Facilitated diffusion in a crowded environment: from kinetics to stochastics

2009 J. Phys. A: Math. Theor. 42 434012

(http://iopscience.iop.org/1751-8121/42/43/434012)

IOP PUBLISHING JOURNAL OF PHYSICS A: MATHEMATICAL AND THEORETICAL

J. Phys. A: Math. Theor. **42** (2009) 434012 (9pp) doi:10.1088/1751-8113/42/43/434012

Facilitated diffusion in a crowded environment: from kinetics to stochastics

Yasmine Meroz[1], Iddo Eliazar[2] and Joseph Klafter[1,3]

[1] School of Chemistry, Raymond & Beverly Sackler Faculty of Exact Sciences, Tel-Aviv University, Tel Aviv 69978, Israel
[2] Department of Technology Management, Holon Institute of Technology, PO Box 305, Holon 58102, Israel
[3] Freiburg Institute for Advanced Studies (FRIAS), University of Freiburg, 79104, Freiburg, Germany

E-mail: merozyas@post.tau.ac.il, eliazar@post.tau.ac.il and klafter@post.tau.ac.il

Received 28 July 2009, in final form 11 August 2009
Published 13 October 2009
Online at stacks.iop.org/JPhysA/42/434012

Abstract

Facilitated diffusion is a fundamental search process used to describe the problem of a searcher protein finding a specific target site over a very large DNA strand. In recent years macromolecular crowding has been recognized to affect this search process. In this paper, we bridge between two different modelling methodologies of facilitated diffusion: the physics-oriented kinetic approach, which yields the reaction rate of the search process, and the probability-oriented stochastic approach, which yields the probability distribution of the search duration. We translate the former approach to the latter, ascertaining that the two approaches yield coinciding results, both with and without macromolecular crowding. We further show that the stochastic approach markedly generalizes the kinetic approach by accommodating a vast array of search mechanisms, including mechanisms having no reaction rates, and thus being beyond the realm of the kinetic approach.

PACS numbers: 02.50Ey, 05.40Fb, 05.40Jc

1. Introduction

DNA binding proteins face the daunting task of pinpointing a specific binding site out of a very large number of DNA sequences. This biological process is of great importance and poses an interesting search problem. The simplest description of this problem is a three-dimensional diffusion-controlled bi-molecular reaction rate, formulated by the reaction

$$S + T \underset{k_d}{\overset{k_a}{\rightleftharpoons}} ST, \tag{1}$$

where S represents the searcher protein, T represents the DNA target site, k_a is the association rate and k_d is the dissociation rate. Experiments have demonstrated that DNA binding proteins find their target sites a couple of orders faster than a three-dimensional diffusion-controlled process [1–3]. On the other hand, first-passage-time reasoning implies that a search performed by one-dimensional diffusion along a DNA strand consisting of n base pairs (the search initiated from a random location on the strand) will take $O(n^2)$ time steps [4]—which is also too long a time scale. Hence, neither a three-dimensional diffusion-controlled process in the cytoplasm nor one-dimensional diffusion along the DNA can yield, on their own, the fast search times observed experimentally.

The discrepancy between the biological search and the search efficiency of diffusive processes is an important theoretical challenge. To resolve this problem various search algorithms based on intermittent search processes have been proposed [5, 6]. One such process considers the search to take place in two alternating phases: (i) *relocation*—a three-dimensional diffusion-controlled motion in the cytoplasm and (ii) *scanning*—landing on a non-specific DNA binding site and thereafter performing a one-dimensional motion along the DNA. If the specific target is found during a scanning phase then a binding reaction takes place and the search process comes to an end. If not, the protein dissociates and relocates in three-dimensional diffusion again. The time during which the protein stays associated with the DNA, t_{1D}, is exponentially distributed. The dissociation rate is defined as $\lambda_{1D} = 1/\langle t_{1D} \rangle$. This two-phase search process, termed *facilitated diffusion* in the seminal paper by Berg *et al* [3], is formulated by the reaction

$$S + D + T \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} SD + T \underset{k_{-2}}{\overset{k_2}{\rightleftharpoons}} ST + D, \tag{2}$$

where S and T are the same as in equation (1), D represents the DNA strand, and k_1, k_{-1}, k_2 and k_{-2} are the appropriate reaction rates. The facilitated diffusion search process is illustrated in figure 1.

In this paper, we discuss two approaches modelling the facilitated diffusion search process. The first is the conventional physical *kinetic approach* [3, 7], whereas the second is the more recent probabilistic *stochastic approach* presented in [8, 9]. We will show that the latter approach generalizes the former, extending the kinetic results to different facilitated diffusion mechanisms (including anomalous mechanisms) and yielding probability distributions of the search durations (rather than rates and means)—thus providing a more general and robust model of DNA search processes.

In recent years it has been realized that macromolecular crowding in the cell may well affect the search process. The conventional kinetic approach has been adjusted to take the crowding effect into account. We show how the stochastic approach can also take the crowding effect into account, incorporating it more flexibly and robustly than the kinetic approach.

The remainder of this paper is organized as follows. The kinetic approach, in both the non-crowded and crowded cases, is described in sections 2 and 3. The stochastic approach is described in section 4. In section 5, we show that the two approaches are in agreement, and in section 6 we explain how the stochastic approach generalizes the kinetic approach.

2. The kinetic approach

The diffusion-controlled reaction rate k between two diffusive reactants in a medium is given by the Debye–Smoluchowski equation:

$$k = (4\pi\kappa N_A) \cdot D_3 \cdot L, \tag{3}$$

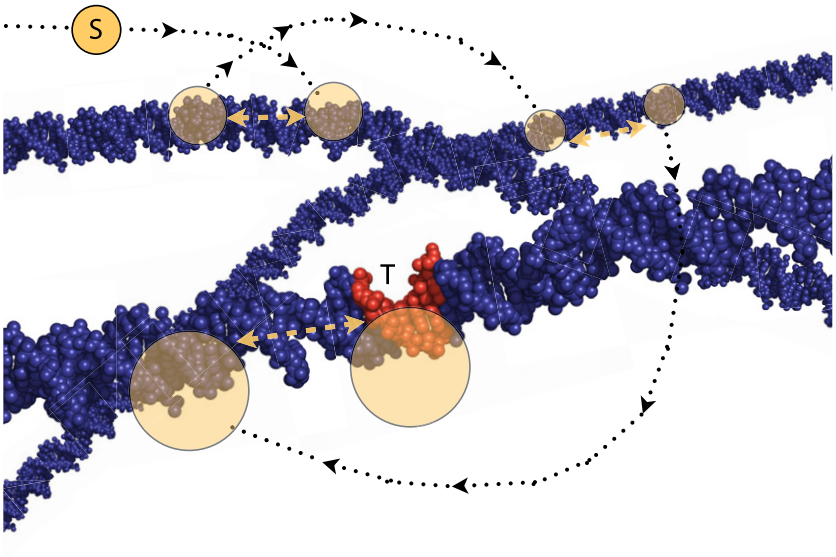

Figure 1. A schematic illustration of the facilitated diffusion search process. A searcher protein S (depicted by a circle) seeks a specific target site T over a long DNA strand (in blue). The searcher initiates with a three-dimensional diffusion-controlled motion in the cytoplasm, associates with the DNA, and performs a scanning phase moving along the DNA (represented by an arrow). The searcher then dissociates into the cytoplasm, relocates and performs a second scanning phase. In the third scanning phase the searcher finds the target, and the search process comes to an end.

(This figure is in colour only in the electronic version)

where κ is a unit-less factor which takes into account possible interactions such as steric or electrostatic, N_A is the Avogadro number, D_3 is the diffusion coefficient of the reactants and L is the interaction radius.

The kinetic approach approximates the facilitated diffusion process by considering the searcher protein and the target site as two diffusive reactants. The reaction rate $\tilde{k}$ of the approximated facilitated diffusion process is given by the following modification of the Debye–Smoluchowski equation [3, 7]:

$$\tilde{k} = C \cdot \tilde{D}_3 \cdot \tilde{L}, \tag{4}$$

(the constant C in equation (4) is the counterpart of the constant $(4\pi\kappa N_A)$ in equation (3)). $\tilde{D}_3$ is the effective diffusion coefficient of the facilitated diffusion process, and it accounts for the relocation phase. $\tilde{L}$ is the effective interaction radius of the facilitated diffusion process, and it accounts for the scanning phase.

In the relocation phase, the three-dimensional diffusion is effectively slowed down by the constant association and dissociation of the protein searcher to the DNA. The diffusion coefficient is therefore normalized according to the concentration of nonspecific DNA sites c_{ns}, and the binding constant between the protein and nonspecific DNA sites K_{RD}, producing the effective diffusion coefficient

$$\tilde{D}_3 \approx \frac{D_3}{1 + K_{RD} c_{ns}}. \tag{5}$$

The average number of DNA base pairs covered by the searcher during the association time t_{1D} is equal, on average, to twice the square root of the scanning motion's mean square displacement (MSD): $2\langle\sqrt{\mathbf{r}^2(\langle t_{1D}\rangle)}\rangle$. This is also the effective interaction radius since if the target site is situated within such a distance from the searcher's initial association site then the searcher will, on average, reach it during its scanning phase. Hence, considering a Brownian motion scanning mechanism with the diffusion coefficient D_1, the effective interaction radius is given by

$$\tilde{L} = l + 2\sqrt{D_1\langle t_{1D}\rangle}, \tag{6}$$

where l is the size of one base pair. Finally, substituting equations (5) and (6) into equation (4) yields the reaction rate $\tilde{k}$ of the facilitated diffusion process.

3. Macromolecular crowding

The effects of macromolecular crowding on the search process, caused by proteins already bound to the DNA, have been incorporated into the conventional kinetic approach [7]. These bound proteins crowd the DNA, making it less accessible for the searcher protein to bind to and thus impeding protein association to the DNA. Setting v to be the DNA vacancy, i.e. the fraction of DNA free of bound proteins, the fraction of accessible DNA to a searcher protein is given by $f_v = v\,\mathrm{e}^{1-\frac{1}{v}}$ [10, 11], and the effective diffusion coefficient now takes the form [7]

$$\tilde{D}_3 \approx \frac{D_3}{1 + K_{RD}c_{ns}f_v}. \tag{7}$$

The bound proteins also obstruct the one-dimensional scanning of the searcher protein along the DNA, and are thus termed *roadblocks* [7]. This hindering of the searcher protein leads to a smaller effective interaction radius [7]:

$$\tilde{L} \approx l + 2\sqrt{\frac{D_1\langle t_{1D}\rangle}{1 + (\langle t_{1D}\rangle/t_x)^{1/2}}}, \tag{8}$$

where $t_x = 1/\pi D_1\rho_v^2$ is the diffusion time across the average gap between roadblocks and $\rho_v = (1-v)/dv$ denotes the density of roadblocks (d being the size of the search protein and roadblocks) [10, 11]. In the limit $\langle t_{1D}\rangle/t_x \ll 1$ the searcher protein is not affected by the roadblocks (since it dissociates before coming into contact with them), and the non-crowded result of equation (6) is recovered. On the other hand, in the limit $\langle t_{1D}\rangle/t_x \gg 1$ the searcher protein is affected by the roadblocks. This case of one-dimensional motion along the DNA where no mutual exchanges of the diffusants are allowed (i.e. the order of the diffusants is kept) is termed *single file diffusion* (SFD). Taking this limit and substituting $t_x = 1/\pi D_1\rho_v^2$ into equation (8) one can recognize the scanning motion's MSD corresponding to SFD [12, 13]:

$$\tilde{L} \approx l + \sqrt{2}\sqrt{\frac{2}{\rho_v}}\left(\frac{D_1}{\pi}\right)^{1/4}\langle t_{1D}\rangle^{1/4}. \tag{9}$$

It should be noted that when roadblocks are not taken into consideration then $v = 1$, yielding $f_v = 1$ and $\rho_v = 0$. In this case, equations (7), (8) reduce, respectively, to equations (5), (6) of the non-crowded case.

4. The stochastic approach

Eliazar *et al* [8, 9] introduced a stochastic model of the DNA search problem on circular DNA strands (plasmids), enabling the calculation of the distributions of search times (rather than

their means and rates), and expanding the range of possible mechanisms of both the relocation and scanning phases. The approach is probabilistic, retaining the facilitated diffusion setting without resorting to reaction-rate equations.

It was shown [8, 9] that the effect of the scanning mechanism on the performance of the search process can be represented by a function $\Psi(\lambda_{1D})$ termed the *scan function* (for further details see the appendix). Linear scanning with constant velocity v results in the scan function $\Psi(\lambda_{1D}) = v\langle t_{1D}\rangle = v/\lambda_{1D}$. Brownian motion scanning, with a diffusion coefficient D_1, results in the scan function

$$\Psi(\lambda_{1D}) = \sqrt{2D_1\langle t_{1D}\rangle} = \sqrt{\frac{2D_1}{\lambda_{1D}}}. \tag{10}$$

Another important scanning mechanism is fractional Brownian motion (fBm), a generalization of Brownian motion introduced by Mandelbrot and Van-Ness [14], and the quintessential example of a continuous random motion with correlated fluctuations. fBm is governed by a Hurst exponent H ($0 < H < 1$) which quantifies the statistical self-similarity of its sample-path trajectories [15]. In the Hurst range $0 < H < 1/2$ fBm has negatively correlated fluctuations, its dynamics are anti-persistent, its memory is short-ranged and its propagation is subdiffusive. In the Hurst range $1/2 < H < 1$ fBm has positively correlated fluctuations, its dynamics are persistent, its memory is long-ranged and its propagation is superdiffusive. The Hurst exponent $H = 1/2$ retrieves Brownian motion. fBm scanning motion, with Hurst exponent H, results in the scan function

$$\Psi(\lambda_{1D}) = c_H\langle t_{1D}\rangle^H = \frac{c_H}{(\lambda_{1D})^H} \tag{11}$$

(c_H is a positive constant depending on the Hurst exponent H).

We note that the three aforementioned scanning mechanisms resulted in power-law scan functions. We further note that power-law scan functions coincide, up to a multiplicative constant, with the root of the MSD of the corresponding scanning motions: $\Psi(\lambda_{1D}) \approx \sqrt{\langle \mathbf{r}^2(\langle t_{1D}\rangle)\rangle} = \sqrt{\langle \mathbf{r}^2(1/\lambda_{1D})\rangle}$.

The DNA coverage rate λ_{cov}, i.e. the rate at which the searcher protein covers the DNA, is defined as the ratio of the average amount of DNA covered per search cycle and the average duration of a search cycle (the search cycle being a three-dimensional cytoplasm relocation phase, followed by a one-dimensional DNA scanning phase). In the case of power-law scan functions the average amount of scanned DNA per search cycle is $l + \Psi(\lambda_{1D})$. The average duration of a search cycle is $\langle t_{3D}\rangle + \langle t_{1D}\rangle$, where t_{3D} denotes the relocation time (i.e. the random time spent during the relocation phase of the searcher protein, from DNA dissociation till the subsequent DNA reassociation). Hence, the DNA coverage rate is given by

$$\lambda_{\text{cov}} \equiv \frac{l + \Psi(\lambda_{1D})}{\langle t_{3D}\rangle + 1/\lambda_{1D}}. \tag{12}$$

The facilitated diffusion process is a sequence of independent search cycles which may be viewed as a series of Bernoulli trials, where *success* is defined as the searcher protein finding the target site. The number of trials till first success is thus geometrically distributed. This observation, in turn, implies that the overall search time, t_{find}, is asymptotically exponentially distributed:

$$\mathbf{P}(t_{\text{find}} > t) \approx \exp(-\lambda_{\text{cov}} \cdot t). \tag{13}$$

A further generalization is the introduction of m searcher proteins, operating independently and simultaneously, which improves the DNA coverage rate λ_{cov} by the factor m, namely

$$\mathbf{P}(t_{\text{find}} > t) \approx \exp(-m\lambda_{\text{cov}} \cdot t). \tag{14}$$

5. The stochastic approach recovers the kinetic results

In this section, we show that the stochastic approach recovers the results of the kinetic approach in both the conventional non-crowded case [3] and the crowded case [7]. To this end we compare the rate constant $\tilde{k}_a$ of the kinetic approach to the coverage rate $\lambda_{\text{cov}} = 1/\langle t_{\text{find}} \rangle$ of the stochastic approach. Both these quantities admit a multiplicative representation composed of two factors—one describing the relocation phase and the other describing the scanning phase. We compare these two factors respectively.

Consider the relocation phase. One can reach the kinetic relocation term from the relocation term defined in the stochastic approach, expressed by the reciprocal of the average cycle time $\langle t_{3D} \rangle + \langle t_{1D} \rangle$. From the kinetic viewpoint the expression for the mean relocation time is given by [3, 7]: $\langle t_{3D} \rangle = 1/\lambda_{1D} K_{RD} c_{ns} f_v$. Using this expression one gets $\langle t_{3D} \rangle + \langle t_{1D} \rangle = \langle t_{3D} \rangle (1 + c_{ns} K_{RD} f_v)$. Also, defining ξ as the average distance between the nearest DNA segments, one can make a rough estimate that $\langle \xi^2 \rangle = D_3 \langle t_{3D} \rangle$. Combining these expressions we obtain

$$\frac{1}{\langle t_{3D} \rangle + \langle t_{1D} \rangle} \approx \frac{D_3}{1 + c_{ns} K_{RD} f_v} = \tilde{D}_3, \tag{15}$$

successfully recovering the kinetic result (both with and without crowding) from the stochastic approach.

Consider now the scanning phase. In the case of no crowding, the coincidence of the scanning phase terms is immediate: substituting the Brownian scan function of equation (10) into the stochastic-approach term $l + \Psi(\lambda_{1D})$ yields the kinetic-approach result of equation (6). In the crowded case, one has to choose an appropriate stochastic-approach scan function in order to meet the corresponding SFD behaviour described by equation (8). It has been shown [16] that SFD is suitably described by fBm with $H = 1/4$. Hence, the corresponding MSD takes the form [12, 13] $\langle r^2(t_{1D}) \rangle = 2\sqrt{D_1 t_{1D}/\pi \rho_v^2}$. This MSD yields the scan function

$$\Psi(\lambda_{1D}) = \sqrt{\frac{2}{\rho_v}} \left(\frac{D_1}{\pi} \frac{1}{\lambda_{1D}} \right)^{1/4}. \tag{16}$$

Finally, substituting the scan function of equation (16) into the stochastic-approach term $l + \Psi(\lambda_{1D})$ yields the kinetic-approach result of equation (9).

6. The stochastic approach generalizes the kinetic approach

In this section we show how the stochastic approach generalizes the kinetic approach, yielding a range of further analyses and directions. We focus on two key issues: anomalous halting of the scanning mechanism, and anomalous relocation mechanisms.

Consider the incorporation of a halting mechanism affecting the scanning phase. The halting occurs randomly in time according to a given rate. Once halted, the scanning motion freezes for a random duration t_{halt}, and thereafter resumes its motion. Such halting can stem from a distribution of energetic traps. Anomalous halting corresponds to the case of infinite-mean halting durations $\langle t_{\text{halt}} \rangle = \infty$, diverging due to heavy-tailed probability laws of the form $\mathbf{P}(t_{\text{halt}} > t) \approx 1/t^{\beta}$ with exponent $0 < \beta < 1$. Anomalous halting is one possible theoretical model explaining the biologically prevalent subdiffusive molecular transport [17].

fBm scanning motion with Hurst exponent H, combined with anomalous halting with exponent β, results in the scan function [8, 9]:

$$\Psi(\lambda_{1D}) = \frac{c_H}{\left(\lambda_{1D} + b\lambda_{1D}^{\beta}\right)^H} \tag{17}$$

(b is a positive constant depending on the halting mechanism). Note that this scan function captures all the aforementioned kinetic results: equation (8) is obtained by setting $H = 1/2$ (Brownian motion scanning), $\beta = 1/2$ (halting with the Lévy–Smirnov exponent) and $b = \sqrt{1/t_x}$; equations (6) and (9) are obtained by setting $H = 1/2$ (Brownian motion scanning) and $H = 1/4$ (fBm scanning representing SFD), respectively, and $b = 0$ (no halting).

An important feature of the stochastic approach is its modularity, i.e. that one can easily change the Hurst exponent H and the halting exponent β (in equation (17)) to fit different diffusion and anomalous diffusion models. Moreover, fBm is just one special class of the much broader class of self-similar stochastic motions [15]. The stochastic approach accommodates scanning mechanisms performed by general self-similar motions: both equations (11) and (17) hold valid for any self-similar scanning motion with continuous sample-path trajectories. This fact, combined with a recent work on the universal generation of statistical self-similarity [18], allows for the incorporation of a large class of highly non-Brownian scanning processes.

Let us now turn to discuss the case of infinite-mean relocation times, which lies beyond the realm of the kinetic approach. This case has been observed experimentally [17] and is therefore of great importance. The kinetic approach is restricted to regular diffusion whereas the stochastic approach is not. Indeed, the stochastic model allows for very general scanning and relocation mechanisms, including various anomalous diffusion motions (both in one and three dimensions). An important application is the case of anomalous relocation mechanisms yielding heavy-tailed relocation times [19]: namely, infinite-mean relocation times $\langle t_{3D} \rangle = \infty$, diverging due to heavy-tailed probability laws of the form $\mathbf{P}(t_{3D} > t) \approx 1/t^{\alpha}$ with exponent $0 < \alpha < 1$. In this case the counterpart of equation (14) is

$$\mathbf{P}(t_{\text{find}} > t) \sim \left(\frac{l + \Psi(\lambda_{1D})}{\alpha} \cdot t^{\alpha} \right)^{-m} . \tag{18}$$

The asymptotically exponential search duration of equation (14) is replaced by the asymptotically Pareto search time of equation (18). Note the marked differences between these two cases: in the infinite-mean case the overall search time t_{find} is non-exponential and thus has no rate. We emphasize that the notion of rate, the very foundation of the kinetic approach, is meaningless when dealing with infinite-mean relocation times. Moreover, in equation (14) the number of searchers m affects only the exponential rate, whereas in equation (18) the number of searchers m determines the order of finite moments that the search duration t_{find} possesses. Specifically (in the infinite-mean case), as the number of searchers m increases, the overall search time t_{find} gains more and more converging moments.

We conclude this section with a discussion emphasizing the central role of the relocation mechanism. To this end, consider the DNA strand to consist of $n \gg 1$ base pairs. In [8, 9] it was shown that, in the case of finite-mean relocation times $\langle t_{3D} \rangle < \infty$, the mean search time $\langle t_{\text{find}} \rangle$ is of order $O(n)$. This result is general, and holds for arbitrary scanning mechanisms. On the other hand, if the search process is performed by the scanning motion alone (with no relocation) then [8, 9]: (i) self-similar scanning motions with Hurst exponent H yield mean search times $\langle t_{\text{find}} \rangle$ of order $O(n^{1/H})$ (e.g. $O(n^2)$ in the case of Brownian motion scanning, and $O(n^4)$ in the case of SFD scanning); (ii) the incorporation of anomalous halting yields infinite-mean search times $\langle t_{\text{find}} \rangle = \infty$ (for whatever scanning motion applied). These observations pinpoint the effect of relocation: in the first case the relocation mechanism is most beneficial since it reduces the mean search times $\langle t_{\text{find}} \rangle$ from the higher order $O(n^{1/H})$ to the lower order $O(n)$; in the second case the relocation mechanism is essential for it reduces the infinite-mean search times $\langle t_{\text{find}} \rangle = \infty$ to the finite order $O(n)$. For an analogous analysis regarding the case of infinite-mean relocation times $\langle t_{3D} \rangle = \infty$, the readers are referred to [8, 9].

7. Conclusions

In this paper, we compared two approaches modelling the biological facilitated diffusion search process—the physically-oriented kinetic approach, and the probabilistically-oriented stochastic approach. We ascertained that the results obtained by the stochastic approach indeed confirm and coincide with the well-established kinetic approach results, both with and without DNA macromolecular crowding. Having expressed the parameters of the stochastic approach in terms of the parameters of the kinetic approach, we generalized the kinetic reaction-rate results to the more informative stochastic probability-distribution results. Moreover, the stochastic approach allows for the accommodation of very general relocation and scanning mechanisms, including anomalous ones, which lay beyond the realm of the kinetic approach. This was exemplified by the important case of infinite-mean relocation times, where the kinetic notion of reaction rates is no longer valid and the overall search times are Paretian.

Bridging between the kinetic approach and the stochastic approach, this paper enables researchers to translate from a physical to a probabilistic nomenclature. This, in turn, provides wide access to the stochastic modelling of the DNA search problem, and to the advantages this approach offers: modularity, generality and robustness.

Acknowledgments

The authors thank Aleksei Chechkin and Alessandro Taloni for fruitful discussions.

Appendix. The scan function appearing in the stochastic approach

Consider a search process initiating from a random location along a circular DNA strand consisting of n base pairs. Let (i) S_n denote the random time it takes a searcher protein to find the target site using scanning alone (i.e. no relocation involved); (ii) T_n denote the random time it takes a searcher protein to find the target site using the two-phase facilitated diffusion process; (iii) t_{3D} denote the relocation time, i.e. the random time spent during the relocation phase of the searcher protein, from DNA dissociation till the subsequent DNA reassociation. Assuming that each relocation phase lands the searcher protein randomly along the DNA strand, the stochastic approach establishes a closed-form formula expressing the Laplace transform of the facilitated diffusion search time T_n as a function of Laplace transforms of the relocation time t_{3D} and the scanning time S_n.

In the case of long DNA strands, an asymptotic analysis of the limit $n \to \infty$ is required. Let $\hat{S}_n(\theta)$ $(\theta \geqslant 0)$ denote the Laplace transform of the scanning time S_n, and let the stochastic limit $t_{\text{find}} = \lim_{n\to\infty} T_n$ denote the overall search time. Analysis shows [8, 9] that the stochastic limit t_{find} exists if and only if the following condition holds: the limit $\Psi(\theta) := \lim_{n\to\infty} n\hat{S}_n(\theta+\epsilon_n)$ exists for all positive θ and for all non-negative valued sequences $\{\epsilon_n\}_{n=1}^{\infty}$ decaying to zero. The limit $\Psi(\theta)$, termed the scan function, captures the effect of the scanning mechanism on the performance of the facilitated diffusion search process. In other words, the scan function $\Psi(\theta)$ codes all the information regarding the scanning mechanism which is relevant to the computation of the overall search time t_{find}.

References

[1] Richter P and Eigen M 1974 Diffusion controlled reaction rates in spheroidal geometry: application to repressor operator association and membrane bound enzymes *Biophys. Chem.* **2** 255–63

[2] Riggs A D, Bourgeois S and Cohn M 1970 The lac repressor–operator interaction: 3. Kinetic studies *J. Mol. Biol.* **53** 401–17
[3] Berg O G, Winter R B and von Hippel P H 1981 Diffusion-driven mechanisms of protein translocation on nucleic acids: 1. Models and theory *Biochemistry* **20** 6929–48
[4] Montroll E W 1969 Random walks on lattices: III. Calculation of first-passage times with application to exciton trapping on photosynthetic units *J. Math. Phys.* **10** 753
[5] Lomholt M A, Koren T, Metzler R and Klafter J 2008 Lévy strategies in intermittent search processes are advantageous *Proc. Natl Acad. Sci.* **105** 11055–9
[6] Coppey M, Benichou O, Voituriez R and Moreau M 2004 Kinetics of target site localization of a protein on DNA: a stochastic approach *Biophys. J.* **87** 1640–9
[7] Li G, Berg O G and Elf J 2009 Effects of macromolecular crowding and DNA looping on gene regulation kinetics *Nat. Phys.* **5** 294–7
[8] Eliazar I, Koren T and Klafter J 2007 Searching circular DNA strands *J. Phys.: Condens. Matter* **19** 065140
[9] Eliazar I, Koren T and Klafter J 2008 Parallel search of long circular strands: modeling, analysis, and optimization *J. Phys. Chem.* B **112** 5905–9
[10] McGhee J D and von Hippel P H 1974 Theoretical aspects of DNA–protein interactions: co-operative and non-co-operative binding of large ligands to a one-dimensional homogeneous lattice *J. Mol. Biol.* **86** 469–89
[11] Flyvbjerg H, Keatch S A and Dryden D T F 2006 Strong physical constraints on sequence-specific target location by proteins on DNA *Nucleic Acids Res.* **34** 2550–7
[12] Kärger J 1992 Straightforward derivation of the long-time limit of the mean-square displacement in one-dimensional diffusion. *Phys. Rev.* A **45** 4173–4
[13] Taloni A and Marchesoni F 2006 Single-file diffusion on a periodic substrate *Phys. Rev. Lett.* **96** 020601
[14] Mandelbrot B B and van Ness J W 1968 Fractional Brownian motions, fractional noises and applications *SIAM Rev.* **10** 422–37
[15] Embrechts P and Maejima M 2002 *Self-similar Processes* (Princeton, NJ: Princeton University Press)
[16] Taloni A and Lomholt M A 2008 Langevin formulation for single-file diffusion *Phys. Rev.* E **78** 051116
[17] Golding I and Cox E C 2006 Physical nature of bacterial cytoplasm *Phys. Rev. Lett.* **96** 098102
[18] Eliazar I and Klafter J 2009 Universal generation of statistical self-similarity: a randomized central limit theorem *Phys. Rev. Lett.* **103** 040602
[19] Metzler R and Klafter J 2000 The random walk's guide to anomalous diffusion: a fractional dynamics approach *Phys. Rep.* **339** 1–77