BioPolymer Statistics I

Basic Theory
DNA  RNA  proteins  lipids  polysaccharides
The properties of bio polymers depend on:

- Chemical composition of the monomer
- Degree of polymerization
- Flexibility
- Architecture
- Homo- vs. hetero- polymer
ARCHITECTURE

(a)  (b)  (c)  (d)  

(e)  (f)  (g)  (h)  

BioPolymers
a disordered/unfolded protein or single-stranded nucleic acid is very dynamic and explores many different conformations.
An ideal chain has **no net interactions** between two elements $A_i$ and $A_j$. 
How can we describe the conformations of an ideal chain?
CONFORMATIONS

HOW CAN WE DESCRIBE THE CONFORMATIONS OF AN IDEAL CHAIN?

WE NEED TO INTRODUCE SOME USEFUL QUANTITIES:

- END-TO-END VECTOR
- RADIUS OF GYRATION
- ASPHERICITY AND SHAPE FACTOR
- DISTANCE DISTRIBUTION
ideal chain

Since a polymer adopts many conformations, we are interested in average properties!

\[ \langle \ldots \rangle : \text{we are going to consider average over all conformations in the ensemble} \]
**end-to-end vector**

The **end-to-end vector** is the sum of all bond vectors $r_m$ in the chain.

$$
R_m = \sum_{i=1}^{m} r_i
$$

**End to End Vector**
end-to-end vector

Since there is no preferred direction for an ideal chain, the mean end-to-end vector is equal to zero:

$$\langle \vec{R}_M \rangle = 0 \quad (\sum_{i=1}^{\infty} \langle \vec{v}_i \rangle)$$

- $\vec{R}_M$ has the same probability of $+\vec{R}_M$
end-to-end vector

What about the "mean-square end-to-end distance"?

\[ \langle R^2 \rangle \equiv \langle R_m^2 \rangle = \langle \vec{R}_m \cdot \vec{R}_m \rangle = \]

assuming the vector bond has constant length

\[ |\vec{\nu}_i| = b \]

\[ = n b^2 + b^2 \sum_{i=1}^{m} \sum_{j \neq i} \langle \cos \theta_{ij} \rangle \]
\[
\langle R^2 \rangle = \langle R^2_m \rangle = \langle \mathbf{R}_m \cdot \mathbf{R}_m \rangle = \\
= \langle (\sum_{i=1}^{m} \mathbf{r}_i) (\sum_{j=1}^{m} \mathbf{r}_j) \rangle = \\
= \langle \sum_{i=1}^{m} \mathbf{r}_i \cdot \mathbf{r}_i + \sum_{i=1 \neq j}^{m} \mathbf{r}_i \cdot \mathbf{r}_j \rangle = \text{MEAN OF THE SUM IS THE SUM OF THE MEANS} \\
= \langle \sum_{i=1}^{m} |\mathbf{r}_i|^2 \rangle + \langle \sum_{i=1 \neq j}^{m} \mathbf{r}_i \cdot \mathbf{r}_j \rangle = \text{ASSUMING THE VECTOR BOND HAS CONSTANT LENGTH} \\
|\mathbf{r}_i| = b \\
= \langle \sum_{i=1}^{m} b^2 \rangle + \langle \sum_{i=1 \neq j}^{m} b^2 \cos \theta_{ij} \rangle = \\
= mb^2 + b^2 \sum_{i=1 \neq j}^{m} \langle \cos \theta_{ij} \rangle \\
\]
\[ |r_i| = b \quad \text{fixed} \]

\[ \theta, \phi \quad \text{free} \]
\[ \langle R^2 \rangle = m b^2 + b^2 \sum_{i=1}^{n} \sum_{j \neq i} \langle \cos \theta_{ij} \rangle \]

For a "freely jointed chain" there is no correlation between the direction of different bonds, so \( \langle \cos \theta_{ij} \rangle = 0 \) and \( \langle R^2 \rangle = m b^2 \)
Flory’s characteristic ratio

In a typical polymer, \( \langle \cos \theta_{ij} \rangle \neq 0 \)
and only when \( \lim_{|i-j| \to \infty} \langle \cos \theta_{ij} \rangle = 0 \).

In this case:

\[
\langle R^2 \rangle = b^2 \sum_{i=1}^{m} \sum_{j=1}^{m} \langle \cos \theta_{ij} \rangle = b^2 \sum_{i=1}^{m} C_i = C_m \, m \, b^2
\]

\[
C_m = \frac{1}{m} \sum_{i=1}^{m} C_i
\]

Flory’s characteristic ratio
Flory’s characteristic ratio

For sufficiently long chains

\[
\langle R^2 \rangle \approx C_\infty M b^2
\]
Let's define a different segment:

\[
N_k b_k = l_c
\]

Number of Kuhn segments \( N_k \)

Length of Kuhn segment \( b_k \)

Contour length \( l_c \)

\[
\begin{align*}
\langle R^2 \rangle &= N_k b_k^2 = C_0 m b^2 \\
N_k b_k &= l_c = m b
\end{align*}
\]

Therefore:

\[
N_k = \frac{m^2 b^2}{C_0 m b^2} = \frac{l_c^2}{C_0 m b^2}
\]

\[
b_k = \frac{C_0 m b^2}{l_c}
\]
How to use Kuhn segments?

For a given system we usually know the contour length and we can access the end-to-end distance.

\[ \langle R^2 \rangle, \ l_c \ known \Rightarrow \ \begin{cases} \ l_c = Nk \ bk \\ \langle R^2 \rangle = Nk \ bk^2 \end{cases} \]

\[ bk = \frac{\langle R^2 \rangle}{l_c} \]

\[ Nk = \frac{l_c}{bk} = \frac{l_c^2}{\langle R^2 \rangle} \]
Kuhn segment

* How to use Kuhn segments?

Any segment of an ideal chain $R_{ij}$ for $|i-j| \gg \sigma$ is "ideal" and follows the same statistics.

$\langle R_{ij}^2 \rangle$ unknown / $\langle R^2 \rangle$, known

$$\frac{\langle R_{ij}^2 \rangle}{\langle R^2 \rangle} = \frac{Nk^3 b^2}{Nk^2 b^2} = \frac{(i-j)^2}{\sigma^2}$$

0-end distance scales with number of
Kuhn segment

How to use Kuhn segments?

Any segment of an ideal chain $R_{ij}$ for $|i-j| \gg 0$ is "ideal" and follows the same statistics

\[
\frac{\langle R_{ij}^2 \rangle}{\langle R^2 \rangle} = \frac{NK^2}{NK} = \frac{(i-j)N}{mN} \quad * 
\]

The mean-square end-to-end distance scales with number of bonds.
Freely rotating chain

Bond length $b$
Bond angle $\theta$ (fixed)
Torsional angle $\phi$ (free)

$b \cos \theta \vec{r}_j$
$b \cos^2 \theta \vec{r}_{j-1}$
$b \cos^3 \theta \vec{r}_{j-2}$

$b \cos |j-i| \theta$
\[ \langle R^2 \rangle = \sum_{i=1}^{N} \sum_{j=1}^{N} \langle \vec{r}_i \cdot \vec{r}_j \rangle \]

\[ \langle \vec{r}_i \cdot \vec{r}_j \rangle = b^2 \cos \theta |i-j| \]

\[ e^{j\delta \log(\cos \theta)} = e^{-1 |i-j|} \]

After \( Sp \) segments the chain starts to lose memory of the previous configuration!

\[ [Sp \cdot b] \] defines the "persistance length".
\[ \langle R^2 \rangle = b^2 \sum_{i=1}^{M} \sum_{j=1}^{M} e^{-\frac{|i-j|b}{\ell_p}} \]

\[ \langle R^2 \rangle = 2\ell_p l c - 2\ell_p^2 \left(1 - e^{-\frac{l c}{\ell_p}}\right) \]
Freely rotating chain

\[<R^2> = 2l_p l_c - 2l_p^2 \left( 1 - e^{-l_c/l_p} \right)\]

**TWO LIMITS:**

1. \(l_c \gg l_p\)
   - Flexible chain
   - \(<R^2> = 2l_p l_c\)
   - \(= N_k b_K\)
   - \(= b_K N_k b_K\)
   - \(b_K = 2l_p\)

   *From Kuhn segment we can compute the persistence length and vice versa*

2. \(l_p \gg l_c\)
   - Rod-like
   - Taylor series of \(e^{-x}\)
   - \(= 1 - x + \frac{1}{2} x^2 - \frac{1}{6} x^3 + \ldots\)
   - \(<R^2> = l_c^2 - \frac{l_c^3}{3l_p} + \ldots\)