow Newtonian dynamics

\[ m \ddot{r} = F_{FF} \]
Solution of the Heat Equation

<table>
<thead>
<tr>
<th>t</th>
<th>T_{min}</th>
<th>t</th>
<th>T_{min}</th>
<th>t</th>
<th>T_{min}</th>
<th>t</th>
<th>T_{min}</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.05</td>
<td>266.50</td>
<td>1.05</td>
<td>276.00</td>
<td>1.95</td>
<td>267.00</td>
<td>3.25</td>
<td>261.00</td>
</tr>
<tr>
<td>0.15</td>
<td>266.50</td>
<td>1.15</td>
<td>276.50</td>
<td>2.05</td>
<td>268.50</td>
<td>3.45</td>
<td>258.50</td>
</tr>
<tr>
<td>0.35</td>
<td>265.50</td>
<td>1.25</td>
<td>275.50</td>
<td>2.25</td>
<td>266.50</td>
<td>3.55</td>
<td>258.50</td>
</tr>
<tr>
<td>0.55</td>
<td>262.50</td>
<td>1.35</td>
<td>271.00</td>
<td>2.35</td>
<td>264.50</td>
<td>3.95</td>
<td>256.50</td>
</tr>
<tr>
<td>0.65</td>
<td>262.75</td>
<td>1.45</td>
<td>271.25</td>
<td>2.55</td>
<td>265.00</td>
<td>4.05</td>
<td>257.25</td>
</tr>
<tr>
<td>0.75</td>
<td>270.00</td>
<td>1.65</td>
<td>268.50</td>
<td>2.85</td>
<td>264.50</td>
<td>4.45</td>
<td>254.00</td>
</tr>
<tr>
<td>0.85</td>
<td>277.50</td>
<td>1.85</td>
<td>268.00</td>
<td>2.85</td>
<td>262.00</td>
<td>4.85</td>
<td>255.25</td>
</tr>
<tr>
<td>1.00</td>
<td>277.50</td>
<td>1.85</td>
<td>268.00</td>
<td>3.05</td>
<td>262.50</td>
<td>4.85</td>
<td>252.00</td>
</tr>
</tbody>
</table>

Table 1: Mean temperature ($T_{min}$) [K] of the protein as a function of time $t$ [ps].

---

Heat Conduction Equation

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

- thermal diffusion coefficient $D = \frac{K}{\rho c}$
- approximate the protein with a homogeneous sphere of radius $R \sim 20$ Å
- calculate $T(r,t)$ assuming initial and boundary conditions:
  \[
t(r,0) = T_0 \quad \text{for} \quad r < R \\
T(R,t) = T_{\text{bath}}
\]
Solution of the Heat Equation

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

\[ D = \frac{K}{\rho c} \]

Initial condition

\[ T(r, 0) = \langle T_{sim}(0) \rangle \quad \text{for} \quad r < R, \]

Boundary condition

\[ T(R, t) = T_{bath} \]

Solution of the Heat Equation

Spherical symmetry

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

We assume

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

Here \( u_n \) are the eigenfunctions of the spherical diffusion operator

\[ L = \frac{D}{r \frac{d^2}{dr^2}} \]

\[ \frac{D}{r \frac{d^2}{dr^2}} r u_n(r) = \lambda_n u_n(r), \quad u_n(0) = \text{finite}, \quad u_n(R) = 0 \]
Solution of the Heat Equation

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

\[ a_n = \frac{2R}{n\pi} \Delta T (-1)^{m+1} \]

\[ \frac{\Delta T(t, r)}{\Delta T(0, r)} = \frac{1}{2} \int_0^1 r^2 dr \sin \left( \frac{n\pi r}{R} \right) \]

Temperature averaged over volume

\[ \langle T \rangle(t) = \frac{\int d^3 r T(r, t)}{3} = \frac{3}{R^3} \int_0^R r^2 dr T(r, t) \]

\[ = T_{\text{bath}} + \sum_{n=1}^{\infty} a_n \exp \left[ -\frac{(n\pi)^2}{R^2} \right] D t \int_0^R r dr \sin \left( \frac{n\pi r}{R} \right) \]

\[ = T_{\text{bath}} + \frac{6\Delta T}{\pi^2} \sum_{n=1}^{\infty} \frac{1}{n^2} \exp \left[ -\frac{(n\pi)^2}{R^2} \right] D t \]

\[ D \approx 0.38 \times 10^{-3} \text{cm}^2\text{s}^{-1} \]

Water: \[ 1.4 \times 10^{-3} \text{cm}^2\text{s}^{-1} \]