http://dx.doi.org/10.5012/bkcs.2012.33.3.971

Facilitated Protein-DNA Binding: Theory and Monte Carlo Simulation†

Kihyun Park,* Taejun Kim,‡ and Hyojoon Kim‡,*

*School of Computational Sciences, Korea Institute for Advanced Study, Seoul 130-722, Korea. *E-mail: real@kias.re.kr*
*‡Department of Chemistry, Dong-A University, Busan 604-714, Korea. *E-mail: hkim@donga.ac.kr*
Received December 5, 2011, Accepted January 12, 2012

The facilitated diffusion effect on protein-DNA binding is studied. A rigorous theoretical approach is presented to deal with the coupling between one-dimensional and three-dimensional diffusive motions. For a simplified model, the present approach can provide numerically exact results, which are confirmed by the lattice-based Monte Carlo simulations.

Key Words : Diffusion-influenced reaction, Facilitated diffusion, Protein-DNA binding, Dimension reduction, Exact solution

Introduction

One of central questions in molecular biology is how a DNA-binding protein searches for a specific binding site on DNA. The protein can find its target much faster than simple diffusion-reaction theories predict. Numerous experimental and theoretical studies[1-19] make us believe that diffusive sliding along DNA greatly facilitates the binding process between DNA and protein.

From the viewpoint of diffusion-influenced reaction theories, the facilitated diffusion problem has two coupled mechanisms of the three-dimensional (3D) nonspecific binding to a random site on DNA and the one-dimensional (1D) sliding along DNA to the specific binding site, as illustrated in Figure 1. Even if each mechanism can be solved analytically, the coupling between the two prohibits us from obtaining the exact analytical results. Most previous theoretical approaches treated the diffusion-reaction process on the basis of the phenomenological rate law. In this paper, a rigorous theoretical approach is presented to provide numerically exact solutions for a simplified model of the facilitated diffusion-reaction problem. The lattice-based Monte Carlo simulations are also performed to confirm our theoretical predictions.

Theory

As an archetype of the facilitated diffusion-influenced reaction between a protein and a specific binding site on DNA, we assume the DNA-binding protein as a sphere and the DNA strand as an infinite cylinder. When the protein encounters the DNA within the distance R_r by diffusion in the bulk cytoplasm, the protein is assumed to be trapped on the DNA strand. Without loss of generality, we can regard the protein as a point particle and the DNA as a cylinder with radius R_r along z-axis. The specific binding site is modeled with a reactive cylinder in DNA located at origin. The trapped protein searches for the binding site by 1D sliding. The search ends when the protein encounters the reactive cylinder within the distance R_z. The simplified model is illustrated in Figure 1. The reaction can occur in two ways. One is the direct encounter with the reactive cylinder by diffusion in 3D and the other is the two-step reaction by coupling of 1D and 3D diffusive motions.

Let us denote the probability density to find a free protein at time t as $p(r,z,t)$ in cylindrical coordinates r and z. Similarly, $p(*_r,z,t)$ is the probability density for a protein trapped on the DNA strand and $p(*_r,*_z,t)$ is the specific binding reaction probability. Here, $*_r$ denotes the trapped state and $*_z$ denotes the bound state. Then, the evolution of reaction-diffusion equations are as follows:

$$\frac{\partial p(r,z,t)}{\partial t} = \mathcal{L}_{\text{free}} p(r,z,t)\,, \tag{1}$$

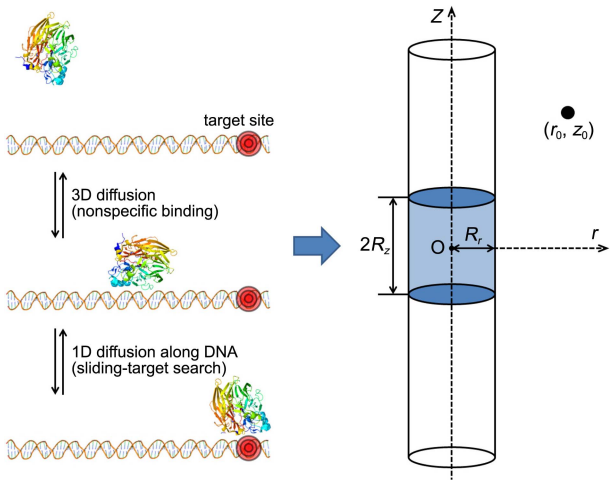

Figure 1. Schematic representation of protein-DNA binding by the facilitated diffusion. In the left figure, the searching process consists of 3D and 1D diffusive motions. In the right figure, our simplified model is presented. The protein is regarded as a point particle, DNA is an infinite cylinder along z-axis, and the specific binding site is a color cylinder at origin in the cylindrical coordinate.

†This paper is to commemorate Professor Kook Joe Shin's honourable retirement.

$$\frac{\partial p(*_r,z,t)}{\partial t} = H(|z|-R_z)\mathcal{L}_{\text{trap}}p(*_r,z,t)\,, \tag{2}$$

where $r \geq R_r$, $-\infty < z < \infty$, and $H(z)$ is the Heaviside step function. The diffusion operators without an interaction potential are defined as

$$\mathcal{L}_{\text{free}} = D_r\left(\frac{\partial^2}{\partial r^2}+\frac{1}{r}\frac{\partial}{\partial r}\right) + D_z\frac{\partial^2}{\partial z^2}\,, \tag{3}$$

$$\mathcal{L}_{\text{trap}} \equiv D_{\text{trap}}\frac{\partial^2}{\partial z^2}\,, \tag{4}$$

with the corresponding *mutual* diffusion constants D_r, D_z, and D_{trap}.

To make the problem tractable, here, we consider only the perfect irreversible trapping and binding, which means that every encounter leads to the trapping and binding and reversible untrapping and unbinding are not allowed. Then, for an initially free state, namely,

$$p(r,z,0|r_0,z_0) = \frac{\delta(r-r_0)\delta(z-z_0)}{2\pi r}\,, \tag{5}$$

we can solve Eq. (1) by using separation of variables and obtain the Green function as a product solution[20] of well-known functions,

$$p(r,z,t|r_0,z_0) = p_r(r,t|r_0)G_z(z,t|z_0)\,, \tag{6}$$

where $G_z(z,t|z_0)$ is the 1D Green function of a free particle[21]

$$G_z(z,t|z_0) = \frac{1}{\sqrt{4\pi D_z t}}\exp\left[-\frac{(z-z_0)^2}{4D_z t}\right], \tag{7}$$

and the two-dimensional (2D) function $p_r(r,t|r_0)$ can be obtained in the Laplace-transformed $[\tilde{f}(s)=\int_0^\infty f(\tau)e^{-s\tau}d\tau]$ domain as[21]

$$\tilde{p}_r(r,s|r_0)=\frac{K_0(\zeta^{\max}\sqrt{s\tau_r})}{k_{D_r}K_0(\sqrt{s\tau_r})}[K_0(\sqrt{s\tau_r})I_0(\zeta^{\min}\sqrt{s\tau_r})-I_0(\sqrt{s\tau_r})K_0(\zeta^{\min}\sqrt{s\tau_r})]. \tag{8}$$

Here, $\zeta^{\max}=\max(r,r_0)/R_r$ and $\zeta^{\min}=\min(r,r_0)/R_r$ and the diffusion control rate constant and time scale are defined as $k_{D_r}=2\pi D_r$ and $\tau_r \equiv R_r^2/D_r$, respectively. $K_n(x)$ and $I_n(x)$ are the modified Bessel functions of order n. Integrating Eq. (8) gives the survival probability

$$\tilde{S}_r(s|r_0) \equiv \int_{R_r}^\infty 2\pi r dr \tilde{p}_r(r,s|r_0) = \frac{1}{s}\left[1-\frac{K_0(\sqrt{s\tau_r}\,r_0/R_r)}{K_0(\sqrt{s\tau_r})}\right]. \tag{9}$$

We can use well-known numerical inverse Laplace-transform algorithms[22] for Eqs. (8) and (9).

The solution for an initially trapped state at z_0 is also well-known as

$$p(*_r,z,t|*_r,z_0) = \frac{H(|z_0|-R_z)H(|z|-R_z)}{\sqrt{4\pi D_{\text{trap}}t}} \times \left\{\exp\left[-\frac{(z-z_0)^2}{4D_{\text{trap}}t}\right]-\exp\left[-\frac{(|z+z_0|-2R_z)^2}{4D_{\text{trap}}t}\right]\right\}, \tag{10}$$

and integrating Eq. (10) over z leads to

$$S_{\text{trap}}(t|*_r,z_0) = H(|z_0|-R_z)\left[1-\text{erfc}\left(\frac{|z_0|-R_z}{\sqrt{4D_{\text{trap}}t}}\right)\right], \tag{11}$$

where erfc(x) is the complementary error function.

Now, we suppose that a free diffusing particle is trapped on the nonspecific position z' at time t' and slides until time t. Then, the probability density function of the particle is proportional to $k_r(t'|r_0)G_z(z',t'|z_0)p(*_r,z,t-t'|*_r,z')$, where the time-dependent trapping rate is defined as

$$k_r(t|r_0) \equiv -\frac{dS_r(t|r_0)}{dt} = k_{D_r}R_r\frac{\partial p_r(r,t|r_0)}{\partial r}\bigg|_{r=R_r}. \tag{12}$$

Since $zz' > 0$ for the irreversible trapping and specific binding reactions, we have

$$p(*_r,z,t|r_0,z_0) = \int_0^t dt' k_r(t'|r_0) \times \int_{-\infty}^\infty dz' H(|z'|-R_z)H(zz')G_z(z',t'|z_0)p(*_r,z,t-t'|*_r,z')\,. \tag{13}$$

When $z_0 = 0$, this equation, by symmetry, reduces to

$$p(*_r,z,t|r_0,0)=2\int_{R_z}^\infty\int_0^t dz'dt'k_r(\tau|r_0)G_z(z',t'|0)p(*_r,z,t-t'|*_r,z'). \tag{14}$$

The total survival probability function without the specific binding can be obtained as

$$S(t|r_0,z_0) = S_{\text{free}}(t|r_0,z_0) + S_{\text{trap}}(t|r_0,z_0), \tag{15}$$

where

$$S_{\text{free}}(t|r_0,z_0) = \int_{R_r}^\infty\int_{-\infty}^\infty dz\, dr 2\pi r\, p(r,z,t|r_0,z_0) = S_r(t|r_0)\,, \tag{16}$$

$$S_{\text{trap}}(t|r_0,z_0) = \int_{-\infty}^\infty dz\, p(*_r,z,t|r_0,z_0) = \int_0^t dt' k_r(t'|r_0)\int_{R_z}^\infty dz'[G_z(z',t'|z_0)+G_z(-z',t'|z_0)]S_{\text{trap}}(t-t'|*_r,z'). \tag{17}$$

Note that these equations are formally exact. However, since $S_r(t|r_0)$ and therefore $k_r(t|r_0)$ are known only in the Laplace-transformed domain, we cannot obtain the closed-form expression of $S(t|r_0,z_0)$. By integrating Eq. (17) numerically, we can obtain the numerically exact results. For the numerical integration of Eq. (17), the following recent approximation for $S_r(t|r_0)$ in the time domain, which was reported to be accurate for all times,[23] may be useful:

$$S_r(t|r_0) \cong \min\left[1-\sqrt{\frac{R_r}{r_0}}\,\text{erfc}\left(\frac{r_0-R_r}{2\sqrt{D_r t}}\right), \frac{\ln(r_0/R_r)}{\ln(\sqrt{\pi t/\tau_r}+1)}\right]. \tag{18}$$

Normalization conditions are given by

$$S_{\text{free}}(t|r_0,z_0)+S_{\text{trap}}(t|r_0,z_0)+p(*_r,*_z,t|r_0,z_0)=1\,, \tag{19}$$

$$S_{\text{trap}}(t|*_r,z_0)+p(*_r,*_z,t|*_r,z_0)=1\,. \tag{20}$$

Inserting Eq. (11) into Eq. (20), we have

$$p(*_r,*_z,t|*_r,z_0) = H(|z_0|-R_z)\text{erfc}\left(\frac{|z_0|-R_z}{\sqrt{4D_{\text{trap}}t}}\right) + H(R_z-|z_0|)\,. \tag{21}$$

Results and Discussions

To confirm our theoretical predictions, we perform lattice-

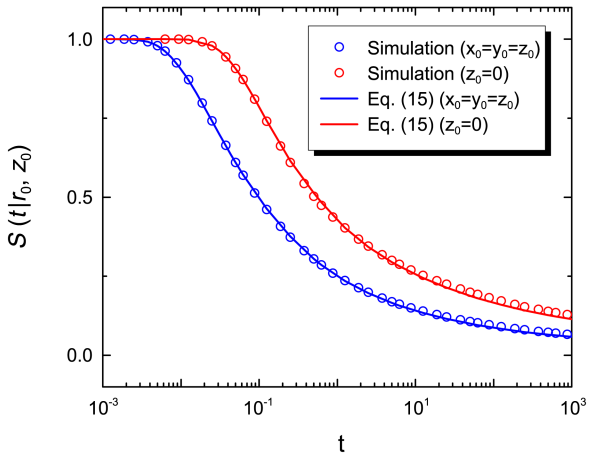

Figure 2. The time-dependent survival probability of a protein. Simulation results (open circles) are compared with those (solid lines) of Eq. (15) for two different initial conditions: at $x_0 = y_0 = z_0$ (blue) and on the xy-plane (red).

based Monte Carlo simulations for the present model system.[24] An initially implanted particle moves by random walk in 3D lattice. When the particle reaches the cylindrical trap, the trapping occurs if $|z| > R_z$ or the binding reaction with the reactive cylinder occurs if $|z| \le R_z$. Once trapped particle diffuses along the cylinder until the binding reaction occurs. The lattice constant should be small enough to reproduce the known theoretical results. All simulations are done for more than 10^8 time-steps ($t \sim 10^3$) and at least 10^5 independent trajectories are averaged.

Since survival probabilities depend on the initial position, we select two distinct initial positions with the same $\sqrt{r_0^2+z_0^2} = 1.5$: at $x_0 = y_0 = z_0$ ($r_0 = 1.22$ and $z_0 = 0.87$) and on the xy-plane ($r_0 = 1.5$ and $z_0 = 0$). In both cases, diffusion constants are the same ($D_r = D_z = D_{\text{trap}} = 1$) and $R_r = R_z = 1$. In Figure 2, simulation results for $S(t|r_0,z_0)$ are compared with numerical solutions of Eqs. (15)-(17). One can see that two results are nearly identical. We believe that the small deviations at long times come from the numerical integration of Eq. (17). As expected, the survival probability decreases more rapidly when the initial position is closer to the cylindrical trap.

To study effects of the facilitated diffusion, we compare the theoretical predictions of above two cases with other data in Figure 3. First, simulation results without the cylindrical trap are compared since no analytical result is known. The difference shows the facilitated diffusion effects. The biggest difference is the existence of an escape probability. When the dimensionality is larger than 2, there is always an escape probability, which is realized by the plateau region in the figure. Adding an infinite cylindrical trap to the system, the dimensionality effectively reduces to 2 and the protein always reaches the DNA strand and eventually the specific binding site at long times. Therefore, the ratio of the reaction rate in the normal diffusion case to that in the facilitated diffusion case goes up to infinity in the macroscopic time limit.

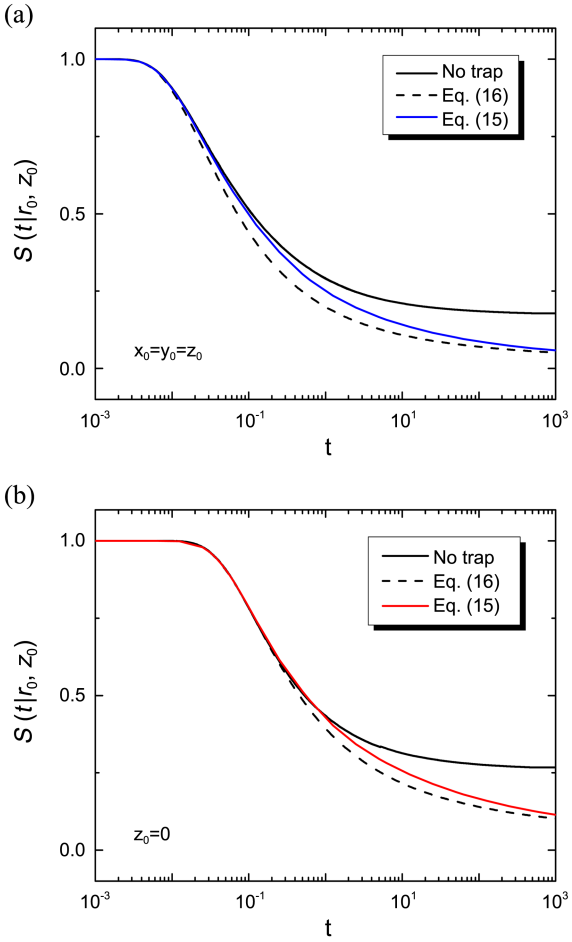

Figure 3. The time-dependent survival probability of a protein for two initial conditions: (a) $x_0 = y_0 = z_0$ (b) $z_0 = 0$. The blue and red solid lines are the same as those in Figure 2. The black solid lines are obtained from the simulation without an infinite cylindrical trap. The dashed lines are obtained numerically from Eq. (16).

For comparison, we also plot $S_r(t|r_0)$ or $S_{\text{free}}(t|r_0,z_0)$, which can be evaluated from Eq. (9) or from simulation results. We confirm that the simulation results reproduce Eq. (9). At short times, the results of $S(t|r_0,z_0)$ seems similar to those of $S_r(t|r_0)$. In the intermediate time region, the difference between $S(t|r_0,z_0)$ and $S_r(t|r_0)$, namely, $S_{\text{trap}}(t|r_0,z_0)$ becomes larger. In the long time limit, $S(t|r_0,z_0)$ apparently converges to $S_r(t|r_0)$ again. Therefore, $S(t|r_0,z_0)$ strongly depends on $S_r(t|r_0)$, in other words, the specific binding rate strongly depends on the nonspecific binding rate. This can be understood by the fact that the 1D searching rate of the trapped protein is much faster than the 2D trapping rate.

Conclusion

We have presented a rigorous theoretical approach for the celebrated facilitated diffusion problem. The theoretical method can provide numerically exact results for a simplified model by solving the coupling effect between one- and three-dimensional diffusive motions. The theoretical results are confirmed by the lattice-based Monte Carlo simulations.

The rate enhancement by the facilitated diffusion results from the dimensionality reduction. The three-dimensional diffusive motion reduces to two-dimensional plus one-dimensional motions. The escape probability in three dimensions goes to zero in the reduced dimensions. Therefore, the ratio of the reaction rate in the normal diffusion case to that in the facilitated diffusion case goes up to infinity in the macroscopic time scale. Even when the one-dimensional diffusion is much slower than three-dimensional diffusion,[17] the facilitated effect will appear in a different time region.

In reality, the cell environment is much more complicated. The cytoplasm is crowded and one-dimensional diffusion along the DNA strand suffers from obstacles and different short-range interactions. However, the key feature of the dimensionality reduction should not vanish. Diffusion usually results from many kinds of collisions in the presence of obstacles and if one-dimensional and three-dimensional "diffusions" can be assumed to include complicated effects, our theoretical predictions should be useful.

More thorough studies may include effects of the finite reactivity, the finite cylinder size, and so forth. Other kinds of dimensionality reductions can be studied in a similar way. Since more complicated two-dimensional diffusion-influenced reactions were reported,[25,26] these results can be applied to the present problem dealing with different mechanisms. These will be reported elsewhere.

Acknowledgments. K.P. thanks the Korea Institute for Advanced Study for providing computing resources (KIAS Center for Advanced Computation Linux Cluster System). This work was supported by research funds from Dong-A university.

References

1. Berg, O. G.; Winter, R. B.; von Hippel, P. H. *Biochemistry* **1981**, *20*, 6929.
2. von Hippel, P. H.; Berg, O. G. *Biophys. J.* **1987**, *51*, A360.
3. von Hippel, P. H.; Berg, O. G. *J. Biol. Chem.* **1989**, *264*, 675.
4. Zhou, H. X.; Szabo, A. *Phys. Rev. Lett.* **2004**, *93*, 178101.
5. Halford, S. E.; Marko, J. F. *Nucleic Acids Res.* **2004**, *32*, 3040.
6. Zhou, H. X. *Phys. Biol.* **2005**, *2*, R1.
7. Zhou, H. X. *Biophys. J.* **2005**, *88*, 1608.
8. Blainey, P. C.; Van Oijen, A. M.; Banerjee, A.; Verdine, G. L.; Xie, X. S. *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103*, 5752.
9. Klenin, K. V.; Merlitz, H.; Langowski, J.; Wu, C. X. *Phys. Rev. Lett.* **2006**, *96*, 018104.
10. Elf, J.; Li, G.-W.; Xie, X. S. *Science* **2007**, *316*, 1191.
11. Alsallaq, R.; Zhou, H.-X. *J. Chem. Phys.* **2008**, *128*, 115108.
12. Florescu, A. M.; Joyeux, M. *J. Chem. Phys.* **2009**, *130*.
13. Li, G.-W.; Berg, O. G.; Elf, J. *Nat. Phys.* **2009**, *5*, 294.
14. Mirny, L.; Slutsky, M.; Wunderlich, Z.; Tafvizi, A.; Leith, J.; Kosmrlj, A. *J. Phys. A: Math. Gen.* **2009**, *42*.
15. Lomholt, M. A.; van den Broek, B.; Kalisch, S. M. J.; Wuite, G. J. L.; Metzler, R. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106*, 8204.
16. Bénichou, O.; Grebenkov, D.; Levitz, P.; Loverdo, C.; Voituriez, R. *Phys. Rev. Lett.* **2010**, *105*, 150606.
17. Florescu, A. M.; Joyeux, M. *J. Phys. Chem. A* **2010**, *114*, 9662.
18. Bénichou, O.; Chevalier, C.; Klafter, J.; Meyer, B.; Voituriez, R. *Nat. Chem.* **2010**, *2*, 472.
19. Bénichou, O.; Chevalier, C.; Meyer, B.; Voituriez, R. *Phys. Rev. Lett.* **2011**, *106*, 038102.
20. Duffy, D. G. *Green's Functions with Applications*; CRC Press: 2001; Vol. 38.
21. Carslaw, H. S.; Jaeger, J. C. *Conduction of Heat in Solids*; 2nd ed.; Oxford University Press: New York, 1986.
22. *IMSL FORTRAN Library*; IMSL: Houston, 1989.
23. Kim, H. *Chem. Phys. Lett.* **2011**, *507*, 265.
24. Kim, H.; Shin, S.; Lee, S.; Shin, K. J. *J. Chem. Phys.* **1996**, *105*, 7705.
25. Park, K.; Shin, K. J.; Kim, H. *J. Chem. Phys.* **2009**, *131*, 154105.
26. Park, K.; Shin, K. J.; Kim, H. *Chem. Asian J.* **2010**, *5*, 1213.