Doctoral Dissertation

Data-driven System Identification of Signal Conversion from cGMP to Membrane Potential in Neuronal Growth Cone

Tatsuya Yamada

February 1, 2018

Department of Bioinformatics and Genomics
Graduate School of Information Science
Nara Institute of Science and Technology

Tatsuya Yamada

Data-driven System Identification of Signal Conversion from cGMP to Membrane Potential in Neuronal Growth Cone[*]

Tatsuya Yamada

Abstract

Growth cone, at the tips of growing neurites (axons and dendrites) of developing neurons are guided to their synaptic targets by various external guidance molecules to make functional neuronal connections. In the presence of guidance factor, Sema3A, repulsive movement of the growth cone from Sema3A is accompanied by a decrease of membrane potential (MP) of the growth cone. The growth cone turns to an attractive moving with MP elevation when the intracellular cGMP concentration is increased. This fact indicates that the biochemical signal encoded by cGMP concentration is converted to the electrical signal of MP. However, the mechanism of signal conversion among such different physical modalities remains unknown. In this study, we aim to identify the system which quantitatively realizes the signal conversion from cGMP to MP in growth cone. We described MP time series by deterministic computational model and expressed individual experimental variation and cell-to-cell variation by the probability distribution of the model parameters. By applying Markov Chain Monte Carlo (MCMC) method to Bayes' theorem, we estimated the posterior distribution of the model parameters, and then we extracted the most plausible system by comparing models based on logarithmic evidence. Our quantitative analysis revealed that the inhibition of Chloride ion channel by the downstream factor of Protein Kinase G is necessary for a core system. In addition, the identified model

quantitatively reproduced the unused time series data and the cGMP dependency of steady-state MP.

神経成長円錐における cGMP から膜電位へのシグナル変換のデータ駆動型システム同定*

山田 達也

内容梗概

神経突起先端にある成長円錐は，細胞外の環境を検知し，神経突起を適切なシナプス結合のターゲットへと導く．例えば，細胞外に誘導因子 Sema3A の濃度勾配がある場合，通常状態の成長円錐は Sema3A に対して膜電位の下降を伴いながら忌避性の運動を行う．しかし，成長円錐内の cGMP 濃度を増加させると，膜電位の上昇を伴いながら誘引性の運動を行うようになる．この事実は，cGMP 濃度による生化学的シグナルが膜電位による電気的シグナルに変換されていることを意味する．しかしながら，こうした異なる物理量間のシグナル変換を実現するメカニズムは未解明である．本研究は，成長円錐における cGMP から膜電位へのシグナル変換を実現するシステムの同定を目的とする．そのために，データ点が豊富な膜電位時系列を決定論的数理モデルで表し，個々の実験や細胞に含まれるばらつきをモデルパラメータの確率分布で表現した．ベイズの定理にマルコフ連鎖モンテカルロ（MCMC）法を適用することによりモデルパラメータの事後分布を推定し，対数エビデンスを用いたモデル比較を行うことで細胞内システムの同定を行った．その結果，既知の相互作用に加え，Protein Kinase G (PKG) の下流による塩素イオンチャネル抑制が必要であることが推定された．妥当性検証を行ったところ，推定されたモデルが他の実験条件の膜電位時系列を定量的に再現するとともに，定常状態後の膜電位においても実験定量データに酷似した cGMP 依存性をみせた．

キーワード

成長円錐，膜電位，ベイズ推定，シグナル変換，データ駆動型システム同定

Contents

List of Figures

List of Tables

1. Introduction

The structural plasticity of neuronal cells have an important role to construct the neuronal network in the developmental stage of animals. In the stage, the tip of neurite (axon and dendrites) called growth cone has the central role of structural plasticity. The growth cone guides its neuronal fiber to an appropriate neuronal target by responding to various external signals (Fig. 1). This phenomena is called growth cone guidance what we deal in this study.
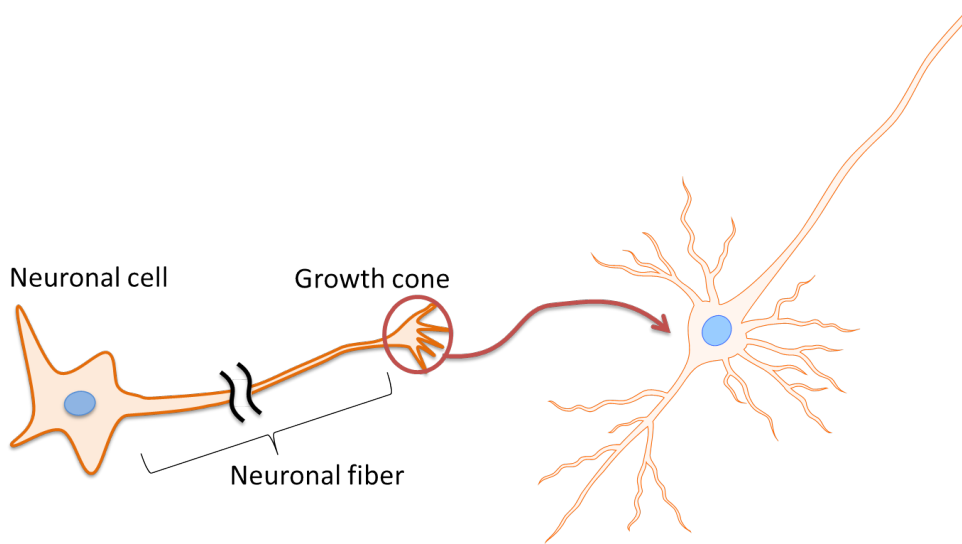
Figure 1. Growth cone guidance. To constructing a neuronal network, growth cone makes neuronal connection by guiding its neuronal fiber to the appropriate target neuron.

Growth cone guidance depends on various molecule signals. The main factor of growth cone guidance is the molecules called guidance factors. Growth cone receives these molecules as a signal by the receptor on the cell membrane, and then guides axon or dendrites to an appropriate target [1]. Depending on the extracellular guidance signal, growth cone behaves differently in elongation speed and in turning direction. In the case of turning direction, there are two types of guidance molecules. "Attractants" such as NGF and BDNF attract the growth cone, whereas "Repellents" such as EphrinA2 and Semaphorin3A (Sema3A) repel the growth cone [2] (Fig. 2).

(a) Attractant   (b) Repellent
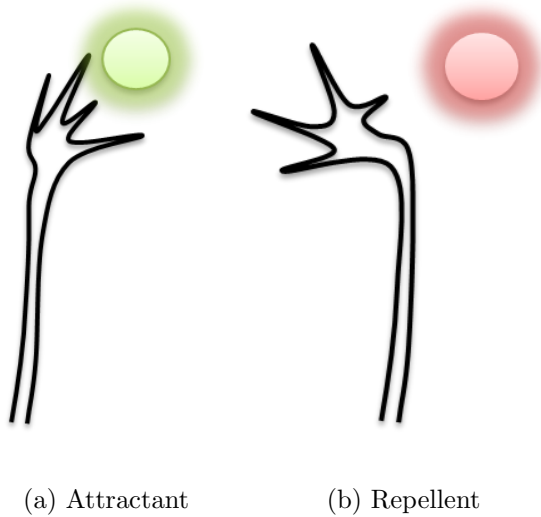
Figure 2. Growth factor. (a) In the attractant gradient, growth cone moves to the high concentration area. (b) In the repellent gradient, growth cone moves to the low concentration area.

Although the guidance molecules are well classified into these two groups, it has become known that growth cone switches behavior depending on the concentration of the second messenger such as cyclic-AMP (cAMP) and cyclic-GMP (cGMP) [3, 4]. For instance, a growth cone, in a control condition, exhibits repulsive response to a Sema3A gradient, but the repulsion is inverted to attraction if the normal cGMP level is elevated [5, 6] (Fig. 3). This suggests the intracellular biochemical signaling pathways instruct the migration direction of the growth cone during its navigation [5, 7, 8, 9]. Furthermore, Ca^{2+} signal is also important signal for growth cone guidance, which leads to the formation of cellular skeleton [10]. In addition, baseline of Ca^{2+} signal switches the direction of growth cone turning even if the elevation of local Ca^{2+} signal induced by guidance molecule is same level [9]. Therefore, unveiling the relationship between the molecular interaction via cAMP/cGMP and Ca^{2+} signaling is an important task to understand the mechanism of growth cone guidance [6].

Previous study showed that membrane potential (MP) shifts regulated by cGMP signaling [11] regulate Ca^{2+} signals through activation of MP-dependent

(a) Low cGMP   (b) Normal cGMP   (c) High cGMP
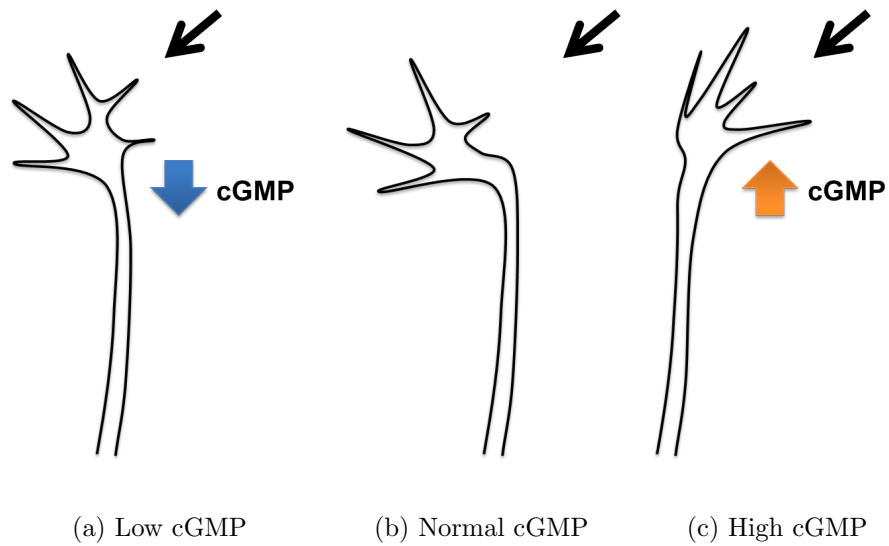
Figure 3. Sema3A-induced cGMP-dependent growth cone turning. Under the condition of Sema3A exposure (from right-upper arrow for each panel), growth cone (a) moves straight, i.e., no response to the Sema3A gradient, (b) repels from the Sema3A, (c) attracts to the Sema3A.

calcium channels [12] that governs the growth cone turning direction. In other words, biochemical signals are converted to electrical signal, which are reconverted into other biochemical signals through various ion channels [13, 14]. In short, growth cone guidance is regulated not only by the molecule signals but also by electrical signal of MP shifts [11].

1.1 MP shifts dependent growth cone turning

In the previous study, the relationship between growth cone turning angle and MP shifts were examined [11] (Fig. 4). As we described, growth cone turning angle depends on the concentration of the second messenger of guidance molecules. The growth cone turning direction is also modulated by state of growth cone MP; a hyperpolarized and depolarized state induces repulsion and attraction, respectively, in response to many diffusible guidance molecules [11]. Moreover, a low level of cGMP causes growth cone hyperpolarization whereas a high level of cGMP causes depolarization [11], demonstrating that a cGMP signal regulates MP shift which eventually determines growth cone turning direction. This indicates that MP has the information to determine the turning direction of the growth cone.

The signal conversion in the Sema3A growth cone guidance, cGMP and MP shifts are the input and the output of the intracellular molecular system, respectively. This input-output relation shows bidirectional behavior (Blue square in Fig. 5). Furthermore, turning angle of the growth cone also shows bidirectional behavior (Yellow circle in Fig. 5). Surprisingly, these two output of growth cone signaling are highly correlated. Taking into account the causality from cGMP molecular signaling to electrical signaling of MP shifts, these facts implies MP directs the growth cone turning behavior [11] (Fig. 6).

Guidance molecules are diffused
from the direction of arrow

0h   1h

Sema 3A

Right panel shows superimposed
plot of growth cone trajectories
(during 0h-1h)

Sema3A

Growth Cone

Record

20 μm   Soma

V

ΔE

0h   1h

ΔE: Membrane potential shifts
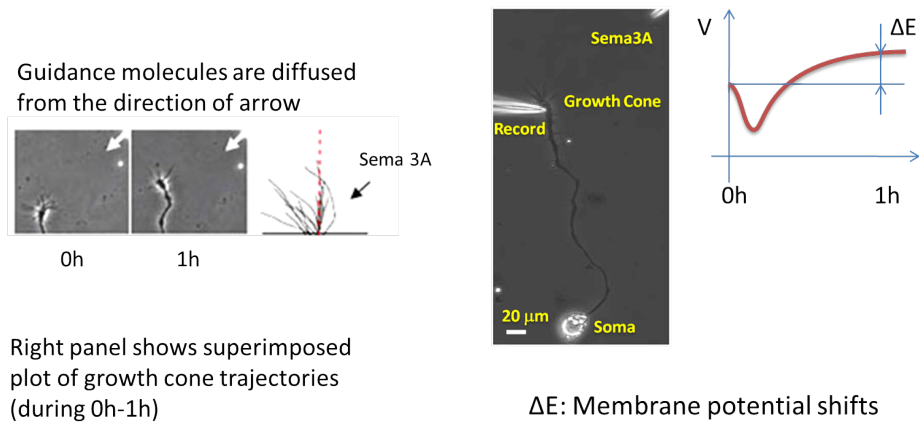
(a) Turning assay

(b) Voltage recording

Figure 4. Experimental configuration of growth cone turning. (a) In turning assay, growth cone is exposed to the Sema3A gradient with/without 8-Br-cGMP bath application. Turning angle is calculated by the difference of the position after 1 hour from exposure and the one of beginning of exposure. (b) In voltage recording, the experimental configuration is same as turning assay. MP shifts is calculated by the difference of same period of turning assay. (Modified from [11])
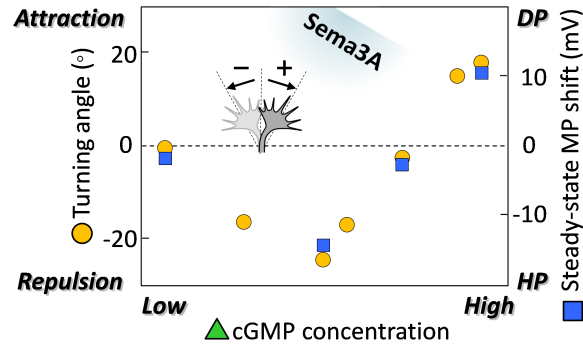
Figure 5. cGMP dependency of turning angle and MP shift. Change of intracellular cGMP level (green triangle) results in MP shift (blue square) and growth cone steering (yellow circle) [6]. Growth cone shows bidirectional turning in response to Sema3A and the direction has a high correlation with steady-state MP shift from the resting potential (before cGMP analogue stimulation); attraction for depolarization (DP) and repulsion for hyperpolarization (HP).
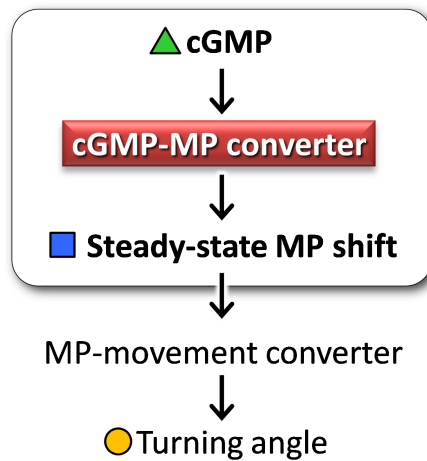
Figure 6. Signaling conversion from cGMP to MP signal. Block diagram of signal conversions among biochemical, electrophysiological, and morphological factors. Intracellular cGMP shifts MP first, and then MP shift steers growing angle. Signal conversion from cGMP to MP (highlighted in red) is analyzed by the model in Fig. 13.

6

1.2 Intracellular molecular system

It is previously demonstrated that activation of Cyclic Nucleotide-Gated ion Channel (CNGC) and cGMP-dependent protein kinase (Protein Kinase G; PKG) are required for MP shifts [11]. The experiment using PKG inhibitor, KT5823, suggested the downstream of PKG could activate sodium channel (NaC). In addition, a high probability was suggested so that CNGC activates chloride channels (ClC). We confirmed increase of Cl^- current through ClC by CNGC activation during 8-Br-cGMP injection (Fig. 7).
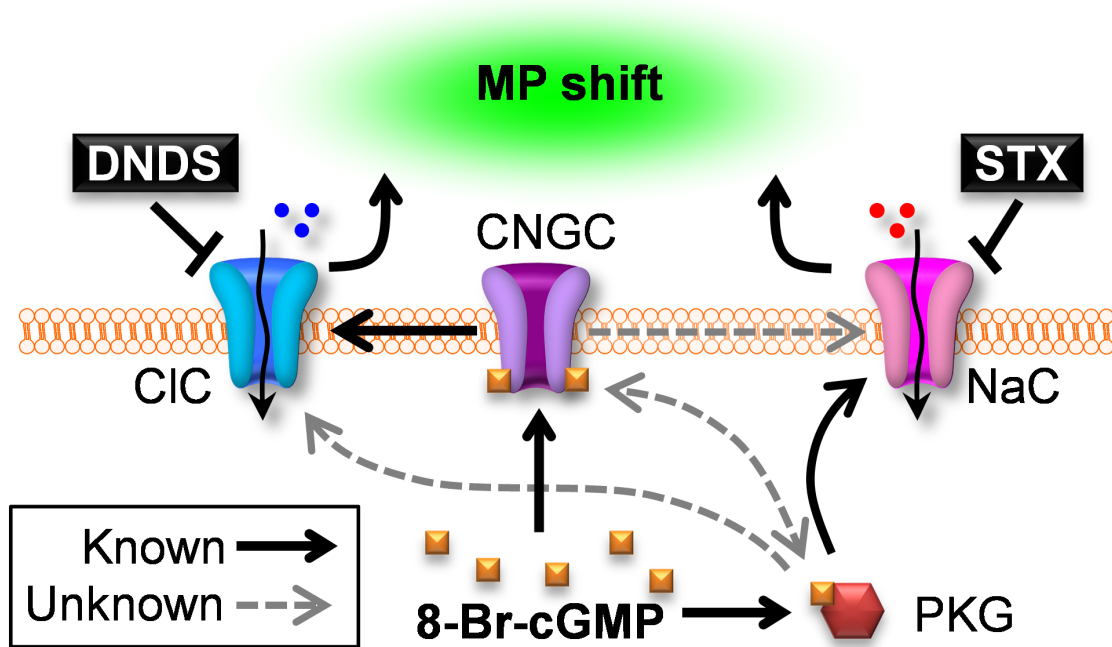
Figure 7. Molecular system which regulates growth cone MP during cGMP stimulation. cGMP diffuses toward the vicinity of growth cone membrane and is bound by PKG and CNGC, which causes activation of the downstream factor. As a result, ion channels (NaC and ClC) on the membrane are modified via intracellular molecular interaction and corresponding ions inflow which cause MP shift.

7

1.3 Purpose of this study

In this study, we aim to identify the intracellular molecular system from different physical modality, MP time series, by quantitative modeling and Bayesian approach. Previously, the biochemical experiment is the only way to investigate the interaction of intracellular molecules, however, it is time consuming. From this reason, it becomes wanted that the information science approach to extract the molecular system from the available data [15]. Several researches provide the solution to extract molecular system from the biochemical data, by using Bayesian approach [15]. The approach is utilized in the present study, however, the particular task is to estimate the molecular system from the MP data.

1.4 Approach of this study

Here, we demonstrate the computational derivation of the significant molecular interaction from a set of time series of electrophysiology recordings obtained from the palm of single growth cones, by a combination of deterministic modeling and Bayesian theory. We used the electrophysiology data, i.e., MP shifts recorded from the growth cone during its movement in response to an external signal, in which each time series contains over 10,000 data points, allowing us to perform quantitative analysis based on the deterministic modeling. We addressed the experimental and cell-to-cell variation by probability distributions of the model parameters and estimated their posterior distributions with datasets obtained over several experimental conditions. Specifically, in our case, data points are sufficient to estimate the model parameters of a time series but number of dataset is insufficient for taking into account the variation of data (a set of time series). A Bayesian approach is especially effective approach when an insufficient number of data are available and there are uncontrollable experimental variations. Applications of Bayesian framework have been well developed in neuroscience [16, 17, 18, 19]. Reverse engineering technique has a potential to deduce unknown biochemical interactions from known biochemical dataset [15, 20], whereas the particular task in the current study is to estimate unknown biochemical interactions from electrophysiology dataset while taking into consideration the experimental variations.

1.5 Organization of the dissertation

In chapter 2, biological experiments and observed data are explained. In chapter 3, deterministic mathematical model is defined. In chapter 4, employed Bayesian approach and model criteria are introduced. In chapter 5, the model selection analysis and the several validation analysis were described. In chapter 6, the supporting studies are glanced and our studies' limitation and future direction are described. Finally, in chapter 7, we conclude this study.

2. Biological experiment and experimental data

In this chapter, we explain the experimental procedure of MP time series (MPTS) recording performed by previous study [11]. The experimental data of the MPTS during growth cone guidance which is analyzed in the present study is also described.

2.1 Experimental procedure and experimental data

The experimental dataset analyzed in this study were obtained by the electro-physiological experiment (Fig. 8). The data is measured by attaching a micro pipette on the palm of growth cone (Left-upper inset of Fig. 8), with intracel-lular 8-Br-cGMP perfusion under the gradient of Sema3A. The experiments are performed with different conditions of the 8-Br-cGMP concentration (5 μM and 10 μM) and the pharmacological reagents (Control, DNDS, STX and KT5823). Pharmacological effects of these reagents are as follows: DNDS and STX block ClC and NaC, respectively. And KT5823 inhibits PKG activity.

We analyzed these data for identifying the intracellular molecular system. To make a meaningful discussion, we used the preprocessed 10 μM dataset (Control/+DNDS/+STX in Fig. 9) only for the identification step and then validated with 5 μM dataset (Control/+DNDS/+STX in Fig. 10) and 10 μM +KT5823 dataset (Fig. 11).
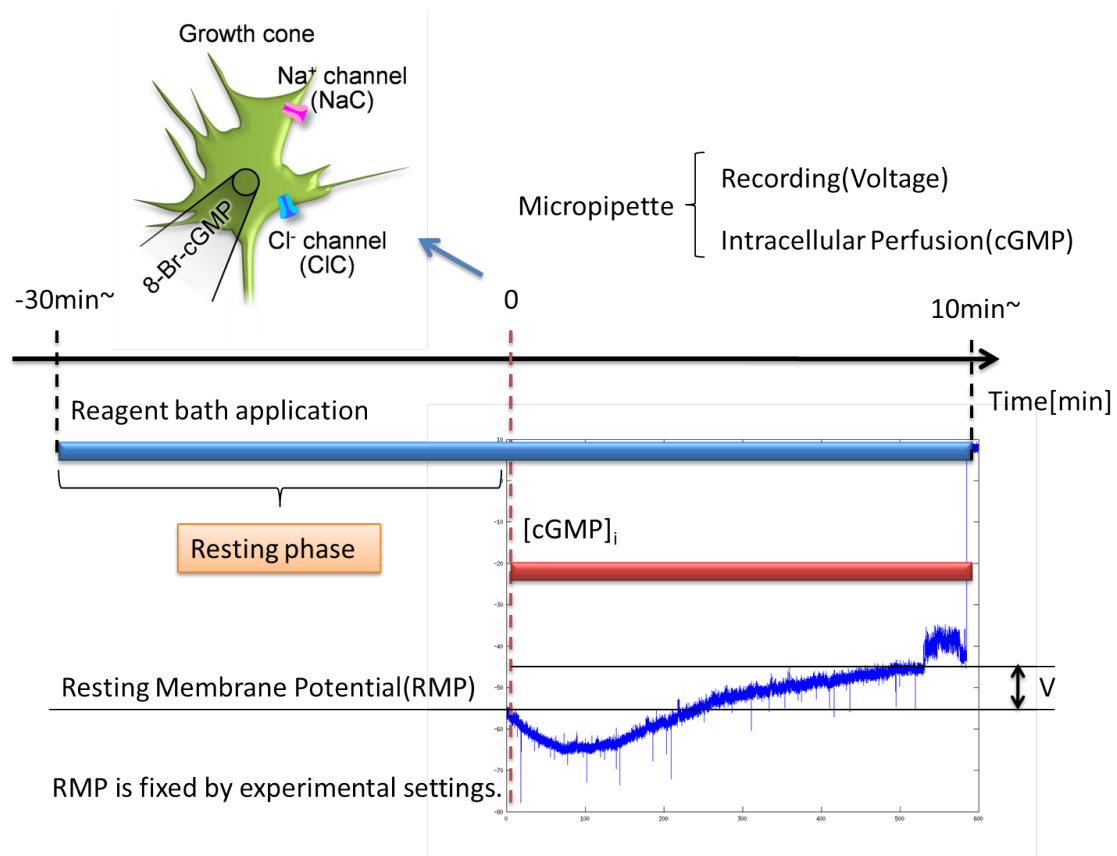
Figure 8. Experimental procedure of MPTS recording. Growth cone is rested before the experiment more than 30 minutes. In the pharmacological condition (in the presence of DNDS, STX or KT5823), reagent is bath-applied from this experimental phase, whereas none of them is applied in the control condition. The MPTS is recorded from the time of micro pipette injection with intracellular 8-Br-cGMP perfusion under the gradient of Sema3A.
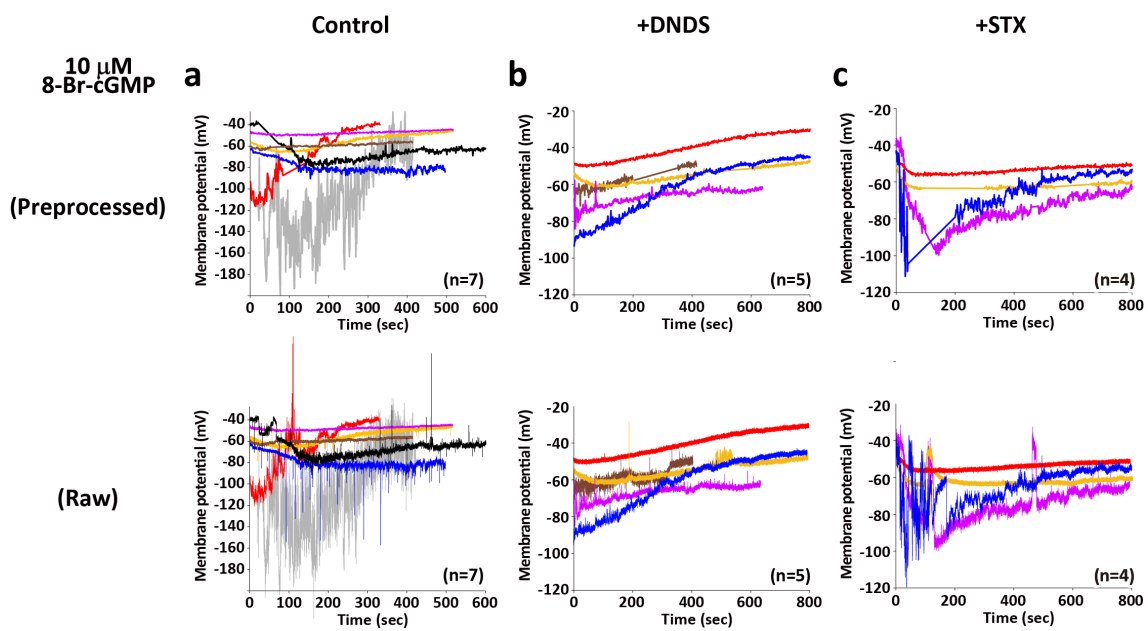
Figure 9. Experimental data. Growth cone MPTS after attachment of the pipette filled with 10 μM 8-Br-cGMP under three conditions; (a) Control (n=7), (b) DNDS (n=5) and (c) STX (n=4). Upper panels show preprocessed MPTSs and lower panels show raw MPTSs
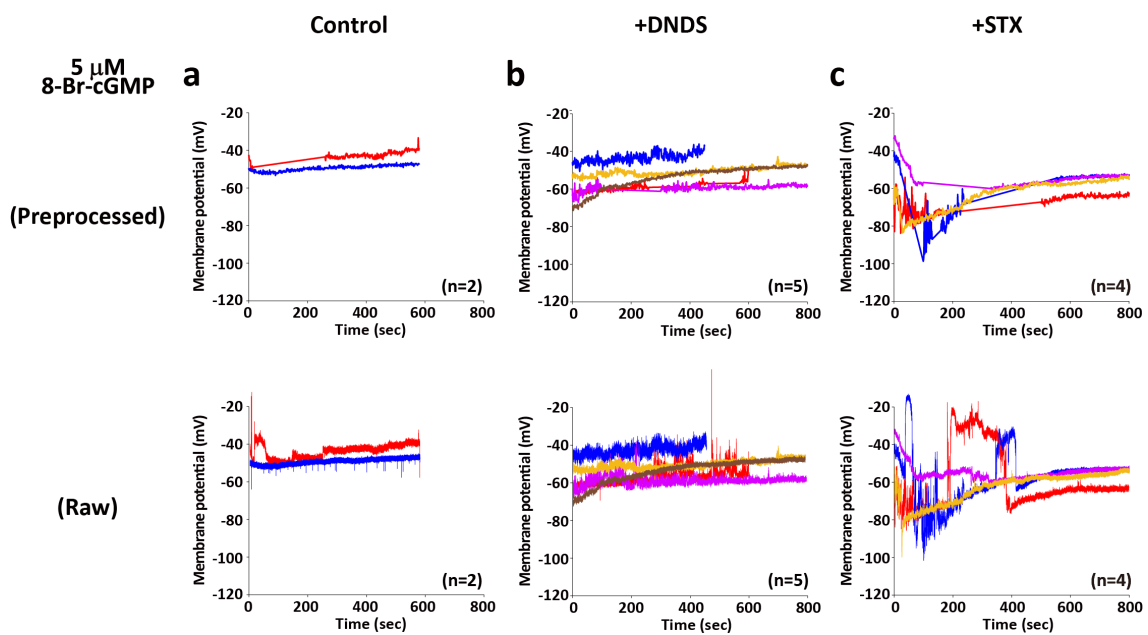
Figure 10. Experimental dataset of MP time series by 5 μM 8-Br-cGMP stimulation with in three different conditions; (a) Control, in the presence of (b) DNDS, and in the presence of (c) STX. These profiles showed relatively small changes than those by 10 μM 8-Br-cGMP stimulation.
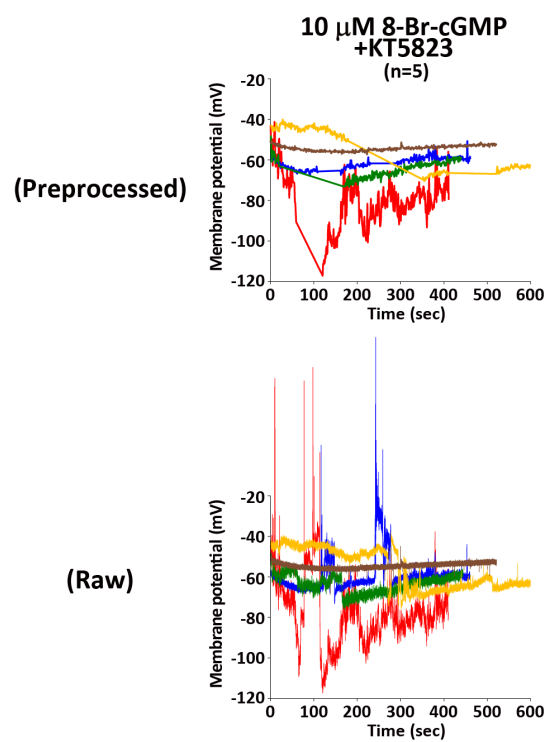
Figure 11. Experimental dataset of MPTS by 10 μM 8-Br-cGMP stimulation with KT5823 (PKG inhibitor) application.

2.2 Preprocessing

In contrast with long time phenomena of growth cone turning and MP shifts (in 1h time scale) what the present study interests, raw experimental MPTS shows several short time events which are dealt as some kind of noise. Specifically, in the analysis, we used the preprocessed MPTS of 300–900 data points (one sec interval) resampled from raw data ($>$10,000 data points). The original time series contained various noises assumed to be of three different kinds (Fig. 12). The first kind had a spike-like form such as a post-synaptic action potential. It was removed by simply applying a smoothing filter with a window of one-second time interval. The second kind was assumed to be observation noises generated by experimental conditions like equipment, and was modeled as white Gaussian in the analysis for introducing the noise into Bayesian framework. The third kind included abrupt and temporal changes in the time series, probably due to experimental artifacts. We removed data points during such temporal change period and complemented with a straight line which connected the onset point and the end point of the temporal change.
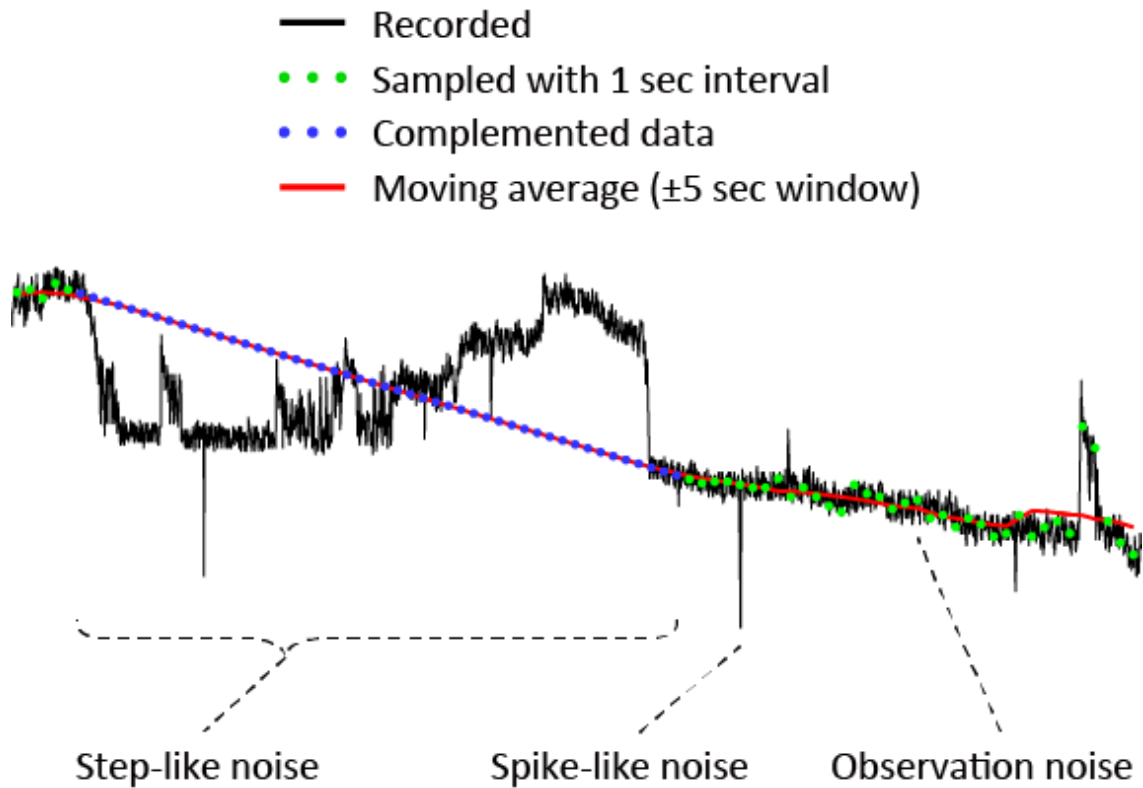
Figure 12. Sample time series of the MP observed in experiment (10 μM 8-Br-cGMP). Black line represent raw MPTS and dots are preprocessed time series, Preprocessing was performed by removing abrupt and temporal changes manually and inserting a straight line to the removed regions (blue dots), and then spike-like noise is also removed by resampling with 1 second interval (i.e., smoothing with 1 second window). Observation noise is considered as Gaussian noise and its standard deviation is estimated as the difference of the preprocessed time series and the smoothed time series (red line) which were obtained by applying moving average with ± 5 second window to each point in the series.

3. Deterministic model of signaling pathway

In this chapter, we model the growth cone MP regulation incorporating intracellular molecular system and membrane electrophysiology. And we show 81 possible model candidates for comparing in the present study.

3.1 Summary of model formulation

To estimate unknown signal interactions (Fig. 13), a quantitative mathematical model was developed based on biological knowledge and physical constraints. The cGMP-MP conversion comprises multiple biochemical and biophysical processes with various time scales (Fig. 9); diffusion of 8-Br-cGMP from an application-recording pipette, biochemical signal transductions of 8-Br-cGMP downstream factors (DFs) including CNGC and PKG, modification of effective densities of ion channels, ClC and NaC, on plasma membrane, and open-close dynamics of the voltage-dependent ion channels. By quantitative simulation analysis, we found that the concentration of 8-Br-cGMP in the growth cone attached with the stimulant pipette rapidly increases in the first phase and then reaches a steady state for several minutes (Fig. 17). Although typical biochemical reactions proceed on a time scale of seconds, it is unknown how fast the effective densities of ion channels are modified. In addition, open-close dynamics of the voltage-dependent ion channels occur on a millisecond time scale [21]. According to the differences in time scale of these processes, it is suggested that the 8-Br-cGMP diffusion and the modification of effective channel densities are possible rate-limiting steps for MPTS, which also shows a fast change within several minutes and a slow change thereafter (Fig. 9 and 14). The rate-limiting step allows us to introduce quasi-steady state approximation (Fig. 15) into the model, in which the fast processes are in steady states at all times and therefore can be formulated by a time-independent function, while the rate-limiting steps are described by time-dependent functions or ordinary differential equations (ODEs).

To identify the interactions which regulate MP shift on the growth cone, 81 possible models are considered based on the molecular-molecular interactions (Fig. 13) as formulated in this chapter.
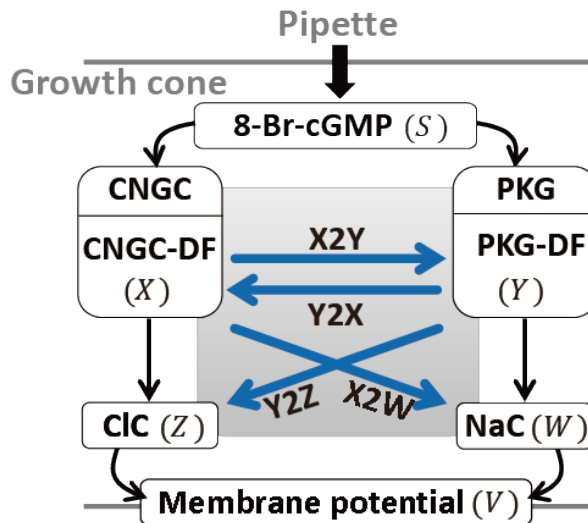
Figure 13. Signaling pathways from cGMP to MP. Black arrows are known as positive pathways, whereas blue arrows represent possible pathways which are examined. cGMP binding to CNGC and PKG activate their downstream factors (DFs) first. CNGC- and PKG-DFs regulate efficacies of ion channels, ClC and NaC, respectively, which cause MP shift. Parallel blue arrows, (X2Y) and (Y2X), indicate possible interactions between DFs whereas cross blue arrows, (X2W) and (Y2Z) indicate possible interactions from DFs to the ion channels. Each of the interactions has one of three types; activation, inhibition, and no interaction. These interaction types of blue arrows in the gray box are interactions to be identified. Italic characters indicate the model variables; S, X, Y, Z, W, and V are 8-Br-cGMP concentration, relative concentrations of CNGC- and PKG-DFs, relative changes of ClC and NaC densities, and MP shift from the potential before 8-Br-cGMP stimulation, respectively.
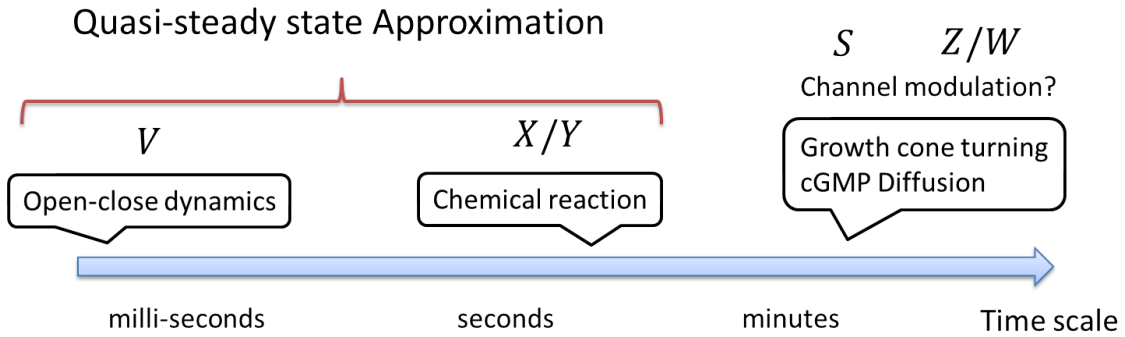
Figure 14. Time scale of the membrane potential regulation. The phenomenon includes several processes which have different time scale. Open-close dynamics of ion channels and chemical reactions are faster than ion channel modulation by the regulator. Notations of each variable are same as Fig. 13
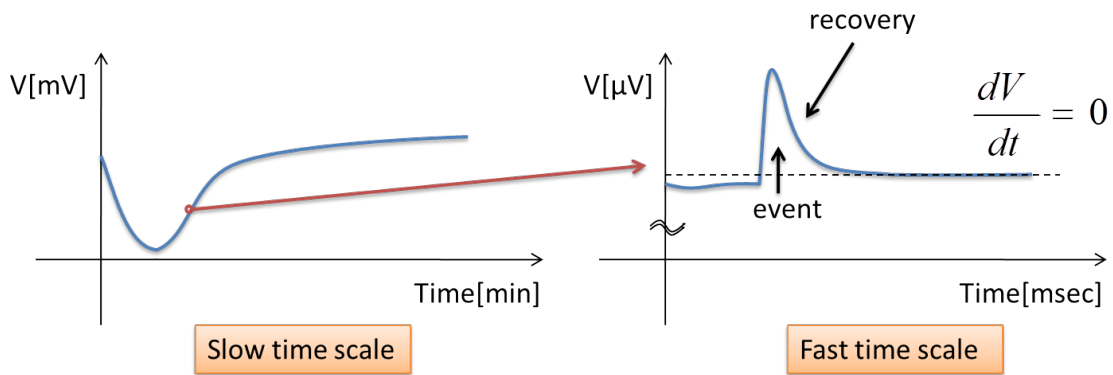
Figure 15. Quasi-steady state approximation. In the slow time scale, the fast event can be ignored. Mathematically, it is equivalent to setting zero to the derivative of the fast time scale variable.

3.2 Formulation of 8-Br-cGMP

In the experiment, the pipette filled with solution of 8-Br-cGMP was attached near the center of the growth cone, from which cGMP analogues (8-Br-cGMP) diffused over the whole growth cone (Fig. 16). The averaged time series of the 8-Br-cGMP concentration, S, was approximated as the simple exponential function of time t,

$$S = S_{\mathrm{max}}(1 - e^{-t/\tau_S}), \tag{1}$$

where S_{max} was the same as the 8-Br-cGMP concentration in the pipette.
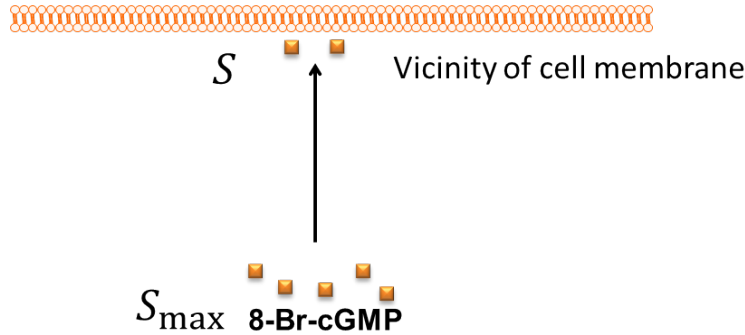
Figure 16. 8-Br-cGMP diffusion. By the intracellular perfusion, 8-Br-cGMP achieves the vicinity of the growth cone from the attached patch pipette.

The time constant, τ_S, was estimated from numerical calculation of a model growth cone consisting of three shell compartments (Fig. 17(a)). The inner two shell compartments were of 1 μm thickness and the most outside shell was of a variable thickness given by 1 μm added by Gaussian noise, $N(0, 1^2)$. In addition to these three compartments, we added one pipette compartment, indexed by 0, in which the cGMP analogue is abundant with a fixed concentration. The 8-Br-cGMP concentration in the n-th compartment, C_n (μM), obeys the Fick's law,

$$\frac{d}{dt}C_n = D\frac{A_{n-1,n}}{d_{n-1,n}V_n}(C_{n-1} - C_n) + D\frac{A_{n,n+1}}{d_{n,n+1}V_n}(C_{n+1} - C_n), \tag{2}$$

where D, $A_{n-1,n}$, $d_{n-1,n}$, and V_n are the diffusion coefficient of 8-Br-cGMP (1 m^2/s, [22]), the surface area between the $(n-1)$-th and the n-th compartments,

20

the distance between the compartment midpoints, and the volume of the n-th compartment, respectively. We solved Eq. (2) numerically with the boundary condition, $C_0 = S_{\max}$, and confirmed that the time course in the third compartment (the outmost shell) can be well approximated by the exponential function in Eq. (1) (Fig. 17(b)). Monte Carlo simulation with various growth cone sizes gave the distribution of the time constant, τ_S, whose mean was 40 sec.

(a) Model

(b) Time course
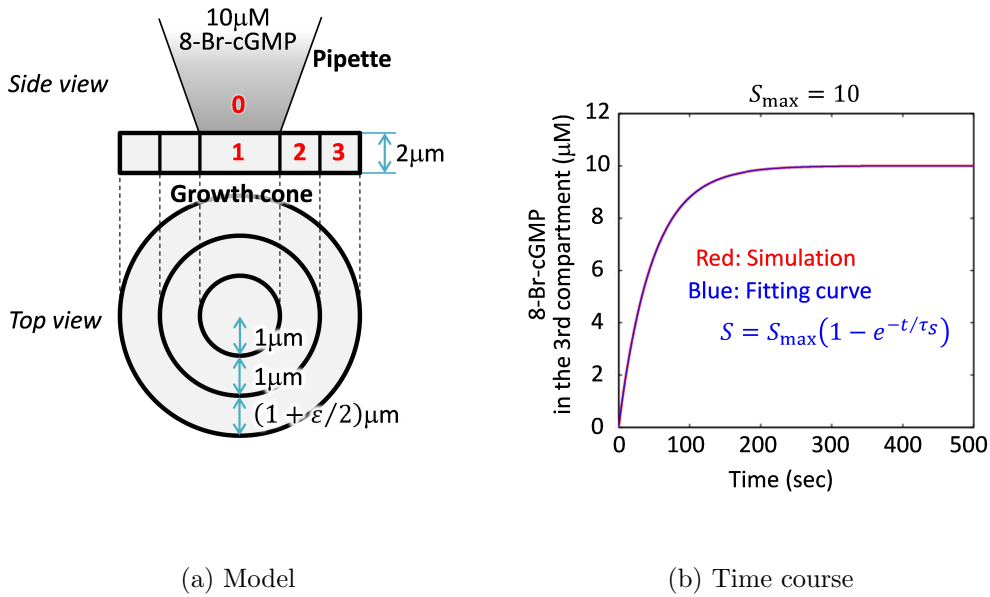
Figure 17. Computation of the 8-Br-cGMP diffusion. (a) Growth cone model for computing time series of diffusing 8-Br-cGMP. It is composed of three shell-shaped compartments and the 0-th compartment is a pipette which is filled with abundant 10 μM 8-Br-cGMP. To express size variation of growth cone, Gaussian noise $\epsilon \ N(0, 1^2)$ was added to the thickness of the outmost shell compartment (compartment #3). (b) Sample time series of the concentration in the compartment #3 (red line) by solving the diffusion equation (Eq. (2)). Blue line is the exponential function (Eq. (1)) fitted to the red one.

As we described, injection of 8-Br-cGMP from a pipette raises intracellular 8-Br-cGMP concentration in the growth cone by diffusion. This process is rate-limiting and its time series is well fitted to simple exponential function of time (Eq. (1)). We used this simple function instead of calculating the diffusion

equation.

3.3 Formulation of 8-Br-cGMP binding and downstream factors

Diffused 8-Br-cGMP first binds to CNGC and PKG, which transduce the signal to their DFs (Fig. 13 and 18) by the chain of chemical reaction (Appendix A). If we express the normalized activity levels of CNGC- and PKG-DFs as X and Y, respectively, their reaction rates are faster than the change rate of MPTS. By applying quasi-steady state approximation, they can be written as the Hill-type functions of S,

$$X = \frac{S^n}{K_X + S^n}, \qquad\qquad K_X = \frac{k_{Xb}}{k_{Xf}} \qquad\qquad (3)$$

$$Y = \frac{S^m}{K_Y + S^m}, \qquad\qquad K_Y = \frac{k_{Yb}}{k_{Yf}}, \qquad\qquad (4)$$

where Hill coefficients, n and m, are effective ones incorporating influence of the downstream processes, rather than the numbers of 8-Br-cGMP binding sites. K_X and K_Y are dissociation parameters defined by the ratio of backward rate (k_{Xb} and k_{Yb}) to forward rate (k_{Xf} and k_{Yf}) of the reactions.
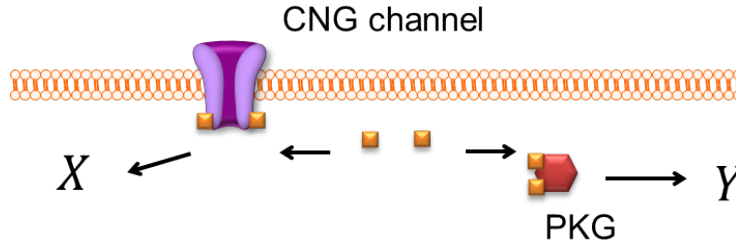
Figure 18. 8-Br-cGMP binding and DFs. 8-Br-cGMP is bound by CNGC and PKG, and then the signals are transduced to the DFs by chain of chemical reactions

To describe interactions between CNGC-DF and PKG-DF (parallel blue arrows in Fig. 13), the linear approximation was applied to forward and backward reaction rates in Eqs. (3) and (4). Such a linear approximation is acceptable

22

because the MP shows a small shift (within several tens mV) from the potential before the 8-Br-cGMP stimulation and gradual change during the entire observation period, implying that strong nonlinearities are unlikely to exist in these interactions. The effective dissociation parameters, K_X and K_Y, can be written as,

$$K_X(Y) = \frac{k_{Xb0} + k_{XbY}Y}{k_{Xf0} + k_{XfY}Y} \tag{5}$$

$$K_Y(X) = \frac{k_{Yb0} + k_{YbX}X}{k_{Yf0} + k_{YfX}X}, \tag{6}$$

where the denominator and numerator represent the forward and backward reactions, respectively. The values of the eight parameters reflect the structure of model pathways; for instance, only k_{XbY} takes zero in Eq. (5) if PKG-DF accelerates the forward reaction of CNGC-DF, and both of k_{XfY} and k_{XbY} are zero if PKG-DF has no effect on CNGC-DF. Therefore, each of these equations can represent three possible interactions, activation, inhibition, and no interaction from one to another, and thus the total of nine interactions can be considered (parallel blue pathways in Fig. 13). In the estimation step, we normalized these parameters by dividing them by k_{Xf0} or k_{Yf0} to reduce the number of parameters (see Table 1).

3.4 Formulation of ion channel regulation

CNGC- and PKG-DFs regulate the effective densities of chloride and sodium channels (ClC and NaC, respectively; Fig. 13 and 19). The regulation of both ClC by CNGC-DF and NaC by PKG-DF are known (black arrows in Fig. 13) whereas their cross interactions are unknown (cross blue arrows in Fig. 13). Here we modeled the dynamics of the difference in the effective channel densities between with and without regulation, by defining the normalized change ratios of the channel densities, Z for ClC and W for NaC. They take constant values if no regulations occur (no 8-Br-cGMP stimulation) but variable values if they are modulated when 8-Br-cGMP stimulation occurs. Due to rate-limitations, we

described their dynamics by ODEs of the biochemical reactions,

$$\frac{dZ}{dt} = k_{Zf}(1 - Z) - k_{Zb}Z \tag{7}$$

$$\frac{dW}{dt} = k_{Wf}(1 - W) - k_{Wb}W, \tag{8}$$

where k with a subscript is the effective forward or backward reaction rate which is either a constant or a function of the normalized variables, X and/or Y, similar to the case of the interactions between X and Y.
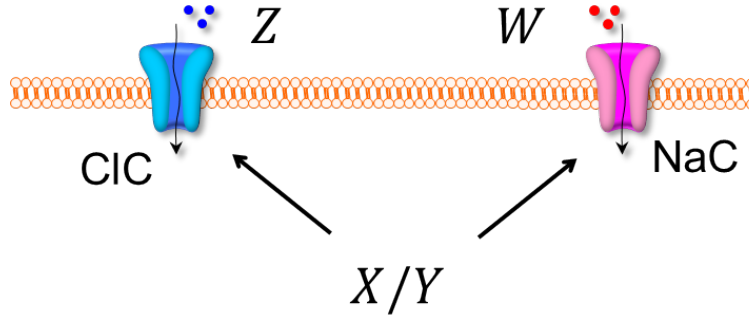
Figure 19. Ion channel regulation. CNGC- and PKG-DFs modify the conductances of ion channels on the membrane.

The reaction rates in Eqs. (7) and (8) are regulated by CNGC- and PKG-DFs (cross blue arrows in Fig. 13). By the linear approximations, these regulations are described as,

$$k_{Zf}(X, Y) = k_{Zf0} + k_{ZfX}X + k_{ZfY}Y \tag{9}$$

$$k_{Zb}(X, Y) = k_{Zb0} + k_{ZbX}X + k_{ZbY}Y \tag{10}$$

$$k_{Wf}(X, Y) = k_{Wf0} + k_{WfX}X + k_{WfY}Y \tag{11}$$

$$k_{Wb}(X, Y) = k_{Wb0} + k_{WbX}X + k_{WbY}Y, \tag{12}$$

where the reaction rates, k_{**0}, are constant (X/Y independent) while the others are the maximum rates contributed by X or Y. In these four interactions, only ClC activation by CNGC-DF (k_{ZfX}) and NaC activation by PKG-DF (k_{WfY}) are known regulations, and the other interactions are unknown (Fig. 13). We excluded the simultaneous combinations of forward and backward regulations on

24

the same effectors. There are three possible effects (activation, inhibition, and no interaction) in each regulation, and then the total of nine interactions can be considered (3×3 interactions in cross blue arrows in Fig. 13).

In model equations (3)-(12), we replaced positive feedback loop pathway among the molecules to an effectively alternative feedforward pathway in the concerned system (Fig. 13). The reason for this is that positive feedback with some conditions leads the system's state to oscillation, bistability, or divergence. However, none of such highly unstable responses were observed in the trends of MPTS (Fig. 9 and Fig. 10). This implies that even if the system contains a feedback loop, it can be modeled effectively as a feedforward process.

3.5 Formulation of MP regulation

The signal conversion from the ion channel densities to the MP is well established by Hodgkin and Huxley [21] (Fig. 20 and Appendix B),

$$C_m \frac{dV}{dt} = -g_{\mathrm{Na}}(V - V_{\mathrm{Na}}) - g_{\mathrm{K}}(V - V_{\mathrm{K}}) - g_{\mathrm{Cl}}(V - V_{\mathrm{Cl}}), \tag{13}$$

with the membrane capacitance, C_m, the membrane potential (MP), V, the reversal potentials, V_{Na}, V_{K}, and V_{Cl}, and the channel conductances, g_{Na}, g_{K}, and g_{Cl}, which depends on channel density and/or MP. The potassium conductance at the resting state has a dominant permeability ($g_{\mathrm{K}} \gg g_{\mathrm{Cl}}, g_{\mathrm{Na}}$). By setting the left-hand side of Eq. (13) to zero, the MP at the steady state, including resting state before the 8-Br-cGMP stimulation, can be approximately given by

$$V = \frac{\bar{g}_{\mathrm{K}}V_{\mathrm{K}} + \bar{g}_{\mathrm{Cl}}V_{\mathrm{Cl}} + \bar{g}_{\mathrm{Na}}V_{\mathrm{Na}}}{\bar{g}_{\mathrm{K}} + \bar{g}_{\mathrm{Cl}} + \bar{g}_{\mathrm{Na}}} \approx \frac{\bar{g}_{\mathrm{K}}V_{\mathrm{K}} + \bar{g}_{\mathrm{Cl}}V_{\mathrm{Cl}} + \bar{g}_{\mathrm{Na}}V_{\mathrm{Na}}}{\bar{g}_{\mathrm{K}}}, \tag{14}$$

where $\bar{g}_{\mathrm{Na}}$, $\bar{g}_{\mathrm{K}}$, and $\bar{g}_{\mathrm{Cl}}$ are the conductance at the steady state. In our model, densities of NaC and ClC are modulated by DFs of 8-Br-cGMP. Approximating change of V around the resting potential, the MP with 8-Br-cGMP stimulation obeys the following formula,

$$\hat{V} = V_{\mathrm{K}} - A_Z Z + A_W W, \tag{15}$$

where $A_Z = (\bar{g}_{\mathrm{Cl}}|V_{\mathrm{Cl}}|)/\bar{g}_{\mathrm{K}}$ and $A_W = (\bar{g}_{\mathrm{Na}}V_{\mathrm{Na}})/\bar{g}_{\mathrm{K}}$ representing the maximum amplitude of MP shifts. The parameters, V_{K}, A_Z, and A_W, depend on cell characteristics, so each of them takes different values between cells. We define Z and

25

W as the normalized change of the effective densities of ClC and NaC induced by the 8-Br-cGMP stimulation, respectively; they take one if their changes reach their maximum. We ignored the dependency of Z and W on the MP because their changes are little with small MP shift around the resting potential [21].
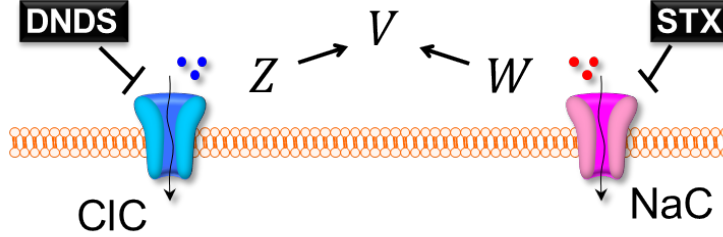
Figure 20. Membrane potential regulation and pharmacological agents. Modified ion channels regulate the inward current of the corresponding ionic molecules which lead to the change in MP.

In the experiments, phormacological agents are employed to measure the individual output of the ion channels (Fig. 20). This effects can be formulated as

$$\hat{V} = V_{\text{K}} - \eta_Z A_Z Z + \eta_W A_W W, \tag{16}$$

where η_Z (DNDS) and η_W (STX) are the effects of the channel blockers which are implemented as the gain parameters. Here, we assumed that the blockers completely suppress the channel functions, by setting these parameters to zero (blocker applied) or one (control).

3.6 Model candidates

We considered all possible interactions caused by CNGC- and PKG-DFs. There are four possible pathways (blue arrows in Fig. 13), where each pathway consists of three possible reactions, activation, inhibition, and no interaction. Therefore total of 81 possible models were examined; the simplest model has no interactions while the most complex model has four interactions, other than known interactions (Fig. 21 and 22). A complex model is compatible with simpler models.
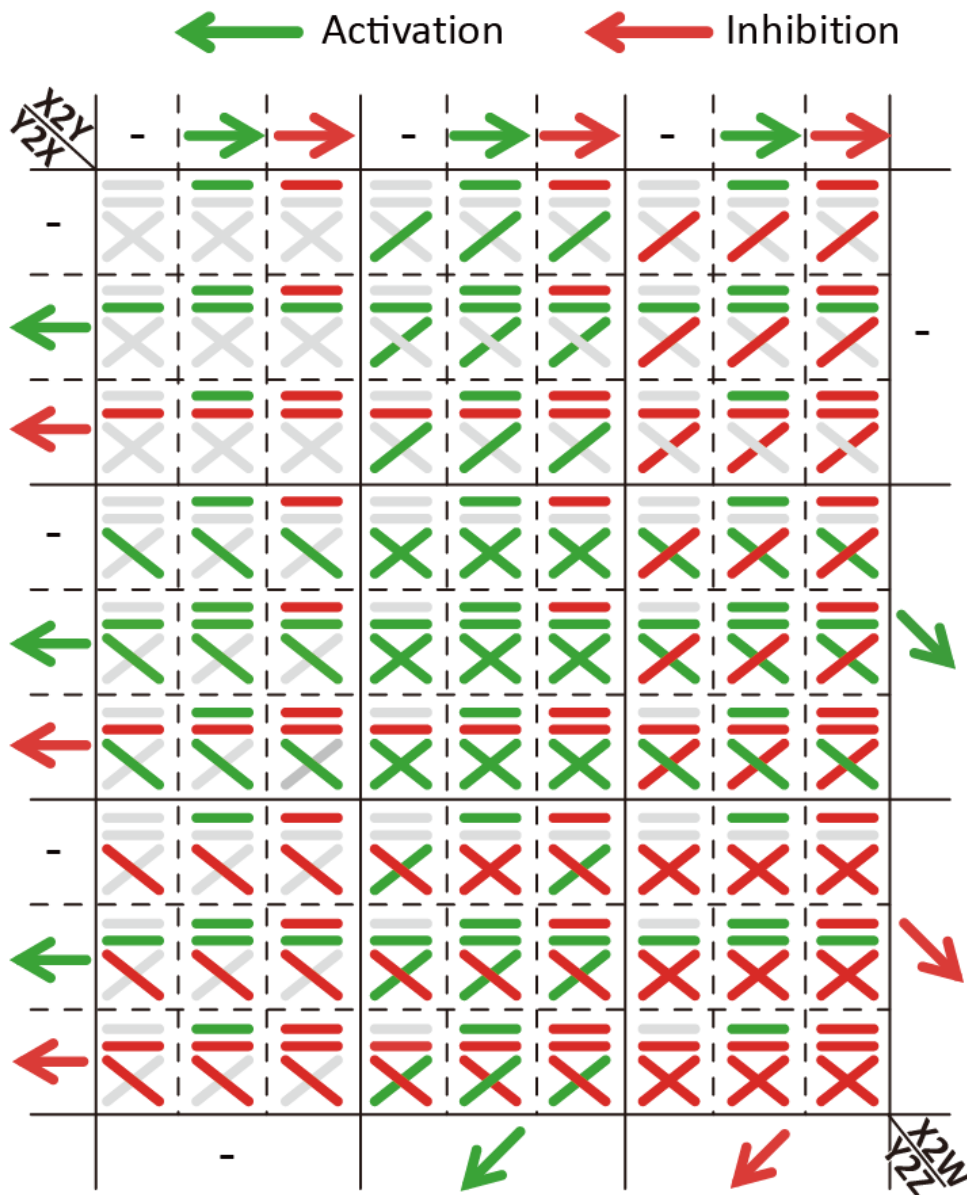
Figure 21. Summary of all the possible models incorporating or not incorporating blue pathways in the gray box in Fig. 13. For easy visibility, activation, inhibition, and no interaction are represented by green, red, and gray lines, respectively.
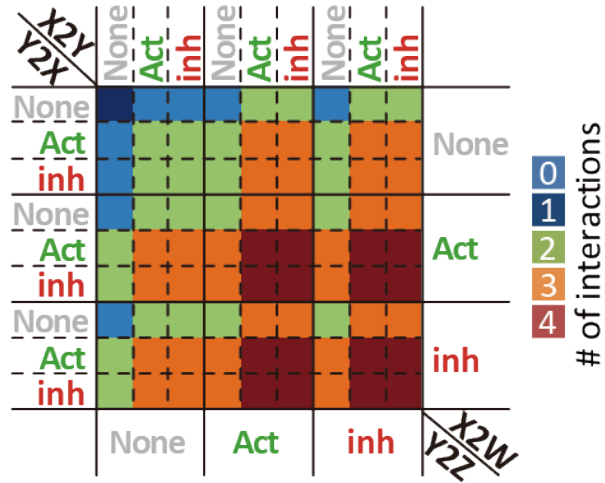
Figure 22. Summary of model complexity. Graded five colors indicate complexity levels of the system, which are defined as the number of possible interactions ranging from 0 to 4.

4. Bayesian formulation and model evaluation criteria

In this chapter, we formulate the posterior probability based on the Bayesian theorem. Then we introduce model evaluation criterion and model comparison criteria for model estimation and model validation, respectively.

4.1 Model parameters and posterior distribution

Since the actual MPTS is significantly affected by parameters such as growth cone volumes and ion channel densities, we adopted a Bayesian framework to model growth cone MPTS with the parameters categorized to three attributes; cell-dependent, cell-independent, and experimental condition (Table 1 and Fig. 23). The cell-dependent parameter set, $\phi = \{V_\mathrm{K}, A_Z, A_W, \tau_S\}$, is highly dependent on characteristics of individual neurons (e.g., growth cone volume, expression levels of ion channels and other physical conditions), which expresses the peripheral system of the intracellular molecular system. The parameters in this set were estimated from each MPTS. On the other hand, the parameters, in the cell-independent parameter set, θ (reaction rates, Hill coefficients, and means of A_Z and A_W), take a common value for any neurons, which expresses the core system of the intracellular molecular system. And they were estimated from all of the MP dataset. The experimental condition parameter set, $c = \{\eta_Z, \eta_W, S_\mathrm{max}\}$, represent the experimentally controllable parameters, such as pharmacological condition and 8-Br-cGMP concentration, and then were given at the initial setting (not necessary to be estimated).

Due to our configuration of model candidates (Fig. 22), complex models have more capability to fit the given dataset than the nested structures which correspond simpler ones. This makes it difficult to identify what part of the model structure is really important. To deal this problem, we designed the prior distributions as the complex model exhibits the capability of its nested structures with high probability when extra parameters with regards to its simpler model become 0. We implemented this strategy by utilizing the left-truncated Gaussians for most of the parameter priors (Table 1) with appropriate parameter ranges (σ). Then we employed Bayes' theorem to estimate the posterior distribution of
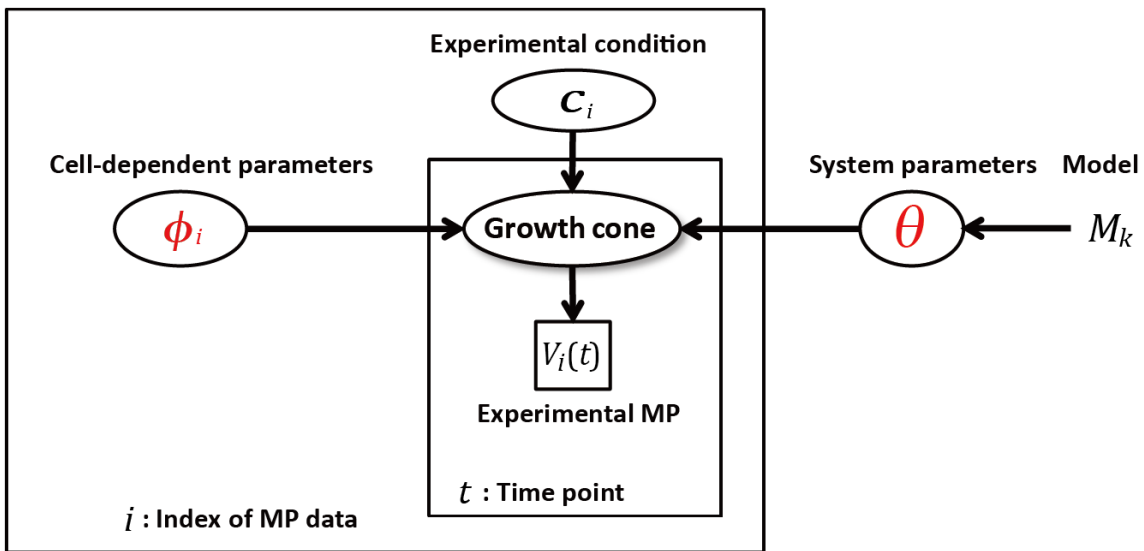
Figure 23. Graphical representation of relationship among MPTS, model variables, and parameters. Preprocessed MP value of the i-th data at time t ($V_{i,t}$) is characterized by three parameter sets, which are dependent on experimental condition (c_i), dependent on cell (ϕ_i), independent of cell (θ), in each deterministic model (M_k).

them as,

$$p(\phi, \theta | V, c) = \frac{p(V|\phi, \theta, c)p(\phi|\theta, c)p(\theta)}{\int\int p(V|\phi, \theta, c)p(\phi|\theta, c)p(\theta)d\phi d\theta}, \tag{17}$$

where the likelihood, fitness to the given dataset, is the product of the Gaussians,

$$p(V|\phi, \theta, c) = \prod_{i=1}^{I} p(V_i|\phi, \theta, c_i) \tag{18}$$

$$= \prod_{i=1}^{I}\prod_{t=1}^{T_i} \frac{1}{\sqrt{2\pi\sigma_i^2}} \exp\left(-\frac{\left(V_{i,t} - \hat{V}_t(\phi_i, \theta, c_i)\right)^2}{2\sigma_i^2}\right),$$

where $V = \{V_{i,t}\}$ is the set of observed MPTS and $V_{i,t}$ is the MP value of the i-th MPTS at time t ($I = 16$ and T_i=300–900), while $\hat{V}_t(\phi_i, \theta, c_i)$ is the MP value computed from the deterministic model above. The s.d. of the i-th MPTS, σ_i, was determined from the differences between the preprocessed and the smoothed data (Fig. 12).

To estimate posterior distribution (Eq. (17)), we determine the parameters of prior distribution by taking biophysical and phenomenological information into account. A left-truncated Gaussian is used for a prior distribution of the time constant, τ_S. The s.d. of the Gaussian (100 sec) is larger than the mean obtained by the Monte Carlo simulation. Similarly, all the other parameters in ϕ and θ except for V_K have the non-negative probability distributions were given as summarized in Table 1. The hyperparameters of the prior distributions (μ, σ) were determined by following accounts:

1. Reversal potential of potassium channel (V_K; ϕ) was set to around -60mV in the experiment.

2. According to the MPTS (Fig. 9), maximum MP shifts by ClC (A_Z; ϕ) and NaC (A_W; ϕ) depends on each cell, but we assumed their means take cell-independent value (A_{Z0} and A_{W0}; θ). The time constants of the MP shift ($1/k_Z$ and $1/k_W$; θ), can be more than a second according to the data.

3. The numbers of binding sites of cGMP to CNGC and PKG are four and four, respectively [23, 24]. These values are not corresponding to the Hill

31

coefficients of Eqs. (3) and (4) in the text (n and m; θ). Generally, if the relaxation time of a reaction (backward reaction) is slow compared to an input rate (forward reaction), the system acts as a low-pass filter and the input signal is degraded [25], which corresponds to input-output relation with small Hill coefficient. Here we set the prior distributions for n and m to the non-negative part of the Gaussians, $N(0, 1^2)$, whose s.d. is smaller than the number of the binding sites.

4. The distributions of the ratio, k_{X*}/k_{Xf0} and k_{Y*}/k_{Yf0} in Eqs. (5) and (6), become Cauchy distributions because all the rates are given as Gaussians.

Table 1. Summary of prior distributions of model parameters. Parameter types; ϕ (cell-dependent), θ (cell-independent), and c (experimental condition). Distributions; N (Gaussian), N_+ (left-truncated Gaussian), and N_+/N_+ (left-truncated standard Cauchy distribution). Asterisk(*) represents f, b, X or Y.

Parameter	Description	Unit	Equation	Type	Distribution	μ	σ
V_{K}	Reversal potential of K^+ channel	mV		ϕ	N	-60	5
A_{Z0}	Mean of A_Z	mV	-	θ	N_+	0	10
A_{W0}	Mean of A_W	mV	-	θ	N_+	0	10
A_Z	Maximum MP shift by ClC	mV		ϕ	N_+	A_{Z0}	5
A_W	Maximum MP shift by NaC	mV		ϕ	N_+	A_{W0}	5
k_{Z*}, k_{W*}	Reaction rates of Z, W	$\sec^{-1}$		θ	N_+	0	1
n, m	Hill coefficient	-		θ	N_+	0	1
k_{Xf*}/k_{Xf0} k_{Xb*}/k_{Xf0}	Normalized reaction rates of X	- $(\mu M)^n$		θ	N_+/N_+	-	-
k_{Yf*}/k_{Yf0} k_{Yb*}/k_{Yf0}	Normalized reaction rates of Y	- $(\mu M)^m$		θ	N_+/N_+	-	-
τ_S	Time constant	sec		ϕ	N_+	0	100

Parameter	Description	Unit	Equation	Type	Value		
η_Z	DNDS application	-		c	0 (applied) or 1 (control)		
η_W	STX application	-		c	0 (applied) or 1 (control)		
$S_{\max}$	Maximum conc. of 8-Br-cGMP	μM		c	10 or 5		

33

4.2 Model evaluation criteria

In the system identification, we need the criterion to evaluate some index such as the plausibility by the subjectve view point. For this reason, we introduce criteria for both identification step and validation step based on Bayesian perspective.

4.2.1 Evaluation criterion in model selection step

In the Bayesian formulation above, model M_k is omitted because it is trivial. If we denote the denominator of Eq. (17) by $p(V|M_k)$, then,

$$p(M_k|V) \propto p(V|M_k)p(M_k), \tag{19}$$

represents model plausibility for the given dataset V according to the Bayesian theorem. Here, considering $p(M_k)$ is uniform distribution, i.e., all the models are equally plausible beforehand, $p(V|M_k)$ also represents model plausibility. This is the reason why the denominator of Bayesian formulation is called evidence.

Therefore, A log-evidence (logarithm of marginal likelihood) was employed for a criterion to compare the model plausibility. For estimating the log-evidence by sampling from the posterior distribution, we practically carried out Markov Chain Monte Carlo (MCMC) simulation (Appendix C) [19, 26], in which resultant posterior samples reflect the bias stemming from the priors. Using the parameters given as MCMC samples, we evaluated the model plausibility (model criterion, see below) for 81 possible models by regarding larger log-evidence as more plausible. We introduced Ogata's method [27, 28] for calculating the log-evidence (Appendix D). The log-evidence of the k-th model, $E(M_k)$, is described as follows with the MCMC-based posterior parameters plugged-in over the entire MP time-series,

$$E(M_k) = \sum_{l=1}^{L-1} \log \sum_{b=1}^{B} \exp \left\{ (\beta^{l+1} - \beta^l) \sum_{i=1}^{I} \log p\left(V_i|\theta_i^{b,l}\right) \right\} - (L-1)\log B, \tag{20}$$

where $\theta_i^{(b,l)} = \{\phi_i^{(b,l)}, \theta^{(b,l)}, c_i\}$ represents the b-th set of parameter samples from MCMC simulation (total $B > 10^5$)(Appendix D) and β^l is l-th inverse temperature of Ogata's method ($L = 16$) with equal interval. We also computed mean values of the samples leading to mean a posteriori (MAP) parameters, which were used for reconstructing MP time series in the section of model validation.

4.2.2 Comparison criterion in model validation step

Logarithm of likelihood (log-likelihood) is also employed as a criterion for comparing the fitness or expressiveness for the given dataset among the model candidates,

$$
e(M_k) = \sum_{i=1}^{I} \left\{ \frac{1}{T_i} \frac{1}{B} \sum_{b=1}^{B} \log p(V_i|\theta_i^{(b)}) \right\}, \tag{21}
$$

where $\theta_i^{(b)} = \{\phi_i^{(b)}, \theta^{(b)}, c_i\}$ represents MCMC sample. The log-likelihood takes larger value if the given model outperforms to represent the given dataset than the other models. We confirmed whether the model criterion depends on the dataset or not, by performing Leave One Out Cross Validation (LOO-CV; see Appendix E) procedure, which evaluates the reproduction ability of a single MPTS by our model that employed the MAP parameters estimated from the remaining set of time series. We also examined generalization capability of each model by estimating the log-likelihood for 5 μM 8-Br-cGMP dataset.

5. Result

5.1 PKG-mediated ClC inhibition is essential for MP-shifts

We evaluated the model criterion, the log-evidence Eq. (20), for all the 81 model candidates based on 10 μM 8-Br-cGMP dataset (Control/+STX/+DNDS conditions) (Fig. 24). According to the log-evidence, models are clearly separated into two groups (a and b in Fig. 24). If we term the simplest model in group a as M_a and that in group b as M_b (Fig. 24), both of M_a and M_b are minimal in the structure within their respective model groups. This suggests that M_a and M_b are core models in respective groups and that PKG-mediated ClC inhibition (inset of Fig. 24 and 25) is the most essential pathway for reconstructing the given dataset. In fact, the MPTS were well reconstructed by the model M_b with the MAP parameters (Fig. 26 and Table 2).

Checking the stability of the identification by the LOO-CV procedure (Fig. 27), we also confirmed that the models in group b are superior to those in group a, irrespective of the applied MPTS dataset (Fig. 28). We also examined whether the estimated parameters of the model M_b do not over-fit to the employed MP dataset by comparison with the MAP parameters estimated during the LOO-CV procedure (Table 2); the MAP parameters estimated based on whole 10 μM 8-Br-cGMP dataset were all within the mean $\pm$s.d. of the set of MAP parameters over the LOO-CV procedure, suggesting the robustness of the parameter estimation by system-identification method regardless of the employed MP dataset.

(a) Log-evidence map   (b) Histogram
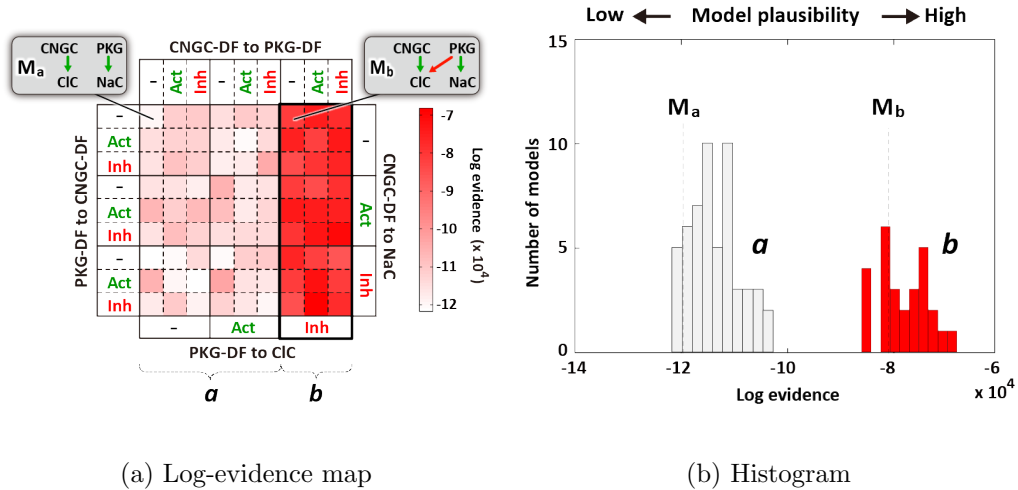
Figure 24. Comparison of model plausibility (10 μM). (a) Model plausibility was represented in terms of log-evidence (in color) computed from the MPTS given by 10 μM 8-Br-cGMP stimulation in all the conditions (control/+DNDS/+STX). The matrix representation of model configurations follows Fig. 21. A model with a large log-evidence (red) is more plausible than that with a small log-evidence (white). Models with no interaction or activation (rather than inhibition) from Y to Z are labeled as model group a, while those with inhibition (rather than activation) from Y to Z as model group b. Models indicated by M_a and M_b (insets) contain fewest and common interactions in groups a and b, respectively; that is, they are minimal models in respective model groups. (b) Histograms of log-evidence in the model groups a and b, showing that log-evidence in group b is substantially larger than that in group a. Vertical dashed lines indicate the log-evidence values of the two representative models, M_a and M_b.
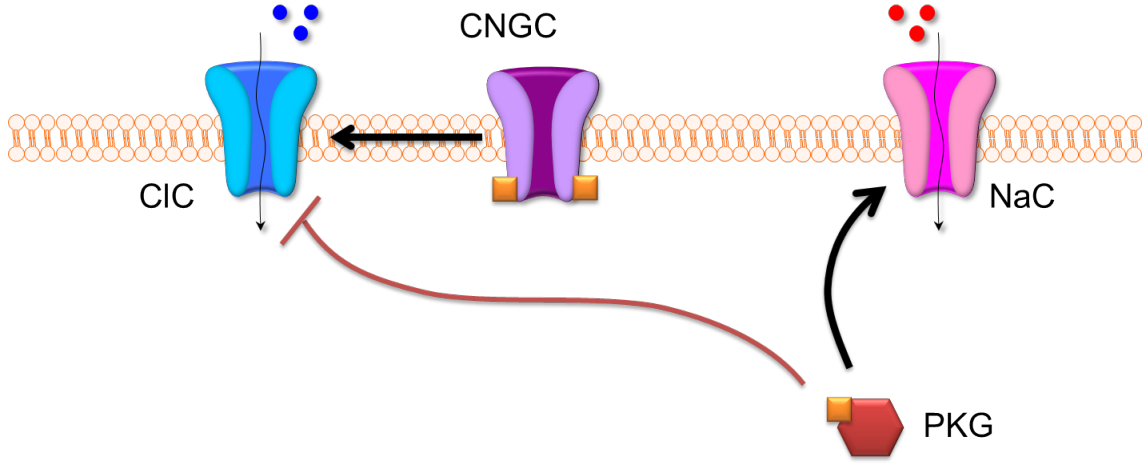
Figure 25. Identified minimal core model M_b. The model is consisted with known interaction (ClC activation by CNGC-DF and NaC activation by PKG-DF) and estimated interaction, ClC inhibition by PKG-DF.

Table 2. Comparison of estimated parameters θ for model M_b; mean a posteriori values by all the dataset and mean±s.d. of mean a posteriori values by LOO-CV.

Parameter	All dataset	LOO-CV
n	2.26	2.19±0.59
m	2.61	2.58±0.18
$\sqrt[n]{K_X}$	1.15	1.07±0.61
$\sqrt[m]{K_Y}$	16.61	15.87±4.49
k_{Zf0}	0.04	0.04±0.03
k_{Zb0}	0.16	0.15±0.11
k_{Wf0}	0.00	0.04±0.15
k_{Wb0}	0.18	0.60±0.90
k_{ZfX}	1.06	1.69±2.62
k_{ZbY}	8.22	6.72±3.66
k_{WfY}	5.74	6.02±4.00
A_{Z0}	36.87	41.46±11.98
A_{W0}	30.41	40.99±26.27
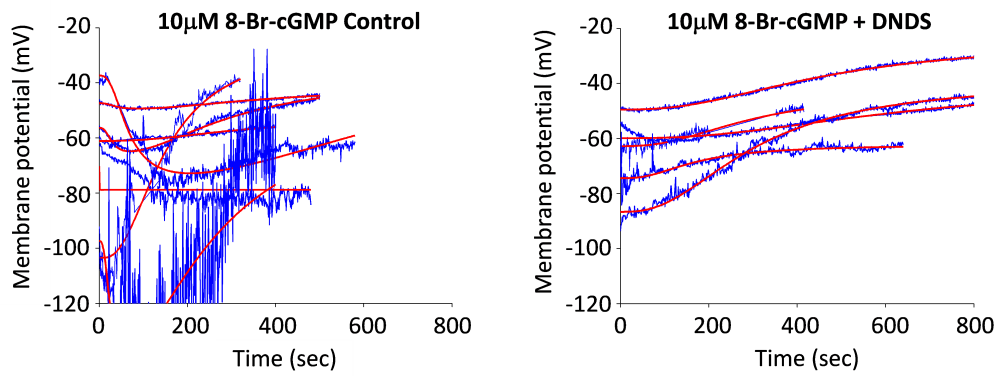
(a) Control

(b) DNDS
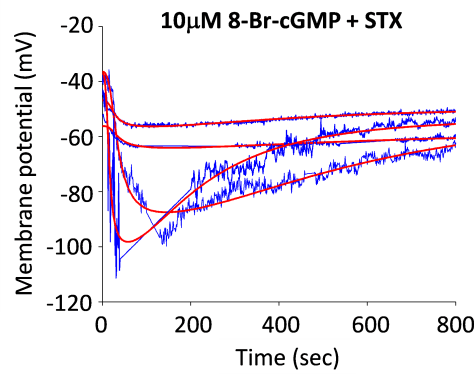
(c) STX

Figure 26. Reconstructed MPTS. Reconstruction of MP time series generated by the model with 10 μM 8-Br-cGMP in control (a), in presence of DNDS (b), and in presence of STX (c). Reconstructed data (red) showed good agreement with experimental data (blue) in Fig. 9. Mean a posteriori parameters were used (Table. 2).
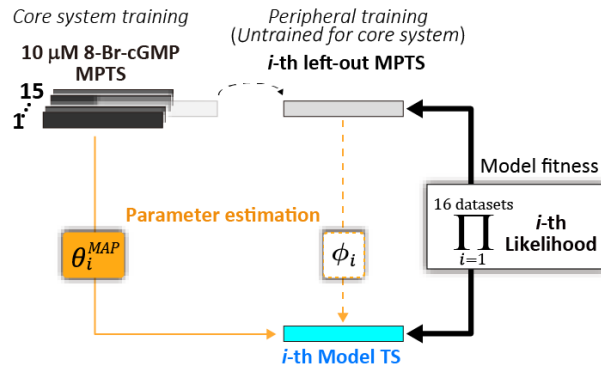
Figure 27. Schematic procedure of the leave-one-out (LOO) data selection for cross validation. Black and gray horizontal bars represent experimental MPTS. When the i-th MPTS is left out from the 10 μM 8-Br-cGMP datasets (n=16), it is used to estimate peripheral parameters (ϕ_i) and validate model fitness (product of all the likelihood in white box), whereas the remaining MPTS (n=15) are completely separated from the left-out MPTS to estimate the MAP core system parameters (θ_i^{MAP}). Repeating LOO for all the datasets, 16 system parameter sets are obtained (θ_i^{MAP};i=1,$\cdots$,16). The model fitness to the left-out MPTS was taken the logarithm and averaged over all the data points.
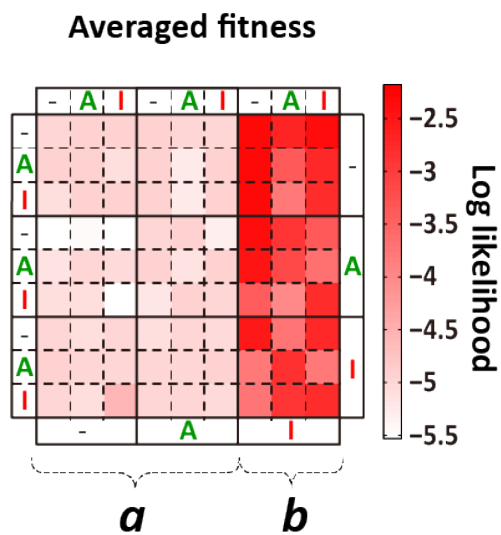
Figure 28. Fitness (Log-likelihood) map of leave one out cross validation for 10 μM 8-Br-cGMP dataset. This result supported that the model group b is superior to the model group a, showing stability of our system-identification methodology.

5.2 Specificity test for different conditions

Next, we tested the specificity of the identified model to the given 10 μM 8-Br-cGMP-induced MPTS datasets (Fig. 9). To test the degree of specificity, we introduced untrained datasets of MPTSs induced under different experimental conditions (Fig. 29). When we introduced the MPTS induced by a different concentration of stimulant 5 μM, instead of 10 μM 8-Br-cGMP, the model fitness matrix shows a similar pattern of model groups, as in Fig. 30 (a). This might indicate that core system parameters have a small dependency on the stimulus intensity between 5 μM and 10 μM concentration of 8-Br-cGMP. In contrast, when we introduced the MPTS induced by 10 μM 8-Br-cGMP in the presence of KT5823 (n=5), a PKG activity inhibitor that abolishes the PKG activation in the core system, the model fitness matrix shows the total absence of the distinct segregation between the model group a and b (Fig. 30 (b)). This supports that PKG activation is required for the pattern of model groups observed in Fig. 24, at least partially. Taken together, these tests indicate that our system identification method has high specificity.
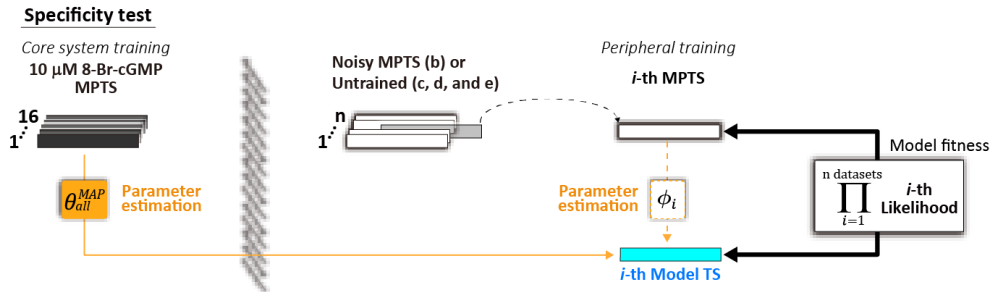
Figure 29. Schematic procedure of model specificity test using MPTSs in different conditions using MAP core system parameters estimated from all the 10 μM 8-Br-cGMP datasets (θ_{all}^{MAP}; n=16; black horizontal bars in the left side of the shaded band). The model likelihood to the i-th MPTS from other testing datasets (white bars) was computed while computing the cell-dependent parameters, ϕ_i, and by repeating this the model fitness to all the n datasets were obtained. Testing datasets are the datasets untrained by the core system (5 μM 8-Br-cGMP and 10 μM 8-Br-cGMP+KT5823).

(a) 5 μM 8-Br-cGMP

(b) 10 μM 8-Br-cGMP +KT5823
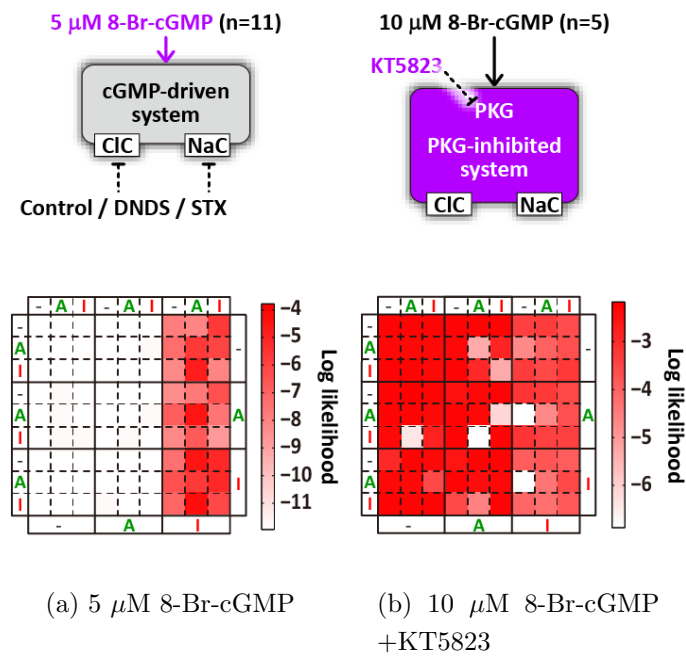
Figure 30. (a) Matrix representations of model fitness by the datasets derived by the control core system with 5 μM 8-Br-cGMP-induced MPTS (c; total n=11; control: n=2; with DNDS blocking ClC: n=5; with STX blocking NaC: n=4). (b) Same as a, but the datasets derived by the different core systems; 10 μM 8-Br-cGMP-induced MPTS under the PKG-inhibited condition (KT5823).

5.3 Predictability test by data assimilation

If we know the model parameters, we can predict how the MP behaves toward the future with high certainty because of the system noise-free character of the deterministic model, Eqs. (3)-(16). In our model, however, there are cell-dependent parameters (ϕ_i), which can be estimated only from the data from the cell. To perform behavior prediction even in such case, we used an idea from the technique called data assimilation which has been getting popular in earth science [26]. More specifically, when predicting MPTS, the cell-independent parameters were estimated from the LOO-CV (for 10 μM) or whole 10 μM dataset (for 5 μM), and the cell-dependent parameters were estimated based on the data points in the early phase of the time series to be predicted (Fig. 31 and 33). The surprisingly good prediction ability (red lines in Fig. 32(a) and 34(a)) suggests the reliability of the model we estimated, and moreover, significantly better prediction ability of the model group b than that of the group a (Fig. 32(b) and 34(b)) again shows the plausibility of the model configuration in b.
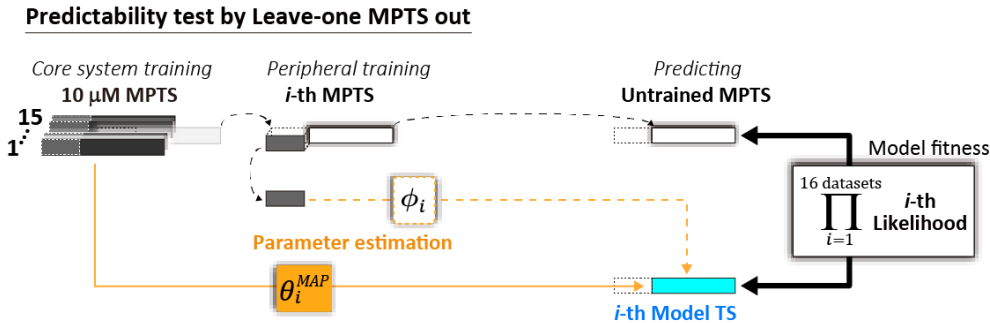
Figure 31. Schematic illustration of the procedure for testing model predictability for the late-phase of MPTS induced by 10 μM 8-Br-cGMP. The initial-phase (250 sec) of the left-out i-th MPTS was extracted to estimate the MAP peripheral parameters (ϕ_i), and the remaining i-th MPTS (after 250 sec, late-phase) was used to test the model predictability. The MAP core system parameters were estimated from the initial and late phase of the remaining MPTS (θ_i^{MAP}; n=15).
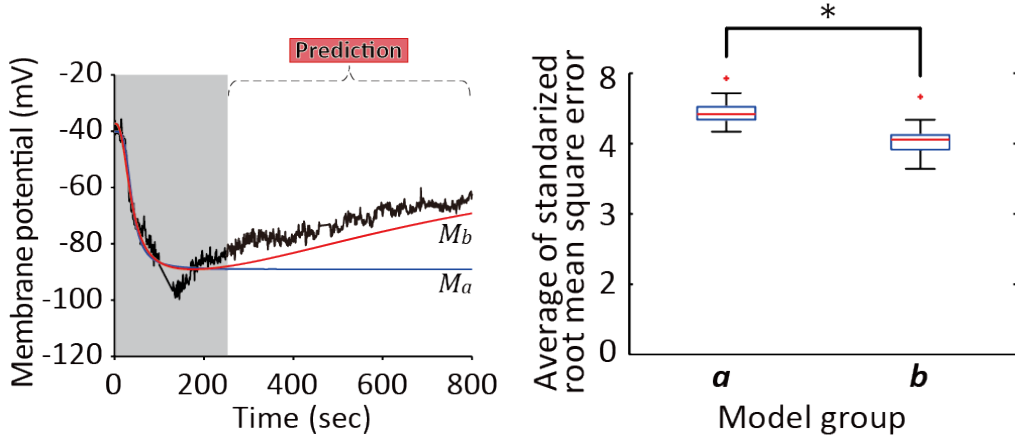
(a) Time course   (b) Comparison

Figure 32. (a) Sample time series predicted by model group a (white) and b (red) for 10 μM 8-Br-cGMP stimulations. Black lines represent experimental data. MP time series in the gray regions (up to 250 seconds) were used for cell-dependent parameter estimation, but those in the non-gray regions (after 250 seconds) were never used for the parameter estimation. (b) Box plots of prediction error for 10 μM 8-Br-cGMP stimulations. The prediction error was given by standardized root mean square error, $(1/I) \sum_{i=1}^{I} \left\{ (1/T_p) \sum_{t=251}^{T_p+250} \sqrt{\left(V_{i,t} - \hat{V}_t(\phi_i, \theta, c_i) \right)^2} / \sigma_i \right\}$, where T_p is the number of data points in later phase of the MPTS to be predicted (i.e., after 250 seconds). The errors depending on the models were distributed in each group and thus represented by box plots to show the difference between the two groups (red bar); student's t-test with unequal variances, *P¡0.01.
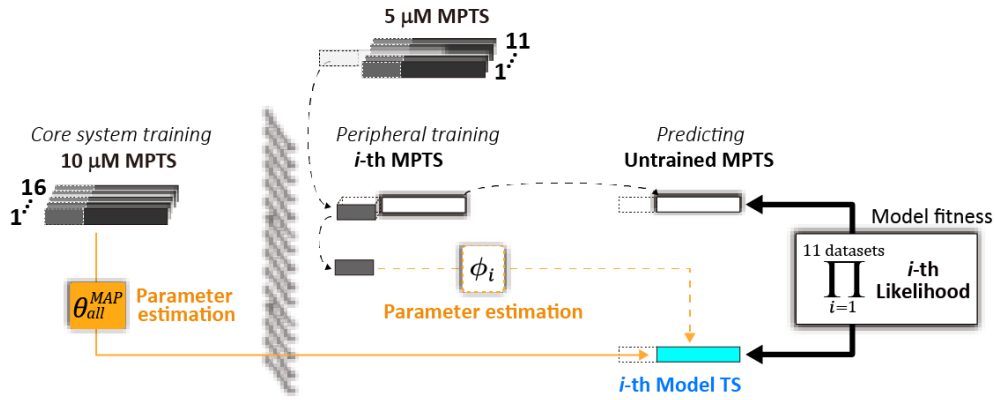
Figure 33. Same as Fig. 31, but 5 μM 8-Br-cGMP MPTS was predicted with ϕ_i given the initial-phase MPTS (up to 250 sec) and θ_{all}^{MAP} by all the 10 μM 8-Br-cGMP MPTS. Black and white bars represent experimental MPTS. MPTS in the white regions (after 250 sec, the late-phase) were never used for the parameter estimation; these MPTSs were completely untrained data sets even for the peripheral parameters.

(a) Time course

(b) Comparison
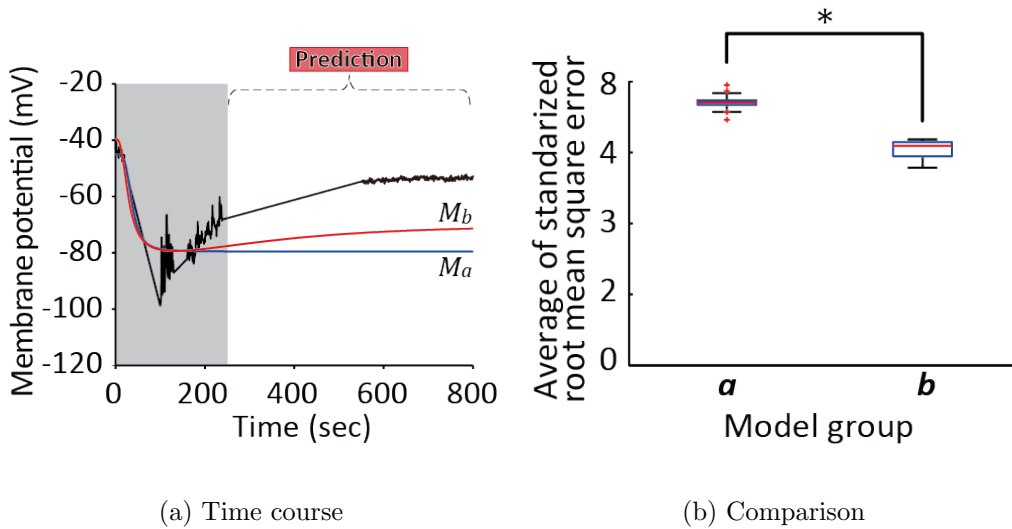
Figure 34. Same as Fig. 32 except for using 5 μM 8-Br-cGMP dataset

5.4 Reproduction of bidirectional MP shift responding to cGMP level

Here, we examined how the estimated model shows MP shift in its steady state (60 min after 8-Br-cGMP stimulation in the experiment). Experiments showed that under steady state of Sema3A application, the growth cone produces bidirectional MP shift that is dependent on the cGMP level [11].

To examine whether the core model M_b in group b can show the bidirectional MP shift in the steady state, we computed time series when the model M_b was stimulated by various levels of 8-Br-cGMP (Fig. 35). The model time series showed different patterns; (i) gradual hyperpolarization (0.5 μM), (ii) sharp hyperpolarization and gradual recovery to the resting level (6 μM), and (iii) sharp hyperpolarization and depolarization (20 μM). We defined the model's steady-state MP shift as the difference between the MPs at times zero and infinity, namely, $V(t \to \infty | \theta^{\mathrm{MAP}}) - V(t = 0 | \theta^{\mathrm{MAP}})$, where the cell-independent parameters (θ^{MAP}) were given as the MAPs estimated from the 10 μM 8-Br-cGMP dataset. For the cell-dependent parameters, V_K is a mere constant bias in $V(t)$ (Eq. (16)) and therefore was canceled out, τ_S was ignorable at times zero and infinity, and A_Z and A_W were replaced by their means in θ^{MAP}. Using this definition, the cGMP level-dependent bimodal steady-state MP shifts (Fig. 36(a)) appear, as observed in experiment [11]. Our identified model M_b also expresses that the decomposition of the MP shift into those induced by ClC and NaC. In the range of lower cGMP, MP shift is induced by the ClC-dependent hyperpolarization. On the other hand, the MP shift in the range of higher cGMP level is formed by the NaC-dependent depolarization, as summarized in Fig. 36(b). These results suggest that the model, M_b, with the estimated model parameters expresses a general mechanism of this system in a satisfactory manner.
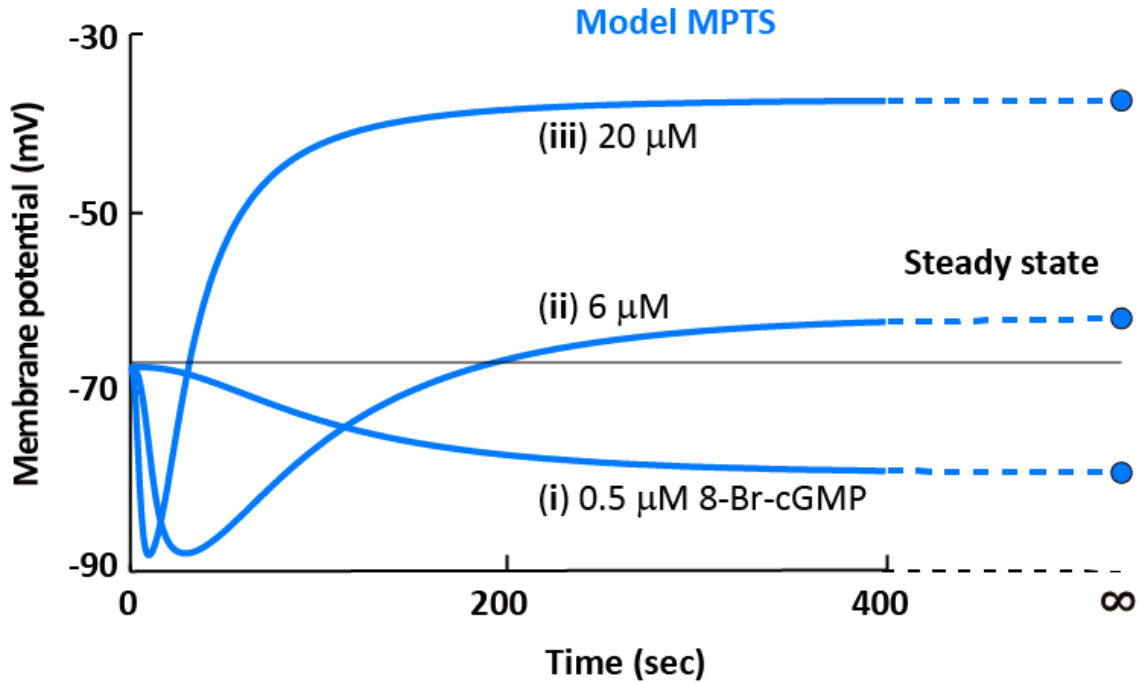
Figure 35. Reconstructed MPTS with several 8-Br-cGMP levels; (i) 0.5, (ii) 6, and (iii) 20 μM. $V_K = -60$, $\tau_S = 100$, and mean a posteriori cell-independent parameters (A_Z and A_W in the cell-dependent parameters are replaced by its mean, A_{Z0} and A_{W0} in the cell-independent parameters, respectively.) were used in control condition of model M_b in Fig. 24. Each time series reaches the steady-state level at about 300 sec which is three times of the diffusion time constant of 8-Br-cGMP ($\tau_S = 100$ (sec)).

(a) MPS

(b) Mechanism
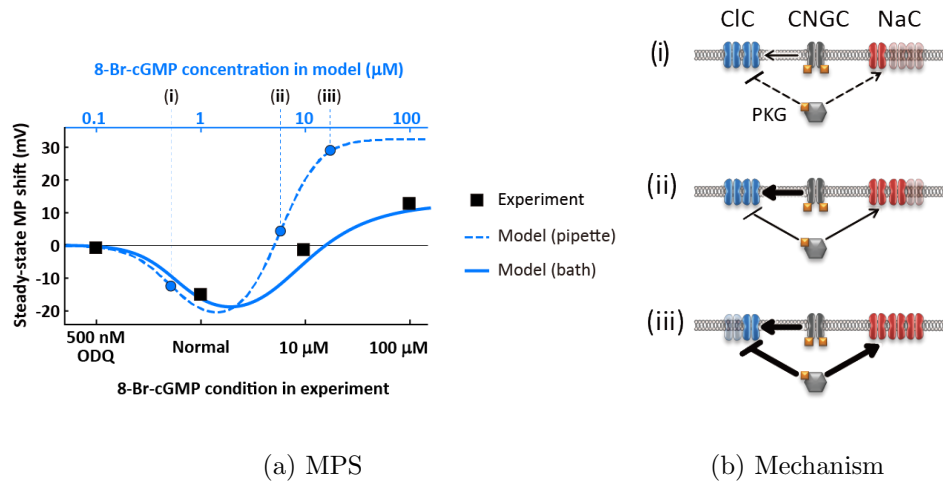
Figure 36. (a) Predicted dependency of steady-state MP shift (blue) on cGMP levels with the experimental data (black square). Dashed blue line shows the model prediction with the intracellular cGMP perfusion. On the other hand, solid blue line shows the modified model prediction with the cGMP permeation model (Appendix F). Vertical dashed lines are corresponding to the three conditions shown in Fig. 35. (b) Schematic figures describing mechanism of cGMP-MP converter; (i) MP is hyperpolarized by high affinity CNGC-mediated ClC activation in a low cGMP level, (ii) MP is recovered to around resting level by PKG-mediated ClC inhibition and NaC activation in the middle range of cGMP level where CNGC-mediated ClC activation is saturated, and (iii) in a higher cGMP level, MP is depolarized by low affinity PKG-mediated ClC inhibition and NaC activation.

6. Discussion

By computationally analyzing the experimental MP shift time series recorded from the growth cone responding to an external signal, we were able to deduce a core molecular system that converts a biochemical signal to an electrophysiological one. Specifically, our system-identification methodology revealed that PKG-mediated ClC inhibition is an essential pathway that acts in concert with CNGC-mediated ClC activation and PKG-mediated NaC activation (Fig. 36(b)). Actually, even the known interactions are putative signaling pathway demonstrated by the experiments [11], i.e., the cascades are unknown. A second messenger, cGMP-activated PKG is known as a regulator of the Mitogen-activated Protein Kinase (MAPK) such as ERK and p38 [29]. Furthermore, MAPK acts as not only a regulator of transcription but also a regulator of membrane protein including ion channels. On the growth cone, a TTX-resistant sodium channel, Nav1.8 [30], we called NaC in this study, is expressed. In addition, consistent with our speculation from the experimental data, Nav1.8 is activated by p38 as a downstream factor of PKG [31]. Surprisingly, it is also reported that ERK and p38 inhibit calcium-dependent ClC [32] which is consistent with the interaction identified by our present study on the growth cone.

Our computational analysis suggests that these three interactions form the core system which converts a biochemical signaling of the cGMP pathway to an electrophysiological signal of MP shifts. The validation of the system stability by LOO-CV method (Fig. 28 and Table 2) proves that the resultant model must represent a minimal and core system. It is also proved that the identified model has an ability to predict time series with the cell-dependent parameters estimated from the data in initial phase of the time series. However, although our computational analysis revealed that model M_b represents a core and minimum required system for converting the biochemical signal to the electrophysiological signal (Fig. 24), it does not completely eliminate the existence of other potential signaling cascades that facilitate the MP shifts in the growth cone.

Validating the model with $5\mu M$ 8-Br-cGMP demonstrated that some models in model group b with a pathway from X to Y (activation or inhibition) had larger log-likelihoods than their minimal model M_b (Fig. 30a). This may suggest that another, a more complicated model could be better fitted than the

model M_b, which cannot be excluded by our current computational analysis. Such large log-likelihood by more complicated models than the model M_b might have occurred because larger number of model parameters would simply lead better expressiveness of the MP time series for $5\mu M$ 8-Br-cGMP.

The bidirectional cGMP-dependency of MP shifts (hyperpolarization to depolarization, Fig. 36(a)) showed that CNGC-mediated ClC activation is due to a high-affinity of cGMP ($\sqrt[n]{K_X} = 1.15$), while PKG-mediated NaC activation is due to a low-affinity of cGMP ($\sqrt[m]{K_Y} = 16.61$). The bidirectional phenomenon based on a difference of dissociation constants can be seen in the molecular system of synaptic plasticity, in which the phosphorylation of α-amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid (AMPA) receptors occurs as a competition between the kinase and phosphatase [33, 34]. Interestingly, unlike for synaptic plasticity, for the bidirectional cGMP-dependency of MP shifts, PKG not only facilitates membrane depolarization, but also inhibits hyperpolarization. This means that the PKG-mediated ClC inhibition works to expand the dynamic range of MP shift against the input of 0.1–$10\mu M$ cGMP stimulation, which may contribute to widening the dynamic range of the growth cone turning driven by external guidance cues. ClC inhibition may also be explained by ClC self-adaptation to cGMP level, which is independent of the status of other system components. In the case of ClC self-adaptation, however, due to the lack of harmony with other system components, the output of the whole system becomes relatively unstable in response to identical cGMP stimulation. This suggests that the adaptation by the PKG-mediated ClC inhibition is more feasible than ClC self-adaptation.

The computational analysis of the electrophysiological dataset, e.g. the growth cone MP shifts, in conjunction with well-established experimental methods provides a significant advantage because it allows the identification of system components that cannot be otherwise determined by solely molecular or biochemical methods. Moreover, electrophysiological techniques such as a patch clump method are well established [35], and the MP can be measured with high temporal resolution (up to 40 msec interval in our case). Although electrical data contain observable noise, the time series of MP measurements provides greater information about intracellular biochemical processes (on a time scale of minutes) that is otherwise difficult to obtain. The biophysical mechanism connecting the rele-

vant biochemical molecules to the establishment of the MP has been extensively studied [21]. As a result, it has become possible, with the use of computational analysis, to unveil the mechanism of signal conversion from biochemical molecular signals to biophysical MP shifts.

The present data-driven system identification analysis employed empirical and phenomenological equations to describe a quantitative relationship between input stimulus (cGMP) and output response (MP shift) of the growth cone, but did not explicitly incorporate all possible signaling molecules into the model. Equation converting between different physical quantities has been often derived from empirical-quantitative laws; for instance, general biochemical equation (e.g., Michaelis-Menten kinetics) quantitatively describe dynamics of molecular concentrations without explicitly incorporating detailed biophysics such as thermo-diffusion, collision, structural change of proteins, and bindings. The Hodgkin-Huxley equations [21] is a well-established successful example of such empirical-quantitative modeling of biological functions. Such simplified but still biologically plausible empirical-quantitative modeling is indeed effective to know essential biological structure underlying the observed time-series in a data-driven fashion.

The experimental and cell-to-cell variation of the data is one of the difficulties of the present study. Such a variation easily hides a core mechanism of the system because it is impossible to infer which parameter is the source of variation when the model parameter has redundancy. Thus, we employed Bayesian framework to constraint the expressiveness of the model by introducing the cell-independent parameters which are common among the given dataset. This constraint allows us to access the intracellular system parameter such as reaction rate and Hill coefficient by the estimation with the whole dataset, even if the number of data is insufficient. In our present analysis, the approach succeeded due to the relatively small number of the cell-dependent parameters, i.e., the variation of a set of time series is well expressed by the variation of the small number of parameters. Although the current study doesn't present a systematic approach to determine parameter's attribute such as cell-independent and cell-dependent, it suggests that it is important to characterize each parameter as common or uncommon among the data (Fig. 23) and also to incorporate the prior based on phenomenologically or physically consistent manner (Table 1).

7. Conclusion

We identified the molecular system which regulates MP shifts during the growth cone turning, from MP time series dataset. The result shows ClC inhibition by PKG-DFs likely exists. Moreover, the earlier biochemical studies about MAPK (ERK and p38) interaction are consistent with the result. This probably contribute to the successive studies of the growth cone guidance.

The present study provide the system-identification methodology which is consisted with empirical-quantitative modeling and parameter characterization. In the modeling part, we bridged electrophysiology and intracellular molecular system based on the well quantified equations such as Hodgkin-Huxley equation. In the estimation part, we modeled the experimental and cell-to-cell variation based on the Bayesian formulation. Based on the two main attempt, we successfully applied the method for electrophysiology data during growth cone turning. This tool benefit the biological experiments which are very time consuming in the measurements.

During navigation, the growth cone must reconvert an electrical signal of MP shifts into other type of signals, i.e., biochemical and molecular signals, to achieve a morphological change, e.g. turning direction. Our understanding of the process of signal conversion is substantially incomplete. The computational analysis we describe here will facilitate a better understanding of those mechanisms of signal conversion. As the growth cone bimodal behavior depends on Ca^{2+} concentration [36, 37, 38, 39], voltage-dependent calcium channels are one of the major components of the MP-Ca^{2+} conversion mechanism. A recent biophysical computational study showed that the growth cone bimodal behavior depends on CaMKII and CaN, the downstream factors of Ca^{2+} signal [40]. It should subsequently be possible to integrate these findings into the current model to further elucidate the mechanism of growth cone motility.

Acknowledgements

References

[1] Corey S Goodman. Mechanisms and Molecules that Control Growth Cone Guidance. *Annual Review of Neuroscience*, Vol. 19, No. 1, pp. 341–377, 1996.

[2] Bernhard K Mueller. GROWTH CONE GUIDANCE: First Steps Towards a Deeper Understanding. *Annual Review of Neuroscience*, Vol. 22, No. 1, pp. 351–388, 1999.

[3] Hong-jun Song, Guo-li Ming, and Mu-ming Poo. cAMP-induced switching in turning direction of nerve growth cones. *Nature*, Vol. 388, No. 6639, pp. 275–279, 1997.

[4] H Song, G Ming, Z He, M Lehmann, L McKerracher, M Tessier-Lavigne, and M Poo. Conversion of neuronal growth cone responses from repulsion to attraction by cyclic nucleotides. *Science (New York, N.Y.)*, Vol. 281, No. 5382, pp. 1515–8, September 1998.

[5] Guo-li Ming, Hong-jun Song, Benedikt Berninger, Christine E Holt, Marc Tessier-Lavigne, and Mu-ming Poo. cAMP-Dependent Growth Cone Guidance by Netrin-1. *Neuron*, Vol. 19, No. 6, pp. 1225–1235, December 1997.

[6] Makoto Nishiyama, Akemi Hoshino, Lily Tsai, John R Henley, Yoshio Goshima, Marc Tessier-Lavigne, Mu-Ming Poo, and Kyonsoo Hong. Cyclic AMP/GMP-dependent modulation of Ca2+ channels sets the polarity of nerve growth-cone turning. *Nature*, Vol. 423, No. 6943, pp. 990–5, June 2003.

[7] Hong-jun J Song and Mu-ming M Poo. Signal transduction underlying growth cone guidance by diffusible factors. *Current opinion in neurobiology*, Vol. 9, No. 3, pp. 355–363, June 1999.

[8] Kyonsoo Hong, Lindsay Hinck, Makoto Nishiyama, Mu-ming Poo, Marc Tessier-Lavigne, and Elke Stein. A Ligand-Gated Association between Cytoplasmic Domains of UNC5 and DCC Family Receptors Converts Netrin-Induced Growth Cone Attraction to Repulsion. *Cell*, Vol. 97, No. 7, pp. 927–941, June 1999.

[9] Timothy M Gomez and James Q Zheng. The molecular basis for calcium-dependent axon pathfinding. *Nature reviews. Neuroscience*, Vol. 7, No. 2, pp. 115–25, February 2006.

[10] Paul A Janmey. Phosphoinositides and Calcium as Regulators of Cellular Actin Assembly and Disassembly. *Annual Review of Physiology*, Vol. 56, No. 1, pp. 169–191, 1994.

[11] Makoto Nishiyama, Melanie J von Schimmelmann, Kazunobu Togashi, William M Findley, and Kyonsoo Hong. Membrane potential shifts caused by diffusible guidance signals direct growth-cone turning. *Nature neuroscience*, Vol. 11, No. 7, pp. 762–71, July 2008.

[12] Kazunobu Togashi, Melanie J von Schimmelmann, Makoto Nishiyama, Chae-Seok Lim, Norihiro Yoshida, Bokyoung Yun, Robert S Molday, Yoshio Goshima, and Kyonsoo Hong. Cyclic GMP-gated CNG channels function in Sema3A-induced growth cone repulsion. *Neuron*, Vol. 58, No. 5, pp. 694–707, June 2008.

[13] X Gu and N C Spitzer. Low-threshold Ca2+ current and its role in spontaneous elevations of intracellular Ca2+ in developing Xenopus neurons. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, Vol. 13, No. 11, pp. 4936–48, November 1993.

[14] Cristina Jiménez-González, Gerald J McLaren, and Nicholas Dale. Development of Ca2+-channel and BK-channel expression in embryos and larvae of Xenopus laevis. *The European journal of neuroscience*, Vol. 18, No. 8, pp. 2175–87, October 2003.

[15] Junichiro Yoshimoto and Kenji Doya. Bayesian System Identification of Molecular Cascades. In *ICONIP (1)*, pp. 614–624, 2007.

[16] E N Brown, L M Frank, D Tang, M C Quirk, and M A Wilson. A statistical paradigm for neural spike train decoding applied to position prediction from ensemble firing patterns of rat hippocampal place cells. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, Vol. 18, No. 18, pp. 7411–25, September 1998.

[17] Liam Paninski, Yashar Ahmadian, Daniel Gil Ferreira, Shinsuke Koyama, Kamiar Rahnama Rad, Michael Vidne, Joshua Vogelstein, and Wei Wu. A new look at state-space models for neural data. *Journal of computational neuroscience*, Vol. 29, No. 1-2, pp. 107–26, August 2010.

[18] Greg D Field, Jeffrey L Gauthier, Alexander Sher, Martin Greschner, Timothy A Machado, Lauren H Jepson, Jonathon Shlens, Deborah E Gunning, Keith Mathieson, Wladyslaw Dabrowski, Liam Paninski, Alan M Litke, and E J Chichilnisky. Functional connectivity in the retina at the resolution of photoreceptors. *Nature*, Vol. 467, No. 7316, pp. 673–7, October 2010.

[19] Yashar Ahmadian, Jonathan W Pillow, and Liam Paninski. Efficient Markov chain Monte Carlo methods for decoding neural spike trains. *Neural computation*, Vol. 23, No. 1, pp. 46–96, January 2011.

[20] Vladislav Vyshemirsky and Mark A Girolami. Bayesian ranking of biochemical system models. *Bioinformatics (Oxford, England)*, Vol. 24, No. 6, pp. 833–9, March 2008.

[21] Alan L HODGKIN and Andrew F HUXLEY. A quantitative description of membrane current and its application to conduction and excitation in nerve. *The Journal of physiology*, Vol. 117, No. 4, pp. 500–544, August 1952.

[22] Y Koutalos, R L Brown, J W Karpen, and K W Yau. Diffusion coefficient of the cyclic GMP analog 8-(fluoresceinyl)thioguanosine 3',5' cyclic monophosphate in the salamander rod outer segment. *Biophysical journal*, Vol. 69, No. 5, pp. 2163–7, November 1995.

[23] U Benjamin Kaupp and Reinhard Seifert. Cyclic nucleotide-gated ion channels. *Physiological reviews*, Vol. 82, No. 3, pp. 769–824, July 2002.

[24] J A Smith, R B Reed, S H Francis, K Grimes, and J D Corbin. Distinguishing the roles of the two different cGMP-binding sites for modulating phosphorylation of exogenous substrate (heterophosphorylation) and autophosphorylation of cGMP-dependent protein kinase. *The Journal of biological chemistry*, Vol. 275, No. 1, pp. 154–8, January 2000.

[25] Adam Arkin and John Ross. Statistical Construction of Chemical Reaction Mechanisms from Measured Time-Series. *The Journal of Physical Chemistry*, Vol. 99, No. 3, pp. 970–979, January 1995.

[26] Roger Daley. *Atmospheric Data Analysis*. Cambridge University Press, 1993.

[27] Yosihiko Ogata. A Monte Carlo method for an objective Bayesian procedure. *Annals of the Institute of Statistical Mathematics*, Vol. 42, No. 3, pp. 403–433, September 1990.

[28] Kenji Nagata, Seiji Sugita, and Masato Okada. Bayesian spectral deconvolution with the exchange Monte Carlo method. *Neural networks : the official journal of the International Neural Network Society*, Vol. 28, pp. 82–9, April 2012.

[29] Zhenyu Li, Guoying Zhang, Robert Feil, Jiahuai Han, and Xiaoping Du. Sequential activation of p38 and ERK pathways by cGMP-dependent protein kinase leading to activation of the platelet integrin alphaIIb beta3. *Blood*, Vol. 107, No. 3, pp. 965–72, February 2006.

[30] Maya Mikami and Jay Yang. Short hairpin RNA-mediated selective knockdown of NaV1.8 tetrodotoxin-resistant voltage-gated sodium channel in dorsal root ganglion neurons. *Anesthesiology*, Vol. 103, No. 4, pp. 828–36, October 2005.

[31] Andy Hudmon, Jin-Sung Choi, Lynda Tyrrell, Joel A Black, Anthony M Rush, Stephen G Waxman, and Sulayman D Dib-Hajj. Phosphorylation of sodium channel Na(v)1.8 by p38 mitogen-activated protein kinase increases current density in dorsal root ganglion neurons. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, Vol. 28, No. 12, pp. 3190–201, March 2008.

[32] Stephen J Keely and Kim E Barrett. p38 mitogen-activated protein kinase inhibits calcium-dependent chloride secretion in T84 colonic epithelial cells. *American journal of physiology. Cell physiology*, Vol. 284, No. 2, pp. C339–48, February 2003.

[33] J Lisman. A mechanism for the Hebb and the anti-Hebb processes underlying learning and memory. *Proceedings of the National Academy of Sciences of the United States of America*, Vol. 86, No. 23, pp. 9574–8, December 1989.

[34] Pablo D'Alcantara, Serge N Schiffmann, and Stéphane Swillens. Bidirectional synaptic plasticity as a consequence of interdependent Ca2+-controlled phosphorylation and dephosphorylation pathways. *The European journal of neuroscience*, Vol. 17, No. 12, pp. 2521–8, June 2003.

[35] O P Hamill, A Marty, E Neher, B Sakmann, and F J Sigworth. Improved patch-clamp techniques for high-resolution current recording from cells and cell-free membrane patches. *Pflügers Archiv : European journal of physiology*, Vol. 391, No. 2, pp. 85–100, August 1981.

[36] Zhexing Wen, Carmine Guirland, Guo-Li Ming, and James Q Zheng. A CaMKII/calcineurin switch controls the direction of Ca(2+)-dependent growth cone guidance. *Neuron*, Vol. 43, No. 6, pp. 835–46, September 2004.

[37] John Henley and Mu-ming Poo. Guiding neuronal growth cones using Ca2+ signals. *Trends in cell biology*, Vol. 14, No. 6, pp. 320–30, June 2004.

[38] Takuro Tojima, Jacob H Hines, John R Henley, and Hiroyuki Kamiguchi. Second messengers and membrane trafficking direct and organize growth cone steering. *Nature reviews. Neuroscience*, Vol. 12, No. 4, pp. 191–203, April 2011.

[39] Daniel J Sutherland, Zac Pujic, and Geoffrey J Goodhill. Calcium signaling in axon guidance. *Trends in neurosciences*, Vol. 37, No. 8, pp. 424–32, August 2014.

[40] Elizabeth M Forbes, Andrew W Thompson, Jiajia Yuan, and Geoffrey J Goodhill. Calcium and cAMP levels interact to determine attraction versus repulsion in axon guidance. *Neuron*, Vol. 74, No. 3, pp. 490–503, May 2012.

[41] Junichiro Yoshimoto and Kenji Doya. Bayesian system identification method for biochemical reaction systems (in japanese). *Information Processing Society of Japan Research Report. BIO, Bioinformatics*, Vol. 2006, No. 64, pp. 35–40, jun 2006.

Appendix

A. Biochemical reaction and its regulation

Basic of biochemical reaction is the state transition between active and inactive state (Fig. A.1(a)). This dynamics can be written as

$$\frac{d[\text{A}^*]}{dt} = k_f[\text{A}] - k_b[\text{A}^*],\tag{A.1}$$

which is known as the law of mass-action, where the parentheses represent the concentration of the corresponding molecule. Constants, k_f and k_b, represent the rate of chemical reaction, which are called reaction rate of forward and backward reaction, respectively. Thus, the equation expresses auto-activation (1st term of right-hand side) and self-inactivation (2nd term of right-hand side).

Modulation of the chemical reaction is often modeled by the linear effect of the concentration of regulator molecules. If forward reaction is positively regulated (Fig. A.1(b)), the dynamics become

$$\frac{d[\text{A}^*]}{dt} = (k_{f0} + k_{fB}[\text{B}])[\text{A}] - k_b[\text{A}^*],\tag{A.2}$$

where k_{f0} and k_{fB} are constant of forward reaction rate. In the similar way, if both forward and backward reaction are positively regulated (Fig. A.1(c)), the dynamics can be described as

$$\frac{d[\text{A}^*]}{dt} = (k_{f0} + k_{fC}[\text{C}])[\text{A}] - (k_{b0} + k_{bDxo}[\text{D}])[\text{A}^*],\tag{A.3}$$

where k_{f0} and k_{fB} are constants of forward reaction rate and k_{b0} and k_{bD} are constants of backward reaction rate.

(a) No modulation   (b) Partial modulation   (c) Both-side modulation
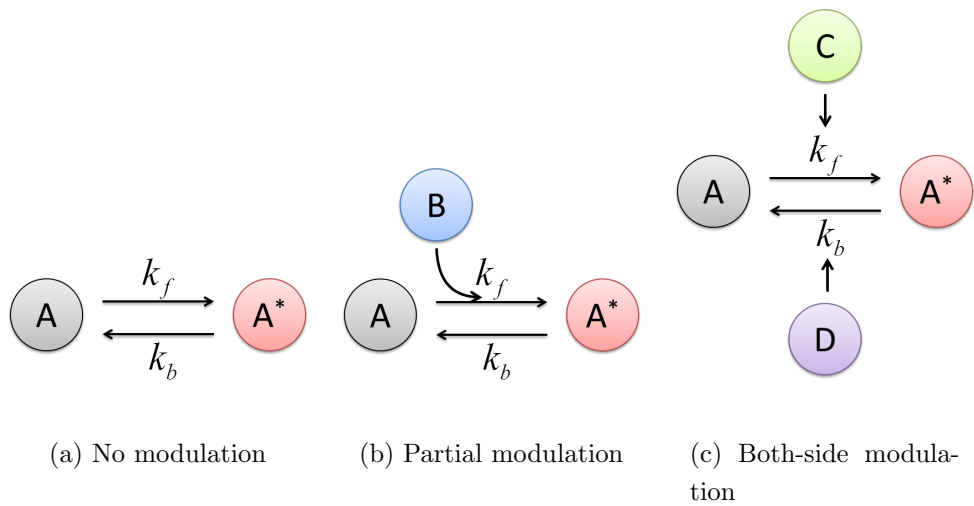
Figure A.1. Biochemical reaction and its regulation. A, B, C and D represent molecules and asterisk (*) represents active state of molecule. (a) Molecule A is activated/inactivated by itself or by physical state such as temperature. (b) Molecule A is activated by molecule B. (c) Molecule A is activated by molecule C and inactivated by molecule D, simultaneously.

B. Parallel conductance model

Membrane potential of a cell is modeled by electrical circuit which is known as parallel conductance model. Each ion channel on the cell membrane can be modeled as a battery with internal conductance (Fig. A.2(a)). Therefore, cell itself forms electrical circuit (Fig. A.2(b)) and it can be described by

$$C_m \frac{dV}{dt} = -g_{\text{Na}}(V - V_{\text{Na}}) - g_{\text{K}}(V - V_{\text{K}}) - g_{\text{Cl}}(V - V_{\text{Cl}}), \qquad \text{(A.4)}$$

which is same as Hodgkin-Huxley type equation (Eq. (13)).
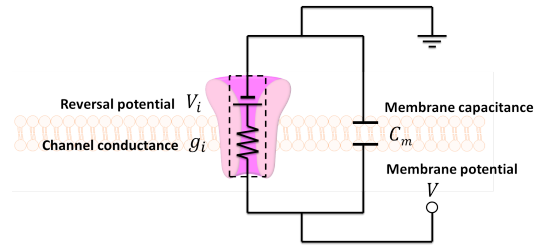
(a) Conductance model of ion channel
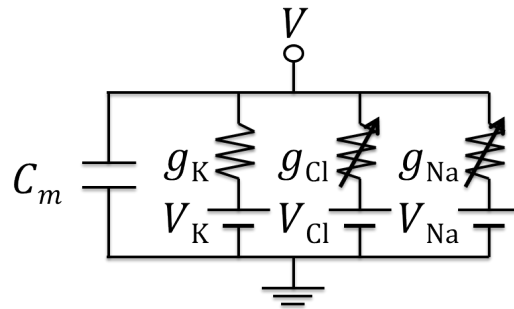
(b) Parallel conductance model

Figure A.2. Parallel conductance model. (a) Ion channel acts as battery on the cell membrane (V_i with internal conductance g_i). Electrical circuit is formed with cell membrane capacity (C_m), which drives electrical potential called membrane potential (V). (b) Cell contains several ion channels such as potassium, chloride and sodium channels. These make parallel circuit and modify its cell membrane potential.

C. MCMC algorithm

To obtain Markov Chain Monte Carlo (MCMC) sample, we utilized the modified Metropolis-Heisting algorithm [41]:

Algorithm 1 Modified Metropolis-Heisting algorithm

Set appropriate number of sample candidates L and iteration H

Initialize sample candidates set $\{\bar{\Theta}^{(l)}| = 1, \ldots, L\}$ appropriately

for $h = 1$ to H **do**

  for $l = 1$ to L **do**

    $\circ$ Generate new sample candidate

    Randomly choose l_1 and l_2 from the set $\{1, \ldots, L\}$

    Randomly choose η from standard uniform distribution $U(0,1)$

    $\tilde{\Theta}^{(l)} \leftarrow \bar{\Theta}^{(l)} + \eta \left(\bar{\Theta}^{(l_1)} - \bar{\Theta}^{(l_2)} \right)$

    $\circ$ Decide to accept or reject of the sample candidate

    Randomly choose r from $U(0,1)$

    if $r < \min \left\{ 1, \frac{p(\tilde{\Theta}^{(l)}|D)}{p(\bar{\Theta}^{(l)}|D)} \right\}$, **then**

      $\circ$ Accept the $\tilde{\Theta}^{(l)}$

      $\tilde{\Theta}^{(l)} \leftarrow \bar{\Theta}^{(l)}$

    else

      $\circ$ Reject the $\tilde{\Theta}^{(l)}$

    end if

    $\Theta^{(hL+l)} \leftarrow \bar{\Theta}^{(l)}$

  end for

end for

D. Calculation of log-evidence by Ogata's method

Practically, simple MCMC algorithm such as Metropolis-Heisting tends to accept proposals from regions of high probability density rather than from the entire space where the posterior density resides. Such an inadequacy in sampling process may make marginalization practically inaccurate. Ogata's method [27, 28] attempts to solve this practical issue by defining a modified marginal likelihood function weighted by an inverse temperature β,

$$z(\beta) = \int p(x|\Theta)^\beta p(\Theta)d\Theta, \tag{A.5}$$

where $z(1)$ is equivalent to the marginal likelihood (evidence) and $z(0) = 1$. From these observation, it is apparent that the marginal likelihood, $p(x) = z(1)$, is simply given by the product of quotients, each given by two adjacent modified marginal likelihood values as,

$$p(x) = \frac{z(\beta_L)}{z(\beta_{L-1})} \times \cdots \times \frac{z(\beta_2)}{z(\beta_1)} = \prod_{l=1}^{L-1} \frac{z(\beta_{l+1})}{z(\beta_l)}, \tag{A.6}$$

where $0 = \beta_1 < \beta_2 < \cdots < \beta_L = 1$. Using the definition of Eq. (A.5), the fraction term can be transformed as,

$$\frac{z(\beta_{l+1})}{z(\beta_l)} = \int p(x|\Theta)^{\beta_{l+1}-\beta_l} \frac{p(x|\Theta)^{\beta_l}p(\Theta)}{z(\beta_l)}d\Theta = \int p(x|\Theta)^{\beta_{l+1}-\beta_l}p(\Theta|x;\beta_l)d\Theta. \tag{A.7}$$

The most right-hand side can be approximated as the mean of $p(x|\Theta^{(b,l)})^{\beta_{l+1}-\beta_l}$, with the b-th MCMC sample for modified posterior with the l-th inverse temperature β_l, $\Theta^{(b,l)}$. Therefore, Eq. (A.6) can be computed as,

$$p(x) = \prod_{l=1}^{L-1} \frac{1}{B} \sum_{b=1}^{B} p(x|\Theta^{(b,l)})^{\beta_{l+1}-\beta_l}, \tag{A.8}$$

with MCMC samples from the modified posterior. Logarithm of Eq. (A.8) is the same as Eq. (20) in the main text.

E. Leave one out cross validation (LOO-CV) algorithm

Leave one out cross validation is performed as the following algorithm with $B \geq 10^5$. Briefly, in this process, generalization fitness (negative of generalization error) can be estimated by the test phase which is performed as calculating the fitness with the estimated cell-independent parameter in the training phase:

Algorithm 2 LOO-CV algorithm

---

for $i = 1$ to #data **do**

Let LOO(i) a dataset omitting the i-th data

Training phase
Obtain MCMC samples $\{\phi^{(b)}, \theta^{(b)}\}$, $b = 1$:B using LOO(i)

Test phase
Obtain MCMC samples $\{\phi_i^{(b)}\}$, $b = 1$:B using mean a posteriori (MAP) parameter θ^{MAP} and i-th data V_i

Calculate Fitness (Log-likelihood) for V_i

$$e(i) = \frac{1}{T_i} \frac{1}{B} \sum_{b=1}^{B} \log p(V_i | \phi_i^{(b)}, \theta^{\mathrm{MAP}}, c_i)$$

end for

Finally, total leave one out fitness, $\sum_{i=1}^{\#data} e(i)$, is calculated.

F. 8-Br-cGMP permeation model

We modeled the 8-Br-cGMP permeation into the cytoplasm in a growth cone to convert the extracellular 8-Br-cGMP concentration (bath-applied) to its intracellular concentration. We utilized a common enzymatic reaction, in which the membrane permeated 8-Br-cGMPs are in two forms in the cytoplasm upon permeation through the plasma membrane: bounded to the target proteins and in free forms. The model considers the permeation of 8-Br-cGMP into the cytoplasm with a fixed rate (Fig. A.A.3 (a)). The model also considers that the intracellular 8-Br-cGMP concentration decreases proportionally at a rate at which the target bounded 8-Br-cGMP increases. Thus, the intracellular 8-Br-cGMP concentration (S_{in}) is expressed by the following ODE,

$$\frac{dS_{in}}{dt} = \frac{aS_{bath}}{K + S_{bath}} - bS_{in}, \tag{A.9}$$

where S_{bath} is the extracellularly applied concentration that corresponds to substrate in the Michaelis-Menten equation. By taking a steady state condition ($dS_{in}/dt = 0$), the following relationship is obtained,

$$S_{in} = \frac{AS_{bath}}{K + S_{bath}}, \tag{A.10}$$

We set $A = a/b = 7.8$ and $K = 8.3$, which S_{in} saturates at $S_{bath} = 100 \ \mu$M (Fig. A.A.3 (b)). These parameters were estimated to minimize the error between the model and experimental data (Fig. 36(a)).
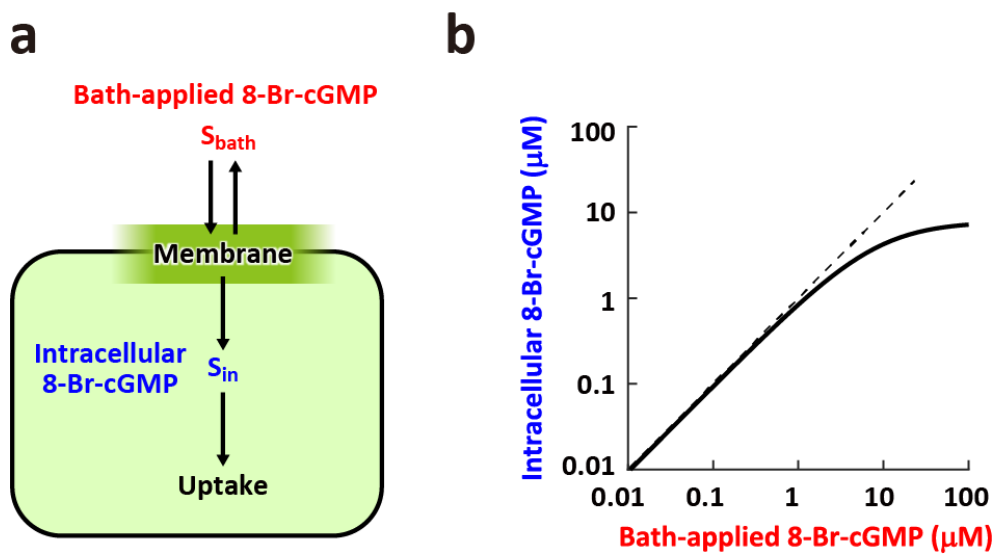
Figure A.3. (a) Bath-applied extracellular 8-Br-cGMPs permeates through the plasma membrane into the cytoplasm with a fixed diffusion rate. Simultaneously, the level of the intracellular free 8-Br-cGMPs decreases as the uptake by bio-molecular reactions increases. (b) Computed intracellular 8-Br-cGMP concentration using the permeation model in (a).

G. Publication list

G.1 Journal paper

- <u>TatsuyaYamada</u>, Makoto Nishiyama, Shigeyuki Oba, Henri Claver Jimbo, Kazushi Ikeda, Shin Ishii, Kyonsoo Hong, Yuichi Sakumura, "Computational Methods for Estimating Molecular System from Membrane Potential Recordings in Nerve Growth Cone" (under submission).

G.2 International conference

- <u>TatsuyaYamada</u>, Henri Claver Jimbo, Shin Ishii, Makoto Nishiyama, Kyonsoo Hong and Yuichi Sakumura, "Identification of a molecular system that regulates growth cone membrane potential during growth cone guidance", OCNS 2011, July 2011.