Supplementary Material for "Optimal length of conformational transition region in protein search for targets on DNA"

Maria P. Kochugaeva, Alexander M. Berezhkovskii and Anatoly B. Kolomeisky
Department of Chemistry and Center for Theoretical Biological Physics, Rice University, Houston, TX 77005, USA; Mathematical and Statistical Computing Laboratory, Division of Computational Bioscience, Center for Information Technology, National Institutes of Health, Bethesda, Maryland 20892, USA
(Dated: July 7, 2017)

In this supplementary material, we provide a derivation of the expression for the mean search time in a more general case of arbitrary k_t in the continuum model framework, presented in Fig. 2, assuming that the size of the target may also vary. We assume that the system is in the stationary state with a constant flux J_0 of proteins injected into the solution and removed at the target. Then the mean search time, T_0, can be calculated from

$$T_0 = \frac{N_b + N_s + N_r}{J_0}, \tag{1}$$

where N_b, N_s and N_r are the steady-state numbers of proteins in the bulk, in the search mode, and in the recognition mode, respectively. Because the conformational transition is irreversible, the overall mean search time, as in the discrete-state case, can be written as the sum of two terms, $T_0 = t_1 + t_2$, corresponding to the mean times required for entering the recognition mode and finding the target after entering this mode. These two times are

$$t_1 = \frac{N_b + N_s}{J_0}, \quad t_2 = \frac{N_r}{J_0}. \tag{2}$$

Each of these times is evaluated separately as explained below.

The time to reach the recognition mode can be calculated in the following way. The number of proteins in the search mode is given by

$$N_s = \int_{L/2}^{L/2} c_s(x)dx, \tag{3}$$

where $c_s(x)$ is the stationary protein concentration in this mode. It is convenient to introduce the following notations, $c_s(x) = c_1(x)$ for $-L/2 < x < -l_r/2$ and $l_r/2 < x < L/2$, and $c_s(x) = c_2(x)$ for $-l_r/2 < x < l_r/2$. In addition, because of the symmetry, $c_s(x) = c_s(-x)$, we may consider only the region $0 < x < L/2$. The concentrations $c_1(x)$ and $c_2(x)$ can be explicitly calculated by solving the system of two reaction-diffusion equations,

$$D_s \frac{d^2 c_1}{dx^2} - k_{\text{off}} c_1 + (k_{\text{on}}/L) N_b = 0, \tag{4}$$

$$D_s \frac{d^2 c_2}{dx^2} - (k_{\text{off}} + k_t) c_2 + (k_{\text{on}}/L) N_b = 0, \tag{5}$$

subject to the boundary conditions:

$$\frac{dc_1}{dx}|_{x=L/2} = \frac{dc_2}{dx}|_{x=0} = 0, \tag{6}$$

$$c_1(l_r/2) = c_2(l_r/2), \tag{7}$$

,

$$\frac{dc_1}{dx}|_{x=l_r/2} = \frac{dc_2}{dx}|_{x=l_r/2}. \tag{8}$$

The solutions, satisfying the boundary conditions in Eq. (6), are

$$c_1(x) = b_1 - A_1 \cosh\left[(L/2 - x)\lambda_1\right], \tag{9}$$

for $l_r/2 < x < L/2$, and

$$c_2(x) = b_2 + A_2 \cosh(x\lambda_2), \tag{10}$$

for $0 < x < l_r/2$, where

$$\lambda_1 = \sqrt{k_{\rm off}/D_s}, \quad \lambda_2 = \sqrt{(k_{\rm off}+k_t)/D_s}, \tag{11}$$

and

$$b_1 = \frac{k_{\rm on}N_b}{k_{\rm off}L}, \quad b_2 = \frac{k_{\rm on}N_b}{(k_{\rm off}+k_t)L}. \tag{12}$$

It is important to note that

$$b_1 - b_2 = \frac{k_t}{k_{\rm off}+k_t}b_1 = \frac{k_t}{k_{\rm off}}b_2 = \frac{k_t k_{\rm on}N_b}{k_{\rm off}(k_{\rm off}+k_t)L}. \tag{13}$$

Factors A_1 and A_2 can be found from the matching conditions in Eqs. (7) and (8),

$$A_1 = \frac{(b_1-b_2)\lambda_2\tanh(z_2)}{\cosh(z_1)[\lambda_1\tanh(z_1)+\lambda_2\tanh(z_2)]}, \quad A_2 = \frac{(b_1-b_2)\lambda_1\tanh(z_1)}{\cosh(z_2)[\lambda_1\tanh(z_1)+\lambda_2\tanh(z_2)]}, \tag{14}$$

where

$$z_1 = (L-l_r)\lambda_1/2, \quad z_2 = l_r\lambda_2/2. \tag{15}$$

In addition, we have a stationary balance condition

$$J_0 = k_{\rm on}N_b - k_{\rm off}N_s. \tag{16}$$

Using this and the above relations, we obtain

$$N_b = \frac{k_{\rm off}+k_t}{k_{\rm on}k_t}\frac{L}{l_r+S}J_0 \tag{17}$$

and

$$N_s = \frac{1}{k_{\rm off}}\left[\frac{k_{\rm off}+k_t}{k_t}\frac{L}{l_r+S} - 1\right]J_0, \tag{18}$$

which can be used to get the explicit expression for t_1, namely,

$$t_1 = \frac{N_s+N_b}{J_0} = \frac{1}{k_{\rm off}}\left[\frac{(k_{\rm off}+k_t)(k_{\rm off}+k_{\rm on})L}{k_t k_{\rm on}(l_r+S)} - 1\right]. \tag{19}$$

In these expressions, the parameter S is given by

$$S = \frac{2k_t}{\sqrt{k_{\rm off}(k_{\rm off}+k_t)}}\frac{\tanh(z_1)\tanh(z_2)}{\lambda_1\tanh(z_1)+\lambda_2\tanh(z_2)}. \tag{20}$$

A similar analysis can be done to calculate the time needed to reach the target while being in the recognition mode. For that we need to know the steady-state number of proteins in the recognition mode, N_r, which is given by

$$N_r = \int_{-l_r/2}^{l_r/2} c_r(x)dx = 2\int_{l_t/2}^{l_r/2} c_r(x)dx, \tag{21}$$

where $c_r(x)$ is the steady-state concentration of the protein in the recognition mode, and we have used the fact that $c_r(x) = c_r(-x)$ and $c_r(x)|_{|x|<l_t/2} = 0$. The concentration $c_r(x)$ can be calculated from the corresponding reaction-diffusion equation for $l_t/2 < x < l_r/2$:

$$D_r\frac{d^2c_r}{dx^2} + k_t c_2(x) = 0, \tag{22}$$

subject to the boundary conditions,

$$c_r(l_t/2) = \frac{dc_r}{dx}|_{x=l_r/2} = 0. \tag{23}$$

After substitution $c_2(x)$ from Eq. (10) into Eq. (22) and two integrations, we find $c_r(x)$. Next, we substitute the obtained $c_r(x)$ in Eq. (21) to find the number of proteins in the recognition mode, N_r,

$$N_r = 2\lambda_r^2 b_2 \left\{ \frac{(l_r - l_t)^3}{3} + \frac{S}{2\lambda_2^2} \left[\frac{(l_r - l_t)^2 \lambda_2^2}{2} - 1 + \frac{1}{\sinh(z_2)} \left(\sinh(l_t\lambda_2/2) + (l_r - l_t)\lambda_2 \cosh(l_t\lambda_2/2) \right) \right] \right\}, \tag{24}$$

where $\lambda_r = \sqrt{k_t/D_r}$.

To finish the derivation of time t_2, we need to know the ratio b_2/J_0. According to Eq. (12), we have

$$\frac{b_2}{J_0} = \frac{k_{\rm on} N_b}{(k_{\rm off} + k_t) L J_0} = \frac{1}{k_t(l_r + S)}, \tag{25}$$

where we have taken advantage of Eq. (17). Thus, the final result for $t_2 = N_r/J_0$ is

$$t_2 = \frac{2}{D_r(l_r + S)} \left\{ \frac{(l_r - l_t)^3}{3} + \frac{S}{2\lambda_2^2} \left[\frac{(l_r - l_t)^2 \lambda_2^2}{2} - 1 + \frac{1}{\sinh(z_2)} \left(\sinh(l_t\lambda_2/2) + (l_r - l_t)\lambda_2 \cosh(l_t\lambda_2/2) \right) \right] \right\}. \tag{26}$$

The results of the calculations of the mean search time T_0 in Eq. (1) for the continuum model are presented in Figs. 5 and 6.