Biomolecules as polymers

– at least two of the primary classes of biomolecules are polymers (proteins and DNA/RNA, often saccharides as well)

– physicists have developed a rich set of theories for polymers

– How can we use them to gain new biological understanding?
Do atoms always matter?

For understanding the function of an enzyme, we need an atomic-detailed structural description.

The bacterial cell wall (PBP’s substrate) is often best described from a statistical (average) point of view.
Random walks

A simple model of a polymer is a random walk in space

Fixed segments of equal length are connected at hinges, can go in any direction

length of each segment is defined as the Kuhn length \( (a) \)

also known as the Freely Jointed Chain (FJC) model

DNA as a random walk
Measurable properties

What do we want to measure for these random walks?

“random” implies only average properties are useful, e.g., $<R>$, $<R^2>$, $p(R,N)$

Apply the tools of statistical mechanics to calculate overall length (macrostate) from its underlying microstates

counting # states ($W$) for each $n_r$ or $n_l$
Probability distributions

Binomial distribution

\[ P(n_r; N) = \frac{N!}{n_r!(N - n_r)!} \left( \frac{1}{2} \right)^N \]

Gaussian distribution

\[ p(R; N) = \frac{2}{\sqrt{2\pi N}} e^{-R^2/2Na^2} \]

What is \( P(R; N) \)?

\[ R = (n_r - n_l) a \quad \text{end-end distance} \]
Probability distributions

\[ p(R; N) = \frac{2}{\sqrt{2\pi N}} e^{-R^2/2Na^2} \]

Normalization \((\int_{-\infty}^{\infty} p(R; N) dR = 1)\) implies that

\[ P(R; N) = \frac{1}{\sqrt{2\pi Na^2}} e^{-R^2/2Na^2} \]

and in 3D:

\[ P(R; N) = \left(\frac{3}{2\pi Na^2}\right)^{3/2} e^{-3R^2/2Na^2} \]

Distribution is **Gaussian** with mean \(<R> = 0\) and variance \(<R^2> = Na^2\)
Central limit theorem

For a set of $N$ random variables \( \{x_n\} \) with finite mean and variance (e.g., the individual segments of a polymer), the sum $X = x_1 + x_2 + \ldots + x_N$ (e.g., the end-to-end distance) will tend towards a **Gaussian distribution** regardless of the distribution of $x_n$.

\[
P(R; N) = \frac{1}{\sqrt{2\pi Na^2}} e^{-R^2 / 2Na^2}
\]

sum of $n$ dice throws
Persistence length ($L_p$) is defined microscopically by the correlation between the directions of successive segments of our chain (from Flory\(^1\):)

$$L_p = \lim_{N \to \infty} \langle \sum_{j \geq i} \vec{a}_i \cdot \frac{\vec{a}_j}{a} \rangle$$

$$L_p = \frac{\langle R^2 \rangle + Na^2}{2Na} \quad \langle R^2 \rangle = Na^2$$

Therefore $L_p = a$ (the Kuhn length)

The random walk model we just developed is also known as the Freely Jointed Chain model

---

Persistence length (continuous)

Persistence length ($L_p$) is defined microscopically by the correlation between the directions of successive segments of our chain:

$$\langle \vec{t}(s) \cdot \vec{t}(u) \rangle = e^{-|s-u|/L_p}$$

$$\vec{R} = \int_0^L ds \, \vec{t}(s)$$

$$\langle \vec{R}^2 \rangle = \langle \int_0^L ds \, \vec{t}(s) \cdot \int_0^L du \, \vec{t}(u) \rangle$$

$$= 2 \int_0^L ds \int_s^L du \, e^{-|u-s|/L_p} \approx 2LL_p = Na^2 = aL$$

Therefore $L_p = a/2$ (a is the Kuhn length)

This is the start of the worm-like chain model

$$a = \langle R^2 \rangle / R_{\text{max}}$$ in the continuous model
Radius of gyration

\[ \langle R_G^2 \rangle = \frac{1}{N} \sum_{i=1}^{N} \langle (\vec{R}_i - \vec{R}_{CM})^2 \rangle \]

Radius of gyration \((R_G)\): average distance between links and center-of-mass

\[ \vec{R}_{CM} = \frac{1}{N} \sum_{i=1}^{N} \vec{R}_i \]

\[ \sqrt{\langle R_G^2 \rangle} = \sqrt{\frac{LL_p}{3}} \]

E. coli genome is \(~ 4.6 \times 10^6\) base pairs

\[ \sqrt{\langle R_G^2 \rangle} \approx 5 \mu m \]

About 2x as big as seen in the picture, but it has some constraints (e.g., circular, so it has to return to the cell body often)

Clearly strong forces are necessary to pack it in the cell!
Entropic force

Recall that our free energy $G = U - TS + PV$ (Gibbs)

A change in energy (e.g., from applied tension) generates a restoring force, $F_{\text{restore}} = -\frac{\partial G}{\partial L}$

$U, T, P$ are constant for freely-jointed chain, making

$$F = -T\frac{\partial S}{\partial L}$$

Even though the chain does not store internal energy, it still can exert a force as an entropic spring

$$F = kTL/Na^2$$
Improvements to the FJC model

FJC model (fixed segments)

Worm-like chain (elastic, with an energetic bending cost)

modified WLC, now with a stretching cost as well


“Molecular origin of the hierarchical elasticity of titin: simulation, experiment and theory.”
Worm-like chain (WLC) model

Our polymer is now characterized by a unit tangent vector $\mathbf{t}(s)$, where $s$ is the position along the chain.

**How much energy does it cost to bend the chain?**

$$W(\epsilon) = \frac{1}{2} E \epsilon^2 = \frac{1}{2} E \left( \frac{\Delta L}{L_0} \right)^2$$

$E$ is Young’s modulus (units of force/area)

$\epsilon$ is strain; $W(\epsilon)$ is strain energy density (energy/volume)

One way of visualizing the strain

neutral axis
Worm-like chain (WLC) model

\[ \theta = \frac{s}{(R+z)} = \frac{L_0}{R} \]

\[ \Delta L = s - L_0 = (R+z)\frac{L_0}{R} - L_0 = (z/R)\frac{L_0}{R} \]

\[ W(\epsilon) = \frac{1}{2} E \left( \frac{\Delta L}{L_0} \right)^2 = \frac{1}{2} E \left( \frac{z}{R} \right)^2 \]

\[ E_b = \int W(\epsilon) \, dV = L_0 \int \frac{E}{2R^2} z^2 \, dA \]

\[ = \frac{E L_0}{2R^2} \int z^2 \, dA = \frac{E I L_0}{2R^2} \]

\( I \) is called the “geometric moment” (don’t have to calculate it!)

For a circle?

\[ E_b = \frac{E I L_0}{2R^2} = \frac{E I (2\pi R)}{2R^2} = \frac{\pi E I}{R} \]
Persistence length

The WLC model represents a balance between internal energy (from the resistance to bending) and entropy.

How does $L_p$ vary with temperature?

$$E_b = \frac{EI L}{2R^2} = \frac{EI L}{2} \left( \frac{\theta^2}{s^2} \right)$$

Let $L = s$ for a short segment.

$$s = \theta R$$

$$= \frac{EI}{2s} \theta^2$$
Persistence length

\[ \langle \vec{t}(s) \cdot \vec{t}(u) \rangle = e^{-|s-u|/L_p} \]
\[ \langle \cos(\theta) \rangle = e^{-s/L_p} \]

Assume \( t(u) = t(0) = z \) axis

\[ \langle \cos(\theta) \rangle \approx \langle 1 - \frac{\theta^2}{2} \rangle = 1 - \frac{1}{2}\langle \theta^2 \rangle \]

\[ e^{-s/L_p} \approx 1 - s/L_p \]

\[ \langle E_b \rangle = \frac{EI}{2s}\langle \theta^2 \rangle = 2 \left( \frac{kT}{2} \right) = kT \]

2 DoF (one constrained by total length)

\[ 1 - s/L_p = 1 - \frac{1}{2}\langle \theta^2 \rangle = 1 - \frac{1}{2} \left( \frac{2skT}{EI} \right) \]

\[ L_p = \frac{EI}{kT} \]

\( L_p \) increases with increasing Young's modulus (stiffer)

\( L_p \) decreases with increasing temperature (more flexible)
Force-extension relationship for WLC

\[ E_{\text{bend}} = \frac{EI}{2} \int_0^L \frac{ds}{R(s)^2} = \frac{kTL_p}{2} \int_0^L \left| \frac{dt}{ds} \right|^2 ds \]

\( R \) is constant

don’t assume

partition function is summed over all possible curves (a path integral)

\[ Z = \int Dt(s) e^{-E_{\text{bend}}/kT} = \int Dt(s) \exp\left(-\frac{L_p}{2} \int_0^L \left| \frac{dt}{ds} \right|^2 ds \right) \]

applying a force adds a term to the energy:

\[ E_{\text{app.}} = -Fz = -F \int_0^L t_z ds \]

\[ \langle z \rangle = \frac{1}{Z(F)} \int Dt(s) z\left[e^{-\left(E_{\text{bend}} + E_{\text{app}}(F)\right)/kT}\right] \]

Good luck calculating this!

\[ = kT \frac{d \ln Z(F)}{dF} \]
A closed form solution for $\langle z \rangle$ does not exist!

$$\frac{F L_p}{kT} \approx \frac{z}{L} + \frac{1}{4(1 - z/L)^2} - \frac{1}{4}$$

High-force limit

Low-force limit

deviates by up to 10% from numerical solution (not shown)

Force spectroscopy

- Different force profiles are like molecular signatures.
- Deviations from WLC imply a change in structure.
Single-molecule techniques

AFM

optical tweezers

magnetic tweezers

pipette-based

Figure 8.21  Physical Biology of the Cell, 2ed. (© Garland Science 2013)
Atomic-force microscopy

unfolding of titin kinase + Ig domains

fits from WLC model

20-ns SMD Simulation of **fibrinogen**, 1.06 million atoms


A Blood Clot
Red blood cells within a network of fibrin, composed of polymerized fibrinogen molecules.
High-speed AFM show agreement with (relatively) slow simulations

optical tweezers

unfocused laser

- Gradient induces lateral force

focused laser

- Gradient induces axial force