

Series: "Encyclopedia of the Molecular Life Sciences"

Volume: "Biological Signaling"

Section: "Coupling of Recognition and Effect"

Title: "Coupling of recognition and effect in GPCR signaling: computational modelling approaches"

Authors: Eric Reiter[#], Romain Yvinec, Pascale Crépieux and Anne Poupon

From BIOS group, INRA, UMR85, Unité Physiologie de la Reproduction et des Comportements, F-37380 Nouzilly, France ; CNRS, UMR7247, F-37380 Nouzilly, France ; Université François Rabelais, F-37041 Tours, France

#Corresponding author:

Eric Reiter, UMR PRC, F-37380 Nouzilly, France.

Phone : + 33 2 47 42 77 83 ; Fax : + 33 2 47 42 77 43 ; e-mail : Eric.Reiter@tours.inra.fr

Synopsis (250 words)

Understanding how cells process the plethora of signals transduced by G protein-coupled receptors (GPCRs) is of paramount importance, not only to decipher the molecular basis of life, but also to facilitate innovative approaches in drug discovery. Over the last few years, GPCRs structural biology has made huge strides, strengthening the concept that these receptors exist in multiple conformations. Importantly, the nature of the conformers that are stabilized and therefore the effectors they recruit directly depend on the chemical structure of the bound ligand. Selective control of intracellular signaling pathways activated downstream of GPCRs using particular compounds that are referred to as “biased ligands” is becoming a reality, provoking a revision of the classical concept of pharmacological efficacy. Inside the cell, the networks triggered by GPCRs and the way they ultimately generate context-dependent biological responses are extremely intricate; activated GPCRs can couple to multiple transducers; homo- and hetero- receptor dimers can recruit distinct effectors in response to the same ligand; intracellular signaling effectors can be shared by different networks and their local concentrations and trafficking are tightly regulated; cellular outcomes generally require a combination of upstream regulatory events; finally different extracellular signals co-exist, eliciting composite responses. To face this huge complexity, computational modeling has become critical. This chapter reviews the existing modeling strategies and how they can be combined with experimental biology in order to build quantitative and predictive computational models of GPCR-induced transduction mechanisms and intracellular networks.

Introduction

GPCR pharmacology has profoundly evolved over the last decade as it is now admitted that particular ligands, generally referred to as “biased”, can selectively modulate a subset of the signaling events triggered by the full agonist while having no or even negative effects on the other signaling events. This concept is consistent with the accumulating structural data demonstrating the existence of multiple inactive and active conformations of GPCRs. This raises the possibility to develop new classes of drugs with more selective actions hence fewer side effects. However, this also demands a redefinition of the classical representation of pharmacological efficacy and of the methods that allow its estimation ([1](#)).

The general representation of how GPCRs activate intracellular signaling and eventually modify cell fate has also dramatically evolved over the past decades. Discrete linear signaling cascades have long been thought to connect GPCRs to transcription factors. This simplistic representation has progressively moved towards a paradigm in which communication networks made of multi-protein ensembles convey signals and integrate them into adapted biological outcomes. Several layers of complexity have been identified and contribute to the remarkable processing capability of these signaling networks: i) multiple ligands hit a target cell at the same time, some of them in a pulsatile manner, and must be processed simultaneously, directly impacting the activation dynamics of the intracellular networks; ii) any given plasma membrane GPCR, acting as a monomer or participating to a homodimer, a heterodimer or a higher order oligomer, can connect to multiple transduction mechanisms; iii) receptors belonging to different classes can be associated into signalosomes which then trigger intracellular signals that are spatially and temporally encoded; iv) feedbacks, feedforward controls and crosstalks also contribute to the processing of signals through

networks; v) signaling pathways convey quantitative information reflecting the strength of the stimulus by taking advantage of their nonlinear nature to convert stimulus intensity into signal duration, even leading to oscillations in some cases (2).

Deciphering how all these entities operate together at the molecular level clearly represents a major challenge for the whole field and could open novel avenues towards the development of pathway-selective drugs. It is becoming increasingly clear that mathematical and computational approaches are necessary to tackle the high level of complexity involved.

Intrinsic efficacy and its measure

GPCR ligands have long been characterized by two fundamental parameters: i) affinities which reflect the tightness of the ligand/receptor interaction and ii) intrinsic efficacies which account for the ability of a receptor-bound ligand to elicit a biological response by recruiting effector proteins. The combination of these two parameters defines the potency.

Accordingly, two agonists can be equipotent while having different intrinsic efficacies and affinities. In the classical two-state model, ligands were considered to have linear efficacy, a receptor existing in either an inactive or an active conformation. In this view, the magnitudes of responses of different signaling pathways activated by a GPCR/ligand pair were expected to be always proportional to the intrinsic efficacy of this ligand. Practically, this means that for any given GPCR/ligand pair, all functional read-outs were predicted to lead to the same pharmacological profile (1).

An increasing number of observations do not fit with this theory. Biophysical and structural data have recently accumulated for several GPCRs and provide evidence that unliganded as

well as activated GPCRs actually exist in multiple active and inactive conformations.

Extending this idea, the existence of ligand-specific conformations has been substantiated using a vast array of approaches such as molecular dynamics, nuclear magnetic resonance spectroscopy, fluorescence spectroscopy, quantitative mass spectrometry or intra-molecular FRET. Collectively, all these findings lead to a novel paradigm in which each ligand stabilizes a defined set of active conformations of the receptor. Importantly, a receptor conformer reaches optimal stability when it simultaneously binds a ligand on the extracellular side and an effector protein in the cytosol (3). If different (biased) ligands can stabilize different conformers of a receptor, they can also differentially recruit and activate intracellular effectors (1, 3). Defining and calculating intrinsic efficacies in this paradigm becomes much more challenging because of the multidimensional nature of the problem (4).

Despite the wide acceptance of the concept of ligand bias, the methodology allowing to rigorously quantify it remains a subject of debate. Experimentally, the ideal situation would be, for each ligand/receptor pair, to be able to measure the kinetics of association with all the effector proteins recruited as well as their respective degree of activation. This task is extremely difficult to achieve with the current experimental methods. Instead, downstream read-outs (second messengers, kinase activities, gene expression, etc) are often measured. Measurements are not always in real-time, many end-point assays are used with reaction times ranging from seconds (calcium) to minutes (cAMP, IP, kinase activities) or hours (reporter genes). One of the main issues is the difficulty to compare signaling events that dramatically differ in their degree of amplification (e.g.: second messenger or reporter genes *versus* β -arrestin recruitment), thereby differentially reflecting the amount of spare receptors (*i.e.*: a small fraction of ligand-bound receptor generates a maximal second

messenger response whereas full β -arrestin recruitment necessitates full receptor occupancy). Moreover, the activation level of downstream read-outs only partially depends on the intrinsic efficacy of the ligand. Many intermediate reactions and cross-talks also impact on the end-point result. Inferring intrinsic efficacies from such heterogeneous data is very difficult and can lead to erroneous estimations.

Several methods, all derived from Black and Leff's operational model (5), have been proposed recently, each presenting assets and drawbacks (4, 6). All of them are based on the acquisition of sigmoidal dose-response curves for all the read-outs, neglecting the kinetics of the different responses that can occur at different time scales (Figure 1A, D).

Novel experimental methods based on different variation of BRET, FRET or protein complementation assays now allow measuring many molecular events (interactions, activities, etc) in real-time in living cells. Tools have been developed to measure the recruitment and activation of many key effectors involved in GPCR transduction.

Paradoxically, in current practice, only the area under the curves is retained and plotted as sigmoidal dose response signal.

For all these reasons, we believe that dynamical modeling of GPCR activation would represent an asset over the existing methods. Indeed, dynamical models take into account both quantitative and temporal aspects of GPCR activation and would allow computing intrinsic efficacies for each ligand/conformer pair compared to a reference (Figure 1B).

When downstream read outs are measured, dynamical models take into account concentrations and dynamics of all the shared effectors as well as cross-talks and complexation with effector proteins (Figure 1C). The mathematical formalisms and the

parameter optimization methods that could be used to infer intrinsic efficacies are similar as those used to build signaling network models and are described below.

From receptor transduction to intracellular signaling networks activation

As discussed in the previous section, intrinsic efficacy is a fundamental characteristic inherent to every GPCR ligand, which describes the way a given GPCR-bound ligand stabilizes a specific subset of receptor conformers and how these conformers are functionally coupled to different transduction mechanisms. All these effectors, activated at the receptor level, can be viewed as delivering the input signals that control intricate signaling networks inside the cell. As these networks process all the extracellular information, they are ultimately responsible for triggering complex cellular responses such as: proliferation, differentiation, apoptosis, gene regulation, etc. Therefore, it is of paramount importance to decipher how intrinsic efficacies at the receptor level convey input signals to intracellular networks, how these signals are propagated, processed and ultimately lead to adapted outcomes. For all the reasons already discussed above, this represents a very challenging task and cannot be tackled without computational modelling.

Generating experimental data to build models

It is important to bear in mind that the quality of the computational model directly depends on the experimental data that can be generated. The experimental methods that are available to study cellular signaling are rapidly evolving, paving the way for better, more predictive computational models.

The analysis of protein phosphorylation profiles in GPCR-stimulated cells is a classical way to assess the activation of intracellular signaling pathways. Typically, phosphorylation dynamics are measured with phosphospecific antibodies combined to Western blotting. However, this approach requires large amounts of biological material and due to its poor throughput, does not allow large-scale analysis. More recently, phosphorylated peptides from complex samples have been analyzed by mass spectrometry, typically allowing the identification of thousands of phosphopeptides per sample. However, these MS-based methods allow comparing only a limited number of biological conditions. In contrast, automated spotting of protein extracts, such as Reverse-Phase Protein Array or microwestern arrays, allow simultaneous analysis of thousands of samples with many different phosphor-specific antibodies. Cell-sorting-based approaches are also capable of acquiring cell signaling data with high throughput and reliability (2). Thanks to this substantial throughput, kinetics, dose-responses, and targeted perturbations can be systematically carried out and analyzed.

The identification and analysis of multi-protein complexes are essential to decipher the architecture of signaling networks. High throughput approaches have long been developed to tackle this problem: yeast two hybrid screens, tandem-affinity purification and mass spectrometry or luminescence-based mammalian interactome mapping have been successfully used at a genome-wide scale (2).

One major breakthrough has been the advent of experimental approaches allowing monitoring GPCR-induced signaling events in real time in living cells (7). In particular, approaches that use biosensors based on fluorescence, bioluminescence and/or protein complementation combine unmatched spatial and/or kinetics resolutions of signaling activities. Protein-protein interactions and changes in protein conformation can also be

monitored in real time in living cells using similar approaches. This obviously represents great assets when developing computational models. Genetically encoded fluorescence resonance energy transfer (FRET)-based sensors have helped capturing the spatiotemporal patterns of second messengers, kinase activities and GPCR activation. As FRET data are generally acquired using fluorescence microscopy, they allow measurements of signaling events in single cells, a feature amenable to stochastic modeling. Bioluminescence resonance energy transfer (BRET) has also been extensively used to monitor signaling, protein-protein interactions and conformational changes in living cells. BRET is typically carried-out in multi-well plate format and ensures the production of huge amounts of dynamic data from cell populations that are very well suited for deterministic dynamical modeling.

Another fundamental aspect when building computational models is the ability to apply targeted perturbations and to measure the consequences on the network behavior. Perturbations are very useful when building the model to reduce the number of possible connections within the network as well as, at a later stage, when testing the validity of model-driven predictions. There are many different options to perturb GPCR-induced signaling networks. The perturbations can be applied at the ligand level by using biased compounds or by applying non-monotonous stimulation patterns. Mutations and/or polymorphisms occurring at the receptor level can also lead to biased intracellular responses. Various targeted perturbation approaches such as specific kinase inhibitors, dominant negative constructs, or interfering RNAs, allow hitting specific nodes within the signaling network and are therefore invaluable tools ([2](#)).

Building computational models

Different types of mathematical formalisms can be used depending on the situation, the available data and the modelling philosophy. Models can be static or dynamic, they can represent time in a discrete or continuous fashion and they can be deterministic or stochastic (*i.e.*: governed by random laws).

Static models are mathematical formalizations of the molecular species involved in signaling networks and the type of relationships they have: catalysis, physical interaction, etc (8). Over the last years, knowledge from literature has been made accessible through databases.

Examples of such databases are Reactome, in which pathways are manually curated and peer-reviewed; WikiPathways, which is based on community curation or NetPath which gives access to the references used to build the pathways. An extensive list of databases can be found on the Pathguide website: <http://www.pathguide.org>. In addition, published models are increasingly deposited in the Biomodels database. Cytoscape or Pathway Studio are tools directly connected to these databases, to enable the automatic reconstruction and the graphical visualization of large networks.

In parallel, various automated methods have been developed in order to build static models directly from data sets. In particular, many methods are available to build transcription regulation networks from microarray data. Methods have also been designed to exploit other type of data such as RNAi or phosphoproteomic. However, these methods are based on the exploitation of a single experimental method whereas experimental exploration of a signaling system generally requires a combination of multiple methods. Attempts to formalize the elementary reasoning elements using first order logic rules are currently being

explored. To infer a given signaling network, these rules, together with elementary experimental facts, feed an inference engine which reconstructs the network (9).

GPCR-induced signaling events involve fast and reversible reactions. It is therefore fundamentally important to take the dynamics into consideration when trying to model and predict the recognition/effect relationships in GPCR signaling. Static models described above are often too detailed, involving hundreds of molecules, and therefore are not amenable to dynamic modeling. Consequently, some molecules have to be eliminated in a process known as “model reduction” in order to obtain a model that retains the essential properties of the complete system with a limited number of variables.

The ordinary differential equations (ODE) formalism remains very popular when building dynamical models. In this formalism, the concentration variation of a given molecule is expressed as a function of time and depends on concentrations of the other molecules within the network. Different rate laws can be chosen. The simplest kinetic law is derived from mass action law. For enzymatic reactions, many authors still prefer Michaëlis-Menten kinetic law (Figure 2). However, for most instances and using the right parameters, the simulations of both kinetic laws are superimposable with mass action law generating less parameters (2).

Another popular formalism is the Bayesian network. In this formalism the time course is discrete, and at each time point, each molecule has a probability to undergo a state transition. This method has the advantage to represent more accurately the reality of molecular reactions and, unlike ODE, allows simulating stochastic events. However, computation time rapidly explodes when modelling large networks.

Once the model is precisely defined and all the parameters are known (kinetic constants, affinities, initial concentrations, etc) the system can be simulated to estimate the evolution of the concentrations of the different molecular species as a function of time. Different software such as CellDesigner or Copasi allow performing these simulations easily.

Estimating parameters

The most challenging part of dynamical modelling is certainly the determination of the parameter values. Indeed, measuring rate constants or local concentrations in living cells is a difficult task. Moreover, the models are often simplified, and successive reactions are aggregated in the model-reduction process. In this case, the kinetic constants cannot be experimentally approached. Some authors approximate parameter values from publications, even though they were often obtained in different biological conditions, or to use “educated guesses”. Another more rigorous option is to search for a parameter set that, when used to simulate the model dynamics, reproduces a set of experimental data. This process is called “parameter estimation” (Figure 3) and has become an important challenge. Classical continuous parameter optimization methods include genetic algorithms, evolutionary programming, Hooke-Jeeves, Levenberg-Marquardt, Praxis, Particle swarm, Steepest descent, SSmGo or BIOCHAM ([2](#)).

The goal when developing dynamical models is to be able to predict the behavior of the system in biological conditions different from those used to build and calibrate the model. For instance, we have recently shown, using the Angiotensin type I receptor-mediated signaling pathway as a benchmark, that a topologically correct model and accurate parameter estimation can lead to valid predictions. Indeed, in this work we were able to

propose multiple predictions using the model that were experimentally validated at a later stage ([10](#)).

Conclusions

Modeling GPCR-induced transduction and the resulting signaling networks is a difficult but important task. Experimental methods now exist to acquire dynamic data with the precision and throughput compatible with modeling. This data acquisition step is essential and has to be carefully planned. The nature of the data to acquire, but also the frequency of measurements and the type of targeted perturbation to apply, directly impact on the quality of the model. Computational tools for model building, reduction and parameter estimation are available to build models that can be predictive. In our opinion, in the near future, computational modelling will improve the way intrinsic efficacies are measured, especially in the case of biased ligands. In addition, models of signaling networks could lead to an unprecedented understanding of the cellular machinery and how the multiple signals triggered by a given receptor are processed in an adapted biological response. Although a lot still has to be done, models can already be used to explore the possible behaviors of the cell as a function of external stimuli, thereby facilitating the rational design of drugs with more selective action.

Figure legends

Figure 1: Determining ligand bias. **A:** In methods derived from Black & Leff's operational models, the ligand L associates with the receptor to give the LR complex, which directly

triggers effects 1 and 2 (E1 and E2). The computation of the bias is valid under the hypotheses that the ligand largely exceeds the receptor, and that at equilibrium the quantity of ligand-receptor complex remains constant. **B**: the ligand-receptor complex can reversibly adopt a spectrum of different conformations, each triggering one of the observed effects. **C**: triggering of both effects relies on the complexation of the ligand-receptor complexes with an effector protein (P1 and P2). **D**: toy example of two different ligands leading to same global effects (area under the curves) but with different dynamics (top curves). If observed at a given time point, they give different dose-response curves (bottom left), but if observed at end-point, when the system is back to equilibrium, they give the same dose-response curves (bottom right).

Figure 2: Formalism for dynamic models. General differential expression of the variation of B's concentration when considering a simple mass action law (1) or a state transition depending on a third molecule (C), for example state transition mediated by an enzyme (2). Alternatively, Michaëlis-Menten model for a reversible state transition mediated by an enzyme (C) can be used (3).

Figure 3: Computing the error between experiment and simulation. **A**: The error is computed as a function of the differences between simulation and experimental measure for each observable (i) and each time point (t). Classical error functions are sums of the square of the δs . **B**: Error landscape for a simple two parameters problem. The red region represents the global minimum, which should be reached when estimating the parameters. This landscape also presents two local minima, in which the optimization algorithms might get trapped. Of course real optimization problems are in n dimensions with n being the number of parameter to estimate which is typically at least a few tenth in signaling network models.

Acknowledgements

Only ten references being allowed in this paper, we apologize to the many authors whose work could not be cited. This work was supported by Region Centre and ANR research grants.

References

1. Reiter E, Ahn S, Shukla AK, Lefkowitz RJ 2012 Molecular mechanism of beta-arrestin-biased agonism at seven-transmembrane receptors. *Annu Rev Pharmacol Toxicol* 52:179-197
2. Poupon A, Reiter E 2013 Computational models to decipher cell signaling pathways. In *Cellular endocrinology in health and disease*, Ulloa-Aguirre A and Conn MP (editors), Elsevier, 269-284
3. Nygaard R, Zou Y, Dror RO, Mildorf TJ, Arlow DH, Manglik A, Pan AC, Liu CW, Fung JJ, Bokoch MP, Thian FS, Kobilka TS, Shaw DE, Mueller L, Prosser RS, Kobilka BK 2013 The dynamic process of beta(2)-adrenergic receptor activation. *Cell* 152:532-542
4. Kenakin T, Christopoulos A 2013 Signalling bias in new drug discovery: detection, quantification and therapeutic impact. *Nature reviews Drug discovery* 12:205-216
5. Black JW, Leff P 1983 Operational models of pharmacological agonism. *Proceedings of the Royal Society of London Series B, Containing papers of a Biological character Royal Society* 220:141-162
6. Landomiel F, Gallay N, Jegot G, Tranchant T, Durand G, Bourquard T, Crepieux P, Poupon A, Reiter E 2014 Biased signalling in follicle stimulating hormone action. *Mol Cell Endocrinol* 382:452-459
7. Lohse MJ, Nuber S, Hoffmann C 2012 Fluorescence/bioluminescence resonance energy transfer techniques to study G-protein-coupled receptor activation and signaling. *Pharmacological reviews* 64:299-336
8. Gloaguen P, Crepieux P, Heitzler D, Poupon A, Reiter E 2011 Mapping the follicle-stimulating hormone-induced signaling networks. *Front Endocrinol (Lausanne)* 2:45
9. Aslaoui-Errafi Z, Cohen-Boulakia S, Froidevaux C, Gloaguen P, Poupon A, Rougny A, Yahiaoui M 2012 Towards a logic-based method to infer provenance-aware molecular networks. *1st ECML/PKDD International workshop on Learning and Discovery in Symbolic Systems Biology (LDSSB):103-110*
10. Heitzler D, Durand G, Gallay N, Rizk A, Ahn S, Kim J, Violin JD, Dupuy L, Gauthier C, Piketty V, Crepieux P, Poupon A, Clement F, Fages F, Lefkowitz RJ, Reiter E 2012 Competing G

protein-coupled receptor kinases balance G protein and beta-arrestin signaling. Mol Syst Biol 8:590