

Homework 3

Jeffrey Moffitt

February 25, 2008

■ Problem 1 (4 points)

Recall that the probability of finding k of the units in a helix for a polymer of N units is

$$p(k) = \frac{(N-k+1)\sigma s^k}{Z}$$

where Z is the partition function

$$Z = 1 + \sigma s \frac{(s^{N+1} - (N+1)s + N)}{(s-1)^2}$$

(Note: this is different than the expression in the notes. This is correct, the one in the notes has a typo)

Now to calculate the equilibrium concentration, recall that the concentration of a given species is equal to the total concentration times the probability of having that species. For example, in the reaction $A \leftrightarrow B$, $p_A = \frac{[A]}{[A]+[B]}$, where p_A is the probability of having the species A. Similarly, $p_B = \frac{[B]}{[A]+[B]}$. This immediately implies that

$$K = \frac{[B]}{[A]} = \frac{[A]+[B]}{[A]+[B]} \frac{[B]}{[A]} = \frac{p_B}{p_A}$$

Thus, we can calculate the equilibrium constants by simply taking the ratio of probabilities. Thus, for 50% helix \leftrightarrow 20% helix we have

$$K = \frac{p(k=20, N=100)}{p(k=50, N=100)} = \frac{(N-k_1+1)}{(N-k_2+1)} = \frac{(100-20+1)}{(100-50+1)} = \frac{81}{51} = \frac{27}{17} = 1.59$$

Similarly for the equilibrium between 75 % helix and 25 % helix, we find that

$$K = \frac{p(k=25, N=100)}{p(k=75, N=100)} = \frac{(N-k_1+1)}{(N-k_2+1)} = \frac{(100-25+1)}{(100-75+1)} = \frac{76}{26} = \frac{38}{13} = 2.92$$

■ Problem 2 (6 points)

When a process is highly cooperative, the probability that the system will be in any of the intermediate states is very small. Thus, when a helix-coil transition is highly-cooperative, i.e. the major energetic cost is in nucleating the helix, the major components of the partition function will be the state in which the entire polymer is a coil and the state in which the entire polymer is a helix.

Thus, we need only consider the probability of having all of the units in helix 1 or helix 2.

The probability of having all N units in each helical state is

$$p_1(N = 100) = \sigma s^N / Z$$

$$p_2(N = 100) = \sigma' s'^N / Z$$

where Z is the partition function that includes both types of possible helices.

Now, if the free energy of these two helical conformations is identical, then the equilibrium constant must be 1 via

$$K = e^{\frac{-\Delta G}{k_B T}} = e^0 = 1$$

Since the equilibrium constant is simply the ratio of the relative probabilities, we don't need to evaluate the partition function, and we find that

$$K = \frac{\sigma s^N}{\sigma' s'^N} = 1 \Rightarrow \sigma' = \left(\frac{s}{s'}\right)^N \sigma = (10^{-0.02})^{100} (10^{-4}) = 10^{-2} 10^{-4} = 10^{-6}$$

What does this result mean physically? Recall that s and s' are determined from the free energy difference between the coil state and each of the helical states, $s = e^{-\Delta G/k_B T}$. Thus, a ratio of s/s' of $10^{-0.02}$ implies more free energy is released when a unit of helix 2 is formed than when a unit of helix 1 is formed. If this is true, then how can helix 1 and helix 2 exist in equal concentrations at equilibrium? The answer is that the nucleation of helix 2 costs 100 times more energy than the nucleation of helix 1. Thus despite the more stable interactions in helix 2, it just as likely as helix 1!

■ Problem 3 (10 points)

■ Zipper Model (5 points)

To calculate the partition function for the Zipper model for circular DNA, let's first recall the partition function for linear DNA.

$$Z = 1 + \sigma \sum_{n=1}^N \Omega_N s^n$$

where Ω_N is the degeneracy of each state, i.e. the number of ways that we can have k contiguous sections of helix. In a linear piece of DNA, where these stretches cannot extend beyond the ends of the DNA, $\Omega_N = (N - k + 1)$. (Note that this degeneracy implies that we can distinguish each part of the polymer, i.e. CCHHHCCC is counted separately from its inversion CCCHHHCC)

In circular DNA, the contiguous stretches of helix can wrap around the DNA. For example, a stretch of 9 helical units in a piece of DNA 10 units long can only have two locations in a linear piece, but can have 10 unique locations in a circular piece of DNA. Thus, for a circular piece of DNA (in which we can distinguish each segment from the next) $\Omega_N = N$. However, when all of the units are in the helix, then there is only one way this chain can be oriented.

Thus, the partition function for circular DNA is

$$Z = 1 + \sigma \sum_{n=1}^{N-1} N s^n + s^N$$

The σ is missing in the last term because σ originates in the boundary between non-helical regions. When the entire molecule is in a helix (and its circular) then there is no boundary; thus, no σ factor. Alternatively, reasonable arguments can be made to keep it in.

This sum is a geometric sum that can be evaluated analytically

$$\sum_{n=1}^{N-1} s^n = \frac{(s^N - s)}{s - 1}$$

Thus, the partition function is

$$Z = 1 + \sigma \frac{N(s^N - s)}{s - 1} + s^N$$

To calculate the fraction of the DNA that is in helix, θ , recall that

$$\theta = \frac{1}{N} \frac{d \ln Z}{d \ln s} = \frac{s}{N} \frac{1}{Z} \frac{dZ}{ds}$$

The derivative of the partition function is

$$D \left[1 + s^N + \frac{N (s^N - s) \sigma}{s - 1}, s \right] // \text{Simplify}$$

$$\frac{N (s^{2+N} + s \sigma + s^{1+N} (-2 + (-1 + N) \sigma) + s^N (1 - N \sigma))}{(-1 + s)^2 s}$$

and thus θ is

$$\theta = \text{FullSimplify} \left[\frac{s}{N} \frac{1}{1 + \sigma \frac{N (s^N - s)}{s - 1} + s^N} \frac{N (s^{2+N} + s \sigma + s^{1+N} (-2 + (-1 + N) \sigma) + s^N (1 - N \sigma))}{(-1 + s)^2 s} \right]$$

$$\frac{s \sigma + s^N (1 - N \sigma + s (-2 + s + (-1 + N) \sigma))}{(-1 + s) (-1 + s - N s \sigma + s^N (-1 + s + N \sigma))}$$

Thus, the full expression is

$$\theta_{\text{circle}} = \frac{s}{N} \frac{1}{Z} \frac{dZ}{ds} = \frac{s^{N+2} + s^{N+1} ((N-1) \sigma - 2) + s^N (1 - N \sigma) + \sigma s}{(s-1) (s^{N+1} + s^N (N \sigma - 1) + s (1 - N \sigma) - 1)}$$

The value for θ_{linear} derived in class is

$$\theta_{\text{linear}} = \frac{\sigma s}{(s-1)^3 N} \frac{N s^{N+2} - (N+2) s^{N+1} + (N+2) s - N}{1 + \frac{\sigma s}{(s-1)^2} (s^{N+1} - (N+1) s + N)}$$

Comparing these expressions directly is not simple or insightful. Instead, let's consider the degeneracy for each state. In the linear model $\Omega_k = (N - k + 1)$ whereas in the circular DNA model, $\Omega_k = N$. Thus, for all values of k , $\Omega_{\text{linear}} \leq \Omega_{\text{circle}}$. Thus, in the circular DNA there are more ways to have longer stretches of helical regions, and since the energy of these helical regions is the same for both the linear and circular cases, this implies that $\theta_{\text{circle}} > \theta_{\text{linear}}$.

Alternatively, if one argued that all k regions of helix are indistinguishable, then the above argument is reversed and $\theta_{\text{circle}} < \theta_{\text{linear}}$.

■ Bragg and Zimm Model (5 points)

Recall that the partition function in the Bragg and Zimm model is

$$Z = \begin{pmatrix} 1 & 0 \end{pmatrix} M^N \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

where in vector notation $\begin{pmatrix} a & b \end{pmatrix}$ and $\begin{pmatrix} a \\ b \end{pmatrix}$ represent the state of the system, a is 1 when the state is in a coil and b is 1 when the state is in a helix, otherwise a and b are zero. The matrix M is

$$M = \begin{pmatrix} 1 & \sigma s \\ 1 & s \end{pmatrix}$$

and represents the statistical weights for the system, given the value of the preceding state. It is important to realize that the above expression for the partition function is actually shorthand for

$$Z = \begin{pmatrix} 1 & 0 \end{pmatrix} M^N \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \begin{pmatrix} 1 & 0 \end{pmatrix} M^N \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

since the state vector can either be $\begin{pmatrix} 0 & 1 \end{pmatrix}$ or $\begin{pmatrix} 1 & 0 \end{pmatrix}$, i.e. either coil or helix, not both. Thus, our partition function is the sum of both possibilities, the final segment is a helix $\begin{pmatrix} 0 & 1 \end{pmatrix}$ or the final segment is a coil $\begin{pmatrix} 1 & 0 \end{pmatrix}$.

In the case of circular DNA, we have periodic boundary conditions, which implies that the first vector and the last vector must be equal. In addition, because there is no segment that does not have a real segment next to it, we do not need to assume the the 0th segment is a coil. Thus, the partition function is the sum of both possible configurations of the n th segment

$$Z = \begin{pmatrix} 1 & 0 \end{pmatrix} M^N \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 & 1 \end{pmatrix} M^N \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

Now to perform these matrix operations, we assume that there is a similarity transformation that converts M^N into a diagonal matrix Λ^N , via the relationship $T^{-1} M T = \Lambda = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$, where $\lambda_{1,2}$ are the eigenvalues of M . Furthermore, one can show that $M = T \Lambda T^{-1}$ and $M^N = (T \Lambda T^{-1})^N = T \Lambda^N T^{-1}$

This last relationship can be proven by writing out the product

$$(T \Lambda T^{-1})^N = (T \Lambda T^{-1})(T \Lambda T^{-1})(T \Lambda T^{-1}) \dots (T \Lambda T^{-1}) = T \Lambda (T^{-1} T) \Lambda (T^{-1} T) \Lambda (T^{-1} T) \dots (T^{-1} T) \Lambda T^{-1} = T \Lambda^N T^{-1}$$

Thus, our partition function becomes

$$Z = \begin{pmatrix} 1 & 0 \end{pmatrix} T \Lambda^N T^{-1} \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 & 1 \end{pmatrix} T \Lambda^N T^{-1} \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

Because the matrix M is the same as for the linear polymer, the eigenvalues and the eigenvectors (the columns of T) are the same as calculated in class. Thus,

$$T = \begin{pmatrix} 1 - \lambda_2 & 1 - \lambda_1 \\ 1 & 1 \end{pmatrix}$$

$$T^{-1} = \frac{1}{\lambda_1 - \lambda_2} \begin{pmatrix} 1 & \lambda_1 - 1 \\ -1 & 1 - \lambda_2 \end{pmatrix}$$

$$\lambda_{1,2} = \lambda_{\pm} = \frac{1+s \pm \sqrt{(1-s)^2 + 4\sigma s}}{2}$$

To calculate the partition function, we need to do some matrix algebra

$$\begin{aligned} Z &= (1 \ 0) T \Lambda^N \frac{1}{\lambda_1 - \lambda_2} \begin{pmatrix} 1 \\ -1 \end{pmatrix} + (0 \ 1) T \Lambda^N \frac{1}{\lambda_1 - \lambda_2} \begin{pmatrix} \lambda_1 - 1 \\ 1 - \lambda_2 \end{pmatrix} \\ &= \frac{1}{\lambda_1 - \lambda_2} (1 \ 0) T \begin{pmatrix} \lambda_1^N \\ -\lambda_2^N \end{pmatrix} + \frac{1}{\lambda_1 - \lambda_2} (0 \ 1) T \begin{pmatrix} \lambda_1^N(\lambda_1 - 1) \\ \lambda_2^N(1 - \lambda_2) \end{pmatrix} \\ &= \frac{1}{\lambda_1 - \lambda_2} (1 \ 0) \begin{pmatrix} \lambda_1^N(1 - \lambda_2) + -\lambda_2^N(1 - \lambda_1) \\ \lambda_1^N - \lambda_2^N \end{pmatrix} + \frac{1}{\lambda_1 - \lambda_2} (0 \ 1) \begin{pmatrix} \lambda_1^N(\lambda_1 - 1)(1 - \lambda_2) + \lambda_2^N(1 - \lambda_2)(1 - \lambda_1) \\ \lambda_2^N(1 - \lambda_2) + \lambda_1^N(\lambda_1 - 1) \end{pmatrix} \\ &= \frac{1}{\lambda_1 - \lambda_2} (\lambda_1^N(1 - \lambda_2) - \lambda_2^N(1 - \lambda_1)) + \frac{1}{\lambda_1 - \lambda_2} (\lambda_2^N(1 - \lambda_2) - \lambda_1^N(1 - \lambda_1)) = \frac{1}{\lambda_1 - \lambda_2} (\lambda_1^N(\lambda_1 - \lambda_2) + \lambda_2^N(\lambda_1 - \lambda_2)) \\ &= \lambda_1^N + \lambda_2^N \end{aligned}$$

Thus, at the end of all of the horrible algebra, we find the simple and elegant solution

$$Z = \lambda_1^N + \lambda_2^N$$

We could have skipped all of this if we had made one simple observation concerning linear algebra. The partition function is

$$Z = (1 \ 0) M^N \begin{pmatrix} 1 \\ 0 \end{pmatrix} + (0 \ 1) M^N \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \text{Tr}(M^N)$$

where Tr represents the trace of the matrix. A general property of traces is that $\text{Tr}(M) = \text{Tr}(\Lambda)$. Thus,

$$Z = \text{Tr}(M^N) = \text{Tr}(\Lambda^N) = \text{Tr} \begin{pmatrix} \lambda_1^N & 0 \\ 0 & \lambda_2^N \end{pmatrix} = \lambda_1^N + \lambda_2^N$$

confirming the more complicated derivation above.

With the partition function we can calculate θ_{circle} as above:

$$\theta_{\text{circle}} = \frac{s}{N} \frac{1}{Z} \frac{dZ}{ds} = \frac{s}{N} \frac{1}{Z} (N \lambda_1^{N-1} \frac{d\lambda_1}{ds} + N \lambda_2^{N-1} \frac{d\lambda_2}{ds}) = \frac{s}{\lambda_1^N + \lambda_2^N} (\lambda_1^{N-1} \frac{d\lambda_1}{ds} + \lambda_2^{N-1} \frac{d\lambda_2}{ds})$$

The derivatives of the eigenvalues are

$$\frac{d\lambda_{1,2}}{ds} = \frac{1}{2} \left(1 \pm \frac{-1+s+2\sigma}{\sqrt{(-1+s)^2 + 4\sigma s}} \right)$$

Thus,

$$\theta_{\text{circle}} = \frac{s}{N} \frac{1}{Z} \frac{dZ}{ds} = \frac{s}{N} \frac{1}{Z} (N \lambda_1^{N-1} \frac{d\lambda_1}{ds} + N \lambda_2^{N-1} \frac{d\lambda_2}{ds}) = \frac{1}{2} \frac{s}{\lambda_1^N + \lambda_2^N} \left(\lambda_1^{N-1} \left(1 + \frac{-1+s+2\sigma}{\sqrt{(-1+s)^2 + 4\sigma s}} \right) + \lambda_2^{N-1} \left(1 - \frac{-1+s+2\sigma}{\sqrt{(-1+s)^2 + 4\sigma s}} \right) \right)$$

Without explicitly comparing θ_{circle} and θ_{linear} we can make the same argument as above: because there are more ways to make large strings of helix in the circular DNA case, the average value of θ will be larger for the circular DNA as opposed to the linear DNA.

The full blown solution is ugly and cumbersome. However, if we recognize that because of the + sign in λ_1 , $\lambda_1 > \lambda_2$, thus when N is large, $\lambda_1^N \gg \lambda_2^N$, and our partition function is $Z \sim \lambda_1^N$.

The value for θ_{circle} is much simpler in this limit

$$\theta_{\text{circle}} = \frac{s}{N} \frac{1}{Z} \frac{dZ}{ds} \sim \frac{s}{N} \frac{1}{\lambda_1^N} N \lambda_1^{N-1} \frac{d\lambda_1}{ds} = s \frac{1}{\lambda_1} \frac{d\lambda_1}{ds} = \frac{s}{2\lambda_1} \left(1 + \frac{-1+s+2\sigma}{\sqrt{(-1+s)^2+4s\sigma}} \right)$$

It is not surprising to note that in this limit θ_{circle} takes exactly the same form as θ_{linear} !

■ Problem 4 (10 points)

■ Part a (5 points)

From the lecture notes, the charge density of a given ion will be governed by the Boltzmann equation:

$$\begin{aligned} n_+ &= c e^{\frac{-E_+}{k_B T}} \\ n_- &= c e^{\frac{-E_-}{k_B T}} \end{aligned}$$

where n_+ is the concentration (number per volume) of the positive ion with a charge e , c is the total concentration of the salt (i.e. the original concentration of the ion), and $E_{+/-}$ are the energies associated with a positive or negative charge a distance r away from the DNA.

From simple electrostatics, if we have a potential $\psi(r)$, then $E_{+/-} = Z_{+/-} e \psi(r)$; thus for a monovalent salt, $E_{+/-} = \pm e \psi(r)$.

In class, we showed that the electrostatic potential can be solved for a distance r away from the linear DNA molecule when r is large and the ionic strength is relatively low. If these conditions are met, then the electrostatic potential is described by

$$\psi(r) = -2 \frac{k_B T}{e} \xi \frac{K_0(\kappa r)}{\kappa a K_1(\kappa a)}$$

where $\xi = \frac{e^2}{\epsilon k_B T b}$, $\kappa = \left(\frac{8\pi e^2}{k_B T \epsilon} c \right)^{1/2}$, and $K_{0,1}$ are modified Bessel functions of the second kind of order 0 or 1. In addition, b is the inverse of the linear charge density of the molecule, i.e. the length between charges, ϵ is the electric permittivity in water, $\epsilon = 4\pi \epsilon_r \epsilon_0$, and a is the radius of DNA, ~ 1 nm. (Note these expressions are valid only for monovalent salts!) Moreover, we have taken into account that the charge on an electron is negative which gives the minus sign not hidden in the expression in the notes.

In the Manning condensation model, positive ions are condensed onto the DNA, neutralizing its charge, until $\xi = 1$ for a monovalent salt. Thus, we can set $\xi = 1$ in the above expression.

Thus, the energy of an ion at a distance r is

$$E_{+/-}(r) = \mp 2 k_B T \frac{K_0(\kappa r)}{\kappa a K_1(\kappa a)}$$

Neglecting the ions lost in condensation, the concentration of each type of ion will be

$$n_+ = c e^{2 \frac{K_0(\kappa r)}{\kappa a K_1(\kappa a)}}$$

$$n_- = c e^{-2 \frac{K_0(\kappa r)}{\kappa a K_1(\kappa a)}}$$

■ Part b (5 points)

Now, to find these concentrations at the given distances with a salt concentration of 100 mM, we must first compute the Debye-length, κ^{-1} .

Recall from above that

$$\kappa = \left(\frac{8\pi e^2}{k_B T \epsilon} c \right)^{1/2}$$

In this case,

$$e = 1.6 \times 10^{-19} \text{ C}$$

$$k_B = 1.38 \times 10^{-23} \text{ J / K}$$

$$T = 273 + 25 = 298 \text{ K}$$

$$\epsilon = 4\pi \epsilon_r \epsilon_0 = 4\pi (78.5) (8.85 \times 10^{-12} \text{ C}^2 / \text{J m})$$

$$= 8.73 \times 10^{-9} \text{ C}^2 / \text{J m}$$

$$n_s = .1 \text{ M} = (0.1 \text{ mol/L}) (6.22 \times 10^{23} \text{ molecules/mol}) (1000 \text{ m}^3 / \text{L})$$

$$= 6.22 \times 10^{25} \text{ molecules/m}^3$$

One specific note, the equations derived in class use CGS units, by setting $\epsilon = 4\pi \epsilon_r \epsilon_0$ we are converting to SI units. Also, ϵ_r is the dielectric constant for water, 78.5 at room temperature.

Thus,

$$\kappa = \left(\frac{8\pi (1.6 \times 10^{-19} \text{ C})^2}{(1.38 \times 10^{-23} \text{ J/mol K}) (298 \text{ K}) 4\pi (78.5) (8.85 \times 10^{-12} \text{ C}^2 / \text{J m})} \right)^{1/2} (6.22 \times 10^{25} \text{ molecules/m}^3)^{1/2} = 1.056 \times 10^9 \text{ m}^{-1}$$

Thus, the Debye length, κ^{-1} , is

$$\kappa^{-1} = 9.5 \text{ \AA}$$

Returning to the expressions derived in part a

$$n_+ = c e^{2 \frac{K_0(\kappa r)}{\kappa a K_1(\kappa a)}}$$

$$n_- = c e^{-2 \frac{K_0(\kappa r)}{\kappa a K_1(\kappa a)}}$$

we can determine if the exponents $2 \frac{K_0(\kappa r)}{\kappa a K_1(\kappa a)}$, can be simplified

In particular, we can determine if $\kappa a \ll 1$.

$$\kappa a = 10 \text{ \AA} / 9.5 \text{ \AA} \sim 1$$

It is not! Calculating it

```
(10 / 9.5) BesselK[1, 10 / 9.5]
0.580003
```

We find that $\kappa a K_1(\kappa a) = 0.58$. Thus, all of our exponents will be scaled by this factor.

Now calculating, $K_0(\kappa r)$, for the different distances: 50 Å, 100 Å, and 200 Å:

```
BesselK[0, 50 / 9.5]
```

```
0.00276815
```

```
BesselK[0, 100 / 9.5]
```

```
0.0000102439
```

```
BesselK[0, 200 / 9.5]
```

```
1.95364 x 10^-10
```

Thus, the exponents for 50 Å, 100 Å, and 200 Å:

```
2 BesselK[0, 50 / 9.5]
  0.58
```

```
0.00954535
```

```
2 BesselK[0, 100 / 9.5]
  0.58
```

```
0.000035324
```

```
2 BesselK[0, 200 / 9.5]
  0.58
```

```
6.73669 x 10^-10
```

All of these values are much less than 1, so we can expand the exponential to first order, $e^x \sim 1 + x$. Thus our final concentrations will be

$$n_+ \approx c \left(1 + 2 \frac{K_0(\kappa r)}{\kappa a K_1(\kappa a)} \right)$$

$$n_- \approx c \left(1 - 2 \frac{K_0(\kappa r)}{\kappa a K_1(\kappa a)} \right)$$

and for 50 Å, 100 Å, and 200 Å: we will have

$$n_+ \approx 100 \text{ mM} (1 + 2.7 \times 10^{-3}) = 100.3 \text{ mM}$$

$$n_- \approx 100 \text{ mM} (1 - 2.7 \times 10^{-3}) = 99.7 \text{ mM}$$

$$n_+ \approx 100 \text{ mM} (1 + 3.5 \times 10^{-5}) = 100 \text{ mM}$$

$$n_- \approx 100 \text{ mM} (1 - 3.5 \times 10^{-5}) = 100 \text{ mM}$$

$$n_+ \approx 100 \text{ mM} (1 + 6.7 \times 10^{-10}) = 100 \text{ mM}$$

$$n_- \approx 100 \text{ mM} (1 - 6.7 \times 10^{-10}) = 100 \text{ mM}$$

Thus, the changes are small even when within 5 nm of the DNA! This is because the Debye length is small, $\kappa^{-1} \sim 1 \text{ nm}$, beyond this the ions in solution don't "know" that there is a negatively charged cylinder (DNA).

■ **Problem 5 (10 points)**