Chapter 28

An Interdisciplinary Approach for Designing Kinetic Models of the Ras/MAPK Signaling Pathway

Marcelo S. Reis, Vincent Noël, Matheus H. Dias, Layra L. Albuquerque, Amanda S. Guimarães, Lulu Wu, Junior Barrera, and Hugo A. Armelin

Abstract

We present in this article a methodology for designing kinetic models of molecular signaling networks, which was exemplarily applied for modeling one of the Ras/MAPK signaling pathways in the mouse Y1 adrenocortical cell line. The methodology is interdisciplinary, that is, it was developed in a way that both dry and wet lab teams worked together along the whole modeling process.

Key words Cell signaling pathways, Ras, MAPK, ERK, Western blot, ELISA, Ordinary differential equation, Differential-algebraic equation, Model-fitting analysis, Model simplification, Nonlinear optimization

1 Introduction

One of the current challenges of molecular cell biology is to unravel mechanisms underlying molecular signaling pathways, especially in the context of how extracellular signals propagate from cell surface to cell nucleus, with implications in cell cycle control and ultimately in cell fate. An important example of such signaling axes is the growth factor-activated Rat sarcoma (Ras)/mitogen-activated protein kinase (MAPK) cascade, which mediates a myriad of cellular processes, from cell growth to proliferation and death [1]. To tackle this challenge, intuitive navigations through static interactome maps are not sufficient, since in cell signaling, it is necessary to know concentration changes of involved chemical species along a given time frame. Thus, designing kinetic models is essential in mechanistic studies of molecular signaling networks. However, the modeling process involves a sequence of complex procedures, and a single mistake at one of them might yield an incorrect kinetic model. Moreover, the complexity of

the kinetic model grows exponentially as a function of the number of considered chemical species, rendering impossible manual fitting of the model. Therefore, there is a need for a systematic approach to design and simulate kinetic models of molecular signaling networks, including the Ras/MAPK pathway.

In this work, we present an approach for designing kinetic models of molecular signaling networks, which was exemplarily applied on the Ras/MAPK signaling pathway. This methodology is intrinsically interdisciplinary, in the sense that it combines in silico procedures with biological experiments, i.e., both dry and wet lab teams work together along the whole modeling process. Additionally, our methodology associates classical approaches for modeling of biochemical reactions [2] with original solutions at critical steps. Namely, for optimal selection of chemical species in time-course experiments, we systematically test different hypotheses involving relevant biochemical reactions of the signaling pathway under focus. To carry out these procedures, we developed and implemented a computational framework, which is available to the scientific community. Finally, we illustrated the usage of present methodology applying it for modeling one of the Ras/MAPK pathways in mouse Y1 adrenal tumor cells stimulated by fibroblast growth factor 2 (FGF2), namely, the Ras/MAPK pathway, whose MAPK component is composed of isoforms of the extracellular signal-regulated kinase (ERK).

2   Materials

2.1   Experimental Data

1. Mouse Y1 adrenocortical carcinoma cell line [3] was originally obtained from the American Type Culture Collection in 1973 and kept frozen in liquid nitrogen ever since.

2. 60 mm cell culture plates.

3. Y1 culture medium: Dulbecco's modified Eagle's medium supplemented with 2 mM glutamine, 100 U/mL penicillin, 100 mg/mL streptomycin, and 10% fetal calf serum.

4. Serum-free medium: Dulbecco's modified Eagle's medium supplemented with 2 mM glutamine, 100 U/mL penicillin, and 100 mg/mL streptomycin.

5. Recombinant FGF2. 1 mg/mL stock solution in PBS pH 7.4.

6. RIPA lysis buffer: 150 mM NaCl, 1.0% Nonidet P-40, 0.5% sodium deoxycholate, 0.1% SDS, and 50 mM Tris, pH 8.0.

7. Halt protease and phosphatase inhibitor cocktail.

8. Eppendorf tubes.

9. Refrigerated microcentrifuge.

10. 10% SDS-PAGE gels.

11. Nitrocellulose membranes.

12. SDS-PAGE apparatus.

13. TBS-T buffer: 150 mM NaCl, 50 mM Tris [pH 8.0], and 0.1% Tween 20.

14. Nonfat dried milk.

15. Bovine serum albumin (BSA).

16. Rocker platform agitator.

17. Antibodies: phospho-ERK1/phospho-ERK2 (Thr202/Tyr204) (#9101) and ERK1/ERK2 (#9102) from Cell Signaling, hypoxanthine-guanine phosphoribosyltransferase (HPRT) (sc-20975) from Santa Cruz Biotechnology, and goat anti-rabbit peroxidase-linked antibody (474-1506) from KPL.

18. SuperSignal West Pico Chemiluminescent Substrate.

19. G-LISA® Ras Activation Assay Biochem kit (BK131) from Cytoskeleton.

20. Western blot imaging system: Uvitec Alliance 9.7 equipment.

21. Additional reagents and equipment for basic cell culture techniques.

2.2 Retrieval of Biological Knowledge

Values for the initial concentrations of the chemical species and/or rate constants for the chemical reactions were obtained from the literature, including repositories such as BioNumbers [4]. We extracted signaling pathway interactomes from databases such as Kyoto Encyclopedia of Genes and Genomes (KEGG) [5]. Additionally, we departed from an already formalized kinetic model, which can be obtained from the BioModels database [6]. Formalized kinetic models can be used either alone or coupled with other models; in either case, modifications in the initial model may also be applied according to prior knowledge of the cellular system being modeled.

2.3 Computational Tools

In order to design and adjust kinetic models to our experimental data, we developed a computational framework called Signaling Network Simulator (SigNetSim). This framework was coded in Python programming language and can be used through a user-friendly web interface [7]. In SigNetSim, we describe a kinetic model in the standard Systems Biology Markup Language (SBML) format, which enables us to reuse models that were already published and also to make easier the usage of our models by other researchers. The web interface of SigNetSim also includes a simple data repository, which can be used either to simulate or to optimize models employing different experimental conditions.

Rate constant adjustment of kinetic models is very demanding in computational resources, since the search space being exponential on the number of rate constants. Hence, we implemented in our framework an efficient simulated annealing algorithm [8]. This algorithm, which was coded in C programming language, works in

parallel and can be used easily on most modern computational servers.

SigNetSim was used successfully for modeling the in vitro competition between the oxidized form of nicotinamide adenine dinucleotide (NAD^+) and telomeric sequences for the binding to glyceraldehyde 3-phosphate dehydrogenase (GAPDH) [9] and also for educational purposes.

3 Methods

The outline of our modeling methodology is presented in Fig. 1. There are six major procedures, which will be explained in details in Subheadings 3.1–3.6. During this whole section, we will use as study case the modeling of the Ras/ERK signaling pathway in Y1 cells. These cells display constitutive high levels of active Ras, that is, Ras bound to guanosine triphosphate (Ras-GTP), due to wild-type K-Ras amplification and overexpression. Despite this, Ras-GTP levels can be further increased by mitogen stimulation, and MAPK pathway activation remains fully dependent on such signaling [10–12]. Therefore, we will be interested in obtaining a phenomenological model to explain the activation of ERK as a function of levels of Ras-GTP, which in turn are modulated by stimulation of FGF2. To this end, we will depart from the whole canonical Ras/MAPK signaling pathway as it is described in *Mus musculus* MAPK pathway at KEGG database (Fig. 2).

3.1 Selection of the Involved Biochemical Species and Interactions

The first procedure is, given an interactome map, to select a minimal set of chemical species whose kinetic interactions among them might be sufficient to explain the observed phenomenon (*see* **Note 1**). In our example, we chose the three chemical species necessary to propagate a signal from Ras-GTP to ERK: rapidly accelerated fibrosarcoma (Raf), MAPK/ERK kinase (MEK), and ERK (Fig. 2, proteins and interactions in bold) (*see* **Note 2**). The remaining chemical species that interact with them (e.g., Ras-GTP, phosphatases, etc.) compose the inputs of our kinetic model and will be discussed in the next subsections.

The next step is to define the biochemical reactions whose kinetics are relevant to the signal propagation through the signaling pathway; they can be one of the following:

1. A first-order reaction, which may or not be reversible. This type of reaction is used to describe, for instance, the binding of a ligand to a receptor (e.g., FGF2 binding to its receptor, FGFR) or the migration of a chemical species from one cell compartment to another (e.g., phosphorylated ERK migrating from cytosol to nucleus).
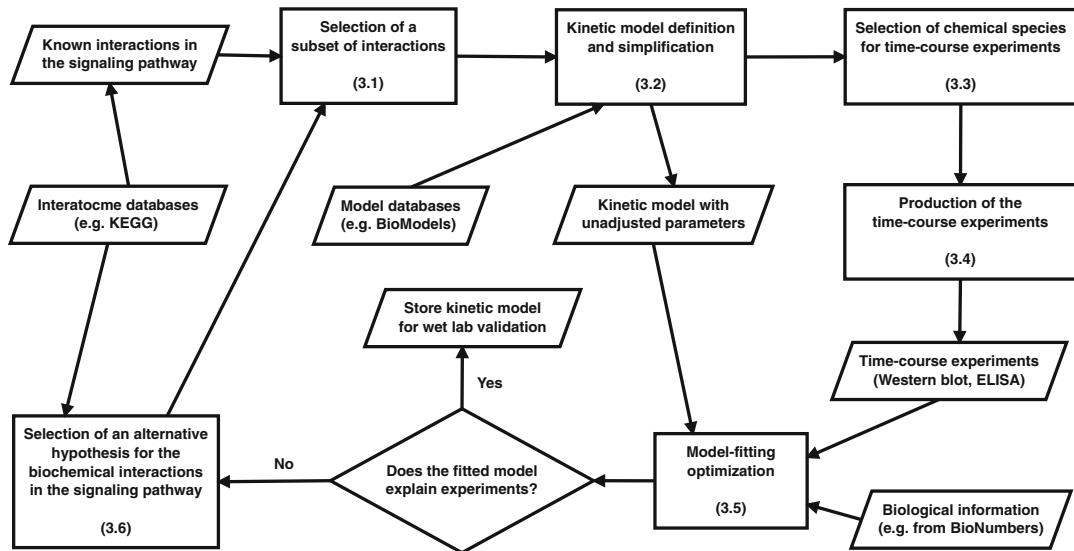
Fig. 1 Fluxogram of the methodology presented in this paper. The equiangular and non-equiangular rectangles represent, respectively, procedures and data, while the diamond denotes decision. The numbers between parentheses assign the major procedures that are described in details in Subheadings 3.1–3.6
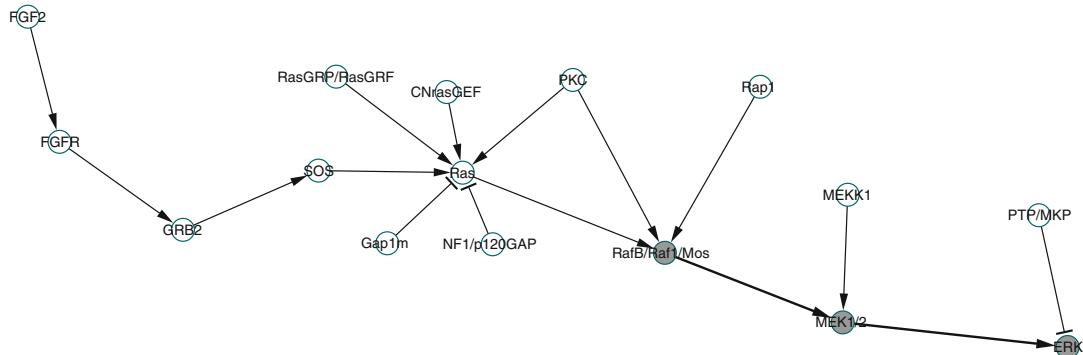
Fig. 2 Part of *Mus musculus* MAPK signaling pathway interactome, as it was described in KEGG database (map 4010, updated in January 18, 2016). This section of the interactome depicts the main signaling pathway, from the binding of FGF2 to its receptor (FGFR), passing through the Ras small GTPase, and finally reaching the cascade Raf-MEK-ERK (in *bold*)
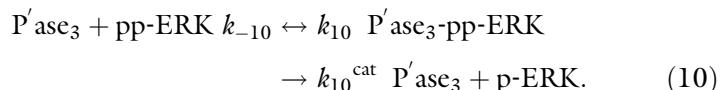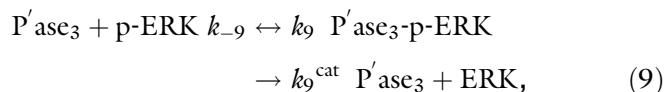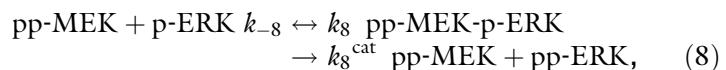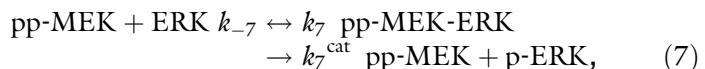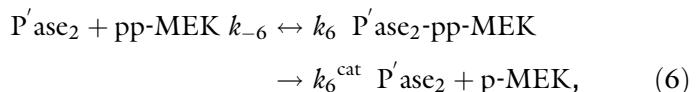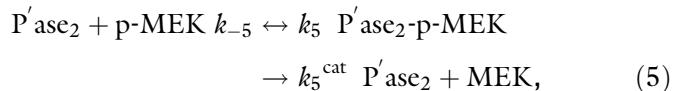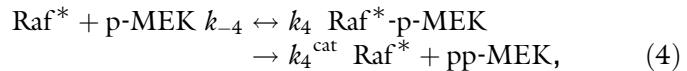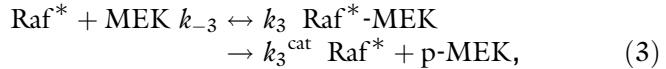
2. A second-order reaction, which also may or not be reversible. This type of reaction is useful to describe, for instance, the association of two monomers, resulting in a protein dimer.

3. An enzymatic reaction, which can be described as a reversible second-order reaction (enzyme binding to substrate, hence forming an enzyme-substrate complex) coupled with an irreversible first-order reaction (complex resulting in product and releasing the enzyme). For instance, in our study case, the enzymatic

reaction that describes the activation of Raf by Ras-GTP (*see* **Note 3**) is given by:

$$\text{Ras-GTP} + \text{Raf } k_{-1} \leftrightarrow k_1 \text{Ras-GTP-Raf}$$
$$\rightarrow k_1{}^{\text{cat}}\text{Ras-GTP} + \text{Raf}^*, \qquad (1)$$

where k_1 and k_{-1} are the rate constants of the formation of the enzyme-substrate complex and the dissociation of the complex, respectively, and $k_1{}^{\text{cat}}$ is the rate constant of the transformation of substrate into product.

In the modeling of signaling pathways, cascade of kinases such as MAPKs is mostly composed by a chain of enzymatic reactions. Hence, coupled with the biochemical reaction depicted in Eq. 1, we can describe the signal flow through the Raf-MEK-ERK system with the following reactions:

$$\text{P}'\text{ase}_1 + \text{Raf}^* \ k_{-2} \leftrightarrow k_2 \ \text{P}'\text{ase}_1\text{-Raf}^* \rightarrow k_2{}^{\text{cat}} \ \text{P}'\text{ase}_1 + \text{Raf}, \quad (2)$$

$$\text{Raf}^* + \text{MEK } k_{-3} \leftrightarrow k_3 \ \text{Raf}^*\text{-MEK}$$
$$\rightarrow k_3{}^{\text{cat}} \ \text{Raf}^* + \text{p-MEK}, \qquad (3)$$

$$\text{Raf}^* + \text{p-MEK } k_{-4} \leftrightarrow k_4 \ \text{Raf}^*\text{-p-MEK}$$
$$\rightarrow k_4{}^{\text{cat}} \ \text{Raf}^* + \text{pp-MEK}, \qquad (4)$$

$$\text{P}'\text{ase}_2 + \text{p-MEK } k_{-5} \leftrightarrow k_5 \ \text{P}'\text{ase}_2\text{-p-MEK}$$
$$\rightarrow k_5{}^{\text{cat}} \ \text{P}'\text{ase}_2 + \text{MEK}, \qquad (5)$$

$$\text{P}'\text{ase}_2 + \text{pp-MEK } k_{-6} \leftrightarrow k_6 \ \text{P}'\text{ase}_2\text{-pp-MEK}$$
$$\rightarrow k_6{}^{\text{cat}} \ \text{P}'\text{ase}_2 + \text{p-MEK}, \qquad (6)$$

$$\text{pp-MEK} + \text{ERK } k_{-7} \leftrightarrow k_7 \ \text{pp-MEK-ERK}$$
$$\rightarrow k_7{}^{\text{cat}} \ \text{pp-MEK} + \text{p-ERK}, \qquad (7)$$

$$\text{pp-MEK} + \text{p-ERK } k_{-8} \leftrightarrow k_8 \ \text{pp-MEK-p-ERK}$$
$$\rightarrow k_8{}^{\text{cat}} \ \text{pp-MEK} + \text{pp-ERK}, \qquad (8)$$

$$\text{P}'\text{ase}_3 + \text{p-ERK } k_{-9} \leftrightarrow k_9 \ \text{P}'\text{ase}_3\text{-p-ERK}$$
$$\rightarrow k_9{}^{\text{cat}} \ \text{P}'\text{ase}_3 + \text{ERK}, \qquad (9)$$

$$\text{P}'\text{ase}_3 + \text{pp-ERK } k_{-10} \leftrightarrow k_{10} \ \text{P}'\text{ase}_3\text{-pp-ERK}$$
$$\rightarrow k_{10}{}^{\text{cat}} \ \text{P}'\text{ase}_3 + \text{p-ERK}. \qquad (10)$$

Having defined the set of biochemical reactions that describe the kinetics of the signaling pathway, the next step is to establish a mathematic description of the kinetics of these reactions.

3.2 Kinetic Model Definition and Simplification

Among different possible mathematical formalisms, in our methodology, we employ systems of ordinary differential equations (ODEs) (*see* **Note 4**). Although we could map directly the biochemical reactions of Eqs. 1–10 as a system of coupled ODEs (system S1.1 in [13]), at this point we can resort to a model already available in the literature to serve as "scaffold" in the modeling process. There is a myriad of MAPK kinetic models in the literature [14]; for a comprehensive survey on MAPK kinetic models, refer to [15]. We started our modeling with the classic MAPK model of Huang and Ferrell [16], whose most of contemporary MAPK kinetic models stem from [15], complemented by the adoption of the quasi-steady-state (QSS) approximation for Michaelis-Menten kinetics [17]. We show in Fig. 3 the set of reactions with the QSS approximation for the biochemical reactions of Eqs. 1–10, which yields the following simplified system of ODEs:

$$d[\text{Raf}]/dt = k_2{}^{\text{cat}}\left[\text{P}'\text{ase}_1\right]\left[\text{Raf}^*\right]/\left(K_2 m + \left[\text{Raf}^*\right]\right)$$
$$- k_1{}^{\text{cat}}[\text{Ras-GTP}][\text{Raf}]/(K_1 m + [\text{Raf}]) \qquad (11)$$

$$d\left[\text{Raf}^*\right]/dt = k_1{}^{\text{cat}}[\text{Ras-GTP}][\text{Raf}]/(K_1 m + [\text{Raf}])$$
$$- k_2{}^{\text{cat}}\left[\text{P}'\text{ase}_1\right]\left[\text{Raf}^*\right]/\left(K_2 m + \left[\text{Raf}^*\right]\right) \qquad (12)$$

$$d[\text{MEK}]/dt = k_5{}^{\text{cat}}\left[\text{P}'\text{ase}_2\right][\text{p-MEK}]/(K_5 m + [\text{p-MEK}])$$
$$- k_3{}^{\text{cat}}\left[\text{Raf}^*\right][\text{MEK}]/(K_3 m + [\text{MEK}]) \qquad (13)$$

$$d[\text{p-MEK}]/dt = k_3{}^{\text{cat}}\left[\text{Raf}^*\right][\text{MEK}]/(K_3 m + [\text{MEK}])$$
$$- k_4{}^{\text{cat}}\left[\text{Raf}^*\right][\text{p-MEK}]/(K_4 m + [\text{p-MEK}])$$
$$- k_5{}^{\text{cat}}\left[\text{P}'\text{ase}2\right][\text{p-MEK}]/(K_5 m + [\text{p-MEK}])$$
$$+ k_6{}^{\text{cat}}\left[\text{P}'\text{ase}2\right][\text{pp-MEK}]/(K_6 m + [\text{pp-MEK}])$$
$$(14)$$

$$d[\text{pp-MEK}]/dt = k_4{}^{\text{cat}}\left[\text{Raf}^*\right][\text{p-MEK}]/(K_4 m + [\text{p-MEK}])$$
$$- k_6{}^{\text{cat}}\left[\text{P}'\text{ase}_2\right][\text{pp-MEK}]/(K_6 m + [\text{pp-MEK}])$$
$$(15)$$

$$d[\text{ERK}]/dt = k_9{}^{\text{cat}}\left[\text{P}'\text{ase}_3\right][\text{p-ERK}]/(K_9 m + [\text{p-ERK}])$$
$$- k_7{}^{\text{cat}}[\text{pp-MEK}][\text{ERK}]/(K_7 m + [\text{ERK}]) \qquad (16)$$

$$d[\text{p-ERK}]/dt = k_7{}^{\text{cat}}[\text{pp-MEK}][\text{ERK}]/(K_7 m + [\text{ERK}])$$
$$- k_8{}^{\text{cat}}[\text{pp-MEK}][\text{p-ERK}]/(K_8 m + [\text{p-ERK}])$$
$$- k_9{}^{\text{cat}}\left[\text{P}'\text{ase}3\right][\text{p-ERK}]/(K_9 m + [\text{p-ERK}])$$
$$+ k_{10}{}^{\text{cat}}\left[\text{P}'\text{ase}3\right][\text{pp-ERK}]/(K_{10} m + [\text{pp-ERK}])$$
$$(17)$$

$$\mathrm{d[pp\text{-}ERK]}/\mathrm{d}t = k_8{}^{\mathrm{cat}}[\mathrm{pp\text{-}MEK}][\mathrm{p\text{-}ERK}]/(K_8 m + [\mathrm{pp\text{-}ERK}])$$
$$- k_{10}{}^{\mathrm{cat}}\left[\mathrm{P'ase_3}\right][\mathrm{pp\text{-}ERK}]/(K_{10}m + [\mathrm{pp\text{-}ERK}])$$

$$(18)$$

Once QSS approximation was also applied in a model of Kholodenko [18], we adopted some of the rate constants and initial concentrations that were used in that model (Tables 1 and 2). Finally, we added a reaction representing the activation of Raf by Ras-GTP, with unknown rate constants. Remark that during a kinetic model simulation, the value of [Ras-GTP] is updated for each experimental data point; hence Ras-GTP levels actually compose the main input of the model whose kinetics is described by the system of ODEs in Eqs. 11–18.

3.3 Selection of Chemical Species for Time-Course Measurements

This procedure is critical in the modeling process, since it has impact on the outcome of the model-fitting optimization (Subheading 3.5). Thus, in order to select the chemical species whose time-course measurements would be more informative to adjust the rate constants of the model, we employ the mass conservation algebraic relations to replace some of the ODEs of the system, hence obtaining a system of differential-algebraic equations (DAEs). For instance, for the set of biochemical reactions depicted in Fig. 3, we have the following mass conservation equations (*see* **Note 5**):

$$[\mathrm{Raf}]_0 = [\mathrm{Raf}] + \left[\mathrm{Raf}^*\right] \tag{19}$$

$$[\mathrm{MEK}]_0 = [\mathrm{MEK}] + [\mathrm{p\text{-}MEK}] + [\mathrm{pp\text{-}MEK}] \tag{20}$$

$$[\mathrm{ERK}]_0 = [\mathrm{ERK}] + [\mathrm{p\text{-}ERK}] + [\mathrm{pp\text{-}ERK}] \tag{21}$$

where $[\mathrm{Raf}]_0$, $[\mathrm{MEK}]_0$, and $[\mathrm{ERK}]_0$ are the total concentration of Raf, MEK, and ERK proteins, respectively, along the whole biological experiment. Using the equations in Eqs. 19–21, we can make different removals of ODEs from the system of Eqs. 11–18: for instance, with Eq. 20, we can remove the differential equation that describes the kinetics of [MEK], or [p-MEK], or [pp-MEK] (*see* **Note 6**). Using this strategy, we do not need to measure chemical species whose values along time can be obtained through the algebraic relations, since those species can be derived from these algebraic relations. After verifying different possibilities of substitution (*see* **Note 7**), we obtain the following system of DAEs:

$$[\mathrm{Raf}] = [\mathrm{Raf}]_0 - \left[\mathrm{Raf}^*\right] \tag{22}$$

$$\mathrm{d}\left[\mathrm{Raf}^*\right]/\mathrm{d}t = k_1{}^{\mathrm{cat}}[\mathrm{Ras\text{-}GTP}][\mathrm{Raf}]/(K_1 m + [\mathrm{Raf}])$$
$$- k_2{}^{\mathrm{cat}}\left[\mathrm{P'ase_1}\right]\left[\mathrm{Raf}^*\right]/(K_2 m + \left[\mathrm{Raf}^*\right]) \tag{23}$$

Table 1
Rate constants of the two models of Ras/ERK pathway in Y1 cells that are presented in this paper

Rate constant	Value before optimization (bad and good fitting)	Adjusted value (bad fitting)	Adjusted value (good fitting)
$k_1{}^{cat}$	N/A	0.00446	0.585
$K_1 m$	N/A	3.99	1.19e−06
V_2	0.25	0.25	0.25
$K_2 m$	8	8	8
$k_3{}^{cat}$	0.025	0.025	0.025
$K_3 m$	15	15	15
$k_4{}^{cat}$	0.025	0.025	0.025
$K_4 m$	15	15	15
V_5	0.75	0.75	0.75
$K_5 m$	15	15	15
V_6	0.75	0.75	0.75
$K_6 m$	15	15	15
$k_7{}^{cat}$	0.025	0.025	0.025
$K_7 m$	15	15	15
$k_8{}^{cat}$	0.025	0.025	0.025
$K_8 m$	15	15	15
V_9	0.5	0.5	0.5
$K_9 m$	15	15	15
V_{10}	0.5	0.5	0.5
$K_{10} m$	15	15	15
$k_{11}{}^{cat}$	N/A	N/A	10.8
$K_{11} m$	N/A	N/A	3.45

In both models, the same initial values were used before the curve-fitting optimization; however, the final adjusted values of the rate constants are slightly different between models. All Michaelis constants units are in nM. The catalytic rate constants ($k_1{}^{cat}$, $k_3{}^{cat}$, $k_4{}^{cat}$, $k_7{}^{cat}$, $k_8{}^{cat}$, $k_{11}{}^{cat}$) and the maximal enzyme rates (V_2, V_5, V_6, V_9, V_{10}) units are in s^{-1} and nM/s, respectively

$$d[\text{MEK}]/dt = k_5{}^{cat}\left[\text{P}'\text{ase}_2\right][\text{p-MEK}]/(K_5 m + [\text{p-MEK}])$$
$$- k_3{}^{cat}\left[\text{Raf}^*\right][\text{MEK}]/(K_3 m + [\text{MEK}]) \qquad (24)$$
$$[\text{p-MEK}] = [\text{MEK}]_0 - [\text{MEK}] - [\text{pp-MEK}] \qquad (25)$$
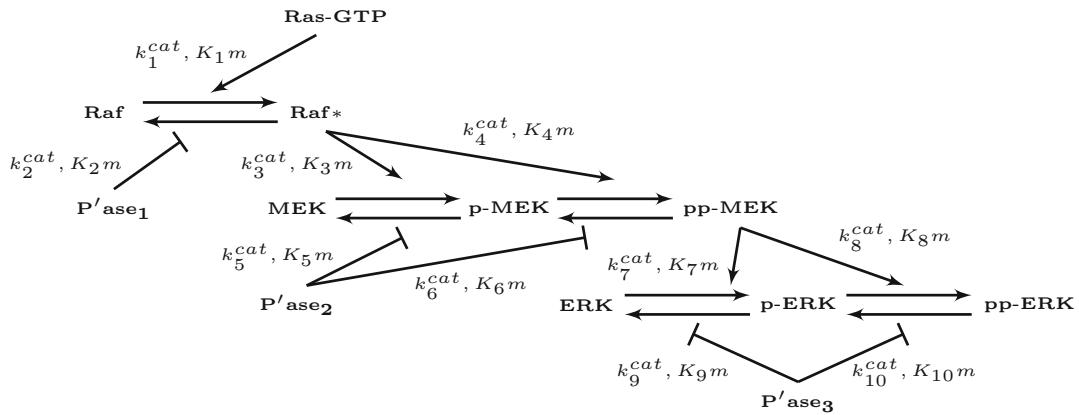
Fig. 3 Our initial hypothesis for the set of biochemical reactions that describe the signal flow through the Raf-MEK-ERK cascade in Y1 cells. Once we applied the QSS approximation to all reactions, each one is accompanied by its respective value of catalytic constants (k_i^{cat}) and of Michaelis constant ($K_i m$)

Table 2
Initial concentrations of the two models of Ras/ERK pathway in Y1 cells that were presented in this paper

Chemical species	Initial concentration (nM)
Ras-GTP	40
Raf	82
Raf*	18
MEK	272
p-MEK	20
pp-MEK	8
ERK	288
p-ERK	9
pp-ERK	3

In both models, the same initial values were used. All the initial concentrations are given in nM

$$\mathrm{d[pp\text{-}MEK]}/\mathrm{d}t = k_4{}^{\text{cat}}\big[\mathrm{Raf}^*\big][\mathrm{p\text{-}MEK}]/(K_4 m + [\mathrm{p\text{-}MEK}])$$
$$- k_6{}^{\text{cat}}\Big[\mathrm{P'ase_2}\Big][\mathrm{pp\text{-}MEK}]/(K_6 m + [\mathrm{pp\text{-}MEK}]) \tag{26}$$

$$\mathrm{d[ERK]}/\mathrm{d}t = k_9{}^{\text{cat}}\Big[\mathrm{P'ase_3}\Big][\mathrm{p\text{-}ERK}]/(K_9 m + [\mathrm{p\text{-}ERK}])$$
$$- k_7{}^{\text{cat}}[\mathrm{pp\text{-}MEK}][\mathrm{ERK}]/(K_7 m + [\mathrm{ERK}]) \tag{27}$$

$$[\text{p-ERK}] = [\text{ERK}]_0 - [\text{ERK}] - [\text{pp-ERK}] \tag{28}$$

$$\mathrm{d}[\text{pp-ERK}]/\mathrm{d}t = k_8{}^{\text{cat}}[\text{pp-MEK}][\text{p-ERK}]/(K_8 m + [\text{p-ERK}])$$
$$- k_{10}{}^{\text{cat}}\left[\text{P}'\text{ase}_3\right][\text{pp-ERK}]/(K_{10} m + [\text{pp-ERK}]). \tag{29}$$

Once both pp-MEK and pp-ERK still had their respective kinetics described by differential equations, we chose to measure phosphorylated ERK, since this protein is at the bottom of the kinase cascade and is also the "output" of this system.

3.4 Production and Normalization of Time-Course Measurements

As it was showed in the previous section, we decided to produce time-course measurements for Ras-GTP (as the main input of the system) and for phosphorylated ERK (as the dependent variable to be adjusted). For both chemical species, we carried out time-course assays and data posttreatment. Those assays were performed for time points within the first 30 min after the stimulation of starved Y1 cells by FGF2; we focused on this time frame because, in our study case, we are interested in modeling the biochemical events that rely on post-translational modifications only (e.g., phosphorylation), without the gene regulatory effects that are eventually caused by ERK activation.

3.4.1 Ras-GTP Enzyme-Linked Immunosorbent Assay (ELISA)

1. Plate cells at 50% confluence in 60 mm cell culture plates using Y1 culture medium and let to adhere overnight.

2. Wash cells twice with PBS and starve in serum-free medium for 48 h.

3. Stimulate starved cells with 10 ng/mL FGF2 for the indicated times.

4. For harvesting, discard culture medium, wash cells twice with 5 mL of 4 °C PBS, and apply 250 μL of RIPA buffer with protease inhibitor cocktail and place the plate on ice (*see* **Note 8**).

5. For cell lysis, scrape the cells from the entire plate, homogenize pipetting up and down, and transfer to Eppendorf tubes pre-chilled in ice (*see* **Note 8**).

6. Centrifuge cell lysates at 20,000 × *g* for 10 min at 4 °C to remove debris.

7. Probe Ras-GTP levels using G-LISA® Ras Activation Assay Biochem kit according to the manufacturer's protocol.

3.4.2 Ras-GTP Time-Course Data Normalization

1. Normalize Ras-GTP data replicates using the average signal strength of each replicate.

2. Once the Ras-GTP assay does not yield absolute quantification, we formulate a hypothesis assuming the peak level as full Ras activation (*see* **Note 9**). In Table 3, we summarize the Ras-GTP time-course experiment produced and normalized.

Table 3
Quantification values of Ras-GTP ELISA time-course experiments

	Raw value (avg)	Active Ras (%)	Concentration (nM)
Starved	0.0976	0.1215	0.45
0.5′	0.2214	0.2756	91.76
1′	0.1879	0.2338	77.85
3′	0.8036	1.0000	333.00
5′	0.5367	0.6679	222.41
15′	0.4237	0.5273	175.60
30′	0.1649	0.2052	68.34

We estimated the proportion of Ras-GTP to total Ras under the assumptions stated in Subheading 3.4, yielding these concentration values in nM

3.4.3 Phosphorylated ERK Western Blot Assay

1. Prepare cell lysates for Western blot assays exactly as described for Ras-GTP ELISA.

2. Remove debris and load 50 µg of each sample for Western blot experiments.

3. Carry out Western blots following the standard protocol [19]. Briefly, supernatant aliquots of proteins were resolved by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE; 10% acrylamide/bis-acrylamide) and transferred onto nitrocellulose membranes. Membranes were blocked for 1 h in TBS-T buffer containing 5% nonfat dried milk.

4. Wash the membrane three times of 5 min each with TBS-T on a rocker platform agitator.

5. Incubate overnight with primary antibodies at 1:2000 dilution in TBS-T buffer containing 5% BSA on a rocker platform agitator at 4 °C.

6. Wash the membrane three times of 10 min each with TBS-T on a rocker platform agitator.

7. Incubate with goat anti-rabbit peroxidase-linked secondary antibody at 1:10,000 in TBS-T buffer for 1 h on a rocker platform agitator.

8. Wash the membrane three times of 5 min each with TBS-T on a rocker platform agitator.

9. Remove the TBS-T and apply chemiluminescent substrate to membranes.

10. Produce the Western blot images (Fig. 4a) using a Western blot imaging system. We used Uvitec Alliance 9.7 equipment. Choose automatic exposure setting (*see* **Note 10**).
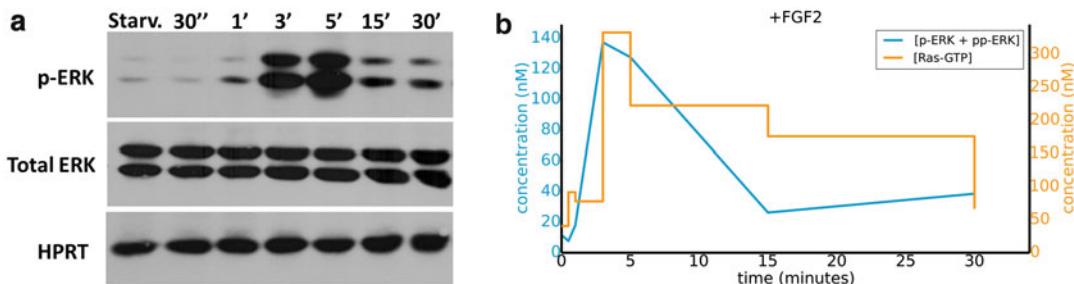
Fig. 4 (**a**) Time-course Western blot assays of total ERK and phosphorylated ERK, where starved Y1 cells were stimulated with FGF2 and time points were collected for up to half hour; HPRT was used as loading. (**b**) Quantifications for phosphorylated ERK in Western blot assays showed in (**a**) (*blue line*) and also for Ras-GTP in ELISA assays (*orange line*)

Table 4
Quantification values of ERK and phosphorylated ERK Western blot time-course experiments

	Raw value	Loading (HPRT)	Adjusted lane value	Active ERK (%)	Concentration (nM)
Starved	4,595,873	18,110,349	4,595,873	3.85	11.54
0.5′	4,588,951	27,392,596	3,033,940	2.54	7.62
1′	7,308,672	18,640,216	7,100,916	5.94	17.83
3′	67,940,866	22,494,710	54,698,762	45.79	137.36
5′	62,105,113	22,135,022	50,812,928	42.53	127.60
15′	17,370,964	30,109,232	10,448,431	8.75	26.24
30′	18,552,417	21,861,616	15,368,980	12.86	38.59

For each time point, the raw [phosphorylated ERK] value was adjusted by the loading protein quantification (HPRT), and its proportion in respect to [total ERK] was estimated using additional data from our lab. Finally, these values were given in nM

11. Background-subtract and quantify Western blot bands using the Uvitec Alliance 9.7 equipment software, with quantification values given in arbitrary units (*see* **Note 11**).

12. Perform data posttreatment through data normalization relative to each lane of the Western blot. For this purpose, use the housekeeping protein HPRT as a loading control for each lane.

3.4.4 Phosphorylated ERK Time-Course Data Normalization

1. Use the same method as Ras-GTP to normalize data replicates.

2. Use the same method as Ras-GTP to convert quantification values to concentrations, also using the hypothesis of a maximum activation. The final result of phosphorylated ERK time-course normalization is showed in Table 4.

3.5 Model-Fitting Optimization

In this procedure, the first step is to prepare the model according to the experimental data. Once the antibody used for ERK quantification targets both single- and double-phosphorylated states, we need a variable representing this specific target. Hence, we added an additional species to the model, whose value is the sum of [p-ERK] and [pp-ERK] (Fig. 4b, blue line). We will then use this variable when comparing with the experimental data. Additionally, once the phosphatases are considered constant in our model, we decided to use the maximum enzyme rate defined as:
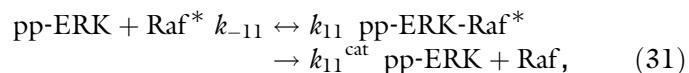
$$V = k^{\text{cat}}[\text{Phosphatase}] \qquad (30)$$

for all right-side terms in the system of DAEs which involve a phosphatase. By doing this, we create classes of equivalence among pairs of ⟨*catalytic rate*, [*Phosphatase*]⟩, which relieves the optimization process through the reduction of the number of variables to be fitted. This procedure is acceptable, since the phosphatases are treated in our kinetic model as constant inputs of the system.

Finally, last missing pieces of our model are the unknown rate constants for Raf activation by the Ras-GTP time-course input (Fig. 4b, orange line) and also for the negative feedback. We used model-fitting optimization to search for values of these unknown parameters for which the model could reproduce the experimental data. The basic method to perform optimizations consists in varying parameter values and selecting those for which the model is the closest to the experimental data. When no better parameter values can be found, then the best set of parameters is returned. Once we did not have any initial guess for these parameters, we decided to give loose bounds to the optimization for them, with values 1e−4 and 1e+8 for the catalytic constants (k_i^{cat}) and 1e−8 and 1e+4 for Michaelis constants ($K_i m$), respectively, for Raf activation and the negative feedback (*see* **Note 12**).

3.6 Test of Alternative Hypotheses

Once our experimental data showed a transient activation of the MAPK cascade, we could not reproduce the observed data without a mechanism to shut down ERK (Fig. 5). When this occurs during the modeling process, it might suggest us the model misses relevant components of the system. Through research on the literature and also on biological databases (*see* **Note 13**), we discovered that it was reported that a negative feedback from activated ERK is responsible to dephosphorylate and thus to deactivate Raf [20]. Therefore, we included a new reaction into the set of reactions of Eqs. 1–10 (Fig. 6), which describes a negative feedback:

$$\text{pp-ERK} + \text{Raf}^* \; k_{-11} \leftrightarrow k_{11} \; \text{pp-ERK-Raf}^* \\ \rightarrow k_{11}{}^{\text{cat}} \; \text{pp-ERK} + \text{Raf}, \qquad (31)$$
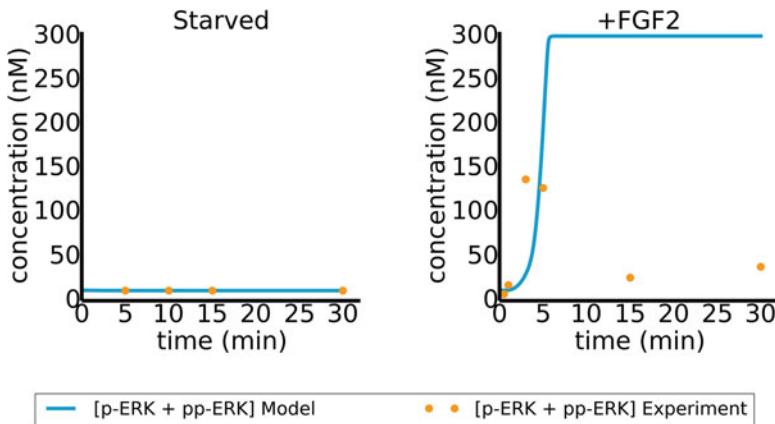
Fig. 5 Result of the model fitting for the first version of the model. *Left*, starved cells; *right*, cells stimulated with FGF2. While the fitting was able to adjust the model for starved cells, it lacks the ability to reproduce the transient response upon FGF2 stimulation
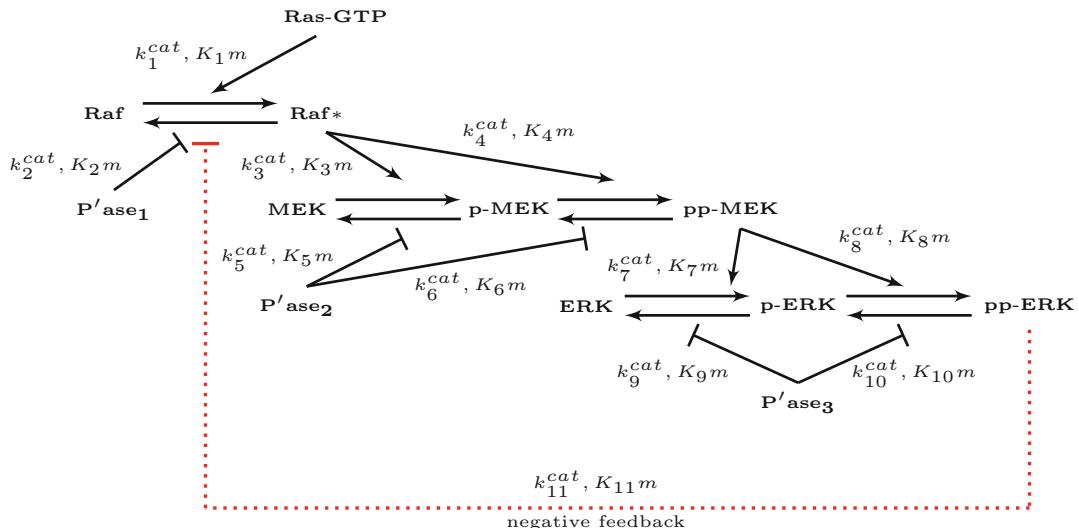
Fig. 6 A new hypothesis for the set of biochemical reactions that describe the signal flow through the Raf-MEK-ERK cascade in Y1 cells. A negative feedback reaction from pp-ERK to Raf*, with its respective catalytic constant and Michaelis constant, was included into the model

and repeated the procedures described in Subheadings 3.2 and 3.3, hence obtaining the following system of DAEs:

$$[\text{Raf}] = [\text{Raf}]_0 - [\text{Raf}^*] \qquad (32)$$

$$\begin{aligned} \mathrm{d}[\text{Raf}^*]/\mathrm{d}t = {}& k_1{}^{\text{cat}}[\text{Ras-GTP}][\text{Raf}]/(K_1 m + [\text{Raf}]) \\ & - k_2{}^{\text{cat}}\left[\text{P}'\text{ase}_1\right][\text{Raf}^*]/\left(K_2 m + [\text{Raf}^*]\right) \\ & - k_{11}{}^{\text{cat}}[\text{pp-ERK}][\text{Raf}^*]/\left(K_{11} m + [\text{Raf}^*]\right) \quad (33) \end{aligned}$$

$$\mathrm{d}[\mathrm{MEK}]/\mathrm{d}t = k_5{}^{\mathrm{cat}}\Big[\mathrm{P}'\mathrm{ase}_2\Big][\mathrm{p\text{-}MEK}]/(K_5 m + [\mathrm{p\text{-}MEK}])$$
$$- k_3{}^{\mathrm{cat}}\big[\mathrm{Raf}^*\big][\mathrm{MEK}]/(K_3 m + [\mathrm{MEK}]) \qquad (34)$$

$$[\mathrm{p\text{-}MEK}] = [\mathrm{MEK}]_0 - [\mathrm{MEK}] - [\mathrm{pp\text{-}MEK}] \qquad (35)$$

$$\mathrm{d}[\mathrm{pp\text{-}MEK}]/\mathrm{d}t = k_4{}^{\mathrm{cat}}\big[\mathrm{Raf}^*\big][\mathrm{p\text{-}MEK}]/(K_4 m + [\mathrm{p\text{-}MEK}])$$
$$- k_6{}^{\mathrm{cat}}\Big[\mathrm{P}'\mathrm{ase}_2\Big][\mathrm{pp\text{-}MEK}]/(K_6 m + [\mathrm{pp\text{-}MEK}])$$
$$(36)$$

$$\mathrm{d}[\mathrm{ERK}]/\mathrm{d}t = k_9{}^{\mathrm{cat}}\Big[\mathrm{P}'\mathrm{ase}_3\Big][\mathrm{p\text{-}ERK}]/(K_9 m + [\mathrm{p\text{-}ERK}])$$
$$- k_7{}^{\mathrm{cat}}[\mathrm{pp\text{-}MEK}][\mathrm{ERK}]/(K_7 m + [\mathrm{ERK}]) \qquad (37)$$

$$[\mathrm{p\text{-}ERK}] = [\mathrm{ERK}]_0 - [\mathrm{ERK}] - [\mathrm{pp\text{-}ERK}] \qquad (38)$$

$$\mathrm{d}[\mathrm{pp\text{-}ERK}]/\mathrm{d}t = k_8{}^{\mathrm{cat}}[\mathrm{pp\text{-}MEK}][\mathrm{p\text{-}ERK}]/(K_8 m + [\mathrm{p\text{-}ERK}])$$
$$- k_{10}{}^{\mathrm{cat}}\Big[\mathrm{P}'\mathrm{ase}_3\Big][\mathrm{pp\text{-}ERK}]/(K_{10} m + [\mathrm{pp\text{-}ERK}]).$$
$$(39)$$

Once just the ODE corresponding to Raf* kinetics (Eq. 33) changed when compared with the system of DAEs of Eqs. 22–29, both pp-MEK and pp-ERK remained in the ODEs, which implies that new time-course measurements (Subheading 3.4) were not necessary. To carry out the curve-fitting optimization (Subheading 3.5), the rate constants introduced into the system by this feedback ($k_{11}{}^{\mathrm{cat}}$ and $K_{11} m$) were also considered unknown. After a new round of fitting, we could produce a model whose simulation has good agreement with experimental data (Fig. 7); the adjusted rate
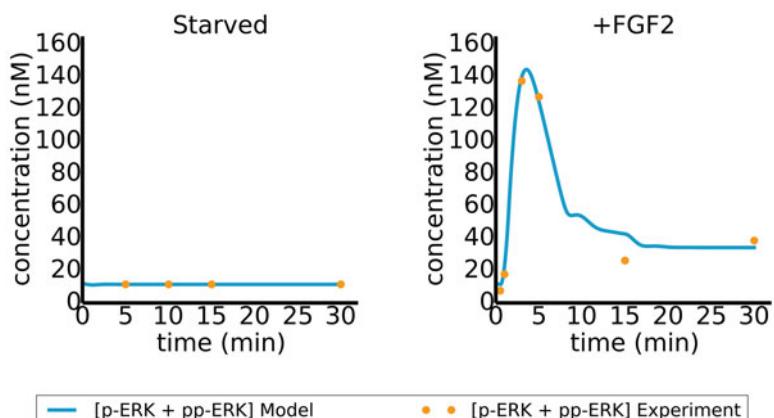
Fig. 7 Result of the model fitting for the second version of the model, including pp-ERK negative feedback on Raf*. *Left*, starved cells; *right*, cells stimulated with FGF2. While still being able to adjust the model for the starved condition, the addition of the negative feedback enables the model to also reproduce the transient response

constants and initial conditions that were used to execute this simulation are showed in Tables 1 and 2.

4   Notes

1. This "bottom-up" approach is preferable than to start from a more comprehensive set of chemical species, since the size of the system identification problem is proportional to the size of such set, which has implications both in the model-fitting optimization problem (making the search space too large) and in the estimation problem (lack of data to estimate a large number of rate constants).

2. For modeling purposes, we must decide whether different protein isoforms should be considered or not. This decision is critical and must be made together with the wet lab team, since measurements for the time-course experiments might not have specificity for distinguishing each isoform. In our example, antibodies for phosphorylated ERK do not distinguish between the isoforms 1 and 2 of this kinase.

3. Although Ras is not a kinase, its interaction with Raf is necessary to allow its activation through phosphorylation [21]. However, for the sake of model simplicity, we assume that (1) [Ras-GTP] levels are proportional to the Raf activation and (2) recruitment of Raf to the membrane by active Ras is a very fast reaction; hence it can be approximated as an instantaneous event.

4. We can adopt systems of ODEs to describe the kinetics of the signaling pathway if two conditions hold: (1) the concentration levels of the involved chemical species are high enough (i.e., we can describe a chemical species as a continuous variable) and (2) there is low uncertainty due to noise (i.e., we can consider the kinetics of these species a deterministic phenomenon). If (1) does not hold, one should consider the usage of a discrete model (e.g., a Boolean model). If (2) does not hold, stochastic approaches should be employed instead.

5. The mass conservation algebraic equations of a given set of biochemical reactions might be more complicated to be extracted than the ones of our study case, depending on the size of the set and also how intertwined are these reactions. Moreover, once there are different possible sets of mass conservation algebraic equations, it is useful to apply computational techniques to obtain the largest possible set of equations that are linearly independent (i.e., nonredundant equations) [22].

6. We must choose a system of DAEs which minimizes the number of ODEs while keeping in ODEs all the chemical species whose time-course experiments are feasible. We call such

system of DAEs as optimal. In our case, an optimal system of DAEs always contains [p-MEK] and [p-ERK], both with feasible time-course experiments.

7. For small models like the one of our study case, an exhaustive search for an optimal system of DAEs is feasible. However, the search space grows exponentially as a function of the model size. Therefore, for larger models the optimal system of DAEs should be searched through a procedure such as the classical branch-and-bound algorithm [23].

8. For all cell lysates, keep the samples in the ice all the time and make the manipulation as fast as possible, especially for Ras-GTP assays, given the intrinsic hydrolytic activity of Ras.

9. Once Ras-GTP is only present in one enzymatic reaction in our model and with unknown catalytic constant, we can choose an arbitrary value for this peak of activation. The value of this unknown parameter will be adjusted by the model-fitting optimization procedure (Subheading 3.5), thus making only the pattern of activation important. Hence, we decided to choose the hypothesis of a full Ras activation, enabling us to estimate the concentration for this data point and then to extrapolate the concentration of the remaining data points of our experiment.

10. Automatic exposition should be used for acquiring Western blot images to avoid overexposure of signals.

11. We made alternative quantifications for all Western blot experiments using the Image J software [24]; however, despite differences on the absolute values, no significant difference on the relative values was observed.

12. An important pitfall about optimization, especially when you have no initial guess about some rate constants, is to find a stable point for the optimization start from. Especially in cases where we are dealing with Michaelis-Menten quasi-stationary (QS) or QSS approximations, some ranges of parameters provide unstable values, which can prevent the optimizer to work. Therefore, choosing the bounds for the parameter values is also a delicate exercise, especially working with enzymatic kinetics whose rate constants can vary a lot. Therefore, the optimization process will be increasingly more complicated with very loose bounds. A way to deal with this trade-off is to execute several optimizations, with increasingly loose bounds, until finding a satisfactory result.

13. For small models like our study case, we can test some hypotheses for the network topology in a manual fashion. However, for larger models the space of hypotheses grows exponentially on the model size. One way to tackle this problem is to carry out the steps of Subheadings 3.1–3.3 and 3.5–3.6 in an automated fashion, using a first iteration of the methodology to

produce a set of time-course data (through the procedure of Subheading 3.4). Within the SigNetSim framework, we implemented such automation using interactome databases such as KEGG to generate a list of chemical species and reactions that are candidates to be included into the kinetic model and a greedy search algorithm to select and assess them in a combinatorial way [25].

Acknowledgment

References

1. Seger R, Krebs EG (1995) The MAPK signaling cascade. FASEB J 9(9):726–735

2. Chen WW, Niepel M, Sorger PK (2010) Classic and contemporary approaches to modeling biochemical reactions. Genes Dev 24(17):1861–1875

3. Yasumura Y, Buonassisi V, Sato G (1966) Clonal analysis of differentiated function in animal cell cultures. I. Possible correlated maintenance of differentiated function and the diploid karyotype. Cancer Res 26:529–535

4. Milo R, Jorgensen P, Moran U, Weber G, Springer M (2010) BioNumbers – the database of key numbers in molecular and cell biology. Nucleic Acids Res 38(1):D750–D753

5. Kanehisa M, Susumu G (2000) KEGG: Kyoto encyclopedia of genes and genomes. Nucleic Acids Res 28(1):27–30

6. Le Novère N, Bornstein N, Broicher A et al (2006) BioModels database: a free, centralized database of curated, published, quantitative kinetic models of biochemical and cellular systems. Nucleic Acids Res 34(1):D689–D691

7. Noël V (2016) SigNetSim (Signaling Network Simulator), an e-Science framework to assist the design of kinetic models of biochemical reactions. http://cetics.butantan.gov.br/signetsim. Accessed 22 Jan 2016

8. Chu KW, Yuefan D, Reinitz J (1999) Parallel simulated annealing by mixing of states. J Comput Phys 148(2):646–662

9. Pariona-Llanos R, Pavani RS, Reis MS et al (2015) Glyceraldehyde 3-phosphate dehydrogenase-telomere association correlates with redox status in *Trypanosoma cruzi*. PLoS One 10(3):e0120896

10. Forti FL, Schwindt TT, Moraes MS et al (2002) ACTH promotion of p27Kip1 induction in mouse Y1 adrenocortical tumor cells is dependent on both PKA activation and Akt/PKB inactivation. Biochemistry 41 (31):10133–10140

11. Costa ET, Forti FL, Matos TGF et al (2008) Fibroblast growth factor 2 restrains Ras-driven proliferation of malignant cells by triggering RhoA-mediated senescence. Cancer Res 68 (15):6215–6223

12. Salotti J, Dias MH, Koga MM et al (2013) Fibroblast growth factor 2 causes G2/M cell cycle arrest in Ras-driven tumor cells through a Src-dependent pathway. PLoS One 8(8): e72582

13. Reis MS (2016) Supplementary material of "An interdisciplinary approach for designing kinetic models of the Ras/MAPK signaling pathway". http://cetics.butantan.gov.br/papers/MiMB. Accessed 27 Jan 2016

14. Shankaran H, Ippolito DL, Chrisler WB et al (2009) Rapid and sustained nuclear—cytoplasmic ERK oscillations induced by epidermal growth factor. Mol Syst Biol 5(1):332

15. Vayttaden SJ, Ajay SM, Bhalla US (2004) A spectrum of models of signaling pathways. Chembiochem 5(10):1365–1374

16. Huang CY, Ferrell JE (1996) Ultrasensitivity in the mitogen-activated protein kinase cascade. Proc Natl Acad Sci U S A 93(19):10078–10083

17. Briggs GE, Haldane JBS (1925) A note on the kinetics of enzyme action. Biochem J 19 (2):338–339

18. Kholodenko BN (2000) Negative feedback and ultrasensitivity can bring about oscillations in the mitogen-activated protein kinase cascades. Eur J Biochem 267(6):1583–1588

19. Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning. Cold Spring Harbor Laboratory Press, New York

20. Dougherty MK, Müller J, Ritt DA et al (2005) Regulation of Raf-1 by direct feedback phosphorylation. Mol Cell 17(2):215–224

21. Marais R, Light Y, Paterson HF et al (1995) Ras recruits Raf-1 to the plasma membrane for activation by tyrosine phosphorylation. EMBO J 14(13):3136–3145

22. Guimarães AM (2016) An algorithm to simplify systems of differential equations that describe the kinetics of chemical reactions. Dissertation, University of São Paulo, Brazil (in Portuguese)

23. Narendra PM, Fukunaga K (1977) A branch and bound algorithm for feature subset selection. IEEE Trans Comput 100(9):917–922

24. Rasband WS (2015) ImageJ. US National Institutes of Health, Bethesda, MD, USA

25. Wu L (2015) A method to modify molecular signaling networks through examination of interactome databases. Dissertation, University of São Paulo, Brazil (in Portuguese)