Journal of Statistical Physics, Vol. 127, No. 3, May 2007 (© 2007)
DOI: 10.1007/s10955-007-9282-4

Modeling DNA and Virus Trafficking in the Cell Cytoplasm[1]

D. Holcman[2,3]

Received September 1, 2006; accepted January 8, 2007
Published Online: February 13, 2007

Cytosol trafficking is a limiting step of viral infection or DNA delivery. Starting from the cell surface, most viruses have to travel through a crowded and risky environment in order to reach a small nuclear pore. This work is dedicated to estimating the probability p_N of a viral arrival success and, in that case, the mean time τ_N it takes. Viral movement is described by a stochastic equation, containing both a drift and a Brownian component. The drift part represents the movement along microtubules, while the Brownian component corresponds to the free diffusion. The success of a viral infection is limited by a killing activity occurring inside the cytoplasm. We model the killing activity by a steady state killing rate k. Because nuclear pores occupy a small fraction of the nuclear area, we use this property to obtain asymptotic estimates of p_N and τ_N as a function of the diffusion constant D, the amplitude of the drift B and the killing rate k.

KEY WORDS: mean first passage time, modeling DNA and virus trafficking, asymptotic analysis, stochastic process, small hole

1. INTRODUCTION

Intracellular trafficking of genome is a key step in viral and non-viral mediated gene transfer.[4,24] After entering into the cell, by endocytosis for example, a viral DNA (or a plasmid DNA) has to escape from endosomes and travel through a highly crowded media before reaching the nucleus.[7] In this study, we are interested for both, viruses and plasmids, in computing the efficiency of the delivery process. We propose to obtain asymptotic computations of the probability p_N they reach

[1] This paper is dedicated to my wife Nathalie Rouach.

[2] Département de Biologie, Ecole Normale Supérieure, 46 rue d'Ulm 75005 Paris, France. D. H. is supported by the program "Chaire d'Excellence."

[3] Department of Mathematics, Weizmann Institute of Science, Rehovot 76100 Israel. D.H is incumbent to the Haas Russell Chair; e-mail: holcman@biologie.ens.fr.

one of the nuclear pore (which occupy only a small fraction of the nucleus surface) and the mean time τ^N it takes. We limit this study to the steps involved in cytosol trafficking i.e. the motion starting after endosome release and ending at the level of the nuclear pore.

I will make the distinction between naked DNA (plasmid DNA) and viral DNA (spherical particle) motion. DNA movement inside a cell depends on its shape and on its size. For small DNAs (<250 base pairs), mobility can be described as a pure Brownian motion,[1] while larger DNAs (>250 base pairs) movement has been described by a different theory.[2] The density of actin filaments and microtubules highly restrict the DNA diffusion, as demonstrated experimentally.[1] In contrast, viruses movement can be seen as a succession of jumps between a pure Brownian behavior and a deterministic transport along microtubules.[5] Active transport of a virus generally involves motor proteins such as Kinesin (to travel in the direction of the cell membrane) or Dynein (to travel toward the nucleus).[22] Once a virus is attached to a Dynein protein, its movement can be modeled as a determinist drift toward the nucleus. In the crowded cell environment, due to DNA trapping into the actin meshwork, modeling a naked DNA fragment (>250 bp) as a material point is somehow simplistic. However, we will see that interesting results can still be obtained under this assumption. We will consider that a permanent immobile DNA, trapped in the network, or a DNA degraded by cytosolic nucleases[12] are killed. In our computations, this killing process will be modeled as a space dependent killing rate $k(\boldsymbol{x})$. For viruses, in the notion of killing rate, I include various phenomena such as cytoplasmic irreversible immobilization, outward movement of viral particle and/or proteasome degradation. Thus, the probability to reach the nucleus depends on all these degradation processes.

In the first part of the paper, to account for the geometrical features of the cell, we present the Fokker–Planck equation satisfied by the survival probability density function $p_t(\boldsymbol{x})$ to find a particle (either a virus or naked DNA) alive inside a volume element $\boldsymbol{x} + d\boldsymbol{x}$ at time t. We obtain various expressions for the probability p_N that a particle reaches a small pore and for the conditioned Mean First Passage Time (MFPT) τ^N. These quantities are related to some solutions of elliptic partial differential equations (PDEs), derived from the probability $p_t(\boldsymbol{x})$. In the second part of the paper, we present an asymptotic analysis of these PDEs, under the following assumptions: 1) the area occupied by the nuclear pores is small compared to the size of the nucleus and the cell and 2) the killing rate is small compared to the diffusion scale. Under these assumptions, we obtain explicit asymptotic expressions of p_N and τ^N as a function of the drift amplitude, the killing rate k, the diffusion constant D and the total number and size of nuclear pores. The estimates are obtained when the cell geometry is well approximated as a limit domain in dimension two and when it is a regular domain in dimension three. To test the asymptotic formula against biological datas, we present some numerical results for the cases of plasmid DNA (zero drift case) and for the case

of a virus, which accounts also for a drift. Finally, we shall discuss the biological implications of these asymptotic formulas.

2. MODELING THE TRAJECTORY OF A GENERIC PARTICLE IN THE CYTOPLASM

The cytoplasm of biological cells defines a geometrical domain which contains a nucleus N, occupying a small portion of the cell volume. We denote by Ω the volume occupied by the cytoplasm. We approximate N as a ball of radius δ, with $\delta \ll |\Omega|^{1/3}$. We assume that the geometry of the domain Ω is such that the isoperimetric ratio is $\frac{|\partial\Omega|}{|\Omega|^{2/3}} = O(1)$ and that there is no bottle neck of size comparable to the size of the small pore located on the nucleus. This assumption is satisfied in most of the cell, except for example neurons.

We model a naked DNA or a virus as a material point and thus neglect the effect of its finite size. As a consequence possible trapping of the DNA by actin filaments or microtubules is neglected. In this paper, we approximate the intermittent dynamics between free diffusion and drift motion along microtubule by an effective movement which accounts for both components at the same time. In that case, the effective drift represents the statistical mean drift and its amplitude is denoted by a constant B, it is radial and attractive. The effective drift depends on the events of attaching and detaching to the microtubule filaments and many other parameters such as the density of microtubules, the binding and unbinding rate (see Ref. 11 unpublished data) for further analysis. Thus, if we denote by $X(t)$ the position of a virus at time t, then it satisfies the standard stochastic equation

$$\frac{dX}{dt} = b(X) + \sqrt{2D}\dot{w} \tag{1}$$

where D is the diffusion constant and the drift b is given by

$$b(\boldsymbol{x}) = -B\frac{\boldsymbol{r}}{|\boldsymbol{r}|}, \tag{2}$$

where $\boldsymbol{r}$ is the radial vector $\boldsymbol{x}$, the origin of which is taken inside the nucleus. A schematic viral trajectory and its mathematical idealization inside a cell are represented in Fig. 1.

2.1. Steady State Equations

We will first recall the expressions for the probability p_N and the mean time τ^N. We denote the external boundary of the cell by $\partial\Omega_{\text{ext}}$ while the boundary of the nucleus is divided into an absorbing part ∂N_a and a reflecting part ∂N_r. ∂N_a is the union of well separated small pores. The total boundary is $\partial\Omega = \partial\Omega_{\text{ext}} \cup \partial N_a \cup \partial N_r$.

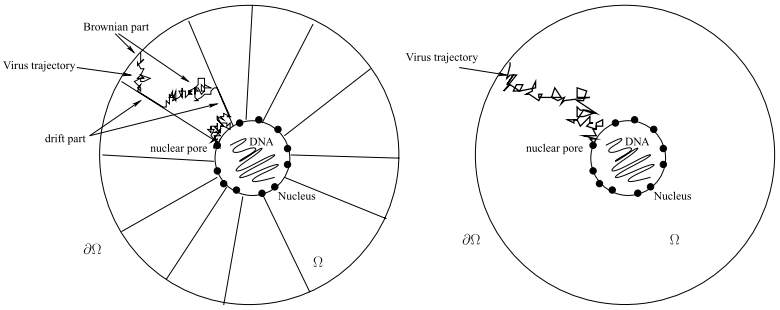

Fig. 1. Virus trafficking inside the cytoplasm of a biological cell. On the left figure, microtubules are represented originating from the cell surface and ending to the nucleus. The small nuclear pores are represented by black dots distributed uniformly on the nucleus surface. A viral trajectory is represented by an intermittent movement: bound to microtubules and thus moving with a constant drift in direction of the nucleus or moving by Brownian motion inside the cell cytoplasm. On the right figure, the virus movement has been approximated by a standard stochastic equation with a constant radial drift.

Let us consider the probability $P_N(\boldsymbol{x})$ that a DNA particle reaches the nucleus alive and conditioned on the initial position $\boldsymbol{x}$. The cell geometry is modeled as a domain Ω and the nucleus is approximated by a small ball $B(a)$ of radius a. We assume that all trajectories of the vector field b originating from the cell boundary $\partial\Omega_{\text{ext}}$ reach the boundary of $B(a)$ in a finite time (no recurrent sets of b are contained in $\Omega - B(a)$).

The survival probability density function (SPDF) is given by

$$p(\boldsymbol{x}, t|\boldsymbol{y})d\boldsymbol{x} = Pr\{X(t) \in \boldsymbol{x} + d\boldsymbol{x}, \tau^k > t, \tau^a > t|X(0) = \boldsymbol{y}\}, \tag{3}$$

where τ^a is the first time that the virus reaches the absorbing boundary ∂N_a alive and τ^k is the first time that it is hydrolyzed. The SPDF $p(\boldsymbol{x}, t|\boldsymbol{y})$ satisfies the forward Fokker–Planck equation[16]

$$\frac{\partial p}{\partial t}(\boldsymbol{x}, 0|\boldsymbol{y}) = D\Delta p(\boldsymbol{x}, 0|\boldsymbol{y}) - \nabla(p(\boldsymbol{x}, 0|\boldsymbol{y})b(\boldsymbol{x})) - k(\boldsymbol{x})p(\boldsymbol{x}, 0|\boldsymbol{y}) \text{ in } \Omega \tag{4}$$

with the initial condition

$$p(\boldsymbol{x}, 0|\boldsymbol{y}) = \delta(\boldsymbol{x} - \boldsymbol{y}) \quad \text{for} \quad \boldsymbol{x}, \boldsymbol{y} \in \Omega \tag{5}$$

and the boundary conditions

$$p(\boldsymbol{x}, t|\boldsymbol{y}) = 0 \quad \text{on} \quad \partial N_a \tag{6}$$

$$\boldsymbol{J}(\boldsymbol{x}, t \mid \boldsymbol{y}).\boldsymbol{n}_{\boldsymbol{x}} = 0 \quad \text{on} \quad \partial N_r \cup \partial\Omega_{\text{ext}}, \tag{7}$$

$\boldsymbol{n}_x$ denotes the normal derivative at a boundary point $\boldsymbol{x}$. The flux density vector $\boldsymbol{J}(\boldsymbol{x}, t \mid \boldsymbol{y})$ is defined as

$$J^i(\boldsymbol{x}, t \mid \boldsymbol{y}) = -D\nabla^i p(\boldsymbol{x}, t \mid \boldsymbol{y}) + b^i(\boldsymbol{x}) p(\boldsymbol{x}, t \mid \boldsymbol{y}). \tag{8}$$

To express the probability $P_N(\boldsymbol{y})$ that a virus arrives to the nucleus N before being killed, conditioned on the initial location $\boldsymbol{y}$, we follow the result of Ref. 9 and we get

$$P_N(\boldsymbol{y}) = Pr\{\tau^a < \tau^k | X(0) = \boldsymbol{y}\}. \tag{9}$$

Similarly, the probability of being killed before arriving at the absorbing part of the nucleus is defined as $Pr\{\tau^k < \tau^a | X(0) = \boldsymbol{y}\}$. These probabilities can be expressed in terms of SPDF by integrating the Fokker–Planck eq. (4) from 0 to infinity first and then over the domain Ω.[9] We get

$$1 = \int_0^\infty \oint_{\partial\Omega} \boldsymbol{J}(\boldsymbol{x}, t \mid \boldsymbol{y}) \cdot \boldsymbol{n}(\boldsymbol{x})\, dS_{\boldsymbol{x}}\, dt + \int_0^\infty \int_\Omega k(\boldsymbol{x}) p(\boldsymbol{x}, t \mid \boldsymbol{y})\, d\boldsymbol{x}\, dt. \tag{10}$$

Thus identifying the probabilities, we obtain that

$$P_N(\boldsymbol{y}) = Pr\{\tau^a < \tau^k | X(0) = \boldsymbol{y}\} = \int_0^\infty \oint_{\partial\Omega} \boldsymbol{J}(\boldsymbol{x}, t \mid \boldsymbol{y}) \cdot \boldsymbol{n}(\boldsymbol{x})\, dS_{\boldsymbol{x}}\, dt \tag{11}$$

and

$$\begin{aligned} Pr\{\tau^k < \tau^a | X(0) = \boldsymbol{y}\} &= \int_0^\infty \int_\Omega k(\boldsymbol{x}) p(\boldsymbol{x}, t \mid \boldsymbol{y})\, d\boldsymbol{x}\, dt. \\ &= \int_\Omega k(\boldsymbol{x}) \tilde{p}(\boldsymbol{x} \mid \boldsymbol{y})\, d\boldsymbol{x}\,, \end{aligned} \tag{12}$$

where

$$\tilde{p}(\boldsymbol{x} \mid \boldsymbol{y}) = \int_0^\infty p(\boldsymbol{x}, t \mid \boldsymbol{y})\, dt \tag{13}$$

is solution of

$$D\Delta\tilde{p} - \nabla(b(\boldsymbol{x})\tilde{p}) - k(\boldsymbol{x})\tilde{p} = -\delta(\boldsymbol{x} - \boldsymbol{y}) \quad \text{for} \quad \boldsymbol{x}, \boldsymbol{y} \in \Omega \tag{14}$$

and the boundary conditions are given by (6). When the initial distribution is given by a smooth function p_i, we can define the averaging probability

$$\tilde{p}(\boldsymbol{x}) = \int_\Omega \tilde{p}(\boldsymbol{x} \mid \boldsymbol{y}) p_i(\boldsymbol{y}) d\boldsymbol{y}, \tag{15}$$

solution of the equation

$$D\Delta\tilde{p}(\boldsymbol{x}) - \nabla(b(\boldsymbol{x})\tilde{p}(\boldsymbol{x})) - k(\boldsymbol{x})\tilde{p}(\boldsymbol{x}) = -p_i(\boldsymbol{x}) \quad \text{for} \quad \boldsymbol{x} \in \Omega. \tag{16}$$

We define the time dependent averaged probability by

$$\tilde{p}(\boldsymbol{x}, t) = \int_{\Omega} \tilde{p}(\boldsymbol{x}, t| \boldsymbol{y}) p_i(\boldsymbol{y}) d\boldsymbol{y}. \tag{17}$$

and the associated flux

$$J^i(\boldsymbol{x}, t \mid p_i) = -D\nabla^i \tilde{p}(\boldsymbol{x}, t) + b^i(\boldsymbol{x})\tilde{p}(\boldsymbol{x}, t). \tag{18}$$

Finally, we define the probability p_N to reach the nucleus as the reaching probability averaged over the initial position, by the formula

$$\begin{aligned} P_N &= \int_{\Omega} Pr\{\tau^a < \tau^k | X(0) = \boldsymbol{y}\} p_i(\boldsymbol{y}) d\boldsymbol{y} \\ &= 1 - \int_{\Omega} k(\boldsymbol{x})\tilde{p}(\boldsymbol{x})\, d\boldsymbol{x}. \end{aligned} \tag{19}$$

For technical reasons, it is more convenient to work with a smooth function p_i rather than the Dirac distribution, which leads to a singularity of the function of $\tilde{p}(\boldsymbol{x}| \boldsymbol{y})$.

We will now define two mean times, averaged over the initial distribution: first the mean time τ_K a virus or a particle DNA is killed before reaching the nucleus. It is the mean first passage time to be killed conditioned on the event that the virus or the DNA is killed before reaching the absorbing boundary condition. Second, the mean time τ_N a virus or a naked DNA reaches the nucleus is defined as the mean first passage time to the absorbing boundary, conditioned on the event that they are not killed somewhere inside the domain Ω. We derive in the next paragraph some analytical expressions for both τ_N and τ_K.

2.2. Mean Time to Reach the Nucleus τ_N

The mean time τ_N a DNA particle reaches the nucleus is by definition the MFPT (Mean First Passage Time) conditioned on success, that is, the virus or the DNA must arrive alive at a nucleus pore. We derive now a set of partial differential equations (see also Ref. 9) satisfied by the MFPT. The probability distribution function (pdf) of the time τ^k that a viral trajectory is killed before reaching an absorbing pore ∂N_a is found by integrating the Fokker–Planck eq. (4) first with respect to $\boldsymbol{x}$ over the cell domain Ω, and second with respect to s from 0 to t.

We first derive an equation for pdf of the killing time $\Pr\{\tau^k < t \mid \tau^a > \tau^k, p_i\}$ conditioned first on the event that the virus will be killed before it escapes and second on the initial distribution of the virus. Starting with Bayes law and using eq. (19), we have that

$$\Pr\{\tau^k < t \mid \tau^a > \tau^k, p_i\} = \frac{\Pr\{\tau^k < t, \tau^a > \tau^k \mid p_i\}}{\Pr\{\tau^a > \tau^k \mid p_i\}} \tag{20}$$

where

$$\Pr\{\tau^a > \tau^k \mid p_i\} = \int_0^\infty \int_\Omega \int_\Omega k(\boldsymbol{x}) p(\boldsymbol{x}, s \mid \boldsymbol{y}) p_i(\boldsymbol{y})\, d\boldsymbol{x}\, ds d\boldsymbol{y}$$
$$= \int_\Omega k(\boldsymbol{x}) \tilde{p}(\boldsymbol{x})\, d\boldsymbol{x}$$

and

$$\Pr\{\tau^k < t, \tau^a > \tau^k \mid p_i\} = \int_0^t \int_\Omega \int_\Omega k(\boldsymbol{x}) p(\boldsymbol{x}, s \mid \boldsymbol{y})\, d\boldsymbol{x} d\boldsymbol{y}\, ds$$
$$= \int_0^t \int_\Omega k(\boldsymbol{x}) \tilde{p}(\boldsymbol{x}, s)\, d\boldsymbol{x}\, ds.$$

Thus expression (20) for the probability density function of the killing time is given by

$$\Pr\{\tau^k < t \mid \tau^a > \tau^k, p_i\} = \frac{\int_0^t \int_\Omega k(\boldsymbol{x}) \tilde{p}(\boldsymbol{x}, s)\, d\boldsymbol{x}\, ds}{\int_\Omega k(\boldsymbol{x}) \tilde{p}(\boldsymbol{x})\, d\boldsymbol{x}}. \tag{21}$$

Hence after integrating by parts, the MFPT is given by

$$\tau_K = E[\tau^k \mid \tau^k < \tau^a, p_i] = \int_0^\infty t \frac{d}{dt} \Pr\{\tau^k < t \mid \tau^a > \tau^k, p_i\} dt \tag{22}$$
$$= \frac{\int_0^\infty \int_t^\infty \int_\Omega k(\boldsymbol{x}) \tilde{p}(\boldsymbol{x}, s)\, d\boldsymbol{x}\, ds\, dt}{\int_\Omega k(\boldsymbol{x}) \tilde{p}(\boldsymbol{x})\, d\boldsymbol{x}}$$
$$= \frac{\int_0^\infty s \int_\Omega k(\boldsymbol{x}) \tilde{p}(\boldsymbol{x}, s)\, d\boldsymbol{x}\, ds}{\int_\Omega k(\boldsymbol{x}) p(\boldsymbol{x})\, d\boldsymbol{x}}. \tag{23}$$

To obtain an analytic expression for (22), we derive a partial differential equation satisfied by the function q defined by

$$q(\boldsymbol{x}) = \int_0^\infty s\, \tilde{p}(\boldsymbol{x}, s) ds. \tag{24}$$

Indeed, integrating the Fokker–Planck eq. (4) after multiplying by t, we get

$$\int_0^\infty t \frac{\partial p(\boldsymbol{x}, t \mid \boldsymbol{y})}{\partial t} dt = D\Delta q(\boldsymbol{x} \mid \boldsymbol{y}) - \nabla(q(\boldsymbol{x} \mid \boldsymbol{y}) b(\boldsymbol{x})) - k(\boldsymbol{x}) q(\boldsymbol{x} \mid \boldsymbol{y}) \quad \text{for} \quad \boldsymbol{x} \in \Omega.$$

Thus the function q satisfies the following boundary value problem

$$\begin{cases} -\tilde{p}(\boldsymbol{x} \mid \boldsymbol{y}) = D\Delta q(\boldsymbol{x} \mid \boldsymbol{y}) - \nabla(b(\boldsymbol{x}) q(\boldsymbol{x} \mid \boldsymbol{y})) - k(\boldsymbol{x} \mid \boldsymbol{y}) q(\boldsymbol{x} \mid \boldsymbol{y}) & \text{for} \quad \boldsymbol{x} \in \Omega \\ q(\boldsymbol{x} \mid \boldsymbol{y}) = 0 & \text{for} \quad \boldsymbol{x} \in \partial N_a \\ \boldsymbol{J}(\boldsymbol{x} \mid \boldsymbol{y}).\boldsymbol{n} = -D\nabla q(\boldsymbol{x}).\boldsymbol{n} + (b(\boldsymbol{x}).\boldsymbol{n}) q(\boldsymbol{x}) = 0 & \text{for} \quad \boldsymbol{x} \in \partial N_r \cup \partial\Omega_{\text{ext}}. \end{cases}$$

Integrating with respect to the initial distribution p_i and using that $\tilde{p}$ is solution of eq. (16), we obtain that

$$\begin{cases} -\tilde{p}(\boldsymbol{x}) = D\Delta q(\boldsymbol{x}) - \nabla(b(\boldsymbol{x})q(\boldsymbol{x})) - k(\boldsymbol{x})q(\boldsymbol{x}) & \text{for} \quad \boldsymbol{x} \in \Omega \\ q(\boldsymbol{x}) = 0 & \text{for} \quad \boldsymbol{x} \in \partial N_a \\ \boldsymbol{J}(\boldsymbol{x}|\, \boldsymbol{y}).\boldsymbol{n} = -D\nabla q(\boldsymbol{x}).\boldsymbol{n} + (b(\boldsymbol{x}).\boldsymbol{n})q(\boldsymbol{x}) = 0 & \text{for} \quad \boldsymbol{x} \in \partial N_r \cup \partial\Omega_{\text{ext}}. \end{cases} \tag{25}$$

Thus, the conditional MFPT to be killed can be expressed as follow:

$$E[\tau^k \mid \tau^k < \tau^a, p_i] = \frac{\int_\Omega k(\boldsymbol{x})q(\boldsymbol{x})\, d\boldsymbol{x}}{\int_\Omega k(\boldsymbol{x})\tilde{p}(\boldsymbol{x})\, d\boldsymbol{x}}. \tag{26}$$

Similarly, by using the pdf of the absorbing time τ^a to the absorbing boundary ∂N_a, for survival trajectories, we obtain an expression for the conditioned MFPT τ^a:

$$\Pr\{\tau^a < t \mid \tau^a < \tau^k, p_i\} = \frac{\int_0^t J(s \mid p_i)\, ds}{1 - \int_0^\infty \int_\Omega k(\boldsymbol{x})\tilde{p}(\boldsymbol{x}, s)\, d\boldsymbol{x}\, ds}, \tag{27}$$

where the flux is by definition

$$J(s \mid p_i) = \oint_{\partial\Omega} \boldsymbol{J}(\boldsymbol{x}, t|\, p_i) \cdot \boldsymbol{n}(\boldsymbol{x})\, dS_{\boldsymbol{x}}. \tag{28}$$

Thus the mean time τ^a to absorption is given by

$$\begin{aligned} \tau_N = E[\tau^a \mid \tau^a < \tau^k,\ p_i] &= \int_0^\infty (1 - \Pr\{\tau^a < t \mid \tau^a < \tau^k, \boldsymbol{y}\})\, dt \\ &= \frac{\int_0^\infty s J(s \mid p_i)\, ds}{1 - \int_\Omega k(\boldsymbol{x})\tilde{p}(\boldsymbol{x})\, d\boldsymbol{x}} = \frac{\int_\Omega \tilde{p}(\boldsymbol{x}) d\boldsymbol{x} - \int_\Omega k(\boldsymbol{x})q(\boldsymbol{x}) d\boldsymbol{x}}{1 - \int_\Omega k(\boldsymbol{x})\tilde{p}(\boldsymbol{x})\, d\boldsymbol{x}}, \end{aligned} \tag{29}$$

where the last identity is obtained by integrating Eq. 25 over the domain Ω.

3. EXPRESSION OF THE PROBABILITY AND THE MFPT A NAKED DNA REACHES A SMALL NUCLEAR PORE

Naked DNA are well described by pure Brownian diffusion. Thus to estimate p_N and τ_N, we set the vector field b to zero. The case $b \neq 0$ will be treated in sec. 4.

3.1. Probability to Reach the Nucleus Alive with a Zero Drift

In this section, we shall compute the probability $P_N(\boldsymbol{y})$ that a virus reaches one of the nuclear pore alive, conditioned on the initial position $\boldsymbol{y}$. We consider both dimensions $n = 2$ and $n = 3$. The difficulty of the present computation arises from the fact that a virus must arrive to only a small fraction of the nucleus surface. Thus the boundary conditions imposed on the nucleus are reflective except for a

small opening $\partial N_a = B_\eta$, which consists of a small ball of radius $\eta \ll 1$ (further on, we include the case of many small openings). Equation (16) for the probability becomes

$$D\Delta\tilde{p} - k(\boldsymbol{x})\tilde{p} = -p_i(\boldsymbol{x}) \quad \text{for} \quad \boldsymbol{x}, \boldsymbol{y} \in \Omega \tag{30}$$

$$\tilde{p}(\boldsymbol{x}) = 0 \quad \text{for} \quad \boldsymbol{x} \in \partial N_a \tag{31}$$

$$\frac{\partial \tilde{p}}{\partial n}(\boldsymbol{x}) = 0 \quad \text{for} \quad \boldsymbol{x} \in \partial N_r \cup \partial\Omega_{\text{ext}}. \tag{32}$$

Similarly to the method proposed in Ref. 10, to determine asymptotically the solution of (30), we shall use the Green function $G_Q(\boldsymbol{x})$, solution of

$$D\Delta G_Q(\boldsymbol{x}) = -\delta_Q(\boldsymbol{x}), \quad \boldsymbol{x} \in \Omega, \tag{33}$$

$$D\frac{\partial G_Q}{\partial n}(\boldsymbol{x}) = \frac{-1}{|\partial\Omega|}, \quad \text{for} \quad \boldsymbol{x} \in \partial\Omega, \tag{34}$$

where δ_Q is the Dirac function centered at point Q. G_Q is defined up to an additive constant that will be specified later on. We have,[6]

$$D\,G_Q(\boldsymbol{x}) = \begin{cases} -\frac{1}{2\pi}\ln|\boldsymbol{x} - Q| + w_Q, & \text{for} \quad n = 2 \\ \frac{1}{4\pi}|\boldsymbol{x} - Q|^{-1} + w_Q, & \text{for} \quad n = 3, \end{cases} \tag{35}$$

where w_Q is a regular harmonic function. We recall that on the regular boundary ∂N_a, due to the image source, the singularity of the Green function is multiplied by a factor 2 (see Ref. 10 for further details).

3.2. Probability to Reach the Nucleus for a Small Killing Rate ($k \ll 1$)

To estimate p_N asymptotically, we use the method developed in Ref. 10 which consists first in deriving an integral equation by using the Green identity and then, to use the Dirichlet condition at the small absorbing boundary. As the size of the absorbing boundary goes to zero, we approximate asymptotically the leading order term of the conditional probability by a constant. The asymptotic expression is obtained by using a Taylor expansion of the probability as a function of the absorbing boundary size.

An integral representation of $\tilde{p}$ is obtained by using eq. (30) and (33) in Green's formula. We get from

$$\int_\Omega \left(G_Q\Delta\tilde{p} - \tilde{p}\Delta G_Q\right)(\boldsymbol{x})d\boldsymbol{x} = \int_{\partial\Omega}\left(G_Q\frac{\partial\tilde{p}}{\partial n} - \tilde{p}\frac{\partial G_Q}{\partial n}\right)(\boldsymbol{x}(S))dS. \tag{36}$$

that

$$\int_\Omega G_Q(\boldsymbol{x})\left(-p_i(\boldsymbol{x}) + k(\boldsymbol{x})\tilde{p}(\boldsymbol{x})\right)d\boldsymbol{x} + \tilde{p}(Q) = D\int_{\partial N_a} G_Q\frac{\partial\tilde{p}}{\partial n}dS + \frac{1}{|\partial\Omega|}\int_{\partial\Omega}\tilde{p}dS.$$

We assume that the absorbing boundary is a small ball B_η (in dimension 3) or a small arclength (in dimension 2) $\partial N_a = B_\eta$, where $\eta \ll 1$. Under this assumption the leading order term $\tilde{p}(\boldsymbol{x})$ outside a boundary layer located near the absorbing boundary can be approximated by a constant (see Ref. 18), $\tilde{p}(\boldsymbol{x}) \approx P_\eta$ when the killing rate k is much smaller, less than the effect of diffusion. Thus, we get

$$\tilde{p}(Q) = D \int_{\partial N_a} G_Q \frac{\partial \tilde{p}}{\partial n} dS + P_\eta \left(1 - \int_\Omega k(\boldsymbol{x}) G_Q(\boldsymbol{x}) d\boldsymbol{x}\right) + \int_\Omega G_Q(\boldsymbol{y}) p_i(\boldsymbol{y}) d\boldsymbol{y}. \tag{37}$$

In order to estimate the constant P_η (as a function of $\boldsymbol{x}$), we use that at the absorbing boundary ∂N_a, the function $\tilde{p}$ has to vanish. If we denote the unknown flux by

$$g(S) = \frac{\partial \tilde{p}}{\partial n}(\boldsymbol{x}(S)), \tag{38}$$

and take a point $Q \in \partial N_a$ where $\tilde{p}(Q) = 0$, then we obtain the relation

$$0 = D \int_{\partial N_a} G_Q(\boldsymbol{x}(S)) g(S) dS + P_\eta \left(1 - \int_\Omega k(\boldsymbol{x}) G_Q(\boldsymbol{x}) d\boldsymbol{x}\right) + \int_\Omega G_Q(\boldsymbol{y}) p_i(\boldsymbol{y}) d\boldsymbol{y}. \tag{39}$$

To compute $g(s)$, we integrate (30) over the domain Ω, using that $\int_\Omega p_i(\boldsymbol{x}) d\boldsymbol{x} = 1$,

$$D \int_{\partial N_a} g(S) dS = -1 + \int_\Omega k(\boldsymbol{x}) \tilde{p}(\boldsymbol{x}) d\boldsymbol{x} = -1 + P_\eta \int_\Omega k(\boldsymbol{x}) d\boldsymbol{x}. \tag{40}$$

To estimate the left-hand side of expression 40, we use a Taylor expansion of the flux $g(S)$ at a fixed point in the interior of the absorbing boundary far from the edges (see Ref. 10), we get

$$\int_{\partial N_a} g(S) dS = \int_{-\eta}^{\eta} g(S) dS = 2\eta g_0 + o(\eta), \tag{41}$$

where g_0 is the first term of the expansion of $g(S)$. Thus, we obtain from relation (40) that

$$g_0 = \frac{-1 + P_\eta \int_\Omega k(\boldsymbol{x}) d\boldsymbol{x}}{2D\eta}. \tag{42}$$

To estimate the constant P_η, we use the expansion of the Green function in dimension 2 and 3 in expression (39),

$$DG_Q(\boldsymbol{x}) = \begin{cases} -\frac{1}{2\pi}\ln|\boldsymbol{x}-Q| + C_0 + h_Q, & \text{for} \quad n=2 \\ \frac{1}{4\pi}|\boldsymbol{x}-Q|^{-1} + C_0 + h_Q, & \text{for} \quad n=3, \end{cases} \tag{43}$$

where C_0 is a constant and h_Q a harmonic function, then in dimension 2, we get (for the details of the computations see Ref. 10)

$$\begin{aligned} 0 &= D\int_{\partial N_a} G_Q(\boldsymbol{x}(S))g(S)dS + P_\eta\left(1-\int_\Omega k(\boldsymbol{x})G_Q(\boldsymbol{x})d\boldsymbol{x}\right)+\int_\Omega G_Q(\boldsymbol{y})p_i(\boldsymbol{y})d\boldsymbol{y} \\ &\approx -\frac{1}{\pi}\int_{-\eta}^{\eta}\ln|s|g_0 dS + 2(C_0-1)g_0\eta + P_\eta\left(1-\int_\Omega k(\boldsymbol{x})G_Q(\boldsymbol{x})d\boldsymbol{x}\right) \\ &\quad + \int_\Omega G_Q(\boldsymbol{y})p_i(\boldsymbol{y})d\boldsymbol{y} \\ &\approx 2g_0\eta\left(-\frac{\ln\eta}{\pi}+(C_0-1)\right)+P_\eta\left(1-\int_\Omega k(\boldsymbol{x})G_Q(\boldsymbol{x})d\boldsymbol{x}\right)+\int_\Omega G_Q(\boldsymbol{y})p_i(\boldsymbol{y})d\boldsymbol{y} \\ &\approx \frac{1}{D}\left(-1+P_\eta\int_\Omega k(\boldsymbol{x})d\boldsymbol{x}\right)\left(-\frac{\ln\eta}{\pi}+(C_0-1)\right)+P_\eta\left(1-\int_\Omega k(\boldsymbol{x})G_Q(\boldsymbol{x})d\boldsymbol{x}\right) \\ &\quad + \int_\Omega G_Q(\boldsymbol{y})p_i(\boldsymbol{y})d\boldsymbol{y}. \end{aligned}$$

It is easy to check that this identity does not depend on the Green function. Finally, far enough from the boundary layer near the window $\partial\Omega_a$, the solution $\tilde{p}(\boldsymbol{x})$ is approximated by,

$$\tilde{p}(\boldsymbol{x}) \approx P_\eta \approx \frac{\frac{\ln\frac{1}{\eta}}{D\pi} + C_0 - 1 - \int_\Omega G_Q(\boldsymbol{y})p_i(\boldsymbol{y})d\boldsymbol{y}}{\frac{\ln\frac{1}{\eta}}{D\pi}\int_\Omega k(\boldsymbol{x})d\boldsymbol{x} + (C_0-1)\int_\Omega k(\boldsymbol{x})d\boldsymbol{x} + \left(1-\int_\Omega k(\boldsymbol{x})G_Q(\boldsymbol{x})d\boldsymbol{x}\right)} \tag{44}$$

Thus using formula (12), we obtain an expression for the probability P_N given by

$$\begin{aligned} P_N &= 1 - \int_\Omega k(\boldsymbol{x})\tilde{p}(\boldsymbol{x})d\boldsymbol{x} \\ &\approx \frac{1+\int_\Omega k(\boldsymbol{x})d\boldsymbol{x}\int_\Omega G_Q(\boldsymbol{y})p_i(\boldsymbol{y})d\boldsymbol{y} - \int_\Omega k(\boldsymbol{x})G_Q(\boldsymbol{x})d\boldsymbol{x}}{\frac{\ln\frac{1}{\eta}}{D\pi}\int_\Omega k(\boldsymbol{x})d\boldsymbol{x} + (C_0-1)\int_\Omega k(\boldsymbol{x})d\boldsymbol{x} + \left(1-\int_\Omega k(\boldsymbol{x})G_Q(\boldsymbol{x})d\boldsymbol{x}\right)}. \end{aligned}$$

In the limit of $k \ll 1$, whatever is the choice of C_0, which can be used to fix the Green function, we can neglect terms containing k (which are all bounded), except the term in the denominator which is multiplied by a large flux. The leading order

term for the probability P_N in dimension 2 is given by

$$P_N \approx \frac{1}{1 + \frac{\ln \frac{1}{\eta}}{D\pi} \int_\Omega k(\boldsymbol{x}) d\boldsymbol{x}} \quad \text{for} \quad k(\boldsymbol{x}) \ll 1,$$

where we recall that in dimension two, η represents the ratio of absorbing to the total nucleus surface. In dimension 3, computations are similar. Let us outline the main differences. The flux through ∂N_a, which is a small disk of radius η, is given by Ref. 18

$$\frac{\partial \tilde{p}}{\partial n}(\boldsymbol{x}(S)) = g(s) = \frac{g_0}{\sqrt{\eta^2 - s^2}} + o(\eta) \tag{45}$$

The leading order term of $g(s)$ is computed by integrating eq. (30)

$$D \int_{\partial N_a} g(S) dS = -1 + P_\eta \int_\Omega k(\boldsymbol{x}) d\boldsymbol{x} \tag{46}$$

and by a direct computation, we obtain

$$\int_{\partial N_a} g(S) dS = \int_0^\eta \frac{g_0}{\sqrt{\eta^2 - s^2}} 2\pi s dS = 2\pi \eta g_0 \tag{47}$$

leading to the estimate:

$$g_0 = \frac{-1 + P_\eta \int_\Omega k(\boldsymbol{x}) d\boldsymbol{x}}{2\pi D\eta}. \tag{48}$$

As in dimension 2, replacing the leading order term in relation (39) by a constant and using the Green function expansion (of G_Q) in dimension 3, we get

$$D \int_{\partial N_a} G_Q(x(S)) g(S) dS = \int_0^\eta \left(\frac{1}{2\pi} |s|^{-1} + C_0 \right) \frac{g_0}{\sqrt{\eta^2 - s^2}} 2\pi s ds \tag{49}$$

Thus Eq. (39) becomes

$$\begin{aligned}
0 &= \frac{1}{2\pi} \int_0^\eta |s|^{-1} \frac{g_0}{\sqrt{\eta^2 - s^2}} 2\pi s ds + 2\pi \eta C_0 g_0 + P_\eta \left(1 - \int_\Omega k(\boldsymbol{x}) G_Q(\boldsymbol{x}) d\boldsymbol{x} \right) \\
&\quad + \int_\Omega G_Q(\boldsymbol{y}) p_i(\boldsymbol{y}) d\boldsymbol{y} \\
&= g_0 \left(\frac{\pi}{2} + 2\pi \eta C_0 \right) + P_\eta \left(1 - \int_\Omega k(\boldsymbol{x}) G_Q(\boldsymbol{x}) d\boldsymbol{x} \right) + \int_\Omega G_Q(\boldsymbol{y}) p_i(\boldsymbol{y}) d\boldsymbol{y} \\
&= \left(-1 + P_\eta \int_\Omega k(\boldsymbol{x}) d\boldsymbol{x} \right) \left(\frac{1}{4D\eta} + C_0 \right) + P_\eta \left(1 - \int_\Omega k(\boldsymbol{x}) G_Q(\boldsymbol{x}) d\boldsymbol{x} \right) \\
&\quad + \int_\Omega G_Q(\boldsymbol{y}) p_i(\boldsymbol{y}) d\boldsymbol{y}.
\end{aligned}$$

Thus

$$P_\eta \approx \frac{\frac{1}{4D\eta} + C_0 - \int_\Omega G_Q(\boldsymbol{y}) p_i(\boldsymbol{y}) d\boldsymbol{y}}{\frac{1}{4D\eta}\int_\Omega k(\boldsymbol{x})d\boldsymbol{x} + C_0 \int_\Omega k(\boldsymbol{x})d\boldsymbol{x} + \left(1 - \int_\Omega k(\boldsymbol{x}) G_Q(\boldsymbol{x})d\boldsymbol{x}\right)}. \quad (50)$$

Finally, using the small k approximation,

$$P_N = 1 - \int_\Omega k(\boldsymbol{x})\tilde{p}(\boldsymbol{x})d\boldsymbol{x} \approx \frac{1}{1 + \frac{1}{4D\eta}\int_\Omega k(\boldsymbol{x})d\boldsymbol{x}} \quad \text{for} \quad k(\boldsymbol{x}) \ll 1. \quad (51)$$

Remark. In dimension 2, η represents the ratio of the absorbing part to the reflective part of the nucleus.

3.3. Conditional Mean Time to Reach the Nucleus

In this section, we shall compute the average time τ_N a nacked DNA reaches a nuclear pore alive. The computations go along the same lines as the ones developed in the previous section, but their require to study a system of coupled partial differential equations.

We start by estimating the solution q of eq. (25). Outside the boundary layer of the small hole, $\tilde{p}$ has been approximated by P_η, thus

$$\begin{cases} -P_\eta = D\Delta q(\boldsymbol{x}) - k(\boldsymbol{x})q(\boldsymbol{x}) & \text{in} \quad \Omega \\ q(\boldsymbol{x}) = 0 & \text{on} \quad \partial N_a \\ \dfrac{\partial}{\partial \boldsymbol{n}} q(\boldsymbol{x}) = 0 & \text{on} \quad \partial N_r \cup \partial\Omega_{\text{ext}}. \end{cases} \quad (52)$$

The analysis of eq. (52) follows the steps of the previous paragraph, where q replaces p_i in eq. 30. Outside a boundary layer of the small absorbing window, the solution q can be approximated by a constant, $q(\boldsymbol{x}) \approx T_\mu$, thus the Green representation gives

$$q(Q) = T_\eta \left(1 - \int_\Omega G_Q(\boldsymbol{x})k(\boldsymbol{x})d\boldsymbol{x}\right) + \int_\Omega P_\eta G_Q(\boldsymbol{x})d\boldsymbol{x} + D\int_{\partial N_a} G_Q \frac{\partial q}{\partial n} dS. \quad (53)$$

The value T_η is obtained first by estimating the flux $g(S) = \frac{\partial q}{\partial n}(\boldsymbol{x}(S))$ and second by using the absorbing boundary condition at $\partial\Omega_a$. Integrating the first equation in (52), we get

$$D\int_{\partial N_a} g(S)dS = -P_\eta|\Omega| + \int_\Omega k(\boldsymbol{x})q(\boldsymbol{x})d\boldsymbol{x} \approx -P_\eta|\Omega| + T_\eta \int_\Omega k(\boldsymbol{x})d\boldsymbol{x}. \quad (54)$$

For $Q \in \partial N_a$, the boundary condition (53) gives that

$$0 = T_\eta \left(1 - \int_\Omega G_Q(\boldsymbol{x})k(\boldsymbol{x})d\boldsymbol{x}\right) + \int_\Omega P_\eta G_Q(\boldsymbol{x})d\boldsymbol{x} + D\int_{\partial N_a} G_Q(\boldsymbol{x}(S))g(S)dS. \quad (55)$$

The flux g is computed by using eq. (54) and the leading expansion

$$g(s) = \begin{cases} g_0 + o(s) & \text{for} \quad n = 2, \\ \dfrac{g_0}{\sqrt{\eta^2 - s^2}} + o(\eta) & \text{for} \quad n = 3 \end{cases} \tag{56}$$

we get

$$\int_{\partial N_a} g(S)dS = \int_{-\eta}^{\eta} g(S)dS = \begin{cases} 2\eta g_0, & \text{for} \quad n = 2 \\ 2\pi\eta g_0, & \text{for} \quad n = 3 \end{cases} \tag{57}$$

Thus,

$$g_0 = \begin{cases} \dfrac{-P_\eta|\Omega| + T_\eta \int_\Omega k(\boldsymbol{x})d\boldsymbol{x}}{2D\eta} & \text{for} \quad n = 2 \\ \dfrac{-P_\eta|\Omega| + T_\eta \int_\Omega k(\boldsymbol{x})d\boldsymbol{x}}{2\pi D\eta} & \text{for} \quad n = 3 \end{cases} \tag{58}$$

Injecting the leading order term of the flux in eq. (55), we get

$$0 = \begin{cases} T_\eta \left(1 - \int_\Omega G_Q(\boldsymbol{x})k(\boldsymbol{x})d\boldsymbol{x}\right) + P_\eta \int_\Omega G_Q(\boldsymbol{x})d\boldsymbol{x} + 2\eta g_0 \\ \qquad\qquad \times\left(-\frac{\ln\eta}{\pi} + C_0 - 1\right), \quad \text{for} \quad n = 2 \\ T_\eta \left(1 - \int_\Omega G_Q(\boldsymbol{x})k(\boldsymbol{x})d\boldsymbol{x}\right) + P_\eta \int_\Omega G_Q(\boldsymbol{x})d\boldsymbol{x} + \pi g_0\left(\frac{1}{2} + 2\eta C_0\right) \quad \text{for} \quad n = 3 \end{cases} \tag{59}$$

Finally,

$$T_\eta = \begin{cases} P_\eta \dfrac{\frac{|\Omega|}{D}\left(\frac{\ln\left(\frac{1}{\eta}\right)}{\pi} + C_0 - 1\right) - \int_\Omega G_Q(\boldsymbol{x})d\boldsymbol{x}}{\frac{1}{D}\left(\frac{\ln\left(\frac{1}{\eta}\right)}{\pi} + C_0 - 1\right)\int_\Omega k(\boldsymbol{x})d\boldsymbol{x} + 1 - \int_\Omega G_Q(\boldsymbol{x})k(\boldsymbol{x})d\boldsymbol{x}}. \\ P_\eta \dfrac{|\Omega|\left(\frac{1}{4D\pi} + C_0\right) - \int_\Omega G_Q(\boldsymbol{x})d\boldsymbol{x}}{\frac{1}{D}\left(\frac{1}{4\pi} + C_0\right)\int_\Omega k(\boldsymbol{x})d\boldsymbol{x} + 1 - \int_\Omega G_Q(\boldsymbol{x})k(\boldsymbol{x})d\boldsymbol{x}}. \end{cases} \tag{60}$$

Using the expression (29) for the reaching time and the expression of the probability P_N, we get

$$\tau_N \approx \frac{P_\eta|\Omega| - T_\eta \int_\Omega k(\boldsymbol{x})d\boldsymbol{x}}{1 - P_\eta \int_\Omega k(\boldsymbol{x})d\boldsymbol{x}} \approx \frac{P_\eta|\Omega| - T_\eta \int_\Omega k(\boldsymbol{x})d\boldsymbol{x}}{P_N}. \tag{61}$$

That is, using expressions (60) and under the assumptions that

$$\int_\Omega \mathbf{k}(\boldsymbol{x})\mathbf{d}\boldsymbol{x}, \int_\Omega \mathbf{G_Q}(\boldsymbol{x})\mathbf{k}(\boldsymbol{x})\mathbf{d}\boldsymbol{x} \ll \mathbf{1},$$

we finally obtain the leading order term of the reaching time

$$\tau_N \approx \begin{cases} \dfrac{|\Omega|\left(\frac{\ln\left(\frac{1}{\eta}\right)}{D\pi}\right) - |\Omega| \int_\Omega G_Q(\boldsymbol{x}) p_i(\boldsymbol{x}) d\boldsymbol{x}}{1 + \left(\frac{\ln\left(\frac{1}{\eta}\right)}{D\pi}\right) \int_\Omega k(\boldsymbol{x}) d\boldsymbol{x}} & \text{for} \quad n = 2 \qquad \text{for} \quad C_0 = 1 \\ \dfrac{|\Omega|\left(\frac{1}{4D\eta}\right) - |\Omega| \int_\Omega G_Q(\boldsymbol{x}) p_i(\boldsymbol{x}) d\boldsymbol{x}}{1 + \left(\frac{\int_\Omega k(\boldsymbol{x}) d\boldsymbol{x}}{4D\eta}\right)} & \text{for} \quad n = 3 \qquad \text{for} \quad C_0 = 0. \end{cases} \tag{62}$$

3.3.1. Many Small Holes

If we take into account n well separated small holes, located on the surface of the nucleus, the previous formulas are modified by summing over the contribution of each flux separately , thus we get

$$P_N = \begin{cases} \dfrac{1}{1 + \frac{\ln \frac{1}{\eta}}{nD\pi} \int_\Omega k(\boldsymbol{x}) d\boldsymbol{x}} & \text{for} \quad n = 2. \\ \dfrac{1}{1 + \frac{1}{4nD\eta} \int_\Omega k(\boldsymbol{x}) d\boldsymbol{x}} & \text{for} \quad n = 3. \end{cases} \tag{63}$$

Neglecting the first order terms in expressions 62, we obtain the following asymptotic expansion for the mean reaching time,

$$\tau_N \approx \begin{cases} \dfrac{\left(\frac{|\Omega| \ln\left(\frac{1}{\eta}\right)}{nD\pi}\right)}{1 + \left(\frac{\ln\left(\frac{1}{\eta}\right)}{nD\pi}\right) \int_\Omega k(\boldsymbol{x}) d\boldsymbol{x}} & \text{for} \quad n = 2 \\ \dfrac{\left(\frac{|\Omega|}{4D\eta n}\right)}{1 + \left(\frac{\int_\Omega k(\boldsymbol{x}) d\boldsymbol{x}}{4nD\eta}\right)} & \text{for} \quad n = 3. \end{cases} \tag{64}$$

To test our model, we propose to estimate the probability and the mean time a plasmid needs to reach a small nuclear pore. The typical diffusion coefficient for a cytosolic plasmid DNA of 5500 base pairs (average size of a gene) is about $D \approx 0.02 \mu m^2/s$. We consider a spherical cell of radius $R = 5\mu m$, having 10 percent of its nucleus surface occupied by nuclear pores.[13] This gives a number of nuclear pores of $n = 160$ in dimension three and $n = 25$ in dimension two. The lifespan of the cytoplasmic DNA is about one hour[12] and we choose $k = 1/3600\, s^{-1}$. Using formula (63), we get for the delivery probability

$$P_N \approx 0.9371, \quad \text{for} \quad n = 2 \tag{65}$$

$$P_N \approx 0.6875, \quad \text{for} \quad n = 3 \tag{66}$$

while the conditional time to arrive alive is given by formula (64) and it is given by

$$\tau_N \approx 226s \qquad \text{for} \quad n = 2 \tag{67}$$

$$\tau_N \approx 1125s \qquad \text{for} \quad n = 3. \tag{68}$$

The computations made in dimension two apply for a flat cell, while the dimension three computation is relevant for round cells. The difference in the numerical results for the reaching probability can be explained by noting that in dimension 3, the space visited by a Brownian molecule before reaching a nuclear pore is much more than in dimension two, thus the probability to be killed is increased.

4. THE CASE OF A NON-ZERO DRIFT

In this section, we study the effect of a drift in the computation of the probability P_N and the mean reaching time τ_N. For simplicity, we treat the case of a gradient drift, given by $b = -\nabla\phi$, for which the origin of coordinates, taken inside the nucleus, is an attractor point. At the cell external surface, we assume that the drift is directed inside so that all trajectories are entering, which translates into the condition $(b(\boldsymbol{x}).\boldsymbol{n}) < 0$ for all $\boldsymbol{x} \in \partial\Omega_{\text{ext}}$. As mentioned in the introduction, the potential of the field is given by

$$\phi(\boldsymbol{x}) = Br \quad \text{for} \quad \boldsymbol{x} \in \Omega, \tag{69}$$

where the radial distance is $r = |\boldsymbol{x}|$. The following computations extend some earlier developments.[21]

4.1. Probability to Reach the Nucleus

To estimate the probability P_N, we now estimate asymptotically the solution $\tilde{p}$ of eq. (16). Outside the boundary layer near the absorbing boundary $\partial\Omega_a$, we choose the ansatz

$$\tilde{p}(\boldsymbol{x}) = C_\eta e^{-\frac{\phi(\boldsymbol{x})}{D}}. \tag{70}$$

To obtain an estimate of the constant C_η, we follow the steps presented in the previous section. First, integrating (16) and using the zero flux boundary conditions, we get

$$D\int_{\partial N_a} g(S)dS = -1 + \int_\Omega k(\boldsymbol{x})\tilde{p}(\boldsymbol{x})d\boldsymbol{x}, \tag{71}$$

where the flux is given by $g(S) = \frac{\partial\tilde{p}}{\partial\boldsymbol{n}}(S)$. Equation (71) can be approximated as follows: for s in $\partial\Omega_a$,

$$g(s) = \begin{cases} g_0 e^{-\frac{\phi(\boldsymbol{x}_0)}{D}} + o(s) & \text{for} \quad n = 2, \\ \dfrac{g_0 e^{-\frac{\phi(\boldsymbol{x}_0)}{D}}}{\sqrt{\eta^2 - s^2}} + o(\eta) & \text{for} \quad n = 3, \end{cases} \tag{72}$$

where $\phi(\boldsymbol{x}_0)$ is a value of ϕ in $\partial\Omega_a$ (we assume indeed that ϕ is a regular not oscillating function and as η goes to zero, $\phi(x(s))$ converges uniformly to $\phi(\boldsymbol{x}_0)$

for all s in the interval $[-\eta, \eta]$). Thus,

$$g_0 = \begin{cases} e^{\frac{\phi(x_0)}{D}} \left(\frac{-1 + C_\eta \int_\Omega k(\boldsymbol{x}) e^{-\frac{\phi(\boldsymbol{x})}{D}} d\boldsymbol{x}}{2D\eta} \right) & \text{for} \quad n = 2 \\ e^{\frac{\phi(x_0)}{D}} \left(\frac{-1 + C_\eta \int_\Omega k(\boldsymbol{x}) e^{-\frac{\phi(\boldsymbol{x})}{D}} d\boldsymbol{x}}{2\pi D\eta} \right) & \text{for} \quad n = 3. \end{cases} \tag{73}$$

Using the Green's representation of the function $\tilde{p}$, we obtain

$$\int_\Omega G_Q(\boldsymbol{x}) (-p_i(\boldsymbol{x}) + k(\boldsymbol{x})\tilde{p}(\boldsymbol{x})) \, d\boldsymbol{x} + \int_\Omega G_Q(\boldsymbol{x}) \nabla(\nabla\phi(\boldsymbol{x})\tilde{p}(\boldsymbol{x})) + \tilde{p}(Q)$$
$$= D \int_{\partial\Omega} G_Q \frac{\partial \tilde{p}}{\partial n} dS + \frac{1}{|\partial\Omega|} \int_{\partial\Omega} \tilde{p} dS.$$

Integrating by parts the gradient term, we get

$$\int_\Omega G_Q(\boldsymbol{x}) (-p_i(\boldsymbol{x}) + k(\boldsymbol{x})\tilde{p}(\boldsymbol{x})) \, d\boldsymbol{x} - \int_\Omega \nabla G_Q(\boldsymbol{x}).\nabla\phi(\boldsymbol{x})\tilde{p}(\boldsymbol{x}) + \tilde{p}(Q)$$
$$= D \int_{\partial\Omega_a} G_Q \frac{\partial \tilde{p}}{\partial n} dS + \frac{1}{|\partial\Omega|} \int_{\partial\Omega} \tilde{p} dS. \tag{74}$$

Using the absorbing boundary condition $\tilde{p}(Q) = 0$ for $Q \in \partial\Omega_a$ and the ansatz (70), we finally obtain the relation

$$D \int_{\partial\Omega_a} G_Q g(S) dS + \frac{1}{|\partial\Omega|} \int_{\partial\Omega} C_\eta e^{-\frac{\phi(\boldsymbol{x})}{D}} dS$$
$$= \int_\Omega G_Q(\boldsymbol{x}) \left(-p_i(\boldsymbol{x}) + k(\boldsymbol{x}) C_\eta e^{-\frac{\phi(\boldsymbol{x})}{D}} \right) d\boldsymbol{x}$$
$$- \int_\Omega \nabla G_Q(\boldsymbol{x}).\nabla\phi(\boldsymbol{x}) C_\eta e^{-\frac{\phi(\boldsymbol{x})}{D}} d\boldsymbol{x} \tag{75}$$

The first term in the left-hand side is computed by using the flux condition (73). We get (as in the previous section)

$$D \int_{\partial\Omega_a} G_Q g(S) dS = \begin{cases} 2D\eta g_0 e^{-\frac{\phi(x_0)}{D}} \left(-\frac{\ln \eta}{\pi} + C_0 - 1 \right) & \text{for} \quad n = 2 \\ \pi D e^{-\frac{\phi(x_0)}{D}} g_0 \left(\frac{1}{2} + 2\eta C_0 \right) & \text{for} \quad n = 3. \end{cases} \tag{76}$$

Choosing $C_0 = 1$ in dimension 2 and $C_0 = 0$ in dimension 3, we get from relation (75),

$$C_\eta \approx \begin{cases} \dfrac{\frac{\ln\frac{1}{\eta}}{D\pi} - \int_\Omega G_Q(\boldsymbol{y}) p_i(\boldsymbol{y}) d\boldsymbol{y}}{\frac{\ln\frac{1}{\eta}}{D\pi}\int_\Omega e^{-\frac{\phi(\boldsymbol{x})}{D}} k(\boldsymbol{x}) d\boldsymbol{x} + \frac{1}{|\partial\Omega|}\int_{\partial\Omega} e^{-\frac{\phi(\boldsymbol{x})}{D}} dS - \int_\Omega e^{-\frac{\phi(\boldsymbol{x})}{D}} k(\boldsymbol{x}) G_Q(\boldsymbol{x}) d\boldsymbol{x}} & \text{for} \quad n=2 \\ \dfrac{\frac{1}{4D\eta} - \int_\Omega G_Q(\boldsymbol{y}) p_i(\boldsymbol{y}) d\boldsymbol{y}}{\frac{1}{4D\eta}\int_\Omega e^{-\frac{\phi(\boldsymbol{x})}{D}} k(\boldsymbol{x}) d\boldsymbol{x} + \frac{1}{|\partial\Omega|}\int_{\partial\Omega} e^{-\frac{\phi(\boldsymbol{x})}{D}} dS - \int_\Omega e^{-\frac{\phi(\boldsymbol{x})}{D}} k(\boldsymbol{x}) G_Q(\boldsymbol{x}) d\boldsymbol{x}} & \text{for} \quad n=3. \end{cases}$$

Under the assumption that the integrals in k are small compared to 1, we get that

$$P_N = 1 - \int_\Omega k(\boldsymbol{x}) \tilde{p}(\boldsymbol{x}) d\boldsymbol{x} = 1 - C_\eta \int_\Omega k(\boldsymbol{x}) e^{-\frac{\phi(\boldsymbol{x})}{D}} d\boldsymbol{x}, \tag{77}$$

thus using the expression of C_η, we obtain the expression of the probability to reach the nucleus:

$$P_N \approx \begin{cases} \dfrac{\frac{1}{|\partial\Omega|}\int_{\partial\Omega} e^{-\frac{\phi(\boldsymbol{x})}{D}} dS}{\frac{\ln\frac{1}{\eta}}{D\pi}\int_\Omega e^{-\frac{\phi(\boldsymbol{x})}{D}} k(\boldsymbol{x}) d\boldsymbol{x} + \frac{1}{|\partial\Omega|}\int_{\partial\Omega} e^{-\frac{\phi(\boldsymbol{x})}{D}} dS} & \text{for} \quad n=2, \\ \dfrac{\frac{1}{|\partial\Omega|}\int_{\partial\Omega} e^{-\frac{\phi(\boldsymbol{x})}{D}} dS}{\frac{1}{4D\eta}\int_\Omega e^{-\frac{\phi(\boldsymbol{x})}{D}} k(\boldsymbol{x}) d\boldsymbol{x} + \frac{1}{|\partial\Omega|}\int_{\partial\Omega} e^{-\frac{\phi(\boldsymbol{x})}{D}} dS} & \text{for} \quad n=3. \end{cases} \tag{78}$$

The probability P_N can be further estimated by using the Laplace method when $D \ll 1$ or when the exponential term is small, which requires that $B\ Diam \ll D$, where $Diam$ is the diameter of the cell.

When the potential ϕ is an increasing function of the distance from the nucleus N, the minimum of the phase is achieved on the boundary of the nucleus. For simplicity, we assume that the geometry of the nucleus is a ball of radius δ. When the diffusion constant is small compared to the potential $D \ll \phi$, we consider two extreme cases: first ϕ achieves its minimum uniformly on the nucleus surface and second, it is achieved at a finite number of isolated points.

In the first case, using the explicit expression of the field $\phi = Br$, we get on the surface of the nucleus

$$\frac{1}{|\partial\Omega|}\int_{\partial\Omega} e^{-\frac{\phi(\boldsymbol{x})}{D}} \approx e^{-\frac{\delta B}{D}}. \tag{79}$$

By successive integrations by parts, if we denote

$$\tilde{k}(r) = \int_{r=\delta} k(r,\theta) d\theta, \tag{80}$$

then we obtain

$$\int_\Omega e^{-\frac{\phi(\boldsymbol{x})}{D}} k(\boldsymbol{x}) d\boldsymbol{x} = \int_\delta^R e^{-\frac{Br}{D}} r^{n-1} dr \int_{r=\delta} k(r,\theta) d\theta$$

$$\approx \left\{ \frac{D}{B}\delta^{n-1}\tilde{k}(\delta) + \left(\frac{D}{B}\right)^2 \left(\delta^{n-1}\tilde{k}'(\delta) + (n-1)\delta^{n-2}\tilde{k}(\delta)\right) \right.$$
$$\left. + \left(\frac{D}{B}\right)^3 \left(\delta^{n-1}\tilde{k}''(\delta) + 2(n-1)\delta^{n-2}\tilde{k}'(\delta) + (n-1)(n-2)\delta^{n-3}\tilde{k}(\delta)\right) \right\} e^{-\frac{\delta B}{D}}.$$

When k is locally constant in a neighborhood of the nucleus $r = \delta$, we obtain the following expression

$$\int_\Omega e^{-\frac{\phi(\boldsymbol{x})}{D}} k(\boldsymbol{x}) d\boldsymbol{x} = \left(\frac{D}{B}\delta^{n-1} + \left(\frac{D}{B}\right)^2 (n-1)\delta^{n-2} \right.$$
$$\left. + \left(\frac{D}{B}\right)^3 (n-1)(n-2)\delta^{n-3} \right) k S_n e^{-\frac{\delta B}{D}}, \tag{81}$$

where

$$S_n = \begin{cases} 2\pi & \text{for} \quad n = 2, \\ 4\pi & \text{for} \quad n = 3. \end{cases} \tag{82}$$

Finally, under the previous approximations, formula (78) reduces to

$$P_N \approx \begin{cases} \frac{1}{\frac{2\ln\frac{1}{\eta}}{D}\left(\frac{D}{B}\delta + 2\left(\frac{D}{B}\right)^2\right)k+1} & \text{for} \quad n = 2, \\ \frac{1}{\frac{\pi}{D\eta}\left(\frac{D}{B}\delta^2 + 2\left(\frac{D}{B}\right)^2\delta + 2\left(\frac{D}{B}\right)^3\right)k+1} & \text{for} \quad n = 3. \end{cases} \tag{83}$$

In first approximation, when the diffusion is small, it is interesting to note that the probability P_N does not depend on the diffusion coefficient D but rather on the mean velocity along microtubules.

For a general field, achieving a global minimum inside the domain Ω at isolated points, by using the Laplace method in formula (78), we can approximate the expression for the probability P_N by

$$\int_{\partial\Omega} e^{-\frac{\phi(\boldsymbol{x})}{D}} dS = \begin{cases} e^{-\frac{\phi_m}{D}} \sum_{k=1}^{M} \sqrt{\frac{2D\pi}{\det Hess'(\phi)_k}} & \text{for} \quad n = 2 \\ e^{-\frac{\phi_m}{D}} \sum_{k=1}^{M} \frac{2D\pi}{\sqrt{\det Hess'(\phi)_k}} & \text{for} \quad n = 3, \end{cases} \tag{84}$$

where ϕ_m is the minimum of ϕ achieved on the boundary of the nucleus and $Hess'(\phi)_k$ denoted the Hessian of the restriction of ϕ to the boundary. Similarly,

$$\int_\Omega e^{-\frac{\phi(x)}{D}} k(x)dx = \begin{cases} e^{-\frac{\phi_N}{D}} \sum_{k=1}^{N} k(x_k) \dfrac{2D\pi}{\sqrt{\det Hess(\phi)_k}} & \text{for} \quad n=2, \\ e^{-\frac{\phi_N}{D}} \sum_{k=1}^{N} k(x_k) \dfrac{(2D\pi)^{3/2}}{\sqrt{\det Hess(\phi)_k}} & \text{for} \quad n=3, \end{cases} \tag{85}$$

where ϕ_N is the minimum ϕ achieved inside the cell at the interior points $x_1, \ldots, x_N$. When the minimum of ϕ is achieved on the surface of the nucleus, $\phi_N = \phi_m$ and the right-hand side of expression (85) is multiplied by a factor $1/2$. We leave to the reader to gather the pieces to obtain the fat expression for the probability P_N.

When the condition $B\ Diam \ll D$ is satisfied, we can use a Taylor expansion of the exponential term to get

$$\frac{1}{|\partial\Omega|} \int_{\partial\Omega} e^{-\frac{\phi(x)}{D}} dS = 1 - \frac{\langle\phi\rangle_{\partial\Omega}}{D} + o\left(\frac{\phi}{D}\right)^2 \tag{86}$$

where $\langle\rangle_{\partial\Omega}$ denotes the average over the boundary. For a radial potential and a round cell of radius $R \rangle\rangle \delta$,

$$< \phi >_{\partial\Omega} \approx BR. \tag{87}$$

Thus,

$$P_N \approx \begin{cases} \dfrac{1-\frac{BR}{D}}{\frac{\ln\frac{1}{\eta}}{D\pi}\left(\int_\Omega k(x)dx - \int_\Omega k(x)\phi(x)dx\right)+1-\frac{BR}{D}} & \text{for} \quad n=2, \\ \dfrac{1-\frac{BR}{D}}{\frac{1}{4D\eta}\left(\int_\Omega k(x)dx - \int_\Omega k(x)\phi(x)dx\right)+1-\frac{BR}{D}} & \text{for} \quad n=3. \end{cases} \tag{88}$$

4.2. Mean Conditional Reaching Time to the Nucleus

The mean reaching time τ_N can be computed following the steps of the no drift case (see paragraph 3.3). To estimate the reaching time, defined by expression (29), we use the asymptotic approximation $q(x) = T_\eta e^{-\phi(x)/D}$ of Eq. (25), whereas $\tilde{p}$ is given by expression (70).

The first step consists of obtaining an expression of T_η as a function of P_η: for this, we use relation (75) where p_i is replaced by $P_\eta e^{-\phi(x)/D}$, we then obtain the integral equation

$$D \int_{\partial\Omega_a} G_Q g(S) e^{-\frac{\phi(x)}{D}} dS + \frac{1}{|\partial\Omega|} \int_{\partial\Omega} T_\eta e^{-\frac{\phi(x)}{D}} ds$$

$$= \int_\Omega G_Q(\boldsymbol{x}) \left(e^{-\frac{\phi(\boldsymbol{x})}{D}} (k(\boldsymbol{x}) T_\eta - P_\eta) \right) d\boldsymbol{x}$$
$$- \int_\Omega \nabla G_Q(\boldsymbol{x}).\nabla \phi(\boldsymbol{x}) T_\eta e^{-\frac{\phi(\boldsymbol{x})}{D}} d\boldsymbol{x} \tag{89}$$

and the flux condition (obtained by integrating Eq. (25)) imposes that

$$D \int_{\partial N_a} g(S) dS \approx -P_\eta \int_\Omega e^{-\frac{\phi(\boldsymbol{x})}{D}} d\boldsymbol{x} + T_\eta \int_\Omega k(\boldsymbol{x}) e^{-\frac{\phi(\boldsymbol{x})}{D}} d\boldsymbol{x}. \tag{90}$$

Using the flux expression (76), we obtain an expression for the constant T_η, in dimension 2:

$$T_\eta \approx \frac{P_\eta \frac{\ln \frac{1}{\eta}}{D\pi} \int_\Omega e^{-\frac{\phi(\boldsymbol{x})}{D}} d\boldsymbol{x}}{\frac{\ln \frac{1}{\eta}}{D\pi} \int_\Omega e^{-\frac{\phi(\boldsymbol{x})}{D}} k(\boldsymbol{x}) d\boldsymbol{x} + \int_{\partial\Omega} e^{-\frac{\phi(\boldsymbol{x})}{D}} \frac{dS}{|\partial\Omega|} + \int_\Omega \nabla G_Q(\boldsymbol{x}).\nabla \phi(\boldsymbol{x}) e^{-\frac{\phi(\boldsymbol{x})}{D}} d\boldsymbol{x} + \int_\Omega e^{-\frac{\phi(\boldsymbol{x})}{D}} k(\boldsymbol{x}) d\boldsymbol{x}}$$

and in dimension 3,

$$T_\eta \approx \frac{P_\eta \frac{1}{4D\eta} \int_\Omega e^{-\frac{\phi(\boldsymbol{x})}{D}} d\boldsymbol{x}}{\frac{1}{4D\eta} \int_\Omega e^{-\frac{\phi(\boldsymbol{x})}{D}} k(\boldsymbol{x}) d\boldsymbol{x} + \int_{\partial\Omega} \frac{e^{-\frac{\phi(\boldsymbol{x})}{D}} dS}{|\partial\Omega|} + \int_\Omega \nabla G_Q(\boldsymbol{x}).\nabla \phi(\boldsymbol{x}) e^{-\frac{\phi(\boldsymbol{x})}{D}} d\boldsymbol{x} + \int_\Omega e^{-\frac{\phi(\boldsymbol{x})}{D}} k(\boldsymbol{x}) d\boldsymbol{x}}.$$

Using the definition (29) of the time τ_N, we get that

$$\tau_N \approx \frac{P_\eta}{P_N} \left(\int_\Omega e^{-\frac{\phi(\boldsymbol{x})}{D}} d\boldsymbol{x} - \frac{T_\eta}{P_\eta} \int_\Omega e^{-\frac{\phi(\boldsymbol{x})}{D}} k(\boldsymbol{x}) d\boldsymbol{x} \right). \tag{91}$$

In the approximation where the killing rate is small (integrals in k are negligible compared to one), we obtain by using the asymptotic expressions (78) for the probability P_N, the following asymptotic formulas for the reaching time:

$$\tau_N \approx \begin{cases} \dfrac{\frac{\ln \frac{1}{\eta}}{D\pi} \int_\Omega e^{-\frac{\phi(\boldsymbol{x})}{D}} d\boldsymbol{x}}{\frac{\ln \frac{1}{\eta}}{D\pi} \int_\Omega e^{-\frac{\phi(\boldsymbol{x})}{D}} k(\boldsymbol{x}) d\boldsymbol{x} + \frac{1}{|\partial\Omega|} \int_{\partial\Omega} e^{-\frac{\phi(\boldsymbol{x})}{D}} dS} & \text{for} \quad n = 2, \\ \dfrac{\frac{1}{4D\eta} \int_\Omega e^{-\frac{\phi(\boldsymbol{x})}{D}} d\boldsymbol{x}}{\frac{1}{4D\eta} \int_\Omega e^{-\frac{\phi(\boldsymbol{x})}{D}} k(\boldsymbol{x}) d\boldsymbol{x} + \frac{1}{|\partial\Omega|} \int_{\partial\Omega} e^{-\frac{\phi(\boldsymbol{x})}{D}} dS} & \text{for} \quad n = 3. \end{cases} \tag{92}$$

In addition, when the diffusion coefficient D is small and the potential ϕ achieves its minimum uniformly on the boundary, by using Laplace's method, when the killing rate k is locally constant near the boundary, we obtain

$$\tau_N \approx \begin{cases} \dfrac{\frac{2\ln \frac{1}{\eta}}{D} \left(\frac{D}{B}\delta + 2\left(\frac{D}{B}\right)^2 \right)}{k \frac{2\ln \frac{1}{\eta}}{D} \left(\frac{D}{B}\delta + 2\left(\frac{D}{B}\right)^2 \right) + 1} & \text{for} \quad n = 2, \\ \dfrac{\frac{\pi}{\eta D} \left(\frac{D}{B}\delta^2 + 2\left(\frac{D}{B}\right)^2 \delta + 2\left(\frac{D}{B}\right)^3 \right)}{\frac{\pi}{D\eta} \left(\frac{D}{B}\delta^2 + 2\left(\frac{D}{B}\right)^2 \delta + 2\left(\frac{D}{B}\right)^3 \right) k + 1} & \text{for} \quad n = 3. \end{cases} \tag{93}$$

It is interesting to note that when the diffusion constant is small, the mean reaching time depends mainly on the time spent in a small region around the nucleus. When $\frac{B\delta}{D} \ll 1$, other limit formulas can be obtained, such as

$$\tau_N \approx \begin{cases} \dfrac{\frac{\ln\frac{1}{\eta}}{D\pi}\left(|\Omega| - \int_\Omega \frac{\phi(\boldsymbol{x})}{D} d\boldsymbol{x}\right)}{\frac{\ln\frac{1}{\eta}}{D\pi}\left(\int_\Omega k(\boldsymbol{x})d\boldsymbol{x} - \int_\Omega \frac{\phi(\boldsymbol{x})}{D}k(\boldsymbol{x})d\boldsymbol{x}\right) + 1 - \frac{1}{|\partial\Omega|}\int_{\partial\Omega}\frac{\phi(\boldsymbol{x})}{D}dS} & \text{for} \quad n = 2, \\ \dfrac{\frac{1}{4D\eta}\left(|\Omega| - \int_\Omega \frac{\phi(\boldsymbol{x})}{D} d\boldsymbol{x}\right)}{\frac{1}{4D\eta}\left(\int_\Omega k(\boldsymbol{x})d\boldsymbol{x} - \int_\Omega \frac{\phi(\boldsymbol{x})}{D}k(\boldsymbol{x})d\boldsymbol{x}\right) + 1 - \frac{1}{|\partial\Omega|}\int_{\partial\Omega}\frac{\phi(\boldsymbol{x})}{D}dS} & \text{for} \quad n = 3. \end{cases} \tag{94}$$

In the case of n well separated small holes, formula (83) and (93) are modified as we did in paragraph 3.3.1 (replacing $\frac{1}{4D\eta}$ by $\frac{1}{(4nD\eta)}$ in dimension 3 and $\ln\frac{1}{\eta}$ by $(\frac{1}{n})\ln(\frac{1}{\eta})$ by in dimension 2).

We now propose to use these formula for numerical evaluations. We consider data coming from the Associated Adeno Virus (AAV), where the diffusion constant is $D \approx 1.3\,\mu\text{m}^2/\text{s}$ and the mean drift is in a range 1.8–3.7 μm/s.[17] Because the effective drift of a virus toward the nucleus is about 5 to 10 percent of the mean drift, we chose, close to the nucleus, a mean effective drift value of $B = 2.5/20 = 0.12\,\mu$m/s.[23] We choose also a killing $k = 1/3600\,\text{s}^{-1}$. The cell characteristics are identical to the case of a plasmid. We obtain that the reaching probability is given by

$$P_N = 0.98 \quad \text{for} \quad n = 2 \tag{95}$$

$$P_N = 0.70 \quad \text{for} \quad n = 3, \tag{96}$$

while the mean reaching time is given by

$$\tau_N = 66\,s \quad \text{for} \quad n = 2 \tag{97}$$

$$\tau_N = 1056\,s \quad \text{for} \quad n = 3. \tag{98}$$

The probability to reach the nucleus is higher for viruses compared to plasmids, which is a consequence here of the drift effect. However, the reaching time in dimension two and three is very different for both plasmids and viruses. For viruses, the difference between dimensions 2 and 3 is observed mainly for the reaching time τ_N, which is multiplied by a factor 16.

5. DISCUSSION AND BIOLOGICAL IMPLICATIONS

We provided here a general theoretical framework to study viral and DNA trafficking in a biological cell. While the viral particle movement is approximated by a Markovian stochastic equation, the degradation activity is modeled by a steady state killing rate. Under the assumption that the killing activity is small compared to the diffusion rate, we provided here explicit asymptotic formula of

the probability and the mean time a virus or a plasmid DNA reaches a small nuclear pore.

Various assumptions need to be discussed: first, contrary to what we assumed, the location of the centrosome, where microtubules converge, does not exactly coincide with the nucleus. Thus a virus can reach a nuclear pore only if it detaches from the microtubules. However, the process by which trafficking molecules located in a neighborhood of the centrosome are detached from microtubules and move near the nucleus, is not clear. At this stage of the model, since we have not made any distinctions between the centrosome and the nucleus, we neglected the transport from the centrosome to the nucleus. However, to interpret our numerical result about the time a virus hits a nuclear pore, we observe that formula 93 accounts mainly for the time spent close to the nucleus surface, inside a boundary layer of size $\frac{D}{B}$. This boundary layer is generated by the competition between the diffusion and the drift. Second, the diffusion process is modeled here in the continuum limit: it might also approximate a discreet diffusion process along microtubules, where a trafficking molecule may jump randomly from one to another microtubule at an intersection. Third, the killing field which represents the effect of enzymatic activity may also represent the direct ubiquitination process leading to the proteasome degradation. For a virus, the degradation may happen only in a neighborhood of the nucleus, where it is partially uncoated, while the degradation may be homogeneously active in all parts of the cell for a naked DNA.

When many viruses are involved, the probability P_N represents the fraction of viruses reaching the nucleus between the initial time and a large time (infinity), while τ_N is the mean time it takes for the viral fraction to reach the nucleus. The computation of the mean time the first virus reaches the nucleus is more involved and is postponed to a future article.

Finally, the present asymptotic computations can be used to estimate the effect of disrupting the microtubule network on the mean reaching time. By using a microtubule disrupting drugs, the delivery process will be affected and it can be modeled here by reducing the value of the drift B. As a consequence, the probability to reach a nuclear pore (given by formula (78)) is reduced, leading the way to a killing activity that might be enough to degrade viral DNA. In that case, the reaching time changes according to formula (93). The numerical results presented here suggest that viruses are more efficient, than plasmids in reaching the nucleus in agreement with experimental results. According to our analysis, successful viruses spent most of their time in a neighborhood of the nucleus, while plasmids keep exploring the cell volume.

ACKNOWLEDGMENTS

I would like to thank E. Dauty for discussions and comments on this work.

REFERENCES

1. E. Dauty and A. S. Verkman, Actin cytoskeleton as the principal determinant of size-dependent DNA mobility in cytoplasm: a new barrier for non-viral gene delivery. *J. Biol. Chem.* **280**(9):7823–7828 (2005).
2. P. G. De Gennes, *Scaling Concepts in Polymer Physics* (Cornell University Press, Ithaca, NY, 1979).
3. A. T. Dinh, T. Theofanous, and S. Mitragotri, A model for intracellular trafficking of adenoviral vectors. *Biophys. J.* **89**(3):1574–1588 (2005).
4. W. Ding, L. Zhang, Z. Yan, and J. F. Engelhardt, Intracellular trafficking of adeno-associated viral vectors. *Gene Ther.* **12**:873–880 (2005).
5. K. Dohner, C. H. Nagel, and B. Sodeik, Viral stop-and-go along microtubules: taking a ride with dynein and kinesins. *Trends Microbiol.* **13**(7):320–327 (2005).
6. P. R. Garabedian, *Partial Differential Equations* (John Wiley Sons, Inc., New York-London-Sydney, 1964).
7. U. F. Greber and M. Way, A superhighway to virus infection. *Cell* **124**(4):741–754 (2006).
8. N. Hirokawa, Kinesin and dynein superfamily proteins and the mechanism of organelle transport. *Science* **279**(5350):519–526 (1998).
9. D. Holcman, A. Marchevska, and Z. Schuss, The survival probability of diffusion with trapping in cellular biology. *Phys. Rev. E Stat, Nonlin. Soft. Matter Phys.* **72**:031910 (2005).
10. D. Holcman and Z. Schuss, Escape through a small opening: receptor trafficking in a synaptic membrane. *J. Stat. Phys.* **117**(5–6):975–1014 (2004).
11. T. Lagache and D. Holcman, Effective drift of a virus trafficking inside a biological cell, pre-print.
12. D. Lechardeur, Metabolic instability of plasmid DNA in the cytosol: a potential barrier to gene transfer. *Gene Ther.* **6**:482–497 (1999).
13. G. G. Maul and L. Deaven, Quantitative determination of nuclear pore complexes in cycling cells with differing DNA content. *J. Cell. Biol.* **73**(3):748–760 (1977)
14. P. Palese, Influenza: Old and new threats. *Nat Med.* **10**(12 Suppl):S82-87 (2004).
15. J. F. Cros and P. Palese, Trafficking of viral genomic RNA into and out of the nucleus: influenza, Thogoto and Borna disease viruses. *Virus Res.* **95**(1–2):3–12 (2003).
16. Z. Schuss, *Theory and Applications of Stochastic Differential Equations*, Wiley Series in Probability and Statistics (John Wiley & Sons, Inc., New York, 1980).
17. G. Seisenberger, M. U. Ried, T. Endress, H. Buning, M. Hallek, and C. Brauchle, Real-time single-molecule imaging of the infection pathway of an adeno-associated virus. *Science* 294(5548):1929–1932 (2001).
18. A. Singer, Z. Schuss, D. Holcman, and R. S. Eisenberg, Narrow escape. I. *J. Stat. Phys.* **122**(3):437–463 (2006).
19. A. Singer, Z. Schuss, and D. Holcman, Narrow escape. II. The circular disk. *J. Stat. Phys.* **122**(3):465–489 (2006).
20. A. Singer, Z. Schuss, and D. Holcman, Narrow escape. III. Non-smooth domains and Riemann surfaces. *J. Stat. Phys.* **122**(3):491–509 (2006).
21. A. Singer and Z. Schuss, Activation through a narrow opening. *Phys. Rev. E Stat. Nonlin. Soft. Matter Phys.* **74**:020103 (2006).
22. G. A. Smith and L. W. Enquist, Break ins and break outs: viral interactions with the cytoskeleton of Mammalian cells. *Annu. Rev. Cell. Dev. Biol.* **18**:135–161 (2002).
23. M. Suomalainen, M. Y. Nakano, S. Keller, K. Boucke, R. P. Stidwill, and U. F. Greber, Microtubule-dependent plus- and minus end-directed motilities are competing processes for nuclear targeting of adenovirus. *J. Cell Biol.* **144**(4):657–672 (1999).
24. G. Zuber, E. Dauty, M. Nothisen, P. Belguise, and J. P. Behr, Towards synthetic viruses. *Adv. Drug Deliv. Rev.* **52**:245–253 (2001).