ELSEVIER

Computers in Biology and Medicine 38 (2008) 574–582

Computers in Biology and Medicine

www.intl.elsevierhealth.com/journals/cobm

Mathematical modeling and application of genetic algorithm to parameter estimation in signal transduction: Trafficking and promiscuous coupling of G-protein coupled receptors

Charin Modchang[a,b], Wannapong Triampo[a,b,*], Yongwimon Lenbury[c]

[a]*R&D Group of Biological and Environmental Physics, Department of Physics, Mahidol University, Bangkok 10400, Thailand*
[b]*Center of Excellence for Vector and Vector-Borne Diseases, Faculty of Science, Mahidol University, Bangkok 10400, Thailand*
[c]*Department of Mathematics, Mahidol University, Bangkok 10400, Thailand*

Received 2 July 2007; accepted 9 February 2008

Abstract

G-protein-coupled receptors (GPCRs) constitute a large and diverse family of proteins whose primary function is to transduce extracellular stimuli into intracellular signals. These receptors play a critical role in signal transduction, and are among the most important pharmacological drug targets. Upon binding of extracellular ligands, these receptor molecules couple to one or several subtypes of G-protein which reside at the intracellular side of the plasma membrane to trigger intracellular signaling events. The question of how GPCRs select and activate a single or multiple G-protein subtype(s) has been the topic of intense investigations. Evidence is also accumulating; however, that certain GPCRs can be internalized via lipid rafts and caveolae. In many cases, the mechanisms responsible for this still remain to be elucidated. In this work, we extend the mathematical model proposed by Chen et al. [Modelling of signalling via G-protein coupled receptors: pathway-dependent agonist potency and efficacy, Bull. Math. Biol. 65 (5) (2003) 933–958] to take into account internalization, recycling, degradation and synthesis of the receptors. In constructing the model, we assume that the receptors can exist in multiple conformational states allowing for a multiple effecter pathways. As data on kinetic reaction rates in the signalling processes measured in reliable *in vivo* and *in vitro* experiments is currently limited to a small number of known values. In this paper, we also apply a genetic algorithm (GA) to estimate the parameter values in our model.

Keywords: G-protein coupled receptors; Promiscuous coupling; Trafficking; Genetic algorithm

1. Introduction

Signal transduction is the process of conversion of external signals, such as hormones, growth factors, neurotransmitters and cytokines, to a specific internal cellular response, such as gene expression, cell division, or even cell suicide. This process begins at the cell membrane where an external stimulus initiates a cascade of enzymatic reactions inside the cell that typically includes phosphorylation of proteins as mediators of downstream processes [1]. Signal transduction consists of three processes. The first is reception—an agonist binds to a specific receptor on the cell membrane that triggers a change in the receptor molecule. The second is transduction. The change in the receptor brings about ordered sequences of biochemical reactions inside the cell that are carried out by enzymes and linked through second messengers. The third is response. After receiving the signal, target protein produces response which can be any of many different cellular activities, such as activation of a certain enzyme, rearrangement of the cytoskeleton, or changes in gene expression.

G-protein-coupled receptors (GPCRs) constitute a large and diverse family of proteins whose primary function is to transduce extracellular stimuli into intracellular signals. GPCRs are among the most heavily investigated drug targets in the pharmaceutical industry. They account for the majority of best-selling

* Corresponding author at: R&D Group of Biological and Environmental Physics, Department of Physics, Mahidol University, Bangkok 10400, Thailand. Tel.: +66 24419816.
E-mail addresses: cmodchang@yahoo.com (C. Modchang), scwtr@mahidol.ac.th, wtriampo@gmail.ac.th (W. Triampo), scylb@mahidol.ac.th (Y. Lenbury).

0010-4825/$ - see front matter © 2008 Elsevier Ltd. All rights reserved.
doi:10.1016/j.compbiomed.2008.02.005

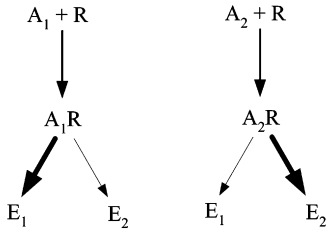

Fig. 1. Illustration of agonist-directed trafficking hypothesis. When an agonist A_1 binds to the receptor R, it activates effector pathway 1 (E_1) more effectively. While, when A_2 binds to the receptor R, E_2 is favored over E_1. The preferences in pathway activations are indicated by the line thickness.

drugs and about 40% of all prescription pharmaceuticals on the market [2]. The inactive form of G-protein is a heterotrimer composed of three subunits, α, β and γ with a molecule of guanosine diphosphate (GDP) bound to the α subunit. The binding of ligands to receptors causes them to interact with the G-protein. The interaction of this inactive G-protein with bound receptor promotes the release of GDP from the α subunit and the binding of nucleotide guanosine triphosphate (GTP) at the same site. The G-protein is then released from the receptor and dissociates into separate $\beta\gamma$ and α-GTP subunits. The α-GTP is the active form of the G-protein. The activated $\beta\gamma$ and α-GTP subunits in turn stimulate the generation of second messengers via intracellular effectors, passing on the signal by altering the activities of selected cellular proteins [3]. Depending on the type of G-protein to which the receptor is coupled, a variety of downstream signalling pathways can be activated [4,5]. There are many examples of a single receptor coupling directly to more than one cellular signal transduction pathway [6,7]. In traditional receptor theory, it is predicted that the relative degree of activation of each effector pathway by an agonist (relative efficacy) must be the same. More recently, however, evidence from a variety of systems suggests that some agonists, acting at a single receptor, may preferentially activate E_1, say, while other agonists, acting at the same receptor in the same system, may preferentially activate E_2 as shown schematically in Fig. 1 [6,7]. These phenomenon termed "agonist-directed trafficking of receptor stimulus" were originally proposed in [8].

Many mathematical models have been proposed to describe agonist-directed trafficking of receptor stimulus and promiscuous coupling of receptors in the signal transduction process. In actual fact, the idea that a receptor can adopt more than one active state was derived from the concept of agonist-directed trafficking of a receptor stimulation to explain the ability of structurally diverse agonists to activate different G-protein-mediated signaling [6–8]. According to this model, each agonist is able to promote its own specific active receptor state, leading to an unlimited number of receptor conformations. Thus, different active states of the receptor may be associated with a particular G-protein. In contrast, Leff et al. suggested a three-state model where the receptor might exist in three states, an inactive (R) and two active formations (R^*, R^{**}), accounting for multiple G-protein coupling but limiting the number of active conformations [9]. They studied the agonist activity in two systems, one being an intact system in which receptors and G-protein are uniformly distributed and the other comprising of isolated pathways each with a distinct G-protein, operating independently from the other. Assuming that the efficacy of each pathway is proportional to the number of receptors activated, this model allows for pathway-dependent agonist efficacy and successfully simulates the differential activation of effector pathways observed previously with 5-HT_{2c} agonists [6]. Leff et al. [9] have; however, neglected the role of G-protein activation and, although different agonist potency has been predicted for the system of isolated pathways, their intact system has failed to predict the pathway-dependent agonist potency which has been observed in various experiments. Chen et al. [10] proposed a mathematical model which demonstrates the role of G-proteins in determining pathway-dependent agonist potency. In their model, the receptor can exist in four conformational states, one inactive and three active states. Chen et al. [10] have; however, neglected many biological factors concerning cell signalling via GPCRs, such as synthesis, degradation, internalization and recycling of receptors. Under normal physiological conditions, however, these dynamic trafficking events take place concurrently with receptor–ligand binding [11–14].

While being a natural activity of receptors linked to signaling, internalization may be a therapeutically useful activity in itself. Ligands that selectively induce receptor internalization may have utility in the prevention of HIV-1 infection. This is because internalization may remove critical co-receptors for membrane fusion and subsequent HIV-1 infection [15–17]. In fact, this approach may be superior to blocking the HIV-1 infection. In many cases, the mechanisms responsible for these dynamic trafficking of receptors still remain to be elucidated. The focus of our present work is to study the effect of receptors trafficking, including receptors internalization, receptors synthesis, recycling of receptors and receptors degradation by extending a mathematical model proposed in [10].

As data on kinetic reaction rates in the signalling processes measured in reliable *in vivo* and *in vitro* experiments is currently limited to a small number of known values. In this paper, we also apply a genetic algorithm (GA) to estimate the parameter values in our model. The GA is an effective stochastic global search algorithm that is inspired by the evolutionary features of biological systems [18]. The GAs are the most popular evolutionary techniques, in virtue of their conceptual simplicity, the ease of programming entailed, and small number of parameters to be defined. Moreover, they have been shown to outperform alternative search techniques on difficult problems involving high dimensional, discontinuous, noisy and multi-modal objective functions [19,20]. It has been successfully applied to various problems, such as function optimizations, parameter estimation in biochemical pathways [21–23], cancer gene search [24] and parameter estimation in mathematical modeling [20]. In the present work, a GA was applied to estimate 18 parameter values in our model and the predictions of the model were compared with the experimental results obtained by the authors of [7].

There are some quantitative pharmacological terms that are useful for our present analysis. Efficacy is defined as the ability

of a drug to produce a stimulus, indicated by the maximum effect that can be produced by that drug. Potency, commonly expressed as the *EC50*, refers to the concentration or amount of an agonist needed to produce a 50% of the maximum effect of that agonist. A full agonist is a ligand that binds to a receptor and leads to a maximum biological response in the system under study while a partial agonist is an agonist that does not elicit as large an effect as a full agonist. An antagonist is a ligand that binds to a receptor, but does not produce a biological response, while blocking the actions of agonists. An inverse agonist means a ligand that binds to a receptor and reduces the constitutive activity of the receptor, thereby producing an effect opposite to that of an agonist.

2. The mathematical model

2.1. Model construction

We extend the model proposed in [10], in which receptors are allowed to exist in multiple conformational states, to include the receptors synthesis, degradation, internalization and recycling. In the absence of agonists, the receptors can exist in four different conformational states, one inactive or resting R and three active R^{j*}states, where the superscript j (and subscript j below) henceforth takes the value of 1, 2 and 3 unless stated otherwise. Each receptor can bind to a different G-protein subtype G_j as shown in Fig. 2(a). The inactive receptors, R, are converted into an active state R^{j*} with rate constants L_j^+. The parameters K^+ and K^- represent ligand association and dissociation rates with receptor R, respectively. The effect of the agonist-directed trafficking of receptor stimulus proposed in [8] is indicated by the values of the parameters μ_j, corresponding to the effect on the various G_j-linked pathways. The numbers of ligand-bound receptors are denoted by R_A while these of the activated ligand-bound receptor are denoted by R_A^{j*}. The ligand dissociation rates from the ligand-bound activated receptors, R_A^{j*}, are assumed to be K^-/μ_j and the deactivation rates of R_A^{j*} to R_A are L_j^-/μ_j, both of which are ligand-dependent. With one effector pathway, a full agonist or partial agonist which preferentially binds to an active receptor has $\mu_j > 1$, while an antagonist which binds equally well to both active and inactive receptor has $\mu = 1$ and an inverse agonist which is more likely to bind to an inactive receptor has $\mu < 1$. This, however, may not always be the case for a system with multiple effector pathways. For example, an agonist with $\mu_j > 1$ for all j and, say, $\mu_1 \gg \mu_2, \mu_3$, can increase the number of active receptors in G_1-linked pathway but reduce the number of active receptors in other conformations and hence behaves as a partial agonist for one pathway and an inverse agonist for others [9,10].

In the model suggested by Chen et al. [10], they assumed that the number and location of receptors on the cell surface are constant, i.e., that no significant synthesis, degradation, internalization, or recycling of receptors occur over the time frame for which the model applies. Under normal physiological conditions; however, these dynamic trafficking events take place concurrently with receptor–ligand binding [11–14,25,26]. The

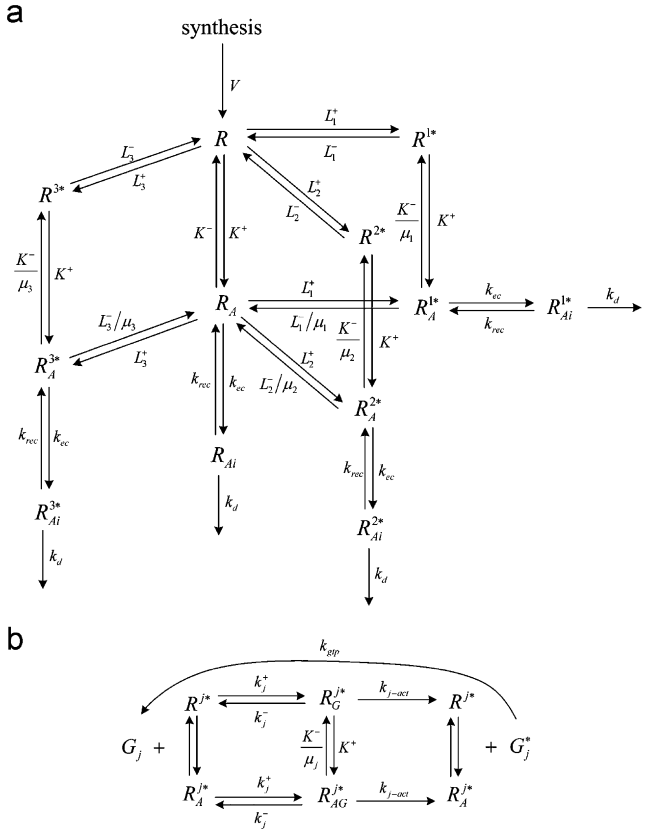

Fig. 2. Extended model structure of (a) receptor–ligand binding with multiple receptor conformation. The free receptors are synthesized with rate V while the ligand–receptor complex internalizes with rate k_{ec} and the internalized ligand–receptor complex degrades with rate k_d. (b) G-protein activation of the G_j -linked pathways, assuming R^{j*} and R_A^{j*} associate with or dissociate from G-protein at the same rate.

important feature of our model is the receptor synthesis, degradation and trafficking. The rate of new receptor synthesis and expression on the cell surface as free receptors is V. The rate constant describing the internalization of receptor–ligand complex is k_{ec}. The parameter k_{rec} represents the rate constant for transport of material *via* vesicles from the endosome back to the cell surface [27] and k_d represents a rate constant for the routing of receptor–ligand complex from the endosome to the lysosome, and degradation in the lysosome. Fig. 2(b) shows the activation of G-protein by the activated receptors. Both ligand-bound and nonligand-bound activated receptors, denoted by R_A^{j*} and R^{j*}, respectively, associate with and dissociate from G-protein with rate constants k_j^+ and k_j^- for G_j. The inactive form of the G-protein consists of α, β and γ subunits with a molecule of GDP bound to the α subunit (G_α). The interaction of this inactive G-protein with an activated receptor promotes the release of GDP from the α subunit and the binding of GTP at the same site. We have assumed that the dissociation of GDP and association of GTP happen instantaneously and the activated receptor activates the G-protein with a rate constant $k_{j-\mathrm{act}}$. Active G-proteins are returned to their inactive state upon hydrolysis of GTP by the GTPase activity found in the α subunit itself, and the α-GDP and $\beta\gamma$ subunits ($G_{\beta\gamma}$) can

then recombine. In our model, we assume that the inactivation of G-protein is so fast for our time scale, and thus we consider it to be a one-step process with rate constant k_{gtp}. Note that only the α-GTP subunits (G_α) are considered as the activated G-protein in our model.

Based on the model structure in Fig. 2, the law of mass action leads to the following set of coupled ordinary differential equations:

$$\frac{\mathrm{d}R}{\mathrm{d}t} = K^- R_A - \sum_{n=1}^{3} L_n^+ R - K^+ AR + \sum_{n=1}^{3} L_n^- R^{n*} + V, \quad (2.1)$$

$$\frac{\mathrm{d}R^{j*}}{\mathrm{d}t} = L_j^+ R - (L_j^- + K^+ A)R^{j*} + \frac{K^-}{\mu_j} R_A^{j*} - k_j^+ G_j R^{j*} + (k_j^- + k_{j-\text{act}})R_G^{j*}, \quad (2.2)$$

$$\frac{\mathrm{d}R_A^{j*}}{\mathrm{d}t} = K^+ AR^{j*} - \left(\frac{L_j^-}{\mu_j} + \frac{K^-}{\mu_j}\right) R_A^{j*} + L_j^+ R_A - k_j^+ G_j R_A^{j*} + (k_j^- + k_{j-\text{act}})R_{AG}^{j*} - k_{\text{ec}} R_A^{j*} + k_{\text{rec}} R_{Ai}^{j*}, \quad (2.3)$$

$$\frac{\mathrm{d}R_{AG}^{j*}}{\mathrm{d}t} = k_j^+ G_j R_A^{j*} - (k_j^- + k_{j-\text{act}})R_{AG}^{j*} + K^+ AR_G^{j*} - \frac{K^-}{\mu_j} R_{AG}^{j*}, \quad (2.4)$$

$$\frac{\mathrm{d}R_G^{j*}}{\mathrm{d}t} = k_j^+ G_j R^{j*} - (k_j^- + k_{j-\text{act}})R_G^{j*} - K^+ AR_G^{j*} + \frac{K^-}{\mu_j} R_{AG}^{j*}, \quad (2.5)$$

$$\frac{\mathrm{d}G_j^*}{\mathrm{d}t} = k_{j-\text{act}}(R_G^{j*} + R_{AG}^{j*}) - k_{\text{gtp}} G_{\beta\gamma} G_j^*, \quad (2.6)$$

$$\frac{\mathrm{d}R_{Ai}}{\mathrm{d}t} = k_{\text{ec}} R_A - k_d R_{Ai} - k_{\text{rec}} R_{Ai}, \quad (2.7)$$

$$\frac{\mathrm{d}R_{Ai}^{j*}}{\mathrm{d}t} = k_{\text{ec}} R_A^{j*} - k_d R_{Ai}^{j*} - k_{\text{rec}} R_{Ai}^{j*}, \quad (2.8)$$

where $G_{\beta\gamma} = G_1^* + G_2^* + G_3^*$ and

$$R_A = R_0 - R - R_{Ai} - \sum_{n=1}^{3} (R^{n*} + R_A^{n*} + R_{AG}^{n*} + R_G^{n*} + R_{Ai}^{n*}), \quad (2.9)$$

$$G_j = g_j - G_j^* - R_{AG}^{j*} - R_G^{j*}. \quad (2.10)$$

In order to obtain the model equations, we assume that the total number of a G-protein subtype on the cell surface and the concentration of ligands, denoted by A, remain constant. Moreover, we also assume that the number of newly synthesized receptors is approximately equal to the number of degraded receptors so that the total number of receptors is conserved. The parameters g_j's and R_0 are the total number of each G-protein subtype and that of the receptors, respectively. The notation R_G^{j*} and R_{AG}^{j*} denote the number of G_j-precoupled active receptors and of ligand–receptor–G_j complexes. The initial conditions for the system of equations are as follows:

$$R = R_0, \quad G_j = g_j, \quad \text{at } t = 0, \quad (2.11)$$

with the concentrations of all other species being zero at $t = 0$.

2.2. Nondimensionalization

We now, in the same manner as in [10], proceed to carry out nondimensionalization by the following rescaling:

$$\begin{array}{lll} t = \bar{t}/K^-, & R = R_0\bar{R}, & R_A = R_0\bar{R}_A, \\ G_{\beta\gamma} = G_0\bar{G}_{\beta\gamma}, & G_j^* = G_0\bar{G}_j, & G_j = G_0\bar{G}_j, \\ A = a_0\bar{A}, & R^{j*} = R_0\bar{R}^{j*}, & R_A^{j*} = R_0\bar{R}_A^{j*}, \\ R_{AG}^{j*} = R_0\bar{R}_{AG}^{j*}, & R_G^{j*} = R_0\bar{R}_G^{j*}, & g_j = G_0\bar{g}_j, \\ \overline{K}^+ = \frac{K^+ a_0}{K^-}, & \overline{L}_j^+ = \frac{L_j^+}{K^-}, & \overline{L}_j^- = \frac{L_j^-}{K^-}, \end{array}$$

where $G_0 = g_1 + g_2 + g_3$ and thus $\bar{g}_1 + \bar{g}_2 + \bar{g}_3 = 1$. The constant a_0 is chosen such that $\overline{K}^+ = K^+ a_0/K^- = \mathrm{O}(1)$. Since the binding of G-protein to an activated receptor often leads to the activation of the G-protein, we assume that the G-protein activation rate constant $k_{j-\text{act}}$ is very large and G-protein dissociation rate constant k_j^- is very small. Now, let us make some approximations by letting

$$\varepsilon = \frac{K^-}{k_{1-\text{act}}}, \quad \bar{k}_{j-\text{act}} = \frac{k_{j-\text{act}}}{k_{1-\text{act}}}, \quad \bar{k}_j^- = \frac{k_j^- k_{1-\text{act}}}{(K^-)^2}, \quad \bar{k}_{\text{ec}} = \frac{k_{\text{ec}}}{K^-},$$
$$N = \frac{R_0}{G_0}, \quad \bar{k}_{\text{gtp}} = \frac{k_{\text{gtp}} G_0}{K^-}, \quad \bar{k}_j^+ = \frac{k_j^+ G_0}{K^-},$$
$$\overline{V} = \frac{V}{R_0 K^-}, \quad \bar{k}_{\text{rec}} = \frac{k_{\text{rec}}}{K^-}, \quad \bar{k}_d = \frac{k_d}{K^-},$$

where $\varepsilon \ll 1$. The overbars will be dropped henceforth for brevity. The nondimensionalized system of equations is then

$$\frac{\mathrm{d}R}{\mathrm{d}t} = R_A - \sum_{n=1}^{3} L_n^+ R - K^+ AR + \sum_{n=1}^{3} L_n^- R^{n*} + V, \quad (2.12)$$

$$\frac{\mathrm{d}R^{j*}}{\mathrm{d}t} = L_j^+ R - (L_j^- + K^+ A)R^{j*} + \frac{R_A^{j*}}{\mu_j} - k_j^+ G_j R^{j*} + \left(\varepsilon k_j^- + \frac{k_{j-\text{act}}}{\varepsilon}\right) R_G^{j*}, \quad (2.13)$$

$$\frac{\mathrm{d}R_A^{j*}}{\mathrm{d}t} = K^+ AR^{j*} - \left(\frac{L_j^-}{\mu_j} + \frac{1}{\mu_j}\right) R_A^{j*} + L_j^+ R_A - k_j^+ G_j R_A^{j*} + \left(\varepsilon k_j^- + \frac{k_{j-\text{act}}}{\varepsilon}\right) R_{AG}^{j*} - k_{\text{ec}} R_A^{j*} + k_{\text{rec}} R_{Ai}^{j*}, \quad (2.14)$$

$$\frac{\mathrm{d}R_{AG}^{j*}}{\mathrm{d}t} = k_j^+ G_j R_A^{j*} - \left(\varepsilon k_j^- + \frac{k_{j-\text{act}}}{\varepsilon}\right) R_{AG}^{j*} + K^+ AR_G^{j*} - \frac{R_{AG}^{j*}}{\mu_j}, \quad (2.15)$$

$$\frac{dR_G^{j*}}{dt} = k_j^+ G_j R^{j*} - \left(\varepsilon k_j^- + \frac{k_{j-act}}{\varepsilon}\right) R_G^{j*} - K^+ A R_G^{j*} + \frac{R_{AG}^{j*}}{\mu_j}, \quad (2.16)$$

$$\frac{dG_j^*}{dt} = N\frac{k_{j-act}}{\varepsilon}(R_G^{j*} + R_{AG}^{j*}) - k_{gtp} G_{\beta\gamma} G_j^*, \quad (2.17)$$

$$\frac{dR_{Ai}}{dt} = k_{ec} R_A - k_d R_{Ai} - k_{rec} R_{Ai}, \quad (2.18)$$

$$\frac{dR_{Ai}^{j*}}{dt} = k_{ec} R_A^{j*} - k_d R_{Ai}^{j*} - k_{rec} R_{Ai}^{j*}, \quad (2.19)$$

and

$$R_A = 1 - R - R_{Ai} - \sum_{n=1}^{3} \left(R^{n*} + R_A^{n*} + R_{AG}^{n*} + R_G^{n*} + R_{Ai}^{n*}\right), \quad (2.20)$$

$$G_j = g_j - G_j^* - N R_{AG}^{j*} - N R_G^{j*}. \quad (2.21)$$

The initial conditions are $R=1$, $G_j = g_j$ with $g_1 + g_2 + g_3 = 1$, with other species having zero initial concentrations.

Since we have assumed that k_{j-act} is very large, it is reasonable to rescale R_G^{j*} and R_{AG}^{j*} by letting

$$R_G^{j*} = \varepsilon \tilde{R}_G^{j*} \quad \text{and} \quad R_{AG}^{j*} = \varepsilon \tilde{R}_{AG}^{j*}. \quad (2.22)$$

Substituting (2.22) in (2.15), (2.16), (2.17), (2.20) and (2.21) leads to

$$\frac{d\tilde{R}_{AG}^{j*}}{dt} = \frac{k_j^+}{\varepsilon} G_j R_A^{j*} - \left(\varepsilon k_j^- + \frac{k_{j-act}}{\varepsilon}\right) \tilde{R}_{AG}^{j*} + K^+ A \tilde{R}_G^{j*} - \frac{\tilde{R}_{AG}^{j*}}{\mu_j}, \quad (2.23)$$

$$\frac{d\tilde{R}_G^{j*}}{dt} = \frac{k_j^+}{\varepsilon} G_j R^{j*} - \left(\varepsilon k_j^- + \frac{k_{j-act}}{\varepsilon}\right) \tilde{R}_G^{j*} - K^+ A \tilde{R}_G^{j*} + \frac{\tilde{R}_{AG}^{j*}}{\mu_j}, \quad (2.24)$$

$$\frac{dG_j^*}{dt} = N k_{j-act} (\tilde{R}_G^{j*} + \tilde{R}_{AG}^{j*}) - k_{gtp} G_{\beta\gamma} G_j^*, \quad (2.25)$$

$$R_A = 1 - R - R_{Ai} - \sum_{n=1}^{3} \left(R^{n*} + R_A^{n*} + \varepsilon \tilde{R}_{AG}^{n*} + \varepsilon \tilde{R}_G^{n*} + R_{Ai}^{n*}\right), \quad (2.26)$$

$$G_j = g_j - G_j^* - \varepsilon N \tilde{R}_{AG}^{j*} - \varepsilon N \tilde{R}_G^{j*}. \quad (2.27)$$

To simplify the calculation, we will also assume that the concentration of $G_{\beta\gamma}$ is constant instead of varying with G_j^*. In the limit $\varepsilon \to 0$, with quasi-steady state analysis, Eqs. (2.23) and (2.24) yield

$$\tilde{R}_{AG}^{j*} = \frac{k_j^+ G_j}{k_{j-act}} R_A^{j*}, \quad \tilde{R}_G^{j*} = \frac{k_j^+ G_j}{k_{j-act}} R^{j*}. \quad (2.28)$$

Using (2.28) to eliminate R_G^{j*} and R_{AG}^{j*} in (2.13) and (2.14) and substituting (2.28) in (2.25), our system of equations becomes

$$\frac{dR}{dt} = R_A - \sum_{n=1}^{3} L_n^+ R - K^+ A R + \sum_{n=1}^{3} L_n^- R^{n*} + V \quad (2.29)$$

$$\frac{dR^{j*}}{dt} = L_j^+ R - (L_j^- + K^+ A) R^{j*} + \frac{R_A^{j*}}{\mu_j} \quad (2.30)$$

$$\frac{dR_A^{j*}}{dt} = K^+ A R^{j*} + L_j^+ R_A - \left(\frac{L_j^-}{\mu_j} + \frac{1}{\mu_j}\right) R_A^{j*} - k_{ec} R_A^{j*} + k_{rec} R_{Ai}^{j*}, \quad (2.31)$$

$$\frac{dR_{Ai}}{dt} = k_{ec} R_A - k_d R_{Ai} - k_{rec} R_{Ai}, \quad (2.32)$$

$$\frac{dR_{Ai}^{j*}}{dt} = k_{ec} R_A^{j*} - k_d R_{Ai}^{j*}, -k_{rec} R_{Ai}^{j*}, \quad (2.33)$$

where

$$R_A = 1 - R - R_{Ai} - \sum_{n=1}^{3} (R^{n*} + R_A^{n*} + R_{Ai}^{n*}),$$

$$G_j^* = \frac{H_j g_j}{1 + H_j}, \quad G_j = g_j - G_j^*,$$

$$H_j \equiv \frac{N k_j^+}{k_{gtp} G_{\beta\gamma}} (R_A^{j*} + R^{j*}). \quad (2.34)$$

Eqs. (2.29)–(2.34) are quasi-steady state equations where $d\tilde{R}_{AG}^{j*}/dt$, $d\tilde{R}_G^{j*}/dt$ and dG_j^*/dt have already been set to zero. These equations can now be used to describe the steady state of ligand–receptors–G-proteins binding, if all of a parameter values were known.

3. Parameters estimation

Our model equations (2.29)–(2.34) contain 18 parameters, L_j^+, L_j^-, μ_j, k_j^+, K^+, $k_{gtp} G_{\beta\gamma}$, V, k_{ec}, k_{rec} and k_d for $j=1, 2, 3$, but the experimental data available from the literatures are limited to only a few parameters and only some types of receptors [28]. When only a few kinetic parameters are available to implement a model of signal transductions, one might resort to attempting a theoretical estimate of these values. The attempt could be performed, in principle, by using an "inverse problem" approach, i.e., by optimizing the unknown parameters of a reaction's model in order to obtain the best possible agreement between simulated and experimental data [20–23]. In the present work we will use GA to estimate these unknown parameters.

The GA is an effective stochastic global search algorithm that mimics biological evolution [18]. As it is robust, i.e., it uses only objective function information and not other auxiliary information, it has been successfully applied to various problems, such as function optimization and combinatorial optimization, especially when a rigorous mathematical model is too complicated to be practically implemented [29]. In our problem, the input to the GA is a set (called a "population") of vectors

(called "individuals") whose elements (called "genomes") are the values of those 18 parameters. A fitness function is defined to be the distance $f(\mathbf{x})$ measured between experimental and predicted values of the steady state activated G-proteins concentration,

$$f(\mathbf{x}) = \left(\sum_{i=1}^{n} \sum_{j=1}^{m} \left\{ \frac{|y_{\text{pred}}(i) - y_{\text{exp}}(i)|}{|y_{\text{exp}}(i)|} \right\}_j \right)^2, \tag{3.1}$$

where n is the number of data points for each experiment, m is the number of G-protein subtypes, y_{exp} represents the known experimental data and y_{pred} is the simulated data of the steady state activated G-proteins concentration obtained by using GA. The purpose of the GA is to produce successive populations of individuals which are generated with the aim of increasing the fitness of their individuals, i.e., their ability to solve the optimization problem by decreasing the distance $f(\mathbf{x})$ between simulated data and experimental data.

In order to estimate the parameter values, we use the forward time scheme $(\mathrm{d}R(t)/\mathrm{d}t \approx (R(t+\Delta t) - R(t))/\Delta t)$ to transform our model equations (2.29)–(2.34) to a system of finite difference equations

$$\vec{u}^{k+1} = \mathbf{A}\vec{u}^{k} + \vec{b}, \tag{3.2}$$

where $\vec{u}^k$ is a vector of all R and G species at time step k, $\mathbf{A} = [a_{ij}]$ is a coefficient matrix whose elements consist of those 18 parameters and time step Δt, and $\vec{b}$ is a constant vector. This iteration can take an excessive time to long run and get into an unstable regime if the parameters in matrix $\mathbf{A}$ and vector $\vec{b}$ are not appropriately chosen (which can be the case for the unconstrained GAs in which the population is chosen stochastically). To avoid this problem, we make use of a theorem namely; if an infinity norm of $\mathbf{A}$, $\|\mathbf{A}\|$, is less than one then the iteration scheme $\vec{u}^{k+1} = \mathbf{A}\vec{u}^{k} + \vec{b}$ will converge to the steady state equations:

$$\vec{u} = (\mathbf{I} - \mathbf{A})^{-1}\vec{b}, \tag{3.3}$$

Table 1
The lower bound and upper bound of the parameter values used in genetic algorithm parameter estimations

Parameters	Lower bound	Upper bound
L_1^+	0.05	20
L_2^+	0.05	20
L_3^+	0.05	20
L_1^-	5	1500
L_2^-	5	1500
L_3^-	5	1500
μ_1	15	2500
μ_2	5	1500
μ_3	1	500
k_1^+	0.01	100
k_2^+	0.01	100
k_3^+	0.01	100
K^+	0.001	10
$k_{\text{gtp}}G_{\beta\gamma}$	0.01	100
V	0	2500
k_{ec}	0	2500
k_d	0	2500
k_{rec}	0	2500

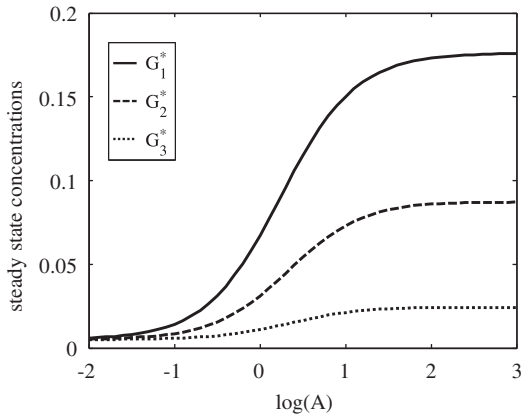

Fig. 3. Simulated concentrations and response curves for the three separate pathways obtained by using genetic algorithm to fit the data in Fig. 3 of Chen *et al.* [10]. For illustrative purposes, the parameter values of RUN1 in Table 2 were used.

Table 2
The parameter values for concentrations and response curves obtained by using genetic algorithm

Parameters	RUN1	RUN2	RUN3
L_1^+	0.0503	9.7134	1.909
L_2^+	0.0511	8.9495	8.160
L_3^+	0.1303	14.278	1.1368
L_1^-	133.25	461.126	198.04
L_2^-	27.742	696.71	1483.5
L_3^-	5.934	1486.8	312.01
μ_1	87.728	2499.9	667.49
μ_2	47.852	1350.9	14.573
μ_3	13.376	361.42	6.5537
k_1^+	45.319	28.16	99.994
k_2^+	9.3357	34.27	99.640
k_3^+	0.7807	42.86	92.338
K^+	0.2974	0.0054	0.3048
$k_{\text{gtp}}G_{\beta\gamma}$	1.731	39.829	28.2727
V	0.00001	0.0001	0.0403
k_{ec}	189.27	1.573	15.927
k_d	6.178	0.00005	1425.9
k_{rec}	2436	2129	785.61
Fitness values	0.7838	5.142	0.1392

Three independent runs were performed to obtain the three sets of parameter values.

where $\mathbf{I}$ is the identity matrix and the infinity norm of matrix $\mathbf{A}$ is defined by [30]

$$\|\mathbf{A}\| \equiv \max_i \sum_j |a_{ij}|.$$

Thus, in generating a parameter estimation, we also impose a nonlinear constraint $\|\mathbf{A}\| < 1$. This constraint makes the problem of parameter estimation a lot easier, e.g., we are not concerned if the population is in the unstable regions and there is no need to use the iteration equation (3.2) to obtain to the steady state. We simply use the steady state equation (3.3). To solve the parameter estimation problem with the constrained GA, we use the Augmented Lagrangian Genetic Algorithm (ALGA) which is provided in Genetic Algorithm and Direct Search Toolbox

in MATLAB® software [31,32]. In addition to the constraint that $\|\mathbf{A}\| < 1$, we also impose lower and upper bounds on each parameter value as shown in Table 1. The GA was performed by using 60 individuals of population and was run up to 100 generations.

4. Results and discussion

For the purpose of testing our model, first it will be used to reproduce a result like that in Fig. 3 of Chen et al. [10]. We use their Eq. (2.33) to generate the data points (50 data points for each curve). After a total 150 data points are obtained, we use our model with the help of the constrained GA to fit that data by minimizing the fitness function (3.1). We run the GA several times, but only the three sets of parameters value obtained are shown in Table 2. By inserting the parameter values in RUN1 into our model equations (2.29)–(2.34), the numerical solutions of the steady state active G-proteins concentrations are shown in Fig. 3. Note, however, that the system under investigation does not guarantee that the inverse problem has one unique solution. We could say only that we have found a good solution but that might not be the best solution. If we look at Table 2, we can see that among all of the three parameter sets, $\mu_1 > \mu_2 > \mu_3$, which is a result that we could expect for the G_1-linked pathway to have a higher efficacy (i.e., inducing higher maximum response) [10]. If the receptor synthesis rate V is large, the receptor degradation rate k_d tends to be large also. Fig. 4 shows a time evolution of the fitness value during the calculation of the three sets of kinetic parameter values in Table 2. The fitness value decreases very fast in the early generations, but in the later generations the fitness value decreases very slowly.

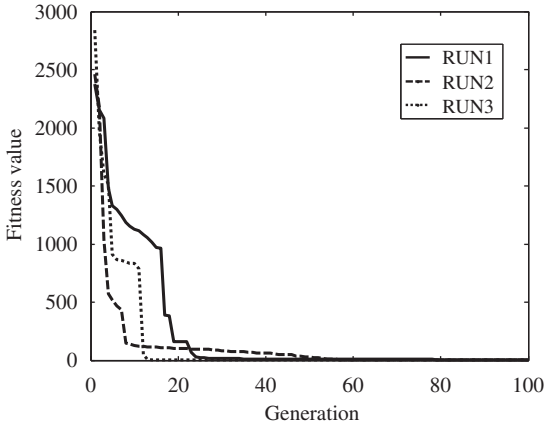

Fig. 4. Time evolution of the fitness value during the calculation of the three sets of kinetic parameters. Notice a fast decrease in the fitness value in the early generations.

To verify the validity of the model, we also qualitatively reproduce the experimental results given by Cordeaux *et al.* [7]. In the experiment of Cordeaux *et al*., they investigated the effect of two agonists, NECA and CPA, on adenosine A_1 receptors which can couple to three different families of G-protein, G_i, G_s and G_q, where each family of G-protein regulates specific classes of effector molecules within the cell. In this study, they found that NECA had a higher efficacy while CPA appeared to have a higher potency (having lower *EC50*). To be consistent with our model, we thus set $G_i = G_1$, $G_q = G_2$ and $G_s = G_3$. We rescale the agonist concentrations by using $a_0 = 10^8\,\mathrm{M}^{-1}$, so that $\overline{A} = A \times 10^8$ and, for the purpose of a good parameter estimation by using GA, we use the steady state solutions given in [10] and the experimental data given in Fig. 9 of [7] to generate the data points (50 data points per curve). The constrained GA was run several times for each agonist; the three

Table 3
The parameter values for agonists NECA and CPA obtained by using genetic algorithm

Parameters	NECA			CPA		
	RUN1	RUN2	RUN3	RUN1	RUN2	RUN3
L_1^+	18.85	19.99	6.68	19.95	14.13	3.54
L_2^+	16.10	9.84	13.92	13.65	13.72	1.42
L_3^+	0.1253	0.05	0.05	8.197	19.99	1.07
L_1^-	1498	458.22	664.44	1499	191.95	528.35
L_2^-	1457	481.01	787.19	1105	622.52	432.37
L_3^-	87.08	684.35	1346.88	1312	162.93	27.57
μ_1	2116	1805.25	2042.98	2499	64.03	75.73
μ_2	1007	37.25	981.49	53.55	5.59	29.53
μ_3	500	6.90	499.99	32.77	180.64	20.19
k_1^+	75.41	99.82	54.92	98.98	99.36	96.65
k_2^+	79.25	95.91	28.60	68.23	72.25	99.99
k_3^+	0.7303	16.53	15.79	0.1328	0.010	0.010
K^+	0.0089	0.265	0.0095	1.155	1.416	1.509
$k_{\mathrm{gtp}}G_{\beta\gamma}$	81.22	13.58	46.72	78.07	23.13	35.67
V	0.0001	0.680	0.0001	0.0021	0.3274	0.0038
k_{ec}	0.0095	577.71	0.4741	53.00	706.65	2.201
k_d	0.00004	1086	9.98	911.7	147.77	145.61
k_{rec}	453.3	2494	2476	2494	2461	2465
Fitness values	13.1123	0.00025	7.48	0.3121	0.0052	2.577

In each agonist types, three independent runs were performed to obtain the three sets of parameter values.

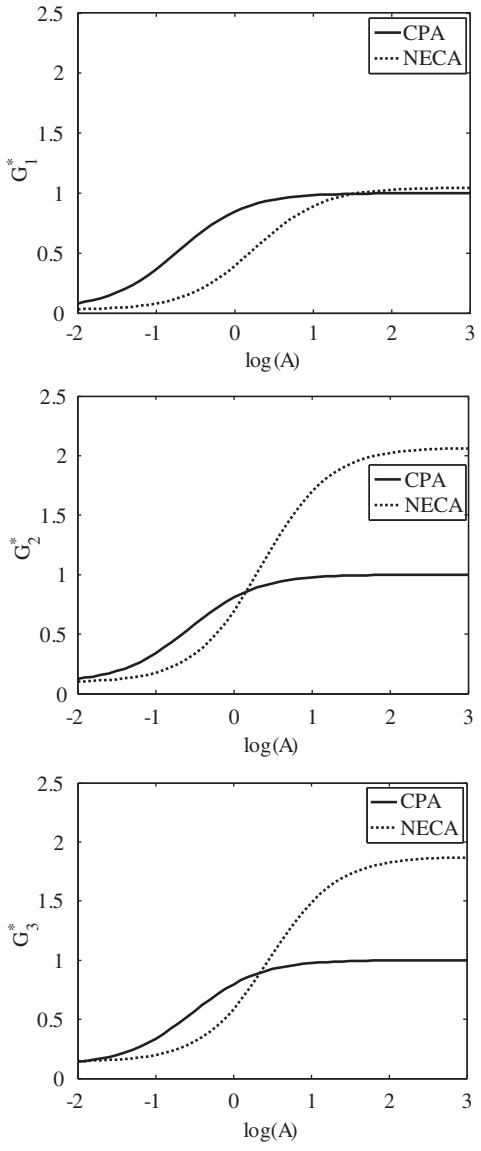

Fig. 5. Simulated data of the effect of CPA and NECA, given relative to the maximum response of CPA, on G-protein activation for G_1, G_2 and G_3. In this figure, for illustrative purposes, the parameter values of both CPA and NECA in RUN1 shown in Table 3 were used.

representative sets of parameter values for CPA and NECA are shown in Table 3. The best fitness value for NECA is 0.00025 while that of CPA is 0.0052. For both NECA and CPA, we found that the inequality $\mu_1 > \mu_2 > \mu_3$ is frequently true but not always (e.g., parameter values for CPA in RUN2). So the parameter values obtained by using the GA is frequently and qualitatively consistent with the results obtained in [7,10] where they concluded that both agonists prefer the G_1-linked pathways. The numerical solutions for the effect of CPA and NECA on the three G-proteins activation are shown in Fig. 5. They are in good agreement, at least qualitatively, with the concentration response curves reported in [7]. From this figure, we can clearly see that CPA appears to be more potent while NECA is a more efficacious drug.

If we consider the trafficking event of the receptors, we found that the internalization rate k_{ec} for the receptors which bind or

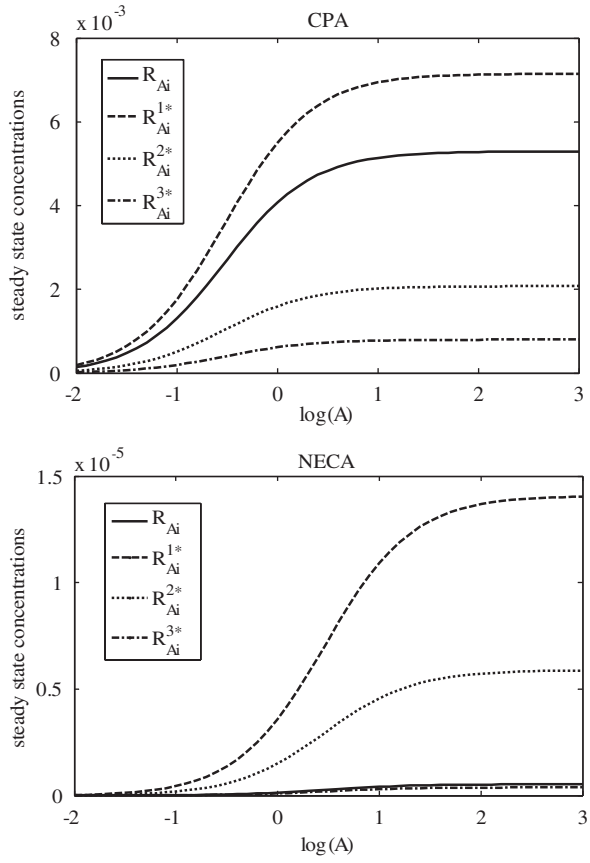

Fig. 6. Steady state concentrations of internalized receptors bound to CPA (above) and NECA (below). The concentrations of all species of internalized receptors increase as the agonist concentration increases and, at the same agonist concentration, $R_{Ai}^{1*} > R_{Ai}^{2*} > R_{Ai}^{3*}$ which is quite obvious when $\mu_1 > \mu_2 > \mu_3$. In this figure, for illustrative purposes, the parameter values of both CPA and NECA in RUN1 shown in Table 3 were used.

couple to CPA is frequently larger than the rate for those which bind or couple to NECA. This finding suggests that not only the efficacy and potency of receptors [6,8], but also the internalization event and other kinds of trafficking events of receptors may depend on the types of agonists. Fig. 6 shows the steady state concentrations of internalized receptors which bound to agonists CPA and NECA. The concentrations of all species of internalized receptors increase as more agonists become available to bind with free receptors R. The curves of steady state concentrations of all internalized activated receptor types have the same trend as that of the activated G-protein concentrations shown in Fig. 3, that is, at the same agonist concentration, $R_{Ai}^{1*} > R_{Ai}^{2*} > R_{Ai}^{3*}$, which is quite obvious when $\mu_1 > \mu_2 > \mu_3$.

5. Conclusions

In this work, we have extended a mathematical model proposed in [10] to include the trafficking events of G-protein coupled receptors (GPCR). The trafficking events we consider here include receptors synthesis, receptors internalization, recycling of receptors and receptors degradation. Taking the number of G-protein subtypes in the system to indicate the number of

receptor conformations, our model assume four receptor states, including one resting, to account for coupling separately to G_1, G_2 and G_3. When the trafficking events of receptors are integrated into the model, we have found extra information which indicates that the internalization event and other kinds of trafficking events of membrane receptors may depend on the types of agonists which bind to them. The lack of kinetic interaction rates measured in reliable *in vivo* and *in vitro* experiments is currently the major limitation to the creation of complex models of signaling pathways. Thus, we have also used the constrained GA to estimate sets of unknown parameters. With the parameter values estimated by the GA, the model is able to predict pathways-dependent agonist potency and efficacy as observed by Cordeaux et al. [7].

Acknowledgments

This work was supported by the National Center for Genetic Engineering and Biotechnology (BIOTEC), the Commission on Higher Education, the Thailand Research Fund (TRF), the Third World Academy of Sciences (TWAS) and the Thai Center of Excellence for Physics (research project in integrated physics).

References

[1] A. Persidis, Signal transduction as a drug-discovery platform, Nat. Biotechnol. 16 (11) (1998) 1082–1083.
[2] D. Filmore, It's a GPCR world, Mod. Drug Discovery 7 (11) (2004) 2002–2005.
[3] D.A. Lauffenburger, J.J. Linderman, Receptors: Models for Binding, Trafficking, and Signaling, Oxford University Press, USA, 1993.
[4] M.J. Marinissen, J.S. Gutkind, G-protein-coupled receptors and signaling networks: emerging paradigms, Trends Pharmacol. Sci. 22 (7) (2001) 368–376.
[5] S.R. Neves, P.T. Ram, R. Iyengar, G protein pathways, Science 296 (5573) (2002) 1636–1639.
[6] K.A. Berg, S. Maayani, J. Goldfarb, C. Scaramellini, P. Leff, W.P. Clarke, Effector pathway-dependent relative efficacy at serotonin type 2A and 2C receptors: evidence for agonist-directed trafficking of receptor stimulus, Mol. Pharmacol. 54 (1) (1998) 94–104.
[7] Y. Cordeaux, S.J. Briddon, A.E. Megson, J. McDonnell, J.M. Dickenson, S.J. Hill, Influence of receptor number on functional responses elicited by agonists acting at the human adenosine A(1) receptor: evidence for signaling pathway-dependent changes in agonist potency and relative intrinsic activity, Mol. Pharmacol. 58 (5) (2000) 1075–1084.
[8] T. Kenakin, Agonist-receptor efficacy. II. Agonist trafficking of receptor signals, Trends Pharmacol. Sci. 16 (7) (1995) 232–238.
[9] P. Leff, C. Scaramellini, C. Law, K. McKechnie, A three-state receptor model of agonist action, Trends Pharmacol. Sci. 18 (10) (1997) 355–362.
[10] C.Y. Chen, Y. Cordeaux, S.J. Hill, J.R. King, Modelling of signalling via G-protein coupled receptors: pathway-dependent agonist potency and efficacy, Bull. Math. Biol. 65 (5) (2003) 933–958.
[11] M.S. Brown, R.G. Anderson, J.L. Goldstein, Recycling receptors: the round-trip itinerary of migrant membrane proteins, Cell 32 (3) (1983) 663–667.
[12] T. Wileman, C. Harding, P. Stahl, Receptor-mediated endocytosis, Biochem. J. 232 (1) (1985) 1–14.
[13] B. van Deurs, O.W. Petersen, S. Olsnes, K. Sandvig, The ways of endocytosis, Int. Rev. Cytol. 117 (1989) 131–177.
[14] A.L. Schwartz, Cell biology of intracellular protein trafficking, Annu. Rev. Immunol. 8 (1990) 195–229.
[15] G. Simmons, P.R. Clapham, L. Picard, R.E. Offord, M.M. Rosenkilde, T.W. Schwartz, R. Buser, T.N. Wells, A.E. Proudfoot, Potent inhibition of HIV-1 infectivity in macrophages and lymphocytes by a novel CCR5 antagonist, Science 276 (5310) (1997) 276–279.
[16] A. Amara, S.L. Gall, O. Schwartz, J. Salamero, M. Montes, P. Loetscher, M. Baggiolini, J.L. Virelizier, F. Arenzana-Seisdedos, HIV coreceptor downregulation as antiviral principle: SDF-1alpha-dependent internalization of the chemokine receptor CXCR4 contributes to inhibition of HIV replication, J. Exp. Med. 186 (1) (1997) 139–146.
[17] T. Kenakin, Drug efficacy at G protein-coupled receptors, Annu. Rev. Pharmacol. Toxicol. 42 (2002) 349–379.
[18] J.H. Holland, Adaptation in Natural and Artificial Systems, MIT Press, Cambridge, MA, USA, 1992.
[19] H. Lin, K. Yamashita, Hybrid simplex genetic algorithm for blind equalization using RBFnetworks, 2000 IEEE International Conference on Systems, Man, and Cybernetics, vol. 1, 2000.
[20] U. Morbiducci, G.M. Andrea Tura, Genetic algorithms for parameter estimation in mathematical modeling of glucose metabolism, Comput. Biol. Med. 35 (10) (2005) 862–874.
[21] L.J. Park, C.H. Park, C. Park, T. Lee, Application of genetic algorithms to parameter estimation of bioprocesses, Med. Biol. Eng. Comput. 35 (1) (1997) 47–49.
[22] C.G. Moles, P. Mendes, J.R. Banga, Parameter estimation in biochemical pathways: a comparison of global optimization methods, Genome Res. 13 (11) (2003) 2467–2474.
[23] I. Arisi, A. Cattaneo, V. Rosato, Parameter estimate of signal transduction pathways, BMC Neurosci. 7 (Suppl. 1) (2006) S6.
[24] S. Shah, A. Kusiak, Cancer gene search with data-mining and genetic algorithms, Comput. Biol. Med. 37 (2) (2007) 251–261.
[25] S.S. Ferguson, Evolving concepts in G protein-coupled receptor endocytosis: the role in receptor desensitization and signaling, Pharmacol. Rev. 53 (1) (2001) 1–24.
[26] C.M. Tan, A.E. Brady, H.H. Nickols, Q. Wang, L.E. Limbird, Membrane trafficking of G protein-coupled receptors, Annu. Rev. Pharmacol. Toxicol. 44 (2004) 559–609.
[27] D.M. Ward, R. Ajioka, J. Kaplan, Cohort movement of different ligands and receptors in the intracellular endocytic pathway of alveolar macrophages, J. Biol. Chem. 264 (14) (1989) 8164–8170.
[28] C. Starbuck, D.A. Lauffenburger, Mathematical model for the effects of epidermal growth factor receptor trafficking dynamics on fibroblast proliferation responses, Biotechnol. Prog. 8 (2) (1992) 132–143.
[29] D.E. Goldberg, Genetic Algorithms in Search, Optimization and Machine Learning, Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA, 1989.
[30] R.E. White, Computational Mathematics: Models, Methods and Analysis with MATLAB and MPI, CRC Press, Boca Raton, 2004.
[31] A.R. Conn, N.I.M. Gould, P.L. Toint, A globally convergent augmented Langrangian algorithm for optimization with general constraints and simple bounds, SIAM J. Numer. Analy. 28 (2) (1991) 545–572.
[32] A.R. Conn, N. Gould, P.L. Toint, A globally convergent Lagrangian barrier algorithm for optimization with general inequality constraints and simple bounds, Math. Comput. 66 (217) (1997) 261–288.

Wannapong Triampo was born in Nakhonratchasima Province, Thailand, in 1970. He received a Ph.D. from Virginia Tech, USA, in 2001. After graduation, he has been with Mahidol University, Thailand, where he is an Assistant Professor in Physics. His research interests focus on biophysics and modeling.

Yongwimon Lenbury was born in Bangkok, Thailand, in 1952. He received a Ph.D. from Vanderbilt University. After graduation, she has been with Mahidol University, Thailand, where she is a Professor in Mathematics. Her research interests focus on applied mathematics and modeling.

Charin Modchang was born in Pichit, Thailand, in 1983. He received a B.Sc. degree from Physics Department, Mahidol University, Bangkok, Thailand, in 2005. Now he is a Ph.D. candidate in Physics Department, in the same university. His research interests focus on computational biophysics and theoretical biology and medicine.