

UNIVERSITÉ PIERRE ET MARIE CURIE

N° attribué par la bibliothèque

--	--	--	--	--	--	--	--	--	--

THÈSE

pour obtenir le grade de

DOCTEUR de l'UPMC

Spécialité : **Mathématiques Appliquées**

préparée au laboratoire **UMR8542**

dans le cadre de l'École Doctorale **École Doctorale de Sciences
Mathématiques de Paris Centre**

présentée et soutenue publiquement

par

Thibault Lagache

le 3 Décembre 2009

Titre:

**Modeling the early steps of viral infection: a stochastic
approach**

Une approche stochastique en virologie quantitative

Directeur de thèse: **David Holcman**

Jury

M. Charles Tier,	Rapporteur
M. Michael J. Ward,	Rapporteur
M. Habib Ammari,	Examineur
M. Olivier Danos,	Examineur
M. Ivan Kupka,	Examineur
M. Marc Yor,	Président
M. David Holcman,	Directeur de thèse

Résumé

L'entrée du virus dans une cellule est un mécanisme complexe qui commence par l'attachement du virus à des récepteurs spécifiques à la surface de la cellule pour finir quand le virus délivre son ADN dans le noyau au travers d'un pore nucléaire. En particulier, la plupart des virus sont endocytés dans des endosomes pour être plus tard relâchés dans le cytoplasme. Leur mouvement alterne alors entre diffusion libre et transport actif le long des microtubules jusqu'à atteindre un pore nucléaire. L'étape endosomale et le déplacement libre dans le cytoplasme limitent particulièrement le transfert de gènes dans le noyau des cellules en thérapie génique et la capacité des virus à délivrer leur ADN avec succès indique que ceux-ci ont développé au cours de l'évolution des outils moléculaires efficaces pour franchir ces deux étapes précoces de l'infection. Pour mieux comprendre les mécanismes sous-jacents à l'efficacité virale, nous construisons des modèles stochastiques à l'échelle moléculaire pour estimer le temps moyen d'échappée de l'endosome et la probabilité qu'un virus relâché dans le cytoplasme atteigne un pore nucléaire avant d'être dégradé, en fonction de la géométrie cellulaire et des paramètres structuraux et dynamiques du virus. L'échappée endosomale est provoquée par le changement de conformation pH-dépendant de protéines actives à la surface du virus et nous développons donc, dans un premier temps, un modèle Markovien à sauts afin d'étudier la cinétique de ce changement de conformation à pH donné. Notre modèle théorique permet d'interpoler les cinétiques mesurées expérimentalement pour l'hémagglutinine de la grippe à différents pH. Dans un second temps, modélisant l'entrée des protons dans le cytoplasme par un processus de Poisson, nous calculons la moyenne et la variance du temps de l'échappée endosomale. En particulier, nous calculons que le temps moyen de l'échappée du virus adéno-associé, un virus très utilisé en thérapie génique, est de 20 minutes, indiquant que le virus s'échappe depuis l'endosome tardif. Afin d'obtenir des résultats asymptotiques généraux concernant le mouvement libre du virus dans le cytoplasme, nous modélisons la dynamique intermittente du virus par une équation continue de Langevin pour la vitesse contenant à la fois un terme de diffusion et un terme de dérive prenant en compte les périodes de transport actif le long des microtubules, et nous présentons une procédure générale pour calibrer ce terme de dérive en fonction de l'organisation du réseau de microtubules, de la constante de diffusion du virus dans le cytoplasme et de la constante d'affinité entre le virus et les microtubules. Grâce à cette description continue de la vitesse du virus, calculer la probabilité et le temps moyen de premier passage du virus à un petit pore nucléaire devient un problème d'"échappée belle". En particulier, du fait que les noyaux des cellules sont recouverts d'un grand nombre de pores nucléaires, nous étendons certains résultats asymptotiques de la théorie de l'"échappée belle" afin de mesurer l'impact de l'interaction entre les différents pores sur la probabilité et le

temps moyen de premier passage du virus à un pore. Utilisant des données expérimentales, nous calculons en particulier que le temps moyen de premier passage du virus adéno-associé à est d'environ 3 minutes, ce qui est en accord avec les observations *in vivo*, et que la probabilité que celui-ci atteigne le pore nucléaire avant d'être dégradé est d'environ 95%, confirmant l'efficacité du virus à délivrer son ADN au noyau.

Abstract

At the cellular level, viral entry is a complex multi-step process that starts from the binding of the virus to specific cellular receptors and ends when the virus delivers its genetic material in the nucleus through a nuclear pore. Most viruses are internalized in endosomes to be later released in the cytoplasm where they alternate between diffusion and ballistic periods along the microtubules network until they reach a nuclear pore. The endosomal step and the free trafficking in the cytoplasm to a nuclear pore particularly limit the genes delivery in gene therapy, and the high infection probability of viruses suggest that they have developed efficient tools in these two steps. Yet the principles underlying these tools are still largely unknown and to unravel viral mechanisms underlying their infection success, we construct stochastic models at a molecular level to estimate the mean escape time from the endosome and the probability that a virus released in the cytoplasm reaches a nuclear pore before being degraded, as a function of the geometry of the cell and some dynamical parameters of the virus. The endosomal escape of viruses is triggered by the pH-dependent conformational change of viral active proteins and, in a first time, we develop a Markov jump model to study the kinetics of the conformational change at a given pH. Our theoretical model interpolates nicely the reported kinetics of the influenza hemagglutinin at several pH. Then, modeling the protons entry in the endosome with a Poissonian process, we compute the mean and the variance of the viral escape time. Interestingly, we found that the mean escape time of the adeno-associated virus, an extensively used virus in gene therapy, is around 20 minutes, suggesting that the virus should escape from the late endosome. To obtain general asymptotic results for the free cytoplasmic step, we first coarse grain the complex intermittent dynamics of the virus into a continuous Langevin equation for the viral velocity that contains both a diffusion term and a drift term that accounts for the active transport along the microtubules, and we provide a general calibration procedure to compute the drift amplitude as a function of the microtubules network organization, the diffusion constant of the virus and its unbinding rate from a microtubule. Using that continuous Langevin description, reaching a small nuclear pore reduces to a narrow escape problem and in particular, because biological cells contain many holes, we extend some previous asymptotic results to quantify the impact of the interactions between the absorbing windows on the probability and the conditional mean first passage time of a virus to a pore. Using biological data, we found that the mean first passage time of an adeno-associated virus to a nuclear pore is approximately equal to 3 minutes, which is coherent with the experimental observations, and that the probability the virus reaches a pore alive is equal to 95%, which confirms the efficiency of the viral particle to deliver its DNA through a nuclear pore.

Remerciements

Tout d'abord, je tiens à remercier David Holcman pour sa disponibilité et pour m'avoir transmis sa curiosité et sa passion scientifiques. Je remercie également Juer-gen Reingruber pour nos discussions et sa relecture constructive de cette thèse, ainsi que tous les autres membres du laboratoire : Godefroy, Stavros, Carlo, Khanh, Nathanaël, Assaf et Diana qui participent au dynamisme et à la bonne ambiance au sein du groupe. Je tiens également à remercier Olivier Danos avec qui nous essayons de comprendre, petit à petit, la vie d'un virus dans une cellule . . . Je remercie également M. Yor pour avoir accepté de présider mon jury de soutenance et pour nos discussions très intéressantes et constructives, ainsi que M. Tier, M. Ward, M.Ammari et M.Kupka pour m'avoir fait l'honneur de faire partie de mon jury. Je remercie également les personnes qui m'ont peu à peu orienté vers cette carrière scientifique : M. Capelle qui m'a initié aux mathématiques, M. Porte et M. Balabane qui m'ont encouragé à commencer ce doctorat après un très court passage en tant qu'analyste quantitatif au Crédit Agricole, ainsi que M. Cancès pour avoir toujours été de très bon conseil. Enfin je remercie Nina pour m'avoir accompagné dans mes longues journées de recherches infructueuses, ainsi que Jérôme Martin pour me vider un peu la tête. . .

À Émeline ...

Table des matières

Résumé	iii
Abstract	v
Remerciements	vii
Table des matières	ix
Introduction (français)	3
1 Le cycle de réplication des virus à ADN	4
2 Modéliser les étapes précoces de l'infection virale	5
3 Le mouvement libre dans le cytoplasme	7
3.1 Calcul du terme de dérive dû au transport le long des micro- tubules	8
3.2 Analyse quantitative du mouvement libre du virus dans le cytoplasme	14
4 L'étape endosomale	18
4.1 Le modèle de changement conformationnel	19
4.2 Modèle de sortie endosomale et résultats	20
5 Conclusion	22
Introduction	23
1 Viral structure	23
2 Application of viruses in gene therapy	25
3 Early steps of viral infection	26
4 Modeling the early steps of viral infection	30
5 Models and Results	32
5.1 The endosomal step	32
5.2 The free cytoplasmic step	35
6 Conclusion	46
1 Effective motion of a virus trafficking inside a biological cell, <i>SIAM Journal of Applied Mathematics</i> 68 (2008)	49
1 Introduction	49
2 Modeling stochastic viral movement inside a biological cell	50
2.1 Modeling viral dynamics in the cytoplasm	51
2.2 Computing the MFPT to reach the nucleus	53
3 Mean First Passage Time and Exit point distribution	54
3.1 The general solution for the MFPT	55
3.2 Exit points distribution	56
3.3 Exit pdf in a Pie Wedge	57
3.4 The Mean Exit Radius (MER)	61

4	Approximation of a virus motion by an effective Markovian stochastic equation	63
4.1	Methodology	63
4.2	Explicit expression of the drift in the limit of $\Theta \ll 1$	65
4.3	Justification of the MFPT-criteria.	66
5	Conclusion	68
2	Quantifying intermittent transport in cell cytoplasm, <i>Phys. Rev. E</i> 77 (2008)	73
1	Introduction	73
2	Mathematical Modeling	74
3	General Methodology	75
4	The steady state drift for a two-dimensional radial cell	75
5	Tests against Brownian simulations	78
6	Computation of P_N and τ_N	78
7	Conclusion	80
3	Mean time and probability to reach a structured target	83
1	Introduction.	83
2	Asymptotic derivations of P_n and τ_n	86
2.1	Asymptotic derivation of P_n	86
2.2	Asymptotic derivation of τ_n	91
3	Tests against Brownian simulations	92
4	Conclusion	92
5	Appendix	94
4	Quantitative analysis of virus and plasmid trafficking in cells, <i>Phys. Rev. E</i> 79 (2009)	97
1	Introduction	97
2	Modeling intracellular viral and DNA trafficking	98
3	Distribution of degraded DNA carriers	105
4	Impact of the degradation density distribution	105
5	Mean first passage time of the first DNA carrier to a nuclear pore. . .	106
6	The large degradation rate limit	110
7	Conclusion	114
8	Appendix	114
5	The endosomal step of viral infection	121
1	Introduction	121
2	Results	122
3	Discussion	131
4	Appendix	133
4.1	Computation of $u(x_0(c))$	133
4.2	Computation of $\langle c_{\tau_e} \rangle$	136
4.3	Computation of $cm(t)$	137
4.4	Computation of the mean escape time $\bar{\tau}_e$ and the associated variance $var(\tau_e)$	139

Conclusion	143
Appendix : Brownian simulations	147
Bibliographie	153

PUBLICATIONS

1. Thibault Lagache and David Holcman, *Effective Motion of a Virus Trafficking Inside a Biological Cell*, SIAM J. Of Applied Maths **68**, 1146 (2008).
2. Thibault Lagache and David Holcman, *Quantifying Intermittent Transport in Cell Cytoplasm*, Phys. Rev. E **77**, 030901 (2008).
3. Thibault Lagache, Emmanuel Dauty and David Holcman, *Quantitative Analysis of virus and plasmid trafficking in cell*, Phys. Rev. E **79**, 011921 (2009).
4. Thibault Lagache, Emmanuel Dauty and David Holcman, *Physical principles and models describing intracellular virus particle dynamics*, Curr Opin Microbiol. **12**, 439 (2009). Review.
5. Thibault Lagache, Olivier Danos and David Holcman, *The endosomal step of viral infection*, Submitted.

ORAL PRESENTATIONS AND OTHERS

1. Thibault Lagache and David Holcman, *Quantifying Intermittent Transport in Cell Cytoplasm*. SIAM Annual Meeting, San Diego, U.S.A. (2008) Contributed Talk.
2. Thibault Lagache and David Holcman, *Quantifying Intermittent Transport in Cell Cytoplasm with Application to Viral Infection Modeling*. SIAM Conference on Dynamical Systems, Salt Lake City, U.S.A. (2009) Minisymposium Speaker.
3. Thibault Lagache, *Quand la guêpe finit par s' échapper*, Découverte **356**, 48 (2008). Popular scientific paper.

POSTERS

1. Thibault Lagache, Emmanuel Dauty and David Holcman, *Modeling Virus Trafficking in the Cytoplasm*. Biophysical Society Annual Meeting, Baltimore, U.S.A. (2007).
2. Thibault Lagache and David Holcman, *Quantitative Analysis of Viral Infection*. Early Steps of the Virus Life Cycle : Molecular and Cellular Insights. Institut Pasteur. Paris, France. (2007).
3. Thibault Lagache, Olivier Danos and David Holcman, *Modeling the Endosomal Step of Viral Infection* Biophysical Society Annual Meeting. Boston, U.S.A. (2009).

Introduction (français)

Les virus utilisent les cellules pour se répliquer et se propager. Étant facilement produits en large quantité, les virus sont de formidables outils pour explorer les mécanismes cellulaires comme l'endocytose ou pour transporter l'ADN en thérapie génique. Même si les différentes étapes du cycle virale sont assez bien comprises, beaucoup de questions importantes restent en suspens. En particulier, l'impact de l'environnement complexe de la cellule sur les voies d'infection est encore mal compris. Pour de nombreux virus, des modèles moléculaires *in vitro* existent pour décrire les changements structuraux du virus lors des différentes étapes de son cycle de répllication et des nouvelles techniques d'imagerie sont développées pour confirmer ou infirmer ces modèles dans l'environnement plus complexe de la cellule. De manière très surprenante, une grande hétérogénéité du comportement viral est observée, révélant des interactions complexes entre la cellule et la structure moléculaire de son hôte. En effet, à chaque étape de l'infection, le virus subit des changements structuraux sous l'influence de l'environnement cellulaire et une part stochastique dans ces métamorphoses conduit à des voies d'infection très différentes, avec des réussites inégales.

L'infection virale est un mécanisme complexe qui commence quand le virus s'attache à la cellule pour finir une fois que celui-ci a délivré son matériel dans le noyau *via* un petit pore nucléaire. Après s'être lié à un récepteur spécifique, la plupart des virus à ADN sont endocytés dans des endosomes pour être plus tard relâchés dans le cytosol. Une fois libres, ils alternent alors entre diffusion libre et transport actif le long des microtubules (MTs) jusqu'à atteindre un pore nucléaire et délivrer leur ADN. Dans l'endosome, les virus peuvent être dégradés par des protéases alors qu'ils peuvent être digérés par des protéasomes ou piégés dans le cytoplasme encombré d'organelles. Pour ces étapes limitant le taux d'infection, nous construisons des modèles biophysiques à l'échelle moléculaire afin de calculer différents paramètres comme le temps de sortie moyen de l'endosome ou encore la probabilité et le premier temps de passage conditionnel pour qu'un virus atteigne un pore nucléaire avant d'être dégradé ou piégé dans le cytoplasme. Pour l'étape endosomale, nous considérons que le virus s'échappe quand un nombre suffisant de ligands se lie à une de ses protéines actives et nous utilisons donc dans le cinquième chapitre des chaînes de Markov à sauts et des solutions asymptotiques pour les équations de Kramers-Moyal afin d'obtenir la cinétique de sortie du virus. Pour obtenir des résultats asymptotiques généraux concernant l'étape cytoplasmique, nous réduisons, dans un premier temps, la dynamique intermittente complexe du virus en une équation de Langevin pour la vitesse contenant à la fois un terme de diffusion et un terme de dérive comptant pour les périodes ballistiques le long des MTs. Les deux premiers chapitres présentent la procédure de calibration du terme de dérive à partir de la géométrie

de la cellule (organisation du réseau de MTs ...) et de paramètres dynamiques du virus (coefficient de diffusion ...). Les pores nucléaires occupant une faible fraction de l'enveloppe du noyau, le calcul de la probabilité et du temps de passage conditionnel du virus à un pore nucléaire devient alors un problème d'"échappée belle". Le troisième chapitre est consacré à l'étude de l'impact de l'interaction entre les nombreux pores nucléaires sur la probabilité et le temps de passage conditionnel, alors que le quatrième chapitre présente quelques nouveaux résultats comme le premier temps de passage du premier virus quand plusieurs virus indépendants infectent une même cellule. Pendant les étapes précoces de l'infection, les virus communiquent en permanence avec l'environnement cellulaire et subissent différents changements structuraux. Ces interactions impactent le comportement du virus et un bruit stochastique au cours des différentes métamorphoses affecte de manière conséquente la voie infectieuse empruntée et donc le succès du virus à se répliquer et à se propager. Pour tenir compte de cette variabilité, nous avons développé des modèles biophysiques pour chacune des étapes précoces de l'infection et ensembles, ces modèles fournissent un cadre mathématique pour modéliser l'infection dans sa globalité.

1 Le cycle de réplication des virus à ADN

La plupart des virus à ADN diffusent dans la matrice extracellulaire avant de se lier à des récepteurs spécifiques à la surface des cellules (voir figure 1). Les récepteurs sont spécifiques à l'espèce virale et même parfois au sérotype (voir la revue de Engelhardt concernant les virus adéno-associés (AAV) [1]). La plupart d'entre eux sont ensuite internalisés dans le cytoplasme de la cellule via une endocytose clathrine dépendante. Ils entrent alors dans le cytoplasme dans des endosomes, qui sont des vésicules de routage vers les différents compartiments cellulaires. La famille des protéines Rab joue un rôle essentielle dans ce triage [1]. L'une des voies les plus empruntées par les virus semble être la voie Rab5-Rab7 (voir [2] pour plus de détails) : les virus entrent dans des endosomes précoces qui mûrissent progressivement pour devenir des endosomes tardifs et finalement fusionner avec les lysosomes, organites assurant la digestion intra-cellulaire. Tout au long de leur maturation, les vésicules endosomales s'acidifient grâce à des pompes à protons ATPases situées sur leur membrane. Les virus doivent alors s'échapper de l'endosome avant d'être dégradés par des protéases pH dépendantes. Bien que le processus de sortie des endosomes est fondamental, il est encore assez mal compris mais il semblerait que la baisse de pH provoque le changement conformationnel des protéines de pénétration de la capsid des virus nus ou des glycoprotéines de la membrane des virus enveloppés initiant la sortie du virus de l'endosome. Une fois sortis de l'endosome, les virus se déplacent alors librement jusqu'à atteindre un pore nucléaire et importer leur ADN. Les vecteurs viraux et synthétiques ont tous deux à se déplacer jusqu'à trouver un pore nucléaire, ce qui est une des étapes limitantes du transfert de gènes. En effet, les déplacements dans le cytoplasme de gros vecteurs sont très fortement limités par des barrières tant physiques qu'électrostatiques [3, 4]. Alors que les molécules de taille inférieure à 500 kDa peuvent encore se déplacer par diffusion libre, les particules plus grosses comme certains virus ou vecteurs synthétiques de gènes ont besoin d'être transportés activement par des moteurs moléculaires le long des microtubules (MTs) [5] (voir figure 1). Les virus sont bien plus efficaces que les vecteurs synthé-

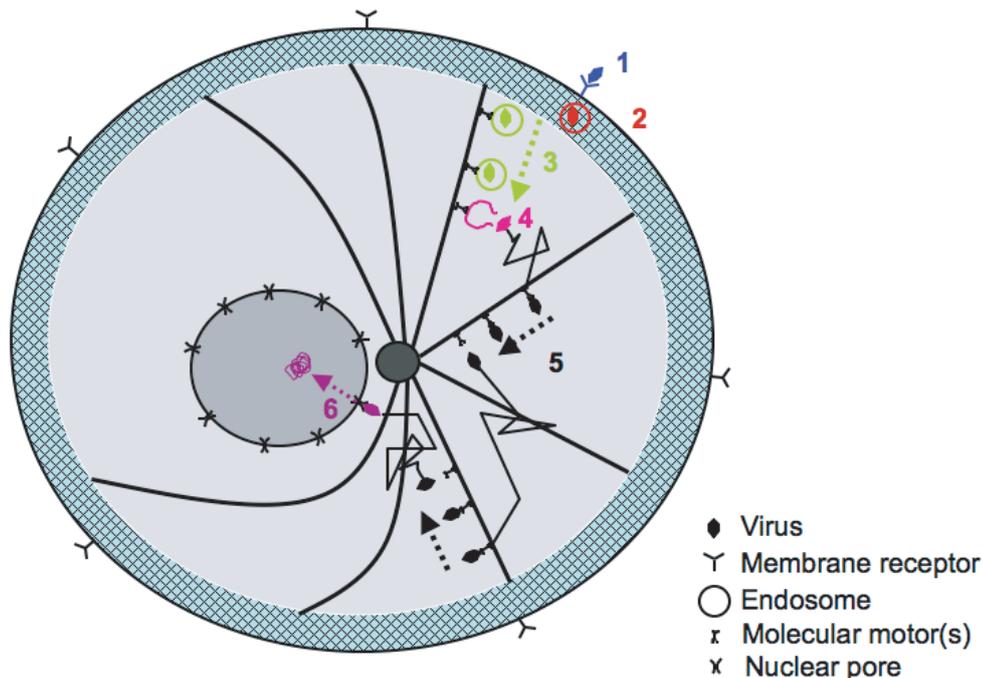


FIGURE 1 – (1) Une particule virale se lie à un récepteur spécifique et est internalisée dans un endosome. (2) Le virus, enfermé dans l’endosome, traverse le réseau d’actine périphérique. (3) Virus et endosome sont activement transportés le long des MTs. (4) Le virus s’échappe de l’endosome. (5) Le virus alterne entre diffusion libre et déplacement actif le long du réseau de MTs. (6) Le virus atteint finalement un pore nucléaire et importe son patrimoine génétique.

tiques à base de polymères ou de lipides : alors qu’un grand nombre ($\approx 100\,000$) de vecteurs synthétiques doit entrer dans le cytoplasme pour que le gène soit exprimé, quelques virus semblent suffire à infecter une cellule. La compréhension et la quantification des différentes étapes de l’entrée virale sont donc essentielles à la conception et à l’optimisation futures des vecteurs de gènes qu’ils soient synthétiques ou viraux.

2 Modéliser les étapes précoces de l’infection virale

Les premiers modèles utilisés pour quantifier le succès du transfert de gènes au niveau cellulaire [6, 7, 8] sont basés sur la loi de conservation entre différents états du virus : le virus peut passer d’un état (enfermé dans l’endosome, diffusant librement dans le cytoplasme ...) à un autre au cours du temps avec une certaine cinétique. Il peut de plus être dégradé. Le problème de ces approches est que les cinétiques de passage entre différents états ne sont pas dérivées à partir de lois biophysiques mais ajustées en fonction des valeurs expérimentales. On ne peut donc pas directement regarder l’impact des différents paramètres dynamiques (constante de diffusion du virus ...) ou géométriques (densité du réseau de MTs ...) sur le succès de l’infection. De plus, elles ne peuvent être utilisés pour analyser le bruit

stochastique intrinsèque à chacune des étapes précoces de l'entrée du virus et ne saisissent donc pas l'individualité structurale des virus au cours du temps. Pour quantifier clairement l'impact des différents paramètres dynamiques et géométriques sur les trajectoires virales, nous proposons donc de modéliser les étapes précoces d'entrée au niveau de la particule virale. Parmi ces étapes nous avons distingué 3 modules fonctionnels au cours desquels les interactions entre la cellule hôte et le virus vont moduler les trajectoires virales et impacter le succès de l'infection. Ces modules peuvent être modélisés séparément par des modèles biophysiques distincts mais interagissent entre eux. En effet, chacun va modifier la structure du virus et donc son comportement dans les modules suivants. Premièrement les virus se lient à des récepteurs spécifiques sur la membrane des cellules et sont ensuite internalisés. Les récepteurs utilisés lors de ce premier module sont les premiers contacts du virus avec son hôte et les changements structurels de la particule commencent alors. Ces changements peuvent ensuite être déterminants pour la voie empruntée par le virus dans le cytoplasme et le succès de l'infection. Ainsi, par exemple, les glycoprotéines de fusion de nombreux virus enveloppés se métamorphosent entièrement au cours de l'entrée virale [9]. En particulier, l'interaction du rétrovirus leukosis aviaire avec son récepteur membranaire permet à certaines des glycoprotéines de son enveloppe de devenir sensibles au pH. Ces protéines sont alors capables de déployer leur activité fusogénique à pH bas [10] et permettent ainsi au virus de sortir de l'endosome avant d'être dégradé dans un lysosome. Dans le cas des AAV, les récepteurs spécifiques des différents sérotypes conduisent à des voies endosomales très variées [1]. Chaque voie endosomale se caractérise par un environnement vésiculaire. Le virus est donc relâché à différents endroits et dans un certain état structural selon la voie endosomale empruntée. De plus, le nombre de virus endocyté dans un même endosome semble être un paramètre important de la cinétique de sortie (voir chapitre 5). Le deuxième module que nous considérons est l'étape endosomale. Bien que la dynamique de sortie puisse être calculée à partir de la cinétique du changement conformationnel de protéines actives (voir chapitre 5), l'endroit et l'état dans lequel le virus sort va dépendre de la voie empruntée et donc de son interaction avec les récepteurs spécifiques à la surface des cellules (premier module). Enfin, L'endroit et le pH de l'endosome quand sort le virus sont des paramètres clés du troisième module : le mouvement libre du virus dans le cytoplasme jusqu'à un pore nucléaire. Par exemple, dans le cas de l'AAV, l'acidité de l'endosome au moment de sa sortie va plus ou moins dénaturer sa capsid, ce qui va ensuite influencer l'ubiquitination de celle-ci une fois dans le cytoplasme [11]. Une ubiquitination qui va accélérer sa dégradation par les protéasomes mais qui, d'autre part, semble aider au démantèlement de la capsid au niveau du pore nucléaire et donc à l'import de l'ADN dans le noyau [11]. Finalement, afin de trouver la voie d'entrée optimale du virus, ces trois modules, avec leur modèles biophysiques distincts, devront être couplés ; l'état du virus à la sortie d'un module servant de paramètre d'entrée au suivant. Par exemple, dans le cas de l'AAV, il serait intéressant de mesurer l'impact du pH à sa sortie de l'endosome sur son ubiquitination et son taux de dégradation par les protéasomes dans le cytoplasme.

3 Le mouvement libre dans le cytoplasme

Les virus ne peuvent se déplacer seuls autrement que par diffusion. Pour traverser la cellule et atteindre un pore nucléaire, ils utilisent donc les mécanismes de transport cellulaire. Le cytoplasme est un milieu très encombré, contenant de nombreux organelles, un cytosquelette et une grande concentration de macromolécules libres [12]. La mobilité d'une particule dans ce milieu va dépendre de nombreux paramètres comme sa taille, sa forme et les nombreuses interactions électrostatiques avec les différents composants cellulaires. Il a été observé que des particules neutres sphérique de taille inférieure à $25nm$ diffusent librement dans le cytoplasme [13]. A partir de $45nm$, leur mouvement est considérablement réduit. De récents progrès en imagerie microscopique permettent à présent de visualiser le comportement *in vivo* de virus et notamment leur trajectoire au sein de la cellule [14, 15, 16]. Il est à présent admis que la plupart des virus alternent entre diffusion libre et transport actif le long des MTs *via* des moteurs moléculaires tels que la kinésine (du noyau vers la périphérie) ou la dynéine (de la périphérie vers le noyau). Ces moteurs sont ATPases, c'est à dire qu'ils tirent leur énergie de la déphosphorylation de l'ATP. De plus, il a été observé que de nombreux cargos sont transportés *in vivo* par plusieurs moteurs de différente polarité. Ainsi des virus comme l'Herpes [17], l'Adénovirus [18] ou encore le HIV [19] se lient à plusieurs dynéines et kinésines et un mouvement bi-directionnel est observé. Il semblerait qu'un mécanisme de régulation favorise tout de même le mouvement vers l'une ou l'autre direction (vers le noyau lors de l'entrée du virus dans la cellule) [20]. L'analyse de ces trajectoires aléatoire peut se traduire en équations stochastiques. La position $\mathbf{X}(t)$ à un temps t de la particule virale est traitée comme un processus stochastique [21, 22] et la dynamique dépend alors des forces exercées sur la particule. Dans le cytoplasme, les nombreuses collisions entre le virus et les autres macromolécules sont modélisées par un bruit blanc gaussien $\sqrt{2D} \frac{d\mathbf{W}}{dt}$, où D est la constante de diffusion du virus et W le mouvement Brownien standard. En l'absence d'autres forces, la vitesse de la particule est régie par l'équation dynamique $\frac{d\mathbf{X}}{dt} = \sqrt{2D} \frac{d\mathbf{W}}{dt}$. Cependant, quand le virus alterne entre diffusion libre et transport actif le long des MTs, sa position $X(t)$ à un temps t est donnée par

$$\dot{\mathbf{X}} = \begin{cases} \sqrt{2D} \dot{\mathbf{w}} & \text{pour une particule libre} \\ \mathbf{V}(s(t)) & \text{pour une particule liée} \end{cases}, \quad (1)$$

avec $s(t)$ une variable décrivant l'état interne du virus et qui tient compte du nombre de kynésines et de dynéines le transportant. \mathbf{V} est la vitesse résultante, qui dépend à la fois de l'organisation du réseau de MTs et du chargement exercé par le virus. La description stochastique intermittente (1) est suffisante pour générer des simulations de Monte Carlo des trajectoires mais ne peut conduire à des résultats théoriques généraux. Dans les deux premiers chapitres, nous présentons donc une procédure mathématique permettant de réduire (1) à une dynamique de Langevin générale où la vitesse $\dot{\mathbf{X}}$ est donnée par la somme d'un terme de dérive $\mathbf{b}(\mathbf{X})$ prenant en compte les périodes ballistiques du virus le long des MTs et d'un terme reprenant les interactions aléatoires du virus avec son environnement $\sqrt{2D} \frac{d\mathbf{W}}{dt}$ (voir figure 2)

$$\frac{d\mathbf{X}}{dt} = \mathbf{b}(\mathbf{X}) + \sqrt{2D} \frac{d\mathbf{W}}{dt}. \quad (2)$$

3.1 Calcul du terme de dérive dû au transport le long des microtubules

Dans le premier chapitre, nous présentons le calcul d'une dérive équivalente constante B dans une cellule radiale 2-dimensionnelle. Dans le deuxième chapitre, nous étendons nos calculs à une dérive radiale-dépendante $b(r)$ et nous considérons de plus le cas d'une géométrie cylindrique représentant les neurites (neurones et dendrites). Pour calculer ces dérivées nous suivons la procédure suivante : nous considérons un virus partant d'une position \mathbf{x}_0 et nous calculons le temps moyen $u(\mathbf{x}_0)$ nécessaire à ce virus pour atteindre un MT par pure diffusion et s'y lier. Nous considérons ensuite qu'il est transporté par les moteurs moléculaires vers le noyau pendant un temps moyen t_m (et donc sur une distance $d_m = Vt_m$) avant d'être relâché à une position finale moyenne \mathbf{x}_f (voir FIG. 3). Pour une diffusion faible $D \ll 1$, nous posons alors

$$\mathbf{b}(\mathbf{x}_0) \approx \frac{|\mathbf{x}_f - \mathbf{x}_0|}{u(\mathbf{x}_0) + t_m} \quad (3)$$

Pour une cellule idéale radiale 2-dimensionnelle, nous considérons le domaine fondamental $\tilde{\Omega}$, défini comme le domaine angulaire d'angle Θ bordé par deux MTs voisines. Nous considérons que les MTs sont uniformément réparties et donc que $\Theta = \frac{2\pi}{N}$, où N est le nombre total de MTs (voir figure 4). Du fait que les MTs soient uniformément répartis, nous considérons que le virus est toujours relâché dans $\tilde{\Omega}$ et étudions donc le mouvement du virus dans $\tilde{\Omega}$. Dans le cytoplasme, le virus bouge par pure diffusion jusqu'à ce qu'il s'attache à un MT qui est à présent la frontière latérale $\partial\tilde{\Omega}_a$ de $\tilde{\Omega}$ (voir figure 4). Nous simplifions le mouvement bidirectionnel du virus le long du MT et considérons qu'une fois lié, le virus est transporté vers le noyau à une vitesse constante V pendant un temps moyen t_m . La longueur moyenne d'une période ballistique est donc de $d_m = Vt_m$. Le virus est ensuite relâché avec un angle uniformément distribué dans $[0; \Theta]$ à une distance fixe du noyau. Une fois le virus libre dans le cytoplasme, le processus de diffusion/transport actif peut de nouveau recommencer. Ce scénario se répète jusqu'à ce que le virus atteigne un pore nucléaire. Le premier temps de passage moyen (PTPM) $u(\mathbf{x} = (r, \theta))$ d'un virus diffusant à partir d'une position initiale $\mathbf{x} = (r, \theta)$ jusqu'à un MT est solution de l'équation de Dynkin [21]

$$D\Delta u(\mathbf{x}) = -1 \text{ pour } \mathbf{x} \in \tilde{\Omega} \quad (4)$$

$$u(\mathbf{x}) = 0 \text{ pour } \mathbf{x} \in \partial\tilde{\Omega}_a$$

$$\frac{\partial u}{\partial \mathbf{n}} = 0 \text{ pour } \mathbf{x} \in \partial\tilde{\Omega}_{ext},$$

où $\partial\tilde{\Omega}_a = \{\theta = 0\} \cup \{\theta = \Theta\}$ et $\tilde{\Omega}_{ext} = \{r = R\}$ (voir figure 4). Pour un virus partant d'un angle θ uniformément distribué dans $[0; \Theta]$, l'espérance du PTPM à un MT est donné par (voir chapitre 1)

$$\bar{u}(r) = \frac{1}{\Theta} \int_{\theta=0}^{\theta=\Theta} u(r, \theta) d\theta = \frac{r^2}{4D} \left(\frac{\tan(\Theta)}{\Theta} - 1 \right) - \sum_{n=0}^{\infty} \frac{16R^{2-\lambda_n} r^{\lambda_n}}{D\Theta^2 \lambda_n^3 (\lambda_n^2 - 4)}, \quad (5)$$

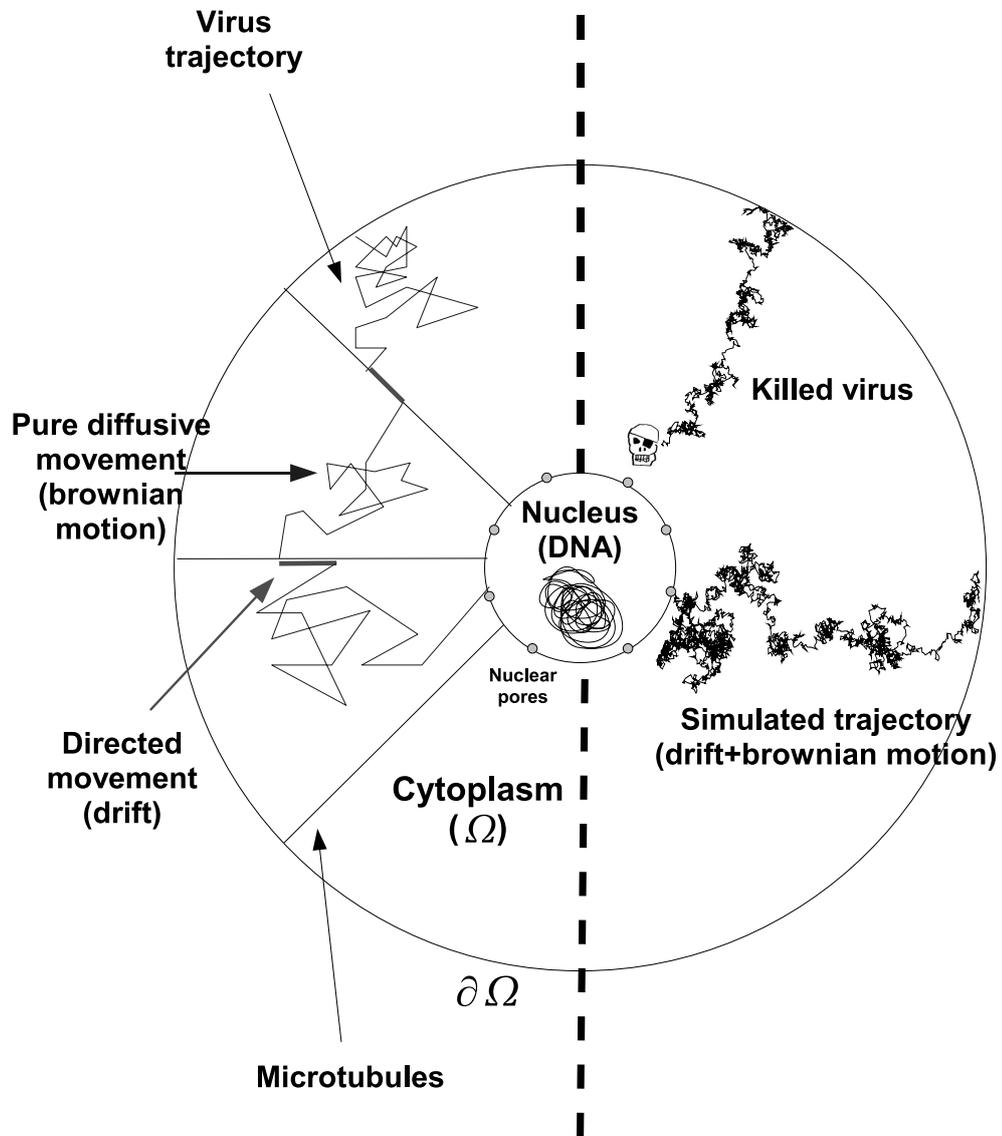


FIGURE 2 – Représentations de la trajectoire d'un virus. Sur le côté gauche de la cellule idéalisée, une trajectoire intermittente réelle reprenant des périodes de diffusion et de mouvement actif le long des MTs est représentée. Sur le côté droit de la cellule, nous avons représenté deux trajectoires simulées à partir de l'équation de Langevin (2). L'une d'entre elles est un succès, le virus parvenant à atteindre un pore nucléaire avant d'être dégradé.

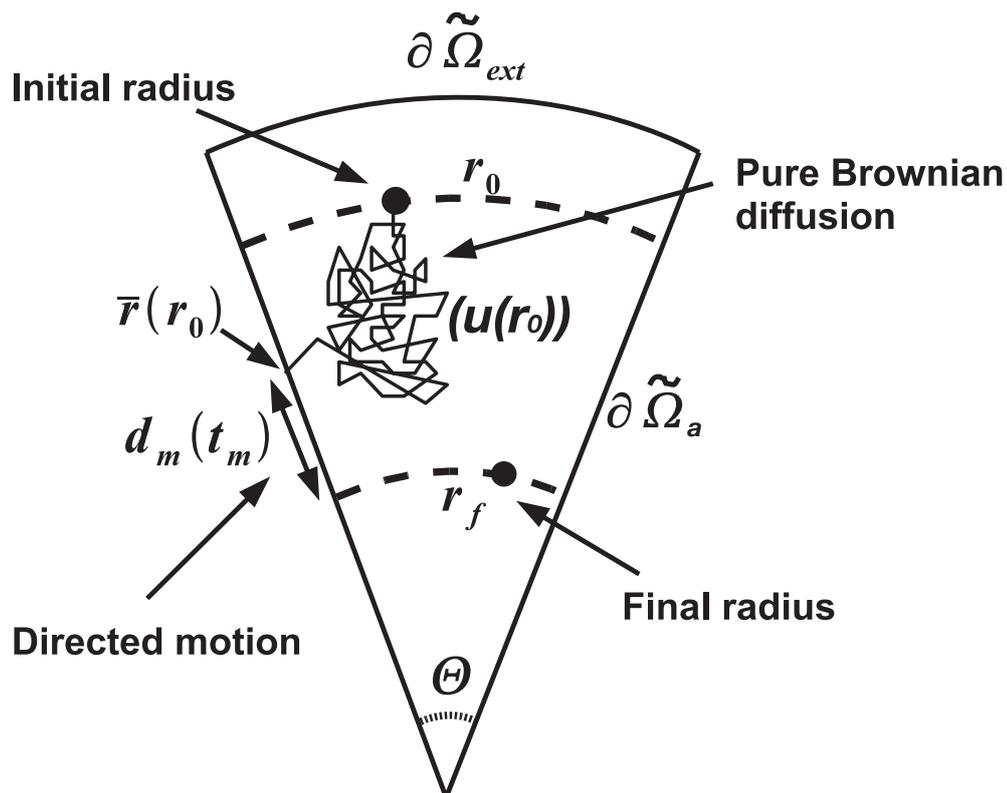


FIGURE 4 – Une étape fondamentale dans $\tilde{\Omega}$. Le virus part d'un rayon initial r_0 , avec un angle uniformément distribué dans $[0; \Theta]$, il diffuse librement durant un temps moyen $u(r_0)$ jusqu'à ce qu'il se lie à un MT ($\partial\tilde{\Omega}_a$) à un rayon moyen $\bar{r}(r_0)$; Il est alors transporté sur une distance moyenne $d_m = Vt_m$ vers le centre de la cellule avant d'être relâché avec un angle aléatoire à un rayon final r_f . Les temps moyens de chaque portion de cette étape fondamentale sont indiqués entre parenthèses.

où le flux j est défini par

$$j(\mathbf{y}, t|\mathbf{x}_0) = -D \frac{\partial p(\mathbf{x}, t|\mathbf{x}_0)}{\partial \mathbf{n}} \Big|_{\mathbf{x} = \mathbf{y}}.$$

Quelques calculs présentés dans le chapitre 1 conduisent à la distribution moyenne des points d'attache $\bar{\epsilon}(r|r_0) = \frac{1}{\Theta} \int_{\Theta_0=0}^{\Theta} \epsilon(r|r_0, \theta_0) d\theta_0$ pour un virus démarrant à une position $\mathbf{x}_0 = (r_0, \theta_0)$ où θ_0 est uniformément distribué entre 0 et Θ :

$$\bar{\epsilon}(r|r_0) = \frac{2}{\Theta \pi r} \left(\ln \left(\frac{r^\nu + r_0^\nu}{|r^\nu - r_0^\nu|} \right) + \ln \left(\frac{R^{2\nu} + (rr_0)^\nu}{R^{2\nu} - (rr_0)^\nu} \right) \right),$$

où $\nu = \frac{\pi}{\Theta}$. Le rayon d'attache moyen $r_m(r_0)$ est alors donné par : $r_m(r_0) = \mathbf{E}(\bar{\epsilon}(r|r_0))$. Ainsi,

$$r_m(r_0) = \mathbf{E}(r|r_0) = \int_0^R r \bar{\epsilon}(r|r_0) dr. \quad (8)$$

Dans la limite $\Theta \ll 1$ le terme d'ordre principal de $r_m(r_0)$ est

$$r_m(r_0) \approx r_0 \left(1 + \frac{\Theta^2}{12} \right). \quad (9)$$

Finalement, dans une géométrie radiale 2-dimensionnelle simplifiée, un virus commençant à diffuser à un rayon initial r_0 sera relâché, après une étape de diffusion/transport fondamentale, à un rayon moyen final $r_f(r_0) = r_m(r_0) - d_m \approx r_0 \left(1 + \frac{\Theta^2}{12} \right) - d_m$ après un temps moyen $\bar{u}(r_0) + t_m \approx \frac{r_0^2 \Theta^2}{12D} + t_m$. La dérive radiale dépendante $b(r)$ comptant pour les périodes ballistiques le long des MTs est alors approximativement donné par

$$b(r) = \frac{r - r_f(r)}{\bar{u}(r) + t_m} \approx \frac{d_m - r \frac{\Theta^2}{12}}{t_m + r^2 \frac{\Theta^2}{12D}}. \quad (10)$$

Dans le premier chapitre, nous obtenons l'amplitude de la dérive constante B grâce à une méthode itérative : nous considérons qu'une fois le virus a été transporté et relâché dans le cytoplasme, celui-ci commence une nouvelle étape de diffusion/transport fondamentale jusqu'à ce qu'il atteigne la surface du noyau. Nous calculons alors le nombre moyen d'étapes fondamentales nécessaires et le temps moyen correspondant pour que le virus partant de la membrane périphérique atteigne le noyau ; la valeur de la dérive B est ensuite calculée pour que le temps moyen d'une particule, dont la vitesse est solution de l'équation de Langevin (2) avec la dérive constante B , soit égale au temps moyen itératif précédent. Dans la limite $\Theta \ll 1$, le terme dominant de B est donnée par

$$B \approx \frac{\frac{d_m}{t_m}}{1 + \left(1 + \frac{R+\delta}{d_m} \right) \frac{\Theta^2}{24} + O(\Theta^4)}. \quad (11)$$

Dans les chapitres 1 et 2, nous imposons que les frontières de $\tilde{\Omega}$ soient entièrement réfléchissantes et nous comparons la densité de probabilité à l'équilibre obtenue avec des simulations Browniennes où le virus peut alternativement diffuser et être

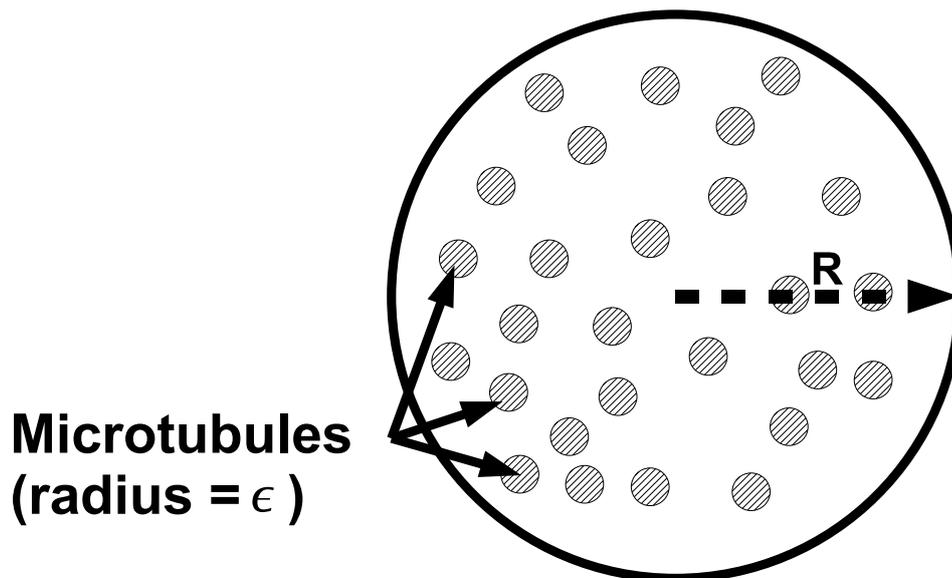


FIGURE 5 – Coupe perpendiculaire à l’axe principal du dendrite. Les N MTs sont de longs et fins cylindres uniformément distribués le long de l’axe principal du neurite.

transporté le long des MTs et la densité de probabilités théorique si les trajectoires du virus sont régies par l’équation de Langevin (2) avec les dérives (constante ou radiale dépendante) calculées précédemment. Dans les deux cas, les densités sont très similaires à celle obtenue par simulation, ce qui représente un résultat central de ces deux chapitres et qui confirme la pertinence de notre approche. De nombreux virus comme l’herpes [23]) infectent les neurones et se déplacent le long des neurites pour atteindre le corps cellulaire et le noyau. Pour calculer ici la dérive équivalente aux périodes ballistiques du virus le long des MTs, nous modélisons les neurites par de long cylindres (rayon R) et nous considérons que les N MTs sont de long cylindres (radius $\epsilon \ll R$) répartis uniformément et parallèlement le long de l’axe principal du neurite. Une coupe perpendiculaire à l’axe principal du neurite est représentée dans la figure 5. Par symétrie, pour toute position \mathbf{x} dans le neurite, nous considérons que la dérive $\mathbf{b}(\mathbf{x})$ est constante et orientée suivant l’axe principal du cylindre : $\mathbf{b}(\mathbf{x}) = B\mathbf{z}$ avec B une constante et \mathbf{z} le vecteur unitaire de l’axe principal. Pour une constante de diffusion suffisamment petite, l’amplitude de la dérive B est égale à la vitesse effective du virus dans le neurite [24, 25] : $B = \frac{d_m}{t_m + \tau}$, où t_m est le temps moyen de liaison du virus à un MT, $d_m = Vt_m$ la distance moyenne d’un transport actif et τ le PTPM d’un virus uniformément distribué dans le cytoplasme à un MT. Dans la limite $\epsilon \ll 1$, en approximant $\tau \approx \frac{1}{\lambda_0(\epsilon)}$ avec $\lambda_0(\epsilon)$ la valeur propre principale de l’opérateur laplacien dans la coupe 2-dimensionnelle du neurite contenant N petits disques absorbants de rayon ϵ nous obtenons dans le chapitre 2 :

$$B = \frac{d_m}{t_m + \tau} = \frac{2NDd_m}{2NDt_m + R^2 \ln\left(\frac{1}{\epsilon}\right)}. \quad (12)$$

3.2 Analyse quantitative du mouvement libre du virus dans le cytoplasme

La probabilité P_n et le temps moyen τ_n pour qu'un virus libre dans le cytoplasme parvienne à un des n pores nucléaires permet de quantifier globalement l'étape de déplacement libre du virus dans le cytoplasme. Du fait que le rayon ϵ des pores nucléaires est petit, les calculs de P_n et τ_n sont des problèmes d'"échappée belle" et se basent sur la théorie des perturbations singulières. Pour calculer ces quantités, nous partons de l'équation de Langevin des trajectoires (2) et introduisons de plus un taux de dégradation $k(\mathbf{x})$ pour tenir compte de la possible dégradation des virus par les protéasomes ou leur immobilisation par le cytosquelette. P_n et τ_n sont alors liés à la densité de probabilité conditionnelle $p(\mathbf{x}, t)$ représentant la probabilité de trouver la particule en vie (non dégradée) dans l'élément de volume $\mathbf{x} + d\mathbf{x}$ au temps t . Cette densité de probabilité est définie par [26],

$$p(\mathbf{x}, t)d\mathbf{x} = Pr\{X(t) \in \mathbf{x} + d\mathbf{x}, \tau^k > t, \tau^a > t | p_i\}, \quad (13)$$

où τ^a est le premier temps de passage d'un virus non dégradé à un pore nucléaire, τ^k le premier temps où le virus est dégradé ou immobilisé et p_i la distribution initiale des particules virales dans le cytoplasme. Un résultat essentiel dans notre démarche est que $p(x, t)$ est solution d'une équation aux dérivées partielles, connue sous le nom d'équation de Fokker-Planck (EFP), décrivant l'évolution de la densité de probabilité conditionnelle du virus au cours du temps [21]

$$\begin{aligned} \frac{\partial p}{\partial t}(\mathbf{x}, t) &= D\Delta p(\mathbf{x}, t) - \nabla \cdot \mathbf{b}(\mathbf{x})p(\mathbf{x}, t) - k(\mathbf{x})p(\mathbf{x}, t) \quad \text{pour } \mathbf{x} \in \Omega \\ p(\mathbf{x}, 0) &= p_i(\mathbf{x}) \quad \text{pour } \mathbf{x} \in \Omega, \end{aligned} \quad (14)$$

Le premier terme du membre de droite indique la contribution de la diffusion (constante de diffusion D), le deuxième correspondant à la dérive équivalente au transport actif le long des MTs et le dernier terme indique la dégradation d'une partie des particules (taux k). De plus, nous avons à la frontière du domaine

$$\begin{aligned} p(\mathbf{x}, t) &= 0 \quad \text{pour } \mathbf{x} \in \partial N_a \\ \mathbf{J}(\mathbf{x}, t) \cdot \mathbf{n}_x &= 0 \quad \mathbf{x} \in \partial\Omega - \partial N_a. \end{aligned} \quad (15)$$

La première condition indique qu'idéalement, le virus est absorbé par le premier pore nucléaire rencontré, l'ensemble des pores formant donc une frontière purement absorbante ∂N_a . La deuxième condition indique que le reste de la frontière cellulaire $\partial\Omega - \partial N_a$ (membrane extérieure et reste de l'enveloppe nucléaire) est purement réfléchissante, \mathbf{n}_x étant le vecteur normal sortant au point \mathbf{x} de la frontière. Le vecteur $\mathbf{J}(\mathbf{x}, t)$ de densité de flux est donné par

$$\mathbf{J}(\mathbf{x}, t) = -D\nabla p(\mathbf{x}, t) + \mathbf{b}(\mathbf{x})p(\mathbf{x}, t). \quad (16)$$

A partir de la densité de probabilité conditionnelle $p(\mathbf{x}, t)$, nous pouvons à présent exprimer la probabilité P_n et le temps moyen conditionnel τ_n pour que le virus atteigne un pore nucléaire [26] :

$$P_n = 1 - \int_{\Omega} k(\mathbf{x})\tilde{p}(\mathbf{x}) d\mathbf{x}, \quad (17)$$

où $\tilde{p}(\mathbf{x}) = \int_0^\infty p(\mathbf{x}, t) dt$ est solution de l'équation

$$D\Delta\tilde{p}(\mathbf{x}) - \nabla \cdot \mathbf{b}(\mathbf{x})\tilde{p}(\mathbf{x}) - k(\mathbf{x})\tilde{p}(\mathbf{x}) = -p_i(\mathbf{x}) \quad \text{for } \mathbf{x} \in \Omega$$

avec les conditions mixtes aux frontières (15). Le temps moyen conditionnel τ_n est quant à lui donné par [26]

$$\tau_n = \frac{\int_\Omega \tilde{p}(\mathbf{x}) d\mathbf{x} - \int_\Omega k(\mathbf{x})q(\mathbf{x}) d\mathbf{x}}{1 - \int_\Omega k(\mathbf{x})\tilde{p}(\mathbf{x}) d\mathbf{x}}, \quad (18)$$

où $q(\mathbf{x}) = \int_0^\infty sp(\mathbf{x}, s) ds$ satisfait [26]

$$-\tilde{p} = D\Delta q(\mathbf{x}) - [\nabla \cdot \mathbf{b}q] - kq \quad \text{for } \mathbf{x} \in \Omega \quad (19)$$

avec les conditions aux frontières (15). Quand la dérive \mathbf{b} provient d'un potentiel $\Phi : \mathbf{b} = -\nabla\Phi$, dans la limite $k \ll 1$, le terme principal asymptotique de P_n et τ_n pour des petits pores nucléaires $\epsilon \ll 1$ est donné par [26]

$$\begin{cases} P_n = \frac{e^{-\frac{\Phi_0}{D}}}{\frac{1}{4Dn\epsilon} \int_\Omega e^{-\frac{\Phi(\mathbf{x})}{D}} k(\mathbf{x}) d\mathbf{x} + e^{-\frac{\Phi_0}{D}}} \\ \tau_n = \frac{\frac{1}{4Dn\epsilon} \int_\Omega e^{-\frac{\Phi(\mathbf{x})}{D}} d\mathbf{x}}{\frac{1}{4Dn\epsilon} \int_\Omega e^{-\frac{\Phi(\mathbf{x})}{D}} k(\mathbf{x}) d\mathbf{x} + e^{-\frac{\Phi_0}{D}}} \end{cases} \quad (20)$$

où Φ_0 est la valeur constante du drift radial $\Phi(\mathbf{x})$ sur le noyau quand celui ci est pris au centre de la cellule, les n pores nucléaires étant uniformément répartis à sa surface. Ces formules ne sont valables que dans le cas où le drift est constant à la surface du noyau et surtout, elles ne tiennent pas compte des possibles interactions entre les différents pores. Ainsi, pour un nombre important de pores occupant une surface restreinte du noyau comme c'est le cas dans de nombreuses cellules animales, ces formules ne sont plus valides ; en effet, $\lim_{n \rightarrow \infty, n\epsilon^2 \ll 1} \tau_n = 0$. De récents travaux [27, 28] ont commencé à étudier finement ces interactions. Dans le chapitre 3, nous quantifions ces interactions quand un grand nombre $n \gg 1$ de pores absorbants sont uniformément répartis sur une structure Σ (Σ étant par exemple le noyau), leur surface $n\pi\epsilon^2$ restant petite par rapport à celle de Σ . En particulier, nous trouvons que dans la limite $\frac{|\Sigma|}{|\Omega|} \ll 1$

$$\begin{cases} P_n = \frac{\frac{\int_{\partial\Sigma} e^{-\frac{\Phi(\mathbf{x})}{D}} d\mathbf{x}}{|\partial\Sigma|}}{\left(\frac{1}{4nD\epsilon} + \frac{1}{DC_\Sigma}\right) \int_\Omega k(\mathbf{x}) e^{-\frac{\Phi(\mathbf{x})}{D}} d\mathbf{x} + \frac{\int_{\partial\Sigma} e^{-\frac{\Phi(\mathbf{x})}{D}} d\mathbf{x}}{|\partial\Sigma|}}, \\ \tau_n = \frac{\left(\frac{1}{4nD\epsilon} + \frac{1}{DC_\Sigma}\right) \int_\Omega e^{-\frac{\Phi(\mathbf{x})}{D}} d\mathbf{x}}{\left(\frac{1}{4nD\epsilon} + \frac{1}{DC_\Sigma}\right) \int_\Omega k(\mathbf{x}) e^{-\frac{\Phi(\mathbf{x})}{D}} d\mathbf{x} + \frac{\int_{\partial\Sigma} e^{-\frac{\Phi(\mathbf{x})}{D}} d\mathbf{x}}{|\partial\Sigma|}}, \end{cases} \quad (21)$$

où C_Σ est la capacitance électrostatique de la surface conductive $\partial\Sigma$. Par exemple, si la structure est un noyau sphérique de rayon δ alors $C_\Sigma = 4\pi\delta$. Quand le ratio

$\sigma = \frac{n\pi\epsilon^2}{|\partial\Sigma|}$ de $\partial\Sigma$ couverts par les pores absorbant est négligeable, pour des particules purement diffusives ($\Phi(\mathbf{x}) = 0$) ne pouvant être dégradées ($k = 0$), le temps moyen (21) est identique à celui trouvé par des considérations électrostatiques dans [29] (formule (54)). Un unique virus peut être suffisant à infecter une cellule. Par conséquent, quand M virus indépendants entrent dans une cellule, il est important de connaître la dynamique du premier virus qui réussira à atteindre vivant un pore nucléaire pour délivrer son patrimoine génétique. Dans le chapitre 4, nous calculons donc le temps de passage moyen conditionnel $\tau_{first}(M)$ à un pore nucléaire du premier virus parmi M . Les M trajectoires des virus sont considérées indépendantes et nous introduisons le temps de passage conditionnel τ_N^j du j^{ime} virus à un pore. Comme dans [26], nous considérons le temps $\tau_{first}^a(M)$ auquel le premier virus atteint la frontière absorbante ∂N_a et le temps $\tau_{first}^k(M)$ auquel il est dégradé. La probabilité $P(t)$ que le premier virus atteigne un pore nucléaire, en vie, avant un temps t est donnée par :

$$P(t) = Pr\{\tau_{first}^a(M) < t | \tau_{first}^a(M) < \tau_{first}^k(M), p_i\}. \quad (22)$$

Le premier temps de passage conditionnel $\tau_{first}(M)$ du premier virus à un pore est alors défini par

$$\tau_{first}(M) = \int_0^\infty t \frac{dP(t)}{dt} dt = \int_0^\infty (P(\infty) - P(t)) dt. \quad (23)$$

Les calculs présentés dans le chapitre 4 conduisent à

$$P(t) = \frac{1 - \left(1 - \int_0^t J(s) ds\right)^M}{1 - (1 - P_N)^M}. \quad (24)$$

où $J(s) = \oint_{\partial\Omega} \mathbf{J}(\mathbf{x}, t) \cdot \mathbf{n}_x dS_x$, avec \mathbf{n}_x le vecteur normal sortant au point \mathbf{x} et $\mathbf{J}(\mathbf{x}, t)$ le vecteur de densité de flux défini par (16). Pour estimer le terme d'ordre principal de $\tau_{first}(M)$ dans un asymptotique en temps long ($k \ll 1$ et $\epsilon \ll 1$), nous approximations $p(\mathbf{x}, t)$ par son premier terme exponentiel

$$p(\mathbf{x}, t) \approx p(\mathbf{x}, 0) e^{-\lambda_0 t}, \text{ avec } \int_{\Omega} p(\mathbf{x}, 0) d\mathbf{x} = 1. \quad (25)$$

où $\lambda_0 = \frac{1}{\tau_n}$ ([21] p.175), est la première valeur propre (voir aussi [30]). Cet asymptotique en temps long conduit alors à l'expression du flux suivante

$$J(t) = \frac{P_n}{\tau_n} e^{-\frac{t}{\tau_n}}. \quad (26)$$

En remplaçant $\int_0^t J(s) ds$ par son approximation (26) dans (23) nous obtenons

$$\tau_{first}(M) = \int_0^\infty \frac{\left(1 - P_n \left(1 - e^{-\frac{t}{\tau_n}}\right)\right)^M - (1 - P_n)^M}{1 - (1 - P_n)^M} dt. \quad (27)$$

Et quelques simplifications conduisent finalement à (voir chapitre 4)

$$\tau_{first}(M) = \frac{\tau_n(\xi)}{1 - \xi^M} \left(\sum_{k=0}^{M-1} (\xi^k - \xi^M) \frac{1}{M - k} \right). \quad (28)$$

où $\xi = 1 - P_n$ ($0 \leq \xi \leq 1$). Quand le nombre de virus atteignant un pore nucléaire avant d'être dégradé augmente, le premier temps de passage moyen conditionnel du premier survivant décroît assez logiquement. Pour ξ petit, le terme d'ordre principal de $\tau_{first}(M)$ est donné par

$$\frac{\tau_{first}(M)}{\tau_n(\xi)} \approx \frac{1}{M}, \quad (29)$$

alors que quand ξ tend vers 1, c'est à dire quand la plupart des virus sont dégradés avant d'avoir atteint leur but, nous obtenons l'approximation

$$\tau_{first}(M) \approx \tau_n(\xi). \quad (30)$$

Dans tous les calculs ci-dessus, nous avons considéré que le taux de dégradation et d'immobilisation des virus était faible comparativement à l'échelle de temps de la diffusion et nous avons donc pu utilisé des asymptotiques en temps long pour calculer P_n et τ_n . Cependant pour certains vecteurs de gènes, et notamment les vecteurs synthétiques purement diffusifs ($\mathbf{b} = \mathbf{0}$) qui sont fréquemment piégés dans le cytosquelette ou dégradés par des protéases, le taux de dégradation ne peut plus être considéré petit. Dans le chapitre 4, nous calculons donc les asymptotiques de P_n et τ_n dans la limite $k \gg 1$. L'analyse n'est plus basé sur des asymptotiques en temps long mais sur des développements asymptotiques des solutions avec recollement au voisinage des pores absorbants. Nous considérons que les vecteurs synthétiques (ou plasmides) sont initialement uniformément répartis dans le cytoplasme $p_i(\mathbf{x}) = p_0 = \frac{1}{|\Omega|}$. Pour calculer la probabilité P_n donnée par (17), nous commençons par résoudre l'équation (18)

$$D\Delta\tilde{p}(\mathbf{x}) - k(\mathbf{x})\tilde{p}(\mathbf{x}) = -p_0 = -\frac{1}{|\Omega|}, \quad (31)$$

avec les conditions au bord (15). Quand $\frac{D}{|\Omega|}$ est très petit devant k , pour une particule partant loin des pores nucléaires absorbants, nous approximons la solution de l'équation (31) par

$$p_{outer}(\mathbf{x}) = \frac{p_0}{k(\mathbf{x})} + O(D). \quad (32)$$

Cependant cette solution ne satisfait pas les conditions au bord et notamment les conditions absorbantes sur ∂N_a . Nous construisons donc deux solutions près du bord p_{inner}^1 près de ∂N_a et p_{inner}^2 près de $\partial N_r = \partial\Omega - \partial N_a$ et recollons ensuite ces deux solutions avec la solution loin du bord p_{outer} . Dans un système de coordonnées locales (ρ, s) près de $\partial\Omega$, où ρ mesure la distance normale par rapport à $\partial\Omega$, mesurée positivement dans Ω , et s sont les variables tangentielles dans le plan $\rho = 0$, nous effectuons un développement de Taylor du taux de dégradation le long de la coordonnée normale ρ ,

$$k(\rho, s) = k_0(s) + k_1(s)\rho + O(\rho^2). \quad (33)$$

où $k(\rho = 0, s) = k_0(s)$ et $\frac{dk}{d\rho}(\rho = 0, s) = k_1(s)$. Nous utilisons ensuite le changement de variable

$$u = u(\rho, s) = \frac{k_0(s) + k_1(s)\rho}{\beta(s)D}, \quad \text{where } \beta(s) = \left(\frac{k_1(s)}{D}\right)^{\frac{2}{3}}. \quad (34)$$

et quelques calculs présentés dans le chapitre 4 conduisent à

$$p_{inner}^1(u, s) = \frac{\pi}{|\Omega|\beta(s)D} \left(Gi(u) - \frac{Gi\left(\frac{k_0(s)}{\beta(s)D}\right)}{Ai\left(\frac{k_0(s)}{\beta(s)D}\right)} Ai(u) \right), \quad (35)$$

et

$$p_{inner}^2(u, s) = \frac{\pi}{|\Omega|\beta(s)D} \left(Gi(u) - \frac{Gi'\left(\frac{k_0(s)}{\beta(s)D}\right)}{Ai'\left(\frac{k_0(s)}{\beta(s)D}\right)} Ai(u) \right). \quad (36)$$

où A_i and G_i sont respectivement les fonctions d'Airy et de Scorer ([31], p.446 et 448). Ces solutions près du bord satisfont les conditions d'absorption et de réflexion imposées et tendent vers p_{outer} loin du bord. Pour un taux de dégradation suffisamment régulier, quand ∂N_a est constituée de n pores absorbants (rayon $(\epsilon_q)_{1 \leq q \leq n}$) centrés en $(\mathbf{x}_q)_{1 \leq q \leq n}$ sur $\partial\Omega$, en injectant (32), (35) et (36) dans (17), nous montrons que le terme d'ordre principal de P_n est alors donné par :

$$P_n \approx \sum_{q=1}^n \frac{\pi \eta_q^2}{|\Omega|} \sqrt{\frac{D}{k(x_q)}} + O\left(e^{-\sqrt{\frac{k_0}{D}} \rho_0}\right). \quad (37)$$

où $k_0 = \inf_q k(\mathbf{x}_q)$ et $\rho_0 = \inf_{s \in \partial N_a} \rho_0(s)$. Concernant le temps moyen de premier passage conditionnel τ_n , des calculs très similaires (voir les détails dans le chapitre 4) conduisent à l'expression asymptotique

$$\tau_n \approx \frac{1}{n} \sum_{q=1}^n \frac{1}{2k(x_q)} + O\left(e^{-\sqrt{\frac{k_0}{D}} \rho_0}\right). \quad (38)$$

Dans le chapitre 4 nous comparons les formules asymptotiques obtenues (37) et (38) contre des simulations Browniennes et nous observons une concordance pour des taux de dégradation très grands (de l'ordre de 200 fois le taux observé expérimentalement [32]) qui pourraient caractériser des cellules anormales ou des vecteurs particulièrement vulnérables.

4 L'étape endosomale

Le séjour du virus dans l'endosome est une autre étape fondamentale dans l'entrée virale. Pour sortir de l'endosome avant un temps critique, le virus est aidé de glycoprotéines ou de protéines de pénétration suivant qu'il est enveloppé ou non. Pour remplir leur fonction, la plupart de ces protéines actives changent de conformation en parallèle de l'acidification de l'endosome. Dans le chapitre 5, nous développons un modèle biophysique de la sortie endosomale dont la cinétique semble déterminante dans l'infectivité du virus. En utilisant des processus de Markov à sauts [33, 34], nous estimons dans un premier temps et à pH donné le temps moyen nécessaire à ce que le nombre de ligands (des protons par exemple) liés à la protéine active dépasse un certain seuil critique et provoque son changement conformationnel. En utilisant le nombre moyens de protons liés à HA1 une sous-unité de la glycoprotéine

hémagglutinine (HA) de la grippe à différents pH [35], nous retrouvons grâce à notre analyse les cinétiques de changement conformationnel observées expérimentalement [36]. Ce qui en outre confirme l'hypothèse que seuls les changements structurels de HA1 sont pH dépendants, les autres changements en découlant spontanément [35, 37]. Combinant ce modèle biophysique de changement conformationnel avec une dynamique Poissonnienne d'entrée des ligands, nous développons ensuite un modèle dynamique pour la sortie de l'endosome de virus contenant un faible nombre de protéines actives comme par exemple l'AAV dont la capside ne contient que 7 protéines de pénétration VP1 [38]. Nous considérons en effet que ce type de virus lyse la membrane endosomale et sort dès qu'une de ses protéines actives change de conformation. En particulier nous calculons que le temps de sortie moyen d'un AAV serait aux alentours de 20 ± 5 minutes (ce qui reste cohérent avec les 10 minutes observées expérimentalement). Lorsque plusieurs virus sont endocytés dans un même endosome, nous considérons que la sortie du premier virus provoque la lyse de l'endosome et donc la sortie de tous les virus. Le nombre de virus présents dans l'endosome devient donc un paramètre important dans la cinétique de sortie. Ainsi, si les virus doivent sortir dans un intervalle de pH compris entre 6.1 et 6.3, nous avons calculé qu'il fallait idéalement 5 virus dans l'endosome. Enfin, grâce à notre modèle biophysique, nous démontrons que la taille de l'endosome qui peut varier au cours du temps *via* des événements de fusion et de division des endosomes [2] n'impacte pas significativement la dynamique de sortie des virus.

4.1 Le modèle de changement conformationnel

La durée du séjour endosomal du virus dépend de sa capacité à lyser la membrane de l'endosome pour s'échapper dans le cytoplasme. La lyse est induite par le changement conformationnel de glycoprotéines ou de protéines de pénétration. Nous considérons qu'un virus contient n_P protéines actives (glycoprotéines ou protéines de pénétration) formées de n_s sites où peuvent se lier des ligands (protons, protéases endosomales ...). Quand le nombre de ligands liés à une protéine atteint un seuil critique n_c , le changement conformationnel se produit. Pour analyser le changement conformationnel de la protéine active, nous étudions donc le nombre de sites occupés $X(t, c)$ au temps t , pour une concentration donnée c de ligands. Entre t et $t + \Delta t$, le nombre de sites occupés peut soit croître avec une probabilité $r(X, c)\Delta t$ quand un ligand supplémentaire se lie à la protéine, soit décroître avec une probabilité $l(X, c)\Delta t$ quand un ligand est relâché, soit rester inchangé avec une probabilité $1 - l(X, c)\Delta t - r(X, c)\Delta t$. En introduisant la variable normalisée $x(t, c) = \epsilon X(t, c)$ où $\epsilon = \frac{1}{n_s}$ et $\Delta x = x(t + \Delta t, c) - x(t, c)$, nous obtenons les probabilités de transition

$$\begin{aligned} \text{Prob}\{\Delta x = \epsilon | x(t, c) = x\} &= r(x, c)\Delta t, \\ \text{Prob}\{\Delta x = -\epsilon | x(t, c) = x\} &= l(x, c)\Delta t, \\ \text{Prob}\{\Delta x = 0 | x(t, c) = x\} &= (1 - r(x, c) - l(x, c)) \Delta t. \end{aligned}$$

Quand la concentration de ligands c est fixe, la probabilité $p(x, y, t, c)$ que le ratio $x(t, c)$ de sites occupés soit égal à y au temps t , sachant qu'à $t = 0$, $x(t = 0, c) = x$,

est solution de l'équation de Kramers-Moyal [33] :

$$\begin{aligned} \frac{\partial p}{\partial t} &= L_x p = r(x, c) \sum_{n=1}^{\infty} \frac{\epsilon^n}{n!} (\partial_x)^n p(x, y, t, c) \\ &+ l(x, c) \sum_{n=1}^{\infty} \frac{(-\epsilon)^n}{n!} (\partial_x)^n p(x, y, t, c). \end{aligned} \quad (39)$$

Le moment où la protéine active change de conformation correspond au premier temps de passage moyen $\tau(x, c)$ de $x(t, c)$ au seuil $x_c = \frac{n_c}{n_s}$ sachant que $x(0, c) = x$. $\tau(x, c)$ satisfait l'équation [39] :

$$\begin{aligned} L_x \tau(x, c) &= -1 \text{ pour } x \text{ in } [0, x_c], \\ \tau(x, c) &= 0 \text{ pour } x = x_c \text{ et } \frac{\partial \tau(x, c)}{\partial x} = 0 \text{ pour } x = 0. \end{aligned}$$

Pour estimer $\tau_0(c) = \tau(x_0(c), c)$, nous considérons le nombre moyen $0 < x_0(c) < x_c$ de ligands liés à la protéine active pour une concentration c de ligands. Pour $\epsilon \ll 1$, le terme d'ordre principal de $\tau_0(c)$ est [33] :

$$\tau_0(c) \approx C(\epsilon, c) \left(1 - \left(\frac{l(x_c, c)}{r(x_c, c)} \right)^{-\frac{x_c - x_0(c)}{\epsilon}} \right), \quad (40)$$

où

$$C(\epsilon, c) \approx \frac{1}{r(x_0(c), c)} \frac{\sqrt{\frac{2\pi}{\epsilon \frac{d}{dx} \left(\frac{l}{r} \right) (x_0(c), c)}}}{\phi(x_c, c)}$$

et

$$\phi(x, c) = \frac{e^{-\frac{1}{\epsilon} \int_{x_0(c)}^x \log \left(\frac{l(s, c)}{r(s, c)} \right) ds}}{\sqrt{\frac{l(x, c)}{r(x, c)}}} \left(\frac{l(x, c)}{r(x, c)} - 1 \right).$$

La formule (40) relie les affinités relatives des ligands (concentration c) pour les sites des protéines actives avec le temps moyen $\tau_0(c)$ au bout duquel la protéine va changer de conformation. Dans le chapitre 4, nous validons notre modèle de changement conformationnel avec les données expérimentales rapportées pour l'hémagglutinine (HA) de la grippe. En utilisant le nombre moyen de ligands (des protons ici) liés à HA1 une sous unité de HA à différents pH [35], nous retrouvons les cinétiques expérimentales du changement conformationnelle de la protéine [36]. Notre analyse confirme l'hypothèse que seul les changements structuraux de HA1 sont pH dépendants, les autres changements intervenant au cours de la métamorphose de HA ayant alors lieu spontanément [35, 37].

4.2 Modèle de sortie endosomale et résultats

Notre modèle s'applique particulièrement aux petits virus nus contenant peu de protéines actives (≈ 10) comme l'AAV. Dans ce cas, nous considérons que le virus

sort de l'endosome dès qu'une de ses protéines actives a changé de conformation. Une fois la membrane lysé par un virus, nous supposons que tous les virus contenus dans l'endosome sont libérés dans le cytoplasme. Il a en effet été observé que des virus sauvages permettaient à des virus mutants dépourvus de protéines actives de redevenir infectieux [40]. Les protons entrent activement grâce à des pompes ATPases uniformément distribuées dans la membrane de l'endosome. Ayant peu d'information sur l'activité de ces pompes, nous modélisons l'arrivée de protons (ou l'activation des autres ligands souvent pH dépendante) par un processus Poissonien d'intensité λ . Pour estimer la concentration de ligands c_{τ_e} dans l'endosome au moment τ_e où s'échappe le virus, nous calculons dans un premier temps la probabilité $P_e^0(c)$ qu'une protéine active change de conformation avant qu'un nouveau ligand n'entre ou soit activé dans l'endosome (c est fixée). Dans le chapitre 5, nous montrons que

$$P_e^0(c) = 1 - u(x_0(c)), \quad (41)$$

avec u la solution de

$$\begin{aligned} (L_x - \lambda) u(x) &= -\lambda \text{ pour } 0 < x < x_c, \\ \frac{du(x)}{dx} &= 0 \text{ pour } x = 0, \\ u(x) &= 0 \text{ pour } x = x_c. \end{aligned} \quad (42)$$

Pour résoudre (42), nous appliquons la méthode asymptotique de [33] et obtenons

$$P_e^0(c) = 1 - \frac{\lambda}{\lambda + (C(\epsilon, c))^{-1}} + \frac{\lambda}{\lambda + (C(\epsilon, c))^{-1}} \left(\frac{l}{r}(x_c, c) \right)^{-\frac{x_c - x_0(c)}{\epsilon}}. \quad (43)$$

Le changement conformationnel d'une protéine active suffit pour que tous les virus présent dans l'endosome soient relâchés dans le cytoplasme. En utilisant la probabilité $P_e(j) (1 - (1 - P_e^0(c(j)))^{n_v n_P}) \prod_{i=0}^{j-1} (1 - P_e^0(c(i)))^{n_v n_P}$ qu'un changement conformationnel survienne après qu'exactly j ligands soient entrés (ou aient été activés) dans l'endosome, nous calculons finalement la concentration moyenne $\langle c_{\tau_e} \rangle = \frac{\sum_{j=1}^{\infty} j P_e(j)}{N V_0}$ de ligands dans l'endosome quand sortent les virus (V_0 est le volume de l'endosome)

$$\langle c_{\tau_e} \rangle = \frac{1}{N V_0} \sum_{j=0}^{\infty} \prod_{i=0}^j \left(\frac{\lambda}{\lambda + (C(\epsilon, c(i)))^{-1}} \right)^{n_v n_P}. \quad (44)$$

Dans le chapitre 5, nous calculons aussi le temps moyen $\bar{\tau}_e$ de sortie des virus

$$\bar{\tau}_e = \frac{1}{\lambda} \left(1 + \sum_{k=1}^{\infty} \frac{1}{\prod_{i=1}^k (1 + \lambda_i / \lambda)} \right). \quad (45)$$

où $\lambda_i = \frac{n_v n_P}{C(\epsilon, c(i))}$. En utilisant les paramètres dynamiques relevés pour l'hémagglutinine, nous trouvons que pour un taux d'entrée des ligands $\lambda = 0.15 s^{-1}$ (voir chapitre 5), pour un virus contenant $n_P = 7$ protéines actives (comme l'AAV), le temps moyen $\bar{\tau}_e$ est alors approximativement de $20 min.$, ce qui reste cohérent avec les $10 min.$ observées expérimentalement [14]. De façon intéressante, quand 10 virus

au lieu d'un sont présents dans l'endosome, le temps moyen de sortie décroît de 35% . Enfin, nous remarquons que, dans notre modèle, la taille de l'endosome n'impacte que peu la cinétique de sortie du virus. Ce qui laisse à penser que les fusions entre endosomes et donc leur variation de volume au cours du temps observé *in vivo* [2] pourrait être négligée.

5 Conclusion

Nous présentons ici un des premiers modèles biophysiques à l'échelle moléculaire de l'entrée des virus dans les cellules. A partir de la géométrie cellulaire et de données dynamiques mesurables du virus comme son coefficient de diffusion ou l'affinité des protéines de sa capsid pour certains ligands, notre approche permet d'obtenir des paramètres dynamiques globaux de l'infection comme le temps de sortie du virus de l'endosome ou la probabilité que celui-ci, une fois relâché dans le cytoplasme, atteigne un pore nucléaire et délivre son patrimoine génétique avant d'être dégradé par les protéasomes. Notre modèle de l'étape cytoplasmique libre du virus s'applique particulièrement à certains vecteurs synthétiques purement diffusifs. La constant de diffusion du complexe (ADN+vecteur) et son taux de dégradation dans le cytoplasme dépendent crucialement du vecteur utilisé. Notre analyse permettrait donc de concevoir des vecteurs optimisant le taux de transfection. Chaque étape dans l'entrée du virus modifie sa structure et donc ses interactions futures avec la cellule. Un bruit stochastique intrinsèque à chacune de ces modifications structurelles conduit à des comportements très différents au sein d'un même groupe de virus. Nous avons développé des modèles biophysiques distincts pour chacune des étapes précoces d'entrée du virus. Ainsi, pour tenir compte de la singularité de chaque virus, nous pourrions coupler ces différents modèles biophysiques, l'état du virus calculé à la fin d'une étape servant de paramètre d'entrée à la suivante.

Introduction

A virus is small entity that hijack the cell machinery to reproduce and that can infect all types of organisms from the bacteria to animals and plants. While some viruses reproduce without causing disease, in most cases they trigger an immune response that destroys them. Since first observation of the tobacco mosaic virus by Martinus Beijerinck in 1898, more than 5,000 species of viruses have been identified and many species shall remain undiscovered. Viruses can be roughly decomposed into three major parts : genetic payload, capsid and, for some of them, envelope. The genetic payload contains the genetic information of the virus and is coded by genes, that are composed of nucleic acids (RNA or DNA). The genetic payload carries the information to replicate the virus and is contained in a capsid which is formed with proteins coded by the viral genes. The capsid protects the genetic payload and facilitates the viral replication. Finally, a fat envelope surrounds some viruses when they are outside a cell (see figure 1). Although viruses share that same organization, they greatly vary in shape and spread in many different ways. For example, the ebola virus has a filamentous shape and spreads person-to-person through body fluids (blood, sperm, saliva . . .) while the HIV has a quasi spherical structure and is transmitted through sexual contact.

1 Viral structure

A viral particle is composed of a nucleic acid coated by a protective shell, the capsid, that is built of polypeptides. Some viruses are moreover surrounded by a fat envelope. As the capsid, the envelope plays a key role in viral infection by helping the virus to deliver its nucleic acid to the cell replication machinery. The polypeptides that form the capsid are encoded by the viral genome and self assemble to form the capsid in the presence of the viral genome. The nucleic acid and its protective capsid are called the nucleocapsid. Depending on the number of polypeptides that compose it, the capsid varies in shape from an helical shape when a single polypeptide is stacked around a central axis, to more complex icosahedral shapes when many polypeptides are recruited (see figure 2). The polypeptides that compose the capsid are organized in morphological units called capsomers. The capsomers of helical viruses are composed of a single polypeptide and are also called protomers while for icosahedral viruses there are two kind of capsomers : the pentamers composed by 5 identical subunits which occupy the 12 corner positions of the icosahedron and the hexamers formed by 6 identical subunits which occupy the faces and edges (see figure 3).

The capsid protects the genetic information of the virus coded by a nucleic acid (RNA or DNA). The size of the genome varies with species : the smallest code for

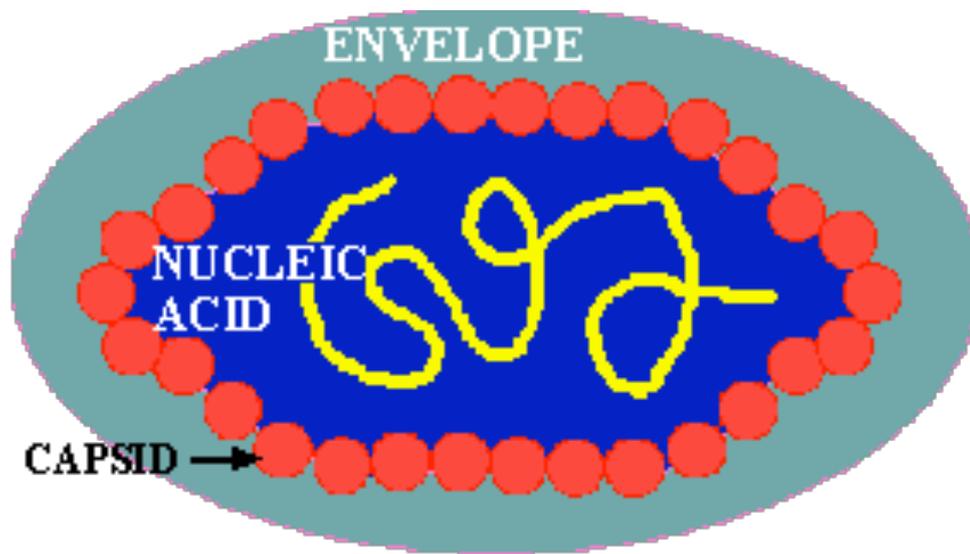


FIGURE 1 – Scheme of an enveloped virus : the nucleic acid is coated by a capsid that is in turn surrounded by a lipid bilayer derived from the host cell membrane

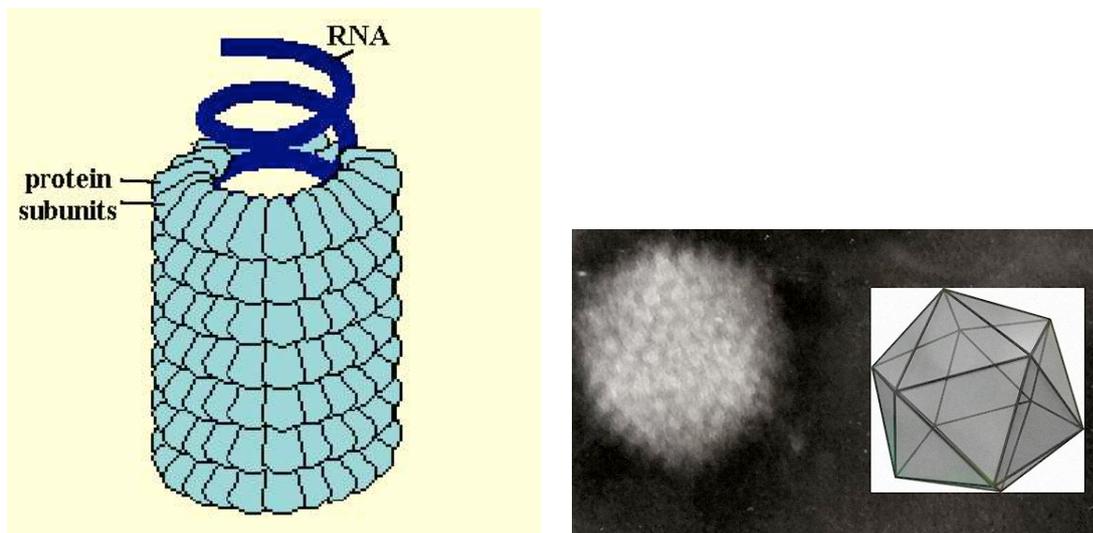


FIGURE 2 – Viruses greatly vary in shape. The two main type of shapes that depend on the number of polypeptides that compose the capsid are the helical shape (left-hand side) when a single polypeptide is stacked around a central axis to form a helical structure and the icosahedral shape when many polypeptides are recruited (right-hand side).

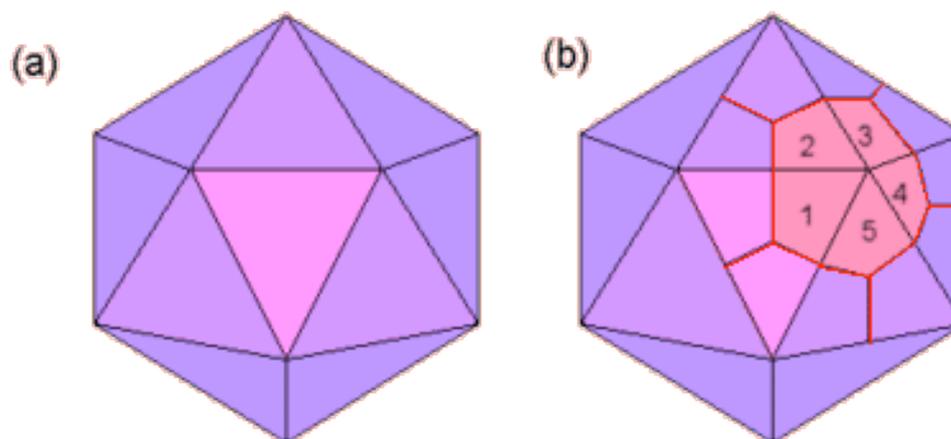


FIGURE 3 – (a) An icosahedron has 20 identical equilateral triangular faces. (b) In most icosahedral capsids, each triangular face is made up of three identical subunits. Hence, a capsid contains 60 subunits. The five subunits surrounding each vertex are arranged in a five-fold symmetry : a pentamer.

only 4 proteins and have a mass of 10^6 daltons while the biggest have a mass of 10^8 daltons and code for more than one hundred of proteins (a dalton is the approximate mass of a hydrogen atom that is $1.66E^{-27}kg$).

2 Application of viruses in gene therapy

The viruses use the cell machinery to reproduce and have thus developed evolutionary tools to enter the cell and transport towards the different compartment of the cell such as the nucleus where DNA viruses deliver their genetic information to reproduce. Consequently, they are interesting tools to transfer genetic material and treat diseases in gene therapy on the one hand, and to study fundamental aspects of cellular biology, like endocytosis or DNA replication [41]. Here, viruses are labelled by a fluorescent dye and are tracked using novel imaging techniques. In particular, tracking of single simian viruses 40 entry has revealed the presence of a new intracellular organelle, the caveosome [42]. In viral gene vectors, the genetic payload is modified to include the transferred genes that will make the cell produce a foreign or missing substance. The viruses have developed evolutionary tools to reproduce and consequently the viral vectors efficiently transfer the therapeutic DNA to the nucleus. But the use of viruses presents several limitations : viruses are not always safe (viral vectors have been implicated in the death of at least one patient, leading to the suspension of clinical trials [43, 44]), they trigger an immune response and thus cannot be administrated repeatedly, and the size of the transferred DNA is limited by the nucleocapsid architecture. To overcome the difficulties impose by the use of viruses, gene therapy with non-viral vectors, such as naked DNA or complexes of polymers/lipids with DNA, is important. Yet non-viral vectors present low levels of transfection and expression of the gene, and to combine the high efficiency of viruses with the advantages of non-viral vectors, virus-like particles that integrate the evolutionary tools developed by viruses such as the protecting capsid or the active

proteins that permit the endosomal escape are now investigated. For example, virus-like particles build out of hepatitis B virus surface proteins have been developed [45] to efficiently transfer genes into human hepatocytes. Developing such hybrid vectors require a molecular understanding of the early steps of infection, that start when the virus binds to the cell membrane and that end when the nucleic acid reaches its replication site, that is the nucleus for DNA viruses.

3 Early steps of viral infection

Replication cycle starts with the attachment of viruses to specific receptors at the cell surface. Because the type of receptors present on the membrane depend on the cell population, each virus infect only a certain type of cells, which determines its tropism. A striking example is the HIV that only infects the human T cells, coated by the CD4 receptors that interact with the glycoprotein gp120 of the viral envelope [46]. After the virus attaches to the cell surface, it enters into the cell through the receptor mediated endocytosis or by membrane fusion for certain enveloped viruses.

The viruses that are endocytosed begin their cytoplasmic journey in an endosome that ferry them as passive luminal cargo. Endosomes are sorting vesicles towards the different compartment of the cell and the Rab family of proteins plays a crucial role in determining endosomal localization [1]. One of the principal pathway is the Rab5-Rab7 mediated one (see [2] for details) where viruses are endocytosed in early endosomes that matures into late endosomes. During that maturation process, the cell activates proton pumps located on the endosomal membrane to decrease the pH and viruses have to escape the endosome before being digested low-pH activated proteases. Although the escape process is fundamental, it is not yet clearly understood but following acidification and depending on whether the virus is enveloped or naked, endosomal escape is triggered by the conformational changes of glycoproteins or penetration proteins, involved in membrane disruption. The conformational change of these active proteins is initiated by the binding of pH activated ligands or protons up to a critical threshold [35]. Because viruses are degraded by low-pH activated proteases, and that their partial denaturation could enhance the nuclear import of their genome [11], the time dependent pH of the endosome plays a critical role in the infection process. Indeed, the viral escape triggered by the pH-dependent conformational change of active proteins must occur in a certain pH range, that is after the virus is sufficiently denatured or tagged by certain low-pH activated proteases, and before it is completely degraded.

To produce proteins and replicate themselves viruses must generate messenger RNAs (mRNAs) from their genomes, and the mRNA production mechanism depends on their genome organization. The nucleic acid (RNA or DNA) can be either single stranded (ss) or double stranded (ds) and the ssRNA can be either positive-sense or negative-sense, depending on its complementary to the viral mRNA : while a negative-sense RNA strand is identical to mRNA and can be immediately translated, a positive-sense strand is complementary to mRNA and must be first converted to a negative-sense strand by an RNA polymerase. In addition, retroviruses, such as the HIV, use a reverse transcriptase that transcribes single-stranded RNA into single-stranded DNA. The mRNA production strategy is then driven by the genome organization : for example, the single RNA strand of a (+)sense RNA virus is first

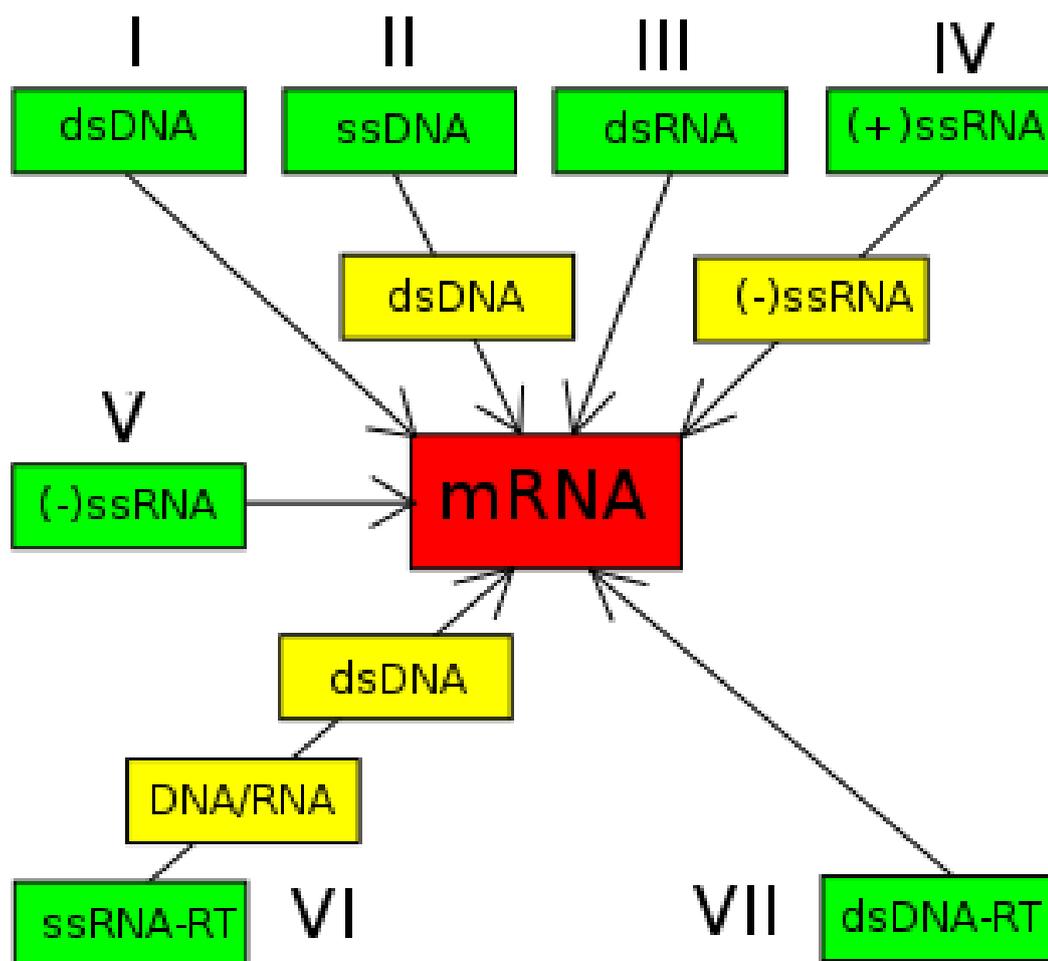


FIGURE 4 – Baltimore classification of viruses : depending on their genome organization and their subsequent mRNA production mechanism, viruses are divided in seven categories.

converted to a (-) strand to be translated while a the double strands of a dsDNA virus is transcribed by the cell into the mRNA molecule that will produce the proteins. The Baltimore classification of viruses is based on the mechanism of mRNA production and thus divides viruses in seven categories (see figure 4)

- I : dsDNA viruses (e.g. Adenoviruses, Herpesviruses, Poxviruses)
- II : (+)ssDNA viruses (e.g. Parvoviruses)
- III : dsRNA viruses (e.g. Reoviruses)
- IV : (+)ssRNA viruses (e.g. Picornaviruses, Togaviruses)
- V : (-)ssRNA viruses (e.g. Orthomyxoviruses, Rhabdoviruses)
- VI : ssRNA-RT viruses : (+)sense RNA with DNA intermediate in life-cycle (e.g. Retroviruses)
- VII : dsDNA-RT viruses (e.g. Hepadnaviruses)

In particular, to transcribe and replicate their genome, DNA viruses depend on the cell transcription machinery and the DNA polymerase that are located in the cell nucleus. Consequently, once they are released in the cell cytoplasm, DNA viruses have to reach a small nuclear pore where they either enter the cell nucleus when they are sufficiently small, such as the parvovirus, or dock and disassemble at the

nuclear pore to deliver the genetic information, such as the hepatitis B virus, the herpesvirus or the adenovirus (see figure 5). On the other hand RNA or reverse transcribing viruses, such as Influenza or HIV respectively, uncoat in the cytoplasm before translocating into the nucleus. To reach a nuclear pore, DNA viruses entirely rely on diffusion and cellular transport systems. The cell cytoplasm is a highly crowded fluid containing many organelles, a cytoskeleton and a high concentration of diffusible macromolecules [12]. The mobility depends on many parameters such as the size, the shape and the nature of the interactions between viral particle and the surrounding cellular components. Non interacting spherical particles with size up to $\approx 25nm$ are freely diffusible in the cell cytoplasm [13]. Increasing the size above $45nm$ reduces considerably the motion. Recent progress using single particle tracking has revealed the complexity of viral trajectories [14, 15, 16]. It has now been recognized that such trajectories consists of a succession of free or confined diffusion and/or ballistic periods. These later involve transport along MTs or actin networks which requires energy. Motors are ATPase : they move along MTs powered by the dephosphorylation of ATP. There are two types of motors : the kinesins that move from the centrosome, an organelle located nearby the cell nucleus (minus end), to the periphery of the cell (plus end) and the dyneins that move from the plus end to the minus end. There is significant evidence that cargoes in-vivo are transported by multiple motor. Indeed, many viruses such as herpes virus [17], adenovirus [18] or HIV [19] bind motors of different polarities which lead to a bidirectional transport. A regulatory mechanism should favor the switch dynamics in one direction leading to a net velocity in that particular direction [20]. The diffusion periods and the random organization of the MTs network lead to a broad class of viral trajectories. Because the viruses can be trapped in the crowded cytoplasm or digested through the ubiquitin-proteasome machinery, that is they are tagged for degradation with a small protein called ubiquitin to be later degraded by proteolysis in the large protein complexes proteasomes, the random trajectory of the virus and consequent resident time in the cell cytoplasm impacts the infection rate. In particular, the viral shape impacts the diffusion constant and the probability the virus is trapped, and filamentous viruses are particularly vulnerable. Finally we can schematically decompose the early steps of DNA viruses infection into 6 steps (see figure 6)

1. The virus binds to a specific receptor and is internalized in an endosome
2. It traffics inside the endosome through the cortical actin network
3. It is transported actively in the endosome along the MTs
4. It escapes from the endosome
5. The viral motion alternates between diffusion and active transport along the MTs
6. Virus finally reaches a nuclear pore and delivers its genetic material.

After their assembly, the new viral particles are released from the host cell to infect other cells. While enveloped viruses bud from the host cell, a process where they acquire their host cell derived envelope, most of viruses destroy the cell membrane and kill it to escape.

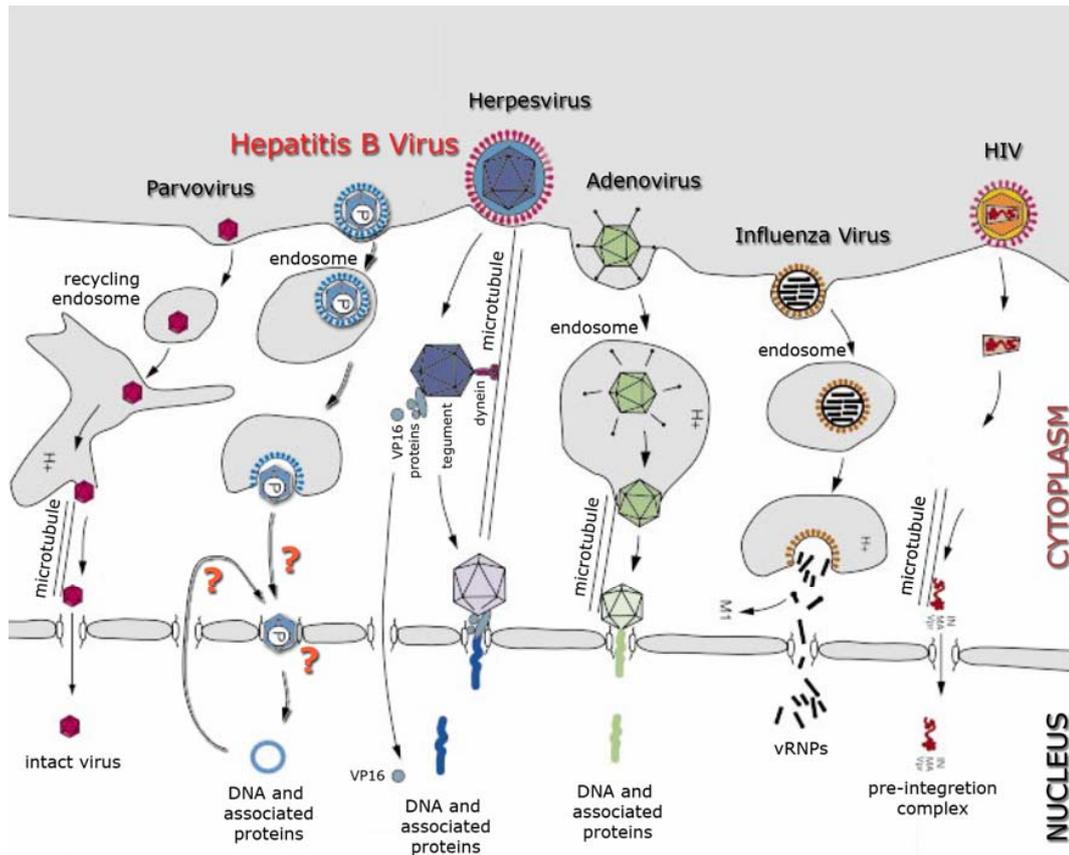


FIGURE 5 – Early infections steps of some nuclear-replicating viruses. While RNA or reverse transcribing viruses such as Influenza or HIV respectively uncoat in the cytoplasm before translocating into the nucleus, DNA viruses dock to the nuclear pore complex and deliver their genetic payload or enter the nucleus if they are sufficiently small, such as parvoviruses. This figure has been adapted from [47].

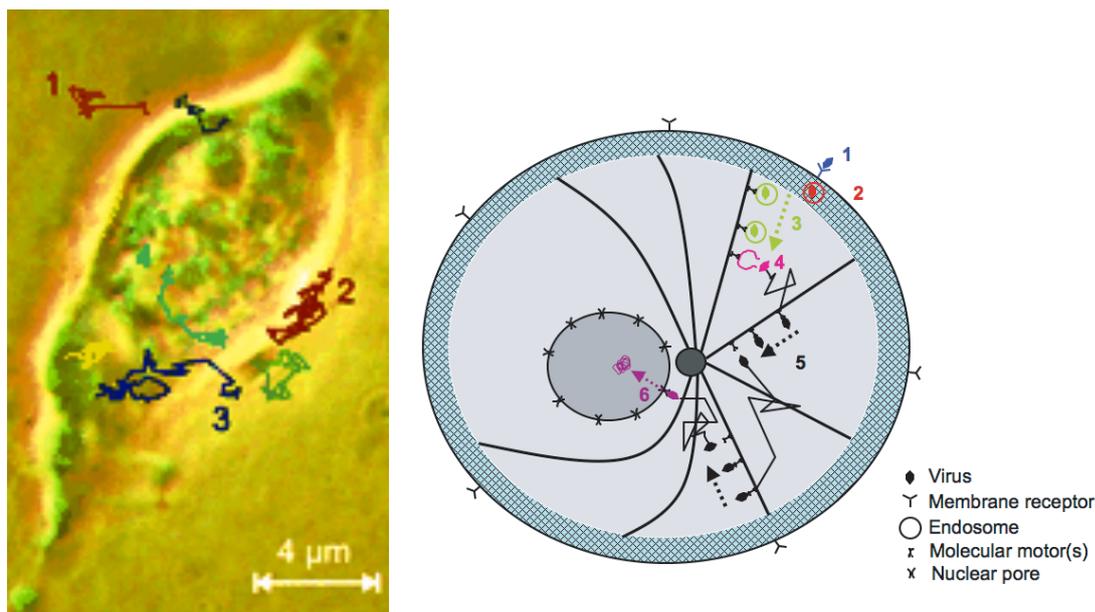


FIGURE 6 – **Left-Hand Side** : Single particle tracking of a small DNA virus (AAV) in a living cell (from [14]). (1) The virus diffuses in the extracellular matrix. (2) The Virus searches for a specific receptor at the cell membrane. (3) The virus enters the cell in an endosome. It later escapes and alternates between diffusion and active transport along MTs. (4) The Virus finally enters the nucleus through a nuclear pore to deliver its DNA. **Right-Hand Side** Schematic representation of the early steps of infection for a DNA virus. (1) A viral particle binds to a specific receptor and is internalized in an endosome. (2) Virus traffics inside the endosome through the cortical actin network. (3) The virus is transported actively in the endosome along the MTs. (4) Virus escapes from the endosome. (5) The viral motion alternates between diffusion and active transport along the MTs. (6) Virus finally reaches a nuclear pore and delivers its genetic material.

4 Modeling the early steps of viral infection

Viruses have developed evolutionary tools to efficiently deliver their genes and reproduce, and their *in vivo* tracking reveal unexpected aspects of cellular biology such as new ways of endocytosis [42]. In particular, the endosomal escape and the cytoplasmic trafficking to the nucleus remain a major obstacle to gene delivery. Indeed, many non-viral gene vectors fail to escape the endosome and are routed toward a lysosome where they are digested, and the cytosolic motion of large DNA molecules is limited by physical and chemical barriers of the crowded cytoplasm [3, 4]. Whereas molecules smaller than 500 kDa can diffuse, larger cargos such as viruses or non-viral DNA particles, require an active transport system [5] such as the microtubules. To understand the viral tools at a molecular level and thus design efficient synthetic gene carriers, quantitative models of the early steps of viral infection are needed : while modeling the endosomal escape of viruses will help to understand the molecular mechanisms underlying their reliable escape in a certain pH range, a model of the free cytoplasmic step will measure, for example, how the active transport along microtubules increases the probability the virus reaches a nuclear pore compared to a pure diffusive non-viral gene vector.

TABLE 1 Nomenclature of mathematical symbols

$c_s^E(r, t)$	Areal density of particles at state S and substate s (no. of viruses per unit area)
$c(r, t)$	Areal density of particles at distance r from cell center
D_0	Free diffusion constant of the particle
D_{eff}^S	Effective diffusion coefficient for state S
DNA	Number of exogenous DNA entering the nucleus
δ_{MT}	Average spacing between microtubules
f_{\pm}	Fraction of actively transported particles in \pm directions
f_{cat}	Frequency of catastrophe
k_{\pm}^b	Rates of binding of particles to microtubules
k_{\pm}^d	Rates of detachment of particles from microtubules
k_{bind}/k_{unbind}	Rates of binding/detachment rates of viruses to/from cell membrane
$k_{disassembly}$	Rates of capsid disintegration once virus is bound to nuclear membrane
$k_{transcription}/k_{translation}$	Rate of transcription and translation of viral DNA
$mRNA$	Number of transcripts
$n_s^E(r, t)$	Radial density of particles at state S and substate s at distance r from the cell center (no. of viruses per unit length)
$n(r, t)$	Radial density of particles at distance r from cell center

TABLE 2 Governing equations for intracellular trafficking of viral particles

Biophysical state	Diffusion-advection-reaction equations
Membrane-bound (Eq. 1a)	$\frac{\partial n^M}{\partial t} = D^M \frac{\partial}{\partial r} \left[\frac{\partial}{\partial r} \left(\frac{n^M}{r} \right) \right] - k_{un} n^M + k_{bind} \frac{2V_{cell} r}{R_C} - k_{inact} n^M.$
Endosomal (Eq. 1b)	a), $R_N < r < R_C$ $\frac{\partial n_0^E}{\partial t} = (k_{-}^E n_{-1}^E + k_{+}^E n_{+1}^E) - (k_{-}^E n_0^E + k_{+}^E n_0^E) + D^E \frac{\partial}{\partial r} \left[r \frac{\partial}{\partial r} \left(\frac{n_0^E}{r} \right) \right] + k_{un} n^M - k_{escape} n_0^E$ $\frac{\partial n_{\pm 1}^E}{\partial t} = -k_{\pm}^E n_{\pm 1}^E + k_{\pm}^E n_0^E - V_{\pm}^E \frac{\partial n_{\pm 1}^E}{\partial r}$ $\frac{\partial n_{\pm 1}^E}{\partial t} = -k_{\pm}^E n_{\pm 1}^E + k_{\pm}^E n_0^E - V_{\pm}^E \frac{\partial n_{\pm 1}^E}{\partial r}$ b), $r < R_N$ $\frac{\partial n_0^E}{\partial t} = D^E \frac{\partial}{\partial r} \left[\frac{\partial}{\partial r} \left(\frac{n_0^E}{r} \right) \right] + k_{un} n^M - k_{escape} n_0^E.$

FIGURE 7 – Intracellular trafficking model based on mass action law (extracted from [48]). The viral particle transits between different states (right-hand side) such as endosomal, cytoplasmic or nuclear bound state. Transition rates and other parameters of the model (table on the left-hand side of the figure) are mostly fitted to experimental data.

The early mathematical models to quantify the success of gene delivery [6, 7, 8] divide the infection process into various states, such as as being transported inside the endosome, freely trafficking in the cytoplasm or entering the cell nucleus, and use mass action-law law to account for the transition between the states and the possible degradation of the virus (see figure 7). The principal limitation of these approaches is that the transition rates between the different viral states are not derived from underlying viral dynamics but mostly fitted to experimental data. Consequently, this type of analysis does not explain how general quantitative parameters of the infection such as the endosomal escape mean time or the probability a virus reaches a nuclear pore before being degraded depend on the cellular geometry (distribution of MTs ...) and dynamical parameters of the virus (diffusion constant, affinities of viral active proteins with ligands ...). Although an alternative approach in [48] using the cellular geometry allowed the authors to give a macroscopic description of the adenovirus concentration so that they can analyze the effect of varying the number of MTs on the invasion process, the authors do not provide explicit analytical formulas that measure the impact of the cell and virus parameters on general infection outcomes : For example, the impacts of the MTs network organisation, the diffusion constant of the virus or the number and size of nuclear pores on the probability and the mean time the virus reaches a nuclear pore to deliver its DNA cannot be analytically quantified. In addition, each early step of viral infection modifies the viral molecular structure and modulates the viral behavior in the following steps of infection. For example, the capsid denaturation of the adeno-associated virus (AAV) in the endosome that depends on the escape time impact its cytoplasmic ubiquitination [11] that in turn competitively increases its proteasome-mediated degradation and enhances capsid disassembly and subsequent nuclear import [11]. Yet, the early models are only valid at a population level and cannot account for these individual variations in activity. Thus, to account for the individual structural changes that impact the behavior of each particle, biophysical models at the single unit are required.

To quantify the impact of the cell geometry and the dynamical parameters on viruses trajectories, we propose a stochastic approach to model compartment of an individual particle in each early step of infection. In these infection steps, we identify three functional modules in which the host-virus interaction modulates the viral trajectory and the infection process : the viral binding to the cell membrane, the endosomal step and the free cytoplasmic trafficking after the viral release from the endosome. These modules interact with each other through a complex host-virus communication and require specific biophysical models. In the first module, viruses interact with specific cell surface receptors, that will determine the fate and/or the viral pathway in the cytoplasm. For example, fusion proteins of widely disparate enveloped viruses completely metamorphose during viral entry [9]. In particular, the interaction of the retrovirus avian leukosis virus with the cell membrane specific receptor, transforms its pH-independent glycoproteins to pH-dependent ones. As a consequence the fusogenic activity at low pH [10] is deployed which is necessary for the endosomal escape. In the case of AAV, cells and serotype specific receptors lead to a broad class of endocytic pathways [1]. Each pathway is characterized by a specific endosomal environment and an escape dynamical process, both lead to different viral escape locations. In addition, the viral binding to the cell membrane determines the number of viruses per endosome which is crucial for the escape dynamics (see chapter 5). The second module consists of the endosomal step. Although, the escape time from endosomes can be computed from the conformational changes of viral active proteins, the escape location depends on the surface receptor interactions (first module). Both the escape location and the associated pH are key input parameters for the third module, which consists of the free cytoplasmic step starting from the endosomal escape and ending at a nuclear pore where DNA viruses deliver their genetic information. The first two modules impact that free cytoplasmic step : the endosomal pathway and the escape location depends on the membrane receptor the virus binds and we have seen above that in the case of AAV, the pH dependent capsid denaturation in the endosome impacts the cytoplasmic degradation and nuclear import rates [11]. Finally, to find the optimal infection pathways, all three modules should be coupled and the output parameters of one will serve as the inputs for the next one. For example, in the case of AAV, it would be interesting to determine how the escape pH and subsequent capsid denaturation impacts cytoplasmic degradation rate through the proteasome-mediated digestion of the capsid. Because each virus is routed to a specific pathway through a complex host-cell communication, a quantitative analysis of each single pathway would be needed. More fascinating, as viruses infect cells and the host cell interaction start to change, viruses should see a different cell environment depending on their arrival time at the surface. Hereafter we present the biophysical models we have constructed for the endosomal step and the free cytoplasmic step, and the principal quantitative results we have obtained

5 Models and Results

5.1 The endosomal step

A fundamental aspect of viral trafficking concerns the sojourn time in the endosomal compartment. To escape endosomes, before a critical time, the viral payload

is assisted by active proteins, that are glycoproteins for enveloped viruses or penetration proteins for non enveloped viral particles. To fulfill their goal, these proteins have to undergo a conformational change often resulting from endosome acidification. Because the exit time plays a critical role in the viral infectivity process, we develop in chapter 5 a model to estimate the escape time. Using a discrete Markov jump analysis [33, 34], we first estimate, at a given pH, the mean time the number of bound protons (or other pH-activated ligands) reaches a critical threshold, which triggers the conformational change of a given active protein. Combining these computations with experimental data [35] on the mean number of protons bound to HA1 (a subunit of the influenza hemagglutinin (HA)), we recover measured conformational change kinetics [36] and confirm the hypothesis that only HA1 conformational change is pH-dependent and other rearrangements in HA proceed spontaneously [35, 37]. Combining the conformational change discrete model described above with an endosomal Poissonian entry of ligands, we derive for non enveloped viruses that contain a small number of penetration proteins, such as AAV with 7 VP1 penetration proteins [38], the mean escape time from the endosome and the associated pH. In that computations, we consider that viruses escape from the endosome when at least one conformational change occurs. In particular, we find that for AAV the mean time to escape is around 20 ± 5 minutes (which is consistent with the observed 10 minutes) and when the virus has to escape in a pH range of 6.1-6.3, this is optimally achieved when 5 viral particles are inside an endosome. Finally, this biophysical model predicts that the size of the endosome, which may vary following endosomal fusion or split [2], does not impact much the escape.

Conformational change model The resident time of a virus inside an endosome depends on its ability to disrupt and escape the membrane. Disruption is induced by glycoproteins or penetration protein conformational change. We consider a virus carry n_p independent proteins (glycoproteins or penetration proteins) formed of n_s sites that can bind competitively ligands such as protons or endosomal proteases. When the number of bound sites at a single protein reaches a critical threshold n_c , a conformational change occurs that lead to membrane disruption and viral escape. To follow the conformational change for a single glycoprotein or a penetration protein, we count the amount of occupied sites $X(t, c)$ at time t , for a given ligand concentration c . During time t and $t + \Delta t$, the number of bound sites can either increase with a probability $r(X, c)\Delta t$ when a ligand arrives to a free site, decreases with probability $l(X, c)\Delta t$ when a ligand unbinds or remains unchanged with probability $1 - l(X, c)\Delta t - r(X, c)\Delta t$. Using the scaled variable $x(t, c) = \epsilon X(t, c)$ where $\epsilon = \frac{1}{n_s}$ and $\Delta x = x(t + \Delta t, c) - x(t, c)$, we obtain the transition probabilities

$$\begin{aligned} \text{Prob}\{\Delta x = \epsilon | x(t, c) = x\} &= r(x, c)\Delta t, \\ \text{Prob}\{\Delta x = -\epsilon | x(t, c) = x\} &= l(x, c)\Delta t, \\ \text{Prob}\{\Delta x = 0 | x(t, c) = x\} &= (1 - r(x, c) - l(x, c)) \Delta t. \end{aligned}$$

When the ligand concentration is fixed, the probability $p(x, y, t, c)$ that the number of bound is equal to y at time t $x(t, c) = y$, given that initially the number of bound

is x ($x(t=0, c) = x$) is solution of the backward Kramers-Moyal equation [33] :

$$\begin{aligned} \frac{\partial p}{\partial t} &= L_x p = r(x, c) \sum_{n=1}^{\infty} \frac{\epsilon^n}{n!} (\partial_x)^n p(x, y, t, c) \\ &+ l(x, c) \sum_{n=1}^{\infty} \frac{(-\epsilon)^n}{n!} (\partial_x)^n p(x, y, t, c). \end{aligned} \quad (46)$$

The first time a glycoprotein or a penetration protein is filled up to a critical threshold $x_c = \frac{n_c}{n_s}$ is the mean first passage time $\tau(x, c)$ that the level of bound ligands x reaches the level x_c , starting at a point x for a given concentration c , and satisfies [39] :

$$\begin{aligned} L_x \tau(x, c) &= -1 \text{ for } x \text{ in } [0, x_c], \\ \tau(x, c) &= 0 \text{ for } x = x_c \text{ and } \frac{\partial \tau(x, c)}{\partial x} = 0 \text{ for } x = 0. \end{aligned}$$

To estimate $\tau_0(c) = \tau(x_0(c), c)$, we consider the number of bound ligands at equilibrium $0 < x_0(c) < x_c$. For $\epsilon \ll 1$, $\tau_0(c)$ is approximated by [33] :

$$\tau_0(c) \approx C(\epsilon, c) \left(1 - \left(\frac{l(x_c, c)}{r(x_c, c)} \right)^{-\frac{x_c - x_0(c)}{\epsilon}} \right), \quad (47)$$

where

$$C(\epsilon, c) \approx \frac{1}{r(x_0(c), c)} \frac{\sqrt{\frac{2\pi}{\epsilon \frac{d}{dx} \left(\frac{l}{r} \right) (x_0(c), c)}}}{\phi(x_c, c)}$$

and

$$\phi(x, c) = \frac{e^{-\frac{1}{\epsilon} \int_{x_0(c)}^x \log \left(\frac{l(s, c)}{r(s, c)} \right) ds}}{\sqrt{\frac{l(x, c)}{r(x, c)}}} \left(\frac{l(x, c)}{r(x, c)} - 1 \right).$$

Formula (47) links the affinities between the ligand (concentration c) and the binding sites of glycoproteins or penetration proteins to its conformational change mean time $\tau_0(c)$. We validate our model with experimental data for the influenza hemagglutinin (HA) : using the mean number of protons bound to the subunit HA1 of HA at different pH [35], we recover the reported conformational change kinetics of the protein. We also confirm the hypothesis that only HA1 conformational change is pH-dependent and other rearrangements in HA proceed spontaneously [35, 37]

Endosomal escape model and results We consider that all viruses escape when at least one of them disrupts the membrane as suggested for instance by experiments where a wild type parvovirus was shown to rescue the infectivity of a mutant devoid of escape protein [40]. In addition, our model apply to naked viruses that contain few (≈ 10) penetration proteins and we thus assume viruses escape when at least one of the active proteins changes of conformation. To model the proton influx through pumps uniformly distributed over the endosomal membrane [49], we use

a Poisson process of rate λ . To estimate the mean ligands concentration c_{τ_e} at which viruses escape the endosome, we first estimate the probability $P_e^0(c)$ that a penetration protein changes conformation before a new ligand enters or is activated (the concentration c is fixed). In chapter 5, we show that

$$P_e^0(c) = 1 - u(x_0(c)), \quad (48)$$

with u solution of

$$\begin{aligned} (L_x - \lambda) u(x) &= -\lambda \text{ for } 0 < x < x_c, \\ \frac{du(x)}{dx} &= 0 \text{ for } x = 0, \\ u(x) &= 0 \text{ for } x = x_c. \end{aligned} \quad (49)$$

To solve (49), we apply the methods of [33] and we obtain

$$P_e^0(c) = 1 - \frac{\lambda}{\lambda + (C(\epsilon, c))^{-1}} + \frac{\lambda}{\lambda + (C(\epsilon, c))^{-1}} \left(\frac{l}{r}(x_c, c) \right)^{-\frac{x_c - x_0(c)}{\epsilon}}. \quad (50)$$

One protein conformational change is enough to induce viral escape. Using the probability $P_e(j) (1 - (1 - P_e^0(c(j)))^{n_v n_P}) \prod_{i=0}^{j-1} (1 - P_e^0(c(i)))^{n_v n_P}$ that at least one conformation changes occurs after exactly j ligands have entered the endosome (or have been activated), we finally compute the mean concentration $\langle c_{\tau_e} \rangle = \frac{\sum_{j=1}^{\infty} j P_e(j)}{\mathcal{N}V_0}$ for which viral particles escape the endosome (V_0 is the volume of the spherical endosome)

$$\langle c_{\tau_e} \rangle = \frac{1}{\mathcal{N}V_0} \sum_{j=0}^{\infty} \prod_{i=0}^j \left(\frac{\lambda}{\lambda + (C(\epsilon, c(i)))^{-1}} \right)^{n_v n_P}. \quad (51)$$

In the chapter 5, we also estimate the mean escape time $\bar{\tau}_e$,

$$\bar{\tau}_e = \frac{1}{\lambda} \left(1 + \sum_{k=1}^{\infty} \frac{1}{\prod_{i=1}^k (1 + \lambda_i/\lambda)} \right). \quad (52)$$

where $\lambda_i = \frac{n_v n_P}{C(\epsilon, c(i))}$. Using the parameters fitted with the conformational change of the influenza HA, we find that for a ligand entry rate $\lambda = 0.15s^{-1}$ and $n_P = 7$ penetration proteins, the mean time $\bar{\tau}_e$ is approximately equal to $20min.$, which is comparable to the $10min.$ reported experimentally [14]. Interestingly, for 10 viruses, the mean escape time decays by 35% compared to a single one. In addition, the endosomal radius impacts not that much the escape dynamics. Consequently, endosomal fusion or split [2] observed *in vivo* can be neglected in the biophysical modeling.

5.2 The free cytoplasmic step

Once the DNA viruses are released from the endosome, they entirely rely on diffusion and cellular transport systems to reach a nuclear pore. The analysis of

random trajectories of viral particles is formulated in terms of stochastic equations. The position $\mathbf{X}(t)$ at time t of a particle is treated as a stochastic process [21, 22] and the dynamics depends on the forces applied on the particle. In the cytoplasm, the high frequency collisions between a viral particle and the rest of the molecules are modeled by the classical noise term $\sqrt{2D}\frac{d\mathbf{W}}{dt}$, where D is the diffusion constant and W is the standard Brownian motion. In the absence of any other forces, the equation for the velocity is simply $\frac{d\mathbf{X}}{dt} = \sqrt{2D}\frac{d\mathbf{W}}{dt}$. When the virus switches between diffusion and an active transport, the physical description of the position $X(t)$ at time t is

$$\dot{\mathbf{X}} = \begin{cases} \sqrt{2D}\dot{\mathbf{w}} & \text{for a free particle} \\ \mathbf{V}(s(t)) & \text{for a bound particle} \end{cases}, \quad (53)$$

where the variable $s(t)$ describes the internal state, accounting for the nature and the number of bound kinesins and dyneins. \mathbf{V} is the resulting transport field defined by the MTs network organization, it depends on the load exerted by the transported virus on motors.

Although the previous description using equation (53) may allow to generate simulations of trajectories, we can not use it for a general analysis. In the chapters 1 and 2, we present a mathematical procedure to coarse grain the switching dynamics (53) so that the velocity $\dot{\mathbf{X}}$ can be written as the sum of a permanent effective drift term $\mathbf{b}(\mathbf{X})$ that accounts for the ballistic periods along the MTs and the random interactions noise term $\sqrt{2D}\frac{dW}{dt}$ (see figure 8)

$$\frac{d\mathbf{X}}{dt} = \mathbf{b}(\mathbf{X}) + \sqrt{2D}\frac{d\mathbf{W}}{dt}. \quad (54)$$

Computation of the drift accounting for ballistic periods along MTs

In the first chapter, we present the procedure to derive a constant drift B in a simplified two-dimensional radial cell while in the second chapter, we extend the computations to a radial dependent drift $b(r)$. In that chapter we also compute the drift in a cylindrical geometry that models the viral motion in neurites. Derivation is based on the following procedure : we consider a virus that diffuses (diffusion coefficient D) from a position \mathbf{x}_0 in the cell cytoplasm and we compute the mean first passage time $u(\mathbf{x}_0)$ and the mean location $\bar{\mathbf{x}}(\mathbf{x}_0)$ where it binds to a MT and begins active motion. We also consider the mean release location \mathbf{x}_f (see FIG. 9) from MTs network and in the small diffusion approximation $D \ll 1$ we have

$$\mathbf{b}(\mathbf{x}_0) \approx \frac{|\mathbf{x}_f - \mathbf{x}_0|}{u(\mathbf{x}_0) + t_m} \quad (55)$$

where t_m is the mean time of an active run along MTs. In a two dimensional radial geometry, we consider the fundamental domain $\tilde{\Omega}$ defined as the two dimensional slice of angle Θ between two neighboring microtubules. We consider here that microtubules are uniformly distributed and thus $\Theta = \frac{2\pi}{N}$, where N is the total number of microtubules (see figure 10). Because microtubules are taken uniformly distributed, we can always release the virus inside the slice $\tilde{\Omega}$, between two neighboring microtubules. Thus the movement of the virus will be studied in $\tilde{\Omega}$: inside the cytosol, the viral movement is purely Brownian until it hits a microtubule which is now the

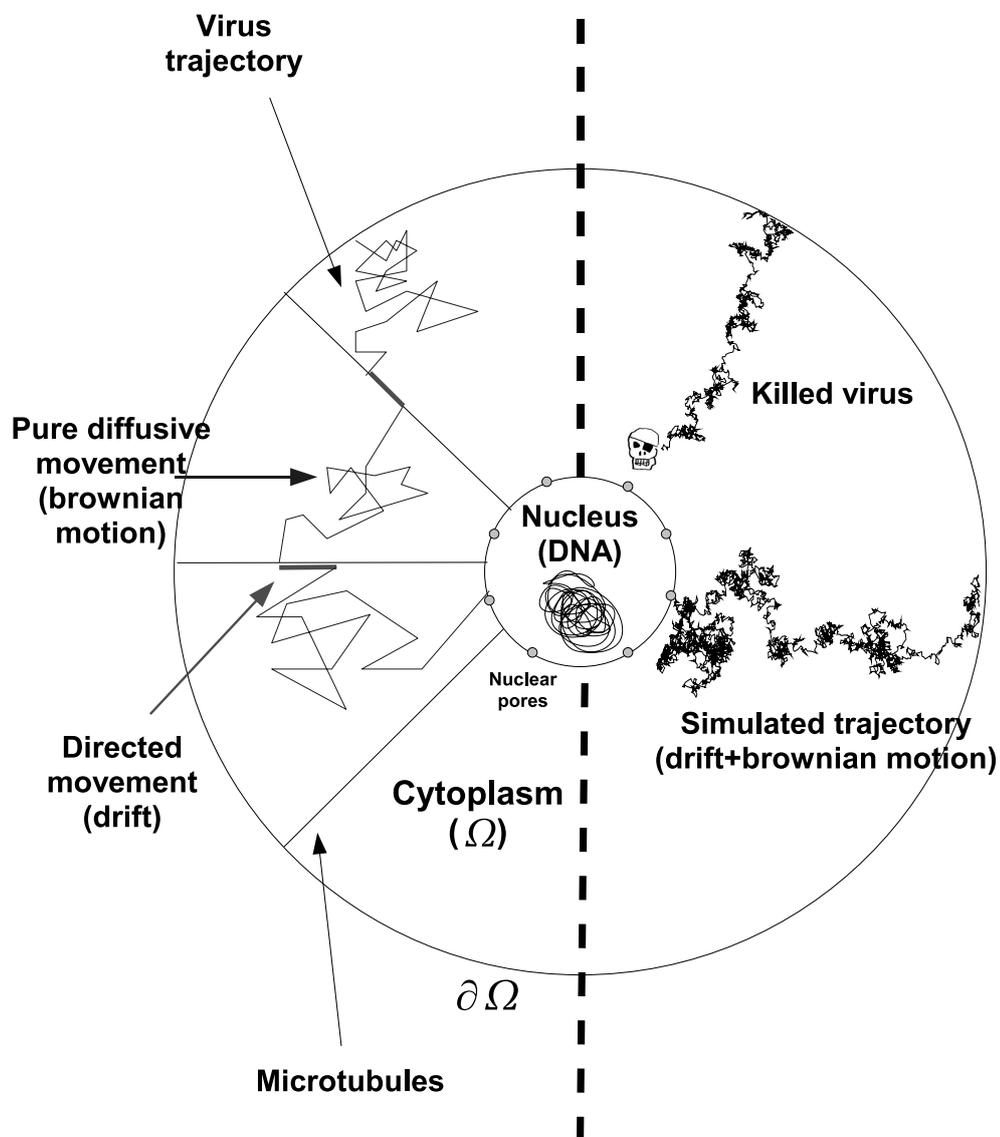


FIGURE 8 – Schematic representation of the viral trajectory approximation : on the left-side of the idealized cell, a real trajectory consists of intermittent Brownian and drift epochs, whereas on the right-side, we show two simulated trajectories obtained by equation (54). In one of them, the viral particle arrives alive to a nuclear pore, while in the other, it is killed inside the cytoplasm. The round dots on the nucleus surface represent nuclear pores.

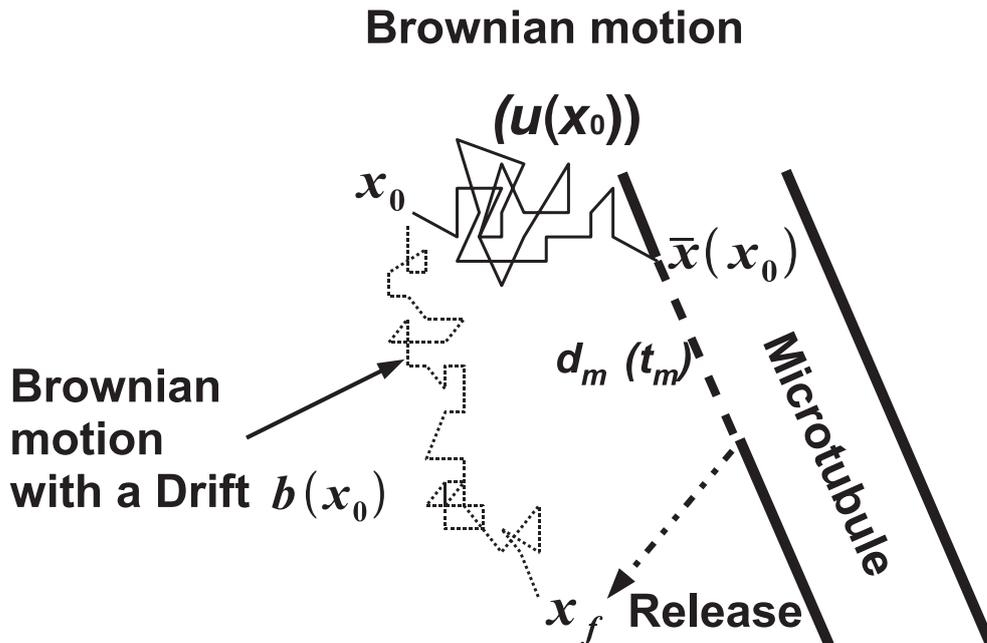


FIGURE 9 – The fundamental step is represented with a dotted line ; a virus starts at a position \mathbf{x}_0 , diffuses freely, binds to a MT over a distance d_m and is then released at a final position \mathbf{x}_f . The solid line represents a trajectory generated by the steady state equation (54). In the parenthesis, we point out the mean times for each portion of trajectories.

lateral boundary $\partial\tilde{\Omega}_a$ of $\tilde{\Omega}$ (see figure 10). When the virus binds to a MT, we simplify its complex bidirectional motion and processivity and we assume the bound particle moves towards the nucleus with the mean constant velocity V during a mean time t_m . The mean length of an active run is thus $d_m = Vt_m$. When the particle is released away from the microtubule, inside the domain, the process can start afresh and the particle diffuses freely. Because the Smoluchowski limit of the Langevin equation does not account for the change in velocity, we release the particle at a certain distance away from the microtubule, but at a fixed distance from the nucleus (at an angle chosen uniformly distributed), see figure 10. To summarize, the virus trajectory is a succession of diffusion steps mixed with some periods of attaching and detaching to microtubules. Thus scenario repeats until the virus hits a nuclear pore. The MFPT to a microtubule $u(\mathbf{x} = (r, \theta))$ of a virus starting initially at position $\mathbf{x} = (r, \theta)$ is solution of the Dynkin's equations [21]

$$D\Delta u(\mathbf{x}) = -1 \text{ for } \mathbf{x} \in \tilde{\Omega} \quad (56)$$

$$u(\mathbf{x}) = 0 \text{ for } \mathbf{x} \in \partial\tilde{\Omega}_a$$

$$\frac{\partial u}{\partial \mathbf{n}} = 0 \text{ for } \mathbf{x} \in \partial\tilde{\Omega}_{ext} ,$$

where $\partial\tilde{\Omega}_a = \{\theta = 0\} \cup \{\theta = \Theta\}$ and $\tilde{\Omega}_{ext} = \{r = R\}$ (see figure 10). For a virus that starts with an angle θ uniformly distributed in $[0; \Theta]$, the averaged MFPT to a

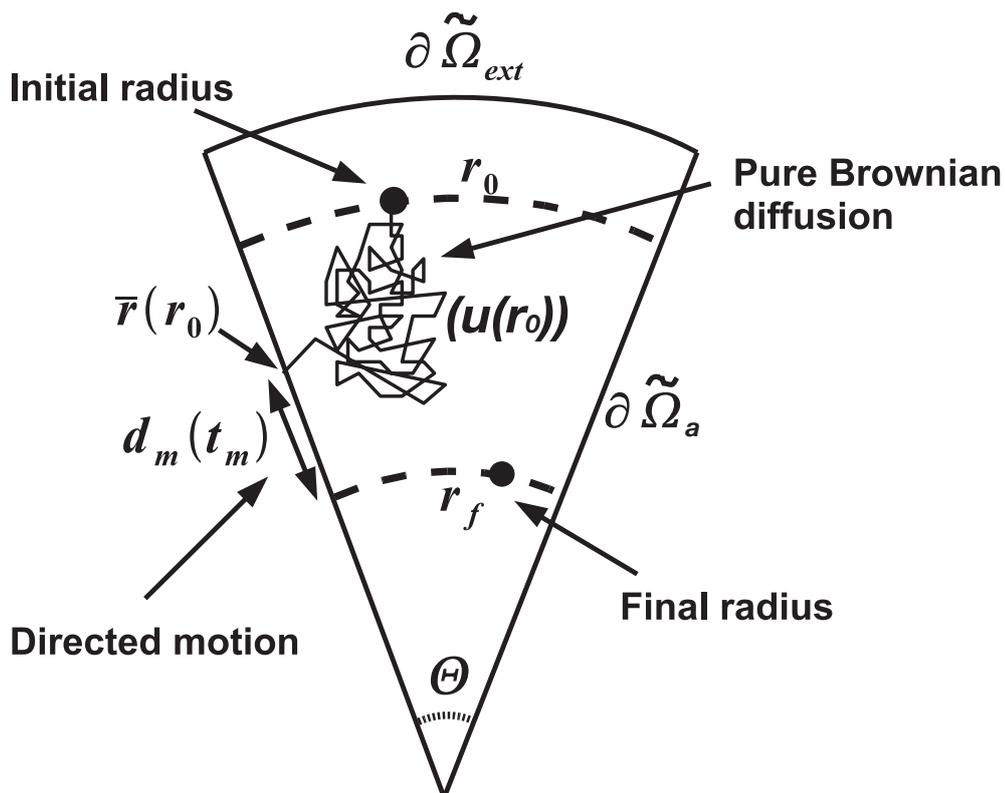


FIGURE 10 – A fundamental step in $\tilde{\Omega}$. The virus starts at a radius r_0 , with an angle uniformly distributed in $[0; \Theta]$, it diffuses freely during a time $u(r_0)$ until it binds to a MT ($\partial \tilde{\Omega}_a$) at a mean radius $\bar{r}(r_0)$; it has then a directed motion over a distance $d_m = Vt_m$ before being released randomly at a final radius r_f . Mean times of each piece of the fundamental step are written inside parenthesis.

one of the wedge is given by (see chapter 1)

$$\bar{u}(r) = \frac{1}{\Theta} \int_{\theta=0}^{\theta=\Theta} u(r, \theta) d\theta = \frac{r^2}{4D} \left(\frac{\tan(\Theta)}{\Theta} - 1 \right) - \sum_{n=0}^{\infty} \frac{16R^{2-\lambda_n} r^{\lambda_n}}{D\Theta^2 \lambda_n^3 (\lambda_n^2 - 4)}, \quad (57)$$

where $\lambda_n = (2n + 1) \frac{\pi}{\Theta}$. For Θ small, equation (57) can be approximated by

$$\bar{u}(r) = \frac{r^2 \Theta^2}{12D}. \quad (58)$$

To estimate the position $r_m(r_0)$ a virus that starts uniformly distributed at a distance r_0 from the cell center will attach preferentially to the microtubule, we determine the distribution of exit points $\epsilon(r|r_0, \theta_0)$. The probability density function (pdf) $p(\mathbf{x}, t|\mathbf{x}_0)$ to find a diffusing particle in a volume element $d\mathbf{x}$ at time t inside the wedge $\tilde{\Omega}$, conditioned on the initial position $\mathbf{x} = \mathbf{x}_0$ is the solution of the diffusion equation

$$\begin{aligned} \frac{\partial p(\mathbf{x}, t|\mathbf{x}_0)}{\partial t} &= D\Delta p(\mathbf{x}, t|\mathbf{x}_0) \text{ for } \mathbf{x} \in \tilde{\Omega} \\ p(\mathbf{x}, t|\mathbf{x}_0) &= 0 \text{ for } \mathbf{x} \in \partial\tilde{\Omega}_a \\ \frac{\partial p(\mathbf{x}, t|\mathbf{x}_0)}{\partial n} &= 0 \text{ for } \mathbf{x} \in \partial\tilde{\Omega}_{ext}, \end{aligned}$$

where the initial condition is $p(\mathbf{x}, 0|\mathbf{x}_0) = \delta(\mathbf{x} - \mathbf{x}_0)$. The distribution of exit points $\epsilon(\mathbf{y}|\mathbf{x}_0)$ is given by

$$\epsilon(\mathbf{y}|\mathbf{x}_0) = \int_0^{\infty} j(\mathbf{y}, t|\mathbf{x}_0) dt, \quad (59)$$

where the flux j is defined by

$$j(\mathbf{y}, t|\mathbf{x}_0) = -D \frac{\partial p(\mathbf{x}, t|\mathbf{x}_0)}{\partial n} \Big|_{\mathbf{x} = \mathbf{y}}.$$

Some computations lead to the averaged exit point distribution $\bar{\epsilon}(r|r_0) = \frac{1}{\Theta} \int_{\theta_0=0}^{\Theta} \epsilon(r|r_0, \theta_0) d\theta_0$ for a viral particle starting initially at position $\mathbf{x}_0 = (r_0, \theta_0)$ where θ_0 is uniformly distributed between 0 and Θ :

$$\bar{\epsilon}(r|r_0) = \frac{2}{\Theta\pi r} \left(\ln \left(\frac{r^\nu + r_0^\nu}{|r^\nu - r_0^\nu|} \right) + \ln \left(\frac{R^{2\nu} + (rr_0)^\nu}{R^{2\nu} - (rr_0)^\nu} \right) \right),$$

where $\nu = \frac{\pi}{\Theta}$. We define the mean exit radius as $r_m(r_0) = \mathbf{E}(\bar{\epsilon}(r|r_0))$. Thus,

$$r_m(r_0) = \mathbf{E}(r|r_0) = \int_0^R r \bar{\epsilon}(r|r_0) dr. \quad (60)$$

In the limit $\Theta \ll 1$ the leading order asymptotics of $r_m(r_0)$ is

$$r_m(r_0) \approx r_0 \left(1 + \frac{\Theta^2}{12} \right). \quad (61)$$

Finally, in a two-dimensional geometry, a virus starting at a distance r_0 from the center will be released at a mean final radius $r_f(r_0) = r_m(r_0) - d_m \approx r_0 \left(1 + \frac{\Theta^2}{12}\right) - d_m$ after a mean time $\bar{u}(r_0) + t_m \approx \frac{r_0^2 \Theta^2}{12D} + t_m$. The radius dependent drift $b(r)$ that accounts for ballistic periods along MTs is then approximately given by

$$b(r) = \frac{r - r_f(r)}{\bar{u}(r) + t_m} \approx \frac{d_m - r \frac{\Theta^2}{12}}{t_m + r^2 \frac{\Theta^2}{12D}}. \quad (62)$$

In the first chapter, the constant amplitude B of the radial steady state drift is obtained by an iterative method. After a virus has moved a certain distance along a microtubule, we assume it is released at a point uniformly distributed on the final radial distance from the nucleus, ready for a new random walk. This scenario repeats until the virus reaches the nucleus surface. We compute the mean number of steps and the subsequent mean time a virus reaches the nucleus. Finally, we deduce the amplitude of the effective drift using the following criteria : the Mean First Passage Time (MFPT) to the nucleus of the iterative approximation is equal to the MFPT obtained by solving directly an Ornstein-Uhlenbeck stochastic equation. In the limit $\Theta \ll 1$, the leading order term of the drift amplitude B is

$$B \approx \frac{\frac{d_m}{t_m}}{1 + \left(1 + \frac{R+\delta}{d_m}\right) \frac{\Theta^2}{24} + O(\Theta^4)}. \quad (63)$$

In chapter 1 and 2, we impose reflecting boundaries in the domain $\tilde{\Omega}$ and compare the steady state distribution obtained by running intermittent Brownian trajectories solutions of (53) with the theoretical ones in a Langevin description (54) for both the constant (63) and radial dependent (62) drifts we computed. In both cases, curves match very nicely, which is a central result of these chapters.

Many viruses such as herpes virus [23] travel in long axons or dendrites which can be approximated as thin cylinders (radius R and length L). We compute the steady state drift that accounts for the directed motion along MTs by modeling the N MTs parallel to the dendrite principal axis as cylinders (radius $\epsilon \ll R$, Length L). The cross-section Ω of the dendrite is shown in FIG. 11.

Due to the cylindrical symmetry, for any position \mathbf{x} , the steady state drift $\mathbf{b}(\mathbf{x})$ is equal to $B\mathbf{z}$ where B is a constant and \mathbf{z} the principal axis unit vector along the dendrite. In a small diffusion approximation ($D \ll 1$), the leading order term of B is equal to the effective velocity [24, 25] : $B = \frac{d_m}{t_m + \tau}$, where t_m is the mean time the virus binds to a MT, $d_m = Vt_m$ the mean length of a run and τ the MFPT to a MT. For small MTs radius $\epsilon \ll 1$, using that $\tau \approx \frac{1}{\lambda_0(\epsilon)}$ with $\lambda_0(\epsilon)$ the principal eigenvalue of the two-dimensionnal cross section of the dendrite with the N small absorbing MTs, we obtain that

$$B = \frac{d_m}{t_m + \tau} = \frac{2NDd_m}{2NDt_m + R^2 \ln\left(\frac{1}{\epsilon}\right)}. \quad (64)$$

Quantitative analysis of the free cytoplasmic step The probability P_n and the mean time τ_n to reach one of the n nuclear pores provide a global quantification of the cytoplasmic viral infection step. Because the radius of the absorbing pores

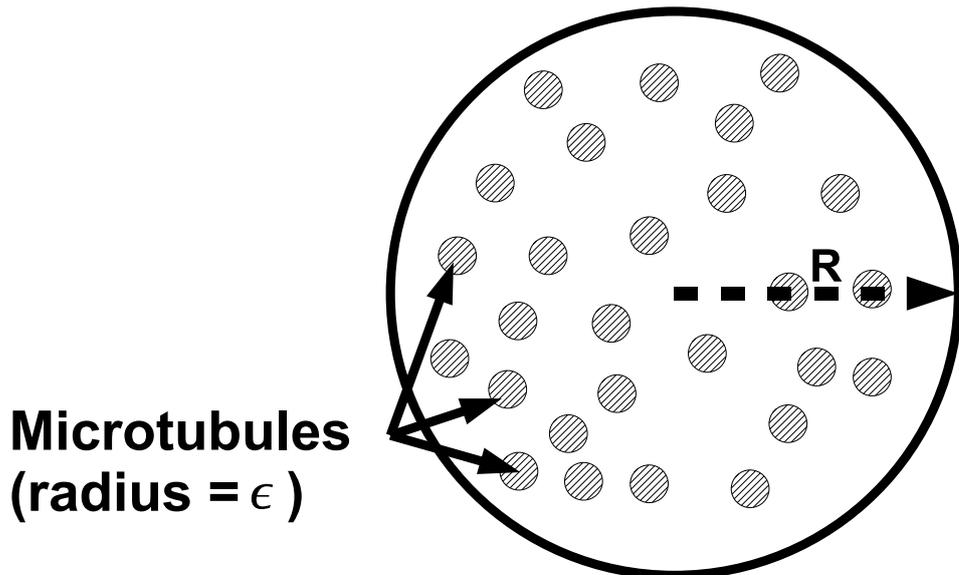


FIGURE 11 – Dendrite cross-section. The N MTs are thin cylinders uniformly distributed inside the dendrite.

ϵ is small, the derivation of τ_n and P_n is a narrow escape problem and derivations are based on singular perturbation theory. To derive such estimates, it is rational to start from the individual description equation (54). To describe the viral journey in the cytoplasm, we shall account for the viral degradation or immobilization, modeled by a steady state degradation rate $k(\mathbf{x})$. To describe the probability P_n that a single virus arrives to a small nuclear pore alive and the associated mean time τ_n , we shall first introduce the survival probability density function (SPDF) $p(\mathbf{x}, t)$. This is the probability to find the viral particle alive (not degraded) inside a cytoplasmic volume element $\mathbf{x} + d\mathbf{x}$ at time t . It is defined by [26],

$$p(\mathbf{x}, t)d\mathbf{x} = Pr\{X(t) \in \mathbf{x} + d\mathbf{x}, \tau^k > t, \tau^a > t | p_i\}, \quad (65)$$

where τ^a is the first time for a live virus to arrive to one of the nuclear pores area, denoted ∂N_a , τ^k the first time that it is degraded, and p_i is the viral initial distribution. The important and deep result [21] is that the SPDF $p(x, t)$ satisfies a partial differential equation, known as the Fokker-Planck equation (FPE)

$$\begin{aligned} \frac{\partial p}{\partial t}(\mathbf{x}, t) &= D\Delta p(\mathbf{x}, t) - \nabla \cdot \mathbf{b}(\mathbf{x})p(\mathbf{x}, t) - k(\mathbf{x})p(\mathbf{x}, t) \quad \text{for } \mathbf{x} \in \Omega \\ p(\mathbf{x}, 0) &= p_i(\mathbf{x}) \quad \text{for } \mathbf{x} \in \Omega, \end{aligned} \quad (66)$$

which describes how the probability to find a random particle evolves in time. The first term in the right-hand side is the contribution of the pure diffusion, the second term corresponds to the drift and the last term is coming from the degradation and says that at each moment of time, the particle can potentially be destroyed. To account for the boundary effect, we add the conditions

$$\begin{aligned} p(\mathbf{x}, t) &= 0 \quad \text{for } \mathbf{x} \in \partial N_a \\ \mathbf{J}(\mathbf{x}, t) \cdot \mathbf{n}_{\mathbf{x}} &= 0 \quad \mathbf{x} \in \partial\Omega - \partial N_a, \end{aligned} \quad (67)$$

where the first condition says that probability to find the particle on ∂N_a is zero. This is the part of the boundary where it is absorbed irreversibly. This condition is an idealized description of a nuclear pore where upon hitting this surface, the particle is instantaneously translocated to the nucleus with probability one and thus disappears from the cytoplasm. The second condition given on $\partial\Omega - \partial N_a$ is the remaining reflecting area of cell surface, described a reflected particle, \mathbf{n}_x is the unit outer normal at a boundary point \mathbf{x} . This second condition is defined by the flux density vector $\mathbf{J}(\mathbf{x}, t)$ as

$$\mathbf{J}(\mathbf{x}, t) = -D\nabla p(\mathbf{x}, t) + \mathbf{b}(\mathbf{x})p(\mathbf{x}, t). \quad (68)$$

and is zero when no viral particle penetrates the membrane surface. The interesting property is that the probability P_n that a live virus arrives at the nucleus and the conditional mean time τ_n can be expressed using the SPDF [26] :

$$P_n = 1 - \int_{\Omega} k(\mathbf{x})\tilde{p}(\mathbf{x}) d\mathbf{x}, \quad (69)$$

where $\tilde{p}(\mathbf{x}) = \int_0^{\infty} p(\mathbf{x}, t) dt$ is the solution of equation

$$D\Delta\tilde{p}(\mathbf{x}) - \nabla \cdot \mathbf{b}(\mathbf{x})\tilde{p}(\mathbf{x}) - k(\mathbf{x})\tilde{p}(\mathbf{x}) = -p_i(\mathbf{x}) \quad \text{for } \mathbf{x} \in \Omega$$

with the boundary conditions (67). While the conditional mean time τ_n is given by

$$\tau_n = \frac{\int_{\Omega} \tilde{p}(\mathbf{x}) d\mathbf{x} - \int_{\Omega} k(\mathbf{x})q(\mathbf{x}) d\mathbf{x}}{1 - \int_{\Omega} k(\mathbf{x})\tilde{p}(\mathbf{x}) d\mathbf{x}}, \quad (70)$$

where

$$q(\mathbf{x}) = \int_0^{\infty} sp(\mathbf{x}, s) ds \quad (71)$$

satisfies [26]

$$-\tilde{p} = D\Delta q(\mathbf{x}) - [\nabla \cdot \mathbf{b}q] - kq \quad \text{for } \mathbf{x} \in \Omega \quad (72)$$

with boundary conditions (67). When the drift \mathbf{b} is derived from a potential Φ : $\mathbf{b} = -\nabla\Phi$, for a small degradation rate $k \ll 1$, leading order asymptotics of P_n and τ_n for a small radius $\epsilon \ll 1$ have been computed in [26]

$$\left\{ \begin{array}{l} P_n = \frac{e^{-\frac{\Phi_0}{D}}}{\frac{1}{4Dn\epsilon} \int_{\Omega} e^{-\frac{\Phi(\mathbf{x})}{D}} k(\mathbf{x}) d\mathbf{x} + e^{-\frac{\Phi_0}{D}}} \\ \tau_n = \frac{\frac{1}{4Dn\epsilon} \int_{\Omega} e^{-\frac{\Phi(\mathbf{x})}{D}} d\mathbf{x}}{\frac{1}{4Dn\epsilon} \int_{\Omega} e^{-\frac{\Phi(\mathbf{x})}{D}} k(\mathbf{x}) d\mathbf{x} + e^{-\frac{\Phi_0}{D}}} \end{array} \right. \quad (73)$$

where Φ_0 is the constant value of the radial potential $\Phi(\mathbf{x})$ on the centered nucleus where the nuclear pores are uniformly distributed. These formulas do not

account for the possible interactions between the small absorbing pores. Because $\lim_{n \rightarrow \infty, n\epsilon^2 \ll 1} \tau_n = 0$, when the number of pores become too large, these expressions are no more valid and a correction term accounting for the holes interactions is needed. Recent studies have begun to quantify the interactions between the absorbing windows [27, 28]. In the chapter 3, we quantify how these interactions impact P_n and τ_n . In particular, when $n \gg 1$ holes are uniformly distributed over a structure Σ such as the nucleus in the cytoplasm, we found that in the limit $\frac{|\Sigma|}{|\Omega|} \ll 1$

$$\left\{ \begin{array}{l} P_n = \frac{\int_{\partial\Sigma} e^{-\frac{\Phi(\mathbf{x})}{D}} d\mathbf{x}}{|\partial\Sigma|} \\ \left(\frac{1}{4nD\epsilon} + \frac{1}{DC_\Sigma} \right) \int_{\Omega} k(\mathbf{x}) e^{-\frac{\Phi(\mathbf{x})}{D}} d\mathbf{x} + \frac{\int_{\partial\Sigma} e^{-\frac{\Phi(\mathbf{x})}{D}} d\mathbf{x}}{|\partial\Sigma|} \\ \tau_n = \frac{\left(\frac{1}{4nD\epsilon} + \frac{1}{DC_\Sigma} \right) \int_{\Omega} e^{-\frac{\Phi(\mathbf{x})}{D}} d\mathbf{x}}{\left(\frac{1}{4nD\epsilon} + \frac{1}{DC_\Sigma} \right) \int_{\Omega} k(\mathbf{x}) e^{-\frac{\Phi(\mathbf{x})}{D}} d\mathbf{x} + \frac{\int_{\partial\Sigma} e^{-\frac{\Phi(\mathbf{x})}{D}} d\mathbf{x}}{|\partial\Sigma|}} \end{array} \right. \quad (74)$$

where C_Σ is the electrostatic capacitance of the conducting surface $\partial\Sigma$. For example, if the structure is a spherical nucleus of radius δ , we have $C_\Sigma = 4\pi\delta$. When the ratio $\sigma = \frac{n\pi\epsilon^2}{|\partial\Sigma|}$ of $\partial\Sigma$ covered by the absorbing windows is neglected, for pure diffusing particles ($\Phi(\mathbf{x}) = 0$) with no degradation activity ($k = 0$), the mean time (74) reduces to the asymptotic formula (54) in [29] derived with electrostatic considerations. A single virus is sufficient to infect a cell and replicate. Consequently, when M independent viruses enter a cell, it is important to have quantitative information about the first virus that reaches a nuclear pore to deliver its payload. In the chapter 4, we thus compute the conditioned MFPT $\tau_{first}(M)$ for the first virus among M to reach a nuclear pore. The M -viruses trajectories are independent and we shall use the conditioned MFPT τ_N^j of the j^{th} carrier to a nuclear pore. As in [26], we consider the absorbing time $\tau_{first}^a(M)$ of the first DNA carrier to the absorbing boundary ∂N_a and the first time $\tau_{first}^k(M)$ it is degraded. The probability the first DNA carrier arrives to the absorbing boundary before time t conditioned on not been killed is then given by :

$$P(t) = Pr\{\tau_{first}^a(M) < t | \tau_{first}^a(M) < \tau_{first}^k(M), p_i\}. \quad (75)$$

The conditional MFPT $\tau_{first}(M)$ is defined by

$$\tau_{first}(M) = \int_0^\infty t \frac{dP(t)}{dt} dt = \int_0^\infty (P(\infty) - P(t)) dt. \quad (76)$$

Some computations lead to

$$P(t) = \frac{1 - \left(1 - \int_0^t J(s) ds\right)^M}{1 - (1 - P_N)^M}. \quad (77)$$

where $J(s) = \oint_{\partial\Omega} \mathbf{J}(\mathbf{x}, t) \cdot \mathbf{n}_x dS_x$, with \mathbf{n}_x the normal derivative at the boundary point \mathbf{x} and $\mathbf{J}(\mathbf{x}, t)$ the flux density vector defined in (68). Hereafter, we shall estimate the

leading order term for $\tau_{first}(M)$. In the long time asymptotic ($k \ll 1$ and $\epsilon \ll 1$), we approximate $p(\mathbf{x}, t)$ by its first exponential term

$$p(\mathbf{x}, t) \approx p(\mathbf{x}, 0)e^{-\lambda_0 t}, \text{ with } \int_{\Omega} p(\mathbf{x}, 0)d\mathbf{x} = 1. \quad (78)$$

where $\lambda_0 = \frac{1}{\tau_n}$ ([21] p.175), is the first eigenvalue (this implies that there is no contribution of the initial condition on the other eigenfunctions, see also [30]). This long time asymptotic lead to the flux expression

$$J(t) = \frac{P_n}{\tau_n} e^{-\frac{t}{\tau_n}}. \quad (79)$$

Replacing $\int_0^t J(s)ds$ by its approximation (79) in relation (76) we get

$$\tau_{first}(M) = \int_0^{\infty} \frac{\left(1 - P_n \left(1 - e^{-\frac{t}{\tau_n}}\right)\right)^M - (1 - P_n)^M}{1 - (1 - P_n)^M} dt. \quad (80)$$

And some simplifications lead to the concise expression

$$\tau_{first}(M) = \frac{\tau_n(\xi)}{1 - \xi^M} \left(\sum_{k=0}^{M-1} (\xi^k - \xi^M) \frac{1}{M - k} \right). \quad (81)$$

where $\xi = 1 - P_n$ ($0 \leq \xi \leq 1$). When the number of viruses reaching alive a nuclear pore decreases, the MFPT of the first survivor increases. For small ξ , the leading order term of $\tau_{first}(M)$ is

$$\frac{\tau_{first}(M)}{\tau_n(\xi)} \approx \frac{1}{M}, \quad (82)$$

whereas when ξ tends to 1, that is when most of viruses are killed before reaching their goal, we get the approximation

$$\tau_{first}(M) \approx \tau_n(\xi). \quad (83)$$

In the computations above, we assumed the degradation activity was small compared to the diffusion time scale. We thus solved narrow escape problems with long time asymptotics. Pure diffusing synthetic gene vectors ($\mathbf{b} = \mathbf{0}$) are frequently embedded in the crowded cytoplasm and the killing activity due to proteases is not that small. In the chapter 4, we thus derive leading order asymptotics of P_n and τ_n in the limit $k \gg 1$. The analysis is quite different from above and is based on matched asymptotic expansions. We consider a uniform initial plasmid distribution over the cytoplasm $p_i(\mathbf{x}) = p_0 = \frac{1}{|\Omega|}$. To compute the probability P_N given by (69), we shall solve equation (70)

$$D\Delta\tilde{p}(\mathbf{x}) - k(\mathbf{x})\tilde{p}(\mathbf{x}) = -p_0 = -\frac{1}{|\Omega|}, \quad (84)$$

with the boundary conditions (67). When $\frac{D}{|\Omega|}$ is much smaller compared to k and for a particle starting far from nuclear pores, we approximated the solution of Eq.(84) by

$$p_{outer}(\mathbf{x}) = \frac{p_0}{k(\mathbf{x})} + O(D). \quad (85)$$

However, this outer solution does not match the absorbing conditions. We thus construct two inner solutions : the first one p_{inner}^1 near ∂N_a and the second p_{inner}^2 near $\partial N_r = \partial\Omega - \partial N_a$ and match the outer solution. In a local coordinates (ρ, s) near ∂N_a , where ρ measures distance from ∂N_a , measured positively into Ω , and s are tangential variables in the plane $\rho = 0$, we expand the steady state radial killing measure along the radial ρ -coordinate,

$$k(\rho, s) = k_0(s) + k_1(s)\rho + O(\rho^2). \quad (86)$$

where $k(\rho = 0, s) = k_0(s)$ and $\frac{dk}{d\rho}(\rho = 0, s) = k_1(s)$. We use the change of variable

$$u = u(\rho, s) = \frac{k_0(s) + k_1(s)\rho}{\beta(s)D}, \text{ where } \beta(s) = \left(\frac{k_1(s)}{D}\right)^{\frac{2}{3}}. \quad (87)$$

and some computations lead to

$$p_{inner}^1(u, s) = \frac{\pi}{|\Omega|\beta(s)D} \left(Gi(u) - \frac{Gi\left(\frac{k_0(s)}{\beta(s)D}\right)}{Ai\left(\frac{k_0(s)}{\beta(s)D}\right)} Ai(u) \right), \quad (88)$$

and

$$p_{inner}^2(u, s) = \frac{\pi}{|\Omega|\beta(s)D} \left(Gi(u) - \frac{Gi'\left(\frac{k_0(s)}{\beta(s)D}\right)}{Ai'\left(\frac{k_0(s)}{\beta(s)D}\right)} Ai(u) \right). \quad (89)$$

where A_i and G_i are respectively the Airy and Scorer functions ([31], p.446 and 448). These inner solutions satisfy boundary conditions and match p_{outer} . For a sufficiently smooth killing field, when ∂N_a consists of n well separated small absorbing nuclear pore (radius $(\epsilon_q)_{1 \leq q \leq n}$) centered at $(\mathbf{x}_q)_{1 \leq q \leq n}$ on $\partial\Omega$, injecting (85), (88) and (89) in (69), we show that the leading order asymptotic of P_n is :

$$P_n \approx \sum_{q=1}^n \frac{\pi\eta_q^2}{|\Omega|} \sqrt{\frac{D}{k(x_q)}} + O\left(e^{-\sqrt{\frac{k_0}{D}}\rho_0}\right). \quad (90)$$

where $k_0 = \inf_q k(\mathbf{x}_q)$ and $\rho_0 = \inf_{s \in \partial N_a} \rho_0(s)$. Concerning the conditioned MFPT τ_n , very similar computations lead to (see details in chapter 4)

$$\tau_n \approx \frac{1}{n} \sum_{q=1}^n \frac{1}{2k(x_q)} + O\left(e^{-\sqrt{\frac{k_0}{D}}\rho_0}\right). \quad (91)$$

In the chapter 4, we test the asymptotic formulas (90) and (91) against Brownian simulations. The matching occurs for a very large degradation rate (more than 200 times the rate observed experimentally [32]) that shall characterize abnormal cells.

6 Conclusion

DNA viruses have developed molecular tools to efficiently transfer their genes to the cell nucleus and a quantitative description of the early steps of viral infection

at the single molecule level will help to design new drugs and optimize non viral gene carriers. Due to a complex host cell-virus communication, the viral particle continuously undergoes structural rearrangements that modulate its behavior in the different compartments of the cell, and consequently, models at the single virus level rather than at a population level are required to account for these individual variations in activity. We thus propose to divide the early steps of infection in three functional modules that impact the viral structure and trajectory : the receptor-mediated entry of the virus, the endosomal routing and the free cytoplasmic motion to a nuclear pore. While viruses reliably escape the endosomes and efficiently traffic to a nuclear pore once they are released in the cytoplasm, non-viral gene vectors mostly failed to escape the endosome and in addition, the vectors that reach the cell cytoplasm are mainly degraded by proteases or trapped in the crowded cytoplasm. To understand the high efficiency of viruses in both of these steps, we have constructed biophysical models at the molecular level to estimate how general quantitative outputs such as the mean escape time from the endosome or the probability a virus released in the cytoplasm reaches a nuclear pore before being degraded by proteasomes depend on the geometry of the cell (ligands influx rate in the endosome, organization of the MTs network, size and number of nuclear pores ...) and on the dynamical parameters of the virus (affinities of its active proteins to endosomal ligands, diffusion constant, unbinding rate from MTs ...). Interestingly, we found that the escape dynamics does not depend on the endosomal size and that the fusion and fission events that have been observed *in vivo* [2] should be neglected in the modeling, that the optimal number of viruses in the endosome to escape in a certain range of pH is 5 and that the number of nuclear pores more than the surface covered by all the pores strongly affect the probability and the conditioned MFPT to a nuclear pore. Finally, to quantify infection success, the three functional modules shall be coupled, the state of the virus at the end of a module serving as input parameter for next. For example, the degradation rate used in module 3 is an input parameter that depends on the state of viral ubiquitination in the cytoplasm, which depends on the pH course during the endosomal phase, an output of module 2.

Chapitre 1

Effective motion of a virus trafficking inside a biological cell, *SIAM* *Journal of Applied Mathematics* 68 (2008)

1 Introduction

Because cytosolic transport has been identified as a critical barrier for synthetic gene delivery [50], plasmids or viral DNAs delivery from the cell membrane to the nuclear pores has attracted the attention of many biologists. The cell cytosol contains many types of organelles, actin filaments, microtubules and many others, so that to reach the nucleus, a viral DNA has to travel through a crowded and risky environment. We are interested here in studying the efficiency of the delivery process and we present a mathematical model of virus trafficking inside the cell cytoplasm. We model the viral movement as a Brownian motion. However, the density of actin filaments and microtubules, inside the cell, can hinder diffusion, as demonstrated experimentally [4]. In a crowded environment, we will model the virus as a material point. This reduction is simplistic for several reasons : actin filament network can trapped a diffusing object and beyond a certain size, as observed experimentally, a DNA fragment cannot find its way across the actin filaments [4]. Active directional transport along microtubules or actin filaments seems then the only way to deliver a plasmid to the nucleus. The active transport of the virus involves in general motor proteins, such as Kinesin (to travel in the direction of the cell membrane) or Dynein (to travel toward the nucleus). Once a virus is attached to a Dynein protein, its movement can be modeled as a determinist drift toward the nucleus.

Recently, a macroscopic modeling has been developed to describe the dynamics of adenovirus concentration inside the cell cytoplasm [48]. This approach offers very interesting results about the effect of microtubules, but neglects the complexity of the geometry and cannot be used to describe the movement of a single virus, which might be enough to cause cellular infection. Modeling a virus trafficking imposes to use a stochastic description. We model here the motion of a virus as that of a material point, so the probability of its trapping by actin filaments or microtubules is neglected. In the present approximation, the viral movement has two main com-

ponents : a Brownian one, which accounts for its free movement, and a drift directed towards the centrosome or MTOC (Microtubules Organization Center), an organelle located near the nucleus. The magnitude of the drift along microtubules depends on many parameters, such as the binding and unbinding rates and the velocity of the motor proteins [51].

In the present approach, we present a method to approximate a time dependent dynamics of virus trafficking by an effective stochastic equation with a radial steady state drift. The main difficulties we have to overcome arise from the time dependent nature of the trajectories which consists of intermittent epochs of drifts and free diffusion. We propose to derive an explicit expression for the steady state drift amplitude. In this approximation, the effective drift will gather the mean properties of the cytoplasmic organization such as the density of microtubules and its off binding rate.

Our method to find the effective drift can be described as follow : first, we approximate the cell geometry as a two dimensional disk and use a pure Brownian description to approximate the virus diffusion step. This geometrical approximation is valid, for any two dimensional cell such as the *in vitro* flat skin fibroblast culture cells [48] : indeed, due to their adhesion to the substrate, the thickness of these cells can be neglected in first approximation. Second, when the distribution of the initial viral position is uniform on the cell surface, we will estimate, during the diffusing period, the hitting position on a microtubule. By solving a partial differential equation, inside a sliced shape domain, delimited by two neighboring microtubules, we will provide an estimate of the mean time to the most likely hitting point. Finally, the amplitude of the radial steady state drift will be obtained by an iterative method which assumes that, after a virus has moved a certain distance along a microtubule, it is released at a point uniformly distributed on the final radial distance from the nucleus, ready for a new random walk. This scenario repeats until the virus reaches the nucleus surface. Finally, we will compute the mean time, the mean number of steps before a virus reaches the nucleus and the amplitude of the effective drift by using the following criteria : the Mean First Passage Time (MFPT) to the nucleus of the iterative approximation is equal to the MFPT obtained by solving directly an Ornstein-Uhlenbeck stochastic equation. The explicit computation of the effective drift is a key result in the estimation of the probability and the mean time a single virus or DNA molecule takes to reach a small nuclear pore [26].

2 Modeling stochastic viral movement inside a biological cell

We approximate the cell as a two dimensional geometrical domain Ω , which is here a disk of radius R and the nucleus located inside is a concentric disk of much smaller radius $\delta \ll R$. We model the motion of an unattached DNA fragment as a material point, so that the probability of its trapping by actin filaments or microtubules is neglected. The motion of a (DNA) molecule of mass m is described by the overdamped limit of the Langevin equation (Smoluchowski's limit) [21] for the position $\mathbf{X}(t)$ of the molecule at time t . When the particle is not bound to a microtubule filament, its movement is described as pure Brownian with a diffusion

constant D . When the particle hits a filament, it binds for a certain random time and moves along with a determinist drift. We only take into account the movement toward the nucleus, which is confound here with the MTOC (Microtubule organization center), an organelle where all microtubules converge (see figure (1.1)). For $\delta < |\mathbf{X}(t)| < R$, we describe the overall movement by the stochastic rule

$$\dot{\mathbf{X}} = \begin{cases} \sqrt{2D}\dot{\mathbf{w}} & \text{for } \mathbf{X}(t) \text{ free} \\ V \frac{\mathbf{r}}{|\mathbf{r}|} & \text{for } \mathbf{X}(t) \text{ bound} \end{cases} \quad (1.1)$$

where V is a constant velocity, $\dot{\mathbf{w}}$ a δ -correlated standard white noise and \mathbf{r} the \mathbf{X} radial coordinate, the origin of which is the center of the cell. We assume that all filaments starting from the cell surface end on the nucleus surface. The binding time corresponds to a chemical reaction event and we assume that it is exponentially distributed and for simplicity we approximate it by a constant t_m .

Once a virus enters the cell membrane, its moves according to the rule (1.1), until it hits a nuclear pore. Although nuclear pores occupy a small portion of the nuclear surface, we only consider the virus movement until it hits the nuclear surface $D(\delta)$. In this article, our goal is to replace equation (1.1) by a steady state stochastic equation

$$\dot{\mathbf{X}} = \mathbf{b}(\mathbf{X}) + \sqrt{2D}\dot{\mathbf{w}}, \quad (1.2)$$

where the drift \mathbf{b} is radially symmetric. In a first approximation, we consider a constant radial drift $\mathbf{b}(\mathbf{X}) = -B \frac{\mathbf{r}}{|\mathbf{r}|}$ and compute hereafter the value of the constant amplitude B such that the MFPT of the process (1.2) and (1.1) to the nucleus are equal.

2.1 Modeling viral dynamics in the cytoplasm

Inside the cytosol, microtubules are distributed on the cell surface and converging radially to the MTOC. We denote by ρ this distribution (see figure (1.1)). We do not take into account in the present analysis, the effect of organelle crowding due to the endoplasmic reticulum, the Golgi apparatus and many others. However, it is always possible to include them indirectly by using an apparent diffusion constant. We consider the fundamental domain $\tilde{\Omega}$ defined as the two dimensional slice of angle Θ between two neighboring microtubules. We consider here that microtubules are uniformly distributed and thus $\Theta = \frac{2\pi}{N}$, where N is the total number of microtubules.

Although a virus can drift along microtubules in both directions by using dynein (resp. kinesin) motor proteins for the inward (resp. forward) movement, we only take into account the drift toward the nucleus [52]. It is still unclear what is the precise mechanism used by a virus to select a direction of motion. Attached to a dynein molecule, the virus transport consists in several steps of few nanometers : the length of each step depends on the load of the transported cargo and ATP-concentration [53]. We neglect here the complexity of this process, assuming that ATP molecules are abundant, uniformly distributed over the cell and is not a limiting factor. We thus assume the bound particle moves towards the nucleus with the mean constant velocity V . When the particle is released away from the microtubule, inside the

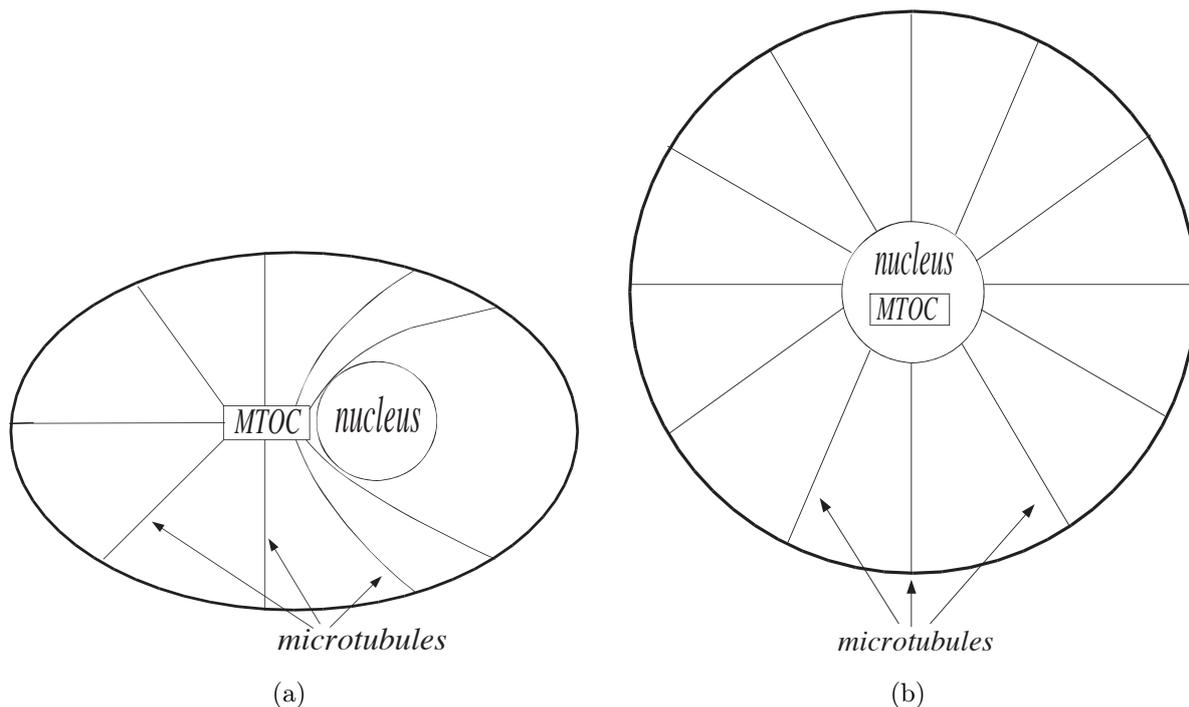


FIGURE 1.1 – **Cell geometry.** (a) Cell’s microtubules network. All microtubules starting from the cell membrane converge to the Microtubule Organization center (MTOC), located near the nucleus. (b) simplified cell’s microtubules network organization. The MTOC coincides with the nucleus.

domain, the process can start afresh and the particle diffuses freely. Because the Smoluchowski limit of the Langevin equation does not account for the change in velocity, we release the the particle at a certain distance away from the microtubule, but at a fixed distance from the nucleus (at an angle chosen uniformly distributed), see figure 1.2.

Because microtubules are taken uniformly distributed, we can always release the virus inside the slice $\tilde{\Omega}$, between two neighboring microtubules. Thus the movement of the virus will be studied in $\tilde{\Omega}$: inside the cytosol, the viral movement is purely Brownian until it hits a microtubule which is now the lateral boundary of $\tilde{\Omega}$ (see figure (1.2)). We assume that once a virus hits a microtubule, with probability one, the dynamics switches from diffusion to a determinist motion with a constant drift. A virus spends on a microtubule a time that we consider to be exponentially distributed, since this time is the sum of escape time from deep potential wells. We approximate the total time on a microtubule by the mean time t_m . Thus a virus moves to a distance $d_m = Vt_m$ along microtubule, which depends only on the characteristic of the virus-microtubule interactions. To summarize, the virus trajectory is a succession of diffusion steps mixed with some periods of attaching and detaching to microtubules. Thus scenario repeats until the virus hits the nucleus surface (Figure (1.2)).

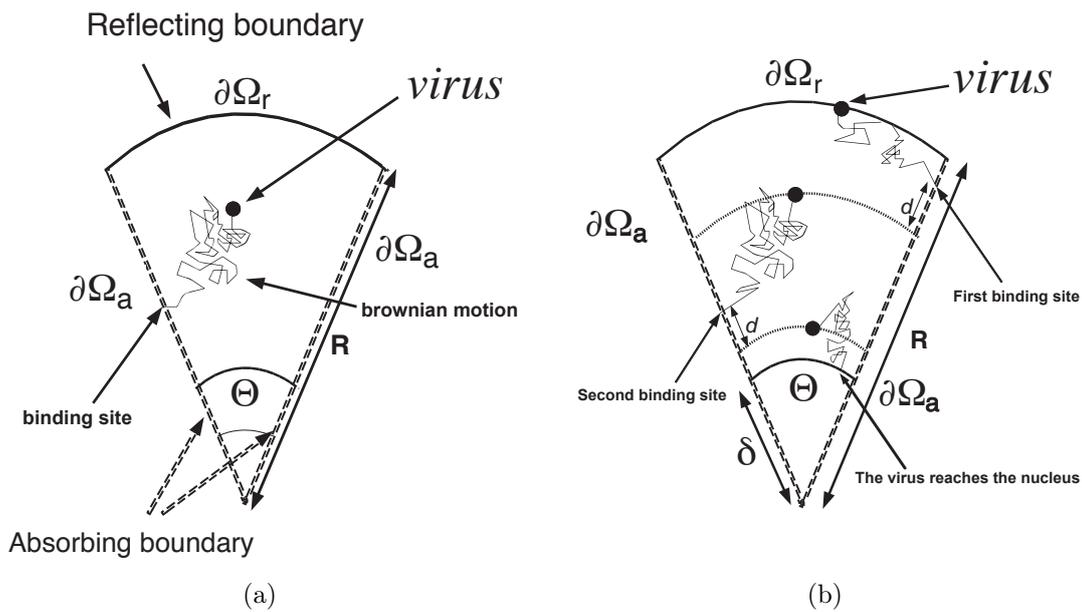


FIGURE 1.2 – **Virus trafficking inside a cell.** (a) Representation of the cell portion between two microtubules. (b) Transport along microtubules : Two fundamental steps are represented. A fundamental step is made of the two intermediate step which are first the diffusion inside the domain followed by the directed motion along the microtubule.

2.2 Computing the MFPT to reach the nucleus

We define the *mean time to infection* as the MFPT a virus reaches the surface of the disk $D(\delta)$ inside the domain $\tilde{\Omega}$ (see figure (1.2)).

To estimate the mean time to infection, we note that we can decompose the overall motion as a repeated fundamental step. This step consists of the free diffusion of the particle inside the domain followed by the motion along the microtubule. The total time of infection τ_i is then the sum of times the particle spends in each step. Although the time on the microtubule is deterministic and equal to t_m , the diffusing time is not easy to compute and depend on the initial condition. Ultimately τ_i depends on the number of times the fundamental step repeats before the particle reaches the nucleus.

Let us now described each step : the first step starts when the virus enter the cell at the periphery $r = R = R_0$ (at a random angle $\theta \in [0; \Theta]$) and ends when the virus hits either the lateral boundary or the nucleus. We now consider the first passage time $u(R_0)$ to the absorbing boundary and by $r(R_0)$ the hitting position. To account for the determinist drift, we move during a determinist time t_m the virus from a distance d_m along the microtubule. In that case, the initial random position for the next step is given by $r = R_1 = r(R_0) - d_m$ and the total time in step 1 is $u(R_0) + t_m$.

We iterate the process as follow and consider in each step k the distance $R_k = r(R_{k-1}) - d_m$ from which the particle starts and the time $u(R_k) + t_m$ it spends inside the step. If we denote by n_s the random number of steps necessary to reach

the nucleus $r = \delta$, the time to infection τ_i is given by

$$\tau_i = \sum_{k=0}^{n_s-1} u(R_k) + n_s t_m + t_r, \quad (1.3)$$

where t_r is a residual time, which is the time to reach the nucleus before a full step is completed.

We are interested in the estimating the mean first passage MFPT τ of τ_i , given by

$$\tau = E(\tau_i) = E\left(\sum_{k=0}^{n_s-1} u(R_k)\right) + \langle n_s \rangle t_m + \langle t_r \rangle, \quad (1.4)$$

where $\langle n_s \rangle$ is the mean number of steps and $\langle t_r \rangle$ is the mean residual time. If we introduce the probability density function $p_m = Pr\{n_s = m\}$ that the number of step is exactly equal to m , we can write

$$\tau = E(\tau_i) = \sum_{m=1}^{\infty} E\left(\sum_{k=0}^{n_s-1} u(R_k) | n_s = m\right) p_m + \langle n_s \rangle t_m + \langle t_r \rangle, \quad (1.5)$$

To estimate the MFPT τ , we shall approximate the previous sum by using the mean first passage time $\bar{u}(R_k)$ in each step k . To estimate $\bar{u}(R_k)$, we will solve (in the next paragraph) the Dynkin's equation with the following boundary conditions : inside $\tilde{\Omega}$, the particle is reflected at the periphery $r = R$, absorbed at the nucleus $\partial\tilde{\Omega}_a$ and at $\theta = 0$ and $\theta = \Theta$. We will also estimate the mean distance \bar{d}_k covered during step k . For that purpose we will estimate the mean exit position $r_m(R_k)$, conditioned on the initial position $r = R_k$. Indeed, we will thus get $\bar{d}_k = R_k - r_m(R_k) - d_m$. The estimates of the mean distances covered for each fundamental step will ultimately lead to an approximation of the mean number of step $n = \langle n_s \rangle$: n will be computed such that $R_n \geq \delta$ and $R_{n+1} < \delta$ (where $R_n = r_m(R_{n-1}) - d_m$ is defined recursively). Finally, we will obtain the following approximation for the infection time

$$\tau \approx \sum_{k=0}^{n-1} \bar{u}(R_k) + n t_m + \langle t_r \rangle, \quad (1.6)$$

The mean residual time $\langle t_r \rangle$ can be equal either to $\bar{u}(R_n) + \alpha t_m$, where $0 \leq \alpha < 1$ if the virus binds to a microtubule in the last step and travels a distance αd_m on the microtubule, or to the MFPT to the nuclear boundary if $r_m(R_n) < \delta$.

3 Mean First Passage Time and Exit point distribution

In first approximation, under the assumptions of a sufficiently small radius $\delta \ll R$ and an angle $\Theta \ll 1$, for the computation of the MFPT and the distribution of exit points, we neglect the nuclear area. We define the full pie wedge Ω^R domain of angle Θ . Inside Ω^R , we use the boundary conditions described above. Consequently,

the MFPT to a microtubule $u = u(r, \theta)$ of a virus starting initially at position (r, θ) is solution of the Dynkin's equations [21]

$$D\Delta u(\mathbf{x}) = -1 \text{ for } \mathbf{x} \in \Omega^R \tag{1.7}$$

$$u(\mathbf{x}) = 0 \text{ for } \mathbf{x} \in \partial\Omega_a^R$$

$$\frac{\partial u}{\partial \mathbf{n}} = 0 \text{ for } \mathbf{x} \in \partial\Omega_r^R,$$

where $\partial\Omega_a^R = \{\theta = 0\} \cup \{\theta = \Theta\}$ and $\Omega_r^R = \{r = R\}$.

3.1 The general solution for the MFPT

In this paragraph only we reparametrize the domain by $-\Theta/2 \leq \theta \leq \Theta/2$. By writing equation (1.7) in polar coordinates and using the separation of variables, the general solution of equation

$$\left(\frac{\partial^2 u}{\partial r^2} + \frac{1}{r} \frac{\partial u}{\partial r} + \frac{1}{r^2} \frac{\partial^2 u}{\partial \theta^2} \right) (r, \theta) = -1 \text{ for } (r, \theta) \in \Omega^R \tag{1.8}$$

$$u(r, \theta) = 0 \text{ for } (r, \theta) \in \partial\Omega_a^R. \tag{1.9}$$

is given by [54]

$$u(r, \theta) = \frac{r^2}{4D} \left(\frac{\cos(2\theta)}{\cos(\Theta)} - 1 \right) + \sum_{n=0}^{\infty} A_n r^{\lambda_n} \cos(\lambda_n \theta), \text{ for } -\Theta/2 \leq \theta \leq \Theta/2 \tag{1.10}$$

where the edge boundary is here located at position $\theta = \pm\Theta/2$. The sum in the right-hand side is the general solution of the homogeneous problem $\Delta u = 0$ in Ω^R . The boundary conditions on the sides of the wedge impose that

$$\lambda_n = (2n + 1) \frac{\pi}{\Theta}, \tag{1.11}$$

while the reflecting condition for $r = R$ reads

$$\frac{\partial u}{\partial r}(R, \theta) = 0 \text{ for all } \theta \in [-\Theta/2, \Theta/2]. \tag{1.12}$$

Using the uniqueness of Fourier decomposition and the boundary condition (1.12), we obtain that

$$A_n = \frac{(-1)^{n+1} 8R^{2-\lambda_n}}{D\Theta\lambda_n^2 (\lambda_n^2 - 4)}. \tag{1.13}$$

By averaging formula (1.10) over an initial uniform distribution, the MFPT of the particle to a microtubule is given by

$$\bar{u}(r) = \frac{1}{\Theta} \int_{\theta=0}^{\theta=\Theta} u(r, \theta) d\theta = \frac{r^2}{4D} \left(\frac{\tan(\Theta)}{\Theta} - 1 \right) - \sum_{n=0}^{\infty} \frac{16R^{2-\lambda_n} r^{\lambda_n}}{D\Theta^2 \lambda_n^3 (\lambda_n^2 - 4)}, \tag{1.14}$$

where $\lambda_n = (2n + 1) \frac{\pi}{\Theta}$. For Θ small, equation (1.14) can be approximated by

$$\bar{u}(r) = \frac{r^2}{4D} \left(\frac{\tan(\Theta)}{\Theta} - 1 \right) - \frac{16\Theta R^2 \left(\frac{r}{R}\right)^{\pi/\Theta}}{D\pi^3 \left((\pi/\Theta)^2 - 4\right)}. \tag{1.15}$$

3.2 Exit points distribution

To estimate the position a virus will attach preferentially to the microtubule, we determine the distribution of exit points, when the viral particle initially started at a radial distance from the nucleus. We recall that the probability density function (pdf) $p(\mathbf{r}, t|\mathbf{r}_0)$ to find a diffusing particle in a volume element $d\mathbf{r}$ at time t inside the wedge Ω , conditioned on the initial position $\mathbf{r} = \mathbf{r}_0$ is solution of the diffusion equation

$$\frac{\partial p(\mathbf{r}, t|\mathbf{r}_0)}{\partial t} = D\Delta p(\mathbf{r}, t|\mathbf{r}_0) \text{ for } \mathbf{r} \in \Omega^R$$

$$p(\mathbf{r}, t|\mathbf{r}_0) = 0 \text{ for } \mathbf{r} \in \partial\Omega_a^R$$

$$\frac{\partial p(\mathbf{r}, t|\mathbf{r}_0)}{\partial n} = 0 \text{ for } \mathbf{r} \in \partial\Omega_r^R,$$

where the initial condition is $p(\mathbf{r}, 0|\mathbf{r}_0) = \delta(\mathbf{r} - \mathbf{r}_0)$. The distribution of exit points $\epsilon(\mathbf{y})$ is given by

$$\epsilon(\mathbf{y}) = \int_0^\infty j(\mathbf{y}, t) dt, \tag{1.16}$$

where the flux j is defined by

$$j(\mathbf{y}, t) = -D \frac{\partial p(\mathbf{r}, t)}{\partial \mathbf{n}} \Big|_{\mathbf{r} = \mathbf{y}}.$$

If we denote $C(\mathbf{r}_0, \mathbf{r}) = \int_0^\infty p(\mathbf{r}, t|\mathbf{r}_0) dt$ then C is solution of

$$-D\Delta C(\mathbf{r}_0, \mathbf{r}) = \delta(\mathbf{r} - \mathbf{r}_0), \tag{1.17}$$

and

$$\epsilon(\mathbf{y}) = -D \frac{\partial C}{\partial n}(\mathbf{r}_0, \mathbf{y}) \text{ for } \mathbf{y} \in \Omega_a^R. \tag{1.18}$$

Consequently, to obtain the pdf of exit points ϵ , we use the Green function in the wedge domain Ω^R . By using a conformal transformation, we hereafter solve a simplified case of an open wedge (*i.e.* without a reflecting boundary at $r = R$). This computation could be compared with the general one that will be derived in the next section.

To compute the exit points distribution, we consider the solution of equation (1.17), obtained by the image method and a conformal transformation from the open wedge to the upper complex half-plane. The Green function, solution of equation(1.17) in the upper complex half-plane is given by

$$C(z) = \frac{1}{2\pi D} \ln \frac{z - z_0}{z - z_0^*}, \tag{1.19}$$

where z_0^* the complex conjugate of z_0 . Using the conformal transformation $\omega = f(z) = z^{\frac{\pi}{\Theta}}$ [55], that maps the interior of the wedge of opening angle Θ to the upper half plane, the Green function in the wedge is given by

$$C(z) = \frac{1}{2\pi D} \ln \left(\frac{z^{\frac{\pi}{\Theta}} - z_0^{\frac{\pi}{\Theta}}}{z^{\frac{\pi}{\Theta}} - (z_0^*)^{\frac{\pi}{\Theta}}} \right). \tag{1.20}$$

The flux to the line θ is given by

$$\begin{aligned} \epsilon_\theta(r) &= -\frac{D}{r} \frac{\partial C}{\partial \theta} (re^{i\theta}) = \frac{1}{2\pi r} \frac{i\nu (re^{i\theta})^\nu \cdot (k_0 - k_0^*)}{((re^{i\theta})^\nu - k_0) ((re^{i\theta})^\nu - k_0^*)} \\ &= \frac{1}{2\pi r} \frac{-2\nu (re^{i\theta})^\nu r_0^\nu \sin(\nu\theta_0)}{(re^{i\theta})^{2\nu} + r_0^{2\nu} - 2(re^{i\theta})^\nu r_0^\nu \cos(\nu\theta_0)}, \end{aligned}$$

where $\nu = \frac{\pi}{\Theta}$, $k_0 = z_0^\nu = (r_0 e^{i\theta_0})^\nu$. Finally, the exit point distribution for $\theta = \Theta$ is given by

$$\epsilon_\Theta(r) = \frac{r_0}{\Theta} \frac{(rr_0)^{(\nu-1)} \sin(\nu\theta_0)}{r^{2\nu} + r_0^{2\nu} + 2(rr_0)^\nu \cos(\nu\theta_0)}, \quad (1.21)$$

while for $\theta = 0$ it is given by

$$\epsilon_0(r) = \frac{r_0}{\Theta} \frac{(rr_0)^{(\nu-1)} \sin(\nu\theta_0)}{r^{2\nu} + r_0^{2\nu} - 2(rr_0)^\nu \cos(\nu\theta_0)}. \quad (1.22)$$

A matlab check guarantees that

$$\int_0^\infty \{\epsilon_\Theta(r) + \epsilon_0(r)\} dr = 1. \quad (1.23)$$

This simple computation is instructive and shall be compared to the full one given in section 3.3.

3.3 Exit pdf in a Pie Wedge

To compute the exit points distribution in a pie wedge with a reflecting boundary at $r = R$, we search for an explicit solution of the diffusion equation in polar coordinates inside the pie wedge. We first consider the general diffusion equation

$$\begin{aligned} \frac{\partial p}{\partial t}(\mathbf{x}, t|\mathbf{y}) &= D \left(\frac{\partial^2 p}{\partial r^2} + \frac{1}{r} \frac{\partial p}{\partial r} + \frac{1}{r^2} \frac{\partial^2 p}{\partial \theta^2} \right) (\mathbf{x}, t|\mathbf{y}) \\ p(\mathbf{x}, 0|\mathbf{y}) &= \delta(\mathbf{x} - \mathbf{y}) \end{aligned} \quad (1.24)$$

where the boundary conditions are given in (1.7). We may often use the change of variable $\forall n \in \mathbf{N}^*$:

$$k = \frac{n\pi}{\Theta}.$$

The initial condition is given by

$$p(\mathbf{x}, 0|\mathbf{y}) = p(r, \theta, 0|r_0, \theta_0) = \frac{2}{\Theta r_0} \delta(r - r_0) \sum_k \sin(k\theta) \sin(k\theta_0),$$

for $\theta < \theta_0$ (if $\theta > \theta_0$, θ_0 must be replaced by $\Theta - \theta_0$). To compute the solution of equation (1.24), we consider the Laplace transform \hat{p} of the probability p

$$\begin{aligned} s\hat{p}(r, \theta, s|r_0, \theta_0) &= \frac{2}{\Theta r_0} \delta(r - r_0) \sum_k \sin(k\theta) \sin(k\theta_0) \\ &= D \left(\frac{\partial^2 \hat{p}}{\partial r^2} + \frac{1}{r} \frac{\partial \hat{p}}{\partial r} + \frac{1}{r^2} \frac{\partial^2 \hat{p}}{\partial \theta^2} \right) (r, \theta, s|r_0, \theta_0). \end{aligned}$$

Using the separation of variables, we have

$$\hat{p}(r, \theta, s | r_0, \theta_0) = \sum_k R_k(r, s) \sin(k\theta) \sin(k\theta_0),$$

Using the change of variable, $x(s) = r\sqrt{\frac{s}{D}}$ and $x_0(s) = r_0\sqrt{\frac{s}{D}}$, we get for all k that

$$\begin{aligned} R_k''(x(s), s) + \frac{1}{x(s)} R_k'(x(s), s) - \left(1 + \frac{k^2}{x(s)^2}\right) R_k(x(s), s) \\ = -\frac{2}{\Theta D x_0(s)} \delta(x(s) - x_0(s)). \end{aligned} \quad (1.25)$$

$R_k(x(s), s)$ is a superposition of modified Bessel functions of order k : $I_k(x(s))$ and $K_k(x(s))$ for $x(s) \neq x_0(s)$:

$$R_k(x(s), s) = A_k I_k(x(s)) + B_k K_k(x(s)),$$

where A_k and B_k are real constants. Since K_k diverges as $x(s) \rightarrow 0$, the interior solution for $(x(s) < x_0(s))$ depends only on I_k . We denote by D_k the exterior solution for $(x(s) > x_0(s))$. We use the general notation $x \wedge y = \min(x, y)$ and $x \vee y = \max(x, y)$, thus

$$R_k(x(s), s) = A_k I_k(x(s) \wedge x_0(s)) D_k(x(s) \vee x_0(s)).$$

To determine $D_k = a_k I_k + b_k K_k$, we use the reflecting condition at $x(s) = x_+(s) = R\sqrt{\frac{s}{D}}$ and we get that

$$A_k I_k(x_0(s)) \cdot \left(a_k I_k'(x_+(s)) + b_k K_k'(x_+(s)) \right) = 0.$$

We choose

$$a_k = -K_k'(x_+(s)) \text{ and } b_k = I_k'(x_+(s)).$$

Thus

$$R_k(x(s), s) = A_k I_k(x(s) \wedge x_0(s)) \left(I_k'(x_+(s)) K_k - K_k'(x_+(s)) I_k \right) (x(s) \vee x_0(s)).$$

The constants A_k are determined by integrating equation (1.25) over an infinitesimal interval that includes r_0 . Using the continuity of R_k , we get

$$(R_k)'_{x(s) > x_0(s)} \Big|_{x(s)=x_0(s)} - (R_k)'_{x(s) < x_0(s)} \Big|_{x(s)=x_0(s)} = -\frac{2}{\Theta D x_0(s)},$$

that is

$$\begin{aligned} A_k \left(I_k \left(I_k'(x_+(s)) K_k' - K_k'(x_+(s)) I_k' \right) - I_k' \left(I_k'(x_+(s)) K_k - K_k'(x_+(s)) I_k \right) \right) (x_0(s)) \\ = -\frac{2}{\Theta D x_0(s)}, \end{aligned}$$

after some simplifications, we get

$$A_k I_k'(x_+(s)) \left(I_k K_k' - I_k' K_k \right) (x_0(s)) = -\frac{2}{\Theta D x_0(s)}.$$

Using the recurrent relation between modified Bessel functions (see [31] or page 489 [56]),

$$\begin{aligned} I'_k(x_0(s)) &= \left(I_{k-1} - \frac{k}{x_0(s)} I_k \right) (x_0(s)) \text{ and } K'_k(x_0(s)) \\ &= \left(-K_{k-1} - \frac{k}{x_0(s)} K_k \right) (x_0(s)), \end{aligned}$$

we get

$$\begin{aligned} A_k I'_k(x_+(s)) \left(I_k \left(-K_{k-1} - \frac{k}{x_0(s)} K_k \right) - \left(I_{k-1} - \frac{k}{x_0(s)} I_k \right) K_k \right) (x_0(s)) \\ = -\frac{2}{\Theta D x_0(s)}, \end{aligned}$$

that is

$$A_k I'_k(x_+(s)) (I_k K_{k-1} + I_{k-1} K_k) (x_0(s)) = \frac{2}{\Theta D x_0(s)}.$$

Finally, using this relation and the following Wronskian relation (page 489 [56]),

$$(I_k K_{k-1} + I_{k-1} K_k) (x_0(s)) = \frac{1}{x_0(s)},$$

we obtain that

$$A_k = \frac{2}{\Theta D I'_k(x_+(s))}.$$

thus

$$\begin{aligned} R_k(x(s), s) &= \frac{2}{\Theta D I'_k(x_+(s))} I_k(x(s) \wedge x_0(s)) \\ &\left(I'_k(x_+(s)) K_k - K'_k(x_+(s)) I_k \right) (x(s) \vee x_0(s)). \end{aligned}$$

We can now express the solution \hat{p} for $\theta < \theta_0$ by

$$\hat{p}(r, \theta, s) = \frac{2}{\Theta D} \sum_k \frac{I_k(x(s) \wedge x_0(s)) \left(I'_k(x_+(s)) K_k - K'_k(x_+(s)) I_k \right) (x(s) \vee x_0(s))}{I'_k(x_+(s))} \sin(k\theta) \sin(k\theta_0).$$

The exit point distribution $\epsilon^0(r)$ is given by

$$\epsilon^0(r) = - \left(\frac{D}{r} \frac{\partial}{\partial \theta} \left(\int_0^\infty p(r, \theta, t) dt \right) \right) (\theta = 0). \quad (1.26)$$

To obtain an analytical expression for expression (1.26), we use the Laplace relation :

$$\mathcal{L} \left(\int_0^t f(u) du \right) = \frac{F(z)}{z},$$

where $F = \mathcal{L}(f)$ is the Laplace transform of the function f . We have

$$\begin{aligned} \int_0^t p(r, \theta, u) du &= \mathcal{L}^{-1} \left(\frac{\hat{p}(r, \theta, s)}{s} \right) \\ &= \mathcal{L}^{-1} \left(\frac{2}{\Theta D} \sum_k \sin(k\theta) \sin(k\theta_0) \right. \\ &\quad \left. \frac{I_k(x(s) \wedge x_0(s)) (I'_k(x_+(s)) K_k - K'_k(x_+(s)) I_k)(x(s) \vee x_0(s))}{s I'_k(x_+(s))} \right). \end{aligned}$$

The computation of the integral

$$\begin{aligned} I(r, \theta, t) &= \frac{1}{\Theta \pi D i} \sum_k \sin(k\theta) \sin(k\theta_0) \\ &\quad \int_{-i\infty}^{+i\infty} \frac{I_k(x(s) \wedge x_0(s)) (I'_k(x_+(s)) K_k - K'_k(x_+(s)) I_k)(x(s) \vee x_0(s))}{s I'_k(x_+(s))} e^{st} ds \end{aligned} \quad (1.27)$$

uses the residue theorem and the details are given in the Appendix. We have

$$I(r, \theta, t) = \int_0^t p(r, \theta, u) du = \frac{2}{\Theta D} (S_1(r, \theta, t) + S_2(r, \theta, t)),$$

where

$$\begin{aligned} S_1(r, \theta, t) &= \sum_k \sin(k\theta) \sin(k\theta_0) \frac{r^k (r_0^{2k} + R^{2k})}{2k R^{2k} r_0^k}, \\ S_2(r, \theta, t) &= -2 \sum_k \sin(k\theta) \sin(k\theta_0) \sum_{j=1}^{\infty} e^{-D\alpha_{j,k}^2 t} \frac{J_k(r\alpha_{j,k}) J_k(r_0\alpha_{j,k})}{(R^2\alpha_{j,k}^2 - k^2) J_k^2(R\alpha_{j,k})}, \end{aligned}$$

and J_k are the k -order Bessel's function and $\alpha_{j,k}$ are the roots of the equation :

$$J'_k(R\alpha) = 0.$$

Consequently, for $r < r_0$, using (1.26), we get the following exit distribution (for $\Theta = 0$) :

$$\epsilon^0(r) = \frac{2}{\Theta} \frac{\partial}{r \partial \theta} \left(\lim_{t \rightarrow \infty} (S_1(r, \theta, t) + S_2(r, \theta, t)) \right)_{\theta=0}.$$

Because :

$$\lim_{t \rightarrow \infty} S_1(r, \theta, t) = S_1(r, \theta) \text{ and } \lim_{t \rightarrow \infty} S_2(r, \theta, t) = 0,$$

we finally obtain that

$$\epsilon^0(r) = \frac{1}{\Theta} \sum_k \sin(k\theta_0) \frac{r^{k-1} (r_0^{2k} + R^{2k})}{R^{2k} r_0^k}, \quad (1.28)$$

and, for $r > r_0$, a similar computation leads to :

$$\epsilon^0(r) = \frac{1}{\Theta} \sum_k \sin(k\theta_0) \frac{r_0^k (r^{2k} + R^{2k})}{R^{2k} r^{k+1}}. \quad (1.29)$$

These expressions can be further simplified. Indeed, we rewrite them as follows (for $r < r_0$) :

$$\epsilon^0(r) = \frac{1}{\Theta r} \sum_k \sin(k\theta_0) \left(\frac{r}{r_0}\right)^k \left(1 + \left(\frac{r_0}{R}\right)^{2k}\right),$$

thus,

$$\epsilon^0(r) = \frac{1}{\Theta r} \Im \left(\sum_{n \geq 1} e^{in\nu\theta_0} \left(\frac{r}{r_0}\right)^{n\nu} \left(1 + \left(\frac{r_0}{R}\right)^{2n\nu}\right) \right),$$

where \Im denotes the imaginary part of the expression. We obtain two geometrical series that can be summed. We get :

$$\epsilon^0(r) = \frac{1}{\Theta r} \Im \left(\frac{e^{i\nu\theta_0} \left(\frac{r}{r_0}\right)^\nu}{1 - e^{i\nu\theta_0} \left(\frac{r}{r_0}\right)^\nu} + \frac{e^{i\nu\theta_0} \left(\frac{r}{r_0}\right)^\nu \left(\frac{r_0}{R}\right)^{2\nu}}{1 - e^{i\nu\theta_0} \left(\frac{r}{r_0}\right)^\nu \left(\frac{r_0}{R}\right)^{2\nu}} \right),$$

that is :

$$\epsilon^0(r) = \frac{1}{\Theta r} \Im \left(e^{i\nu\theta_0} \left(\frac{\left(\frac{r}{r_0}\right)^\nu}{1 - e^{i\nu\theta_0} \left(\frac{r}{r_0}\right)^\nu} + \frac{\left(\frac{rr_0}{R^2}\right)^\nu}{1 - e^{i\nu\theta_0} \left(\frac{rr_0}{R^2}\right)^\nu} \right) \right).$$

After some rearrangements, we obtain the following exit point distribution on $\theta = 0$, conditioned on the initial position (r_0, θ_0) :

$$\begin{aligned} \epsilon^0(r) = \epsilon^0(r|r_0, \theta_0) &= \frac{1}{\Theta r} \left(\frac{(rr_0)^\nu \sin(\nu\theta_0)}{r^{2\nu} + r_0^{2\nu} - 2(rr_0)^\nu \cos(\nu\theta_0)} \right. \\ &\quad \left. + \frac{(rr_0 R^2)^\nu \sin(\nu\theta_0)}{(rr_0)^{2\nu} + R^{4\nu} - 2(rr_0 R^2)^\nu \cos(\nu\theta_0)} \right), \end{aligned} \quad (1.30)$$

for $0 \leq r \leq R$. Similarly, for $\theta = \Theta$, we obtain

$$\begin{aligned} \epsilon^\Theta(r) = \epsilon^\Theta(r|r_0, \theta_0) &= \frac{1}{\Theta r} \left(\frac{(rr_0)^\nu \sin(\nu\theta_0)}{r^{2\nu} + r_0^{2\nu} + 2(rr_0)^\nu \cos(\nu\theta_0)} \right. \\ &\quad \left. + \frac{(rr_0 R^2)^\nu \sin(\nu\theta_0)}{(rr_0)^{2\nu} + R^{4\nu} + 2(rr_0 R^2)^\nu \cos(\nu\theta_0)} \right). \end{aligned} \quad (1.31)$$

We notice that letting R tends to ∞ , we recover the expressions computed in the open wedge case ((1.21) and (1.22)).

3.4 The Mean Exit Radius (MER)

To determine the mean exit distribution radius $\bar{\epsilon}(r|r_0)$ for a viral particle starting initially at position r_0, θ_0 where θ_0 is uniformly distributed between 0 and Θ , we consider $\epsilon(r|r_0, \theta_0) = \epsilon^0(r|r_0, \theta_0) + \epsilon^\Theta(r|r_0, \theta_0)$ and estimate the integral

$$\bar{\epsilon}(r|r_0) = \frac{1}{\Theta} \int_{\theta_0=0}^{\Theta} \epsilon(r|r_0, \theta_0) d\theta_0. \quad (1.32)$$

Exit radius distribution

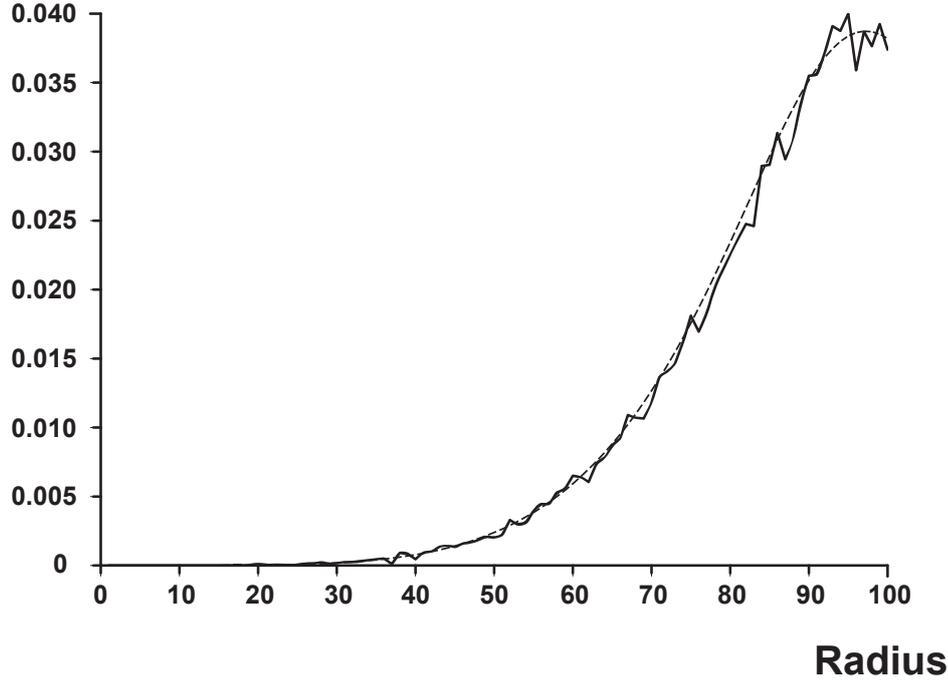


FIGURE 1.3 – **Mean exit points distribution.** The theoretical distribution (dashed line) is tested against the empirical one (solid line) obtained by running a simulation of 20 000 Brownian particles, starting on the wedge bisectrix ($\theta_0 = \frac{\Theta}{2}$ at $r_0 = R = 100$ for $\Theta = \frac{\pi}{6}$). Because the starting point is located on the bisectrix, $\epsilon^0(x) = \epsilon^\Theta(x)$, and thus the analytical curve is given by $\epsilon(r) = \epsilon^0(r) + \epsilon^\Theta(r) = \frac{2}{\Theta r} \left(\frac{(rr_0)^\nu}{r^{2\nu} + r_0^{2\nu}} + \frac{(rr_0 R^2)^\nu}{(rr_0)^{2\nu} + R^{4\nu}} \right)$. In that case, the maximum of the function $\epsilon(r)$ is achieved at $r = r_0 e^{\frac{1}{2\nu} \ln\left(\frac{\nu-1}{\nu+1}\right)}$.

Integrating expressions ((1.30) and (1.31)) we get :

$$\bar{\epsilon}(r|r_0) = \frac{2}{\Theta \pi r} \left(\ln \left(\frac{r^\nu + r_0^\nu}{|r^\nu - r_0^\nu|} \right) + \ln \left(\frac{R^{2\nu} + (rr_0)^\nu}{R^{2\nu} - (rr_0)^\nu} \right) \right).$$

We define the mean exit point as $r_m(r_0) = \mathbf{E}(r|r_0)$ conditioned on the initial radius r_0 . Thus,

$$r_m(r_0) = \mathbf{E}(r|r_0) = \int_0^R r \bar{\epsilon}(r|r_0) dr. \tag{1.33}$$

Using the expansion $\ln(1+x) = \sum_{n \geq 1} (-1)^{n+1} \frac{x^n}{n}$ for $x < 1$, we obtain by a direct integration that

$$r_m(r_0) = \frac{8}{\pi^2} \left(r_0 \sum_{n=0}^{\infty} \frac{1}{(2n+1)^2} \left(\frac{1}{1 - \frac{1}{(2n+1)^2 \left(\frac{\pi}{\Theta}\right)^2}} \right) - R \sum_{n=0}^{\infty} \frac{\left(\frac{r_0}{R}\right)^{(2n+1)\frac{\pi}{\Theta}} \frac{\pi}{\Theta}}{(2n+1) \left(\left((2n+1) \frac{\pi}{\Theta} \right)^2 - 1 \right)} \right), \quad (1.34)$$

using the expansion in the first part,

$$\frac{1}{1 - \frac{1}{(2n+1)^2 \left(\frac{\pi}{\Theta}\right)^2}} = \sum_{p=0}^{\infty} \left(\frac{\Theta}{(2n+1)\pi} \right)^{2p} \quad (1.35)$$

and the approximation $\Theta \ll 1$, we obtain using the value of the Riemann ζ -function, $\zeta(2) = \frac{\pi^2}{6}$ and $\zeta(4) = \frac{\pi^4}{90}$, $r_0 \leq R$, that

$$r_m(r_0) \approx r_0 \left(1 + \frac{\Theta^2}{12} \right) - \frac{8R}{\pi^2} \left(\frac{r_0}{R} \right)^{\pi/\Theta} \frac{\pi/\Theta}{(\pi/\Theta)^2 - 1}. \quad (1.36)$$

For Θ small, the second term in the right-hand side of (1.36) is exponentially small.

4 Approximation of a virus motion by an effective Markovian stochastic equation

We replace the successive steps of viral dynamics with an effective stochastic equation containing a constant steady state drift.

4.1 Methodology

Virus motion described in paragraph (2.2) consists of a succession of drift and diffusing periods. We start with the stochastic equation

$$\dot{\mathbf{X}} = -B \frac{\mathbf{r}}{|\mathbf{r}|} + \sqrt{2D} \dot{\mathbf{w}}, \quad (1.37)$$

where \mathbf{r} is the radial component of \mathbf{X} , B is the amplitude of the drift. The MFPT of the process (1.37) to the nucleus located $r = \delta$, when the initial position is located on the cell surface $r = R$ is solution of

$$\begin{aligned} D \left(\frac{d^2 t}{dr^2} + \frac{1}{r} \frac{dt}{dr} \right) (r, \theta) - B \frac{dt}{dr} (r, \theta) &= -1 \text{ for } (r, \theta) \in \Omega \\ t(r, \theta) &= 0 \text{ for } r = \delta \\ \frac{dt}{dr} (r, \theta) &= 0 \text{ for } r = R. \end{aligned}$$

A similar equation can be written in the domain $\tilde{\Omega}$ with reflective boundary conditions of the wedge. Both processes in the full domain or in $\tilde{\Omega}$ lead to the same MFPT. The solution $t(B, r)$ is given by

$$t(B, r) = C - \int_r^R \left(\int_v^R \frac{ue^{-\alpha(u-v)}}{Dv} du \right) dv, \quad (1.38)$$

where $\alpha = \frac{B}{D}$ and

$$t(B, R) = C = \int_\delta^R \left(\int_v^R \frac{ue^{-\alpha(u-v)}}{Dv} du \right) dv. \quad (1.39)$$

For a fixed radius R , the derivative of the function $t(B, R)$ with respect to B is strictly negative, which shows that $B \rightarrow t(B, R)$ is strictly decreasing. To determine the value of the amplitude B , we equal the mean time $t(B, R)$ with the MFPT to reach the nucleus within the iterative procedure as described in paragraph (2.2) : at time zero, the virus starts at a position $r = R = R_0$ and reaches the edge boundary in a mean time $\bar{u}(R_0)$ and at a mean position $r_m(R_0)$. The viral particle is then transported toward the nucleus over a distance d_m during a time t_m . Either the particle reaches the nucleus before time t_m and then the algorithm is terminated or in a second step, it starts at a position $R_1 = r_m(R_0) - d_m$. The process iterates until the particle reaches the nucleus. We consider the mean number of fundamental steps (diffusion step and directed motion along a MT step) the virus needs to reach the nucleus is equal to $n \geq 0$. The mean time to reach the nucleus computed by equation (1.38) has thus to be equal to the mean time $\tau = \sum_{k=0}^{n-1} \bar{u}(R_k) + nt_m + \langle t_r \rangle$ of the iterative trajectory. In a first approximation, we neglect the mean residual time $\langle t_r \rangle$ and we thus get the equality :

$$t(B, R) = \tau = \sum_{k=0}^{n-1} \bar{u}(R_k) + nt_m \quad (1.40)$$

$$R_{k+1} = r_m(R_k) - d_m \quad (1.41)$$

$$R_0 = R. \quad (1.42)$$

For a fix radius R , equation (1.40) has a unique solution B , which can be found in practice by any standard numerical method.

Remark

The MFPT of a particle where the trajectory consists of alternating drift (traveling along microtubules) and diffusion periods can either be higher or lower than the MFPT of a pure Brownian particle. Indeed when $B < 0$, the drift effect is less efficient than pure diffusion. For example, for $\Theta = \frac{\pi}{6}$, $R = 100\mu m$, $\delta = \frac{R}{4} = 25\mu m$, a large diffusion constant $D = 10\mu m^2 s^{-1}$ with the dynamical parameters $t_m = 1s$ and $d_m = 1\mu m$, leads to a negative mean drift

$$B \approx -0.14\mu m s^{-1}. \quad (1.43)$$

On the other hand, for a small diffusion constant $D = 1\mu m^2 s^{-1}$, an efficient microtubules transport obtained for $t_m = 1s$ and $d_m = 5\mu m$ leads to a mean positive drift

$$B \approx 0.13\mu m s^{-1}. \quad (1.44)$$

4.2 Explicit expression of the drift in the limit of $\Theta \ll 1$

When the number of microtubules is large enough, the condition $\Theta \ll 1$ is satisfied. Moreover, because a virus entering a cell surface has a determinist motion, we can assume that the initial position satisfies $R_0 < R$ so that we can neglect any boundary effects and use the open wedge approximation which consists of using formula (1.36) without the boundary layer term. Actually, this approximation is not that restrictive because after the first iteration process (movement along the microtubule followed by the particle release), the boundary layer term is negligible compared to the other term.

To obtain an explicit expression for the amplitude B, we consider the successive approximations

$$r_m(R_0) \approx R_0 \left(1 + \frac{\Theta^2}{12}\right), \quad (1.45)$$

and

$$\begin{aligned} R_0 &= R_0; \\ R_1 &\simeq R_0 \left(1 + \frac{\Theta^2}{12}\right) - d_m; \\ R_2 &\simeq R_0 \left(1 + \frac{\Theta^2}{12}\right)^2 - d_m \left(1 + \left(1 + \frac{\Theta^2}{12}\right)\right); \\ &\vdots \\ R_i &\simeq R_0 \left(1 + \frac{\Theta^2}{12}\right)^i - d_m \left(\sum_{k=0}^{i-1} \left(1 + \frac{\Theta^2}{12}\right)^k\right); \end{aligned}$$

that is

$$R_i \simeq \left(R_0 - \frac{12d_m}{\Theta^2}\right) \left(1 + \frac{\Theta^2}{12}\right)^i + \frac{12d_m}{\Theta^2}. \quad (1.46)$$

Thus the particle reaches the nucleus after n iteration steps which approximatively satisfies $R_n = \delta$,

$$n \simeq \frac{\ln\left(\frac{1 - \frac{\delta\Theta^2}{12d_m}}{1 - \frac{R_0\Theta^2}{12d_m}}\right)}{\ln\left(1 + \frac{\Theta^2}{12}\right)} \approx \frac{R_0 - \delta}{d_m} + o(1). \quad (1.47)$$

If T_n denotes the mean time a viral particle takes to reach the nucleus, then using formula (1.15), we obtain

$$T_n \simeq n \cdot t_m + \frac{\left(\frac{\tan(\Theta)}{\Theta} - 1\right)}{4D} \sum_{i=0}^{n-1} R_i^2, \quad (1.48)$$

that is

$$\begin{aligned}
 t &\simeq n.t_m + \frac{\left(\frac{\tan(\Theta)}{\Theta} - 1\right)}{4D} \\
 &\quad \sum_{i=0}^{n-1} \left(\left(\frac{12d_m}{\Theta^2}\right)^2 + 2\left(\frac{12d_m}{\Theta^2}\right) \left(R_0 - \frac{12d_m}{\Theta^2}\right) \left(1 + \frac{\Theta^2}{12}\right)^i \right. \\
 &\quad \left. + \left(R_0 - \frac{12d_m}{\Theta^2}\right)^2 \left(1 + \frac{\Theta^2}{12}\right)^{2i} \right), \\
 T_n &\simeq nt_m + \frac{\left(\frac{\tan(\Theta)}{\Theta} - 1\right)}{4D} \\
 &\quad \left(n \left(\frac{12d_m}{\Theta^2}\right)^2 - \left(\frac{24d_m}{\Theta^2}\right) \left(R_0 - \frac{12d_m}{\Theta^2}\right) \frac{1 - \left(1 + \frac{\Theta^2}{12}\right)^n}{\frac{\Theta^2}{12}} \right. \\
 &\quad \left. + \left(R_0 - \frac{12d_m}{\Theta^2}\right)^2 \frac{1 - \left(1 + \frac{\Theta^2}{12}\right)^{2n}}{1 - \left(1 + \frac{\Theta^2}{12}\right)^2} \right).
 \end{aligned}$$

For $\Theta \ll 1$, a Taylor expansion gives that

$$\begin{aligned}
 T_n &\simeq \left(\frac{R_0 - \delta}{d_m}\right) t_m + \frac{t_m (R_0 - \delta)}{24d_m} \left(1 + \frac{R_0 + \delta}{d_m}\right) \Theta^2 \\
 &\quad + \frac{(R_0 - \delta)}{72D} \left(d_m + 3(R_0 + \delta) + \frac{2(R_0^2 + R_0\delta + \delta^2)}{d_m} \right) \Theta^4 + o(\Theta^4).
 \end{aligned}$$

In small diffusion limit $D \ll 1$, $\Theta \ll 1$, the velocity is $B \simeq \frac{R_0 - \delta}{T_n}$ and consequently we obtain for $R_0 \approx R$, a second order approximation

$$B \approx \frac{\frac{d_m}{t_m}}{1 + \left(1 + \frac{R + \delta}{d_m}\right) \frac{\Theta^2}{24} + O(\Theta^4)}, \quad (1.49)$$

where d_m, t_m are the mean distance and the mean time a virus stays on the microtubule, R (resp. δ) is the radius of the cell (resp. nucleus) and $\Theta = \frac{2\pi}{N}$, where N is the total number of microtubules.

4.3 Justification of the MFPT-criteria.

To justify the use of the MFPT-criteria to estimate the steady state drift, we run numerical simulations of 1,000 viruses inside a two dimensional domain Ω ($\delta < r < R$) with intermittent dynamics, alternating between epochs of free diffusion and directed motion along microtubules and compare the steady state distribution with the one obtained by solving the Fokker-Planck equation for viruses whose trajectories are described by the effective stochastic equation (1.2) with our computed constant drift

$$\mathbf{b}(\mathbf{X}) = -\frac{\frac{d_m}{t_m}}{1 + \left(1 + \frac{R + \delta}{d_m}\right) \frac{\Theta^2}{24}} \frac{\mathbf{r}}{|\mathbf{r}|} = -B \frac{\mathbf{r}}{|\mathbf{r}|}. \quad (1.50)$$

We imposed reflecting boundary conditions at the nuclear and the external membrane. The theoretical normalized steady state distribution ρ satisfies

$$\begin{aligned} D\Delta\rho - \nabla \cdot [\mathbf{b}\rho] &= 0 \text{ in } \Omega \\ \frac{d\rho}{dr}(R) = \frac{d\rho}{dr}(\delta) &= 0. \end{aligned}$$

and the solution ρ is given by

$$\rho(r) = \frac{e^{-\frac{Br}{D}}}{\int_{\delta}^R e^{-\frac{Br}{D}} 2\pi r dr} = \frac{e^{-\frac{Br}{D}}}{2\pi \frac{D}{B} \left(\delta e^{-\frac{B\delta}{D}} - R e^{-\frac{BR}{D}} + \frac{D}{B} \left(e^{-\frac{B\delta}{D}} - e^{-\frac{BR}{D}} \right) \right)}. \quad (1.51)$$

The result of both distributions is presented in figure 1.4 where we can observe that both curves match very nicely. This result shows that the criteria we have used is at least enough to recover the distribution. For the simulations, we consider the directed run of the virus along a MT (loaded by dynein) lasts $t_m = 1s$ and covers a mean distance $d_m = 0.7\mu m$ [57]. The diffusion constant is $D = 1.3\mu m^2 s^{-1}$ as observed for the Adeno Associated Virus [14]. The two curves in figure 1.4 fit very

Steady State Distributions

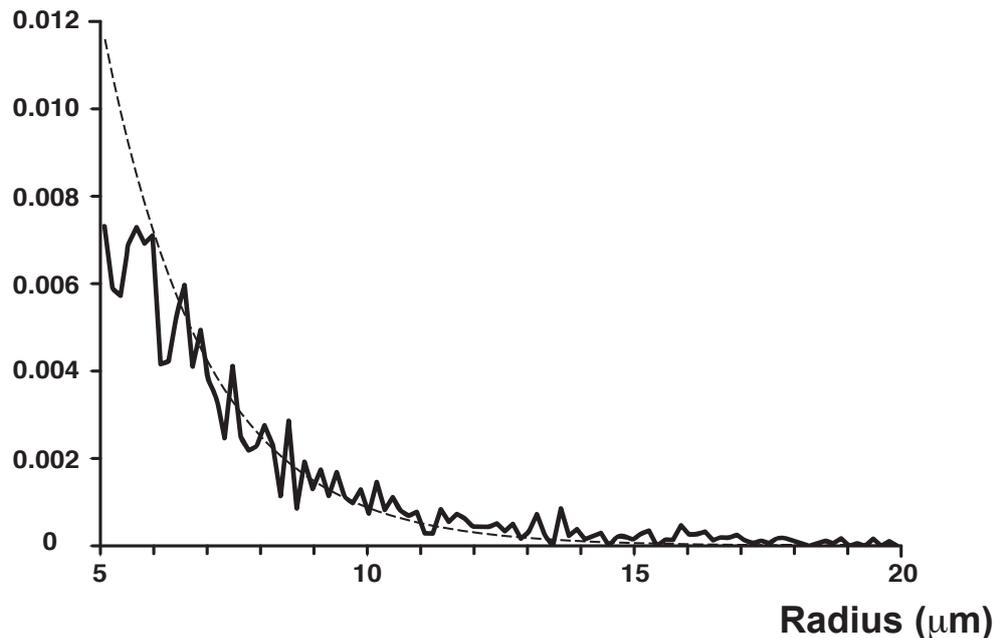


FIGURE 1.4 – **Steady State distributions.** We show the empirical steady state distribution for 1,000 viral trajectories with an intermittent dynamic (solid line). The theoretical distribution of viruses whose trajectories are described by the stochastic equation (1.2) is given in dashed line. Geometrical parameters are : $R = 20\mu m$, $\delta = 5\mu m$ and $\Theta = \frac{\pi}{24}$.

nically except at the neighborhood of the nuclear membrane, where the simulation of the empirical distribution is plagued with a possible boundary layer. Another source of discrepancy comes from the difference of behavior of viruses far and close to the

nucleus : viruses far from the nucleus do not bind as often as those located in its neighborhood. Consequently, a constant effective drift cannot account for the radial geometry near the nucleus. A theory for radius dependent effective is derived in the chapter 2. Interestingly, we show in that chapter that the $\Theta \ll 1$ approximation is still very good up to $\Theta = \frac{\pi}{6}$.

5 Conclusion

In the limit of a cell containing an excess of microtubules, we have presented here a model to describe the motion of biological particles such as viruses, vesicles and many others moving inside the cell cytoplasm by a complex combination of Brownian motion and determinist drift. Our procedure consists mainly in approximating an alternative switching mode between diffusion and determinist drift epochs by a steady state stochastic equation. This procedure consists of estimating the amplitude of the effective drift and is based on the criteria that the MFPTs to the nucleus, computed in both cases are equal. In that case, this amplitude account for the directed transport along microtubules, the cell geometry and the binding constants. The model has however several limitations. First, we do not take into account directly the backward movement of the virus along the microtubules [58, 59], which can affect the mean time and the amplitude of the drift. Second, the present computations are given for two dimensional cell geometry only. It can still be applied to many *in vitro* culture cells, however it is not clear how to generalize our approach to a three dimensional cell geometry. For example, to study the trafficking inside cylindrical axons or dendrites of neuronal cells, a different approach should include this geometrical features. However despite these real difficulties, the present model may be used to analyze plasmid transport in an host cell, at the molecular level, which is one of the fundamental limitation of gene delivery [60, 61, 62, 63].

Appendix

In this appendix, we provide an explicit computation of integral (1.28) using the method of the residues. This method was previously used in a similar context in ([56] p 386). We denote by $(p_j^k)_{j \geq 0}$ the poles of the function

$$\Phi : s \rightarrow \frac{I_k(x(s) \wedge x_0(s)) (I'_k(x_+(s)) K_k - K'_k(x_+(s)) I_k)(x(s) \vee x_0(s))}{s I'_k(x_+(s))} e^{st}.$$

where $(x(s) = r\sqrt{\frac{s}{D}}$, $x_0(s) = r_0\sqrt{\frac{s}{D}}$ and $x_+(s) = R\sqrt{\frac{s}{D}}$). The associated residues are $(r_j^k)_{j \geq 0}$. We now compute the residues explicitly.

To identify the poles, we recall the relation between the k -order Bessel's function J_k (that is true for z such that $-\pi < \arg(z) < \frac{\pi}{2}$) and the modified Bessel functions I_k (p 375 [31]) :

$$I_k(z) = e^{-\frac{1}{2}k\pi i} J_k\left(ze^{\frac{1}{2}\pi i}\right). \quad (1.52)$$

All roots $\alpha_{j,k}$ of the equations

$$J'_k(R\alpha) = 0,$$

are real, simple and strictly positive (p 370 [31]) because k is real and

$$k \leq \alpha_{1,k} < \alpha_{2,k} \dots$$

Thus,

$$I'_k(-iR\alpha_{j,k}) = 0.$$

Finally the poles of Φ are simple given by $p_0^k = 0$ and $\forall j \geq 1, p_j^k = -D\alpha_{j,k}^2$. Consequently the associated residues are given for each k for all $j \geq 0$ by

$$r_j^k = \lim_{s \rightarrow p_j^k} (s - p_j^k) \Phi(s). \quad (1.53)$$

Then using the residues, integral (1.28) is given by

$$I(r, \theta, t) = \frac{1}{\Theta \pi D i} \sum_k \sin(k\theta) \sin(k\theta_0) (2\pi i) \sum_{j \geq 0} r_j^k = \frac{2}{\Theta D} \sum_k \sin(k\theta) \sin(k\theta_0) \sum_{j \geq 0} r_j^k.$$

We now compute the residues r_j^k . The residue r_0^k is associated with the pole $p_0^k = 0$ and given by

$$r_0^k = \lim_{s \rightarrow 0} s \Phi(s)$$

Using the following identities on the modified Bessel functions (p 489 [56])

$$I'_k(z) = I_{k+1}(z) + \frac{k}{z} I_k(z) \text{ and } K'_k(z) = -K_{k-1}(z) - \frac{k}{z} K_k(z),$$

substituting the derivatives I'_k and K'_k in the expression of Φ , we get

$$\begin{aligned} r_0^k &= \lim_{s \rightarrow 0} \frac{I_k(x(s) \wedge x_0(s))}{\left(I_{k+1} + \frac{k}{x_+(s)} I_k \right) (x_+(s))} \\ &\quad \left(\left(\left(I_{k+1} + \frac{k}{x_+(s)} I_k \right) (x_+(s)) K_k \right) \right. \\ &\quad \left. + \left(\left(K_{k-1} + \frac{k}{x_+(s)} K_k \right) (x_+(s)) I_k \right) \right) (x(s) \vee x_0(s)), \end{aligned}$$

Taking into account the dominant terms only, we get

$$r_0^k = \lim_{s \rightarrow 0} \frac{I_k(x(s) \wedge x_0(s)) (I_k(x_+(s)) K_k + K_k(x_+(s)) I_k) (x(s) \vee x_0(s))}{I_k(x_+(s))}.$$

To further compute this limit, we use the Taylor expansions of I_k and K_k (p 375 [31]) expressed in terms of the Γ function :

$$I_k(z) \approx \frac{\left(\frac{1}{2}z\right)^k}{\Gamma(k+1)} \text{ and } K_k(z) \approx \frac{1}{2}\Gamma(k) \left(\frac{1}{2}z\right)^{-k}.$$

For $r < r_0$, we get

$$r_0^k = \lim_{s \rightarrow 0} \frac{\frac{\left(\frac{1}{2}(x(s))\right)^k}{\Gamma(k+1)} \left(\frac{\left(\frac{1}{2}(x_+(s))\right)^k}{\Gamma(k+1)} \frac{1}{2} \Gamma(k) \left(\frac{1}{2}(x_0(s))\right)^{-k} + \frac{1}{2} \Gamma(k) \left(\frac{1}{2}(x_+(s))\right)^{-k} \frac{\left(\frac{1}{2}(x_0(s))\right)^k}{\Gamma(k+1)} \right)}{\frac{\left(\frac{1}{2}(x_+(s))\right)^k}{\Gamma(k+1)}}.$$

Finally, using the relation $\Gamma(k+1) = k\Gamma(k)$, and the expressions of $x(s)$, $x_0(s)$ and $x_+(s)$ we get

$$r_0^k = \frac{r^k (r_0^{2k} + R^{2k})}{2kR^{2k}r_0^k}.$$

The computation of the other residues $(r_j^k)_{j \geq 1}$, is slightly different

$$r_j^k = \lim_{s \rightarrow p_j^k} (s - p_j^k) \Phi(s),$$

where $p_j^k = -D\alpha_{j,k}^2$. Using the Wronskian relation (p 489 [56]) :

$$I_k(z) K_k'(z) - K_k(z) I_k'(z) = -\frac{1}{z},$$

we now substitute

$$K_k'(z) = \frac{-\frac{1}{z} + K_k(z) I_k'(z)}{I_k(z)}.$$

in the expression of Φ , we get

$$r_j^k = \lim_{s \rightarrow p_j^k} \frac{(s - p_j^k) e^{st} I_k(x(s)) \left(I_k'(x_+(s)) K_k - \left(\frac{-\frac{1}{x_+(s)} + K_k I_k'}{I_k} \right) (x_+(s)) I_k \right) (x_0(s))}{s I_k'(x_+(s))}.$$

Because

$$\lim_{s \rightarrow p_j^k} I_k'(x_+(s)) = I_k'(x_+(p_j^k)) = 0,$$

we obtain the expression for the residues :

$$r_j^k = \frac{e^{p_j^k t} I_k(x(p_j^k)) I_k(x_0(p_j^k))}{p_j^k I_k(x_+(p_j^k)) x_+(p_j^k)} \lim_{s \rightarrow p_j^k} \frac{(s - p_j^k)}{I_k'(x_+(s))}.$$

Finally, since

$$\lim_{s \rightarrow p_j^k} \frac{(s - p_j^k)}{I_k'(x_+(s))} = \frac{2\sqrt{Dp_j^k}}{R} \lim_{s \rightarrow p_j^k} \frac{x_+(s) - x_+(p_j^k)}{I_k'(x_+(s)) - I_k'(x_+(p_j^k))} = \frac{2\sqrt{Dp_j^k}}{RI_k''(x_+(p_j^k))},$$

we obtain

$$r_j^k = \frac{e^{p_j^k t} I_k(x(p_j^k)) I_k(x_0(p_j^k))}{p_j^k I_k(x_+(p_j^k)) x_+(p_j^k)} \frac{2\sqrt{Dp_j^k}}{RI_k''(x_+(p_j^k))}.$$

To simplify this expression, we use that I_k satisfies the differential equation (p 374 [31]) :

$$I_k''(z) + \frac{1}{z} I_k'(z) - \left(1 + \frac{k^2}{z^2}\right) I_k(z) = 0,$$

thus for $z = x_+(p_j^k)$:

$$I_k''(x_+(p_j^k)) = \frac{p_j^k R^2 + Dk^2}{p_j^k R^2} I_k(x_+(p_j^k)),$$

we get

$$r_j^k = \frac{2De^{p_j^k t}}{R^2 p_j^k + Dk^2} \frac{I_k(x(p_j^k)) I_k(x_0(p_j^k))}{I_k^2(x_+(p_j^k))},$$

and finally, using (1.52), we get

$$r_j^k = \frac{2e^{-D\alpha_{j,k}^2 t}}{-R^2\alpha_{j,k}^2 + k^2} \frac{J_k(r\alpha_{j,k}) J_k(r_0\alpha_{j,k})}{J_k^2(R\alpha_{j,k})}.$$

Integral (1.28) is given by

$$I(r, \theta, t) = \frac{2}{\Theta D} \sum_k \sin(k\theta) \sin(k\theta_0) \sum_{j \geq 0} r_j^k = \frac{2}{\Theta D} (S_1(r, \theta, t) + S_2(r, \theta, t)). \quad (1.54)$$

where

$$S_1(r, \theta, t) = \sum_k \sin(k\theta) \sin(k\theta_0) \frac{r^k (r_0^{2k} + R^{2k})}{2kR^{2k}r_0^k},$$

$$S_2(r, \theta, t) = -2 \sum_k \sin(k\theta) \sin(k\theta_0) \sum_{j=1}^{\infty} e^{-D\alpha_{j,k}^2 t} \frac{J_k(r\alpha_{j,k}) J_k(r_0\alpha_{j,k})}{(R^2\alpha_{j,k}^2 - k^2) J_k^2(R\alpha_{j,k})},$$

Chapitre 2

Quantifying intermittent transport in cell cytoplasm, *Phys. Rev. E* 77 (2008)

1 Introduction

Cell transport, which may involve vesicles or proteins is essential for cellular function and homeostasis. In general free diffusion in the cell cytoplasm is not efficient and many particles such as large viruses cannot pass the crowded cytoplasm [5] without hijacking the complex cellular transport machinery and use molecular motors, such as dyneins, to travel along microtubules (MTs) toward the nucleus. Both vesicular and viral motions alternate intermittently between periods of free diffusion and directed motion along MTs [64]. Such viral trajectories have been recently monitored by using new imaging techniques *in vivo* [15, 14].

The switch nature of the motion, imposes a complex behavior of the particle trajectories which depends on the number and distribution of MTs, the rate of binding and unbinding and the diffusion constant of the free particle. Some physical properties, such as the mean velocity of trajectories has been obtained for the motion in domain made of parallel strips, in which a random particle has a deterministic motion on the stripes and pure diffusion outside [24]. In case of a population of motors, at equilibrium between free diffusion and bound on MTs, the motor distribution has been studied in cylindrical and radial geometries in [65, 25]; the authors estimate the forward binding rate using Brownian simulations in [25] and experimentally in [65].

We consider here a particle $\mathbf{x}(t)$, which can be described using the stochastic rule :

$$d\mathbf{x} = \begin{cases} \sqrt{2D}d\mathbf{w} & \text{for } \mathbf{x}(t) \text{ free} \\ \mathbf{V} & \text{for } \mathbf{x}(t) \text{ bound} \end{cases} \quad (2.1)$$

where \mathbf{w} is a standard Brownian motion, D the diffusion constant and \mathbf{V} the velocity of the directed motion along MTs.

In this communication, we compute the mean first passage time of a single particle to a population of MTs. We thus provide an analytical expression of the forward

binding rate of a motor to MTs in both radial and cylindrical geometries. Using the analytical expression of the forward binding rate, we propose a coarse-grained description of a switch dynamical motion of a particle, which can either be a virus, a vesicle or a molecular motor. This description, which is the main result of our paper, is a fundamental step to estimate the probability and the mean time to arrive at a small target. Moreover, using this description, we obtain the steady state distribution of virus in MTs network without resorting to the assumption of a two-state model [65, 66].

We thus compute an effective steady state drift $\mathbf{b}(\mathbf{x})$ such that the particle motion (2.1) can be coarse-grained by the stochastic equation :

$$d\mathbf{x} = \mathbf{b}(\mathbf{x})dt + \sqrt{2D}d\mathbf{w}. \quad (2.2)$$

Using results derived in [26], equation (2.2) and the degradation activity in the cytoplasm due to protease or lysosome, we obtain asymptotic estimates of the probability and the mean time for a virus to reach a nuclear pore. The problem of finding a small target is ubiquitous in cellular biology and recent theoretical studies [30, 67] suggest that the geometrical organization of the medium play a fundamental role in this search process.

2 Mathematical Modeling

We represent the cell cytoplasm as a bounded domain Ω , whose boundary $\partial\Omega$ consists of the external membrane $\partial\Omega_{ext}$ and the nuclear envelope, both of which form a reflecting boundary ∂N_r for the trajectories of (2.2), except for small nuclear pores ∂N_a , where they are absorbed. The ratio of boundary surface areas satisfies $\varepsilon = \frac{|\partial N_a|}{|\partial\Omega|} \ll 1$. We model the virus degradation activity in the cell cytoplasm as a steady state killing rate $k(\mathbf{x})$ for the trajectories of (2.2), so the survival probability density function (SPDF) is the solution of the Fokker-Planck equation [68]

$$\begin{aligned} \frac{\partial p}{\partial t} &= D\Delta p - \nabla \cdot \mathbf{b}p - kp \\ p(\mathbf{x}, 0) &= p_i(\mathbf{x}) \end{aligned}$$

with the boundary conditions :

$$p(\mathbf{x}, t) = 0 \text{ on } \partial N_a \text{ and } \mathbf{J}(\mathbf{x}, t) \cdot \mathbf{n}_x = 0 \text{ on } \partial N_r \cup \partial\Omega_{ext} \quad (2.3)$$

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

$$J(\mathbf{x}, t) = -D\nabla p(\mathbf{x}, t) + \mathbf{b}(\mathbf{x})p(\mathbf{x}, t). \quad (2.4)$$

The probability P_N and the mean time τ_N that a trajectory of (2.2) reaches ∂N_a are given by the small hole theory for two-dimensional domains Ω and drifts $b(x) =$

$-\nabla\Phi(x)$, as [26]

$$\left\{ \begin{array}{l} P_N = \frac{\frac{1}{|\partial\Omega|} \int_{\partial\Omega} e^{-\frac{\Phi(\mathbf{x})}{D}} dS_{\mathbf{x}}}{\frac{\ln\left(\frac{1}{\epsilon}\right)}{D\pi} \int_{\Omega} e^{-\frac{\Phi(\mathbf{x})}{D}} k(\mathbf{x}) d\mathbf{x} + \frac{1}{|\partial\Omega|} \int_{\partial\Omega} e^{-\frac{\Phi(\mathbf{x})}{D}} dS_{\mathbf{x}}}, \\ \tau_N = \frac{\frac{\ln\left(\frac{1}{\epsilon}\right)}{D\pi} \int_{\Omega} e^{-\frac{\Phi(\mathbf{x})}{D}} d\mathbf{x}}{\frac{\ln\left(\frac{1}{\epsilon}\right)}{D\pi} \int_{\Omega} e^{-\frac{\Phi(\mathbf{x})}{D}} k(\mathbf{x}) d\mathbf{x} + \frac{1}{|\partial\Omega|} \int_{\partial\Omega} e^{-\frac{\Phi(\mathbf{x})}{D}} dS_{\mathbf{x}}}, \end{array} \right. \quad (2.5)$$

Hereafter we derive explicitly the steady state drift $\mathbf{b}(\mathbf{x})$ as a function of some geometrical and dynamical parameters of the cell (number of MTs) and the virus (binding and unbinding rates, the mean velocity \mathbf{V} of the directed motion and the diffusion constant D).

3 General Methodology

To derive an expression for $\mathbf{b}(\mathbf{x})$, we consider the motion of a virus between the moment it enters the cell at the outer membrane and the moment it reaches the absorbing boundary ∂N_a . Its motion alternates between free diffusion, for a random time τ , until it hits a MT and binds. It continues in a directed motion along the MT for a mean time t_m and a mean distance $d_m = \|\mathbf{V}\| t_m$, until it is released and resumes free diffusion. The steady state drift is chosen to be constant for a sufficiently small step, such that the mean time $\tau + t_m$ to the first release at a point \mathbf{x}_f is the same as that predicted by (2.2) (see FIG. 2.1). This approach leads to explicit expressions for the steady state drift for two-dimensional radial and cylindrical geometries.

4 The steady state drift for a two-dimensional radial cell

We consider a two-dimensional cell cytoplasm which is an annulus Ω of outer radius R and inner radius δ (nuclear surface) with N MTs radially uniformly distributed. They irradiate from the nucleus to the external membrane and the angle between two neighboring ones is $\Theta = 2\pi/N$. The two-dimensional approximation applies for culture cells which are flat [48] due to the adhesion to the substrate. In that case the thickness can be neglected in the computation. Before reaching a small nuclear pore, a virus has an intermittent dynamics, alternating between diffusing and bound periods (see FIG. 2.2). Because the MTs are uniformly distributed, we consider the fundamental domain $\tilde{\Omega}$ defined as the two-dimensional slice of angle Θ between two neighboring ones. In $\tilde{\Omega}$, the fundamental step described above is as follows : the virus starts at a radius r_0 with an angle uniformly distributed in $[0; \Theta]$, it binds to a MT at a time $\tau(r_0)$ and at a radius $\bar{r}(r_0)$. On the MT, it has a radially directed movement towards the nucleus during a mean time t_m and over a distance $d_m = \|\mathbf{V}\| t_m$. Finally, the virus is released with a Θ -uniformly distributed angle at a

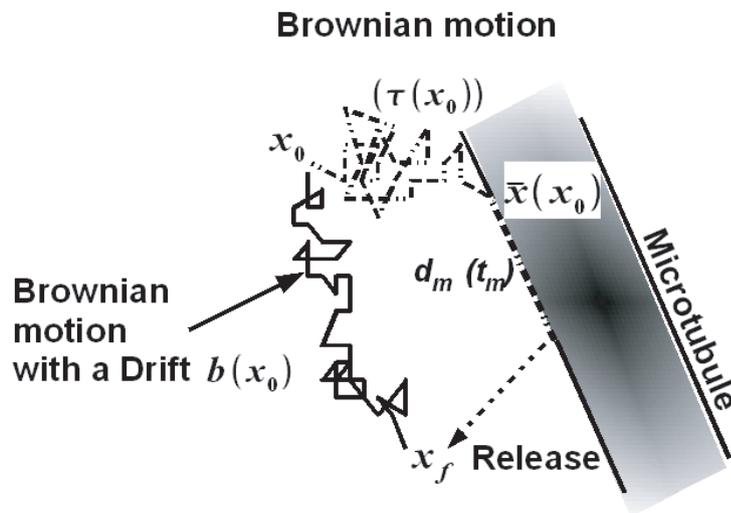


FIGURE 2.1 – The fundamental step is represented with a dotted line; a virus starts at a position \mathbf{x}_0 , diffuses freely, binds to a MT over a distance d_m and is then released at a final position \mathbf{x}_f . The solid line represents a trajectory generated by the steady state equation (2.2). In the parenthesis, we point out the mean times for each portion of trajectories.

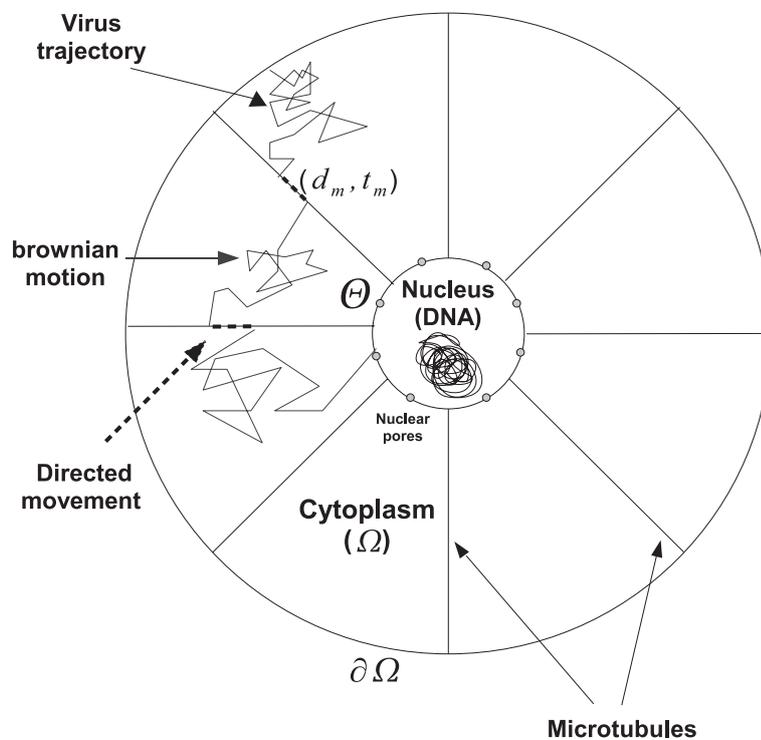


FIGURE 2.2 – Two dimensional radial cell with radially equidistributed MTs. We show a virus trajectory alternating between bound and diffusive periods in cytoplasm.

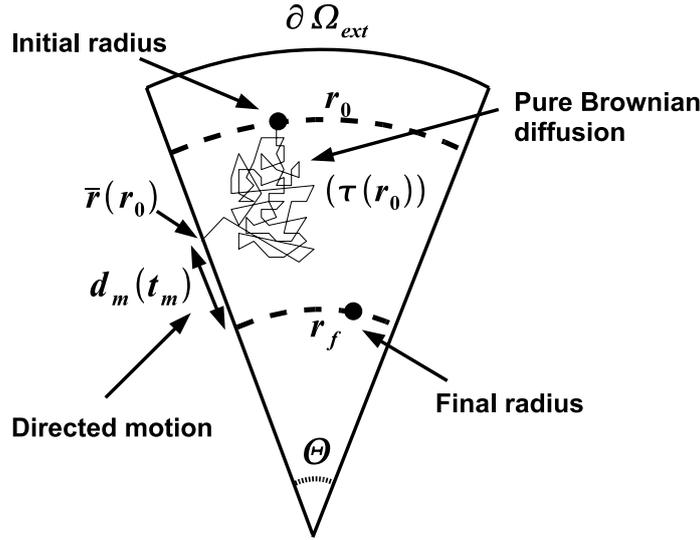


FIGURE 2.3 – A fundamental step in $\tilde{\Omega}$. The virus starts at a radius r_0 , with an angle uniformly distributed in $[0; \Theta]$, it diffuses freely during a time $\tau(r_0)$ until it binds to a MT at a mean radius $\bar{r}(r_0)$; it has then a directed motion over a distance $d_m = \|\mathbf{V}\|t_m$ before being released randomly at a final radius r_f . Mean times of each piece of the fundamental step are written inside parenthesis.

final radius $r_f = \bar{r}(r_0) - \|\mathbf{V}\|t_m$ (see FIG. 2.3). In most eukaryotic cell large asters, there are from 600 to 1000 MTs [65]. We can estimate the average number N of MTs per cell cross section as follow : for a cell thickness $h \approx 9\mu m$, [65], an interaction range $\gamma \approx 50nm$ between the MTs and the molecular motors [69], and for the AAV diameter $d = 30nm$ [14], we obtain for a radial MT organization in a thin cylindrical cell, that the range of N is between $600(2\gamma + d)/h$ and $1000(2\gamma + d)/h$, that is 9 to 15. We are thus in a regime where $\Theta \ll 1$. For $r_0 < R$, by neglecting the reflecting external boundary at $r = R$, $\tilde{\Omega}$ becomes an open wedge and thus using the standard methods from [54] (see also chapter 1), we obtain

$$\tau(r_0) \approx r_0^2 \frac{\Theta^2}{12D} \text{ and } \bar{r}(r_0) \approx r_0 \left(1 + \frac{\Theta^2}{12}\right). \quad (2.6)$$

In radial geometry, $\mathbf{b}(\mathbf{x}) = b(r) \frac{\mathbf{r}}{\|\mathbf{r}\|}$ and the MFPT $u(r_0)$ of a virus starting at r_0 and ending at position r_f , described by equation (2.2) satisfies [21] :

$$\begin{aligned} D\Delta u - b(r_0)\nabla u &= -1 \\ \frac{du}{dr}(R) &= 0 \text{ and } u(r_f) = 0, \end{aligned} \quad (2.7)$$

where we approximated $b(r)$ by $b(r_0)$. The solution of equation (2.7) is

$$u(r_0) = \int_{r_f}^{r_0} \left(\int_v^R \frac{ue^{-\frac{b(r_0)}{D}(u-v)}}{Dv} du \right) dv. \quad (2.8)$$

For $D \ll 1$, using the Laplace method,

$$\int_v^R \frac{ue^{-\frac{b(r_0)}{D}(u-v)}}{Dv} du \approx \frac{1}{b(r_0)}. \quad (2.9)$$

Thus, in first approximation, $u(r_0) \approx \frac{r_0 - r_f}{b(r_0)}$. To obtain the value $b(r_0)$, we equal the MFPT $u(r_0)$ from r_0 to r_f computed from equation (2.2) with the one obtained from an intermittent dynamic : $\tau(r_0) + t_m$. Consequently, we get :

$$b(r_0) = \frac{r_0 - r_f}{\tau(r_0) + t_m} = \frac{d_m - r_0 \frac{\Theta^2}{12}}{t_m + r_0^2 \frac{\Theta^2}{12D}}. \quad (2.10)$$

5 Tests against Brownian simulations

We impose reflecting boundaries at the external membrane $r = R$ and we tested the theoretical steady state distribution against the one obtained by running empirical intermittent Brownian trajectories in the pie wedge domain. For a potential field, the steady state distribution satisfies $D\Delta p - \nabla[\mathbf{b}p] = 0$ in Ω with reflecting boundary condition $\mathbf{J}(\mathbf{x}, t) \cdot \mathbf{n}_x = 0$ on $\partial\Omega$. The distribution p in a two-dimensional radial geometry is :

$$p(r) = \frac{e^{-\frac{\Phi(r)}{D}}}{\int_0^R e^{-\frac{\Phi(r)}{D}} 2\pi r dr}, \quad (2.11)$$

which should be compared to the distribution of [65]. The potential Φ of $b = -\nabla\Phi$ is obtained by integrating equation (2.10) with respect to r ,

$$\begin{aligned} \Phi(r) = & \frac{d_m \sqrt{12Dt_m}}{t_m \Theta} \arctan\left(\frac{\Theta r}{\sqrt{12Dt_m}}\right) \\ & - \frac{D}{2} \ln(12Dt_m + r^2\Theta^2) \end{aligned} \quad (2.12)$$

In FIG. 2.4, we plotted the steady state distribution given in (2.11) against the distribution obtained by the intermittent empirical equation (2.1). The parameters are chosen such that the viruses move towards the nucleus (observed *in vitro*, loaded dynein moves during 1s over a distance of $0.7\mu\text{m}$ [57]), we thus take $t_m = 1\text{s}$ and $d_m = 0.7\mu\text{m}$; furthermore, the diffusion constant is $D = 1.3\mu\text{m}^2\text{s}^{-1}$ as observed for the Associated-Adeno-Virus [14]. The nice agreement of both curves, which is the central result of this communication, confirms that our coarse grained method accounts well for the switch system (2.1).

6 Computation of P_N and τ_N

We derive now asymptotic expressions in the small diffusion limit $D \ll 1$, for the probability P_N and the mean time τ_N a virus arrives at a small nuclear pore. We apply Laplace's method in formulas (2.5) for a radial geometry. When the degradation rate $k(r)$ is taken constant, equal to k_0 in the neighborhood of the nucleus $r = \delta$ and when $12d_m > r\Theta^2$, $b(r) > 0$ so that Φ reaches its minimum at $r = \delta$, we get

$$P_N = \frac{b(\delta)}{\ln\left(\frac{1}{\epsilon}\right) 2\delta k_0 + b(\delta)} \quad \text{and} \quad \tau_N = \frac{\ln\left(\frac{1}{\epsilon}\right) 2\delta}{\ln\left(\frac{1}{\epsilon}\right) 2\delta k_0 + b(\delta)}.$$

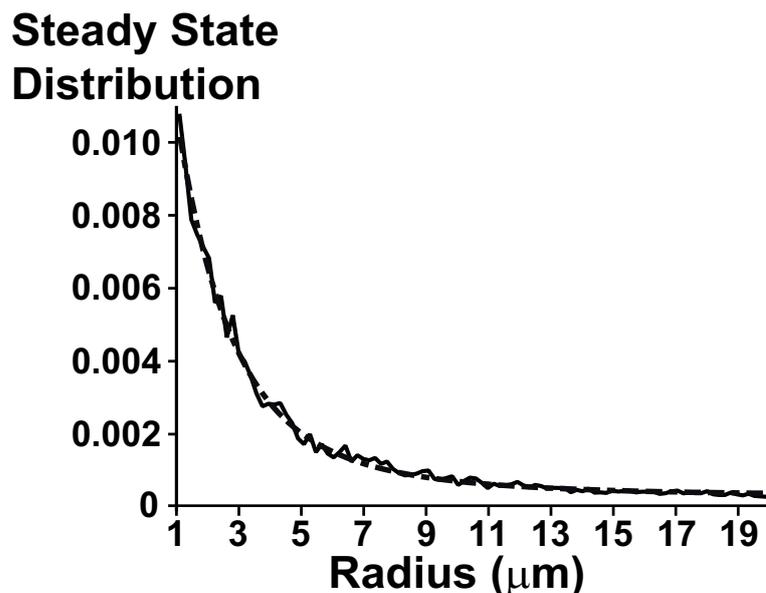


FIGURE 2.4 – Steady state distributions. Dashed line : virus distribution (2.11) with the effective drift $b(r)$ (2.10); solid line : Empirical steady state distribution obtained by running 10,000 intermittent Brownian trajectories. The cell radius is $R = 20\mu m$ and $\Theta = \frac{\pi}{6}$.

A Taylor expansion for $\Theta \ll 1$ gives that

$$P_N \approx \frac{d_m}{d_m + K} \left(1 - \frac{K\delta(d_m\delta + Dt_m)}{12Dt_md_m(d_m + K)}\Theta^2 \right) \quad (2.13)$$

$$\tau_N \approx \frac{K}{k(d_m + K)} \left(1 + \frac{\delta(d_m\delta + Dt_m)}{12Dt_m(d_m + K)}\Theta^2 \right) \quad (2.14)$$

where $K = 2k_0\delta t_m \ln\left(\frac{1}{\epsilon}\right)$.

We can now propose the following predictions : because nuclear pores occupy a fraction $\epsilon = 2\%$ [70] of the nucleus surface (radius $\delta = 8\mu m$) and the measured degradation rate for plasmids [32] is $k = \frac{1}{3600}s^{-1}$, we obtain from formula 2.13-2.14 that

$$P_N \approx 94.3\%, \tau_N \approx 205s. \quad (2.15)$$

We conclude that the infection efficiency is very high, while the mean time to reach a nuclear pore is of the order of 3 minutes. It is interesting to compare this time with the 15 minutes reported in [14], which accounts for all the viral infection steps from the entry to the final nuclear import. This difference between the two times indicates that the phase where the virus is inside an early endosome (EE) may last 10 minutes. Indeed, the endosomal phase ends once the EE has matured into a late endosome (LE) [1], which lasts approximately 10 minutes [2]. To finish, we shall note that a free diffusing virus would reach a nuclear pore in about 15 minutes [30].

The cylindrical geometry. Many transports mechanisms such as viral (herpes virus [23]) and vesicular occur in long axons or dendrites, which can be approximated as thin cylinders (radius R and length L). To derive a quantitative analysis of viral infection in that case, we follow the method described above and compute the steady

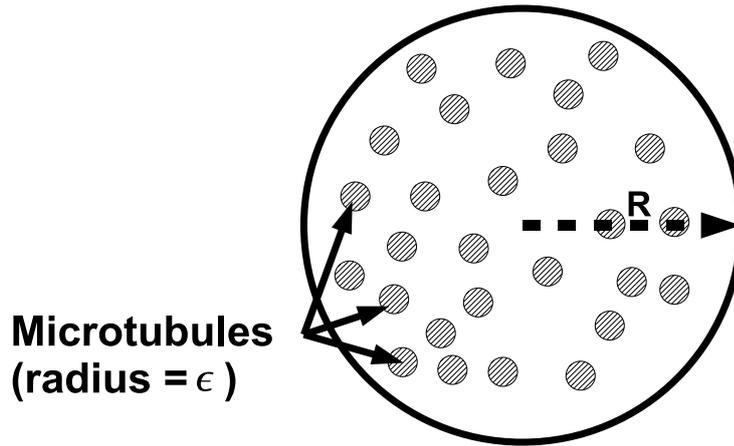


FIGURE 2.5 – Dendrite cross-section. The N MTs are thin cylinders uniformly distributed inside the dendrite.

state drift that accounts for the directed motion along MTs. We model the N MTs parallel to the dendrite principal axis as cylinders (radius $\epsilon \ll R$, Length L). The cross-section Ω of the dendrite is shown in FIG. 2.5. Due to the cylindrical symmetry, for any position \mathbf{x} , the steady state drift $\mathbf{b}(\mathbf{x})$ is equal to $B\mathbf{z}$ where B is a constant and \mathbf{z} the principal axis unit vector along the dendrite. In a small diffusion approximation, the leading order term of B is equal to the effective velocity [24, 25] :

$B = \frac{d_m}{t_m + \tau}$, where t_m is the mean time the virus binds to a MT, $d_m = \|\mathbf{V}\|t_m$ the mean distance of a run and τ the MFPT to a MT. To derive an expression for τ , we consider the cross-section Ω and impose reflecting boundary condition at the external membrane of the dendrite ($r = R$) and absorbing ones at the MTs surfaces. In long time approximation, for a MTs radius $\epsilon \ll 1$, τ is asymptotically equal to $1/(\lambda D)$ where λ is the first eigenvalue of the Laplace operator in Ω with the boundaries conditions described above ([21] p.175). The leading order term of λ as a function of ϵ is [71] $\lambda = \frac{2\pi N}{|\Omega| \ln(\frac{1}{\epsilon})}$, where $|\Omega| = \pi R^2$. Thus, the MFPT to a MT is

$\tau = \frac{1}{\lambda D} = \frac{R^2 \ln(\frac{1}{\epsilon})}{2ND}$, and the steady state drift amplitude B is given by

$$B = \frac{d_m}{t_m + \tau} = \frac{2NDd_m}{2NDt_m + R^2 \ln(\frac{1}{\epsilon})}. \quad (2.16)$$

We conclude that in the limit $t_m \ll \tau$, the effective velocity is proportional to the number of MTs : $B \approx N \frac{2Dd_m}{R^2 \ln(\frac{1}{\epsilon})}$, as already observed in [65].

7 Conclusion

Intermittent dynamics with alternative periods of free diffusion and directed motion along MTs characterizes many cellular transports. We have developed a model to estimate a steady state drift such that the intermittent dynamic can be described by an over-damped limit of the Langevin equation. Our method gives

explicit results in two-dimensional radial cell and in a cylindrical dendrite or axon. The steady state description of the movement enables us to estimate the probability a virus reaches alive a small nuclear pore and its mean time. Because viruses are very efficient DNA carriers, understanding and quantifying their movement in the cell cytoplasm would be very helpful for designing synthetic vectors [72]. In a future work, it would be interesting to derive steady state drifts for three dimensional geometries.

Chapitre 3

Mean time and probability to reach a structured target

1 Introduction.

Cell communication is essential for cellular functions and homeostasis. Diffusive particles such as vesicles or RNA granules [73] have to reach specific targets to deliver their payload or trigger protein synthesis. In some cases, to efficiently pass through the crowded cytoplasm, large particles are intermittently transported on the cytoskeleton by molecular motors such as dyneins that travel along microtubules (MTs) toward the nucleus. In particular, many DNA viruses hijack the cellular transport machinery to reach a nuclear pore and deliver their genetic material [5, 64]. Recently new imaging techniques have allowed to monitor such viral intermittent trajectories *in vivo* [15, 14]. Many of these particles (viruses, granules ...) can be trapped in the crowded cytoplasm or degraded through the ubiquitin-proteasome machinery before reaching their goal. To understand quantitatively the cell biology at a molecular level [28, 74], derive kinetics of chemical reactions that involve a small number of components [75] or quantify early steps of viral infection (see chapter 4), it is important to have precise asymptotics for the conditioned mean first passage time (MFPT) τ_n and the probability P_n a diffusive particle, that can intermittently travel along MTs, reaches a specific target among n . The complex intermittent trajectories of a particle $\mathbf{x}(t)$ can be described using the stochastic rule

$$d\mathbf{x} = \begin{cases} \sqrt{2D}d\mathbf{w} & \text{for } \mathbf{x}(t) \text{ free} \\ \mathbf{V} & \text{for } \mathbf{x}(t) \text{ bound} \end{cases} \quad (3.1)$$

where \mathbf{w} is a standard Brownian motion, D the diffusion constant and \mathbf{V} the velocity of the directed motion along MTs. To pursue the analysis, we coarse grain this complex behavior in a Langevin description of trajectories

$$d\mathbf{x} = \mathbf{b}(\mathbf{x})dt + \sqrt{2D}d\mathbf{w}, \quad (3.2)$$

where the drift $\mathbf{b}(\mathbf{x})$ depends on the cell geometry, the number and distribution of MTs and the rates of binding and unbinding of the free particle to MTs (see chapter 1 and 2). We consider the particles have to reach one over n small partially absorbing disks (radius ϵ) located on the cell ($\partial\Omega_{ext}$) or nucleus membrane ($\partial\Sigma$),

such as nuclear pores. We thus represent the cell cytoplasm as a three-dimensional bounded domain Ω , whose boundary $\partial\Omega = \partial\Omega_{ext} \cup \partial\Sigma$ is reflecting except for n small partially absorbing windows ∂N_a where particles can be absorbed (see FIG.3.1 (left)). We model the degradation activity in the cell cytoplasm as a steady state killing rate $k(\mathbf{x})$ for the trajectories of (3.2), so the survival probability density function (SPDF) is the solution of the Fokker-Planck equation [68]

$$\begin{aligned} \frac{\partial p}{\partial t} &= D\Delta p - \nabla \cdot \mathbf{b}p - kp \\ p(\mathbf{x}, 0) &= p_i(\mathbf{x}) \end{aligned}$$

with the boundary conditions :

$$\mathbf{J}(\mathbf{x}, t) \cdot \mathbf{n}_x = \kappa p(\mathbf{x}, t) \text{ on } \partial N_a \text{ and } \mathbf{J}(\mathbf{x}, t) \cdot \mathbf{n}_x = 0 \text{ on } \partial\Omega - \partial N_a \quad (3.3)$$

where \mathbf{n}_x denotes the normal derivative at a boundary point \mathbf{x} and κ is a negative real constant. The flux density vector $\mathbf{J}(\mathbf{x}, t)$ is defined as

$$\mathbf{J}(\mathbf{x}, t) = -D\nabla p(\mathbf{x}, t) + \mathbf{b}(\mathbf{x})p(\mathbf{x}, t). \quad (3.4)$$

If $\tilde{p}(\mathbf{x}) = \int_0^\infty p(\mathbf{x}, t)dt$ and $q(\mathbf{x}) = \int_0^\infty tp(\mathbf{x}, t)dt$ the probability P_n and the conditioned MFPT τ_n the particle reaches ∂N_a before being killed are [26] :

$$P_n = 1 - \int_{\Omega} k(\mathbf{x})\tilde{p}(\mathbf{x})d\mathbf{x}, \quad (3.5)$$

$$\tau_n = \frac{\int_{\Omega} \tilde{p}(\mathbf{x})d\mathbf{x} - \int_{\Omega} k(\mathbf{x})q(\mathbf{x})d\mathbf{x}}{1 - \int_{\Omega} k(\mathbf{x})\tilde{p}(\mathbf{x})d\mathbf{x}}. \quad (3.6)$$

For pure absorbing disks ($\kappa \rightarrow \infty$), when the drift \mathbf{b} derives from a potential Φ ($\mathbf{b}(\mathbf{x}) = -\nabla\Phi(\mathbf{x})$), asymptotics in ϵ of P_n and τ_n have been derived in [26]

$$\left\{ \begin{aligned} P_n &= \frac{e^{-\frac{\Phi_0}{D}}}{\frac{1}{4Dn\epsilon} \int_{\Omega} e^{-\frac{\Phi(\mathbf{x})}{D}} k(\mathbf{x})d\mathbf{x} + e^{-\frac{\Phi_0}{D}}} \\ \tau_n &= \frac{\frac{1}{4Dn\epsilon} \int_{\Omega} e^{-\frac{\Phi(\mathbf{x})}{D}} d\mathbf{x}}{\frac{1}{4Dn\epsilon} \int_{\Omega} e^{-\frac{\Phi(\mathbf{x})}{D}} k(\mathbf{x})d\mathbf{x} + e^{-\frac{\Phi_0}{D}}} \end{aligned} \right. \quad (3.7)$$

where Φ_0 is the constant value of the radial potential $\Phi(\mathbf{x})$ on the membrane of the centered nucleus. In the chapter 4, we test these asymptotic expressions against Brownian simulations for a single hole ($n = 1$). These formulas do not account for the possible interactions between the small absorbing pores. Because

$$\lim_{n \rightarrow \infty, n\epsilon^2 \ll 1} \tau_n = 0,$$

when the number of pores becomes too large, these expressions are no more valid and a correction term accounting for the holes interactions is needed. Recent studies have begun to quantify the interactions between the absorbing windows [27, 28]. When the $n \gg 1$ windows are uniformly distributed over a small structure Σ , the

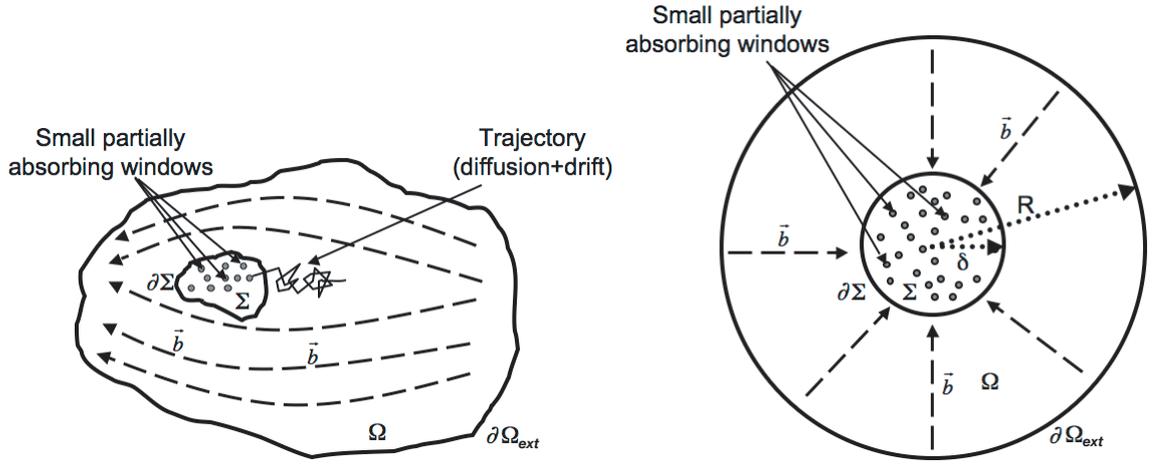


FIGURE 3.1 – Scheme of the general 3-dimensional domain Ω (left-hand side) : Langevin trajectories, solutions of (3.2), can be killed or absorbed at the small windows located on $\partial\Sigma$. On the right-hand side, the simplified spherical cell (radius R) with a spherical centered nucleus (radius δ) is represented.

leading order term of the narrow escape time τ_n of a pure diffusive particle in the limit $\frac{|\Sigma|}{|\Omega|} \ll 1$ has been derived in [29] with electrostatic considerations :

$$\tau_n = \frac{|\Omega|}{D} \left(\frac{1}{C_\Sigma} + \frac{f(\sigma, \kappa)}{4n\epsilon} - \frac{D}{\kappa|\partial\Sigma|} \left(1 + \frac{f(\sigma, \kappa)}{\sigma} \right) \right), \quad (3.8)$$

where C_Σ is the capacitance of the conducting surface $\partial\Sigma$ and $\sigma = \frac{n\pi\epsilon^2}{|\partial\Sigma|}$ is the fraction of the structure covered by the partially absorbing disks. For example, if the structure is a sphere of radius r , we have $C_\Sigma = 4\pi r$ (and $C_\Sigma = 2\pi r$ for a hemispherical spine located on the membrane). We can approximate [29]

$$f(\sigma, \kappa) \approx f(\sigma) + \frac{4D}{\pi\epsilon\kappa} \frac{1 - \sigma - f(\sigma)}{1 + \frac{4D}{\pi\epsilon\kappa}} \quad (3.9)$$

where $f(\sigma)$ has been successively approached by

$$f(\sigma) = \begin{cases} 1 & [76] \\ 1 - \sigma & [77] \\ \frac{\sigma}{F(\sigma)} = \frac{1 - \sigma}{1 + 3.8\sigma^{1.25}} & [78] \end{cases} \quad (3.10)$$

By neglecting the geometrical interactions between the holes ($\frac{1}{C_\Sigma} = 0$), for a small coverage ($\sigma \ll 1$) and perfectly absorbing disks ($\kappa \rightarrow \infty$), the mean time expression (3.8) reduces to (3.7) for pure diffusive particles ($\mathbf{b} = \mathbf{0}$) with no killing activity ($k = 0$).

In this chapter, accounting for the interactions between the partially absorbing windows, we compute the probability P_n and the conditioned MFPT τ_n a diffusive particle that can be degraded ($k \neq 0$) and actively transported ($\mathbf{b} \neq \mathbf{0}$) reaches a window. P_n and τ_n are solutions of a simple linear system and when the small holes concentrate on a structure Σ such as the nucleus, we provide concise leading order

asymptotics in the limit $\frac{|\Sigma|}{|\Omega|} \ll 1$. These asymptotics generalizes formula (3.8) for a small coverage $\sigma \ll 1$. In particular, these new results apply to viruses ($\mathbf{b} \neq \mathbf{0}$) and non-viral gene vectors (pure diffusive $\mathbf{b} = \mathbf{0}$) that have to reach one over the 2,000 nuclear pores [70] that cover the nucleus to deliver their genetic material before being trapped or degraded through the ubiquitin-proteasome machinery.

2 Asymptotic derivations of P_n and τ_n

Because the 2,000 nuclear pores of radius $25nm$ occupy only 2% of the nuclear surface (radius $\delta = 5\mu m$ (Chinese hamster ovary cells [70])), we neglect the fraction σ of $\partial\Sigma$ that is covered by the partially absorbing windows. Without any loss of generality, we consider the n partially absorbing windows $\partial\Omega_i$ of identical radius ϵ are centered at positions $(\mathbf{x}_i)_{i=1}^n$ on $\partial\Sigma$. The SPDF is thus solution of [68]

$$\begin{aligned} \frac{\partial p}{\partial t} &= D\Delta p - \nabla \cdot \mathbf{b}p - kp \\ p(\mathbf{x}, 0) &= p_i(\mathbf{x}) \end{aligned}$$

with the boundary conditions :

$$\begin{aligned} \mathbf{J}(\mathbf{x}, t) \cdot \mathbf{n}_{\mathbf{x}} &= \kappa p(\mathbf{x}, t) \text{ on } \partial N_a = \bigcup_{i=1}^n \partial\Omega_i \\ \text{and } \mathbf{J}(\mathbf{x}, t) \cdot \mathbf{n}_{\mathbf{x}} &= 0 \text{ on } \partial\Omega_{ext} \bigcup (\partial\Sigma - \partial N_a) \end{aligned} \quad (3.11)$$

where $\partial\Sigma$ is the boundary of the microstructure Σ .

2.1 Asymptotic derivation of P_n

We introduce the Neumann function $\mathcal{N}(\mathbf{x}, \mathbf{x}_0)$ solution of the differential equation [26]

$$D\Delta\mathcal{N}(\mathbf{x}, \mathbf{x}_0) = -\delta_{\mathbf{x}_0}(\mathbf{x}), \quad \mathbf{x} \in \Omega, \quad (3.12)$$

$$D\frac{\partial\mathcal{N}}{\partial n}(\mathbf{x}, \mathbf{x}_0) = -\frac{1}{|\partial\Omega|} \quad \mathbf{x} \in \partial\Omega, \quad (3.13)$$

and we first compute

$$\begin{aligned} I &= \int_{\Omega} (D\Delta\tilde{p}(\mathbf{x}) - \nabla \cdot \mathbf{b}\tilde{p}(\mathbf{x}) - k\tilde{p}(\mathbf{x})) \mathcal{N}(\mathbf{x}, \mathbf{x}_0) d\mathbf{x} \\ &\quad - \int_{\Omega} D\Delta\mathcal{N}(\mathbf{x}, \mathbf{x}_0) \tilde{p}(\mathbf{x}) d\mathbf{x}, \end{aligned} \quad (3.14)$$

where $\tilde{p}(\mathbf{x}) = \int_0^{\infty} p(\mathbf{x}, t) dt$ is solution of the differential equation

$$D\Delta\tilde{p}(\mathbf{x}) - \nabla \cdot \mathbf{b}\tilde{p}(\mathbf{x}) - k(\mathbf{x})\tilde{p}(\mathbf{x}) = -p_i(\mathbf{x}) \quad (3.15)$$

with the boundary conditions

$$\begin{aligned} \tilde{\mathbf{J}}(\mathbf{x}) \cdot \mathbf{n}_{\mathbf{x}} &= \kappa\tilde{p}(\mathbf{x}) \text{ on } \partial N_a = \bigcup_{i=1}^n \partial\Omega_i \\ \text{and } \tilde{\mathbf{J}}(\mathbf{x}) \cdot \mathbf{n}_{\mathbf{x}} &= 0 \text{ on } \partial\Omega_{ext} \bigcup (\partial\Sigma - \partial N_a) \end{aligned} \quad (3.16)$$

where $\tilde{\mathbf{J}}(\mathbf{x}) = -D\nabla\tilde{p}(\mathbf{x}) + \mathbf{b}(\mathbf{x})\tilde{p}(\mathbf{x})$. Consequently, on the one hand, (3.13) and (3.15) yields $I = -\int_{\Omega} p_i(\mathbf{x})\mathcal{N}(\mathbf{x}, \mathbf{x}_0) + \tilde{p}(\mathbf{x}_0)$ and, on the other hand, Green's identity leads to

$$\begin{aligned} I &= -\int_{\partial N_a} \tilde{\mathbf{J}}(\mathbf{x}) \cdot \mathbf{n}_{\mathbf{x}} \mathcal{N}(\mathbf{x}, \mathbf{x}_0) d\mathbf{x} + \int_{\Omega} \mathbf{b}(\mathbf{x}) \cdot \nabla \mathcal{N}(\mathbf{x}, \mathbf{x}_0) \tilde{p}(\mathbf{x}) d\mathbf{x} \\ &\quad - \int_{\Omega} k(\mathbf{x}) \tilde{p}(\mathbf{x}) \mathcal{N}(\mathbf{x}, \mathbf{x}_0) d\mathbf{x} + \frac{1}{|\partial\Omega|} \int_{\partial\Omega} \tilde{p}(\mathbf{x}) d\mathbf{x}. \end{aligned} \quad (3.17)$$

Thus, we have

$$\begin{aligned} \int_{\Omega} (k(\mathbf{x})\tilde{p}(\mathbf{x}) - p_i(\mathbf{x})) \mathcal{N}(\mathbf{x}, \mathbf{x}_0) d\mathbf{x} &= -\int_{\partial N_a} \tilde{\mathbf{J}}(\mathbf{x}) \cdot \mathbf{n}_{\mathbf{x}} \mathcal{N}(\mathbf{x}, \mathbf{x}_0) d\mathbf{x} \\ &\quad + \int_{\Omega} \mathbf{b}(\mathbf{x}) \cdot \nabla \mathcal{N}(\mathbf{x}, \mathbf{x}_0) \tilde{p}(\mathbf{x}) d\mathbf{x} \\ &\quad + \frac{1}{|\partial\Omega|} \int_{\partial\Omega} \tilde{p}(\mathbf{x}) d\mathbf{x} - \tilde{p}(\mathbf{x}_0). \end{aligned} \quad (3.18)$$

When the degradation rate is small ($k \ll 1$), for $\mathbf{x} \in \Omega$ at a distance $O(1)$ (with respect to ϵ) away from the windows, we consider the long time approximation [26, 79],

$$\tilde{p}(\mathbf{x}) \approx C_{\epsilon} e^{-\frac{\Phi(\mathbf{x})}{D}}, \quad \mathbf{x} \in \Omega, \quad (3.19)$$

where

$$\lim_{\epsilon \rightarrow 0} C_{\epsilon} = +\infty. \quad (3.20)$$

In particular, we approximate (see [79] for details)

$$\frac{1}{|\partial\Omega|} \int_{\partial\Omega} \tilde{p}(\mathbf{x}) d\mathbf{x} + \int_{\Omega} \mathbf{b}(\mathbf{x}) \cdot \nabla \mathcal{N}(\mathbf{x}, \mathbf{x}_0) \tilde{p}(\mathbf{x}) d\mathbf{x} \approx C_{\epsilon} e^{-\frac{\Phi(\mathbf{x}_0)}{D}}. \quad (3.21)$$

Consequently, (3.18) reduces to

$$\begin{aligned} \int_{\Omega} (k(\mathbf{x})\tilde{p}(\mathbf{x}) - p_i(\mathbf{x})) \mathcal{N}(\mathbf{x}, \mathbf{x}_0) d\mathbf{x} &= -\int_{\partial N_a} \tilde{\mathbf{J}}(\mathbf{x}) \cdot \mathbf{n}_{\mathbf{x}} \mathcal{N}(\mathbf{x}, \mathbf{x}_0) d\mathbf{x} \\ &\quad + C_{\epsilon} e^{-\frac{\Phi(\mathbf{x}_0)}{D}} - \tilde{p}(\mathbf{x}_0). \end{aligned} \quad (3.22)$$

The leading order term $g_i(s = |\mathbf{x} - \mathbf{x}_i|)$ of the flux $\tilde{\mathbf{J}}(s) \cdot \mathbf{n}_s$ in the small absorbing disk $\partial\Omega_i$ of radius ϵ has the form [80]

$$g_i(s) = \frac{g_0^i}{\sqrt{\epsilon^2 - s^2}}, \quad (3.23)$$

where g_0^i is a constant. Integrating (3.15) over Ω , we obtain

$$\int_{\bigcup_{i=1}^n \partial\Omega_i} \tilde{\mathbf{J}}(\mathbf{x}) \cdot \mathbf{n}_{\mathbf{x}} d\mathbf{x} = \int_{\Omega} p_i(\mathbf{x}) - k(\mathbf{x})\tilde{p}(\mathbf{x}) d\mathbf{x} = 1 - \int_{\Omega} k(\mathbf{x})\tilde{p}(\mathbf{x}) d\mathbf{x}. \quad (3.24)$$

Injecting the long time asymptotic (3.19) and the leading order form (3.23) of the flux in the expression above, we obtain

$$\sum_{i=1}^n g_0^i \int_{s=0}^{\epsilon} \frac{2\pi s ds}{\sqrt{\epsilon^2 - s^2}} = 2\pi\epsilon \sum_{i=1}^n g_0^i = 1 - C_\epsilon \int_{\Omega} k(\mathbf{x}) e^{-\frac{\Phi(\mathbf{x})}{D}} d\mathbf{x}. \quad (3.25)$$

Without any loss of generality, we choose $\mathbf{x}_0 = \mathbf{x}_i$, $1 \leq i \leq n$ and, due to the partially absorbing boundary condition (3.16), we approximate

$$\tilde{p}(\mathbf{x}_i) = \frac{-1}{\kappa|\partial\Omega_i|} \int_0^\epsilon \frac{2\pi s g_0^i ds}{\sqrt{\epsilon^2 - s^2}} = \frac{-2g_0^i}{\kappa\epsilon}. \quad (3.26)$$

Injecting (3.26) and the flux form (3.23) in (3.22), we obtain

$$\begin{aligned} \int_{\Omega} (k(\mathbf{x})\tilde{p}(\mathbf{x}) - p_i(\mathbf{x})) \mathcal{N}(\mathbf{x}, \mathbf{x}_i) d\mathbf{x} &= - \int_0^\epsilon \frac{2\pi s g_0^i}{\sqrt{\epsilon^2 - s^2}} \mathcal{N}(s, \mathbf{x}_i) ds - \sum_{j=1, j \neq i}^n \int_{\partial\Omega_j} g_j(\mathbf{x}) \mathcal{N}(\mathbf{x}, \mathbf{x}_i) d\mathbf{x} \\ &+ C_\epsilon e^{-\frac{\Phi(\mathbf{x}_i)}{D}} + \frac{2g_0^i}{\kappa\epsilon}. \end{aligned} \quad (3.27)$$

For \mathbf{x}_0 on the domain boundary [81]

$$\mathcal{N}(\mathbf{x}, \mathbf{x}_0) = \frac{1}{2\pi D |\mathbf{x} - \mathbf{x}_0|} + \omega_{\mathbf{x}_0}(\mathbf{x}), \quad (3.28)$$

where $\omega_{\mathbf{x}_0}$ is a regular harmonic function. For $i \neq j$, we assume $|\mathbf{x}_i - \mathbf{x}_j| \gg \epsilon$ and that for $\mathbf{x} \in \partial\Omega_j$, $\mathcal{N}(\mathbf{x}, \mathbf{x}_i) \approx \mathcal{N}(\mathbf{x}_j, \mathbf{x}_i)$. Consequently, we asymptotically have

$$\begin{aligned} \int_{\Omega} (k(\mathbf{x})\tilde{p}(\mathbf{x}) - p_i(\mathbf{x})) \mathcal{N}(\mathbf{x}, \mathbf{x}_i) d\mathbf{x} &= - \int_0^\epsilon \frac{g_0^i}{D\sqrt{\epsilon^2 - s^2}} ds - \sum_{j=1, j \neq i}^n \mathcal{N}(\mathbf{x}_j, \mathbf{x}_i) \int_0^\epsilon \frac{g_j^0 2\pi s}{\sqrt{\epsilon^2 - s^2}} ds \\ &+ C_\epsilon e^{-\frac{\Phi(\mathbf{x}_i)}{D}} + \frac{2g_0^i}{\kappa\epsilon}, \end{aligned} \quad (3.29)$$

that is

$$\begin{aligned} \int_{\Omega} (k(\mathbf{x})\tilde{p}(\mathbf{x}) - p_i(\mathbf{x})) \mathcal{N}(\mathbf{x}, \mathbf{x}_i) d\mathbf{x} &= \left(\frac{2}{\kappa\epsilon} - \frac{\pi}{2D} \right) g_0^i \\ &- 2\pi\epsilon \sum_{j=1, j \neq i}^n \mathcal{N}(\mathbf{x}_j, \mathbf{x}_i) g_j^0 + C_\epsilon e^{-\frac{\Phi(\mathbf{x}_i)}{D}}. \end{aligned} \quad (3.30)$$

The integral $\int_{\Omega} p_i(\mathbf{x}) \mathcal{N}(\mathbf{x}, \mathbf{x}_i) d\mathbf{x}$ is uniformly bounded for smooth initial distributions p_i as $\epsilon \rightarrow 0$ (and is an integrable singularity for non smooth distributions (see [79] for details)) while all other terms in (3.31) are unbounded in view of (3.20). Consequently, in the small degradation rate $k \ll 1$ limit, we have

$$\left(\frac{\pi}{2D} - \frac{2}{\kappa\epsilon} \right) g_0^i + 2\pi\epsilon \sum_{j=1, j \neq i}^n \mathcal{N}(\mathbf{x}_j, \mathbf{x}_i) g_j^0 = C_\epsilon e^{-\frac{\Phi(\mathbf{x}_i)}{D}}. \quad (3.31)$$

Finally, by solving the linear system

$$\begin{aligned} \left(\frac{\pi}{2D} - \frac{2}{\kappa\epsilon}\right) g_0^i + 2\pi\epsilon \sum_{j=1, j \neq i}^n \mathcal{N}(\mathbf{x}_j, \mathbf{x}_i) g_j^0 &= C_\epsilon e^{-\frac{\Phi(\mathbf{x}_i)}{D}}, \text{ for } 1 \leq i \leq n \\ 2\pi\epsilon \sum_{i=1}^n g_0^i &= 1 - C_\epsilon \int_{\Omega} k(\mathbf{x}) e^{-\frac{\Phi(\mathbf{x})}{D}} d\mathbf{x}. \end{aligned} \quad (3.32)$$

we can get the $n + 1$ unknowns $(C_\epsilon, g_1^0, \dots, g_n^0)$ and compute

$$P_n = 1 - \int_{\Omega} k(\mathbf{x}) \tilde{p}(\mathbf{x}) d\mathbf{x} = 1 - C_\epsilon \int_{\Omega} k(\mathbf{x}) e^{-\frac{\Phi(\mathbf{x})}{D}} d\mathbf{x}. \quad (3.33)$$

We point out that the linear system (3.32) derived with an integral method is very similar to the linear system (3.16) obtained with a method of matched asymptotic expansions in [28].

Leading order asymptotic of P_n in the limit $\frac{|\Sigma|}{|\Omega|} \ll 1$

Summing the n equations that compose the linear system (3.32) we have

$$\left(\frac{\pi}{2D} - \frac{2}{\kappa\epsilon}\right) \sum_{i=1}^n g_0^i + 2\pi\epsilon \sum_{i=1}^n \sum_{j=1, j \neq i}^n \mathcal{N}(\mathbf{x}_j, \mathbf{x}_i) g_j^0 = C_\epsilon \sum_{i=1}^n e^{-\frac{\Phi(\mathbf{x}_i)}{D}}. \quad (3.34)$$

Because for all $j \neq i$, $\mathcal{N}(\mathbf{x}_j, \mathbf{x}_i) = \mathcal{N}(\mathbf{x}_i, \mathbf{x}_j)$, we get from the equation above :

$$\begin{aligned} \left(\frac{\pi}{2D} - \frac{2}{\kappa\epsilon}\right) \frac{1 - C_\epsilon \int_{\Omega} k(\mathbf{x}) e^{-\frac{\Phi(\mathbf{x})}{D}} d\mathbf{x}}{2\pi\epsilon} + 2\pi\epsilon \sum_{i=1}^n g_0^i \sum_{j=1, j \neq i}^n \mathcal{N}(\mathbf{x}_j, \mathbf{x}_i) \\ = C_\epsilon \sum_{i=1}^n e^{-\frac{\Phi(\mathbf{x}_i)}{D}}. \end{aligned} \quad (3.35)$$

When the $n \gg 1$ holes are uniformly distributed over $\partial\Sigma$ we approximate

$$\frac{1}{n} \sum_{j=1, j \neq i}^n \mathcal{N}(\mathbf{x}_j, \mathbf{x}_i) \approx \frac{\int_{\partial\Sigma} \mathcal{N}(\mathbf{x}, \mathbf{x}_i) d\mathbf{x}}{|\partial\Sigma|}, \quad (3.36)$$

and

$$\frac{1}{n} \sum_{i=1}^n e^{-\frac{\Phi(\mathbf{x}_i)}{D}} \approx \frac{\int_{\partial\Sigma} e^{-\frac{\Phi(\mathbf{x})}{D}} d\mathbf{x}}{|\partial\Sigma|}. \quad (3.37)$$

Consequently, (3.35) reduces to

$$\begin{aligned} \left(\frac{\pi}{2D} - \frac{2}{\kappa\epsilon}\right) \frac{1 - C_\epsilon \int_{\Omega} k(\mathbf{x}) e^{-\frac{\Phi(\mathbf{x})}{D}} d\mathbf{x}}{2\pi\epsilon} + 2\pi n\epsilon \sum_{i=1}^n g_0^i \frac{\int_{\partial\Sigma} \mathcal{N}(\mathbf{x}, \mathbf{x}_i) d\mathbf{x}}{|\partial\Sigma|} \\ = n C_\epsilon \frac{\int_{\partial\Sigma} e^{-\frac{\Phi(\mathbf{x})}{D}} d\mathbf{x}}{|\partial\Sigma|}. \end{aligned} \quad (3.38)$$

For $\mathbf{x}_i \in \partial\Sigma$, the leading order of $\frac{\int_{\partial\Sigma} \mathcal{N}(\mathbf{x}, \mathbf{x}_i) d\mathbf{x}}{|\partial\Sigma|}$ in the limit $\frac{|\Sigma|}{|\Omega|} \ll 1$ is (see the appendix for the detailed computations)

$$\frac{\int_{\partial\Sigma} \mathcal{N}(\mathbf{x}, \mathbf{x}_i) d\mathbf{x}}{|\partial\Sigma|} \approx \tau = \frac{\int_{\Omega} u(\mathbf{x}) d\mathbf{x}}{|\Omega|}, \quad (3.39)$$

where $u(\mathbf{x})$ is the MFPT to $\partial\Sigma$ of a diffusing particle (diffusion constant D) that starts from a position \mathbf{x} in Ω . If C_{Σ} is the capacitance of the conducting surface $\partial\Sigma$, then (equation (27) in [29])

$$\tau = \frac{|\Omega|}{DC_{\Sigma}}. \quad (3.40)$$

Consequently, by identifying (3.39) and (3.40) we obtain, for $\mathbf{x}_i \in \partial\Sigma$

$$\frac{\int_{\partial\Sigma} \mathcal{N}(\mathbf{x}, \mathbf{x}_i) d\mathbf{x}}{|\partial\Sigma|} = \frac{1}{DC_{\Sigma}}. \quad (3.41)$$

Replacing (3.41) in (3.35) we have

$$\left(\frac{\pi}{2D} - \frac{2}{\kappa\epsilon} \right) \frac{1 - C_{\epsilon} \int_{\Omega} k(\mathbf{x}) e^{-\frac{\Phi(\mathbf{x})}{D}} d\mathbf{x}}{2\pi\epsilon} + 2\pi n\epsilon \sum_{i=1}^n g_i^0 \frac{1}{DC_{\Sigma}} = nC_{\epsilon} \frac{\int_{\partial\Sigma} e^{-\frac{\Phi(\mathbf{x})}{D}} d\mathbf{x}}{|\partial\Sigma|}, \quad (3.42)$$

that is

$$\begin{aligned} \left(\frac{\pi}{2D} - \frac{2}{\kappa\epsilon} \right) \frac{1 - C_{\epsilon} \int_{\Omega} k(\mathbf{x}) e^{-\frac{\Phi(\mathbf{x})}{D}} d\mathbf{x}}{2\pi\epsilon} + \frac{n}{DC_{\Sigma}} \left(1 - C_{\epsilon} \int_{\Omega} k(\mathbf{x}) e^{-\frac{\Phi(\mathbf{x})}{D}} d\mathbf{x} \right) \\ = nC_{\epsilon} \frac{\int_{\partial\Sigma} e^{-\frac{\Phi(\mathbf{x})}{D}} d\mathbf{x}}{|\partial\Sigma|}. \end{aligned} \quad (3.43)$$

Finally, we find that

$$C_{\epsilon} = \frac{\frac{1}{4nD\epsilon} - \frac{1}{\kappa\pi n\epsilon^2} + \frac{1}{DC_{\Sigma}}}{\frac{\int_{\partial\Sigma} e^{-\frac{\Phi(\mathbf{x})}{D}} d\mathbf{x}}{|\partial\Sigma|} + \left(\frac{1}{4nD\epsilon} - \frac{1}{\kappa\pi n\epsilon^2} + \frac{1}{DC_{\Sigma}} \right) \int_{\Omega} k(\mathbf{x}) e^{-\frac{\Phi(\mathbf{x})}{D}} d\mathbf{x}}. \quad (3.44)$$

Reinjecting (3.44) in (3.33), we find

$$\begin{aligned} P_N &= \frac{\frac{\int_{\partial\Sigma} e^{-\frac{\Phi(\mathbf{x})}{D}} d\mathbf{x}}{|\partial\Sigma|}}{\frac{\int_{\partial\Sigma} e^{-\frac{\Phi(\mathbf{x})}{D}} d\mathbf{x}}{|\partial\Sigma|} + \left(\frac{1}{4nD\epsilon} - \frac{1}{\kappa\pi n\epsilon^2} + \frac{1}{DC_{\Sigma}} \right) \int_{\Omega} k(\mathbf{x}) e^{-\frac{\Phi(\mathbf{x})}{D}} d\mathbf{x}} \\ &= \frac{\frac{\int_{\partial\Sigma} e^{-\frac{\Phi(\mathbf{x})}{D}} d\mathbf{x}}{|\partial\Sigma|}}{\frac{\int_{\partial\Sigma} e^{-\frac{\Phi(\mathbf{x})}{D}} d\mathbf{x}}{|\partial\Sigma|} + \alpha \int_{\Omega} k(\mathbf{x}) e^{-\frac{\Phi(\mathbf{x})}{D}} d\mathbf{x}}, \end{aligned} \quad (3.45)$$

where

$$\alpha = \frac{1}{4nD\epsilon} - \frac{1}{\kappa\pi n\epsilon^2} + \frac{1}{DC_{\Sigma}}. \quad (3.46)$$

2.2 Asymptotic derivation of τ_n

We introduce $q(\mathbf{x}) = \int_0^\infty tp(\mathbf{x}, t)dt$, solution of the partial differential equation

$$D\Delta q(\mathbf{x}) - \nabla \cdot \mathbf{b}q(\mathbf{x}) - k(\mathbf{x})q(\mathbf{x}) = -\tilde{p}(\mathbf{x}) \quad (3.47)$$

with the boundary condition

$$\begin{aligned} \mathbf{J}_q(\mathbf{x}) \cdot \mathbf{n}_x &= \kappa q(\mathbf{x}) \text{ on } \partial N_a = \bigcup_{i=1}^n \partial \Omega_i \\ \text{and } \mathbf{J}_q(\mathbf{x}) \cdot \mathbf{n}_x &= 0 \text{ on } \partial \Omega_{ext} \bigcup (\partial \Sigma - \partial N_a) \end{aligned} \quad (3.48)$$

where $\mathbf{J}_q(\mathbf{x}) = -D\nabla q(\mathbf{x}) + \mathbf{b}(\mathbf{x})q(\mathbf{x})$. We compute

$$\int_{\Omega} (D\Delta q(\mathbf{x}) - \nabla \cdot \mathbf{b}q(\mathbf{x}) - k\tilde{p}(\mathbf{x})) \mathcal{N}(\mathbf{x}, \mathbf{x}_0) d\mathbf{x} - \int_{\Omega} D\Delta \mathcal{N}(\mathbf{x}, \mathbf{x}_0) q(\mathbf{x}) d\mathbf{x} \quad (3.49)$$

and we consider the long time approximation [26],

$$q(\mathbf{x}) \approx T_\epsilon e^{-\frac{\Phi(\mathbf{x})}{D}}, \quad \mathbf{x} \in \Omega. \quad (3.50)$$

We consider the $n \gg 1$ small windows are uniformly distributed over $\partial \Sigma$ and very similar computations to those performed for the derivation of P_N lead to the leading order asymptotic of T_ϵ in the limit $\frac{|\Sigma|}{|\Omega|} \ll 1$

$$T_\epsilon \approx \frac{\alpha^2 \int_{\Omega} e^{-\frac{\Phi(\mathbf{x})}{D}} d\mathbf{x}}{\left(\frac{\int_{\partial \Sigma} e^{-\frac{\Phi(\mathbf{x})}{D}} d\mathbf{x}}{|\partial \Sigma|} + \alpha \int_{\Omega} k(\mathbf{x}) e^{-\frac{\Phi(\mathbf{x})}{D}} d\mathbf{x} \right)^2}, \quad (3.51)$$

where α is given by (3.46). Thus, an asymptotic for the mean time

$$\tau_n = \frac{C_\epsilon \int_{\Omega} e^{-\frac{\Phi(\mathbf{x})}{D}} d\mathbf{x} - T_\epsilon \int_{\Omega} k(\mathbf{x}) e^{-\frac{\Phi(\mathbf{x})}{D}} d\mathbf{x}}{1 - C_\epsilon \int_{\Omega} k(\mathbf{x}) e^{-\frac{\Phi(\mathbf{x})}{D}} d\mathbf{x}}$$

is

$$\tau_n \approx \frac{\alpha \int_{\Omega} e^{-\frac{\Phi(\mathbf{x})}{D}} d\mathbf{x}}{\frac{\int_{\partial \Sigma} e^{-\frac{\Phi(\mathbf{x})}{D}} d\mathbf{x}}{|\partial \Sigma|} + \alpha \int_{\Omega} k(\mathbf{x}) e^{-\frac{\Phi(\mathbf{x})}{D}} d\mathbf{x}}. \quad (3.52)$$

Replacing α by its value (3.46), we finally have

$$\tau_n \approx \frac{\left(\frac{1}{4nD\epsilon} - \frac{1}{\kappa\pi n\epsilon^2} + \frac{1}{DC_\Sigma} \right) \int_{\Omega} e^{-\frac{\Phi(\mathbf{x})}{D}} d\mathbf{x}}{\frac{\int_{\partial \Sigma} e^{-\frac{\Phi(\mathbf{x})}{D}} d\mathbf{x}}{|\partial \Sigma|} + \left(\frac{1}{4nD\epsilon} - \frac{1}{\kappa\pi n\epsilon^2} + \frac{1}{DC_\Sigma} \right) \int_{\Omega} k(\mathbf{x}) e^{-\frac{\Phi(\mathbf{x})}{D}} d\mathbf{x}}. \quad (3.53)$$

When the ratio $\sigma = \frac{n\pi\epsilon^2}{|\Sigma|}$ of the nuclear surface that is occupied by the small partially absorbing holes tends to 0, the function $f(\sigma, \kappa)$ given by (3.9) tends to 1. Consequently, for pure diffusing particles ($\mathbf{b}(\mathbf{x}) = 0$) with no degradation activity ($k = 0$), the mean time (3.53) reduces to the asymptotic formula (54) in [29].

3 Tests against Brownian simulations

We consider a spherical cell (radius R) with a centered spherical nucleus Σ (radius δ) uniformly covered by n small pure absorbing pores (radius ϵ) (see FIG. 3.1 (right)). We impose reflecting boundaries at the external membrane $r = R$ and at the nuclear surface, excepting at nuclear pores $\partial N_a = \bigcup_{i=1}^n \partial \Omega_i$, centered at uniformly distributed locations $(\mathbf{x}_i)_{i=1}^n$. We consider a constant radial drift B directed toward the nucleus (potential $\Phi(r) = -Br$). Because Σ is a sphere of radius δ , we have that $C_\Sigma = 4\pi\delta$ and

$$\alpha = \frac{1}{4nD\epsilon} + \frac{1}{4\pi D\delta}. \quad (3.54)$$

To have concise expressions, we assume the killing rate is constant $k(\mathbf{x}) = k_0$ and consequently (3.45) and (3.53) lead to

$$P_n = \frac{e^{-\frac{B\delta}{D}}}{\alpha \left(e^{-\frac{B\delta}{D}} \left(\frac{D}{B}\delta^2 + 2 \left(\frac{D}{B} \right)^2 \delta + 2 \left(\frac{D}{B} \right)^3 \right) - e^{-\frac{BR}{D}} \left(\frac{D}{B}R^2 + 2 \left(\frac{D}{B} \right)^2 R + 2 \left(\frac{D}{B} \right)^3 \right) \right) k + e^{-\frac{B\delta}{D}}}, \quad (3.55)$$

and

$$\tau_n = \frac{\alpha \left(e^{-\frac{B\delta}{D}} \left(\frac{D}{B}\delta^2 + 2 \left(\frac{D}{B} \right)^2 \delta + 2 \left(\frac{D}{B} \right)^3 \right) - e^{-\frac{BR}{D}} \left(\frac{D}{B}R^2 + 2 \left(\frac{D}{B} \right)^2 R + 2 \left(\frac{D}{B} \right)^3 \right) \right)}{\alpha \left(e^{-\frac{B\delta}{D}} \left(\frac{D}{B}\delta^2 + 2 \left(\frac{D}{B} \right)^2 \delta + 2 \left(\frac{D}{B} \right)^3 \right) - e^{-\frac{BR}{D}} \left(\frac{D}{B}R^2 + 2 \left(\frac{D}{B} \right)^2 R + 2 \left(\frac{D}{B} \right)^3 \right) \right) k + e^{-\frac{B\delta}{D}}}. \quad (3.56)$$

In Fig. 3.2, we test these theoretical asymptotics as well as (3.7) against Brownian simulations for an increasing number of holes. The ratio $\sigma = \frac{n\pi\epsilon^2}{4\pi\delta^2}$ of the nucleus surface covered by the absorbing windows is constant $\sigma = 2\%$ (surface covered by 2,000 pores of 25nm diameter on the nucleus of a chinese hamster ovary cell [70]). Numerical parameters are summarized in table 3.1.

The nice agreement between the Brownian simulations and the new asymptotics (3.55) and (3.56) of P_n and τ_n respectively is the central result of this article. We point out that the additive term that accounts for the interactions between the windows is crucial : for 100 windows, the new asymptotic for the conditioned MFPT $\tau_n \approx 2min.$, which is very close to simulations, is twice as large as the one derived in [26] $\tau_n \approx 1min.$ In addition, $\frac{|\Sigma|}{|\Omega|} = \left(\frac{\delta}{R}\right)^3$ is not that small in our simulations, which confirms a large range of validity for the computed asymptotics. Finally, even if the surface covered by the n absorbing windows is unchanged, the conditioned MFPT to a window significantly decreases with n : for $n = 100$ windows, the MFPT $\tau_n \approx 2min.$ is divided by 2 compared to the single window case $\tau_n \approx 4min.$

4 Conclusion

Intermittent dynamics with alternative periods of free diffusion and directed motion along MTs characterizes many cellular transports. When the intermittent

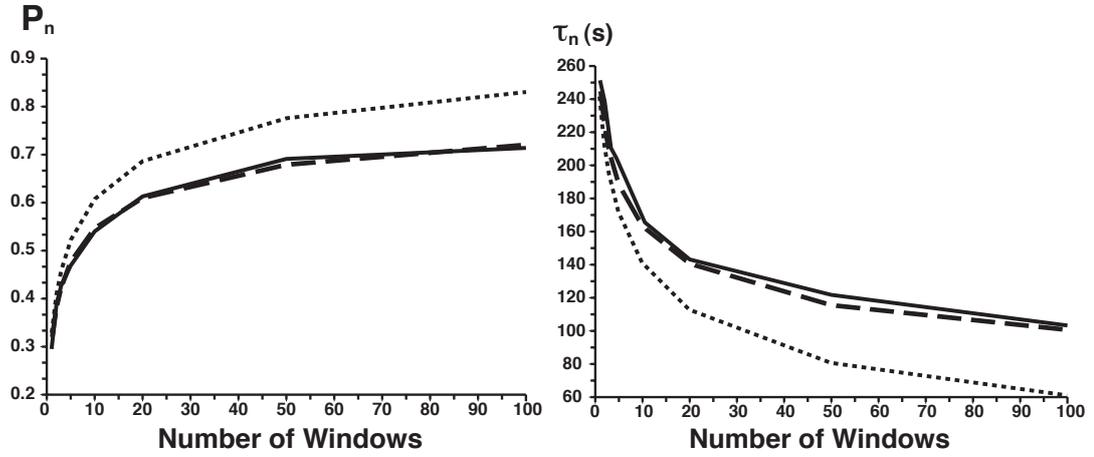


FIGURE 3.2 – New theoretical asymptotics (3.55) and (3.56) (dashed line) for the probability P_n (left) and the conditioned MFPT τ_n (right) are compared against Brownian simulations (solid line) for an increasing number of absorbing windows (the ratio $\sigma = \frac{n\pi\epsilon^2}{4\pi\delta^2}$ of the nuclear surface covered by nanopores is constant $\sigma = 2\%$ [70]). Asymptotics (3.7) that do not account for the interactions between the windows are also drawn (dotted line). 1000 random trajectories are simulated. The parameters are summarized in table 3.1.

TABLE 3.1 – Numerical parameters used for Brownian simulations

Parameters	Description	Value
D	Diffusion constant of the virus	$D = 1.3\mu m^2 s^{-1}$ (as observed for the Associated-Adeno-Virus [14])
B	Drift	$B = 0.2\mu m s^{-1}$ (see chapter 4)
σ	Surface covered by the n nuclear pores	$\sigma = 2\%$ [70]
k	Degradation rate	$k = 1/360s^{-1}$ (10 times the rate observed for synthetic gene vectors [32])
R	Radius of the cell	$R = 15\mu m$ (for a chinese hamster ovary cell [82])
δ	Radius of the nucleus	$\delta = 5\mu m$ [70]

particle can be degraded through the ubiquitin-proteasome machinery or trapped by the crowded cytoplasm, we derived here new asymptotics for the probability P_n and the mean time τ_n the particle reaches a small partially absorbing target among n . These new asymptotics account for the geometrical interactions between the windows. When the targets colocalize on a small structure Σ , asymptotics of P_n and τ_n are obtained in the limit $\frac{|\Sigma|}{|\Omega|} \ll 1$. In particular these formulas apply for DNA viruses or synthetic gene vectors that have to reach a small nuclear pore among the 2,000 that are distributed on the nucleus to deliver their DNA. These theoretical results are tested against Brownian simulations and we observe a very nice agreement between curves. In a future work, it would be very interesting to explore deeper the interactions between the small holes and get the dependency of both P_n and τ_n to the coverage σ intuited [77] or observed with simulations [78]. Quantifying viral movement in the cell cytoplasm would be very helpful for understanding the key limiting steps of infection to design optimal drugs and viral gene vectors [72].

5 Appendix

In that appendix we show that in the limit $\frac{|\Sigma|}{|\Omega|} \ll 1$, for $\mathbf{x}_0 \in \partial\Sigma$, $\frac{\int_{\partial\Sigma} \mathcal{N}(\mathbf{x}, \mathbf{x}_0) d\mathbf{x}}{|\partial\Sigma|}$ tends to $\tau = \frac{\int_{\Omega} u(\mathbf{x}) d\mathbf{x}}{|\Omega|}$, where $u(\mathbf{x})$ is the MFPT to $\partial\Sigma$ of a diffusing particle (diffusion constant D) that starts from a position \mathbf{x} in Ω . We consider $\partial\Sigma$ is purely absorbing and the mean first passage time $u(\mathbf{x})$ to $\partial\Sigma$ of a pure diffusing particle (diffusion constant D) that starts at a position \mathbf{x} in Ω is solution of the Dynkin system [21]

$$D\Delta u(\mathbf{x}) = -1 \text{ for } \mathbf{x} \in \Omega, \quad (3.57)$$

with the boundary conditions

$$\begin{aligned} u &= 0 \text{ on } \partial\Sigma \\ D \frac{\partial u}{\partial n} &= 0 \text{ on } \partial\Omega. \end{aligned} \quad (3.58)$$

We compute

$$\int_{\Omega} D\Delta u(\mathbf{x}) \mathcal{N}(\mathbf{x}, \mathbf{x}_0) d\mathbf{x} - D\Delta \mathcal{N}(\mathbf{x}, \mathbf{x}_0) u(\mathbf{x}) d\mathbf{x}, \quad (3.59)$$

where $\mathcal{N}(\mathbf{x}, \mathbf{x}_0)$ is the normalized Neumann function solution of (3.13). Green identity then yields

$$- \int_{\Omega} \mathcal{N}(\mathbf{x}, \mathbf{x}_0) d\mathbf{x} + u(\mathbf{x}_0) = \int_{\partial\Sigma} D \frac{\partial u}{\partial n}(\mathbf{x}) \mathcal{N}(\mathbf{x}, \mathbf{x}_0) d\mathbf{x} + \frac{\int_{\partial\Omega} u(\mathbf{x}) d\mathbf{x}}{|\partial\Omega|}. \quad (3.60)$$

In the limit $\frac{|\Sigma|}{|\Omega|} \ll 1$, for \mathbf{x} in $\partial\Sigma$, we approximate $D \frac{\partial u}{\partial n}(\mathbf{x}) \approx h_0 = \frac{\int_{\partial\Sigma} D \frac{\partial u}{\partial n}(\mathbf{x}) d\mathbf{x}}{|\partial\Sigma|}$ and integrating (3.57) over Ω , we obtain

$$\int_{\partial\Sigma} D \frac{\partial u}{\partial n}(\mathbf{x}) d\mathbf{x} = -|\Omega|, \quad (3.61)$$

that is $h_0 = -\frac{|\Omega|}{|\partial\Sigma|}$. We introduce the long-time asymptotics $u(\mathbf{x}) \approx \tau = \frac{\int_{\Omega} u(\mathbf{x})d\mathbf{x}}{|\Omega|}$ in (3.60) and we choose \mathbf{x}_0 in $\partial\Sigma$:

$$-\int_{\Omega} \mathcal{N}(\mathbf{x}, \mathbf{x}_0)d\mathbf{x} = -\frac{|\Omega|}{|\partial\Sigma|} \int_{\partial\Sigma} \mathcal{N}(\mathbf{x}, \mathbf{x}_0)d\mathbf{x} + \tau. \quad (3.62)$$

Because $\int_{\Omega} \mathcal{N}(\mathbf{x}, \mathbf{x}_0)d\mathbf{x} = 0(1)$ [26], we finally have

$$\tau = \frac{|\Omega|}{|\partial\Sigma|} \int_{\partial\Sigma} \mathcal{N}(\mathbf{x}, \mathbf{x}_0)d\mathbf{x}. \quad (3.63)$$

Chapitre 4

Quantitative analysis of virus and plasmid trafficking in cells, *Phys. Rev. E* 79 (2009)

1 Introduction

The study of the motion of many particles inside a biological cell is a problem with many degrees of freedom and a large parameter space. The latter may include the different diffusion constants of the different species, velocities along microtubules, their number, the geometry of cell and nucleus, the number and sizes of nuclear pores, the various degradation factors, and so on. The experimental and numerical exploration of this multi-dimensional parameter space is limited perforce to a small part thereof, due to the great complexity of the biological cell. A great reduction in complexity is often achieved by coarse-graining the complex motion by means of effective equations and their explicit analytical solutions, which is the approach we adopt here. We are specifically concerned with finding a concise description of virus and plasmid trafficking in cell cytoplasm.

Recent studies of natural viruses [64, 15, 14] and synthetic (amphiphiles) DNA carriers [72] uncover details of the cellular pathways and the complexity of cellular infection. Viruses invade mammalian cells through multistep processes, which begin with the uptake of particles in endosomal compartments. After escape, the particle move inside the cytoplasm, and the journey ends at a nuclear pore where its DNA is imported. We focus here only on the free cytoplasmic trafficking, a step that both natural and synthetic DNA carriers share. Cytoplasmic trafficking remains a major obstacle to gene delivery, because the cytosolic motion of large DNA molecules is limited by physical and chemical barriers of the crowded cytoplasm [3, 4]. Whereas molecules smaller than 500kDa can diffuse, larger cargos such as viruses or non-viral DNA particles, require an active transport system [5]. Viral infection is much more efficient than gene transfer using polymers- or lipids-based vectors, where a large amount of endocytosed DNA (typically over 100.000 copies of the gene) is required to produce a cellular response, while only a few copies seem to be necessary in the case of viruses.

A recent study [65] showed that microtubules shape the distribution of molecular motors and vesicle trafficking inside the cell cytoplasm by means of a combination

of experiments and numerical simulations. The distribution of viral species was analyzed in [48, 83] by means of the mass-action law and Brownian simulations, but not at a single particle level. In addition, the problem of a viral particle reaching a small nuclear pore was not considered there and this question is central here. In general, the mechanism of a single DNA and viral delivery to a small nuclear pore in the cytoplasm is still an open question. The mean time for a random particle to arrive to a small target has been studied in [84] and in the context biophysical questions and cell biology in [85, 27, 30]. We propose here a coarse-grained reduced description of viral trafficking in the cytoplasm and compare it to plasmid diffusion. Specifically, we are interested in the probability p_N and the mean time τ_N for a DNA carrier or a virus to arrive to a small nuclear pore. The evaluation of these quantities calls for a quantitative approach to the description of particle trajectories at an individual level and also, to quantify the role of the cell organization.

The paper is organized as follow : we start with the observations that a viral movement can be described as a combination of intermittent switches between pure Brownian diffusion and active transport along microtubules [24] (figure 4.1), while DNA motion can be characterized as pure Brownian. We also account for multiple factors involved in degradation, such as hydrolyzation, destruction by proteasomes, or any other factors that prevent irreversibly the particle from reaching a nuclear pore such as entanglements in the cytoskeleton that definitively trap plasmids. This degradation process is modeled as killing with a time-independent rate $k(\mathbf{x})$. We use the overdamped Langevin dynamics with a degradation rate to describe the viral and DNA motion. We first recall the Fokker-Planck-type equations [21, 68, 26] and run Brownian simulations to compare with the asymptotic approximations of p_N and τ_N derived analytically [21, 26] and use these results to estimate the range of validity of our analytical formula. We further compare our numerical simulations and the new analytical formula for the distribution of killed viral particles. The second part of the paper is dedicated to study for many independent viral or DNA particles, the mean time for the first particle to reach a nuclear pore. This mean time is much faster than the time for a single particle to reach a nuclear pore and we obtain here an analytical expression which we then compare to Brownian simulations. In the last part, using a new asymptotic analysis, we obtain novel estimates for p_N and τ_N in the large k limit.

The present approach is a first attempt to develop a theoretical tool for the analysis of virus and DNA particle dynamics at the single molecule level and, hopefully, for the study of trafficking of synthetic vectors, a necessary step toward gene delivery.

2 Modeling intracellular viral and DNA trafficking

Modeling DNA carriers trajectories. We model viral trajectories as a collection of pieces, each of which is characterized either as directed movement along microtubules or pure Brownian motion [64, 15, 14]. In contrast, DNA motion in the cytoplasm can be adequately described as pure Brownian motion [4]. Particles moving inside the cell are reflected at impermeable surfaces and are absorbed at nuclear pores. A virus travels on microtubules as long as it binds to a motor. The three- or two-dimensional position of a particle, $\mathbf{X}(t)$, is described by the coarse-grained

stochastic dynamics

$$\dot{\mathbf{X}} = \begin{cases} \sqrt{2D}\dot{\mathbf{w}} & \text{for a free particle} \\ \mathbf{V}(t) & \text{for a bound particle} \end{cases}, \quad (4.1)$$

where \mathbf{w} is a δ -correlated standard white noise and $\mathbf{V}(t)$ is a time-dependent velocity along a microtubule. The velocity $\mathbf{V}(t)$ can be either positive or negative, depending on whether a viral particle binds to a dynein or to a kinesin motor. However, it is not clear what regulatory mechanisms is involved in such a choice [86].

Mathematical description of a viral trajectory in the cytoplasm. We consider the trafficking of a viral particle from an endosome or the cell membrane to a small nuclear pore. The cell cytosol is a bounded spatial domain Ω , whose boundary $\partial\Omega$ is the external membrane $\partial\Omega_{ext}$ and the nuclear envelope (figure 4.1). Most of the nuclear membrane consists of a reflecting boundary ∂N_r , except for small nuclear pores ∂N_a , where a viral particle can enter the nucleus. We assume that a viral particle that reaches a pore is instantly absorbed, so that this boundary is purely absorbing for trajectories. The ratio of the surface areas is assumed small,

$$\varepsilon = \frac{|\partial N_a|}{|\partial\Omega|} \ll 1. \quad (4.2)$$

Homogenization of viral trajectory. To replace the intermittent dynamics between free diffusion and the drift motion along microtubules, described in equation (4.1), we use a calibration procedure described in chapters 1 and 2. In this homogenization procedure, the motion is described by the overdamped limit of the Langevin equation

$$d\mathbf{X} = \mathbf{b}(\mathbf{X}) dt + \sqrt{2D} d\mathbf{W}, \quad (4.3)$$

where D is the diffusion constant and $\mathbf{b}(\mathbf{X})$ represents the steady state drift that account for the microtubules density, the forward and backward binding rate and the velocity along the microtubules (see chapters 1 and 2). Because most of the microtubules starting from the cell surface converge to the centrosome, a specialized organelle located nearby the cell nucleus (figure 4.1), we choose in a first approximation a radially symmetric effective drift $\mathbf{b}(\mathbf{X})$ converging to the nucleus surface. We thus neglected the minor contribution of microtubules that are not oriented along the radial direction. This radial geometry approximation is actually common in biophysical modelings of *in vitro* experiments [65, 25]. Thus, although viruses move bidirectionally on microtubules, the overall movement is directed toward the nucleus, and we only consider here this average component [86]. The drift component (4.3) can be written as

$$\mathbf{b}(\mathbf{X}) = -b(r) \frac{\mathbf{X}}{|\mathbf{X}|}, \quad (4.4)$$

with $r = |\mathbf{X}|$ is the radial distance to the cell center. In first approximation, we approximate $b(r)$ as a constant $b(r) = B$, which depends on many parameters, such as the density of microtubules, the binding and unbinding rates and the averaged velocity of the directed motion along microtubules (see chapters 1 and 2). Because

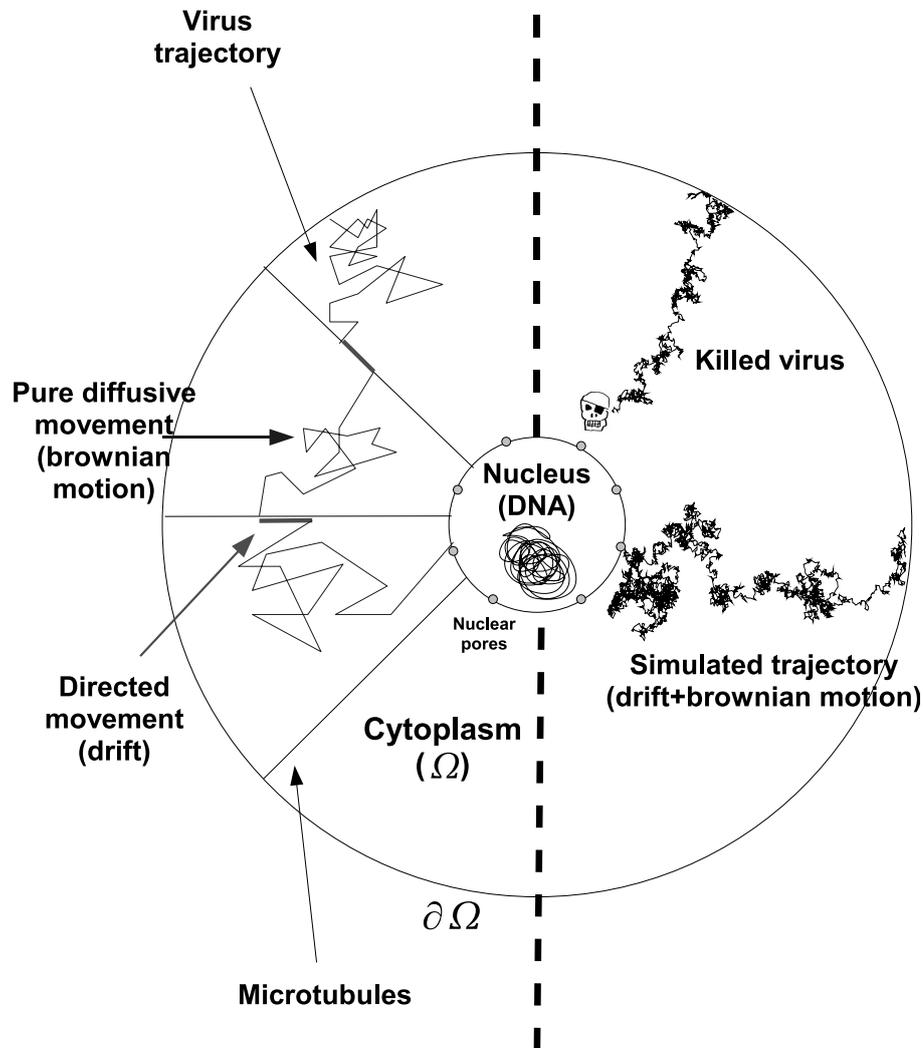


FIGURE 4.1 – Schematic representation of the viral trajectory approximation : on the left-side of the idealized cell, a real trajectory consists of intermittent Brownian and drift epochs, whereas on the right-side, we show two simulated trajectories obtained by equation (4.3). In one of them, the viral particle arrives alive to a nuclear pore, while in the other, it is killed inside the cytoplasm. The round dots on the nucleus surface represent nuclear pores.

the microtubule density increases near the nucleus, a radial dependent drift is more accurate (see chapter 2), but we already show in the first chapter that the constant approximation is good enough and it leads to more concise analytical expressions.

From trajectory description to the probability and mean arrival time. Viral killing or immobilization and naked DNA degradation by nucleases, are coarse-grained into a steady state degradation or killing rate $k(\mathbf{X})$. We briefly recall (see [26] for the details) how to derive the asymptotic expressions for the probability P_N , that a DNA carrier (single virus or DNA) arrives to a small nuclear pore alive and for the mean time τ_N , using approximation (4.2). The asymptotic estimates depend on the diffusion constant D , the amplitude of the drift B , and k . These computations are based on the small hole theory [27], which describes a Brownian particle confined to a bounded domain by a reflecting boundary, except for a small absorbing window, through which it escapes. The domain Ω contains a spherical nucleus of radius δ (a disk in the two-dimensional case). The survival probability density function (SPDF) $p(\mathbf{x}, t)$ to find the virus or naked DNA alive inside the volume element $\mathbf{x} + d\mathbf{x}$ at time t is given by [26]

$$p(\mathbf{x}, t)d\mathbf{x} = Pr\{X(t) \in \mathbf{x} + d\mathbf{x}, \tau^k > t, \tau^a > t | p_i\}, \quad (4.5)$$

where τ^a is the first passage time of a live DNA carrier to the absorbing boundary ∂N_a , τ^k is the first time it is hydrolyzed or immobilized, and p_i is the initial distribution. The SPDF $p(\mathbf{x}, t)$ of the motion (4.3) is the solution of the mixed initial boundary value problem for the Fokker-Planck equation (FPE) [21]

$$\begin{aligned} \frac{\partial p}{\partial t}(\mathbf{x}, t) &= D\Delta p(\mathbf{x}, t) - \nabla \cdot \mathbf{b}(\mathbf{x})p(\mathbf{x}, t) - kp(\mathbf{x}, t) \\ p(\mathbf{x}, 0) &= p_i(\mathbf{x}) \quad \text{for } \mathbf{x} \in \Omega \end{aligned} \quad (4.6)$$

with the boundary conditions

$$\begin{aligned} p(\mathbf{x}, t) &= 0 \quad \text{for } \mathbf{x} \in \partial N_a \\ \mathbf{J}(\mathbf{x}, t) \cdot \mathbf{n}_x &= 0 \quad \mathbf{x} \in \partial N_r \cup \partial \Omega_{ext}, \end{aligned} \quad (4.7)$$

where \mathbf{n}_x is the unit outer normal at a boundary point \mathbf{x} . The flux density vector $\mathbf{J}(\mathbf{x}, t)$ is defined as

$$\mathbf{J}(\mathbf{x}, t) = -D\nabla p(\mathbf{x}, t) + \mathbf{b}(\mathbf{x})p(\mathbf{x}, t). \quad (4.8)$$

The survival probability P_N that a live DNA carrier arrives at the nucleus is $P_N = Pr\{\tau^a < \tau^k\}$ [68]. This probability can be expressed in terms of the SPDF [68] by

$$P_N = 1 - Pr\{\tau^a > \tau^k\} = 1 - \int_{\Omega} k(\mathbf{x})\tilde{p}(\mathbf{x}) d\mathbf{x}, \quad (4.9)$$

where $\tilde{p}(\mathbf{x}) = \int_0^{\infty} p(\mathbf{x}, t) dt$ is the solution of equation

$$D\Delta\tilde{p}(\mathbf{x}) - \nabla \cdot \mathbf{b}(\mathbf{x})\tilde{p}(\mathbf{x}) - k(\mathbf{x})\tilde{p}(\mathbf{x}) = -p_i(\mathbf{x}) \quad \text{for } \mathbf{x} \in \Omega$$

with the boundary conditions (4.7). Using the pdf of the time to absorption, conditioned on the event that the DNA carrier escapes alive $Pr\{\tau^a < t | \tau^a < \tau^k\}$, we define the conditional mean time to absorption as

$$\tau_N = E[\tau^a | \tau^a < \tau^k] = \int_0^{\infty} (1 - Pr\{\tau^a < t | \tau^a < \tau^k\}) dt.$$

Following the computations of [26], we get

$$\tau_N = \frac{\int_{\Omega} \tilde{p}(\mathbf{x}) d\mathbf{x} - \int_{\Omega} k(\mathbf{x}) q(\mathbf{x}) d\mathbf{x}}{1 - \int_{\Omega} k(\mathbf{x}) \tilde{p}(\mathbf{x}) d\mathbf{x}}, \quad (4.10)$$

where

$$q(\mathbf{x}) = \int_0^{\infty} sp(\mathbf{x}, s) ds \quad (4.11)$$

satisfies [26]

$$-\tilde{p} = D\Delta q(\mathbf{x}) - [\nabla \cdot \mathbf{b}q] - kq \quad \text{for } \mathbf{x} \in \Omega \quad (4.12)$$

with boundary conditions (4.7).

Comparison of the Brownian simulations with the asymptotic analytical formula : the plasmid case. The two extreme cases where the previous equations can be developed into analytical formula are a high and small degradation rate compared to the cytoplasm exploring rate defined as $\frac{D}{|\Omega|}$, with $|\Omega|$ the volume of cell cytoplasm. For small k , we obtained in [26] explicit expressions for P_N and τ_N , for a nucleus containing n well separated small holes (nuclear pores) on its surface. In a three dimensional cell, the asymptotic analysis for naked DNA ($\mathbf{b} = \mathbf{0}$) leads to

$$P_N = \frac{1}{1 + \frac{|\Omega|\tilde{k}}{4nD\eta}} \quad \text{and} \quad \tau_N = \frac{\left(\frac{|\Omega|}{4D\eta n}\right)}{1 + \left(\frac{|\Omega|\tilde{k}}{4nD\eta}\right)}, \quad (4.13)$$

where $\tilde{k} = \frac{1}{|\Omega|} \int_{\Omega} k(\mathbf{x}) d\mathbf{x}$ and η is the radius of a small absorbing disk (a nuclear pore). Formula (4.13) does not depend on the specific shape of the degradation rate k , but rather on its integral.

We compare here this asymptotic formula with pure Brownian simulations (no drift), as schemed in the right side of figure 4.1, for eq. (4.3) with the parameters $R = 20\mu m$; $\delta = R/5$; $\eta = \delta\pi/12 = 1.05\mu m$; $k = 1/3600s^{-1}$ [32]; $D = 0.02\mu m^2s^{-1}$ [4]; $n = 1$, (a single big hole), which corresponds to a cell with 2% of the nuclear surface occupied by a large nuclear pore (the $n = 2000$ pores of radius $25nm$ [70] observed experimentally occupy exactly 2% of the nuclear membrane). Numerical simulations using an effective big hole actually leads to an over estimation of the mean time compared with many holes : Formulas (4.13) are only valid for few well-separated holes and the third chapter reveals the true formula for the narrow escape time with many holes. Finally, the small diffusion constant $D = 0.02\mu m^2s^{-1}$ accounts for electrostatics binding and entanglements that slow down processing of plasmids. The results are summarized in the table below, where we observe a nice agreement between the analytical formula and our Brownian simulations.

Time and Probability	τ_N	P_N
Theoretical values	3567s	0.90%
Simulated values (2000 particles.)	3564s	0.97%

Comparison of the Brownian simulations with the asymptotic analytical formula : the virus case.

For a virus trajectory governed by equation (4.3) with a potential drift $\mathbf{b} = -\nabla\Phi$, when the degradation rate is small compare to the diffusion rate, the leading order term of the probability and the mean time are given by [26]

$$P_N = \frac{\frac{1}{|\partial\Omega|} \int_{\partial\Omega} e^{-\frac{\Phi(\mathbf{x})}{D}} dS_{\mathbf{x}}}{\frac{1}{|\partial\Omega|} \int_{\partial\Omega} e^{-\frac{\Phi(\mathbf{x})}{D}} dS_{\mathbf{x}} + \frac{1}{4nD\eta} \int_{\Omega} e^{-\frac{\Phi(\mathbf{x})}{D}} k(\mathbf{x}) d\mathbf{x}} \quad (4.14)$$

$$\tau_N = \frac{\frac{1}{4nD\eta} \int_{\Omega} e^{-\frac{\Phi(\mathbf{x})}{D}} d\mathbf{x}}{\frac{1}{|\partial\Omega|} \int_{\partial\Omega} e^{-\frac{\Phi(\mathbf{x})}{D}} dS_{\mathbf{x}} + \frac{1}{4nD\eta} \int_{\Omega} e^{-\frac{\Phi(\mathbf{x})}{D}} k(\mathbf{x}) d\mathbf{x}} \quad (4.15)$$

where $S_{\mathbf{x}}$ is the surface element corresponding to the boundary position \mathbf{x} . For a scalar drift B and degradation rate k that are both constant, in an idealized spherical cell (radius R), a direct estimation of (4.14) and (4.15) gives

$$P_N = \frac{e^{-\frac{B\delta}{D}}}{\frac{\pi}{nD\eta} \left(e^{-\frac{B\delta}{D}} \left(\frac{D}{B} \delta^2 + 2 \left(\frac{D}{B} \right)^2 \delta + 2 \left(\frac{D}{B} \right)^3 \right) - e^{-\frac{BR}{D}} \left(\frac{D}{B} R^2 + 2 \left(\frac{D}{B} \right)^2 R + 2 \left(\frac{D}{B} \right)^3 \right) \right) k + e^{-\frac{B\delta}{D}}}, \quad (4.16)$$

$$\tau_N = \frac{\frac{\pi}{nD\eta} \left(e^{-\frac{B\delta}{D}} \left(\frac{D}{B} \delta^2 + 2 \left(\frac{D}{B} \right)^2 \delta + 2 \left(\frac{D}{B} \right)^3 \right) - e^{-\frac{BR}{D}} \left(\frac{D}{B} R^2 + 2 \left(\frac{D}{B} \right)^2 R + 2 \left(\frac{D}{B} \right)^3 \right) \right)}{\frac{\pi}{nD\eta} \left(e^{-\frac{B\delta}{D}} \left(\frac{D}{B} \delta^2 + 2 \left(\frac{D}{B} \right)^2 \delta + 2 \left(\frac{D}{B} \right)^3 \right) - e^{-\frac{BR}{D}} \left(\frac{D}{B} R^2 + 2 \left(\frac{D}{B} \right)^2 R + 2 \left(\frac{D}{B} \right)^3 \right) \right) k + e^{-\frac{B\delta}{D}}}. \quad (4.17)$$

Contrary to the formula given in [26], to match the Brownian simulations, we have kept in those expressions the dependency in R . In figure 4.2, we compare (4.16) and (4.17) with Brownian simulations for several values of the drift and a constant degradation rate. Further more, equations(4.16) and (4.17) show that the main contribution to the probability and the mean time comes from a boundary layer located near the nucleus surface.

To see the efficiency of formula (4.16) and (4.17), we can now predict the effect of changing the effective drift $B = 0.2$ by $\pm 30\%$. We recall that value of the drift come from the following rational : for a large number of microtubules, the drift B equals the apparent velocity (which is about 10% [87] of the minus end velocity, approximatively equal to $2\mu m/s$ [14]). We found that increasing the drift leads to a probability $P_N^{+30\%} = 0.80$ and a mean time $\tau_N^{+30\%} = 731s$, while reducing the drift gives $P_N^{-30\%} = 0.64$ and $\tau_N^{-30\%} = 1293s$.

We conclude that decreasing the drift amplitude by 30% increases the time by 33% ($\tau_N = 974s$) and decreases the probability by 12% ($P_N = 0.73$), while increasing the drift by 30%, reduces the time by 22% and increases the probability by 10%. These results show the nonlinear effect of the drift. In a biological context, decreasing the drift can be implemented by disrupting the microtubule network. Moreover, using formula (4.13) with the viral parameters given above, we obtain for zero drift ($B = 0$), a mean arrival time equal to $\tau_N = 2262s$. We conclude that the drift due

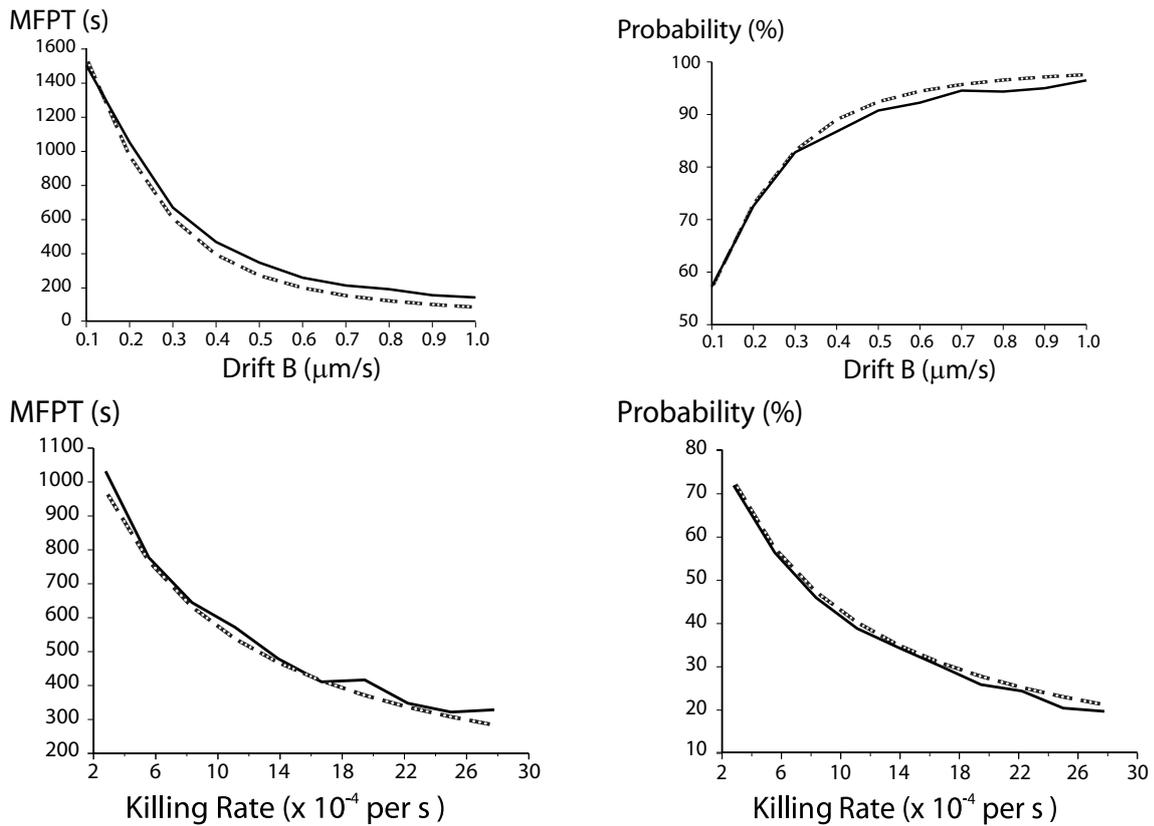


FIGURE 4.2 – MFPT (top left) and the arrival probability (top right) for increasing values of the drift ($k = \frac{1}{3600} s^{-1}$) and for increasing values of the steady state degradation rate ($B = 0.2 \mu\text{m} s^{-1}$) (bottom). 2000 random trajectories are simulated, theoretical and simulated graphs are respectively drawn with dashed and solid lines. The parameters are $R = 20 \mu\text{m}$; $\delta = 4 \mu\text{m}$; $\eta = \frac{\pi}{12} \delta = 1.05 \mu\text{m}$; $D = 1.3 \mu\text{m}^2 s^{-1}$ [14]; $n = 1$.

to active transport along microtubules decreases τ_N of a virus to a nuclear pore by a factor 2.5.

3 Distribution of degraded DNA carriers

Gene delivery using viral vectors, such as AAV, needs the most efficient virus *i.e.* the one reaching the nucleus alive with the highest efficiency. Following endocytosis, viruses can be destroyed either in lysosomes or in the cytoplasm (somewhere between its endosomal release and nuclear pore binding). The distribution of killed viruses can give insights on the cytoplasm degradation activity. For a given steady state degradation rate $k(\mathbf{x})$, the probability $p_k(\mathbf{x})d\mathbf{x}$ that a DNA carrier is degraded in the ball $B(\mathbf{x}, d\mathbf{x})$ of center \mathbf{x} and radius $d\mathbf{x}$ is given by

$$p_k(\mathbf{x})d\mathbf{x} = \tilde{p}(\mathbf{x})k(\mathbf{x})d\mathbf{x}. \quad (4.18)$$

where for sufficiently small nuclear pores and degradation rate, the leading order term of $\tilde{p}(\mathbf{x})$ is given by [26] :

$$\tilde{p}(\mathbf{x}) \approx \frac{e^{-\frac{\Phi(\mathbf{x})}{D}}}{4Dn\eta} \frac{1}{|\partial\Omega|} \int_{\partial\Omega} e^{-\frac{\Phi(\mathbf{x})}{D}} dS_{\mathbf{x}} + \frac{\int_{\Omega} e^{-\frac{\Phi(\mathbf{x})}{D}} k(\mathbf{x})d\mathbf{x}}{4Dn\eta}. \quad (4.19)$$

In a spherical geometry with a constant degradation rate k and a constant radial drift $\mathbf{b}(\mathbf{x}) = -B\frac{\mathbf{r}}{|\mathbf{r}|} \neq 0$ (*i.e.* for a potential $\Phi(r) = Br$), we get :

$$p_k(r) = \frac{ke^{-\frac{Br}{D}}}{4Dn\eta e^{-\frac{B\delta}{D}} + 4\pi k \left(e^{-\frac{B\delta}{D}} \left(\delta^2 \frac{D}{B} + 2\delta \frac{D^2}{B^2} + \frac{D^3}{B^3} \right) - e^{-\frac{BR}{D}} \left(R^2 \frac{D}{B} + 2R \frac{D^2}{B^2} + \frac{D^3}{B^3} \right) \right)}. \quad (4.20)$$

For $n\eta \ll 1$, we obtain

$$p_k(r) = \frac{e^{-\frac{Br}{D}}}{4\pi \left(e^{-\frac{B\delta}{D}} \left(\delta^2 \frac{D}{B} + 2\delta \frac{D^2}{B^2} + \frac{D^3}{B^3} \right) - e^{-\frac{BR}{D}} \left(R^2 \frac{D}{B} + 2R \frac{D^2}{B^2} + \frac{D^3}{B^3} \right) \right)}. \quad (4.21)$$

In FIG.4.3, we compare the theoretical distribution (4.21) with the killed viruses distribution obtained with Brownian simulations (in spherical geometry). The simulations and the analytical formula agree nicely and the maximum of the degradation density probability (equal to $p_k(r)4\pi r^2 dr$) is obtained by a direct computation using formula (4.21). We found that it is achieved for a radius $r = 2D/B = 13\mu m$.

4 Impact of the degradation density distribution

In the plasmid case, because P_N depends only on the integral of $k(\mathbf{x})$ over the cytoplasmic domain, the degradation distribution does not impact the arrival probability. However, because viral particles spend most of their time in the nuclear

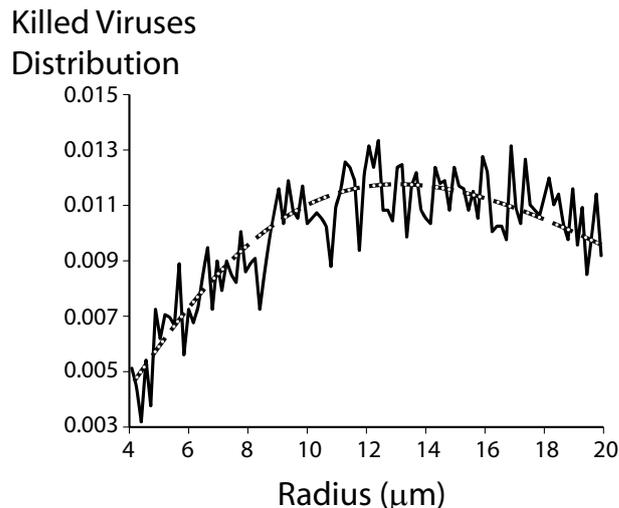


FIGURE 4.3 – Distribution of killed virus $4\pi r^2 p_k(r)$. The simulation is obtained for 30000 Brownian trajectories (solid line). Parameters : $R = 20\mu m$; $\delta = 4\mu m$; $\eta = \frac{\pi}{12}\delta = 1.05\mu m$; $D = 1.3\mu m^2 s^{-1}$; $n = 1$ and $B = 0.2\mu m s^{-1}$.

neighborhood, a large concentration of killing factors such as proteasomes in that area could substantially decrease the arrival probability P_N . To study the impact of the degradation distribution, we compare the virus arrival probability $P_N = 73\%$ obtained with a constant degradation rate k with the one obtained with an exponentially distributed in a nuclear neighborhood. We chose $k(r) = \alpha e^{-\lambda r}$ where

$\alpha = \frac{k|\Omega|}{\int_{\delta}^R e^{-\lambda r} 4\pi r^2 dr}$ is a normalization factor and λ a constant. A direct computation gives $\alpha = \frac{k|\Omega|}{h(\lambda)}$ with $h(\lambda) = 4\pi \left(e^{-\lambda\delta} \left(\frac{\delta^2}{\lambda} + \frac{2\delta}{\lambda^2} + \frac{2}{\lambda^3} \right) - e^{-\lambda R} \left(\frac{R^2}{\lambda} + \frac{2R}{\lambda^2} + \frac{2}{\lambda^3} \right) \right)$.

In that case, we obtain

$$P_N = \frac{e^{-\frac{B\delta}{D}}}{\frac{k|\Omega|}{4nD\eta} \frac{h(\lambda + \frac{B}{D})}{h(\lambda)} + e^{-\frac{B\delta}{D}}}. \quad (4.22)$$

In FIG. 4.4, we plotted P_N as a function of λ : When degradation factors and virions colocalize, which happens for $\lambda \approx \frac{B}{D}$, we obtain that $P_N = 64\%$, which gives a 9% decay compared to the constant killing field case ($P_N = 73\%$). We conclude that the degradation factor distribution does not impact drastically the virions arrival probability.

5 Mean first passage time of the first DNA carrier to a nuclear pore.

Hereafter, we compute the conditioned MFPT $\tau_{first}(M)$ for the first DNA carrier to attain a nuclear pore. The M -DNA carriers trajectories are independent and we shall use the conditioned MFPT τ_N^j of the j^{th} carrier to a nuclear pore. As in [26], we consider the absorbing time $\tau_{first}^a(M)$ of the first DNA carrier to the absorbing

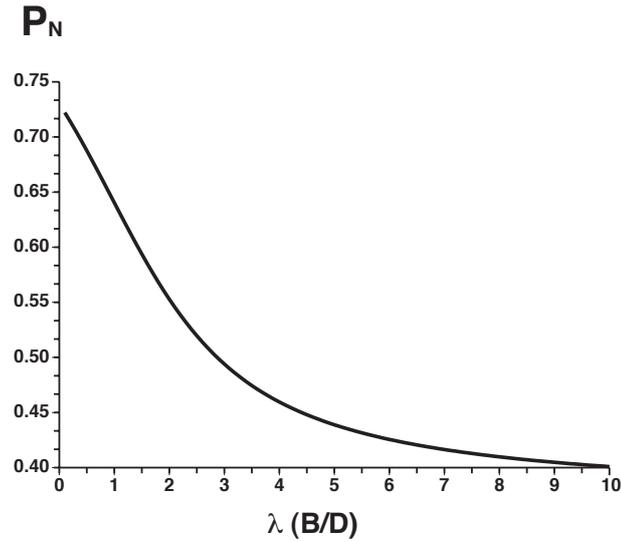


FIGURE 4.4 – The arrival probability P_N is plotted as a function of the characteristic length $\lambda = \frac{B}{D}$. We choose an exponential distribution for the degradation rate $k(r) = \alpha e^{-\lambda r}$, concentrated in a neighborhood of the nucleus, where viruses accumulate.

boundary ∂N_a and the first time $\tau_{first}^k(M)$ it is degraded. The probability the first DNA carrier arrives to the absorbing boundary before time t conditioned on not been killed is then given by :

$$P(t) = Pr\{\tau_{first}^a(M) < t | \tau_{first}^a(M) < \tau_{first}^k(M), p_i\}. \quad (4.23)$$

The conditional MFPT $\tau_{first}(M)$ is defined by

$$\tau_{first}(M) = \int_0^\infty t \frac{dP(t)}{dt} dt = \int_0^\infty (P(\infty) - P(t)) dt. \quad (4.24)$$

To derive an expression for $\tau_{first}(M)$, we shall compute $P(t)$ by using Bayes law :

$$P(t) = \frac{Pr\{\tau_{first}^a(M) < t, \tau_{first}^a(M) < \tau_{first}^k(M), p_i\}}{Pr\{\tau_{first}^a(M) < \tau_{first}^k(M), p_i\}}. \quad (4.25)$$

To estimate the numerator $N(t) = Pr\{\tau_{first}^a(M) < t, \tau_{first}^a(M) < \tau_{first}^k(M), p_i\}$, we use that

$$\begin{aligned} & Pr\{\tau_{first}^a(M) < t, \tau_{first}^a(M) < \tau_{first}^k(M), p_i\} = \\ & 1 - Pr\{\tau_{first}^a(M) > t \text{ or } \tau_{first}^a(M) > \tau_{first}^k(M), p_i\}. \end{aligned} \quad (4.26)$$

The event $\{\tau_{first}^a(M) > t \text{ or } \tau_{first}^a(M) > \tau_{first}^k(M)\}$ means that, at time t , none of the M – DNA carriers have reached alive a small nuclear pore. Since the particles are independent, we obtain

$$\begin{aligned} & Pr\{\tau_{first}^a(M) > t \text{ or } \tau_{first}^a(M) > \tau_{first}^k(M), p_i\} = \\ & \prod_{j=1}^{j=M} (1 - Pr\{\tau_j^a < t, \tau_j^a < \tau_j^k, p_i\}), \end{aligned} \quad (4.27)$$

where τ_j^a (reps. τ_j^k) is the first time the j^{th} particle is absorbed (resp. killed). Using the interpretation of the flux [26], we get that for any of the particles

$$Pr\{\tau_j^a < t, \tau_j^a < \tau_j^k, p_i\} = \int_0^t \oint_{\partial\Omega} \mathbf{J}(\mathbf{x}, t) \cdot \mathbf{n}_x dS_x = \int_0^t J(s) ds, \quad (4.28)$$

where \mathbf{n}_x denotes the normal derivative at the boundary point \mathbf{x} and the flux is defined in (4.8). Finally, we obtain the following expression for the numerator :

$$N(t) = Pr\{\tau_{first}^a(M) < t, \tau_{first}^a(M) < \tau_{first}^k(M), p_i\} = 1 - \left(1 - \int_0^t J(s) ds\right)^M \quad (4.29)$$

Similarly the denominator $D(t)$ of $P(t)$ is given by :

$$D(t) = Pr\{\tau_{first}^a(M) < \tau_{first}^k(M), p_i\} = 1 - Pr\{\tau_{first}^a(M) > \tau_{first}^k(M), p_i\}, \quad (4.30)$$

and because the particles are independent :

$$D(t) = 1 - \prod_{j=1}^{j=M} Pr\{\tau_j^a > \tau_j^k, p_i\}. \quad (4.31)$$

Using the definition of the probability P_N that a particle is killed before reaching the nucleus [26], we get

$$D(t) = 1 - (1 - P_N)^M. \quad (4.32)$$

Finally, the probability density function is given by

$$P(t) = \frac{N(t)}{D(t)} = \frac{1 - \left(1 - \int_0^t J(s) ds\right)^M}{1 - (1 - P_N)^M}. \quad (4.33)$$

and the conditional MFPT $\tau_{first}(M)$ of the first particle is equal to (4.24) :

$$\tau_{first}(M) = \int_0^\infty \frac{\left(1 - \int_0^t J(s) ds\right)^M - \left(1 - \int_0^\infty J(s) ds\right)^M}{1 - (1 - P_N)^M} dt. \quad (4.34)$$

Hereafter, we shall estimate the leading order term for $\tau_{first}(M)$. In the long time asymptotic, we approximate the pdf by its first exponential term : The leading order term of $p(\mathbf{x}, t)$ is given by

$$p(\mathbf{x}, t) \approx p(\mathbf{x}, 0) e^{-\lambda_0 t}, \quad \text{with} \quad \int_{\Omega} p(\mathbf{x}, 0) d\mathbf{x} = 1. \quad (4.35)$$

where $\lambda_0 = 1/\tau_N$ ([21] p.175), is the first eigenvalue (this implies that there is no contribution of the initial condition on the other eigenfunctions, see also [30]). Replacing $p(\mathbf{x}, t)$ by its long time approximation in the equation (4.6), we obtain the following equation for $p(\mathbf{x}, 0)$

$$-\frac{1}{\tau_N} p = D\Delta p - \nabla[bp] - kp. \quad (4.36)$$

Using (4.28), we obtain an explicit expression for the flux $J(t)$ by integrating equation (4.36) over the domain Ω , with $\int_{\Omega} p(\mathbf{x}, 0) d\mathbf{x} = 1$, we obtain

$$J(t) = \frac{e^{-\frac{t}{\tau_N}}}{\tau_N} \left(\int_{\Omega} p(\mathbf{x}, 0) d\mathbf{x} - \tau_N \int_{\Omega} k(\mathbf{x}) p(\mathbf{x}, 0) d\mathbf{x} \right) = \frac{e^{-\frac{t}{\tau_N}}}{\tau_N} \left(1 - \tau_N \int_{\Omega} k(\mathbf{x}) p(\mathbf{x}, 0) d\mathbf{x} \right) \quad (4.37)$$

Using the probability P_N (4.9) and $\tilde{p}(\mathbf{x}) = \int_0^{\infty} p(\mathbf{x}, t) dt = \frac{\tilde{p}(\mathbf{x})}{\tau_N}$, we get an expression for the flux,

$$J(t) = \frac{e^{-\frac{t}{\tau_N}}}{\tau_N} \left(1 - \int_{\Omega} k(\mathbf{x}) \tilde{p}(\mathbf{x}) d\mathbf{x} \right) = \frac{P_N}{\tau_N} e^{-\frac{t}{\tau_N}}. \quad (4.38)$$

Replacing $\int_0^t J(s) ds$ by its approximation (4.38) in relation (4.34) we get :

$$\tau_{first}(M) = \int_0^{\infty} \frac{\left(1 - P_N \left(1 - e^{-\frac{t}{\tau_N}} \right) \right)^M - (1 - P_N)^M}{1 - (1 - P_N)^M} dt. \quad (4.39)$$

With the notation $\xi = 1 - P_N$ ($0 \leq \xi \leq 1$) we have

$$\tau_{first}(M) = \frac{1}{1 - \xi^M} \int_0^{\infty} \left(\left(e^{-\frac{t}{\tau_N}} + \xi \left(1 - e^{-\frac{t}{\tau_N}} \right) \right)^M - \xi^M \right) dt. \quad (4.40)$$

Thus,

$$\begin{aligned} \tau_{first}(M) &= \frac{1}{1 - \xi^M} \sum_{k=0}^{M-1} \binom{M}{k} \xi^k \int_0^{\infty} \left(1 - e^{-\frac{t}{\tau_N}} \right)^k \left(e^{-\frac{t}{\tau_N}} \right)^{M-k} dt \\ &+ \frac{\xi^M}{1 - \xi^M} \int_0^{\infty} \left(\left(1 - e^{-\frac{t}{\tau_N}} \right)^M - 1 \right) dt. \end{aligned} \quad (4.41)$$

An iterative integration by parts yields for $0 \leq k \leq M - 1$:

$$\int_0^{\infty} \left(1 - e^{-\frac{t}{\tau_N}} \right)^k \left(e^{-\frac{t}{\tau_N}} \right)^{M-k} dt = \frac{\tau_N}{(M - k) \binom{M}{k}}. \quad (4.42)$$

Consequently, we have :

$$\tau_{first}(M) = \frac{\tau_N}{1 - \xi^M} \sum_{k=0}^{M-1} \frac{\xi^k}{M - k} + \frac{\xi^M}{1 - \xi^M} \int_0^{\infty} \left(\left(1 - e^{-\frac{t}{\tau_N}} \right)^M - 1 \right) dt.$$

Concerning right-hand side of equation above, polynomial identity : $X^M - 1 = (X - 1) \sum_{k=0}^{M-1} X^k$ leads to :

$$\int_0^{\infty} \left(\left(1 - e^{-\frac{t}{\tau_N}} \right)^M - 1 \right) dt = - \sum_{k=0}^{M-1} \int_0^{\infty} e^{-\frac{t}{\tau_N}} \left(1 - e^{-\frac{t}{\tau_N}} \right)^k dt. \quad (4.43)$$

Replacing M by $k + 1$ in (4.42), we get :

$$\int_0^{\infty} \left(\left(1 - e^{-\frac{t}{\tau_N}} \right)^M - 1 \right) dt = - \sum_{k=0}^{M-1} \frac{\tau_N}{k + 1} = - \sum_{l=0}^{M-1} \frac{\tau_N}{M - l}. \quad (4.44)$$

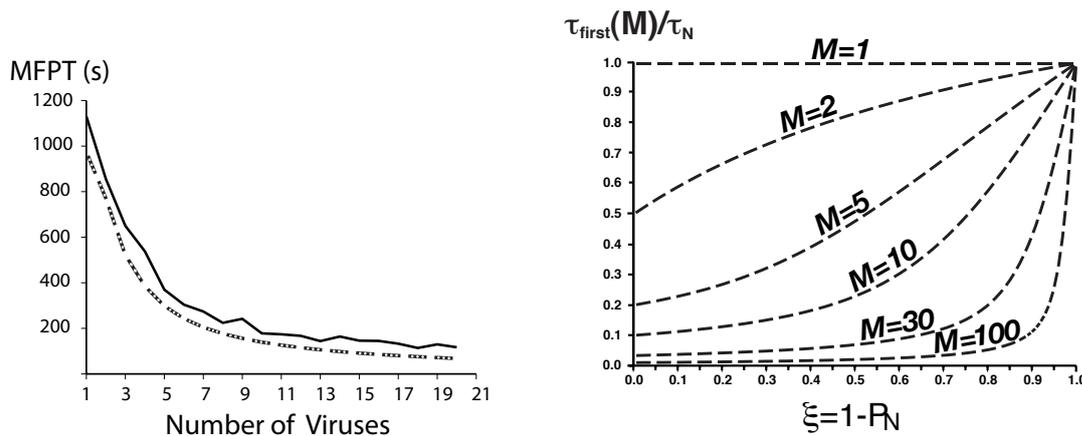


FIGURE 4.5 – **Left** : MFPT of the first virus to a nuclear pore. We generate 300 Brownian trajectories (solid line). The geometry is given by $R = 20\mu m$; $\delta = 4\mu m$; $\eta = \frac{\pi}{12}\delta = 1.05\mu m$; $D = 1.3\mu m^2 s^{-1}$; $n = 1$ and $B = 0.2\mu m s^{-1}$. **Right** : Normalized MFPT of the first virus to the MFPT of a single virus, as a function of the probability $\xi = 1 - P_N$ to be killed before arriving to the nucleus. As ξ tends to 0, τ_{first} tends to $\frac{\tau_N}{M}$ ($\tau_N = 974s$ here); whereas τ_{first} tends to τ_N when almost all DNA carriers are degraded.

Finally we have the concise expression (note that τ_N is a function of ξ) :

$$\tau_{first}(M) = \frac{\tau_N(\xi)}{1 - \xi^M} \left(\sum_{k=0}^{M-1} (\xi^k - \xi^M) \frac{1}{M - k} \right). \quad (4.45)$$

We compare in FIG. 4.5-left the analytical curves with the Brownian simulations. Both curves match very nicely, which confirms the validity of the long time asymptotic approximation.

In FIG. 4.5-right, we plotted $\tau_{first}(M)/\tau_N$ as a function of ξ , which is an increasing function of ξ : when the number of DNA carriers reaching alive a nuclear pore decreases, the MFPT of the first survivor increases. Moreover, the curves confirm that for small ξ , the leading order term of $\tau_{first}(M)$ is

$$\frac{\tau_{first}(M)}{\tau_N(\xi)} \approx \frac{1}{M}, \quad (4.46)$$

whereas when ξ tends to 1, (*i.e.* almost all DNA carriers are killed before reaching nuclear pores) we get the approximation :

$$\tau_{first}(M) \approx \frac{\tau_N}{M(1 - \xi)} \left(\sum_{k=0}^{M-1} (M - k) (1 - \xi) \frac{1}{M - k} \right) = \tau_N(1). \quad (4.47)$$

It would be interesting to find the general expression for $\tau_{first}(M)$ as a function of ξ .

6 The large degradation rate limit

Because the previous analysis [26] does not give any range of validity of the asymptotic formula for the probability and the mean time to reach a nuclear pore,

we decided to investigate more carefully the case where the degradation rate is large $k \gg 1$. We computed in [26] P_N and τ_N in the limit of a small degradation rate limit $k(\mathbf{x}) \ll 1$, however in the plasmid case, the killing activity due to the protease could be much larger than the diffusion time scale. Thus, we derive hereafter new asymptotics in the large degradation rate limit. The analysis is quite different from [26]. We start with a constant degradation rate $k(\mathbf{x}) = k$ (the computations for a general radial degradation rate are given in the appendix). We consider a uniform initial plasmid distribution over the cytoplasm $p_i(\mathbf{x}) = p_0 = \frac{1}{|\Omega|}$. To compute the probability P_N , we shall solve equation

$$D\Delta\tilde{p}(\mathbf{x}) - k(\mathbf{x})\tilde{p}(\mathbf{x}) = -p_0 = -\frac{1}{|\Omega|}, \quad (4.48)$$

with the boundary conditions (4.7). When $\frac{D}{|\Omega|}$ is much smaller compared to k and for a particle starting far from nuclear pores, we approximated the solution of Eq.(4.48) by

$$p_{outer}(\mathbf{x}) = \frac{1}{k|\Omega|} + O(D). \quad (4.49)$$

However, this outer solution does not match the absorbing conditions. We now construct an inner solution $p_{inner}(\mathbf{x})$ near the nuclear pores that will satisfy the absorbing conditions and match the outer solution. In a local coordinates (ρ, s) near ∂N_a , where ρ measures distance from ∂N_a , measured positively into Ω , and s are tangential variables in the plane $\rho = 0$ (see for example [88] and figure 4.6 where the local coordinate system is represented in a two dimensional geometry with a single nuclear pore). Projecting equation (4.48) on the ρ -coordinate (the variations

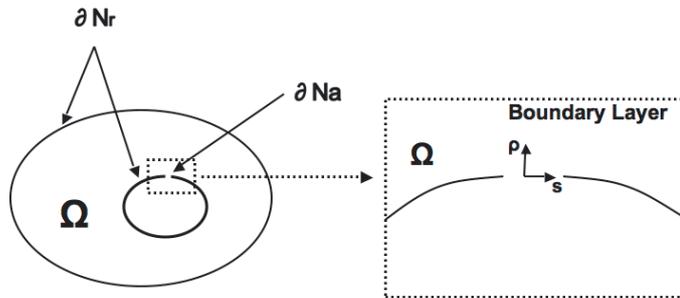


FIGURE 4.6 – Schematic representation of the boundary layer in a local coordinate system (ρ, s) near the boundary ∂N_a , where ρ is the distance from ∂N_a , measured positively and s is the arc length.

of \tilde{p} with respect to s are small compared to the variation in ρ), we obtain for the leading order term p_{inner} :

$$\frac{d^2 p_{inner}(\rho)}{d\rho^2} - \frac{k}{D} p_{inner}(\rho) = -\frac{1}{D|\Omega|}, \quad (4.50)$$

satisfying the absorbing condition on the nuclear pore

$$p_{inner}(0) = 0. \quad (4.51)$$

Far from the boundary layer [88], the matching condition is

$$\lim_{\frac{\rho}{\sqrt{D}} \rightarrow \infty} p_{inner}(\rho) = p_{outer} = \frac{1}{|\Omega|k}. \quad (4.52)$$

Consequently, near the boundary we get

$$p_{inner}(\rho, s) = \frac{1}{|\Omega|k} \left(1 - e^{-\sqrt{\frac{k}{D}}\rho} \right). \quad (4.53)$$

To compute P_N , we use formula (4.9)

$$P_N = 1 - \int_{\Omega} k(\mathbf{x})\tilde{p}(\mathbf{x})d\mathbf{x}, \quad (4.54)$$

which can be rewritten as

$$P_N = 1 - \left(\int_{\Omega \setminus BL} kp_{outer}d\mathbf{x} + \int_{BL} kp_{inner}(\rho)d\rho \right), \quad (4.55)$$

where BL is the boundary layer. Using expression (4.49) for p_{outer} , we get

$$\int_{\Omega \setminus BL} kp_{outer}(\mathbf{x})d\mathbf{x} = \frac{|\Omega \setminus BL|}{|\Omega|} \quad (4.56)$$

and finally

$$\begin{aligned} \int_{BL} kp_{inner}(\rho)d\rho &= \frac{1}{|\Omega|} \left(|BL| + |\partial N_a| \int_0^{\rho_0} -e^{-\sqrt{\frac{k}{D}}\rho}d\rho \right) \\ &= \frac{1}{|\Omega|} \left(|BL| - |\partial N_a| \sqrt{\frac{D}{k}} \left(1 - e^{-\sqrt{\frac{k}{D}}\rho_0} \right) \right), \end{aligned} \quad (4.57)$$

where $\rho_0 \gg \sqrt{\frac{D}{k}}$ is the thickness of the boundary layer. Finally,

$$P_N = \frac{|\partial N_a|}{|\Omega|} \sqrt{\frac{D}{k}} + O\left(e^{-\sqrt{\frac{k}{D}}\rho_0}\right). \quad (4.58)$$

In a three-dimensional cell, when the boundary consists of n well separated small holes of radius η , we obtain that

$$P_N = \frac{n\pi\eta^2}{|\Omega|} \sqrt{\frac{D}{k}} + O\left(e^{-\sqrt{\frac{k}{D}}\rho_0}\right). \quad (4.59)$$

Because our analysis is local, it can be extended to any degradation rate, large compared to the exploring rate. In that case, when for n well separated narrow pores of size η_q , $1 \leq q \leq n$, located at position x_1, \dots, x_n , the asymptotic formula is

$$P_N \approx \sum_{q=1}^n \frac{\pi\eta_q^2}{|\Omega|} \sqrt{\frac{D}{k(x_q)}} + O\left(e^{-\sqrt{\frac{k_0}{D}}\rho_0}\right), \quad (4.60)$$

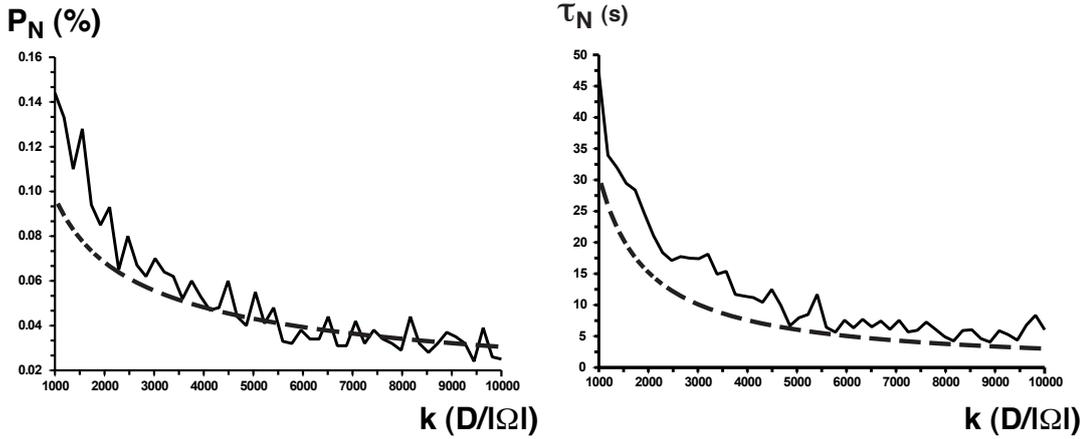


FIGURE 4.7 – The probability and mean time for a plasmid to reach a small nuclear pore plotted as a function of the constant degradation rate for a two dimensional flat cell. The Brownian simulations match the analytic solutions (4.58) and (4.68) only after a rate of $3000 \frac{D}{|\Omega|} = 3000 \frac{D}{\pi(R^2 - \delta^2)} \approx 0.05 s^{-1}$, around 200 higher than the normal rate $1/3600 \approx 2.8 \times 10^{-4} s^{-1}$.

where k_0 is the minimum value of $k(\mathbf{x})$ among the pores. More detailed computations are given in the appendix. From the fitting in figure 4.7 of the Brownian simulations with the analytical formula (4.58), we conclude that the matching occurs for a very large degradation rate (more than 200 times the normal rate [32]) and thus the large case limit might only be useful to characterize gene delivery for abnormal cells, where the degradation rate is large. The MFPT τ_N to a small pore for a live virus is [26] :

$$\tau_N = \frac{\int_{\Omega} \tilde{p}(\mathbf{x}) d\mathbf{x} - \int_{\Omega} k q(\mathbf{x}) d\mathbf{x}}{P_N}. \quad (4.61)$$

where $q(\mathbf{x}) = \int_0^{\infty} s p(\mathbf{x}, s) ds$ satisfies (4.12) with boundary conditions (4.7). To estimate τ_N , we consider for a small diffusion, an outer approximation of q given by

$$q_{outer} = \frac{p_{outer}}{k} = \frac{1}{|\Omega| k^2}. \quad (4.62)$$

The leading term of the inner solution q_{inner} in the boundary layer expansion of q satisfies :

$$\frac{d^2 q_{inner}(\rho)}{d\rho^2} - \frac{k}{D} q_{inner}(\rho) = -\frac{p_{inner}}{D} = -\frac{1}{D|\Omega|k} \left(1 - e^{-\rho\sqrt{\frac{k}{D}}}\right) \quad (4.63)$$

$$q_{inner}(0) = 0 \quad (4.64)$$

$$\lim_{\frac{\rho}{\sqrt{D}} \rightarrow \infty} q_{inner}(\rho) = q_{outer} = \frac{1}{|\Omega| k^2}. \quad (4.65)$$

Consequently, we get :

$$q_{inner} = \frac{1}{|\Omega| k^2} \left(1 - e^{-\sqrt{\frac{k}{D}} \rho}\right) - \frac{\rho}{2\sqrt{D} k^{\frac{3}{2}} |\Omega|} e^{-\sqrt{\frac{k}{D}} \rho}. \quad (4.66)$$

We get then that :

$$\int_{\Omega} \tilde{p}(\mathbf{x}) d\mathbf{x} - \int_{\Omega} k q(\mathbf{x}) d\mathbf{x} = |\partial N_a| \frac{\sqrt{D}}{2|\Omega| k^{\frac{3}{2}}} + O\left(e^{-\sqrt{\frac{k}{D}} \rho_0}\right). \quad (4.67)$$

Finally,

$$\tau_N = |\partial N_a| \frac{\sqrt{D}}{2|\Omega|k^{\frac{3}{2}}P_N} + O\left(e^{-\sqrt{\frac{k}{D}}\rho_0}\right) = \frac{1}{2k} + O\left(e^{-\sqrt{\frac{k}{D}}\rho_0}\right). \quad (4.68)$$

For a large degradation rate, our analytical results match the Brownian simulations (see figure 4.7). Moreover, the present local analysis can be extended to any degradation rate and for n well separated narrow pores, located at position x_1, \dots, x_n . We anticipate the following asymptotic formula,

$$\tau_N \approx \frac{1}{n} \sum_{q=1}^n \frac{1}{2k(x_q)} + O\left(e^{-\sqrt{\frac{k_0}{D}}\rho_0}\right), \quad (4.69)$$

where $k_0 = \min_q k(x_q)$ is the minimum value of $k(\mathbf{x})$ among the pores. k_0 is the minimum concentration of killing factors among nuclear pores.

7 Conclusion

By describing the intermittent dynamics of a DNA carrier inside the cytoplasm with an effective stochastic description (4.3), we derived a quantitative analysis of the nuclear DNA carrying at the single unit level. Modeling the DNA degradation, as protease activity, that occurs in the cell cytoplasm with a steady state degradation rate $k(\mathbf{x})$, we also derived expressions for the probability a DNA carrier hits a small nuclear pore and the mean time it takes (in both cases of small and large degradation rate). We also provided here the distribution of degraded particles.

When many independent viruses are involved, we computed the mean time to a nuclear pore for the first one. We tested our analytical results against Brownian simulations and we obtained that our curves match nicely. Our analysis provides a tool to explore the multi-dimensional parameter space of nuclear DNA carrying. Cytoplasmic trafficking is a limiting step of gene delivery and elucidating viral motion in the cytoplasm may provide a quantitative tool for the improvement and optimization of delivery of synthetic vectors.

8 Appendix

We compute hereafter the probability P_N for a carrier moving by random motion to hit a small nuclear pore for a large (compared to the exploring rate) degradation rate $k(\mathbf{x})$. We use method based on a boundary layer analysis, similar as the one produce in this manuscript for a constant k : far from the nuclear pore, the leading order term of the outer solution is no longer constant and it is given by

$$p_{outer}(\mathbf{x}) = \frac{p_0}{k(\mathbf{x})} + O(D). \quad (4.70)$$

The initial uniform distribution of DNA carriers is $p_0 = \frac{1}{|\Omega|}$. To compute the inner solution near the nuclear surface, we expand the steady state radial killing measure along the radial ρ -coordinate,

$$k(\rho, s) = k_0(s) + k_1(s)\rho + O(\rho^2). \quad (4.71)$$

where $k(\rho = 0, s) = k_0(s)$ and $\frac{dk}{d\rho}(\rho = 0, s) = k_1(s)$. Because p_{outer} does not necessary satisfy the reflecting boundary condition anymore, we construct two inner solutions : the first one p_{inner}^1 near ∂N_a and the second p_{inner}^2 near ∂N_r . Projecting equation (4.48) on the ρ -coordinate (the variations of \tilde{p} with respect to s are small compared to the variation in ρ), we obtain that the leading order terms of $p_{inner}^i(\rho, s)$ for $i = 1, 2$ satisfy :

$$\begin{aligned} \frac{\partial^2 p_{inner}^i}{\partial \rho^2} - \frac{k_0(s) + k_1(s)\rho}{D} p_{inner}^i &= -\frac{1}{|\Omega|D} \text{ for } 0 \leq \rho \leq \rho_0(s) & (4.72) \\ p_{inner}^1(\rho = 0, s) &= 0 \text{ on } \partial N_a \\ \frac{\partial}{\partial \rho} p_{inner}^2(\rho = 0, s) &= 0 \text{ on } \partial N_r \\ \text{for } i = 1, 2 \lim_{\frac{\rho}{\sqrt{D}} \rightarrow \infty} p_{inner}^i(\rho, s) &= p_{outer}(\rho = 0, s) = \frac{1}{|\Omega|k_0(s)}, \end{aligned}$$

where $\rho_0(s) \gg \sqrt{\frac{D}{k_0(s)}}$ is the local thickness of the boundary layer. To solve the homogeneous equation :

$$\frac{\partial^2 p_{inner}^i}{\partial \rho^2} - \frac{k_0(s) + k_1(s)\rho}{D} p_{inner}^i = 0, \quad (4.73)$$

we use the change of variable

$$u = u(\rho, s) = \frac{k_0(s) + k_1(s)\rho}{\beta(s)D}, \text{ where } \beta(s) = \left(\frac{k_1(s)}{D}\right)^{\frac{2}{3}}. \quad (4.74)$$

By this substitution in (4.73), we get

$$\frac{\partial^2 p_{inner}^i}{\partial u^2} - u p_{inner}^i = 0, \quad (4.75)$$

and the solution is

$$p_{inner}^i = C_0^i(s)Ai(u) + C_1^i(s)Bi(u), \quad (4.76)$$

where $C_0^i(s)$ and $C_1^i(s)$ are real functions of s and Ai and Bi are the Airy functions ([31], p. 446). In the small diffusion limit $D \ll 1$, $u \geq \frac{k_0(s)}{\beta(s)D} = \frac{k_0(s)}{(k_1(s))^{\frac{2}{3}} D^{\frac{1}{3}}} \gg 1$.

Because either solutions p_{inner}^i are bounded, but $\lim_{u \rightarrow +\infty} Bi(u) = +\infty$, we get that $C_1^i = 0$ and consequently,

$$p_{inner}^i = C_0^i(s)Ai(u). \quad (4.77)$$

To obtain a particular solution \bar{p}_{inner}^i of equation (4.72), we write is as

$$\frac{\partial^2 \bar{p}_{inner}^i}{\partial u^2} - u \bar{p}_{inner}^i = -\frac{1}{|\Omega|\beta(s)D}. \quad (4.78)$$

Using the Scorer's functions ([31], p.448) and because

$$\lim_{u \rightarrow +\infty} Hi(u) = +\infty, \quad (4.79)$$

we obtain that

$$\bar{p}_{inner}^i = \frac{\pi}{|\Omega|\beta(s)D} Gi(u) \quad (4.80)$$

Collecting the results, we obtain that

$$p_{inner}^i(u, s) = C_0^i(s) Ai(u) + \frac{\pi}{|\Omega|\beta(s)D} Gi(u). \quad (4.81)$$

Using the matching boundary conditions, we get

$$\begin{aligned} p_{inner}^1(\rho = 0, s) &= p_{inner}^1\left(u = \frac{k_0(s)}{\beta(s)D}, s\right) = 0 \\ \frac{\partial p_{inner}^2}{\partial \rho}(\rho = 0, s) &= \frac{k_1(s)}{\beta(s)D} \frac{\partial p_{inner}^2}{\partial u}\left(u = \frac{k_0(s)}{\beta(s)D}, s\right) = 0. \end{aligned}$$

Using expression (4.81), we obtain the equations

$$C_0^1(s) Ai\left(\frac{k_0(s)}{\beta(s)D}\right) + \frac{\pi}{|\Omega|\beta(s)D} Gi\left(\frac{k_0(s)}{\beta(s)D}\right) = 0 \quad (4.82)$$

$$C_0^2(s) Ai'\left(\frac{k_0(s)}{\beta(s)D}\right) + \frac{\pi}{|\Omega|\beta(s)D} Gi'\left(\frac{k_0(s)}{\beta(s)D}\right) = 0 \quad (4.83)$$

and thus

$$C_0^1(s) = -\frac{\pi}{|\Omega|\beta(s)D} \frac{Gi\left(\frac{k_0(s)}{\beta(s)D}\right)}{Ai\left(\frac{k_0(s)}{\beta(s)D}\right)} \quad (4.84)$$

$$C_0^2(s) = -\frac{\pi}{|\Omega|\beta(s)D} \frac{Gi'\left(\frac{k_0(s)}{\beta(s)D}\right)}{Ai'\left(\frac{k_0(s)}{\beta(s)D}\right)}. \quad (4.85)$$

Finally the inner solutions p_{inner}^i are given by

$$p_{inner}^1(u, s) = \frac{\pi}{|\Omega|\beta(s)D} \left(Gi(u) - \frac{Gi\left(\frac{k_0(s)}{\beta(s)D}\right)}{Ai\left(\frac{k_0(s)}{\beta(s)D}\right)} Ai(u) \right) \quad (4.86)$$

$$p_{inner}^2(u, s) = \frac{\pi}{|\Omega|\beta(s)D} \left(Gi(u) - \frac{Gi'\left(\frac{k_0(s)}{\beta(s)D}\right)}{Ai'\left(\frac{k_0(s)}{\beta(s)D}\right)} Ai(u) \right). \quad (4.87)$$

with the outer solution, we use the large u -asymptotic of $Gi(u)$, $Gi'(u)$, $Ai(u)$ and $Ai'(u)$ ([31], p.448-450) :

$$\begin{aligned} Gi(u) &\approx \frac{1}{\pi u} & , & \quad Gi'(u) \approx \frac{7}{96\pi u^2} \\ Ai(u) &\approx \frac{e^{-\frac{2}{3}u^{\frac{3}{2}}}}{2\sqrt{\pi}u^{\frac{1}{4}}} & , & \quad Ai'(u) \approx -\frac{u^{\frac{1}{4}}e^{-\frac{2}{3}u^{\frac{3}{2}}}}{2\sqrt{\pi}}. \end{aligned}$$

For $\rho = \rho_0(s)$ and thus $u = \frac{k_0(s) + k_1(s)\rho_0(s)}{\beta(s)D} \gg 1$, using the asymptotic behavior for A_i and a Taylor expansion at order 1 ($k_1(s)\rho_0(s) \ll k_0(s)$), we get

$$A_i \left(\frac{k_0(s) + k_1(s)\rho_0(s)}{\beta(s)D} \right) \approx A_i \left(\frac{k_0(s)}{\beta(s)D} \right) e^{-\sqrt{\frac{k_0(s)}{D}}\rho_0(s)}. \quad (4.88)$$

Consequently, using expressions (4.86),(4.88) and the asymptotics above, we get

$$p_{inner}^1 \left(\frac{k_0(s) + k_1(s)\rho_0(s)}{\beta(s)D}, s \right) \approx \frac{1}{|\Omega| (k_0(s) + k_1(s)\rho_0(s))} - \frac{1}{|\Omega| k_0(s)} e^{-\sqrt{\frac{k_0(s)}{D}}\rho_0(s)}, \quad (4.89)$$

which matches well the outer solution (4.70) :

$$p_{inner}^1(\rho = \rho_0(s), s) = p_{outer}(\rho = \rho_0(s), s) + O \left(e^{-\sqrt{\frac{k_0(s)}{D}}\rho_0(s)} \right). \quad (4.90)$$

Similarly, p_{inner}^2 , matches also very well :

$$p_{inner}^2(\rho = \rho_0(s), s) = p_{outer}(\rho = \rho_0(s), s) + O \left(e^{-\sqrt{\frac{k_0(s)}{D}}\rho_0(s)} \right). \quad (4.91)$$

We will now use the previous asymptotic analysis for the probability density function to estimate the overall probability P_N that a virus hits a small nuclear pore. Using formula (4.9), we get

$$P_N = 1 - \int_{\Omega} k(\mathbf{x})\tilde{p}(\mathbf{x})d\mathbf{x} = 1 - \left(\int_{\Omega \setminus BL} p_{outer}(\mathbf{x})k(\mathbf{x})d\mathbf{x} + \int_{BL} p_{inner}^i(\mathbf{x})k(\mathbf{x})d\mathbf{x} \right) \quad (4.92)$$

Using the outer solution expression and that $k(u, s) = \beta(s)Du$ (see 4.74) in the boundary layer, we have

$$P_N = 1 - \int_{\Omega \setminus BL} \frac{1}{|\Omega|} d\mathbf{x} - \int_{BL^1} \beta(s)Dup_{inner}^1(u, s)duds - \int_{BL^2} \beta(s)Dup_{inner}^2(u, s)duds, \quad (4.93)$$

where BL^1 and BL^2 are the boundary layers at resp. the absorbing and reflecting boundaries ($BL = BL^1 \cup BL^2$). Using expressions (4.86) and (4.87) for p_{inner}^i in (4.93), we obtain that

$$P_N = 1 - \frac{|\Omega \setminus BL|}{|\Omega|} - \int_{BL^1} \frac{\pi u}{|\Omega|} \left(Gi(u) - \frac{Gi \left(\frac{k_0(s)}{\beta(s)D} \right)}{Ai \left(\frac{k_0(s)}{\beta(s)D} \right)} Ai(u) \right) duds - \int_{BL^2} \frac{\pi u}{|\Omega|} \left(Gi(u) - \frac{Gi' \left(\frac{k_0(s)}{\beta(s)D} \right)}{Ai' \left(\frac{k_0(s)}{\beta(s)D} \right)} Ai(u) \right) duds,$$

Equivalently,

$$P_N = 1 - \frac{|\Omega \setminus BL|}{|\Omega|} + \int_{BL^1} \frac{\pi}{|\Omega|} \frac{Gi\left(\frac{k_0(s)}{\beta(s)D}\right)}{Ai\left(\frac{k_0(s)}{\beta(s)D}\right)} u Ai(u) duds + \int_{BL^2} \frac{\pi}{|\Omega|} \frac{Gi'\left(\frac{k_0(s)}{\beta(s)D}\right)}{Ai'\left(\frac{k_0(s)}{\beta(s)D}\right)} u Ai(u) duds - \int_{BL} \frac{\pi}{|\Omega|} u Gi(u) duds.$$

For large u , using the asymptotic expansion for $Gi(u)$ (4.88), we obtain that

$$\int_{BL} \frac{\pi}{|\Omega|} u Gi(u) duds \approx \frac{|BL|}{|\Omega|}. \quad (4.94)$$

Thus

$$P_N = \int_{BL^1} \frac{\pi}{|\Omega|} \frac{Gi\left(\frac{k_0(s)}{\beta(s)D}\right)}{Ai\left(\frac{k_0(s)}{\beta(s)D}\right)} u Ai(u) duds + \int_{BL^2} \frac{\pi}{|\Omega|} \frac{Gi'\left(\frac{k_0(s)}{\beta(s)D}\right)}{Ai'\left(\frac{k_0(s)}{\beta(s)D}\right)} u Ai(u) duds. \quad (4.95)$$

Using expression (4.88), we obtain

$$\frac{\frac{Gi\left(\frac{k_0}{\beta D}\right)}{Ai\left(\frac{k_0}{\beta D}\right)}}{\frac{Gi'\left(\frac{k_0}{\beta D}\right)}{Ai'\left(\frac{k_0}{\beta D}\right)}} = \frac{Gi\left(\frac{k_0}{\beta D}\right) Ai'\left(\frac{k_0}{\beta D}\right)}{Gi'\left(\frac{k_0}{\beta D}\right) Ai\left(\frac{k_0}{\beta D}\right)} = O\left((\beta D)^{-\frac{3}{2}}\right) = O\left(\frac{1}{\sqrt{D}}\right). \quad (4.96)$$

In addition, in the small diffusion approximation $D \ll 1$, we have :

$$P_N \approx \int_{BL^1} \frac{\pi}{|\Omega|} \frac{Gi\left(\frac{k_0(s)}{\beta(s)D}\right)}{Ai\left(\frac{k_0(s)}{\beta(s)D}\right)} u Ai(u) duds. \quad (4.97)$$

Using that $u = \frac{k_0(s) + k_1(s)\rho}{\beta(s)D} > \frac{k_0(s)}{\beta(s)D} \gg 1$ and the asymptotic expansions (4.88), we obtain that

$$P_N \approx \frac{1}{|\Omega|} \int_{BL^1} \frac{\left(\frac{k_0(s) + k_1(s)\rho}{\beta(s)D}\right)^{\frac{3}{4}} e^{-\frac{2}{3}\left(\frac{k_0(s) + k_1(s)\rho}{\beta(s)D}\right)^{\frac{3}{2}}}}{\left(\frac{k_0(s)}{\beta(s)D}\right)^{\frac{3}{4}} e^{-\frac{2}{3}\left(\frac{k_0(s)}{\beta(s)D}\right)^{\frac{3}{2}}}} d\rho ds \quad (4.98)$$

that is :

$$P_N \approx \frac{1}{|\Omega|} \int_{BL^1} \left(1 + \frac{k_1(s)}{k_0(s)}\rho\right)^{\frac{3}{4}} e^{-\frac{2}{3}\left(\frac{k_0(s)}{\beta(s)D}\right)^{\frac{3}{2}} \left(\left(1 + \frac{k_1(s)}{k_0(s)}\rho\right)^{\frac{3}{2}} - 1\right)} d\rho ds. \quad (4.99)$$

Because $0 \leq \rho \leq \rho_0(s)$ with $\rho_0(s) = O\left(\sqrt{\frac{D}{k_0(s)}}\right) \ll 1$, we use a Taylor expansion to obtain

$$P_N \approx \frac{1}{|\Omega|} \int_{BL^1} e^{-\left(\frac{k_0(s)}{\beta(s)D}\right)^{\frac{3}{2}} \frac{k_1(s)}{k_0(s)} \rho} d\rho ds. \quad (4.100)$$

Finally, by replacing $\beta(s)$ by its expression (4.74),

$$P_N \approx \frac{1}{|\Omega|} \int_{BL^1} e^{-\sqrt{\frac{k_0(s)}{D}} \rho} d\rho ds. \quad (4.101)$$

By integrating (4.101) over ρ , we have :

$$P_N \approx \frac{1}{|\Omega|} \int_{\partial N_a} \sqrt{\frac{D}{k_0(s)}} \left(1 - e^{-\sqrt{\frac{k_0(s)}{D}} \rho_0(s)}\right) ds. \quad (4.102)$$

For a sufficiently smooth killing field, when ∂N_a consists of n well separated small absorbing nuclear pore located at the points $(x_q)_{1 \leq q \leq n}$ on $\partial\Omega$, we finally obtain :

$$P_N \approx \frac{|\partial N_a|}{|\Omega|} \sum_{q=1}^n \sqrt{\frac{D}{k(x_q)}} + O\left(e^{-\sqrt{\frac{k_0}{D}} \rho_0}\right) \quad (4.103)$$

with $k_0 = \inf_q k(x_q)$ and $\rho_0 = \inf_{s \in \partial N_a} \rho_0(s)$. In a three dimensional cell with narrow pores of radius η_q , $1 \leq q \leq n$, we obtain

$$P_N \approx \sum_{q=1}^n \frac{\pi \eta_q^2}{|\Omega|} \sqrt{\frac{D}{k(x_q)}} + O\left(e^{-\sqrt{\frac{k_0}{D}} \rho_0}\right). \quad (4.104)$$

Chapitre 5

The endosomal step of viral infection

1 Introduction

Viral infection is a multi-step process that starts for most animal viruses at the cell surface where they are internalized inside an endosomal compartment. Viruses must then successfully escape the endosome, either to undergo cytoplasmic replication, or to reach a nuclear pore and deliver their DNA payload [15, 14]. Highly conserved activities in viral proteins are associated with endosomal escape. In enveloped viruses, membrane-associated glycoproteins mediate the fusion between the viral and endosomal membranes, while non-enveloped viruses possess penetration proteins able to initiate formation of small pores leading to the endosomal membrane lysis [89]. Due to the central role of endosomal escape in the viral entry process, it is important to understand its nature and dynamics. Our goal here is to analyze this step and to investigate how it depends on parameters such as the influx rate, the number of viral particles or the endosome size.

Glycoproteins and penetration proteins are usually activated by a pH dependent conformational change. As protons are pumped into the endosome, they bind the proteins and trigger their conformational changes. This is the case for class I and II glycoproteins of enveloped viruses (e.g. influenza virus hemagglutinin (HA) [90] and flavivirus E protein [91]). Alternatively, conformational changes can be mediated by low pH-activated endosomal proteases as for the Ebola virus [92] or the SARS coronavirus [93]. In the case of non-enveloped viruses such as reoviruses, parvoviruses and papillomaviruses, penetration proteins are activated at low pH, often through the action of endosomal proteases [94, 95, 96]. Consequently, the efficiency of viral infection must critically depend on both the resident time within the endosome and the pH value. Yet, the absence of direct measurements of these parameters makes the endosomal step hard to understand. Here, we propose a biophysical model to analyze the endosomal resident time and the success of endosomal escape. In that model, we consider that changes in the viral escape proteins are triggered by the cumulative discrete events of protons or low pH activated ligand bindings.

Using a Markov jump process, we estimate the resident endosomal time and obtain a description of bound ligand dynamics. We obtain the conformational time as a function of the pH and validate our analysis by comparison with experimental data obtained with the influenza HA [36]. We further confirm that only rearrangements in the subunit HA1 are pH-dependent, other HA changes occur spontaneously [35,

37]. The conformational change of active proteins is a common feature to widely disparate viruses although the escape mechanism depends strongly on the virus type. We are interested here in the release of small naked viruses coated by few active proteins and consider that they escape the endosome when a single protein changes of conformation and this is the limiting event. However, for enveloped viruses, such as influenza, several other steps are required, such as a possible cooperativity of the HA molecules [97] (see discussion). We use a Poissonian ligand entry dynamics to derive the mean resident time before escape. Interestingly, we found that increasing the number of viruses or the ligand influx, decreases the mean and the variance escape time. Surprisingly to escape in the most favorable pH-range of [6.1 – 6.3], we obtain that the number of viral particle should be equal to 5. In addition, the endosomal size, that may vary due to fusion or splitting [2], does not impact much the escape dynamics rate.

2 Results

Modeling the conformational change of glycoproteins and penetration proteins. The resident time of a virus inside an endosome depends on its ability to disrupt and pass through the membrane. Escape mechanism is still unclear but while membranes of enveloped viruses and endosome fuse locally, it seems that naked virus form small pores in the endosomal membrane, leading to osmotic swelling and lysis of that compartment [89]. These escape mechanisms are respectively induced by glycoproteins or penetration protein conformational change. To estimate the resident time τ_e , we start when viruses are already located in a formed endosome, neglecting endosomal fusion or split [2]. In a spherical endosome of $r_0 = 0.45\mu m$ and volume V_0 carrying n_v viruses, viral particles carry n_P independent proteins (glycoproteins or penetration proteins) formed of n_s sites that can bind competitively ligands such as protons or endosomal proteases. Thus, a total of $n_v n_P n_s$ sites can bind ligands. When the number of bound sites at a single protein reaches a critical threshold n_c , a conformational change occurs and initiates viral escape. In multi-step process, such as for the reovirus [94, 98] where activated cathepsins have to remove first the intermediate $\sigma 3$ protein, the general description can be obtained by using successively several time the present model for each intermediate step involved in the metamorphosis of the active protein.

To follow the conformational change for a single glycoprotein or a penetration protein, we count the amount of occupied sites $X(t, c)$ at time t , for a given ligand concentration c . During time t and $t + \Delta t$, the number of bound sites can either increase with a probability $r(X, c)\Delta t$ when a ligand arrives to a free site, decreases with probability $l(X, c)\Delta t$ when a ligand unbinds or remains unchanged with probability $1 - l(X, c)\Delta t - r(X, c)\Delta t$. Using the scaled variable $x(t, c) = \epsilon X(t, c)$ where $\epsilon = \frac{1}{n_s}$ and $\Delta x = x(t + \Delta t, c) - x(t, c)$, we obtain the transition probabilities

$$\begin{aligned} Prob\{\Delta x = \epsilon | x(t, c) = x\} &= r(x, c)\Delta t, \\ Prob\{\Delta x = -\epsilon | x(t, c) = x\} &= l(x, c)\Delta t, \\ Prob\{\Delta x = 0 | x(t, c) = x\} &= (1 - r(x, c) - l(x, c))\Delta t. \end{aligned}$$

When the ligand concentration is fixed, the probability $p(x, y, t, c)$ that the number of bound is equal to y at time t $x(t, c) = y$, given that initially the number of bound

is x ($x(t=0, c) = x$) is solution of the backward Kramers-Moyal equation [33] :

$$\begin{aligned} \frac{\partial p}{\partial t} &= L_x p = r(x, c) \sum_{n=1}^{\infty} \frac{\epsilon^n}{n!} (\partial_x)^n p(x, y, t, c) \\ &+ l(x, c) \sum_{n=1}^{\infty} \frac{(-\epsilon)^n}{n!} (\partial_x)^n p(x, y, t, c). \end{aligned} \quad (5.1)$$

The first time a glycoprotein or a penetration protein is filled up to a critical threshold $x_c = \frac{n_c}{n_s}$ is the mean first passage time $\tau(x, c)$ that the level of bound ligands x reaches the level x_c , starting at a point x for a given concentration c , and satisfies [39] :

$$\begin{aligned} L_x \tau(x, c) &= -1 \text{ for } x \text{ in } [0, x_c], \\ \tau(x, c) &= 0 \text{ for } x = x_c \text{ and } \frac{\partial \tau(x, c)}{\partial x} = 0 \text{ for } x = 0. \end{aligned}$$

To estimate $\tau_0(c) = \tau(x_0(c), c)$, we consider the number of bound ligands at equilibrium $0 < x_0(c) < x_c$. For $\epsilon \ll 1$, $\tau_0(c)$ is approximated by [33] :

$$\tau_0(c) \approx C(\epsilon, c) \left(1 - \left(\frac{l(x_c, c)}{r(x_c, c)} \right)^{-\frac{x_c - x_0(c)}{\epsilon}} \right), \quad (5.2)$$

where

$$C(\epsilon, c) \approx \frac{1}{r(x_0(c), c)} \frac{\sqrt{\frac{2\pi}{\epsilon \frac{d}{dx} \left(\frac{l}{r} \right) (x_0(c), c)}}}{\phi(x_c, c)}$$

and

$$\phi(x, c) = \frac{\frac{1}{\epsilon} \int_{x_0(c)}^x \log \left(\frac{l(s, c)}{r(s, c)} \right) ds}{\sqrt{\frac{l(x, c)}{r(x, c)}} \left(\frac{l(x, c)}{r(x, c)} - 1 \right)}.$$

Formula (5.2) links the affinities between the ligand (concentration c) and the binding sites of glycoproteins or penetration proteins to its conformational change mean time $\tau_0(c)$. We validate our model by comparing our analytical formula for the mean time with the measured HA conformational change kinetics obtained from experimental data for various pH.

Validation of the model for the influenza HA. We derive here the transition rate r and l : r is the standard forward rate to the free binding sites. In the approximation that the binding sites occupy a small fraction of the available endosomal surface, $r(x, c) = (n(c)n_s(1-x))/\tau$ [75] where $n(c) = \mathcal{N}cV_0$ is the number of endosomal protons in excess at concentration c (\mathcal{N} is the Avogadro's number), $n_s(1-x)$ is the amount of free sites and τ is the mean time for a proton to activate a binding site.

To determine $\frac{l(x, c)}{r(x, c)}$, we use the experimental values available for the influenza HA [35] obtained in a pH range of [4 – 7] and at temperature $T = 300K$, where the average number of bound protons to HA1 is approximated by a decreasing linear function of the pH (FIG. 3 of [35]) :

$$x_0(c) \approx 1 + \frac{\log_{10}(c) + 4}{7 - 4} = 1 + \frac{1}{3} (\log_{10}(c) + 4). \quad (5.3)$$

Because HA cannot bind less than the mean number of bound protons at $pH = 7$, we calibrate $x_0(10^{-7}molL^{-1}) = 0$. Interestingly, the relation $x_0(c)$ shows that we are not in the classical framework of the theory of chemical reactions where at equilibrium, the number of bounds x_0 is a linear function of the proton concentration, rather here the dependency is logarithmic. This suggests a possible mechanism of interactions between the HA1 binding sites : a cooperativity mechanism between bound sites that should be further investigated. In addition, when the pH decreases from 7 to 4, the number of bound protons increases approximatively from 123 to 132 (FIG 3, [35]), thus the maximal number of bound protons is $n_s = 9$, which defines the small parameter $\epsilon = 1/9$).

We can now obtain the equilibrium ratio $\frac{l(x, c)}{r(x, c)} = \frac{k_d(x)\epsilon\tau}{\mathcal{N}cV_0(1-x)}$, where $k_d(x)$ is the dissociation rate. Thus $\frac{l(x, c)}{r(x, c)} = \frac{f(x)}{c}$ where $f(x) = \frac{k_d(x)\epsilon\tau}{\mathcal{N}V_0(1-x)}$. Using the concentration $c(x) = 10^{-(3(1-x)+4)}$ for which $x_0(c(x)) = x$, at equilibrium the mass action law reads $\frac{l(x_0(c), c)}{r(x_0(c), c)} = 1$, equivalently $\frac{f(x)}{c(x)} = 1$ and we get

$$\frac{l(x, c)}{r(x, c)} = \frac{f(x) c(x)}{c(x) c} = \frac{10^{-(3(1-x)+4)}}{c}. \quad (5.4)$$

To estimate from (5.2), the mean conformational time $\tau_0(c)$, we use the kinetics information about the HA conformational changes at a given pH to determine the two remaining parameters : the activation threshold x_c and the mean time of a ligand to a free binding site τ . Using the measured datas $(\tau_e^0(pH = 4.9))^{-1} = k_e(pH = 4.9) = 5.78s^{-1}$, $k_e(pH = 5.1) = 0.12s^{-1}, \dots, k_e(pH = 5.6) = 0.017s^{-1}$ [36] and a least squares method for interpolation of the time, we obtain $x_c \approx 0.7017281$ and the activation time $\tau \approx 30.4 * 10^3s = 8h20min$ which is quite high. Comparing this activation time with the narrow escape time NET (time for a proton to find a binding site), we use that [75] $NET \approx \frac{V_0}{4\pi D_p \eta}$, where η is the interacting radius between a proton and a binding site and $D_p \approx 100\mu m^2s^{-1}$ [99] is the apparent cytoplasmic proton diffusion coefficient. In the absence of further data, we approximate the endosomal diffusion constant by the cytoplasmic one. For $\eta = 1nm$, we found $NET = 0.3s$. This result suggests that the binding time is dominated by a high activation barrier, to guarantee that the conformational change is not triggered randomly, but only upon binding, in agreement with the large stability of the HA molecule (see table 2 [36]). As a consequence, endosome viral escape should not occur before a certain amount of protons enter the endosome.

To confirm our analysis, we plotted in Fig. 5.2 the mean rate of HA conformational change $k_e = (\tau_e^0)^{-1}$ as a function of the concentration c and compared it

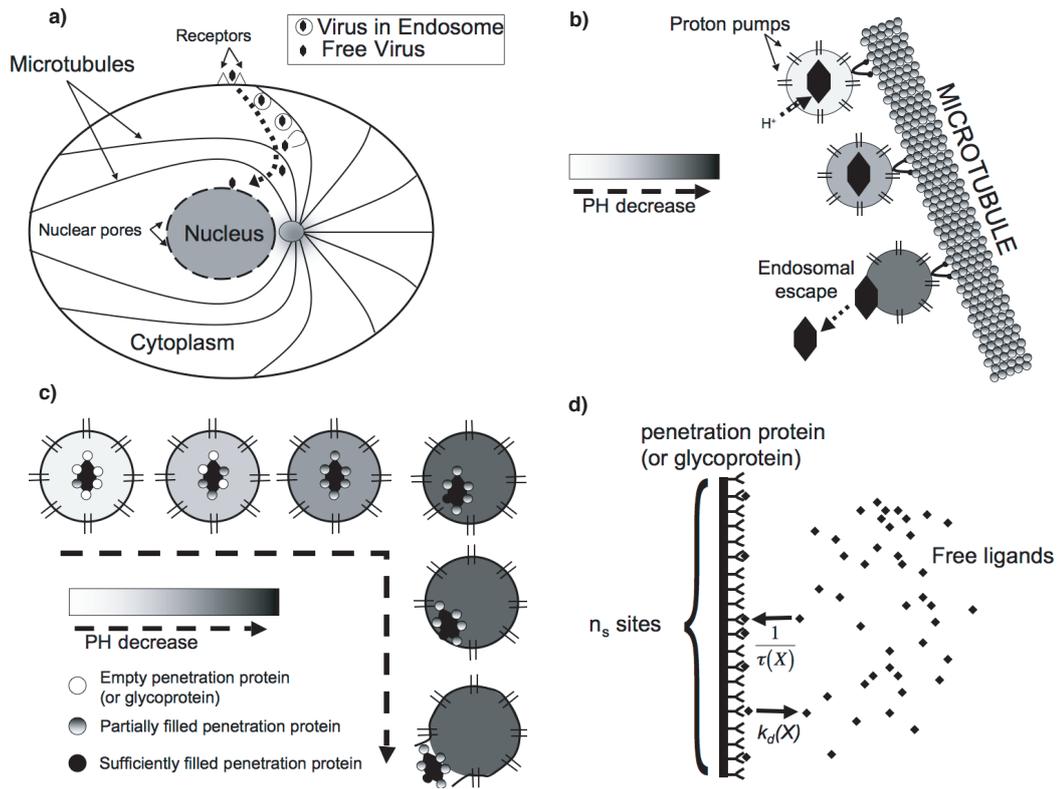


FIGURE 5.1 – (a) Virus is endocytosed by the cell and traffics through the endosomal pathway. It then escapes and, in the case of a DNA-virus, moves freely in cytoplasm to reach a small nuclear pore and delivers its genetic material to the nucleus. (b) Proton pumps located on endosomal membrane lower pH in the endosome. At sufficiently low pH, virus escapes. (c) When pH decreases under proton pumps activity, ligands bind glycoproteins or penetration proteins. When sufficiently sites of the protein are occupied, conformational change occurs and initiate viral escape. (d) Free endosomal ligands bind (forward binding rate $\frac{1}{\tau(X)}$) and undind (dissociation rate $k_d(X)$) to glycoprotein or penetration protein (X is the amount of already bound sites). Once a critical number n_c of sites among the n_s available sites of the protein are occupied, conformational change occurs.

with discret experimental values [36]. The matching between theory and experiments emphasizes that our model accounts well for the mean conformational time based on cumulative discret bindings of ligands reaching an activating threshold. Moreover, our analysis reemphasizes that only the dissociation of HA1 subunit requires an acidic pH and that other rearrangements in HA occur spontaneously [35, 37]. We highlight that the conformational changes occurs when roughly 70% of HA is filled ($x_c \approx 0.7$), that is when 6 protons have been added to the protein. Finally we suggested here that as the HA sites are bound, the activation barrier for the remaining sites should change by a cooperativity process. The parameters of HA are summarized in table 5.1.

Dynamics of ligand influx and viral escape. A first approximation of our model is that endosomes do not contain intraluminal vesicles, thus once a virus disrupts the endosomal membrane, it is released in the cytoplasm. In the class of naked viruses, coated by few (≈ 10) penetration proteins such as parvoviruses (with 7

TABLE 5.1 – Dynamical parameters of HA

Parameters	Description	Value
$r(x, c)$	Binding rate	$r(x, c) = (\mathcal{N}cV_0n_s(1-x))/\tau$
$l(x, c)$	Unbinding rate	$l(x, c) = (r(x, c)/c)10^{-(3(1-x)+4)}$
x_c	Conformational change threshold	$x_c = 0.7017281$
τ	Mean time of a ligand to a free binding site	$\tau = 30.4 * 10^3 s$
$n_s = 1/\epsilon$	Number of binding sites	$n_s = 9$
V_0	Volume of the endosome	$V_0 = 4/3\pi r_0^3 = 0.38\mu m^3$

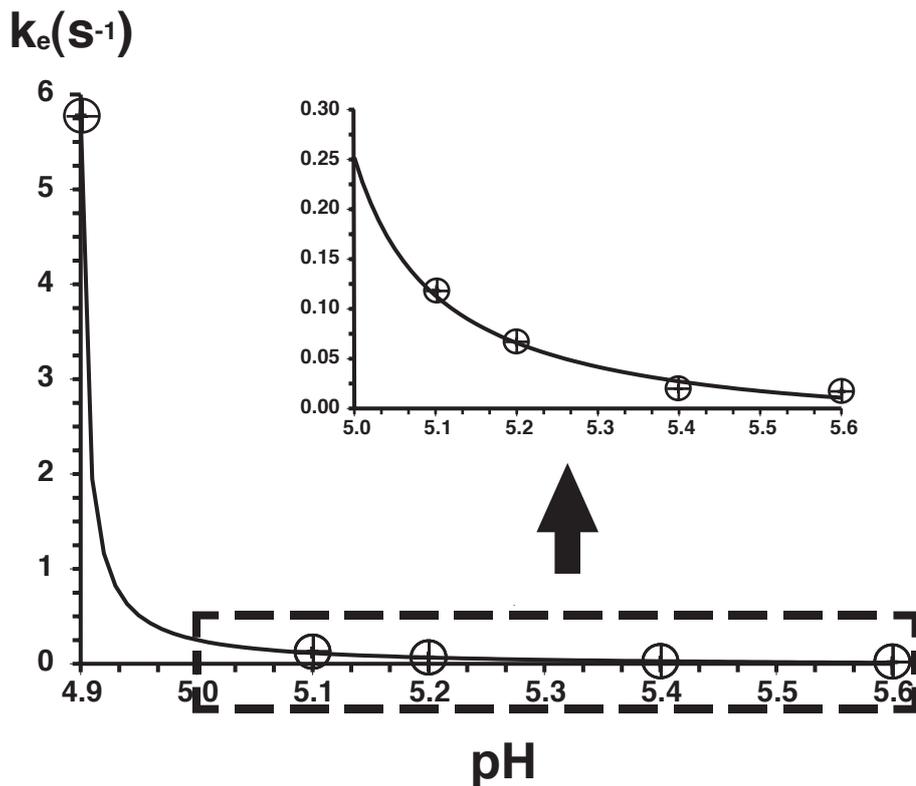


FIGURE 5.2 – Rate of HA conformational changes : $k_e(c) = (\tau_e^0(c))^{-1}$ (dashed line) is compared with the experimental data (circled crosses)[36].

VP1 penetration proteins for adeno-associated viruses (AAV) [38], we consider that endosomal escape occurs when at least one protein changes conformation and forms a pore in the membrane. In addition, as suggested in an experiment where a wild type parvovirus was shown to rescue the infectivity of a mutant devoid of escape protein [40], we consider that all viruses escape when at least one of them disrupts the membrane.

To model the proton influx through pumps uniformly distributed over the endosomal membrane [49], we use a Poisson process of rate λ . More generally, we consider that binding ligands such as proteases, activated at low pH, follow a Poissonian description. The probability $P_q(t)$ of having exactly q ligands in the endosome at time t is given by

$$P_q(t) = e^{-\lambda t} \frac{(\lambda t)^q}{q!}. \quad (5.5)$$

The rate λ can be inferred from recent experiments where ATP addition in endocytic vesicles from kidney proximal tubule, leads to pH decay from 7.2 to 6.6 during a time $T = 5 \text{ min}$ [49], thus $\lambda T = (10^{-6.6} - 10^{-7.2}) V_0 \mathcal{N}$, leading to

$$\lambda \approx 0.15 \text{ s}^{-1}. \quad (5.6)$$

Surprisingly, we conclude that on average, 10 protons enter the endosome per minute and at $pH = 5$, the number of protons in an endosome of volume $V_0 = 0.38 \mu\text{m}^3$ is approximatively equal to 2300. Only $n_s x_c \approx 6$ protons are needed for the conformational change of HA to occur (see table 5.1). We thus assume that bound ligands to the few penetration proteins do not deplete the free ligands pool of the endosome. Consequently, each protein can be treated separately. For viruses coated by a large number of glycoproteins or penetration proteins such as the influenza (400 HA copies [100]) or the reovirus (200 μ_1 trimers [89]), our analysis does not apply and future model should account for the depletion of the free ligands pool due to the binding events.

To estimate the mean ligands concentration c_{τ_e} at which viruses escape the endosome, we estimate the probability $P_e^0(c)$ that a penetration protein changes conformation before a new ligand enters or is activated (the concentration c is fixed). The probability of no conformational change is $\tilde{p}(x_0(c), t, c) = \int_0^{x_c} p(x_0(c), y, t, c) dy$ where initially $x(t=0) = x_0(c)$. Because a ligand is injected during time t and $t + dt$ with a rate $\lambda e^{-\lambda t} dt$, the probability of no conformational changes until a ligand appears is $\int_0^\infty \tilde{p}(x_0(c), t, c) \lambda e^{-\lambda t} dt$. Thus

$$\begin{aligned} P_e^0(c) &= 1 - \int_0^\infty \tilde{p}(x_0(c), t, c) \lambda e^{-\lambda t} dt \\ &= 1 - u(x_0(c)), \end{aligned} \quad (5.7)$$

Multiplying equation (5.1) by $\lambda e^{-\lambda t}$ and integrate over time, we get

$$\begin{aligned} (L_x - \lambda) u(x) &= -\lambda \text{ for } 0 < x < x_c, \\ \frac{du(x)}{dx} &= 0 \text{ for } x = 0, \\ u(x) &= 0 \text{ for } x = x_c. \end{aligned} \quad (5.8)$$

To solve (5.8), we apply the methods of [33] (see the appendix for computational details) and we obtain

$$P_e^0(c) = 1 - \frac{\lambda}{\lambda + (C(\epsilon, c))^{-1}} + \frac{\lambda}{\lambda + (C(\epsilon, c))^{-1}} \left(\frac{l}{r}(x_c, c) \right)^{-\frac{x_c - x_0(c)}{\epsilon}}. \quad (5.9)$$

Since one protein conformational change is enough to induce viral escape, we estimate the probability $P_e(j)$ that at least one conformation changes occurs after exactly j ligands have entered the endosome (or have been activated). $P_e(j)$ is the product of the probabilities that no conformational changes occurs between successive ligand entrance, until the j^{th} one

$$P_e(j) = \left(1 - (1 - P_e^0(c(j)))^{n_v n_P} \right) \prod_{i=0}^{j-1} (1 - P_e^0(c(i)))^{n_v n_P} \quad (5.10)$$

where $c(i) = \frac{i}{\mathcal{N}V_0} + c_0$ is the concentration associated with the entry (or activation) of i ligands in the endosome (c_0 is the initial concentration at $pH = 7$).

Considering an influx rate $\lambda = 0.15s^{-1}$ (see (5.6)) and a number $n_P = 7$ of active proteins as measured for the AAV [38], we plotted in FIG. 5.3 (a), the probability density function (pdf) P_e as a function of the number of ligands j . In the absence of Because specific information on penetration proteins, such as the VP1 of parvoviruses, we use for glycoprotein HA parameters the ones summarized in table 5.1. Interestingly, FIG. 5.3 (a) shows a large dispersion, indicating that escape can happen with few ligands, very early in the endosomal trafficking. Viruses must have escaped when at most 300 protons have entered an endosome (FIG.5.3 (a), which corresponds to a $pH = 5.9$, the ligand concentration c when escape occurs is such

that $x_c - x_0(c) \gg \epsilon$ and thus we neglect $\left(\frac{l}{r}(x_c, c) \right)^{-\frac{x_c - x_0(c)}{\epsilon}} \approx 0$ in formula (5.2) and (5.30).

The mean concentration $\langle c_{\tau_e} \rangle = \frac{\sum_{j=1}^{\infty} j P_e(j)}{\mathcal{N}V_0}$ for which viral particles escape the endosome is (see the appendix)

$$\langle c_{\tau_e} \rangle = \frac{1}{\mathcal{N}V_0} \sum_{j=0}^{\infty} \prod_{i=0}^j \left(\frac{\lambda}{\lambda + (C(\epsilon, c(i)))^{-1}} \right)^{n_v n_P} \quad (5.11)$$

To estimate the escape time, we consider the density function $cm(t) = Pr\{\tau_e < t\}$. Some computations lead to (see the appendix)

$$cm(t) = 1 - e^{-\lambda t} + \sum_{k=1}^{\infty} (-\lambda)^k \sum_{i=1}^k \frac{e^{-\lambda t} - e^{-(\lambda + \lambda_i)t}}{\lambda_i \prod_{j=1, j \neq i}^k (\lambda_i - \lambda_j)},$$

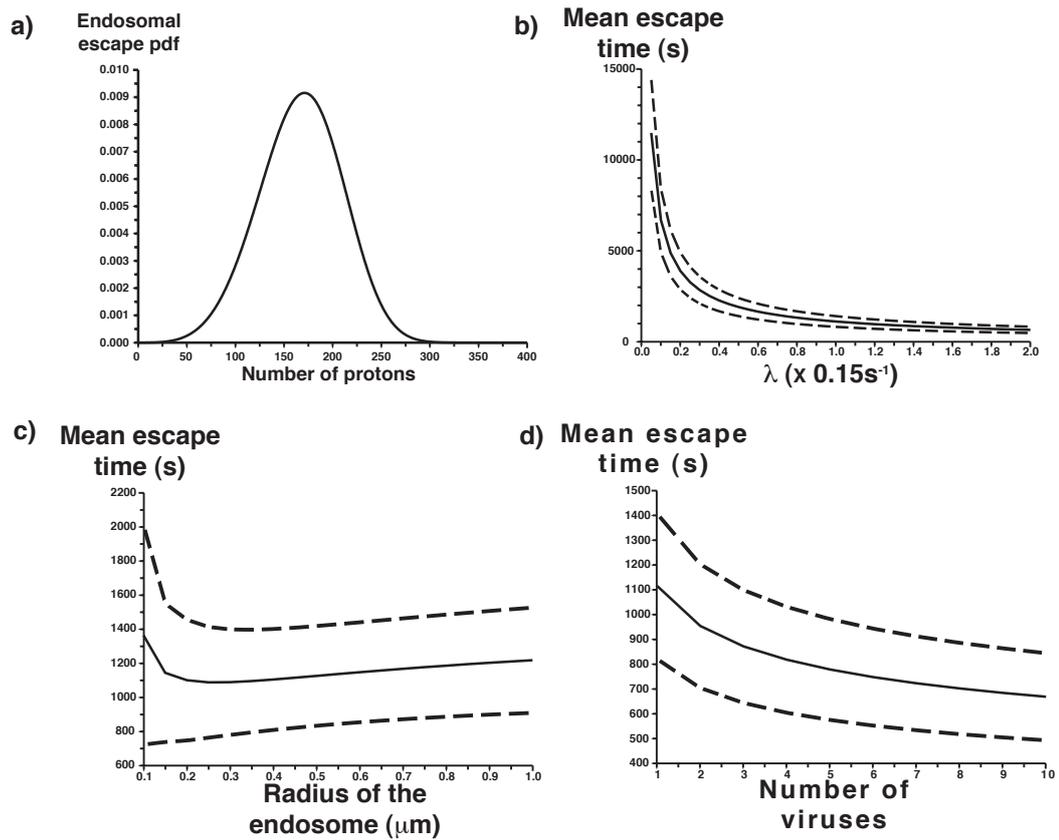


FIGURE 5.3 – (a) Escape time distribution as a function of the protons. Parameters for the HA (table 5.1) are $\lambda = 0.15s^{-1}$, $n_P = 7$, and $n_v = 1$. (b) Mean escape time as a function of the entry rate λ . (c) Mean escape time as a function of the endosome radius when the entry rate is proportional to endosome surface. (d) Mean escape time as a function of the virus number n_v .

where $\lambda_i = \frac{n_v n_P}{C(\epsilon, c(i))}$; the probability density function (pdf) is

$$p_{\tau_e}(t) = \frac{dcm(t)}{dt} = \lambda e^{-\lambda t} - \sum_{k=1}^{\infty} (-\lambda)^k \sum_{i=1}^k \frac{\lambda e^{-\lambda t} - (\lambda + \lambda_i) e^{-(\lambda + \lambda_i)t}}{\lambda_i \prod_{j=1, j \neq i}^k (\lambda_i - \lambda_j)}.$$

The mean escape time $\bar{\tau}_e = \int_0^{\infty} t p_{\tau_e}(t) dt = \int_0^{\infty} (1 - cm(t)) dt$ is given by (see the appendix)

$$\bar{\tau}_e = \frac{1}{\lambda} \left(1 + \sum_{k=1}^{\infty} \frac{1}{\prod_{i=1}^k (1 + \lambda_i/\lambda)} \right). \quad (5.12)$$

Similarly, we obtain for the variance $var(\tau_e) = \int_0^{\infty} t^2 p_{\tau_e}(t) dt - \bar{\tau}_e^2$ (see the appendix) :

$$var(\tau_e) = \frac{2}{\lambda^2} + \frac{2}{\lambda^2} \sum_{k=1}^{\infty} \frac{1}{\prod_{i=1}^k (1 + \lambda_i/\lambda)} + \frac{2}{\lambda^2} \frac{\sum_{j=1}^k \prod_{i=1, i \neq j}^k (1 + \lambda_i/\lambda)}{\left(\prod_{i=1}^k (1 + \lambda_i/\lambda) \right)^2} - \bar{\tau}_e^2. \quad (5.13)$$

For the HA parameters (table 5.1), the protons entry rate (5.6) $\lambda = 0.15s^{-1}$ and $n_P = 7$ penetration proteins, we obtain a mean time $\bar{\tau}_e$ and standard deviation $\sqrt{var(\tau_e)}$ equal to

$$\bar{\tau}_e = 1115s \approx 20min \text{ and } \sqrt{var(\tau_e)} = 294s \approx 5min. \quad (5.14)$$

Using formula (5.12) and (5.13), we evaluate the effect of various parameters on the mean and variance. In FIG. 5.3 (b) and (d), we plot the mean escape time $\bar{\tau}_e$ ($\pm \sqrt{var(\tau_e)}$) as a function of the ligands entry rate λ and the initial number of viruses in the endosome. Interestingly, for 10 viruses, the mean escape time decays by 35% compared to a single one. Moreover a large number of viruses, as well as a higher ligands entry rate, are associated with a more reliable escape time. In a first approximation, we consider that the number of pumps and the influx rate are proportional to the endosomal surface : $\lambda(r) = (r/r_0)^2 \lambda = (r/r_0)^2 0.15s^{-1}$. We plot the mean escape time as a function of the endosome radius (FIG. 5.3 (c)). We observe that the endosomal radius does not impact much the mean escape time and its variance. Consequently, it seems appropriate *a posteriori* to have neglected processes such as endosomal fusion or split [2] in the biophysical modeling of the endosomal step of viral infection. However, if future observations show that endosomal fusion or fission events control the endosomal acidification in a more complex way, our model can be generalized to integrate the time dependency of the influx rate. Another pitfall of our analysis is that depending on the size of the endosome, a non attached individual virion should swim in the endosome, and for larger endosome, the mean position of the virus should impact its escape dynamics. Unless the virions is most of the time attached to the membrane, it seems more likely that it will cause infection when it will stay in average closer to the endosomal membrane, when the pH drops. However, whether or not the virion spends most of its time attached or not to the membrane is still unclear.

3 Discussion

The viral endosomal journey depends crucially on pH and timing : on one hand, viruses must escape before being digested by the late endosomal proteases, while on the other hand, time is required for some of them to be primed by pH-activated components of the endosome such as cathepsins. The pH-dependent conformational change of a glycoprotein (e.g. HA of influenza virus) or a penetration protein (e.g. VP1 capsid protein of parvoviruses) is responsible for the first step of endosomal membrane disruption and viral escape. We first computed the mean time needed for the number of bound ligands to a glycoprotein or a penetration protein to reach a critical threshold and trigger conformational change of the protein. We validated our analysis using experimental data on the pH dependent number of bound protons to the influenza HA protein [35]. Using our model, we were able to recover the kinetics of HA conformational change [36] and confirm that only rearrangements of the HA1 subunit requires an acidic pH [35, 37].

The endosomal escape depends strongly on the number of viral particles. To model endosomal release for naked viruses containing few (about ten) active proteins, we used the mean time for one of these proteins to change conformation. Using HA dynamical parameters (Table 5.1), when the ligand entry rate is $\lambda = 0.15s^{-1}$ (5.6) and $n_P = 7$ penetration proteins, we find that $\bar{\tau}_e \pm var(\tau_e) \approx 20min \pm 5min$. It has been reported that in less than 15 minutes [2], the genetic material is transported from the cell surface through early endosome compartments to the late endosome, we thus predict that viruses should escape from the late phase of the endosome trafficking.

Interestingly, the size of the endosome, that may vary due to fusion or splitting [2], does not impact drastically the dynamics of the viruses escape, while increasing the ligands entry rate or the number of viruses, decreases the mean escape time $\bar{\tau}_e$ and the variance $var(\tau_e)$, making the escape process more reliable (FIG. 5.3 (b)-(d)). When the number of viral particle is increased from 1 to 10, the time $\bar{\tau}_e$ is decreased by 35%. We conclude that the number of endocytosed viral particles and thus the number of glycoproteins or penetration proteins is a key parameter in the escape process. Consequently, the number of active escape proteins that can vary with viral species or serotypes should be considered as a potential target in future genetically modified or chimeric viral vectors [101]. Furthermore, we provided here a general framework to analyze endosomal escape of most naked viruses. When the ligand entry rate λ and the mean conformational change time curve $C(\epsilon, c)$ for proteins involved in the endosomal escape would be obtained experimentally, from our model, we can obtain the mean escape time $\bar{\tau}_e$ (5.12) and the ligand concentration $\langle c_{\tau_e} \rangle$ (5.35). It would be interesting to obtain this conformational rate change for VP1, involved in the parvoviruses endosomal escape.

Limitations and further issues : Our analysis is a first step toward a quantitative approach of both enveloped and naked virus endosomal escape. It provides a tool to explore the multi-dimensional parameter space of endosomal escape and give a measure of global quantitative outputs such as the mean escape time from the endosome depends on fundamental parameters, such as the endosomal size or the influx rate. While the binding of ligands to an active protein up to a critical threshold and the resulting conformational change seems to be a common feature of widely disparate viruses, the endosomal escape mechanism is still unclear and

strongly depends on viral species. Here we presented a simplified model for naked viruses covered by few active proteins in which, all viruses contained in the endosome, escape rapidly when one of the active proteins binds enough ligands and becomes activated.

Our model rely on several biophysical and mathematical assumptions. First, we considered the rate limiting step of the escape process was the binding of ligands to an active protein up to a critical threshold. Future models should account for the other steps participating to the escape process : this includes the pore formation limiting step, involved in membrane disruption [89]. Interestingly, more than one active proteins should be required to the pore formation.

We have also modeled the proton entry as Poissonian. A better knowledge of the proton channel dynamics would certainly reduce the error of this approximation, although we can consider that channels have no memory. Our major mathematical approximation is the asymptotic solution (5.2) of the Kramer Moyal equation. That asymptotic relies on the small parameter $\varepsilon = 1/9$ which is the reciprocal of the number of binding sites. The good agreement of that asymptotic solution with the hemagglutinin data confirms the consistency of our analysis. For enveloped viruses, conformational change of glycoproteins initiate complex mechanisms, at the origin of the local fusion between the virus and the endosome membranes. For the influenza virus, three intermediate and limiting steps have been recognized : (1)the conformational change of HA trimers and their clustering (2) shall precede the formation of a hemi-fusion stalk (3) and subsequent fusion pore [97]. In summary, our model should be considered as a starting building block of future attempts to integrate all these steps together.

Viral fitness and optimality of the delivery process. Although viruses must escape before being digested by the late endosome proteases, priming of other low pH activated components to the capsid should impact the infection efficiency in the later steps, such as the nuclear uncoating and subsequent delivery of the genetic material [102]. To account for these processes, there must be an optimal range of pH, $[pH_{min}; pH_{max}]$ (or equivalently a range in the number of endosomal ligands) in which viruses must escape. When escaping above pH_{max} , a viral particle will be irreversibly damaged by proteases and thus cannot pursue efficiently its infection, however, when escaping below pH_{min} , priming will be insufficient (see FIG. 5.4 (a)) to conduct the next step. Using the probability that viruses escape in a specific pH-range (FIG. 5.4 (b)), we obtain the dependency as a function of the number of viruses in the endosome. Interestingly, we find an optimal number of 5 viral particles, maximizing the probability to escape in that pH-range.

To conclude, we find here that the priming of viral capsids illustrates how virus trafficking involves complex host-cell interactions. Another example is the interaction of the enveloped retrovirus avian leukosis virus with the cell membrane specific receptor that triggers pH sensitivity of glycoproteins [10], essential for endosomal membrane fusion and payload delivery. More generally, as a response to host cell signals, glycoproteins of widely disparate viruses such as the paramyxovirus, the influenza virus or the HIV undergo successive transformations required for infectivity [9]. Future models should account for this complex virus-host communication, where each limiting step impacts viral fate (see FIG. 5.5). As chemically modified viral vectors are emerging concepts for gene delivery [103], a quantitative understanding

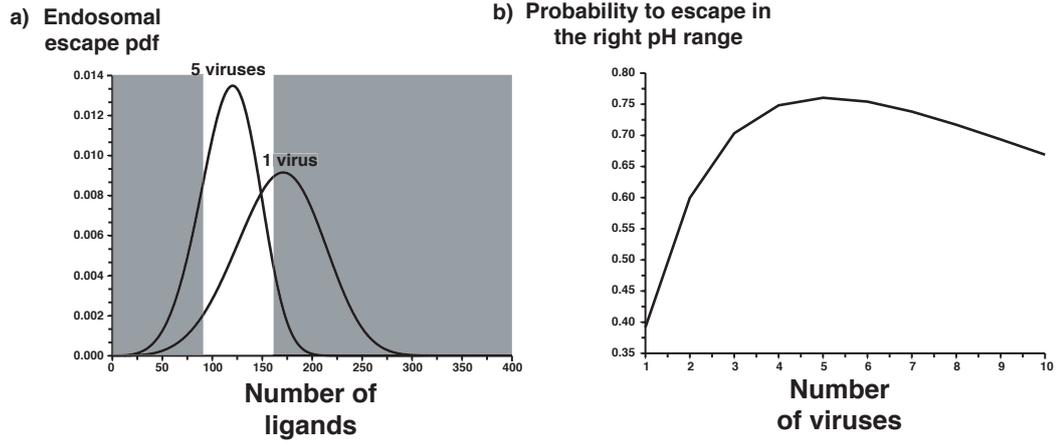


FIGURE 5.4 – (a) The number of endosomal viruses impacts the escape distribution and modifies the probability to escape in a given pH range (grey regions, $\text{pH} \in [6.1; 6.3]$ or equivalently the number of pumped protons is between 90 and 160). We used $\lambda = 0.15\text{s}^{-1}$, $n_P = 7$ and the dynamical parameters of HA (table 5.1) The plots are given for 1 and 5 viruses. (b) We show the probability of escape in a given pH range as function of the number of viruses. 5 viral particles is optimal.

of interactions between viral structural proteins and the cell environment should benefit drug design and optimization.

Acknowledgement : D. H. research is supported by an ERC-Starting Grant.

4 Appendix

4.1 Computation of $u(x_0(c))$

Far from the boundary, in the long time asymptotic [21]

$$\tilde{p}(x, t, c) \approx e^{-\frac{t}{C(\epsilon, c)}}. \quad (5.15)$$

and the leading order term of the outer expansion u_{outer} is

$$u_{outer} \approx Cte = \frac{\lambda}{\lambda + \frac{1}{C(\epsilon, c)}} = 1 + S(\epsilon, c), \quad (5.16)$$

where $S(\epsilon, c) = \sum_{i=1}^{\infty} \frac{(-1)^i}{(\lambda C(\epsilon, c))^i}$ tends to 0 as ϵ goes to 0. However, to satisfy the absorbing boundary condition $u(x_c) = 0$, we study the leading order term of the inner expansion u_{inner} of u that has to match asymptotically the outer solution expansion (5.16). In the neighborhood of $x = x_c$, we stretch the coordinates [33, 34] $\eta = \frac{x_c - x}{\epsilon}$ and consider $F(\eta) = u_{inner}(x)$. For the inner and outer expansions to match, $F(\eta)$ has to satisfy

$$F(0) = 0 \text{ and } \lim_{\eta \rightarrow \infty} F(\eta) = u_{outer} = 1 + S(\epsilon, c). \quad (5.17)$$

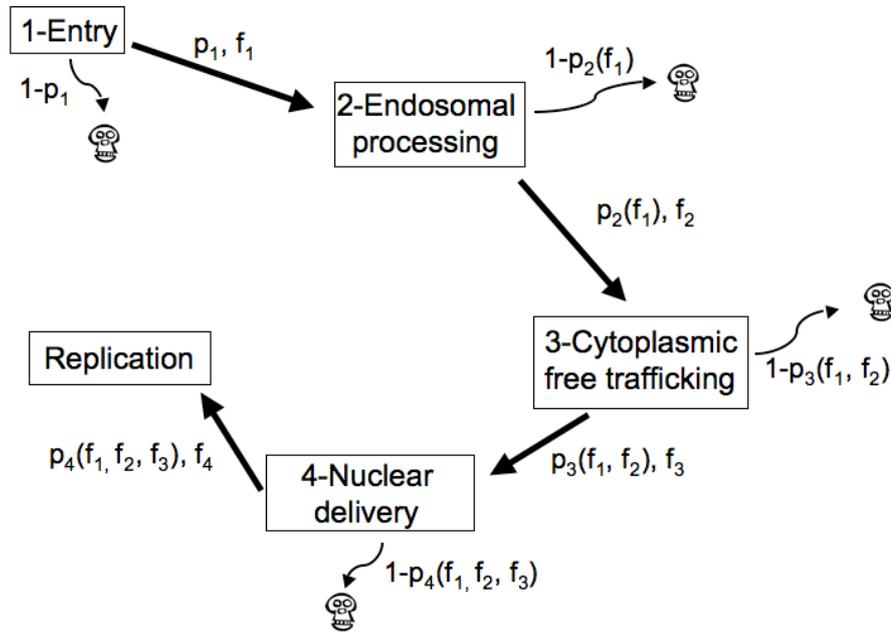


FIGURE 5.5 – Schematic representation of the viral infection steps. Each step is rate limiting (success with probability $(p_i)_{1 \leq i \leq 4}$), the probability of success in each step defines the fitness $(f_i)_{1 \leq i \leq 4}$ that impacts viral efficiency in future steps of infection.

Injecting $F(\eta)$ into the equation

$$\begin{aligned}
 (L_x - \lambda) u(x) &= -\lambda \text{ for } 0 < x < x_c, \\
 \frac{du(x)}{dx} &= 0 \text{ for } x = 0, \\
 u(x) &= 0 \text{ for } x = x_c,
 \end{aligned} \tag{5.18}$$

we get in the leading ϵ -order,

$$\begin{aligned}
 -\lambda &= F(\eta) \left[r(x_c, c) \sum_{n=1}^{\infty} \frac{(-1)^n}{n!} (\partial_\eta)^n \right. \\
 &\quad \left. + l(x_c, c) \sum_{n=1}^{\infty} \frac{1}{n!} (\partial_\eta)^n - \lambda \right].
 \end{aligned} \tag{5.19}$$

We seek solution $F(\eta)$ in the form

$$F(\eta) = (1 + S(\epsilon, c)) (1 - e^{-\gamma\eta})$$

and by reinjecting it in (5.19), we obtain :

$$\begin{aligned}
 \lambda &= \lambda (1 + S(\epsilon, c)) (1 - e^{-\gamma\eta}) + \left[r(x_c, c) \sum_{n=1}^{\infty} \frac{\gamma^n}{n!} \right. \\
 &\quad \left. + l(x_c, c) \sum_{n=1}^{\infty} \frac{(-\gamma)^n}{n!} \right] (1 + S(\epsilon, c)) e^{-\gamma\eta},
 \end{aligned} \tag{5.20}$$

that is :

$$\begin{aligned} \lambda = & \lambda (1 + S(\epsilon, c)) (1 - e^{-\gamma\eta}) + \left[r(x_c, c) (e^\gamma - 1) \right. \\ & \left. + l(x_c, c) (e^{-\gamma} - 1) \right] (1 + S(\epsilon, c)) e^{-\gamma\eta}. \end{aligned} \quad (5.21)$$

For ϵ sufficiently small, because S tends to zero, we get that γ is solution of

$$\begin{aligned} \lambda = & \lambda (1 - e^{-\gamma\eta}) + \left[r(x_c, c) (e^\gamma - 1) \right. \\ & \left. + l(x_c, c) (e^{-\gamma} - 1) \right] e^{-\gamma\eta}, \end{aligned} \quad (5.22)$$

thus

$$r(x_c, c) (e^\gamma - 1) + l(x_c, c) (e^{-\gamma} - 1) - \lambda = 0. \quad (5.23)$$

The quadratic equation in $X = e^\gamma$ is

$$X^2 - \left(1 + \frac{l(x_c, c) + \lambda}{r(x_c, c)} \right) + \frac{l(x_c, c)}{r(x_c, c)} = 0. \quad (5.24)$$

and the root X_1 strictly greater than 1 ($\gamma > 0$) is

$$\begin{aligned} X_1 = & \frac{1}{2} \left(\left(1 + \frac{l(x_c, c) + \lambda}{r(x_c, c)} \right) \right. \\ & \left. + \sqrt{\left(\frac{l(x_c, c) + \lambda}{r(x_c, c)} - 1 \right)^2 + 2 \frac{\lambda}{r(x_c, c)}} \right). \end{aligned} \quad (5.25)$$

Thus $\gamma = \log(X_1)$ is given by,

$$\begin{aligned} \gamma = & \log \left(\frac{1}{2} \left(\left(1 + \frac{l(x_c, c) + \lambda}{r(x_c, c)} \right) \right. \right. \\ & \left. \left. + \sqrt{\left(\frac{l(x_c, c) + \lambda}{r(x_c, c)} - 1 \right)^2 + 2 \frac{\lambda}{r(x_c, c)}} \right) \right). \end{aligned} \quad (5.26)$$

When the proton binding and unbinding rates are much smaller than the protons endosomal entry time scale, $r(x, c) \gg \lambda$ and $l(x, c) \gg \lambda$, then

$$\gamma \approx \alpha = \log \left(\frac{l}{r}(x_c, c) \right). \quad (5.27)$$

We finally obtain

$$u(x_0(c)) = (1 + S(\epsilon, c)) \left(1 - e^{-\gamma \frac{x_c - x_0(c)}{\epsilon}} \right), \quad (5.28)$$

that is

$$u(x_0(c)) = \frac{\lambda}{\lambda + (C(\epsilon, c))^{-1}} \left(1 - \left(\frac{l}{r}(x_c, c) \right)^{-\frac{x_c - x_0(c)}{\epsilon}} \right). \quad (5.29)$$

In particular we obtain for $P_e^0(c) = 1 - u(x_0(c))$

$$P_e^0(c) = 1 - \frac{\lambda}{\lambda + (C(\epsilon, c))^{-1}} \left(1 - \left(\frac{l}{r}(x_c, c) \right)^{-\frac{x_c - x_0(c)}{\epsilon}} \right). \quad (5.30)$$

4.2 Computation of $\langle c_{\tau_e} \rangle$

The mean concentration $\langle c_{\tau_e} \rangle$ at which viral particles escape the endosome is

$$\langle c_{\tau_e} \rangle = \frac{\sum_{j=1}^{\infty} j P_e(j)}{\mathcal{N}V_0}, \quad (5.31)$$

that is

$$\langle c_{\tau_e} \rangle = \frac{1}{\mathcal{N}V_0} \sum_{j=1}^{\infty} j (1 - (1 - P_e^0(c(j)))^{n_v n_P}) \prod_{i=0}^{j-1} (1 - P_e^0(c(i)))^{n_v n_P}. \quad (5.32)$$

Thus,

$$\langle c_{\tau_e} \rangle = \frac{1}{\mathcal{N}V_0} \left(\sum_{j=0}^{\infty} (j+1) \prod_{i=0}^j (1 - P_e^0(c(i)))^{n_v n_P} - \sum_{j=1}^{\infty} j \prod_{i=0}^j (1 - P_e^0(c(i)))^{n_v n_P} \right). \quad (5.33)$$

Finally, we have

$$\langle c_{\tau_e} \rangle = \frac{1}{\mathcal{N}V_0} \sum_{j=0}^{\infty} \prod_{i=0}^j (1 - P_e^0(c(i)))^{n_v n_P}. \quad (5.34)$$

Using expression (5.30), in the approximation

$$\left(\frac{l}{r}(x_c, c) \right)^{-\frac{x_c - x_0(c)}{\epsilon}} \approx 0,$$

we get

$$\langle c_{\tau_e} \rangle = \frac{1}{\mathcal{N}V_0} \sum_{j=0}^{\infty} \prod_{i=0}^j \left(\frac{\lambda}{\lambda + (C(\epsilon, c(i)))^{-1}} \right)^{n_v n_P} \quad (5.35)$$

4.3 Computation of $cm(t)$

The independent events $A_k(t)$ that k protons have entered the endosome at time t follow a Poissonian law $P_k(t) = \frac{(\lambda t)^k}{k!} e^{-\lambda t}$, and

$$\begin{aligned} cm(t) &= \sum_{k=0}^{\infty} Pr\{\tau_e < t | A_k(t)\} P_k(t) \\ &= \sum_{k=0}^{\infty} \frac{(\lambda t)^k}{k!} e^{-\lambda t} (1 - Pr\{\tau_e > t | A_k(t)\}). \end{aligned} \quad (5.36)$$

To estimate $Pr\{\tau_e > t | A_k(t)\}$, we use the long time asymptotics (5.15) $\tilde{p}(x, t, c) \approx e^{-\frac{t}{C(\epsilon, c)}}$ and consider the random times $(T_i)_{i \geq 0}$ when protons enter the endosome. The joint pdf f of (T_1, T_2, \dots, T_k) of k protons ([104] theorem 2.3 page 126) is

$$f(t_1, t_2, \dots, t_k) = \frac{k!}{t^k} \mathbf{1}_{0 \leq t_1 < t_2 < \dots < t_k \leq t} \quad (5.37)$$

Because,

$$\begin{aligned} \tilde{p}^{n_v n_P}(x, t - t_k, c(k)) &\prod_{i=1}^{k-1} \tilde{p}^{n_v n_P}(x, t_{i+1} - t_i, c(i)) \\ &= \prod_{i=1}^{k-1} e^{-\lambda_i(t_{i+1} - t_i)} e^{-\lambda_k(t - t_k)}, \end{aligned} \quad (5.38)$$

where $\lambda_i = \frac{n_v n_P}{C(\epsilon, c(i))}$, we get that

$$\begin{aligned} Pr\{\tau_e > t | A_k(t)\} &= \frac{k!}{t^k} \int_0^t \int_{t_1}^t \dots \int_{t_{k-1}}^t e^{-\lambda_k(t - t_k)} \\ &\quad \prod_{i=1}^{k-1} e^{-\lambda_i(t_{i+1} - t_i)} dt_k \dots dt_1. \end{aligned} \quad (5.39)$$

We now compute $Pr\{\tau_e > t | A_k(t)\}$ by induction. We start with $Pr\{\tau_e > t | A_1(t)\} = \frac{1}{t} \int_0^t e^{-\lambda_1(t-t_1)} dt_1 = \frac{1}{t} \frac{1 - e^{-\lambda_1 t}}{\lambda_1}$. If

$$\begin{aligned} Pr\{\tau_e > t | A_k(t)\} &= \frac{(-1)^{k+1} k!}{t^k} \\ &\quad \sum_{i=1}^k \frac{1 - e^{-\lambda_i t}}{\lambda_i \prod_{j=1, j \neq i}^k (\lambda_i - \lambda_j)}. \end{aligned} \quad (5.40)$$

then $Pr\{\tau_e > t | A_{k+1}(t)\}$ is equal to

$$\begin{aligned} Pr\{\tau_e > t | A_{k+1}(t)\} &= \frac{(k+1)!}{t^{k+1}} \int_0^t \int_{t_1}^t \dots \\ &\quad \int_{t_k}^t e^{-\lambda_{k+1}(t - t_{k+1})} dt_{k+1} \\ &\quad \prod_{i=1}^k e^{-\lambda_i(t_{i+1} - t_i)} dt_k \dots dt_1, \end{aligned} \quad (5.41)$$

that is

$$\begin{aligned}
 Pr\{\tau_e > t | A_{k+1}(t)\} &= \frac{(k+1)!}{t^{k+1}} \int_0^t \int_{t_1}^t \dots \\
 &\quad \int_{t_{k-1}}^t \prod_{i=1}^{k-1} e^{-\lambda_i(t_{i+1}-t_i)} \\
 &\quad \int_{t_k}^t e^{-\lambda_k(t_{k+1}-t_k)} e^{-\lambda_{k+1}(t-t_{k+1})} dt_{k+1} dt_k \dots dt_1.
 \end{aligned} \tag{5.42}$$

Finally, we have

$$\begin{aligned}
 Pr\{\tau_e > t | A_{k+1}(t)\} &= \frac{(k+1)!}{t^{k+1}} \int_0^t \int_{t_1}^t \dots \\
 &\quad \int_{t_{k-1}}^t \prod_{i=1}^{k-1} e^{-\lambda_i(t_{i+1}-t_i)} \\
 &\quad \left(\frac{e^{-\lambda_k(t-t_k)} - e^{-\lambda_{k+1}(t-t_k)}}{\lambda_{k+1} - \lambda_k} \right) dt_k \dots dt_1.
 \end{aligned} \tag{5.43}$$

Using (5.41), we obtain

$$\begin{aligned}
 Pr\{\tau_e > t | A_{k+1}(t)\} &= \frac{(k+1)! (-1)^{k+1}}{t^{k+1} (\lambda_{k+1} - \lambda_k)} \\
 &\quad \left(\sum_{i=1}^k \frac{1 - e^{-\lambda_i t}}{\lambda_i \prod_{j=1, j \neq i}^k (\lambda_i - \lambda_j)} \right. \\
 &\quad \left. - \sum_{i=1}^{k-1} \frac{1 - e^{-\lambda_i t}}{\lambda_i \prod_{j=1, j \neq i}^{k-1} (\lambda_i - \lambda_j) (\lambda_i - \lambda_{k+1})} \right. \\
 &\quad \left. - \frac{1 - e^{-\lambda_{k+1} t}}{\lambda_{k+1} \prod_{j=1}^{k-1} (\lambda_{k+1} - \lambda_j)} \right),
 \end{aligned} \tag{5.44}$$

that is,

$$\begin{aligned}
 Pr\{\tau_e > t | A_{k+1}(t)\} &= \frac{(k+1)! (-1)^{k+1}}{t^{k+1} (\lambda_{k+1} - \lambda_k)} \\
 &\quad \left(\sum_{i=1}^{k-1} \frac{(1 - e^{-\lambda_i t}) (\lambda_k - \lambda_{k+1})}{\lambda_i (\lambda_i - \lambda_{k+1}) \prod_{j=1, j \neq i}^k (\lambda_i - \lambda_j)} \right. \\
 &\quad \left. + \frac{1 - e^{-\lambda_k t}}{\lambda_k \prod_{j=1}^{k-1} (\lambda_k - \lambda_j)} - \frac{1 - e^{-\lambda_{k+1} t}}{\lambda_{k+1} \prod_{j=1}^{k-1} (\lambda_{k+1} - \lambda_j)} \right).
 \end{aligned} \tag{5.45}$$

Thus, we obtain,

$$\begin{aligned}
 Pr\{\tau_e > t | A_{k+1}(t)\} &= \frac{(-1)^{k+2} (k+1)!}{t^{k+1}} \\
 &\quad \sum_{i=1}^{k+1} \frac{1 - e^{-\lambda_i t}}{\lambda_i \prod_{j=1, j \neq i}^{k+1} (\lambda_i - \lambda_j)}.
 \end{aligned} \tag{5.46}$$

Thus (5.41) holds for all $k \geq 1$. Using (5.41) in (5.37), we have

$$cm(t) = \sum_{k=1}^{\infty} \frac{(\lambda t)^k}{k!} e^{-\lambda t} \left(1 - \frac{(-1)^{k+1} k!}{t^k} \sum_{i=1}^k \frac{1 - e^{-\lambda_i t}}{\lambda_i \prod_{j=1, j \neq i}^k (\lambda_i - \lambda_j)} \right). \quad (5.47)$$

Because, $\sum_{k=0}^{\infty} \frac{(\lambda t)^k}{k!} e^{-\lambda t} = 1$, we get

$$cm(t) = 1 - e^{-\lambda t} + \sum_{k=1}^{\infty} (-\lambda)^k \sum_{i=1}^k \frac{e^{-\lambda t} - e^{-(\lambda + \lambda_i)t}}{\lambda_i \prod_{j=1, j \neq i}^k (\lambda_i - \lambda_j)}. \quad (5.48)$$

4.4 Computation of the mean escape time $\bar{\tau}_e$ and the associated variance $var(\tau_e)$

The mean time $\bar{\tau}_e = \int_0^{\infty} t p_{\tau_e}(t) dt = \int_0^{\infty} (1 - cm(t)) dt$ for a virus to escape is

$$\bar{\tau}_e = \frac{1}{\lambda} - \sum_{k=1}^{\infty} (-\lambda)^k \sum_{i=1}^k \frac{\int_0^{+\infty} e^{-\lambda t} - e^{-(\lambda + \lambda_i)t} dt}{\lambda_i \prod_{j=1, j \neq i}^k (\lambda_i - \lambda_j)}, \quad (5.49)$$

that is

$$\bar{\tau}_e = \frac{1}{\lambda} + \sum_{k=1}^{\infty} (-1)^{k+1} \lambda^{k-1} \sum_{i=1}^k \frac{1}{(\lambda + \lambda_i) \prod_{j=1, j \neq i}^k (\lambda_i - \lambda_j)}. \quad (5.50)$$

Using a fraction decomposition analysis (the λ_i are all distinct)

$$\frac{1}{\prod_{i=1}^k (\lambda + \lambda_i)} = \sum_{i=1}^k \frac{(-1)^{k+1}}{(\lambda + \lambda_i) \prod_{j=1, j \neq i}^k (\lambda_i - \lambda_j)}, \quad (5.51)$$

we obtain

$$\begin{aligned} \bar{\tau}_e &= \frac{1}{\lambda} + \sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{\prod_{i=1}^k (\lambda + \lambda_i)} \\ &= \frac{1}{\lambda} \left(1 + \sum_{k=1}^{\infty} \frac{1}{\prod_{i=1}^k (1 + \lambda_i/\lambda)} \right). \end{aligned} \quad (5.52)$$

The variance $var(\tau_e) = \int_0^{\infty} t^2 p_{\tau_e}(t) dt - \bar{\tau}_e^2 = 2 \int_0^{\infty} t (1 - cm(t)) dt - \bar{\tau}_e^2$ is

$$\begin{aligned} var(\tau_e) &= \frac{2}{\lambda^2} - 2 \sum_{k=1}^{\infty} (-\lambda)^k \sum_{i=1}^k \frac{\int_0^{+\infty} t e^{-\lambda t} - t e^{-(\lambda + \lambda_i)t} dt}{\lambda_i \prod_{j=1, j \neq i}^k (\lambda_i - \lambda_j)} - \bar{\tau}_e^2 \end{aligned} \quad (5.53)$$

that is

$$\begin{aligned} \text{var}(\tau_e) &= \frac{2}{\lambda^2} + 2 \sum_{k=1}^{\infty} (-1)^{k+1} \lambda^k \\ &\quad \sum_{i=1}^k \frac{1/\lambda^2 - 1/(\lambda + \lambda_i)^2}{\lambda_i \prod_{j=1, j \neq i}^k (\lambda_i - \lambda_j)} - \bar{\tau}_e^2. \end{aligned} \quad (5.54)$$

Thus, we have

$$\begin{aligned} \text{var}(\tau_e) &= \frac{2}{\lambda^2} + 2 \sum_{k=1}^{\infty} (-1)^{k+1} \lambda^{k-2} \\ &\quad \sum_{i=1}^k \frac{2\lambda + \lambda_i}{(\lambda + \lambda_i)^2 \prod_{j=1, j \neq i}^k (\lambda_i - \lambda_j)} \\ &\quad - \bar{\tau}_e^2. \end{aligned} \quad (5.55)$$

The partial fraction decomposition of $\frac{-\lambda}{\prod_{i=1}^k (\lambda + \lambda_i)}$ gives $\sum_{i=1}^k \frac{(-1)^{k+1} \lambda_i}{(\lambda + \lambda_i) \prod_{j=1, j \neq i}^k (\lambda_i - \lambda_j)}$. Because

$$\sum_{i=1}^k \frac{1}{(\lambda + \lambda_i)^2 \prod_{j=1, j \neq i}^k (\lambda_i - \lambda_j)}$$

is the derivative of

$$\sum_{i=1}^k \frac{-1}{(\lambda + \lambda_i) \prod_{j=1, j \neq i}^k (\lambda_i - \lambda_j)}$$

we have

$$\begin{aligned} \text{var}(\tau_e) &= \frac{2}{\lambda^2} + 2 \sum_{k=1}^{\infty} \lambda^{k-2} \left(2\lambda \left(\frac{-1}{\prod_{i=1}^k (\lambda + \lambda_i)} \right)' \right. \\ &\quad \left. + \left(\frac{\lambda}{\prod_{i=1}^k (\lambda + \lambda_i)} \right)' \right) - \bar{\tau}_e^2. \end{aligned} \quad (5.56)$$

A mathematical induction yields

$$\left(\prod_{i=1}^k (\lambda + \lambda_i) \right)' = \sum_{j=1}^k \prod_{i=1, i \neq j}^k (\lambda + \lambda_i). \quad (5.57)$$

Consequently, we have

$$\begin{aligned} \text{var}(\tau_e) &= \frac{2}{\lambda^2} + 2 \sum_{k=1}^{\infty} \lambda^{k-2} \frac{1}{\prod_{i=1}^k (\lambda + \lambda_i)} \\ &\quad + \frac{\lambda \sum_{j=1}^k \prod_{i=1, i \neq j}^k (\lambda + \lambda_i)}{\left(\prod_{i=1}^k (\lambda + \lambda_i) \right)^2} - \bar{\tau}_e^2, \end{aligned} \quad (5.58)$$

that is

$$\begin{aligned}
 \text{var}(\tau_e) = & \frac{2}{\lambda^2} \left(1 + \sum_{k=1}^{\infty} \frac{1}{\prod_{i=1}^k (1 + \lambda_i/\lambda)} \right. \\
 & \left. + \frac{\sum_{j=1}^k \prod_{i=1, i \neq j}^k (1 + \lambda_i/\lambda)}{\left(\prod_{i=1}^k (1 + \lambda_i/\lambda) \right)^2} \right) - \bar{\tau}_e^2.
 \end{aligned} \tag{5.59}$$

1) Recapitulate the key steps of viral infection.

2) State for which type of virus your work applies, and give main results, approximations and mathematical tools used, and maybe add in which chapter the results can be found.

3) then give an outlook. Discuss the most important approximations (biological and mathematical) you made and how they can be improved. And then show how your model can be extended to other virus types.

Conclusion

The outcome whether the entry of a viral particle into a cell leads to an infection or not depends on various factors. In particular, most DNA viruses have to escape the endosome in a certain pH range and once they are released in the cytoplasm, they have to reach a nuclear pore before being trapped in the crowded cytoplasm or degraded through the ubiquitin-proteasome machinery. While DNA viruses have developed evolutionary tools to reliably escape the endosome and hijack the cell transport machinery to efficiently reach the nucleus [5], non viral genes vectors often fail to escape the endosome before being routed to the degradative lysosomes, and the cytosolic motion of pure diffusive DNA complexes is limited by physical and chemical barriers of the cytoplasm [3, 4]. Yet the principles underlying the viral tools are still largely unknown. This lack of knowledge is also the reason the efficiency of infection with non-viral gene vectors is still very low. To unravel and better understand the mechanisms underlying viral infection, we developed in this work mathematical models for the endosomal step and the free cytoplasmic step, which are both key early steps of infection. Early models describing viral infection [6, 7, 8, 48] divide the infection process into several states, eg. endosomal state, diffusing state . . . and use the mass-action law to describe the evolution in time of the viral state. The main limitation of these models is that the transition rates between the different states of the virus are mostly fitted to experimental data instead of being derived from biophysical models at the molecular level. To quantify how cell and viruses parameters such as the endosome size, the MTs network organization or the diffusion constant of the virus impact general quantitative outputs such as the endosomal escape kinetics or the probability the virus reaches alive a nuclear pore, we constructed stochastic models starting from the single molecule level. We summarize these models hereafter.

In the fifth chapter, we constructed a biophysical model for the endosomal escape of viruses coated by few active proteins (glycoproteins for enveloped viruses and penetration proteins for non-enveloped ones), and we computed relative key quantitative outputs such as the mean escape time from the endosome. The viral escape is triggered by the conformational change of active proteins that initiate the membrane disruption. Consequently, to analyze the escape process, we first modeled the conformational change kinetics of the active proteins by computing the mean time the number of ligands bound to the active proteins reaches a given threshold and triggers the conformational change. To model the probability $p(x, y, c, t)$, that the number of bound ligands is equal to x at time t for a concentration c of ligands and an initial number of bound ligands equal to y at time $t = 0$, we started using Markov jump process and showed that $p(x, y, c, t)$ is solution of a Kramers-Moyal equation. Using asymptotic analysis, we computed the conformational change

kinetics of the active protein as a function of the ligand concentration. Our theoretical results agree with the reported kinetics for the influenza hemagglutinin, which confirms that only the rearrangements in the HA1 subunit of the hemagglutinin are pH dependents [35, 37]. We considered that the conformational change of a single active protein was the limiting event to the endosomal escape of viruses coated by few active proteins such as parvoviruses. Then, modeling the ligands influx rate with a Poissonian process, we computed the mean time and pH the virus escapes the endosome. Interestingly, we found that the endosomal size had no effect on the escape dynamics and that the optimal number of viruses in the endosome to escape in the right pH range was 5.

Concerning the free cytoplasmic step of the DNA viruses after their release in the cytoplasm, we described their intermittent trajectories, oscillating between diffusion and active motion along MTs, with a Langevin equation containing a diffusion term and a drift accounting for the active transport periods. The first two chapters were dedicated to calibrate the drift amplitude with respect to the MTs network organization and some dynamical parameters of the virus such as its diffusion constant and its unbinding rate from MTs. Then, using this Langevin stochastic description, we computed the probability the virus reaches a nuclear pore before being trapped or degraded, and the associated conditional mean time. Because the pores occupy a small fraction of the nuclear envelope, these computations are based on the narrow escape theory [30] and we studied in the third chapter the impact of the pores distribution on both the probability and the MFPT. Interestingly a key parameter in the gene delivery dynamics is the number of nuclear pores rather than the surface covered by the nuclear pores : The mean arrival time to a single big absorbing pore that covers 2% of the nuclear surface is twice as large as the mean arrival time to one over 100 pores that cover the same surface (see chapter 3). Finally, in the fourth chapter we computed the MFPT of the first virus when many viruses traffics in the cytoplasm and we extended previous results in the large degradation rate limit.

Viral infection is an extremely diversified and complex process, and to obtain first and tractable molecular models of this process we had to introduce many biophysical approximations that can be further improved in future work. For example, a key approximation in the endosomal step was the Poissonian description of the protons influx and a better knowledge of the proton channel dynamics would certainly reduce the error of this approximation, although we can consider that channels have no memory. We also assumed the conformational change of the active protein was the limiting event in the escape process, and future models shall account for the other steps participating to that process, particularly the pore formation limiting step, involved in membrane disruption [89]. Concerning the free cytoplasmic step, we assumed the free diffusion step to be Brownian motion. Yet, the cell cytoplasm is very crowded and a fractional subdiffusive type process leading to fractional diffusion effects should be considered in a future analysis.

To get concise and readable formulas that link the different key parameters of the infectious process, we also made mathematical approximations by solving asymptotically most of the equations, that is we provided the first leading order term of solutions in a certain small parameter. For example, in the endosomal step, we solved the Kramers-Moyal equation using the WKB theory with the small parameter $\epsilon = 1/n_s$ where $n_s \gg 1$ is the number of sites of the active protein that bind ligands,

while in the free cytoplasmic step, the nuclear pores occupying a small fraction of the nuclear envelope, we used the narrow escape theory to get concise results for the probability and the mean time the DNA viruses reach one of these small absorbing pores. In a future work, it should be interesting to pursue the asymptotic expansion of the principal quantitative outputs by computing lower order terms. In particular, it should be interesting to quantify how the conditional MFPT to a small absorbing pore depends on the pores organization over the nucleus surface by expanding in ϵ the linear system (3.32) in the chapter 3. Finally, the mathematical analysis of the free trafficking step is based on the continuous Langevin description of trajectories and the effective drift calibration in a simplified two-dimensional radial geometry. Although the radial geometry is a good approximation of flat culture cells, the calibration computations strongly rely on the conformal mapping of a wedge domain into the upper half plane and it is not clear how to generalize these computations in a three-dimensional cell. Consequently, a new calibration methodology for three dimensional domains is required and should be investigated carefully.

The biophysical models we constructed concern a certain class of viruses : the endosomal escape model applies for viruses coated by few active proteins such as parvoviruses and the free trafficking model deals with DNA viruses that have to reach a nuclear pore to deliver their genome into the nucleus. In a future work, these models shall be extended to other viral species and new models shall be developed for the other steps of infection such as the nuclear import of the genome. In particular it will be interesting to analyze the endosomal escape of enveloped viruses coated by many active proteins such as the influenza (400 HA copies [100]) where the binding of ligands depletes significantly the free ligands pool in the endosome, and where the conformational change of glycoproteins initiate complex mechanisms, at the origin of the local fusion between the virus and the endosome membranes. For the influenza virus, three intermediate and limiting steps have been recognized and shall be integrated in the modeling : (1) the conformational change of HA trimers and their clustering (2) the formation of a hemi-fusion stalk (3) the final fusion pore formation [97].

During the infectious process, the cellular environment triggers molecular rearrangements of the viral capsid or envelope that in turn modify the viral behavior. For example, the capsid denaturation of the AAV capsid in the endosome increases its cytoplasmic ubiquitination [11] that competitively triggers its proteasome-mediated degradation and helps capsid disassembly and subsequent nuclear import [11]. It is not clear how to integrate this complex host-cell communication in quantifying the infection success, but the models we developed for each early step of infection shall be coupled, the state of the virus computed at the end of a step, such as its capsid denaturation when it escapes the endosome, serving as an input parameter for the next step. Experimental data on the impact of the viral state on its dynamical parameters are thus needed and, in particular, it will be interesting to quantify how the viral capsid denaturation influences its degradation rate in the cytoplasm. Finally, future models will have to integrate the cell specificity through its complex geometry and environment. In particular, by controlling the level of tau proteins, the cell regulates the binding rate of molecular motors to the MTs and thus the velocity of the transported viruses [105].

Appendix : Brownian simulations

We performed the Brownian simulations using Scilab. Hereafter, we present the code used in the chapter 3 to compute the mean first passage time of a Brownian particle to an absorbing pore when many pores are uniformly distributed over the spherical surface of the nucleus. In a first time, we tested the convergence of the algorithm for decreasing time steps and an increasing number of Brownian trajectories and we finally calibrated the time step dt such that $\sqrt{Ddt} = 5\%\epsilon$ where D is the diffusion constant of the virus and ϵ the radius of a nuclear pore, and we ran $k = 1,000$ Brownian simulations. We used the following rule to test the convergence of the algorithm : we chose an initial time step $dt_0 = (\epsilon/D)^2$ and we initially ran $k_0 = 100$ Brownian simulations. We then computed the mean arrival time τ_n for $k_0 = 100$ Brownian trajectories and a decreasing time step $dt = dt_0, dt = \frac{dt_0}{2}, dt = \frac{dt_0}{5}, dt = \frac{dt_0}{10}$ and $dt = \frac{dt_0}{20}$. Because $\frac{|\tau_n(dt=\frac{dt_0}{10})-\tau_n(dt=\frac{dt_0}{20})|}{\tau_n(dt=\frac{dt_0}{20})} > 1\%$, we then ran 1,000 Brownian simulations and our convergence rule was then verified : $\frac{|\tau_n(dt=\frac{dt_0}{10})-\tau_n(dt=\frac{dt_0}{20})|}{\tau_n(dt=\frac{dt_0}{20})} < 1\%$.

Scilab code

```
pi = %pi ;
//parameters
//cell radius
r_c = 15;
//viral diffusion constant
d_v = 1.3;
//nuclear radius
r_n = 5;
//number of Brownian trajectories
n_v = 1000;
//Constant drift
B=0.2;
//number of absorbing pores (table)
nb_trou_tab=[1,2,3,5,10,20,50,100,300];
//size of the table above
temp=size(nb_trou_tab);
ite=temp(1,2);
//degradation rate
k=10/3600;
```

```
//initialisation
//number of killed viruses
tue_drift=zeros(1,ite);
//conditional MFPT to an absorbing pore
temps_drift=zeros(1,ite);
//theoretical MFPT without interaction
theo_temps=zeros(1,ite);
//theoretical MFPT with interaction
theo_tempsbis=zeros(1,ite);
//theoretical probability without interaction
theo_proba=zeros(1,ite);
//theoretical probability with interaction
theo_probabis=zeros(1,ite);

//loop
for l=1:1:ite,

//number of pores
nb_trou=nb_trou_tab(1,l);
//radius of a pore (computed such that the surface covered by the pores=$2\%$ of the r
epsilon = 4*sqrt(0.02/(nb_trou));

dt0=((1/d_v)*(0.05*r_n*epsilon)^2);

//the pores are uniformly distributed over the nuclear surface
test=1;
while (test>0) do,
phis_trou=pi.*rand(1,nb_trou,'uniform');
thetas_trou=-alpha/2.*ones(1,nb_trou) + alpha.*rand(1,nb_trou,'uniform');
distances=100.*ones(nb_trou,nb_trou);
x_trou=r_n*sin(phis_trou).*cos(thetas_trou);
y_trou=r_n*sin(phis_trou).*sin(thetas_trou);
z_trou=r_n*cos(phis_trou);
for i=1:1:nb_trou,
for j=i+1:1:nb_trou,
distances(i,j)=sqrt((x_trou(1,i)-x_trou(1,j))^2+(y_trou(1,i)-y_trou(1,j))^2+(z_trou(1,i)-z_trou(1,j))^2);
distances(j,i)=distances(i,j);
end;
end;
temp1=bool2s(distances<r_n*epsilon/2);
test=sum(temp1);
end;

//initial positions of the viruses (uniformly distributed over the cell surface)
```

```

theta =-pi.*ones(1,n_v) + 2*pi.*rand(1,n_v,'uniform');
phi = pi.*rand(1,n_v,'uniform');
r=r_c.*ones(1,n_v);
dt=dt0.*ones(1,n_v);
t=dt;

//killed viruses counter
compteur_tue=0;

//initialisation of MFPT
time=0;

//loop
taille = n_v;
while (taille>=1) do
//increment
temp2=size(r);
taille=temp2(1,2);

//diffusion...
dx = sqrt(2*d_v.*dt).*rand(1,taille,'normal');
dy = sqrt(2*d_v.*dt).*rand(1,taille,'normal');
dz = sqrt(2*d_v.*dt).*rand(1,taille,'normal');

temp_x = dy + (r.*sin(phi)).*cos(theta);
temp_y = dx + (r.*sin(phi)).*sin(theta);
temp_z = dz + r.*cos(phi);

r = sqrt(temp_x.^2+temp_y.^2+temp_z.^2);
theta = atan(temp_y,temp_x);
phi = acos(temp_z./r);

//+drift
drift = (B.*dt).*ones(1,taille);
r=r-drift;

//External reflecting condition (r=R)
correction0 = zeros(1,taille);
correction0 = (r-r_c.*ones(1,taille));
correction1 = bool2s(correction0>0.0).*correction0;
correction0 = 2.*correction1;
r = r-correction0;

correction4=ones(1,taille);
for j=1:1:nb_trou,
phi_trou=phis_trou(1,j);

```

```
theta_trou=thetas_trou(1,j);

//Reflecting condition on the nuclear surface outside the absorbing pores
correction2 = zeros(1,taille);
correction2=(sin(phi_trou)*cos(theta_trou)).*sin(phi).*cos(theta)+(sin(phi_trou)*sin(theta).*cos(phi_trou)*sin(theta));
correction4 = correction4.*bool2s(correction2<0.0);
end;
correction7 = correction4.*r;
correction8 = r_n.*correction4;
correction9 = correction8-correction7;
correction10 = bool2s(correction9>0.0).*correction9;
r = r+2.*correction10;

//MFPT
temp_time=zeros(1,taille);
temp_time=r-r_n.*ones(1,taille);
temp_time_bis=bool2s(temp_time<0.0);
time=time+t*temp_time_bis';

//we remove the absorbed viruses
pluto = r;
r = r(pluto>r_n);
theta = theta(pluto>r_n);
phi = phi(pluto>r_n);
t=t(pluto>r_n);
dt=dt(pluto>r_n);

//increment
temp2=size(r);
taille=temp2(1,2);

t=t+dt;

//killing rate
kill=k.*dt;
//some viruses are degraded..
killing = rand(1,taille,'uniform');
temp=taille;
if taille>0 then,
r = r(killing>kill);
theta = theta(killing>kill);
phi = phi(killing>kill);
t=t(killing>kill);
dt=dt(killing>kill);

//increment
temp2=size(r);
```

```

taille=temp2(1,2);
end;
//total number of killed viruses
compteur_tue=compteur_tue+temp-taille;
end;

//simulated conditional MFPT
temps_drift(1,1)=time/(n_v-compteur_tue);
//simulated probability
tue_drift(1,1) = compteur_tue/(n_v);

eta=epsilon/2*r_n*nb_trou;
vol=(4/3)*pi*(r_c^3-r_n^3);

temp=(pi/(d_v*eta)).*(exp(-(r_n/d_v).*B).*((d_v*r_n^2)./B+(2*d_v^2*r_n)./(B.^2))+
temp1=exp(-(r_n/d_v).*B);
temp2=(pi/(d_v*eta)+pi/(d_v*pi*r_n)).*(exp(-(r_n/d_v).*B).*((d_v*r_n^2)./B+(2*d_v
tempsss1=temp./temp1);
tempsss2=temp2./temp1;

//theoretical values
theo_proba(1,1)=1/(1+k*tempsss1);
theo_probabis(1,1)=1/(1+k*tempsss2);
theo_temps(1,1)=theo_proba(1,1)*tempsss1;
theo_tempsbis(1,1)=theo_probabis(1,1)*tempsss2;
end;

```


Bibliographie

- [1] Yan Z Engelhardt JF Ding W, Zhang L. Intracellular trafficking of adeno-associated viral vectors. *Gene Ther*, 12 :873–80, 2005. Review.
- [2] Kalaidzidis Y Zerial M Rink J, Ghigo E. Rab conversion as a mechanism of progression from early to late endosomes. *Cell*, 122 :735–49, 2005.
- [3] Verkman AS. Solute and macromolecule diffusion in cellular aqueous compartments. *Trends Biochem Sci*, 27 :27–33, 2002. Review.
- [4] Verkman AS Dauty E. 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 :7823–8, 2005.
- [5] Sodeik B. Mechanisms of viral transport in the cytoplasm. *Trends Microbiol*, 8 :465–72, 2000. Review.
- [6] Shuler ML Dee K. Optimization of an assay for baculovirus titer and design of regimens for the synchronous infection of insect cells. *Biotechnol Prog*, 13 :14–24, 1997.
- [7] Lauffenburger DA Varga CM, Hong K. Quantitative analysis of synthetic gene delivery vector design properties. *Mol Ther*, 5 :438–46, 2001.
- [8] Chen R Giorgio TD Banks GA, Roselli RJ. A model for the analysis of nonviral gene therapy. *Gene Ther*, 20 :1766–75, 2003.
- [9] Lamb RA Dutch RE, Jardetzky TS. Virus membrane fusion proteins : biological machines that undergo a metamorphosis. *Biosci Rep*, 20 :597–612, 2000.
- [10] Mothes W et al. Retroviral entry mediated by receptor priming and low ph triggering of an envelope glycoprotein. *Cell*, 103 :679–89, 2000.
- [11] Luxton GW Ritchie TC Bantel-Schaal U Engelhardt JF Yan Z, Zak R. Ubiquitination of both adeno-associated virus type 2 and 5 capsid proteins affects the transduction efficiency of recombinant vectors. *J Virol*, 76 :2043–53, 2002.
- [12] Frangakis AS Nicastro D Gerisch G Baumeister W Medalia O, Weber I. Macromolecular architecture in eukaryotic cells visualized by cryoelectron tomography. *Science*, 298 :1209–13, 2002.
- [13] Verkman AS Seksek O, Biwersi J. Translational diffusion of macromolecule-sized solutes in cytoplasm and nucleus. *J Cell Biol*, 138 :131–42, 1997.

- [14] Endress T Bøning H Hallek M BrLuchle C Seisenberger G, Ried MU. Real-time single-molecule imaging of the infection pathway of an adeno-associated virus. *Science*, 294 :1929–32, 2001.
- [15] Kim KA Miko S Perret E Olivo-Marin JC Shorte S Charneau P Arhel N, Genovesio A. Quantitative four-dimensional tracking of cytoplasmic and nuclear hiv-1 complexes. *Nat Methods*, 10 :817–24, 2006.
- [16] Zhuang X Brandenburg B. Virus trafficking - learning from single-virus tracking. *Nat Rev Microbiol*, 5 :197–208, 2007.
- [17] Enquist LW Smith GA, Gross SP. Herpes viruses use bidirectional fast-axonal transport to spread in sensory neurons. *Proc Natl Acad Sci U S A*, 98 :3466–70, 2001.
- [18] Keller S Boucke K Stidwill RP Greber UF Suomalainen M, Nakano MY. Mt-dependent plus- and minus end-directed motilities are competing processes for nuclear targeting of adenovirus. *J Cell Biol*, 144 :657–72, 1999.
- [19] Lucero G Svitkina TM Borisy GG Emerman M Hope TJ McDonald D, Voldicka MA. Visualization of the intracellular behavior of hiv in living cells. *J Cell Biol*, 159 :441–52, 2002.
- [20] Gross SP. Hither and yon : a review of bi-directional mt-based transport. *Phys Biol*, 1 :R1–11, 2004. Review.
- [21] Schuss Z. *Theory and Applications of Stochastic Differential Equations*. John Wiley & Sons Inc, New-York, 1980.
- [22] Berg HC. *Random walks in Biology*. Princeton University Press, 1983.
- [23] Gross SP Enquist LW Smith GA, Pomeranz L. Local modulation of plus-end transport targets herpesvirus entry and egress in sensory axons. *Proc Natl Acad Sci U S A*, 101 :16034–9, 2004.
- [24] Ajdari A. Transport by active filaments. *Europhys Lett*, 31 :341, 1995.
- [25] Nieuwenhuizen TM Lipowsky R, Klumpp S. Random walks of cytoskeletal motors in open and closed compartments. *Phys Rev Lett*, 87 :108101, 2001.
- [26] Holcman D. Modeling trafficking of a virus and a dna particle in the cell cytoplasm. *J of Statistical Physics*, 127 :471–94, 2007.
- [27] Schuss Z Holcman D. Diffusion escape through a cluster of small absorbing windows. *J Phys A : Math Theor*, 41 :155001–15, 2008.
- [28] Ward MJ Coombs D, Straube R. Diffusion on a sphere with localized traps : Mean first passage time, eigenvalue asymptotics, and fekete points. *SIAM J Appl Math*, 70 :302, 2009.
- [29] Holcman D Reingruber J, Abad E. Narrow escape time to a structured target located on the boundary of a microdomain. *J Chem Phys*, 130 :094909, 2009.

- [30] Holcman D Schuss Z, Singer A. The narrow escape problem for diffusion in cellular microdomains. *Proc Natl Acad Sci U S A*, 104 :16098–103, 2007.
- [31] Stegun IA Abramowitz M. *Handbook of Mathematical Functions*. Dover, New York, 1972.
- [32] Lechardeur D et al. Metabolic instability of plasmid dna in the cytosol : a potential barrier to gene transfer. *Gene Therapy*, 6 :482–97, 1999.
- [33] Knessl C et al. Solution of kramers-moyal equations for problems in chemical physics. *J Chem Phys*, 81 :1285–93, 1984.
- [34] Matkowsky BJ et al. Asymptotic solution of the kramers-moyal equation and first-passage times for markov jump processes. *Phys Rev A*, 29 :3359–69, 1984.
- [35] Knapp EW Herrmann A Huang Q, Opitz R. Asymptotic solution of the kramers-moyal equation and first-passage times for markov jump processes. *Biophys J*, 82 :1050–58, 2002.
- [36] Blumenthal R Krumbiegel M, Herrmann A. Kinetics of the low ph-induced conformational changes and fusogenic activity of influenza hemagglutinin. *Biophys J*, 67 :2355–60, 1994.
- [37] Huang Q et al. Energetics of the loop-to-helix transition leading to the coiled-coil structure of influenza virus hemagglutinin ha2 subunits. *Proteins*, 74 :291–303, 2009.
- [38] Hoggan MD Johnson FB, Ozer HL. Structural proteins of adenovirus-associated virus type 3. *J Virol*, 8 :860–3, 1971.
- [39] Van Kampen NG. *Stochastic Processes in Physics and Chemistry*. North-Holland, Amsterdam, Netherlands, 1992.
- [40] Tattersall P Farr GA, Zhang LG. Parvoviral virions deploy a capsid-tethered lipolytic enzyme to breach the endosomal membrane during cell entry. *Proc Natl Acad Sci U S A*, 102 :17148–53, 2005.
- [41] Zhuang X Brandenburg B. Virus trafficking & learning from single-virus tracking. *Nat Rev Microbiology*, 5 :197–208, 2007. Review.
- [42] Helenius A Pelkmans L, Kartenbeck J. Caveolar endocytosis of simian virus 40 reveals a new two-step vesicular-transport pathway to the er. *Nature Cell Biol*, 3 :473–83, 2001.
- [43] Verma IM. A tumultuous year for gene therapy. *Mol Ther*, 2 :415–26, 2000.
- [44] Gansbacher B. Report of a second serious adverse event in a clinical trial of gene therapy for x-linked severe combined immune deficiency (x-scid). position of the european society of gene therapy (esgt). *J Gene Med*, 5 :261–2, 2003.
- [45] Tada H Iwabuki H Chuah MK VandenDriessche T et al Yamada T, Iwasaki Y. Nanoparticles for the delivery of genes and drugs to human hepatocytes. *Nat Biotechnol*, 21 :885–90, 2003.

- [46] Berger JM Kim PS Chan DC, Fass D. Core structure of gp41 from the hiv envelope glycoprotein. *Cell*, 89 :263D73, 1997.
- [47] Brandenburg B. *Cell permeable nucleocapsids as a novel tool for efficient gene transfer and HBV biology*. PhD thesis, Freie Universitat Berlin, 2005.
- [48] Mitragotri S Dinh AT, Theofanous T. A model for intracellular trafficking of adenoviral vectors. *Biophysical J*, 89 :1574–88, 2005.
- [49] Bae HR Verkman AS Shi LB, Fushimi K. Heterogeneity in atp-dependent acidification in endocytic vesicles from kidney proximal tubule. measurement of ph in individual endocytic vesicles in a cell-free system. *Biophys J*, 59 :1208–17, 1991.
- [50] Middaugh CR Wiethoff CM. Barriers to non-viral gene delivery. *Journal of Pharmaceutical Sciences*, 92 :203–17, 2003.
- [51] Lewis J Raff M Roberts K Walter P Alberts B, Johnson A. *Molecular Biology of the Cell*. 4th Edition, Garland, New-York, 2002.
- [52] Hirokawa N. Kinesin and dynein superfamily proteins and the mechanism of organelle transport. *Science*, 279 :519–26, 1998.
- [53] Mallick R. Cytoplasmic dynein functions as a gear in response to load. *Nature*, 427 :649–52, 2004.
- [54] Redner S. *A Guide to First Passage Processes*. Cambridge University Press, Cambridge, Massachussets, 2001.
- [55] Henrici P. *Applied and Computational Complex Analysis. Vol. 3*. John Wiley & Sons Inc, New-York, 1977.
- [56] Jaegger JC Carslaw HS. *Conduction of Heat in Solids*. Oxford University Press, Oxford, U.K., 1959.
- [57] Schroer TA King SJ. Dynactin increases the processivity of the cytoplasmic dynein motor. *Nat Cell Biol*, 2 :20–4, 2000.
- [58] Sodeik B Katinka D, Claus-Henning N. Viral stop-and-go along microtubules : taking a ride with dynein and kinesins. *Trends Microbiol*, 13 :320–7, 2005.
- [59] Block SM Wieschaus EF Gross SP, Welte MA. Dynein-mediated cargo transport in vivo : a switch controls travel distance. *J Cell Biol*, 5 :945–55, 2000.
- [60] Whittaker GR. Virus nuclear import. *Adv Drug Deliv Rev*, 55 :733–47, 2003.
- [61] Zhou R Dean DA, Geiger RC. Intracellular trafficking of nucleic acids. *Expert Opinion Drug Delivery*, 1 :127–40, 2004.
- [62] Hope TJ Campbell EM. Gene therapy progress and prospects : viral trafficking during infection. *Gene Therapy*, 12 :1353–9, 2005.

- [63] Saltzman WM Luo D. Synthetic dna delivery systems. *Nat Biotech*, 18 :33–7, 1999.
- [64] Way M Greber UF. A superhighway to virus infection. *Cell*, 124 :741–54, 2006.
- [65] Maggs AC Nedelec F, Surrey T. Dynamic concentration of motors in microtubule arrays. *Phys Rev Lett*, 86 :3192–5, 2001.
- [66] Lipowsky R Klumpp S, Nieuwenhuizen TM. Self-organized density patterns of molecular motors in arrays of cytoskeletal filaments. *Biophys J*, 88 :3118–32, 2005.
- [67] Tejedor V Voituriez R Klafter J Condamin S, Bénichou O. First-passage times in complex scale-invariant media. *Nature*, 450 :77–80, 2007.
- [68] Schuss Z Holcman D, Marchewka A. Survival probability of diffusion with trapping in cellular neurobiology. *Phys Rev E*, 72 :031910, 2005.
- [69] Howard JJ Coy D, Wagenbach M. Kinesin takes one 8-nm step for each atp that it hydrolyzes. *J Biol Chem*, 274 :3667–71, 1999.
- [70] Deaven L Maul GG. Quantitative determination of nuclear pore complexes in cycling cells with differing dna content. *J Cell Biol*, 73 :748–60, 1977.
- [71] Ozawa S. Surgery of domain and the green’s function of the laplacian. *Proc Japan Acad*, 56 :459–461, 1980.
- [72] Nothisen M Belguise P Behr JP Zuber G, Dauty E. Towards synthetic viruses. *Adv Drug Deliv Rev*, 52 :245–53, 2001.
- [73] Steward O Dynes JL. Dynamics of bidirectional transport of arc mrna in neuronal dendrites. *J Comp Neurol*, 500 :433–447, 2007.
- [74] Newby J Bressloff PC. Directed intermittent search for hidden targets. *New J Phys*, 11 :023033, 2009.
- [75] Schuss Z Holcman D. Stochastic chemical reactions in microdomains. *J Chem Phys*, 122 :114710, 2005.
- [76] Purcell M Berg HC. Physics of chemoreception. *Biophys J*, 20 :193–219, 1977.
- [77] Zwanzig R. Diffusion-controlled ligand binding to spheres partially covered by receptors : an effective medium treatment. *Proc Natl Acad Sci U.S.A.*, 87 :5856–7, 1990.
- [78] Monine MI Zitserman VY Shvartsman SY Berezhkovskii AM, Makhnovskii YA. Boundary homogenization for trapping by patchy surfaces. *J Chem Phys*, 121 :11390–4, 2004.
- [79] Schuss Z Singer A. Activation through a narrow opening. *SIAM J Appl Math*, 68 :98, 2007.

- [80] von Helmholtz HLF. Theorie der luftschwingungen in röhren mit offenen enden. *Crelle*, 57 :1–72, 1860.
- [81] Jackson JD. *Classical Electrodynamics*. John Wiley & Sons Inc, New-York, 1975.
- [82] Han Y et al. Cultivation of recombinant chinese hamster ovary cells grown as suspended aggregates in stirred vessels. *J Biosci Bioeng*, 102 :430–5, 2006.
- [83] Theofanous T Mitragotri S Dinh AT, Pangarkar C. Understanding intracellular transport processes pertinent to synthetic gene delivery via stochastic simulations and sensitivity analyses. *Biophys J*, 92 :831–46, 2007.
- [84] Keller JB Ward MJ. Strong localized perturbations of eigenvalue problems. *SIAM J of Appl Math*, 53 :770, 1993.
- [85] Berezhkovskii AM Zitserman VY Grigoriev IV, Makhnovskii YA. Kinetics of escape through a small hole. *J Chem Phys*, 116 :9574, 2002.
- [86] Welte MA. Bidirectional transport along microtubules. *Curr Biol*, 14 :R525–37, 2004.
- [87] Suomalainen M et al. Microtubule-dependent plus- and minus end-directed motilities are competing processes for nuclear targeting of adenovirus. *J Cell Biol*, 144 :657–72, 1999.
- [88] Schuss Z Matkowski BJ. Diffusion across characteristic boundaries. *SIAM J of Appl Math*, 42 :822, 1982.
- [89] Nibert ML Agosto MA, Ivanovic T. Mammalian reovirus, a nonfusogenic nonenveloped virus, forms size-selective pores in a model membrane. *Proc Natl Acad Sci U S A*, 103 :16496–501, 2006.
- [90] Wiley DC Skehel JJ. Receptor binding and membrane fusion in virus entry : the influenza hemagglutinin. *Annu Rev Biochem*, 69 :531–69, 2000.
- [91] Allison SL Heinz FX. The machinery for flavivirus fusion with host cell membranes. *Curr Opin Microbiol*, 4 :450–5, 2001.
- [92] Chandran K et al. Endosomal proteolysis of the ebola virus glycoprotein is necessary for infection. *Science*, 308 :1643–5, 2005.
- [93] Whittaker GR Belouzard S, Chu VC. Activation of the sars coronavirus spike protein via sequential proteolytic cleavage at two distinct sites. *Proc Natl Acad Sci U S A*, 106 :5871–6, 2009.
- [94] Peters C Dermody TS Ebert DH, Deussing J. Cathepsin l and cathepsin b mediate reovirus disassembly in murine fibroblast cells. *J Biol Chem*, 277 :24609–17, 2002.

- [95] Tattersall P Farr GA, Cotmore SF. Vp2 cleavage and the leucine ring at the base of the fivefold cylinder control ph-dependent externalization of both the vp1 n terminus and the genome of minute virus of mice. *J Virol*, 80 :161–71, 2006.
- [96] Schiller JT Day PM Richards RM, Lowy DR. Cleavage of the papillomavirus minor capsid protein, l2, at a furin consensus site is necessary for infection. *Proc Natl Acad Sci U S A*, 103 :1522–7, 2006.
- [97] Floyd DL et al. Single-particle kinetics of influenza virus membrane fusion. *Proc Natl Acad Sci U S A*, 105 :15382–7, 2008.
- [98] Nibert ML Chandran K. Animal cell invasion by a large nonenveloped virus : reovirus delivers the goods. *Trends Microbiol*, 11 :374–82, 2003.
- [99] Abercrombie RF al Baldawi NF. Cytoplasmic hydrogen ion diffusion coefficient. *Biophys J*, 61 :1470–9, 1992.
- [100] Kawasaki K Masaki I, Mizuno T. Membrane fusion by single influenza hemagglutinin trimers : kinetic evidence from image analysis of hemagglutinin-reconstituted vesicles. *J Biol Chem*, 281 :12729–35, 2006.
- [101] Kohlbrenner E et al. Successful production of pseudotyped raav vectors using a modified baculovirus expression system. *Mol Ther*, 12 :1217–25, 2005.
- [102] Akache B et al. A two-hybrid screen identifies cathepsins b and l as uncoating factors for adeno-associated virus 2 and 8. *Mol Ther*, 15 :330–9, 2007.
- [103] Wagner E Boeckle S. Optimizing targeted gene delivery : chemical modification of viral vectors and synthesis of artificial virus vector systems. *AAPS J*, 8 :E731–42, 2006.
- [104] Taylor HM Karlin S. *A First Course in Stochastic Processes*. Academic Press, 1975.
- [105] Razafsky DS King SJ Gross SP Vershinin M, Carter BC. Multiple-motor based transport and its regulation by tau. *Proc Natl Acad Sci U S A*, 104 :87–92, 2007.

