Special Topics in Physical Chemistry. Homework 4 (Fall 2008)

In the class we learned about Levinthal's paradox, which states that finding the native state of a protein by a random search among all possible configurations can take an enormously long time, yet proteins can fold in seconds or less. Mathematical analysis can show that a small energetic bias ($\sim k_BT$) against a locally unfavorable configurations can reduce the search time to a biologically significant time scale.

Suppose a protein consisting of N-bond, each of which can take either a correct bond or ν incorrect bonds. k_0 defines the rate of transition from correct to incorrect bond ($c \to i$), while k_1 defines the rate of transition from incorrect to correct bond, and also there are ν states for incorrect bond ($i \to c$). Hence, the chemical equilibrium between correct and incorrect bond ($c \rightleftharpoons i$) is defined as

$$K = [i]_{eq}/[c]_{eq} = k_0/k_1 = \nu e^{-U/k_BT}. \tag{1}$$

Let us define $P(S,t)$ as a probability of our polypeptide chain having S incorrect bonds.

1. Show that the rate equation can be written as follows

$$\begin{aligned}\frac{d}{dt}P(S,t) &= \sum_{S'} W(S,S')P(S',t)\\ &= W(S,S+1)P(S+1,t) + W(S,S)P(S,t) + W(S,S-1)P(S-1,t)\end{aligned} \tag{2}$$

where $W(S,S) = -(N-S)k_0 - Sk_1$, $W(S,S-1) = (N-S+1)k_0$, $W(S,S+1) = (S+1)k_1$ with boundary value $P(-1,t) = P(S+1,t) = 0$.

2. We can understand the above transition matrix $W(S,S')$ as a Fokker Planck operator ($\mathcal{L}_{FP}(x)$). In parallel to the equation for mean first passage time that we derived in the class ($\mathcal{L}^{\dagger}_{FP}(x_0)\tau(x_0) = -1$), the mean first passage time from S_0 satisfies

$$\sum_{S_0} \tau(S_0)W(S_0,S) = -1 \tag{3}$$

for all S when an absorbing boundary condition at $S = 0$ is imposed, so that only the states $S = 1$ to N are involved.

Show that Eq.(3) leads to the difference equation

$$(N-S)k_0[\tau(S+1) - \tau(S)] - Sk_1[\tau(S) - \tau(S-1)] = -1 \tag{4}$$

with boundary condition $\tau(0) = \tau(N+1) = 0$.

3. Show that the solution of Eq.(4) is

$$\tau(S) = \frac{1}{Nk_0} \sum_{n=0}^{S-1} \binom{N-1}{n}^{-1} \sum_{m=n+1}^{N} \binom{N}{m} K^{m-n} \tag{5}$$

(*Hint*) Eq.(4) is called a "difference equation". A difference equation of the following form

$$a_n - p(n)a_{n-1} = q(n) \tag{6}$$

can be solved by multiplying a integrating factor $\left[\prod_{j=1}^{n} p(j)\right]^{-1}$ on both sides of the equation. Once the integrating factor is multiplied, Eq.(6) is converted to

$$\frac{a_n}{\left[\prod_{j=1}^{n} p(j)\right]} - \frac{a_{n-1}}{\left[\prod_{j=1}^{n-1} p(j)\right]} = \frac{q(n)}{\left[\prod_{j=1}^{n} p(j)\right]} \tag{7}$$

Summing both sides from 1 to $n-1$ gives the solution to Eq.(6).

$$a_n = \prod_{j=1}^{n} p(j) \left[a_0 + \sum_{k=1}^{n} \frac{q(k)}{\prod_{j=1}^{k} p(j)} \right] \tag{8}$$

Let us rewrite Eq.(4) by defining $T(S) = \tau(S+1) - \tau(S)$.

$$(N-S)k_0 T(S) - Sk_1 T(S-1) = -1 \tag{9}$$

Note that Eq.(4) and (6) should be valid for $S = 1, 2, \ldots N$. Show that

$$T(n) = \binom{N-1}{n}^{-1} \frac{1}{K^n} \left\{ T(0) - \sum_{j=1}^{n} \binom{N-1}{j} K^j \frac{1}{(N-j)k_0} \right\} \tag{10}$$

Although it first appears that $T(0)$ is not determined, one can show that

$$T(0) = \tau(1) = \frac{1}{Nk_0}[(1+K)^N - 1] \tag{11}$$

using $T(N-1) = 1/(Nk_1)$ from Eq.(9).

4. By using the following integral identity

$$\sum_{m=n+1}^{N} \binom{N}{m} K^{m-n} = K(n+1) \binom{N}{n+1} \int_0^1 dx (1-x)^n (1+Kx)^{N-n-1} \tag{12}$$

and by changing x to the new variable $y = (1-x)/(1+Kx)$. Show that Eq.(5) is transformed into

$$\tau(S) = \frac{1}{k_0}(1+K)^N K \int_0^1 dy \frac{1-y^S}{1-y} (1+Ky)^{-N-1} \tag{13}$$

5. For large N, the integral is dominated by the contribution from small y. It is very weakly dependent on S. Show that its asymptotic form for large N is given by

$$\tau(S) \to \frac{1}{Nk_0}(1+K)^N \left\{ 1 + 1!(NK)^{-1} + 2!(NK)^{-2} + \cdots \right\}. \tag{14}$$

(*Hint* $(1+Ky)^{-N-1} \approx e^{-KNy}$ when $Ky \ll 1$ and $N \to \infty$).

Therefore, the mean folding time $\tau(S)$ is

$$\tau(S) \approx \frac{1}{Nk_0}(1+K)^N \tag{15}$$

6. Show that if $k_0 \approx 0$ i.e. no transition from c to i (correct to incorrect bond)

$$\tau(S) \approx \tau_M = \frac{1}{k_1}\sum_{i=1}^{S}\frac{1}{i}. \tag{16}$$

where the subscript of τ_M refers to the situation when the correctly typed letter cannot be changed in the random typing game for a monkey. (*Hint* Use Eq.(13)).

Meanwhile, if there is no bias the mean-first passage time estimated by Levinthal is

$$\tau_L \approx \frac{1}{Nk_0}(1+\nu)^N \tag{17}$$

The *enhancement of the rate due to native bias relative to the Levinthal scenario* is

$$\frac{\tau(S)}{\tau_L} = \left(\frac{1+\nu e^{-U/k_BT}}{1+\nu}\right)^N \tag{18}$$

7. Plot the graph of the ratio $\tau(S)/\tau_L$ for $N = 100$, $\nu = 10$ as a function of $U/k_BT (> 0)$ value.