THE JOURNAL OF PHYSICAL CHEMISTRY Letters

Letter

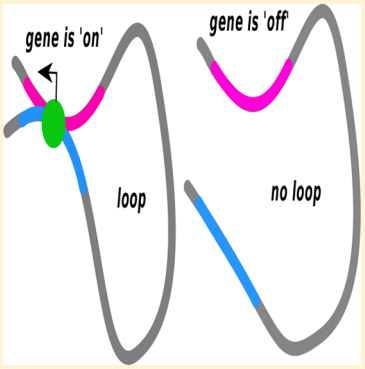
pubs.acs.org/JPCL

The Role of DNA Looping in the Search for Specific Targets on DNA by Multisite Proteins

Alexey A. Shvets and Anatoly B. Kolomeisky*

Department of Chemistry and Center for Theoretical Biological Physics, Rice University, Houston, Texas 77005, United States

ABSTRACT: Many cellular processes involve simultaneous interactions between DNA and protein molecules at several locations. They are regulated and controlled by special protein−DNA complexes, which are known as synaptic complexes or *synaptosomes*. Because of the multisite nature of involved proteins, it was suggested that during the formation of synaptic complexes DNA loops might appear, but their role is unclear. We developed a theoretical model that allowed us to evaluate the effect of transient DNA loop formation. It is based on a discrete-state stochastic method that explicitly takes into account the free-energy contributions due to the appearance of DNA loops. The formation of the synaptic complexes is viewed as a search for a specific binding site on DNA by the protein molecule already bound to DNA at another location. It was found that the search might be optimized by varying the position of the target and the total length of DNA. Furthermore, the formation of transient DNA loops leads to faster dynamics if it is associated with favorable enthalpic contributions to nonspecific protein−DNA interactions. It is also shown that DNA looping might reduce stochastic noise in the system.



It is known that multiple genetic modifications in cells require interactions between spatially distant DNA regions, and they are accomplished by a process known as a *site-specific recombination*.[1−3] These interactions are controlled by specialized proteins, or protein assemblies, that are responsible for the formation of site-specific protein−DNA *synaptic complexes* (or *synaptosomes*).[2,3] In fact, the formation of synaptosomes is a general biological phenomenon that is found in multiple important processes such as gene regulation, genome rearrangements via various mechanisms of site-specific recombination, and insertion of foreign genomes (e.g., genome integration).[2−6] Despite the fundamental importance of these processes, very little is known about the molecular mechanisms that lead to the formation of protein−DNA synaptic complexes.[2]

Although the protein sizes and the complexity of the site-specific recognition systems might vary, there are some common features in the formation of synaptosomes in different systems. The proteins involved in the process have several different sites to associate to different regions on DNA, and initially the protein molecule binds only to one DNA duplex, forming a so-called *presynaptic complex*. In the next steps, the other specific sites must be found, most probably sequentially.[1,3] This dynamic process is taking place *with* the protein molecule already bound to DNA, as illustrated in a Figure 1, and topologically complex protein−DNA structures might be involved. For example, as illustrated in Figure 1, DNA loops might appear and disappear.

These arguments indicate that the formation of a synaptic complex can be viewed as a multisite protein search for a specific site on DNA, while being already associated with the same DNA chain at another location. Protein target search phenomena have been intensively investigated in the last 40 years using a variety of experimental and theoretical

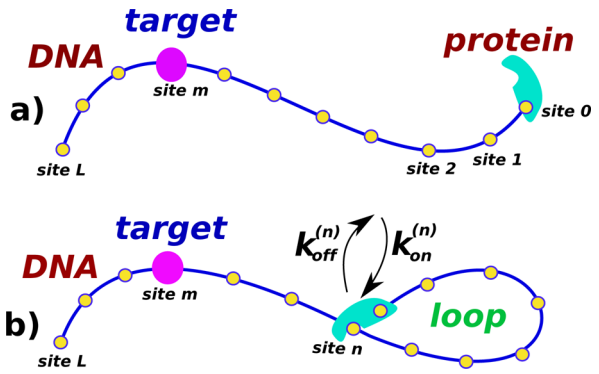


Figure 1. Schematic view of the search process with protein-assisted DNA looping. The second target already bound to the protein molecule is at the beginning of the DNA chain. The search is taking place via the formation of loops. No slidings in the loop configurations are allowed. (a) The conformation without DNA loop. (b) The conformation with the DNA loop.

methods.[3,7−9,12−27] It is now well-accepted that proteins find their specific sites on DNA by alternating between three-dimensional bulk solution diffusion and one-dimensional slidings along the DNA chains.[3,7−9] However, existing theoretical models for the protein search for targets on DNA cannot be directly applied for explaining the development of synaptosomes because they were developed for single-site proteins.[3,7,8] The possibility of different topological structures that might affect protein−DNA interactions has not been taken

Received: October 12, 2016
Accepted: November 18, 2016
Published: November 18, 2016

DOI: 10.1021/acs.jpclett.6b02371
J. Phys. Chem. Lett. 2016, 7, 5022−5027

into account in current theoretical approaches. The formation of a synaptic complex is accomplished by multisite proteins, and this should lead to a very different dynamics.

In this Letter, we develop a new theoretical approach that explicitly accounts for the formation of the simplest topological structures, DNA loops, during the development of the protein−DNA synaptic complexes. Our idea is to evaluate the free-energy change in the system due to the polymer loop creation, which should modify the transition rates for all other relevant processes. Our analytical calculations, which are also supported by computer simulations, indicate that the appearance of DNA loops might strongly influence the search dynamics and lower the stochastic noise in the system.

In our theoretical approach, we start by analyzing a discrete-state stochastic model, as shown in Figure 1. Specifically, we consider a single DNA molecule with $L + 1$ binding sites, with two of them being the target sites for the formation of the synaptosome. One of the targets is at the site m, while the second one generally can be anywhere on the DNA chain (in Figure 1, it is at the beginning of the DNA chain labeled as site 0). To simplify calculations, we assume that the protein molecule has two binding sites: one of them is free, and the other one is already bound (see Figure 1). The analysis can be easily extended to proteins with more than two DNA-binding sites. The protein bound to DNA at the site 0 is trying to find the second target at the position m. This process can be associated with first-passage phenomena, providing us with the analytical tools to obtain a full dynamic description of the system.[9,21,22]

To analyze the dynamics of the multisite protein search, the following simple model is considered. We assume that DNA transient loops can form, but the sliding in the looped configurations is not possible (see Figure 1). The dynamic behavior of the system is fully described by first-passage probability density functions $F_n(t)$ defined as probabilities to reach the target at the site m for the first time at time t given that at $t = 0$ the protein was in the state n. The state n corresponds to a configuration when the transient loop is formed by the protein binding to the site n (Figure 1b). The state 0 corresponds to the situation of no loops (Figure 1a). The temporal evolution of the first-passage probabilities is governed by a set of backward master equations

$$\frac{dF_n(t)}{dt} = k_{off}^{(n)}F_0(t) - k_{off}^{(n)}F_n(t) \tag{1}$$

for $n \neq 0$, m, and

$$\frac{dF_0(t)}{dt} = \sum_{i=1}^{L} k_{on}^{(i)}F_i(t) - F_0(t)\sum_{i=1}^{L} k_{on}^{(i)} \tag{2}$$

In addition, the initial condition implies that

$$F_m(t) = \delta(t) \tag{3}$$

The rates $k_{on}^{(n)}$ and $k_{off}^{(n)}$ describe the formation and breaking of the transient loop at the site n, and they are generally position-dependent. Because there is no sliding, from the state n the protein can only dissociate back into the state 0 (no loops). From the state 0, the protein can bind to any site on DNA, forming the loop. The important factor here is that the formation of the loop has a free-energy cost ΔG_n, which allows us to relate the binding and unbinding rates via the detailed balancelike arguments[9,22]

$$\frac{k_{on}^{(n)}}{k_{off}^{(n)}} = \frac{k_{on}}{k_{off}}\exp(-\Delta G_n) \tag{4}$$

where k_{on} and k_{off} describe the binding and unbinding processes for a hypothetical situation in which the loop formation does not lead to the free-energy change. One can uncouple the transition rates by introducing a parameter $0 \leq \theta \leq 1$ that specifies how the protein−DNA interaction energy is distributed between the association and dissociation transitions.[22] Thus, we have

$$k_{on}^{(n)} = k_{on}e^{-\theta\Delta G_n}, \qquad k_{off}^{(n)} = k_{off}e^{(1-\theta)\Delta G_n} \tag{5}$$

To proceed further, we need to evaluate the free energy associated with creating DNA loops. It can be argued that this free energy can be written as

$$G_n = \varepsilon_n + E_n^{elast} + E_n^{entrop} = \varepsilon_n + g_n \tag{6}$$

where ε_n is an enthalpic term due to making the nonspecific bond after the loop is formed at the site n; the second term, E_n^{elast}, is the elastic contribution to the free energy; and the last term, E_n^{entrop}, is of the entropic nature. We also define g_n as a sum of elastic and entropic contributions to the free energy.

For the enthalpic term, we assume that it does not depend on the DNA sequence, i.e., $\varepsilon_n = \varepsilon$. There are many factors that contribute to the elastic energy of the loop, but we assume that this is mainly the result of bending of the polymer. Then for DNA at typical physiological conditions it can be well-approximated as

$$E_n^{elast} \simeq 3000/n \tag{7}$$

in units of k_BT.[10] Here n is the size of the formed DNA loop (see Figure 1b). The last term in the expression for the free energy is directly related to the conformational entropy, S, of a polymer chain (in units of k_B). It can be estimated using the following arguments. The total number of possible conformations of the Gaussian chain of length $L + 1$ monomers is Ω, and the corresponding entropy is equal to $S = \ln \Omega$. Correspondingly, ΔS_{loop} is the entropy change due to the formation of the loop, and Ω_{loop} is the number of possible conformations when the polymer is looped. Hence

$$\Delta S_{loop} = \ln\frac{\Omega_{loop}}{\Omega} \tag{8}$$

For a random coil, the end-to-end length distribution for the long chain of length $L + 1 \approx L$ is given by the Gaussian distribution[11]

$$P(R) = \frac{1}{(2\pi L)^{3/2}}\exp\left(\frac{-3\mathbf{R}^2}{2L}\right) \tag{9}$$

where we took into account that the monomer size is unity, and $\mathbf{R}$ is the end-to-end distance vector. Integrating the last expression over the all space for the free chain and over the volume of the loop formation region, $V_d \simeq 1$ (ends are in the proximity with each other), we obtain

$$\frac{\Omega_{loop}}{\Omega} = \frac{\int d\mathbf{R}P(R)}{1} = \frac{V_d}{(2\pi n)^{3/2}} \tag{10}$$

Thus, the entropy change can be written in the form

DOI: 10.1021/acs.jpclett.6b02371
J. Phys. Chem. Lett. 2016, 7, 5022−5027

$$\Delta S_{\text{loop}} \simeq -\frac{3}{2}\ln n \tag{11}$$

where we neglect the constant term.

Combining all results for the different contributions to the free energy, one can obtain the final expression for the free-energy change due to the formation of DNA loop

$$\Delta G_n \simeq \varepsilon + \frac{3000}{n} + \frac{3}{2}\ln n \tag{12}$$

This equation is also closely related with a Jacobson–Stockmayer factor, or *J*-factor, that is frequently used to describe polymer looping phenomena.[10,30,31] The calculated free-energy change as a function of the DNA loop size is presented in Figure 2 for realistic sets of parameters. The crucial observation here is that there is an optimal size of the loop which is associated with the minimal free-energy cost. This will be important in analyzing the dynamics of the system as we discuss below.

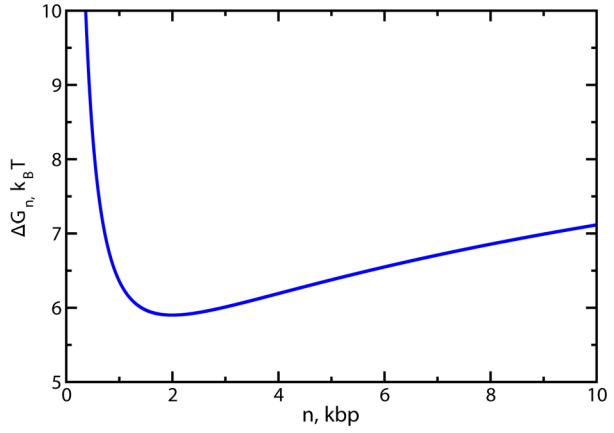


Figure 2. Free-energy cost of the loop formation as a function of the loop contour length, n. Parameters used for calculations are $\varepsilon = -7k_BT$ and the DNA chain length is equal to $L = 10$ kbp.

Now we can solve eqs 1 and 2 by using the Laplace transformation, $\tilde{F}_n(s) = \int_0^\infty F_n(t)e^{-st}\,dt$, which leads to a simple expression

$$\tilde{F}_n(s) = \frac{k_{\text{off}}^{(n)}\tilde{F}_0(s)}{s + k_{\text{off}}^{(n)}} \tag{13}$$

for $n > 0$. For $n = 0$, we utilize the fact that at the target $\tilde{F}_m(s) = 1$ (see eq 3), yielding

$$\tilde{F}_0(s) = \frac{k_{\text{on}}^{(m)}}{s + f(s)} \tag{14}$$

Here, an auxiliary function was introduced

$$f(s) = k_{\text{on}}^{(m)} + \sum_{i=0}^{m-1} k_{\text{on}}^{(i)}\left(1 - \frac{k_{\text{off}}^{(i)}}{s + k_{\text{off}}^{(i)}}\right) + \sum_{i=m+1}^{L} k_{\text{on}}^{(i)}\left(1 - \frac{k_{\text{off}}^{(i)}}{s + k_{\text{off}}^{(i)}}\right) \tag{15}$$

From the knowledge of the first-passage probability functions, all dynamic properties of the system can be easily computed. More specifically, the mean search time of finding the target is associated with T_0, which is given by

$$T_0 = -\frac{\partial F_0(s)}{\partial s}\bigg|_{s=0} = \frac{1 + \sum_{i=1}^{L}(k_{\text{on}}^{(i)}/k_{\text{off}}^{(i)}) - k_{\text{on}}^{(m)}/k_{\text{off}}^{(m)}}{k_{\text{on}}^{(m)}} \tag{16}$$

Using eq 5, we can rewrite this expression as (assuming $k_{\text{on}} = k_{\text{off}}$)

$$T_0 = \frac{1 + \exp(-\varepsilon)[\sum_{i=1}^{L} e^{-\Delta g_i} - e^{-\Delta g_m}]}{k_{\text{on}}\, e^{-\theta(\varepsilon + \Delta g_m)}} \tag{17}$$

If the binding energy is strongly attractive, $\varepsilon \to -\infty$, and the elastic and entropic terms are not large, $\Delta g_i \simeq 1$, then the mean search time can be approximated as

$$T_0 \simeq \frac{L\, e^{-\varepsilon}}{k_{\text{on}}\, e^{-\theta\varepsilon}} \tag{18}$$

The physical meaning of this expression is that the target search time is given by the time to create one loop $(1/k_{\text{on}}e^{-\theta\varepsilon})$ multiplied by the number of attempts (L) and the Boltzmann's factor of being found in the looped states. In this limit, the search time is very large because the protein is frequently trapped in various looped states. In another limit of strong repulsions, $\varepsilon \to \infty$, and weak elastic and entropic contributions, we obtain

$$T_0 \simeq \frac{1}{k_{\text{on}}}\, e^{\theta\varepsilon} \tag{19}$$

Here the search is equal to the dwell time to be in the free unlooped state, which is also very large. The formation of loops is energetically unfavorable, and this prevents the formation of the loop and eventually finding the target. These arguments suggest that there is an intermediate binding strength that minimize the search times.

The analysis of the search dynamics as a function of the position of the target can now be accomplished. The results are presented in Figure 3. One can see that there is an optimal target position with the fastest search dynamics. It directly corresponds to the minimum in the free-energy cost of the

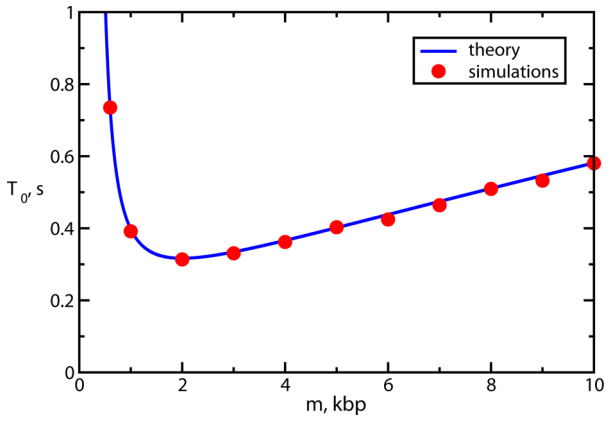


Figure 3. Average search times as a function of the target position. Parameters used for calculations are $k_{\text{on}} = 10^2\ \text{s}^{-1}$ and $\varepsilon = -7k_BT$, and the DNA chain length is $L = 10$ kbp. Symbols correspond to Monte Carlo computer simulations, and the solid curve is the theoretical prediction.

DOI: 10.1021/acs.jpclett.6b02371
J. Phys. Chem. Lett. 2016, 7, 5022–5027

creating DNA loops. The following arguments can be presented to explain this connection. If the target is close to the position of the protein on DNA ($n = 0$), there is a significant penalty for creating such short loops, and this makes the search times very large. This penalty is due to the elastic term in the free energy. It costs a lot of energy to make very short loops in the semiflexible polymers like DNA. If the target is far from the protein ($n \approx L$), then it is entropically unfavorable to make long loops, and again the search is slow. Only for the target positions that are associated with the smallest free-energy cost is the dynamics fast.

A qualitatively similar behavior is observed when we investigate the search dynamics for varying lengths of DNA molecules, as shown in Figure 4. It is difficult for the protein to find the target for very short DNA chains because it is energetically very unfavorable to make short loops because of the elastic term in the free energy. A slow search is also observed for very long DNA chains because of entropic considerations for creating long loops. The most optimal dynamics is observed for intermediate DNA chains when the loop formation is more favorable.

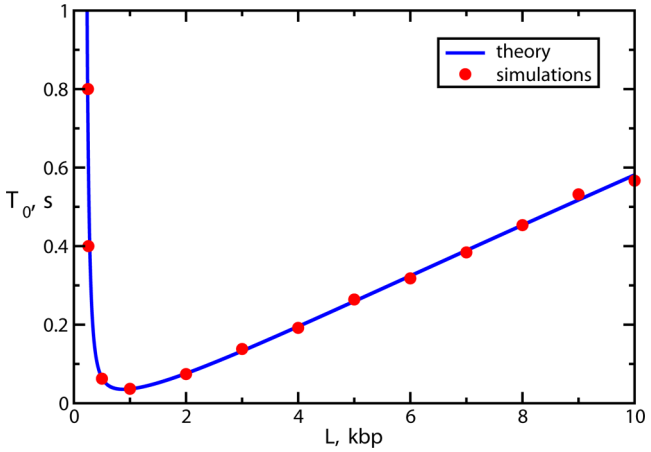


Figure 4. Average search times as a function of the DNA chain size. Parameters used for calculations are $k_{on} = 10^2$ s^{-1}, and the target position is at $m = L$.

The important question is how explicitly the formation of DNA loops affects the dynamics of locating the specific targets. We can analyze it by comparing the multisite protein search process, which involves DNA loops, with the single-site protein search process without the formation of loops. A quantitative comparison can be done because in our theoretical framework the case without loops corresponds to $g_i = 0$, i.e., no elastic or entropic contributions to the free energy. But the enthalpic term is still there because the single-site protein might nonspecifically bind to the DNA chain during the search. Dynamic properties for the systems with loops and without loops are illustrated in Figures 5 and 6. One can see that the search is generally faster with DNA loops for attractive nonspecific interactions between protein and DNA ($\varepsilon < 0$) as compared to the search without loops. The effect is stronger for longer DNA chains (see Figure 6). However, the search without loops is more efficient for repulsive interactions (Figure 5). There is also the optimal interactions strength that minimizes the search dynamics for all systems. In addition, the behavior is independent of the position of the target site.

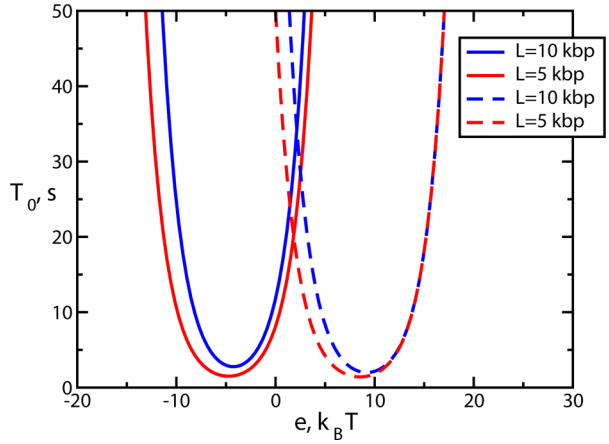


Figure 5. Average search times as a function of enthalpic binding energy for the search with loops (solid lines) and without loops (dashed lines). Parameters used for calculations are $k_{on} = 10^2$ s^{-1}, $L = 10$ kbp, and the target position $m = L$.

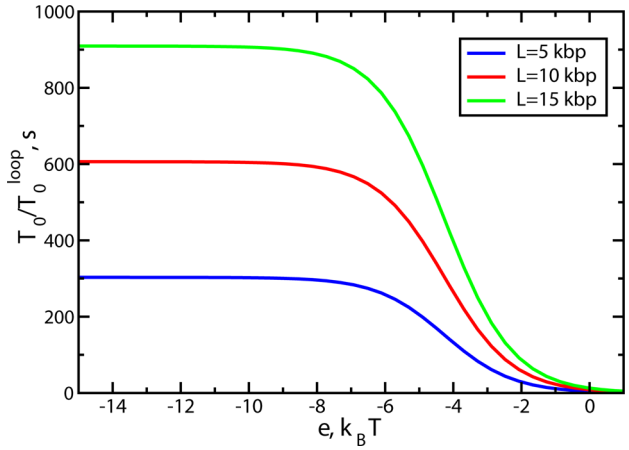


Figure 6. Ratio of search times for the system with loops and without loops as a function of the enthalpic binding energy. Parameters used for calculations are $k_{on} = 10^2$ s^{-1}, $L = 10$ kbp, and the target position $m = L$.

The following physical explanations for these observations can be presented. To find the specific site fast, the protein must quickly associate to DNA and dissociate back into the solution in order to fully scan the DNA chain. In the system with DNA loops, the negative enthalpic energies compensate for positive contributions from the elastic and entropic terms (see eq 11), and the association−dissociation dynamics is fast. For the system without loops, attractive interactions decrease the mobility of protein molecules, slowing the search. The effect is reversed for repulsive interactions when the highest mobility is observed for the protein in the system without loops. In the system with loops it takes a lot of time for the protein to form the loop, which is not efficient for the search dynamics. For realistic conditions in cells, the nonspecific protein−DNA interactions are weakly attractive, and these arguments suggest that transient DNA loops can significantly accelerate the formation of essential protein−DNA complexes.

It is widely accepted that stochastic noise strongly influences many biological processes.[1,10] In our system, this might be associated with a width in the distribution of the search times to find the target. The larger the noise, the wider the distribution. It is reasonable to expect that cells are trying to minimize this

DOI: 10.1021/acs.jpclett.6b02371
J. Phys. Chem. Lett. 2016, 7, 5022−5027

distribution in order to better regulate the following biochemical and biophysical processes. Because our theoretical method provides a comprehensive description of the dynamics, we can analyze the effect of DNA loop formation on the stochastic noise for multisite protein search. To do this quantitatively, one can introduce a dimensionless function

$$\delta = \frac{T_{02} - T_0^2}{T_0^2} \tag{20}$$

where the second moment of the search time, T_{02}, is determined via

$$T_{02} = \int_0^\infty t^2 F_0(t)\,dt = \left.\frac{\partial^2 \tilde{F}_0}{\partial s^2}\right|_{s=0} \tag{21}$$

The function δ is a normalized variance, and it is related to a width of the distribution of search times around the mean search time. The noise is stronger for larger values of δ.

To evaluate explicitly the normalized variance, one can use eq 14, yielding

$$T_{02} = \frac{2}{k_{on}^{(m)2}}(1 + f'(0))^2 - \frac{1}{k_{on}^{(m)}} f''(0) = 2T_0^2 - \frac{1}{k_{on}^{(m)}} f''(0) \tag{22}$$

Therefore

$$\delta = 1 - \frac{1}{k_{on}^{(m)} T_0^2} f''(0) \tag{23}$$

The auxiliary function $f(s)$ is defined in eq 15, and its second derivative at $s = 0$ is given by

$$f''(0) = -2\left(\sum_{i=1}^{L} \frac{k_{on}^{(i)}}{k_{off}^{(i)2}} - \frac{k_{on}^{(m)}}{k_{off}^{(m)2}}\right) \tag{24}$$

Combining all results together, we obtain the explicit expression

$$\delta = 1 + 2\frac{\sum_{i=1}^{L} (k_{on}^{(m)} k_{on}^{(i)})/k_{off}^{(i)2} - (k_{on}^{(m)}/k_{off}^{(m)})^2}{(1 + \sum_{i=1}^{L} k_{on}^{(i)}/k_{off}^{(i)} - k_{on}^{(m)}/k_{off}^{(m)})^2} \tag{25}$$

Finally, this equation can be rewritten as

$$\delta = 1 + 2e^{-2\varepsilon}\frac{[\sum_{i=1}^{L} e^{-2\Delta g_i - \theta(\Delta g_m - \Delta g_i)} - e^{-2\Delta g_m}]}{(1 + e^{-\varepsilon}(\sum_{i=1}^{L} e^{-\Delta g_i} - e^{-\Delta g_m}))^2} \tag{26}$$

If the enthalpic contributions to the free energy dominate, while elastic and entropic terms are small, it can be shown that this expression takes the following form:

$$\delta = 1 + 2\frac{e^{-2\varepsilon}(L-1)}{(1 + e^{-\varepsilon}(L-1))^2} \tag{27}$$

Then, for very strong attractions, $\varepsilon \to -\infty$, the normalized variance is $\delta \simeq 1 + 2/(L-1)$. For strong repulsions, $\varepsilon \to \infty$, the noise is also small, $\delta \simeq 1$. The normalized variance for different sets of conditions is presented in Figure 7. One can see that for realistic conditions of weak attractive protein–DNA nonspecific interactions the formation of DNA loops strongly reduces the stochastic noise in the system (see Figure 7).

We developed a theoretical approach that allowed us to analyze the role of DNA loops during the multisite protein search for specific targets on DNA. Our discrete-state stochastic

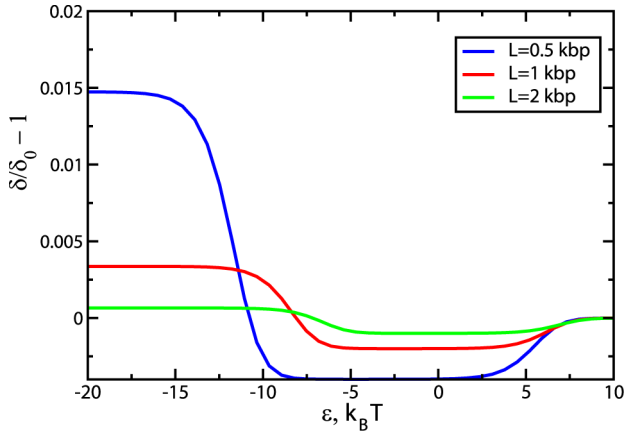


Figure 7. Normalized variance as a function of the enthalpic binding energy for the search with loops and for the search without loops for different DNA chain lengths. Parameters used for calculations are the target position $m = L$. δ_0 is the variance for the search without loops.

model explicitly takes into account the free-energy changes due to the formation of DNA loops, which modifies accordingly the transitions in the system. It is found that the possibility of creating transient DNA loops might have a very strong effect on the dynamics of formation of protein–DNA complexes. Theoretical calculations show a nonmonotonic behavior of search times as a function of the target position and the size of the DNA chain. This behavior is explained by the formation of DNA loops with lowest free-energy cost. Our analysis also indicates that for realistic cellular conditions protein search is faster when DNA loops can form. In addition, the creation of loops significantly reduces the level of the stochastic noise in the system, which might be beneficial for many biological processes.

Although a presented theoretical method probably captures main features of the formation of DNA loops during the search of multisite proteins, it is important to note the approximate nature of our theoretical approach. Many essential processes, such as protein slidings in the looped configurations and the effect of DNA sequence on dynamics, are not taken into account. There are also more advanced descriptions of the free-energy changes associated with DNA looping.[28,29] However, our theoretical model can be viewed as a starting point in investigations on the role of the topological structures during the formation of protein–DNA complexes. It will be critically important to test our theoretical predictions with more advanced theoretical treatments as well as in experimental studies.

AUTHOR INFORMATION

Corresponding Author

*E-mail: tolya@rice.edu.

ORCID

Anatoly B. Kolomeisky: 0000-0001-5677-6690

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

The work was supported by the Welch Foundation (Grant C-1559), by the NSF (Grant CHE-1360979), and by the Center for Theoretical Biological Physics sponsored by the NSF

DOI: 10.1021/acs.jpclett.6b02371
J. Phys. Chem. Lett. 2016, 7, 5022–5027

(Grant PHY-1427654). We also would like to thank Dr. Y. Lyubchenko for useful comments and suggestions.

REFERENCES

(1) Alberts, B. et al. *Molecular Biology of the Cell*, 6th ed.; Garland Science: New York, 2014.
(2) Grindley, N. D.; Whiteson, K. L.; Rice, P. A. Mechanisms of Site-Specific Recombination. *Annu. Rev. Biochem.* **2006**, *75*, 567−605.
(3) Halford, S. E.; Marko, J. F. How Do Site-Specific DNA-Binding Proteins Find Their Targets? *Nucl. Acid Res.* **2004**, *32*, 3040−3052.
(4) Vetcher, A. A.; et al. DNA Topology and Geometry in Flp and Cre Recombination. *J. Mol. Biol.* **2006**, *357*, 1089−1104.
(5) Mani, R.-S.; Chinnaiyan, A. M. Triggers for Genomic Rearrangements: Insights into Genomic, Cellular and Environmental Influences. *Nat. Rev. Genet.* **2010**, *11*, 819−829.
(6) Bushman, F.; et al. Genome-wide Analysis of Retroviral DNA Integration. *Nat. Rev. Microbiol.* **2005**, *3*, 848−858.
(7) Mirny, L. A.; Slutsky, M.; Wunderlich, Z.; Tafvizi, A.; Leith, J. S.; Kosmrlj, A. How a Protein Searches for Its Site on DNA: The Mechanism of Facilitated Diffusion. *J. Phys. A: Math. Theor.* **2009**, *42*, 434013.
(8) Kolomeisky, A. B. Physics of Protein-DNA Interactions: Mechanisms of Facilitated Target Search. *Phys. Chem. Chem. Phys.* **2011**, *13*, 2088−2095.
(9) Veksler, A.; Kolomeisky, A. B. Speed-Selectivity Paradox in the Protein Search for Targets on DNA: Is It Real or Not? *J. Phys. Chem. B* **2013**, *117*, 12695−12701.
(10) Phillips, R.; Kondev, J.; Theriot, J. *Physical Biology of the Cell*, 2nd ed.; Garland Science: New York, 2013.
(11) Rubinstein, M.; Colby, R. H. *Polymer Physics*; Oxford University Press: New York, 2003.
(12) Berg, O. G.; Winter, R. B.; von Hippel, P. H. Diffusion-Driven Mechanisms of Protein Translocation on Nucleic Acids. 1. Models and Theory. *Biochemistry* **1981**, *20*, 6929−6948.
(13) Winter, R. B.; Berg, O. G.; von Hippel, P. H. Diffusion-Driven Mechanisms of Protein Translocation on Nucleic Acids. 3. The Escherichia Coli Lac Repressor-Operator Interaction: Kinetic Measurements and Conclusions. *Biochemistry* **1981**, *20*, 6961−6977.
(14) Gowers, D. M.; Wilson, G. G.; Halford, S. E. Measurement of the Contributions of 1D and 3D Pathways to the Translocation of a Protein along DNA. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 15883.
(15) Kolesov, G.; Wunderlich, Z.; Laikova, O. N.; Gelfand, M. S.; Mirny, L. A. How Gene Order is Influenced by the Biophysics of Transcription Regulation. *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104*, 13948.
(16) Wang, Y. M.; Austin, R. H.; Cox, E. C. Single Molecule Measurements of Repressor Protein 1D Diffusion on DNA. *Phys. Rev. Lett.* **2006**, *97*, 048302.
(17) Elf, J.; Li, G. W.; Xie, X. S. Probing Transcription Factor Dynamics at the Single-Molecule Level in a Living Cell. *Science* **2007**, *316*, 1191−1194.
(18) Tafvizi, A.; Huang, F.; Leith, J. S.; Fersht, A. R.; Mirny, L. A.; van Oijen, A. M. Tumor Suppressor p53 Slides on DNA With Low Friction and High Stability. *Biophys. J.* **2008**, *95*, L01−L03.
(19) Hammar, P.; Leroy, P.; Mahmutovic, A.; Marklund, E. G.; Berg, O. G.; Elf, J. The Lac Repressor Displays Facilitated Diffusion in Living Cells. *Science* **2012**, *336*, 1595−1598.
(20) Zandarashvili, L.; Vuzman, D.; Esadze, A.; Takayama, Y.; Sahu, D.; Levy, Y.; Iwahara, J. Asymmetrical Roles of Zinc Fingers in Dynamic DNA-Scanning Process by the Inducible Transcription Factor Egr-1. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109*, E1724.
(21) Shvets, A. A.; Kochugaeva, M.; Kolomeisky, A. B. The Role of Static and Dynamic Obstacles in the Protein Search for Targets on DNA. *J. Phys. Chem. B* **2016**, *120*, 5802−5809.
(22) Shvets, A. A.; Kolomeisky, A. B. Sequence heterogeneity accelerates protein search for targets on DNA. *J. Chem. Phys.* **2015**, *143*, 245101.
(23) Marcovitz, A.; Levy, Y. Obstacles May Facilitate and Direct DNA Search by Proteins. *Biophys. J.* **2013**, *104*, 2042−2050.
(24) Koslover, E. F.; de la Rosa, M. A. D.; Spakowitz, A. J. Theoretical and Computational Modeling of Target-Site Search Kinetics in Vitro and in Vivo. *Biophys. J.* **2011**, *101*, 856−865.
(25) Sheinman, M.; Benichou, O.; Kafri, Y.; Voituriez, R. Classes of Fast and Specific Search Mechanisms for Proteins on DNA. *Rep. Prog. Phys.* **2012**, *75*, 026601.
(26) Tafvizi, A.; Huang, F.; Fersht, A. R.; Mirny, L. A.; van Oijen, A. M. A Single-Molecule Characterization of p53 Search on DNA. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108*, 563−568.
(27) Esadze, A.; Kemme, C. A.; Kolomeisky, A. B.; Iwahara, J. Positive and Negative Impacts of Nonspecific Sites During Target Location by a Sequence-Specific DNA-Binding Protein: Origin of the Optimal Search at Physiological Ionic Strength. *Nucleic Acids Res.* **2014**, *42*, 7039.
(28) Chen, Y. J.; Johnson, S.; Mulligan, P.; Spakowitz, A. J.; Phillips, R. Modulation of DNA loop lifetimes by the free energy of loop formation. *Proc. Natl. Acad. Sci. U. S. A.* **2014**, *111*, 17396−17401.
(29) Mulligan, P.; Chen, Y. J.; Phillips, R.; Spakowitz, A. J. Interplay of Protein Binding Interactions, DNA Mechanics, and Entropy in DNA Looping Kinetics. *Biophys. J.* **2015**, *109*, 618−629.
(30) Levene, S. D.; Giovan, S. M.; Hanke, A.; Shoura, M. J. The Thermodynamics of DNA Loop Formation, from J to Z. *Biochem. Soc. Trans.* **2013**, *41*, 513−518.
(31) Zhang, Y.; McEwen, A. E.; Crothers, D. M.; Levene, S. D. Statistical-mechanical theory of DNA looping. *Biophys. J.* **2006**, *90*, 1903−1912.

DOI: 10.1021/acs.jpclett.6b02371
J. Phys. Chem. Lett. 2016, 7, 5022−5027