It has recently

On the Relationship Between the Statistics of SPR Area and the Number of Activated G-proteins/PDE* Produced per R*

Introduction

It has recently been argued that the total number of activated G-proteins (G*) produced per activated rhodopsin molecule (R*) may be as low as 20, on average. We show here that this low number is not compatible with the reproducibility of the single0photon response (SPR) under the multiple-phosporylation hypothesis for reproducibility of the SPR. We use statistical reasoning in combination with a particular feature of the SPR, the SPR area, to argue our case.

Statistical Model for Activated Phosphodiesterase (PDE*)

Assume that for each activated rhodopsin molecule, R*, a random number N of PDE* subunits (just PDE* for short) are produced in a one-to-one fashion by activated G-proteins (G*), at random times T_k. Further assume that the time course of deactivation of PDE* is dominated by a single rate-limiting step, so that it can be assume that the lifetimes of the PDE* molecules are independent, identically distributed, exponential ,random variables, with mean $\tau_{\rm PDE}$.

Define P(t) as the random *PDE activity indicator function*. It is defines in such a way that P(t) is the sum over $X_k(t)$, where $X_k(t) = 1$ during the lifetime of the k$^{\rm th}$ activated PDE subunit and zero otherwise. Then

$$P(t) = \sum_{k=0}^{N} X_k(t) \qquad \text{with} \tag{1}$$

$$X_k(t) = \begin{cases} 0 \text{ for } k = 0 \\ H(t - T_k) - H(t - T_k - L_k) \text{ for } k > 0 \end{cases} \tag{2}$$

where T_k is the random time at which the k$^{\rm th}$ PDE* is produced, and L_k is the random lifetime of the k$^{\rm th}$ PDE* for $k \neq 0$ and $L_0 = 0$ by definition.

Relationship between PDE* and the Single-Photon Response (SPR)

Assume the the random SPR photocurrent (or relative photocurrent), $S(t)$, is a linear functional of the PDE* activity function $P(t)$, so $S(t)$ can be written as

$$S(t) = \int_0^\infty dt'\, P(t')\, g(t - t'), \tag{3}$$

where $g(t)$ is a linear kernel that represents the kinetics of all dynamical processes that link PDE* activity to the SPR.

It follows that the random area A of the SPR is given by

$$\begin{aligned} A &= \int_0^\infty dt\, S(t) && (4) \\ &= \left(\int_0^\infty dt\, g(t)\right) \sum_{k=0}^{N} L_k. && (5) \end{aligned}$$

It can be shown that, for exponentially distributed L_k, the coefficient of variation of A is given by

$$\frac{\sigma_A}{\mu_A} = \sqrt{\frac{\sigma_N^2}{\mu_N^2} + \frac{1}{\mu_N}}. \tag{6}$$

Statistics of the Random Number of PDE* per R*

In the simple multiple-phosphorylation model for R* inactivation, each phosphorylated form of R*, R_k^* for $k = 0, 1, ..m$, produces a random number of activated G-proteins/PDE*, N_k, where the N_k are independent random variables. It is reasonable to assume that the N_k are geometrically distributed. The argument is as follow. Every time and R_k^* molecule is free, it can exit that free state either by catalyzing the production of a G* (which, in turn, produces a PDE*), or it can be phosphorylated to give R_{k+1}^*, or it can be capped by arrestin. There is some probability, call it p_k that a G* will be produced before either of the other 2 possible

events occur, and upon producing a G*, R* is freed up again. The outcome every time R_k^* is free is independent of the the number of times a G* has been produced up to that point in time. Therefore, in this scenario, the probability that n G* are produced by R_k^**before one of the other events occur is given by

$$\Pr\{N_k = n\} = (1 - p_k)\, p_k^n; \tag{7}$$

that is, the probability mass function for N_k is a geometric distribution. A property of a geometric random variable is that its variance is its mean plus the square of its mean, and this feature will play a crucial role below.

It can be shown that the best-case scenario, with regard to producing an SPR area with low coefficient of variation, is achieved when all possible R* phosphorylations take place before arrestin capping, and when the mean of each N_k is the same. Let us specialize to this best-case scenario. If we define N as the random total number of G* produced, in the best case scenario,

$$\begin{aligned} \mu_{N_k} &= \frac{\mu_N}{m+1} \text{ and} && (8)\\ \sigma^2_{N_k} &= \mu_{N_k} + \mu^2_{N_k} && (9)\\ &= \left(\frac{\mu_N}{m+1}\right) + \left(\frac{\mu_N}{m+1}\right)^2. && (10) \end{aligned}$$

It follows that

$$\sigma^2_N = \mu_N + \frac{\mu^2_N}{m+1}, \tag{11}$$

and

$$\frac{\sigma^2_N}{\mu^2_N} = \frac{1}{\mu_N} + \frac{1}{m+1}. \tag{12}$$

If we substitute (**??**) into (**??**), we find the the coefficient of variation of the SPR area is

given by

$$\frac{\sigma_A}{\mu_A} = \sqrt{\frac{1}{m+1} + \frac{2}{\mu_N}}. \tag{13}$$

Figure **??** plots the coefficient of variation of SPR area in the best-case scenario as a function of the assumed mean number of G* produced per R*, and for a family of assumed phopshorylations for R* deactivation. It is cleat that only 20 G* produced per R* is incompatible with the measures reproducibility of the SPR area. The measured coefficient of variation of SPR are in a number of diffferent experiments has been measured as 0.36 ± 0.02.

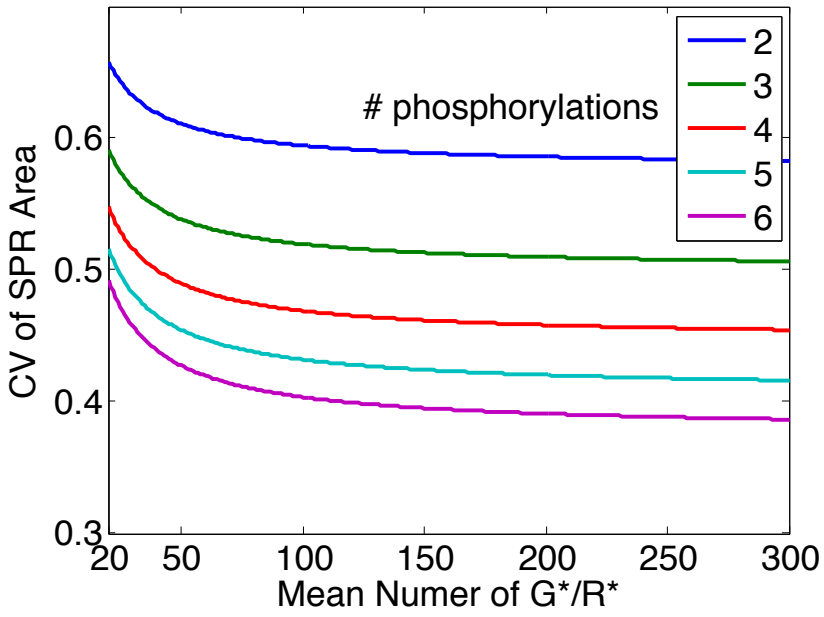

Figure 1: Coefficient of variation of SPR area in best-case scenarios in which R* is deactivated by sequential phophorylations terminating in arrestin capping. Each curve is for a different number of phosphorylation sites as indicated in the figure legend. Each cirve is a plot of the minumum coefficient of variation possible as a function of the average number of G* produced per R*.