Redundancy principle for optimal random search in biology

Z. Schuss*, K. Basnayake†, D. Holcman†

Abstract

Chemical activation rate is traditionally determined by the diffusion flux into an absorbing ball, as computed by Smoluchowski in 1916. Thus the rate is set by the mean first passage time (MFPT) of a Brownian particle to a small target. This paradigm is shifted in this manuscript to set the time scale of activation in cellular biology to the mean time of the first among many arrivals of particles at the activation site. This rate is very different from the MFPT and depends on different geometrical parameters. The shift calls for the reconsideration of physical modeling such as deterministic and stochastic chemical reactions based on the traditional forward rate, especially for fast activation processes occurring in living cells. Consequently, the biological activation time is not necessarily exponential. The new paradigm clarifies the role of population redundancy in accelerating search processes and in defining cellular-activation time scales. This is the case, for example, in cellular transduction or in the nonlinear dependence of fertilization rate on the number of spermatozoa. We conclude that statistics of the extreme set the new laws of biology, which can be very different from the physical laws derived for individuals.

1 Introduction

Why are specialized sensory cells so sensitive and what determines their efficiency? For example, rod photoreceptors can detect a single photon in few tens of milliseconds [1], olfactory cells sense few odorant molecules on a similar time scale, calcium ions can induce calcium release in few milliseconds in neuronal synapses, which is a key process in triggering synaptic plasticity that underlies learning and memory. Decades of research have revealed the molecular processes underlying these cellular responses, that identify molecules and their rates, but in most cases, the underlying physical scenario remains unclear.

*Department of Applied Mathematics, Tel-Aviv University, Tel-Aviv 69978, Israel.

†Computational Bioplogy and Applied Mathematics, IBENS, Ecole Normale Supérieure, Paris, France.

Here, we discuss how the many molecules, which are obviously redundant in the traditional activation theory, define the *in vivo* time scale of chemical reactions. This redundancy is particularly relevant when the site of activation is physically separated from the initial position of the molecular messengers. The redundancy is often generated for the purpose of resolving the time constraint of fast-activating molecular pathways. Activation occurs by the first particle to find a small target and the time scale for this activation uses very different geometrical features (optimal paths) than the one described by the traditional mass-action law, reaction-diffusion equations, or Markov-chain representation of stochastic chemical reactions.

We also discuss the role of the fastest particle to arrive at a small target in the context of fertilization, where the enormous and disproportionate number of spermatozoa, relative to the single ovule, remains an enigma. Yet, when the number of sperms is reduced by four, infertility ensues [2]. The analysis of extreme statistics can explain the waste of resources in so many natural systems. So yes, numbers matter and wasting resources serve an optimal purpose: selecting the most fitted or the fastest.

More specifically, mass-action theory of chemical reactions between two reactants in solution, A and B , is expressed as

where k_1 and k_{-1} are the forward and backward reaction rates, respectively. The computation of the backward rate has long history that begins with Arrhenius' law $k_{-1} = Ae^{E/kT}$, where A is a constant, and E is activation energy, then Kramers' rate, derived from the molecular stochastic Langevin equation, which gives the prefactor A . For the past sixty years, chemical physicists computed the activation energy E and clarified the role of the energy landscape, with extensions to applications in chemistry, signal processing (time to loss of lock in phase trackers [3]), finance (time for binary option price to reach a threshold), and many more [4].

In contrast, the forward rate k_1 represents the flux of three-dimensional Brownian particles arriving at a small ball or radius a . Smoluchowski's 1916 forward rate computation reveals that

$$k_1 = 4\pi Dca, \tag{1}$$

where D is the diffusion coefficient, when the concentration c is maintained constant far away from the reaction site. When the window is in a smooth surface or inside a hidden cusp, the forward rate k_1 is the reciprocal of the MFPT of a Brownian particle to the window. The precise geometry of the activating small windows has been captured by general asymptotics of the mean first arrival time at high activation energy. This mean time is, indeed, sufficient to characterize the rate, because the binding process is

Poissonian and the rate is precisely the reciprocal of the MFPT. These computations are summarized in the narrow escape theory [5–9]. For example, when the small absorbing window represents binding at a surface, the forward rate is given by

$$k_1^{-1} = \frac{|\Omega|}{4aD \left[1 + \frac{L(\mathbf{0}) + N(\mathbf{0})}{2\pi} a \log a + o(a \log a) \right]},$$

with $|\Omega|$ the volume of the domain of Brownian motion, a is the radius of the absorbing window [5], and $L(0)$ and $N(0)$ are the principal mean curvatures of the surface at the small absorbing window.

The forward rate k_1 has been used in almost all representations of chemical reactions: it is the basis of the Gillespie algorithm for generating statistics of stochastic simulations with rate k_1 . It has also been used in coarse-graining stochastic chemical-reactions into a Markov chain. However, the rate k_1 is not used in simulations of Brownian trajectories. In the latter case, the statistics of arrival times can be computed directly. Obviously, diffusion theory computes k_1 from the mean arrival rate of a single particle. However, in cell biology, biochemical processes are often activated by the first (fastest) particle that reaches a small binding target, so that the average time for a single particle does not necessarily represent the time scale of activation. Thus even the Gillespie algorithm would give a rate, sampled with mean k_1 . But, as seen below, its statistics is different from the rate of the fastest arrival. This difference is the key to the determination of the time scale of cellular activation that can be computed from full Brownian simulations and/or asymptotics of the fastest particle. This is an important shift from the traditional paradigm.

2 Biochemical reactions in cell biology

Chemical activation in cell biology starts with the binding of few molecules (Fig 1). The signal is often amplified so that a molecular event is transformed into a cellular signal. How fast is this activation? What defines its time scale? when there is a separation between the site of the first activation and that of amplification. When particle move by diffusion, is it sufficient that the first particle arrives to a receptor to open it and lead to an avalanche either through the entry of ions or the opening of the neighboring receptors. Thus the time scale of activation is not given by the reciprocal of the forward rate, but rather by the extreme statistics, that is, by the mean arrival time of the first particle to the activation site (the target).

The statistics of the first particle to arrive to a target can be computed from the statistics of a single particle when they are all independent and identically distributed [10–13]. With N non-interacting i.i.d. Brownian trajectories (ions) in a bounded domain Ω to a binding site, the shortest arrival time τ^1 is by definition

$$\tau^1 = \min(t_1, \dots, t_N), \quad (2)$$

Figure 1: **(A)** Calcium-induced-calcium-release in a dendritic spine. The first Ryanodine Receptor (RyR) at the base of the spine apparatus that opens is triggered by the fastest calcium ion. An avalanche of calcium release ensues by opening the neighboring receptors. This leads to rapid amplification at a much shorter time than the MFPT of the diffusing calcium ions. **(B)** Activation of calcium release by IP3 receptors, which are calcium channels gated by IP3 molecules, which function as secondary messengers. When the first IP3 molecules arrives at the first IP3R, its calcium release induces an avalanche due to the opening of subsequent IP3 receptors. **(C)** In the post-synaptic terminal, the influx of ions due to the opening of NMDA/AMPA receptors is the amplification process. The signal of the pre-synaptic signal transmitted by the neurotransmitter molecules diffusing in the synaptic cleft and the time scale of the amplification are determined by the fastest molecules that arrive at the receptor targets and open them.

where t_i are the i.i.d. arrival times of the N ions in the medium. The narrow escape problem (NEP) is to find the PDF and the MFPT of τ^1 . The complementary PDF of τ^1 ,

$$\Pr\{\tau^1 > t\} = \Pr^N\{t_1 > t\}. \quad (3)$$

Here $\Pr\{t_1 > t\}$ is the survival probability of a single particle prior to binding at the target. This probability can be computed by solving the diffusion equation

$$\begin{aligned} \frac{\partial p(\mathbf{x}, t)}{\partial t} &= D\Delta p(\mathbf{x}, t) \quad \text{for } \mathbf{x} \in \Omega, \quad t > 0 \\ p(\mathbf{x}, 0) &= p_0(\mathbf{x}) \quad \text{for } \mathbf{x} \in \Omega \\ \frac{\partial p(\mathbf{x}, t)}{\partial \mathbf{n}} &= 0 \quad \text{for } \mathbf{x} \in \partial\Omega_r \\ p(\mathbf{x}, t) &= 0 \quad \text{for } \mathbf{x} \in \partial\Omega_a, \end{aligned} \quad (4)$$

where the boundary $\partial\Omega$ contains N_R binding sites $\partial\Omega_i \subset \partial\Omega$ ($\partial\Omega_a = \bigcup_{i=1}^{N_R} \partial\Omega_i$, $\partial\Omega_r = \partial\Omega - \partial\Omega_a$). The single particle survival probability is

$$\Pr\{t_1 > t\} = \int_{\Omega} p(\mathbf{x}, t) d\mathbf{x}, \quad (5)$$

so that $\Pr\{\tau^1 = t\} = \frac{d}{dt} \Pr\{\tau^1 < t\} = N(\Pr\{t_1 > t\})^{N-1} \Pr\{t_1 = t\}$, where $\Pr\{t_1 = t\} = \oint_{\partial\Omega_a} \frac{\partial p(\mathbf{x}, t)}{\partial \mathbf{n}} dS\mathbf{x}$ and $\Pr\{t_1 = t\} = N_R \oint_{\partial\Omega_1} \frac{\partial p(\mathbf{x}, t)}{\partial \mathbf{n}} dS\mathbf{x}$. The pdf of the arrival time is

$$\Pr\{\tau^1 = t\} = N N_R \left[\int_{\Omega} p(\mathbf{x}, t) d\mathbf{x} \right]^{N-1} \oint_{\partial\Omega_1} \frac{\partial p(\mathbf{x}, t)}{\partial \mathbf{n}} dS\mathbf{x}, \quad (6)$$

which gives the MFPT

$$\bar{\tau}^1 = \int_0^{\infty} \Pr\{\tau^1 > t\} dt = \int_0^{\infty} [\Pr\{t_1 > t\}]^N dt. \quad (7)$$

The shortest ray from the source to the absorbing window δ_{min} plays a key role, because the fastest trajectory is as close as possible to that ray. The diffusion coefficient is D , the number of Brownian particles is N , $s_2 = |\mathbf{x} - \mathbf{A}|$, $\mathbf{x}$ the positions of injection, and the center of the window is $\mathbf{A}$. The asymptotic laws for the expected first arrival time of

Brownian particles to a target for large N , are,

$$\bar{\tau}^{d1} \approx \frac{\delta_{min}^2}{4D \ln(\frac{N}{\sqrt{\pi}})}, \text{ in dim 1} \quad (8)$$

$$\bar{\tau}^{d2} \approx \frac{s_2^2}{4D \log\left(\frac{\pi\sqrt{2}N}{8 \log(\frac{1}{\varepsilon})}\right)}, \text{ in dim 2} \quad (9)$$

$$\bar{\tau}^{d3} \approx \frac{\delta^2}{2D \sqrt{\log\left(N \frac{4a^2}{\pi^{1/2}\delta^2}\right)}}, \text{ in dim 3.} \quad (10)$$

These formulas show that expected arrival time of the fastest particles is $O(1/\log(N))$ (see Fig. 2). They should replace the classical forward rate in models of activation in biochemical reactions.

Figure 2: Escape through a narrow opening in a planar disk. **A.** The geometry of the NEP for the fastest particle. **B.** Plot of the expected arrival time of the fastest path vs the number of particles N . The asymptotic solution (red curve) is fit to $\frac{\alpha}{\log(N)+\beta}$.

3 Examples of signal transduction

There are many examples of activation or transduction by the arrival of the first particle at the activation site, which defines the time scale of [1]. First, neuronal connections often occur on a dendritic spine, where the fast calcium increase in dendrites may happen a few milliseconds after initiation in the spine head. This time scale is incompatible with diffusion alone [5]. It was shown recently that such a short time scale is generated by an avalanche reaction triggered by the arrival of the fastest calcium ion to a receptor (see Fig.

1). A similar process occurs in a fly's photoresponse to the absorption of a single photon: the first TRP channel that opens is due to the local diffusion of several IP3 molecules and calcium ions produced after rhodopsin activation.

In another context, the post-synaptic current is generated when the first receptor of a neurotransmitter is activated (Fig. 1). This activation is mediated by the release of thousands of neurotransmitters from the pre-synaptic terminal. Two binding events on the same receptor are required for activation. The time scale of this process is defined by the arrival of the first two neurotransmitters at the same receptor. The asymmetry between the number of neurotransmitters (2000 to 3000) and the low number of receptors (5 to 50) compensates for the low probability of finding the small targets (receptors) [5].

Another transduction pathway is the one of IP3, which begins with the activation of mGluR receptors and leads to IP3 production. These molecules have to diffuse and to bind to the first IP3 receptor in order to trigger calcium release, which leads to the amplification of the response. The number of activated IP3 and the location of IP3 receptors sets the time scale of this transduction pathway.

4 A nonlinear effect of spermatozoa redundancy on fertility

In the key step of fertilization, sperms have to find the ovule within the short time it is fertile (Fig. 3). Thus the arrival of the first few sperms gives them much higher chance to fertilize the ovule. The trajectories of the fastest spermatozoa, which result from the directed motion model, reveal that the extreme trajectories trace the most effective path of that dynamics. These are the optimal (bang-bang) solutions of the classical control problem (Fig. 3 and [14]). This result suggests that linear trajectories might not be generated by any chemotaxis at a distance of a few centimeters, which is too far away from the source. Similarly, it is not clear how other physical scenarios, such as rheotaxis or thermotaxis contribute to sperm guidance [15]. It is thus conceivable that the extreme statistics are responsible for selection of the fastest trajectories determined by sperm dynamics in the uterus and fallopian tubes. The number of spermatozoa is thus the main determinant of the selection and since the mean time for the first one to arrive is $O(1/\log N)$, a large number of them is necessary to affect the search process. In addition, it is well documented that reducing their number by a factor of 4 may cause infertility [2].

5 Morale of the story

Molecular-level simulations of activation processes should avoid the Gillespie algorithm and reaction-diffusion equations especially in the context of cellular fast activation induced by molecular pathways. They should use straightforward Brownian paths and extreme statistics, as given in the formulas above. The large number of paths produces optimal

Figure 3: **(A)** Schematic representation of the uterus and the diffusive trajectory of the sperm that reaches an ovule. The possibility to form a zygote is determined by whether the time of the fastest sperm to reach the ovule is within the fertile period. **(B)** In a Brownian simulation in a schematic uterus, the trajectories with short arrival time are concentrated along two symmetric optimal paths (white dashed line) [14].

trajectories that set the observed time scale. Disproportionate numbers of particles in natural processes should not be considered wasteful, but rather, they serve a clear purpose: they are necessary for generating the fastest response. This property is universal ranging from the molecular scale to the population level. It seems that nature's strategy for optimizing the response time is not necessarily defined by the physics of the motion of an individual particle, but rather by the collective extreme statistics.

References

- [1] G. L. Fain, Molecular and cellular physiology of neurons. Harvard University Press, 1999.
- [2] A. Bendsdorp, B. J. Cohlen, M. J. Heineman, and P. Vandekerckhove, “Intra-uterine insemination for male subfertility,” Cochrane Database Syst Rev, vol. 4, no. 4, 2007.
- [3] Z. Schuss, Nonlinear filtering and optimal phase tracking, vol. 180. Springer Science & Business Media, 2011.
- [4] D. Holcman and Z. Schuss, “100 years after smoluchowski: stochastic processes in cell biology,” Journal of Physics A: Mathematical and Theoretical, vol. 50, no. 9, p. 093002, 2017.
- [5] D. Holcman and Z. Schuss, “Stochastic narrow escape in molecular and cellular biology,” Analysis and Applications. Springer, New York, 2015.
- [6] D. Holcman and Z. Schuss, “Stochastic chemical reactions in microdomains,” The Journal of chemical physics, vol. 122, no. 11, p. 114710, 2005.
- [7] M. J. Ward, W. D. Heshaw, and J. B. Keller, “Summing logarithmic expansions for singularly perturbed eigenvalue problems,” SIAM Journal on Applied Mathematics, vol. 53, no. 3, pp. 799–828, 1993.
- [8] V. Kurella, J. C. Tzou, D. Coombs, and M. J. Ward, “Asymptotic analysis of first passage time problems inspired by ecology,” Bulletin of mathematical biology, vol. 77, no. 1, pp. 83–125, 2015.
- [9] I. V. Grigoriev, Y. A. Makhnovskii, A. M. Berezhkovskii, and V. Y. Zitserman, “Kinetics of escape through a small hole,” The Journal of chemical physics, vol. 116, no. 22, pp. 9574–9577, 2002.
- [10] I. M. Sokolov, R. Metzler, K. Pant, and M. C. Williams, “First passage time of n excluded-volume particles on a line,” Physical Review E, vol. 72, no. 4, p. 041102, 2005.
- [11] S. B. Yuste and K. Lindenberg, “Order statistics for first passage times in one-dimensional diffusion processes,” Journal of statistical physics, vol. 85, no. 3, pp. 501–512, 1996.
- [12] T. Chou and M. Dorsogna, “First passage problems in biology,” First-Passage Phenomena and Their Applications, vol. 35, p. 306, 2014.
- [13] G. Schehr, S. N. Majumdar, G. Oshanin, and S. Redner, “Exact record and order statistics of random walks via first-passage ideas,” First-Passage Phenomena and Their Applications, vol. 35, p. 226, 2014.

- [14] J. Yang, I. Kupka, Z. Schuss, and D. Holcman, “Search for a small egg by spermatozoa in restricted geometries,” Journal of mathematical biology, vol. 73, no. 2, pp. 423–446, 2016.
- [15] H. D. Hoang and M. A. Miller, “Sperm navigation mechanisms in the female reproductive tract,” in Signaling-Mediated Control of Cell Division, pp. 241–267, Springer, 2017.