DNA Looping and DNA Conformational Fluctuations Can Accelerate Protein Target Search

Published as part of The Journal of Physical Chemistry virtual special issue “Dave Thirumalai Festschrift”.

Cayke Felipe, Jaeoh Shin, and Anatoly B. Kolomeisky*

ABSTRACT: Protein searching and binding to specific sites on DNA is a fundamentally important process that marks the beginning of all major cellular transformations. While the dynamics of protein–DNA interactions in in vitro settings is well investigated, the situation is much more complex for in vivo conditions because the DNA molecules in live cells are packed into chromosomal structures where they are undergoing strong dynamic and conformational fluctuations. In this work, we present a theoretical investigation on the role of DNA looping and DNA conformational fluctuations in the protein target search. It is based on a discrete-state stochastic analysis that allows for explicit calculations of dynamic properties, which is also supplemented by Monte Carlo computer simulations. It is found that for stronger nonspecific interactions between DNA and proteins the search occurs faster on the DNA looped conformation in comparison with the unlooped conformation, and the fastest search is observed when the loop is formed near the target site. It is also shown that DNA fluctuations between the looped and unlooped conformations influence the search dynamics, and this depends on the magnitude of conformational transition rates and on which conformation is more energetically stable. Physical–chemical arguments explaining these observations are presented. Our theoretical study suggests that the geometry and conformational changes in DNA are additional factors that might efficiently control the gene regulation processes.

INTRODUCTION

DNA is the molecule that contains all the genetic information to support life. This information is “read” by protein transcription factors that bind to specific sites on DNA, initiating cascades of biochemical and biophysical processes. All these events together form a complex gene regulation system that is specific for each living organism. In recent years, significant progress has been achieved in uncovering dynamic properties of subcellular structures that are relevant for the flow of genetic information. These experimental advances support the idea that not only the DNA sequence but also its organization inside the living cells might play an important role in regulation of genetic processes. It was found, for example, that due to nonspecific interactions the protein molecules can be engaged in different searching modes. For strong interactions, the search is mostly one-dimensional (1D) since once the protein binds to DNA it diffuses along the chain until the target is found. For weak interactions, protein can only occasionally bind to DNA without sliding, and here the search can be viewed as a three-dimensional (3D) process. The most interesting behavior is observed for intermediate strengths of nonspecific interactions when the protein is engaged in a combination of 1D+3D search process. In this regime, the protein molecule binds to DNA, slides some distance on the DNA chain, dissociates back into the solution, and then the search can be viewed as a three-dimensional (3D) process. The most interesting behavior is observed for intermediate strengths of nonspecific interactions when the protein is engaged in a combination of 1D+3D search process. In this regime, the protein molecule binds to DNA, slides some distance on the DNA chain, dissociates back into the solution, and then diffuses in the bulk until the search cycle is repeated after rebinding to DNA. Multiple search cycles might be observed before the target is located, and it was found that in this regime...
the search could be the fastest, corresponding to a so-called facilitated diffusion.  

While the protein search dynamics has been considered before in simplified systems, which are relevant to in vitro situations, the problem has been extended in multiple ways to understand the dynamics in more realistic cellular environment. Several theoretical studies on the role of protein conformational dynamics, the effect of molecular crowding on DNA, and the role of the protein intermolecular interactions have been presented. Moreover, the role of sequence specificity of the protein–DNA interactions has been investigated. Additionally, the role of DNA conformation on the search dynamics has been studied both experimentally and theoretically, which is particularly important as multiple studies have shown that the genomic organization might strongly influence the protein search in live cells.

DNA loops are frequently found in both prokaryotic and eukaryotic cells, being formed actively by different proteins such as CTCF or condensin. While these DNA loops might work as repressors by sterically blocking the protein transcription factors from finding to their target sites or just allowing another repressor molecule to bind easier, most of them are associated with enhancer–promoter activity. Additionally, DNA loops may create either a set of static (fixed) loops or dynamic looping interactions that affect all processes in the system. It was found that the geometry of DNA chains in the chromatin configurations affects the dynamics of interactions with protein transcription factors. However, the microscopic details of how the topological state of DNA and its conformational dynamics are influencing the protein target search remain not well understood.

In this work, we present a theoretical study on the role of the two most important features of DNA in live cells, formation of loops and conformational fluctuations, in the protein finding specific target sites on DNA. Using a discrete-state stochastic model and Monte Carlo computer simulations, we investigate how the search process can be facilitated in the looped configuration in comparison with linear configuration without loops. Our goal is to consider a relatively simple model of protein target search in order to clarify the molecular picture of underlying processes. We found that search dynamics depends on the strength of nonspecific protein–DNA interactions and on the distance between the target and the loop intersection positions. In addition, the effect of conformational transitions between DNA looped and unlooped states has been investigated, and we suggest that the most efficient search is observed for fast DNA looping/unlooping transitions. The importance of our theoretical findings for understanding the mechanisms of genetic regulation is also discussed.

**THEORETICAL MODEL**

**Search on DNA Loop Conformations.** Let us start by investigating a model of the protein search for the target on the looped DNA chain as illustrated in Figure 1a. To simplify theoretical calculations, we consider a single DNA and a single protein molecule that is searching for the target located inside the DNA chain. But the real concentrations of protein and DNA molecules can be easily recovered from our single-molecule approach. The DNA molecule is viewed as a lattice segment consisting of sites. In our theoretical model, one lattice site corresponds to approximately 10 base pairs (bp), which is a typical size of the DNA segment covered by the bound protein molecule, and it is also the average size of the target sequences on DNA. The end of the DNA molecule at the sites is coupled to the site, which yields a DNA loop of size (Figure 1a). For , the DNA molecule becomes circular. The target is located on the site.

![Figure 1](image)

**Figure 1.** (a) Schematic view of the model where the DNA loop is formed by coupling at the sites and . The green square located at the site is the target site. The magenta circle is the searching protein. Black arrows indicate allowed transitions to different states. (b) Simplified view of the conformational transitions in the DNA molecule. The free DNA (left) can form a loop with rate while the looped DNA (right) can go back to its linear configuration with a rate .
label as \( n (0 \leq n \leq L) \). The state \( n = 0 \) corresponds to the protein being in the bulk solution, while the states \( 1 \leq n \leq L \) describe the protein on the site \( n \) of the DNA chain (Figure 1a). One can define then a probability density function, \( F_{\alpha}(t) \), of reaching the target site \( n \) at time \( t \) for the first time given that, initially, it was an the site \( n \). The temporal evolution of these first-passage probability functions is controlled by the following backward master equations,

\[
\frac{\partial F_{\alpha}(t)}{\partial t} = k_{\alpha f} F_{\alpha}(t) + u[F_{\alpha-1}(t) + F_{\alpha+1}(t)]
\]

\[
- (k_{\alpha f} + 2u)F_{\alpha}(t)
\]

for \( n = 2, 3, 4, ..., L - 1 \) and \( n \neq l \). For the boundary and special states \( (n = 1, l, L) \), we have

\[
\frac{\partial F_{\alpha}(t)}{\partial t} = k_{\alpha f} F_{\alpha}(t) + u[F_{\alpha-1}(t) + F_{\alpha+1}(t)] - (k_{\alpha f} + u)F_{\alpha}(t)
\]

The backward master equation that describes the temporal evolution of the protein in the bulk \( (n = 0) \) is given by

\[
\frac{\partial F_{\alpha}(t)}{\partial t} = k_{\alpha f} F_{\alpha}(t) + u[F_{\alpha-1}(t) + F_{\alpha+1}(t)] - (k_{\alpha f} + u + k_\alpha) F_{\alpha}(t)
\]

The first term describes the contribution to the free energy change due to the bending energy with the parameter \( k_\alpha \), while the opposite transition is taking place with a rate \( k_{\alpha f} \). Then the protein search can be effectively described by the search on the DNA configuration with a rate \( k_{\alpha f} \). The search dynamics for intermediate rates \( k_{\alpha f} \) and \( k_{\alpha} \) are related by the free energy difference and the residence time. The search dynamics for intermediate rates \( k_{\alpha f} \) and \( k_{\alpha} \) is expected to be more complex.

The transition rates \( k_{\alpha f} \) and \( k_{\alpha} \) are related by the free energy difference in the system due to looping/unlooping, and they specify the equilibrium constant \( K_{\text{conform}} \) for the conformational fluctuations,

\[
K_{\text{conform}} \equiv \frac{k_{\alpha f}}{k_{\alpha}} = \exp(-G(p))
\]

where \( G(p) \) is the free energy difference (in the unit of thermal energy, \( k_B T \)) between the looped conformation with the rescaled transfer rate, \( k_{\alpha f} \). Then the protein search will be the same as in the always looped conformation. The search dynamics for intermediate rates \( k_{\alpha f} \) and \( k_{\alpha} \) is expected to be more complex.

The transition rates \( k_{\alpha f} \) and \( k_{\alpha} \) are related by the free energy difference in the system due to looping/unlooping, and they specify the equilibrium constant \( K_{\text{conform}} \) for the conformational fluctuations,

\[
K_{\text{conform}} \equiv \frac{k_{\alpha f}}{k_{\alpha}} = \exp(-G(p))
\]

where \( G(p) = A + a \ln p + \epsilon \) is the bending persistence length of the DNA chain.

The first term describes the contribution to the free energy change due to the bending energy with the parameter \( A \) being proportional to the bending stiffness of DNA. It can be written as \( A = 2\pi^2 l_p \), where \( l_p \) is the bending persistence length of the DNA chain.
DNA. In our explicit calculations, we will utilize \( A = 300 \) in dimensionless units, which corresponds to \( l_p \approx 150 \) bp in real units.\(^{40}\) The second term corresponds to the entropic contribution to the free energy change. The parameter \( \alpha \) is the scaling factor, which for ideal polymer chains is equal to \( 3/2 \). The last term describes the enthalpy of the loop formation,\(^{40}\) and for explicit calculations, we take \( \varepsilon = -3k_BT \), unless otherwise specified. In addition, we assume that the sliding rate \( u \), association rate \( k_{\text{on}} \), and dissociation rate \( k_{\text{off}} \) are the same in both DNA conformations.

In the next section, we present specific calculations of protein target search dynamics on looped DNA and on the DNA with looped conformational transitions using realistic parameters of the system.

**RESULTS AND DISCUSSIONS**

Because our theoretical approach could provide analytical results for dynamic properties of the system, it allows us to obtain a comprehensive description of the protein target search processes. We first analyze the search on the always looped DNA molecule. Figure 2 presents the mean search times as a function of the scanning length \( \lambda \) for the always looped DNA molecules. Analytical results are obtained using eq 11, and they are shown by solid lines. Monte Carlo simulations data are shown by symbols for three different values of the association rate \( k_{\text{on}} \). The following parameters were used for calculations: \( L = 500 \), \( m = 250 \), \( l = 200 \), \( u = 6.0 \times 10^3 \) s\(^{-1} \), and \( k_t = 2.0 \times 10^3 \) s\(^{-1} \).

![Figure 2](https://dx.doi.org/10.1021/jacs.jpcb.0c09599)  
*Figure 2. Mean search times as a function of the sliding length \( \lambda \) for the always looped DNA molecules. Analytical results are obtained using eq 11, and they are shown by solid lines. Monte Carlo simulations data are shown by symbols for three different values of the association rate \( k_{\text{on}} \). The following parameters were used for calculations: \( L = 500 \), \( m = 250 \), \( l = 200 \), \( u = 6.0 \times 10^3 \) s\(^{-1} \), and \( k_t = 2.0 \times 10^3 \) s\(^{-1} \).*

The dissociation is controlled by the protein–DNA interactions: for stronger (weaker) protein–DNA interactions, smaller (larger) \( k_{\text{off}} \) is expected. In our study, the dissociation rate \( k_{\text{off}} \) is varied leading to changing \( \lambda \). Note that the scanning length \( \lambda \) is closely related to the parameter \( S(0) \) that describes the average number of visited sites per cycle, but it is not the same. It can be shown that \( 1 \leq S(0) \leq L \), while the scanning length can be smaller than one or larger than the length of the DNA segment \( L \). It is convenient to utilize the scanning length \( \lambda \) because it correlates better with the strength of nonspecific interactions between the protein and DNA molecules. The stronger the interaction, the larger the scanning length.

![Figure 3](https://dx.doi.org/10.1021/jacs.jpcb.0c09599)  
*Figure 3. Mean search times as a function of the scanning length \( \lambda \) for different transfer rates, \( k_t \). Analytical results are obtained from eq 11, and they are shown by solid lines. Monte Carlo simulations data are shown by symbols. The following parameters were used for calculations: \( u = 6.0 \times 10^3 \) s\(^{-1} \), \( k_{\text{on}} = 10 \) s\(^{-1} \), \( L = 500 \), \( l = 100 \), and \( m = 450 \).*

Several values of \( k_t \). It is found that even small values of the transfer rate \( (k_t \ll u) \) can significantly accelerate the search dynamics in the regimes with intermediate and strong nonspecific interactions. An order of magnitude acceleration in the search might be observed (Figure 3). The reason is that the looped conformation provides an additional channel to

\[ \text{specific interactions. When the scanning length } \lambda \text{ is larger than the total DNA length } L \text{ (strong interactions), the protein is mostly involved in 1D search. This is because after binding to DNA from the solution, the protein will only diffuse along the DNA chain until the target is found. This search regime is slow since the protein molecule redundantly visits the same sites multiple times. In this case, the overall search times are determined by the residence times on DNA, and they are essentially independent of the association rates } k_{\text{on}}. \text{ More efficient search dynamics is observed for very small scanning lengths } \lambda < 1 \text{ (weak interactions). In this regime, the protein is able occasionally to bind to DNA but it cannot slide and after some time it dissociates back into the solution. The search times depend strongly on the association rates because } 1/k_{\text{on}} \text{ gives the average time for the protein to be diffusing in the bulk solution for each search cycle. The faster the protein diffuses in 3D, the faster it will explore the whole DNA via binding/unbinding events and locate the target. For the intermediate scanning lengths, } 1 < \lambda < L, \text{ which also correspond to the intermediate strength of interactions, the protein is performing a combined 1D+3D search, and for some range of parameters, this is the most optimal search regime. In this case, the protein is able to scan some length on DNA every search cycle without being trapped on DNA for long periods of time, and the fast diffusion in the bulk allows it to return quickly to DNA to continue the scanning. This explains the nonmonotonic behavior of the mean search times as a function of the scanning length } \lambda \text{ for some range of parameters.}

We can focus now on the effect of the looped conformational state on the search time by varying the transfer rate \( k_t \) that allows the protein to jump reversibly between the sites \( l \) and \( L \) (see Figure 1a). Figure 3 shows the search time for
transfer between the \( l \) and \( L \) sites near the loop intersection (see Figure 1a), allowing the protein to scan the DNA molecule more efficiently and to move faster between distant parts of the DNA chain. Thus, in these regimes, the search dynamics is accelerated for any value of the transfer rate \( k_u \) and the search times are effectively independent of \( k_u \), as long as \( k_u \neq 0 \).

We note that in ref 25 it has been found that the association rate of protein to a target site on the circular DNA is lower, albeit modestly, than that for the linear counterpart. In that case, the authors consider only a special case where the target is located at the center of the DNA chain. The difference in the association rates originates from the excluded interaction between the search protein and the ends of the linear DNA, which will be negligible for a long DNA. For small values of the protein–DNA interactions, the variation in the transfer rate does not affect the search dynamics as much. The reason is that in these situations the protein molecule moves mostly in the bulk and the additional paths to explore DNA are less helpful to the search since the protein is not sliding along the DNA chain. However, we remark here that there could be another mechanism of enhancing the search efficiency, namely, “hopping” between spatially close segments of the DNA even without direct physical connectivity.\(^{14,24,41,44}\) Therefore, the results of our model would describe the lower bound of the search dynamics depends on the relative positions of the target \( m \) and the loop intersection with DNA \( l \), from where the transitions to the site \( L \) might happen (see Figure 1a). The corresponding results are presented in Figures 4 and 5. From these calculations are shown in Figure 5a. One can see again that the changes in the search times are observed for strong nonspecific interactions (large scanning lengths), and the fastest dynamics is found when the loop intersection is close to the target (\( m \approx l \)). Similar results are found if one fixes the size of the loop and varies the position of the target \( m \) as shown in Figure 5b. The search is the most efficient when the target is close to the loop intersection.

The effect of the distance between the target and the loop intersection on the search dynamics is also illustrated in a slightly different way in Figure 5. For a fixed location of the target, one can vary the position of the loop intersection \( l \), which also corresponds to changing the size of the DNA loop. These calculations are shown in Figure 5a. One can see again that the changes in the search times are observed for strong nonspecific interactions (large scanning lengths), and the fastest dynamics is found when the loop intersection is close to the target (\( m \approx l \)). Similar results are found if one fixes the size of the loop and varies the position of the target \( m \) as shown in Figure 5b. The search is the most efficient when the target is close to the loop intersection.

Our calculations show that, when the target is not far away from the loop intersection, it might lead to a significant acceleration in the search dynamics, modifying the search times by up to 1 order of magnitude. As shown in the Figure 6 for the heat map for the acceleration, the fastest search is clearly observed when \( m \approx l \) and \( m \approx L \). We can conclude then that the DNA looped configuration is generally better for the search dynamics, and the most optimal conditions are found when the loop intersection is close to the target sequence. This effect is observed for intermediate and strong nonspecific protein–DNA interactions.

Figure 4. Mean search times as a function of the scanning length \( L \) for different locations of (a) the DNA loop intersection and (b) the target. Analytical results are shown in solid lines, and Monte Carlo simulations data are shown in symbols. The following parameters were used for calculations: (a) \( u = 6 \times 10^3 \text{ s}^{-1}, k_i = 2 \times 10^3 \text{ s}^{-1}, k_{in} = 1 \text{ s}^{-1}, L = 500, l = 250, \) and \( \lambda = 10^4 \).
parameters were used for calculations: \( u = 6 \times 10^3 \text{ s}^{-1} \), \( k_{\text{on}} = 10 \text{ s}^{-1} \), \( L = 500 \), \( \lambda = 10^3 \), and \( k_i = u \).

A more realistic description of the protein search on DNA should include the conformational transitions between the DNA looped and free states, as schematically presented in Figure 1b. It is clear that search dynamics will be affected by looping/unlooping transitions for intermediate and strong protein–DNA interactions when the protein molecule can slide long distances on DNA. The results of computer simulations for this situation are presented in Figure 7. For

![Figure 6](image)

**Figure 6.** Acceleration of the mean search times on the DNA chain with the loop as compared to the free DNA, \( T_{\text{free}}/T_{\text{looped}} \). The following parameters were used for calculations: \( u = 6 \times 10^3 \text{ s}^{-1} \), \( k_{\text{on}} = 10 \text{ s}^{-1} \), \( L = 500 \), \( \lambda = 10^3 \), and \( k_i = u \).

The presented theoretical approach allows us to explicitly evaluate the role of DNA geometry and conformational dynamics on the protein search for specific sequences on DNA. Our results suggest that tuning the DNA geometry and the degree of conformational fluctuations might be an additional efficient method of how the nature might regulate genetic processes. One could imagine that changing the size of DNA loops and the location of the loop intersections might activate or repress the specific genes. Modifying the conformational dynamics might also lead to nonmonotonic changes in the search times, allowing for fine-tuning the gene activation processes. It will be interesting to explore these ideas in more direct experimental studies.

- **SUMMARY AND CONCLUSIONS**

In this work, we developed a theoretical approach to evaluate the role of DNA geometry and conformational dynamics on protein target search phenomena. By constructing a discrete-state stochastic model for the underlying processes, the dynamics of the system is explicitly computed using a method involving the unlooping rate (and proportionally the looping rate via eq 12) has a more complex effect on the search times. For very high unlooping/looping rates, the conformational equilibrium is quickly reached before the target is located, and the system behaves like a single looped state with a new effective transfer rate, \( k_{\text{eff}} \). In this situation, the search dynamics is also fast and it is close to the case of the protein search on the always looped DNA as discussed above. This is because, as we already argued, the magnitude of the transfer rate is essentially not affecting the search, but the topological change of having the loop versus not having the loop is important.

The most interesting behavior is observed for intermediate \( k_{\text{unloop}} \), as shown in Figure 7. When the free conformation state is energetically more preferred (\( \epsilon = -5k_B T, K_{\text{conform}} \approx 0.01 \)) the protein can transfer from the looped into the free state and it would stay there for significant periods of time. The fraction of time when the DNA molecule is in the free (unlooped) state, \( f_{\text{free}} \), can be estimated from the equilibrium constant, \( f_{\text{free}} = 1/(K_{\text{conform}} + 1) \). In this case, the system is found in the free state 99% of the time, and for this reason, the search times are start to increase with \( k_{\text{unloop}} \) because \( k_{\text{unloop}} \gg k_{\text{loop}} \) and the probability to shift the system into the free state is very large.

Then the search times are approaching closer to \( T_{\text{free}} \). The target is typically found before the conformational transition can happen. In other words, for intermediate \( k_{\text{unloop}} \) the search times are faster than the times to establish the conformational equilibrium. However, when the free state becomes energetically less favorable (\( \epsilon = -9k_B T, K_{\text{conform}} \approx 0.58 \)), the fraction of time the DNA chain is free decreases to 63%. In this case, the search times do not increase with \( k_{\text{unloop}} \) because \( k_{\text{unloop}} \approx k_{\text{loop}} \), and now they are getting closer to \( T_{\text{looped}} \). In all situations, however, for very fast looping/unlooping transitions, the equilibrium is reached faster than the typical search times, and the system can be viewed as the single looped conformation.

As the result, a nonmonotonic dependence of the search times as a function of the unlooping transition rates is observed for the cases when the free conformational state is energetically more stable, but not for the cases when the looped conformations starting to become more relevant. These observations show that the DNA conformational dynamics results in a very nontrivial effect on the protein target search. But in all cases, importantly, it is always faster to search on the conformationally fluctuating DNA than on the free DNA.

![Figure 7](image)

**Figure 7.** Mean search times as a function of the unlooping rates for the system when DNA fluctuates between the free and the loop conformations. Initially, the DNA molecule starts with equal probability in any of these conformations. The dashed lines correspond to analytic limiting results for the free and for the looped conformations, respectively, while the symbols correspond to computer simulations for the systems with different enthalpic contributions due to creating the DNA loop \( \epsilon \). The following parameters were used for calculations: \( u = 6 \times 10^3 \text{ s}^{-1} \), \( k_{\text{on}} = 10 \text{ s}^{-1} \), \( L = 500 \), \( m = 450 \), \( l = 200 \), and \( \lambda = 10^3 \).

In all situations, the search times are bounded between the results for the free and those for the looped DNA conformations. This means that the protein search is always faster on the conformationally fluctuating DNA in comparison with the search on the free DNA chain.

When the looping/unlooping transition rates are small, the search is mostly determined by the initial DNA conformation. In Figure 7, the system started with equal probability in one of two conformations, and the search times are essentially given by \( T \approx (T_{\text{unlooped}} + T_{\text{looped}})/2 \). In this case, the target is found before the conformational equilibrium is established. Increasing the unlooping rate (and proportionally the looping rate via eq 12) has a more complex effect on the search times. For very high unlooping/looping rates, the conformational equilibrium is quickly reached before the target is located, and the system behaves like a single looped state with a new effective transfer rate, \( k_{\text{eff}} \). In this situation, the search dynamics is also fast and it is close to the case of the protein search on the always looped DNA as discussed above. This is because, as we already argued, the magnitude of the transfer rate is essentially not affecting the search, but the topological change of having the loop versus not having the loop is important.

The most interesting behavior is observed for intermediate \( k_{\text{unloop}} \), as shown in Figure 7. When the free conformation state is energetically more preferred (\( \epsilon = -5k_B T, K_{\text{conform}} \approx 0.01 \)) the protein can transfer from the looped into the free state and it would stay there for significant periods of time. The fraction of time when the DNA molecule is in the free (unlooped) state, \( f_{\text{free}} \), can be estimated from the equilibrium constant, \( f_{\text{free}} = 1/(K_{\text{conform}} + 1) \). In this case, the system is found in the free state 99% of the time, and for this reason, the search times are start to increase with \( k_{\text{unloop}} \) because \( k_{\text{unloop}} \gg k_{\text{loop}} \) and the probability to shift the system into the free state is very large.

Then the search times are approaching closer to \( T_{\text{free}} \). The target is typically found before the conformational transition can happen. In other words, for intermediate \( k_{\text{unloop}} \) the search times are faster than the times to establish the conformational equilibrium. However, when the free state becomes energetically less favorable (\( \epsilon = -9k_B T, K_{\text{conform}} \approx 0.58 \)), the fraction of time the DNA chain is free decreases to 63%. In this case, the search times do not increase with \( k_{\text{unloop}} \) because \( k_{\text{unloop}} \approx k_{\text{loop}} \), and now they are getting closer to \( T_{\text{looped}} \). In all situations, however, for very fast looping/unlooping transitions, the equilibrium is reached faster than the typical search times, and the system can be viewed as the single looped conformation. As the result, a nonmonotonic dependence of the search times as a function of the unlooping transition rates is observed for the cases when the free conformational state is energetically more stable, but not for the cases when the looped conformations starting to become more relevant. These observations show that the DNA conformational dynamics results in a very nontrivial effect on the protein target search. But in all cases, importantly, it is always faster to search on the conformationally fluctuating DNA than on the free DNA.

The presented theoretical approach allows us to explicitly evaluate the role of DNA geometry and conformational dynamics on the protein search for specific sequences on DNA. Our results suggest that tuning the DNA geometry and the degree of conformational fluctuations might be an additional efficient method of how the nature might regulate genetic processes. One could imagine that changing the size of DNA loops and the location of the loop intersections might activate or repress the specific genes. Modifying the conformational dynamics might also lead to nonmonotonic changes in the search times, allowing for fine-tuning the gene activation processes. It will be interesting to explore these ideas in more direct experimental studies.
of first-passage probabilities. Our theoretical calculations are also supported by Monte Carlo computer simulations. It is found that the formation of the DNA looped configuration might significantly affect the search dynamics, and this depends on the strength of nonspecific protein–DNA interactions. For weak interactions, the DNA geometry does not much influence the protein search, but for stronger interactions it becomes much more important because in this case the protein molecule spends most of the time on DNA. In these dynamic regimes, the effect is stronger depending on how close the target is to the loop intersection. Taking into account the effect of conformational transitions between the free and the DNA looped conformations is also investigated. Here we found that the conformational fluctuations also modify the search, and the effect is more complex. When the conformational transitions are fast, the system behaves like the always looped DNA state and the search is fast. For intermediate values of the looping/unlooping rates, the search dynamics to a large degree depends on what conformational state is energetically more favorable and how fast the conformational equilibrium is reached. In addition, it is argued that varying the geometry and the degree of conformational fluctuations of the DNA molecules might serve as a very efficient and sensitive tool in genetic regulations and manipulations. Our theoretical calculations strongly suggest that topological features of DNA molecules and dynamic transformations between different conformational states are critically important for understanding the molecular details of protein–DNA interactions. Our main idea is that the protein target search is always faster on the looped DNA molecule or on the conformationally fluctuating DNA because of the appearance of additional channels for finding the target due to topological changes in the system.

While the presented theoretical approach provides a reasonable minimal description of DNA geometry and conformational dynamics, it is important to emphasize its limitations. Several important features of DNA molecules in chromatin structures such as DNA supercoiling,43 sequence heterogeneity,43 and the possibility of multiple DNA loops are molecule or on the conformationally protein target search is always faster on the looped DNA chromatin structures such as DNA supercoiling,45 sequence limitations. Several important features of DNA molecules in reasonable minimal description of DNA geometry and due to topological changes in the system. of the appearance of additional channels for dynamic transformations between different conformational states is more complex. When the conformational transitions are fast, the system behaves like the always looped DNA state and the search is fast. For intermediate values of the looping/unlooping rates, the search dynamics to a large degree depends on what conformational state is energetically more favorable and how fast the conformational equilibrium is reached. In addition, it is argued that varying the geometry and the degree of conformational fluctuations of the DNA molecules might serve as a very efficient and sensitive tool in genetic regulations and manipulations. Our theoretical calculations strongly suggest that topological features of DNA molecules and dynamic transformations between different conformational states are critically important for understanding the molecular details of protein–DNA interactions. Our main idea is that the protein target search is always faster on the looped DNA molecule or on the conformationally fluctuating DNA because of the appearance of additional channels for finding the target due to topological changes in the system.

While the presented theoretical approach provides a reasonable minimal description of DNA geometry and conformational dynamics, it is important to emphasize its limitations. Several important features of DNA molecules in chromatin structures such as DNA supercoiling,43 sequence heterogeneity,43 and the possibility of multiple DNA loops are not considered at all. In addition, the description of the free energy associated with the formation of DNA loops is rather oversimplified. Furthermore, the protein diffusion rates on the free and on the looped DNA chains are probably not the same as was assumed in this work. Spatial correlation effects and crowding are also not taken into account in our approach. However, despite these shortcomings our theoretical model provides some clear physical insights on the complex mechanisms of protein target search in live cells. Most importantly, it gives quantitative predictions that can be probed in experiments. It will be important to extend our theoretical approach by taking more realistic features of protein–DNA dynamics in living cells and to compare with experimental observations.

ASSOCIATED CONTENT

Supporting Information
The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jpcb.0c09599.

Details of analytical calculations for mean search times (PDF)

AUTHOR INFORMATION

Corresponding Author
Anatoly B. Kolomeisky — Department of Physics and Astronomy, Rice University, Houston, Texas 77005, United States; Center for Theoretical Biological Physics, Department of Chemistry, and Department of Chemical and Biomolecular Engineering, Rice University, Houston, Texas 77005, United States; orcid.org/0000-0001-5677-6690; Email: tolya@rice.edu

Authors
Cayke Felipe — Department of Physics and Astronomy, Rice University, Houston, Texas 77005, United States; Center for Theoretical Biological Physics, Rice University, Houston, Texas 77005, United States
Jaeoh Shin — Center for Theoretical Biological Physics and Department of Chemistry, Rice University, Houston, Texas 77005, United States

Complete contact information is available at: https://pubs.acs.org/doi/10.1021/acs.jpcb.0c09599

Notes
The authors declare no competing financial interest.

ACKNOWLEDGMENTS

The work was supported by the Welch Foundation (C-1559), by the NSF (CHE-1953453 and MCB-1941106), and by the Center for Theoretical Biological Physics sponsored by the NSF (PHY-2019745).

REFERENCES

(2) Sträringlfoer, A.; Ekwall, K. Reviews in Cell Biology and Molecular Medicine; American Cancer Society, 2011.


(23) Barbi, M.; Place, C.; Popkov, V.; Salerno, M. Base-sequence-dependent sliding of proteins on DNA. *Phys. Rev. E* 2004, 70, 041901.


