Facilitated search of proteins on DNA: correlations are important

Rahul K. Das and Anatoly B. Kolomeisky*

Received 13th October 2009, Accepted 11th January 2010
First published as an Advance Article on the web 5th February 2010
DOI: 10.1039/b921303a

A starting point of many biological processes is protein binding to specific regions on DNA. Although typical concentrations of DNA-binding proteins are low, and target sites are typically buried among huge number of non-specific sites, the search process is frequently achieved at a remarkably fast rate. For some proteins it has been confirmed that association rates might be even larger than the maximal allowed three-dimensional diffusion rates. The current theoretical view of this phenomenon is based on the idea of lowering dimensionality, *i.e.*, the overall search process is viewed as a combination of uncorrelated three-dimensional excursions in the solution and one-dimensional hoppings on DNA. However, some predictions of this theoretical picture contradict recent single-molecule measurements of protein diffusion processes. An alternative theoretical approach points out the importance of correlations during the search process that appear due to non-specific interactions between protein and DNA molecules. To test different theoretical ideas we performed extensive lattice Monte Carlo computer simulations of the facilitated diffusion. Our results revealed that correlations are important, and the acceleration in the search could only be achieved at some intermediate non-specific binding energies and protein concentrations. Physico-chemical aspects and the origins of these correlations are discussed.

I. Introduction

All essential biological functions of DNA largely rely on search and recognition of specific target sites on DNA by several classes of proteins such as repressor proteins, DNA helicases, DNA polymerases and endonucleases/restrictases.[1,2] Earlier experimental studies[3] suggested that some proteins (*e.g. lac* repressors) find their target sites much faster than allowed by ordinary three-dimensional diffusion limit as outlined in the Debye–Smoluchowski theory.[4–7] This phenomenon is known as a *facilitated diffusion*,[8] and it has been well explored by experimental and theoretical studies.[5,6,9–20]

The observation of extremely high rates of protein search for targets on DNA is surprising, and it has been proposed that it might be due to electrostatic interactions between oppositely charged molecules which significantly modify the original Smoluchowski diffusion limits.[18] However, the original association rate measurements[3] have been performed in a buffer containing Tris/HCl, KCl and magnesium acetate, all at 10 mM concentrations, and for this situation one can easily estimate the Debye length to be of the order $\lambda_D \simeq$ 1–2 nm. The Debye length defines the region beyond which electrostatic forces do not play a role, and for this system it is similar to a size of the target on DNA (several base pairs). At the same time, it is known that electrostatics is crucial for protein–DNA interactions.[5,6] It suggests that electrostatic interactions probably cannot accelerate the approach of the protein molecules to DNA from large distances, but when the protein is already near the DNA chain they become very important. Thus electrostatic forces do not increase the diffusion limit, as suggested in ref. 18 and for explaining the phenomenon of the facilitated diffusion requires a different mechanism.

The most popular and widely used theoretical picture of facilitated diffusion has been proposed by Berg, Winter and von Hippel.[5–7] According to this approach, the overall search process consists of three-dimensional motions in the solution and one-dimensional scanning of DNA. It was argued that the main reason for the search acceleration is "lowering of dimensionality", because the protein molecule on DNA has a higher probability to diffuse in the direction of the target. This theoretical view has been strongly supported by experimental studies that directly show the importance of contributions from one-dimensional sliding in the protein searching for specific DNA target sites.[14]

However, despite the ability to describe and explain some features of the facilitated diffusion,[9,10,13] there are several problems with this theoretical picture. First, in most cases it is assumed that three-dimensional and one-dimensional diffusion constants (D_3 and D_1, respectively) are of the same order of magnitude. However, recent single-molecule experiments[21,22] provided better measurements of one-dimensional motions of proteins on DNA, and it was found that D_1 is much smaller than D_3. It should be noted also that under these conditions ($D_1 \ll D_3$) these theoretical models do not predict any acceleration at all.[13] Second, theoretical calculations using this approach predict that the most optimal search is achieved when times that the protein spends in the solution and on DNA are similar. This is again in contradiction with single-molecule experimental results[21,22] which indicate that the protein molecule typically stays mostly on DNA, and not in the solution. In addition, current theoretical models give unphysical predictions in some limiting cases. For example, they predict that *increasing* target sites concentrations or

Department of Chemistry, Rice University, TX 77005, Houston, USA. E-mail: tolya@rice.edu

concentrations of proteins *lowers* the association rate, in clear contradiction of expected chemical kinetic behavior (see Fig. 6 of ref. 13). These models can be viewed as "uncorrelated" in the sense that after dissociating from the non-specific segment of DNA the protein molecule has an equal probability of binding anywhere on the DNA, *i.e.*, one-dimensional and three-dimensional dynamical modes are fully decoupled.

An alternative theoretical picture has been proposed recently.[15,16] It is based on idea of including the effects of non-specific interactions between protein and DNA molecules that lead to increased local concentrations of proteins and correlations in the search process. This approach can be considered as "correlated" since it is argued that the protein molecule that just dissociated from DNA has a higher probability to come back to the DNA segment closer to the unbinding position due to non-specific interactions. Theoretical predictions from these models agree well with all available experimental observations, and this approach eliminates the unphysical behavior observed in the uncorrelated models. However, theoretical predictions have been obtained in the approximate model[15] that assumes very strong unphysical correlations by mapping three-dimensional motions into effective one-dimensional steps. Thus it is still unclear what role correlations play in the protein search for targets on DNA.

In order to understand mechanisms of protein search for targets on DNA and to test different theoretical ideas we performed extensive Monte Carlo computer simulations. Previous computational works[12] described a system consisting of one protein and one DNA molecule, which is not a realistic description of facilitated diffusion phenomena. Our aim here is not to provide a comprehensive computational description of these complex phenomena but rather to analyze important dynamic properties of the system where one can clearly distinguish between different theoretical pictures.

II. Computer simulations

We consider a lattice model of the protein search for the target on DNA as shown in Fig. 1. Our simulations have been carried out in a $100 \times 100 \times 100$ cubic lattice with periodic boundary conditions in all directions, and with multiple protein molecules around a single DNA molecule. The DNA molecule has been modeled as a worm-like chain, and it has been generated as a continuous persistent loop. At every step the random walk direction is governed by the product of a parameter σ and a Gaussian random number. Variation of this parameter controls the persistence length of the DNA. The parameter σ is the standard deviation for the bending between two successive steps. If σ is zero then DNA will be completely straight and rigid. The simulations have been performed for different values of σ. The persistent walk is continued by growing it on one end and shrinking it on the other end, until the orientation becomes almost parallel. Then the periodic box is chosen in such a way that the ends fall on top of each other. The latter involves finding the rotation of the periodic axes (x, y, z) by a rotation matrix, such that the ends or at least their periodic images are nearby. We have assumed that the thickness of the DNA is negligible compared

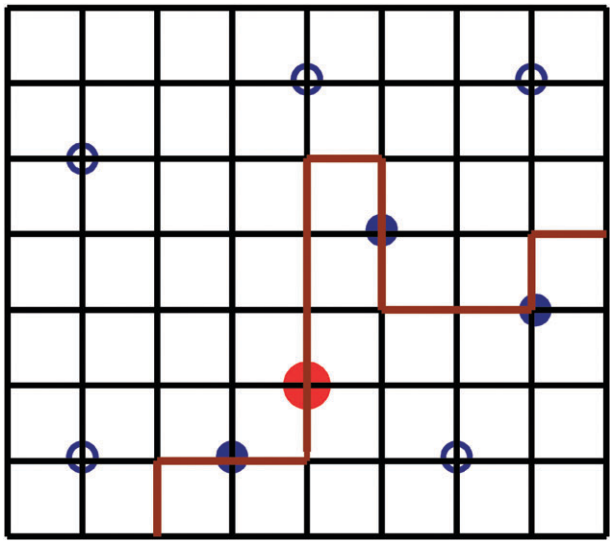

Fig. 1 A general schematic picture of the lattice model of the protein search for the target site on DNA. Brown segments describe the DNA chain and the red sphere locates the position of the target on DNA. Protein molecules are represented by blue spheres where the filled spheres are for proteins bound to DNA, while the unfilled spheres correspond to free proteins in the solution.

to the lateral size (squared) of the volume belonging to a site so that there are no excluded-volume interactions. One of the sites of DNA is randomly chosen to be a target site. The distance between any two lattice sites is the unit length of the problem, and is taken to be equal to one. The results of our computations are averaged out over the position of the target site.

We start our simulation by putting all protein molecules randomly in the system. Every lattice site on the DNA chain is treated as a potential binding site. Any protein molecule has four possibilities: (1) it can diffuse in the bulk of the solution away from the DNA molecule; (2) the protein can come close to the DNA molecule (neighboring sites) and then it might bind to non-specific sites; (3) the protein molecule already non-specifically bound to DNA can slide along the DNA chain by one-dimensional diffusion; (4) the bound protein molecule can also dissociate from DNA back to the solution. There are three types of protein molecules in the system. One group of protein molecules is free in the solution and they perform three-dimensional excursions; another group of the protein molecules is siting at sites neighboring to DNA but not yet bound; and the last group of proteins is non-specifically bound to DNA.

The binding rate constant of the protein molecule to the non-specific DNA site is given by

$$k_{\mathrm{on}} = A \exp[-E_a/k_{\mathrm{B}}T], \qquad (1)$$

and the unbinding rate constant from the DNA to the solution is

$$k_{\mathrm{off}} = A \exp[-(E_a + E_{\mathrm{ads}})/k_{\mathrm{B}}T] \qquad (2)$$

where A is the pre-exponential factor (assumed to be the same for both processes), E_a is the activation energy barrier for

binding to DNA and E_{ads} is the non-specific binding energy that corresponds to a free-energy difference for the protein molecule to be non-specifically bound to DNA in comparison with being free in the solution. The three-dimensional diffusion rate is equal to $r_1 = 6n_1D_3$, where n_1 is the number of free protein molecules in the solution and D_3 is the corresponding three-dimensional diffusion constant. In this system D_3 is equal to the 3D diffusional rate of the individual protein because the length of the lattice unit is one. The rate of non-specific binding to DNA is given by $r_2 = n_2k_{on}$, where n_2 is the number of protein molecules that are one site away from the DNA molecule but not yet bound to it. Note that n_2 is also included in the total number of free proteins in the solution n_1. Similarly, the rate of unbinding from the non-specific sites of DNA is equal to $r_3 = n_3k_{off}$ with n_3 being the number of protein molecules that are currently bound to DNA. Finally, the rate of protein hopping along the DNA chain can be written as $r_4 = 2n_3D_1$ where D_1 is the corresponding one-dimensional diffusion constant. The mechanism of one-dimensional sliding along the DNA molecule is not known, but one expects that increasing non-specific interactions would probably lower hopping rates because each step includes breaking the non-specific bond with the DNA site. To account for this effect we modify hopping rates as $r_4 = 2n_3D_1 \exp[-(\alpha E_{ads})/k_BT]$, where the parameter α can be varied.

The dynamic evolution of our system follows the following rules. At each time step we multiply the sum of all rates $S = r_1 + r_2 + r_3 + r_4$ by a random number r, $0 \leq r \leq 1$. If $rS < r_1$ then one of free proteins in the solution is chosen to be moved to the neighboring site *via* three-dimensional diffusion. In the case of $r_1 < rS < r_1 + r_2$ one of the protein molecules at sites neighboring to DNA is allowed to bind to the closest DNA site. For $r_1 + r_2 < rS < r_1 + r_2 + r_3$ one of the bound protein molecules dissociates back into the solution, and for $r_1 + r_2 + r_3 < rS < r_1 + r_2 + r_3 + r_4$ one of the bound protein molecules jumps to the neighboring DNA sites (with equal probability). After this the time is advanced by $1/S$. It should be noted that some of the events might not take place due to exclusion effects: available sites might be already occupied. In each simulation the target is put randomly at any site on DNA and whenever any of the protein molecules finds it, either through three-dimensional diffusion or through one-dimensional scanning along the DNA chain, the simulation is stopped and the total time is recorded. To quantify the facilitated diffusion and to compare it with purely three-dimensional diffusion search, in parallel we also perform Monte Carlo simulations on an identical system with the target site but *without* the DNA molecule (so only three-dimensional motions are possible). The total number of proteins in the system or the overall concentrations of the proteins is kept same in both the cases. In our calculations we analyze a dimensionless ratio of search times τ/τ_S, where τ is the mean time to reach the target site on DNA in the facilitated diffusion, while τ_S is the mean time to reach the target in the Smoluchowski regime (without DNA). We carry out 1000 simulation cycles for each parameter set both in the presence of DNA and in the absence of DNA. This reduces the error to $\simeq$ 3% for every simulation.

III. Results and discussions

Before analyzing the results it is important to understand what is the search time that is measured in our computer simulations. Starting from the random configuration at $t = 0$ the clock is stopped when *any* of the proteins in the system reach the target site for the first time, and this search time is compared with that of an identical system without DNA. One can argue that the facilitation in the search is achieved when the ratio of the two search times (with and without DNA) is less than one.[15] This exactly corresponds to quantities measured in experiments, but it is very different from the time for a *specific* protein to reach the target for the first time, as frequently analyzed in many theoretical studies.[12]

First, the effect of non-specific binding energies on the search process has been investigated. In typical simulations we started with $N = 100$ protein molecules and with a single DNA molecule with 7500 bases. The parameter α (that couples the non-specific interactions with the protein speed on DNA) was taken to be 0.02, while the ratio of diffusion constants D_1/D_3 was set to be 0.1 to better reflect realistic conditions. These parameters have been varied in our simulations, but the qualitative picture remained the same. The results are displayed in Fig. 2. It shows that the ratio of search times is a non-monotonous function of the binding energy, and there are parameters at which acceleration in the facilitated diffusion might be observed ($\tau/\tau_S < 1$).

This behavior can be understood using the following arguments. At low binding energies bound protein molecules move relatively fast on DNA (but still slower than in the solution), but the concentration of proteins on DNA is small, and the total length of scanned DNA segment is not large. It seems that the ordinary 3D diffusional search is more efficient

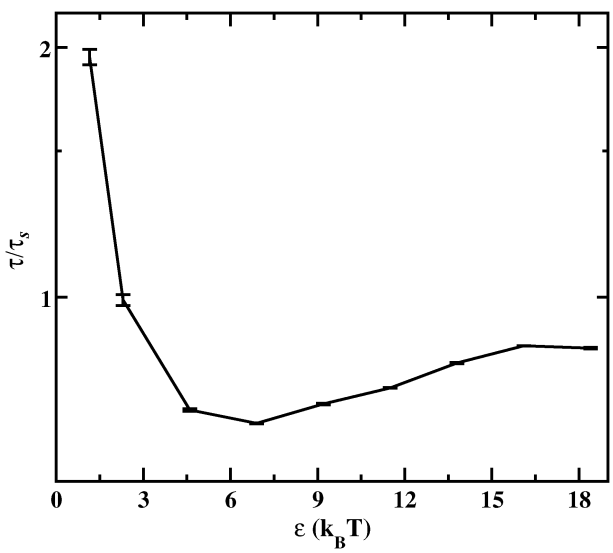

Fig. 2 Effect of non-specific interaction energies on the protein search process. The ratio of search times as a function of binding interactions is calculated from computer simulations. The parameters used in simulations are: $100 \times 100 \times 100$ cubic lattice, $\alpha = 0.02$, $D_1/D_3 = 0.1$, the length of DNA is 7500 sites, the number of proteins in the system is $N = 100$. The details of simulation procedures are described in the text.

in this case. Increasing the strength of non-specific interactions drive more protein molecules to DNA, allowing them to scan larger portions of the DNA chain, lowering the one-dimensional component of the total search time. It leads the facilitated diffusion to become faster than the Smoluchowski diffusion. However, larger interactions also lower the speed of individual proteins on DNA, and it eventually starts to increase the total search time. Thus our simulations predict that the facilitated diffusion might be efficient only for range of intermediate interactions when the speed and the number of bound proteins are such that the fastest search times could be achieved.

The importance of the parameter α in our simulations should be noted. It relates the strength of the non-specific binding to the speed of protein moving on the DNA chain. Although the non-monotonous behavior of the ratio of search times is observed for almost all values of α, the facilitated diffusion is more efficient ($\tau/\tau_S < 1$) only for the limited range of values of α: see Fig. 3. For very low values of α increasing the interaction energy makes the facilitated diffusion more efficient because the concentration of bound proteins also increases but without lowering the diffusional rate along DNA. For relatively large values of α the decrease in the DNA scanning speed is significant at all interactions, and three-dimensional search is always faster at these conditions. We chose $\alpha = 0.02$ to describe the protein search phenomena, but it remains to be determined more precisely from experimental measurements.

We also investigated the effect of the free protein concentration on the search for the target site on the DNA molecule. The results from simulations are presented in Fig. 4. Again, the ratio of search times shows a non-monotonous dependence as a function of the protein concentration. There is also a set of

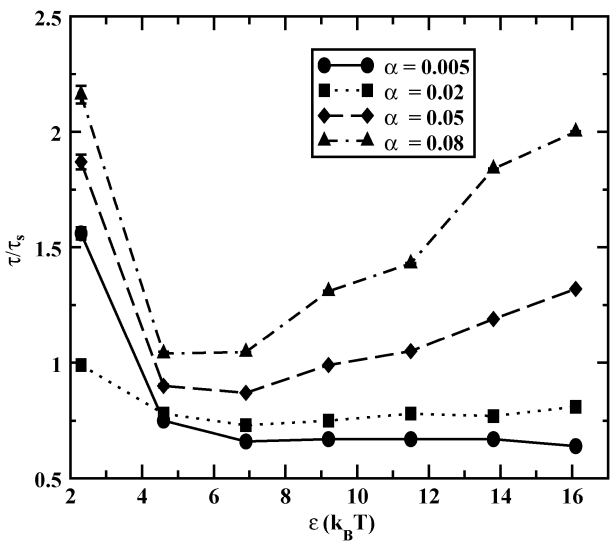

Fig. 3 Effect of parameter α on the protein search process. The ratio of search times as a function of binding interactions is calculated from computer simulations for different values of α. The parameters used in simulations are: $100 \times 100 \times 100$ cubic lattice, $D_1/D_3 = 0.1$, the length of DNA is 7500 sites, and the number of proteins in the system is $N = 100$. The details of simulation procedures are described in the text.

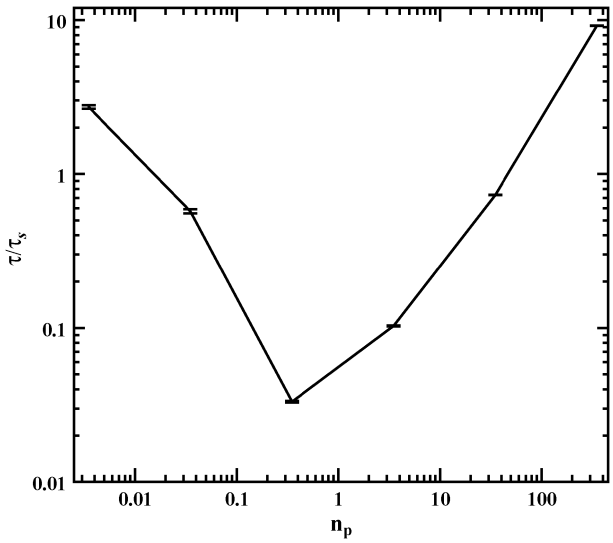

Fig. 4 Effect of free protein concentration on the protein search process. The ratio of search times as the function of number of free proteins around a single DNA chain is calculated from computer simulations. The parameters used in simulations are: $100 \times 100 \times 100$ cubic lattice, $\alpha = 0.02$, $D_1/D_3 = 0.1$, the length of DNA is 7500 sites, and the effective adsorption energy is $E_{ads} = 4.6\ k_BT$. The details of simulation procedures are described in the text.

parameters at which the facilitated diffusion is faster than the normal three-dimensional diffusion. This can be explained by analyzing the details of association and dissociation processes. The free energies of proteins in the solution and proteins bound to DNA depend on the concentrations of those proteins. At very low concentrations any binding to the DNA chain leads to significant lowering of the free energy of proteins in the solution—one can easily see this in the limiting case of only one protein in the system. Then the bound protein will have a higher probability to dissociate back into the solution instead of starting to jump along the DNA. However, without scanning the facilitated mechanism is not working efficiently. Increasing the total concentration of proteins washes out this effect, and the bound proteins can make large excursions along DNA looking for the target. For very high concentrations protein molecules in the solution can be found so close to the target that it can be reached directly through the solution without intermediate non-specific bindings to DNA. These arguments show that only for intermediate protein concentrations might the facilitated diffusion mechanism win over the three-dimensional Smoluchowski diffusion.

It can be argued that the relative magnitudes of the diffusion constants of proteins in the solution and on DNA are important factors in the search process. Although, many current theories assume that they are similar,[13] this assumption contradicts to single-molecule experimental measurements[21,22] which clearly show that $d = D_1/D_3$ is quite small. The importance of this parameter has been investigated in our computer simulations. As shown in Fig. 5, it has a dramatic effect. As expected, when both diffusion constants are comparable the facilitated diffusion that consists of combination of 1D and 3D steps is more efficient for large set of parameters. Lowering d makes the facilitated mechanism much less

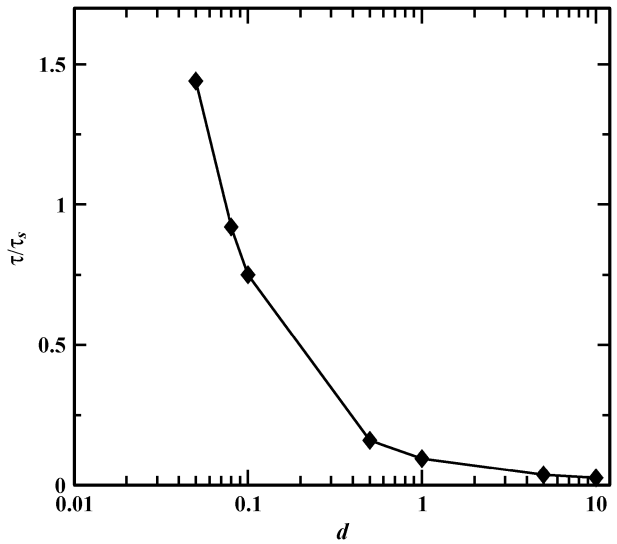

Fig. 5 Effect of diffusion constants on the protein search process. The ratio of search times as the function of the ratio of diffusion constants $d = D_1/D_3$ is calculated from computer simulations. The parameters used in simulations are: $100 \times 100 \times 100$ cubic lattice, $\alpha = 0.02$, the number of proteins in the system is $N = 100$, the length of DNA is 7500 sites, and the effective adsorption energy is $E_{ads} = 6.9\ k_BT$. The details of simulation procedures are described in the text.

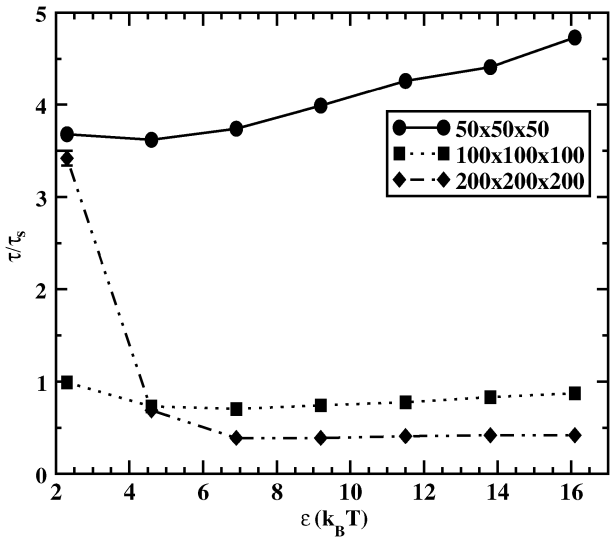

Fig. 6 Effect of lattice size on the protein search process. The ratio of search times as a function of binding interactions is calculated from computer simulations for different lattice sizes. The parameters used in simulations are: $D_1/D_3 = 0.1$, the length of DNA is 7500 sites, the number of proteins in the system is $N = 100$. The details of simulation procedures are described in the text.

efficient. In our simulations we chose $d = 0.1$ as more realistic value which shows that even for small ratios it is still possible to have the acceleration in the facilitated mechanism.

In addition, simulations have been performed for larger lattices, as shown in Fig. 6, and the same general trends have been observed but with larger parameter space where the facilitated diffusion seems to be more effective. Also, for larger lattice sizes the facilitated diffusion search was faster even at much lower values of D_1/D_3 ratio than 0.1. Simulations in the presence of DNA have been carried out with different persistence lengths of DNA as well as considering DNA as a straight chain, but in all cases the trends of the ratio of search times as a function of the binding energies or protein concentrations remained the same. It could be that this observation is the result of periodic boundary conditions that effectively fix DNA density in our simulations. It will be interesting to investigate this issue more carefully in the future.

It is important to compare the results of our computer simulations with predictions from different theoretical models. Let us recall that the uncorrelated models decouple three-dimensional and one-dimensional modes of the motion,[5,7,9,10,13] while the correlated approach argues that non-specific interactions effectively couple 3D and 1D searching segments. Then this fundamental difference should be observed in the behavior of some dynamic properties. Both uncorrelated[5,7,9,10,13] and correlated[15] models suggest that the ratio of search times shows a non-monotonous behavior as a function of non-specific interactions, as found in our simulations. This also agrees with available experimental bulk chemical kinetic measurements[6] where binding energies have been modified by changing the ionic strength of the solution. However, two theoretical approaches differ significantly in the description of search process as a function of protein concentration. The uncorrelated models[9,13] predict that the ratio $\tau/\tau_S < 1$ is *always* an increasing function of free protein concentrations and the most optimal search conditions are achieved in the limit of disappearing proteins in the solution. At the same time, the correlated models[15] suggest that the ratio of search times is a non-monotonous function of protein concentration. It seems that our Monte Carlo computer simulations support more the correlated picture of the facilitated diffusion.

Our computer simulations results imply that the critical role in the fast protein search for the target on DNA is played by non-specific interactions. Because of these interactions the protein molecule moves more slowly on DNA than in the solution, and the molecule that just dissociated from DNA has a higher probability to return back. Both these effects increase the search time for each protein molecule. However, non-specific interactions also increase the local concentration of proteins on DNA that search in parallel, and this reduces the overall search time because the process is stopped when any of the proteins reach the target site. It suggests that, only for some range of parameters (intermediate binding energies and intermediate protein concentrations), the facilitated diffusion might become more efficient and faster than an ordinary three-dimensional search. One might suggest also that theoretical models which do not include correlation effects due to non-specific interactions probably cannot fully capture the complex behavior during the protein target search on DNA.

IV. Summary and conclusions

We have developed a new lattice Monte Carlo computational method to analyze the process of protein search for the target on DNA molecules. It allowed us to test existing theoretical approaches to the problem by analyzing search dynamics for different sets of parameters. Specifically, search times in

the presence of DNA (facilitated mechanism) and in the absence of the DNA chain (Smoluchowski diffusion) have been computed and compared in order to clarify underlying mechanisms.

It is found that: (1) the ratio of search times as a function of the strength of non-specific interactions shows a non-monotonous behavior. This result is consistent with predictions from all theoretical approaches, and it also agrees with available experimental observations. It implies that the facilitated mechanism could be efficient only for the intermediate range of interactions. For large attractive interactions the concentration of proteins adsorbed to DNA is large, but they are too slow and facilitation cannot be reached. At weak interactions proteins move relatively fast on the DNA chain, but their local concentration is too small to support the efficient search in this regime. (2) Monte Carlo simulations also show that the ratio of search times as a function of the free protein concentration is also a non-monotonous function. This observation agrees with the correlated models of the protein search, but not with uncorrelated approaches. It indicates that the facilitated mechanism might efficient only for intermediate range of protein concentrations. For low concentrations proteins mostly involved in binding and unbinding events, and there is not enough scanning of DNA segments. At very large concentrations the average distance between the target and protein in the solution becomes so small that binding to DNA is not required for the search acceleration. (3) Our theoretical and computational results suggest that non-specific interactions are crucial for the fast and efficient protein search. The role of these interactions is a complex one. Interactions slow down proteins moving along the DNA, and the probability for the proteins that just dissociated to return to already scanned segment is larger. However, interactions also increase the local concentration of proteins absorbed to DNA and it accelerates the scanning of DNA for the target. It will be critically important to test this emerging theoretical picture directly in experiments to better quantify these effects.

Acknowledgements

The authors would like to acknowledge the support from the Welch Foundation (Grant No. C-1559), and the U.S. National Science Foundation (Grant No. ECCS-0708765). The authors also would like to thank G. T. Barkema and A. G. Cherstvy for useful discussions and technical help.

References

1 B. Alberts, A. Johnson, J. Lewis, M. Raff, K. Roberts and P. Walter, *Molecular Biology of the Cell*, Garland Science, 4th edn, 2002.
2 H. Lodish, A. Berk, S. L. Zipursky, P. Matsudaira, D. Baltimore and J. Darnell, *Molecular Cell Biology*, W.H. Freeman and Company, New York, 4th edn, 2002.
3 A. Riggs, S. Bourgeois and M. Cohn, *J. Mol. Biol.*, 1970, **53**, 401.
4 H. Gutfreund, *Kinetics for the Life Sciences*, Cambridge University Press, 1995.
5 O. G. Berg, R. B. Winter and P. H. von Hippel, *Biochemistry*, 1981, **20**, 6929.
6 R. B. Winter and P. H. von Hippel, *Biochemistry*, 1981, **20**, 6948.
7 O. G. Berg and P. H. von Hippel, *Annu. Rev. Biophys. Biophys. Chem.*, 1985, **14**, 131.
8 P. H. von Hippel and O. G. Berg, *J. Biol. Chem.*, 1989, **264**, 675.
9 T. Hu, A. Grosberg and B. Shklovskii, *Biophys. J.*, 2006, **90**, 2731.
10 M. Slutsky and L. Mirny, *Biophys. J.*, 2004, **87**, 4021.
11 M. Slutsky, M. Kardar and L. Mirny, *Phys. Rev. E: Stat., Nonlinear, Soft Matter Phys.*, 2004, **69**, 061903.
12 K. Klenin, H. Merlitz, J. Langowski and C.-X. Wu, *Phys. Rev. Lett.*, 2006, **96**, 018104.
13 S. Halford and J. Marko, *Nucleic Acids Res.*, 2004, **32**, 3040.
14 D. M. Gowers, G. G. Wilson and S. E. Halford, *Proc. Natl. Acad. Sci. U. S. A.*, 2005, **102**, 15883.
15 A. G. Cherstvy, A. B. Kolomeisky and A. A. Kornyshev, *J. Phys. Chem. B*, 2008, **112**, 4741.
16 H.-X. Zhou, *Biophys. J.*, 2005, **88**, 1608.
17 B. van den Broek, M. A. Lomholt, S.-M. Kalisch, R. Metzler and G. J. L. Wuite, *Proc. Natl. Acad. Sci. U. S. A.*, 2008, **105**, 15738.
18 S. E. Halford, *Biochem. Soc. Trans.*, 2009, **37**, 343.
19 A. G. Cherstvy, *J. Phys. Chem. B*, 2009, **113**, 4242.
20 M. Coppey, O. Benichou, R. Voituriez and M. Moreau, *Biophys. J.*, 2004, **87**, 1640.
21 Y. M. Wang, R. H. Austin and E. C. Cox, *Phys. Rev. Lett.*, 2006, **97**, 048302.
22 J. Elf, G.-W. Li and X. S. Xie, *Science*, 2007, **316**, 1191.

