Facilitated Diffusion of Proteins on Chromatin

O. Bénichou,[1] C. Chevalier,[1] B. Meyer,[1] and R. Voituriez[1]

[1]*Laboratoire de Physique Théorique de la Matière Condensée CNRS-UPMC, 4 Place Jussieu, 75255 Paris Cedex, France*

(Received 16 June 2010; published 20 January 2011)

We present a theoretical model of facilitated diffusion of proteins in the cell nucleus. This model, which takes into account the successive binding and unbinding events of proteins to DNA, relies on a fractal description of the chromatin which has been recently evidenced experimentally. Facilitated diffusion is shown quantitatively to be favorable for a fast localization of a target locus by a transcription factor and even to enable the minimization of the search time by tuning the affinity of the transcription factor with DNA. This study shows the robustness of the facilitated diffusion mechanism, invoked so far only for linear conformations of DNA.

DOI: 10.1103/PhysRevLett.106.038102 PACS numbers: 87.10.Mn, 05.40.Fb

The now well established theory of facilitated diffusion explains how DNA-binding proteins can in principle find their target sites on DNA efficiently. This model describes search trajectories as alternating phases of free diffusion in the bulk cytoplasm and one-dimensional diffusion along the DNA strand, called sliding, which is made possible by sequence-independent interactions of proteins with DNA. Since the seminal work in Ref. [1], such pathways have been evidenced experimentally both *in vivo* [2] and *in vitro* [3–5] thanks to single molecule technics, and theoretical aspects have been refined [6–9], in particular, highlighting that such strategies can minimize the search time for a target site by a proper tuning of the protein-DNA interaction [10–13].

All these theoretical approaches rely on a schematic description of DNA as a one-dimensional linear chain along which a protein can diffuse, surrounded by an homogeneous medium in which the protein performs regular diffusion. More recently, crowding effects have been incorporated in these models for both the sliding [14] and the bulk cytoplasmic phases [4], leading to more realistic descriptions of gene regulation kinetics in prokaryotes.

However, such models are clearly inapplicable to eukaryotes, in which the DNA is packed in the cell nucleus [15] and forms a complex structure called chromatin which is far from a simple one-dimensional chain. Even if a full bottom-up description of the *in vivo* DNA organization remains out of reach, theoretical ideas [16] and now growing experimental evidences indicate that the chromatin has a hierarchized architecture which displays fractal properties at least over the 100 nm–10 μm range. Indeed, textural image analysis, neutron scattering [17], rheology technics [18], and more recently the Hi-C method [19] revealed independently a fractal structure of the chromatin characterized by a fractal dimension d_f which was found in the range 2.2–3. Despite this complex structure of the chromatin, the switching dynamics of proteins between a DNA bound state and a freely diffusing state, which characterizes facilitated diffusion in prokaryotes, seems to be also at work in the nucleus, as evidenced on the examples of histons, high-mobility group proteins, and more generally chromatin binding proteins [20]. This naturally raises the questions of determining whether the classical facilitated diffusion mechanism can be efficient also in the complex nuclear environment and whether it can be used to regulate and optimize gene expression in eukaryotes. This Letter presents a first theoretical model which quantitatively addresses these two questions.

At the theoretical level, modeling facilitated diffusion in the cell nucleus raises two problems: (i) first, to take into account the switching dynamics of the protein between a state bound to the chromatin and a freely diffusing state in the nucleoplasm and (ii) second, to model the diffusion phase of a protein bound to a complex structure such as chromatin. Point (i) has been studied in the context of intermittent search strategies [21], and general methods to calculate mean search times for intermittent trajectories have been developed. On the other hand, the full distribution of the first-passage time (FPT) for diffusion in fractals has recently been obtained in Ref. [22] and enables us to tackle point (ii) under the assumption, backed by experiments [17–19], that the chromatin is fractal.

In this Letter, we gather and extend these new tools to develop a theoretical model of facilitated diffusion in the cell nucleus. More precisely, we calculate analytically the mean search time for a target for a protein which alternates diffusion phases on the chromatin, which is assumed to have a fractal structure, and free diffusion phases in the nucleoplasm (see Fig. 1). Under these hypotheses, we show quantitatively that facilitated diffusion in eukaryotes can significantly speed up the search process and that it enables us to minimize the search time by tuning the affinity of the protein with DNA. These results are qualitatively similar to the case of prokaryotes and suggest that facilitated diffusion is a robust mechanism. At the theoretical level, this study yields as a by-product the calculation of the

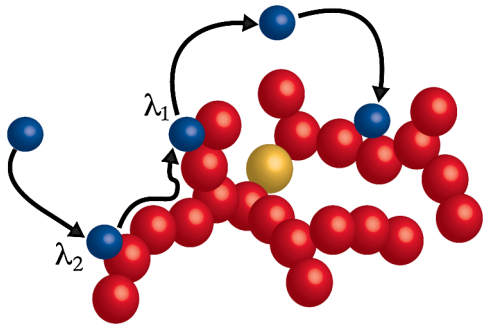

FIG. 1 (color online). Facilitated diffusion on chromatin: A chromatin binding protein (blue) searches for a target locus (yellow) on chromatin (red). The binding (unbinding) rate is denoted by λ_2 (λ_1).

distribution of the FPT averaged over the starting point for a particle diffusing in a fractal structure, which remained a challenge in the field [23–26].

Search time distribution for simple diffusion in a fractal medium.—We first consider a protein which remains in an adsorbed state and diffuses on the chromatin, which is modeled by a discretized domain $\mathcal{D}$ of volume N and, following experimental observations [17–19], is characterized by a fractal dimension d_f and a typical size $R \propto N^{1/d_f}$. We calculate here the distribution of the search time, defined as the FPT at the target, averaged over the starting position of the protein. This first technical step is necessary to address the problem of facilitated diffusion discussed in the next paragraph; besides, it is an important theoretical question. The protein of position $\mathbf{r}(t)$ is assumed to perform a symmetric nearest neighbor random walk on the chromatin with a constant hopping rate (set to 1). To account for the complex organization of DNA-DNA contact points, the chromatin cannot be described as a linear chain: It is effectively branched even if the DNA is linear, yielding a connectivity potentially larger than 2 which takes into account intersegmental transfer. The resulting dynamics is characterized by the walk dimension d_w defined through the scaling of the mean square displacement with time: $\langle \mathbf{r}^2(t) \rangle \propto t^{2/d_w}$, and the nuclear membrane which bounds the chromatin is assumed to act as reflecting walls. As we proceed to show, the search time distribution is independent of these microscopic details of the chromatin conformation and is governed only by its larger scale properties, which are characterized by d_f and d_w.

We denote by $W_{ji}(t)$ the propagator, i.e., the probability that the protein, starting at site i at $t = 0$, is at site j at time t, and write W_j^{stat} for the stationary probability at site j. We will make use of the pseudo-Green function of the walk defined by $H_{ji} = \int_{t=0}^{\infty} [W_{ji}(t) - W_j^{\mathrm{stat}}] dt$, and the Laplace transform of a generic function $f(t)$ will be denoted by $\hat{f}(s)$. We are here interested in the global FPT (GFPT) at a given target site T, which is the FPT at site T *averaged* over the starting site S with weight W_S^{stat}. We thus define the probability density Φ_T of the GFPT by $\Phi_T(t) = \sum_{j=1}^{N} W_j^{\mathrm{stat}} P_{Tj}(t)$, where P_{Tj} is the probability density of the FPT at T starting from *a given* site j.

General expressions have been derived for the first moment of both the FPT [27] and the GFPT [28] (see also [23–26] for specific examples), and more recently, the higher FPT moments have been determined in the large-volume limit $N \gg 1$ [22]. In the case of noncompact exploration ($d_w < d_f$), it reads

$$\langle \tau_{TS}^n \rangle = n! \langle \tau_T \rangle^n \frac{H_{TT} - H_{TS}}{H_{TT}}, \tag{1}$$

where $\langle \tau_T \rangle = H_{TT}/W_T^{\mathrm{stat}}$ is the mean GFPT [28]. Using next that $H_{ji} W_i^{\mathrm{stat}} = H_{ij} W_j^{\mathrm{stat}}$, deduced from detailed balance, and averaging over S, we obtain $\langle \tau_T^n \rangle = n! \langle \tau_T \rangle^n$, from which it can be deduced immediately that the GFPT distribution is a simple exponential of mean $\langle \tau_T \rangle$.

In the compact case ($d_w > d_f$), however, it can be shown that, due to a stronger dependence of $\langle \tau_{TS}^n \rangle$ on S [22,27], the average over S must be taken before the large N limit, which makes the asymptotic form of Ref. [22] unusable for this purpose. Alternatively, one can make use of the renewal equation [29] which reads in Laplace space $\hat{P}_{TS}(s) = \hat{W}_{TS}(s)/\hat{W}_{TT}(s)$. By using the symmetry relation $W_{ji}(t) W_i^{\mathrm{stat}} = W_{ij}(t) W_j^{\mathrm{stat}}$, the average over S can be taken and yields an exact expression of the GFPT distribution: $\hat{\Phi}_T(s) = W_T^{\mathrm{stat}}/[s \hat{W}_{TT}(s)]$. Taking next the large-volume limit, the propagator $\hat{W}_{TT}(s)$ can be evaluated by using the O'Shaughnessy and Procaccia formalism [30], which leads to

$$\hat{\Phi}_T(s) = \left(\frac{4}{A \langle \tau_T \rangle s} \right)^{\nu} \frac{\Gamma(1+\nu)}{\Gamma(1-\nu)} \frac{I_\nu(\sqrt{A \langle \tau_T \rangle s})}{I_{-\nu}(\sqrt{A \langle \tau_T \rangle s})}, \tag{2}$$

where $\nu = d_f/d_w$, $A = 2(1-\nu^2)/\nu$, and I_ν (and later J_ν) denote Bessel functions. Introducing the rescaled variable $\theta = t/\langle \tau_T \rangle$, one finally obtains by inverse-Laplace transforming (2):

$$Q_T(\theta) = \begin{cases} \exp(-\theta) & (d_w < d_f), \\ \frac{2^{2\nu}\nu}{(1-\nu^2)} \frac{\Gamma(1+\nu)}{\Gamma(1-\nu)} \sum_{k=0}^{\infty} \alpha_k^{1-2\nu} \frac{J_\nu(\alpha_k)}{J_{1-\nu}(\alpha_k)} \exp\left(-\frac{\alpha_k^2 \nu}{2(1-\nu^2)} \theta \right) & (d_w > d_f), \end{cases} \tag{3}$$

where the α_k's are the real zeros of $J_{-\nu}$.

This very general result, confirmed by numerical simulations on various fractal sets (see Fig. 2 and [31]) shows that the GFPT distribution takes a universal form indexed by d_f and d_w only for any fractal structure, independently of its microscopic details. In particular, it shows the

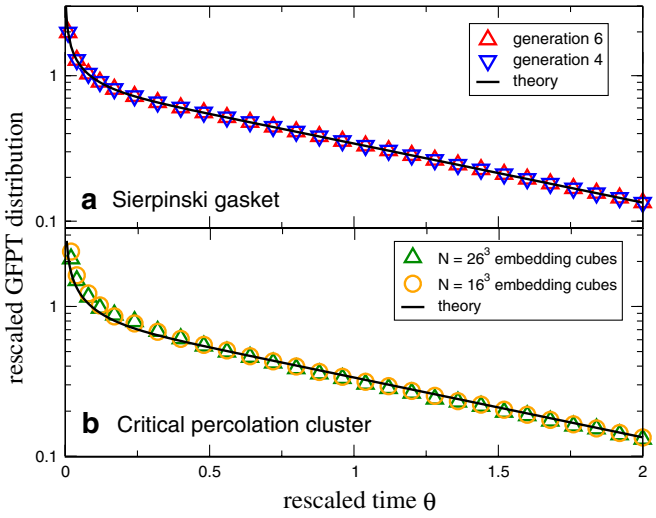

FIG. 2 (color online). Distribution of the GFPT for diffusion in fractal medium in the compact case. Numerical simulations for different system sizes N (symbols) are plotted against the theoretical prediction of Eq. (3) (plain lines). (a) Sierpinski gasket (target at the apex) and (b) critical bond percolation cluster embedded in a 3D cubic lattice (average taken over random targets and cluster realizations).

applicability of our approach to chromatin under the assumption that it is fractal at least at a sufficiently large scale. As illustrative examples, simulations were performed on structures such as critical percolation clusters, which capture the large scale properties of nuclear DNA organization: They are (i) fractals characterized by well defined d_w and d_f, (ii) naturally embedded in Euclidean space, and (iii) disordered. These fractals can therefore be seen as minimal models of chromatin beyond the linear chain description.

Search time distribution for facilitated diffusion in a fractal medium.—Following the classical picture of facilitated diffusion, we now consider that the protein can desorb from the chromatin with rate λ_1 and then freely diffuse in the nucleoplasm before rebinding to the chromatin (see Fig. 1). In this first approach, we adopt a mean field treatment of the phases of free diffusion: We assume that the duration of such a phase is exponentially distributed with mean $\tau_2 = 1/\lambda_2$ and further suppose that the protein rebinds at a position which is uniformly distributed on the chromatin. We determine in this paragraph the mean time necessary for the protein starting at a random position on the chromatin to reach a target locus on the chromatin for the first time. We denote by $\langle \mathbf{T}_T \rangle$ this mean GFPT for facilitated diffusion and by $F_T(t)$ the GFPT probability density. By using tools developed in the context of intermittent search strategies [11,21], it can be shown that the Laplace transform $\hat{F}_T$ is expressed in terms of the distribution $\hat{\Phi}_T$ of the GFPT for simple diffusion on the chromatin, which is given by Eq. (2):

$$\hat{F}_T(s) = \hat{\Phi}_T(\lambda_1 + s)\left[1 - \frac{1 - \hat{\Phi}_T(\lambda_1 + s)}{(1 + \frac{s}{\lambda_1})(1 + \frac{s}{\lambda_2})}\right]^{-1}. \quad (4)$$

Assuming that diffusion on the chromatin is compact (as in most examples of fractals embedded in 3D space [32]) and using expression (2) for $\hat{\Phi}_T$ finally yields an explicit expression of the Laplace transformed distribution of the GFPT:

$$\hat{F}_T(s) = (s + \lambda_1)(s + \lambda_2) I_\nu(x_s) \times \left[\frac{\Gamma(1-\nu)}{4^\nu \Gamma(1+\nu)} x_s^{2\nu} s(s + \lambda_1 + \lambda_2) \times I_{-\nu}(x_s) + \lambda_1 \lambda_2 I_\nu(x_s)\right]^{-1}, \quad (5)$$

where $x_s \equiv \sqrt{A\langle \tau_T \rangle (s + \lambda_1)}$. One can check that the classical result of facilitated diffusion on a one-dimensional DNA [11,21] is recovered for $d_f = 1$ and $d_w = 2$. The mean GFPT $\langle \mathbf{T}_T \rangle$ is then readily obtained by writing

$$\langle \mathbf{T}_T \rangle = -\left(\frac{\partial \hat{F}_T}{\partial s}\right)_{s=0} = \left(\frac{1}{\lambda_1} + \frac{1}{\lambda_2}\right) \frac{1 - \hat{\Phi}_T(\lambda_1)}{\hat{\Phi}_T(\lambda_1)}, \quad (6)$$

which finally yields the central result of this Letter:

$$\langle \mathbf{T}_T \rangle = \left(\frac{1}{\lambda_1} + \frac{1}{\lambda_2}\right)\left[\frac{\Gamma(1-\nu)}{4^\nu \Gamma(1+\nu)} \frac{x_0^{2\nu} I_{-\nu}(x_0)}{I_\nu(x_0)} - 1\right], \quad (7)$$

with $x_0 \equiv x_{s=0}$. This exact expression of the mean search time for facilitated diffusion in a fractal medium, validated by numerical simulations on various examples of fractals which mimic the large scale properties of chromatin (see Fig. 3 and [31]), can be simplified as follows in the large N limit:

$$\langle \mathbf{T}_T \rangle \sim \langle \mathbf{T}_T \rangle_\infty \equiv \left(\frac{1}{\lambda_1} + \frac{1}{\lambda_2}\right) \frac{x_0^{2\nu} \Gamma(1-\nu)}{4^\nu \Gamma(1+\nu)}. \quad (8)$$

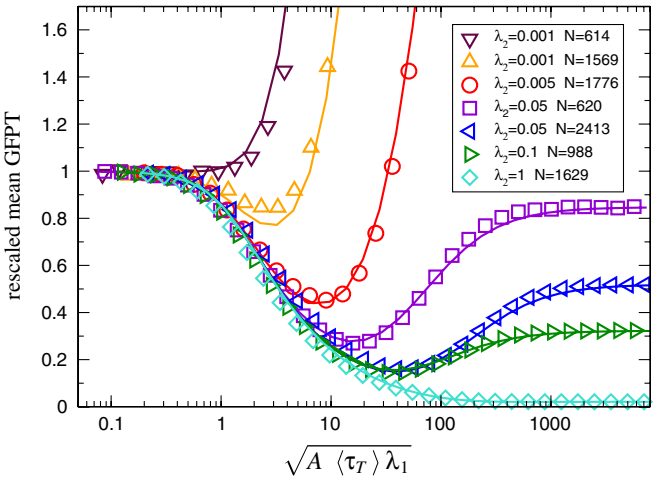

FIG. 3 (color online). Mean search time for facilitated diffusion on critical bond percolation clusters embedded in a 3D cubic lattice with a randomly located target. Numerical simulations (symbols) and theoretical prediction [Eq. (7), plain lines] of the mean GFPT, rescaled by its value for $\lambda_1 = 0$, are plotted as a function of $x_0 = \sqrt{A\langle \tau_T \rangle \lambda_1}$ for various λ_2 and system size N.

Before commenting on this result, note that a limit distribution G_T of the GFPT can be obtained by considering the rescaled time $\Theta = t/\langle \mathbf{T}_T \rangle_\infty$. Denoting $U = s\langle \mathbf{T}_T \rangle_\infty$ the Laplace variable associated to Θ, and writing $\hat{G}_T(U) = \hat{F}_T(U/\langle \mathbf{T}_T \rangle_\infty)$, we find by using (5) in the large-volume limit $\hat{G}_T(U) \sim (1+U)^{-1}$, which yields immediately the simple exponential form $G_T(\Theta) \sim \exp(-\Theta)$.

Optimal search time.—We now comment on previous results and focus on the minimization of the mean search time (7) for facilitated diffusion. First note that $\langle \mathbf{T}_T \rangle_\infty \propto N$ in the large-volume limit, which shows immediately that facilitated diffusion is faster than diffusion alone, which would yield a search time scaling as N^{d_w/d_f} [28]. Actually, the search time can be minimized as a function of the desorption rate λ_1, as soon as the adsorption rate λ_2 is large enough (see Fig. 3 and [31]). Quantitatively, the function $\langle \mathbf{T}_T \rangle$ exhibits a minimum value for some $\lambda_1 = \lambda_1^{\min}$ if the value of the derivative of $\langle \mathbf{T}_T \rangle$ with respect to λ_1 at $\lambda_1 = 0$ is negative. This sets the following condition on λ_2:

$$\lambda_2 \geq \lambda_2^{\min} = \frac{4\nu(4-\nu^2)}{\langle \tau_T \rangle (5+2\nu)(1-\nu)^2}, \tag{9}$$

which is in practice satisfied for a large enough chromatin volume. Under this condition, a direct calculation shows that the optimal value $\lambda_1^{\min}$ can be expanded in the large-volume limit as

$$\frac{\lambda_1^{\min}}{\lambda_2} \simeq \frac{1-\nu}{\nu} - \frac{\Gamma^2(\nu)\sin(\pi\nu)}{\pi} \left(\frac{2\nu^2}{\langle \tau_T \rangle \lambda_2 (1+\nu)(1-\nu)^2} \right)^{\nu}. \tag{10}$$

One recovers, in particular, for $d_f = 1$ and $d_w = 2$ the celebrated result $\lambda_1^{\min} \simeq \lambda_2$ [10,11].

Finally, these results show that facilitated diffusion is a robust mechanism which can speed up the search for a target site even in the case of eukaryotes, under the assumption that the chromatin has a fractal organization, which seems verified experimentally. We add that this approach is independent of the microscopic structure of the DNA and could also be valid to some extent in the case of prokaryotes, where the DNA, even if less densely packed than in eukaryotes, seems to have a rather compact organization significantly departing from a linear chain. Using typical experimental values $d_w \simeq 3$ and $d_f \simeq 2.5$, one finds that the search time is minimized for $\lambda_1^{\min}/\lambda_2 \simeq 1/5$. This suggests that the adsorption time should be significantly larger than the time of free diffusion to minimize the search, in contrast with the classical prediction [10,11]. This result is qualitatively compatible with experimental findings [2] on prokaryotes, even if the fractal properties of the DNA structure need to be determined in this case.

[1] O. G. Berg, R. B. Winter, and P. H. von Hippel, Biochemistry **20**, 6929 (1981).
[2] J. Elf, G.-W. Li, and X. S. Xie, Science **316**, 1191 (2007).
[3] I. Bonnet *et al.*, Nucleic Acids Res. **36**, 4118 (2008).
[4] C. Loverdo *et al.*, Phys. Rev. Lett. **102**, 188101 (2009).
[5] B. van den Broek *et al.*, Proc. Natl. Acad. Sci. U.S.A. **105**, 15 738 (2008).
[6] T. Hu, A. Y. Grosberg, and B. I. Shklovskii, Biophys. J. **90**, 2731 (2006).
[7] M. Sheinman and Y. Kafri, Phys. Biol. **6**, 016003 (2009).
[8] O. Benichou, Y. Kafri, M. Sheinman, and R. Voituriez, Phys. Rev. Lett. **103**, 138102 (2009).
[9] V. Dahirel *et al.*, Phys. Rev. Lett. **102**, 228101 (2009).
[10] M. Slutsky and L. A. Mirny, Biophys. J. **87**, 4021 (2004).
[11] M. Coppey, O. Benichou, R. Voituriez, and M. Moreau, Biophys. J. **87**, 1640 (2004).
[12] M. A. Lomholt, T. Ambjornsson, and R. Metzler, Phys. Rev. Lett. **95**, 260603 (2005).
[13] I. Eliazar, T. Koren, and J. Klafter, J. Phys. Condens. Matter **19**, 065140 (2007).
[14] G.-W. Li, O. G. Berg, and J. Elf, Nature Phys. **5**, 294 (2009).
[15] K. Bystricky *et al.*, Proc. Natl. Acad. Sci. U.S.A. **101**, 16 495 (2004).
[16] A. Grosberg, Y. Rabin, S. Havlin, and A. Neer, Europhys. Lett. **23**, 373 (1993); A. Grosberg, S. Nechaev, and E. Shakhnovich, J. Phys. (Paris) **49**, 2095 (1988).
[17] D. Lebedev *et al.*, Crystallogr. Rep. **53**, 110 (2008).
[18] A. Bancaud *et al.*, EMBO J. **28**, 3785 (2009).
[19] E. Lieberman-Aiden *et al.*, Science **326**, 289 (2009).
[20] T. Misteli, Science **291**, 843 (2001).
[21] O. Benichou, C. Loverdo, M. Moreau, and R. Voituriez, Phys. Chem. Chem. Phys. **10**, 7059 (2008).
[22] O. Bénichou *et al.*, Nature Chem. **2**, 472 (2010).
[23] J. J. Kozak and V. Balakrishnan, Phys. Rev. E **65**, 021105 (2002).
[24] C. P. Haynes and A. P. Roberts, Phys. Rev. E **78**, 041111 (2008).
[25] E. Agliari, Phys. Rev. E **77**, 011128 (2008).
[26] E. Agliari and R. Burioni, Phys. Rev. E **80**, 031125 (2009).
[27] S. Condamin *et al.*, Nature (London) **450**, 77 (2007); O. Benichou, B. Meyer, V. Tejedor, and R. Voituriez, Phys. Rev. Lett. **101**, 130601 (2008); O. Benichou and R. Voituriez, Phys. Rev. Lett. **100**, 168105 (2008).
[28] V. Tejedor, O. Bénichou, and R. Voituriez, Phys. Rev. E **80**, 065104 (2009).
[29] S. Redner, *A Guide to First- Passage Processes* (Cambridge University Press, Cambridge, England, 2001).
[30] B. O'Shaughnessy and I. Procaccia, Phys. Rev. Lett. **54**, 455 (1985).
[31] See supplemental material at http://link.aps.org/supplemental/10.1103/PhysRevLett.106.038102 for other examples.
[32] D. Ben-Avraham and S. Havlin, *Diffusion and Reactions in Fractals and Disordered Systems* (Cambridge University Press, Cambridge, England, 2000).