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

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

NMR structures aligned together to see flexibility

MD simulation

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

Protein sequence exhibits characteristic permanent flexibility!
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

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

structure factor

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

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:

\[ \langle f_j \exp[-i \vec{s} \cdot \vec{r}_j] \rangle = f_j \exp[-i \vec{s} \cdot \vec{x}_j] \langle \exp[-i \vec{s} \cdot \vec{u}_j] \rangle \]
Spatial average for harmonic oscillator: \( \langle (\vec{s} \cdot \vec{u}_j)^2 \rangle = \frac{1}{3}s^2 \langle u_j^2 \rangle \)

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

\[
\frac{1}{2}m\omega^2 u_j^2 = \frac{3}{2}k_BT
\]

one obtains

\[
\langle f_j \exp[-i\vec{s} \cdot \vec{r}_j] \rangle = f_j \exp[-s^2k_BT/2m\omega^2] \exp[-i\vec{s} \cdot \vec{x}_j]
\]
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)

Temperature dependence
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)

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

Energy vs. Temperature [K]

- Kinetic
- Coulomb
- VdW
- Bond
- Angle
- Total
Maxwell Distribution of Atomic Velocities

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

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

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

**Definition of Temperature**

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

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

The atomic velocities of a protein establish a thermometer, but is it accurate?
The atomic velocity thermometer is inaccurate due to the finite size of a protein!
Normal Distribution of Temperatures

Another example, ubiquitin in water bath

$$\sigma^2 = \frac{2T^2}{3N}$$
Specific Heat of a Protein

Total energy of ubiquitin (NVE ensemble)

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

\[ C_{V(\text{ubiquitin})} = 2.24407 \times 10^{-3} \text{ kcal/molxK} \]
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.
  - \( m \ddot{r} = F_{FF} + F_H + F_f + F_L \)
  - \( F_{FF} \rightarrow \text{force field} \)
  - \( F_H \rightarrow \text{harmonic restrain} \)
  - \( F_f \rightarrow \text{friction} \)
  - \( F_L \rightarrow \text{Langevin force} \)

- **Atoms in the inner region follow Newtonian dynamics:**
  - \( m \ddot{r} = F_{FF} \)
Simulated Cooling - Result

<table>
<thead>
<tr>
<th>$t$</th>
<th>$\langle T_{sim} \rangle$</th>
<th>$t$</th>
<th>$\langle T_{sim} \rangle$</th>
<th>$t$</th>
<th>$\langle T_{sim} \rangle$</th>
<th>$t$</th>
<th>$\langle T_{sim} \rangle$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.05</td>
<td>298.75</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>289.25</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>285.50</td>
<td>1.25</td>
<td>275.25</td>
<td>2.25</td>
<td>266.50</td>
<td>3.55</td>
<td>259.50</td>
</tr>
<tr>
<td>0.55</td>
<td>282.25</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>282.75</td>
<td>1.45</td>
<td>271.75</td>
<td>2.55</td>
<td>263.50</td>
<td>4.05</td>
<td>257.25</td>
</tr>
<tr>
<td>0.75</td>
<td>279.00</td>
<td>1.65</td>
<td>269.50</td>
<td>2.65</td>
<td>264.50</td>
<td>4.45</td>
<td>254.00</td>
</tr>
<tr>
<td>0.85</td>
<td>277.75</td>
<td>1.75</td>
<td>271.00</td>
<td>2.85</td>
<td>262.00</td>
<td>4.55</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 $\langle T_{sim} \rangle$ [K] of the protein as a function of time $t$ [ps].

Result from simulation
Heat Conduction Equation

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

- thermal diffusion coefficient
- mass density
- specific heat
- thermal conductivity

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

- approximate the protein with a homogeneous sphere of radius \( R \approx 20 \, \text{Å} \)
- 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

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

Boundary condition

\[ T(R, t) = T_{bath} . \]
Solution of the Heat Equation

Spherical symmetry

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

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

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

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

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

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

Graph showing the temperature distribution with time and radial distance.
Solution of the Heat Equation

Temperature averaged over volume

\[
\langle T \rangle(t) = \left( \frac{4\pi R^3}{3} \right)^{-1} \int d^3 r T(r, t) = \frac{3}{R^3} \int_0^R r^2 dr T(r, t)
\]

\[
= T_{bath} + \sum_{n=1}^{\infty} a_n \exp \left[ - \left( \frac{n\pi}{R} \right)^2 D t \right] \frac{3}{R^3} \int_0^R r dr \sin \left( \frac{n\pi r}{R} \right)
\]

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

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

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