P H Y S I C A L   R E V I E W   L E T T E R S

Quantifying Hopping and Jumping in Facilitated Diffusion of DNA-Binding Proteins

C. Loverdo, O. Bénichou, and R. Voituriez

*Laboratoire de Physique Théorique de la Matière Condensée (UMR 7600), Université Pierre et Marie Curie,
4 Place Jussieu, 75255 Paris Cedex France*

A. Biebricher, I. Bonnet, and P. Desbiolles

Laboratoire Kastler Brossel, ENS, UPMC-Paris6, CNRS UMR 8552, 24 rue Lhomond, F-75005 Paris France

Facilitated diffusion of DNA-binding proteins is known to speed up target site location by combining three dimensional excursions and linear diffusion along the DNA. Here we explicitly calculate the distribution of the relocation lengths of such 3D excursions, and we quantify the short-range correlated excursions, also called hops, and the long-range uncorrelated jumps. Our results substantiate recent single-molecule experiments that reported sliding and 3D excursions of the restriction enzyme EcoRV on elongated DNA molecules. We extend our analysis to the case of anomalous 3D diffusion, likely to occur in a crowded cellular medium.

PACS numbers: 87.10.−e, 05.40.Fb, 05.40.Jc

The kinetics of gene regulation plays a key role for various cell functions such as the response to external signals. A limiting step of this complex process is the location of specific DNA sequences by DNA-binding proteins. This target search process is widely accepted to be speeded up by facilitated diffusion, which combines a linear (1D) diffusion of the protein along the DNA called sliding and three dimensional (3D) excursions named hopping or jumping that can relocate the protein to an unexplored region of the DNA [1–3]. Single-molecule experiments have recently evidenced sliding for a variety of proteins (for a review, see [4]) and, more lately, hopping or jumping have been reported for the processivity factor UL42 [5] and for the restriction enzyme EcoRV [6,7]. The distinction between hopping and jumping, which both stem from free 3D diffusion of the protein after dissociation from DNA, is usually based on the location of the protein reassociation site on the DNA after a 3D excursion. The reassociation site is either close, and thus statistically correlated, to the dissociation site (hopping) or far from it (jumping) [2,5,8]. Following this definition, we hereafter call hops 3D correlated excursions whose span (end-to-end length of the 3D excursion) is shorter than the DNA persistence length ξ (see Fig. 1 left side) and we call jumps the other 3D excursions, leading to random relocations over the DNA, that have been studied extensively [2,7,9–14].

So far an analytical description of hops is still missing, despite its importance for quantifying the efficiency of facilitated diffusion. In the context of colocalization [8,15,16] for instance, it has been observed that in prokaryotes, the target sites of transcription factors can be very close to their coding sequence, making target location by sliding and hopping much more favorable than by jumping. Hops could also play a major role when considering the translocation of a site-specific protein along DNA coated by other DNA-binding proteins that hinder a linear diffusion over large distances.

Another aspect, which has not been considered so far, is the modelling of hopping and jumping in a crowded cellular environment. Crowding can strongly affect the transport properties of tracer molecules in cells, either in the cytoplasm or in the nucleus, causing a dynamics often reported to depart from usual diffusion [17,18]. Though anomalous diffusion is known to drastically modify the kinetics of transport limited reactions [19–21], its impact on facilitated diffusion remains unexplored.

In this Letter, we give the analytical distribution of the length of 3D relocations on an elongated DNA molecule. This result has important merits: (i) It not only provides the distribution of hop lengths, but also gives quantitative access to the relative importance of hops and jumps in the biologically relevant case of a randomly coiled DNA. (ii) It accounts for recent *in vitro* single-molecule experiments [6], and thus confirms the observation of EcoRV 3D relocations on elongated DNA molecules. (iii) It can be generalized to the case of anomalous transport due to crowding relevant to *in vivo* situations.
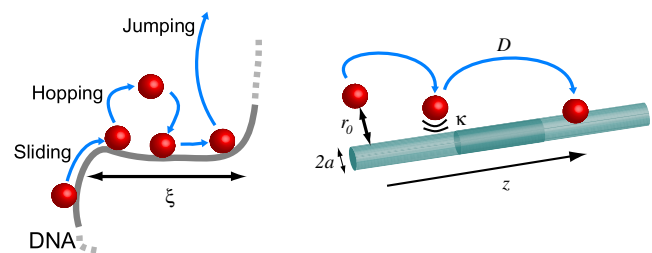

FIG. 1 (color online). Facilitated diffusion of a protein on DNA. *Left:* schematic definition of sliding, hopping, and jumping. *Right:* model parameters.

188101-1

Normal diffusion.—We first introduce the model in the case of diffusive transport. We consider a pointlike protein diffusing in 3D with a diffusion coefficient D, and model the DNA molecule by an infinite cylinder of radius a (see Fig. 1 right side). The density probability that the protein is at time t at $\mathbf{r}$ knowing that it was at $\mathbf{r_0}$ at time 0 satisfies the backward diffusion equation [22]:

$$\partial_t P(\mathbf{r}, t|\mathbf{r_0}) = D\Delta_{\mathbf{r_0}} P(\mathbf{r}, t|\mathbf{r_0}). \tag{1}$$

We use cylindrical coordinates and assume standard radiative boundary conditions,

$$\partial_{r_0} P(\mathbf{r}, t|\mathbf{r_0})(r_0 = a) = \kappa P(\mathbf{r}, t|\mathbf{r_0})(r_0 = a) \tag{2}$$

to describe the adsorption of the protein on DNA. The parameter κ has the dimension of the inverse of a length, and is proportional to the adsorption rate [23]. We first determine the equation satisfied by the density probability $\pi(z|\mathbf{r_0})$ of being adsorbed on the DNA at the longitudinal absciss z, starting from the point $\mathbf{r_0}$. Using the relation between the first-return time probability density at the longitudinal absciss z on the DNA $F(z, t|\mathbf{r_0})$ and the probability current [23] one can show that

$$\pi(z|\mathbf{r_0}) = 2\pi a D \int_0^\infty \partial_r P(a, z, t|\mathbf{r_0})dt. \tag{3}$$

Using Eqs. (1) and (2), it is then easy to obtain that

$$\Delta_{\mathbf{r_0}} \pi(z|\mathbf{r_0}) = 0$$
$$\partial_{r_0} \pi(z|\mathbf{r_0})(r_0 = a) = \kappa \pi(z|\mathbf{r_0})(r_0 = a). \tag{4}$$

After Fourier transform this equation yields finally

$$\pi(z|\mathbf{r_0}) = \frac{1}{\pi} \int_0^\infty \cos(kz) \frac{K_0(kr_0)}{K_0(ka) + K_1(ka)k/\kappa} dk, \tag{5}$$

where K_i (and later J_i) denote Bessel functions. Several comments are in order. (i) First, the behavior of the relocation length distribution at large z is given by $\pi(z|\mathbf{r_0}) \sim [\ln(r_0/a) + (\kappa a)^{-1}]/[2z\ln^2(z/a)]$. This wide tail distribution shows the importance of long-range relocations, reminiscent of the distribution of return times to the DNA [23,24]. More precisely, it can be shown that, starting from the DNA ($r_0 = a$), the median of the distribution (5) is given by $a\exp(1/\kappa a)$ (assuming $\kappa^{-1} > a$), which gives an interpretation of the parameter κ alternative to the previous definition (2). (ii) Most importantly, these results are relevant to *in vivo* situations when the DNA has a random coil conformation, as for $z < \xi$ (where ξ is the DNA persistence length) the DNA can be modeled as a cylinder. Thus, Eq. (5) gives directly the distribution of hops, which could be accessed only numerically so far. On the other hand, since we consider as jumps the relocations longer than ξ, the probability of performing a jump rather than a hop is given by the complementary cumulative distribution:

$$C(z = \xi) = \int_{|z|>\xi} \pi(z|\mathbf{r_0})dz \sim \frac{\ln(r_0/a) + 1/\kappa a}{\ln(\xi/a)}. \tag{6}$$

Hence, as soon as the protein departs the DNA further than its diameter, it has a significant probability of performing a jump. In the context of target localization, Eq. (6) shows that in situations where jumping is favorable, reducing the persistence length makes the search process faster, as was found in [7] where the persistence length could be effectively tuned by stretching the DNA.

Application to single-molecule experiments.—Using our approach, we now give a quantitative interpretation of single-molecule experiments. In [6] we have reported the direct observation by total internal reflection fluorescence microscopy of the interaction of the restriction enzyme EcoRV with elongated DNA molecules. At low salt concentrations [NaCl], we observed a linear diffusion (sliding) of the enzyme along the DNA with a 1D diffusion constant $D_1 \simeq 10^{-2} \, \mu\text{m}^2 \cdot \text{s}^{-1}$. We also observed fast and long-range relocations ($> z_m = 200$ nm) within an observation time $t_{\text{obs}} = 40$ ms, in a proportion significantly larger than expected by sliding. They were interpreted as 3D excursions accomplished with a 3D diffusion constant $D_3 \simeq 50 \, \mu\text{m}^2 \cdot \text{s}^{-1}$, i.e., much larger than D_1. We determined the complementary cumulative distribution of the lengths of these large relocations.

Though our theoretical approach provides the distribution of 3D relocations for an infinitely long DNA, it can be adapted to the experiments. First, to account for the finite observation time t_{obs}, we use the following alternative writing of the relocation length distribution:

$$\pi(z) = \int_0^{t_{\text{obs}}} P_\parallel(z, t) F_\perp(t|\mathbf{a})dt, \tag{7}$$

where $P_\parallel(z, t) = (4\pi Dt)^{-1/2} \exp(-z^2/4Dt)$ is the 1D propagator along the DNA, and $F_\perp(t|\mathbf{a})$ is the first-passage density to DNA in the orthogonal plane, yielded by numerical inversion of its Laplace transform [23]:

$$\hat{F}_\perp(s|\mathbf{a}) = \frac{K_0(x)}{K_0(x) + (\kappa a)^{-1} K_1(x)}, \tag{8}$$

with $x = a\sqrt{s/D}$. Second, we take into account the finite length of the DNA. In the experiments, the DNA was elongated and the DNA ends, separated by $L \simeq 2.2 \, \mu\text{m}$, were chemically bound to a surface. Because of its low nonspecific interactions with the proteins this surface can be considered as reflective. For analytical purposes we consider an effective geometry and model the DNA by a finite cylinder of radius a (given by the sum of the DNA and enzyme radii) and length L standing between two reflective planes perpendicular to the cylinder axis. This approximation is expected to be valid in the regime that we consider here where t_{obs} is small so that multiple reflections are very unlikely. The effect of the reflective planes is dealt with using the method of images. Assuming a homogeneous distribution of the starting point, we obtain through

standard calculations the probability density of observing a relocation of length z within an observation time t_{obs}: $\pi_r(z) = \frac{2}{L} \sum_{n=0}^{\infty} Y_n$, with $Y_0 = \int_z^{\infty} \pi(x)dx + (L - z)\pi(z)$, and $Y_{n>0} = (L - z)[\pi(2nL + z) + \pi(2nL - z)] - \int_{2nL-z}^{2nL+z} \pi(x)dx$.

Experimentally, the cumulative distribution of the relocation lengths z was found to weakly depend on [NaCl]. This is consistent with biochemical studies that reported only a slight dependence on [NaCl] of the binding rate to DNA [25]. Moreover, after normalization by the total number of observed relocations, the data collapse (Fig. 2), demonstrating that [NaCl] does not impact on the shape of the distribution. These normalized cumulative distributions can be compared to the theoretical prediction $C_r(z)/C_r(z_m)$, where $C_r(z) = \int_z^{\infty} \pi_r(z')dz'$. As shown in Fig. 2, the z dependence of the experimental distributions is in good agreement with our model, which strongly supports that the observed large relocations are due to 3D excursions [6]. Since this agreement is obtained for a wide range of values of κ^{-1} (0–40 nm), a quantitative estimation of this parameter cannot be obtained using these normalized distributions.

Subdiffusion.—We now turn to the case where crowding effects in the cellular medium lead to subdiffusion when the protein dissociates from the DNA. Such behavior has been reported in many biological situations [17,18], but an analysis of its quantitative impact on target search is still missing. As a first step in this direction, we calculate the relocation length distribution on DNA.

The subdiffusive behavior is usually characterized by a mean square displacement that scales as [26] $\langle \Delta \mathbf{r}^2 \rangle \sim t^{2/d_w}$, where $d_w > 2$ defines the walk dimension. Such a scaling law can be obtained from a few microscopic models. Here we focus on two possibilities: (i) continuous time
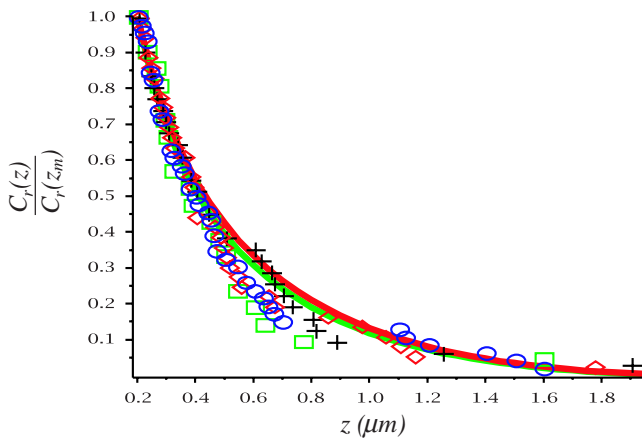

FIG. 2 (color online). Cumulative distribution of hop length z (μm) normalized by the number of hops larger than $z_m = 200$ nm. Data from [6] (in PIPES buffer): black crosses = 10 mM NaCl, red diamonds = 20 mM NaCl, green boxes = 40 mM NaCl, blue circles = 60 mM NaCl. Solid lines: distribution from our model for $\kappa^{-1} = 0.5$ nm (lower green lines), $\kappa^{-1} = 20$ nm (upper red lines).

random-walks models [26,27], in which at each step the walker can land on a trap (such as the free cage around the tracer in a crowded environment) where it can stay for extended periods of time, and (ii) models based on spatial inhomogeneities, as diffusion in random fractals such as critical percolation clusters [28], in which the anomalous behavior is due to fixed obstacles that create numerous dead ends. The relative role of these two scenarios in the subdiffusion observed in cells is still debated [21].

While these two classes of models lead to similar scaling laws for the mean square displacements, their microscopic origins are intrinsically different [21] and thus lead to notable differences in the relocation length distributions. In the case of subdiffusion stemming from continuous time random-walks, the geometry of the trajectories is not affected by crowding, and only the dynamics deviates from normal diffusion. The relocation length distribution, which depends only on the geometry of trajectories, is therefore similar to that obtained above with regular diffusion. As a consequence, Eq. (5) and its implications are still valid in this case. In contrast, in the case of a fractal type subdiffusion, the geometry of trajectories is deeply affected by crowding. We consider here that after dissociation from DNA the protein evolves on a fractal embedded in the 3D space, and keep the notations of Fig. 1. Assuming that the longitudinal and transverse projections of the transport process are independent, we use Eq. (7) to write the relocation length distribution as $\pi(z|\mathbf{r_0}) = \int_0^{\infty} P_{\parallel}(z, t) F_{\perp}(t|\mathbf{r_0})dt$, where $P_{\parallel}$ stands for the longitudinal propagator, characterized by the walk dimension d_w and the fractal dimension $d_f^{\parallel}$ of the projected motion, and $F_{\perp}$ is the transverse first-passage-time density to DNA, characterized by dimensions d_w and $d_f^{\perp}$.

To go further, we follow [29] and we assume that the diffusion current obeys the generalized Ficks's law, which gives an effective transport operator for any quantity c:

$$K\Delta_r c(r) \equiv \frac{K}{r^{d_f-1}} \frac{d}{dr}\left(r^{d_f-d_w+1} \frac{d}{dr} c(r)\right), \quad (9)$$

where K is the generalized diffusion coefficient. Note that for subdiffusion $d_w > 2 \geq d_f^{\perp}$ which means that the exploration is compact and we can take the limit $a \to 0$. After some algebra the first-passage-time density $F_{\perp}$ reads:

$$F_{\perp}(t|\mathbf{r_0}) = \frac{d_w^{-\nu_{\perp}} K r_0^{\nu_{\perp} d_w/2}}{\Gamma(\nu_{\perp})} \int_0^{\infty} e^{-Kut} u^{\nu_{\perp}/2} J_{\nu_{\perp}}(\beta\sqrt{u})du, \quad (10)$$

where $\beta = 2r_0^{d_w/2}/d_w$ and $\nu_i = 1 - d_f^i/d_w$ (here $i = \perp, \parallel$). Using the propagator found in [29], we finally obtain the following explicit result:

$$\pi(z|\mathbf{r_0}) = \alpha \int_0^{\infty} u^{(\nu_{\perp}-\nu_{\parallel})/2} J_{\nu_{\perp}}(\beta\sqrt{u}) K_{\nu_{\parallel}}(\gamma\sqrt{u})du, \quad (11)$$
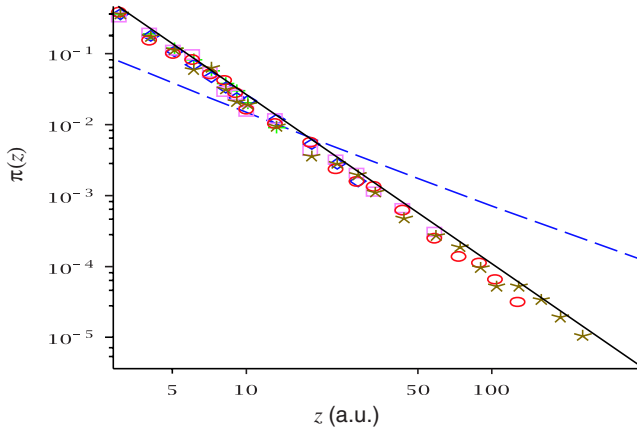
with $\gamma = 2z^{d_w/2}/d_w$ and

FIG. 3 (color online). Distribution of the relocation length $\pi(z|\mathbf{r_0})$ for 3D excursions on a critical percolation cluster embedded in a 3D cubic lattice, for $r_0 = 1$ and $a = 0$. Simulations (normalized for $z \geq 3$) are performed for different system sizes to rule out finite size effects (symbols), and collapse on the theoretical curve (plain line) obtained from Eq. (11). Here $d_w = 3.88\ldots$, $d_f^\perp = 2$, and $d_f^\| = 1$ since $d_f = 2.53\ldots$ [28]. The large z scaling follows $\pi(z|\mathbf{r_0}) \sim 1/z^{2.88\ldots}$, and is compared to the case of normal diffusion (dashed line) with $r_0 = 1$ and $a = 0.1$.

$$\alpha = \frac{2 d_w^{\nu_\| - \nu_\perp - 1} r_0^{\nu_\perp d_w/2} z^{(d_f^\| + d_w)/2 - 1}}{\Gamma(\nu_\perp)\Gamma(1 - \nu_\|)}. \qquad (12)$$

The analytical expression (11) of the relocation length distribution for subdiffusion in fractals is validated by numerical simulations in the representative example of a percolation cluster embedded in a 3D space (see Fig. 3), which justifies our decoupling assumption of longitudinal and transverse projections of motion. The large z behavior can be shown from (11) to be given by $\pi(z|\mathbf{r_0}) \sim r_0^{d_w - d_f^\perp}/z^{1 + d_w - d_f^\perp}$, which decays faster than the diffusion case [$\sim 1/z \ln^2(z/a)$] as illustrated in Fig. 3. Accordingly, the proportion of jumps in the case of a random conformation of DNA now scales like $C(\xi) \sim \xi^{-d_w + d_f^\perp}$ and is therefore significantly decreased in this case of fractal type crowding as compared to the case of regular diffusion. Note that this predicted low proportion of jumps is compatible with recent *in vivo* observations [30], and supports the idea that crowding effects could be important *in vivo*. In the context of colocalization, our results further suggest that crowding could be beneficial for target location since it enhances the local scanning of the DNA by reducing the proportion of jumps.

To conclude, we developed an analytical model that gives the distribution of 3D relocation lengths of a protein dissociating from a DNA molecule. We could quantify the relative proportion of hops and jumps in facilitated diffusion, and we confirmed the direct observation of hops in single-molecule experiments involving the EcoRV restriction enzyme. We extended our approach to the case of anomalous transport and showed quantitatively that crowd-

ing effects can favor hopping and significantly reduce the proportion of jumps. Overall, our study supports the hypothesis that 3D excursions play a crucial role in protein/ DNA interactions.

[1] A. D. Riggs, S. Bourgeois, and M. Cohn, J. Mol. Biol. **53,** 401 (1970).
[2] O. G. Berg, R. B. Winter, and P. H. Von Hippel, Biochemistry **20,** 6929 (1981).
[3] S. E. Halford and J. F. Marko, Nucleic Acids Res. **32,** 3040 (2004).
[4] J. Gorman and E. C. Greene, Nat. Struct. Mol. Biol. **15,** 768 (2008).
[5] G. Komazin-Meredith *et al.*, Proc. Natl. Acad. Sci. U.S.A. **105,** 10 721 (2008).
[6] I. Bonnet *et al.*, Nucleic Acids Res. **36,** 4118 (2008).
[7] B. van den Broek *et al.*, Proc. Natl. Acad. Sci. U.S.A. **105,** 15 738 (2008).
[8] G. Kolesov *et al.*, Proc. Natl. Acad. Sci. U.S.A. **104,** 13 948 (2007).
[9] M. Coppey, O. Bénichou, R. Voituriez, and M. Moreau, Biophys. J. **87,** 1640 (2004).
[10] M. Slutsky and L. A. Mirny, Biophys. J. **87,** 4021 (2004).
[11] M. A. Lomholt, T. Ambjornsson, and R. Metzler, Phys. Rev. Lett. **95,** 260603 (2005).
[12] T. Hu, A. Grosberg, and B. I. Shklovskii, Biophys. J. **90,** 2731 (2006).
[13] I. Eliazar, T. Koren, and J. Klafter, J. Phys. Condens. Matter **19,** 065140 (2007).
[14] M. Sheinman and Y. Kafri, Phys. Biol. **6,** 016003 (2009).
[15] Z. Wunderlich and L. A. Mirny, Nucleic Acids Res. **36,** 3570 (2008).
[16] O. Bénichou, C. Loverdo, and R. Voituriez, Europhys. Lett. **84,** 38 003 (2008).
[17] I. M. Tolic-Norrelykke *et al.*, Phys. Rev. Lett. **93,** 078102 (2004).
[18] E. Golding and E. Cox, Phys. Rev. Lett. **96,** 098102 (2006).
[19] M. A. Lomholt, I. M. Zaid, and R. Metzler, Phys. Rev. Lett. **98,** 200603 (2007).
[20] S. Condamin *et al.*, Nature (London) **450,** 77 (2007).
[21] S. Condamin *et al.*, Proc. Natl. Acad. Sci. U.S.A. **105,** 5675 (2008).
[22] B. Hughes, *Random Walks and Random Environments* (Oxford University Press, New York, 1995).
[23] S. Redner, *A Guide to First Passage Time Processes* (Cambridge University Press, Cambridge, England, 2001).
[24] P. Levitz *et al.*, Phys. Rev. E **78,** 030102 (2008).
[25] T. M. Lohman, CRC Crit. Rev. Biochem. **19,** 191 (1986).
[26] R. Metzler and J. Klafter, Phys. Rep. **339,** 1 (2000).
[27] J. Klafter, A. Blumen, and M. F. Shlesinger, Phys. Rev. A **35,** 3081 (1987).
[28] D. Ben-Avraham and S. Havlin, *Diffusion and Reactions in Fractals and Disordered Systems* (Cambridge University Press, Cambridge, England, 2000).
[29] B. O'Shaughnessy and I. Procaccia, Phys. Rev. Lett. **54,** 455 (1985).
[30] J. Elf, G.-W. Li, and X. S. Xie, Science **316,** 1191 (2007).