

A Mass Action Model of a Fibroblast Growth Factor Signaling Pathway and Its Simplification

E.A. Gaffney^{a,*}, J.K. Heath^b, M.Z. Kwiatkowska^c

^a *Centre for Mathematical Biology, Mathematical Institute, University of Oxford, 24-29 St Giles', Oxford OX1 3LB, UK*

^b *The School of Biosciences, The University of Birmingham, Edgbaston, Birmingham B15 2TT, UK*

^c *Oxford University Computing Laboratory, Wolfson Building, Parks Road, Oxford OX1 3QD, UK*

Received: 11 June 2007 / Accepted: 23 May 2008 / Published online: 9 October 2008
© Society for Mathematical Biology 2008

Abstract We consider a kinetic law of mass action model for Fibroblast Growth Factor (FGF) signaling, focusing on the induction of the RAS-MAP kinase pathway via GRB2 binding. Our biologically simple model suffers a combinatorial explosion in the number of differential equations required to simulate the system. In addition to numerically solving the full model, we show that it can be accurately simplified. This requires combining matched asymptotics, the quasi-steady state hypothesis, and the fact subsets of the equations decouple asymptotically. Both the full and simplified models reproduce the qualitative dynamics observed experimentally and in previous stochastic models. The simplified model also elucidates both the qualitative features of GRB2 binding and the complex relationship between SHP2 levels, the rate SHP2 induces dephosphorylation and levels of bound GRB2. In addition to providing insight into the important and redundant features of FGF signaling, such work further highlights the usefulness of numerous simplification techniques in the study of mass action models of signal transduction, as also illustrated recently by Borisov and co-workers (Borisov et al. in *Biophys. J.* 89, 951–966, 2005, *Biosystems* 83, 152–166, 2006; Kiyatkin et al. in *J. Biol. Chem.* 281, 19925–19938, 2006). These developments will facilitate the construction of tractable models of FGF signaling, incorporating further biological realism, such as spatial effects or realistic binding stoichiometries, despite a more severe combinatorial explosion associated with the latter.

Keywords Signal transduction · FGF · Mathematical modeling · Law of mass action

1. Introduction

Numerous signaling molecules, in the form of secreted hormones or growth factors, exert their influence via receptor tyrosine kinases. Two key examples are Epidermal

*Corresponding author.

E-mail address: gaffney@maths.ox.ac.uk (E.A. Gaffney).

Growth Factor (EGF) and Fibroblast growth factor (FGF). EGF stimulates the proliferation of numerous cell types, and thus plays a key role in tissue homeostasis, as does FGF which also regulates numerous developmental processes. Their importance is highlighted by the diverse range of pathologies associated, at least in part, with their aberrant signaling (Basilico, 2005; Wilkie, 2005; Grose and Dickson, 2005; Eswarakumar et al., 2005; Dailey et al., 2005; Adachi et al., 2004; Augoff et al., 2005; Krueger and Ellis, 2005), including numerous cancers (e.g. Grose and Dickson, 2005; Eswarakumar et al., 2005; Dailey et al., 2005; Yarden and Sliwkowski, 2001; Greenman et al., 2007; Julien et al., 2007). The motivation for the detailed investigation of these signaling proteins is not only heightened by their association with malignant disease, but also because cell signaling pathways can now be targeted in novel drug discovery programs, as showcased with EGF signaling (Sevecka and MacBeath, 2006). In addition, the possibilities of an increased understanding of signaling pathways are highlighted by recent observations that down regulating EGF signaling by inhibiting a tyrosine phosphatase has therapeutic benefits in a mouse model of breast cancer (Julien et al., 2007). However, the mechanisms by which signaling molecules exert their influence on a cell, and how these can be manipulated and controlled, are not intuitive due to multiple feedback processes. Consequently, mathematical and computational modeling have a role in unraveling and simplifying the complexity of these signaling pathways (Citri and Yarden, 2006; Bublil and Yarden, 2007; Yamada et al., 2004; Kiyatkin et al., 2006; Schoeberl et al., 2002); here we focus on FGF.

FGF binds to its receptor which, in complete generality, is one of four receptor tyrosine kinases, FGFR1-FGFR4, with both ligand and receptor classified into different isoforms with varying specificities for each other. The known downstream events arising from the binding of FGF and its receptor is reviewed in Eswarakumar et al. (2005). Here, we will focus on the initiation of the RAS-MAP kinase pathway; this proceeds via the binding of the docking protein FRS2 to activated FGFR. The FRS2 subsequently phosphorylates with the recruitment of multiple binding molecules, including GRB2, the binding of which initiates the desired pathway via subsequent SOS binding. However, other molecules binding FRS2 act to downregulate the FGF signal by a multitude of mechanisms, such as ubiquitination, entailing that the transduced FGF signal is transient.

Signal duration and amplitude are considered to be important features of the response of a cellular signaling pathway (Dailey et al., 2005; Schlessinger, 2000) and yet are particularly hard to grasp intuitively in a system which is as complex and with as many feedbacks as FGF signaling. This illustrates the need for constructing models capable of quantitatively predicting the signal produced by FGFR activation in terms of our current understanding and, given the architecture of FGF signaling is not yet fully defined (Li et al., 2004), plausible hypotheses where current understanding is incomplete. Here, we will be particularly interested in the extent to which knockouts of various binding molecules, or the introduction of an FGFR kinase inhibitor such as SU5402, alter the system dynamics; this will also indicate important and redundant features of the signal transduction process. We will additionally illustrate predictions for the complex behaviors of the dynamics associated with the binding molecule SHP2. In turn, developing predictions for how our model system will behave in the presence of perturbations may ultimately assist in prioritizing which experiments can be performed to critically test our hypotheses and increase our fundamental understanding. Such programs are already developing for EGF

signaling (e.g. Citri and Yarden, 2006; Bublil and Yarden, 2007; Kiyatkin et al., 2006; Schoeberl et al., 2002).

There are a number of possible modeling frameworks one can adopt. In particular, there is the choice between a stochastic framework and a deterministic framework. Both have their respective advantages and disadvantages. For example, stochastic simulations are particularly apt for scenarios with low molecular numbers when the thermodynamic limit implicit in any deterministic framework breaks down. In contrast, deterministic approaches are based upon the law (or rather assumption) of mass action (Kudryavtsev et al., 2001; Billingham and King, 2000) and can readily predict averaged quantities for large molecular numbers. In addition there is a large repertoire of mathematical tools already in existence which frequently enable the simplification of deterministic models such as dynamical systems theory (Tyson et al., 2001) and Tikhonov asymptotics (Tikhonov, 1952), as reviewed by Crampin et al. (2004). A critical evaluation of the choice between the two approaches has been discussed by numerous authors, e.g. Aldridge et al. (2006), Wolkenhauer et al. (2004), and invariably the conclusions as to which is the most appropriate is dependent on the ultimate aims of the modeling study.

Here, one of our aims will be to compare the results of the standard deterministic framework, based on the law of mass action, with previous stochastic simulations for FGF signaling to see if major discrepancies emerge for a system possessing a level of complexity moving toward that found in-situ. In particular, this additionally presents us with the difficulty of combinatorial complexity which has already been noted as a severe problem (Borisov et al., 2005; Aldridge et al., 2006). Thus, another aim will be to utilize mathematically accurate approximations to greatly simplify the simulations, enabling a more in-depth investigation, and also allowing the future development of models with more biological realism despite combinatorial complexity. Hence, the model is not intended to be a final, definitive representation of FGF signaling. While, mathematically and computationally, it is very complex, it nonetheless neglects features of this signal transduction process, such as the existence of different isoforms of ligand and receptor or the possibility of cross-talk (Dailey et al., 2005). It also presents a simplified stoichiometry so that the numerical simulations of the full model are still tractable, thus allowing an explicit testing of the simplification procedure.

We proceed by initially constructing a model of FGF signaling and then proceed to demonstrate how it can be systematically and accurately simplified. We subsequently compare the results of the simplified and full versions of the model with each other and also with results from stochastic simulations. This is followed by a discussion of the insight gained from our simulations, plus the observations of a sensitive dependence on the dynamics of the binding molecule SHP2 and the relative insensitivity to alterations in the FGFR binding rate. Finally, we conclude with a discussion of future applications of the model and its simplifications.

2. Methods

2.1. Model formulation

In Fig. 1, we have a reaction overview. FGF and its receptor, FGFR, can bind and once bound they can undergo phosphorylation at sites 653 and 654 on FGFR. When both these sites are phosphorylated, other FGFR sites can phosphorylate, particularly site 766.

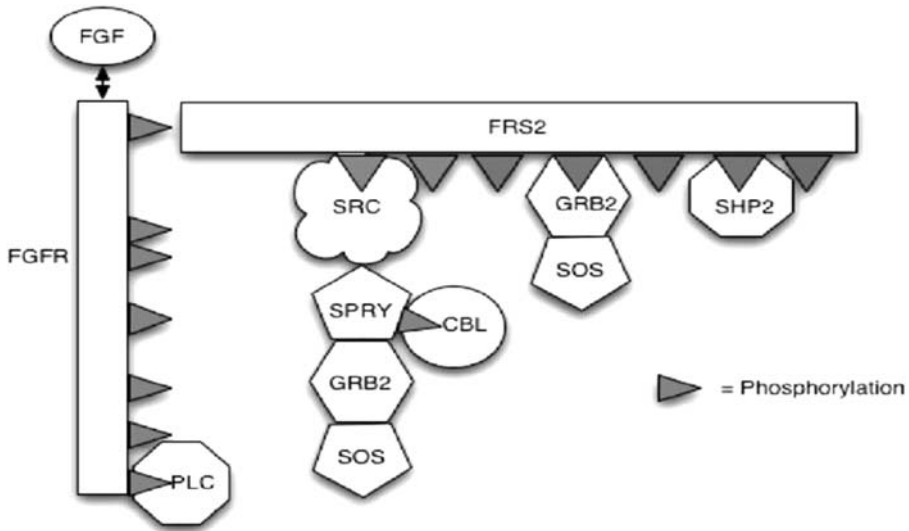


Fig. 1 A summary of the signal transduction model used.

To simplify the model, we assume the FGF:FGFR complex binds to FRS2 in a very fast reaction once phosphorylation at sites 653 and sites 654 has occurred. More general models have allowed FRS2 to bind not only to the complex FGF:FGFR653P654P (where “P” stands for phosphorylated) but also to singly phosphorylated and unphosphorylated FGF:FGFR (Heath et al., 2006; Kwiatkowska et al., 2006). Our motivation for utilizing such an approximation is that it reduces the number of molecular states of the FGF:FGFR:FRS2 complex by one-third. As we will see below, the number of molecular states in the model is becoming prohibitive for a mass action model; increasing the number of states by a factor of three would make simulating the full model impractical. A justification for the validity of such an assumption is given in Section 2.3.2 and Appendix C.

The 766 phosphorylated site on FGF:FGFR can bind PLC, whether or not FGF:FGFR is bound to FRS2.

Numerous sites of the FRS2 complex can phosphorylate once FGF:FGFR653P654P:FRS2 has formed. These phosphorylated sites in turn can bind SRC, GRB2, and SHP2 as indicated in Fig. 1.

Further bindings can occur to augment the bound SRC, GRB2, and SHP2 as indicated in Fig. 1. In addition, SRC binding induces a relocation of the SRC:FRS2 complex, while bound PLC can induce a decay in FGFR. Furthermore, CBL binding leads to a ubiquitination and degradation of FRS2, while the binding of SHP2 to FRS2 leads to the dephosphorylation of bound phosphorylated SPRY and a dephosphorylation of FRS2 binding sites.

In addition, all possible reactions are assumed to proceed in parallel, at least once any constraints are satisfied; one example of such a constraint is the need for sites 653 and 654 to be phosphorylated on activated FGFR before FGFR–FRS2 binding can occur, as considered above. As an illustration of this parallelism, we take it that SPRY can bind to

SRC, and GRB2 can bind to a SRC:SPRY complex irrespective of whether SRC is bound to FRS2. Similarly for all other possibilities.

To simplify the model initially, SOS and RAF are assumed to be absent. This does not alter model predictions as the SOS and RAF bindings do not feedback into any upstream dynamics of the system.

2.1.1. Use of Michaelis–Menten kinetics

It is assumed that the phosphorylation and dephosphorylation reactions are driven by Michaelis–Menten kinetics. These can be understood as the large time asymptotic approximation of an enzyme-catalyzed biochemical reaction, as considered for simple reactions by Murray (1989). There has been recent discussion, for example Millat et al. (2007), about the validity of Michaelis–Menten kinetics in the context of biochemical networks and pathways, emphasizing that such approximations need not be valid when the system is sensitive to the initial, transient, dynamics. This is observed to be especially important within feedback loops. However, such sensitivity is not exhibited by the system behavior here, for the timescales of the biochemical reactions considered, as further discussed in the Appendices.

We note that the context of Michaelis–Menten kinetics required for our model is more complicated than textbook scenarios. We additionally assume, for example, that the phosphorylation of the PLC binding site is such that all molecules with an unphosphorylated PLC binding site are in competition for the same enzyme. Analogous assumptions are made for the other phosphorylation and dephosphorylation reactions. This is sufficient to apply the same reasoning that underlies the standard Michaelis–Menten kinetics, though the resulting equations are more complex, as detailed in Appendix B. In particular, the use of Michaelis–Menten approximations in our model results in an extensive coupling of large subsets of reactants; for example, a single term in the Ordinary Differential Equations (ODEs) describing the system can depend on over 100 different molecular species. This, in turn, makes direct numerical simulation substantially slower compared to reaction schemes using first order, linear, reaction kinetics for phosphorylation and analogous reactions, as considered by Yamada et al. (2004) for example.

2.1.2. Model motivation

A complete reaction scheme, defining the model, plus reference parameter values are given in Appendix A. The basics of the model, such as the binding of FGF, FGFR, and FRS2, together with phosphorylations and the recruitment of PLC to FGFR and the recruitment of GRB2, SHP2 and SRC to FRS2, are commonly noted, as in Eswarakumar et al. (2005), Dailey et al. (2005), Thisse and Thisse (2005), Li et al. (2004). Also commonly noted is the binding of SOS to FRS2 via GRB2 and the subsequent activation of the RAS-MAP kinase pathway (*ibid*) together with CBL induced degradation (Dailey et al., 2005). It is also known that SHP2 can induce a dephosphorylation of SPRY (Jarvis et al., 2006; Hanafusa et al., 2004) and that SPRY associates with GRB2 (Thisse and Thisse, 2005) plus CBL (Wong et al., 2001). Signal attenuation due to SRC binding FRS2 has also been reported, along with the observation that SPRY is a direct physiological substrate for SRC (Li et al., 2004). One should note that although SPRY antagonism of FGF signaling is commonly accepted, the dominant mechanism by which this is realized is unclear and has led to controversy (Thisse and Thisse, 2005; Li et al., 2004).

The molecular details of the action of SHP2 are not clearly defined though recent research does indicate that SHP2 exerts its influence on receptor tyrosine kinase signaling

via its PTPase activity, and thus via inducing dephosphorylation (Agazie and Hayman, 2003). Finally, a slow production of SPRY is incorporated given that its expression is reported to be induced by the FGF pathway (Kramer et al., 1999). Thus, the model is essentially based upon literature reported processes.

Additional details have to be hypothesized, such as all reactions occur in parallel and the validity of both the law of mass action and the use of Michaelis–Menten kinetics. The presence of a slow PLC-induced degradation constitutes the reasonable hypothesis that there is a weak negative feedback control of the PLC γ signaling pathway (as observed in EGF signaling (Chen et al., 1996)). Further hypotheses typically have to be made concerning whether binding molecules are made available following such a degradation; similarly for other degradation events.

2.1.3. *Parameter values*

As discussed in detail by Wolkenhauer et al. (2004), there is a paucity of quantitative stimulus-response time course data sets for realistic pathway models, in contrast to the bioinformatics arena, which is characterized by extensive data. Thus, parameter estimates are not common in the literature and can be subject to large variation. For example, there are three orders of magnitude in variation (Mohammadi et al., 2005) among binding rates for FGF and FGFR according to the isoform of FGF and FGFR. Numerous further complications arise; for example, cytokine membrane reaction rates not need concur with solution reaction rates (Gavutis et al., 2006). Thus, we resort to using values reflecting our current understanding of the system. Modeling can nonetheless be extremely useful in the face of such difficulties. It can still address the above-mentioned questions such as whether current understanding is consistent with observation, what are the important and the redundant features of the system according to current understanding and, ultimately, how can we efficiently test current hypotheses.

2.1.4. *Summary*

The model, therefore, consists of the reactions required to build up the molecular state depicted in Fig. 1, plus the extra details described above, i.e. ubiquitination, PLC induced decay, SRC induced relocation and SHP2 induced dephosphorylation. All phosphorylation and dephosphorylation reactions are assumed to be of Michaelis–Menten form. A detailed summary of these reactions is given in Appendix A. For the model considered, we have 384 non-linear ODEs, with as many as 568 terms within an individual ODE. In addition, the use of Michaelis–Menten kinetics entails that a single term representing the phosphorylation dynamics can depend on over 100 different molecular species.¹

2.2. *Difficulties with the full model. Combinatorial complexity. Numerical methods*

The key difficulty with the full model is the fact the number of equations combinatorially explode as the number of molecular species and interactions increase. This makes an intuitive understanding of the system problematic, especially in the presence of numerous parameters which are not easily estimated. It also makes a detailed numerical study difficult. Indeed doubling the number of GRB2 and SHP2 binding sites, which is consistent with the structural biology of FRS2 (Eswarakumar et al., 2005), would make a full mass

¹ A printout of the equations requires over 165 pages, font size 8.

action model essentially intractable. A further problem arises from the fact the system is, in the context of numerical analysis, stiff. However, numerical algorithms for the *efficient* solution of stiff ordinary differential equations are not designed for large systems as the linear algebra involved at each timestep of a stiff solver will be prohibitively slow once the system is sufficiently large. Consequently, the numerical solver used below is an explicit Runge–Kutta routine rather than a stiff solver.

In the following section, we simplify the model by decoupling subsets of the reacting molecules wherever it is legitimate to do so, including scenarios where this cannot be done exactly, but nonetheless to a high level of accuracy.

2.3. Model simplification

Throughout this section, and the rest of the paper, including the Appendices, our notation does not distinguish between a reactant, for example, FGF, and its concentration. Noting that all equations are in terms of concentrations, it is clear from context whether we are referring to the molecule or its concentration.

2.3.1. Utilizing the different timescales inherent in the model

The model possesses reactions that occur at different timescales. The initial dynamics, related to the production of FGF:FGFR653P654P:FRS2 is taken to occur on what we shall refer to as the ultra-fast timescale. Other reactions, such as the binding of PLC, SRC, GRB2, and SHP2 to sites on the FGFR molecule, or on the FRS2 molecule, occur on a fast timescale. Signal attenuation events, such as PLC induced decays, are taken to occur on a relatively slow timescale. If we can utilize these different timescales to decouple the events occurring on each timescale, we can expect a combinatorial “implosion,” whereby the complexity of the model decreases immensely.

Consider a system of fast variables \mathbf{f} and slow variables \mathbf{s} and the associated ODEs, that is

$$\epsilon \frac{d\mathbf{f}}{dt} = \mathbf{F}(\mathbf{f}, \mathbf{s}), \quad \frac{d\mathbf{s}}{dt} = \mathbf{S}(\mathbf{f}, \mathbf{s}), \quad (1)$$

with $\epsilon \ll 1$ and the components of the vector functions \mathbf{F}, \mathbf{S} are of order unity. The solution to these equations can exhibit a regular behavior such that, asymptotically in ϵ , there is

- I a fast, transient evolution, with \mathbf{s} constant, to the manifold where $\mathbf{F}(\mathbf{f}, \mathbf{s}) = \mathbf{0}$,
- II a subsequent slow evolution governed, asymptotically in ϵ , by $ds/dt = \mathbf{S}(\mathbf{f}, \mathbf{s})$ with \mathbf{f} quasi-stationary, i.e. \mathbf{f} is such that $\mathbf{F}(\mathbf{f}, \mathbf{s}) = \mathbf{0}$.

Technical details aside, this behavior simply requires that the initial condition is within the attracting basin of $\mathbf{F}(\mathbf{f}, \mathbf{s}) = \mathbf{0}$ for the initial value of \mathbf{s} , and the trajectory in II remains within the attracting basin of the manifold $\mathbf{F}(\mathbf{f}, \mathbf{s}) = \mathbf{0}$. We note that a rigorous framework for these observations has been developed by Tikhonov (1952) and Suckley and Biktashev (2003).

Thus, the ODEs governing the fast variables and the slow variables effectively decouple. The above ideas allow us to decouple the ultra-fast, fast, and slow dynamics.² We

²Complicated behavior beyond the scope of the analytical techniques described here does occur if the slow evolution equations drive the solution out of an attracting basin of $\mathbf{F}(\mathbf{f}, \mathbf{s}) = \mathbf{0}$. This is the defining behavior

proceed to decouple the ultra-fast dynamics, before considering the remaining equations in Section 2.3.3.

2.3.2. *Decoupling the initial, ultra-fast, dynamics*

The detailed analysis is presented in Appendix C. Matched asymptotics and a justified use of the quasi-steady state approximation demonstrate that the model effectively decouples from the early FGF, FGFR, and FRS2 binding and phosphorylation events. The system is ultimately driven only by the production of FGF:FGFR653P654P:FRS2. The amount produced in turn is at an excellent level of approximation, equal to the minimum of the initial concentrations of FGF, FGFR, and FRS2. Thus, the initial dynamics is effectively very simple and very robust. This also justifies the simplifications, described at the start of Section 2.1, as discussed further in Appendix C.

2.3.3. *Decoupling the remaining fast and slow dynamics*

Consider the above model, in the unbiological scenario, where we have turned off the decay of FGFR due to PLC binding, the relocation of FRS2 due to SRC, ubiquitination due to CBL binding and the dephosphorylation induced by SHP2 binding. While the number of states are unchanged, the dynamics of the chains of molecules binding at various sites of the FGF:FGFR653P654P:FRS2 complex completely decouples; consequently, there is an enormous simplification. While this level of simplification is not present, it is clear that the limited interaction between different molecular chains emerging from the different binding sites of the FGF:FGFR653P654P:FRS2 complex should nonetheless enable dramatic simplification. This is especially true when coupled with a Tikhonov procedure.

We remark that extensive decoupling of linear combinations of the underlying equations is to be expected in models of signal transduction with large docking proteins and has been analyzed in detail by Borisov and co-workers (Borisov et al., 2005, 2006; Kiyatkin et al., 2006). Our work extensively utilizes such simplifications, combining them with multiple temporal scales and simplifications which accommodate the interaction between the molecular chains emerging from the different binding sites of the FGF:FGFR653P654P:FRS2 complex.

Below, we initially set up our simplified model in the absence of PLC, before proceeding to incorporate the additional difficulties incurred by the presence of this binding molecule.

2.3.4. *The fast dynamics (no PLC present)*

Let FRS2* denote FGF:FGFR653P654P:FRS2 below. A key feature coupling the various molecular chains bound to FRS2* is the fact that bound SHP2 induces the dephosphorylation of the other phosphorylation sites of the FRS2* molecule, plus the dephosphorylation of any SPRY bound via SRC in the model reaction scheme. This all occurs on the fast time scale. Thus, we proceed as follows.

At each time step of the *slow* dynamics, we solve the quasi-steady state of the subsystem depicted in Fig. 2 considering only the fast dynamics. This reaction scheme can be broken down further; for example, one can decouple the SHP2 dynamics. However,

of an excitable system, and thus is observed in electrophysiological models (Suckley and Biktashev, 2003), for example. However, for our model, we do not anticipate the presence of excitability from either a mathematical or a biological perspective; this is confirmed by numerical simulations below.

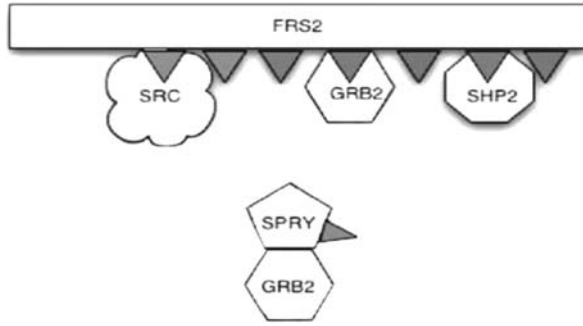


Fig. 2 The core of the signal transduction model, which cannot be easily subdivided. Bound SHP2 can also dephosphorylate SPRY bound to the FGF:FGFR653P654P:FRS2 complex via SRC.

not only does this immediately feed back into the rest of the reaction scheme by the fact that SHP2 binding alters dephosphorylation rates, we have that reducing the model to the extent considered is more than sufficient for our (current) needs. Thus, we do not pursue such model reductions further.

The coupling between SHP2 binding and the dephosphorylation dynamics of SPRY in the above needs further explanation. Once SPRY is bound to FRS2* via SRC, the model reaction scheme assumes that SHP2 binding will also induce SPRY dephosphorylation. However, as can be inferred from the pictorial representation in Fig. 2, the levels of SRC-SPRY binding are not *directly* tracked. Instead, we approximate η , which is *defined* by the fraction of phosphorylated SPRY bound to FRS2*:SRC:SHP2, i.e.

$$\text{FRS2}^*:\text{[SRC:SPRYP]:SHP2} = \eta \text{ SPRYP}.$$

Approximation of η

We take³

$$\eta \sim \underbrace{\left(\frac{\text{FRS2}^*:\text{SRC:SHP2}}{\text{FRS2}^*:\text{SRC}} \right)}_{\text{Fraction of bound SRC also bound to FRS2}^*:\text{SHP2}} \times \underbrace{\left(\frac{\text{FRS2}^*:\text{SRC:SPRYP}}{\text{SPRYP}_{\text{tot}}} \right)}_{\text{Fraction of SPRYP bound to SRC:FRS2}^*}.$$

To evaluate the fraction of bound SRC also bound to FRS2*:SHP2, we consider the dynamics of FRS2*:SRC:SHP2. With FRS2*:SRC:SHP2 denoted by x_{11} below, and the total amount of bound SHP2 denoted by z_1 , we have

$$\begin{aligned} \frac{dx_{11}}{dt} &= \lambda_1(y_1 - x_{11})(\text{SHP2}_{\text{tot}} - z_1) + \lambda_2(z_1 - x_{11})(\text{SRC}_{\text{tot}} - y_1) - (\eta_1 + \eta_2)x_{11} \\ &\quad - \frac{\mu x_{11}}{K + z_1} - \frac{\mu_1 x_{11}}{K_1 + x_{11}}, \end{aligned}$$

³We remark that the use of ratios of known concentrations for estimating concentrations of specific species is also mentioned in the supplemental material of Kiyatkin et al. (2006) for a large model of EGF signal transduction.

$$\frac{dz_1}{dt} = \lambda_1(\text{FRS2}^*_{\text{tot}} - z_1)(\text{SHP2}_{\text{tot}} - z_1) - \eta_1 z_1 - \frac{\mu z_1}{K + z_1},$$

where $y_1 = \text{FRS2}^*:\text{SRC}$ is the total amount of bound SRC. In addition $\mu_1, K_1, \mu, K, \eta_1, \eta_2, \lambda_1, \lambda_2$ are model parameters that represent reaction rate and Michaelis–Menten constants. Such equations can be derived from linear combinations of subsets of the full equations, on neglect of the slow dynamics, and given phosphorylation is fast compared to the binding dynamics. Given $y_1, \text{SHP2}_{\text{tot}}, \text{FRS2}^*_{\text{tot}}, \text{SRC}_{\text{tot}}$ one can readily evaluate the equilibrium level of z_1 , and hence x_{11} . This in turn enables the rapid evaluation of the ratio

$$\left(\frac{\text{FRS2}^*:\text{SRC}:\text{SHP2}}{\text{FRS2}^*:\text{SRC}} \right) = \frac{x_{11}}{y_1}$$

via simple algebraic relations.

Thus, we have one of the required ratios given $y_1, \text{SHP2}_{\text{tot}}, \text{FRS2}^*_{\text{tot}}, \text{SRC}_{\text{tot}}$. All but the first of these are known from the slow dynamics. For the value of $y_1 = \text{FRS2}^*:\text{SRC}$, i.e. the amount of bound SRC present, we use the value from the quasi-steady state of the previous slow timestep; at the initial timestep, it is known from the initial conditions. This is ultimately an approximation, though it is readily justified. In initial simulations, an iterative procedure was used to improve this approximation of $\text{FRS2}^*:\text{SRC}$ at each slow timestep, but it had an imperceptible effect. The approximation can also be justified a posteriori from the accuracy of the simplified model in the results section.

We proceed to consider the remaining ratio, i.e. the fraction of SPRYP bound to $\text{FRS2}^*:\text{SRC}$. Fast phosphorylation dynamics compared to SPRYP production entail that

$$\text{SPRY}_{\text{tot}} \sim \text{SPRYP}_{\text{tot}};$$

i.e. the total amount of phosphorylated SPRYP is approximated by the total amount of SPRYP. On the fast timescale, we also have the amount of SPRYP is constant as its production is on the slow timescale and similarly for the amount of SRC present. Our first aim is to calculate $\text{SPRYP}:\text{SRC}$. We have

$$\begin{aligned} \frac{d}{dt} \text{SPRYP}:\text{SRC} &= \lambda_3(\text{SPRYP}_{\text{tot}} - \text{SPRYP}:\text{SRC})(\text{SRC}_{\text{tot}} - \text{SPRYP}:\text{SRC}) \\ &\quad - \eta_3 \text{SPRYP}:\text{SRC} - \frac{\mu_2 \text{SPRYP}:\text{SRC}}{K_2/\eta_* + \text{SPRYP}:\text{SRC}}, \end{aligned} \quad (2)$$

where η_* gives the fraction of $\text{SPRYP}:\text{SRC}$ complexes which are bound to SHP2 via FRS2. Thus, the final term in the above represents SHP2 induced phosphorylation. In addition, $\mu_2, K_2, \eta_3, \lambda_3$ are model parameters and, as above, such equations can be derived from linear combinations of subsets of the full equations, on neglect of the slow dynamics given the above approximations. We can readily derive an expression for η_* , namely

$$\eta_* = \frac{\text{FRS2}^*:\text{SRC}:\text{SPRYP}:\text{SHP2}}{\text{SPRYP}:\text{SRC}} = \frac{\text{FRS2}^*:\text{SRC}:\text{SHP2}}{\text{SRC}_{\text{tot}}} = \frac{x_{11}}{\text{SRC}_{\text{tot}}}.$$

Thus, the equilibrium value of SPRYP:SRC can be expressed in terms of the smallest positive root, x , of the cubic

$$(\text{SPRYP}_{tot} - x)(\text{SRC}_{tot} - x) \left(\frac{K_2 \text{SRC}_{tot}}{x_{11}} + x \right) = \left(\frac{\eta_3 + \mu_2}{\lambda_3} \right) x$$

obtained by setting the time derivative in Eq. (2) to zero. One can readily implement a numerical solution of the above cubic, which gives us an approximation for $x = \text{SPRYP:SRC}$. The required ratio, i.e. the fraction of SPRYP bound to FRS2*:SRC is then simply given by

$$\underbrace{\frac{\text{SPRYP:SRC}}{\text{SPRYP}_{tot}}}_{\text{Fraction of SPRYP bound to SRC}} \times \underbrace{\frac{\text{FRS2}^*:\text{SRC}}{\text{SRC}_{tot}}}_{\text{Fraction of SRC bound to FRS2}^*}.$$

Use of η

Given η , FRS2^*_{tot} , SRC_{tot} , GRB2_{tot} , SPRY_{tot} , SHP2_{tot} we can then solve for the quasi-steady state of the system of reactions represented by Fig. 2, and thus determine levels of GRB2 bound directly to FRS2* and refine our estimates of any expression used in the estimate of η , for example $\text{FRS2}^*:\text{SRC}:\text{SHP2}$. One can subsequently determine the levels of bound CBL and levels of GRB2 bound to FRS2* via SPRY, as demonstrated in the following.

SPRYP:SRC binding to FRS2* and GRB2

The SPRYP–SRC interaction is assumed independent of the FRS2*–SRC interaction. Hence, we have that the amount of $\text{FRS2}^*:\text{SRC}:\text{SPRYP}$ is given via

$$\frac{\text{FRS2}^*:\text{SRC}:\text{SPRYP}}{\text{SPRYP:SRC}} = \frac{\text{FRS2}^*:\text{SRC}}{\text{SRC}_{tot}}.$$

By similar arguments, we have that the amount of $\text{FRS2}^*:\text{SRC}:\text{SPRYP}:\text{GRB}$ is given by

$$\text{FRS2}^*:\text{SRC}:\text{SPRYP}:\text{GRB2} = \text{SPRYP:GRB2} \times \frac{\text{FRS2}^*:\text{SRC}:\text{SPRYP}}{\text{SPRYP}_{tot}}$$

yielding the amount of GRB2 bound to FRS2* via SPRYP.

CBL dynamics

The CBL dynamics can be captured as follows. Let CBL_{free} denote free CBL, CBL_{bnd} denote bound CBL and S_* denote the SPRYP available to bind to free CBL. We have the reaction



The steady state of the equations which represents reaction (3) gives us that

$$\lambda \text{CBL}_{bnd} = \mu \text{CBL}_{free} S_*,$$

where λ and μ are the forward and back reaction rate constants, respectively. We also know that

$$\text{CBL}_{free} + \text{CBL}_{bnd} = \text{CBL}_{tot},$$

the total amount of CBL present and

$$S_* = \text{SPRYP}_{tot} - \text{CBL}_{bnd}.$$

The latter entails that the amount of SPRYP available for reaction with CBL is the total amount of SPRYP minus the SPRYP bound to CBL.

Thus, we have a quadratic in CBL_{bnd} , with coefficients depending on λ/μ , SPRYP_{tot} , CBL_{tot} , with solution

$$\text{CBL}_{bnd} \stackrel{def}{=} \text{SPRYP:CBL} = \frac{1}{2} \left[\eta_{**} - \sqrt{\eta_{**}^2 - 4\text{CBL}_{tot}\text{SPRYP}_{tot}} \right],$$

$$\text{where } \eta_{**} = (\text{SPRYP}_{tot} + \text{CBL}_{tot} + \lambda/\mu).$$

The amount of SPRYP:CBL bound to $\text{FRS2}^*\text{:SRC}$ is given by

$$\text{FRS2}^*\text{:SRC:SPRYP:CBL} = \text{SPRYP:CBL} \times \frac{\text{FRS2}^*\text{:SRC:SPRYP}}{\text{SPRYP}_{tot}}.$$

2.3.5. The slow dynamics (no PLC present)

Given FRS2^*_{tot} , SHP2_{tot} , SRC_{tot} , SPRY_{tot} , CBL_{tot} , $y_1 = \text{FRS2}^*\text{:SRC}$, we can find η , solve for the quasi-steady state of the reactions in Fig. 2, and hence determine levels of bound GRB2, SHP2, CBL, SRC, SPRYP:CBL and similar molecules, as discussed above. In terms of the Tikhonov procedure, we are locating the manifold $\mathbf{F}(\mathbf{f}, \mathbf{s}) = \mathbf{0}$. The system dynamics is then simply governed by the slow dynamics. This is driven by

- SRC induced relocation of FRS2 and all molecules bound to it.
- Ubiquitination via CBL and the subsequent break up of FRS2, with all molecules bound to FRS2 released back, and available for further participation in the dynamics.
- Slow production of SPRY.

Thus numerically solving for the slow dynamics relies upon timestepping FRS2^*_{tot} , FGFR_{tot} , SRC_{tot} , SHP2_{tot} , GRB2_{tot} , CBL_{tot} , SPRY_{tot} and ubiquitinated protein levels according to the above interactions. This is a very simple timestepping of a small number of ODEs, without disparate timescales.

The results produced by the fast dynamics at the first timestep of the slow dynamics is, in the context of matched asymptotics, the inner solution. The results produced at each timestep of the slow dynamics is, in the same context, the outer solution; these two solutions are numerically combined in the standard manner to generate the uniformly valid composite solution.

2.3.6. A minor complication: incorporating the PLC dynamics

The PLC dynamics introduces an additional complication which we have deferred to after the explanation of the basic structure of our simplification method.

The 766 site phosphorylates rapidly on the fast timescale, but the binding to PLC is really on an intermediate timescale, at least for the reference parameters. The PLC binding occurs much faster than the slow decays induced by CBL and analogous reactions. However, it is significantly slower than SRC binding to FRS2^* and other similar reactions. Treating the PLC binding in the same manner as SRC binding, there is an early time, short lived error, in the predicted PLC binding levels, as described in the results section.

The above mentioned error in the early time predictions of the PLC dynamics can be avoided as follows. On the fast timescale PLC_{tot} , $FRS2^*$ are constant and the dynamics of PLC_{bnd} , the level of bound PLC, is governed by

$$\begin{aligned} \frac{dPLC_{bnd}}{dt} &= \lambda_4(PLC_{tot} - PLC_{bnd})(FRS2^* - PLC_{bnd}) - \eta_4 PLC_{bnd} \\ &\stackrel{def}{=} F(PLC_{bnd}; PLC_{tot}, FRS2^*, \lambda_4, \eta_4). \end{aligned} \quad (4)$$

If the PLC binding dynamics were sufficiently fast this equation would rapidly relax to its equilibrium state, giving a Tikhonov estimate for PLC_{bnd} as the smallest root of

$$F(PLC_{bnd}; PLC_{tot}, FRS2^*, \lambda_4, \eta_4) = 0. \quad (5)$$

However, the timescale of the PLC binding dynamics is not sufficiently fast for this to be very accurate at early times. To circumvent this difficulty, note that during a timestep of the slow dynamics, PLC_{tot} and $FRS2^*$ are approximately constant. One can therefore timestep PLC_{bnd} over the same slow timestep using the analytical solution of Eq. (4), assuming constant PLC_{tot} , $FRS2^*$ which induces only a very small error. Such errors are also not expected to accumulate (as confirmed in the results section) since the qualitative dynamics of the above differential equation is such that PLC_{bnd} will always relax toward the smaller root of Eq. (5).

Thus, to incorporate the PLC dynamics, one proceeds as follows. As previously, given $FRS2^*_{tot}$, $SHP2_{tot}$, SRC_{tot} , $SPRY_{tot}$, CBL_{tot} , $y_1 = FRS2^*:SRC$, one can find η , solve for the quasi-steady state of the reactions in Fig. 2, and hence determine levels of bound GRB2, SHP2, CBL, SRC, $SPRYP:CBL$, and similar molecules. In a standard Tikhonov procedure solving for the smallest root of Eq. (5) gives PLC_{bnd} . In the slow timestep, we proceed as discussed above, and additionally consider

- the PLC induced break up of FGFR. Note that the FRS2 and its bound molecules are released from this decay, and thus available for further participation in the dynamics.

To remove the early time error in the levels of PLC error one supplements the above slow dynamics with a time-stepping of Eq. (4) above, to find the levels of bound PLC rather than solving for the root of Eq. (5) in a standard Tikhonov procedure.

We illustrate the improvement for the full model below, and this improvement is also implemented for the SRC knockout due to the heightened importance of the PLC dynamics in this scenario. However, given the error does not in fact impact on the model's important predictions and the fact it can require a careful choice of the slow dynamics timestep to ensure that the simplified model runtime is not extended to any significant extent while retaining accuracy, it is not used universally. Nonetheless, this example specifically illustrates that this simplification procedure can incorporate processes on an intermediate timescale, thus increasing its applicability.

2.3.7. Model simplification summary

The ultra-fast dynamics of Section 2.3.2 and Appendix C govern the initial conditions of the simplified model in terms of the asymptotic predictions given for FGF:FGFR653P654P:FRS2. The ultra-fast dynamics are otherwise ignored. The remaining fast and slow dynamics are considered as follows.

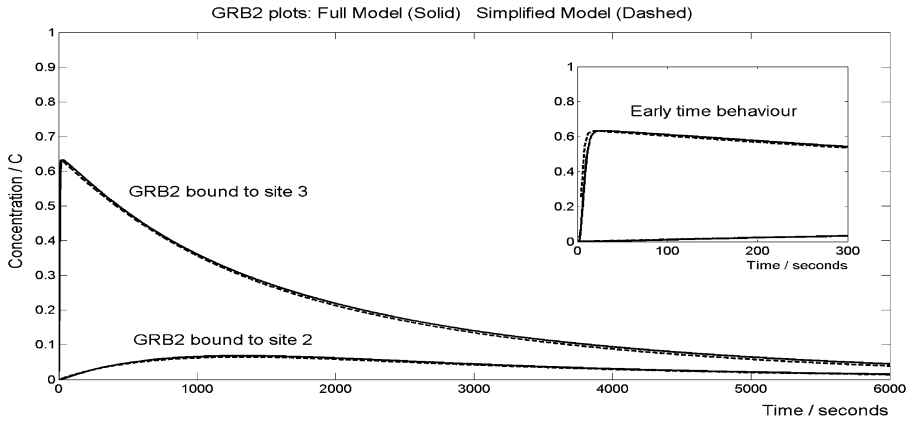


Fig. 3 In the above are results for the GRB2 binding to FRS2* for the full numerical model (solid) and the simplified model (dashed). In the inset is the early time behavior. Here, and for all other figures, C denotes the reference concentration scale. In addition, Site 3 refers to the site where GRB2 binds directly to FRS2* while Site 2 refers to the site where GRB2 binds to FRS2* via SPRY. As indicated in Appendix A, if one assumes a concentration scale of 5×10^4 molecules per cell, one has $C \sim 10^{-7} \text{M}$.

At each timestep of the slow dynamics, we approximate

$$\eta \sim \left(\frac{\text{FRS2}^*:\text{SRC}:\text{SHP2}}{\text{FRS2}^*:\text{SRC}} \right) \times \left(\frac{\text{FRS2}^*:\text{SRC}:\text{SPRYP}}{\text{SPRYP}_{\text{tot}}} \right)$$

using the values of $\text{FRS2}^*:\text{SRC}$, $\text{FRS2}_{\text{tot}}^*$, SRC_{tot} , GRB2_{tot} , SPRY_{tot} , SHP2_{tot} . Then we solve for the quasi-steady state of the dynamics depicted in Fig. 2. Using these quasi-steady values, all remaining quasi-steady state values can be determined as detailed above. The slow dynamics can then be time-stepped numerically in a straightforward manner. In addition, the levels of bound PLC are tracked as described above.

This simplified model is far more amenable to rapid numerical computation. The ODEs on the slow timescale are neither stiff nor excessively numerous while finding the quasi-steady state of the dynamics shown in Fig. 2 benefits from having a very good estimate, namely the solution at the previous slow timestep.

3. Results

3.1. Representative results

In Fig. 3, plots of the predictions for GRB2 binding are given for the full model described above (solid lines), with full details and parameter values specified in Appendix A. Also presented are the predictions from the simplified model (dashed lines). Both for this figure, and the subsequent analogous figures, the presence of only one solid line demonstrates that the predictions of the simplified and full model are identical to within the resolution of the figures presented.

In Fig. 4, plots for other concentrations are given both for the full model (solid lines) and the simplified model (dashed lines). The discrepancy between predicted levels of free

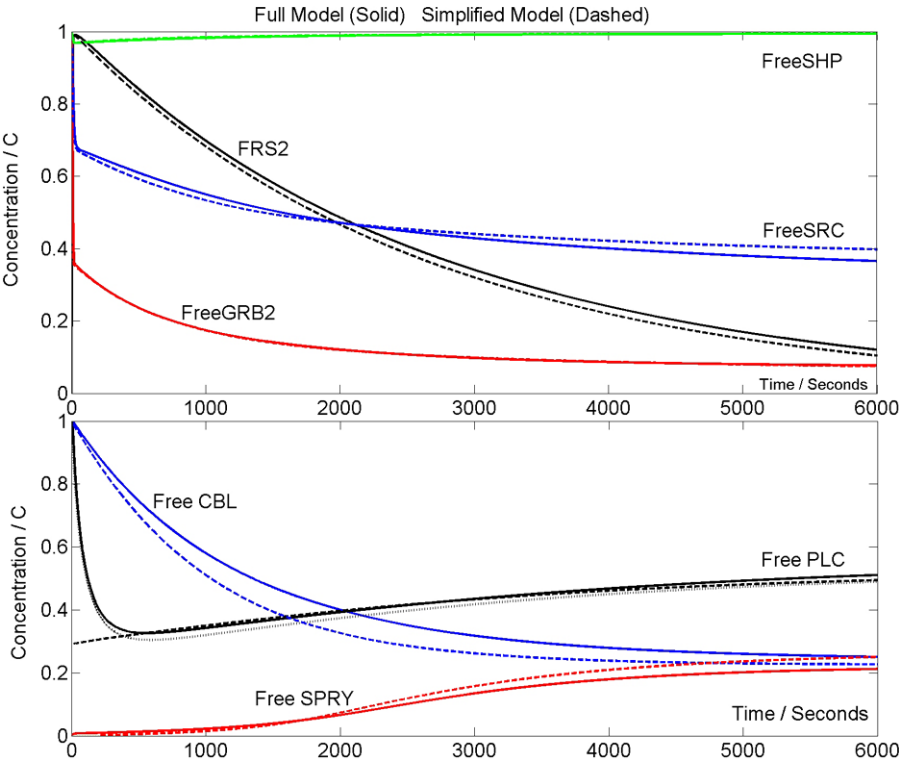


Fig. 4 In the above are results for the levels of FRS2* plus free SHP2, SRC, GRB2, SPRY, CBL, and PLC for the full numerical model (solid) and the simplified model (dashed). See text for an explanation of the dotted line in the lower plot for the levels of free PLC.

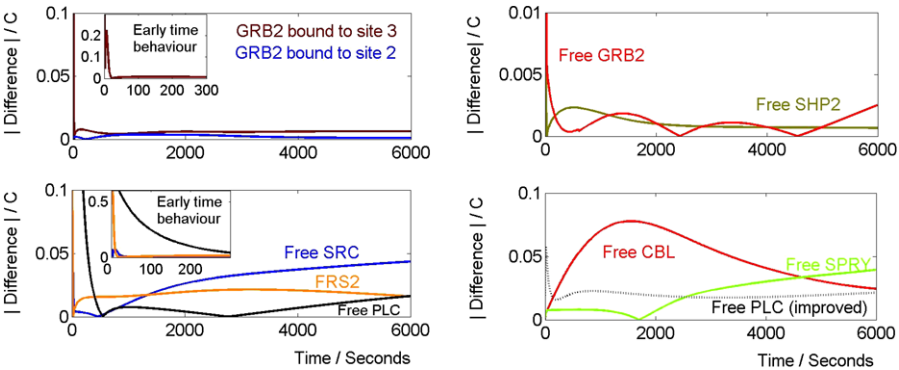


Fig. 5 In the above are plots of the magnitude of the difference between the simplified model on comparison with the full model. The PLC differences are given for the simplified model with (black dotted line) and without (black solid line) the PLC improvement discussed in Section 2.3.6.

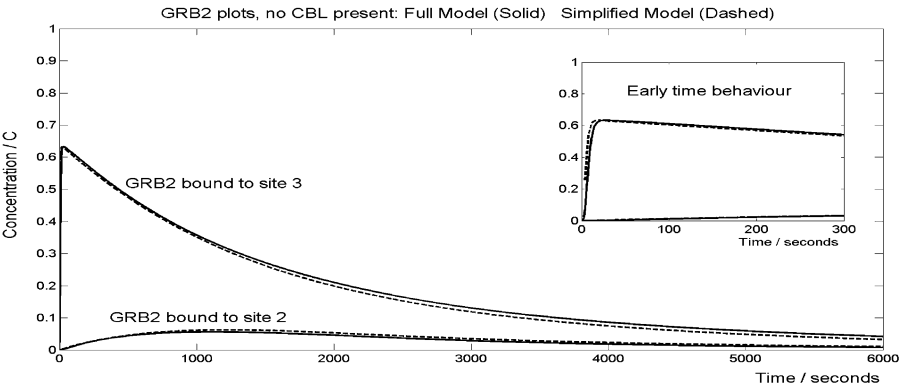


Fig. 6 In the above are results for the GRB2 binding to FRS2* for the full numerical model (solid) and the simplified model (dashed), in the absence of CBL. In the inset is the early time behavior.

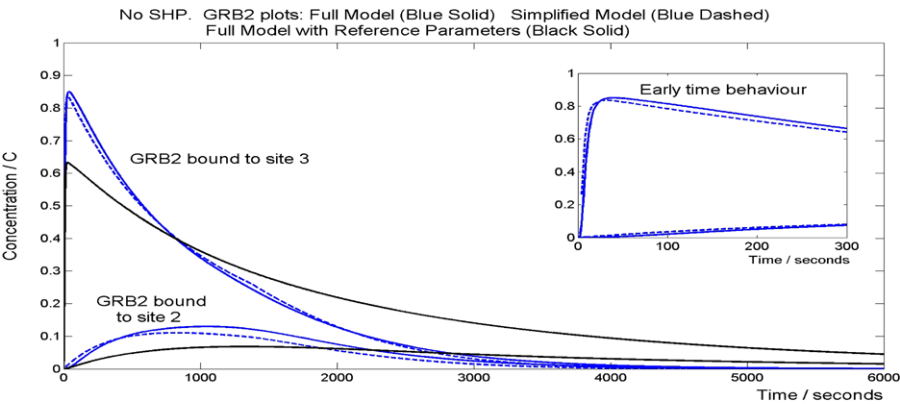


Fig. 7 In the above are results for the GRB2 binding to FRS2* for the full numerical model (blue solid) and the simplified model (blue dashed), in the absence of SHP2. The black solid line is the prediction of the full model and reference parameters for comparison. The early time behavior in the absence of SHP2 is inset.

SRC and free CBL observed in these plots appears to be a characteristic, but a small, inaccuracy of the simplified model. In Fig. 4, three results are presented for the levels of free PLC. The solid and dashed lines are the full and simplified model predictions, respectively, where the root of Eq. (5) is used to track the PLC dynamics for the latter. If one uses the simplified model, solving the differential equation (4) to track PLC, the resulting prediction for the PLC dynamics is the dotted line in the lower plot of Fig. 4. This is discernable, albeit only just, from the full model prediction. The plots of all other concentrations are the same, at the level of resolution of the graphs presented, regardless of whether one uses the simple PLC dynamics, i.e. the root of Eq. (5), or the more complicated PLC dynamics, summarized by Eq. (4). All results for remaining figures have been derived using the simpler PLC dynamics, unless stated otherwise.

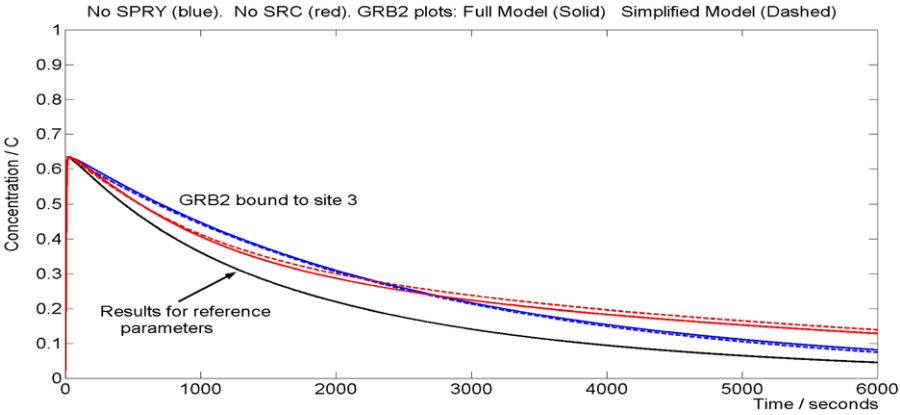


Fig. 8 In the above are results for the GRB2 binding to FRS2* for the full numerical model (solid) and the simplified model (dashed), in the absence of SPRY (blue) and the absence of SRC (red). The curve in black is the prediction of the full model with the reference parameters for comparison.

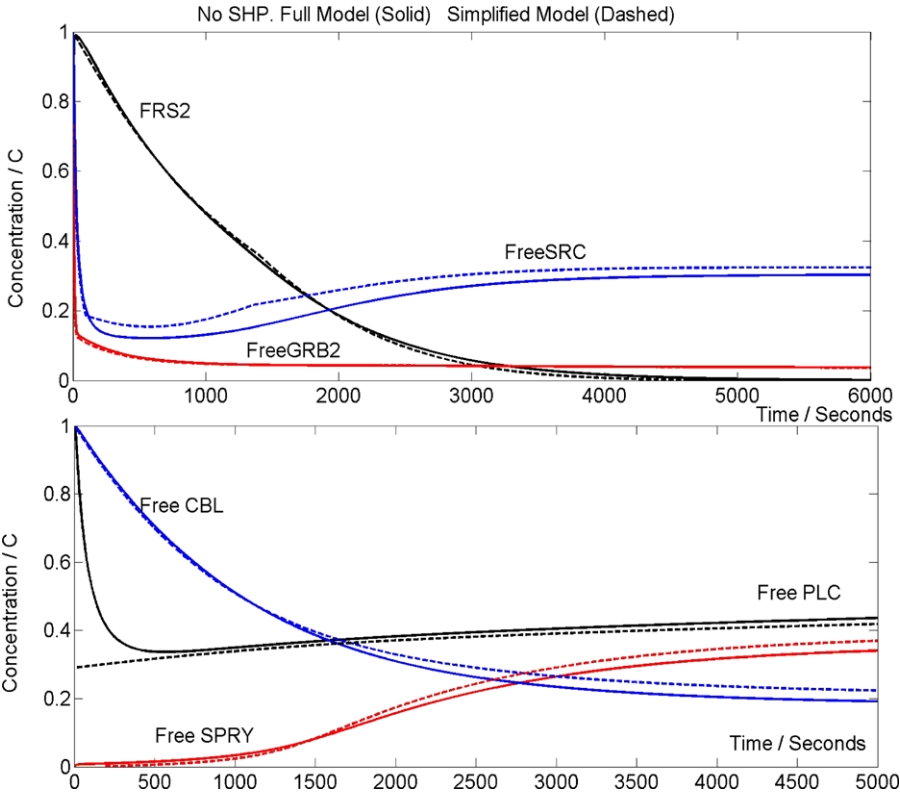


Fig. 9 In the above are results for the levels of FRS2* plus free SRC, GRB2, SPRY, CBL, and PLC for the full numerical model (solid) and the simplified model (dashed) when no SHP2 is present.

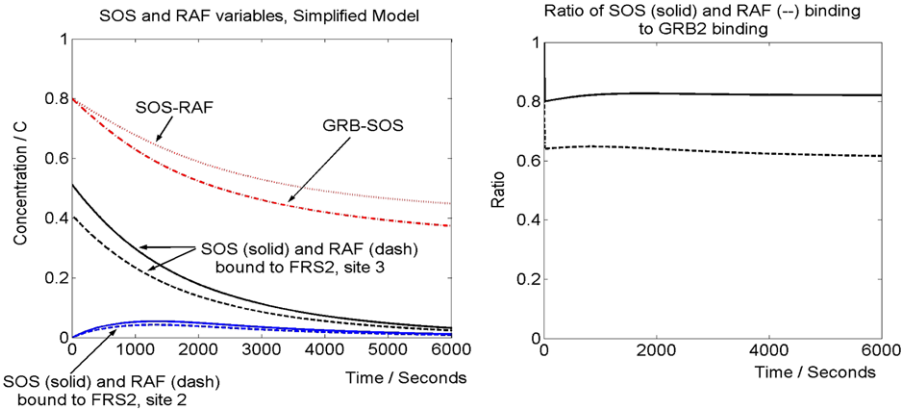


Fig. 10 In the above left are the simplified model's predictions, with the reference parameters, for SOS and RAF binding levels with GRB2 and SOS, respectively. The amounts bound to FRS2* via the two different possible sites are also given. In the above right is the ratio of FRS2*:GRB2 to FRS2*:GRB2:SOS (solid) and the ratio FRS2*:GRB2 to FRS2*:GRB2:SOS:RAF (dash). This ratio is the same (at the level of resolution of the plots) for GRB2 bound to FRS2* directly at site 3 and for GRB2 bound to FRS2* via SRC and SPRY at site 2.

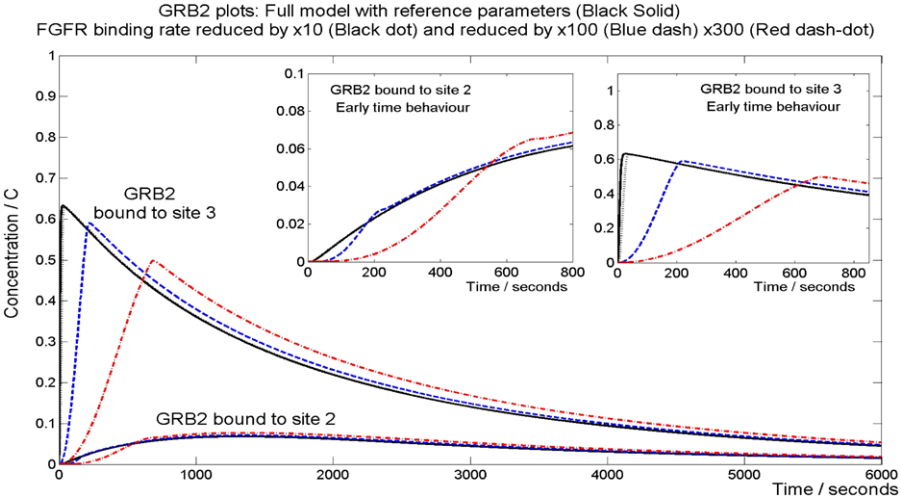


Fig. 11 The effect of varying of the FGF binding rate. The black solid line is the full model with reference parameters. The black dotted line gives the effects of reducing the FGF-FGFR binding rate by a factor of 10. This binding rate is reduced by a factor of 100 for the blue dashed line and 300 for the red dot-dash line.

In Fig. 5, the magnitude of the difference between the simplified model and the full model is plotted for the variables considered in Figs. 3 and 4. The agreement between the full and simplified models is particularly noteworthy especially for bound GRB2, excluding initial transients, which are only present on the ultra-fast timescale. In addition,

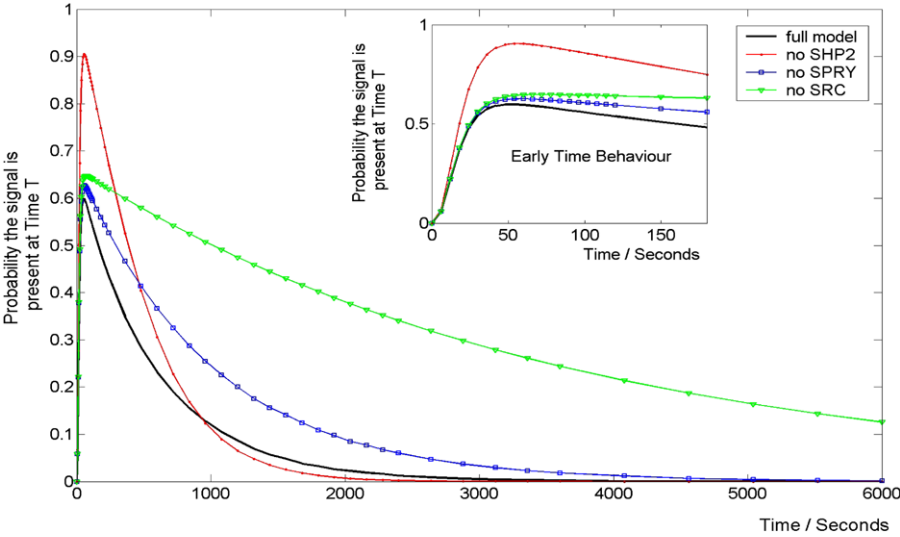


Fig. 12 Results for total bound levels of GRB2 from a previous stochastic model of FGF signaling, with early time results inset.

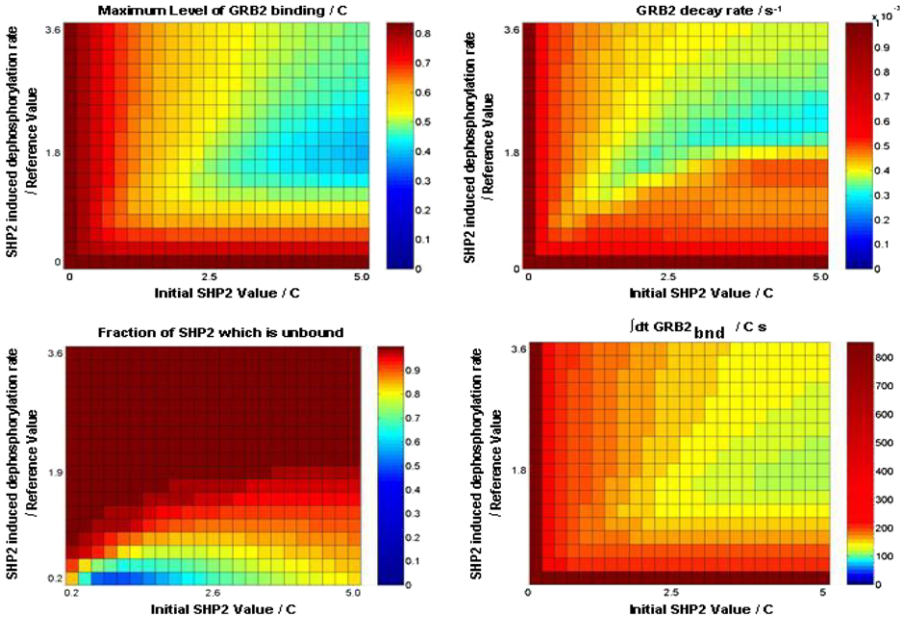


Fig. 13 Plots of (i) the maximum of total bound GRB2 (upper left), (ii) its initial decay rate (upper right), (iii) the relative amount of free SHP2 (lower left) and (iv) $\int_0^T dt \text{ GRB2}_{bnd}$, for $T = 300$ seconds, as functions of the amount of initial SHP2 present and the rate SHP2 induces dephosphorylation. The above graphs (of approx. 400 squares) were generated from 108 simulations, using interpolation to refine the resolution of the results.

the full model simulation took about 5 days to complete, which was over 275 times longer than the simplified model. Similar timescales are observed in all simulations.

3.2. Representative results for knockouts

In Figs. 6, 7, and 8 results for the levels of GRB2 binding are presented for the absence of, respectively, CBL, SHP2, SRC, and SPRY.

The absence of CBL has a barely discernable effect on the levels of GRB2 binding to FRS2*. Similar comments apply for the predicted levels of FRS2*, plus unbound levels of SHP2, GRB2, SPRY, and PLC together with the prediction of a small decline in the levels of free SRC (results not shown). Small discrepancies between the full and simplified models' predictions for levels of free SRC do occur, analogous to those observed in Fig. 4.

The absence of SHP2 is observed to have a substantial effect, with higher initial binding of GRB2 followed by a faster decay rate, again as predicted by both the full and the simplified model. Similarly, the absence of SHP2 induces significant effects on the levels of FRS2*, plus the unbound levels of SPRY, GRB2, and SRC; see Fig. 9. Small differences are induced in the levels of free PLC, and a small reduction in levels of free CBL is also predicted relative to the prediction for the reference parameters of Appendix A. Small differences in the full and simplified models' predictions occur, especially for the levels of free SRC, as can be observed in Fig. 9.

The absence of SPRY also has a significant effect, in that the decay of bound GRB2 is slower; similar observations hold for the absence of SRC. Thus, not surprisingly, one has the model prediction that both SRC and SPRY are antagonists for GRB2 binding to FRS2*.

In the absence of SRC, the decay of bound GRB2 proceeds via PLC degradation; thus, the more complicated PLC estimate, based on Eq. (4) is used. Even in this situation, the choice between the simple or complex representation of the PLC dynamics has no discernable effect on other concentrations, and minimal effect on the PLC predictions after initial transients, at the resolution of the figures.

Again, the level of agreement between the full model and the simplified model is highly noteworthy for all knockouts, as is the improvement in numerical performance.

3.3. Representative results from the simplified model for SOS and RAF binding

The simplified model is readily extended to consider SOS and RAF binding, and results are presented in Fig. 10. Strictly, the presented results will possess errors due to transients at very early times, analogous to the PLC results. However, these errors will only be present for very short times due to the faster reaction rates of SOS and RAF compared to the reaction rates of PLC. From the right-hand plot, we can see that GRB2 binding levels give an accurate measure of the downstream binding levels of SOS and RAF.

3.4. Representative results for the reduction in FGF-FGFR binding rates

We see from Fig. 11 that reducing FGF-FGFR binding rates increases the time it takes for the rise in GRB2 binding to occur (thus decreasing peak levels of GRB2 bound directly to FRS2*). The model is relatively robust to changes in FGF-FGFR binding rates, taking a reduction of $\times 100$ for a significant effect. However, once the GRB2 levels have

peaked, the decay of bound GRB2 proceeds essentially independently of the FGF-FGFR binding rates, as can clearly be seen in the results.⁴ The simplified model predicts that reducing FGF:FGFR binding rates has no effect. This discrepancy is readily understood as discussed below.

3.5. Representative results for a stochastic model

For reference and comparison purposes, the averages of multiple runs for a stochastic model based on the reaction scheme detailed in Appendix A are presented in Fig. 12. A simplification in the reaction scheme is that the phosphorylation and dephosphorylation dynamics are modeled as linear first order kinetics enabling a straightforward use of Gillespie's algorithm; see Heath et al. (2006), Kwiatkowska et al. (2006), Heath et al. (2007) for further details concerning these simulations.

3.6. Representative results given alterations of the SHP2 parameters

It is apparent that the model is *highly* sensitive to SHP2 dynamics; only $O(4\%)$ of the SHP2 molecules ever bind to FRS2* for the reference model, yet this is sufficient to induce a substantial difference between the predictions in the presence and absence of SHP2. Thus, we investigate in detail how the model responds to changes in the levels of SHP2 and the rate of SHP2 induced dephosphorylation. That is, we alter both the amount of SHP2 and the strength of its effect. The peak value and initial decay rate of the total bound GRB2, plus the integral of total bound GRB2 over the first 300 seconds, are given in Fig. 13. The latter is a (crude) measure of the average GRB2 signal induced during the 5 minutes following the initiation of an FGF signal. In addition, the fraction of free SHP2 is also plotted in this figure.

4. Discussion and conclusions

The first thing to note about Figs. 3 and 6, 7, and 8 is that they all show a relatively rapid rise of GRB2 binding, due mainly to binding with FRS2* directly, followed by a decay. This is the qualitative behavior observed in experiments. In addition, Fig. 10 shows that levels of bound SOS and RAF track levels of bound GRB2 demonstrating that the concentration of bound GRB2 gives an excellent measure of the downstream effects of the FGF signaling.

The level of agreement between the full model and the simplified model in Figs. 3, 4, 5, 6, 7, 8, and 9 is highly noteworthy.

Nonetheless, there are discrepancies, though these are readily understood. The first discrepancy concerns errors in the early time predictions for PLC binding, except when the simplified model uses the more complicated PLC dynamics summarized by Eq. (4).

⁴The details of the justification presented in Appendix C for the fact we need only consider FRS2 binding to FGF:FGFR653P654P and not to singly phosphorylated and unphosphorylated FGF:FGFR does strictly breakdown for a greatly reduced FGF-FGFR binding rate. However, the full model simulation is still readily anticipated to be independent of these details due to fact phosphorylation still occurs on a fast timescale compared to downstream reaction timescales even if FGF:FGFR binding is not.

However, these early time errors do not impact on the rest of the model predictions because the PLC decay dynamics act on a much slower timescale compared to the timescale characterizing the existence of these early time errors. Hence, the global error is small, as can be seen from the fact the simplified model tracks other model variables accurately. With the more complicated PLC dynamics, the initial transient errors in PLC levels are eliminated; this latter observation additionally demonstrates that the simplification technique can even accurately incorporate processes occurring on intermediate timescales.

A second discrepancy concerns Fig. 11, but again this has a straightforward explanation. The fact a 10-fold reduction in the FGFR-FGF binding rate does not have much of an effect on the presented results is because the simplified model is completely insensitive to reducing the FGFR-FGF binding rate. Thus, one must reduce this binding rate to a level where the simplified model breaks down, i.e. so that the timescale of the FGFR-FGF binding rate is no longer ultra-fast, only fast, which requires a *very* large change, much greater than an order of magnitude, for the reference parameters. Also, the fact the signal attenuation dynamics is essentially unperturbed is predictable as it is governed by the slow dynamics of the system. Given reducing the FGFR-FGF binding rate is decoupled from the irreducible dynamics of Fig. 2, it also strongly suggests that an efficient, simplified, model can be constructed for this parameter regime, though we have not pursued this. Biologically, we have the prediction that the system can be resilient to the effects of FGF kinase inhibitors, which will have to either completely turn off FGFR activation, or at least reduce its timescale to a level commensurate with the timescales of FRS2* and GRB2 binding, for any significant effect. In addition, this effect will act to reduce the maximum level of GRB2 binding, and to increase the time it takes to reach this peak, with minimal influence on the subsequent decay timescale.

The fact the ultra-fast dynamics is ignored in the simplified model, except that ultra-fast large time asymptotes yield simplified model initial conditions, entails that initial transient deviations of the full model, compared to the simplified model, are expected and indeed observed in Fig. 5. Such transient deviations are known to decay on the timescale of the ultra-fast dynamics and consequently do not affect model predictions on longer timescales; further details are presented in Appendix C.

Finally, given the duration of the run is much larger than the (fast) reaction timescale, even the largest of the remaining discrepancies is not inconsistent with the size of the asymptotically small parameter, namely the ratio of timescales between the fast and slow dynamics. Unlike the errors in the ultra-fast dynamics, which are analytically shown to decay, there is the possibility of an accumulation of error with time for the coupled fast and slow dynamics, and this is seen with the initial behavior of CBL in Figs. 4 and 5, for example. Here, the peak difference between the full and simplified model, which corresponds to a relative error of about 17%, is about an order of magnitude larger than the ratio of fast and slow timescales ($O(1/50)$). However, this error builds up over a timescale which is more than an order of magnitude larger than the typical fast reaction timescale. We anticipate these cumulative errors are ultimately controlled by the global dynamics, particularly the transient nature of the FGF signaling response. This ensures that such errors do not accumulate indefinitely as all variables are slave to the fact the system relaxes to an “unexcited” state over time. This does, however indicate that this simplification procedure should be used with caution when the system’s response is not transient.

The qualitative behavior of the model is the same as a stochastic model based upon the Gillespie algorithm; see Fig. 12. For example, the absence of SHP2 results in a more

extensive rise in GRB2 binding levels followed by a faster decay, while the absence of SPRY or SRC results in a slower decay. From such observations, it is clear that despite the complexity of the full model, it is extremely robust; even different modeling frameworks give the same qualitative behavior, with no indication that this behavior exhibits sensitivity to fluctuations. This is, in fact, is not surprising. There are multiple redundancies in the model to ensure the slow signal attenuation dynamics. The ultrafast dynamics is extremely robust as shown in Appendix C. The qualitative behaviors are also insensitive to the details of the fast dynamics, in that GRB2 binding will take place in the core model of Fig. 2, given the outcome of the ultra-fast dynamics. The accuracy of the simplified model, compared to the full model, also clearly demonstrates that an enormous amount of detail in the full model is redundant and not required, even for quantitative predictions.

However, the quantitative details of signal growth and attenuation can be sensitive to a number of details of the modeling. For example, one *has* to be precise in the definition of the parameter, η , in Section 2.3.3, when capturing the effect of SHP2 on SPRY within the simplified model. Errors in this parameter lead to substantial quantitative discrepancies between the simplified model and the full model (not shown) especially as small errors in the decay rate are magnified into substantial errors at large times. The sensitivity of quantitative predictions is also illustrated by comparing the predictions for GRB2 binding levels in the absence of SRC to predictions in the absence of SPRY. The deterministic model shows the levels of GRB2 binding will cross as in Fig. 8, in contrast to the stochastic model; see Fig. 12.

The investigation of the sensitivity of GRB2 dynamics with respect to the initial levels of SHP2 and the rate SHP2 induces dephosphorylation also reveals a sensitive, complicated, and non-intuitive, system dynamics. From Fig. 13 we can see that simply varying one of these parameters at a time in a sensitivity analysis risks missing much of this complex behavior if one's starting position is toward the lower left of the graphs depicted. We also see sensitivity as either the initial amounts of SHP2 or its induced dephosphorylation rate tends to zero. In addition, at relatively high levels of SHP2 there emerges a non-linear, non-monotonic, dependence on SHP2 induced dephosphorylation rates. In particular, one can readily extract counter-intuitive predictions; for example, increasing the SHP2 dephosphorylation rate can act to either increase or decrease levels of GRB2 binding depending on parameter values.

This complexity ultimately arises because SHP2 induces a negative signal not only on GRB2 binding, but also on its own phosphorylation and other negative regulators of GRB2, such as SPRY and SRC; the detailed behavior thus depends on the relative importance of all these effects. In particular, such observations imply a modeling prediction that the control of FGF signal transduction, via the modulation of the level of SHP2 or its effects, would be difficult due to complex feedbacks. It also indicates that the overexpression of SHP2 may have minimal effect, or a very strong influence on GRB2-binding, depending on the rate of SHP2 induced dephosphorylation, with maximal influence at intermediate dephosphorylation rates. More generally, this illustrates the difficulty of gaining a systems level understanding and of using intuition in the presence of feedback loops within a signal transduction system.

Nonetheless, despite the complexity of the SHP2 dynamics, inspection of Fig. 13 can also reveal robust predictions. For example, the observed increase in both the GRB2 decay rate and the maximum level of GRB2 binding on removing SHP2, as depicted in Fig. 7, is essentially universal. A similarly robust increase in the initial mean level of GRB2 binding

over the first 300 seconds is also observed; the robustness of such predictions entails that they are potentially testable.

The notion of key sets of equations decoupling for models of signal transduction was first raised by Borisov et al. (2005). This, together with a Tikhonov procedure, has allowed an enormous reduction in the complexity of the model. Further simplifications have been implemented allowing, for example, the incorporation of the influence of SHP2 on the dephosphorylation of SPRY bound to FRS2* via SRC, and the competition between SPRY and FRS2* for GRB2. While Michaelis–Menten kinetics have been used for the phosphorylation and dephosphorylation dynamics, this is not a pre-requisite for applying the presented simplification techniques to the current model. It is anticipated one could analogously simplify the modeling equations that would arise if the assumed underlying enzyme dynamics of phosphorylation and dephosphorylation were not approximated a priori with Michaelis–Menten kinetics.

We remark that there is no natural way to speed up the numerical algorithms of the full model due to the combination of stiffness and very large system size. Thus, the fact the simplified model can be simulated hundreds of times faster for only a small loss of accuracy is highly noteworthy. The model simplification not only expedites numerical simulation; it also simplifies one's understanding of the key features of the model. The effects of the ultra-fast dynamics are clear, as are the effects of the slow dynamics. Much of the fast dynamics is simply governed by the equilibrium state of simple interactions plus the core reaction depicted in Fig. 2. This core is still too complicated to solve analytically in general. Nonetheless, it does allow one to see the underlying mechanisms and feedbacks are extremely robust, at least at a qualitative level. It also allows one to see how this signal transduction model can be reduced to distinct modules of interaction. One can also immediately see where the most complicated dynamics will lie, namely in the effects of SHP2 within the core reactions of Fig. 2.

The improved numerical performance of the simplified model allows detailed analyses of parameter variations; the production of Fig. 13 would have been impractical with the full model, for example. It also means that one has much more freedom in investigating different possible biological models, with additional binding molecules and/or additional binding sites despite the further combinatorial explosion this would induce. Similarly, the simplified model gives one the flexibility to start considering modeling downstream of GRB2, SOS, and RAF. It is also sufficiently manageable to allow the future investigation of spatial aspects of signal transduction, while retaining the biological realism of the full model. Similarly, such techniques should be applicable for the process calculi models of FGF signaling (Heath et al., 2006; Kwiatkowska et al., 2006; Heath et al., 2007), as illustrated by the work of Borisov et al. (2006) on stochastic systems. In turn, this may facilitate the further development of stochastic models to more complicated biological scenarios.

There is also the question of whether the simplification process can be automated. A discussion of the difficulties in automating a Tikhonov reduction are discussed in Vallabhajosyula and Sauro (2006) while finding subsets of equations which decouple is anticipated to reduce to a well-posed problem in linear algebra. However, the process of automating the estimation of a parameter such as η in Section 2.3.3 to allow one to deal with interactions between different chains of binding molecules on a large docking protein may prove more challenging. Nonetheless, an automatic simplification procedure which

could generate a simplified model from a suitable template, for example, but not exclusively the Systems Biology Markup Language (Hucka et al., 2003), would be incredibly useful.

In summary, we have extensively simplified a mass-action model of FGF signal transduction, enabling a reduction of over two orders of magnitude in the numerical runtime, while retaining accuracy. We have seen an overall agreement at the qualitative level with stochastic models but there are discrepancies between the detailed predictions of the two modeling frameworks. The model of the signal transduction mechanism, despite its apparent enormous complexity, can be decomposed into units, the most complex of which is relatively straightforward and illustrated in Fig. 2. This in turn yields substantial insight into the behavior of this complex system, illustrating why and how the SHP2 dynamics will be particularly complex. It has also enabled a detailed numerical investigation illustrating, for example, the robustness of model predictions for the effects of removing SHP2, which are thus potentially testable. There is also enormous potential for the further application of the simplified model, especially as a template for other, more complicated, models of FGF signal transduction. This could include the development of models of the complexity studied by Schoeberl et al. (2002) for EGF signal transduction, while incorporating Michaelis–Menten kinetics, a starting point for spatially distributed models or the inclusion of more realistic stoichiometries.

Acknowledgements

E.A.G. and M.Z.K. acknowledge the support of EPSRC grant GR/S72023/01 (Integrative Biology E-science). M.Z.K. and J.K.H. acknowledge the support of Microsoft Research Cambridge contract MRL 2005-44 (Computational tools for advancing science). J.K.H. acknowledges the support of Cancer Research UK and Endotrack. E.A.G. is grateful to Gethin Norman for correspondence concerning the stochastic models.

Appendix A: Complete specification of the model

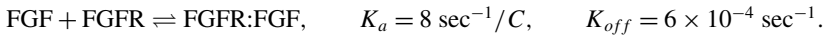
The reference concentration scale is C ; the reference timescale is $T = 1$ second. All concentrations in the results section are in units of C and all times are in units of seconds. A typical value of C , assuming 5×10^4 molecules per cell and a mean cell volume of 10^{-15} m^3 would be $C = 10^{-7} \text{ M}$ (1SF). Please refer to Fig. 1 for a pictorial representation and summary of the following reactions. All reactions are assumed to occur in parallel. As an illustrative example, the ubiquitination of FRS2:SRC:SPRYP:CBL is assumed to proceed whether or not GRB2 is bound to the phosphorylated SPRY molecule.

Note in the following that once FRS2:FGF:FGFR:653P654P has formed, which is the endpoint of the ultrafast dynamics, the Michaelis–Menten phosphorylation and dephosphorylation dynamics never occur on a slower timescale than the binding dynamics. Furthermore, Michaelis–Menten kinetics constitute the large time, outer solution, of a matched asymptotic expansion (Murray, 1989) of the assumed underlying enzyme catalyzed biochemical kinetics. Thus, the binding dynamics is only ever considered to occur on a sufficiently slow timescale to allow accurate neglect of the associated, small time,

inner solution of this enzyme dynamics. Consequently, use of the Michaelis–Menten simplification for phosphorylation and dephosphorylation in the fast dynamics is internally consistent (though an assumption nonetheless).

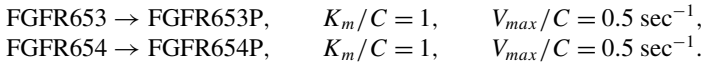
The model reactions are:

Reaction 1. FGF binds FGFR



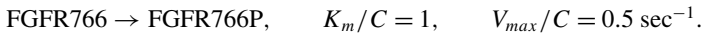
In the above K_a and K_{off} are the dimensional rate constants associated respectively with the above forward and backward reaction; similarly, for the use of K_a and K_{off} in all subsequent reactions. Using C as a concentration scale and $T = 1$ second as a timescale, we have non-dimensional parameters $K_a^{non-dim} = 8$, $K_{off}^{non-dim} = 6 \times 10^{-4}$ in a non-dimensional model.

Reaction 2. Phosphorylation. While FGFR:FGF exists:



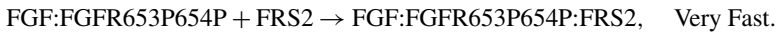
Note for these parameters Michaelis–Menten saturation effects are important for reactant levels around the concentration scale C . Using C as a concentration scale and $T = 1$ second as a timescale, we have non-dimensional parameters $K_m = 1$, $V_{max} = 0.5$. Analogous comments apply below. Thus, the results presented are valid for a two parameter family of models, obtained by varying C and T , which is a standard advantage of non-dimensionalizing.

Reaction 3. Phosphorylation. When FGF:FGFR653P654P exists then:



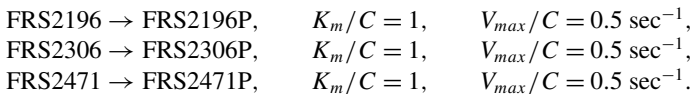
Phosphorylation at other FGFR sites does occur, though only the 766 site has downstream consequences in this model, and hence it is the only one included.

Reaction 4. FGF:FGFR653P654P binds FRS2



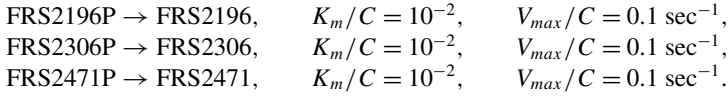
For the model, the timescale of this reaction is taken to be 0.01 seconds, which is sufficient for the reaction to be considered asymptotically fast. The model is robust to the details of such ultra-fast reactions; see Appendix C for further details.

Reaction 5. Phosphorylation of FRS2 once bound to FGFR

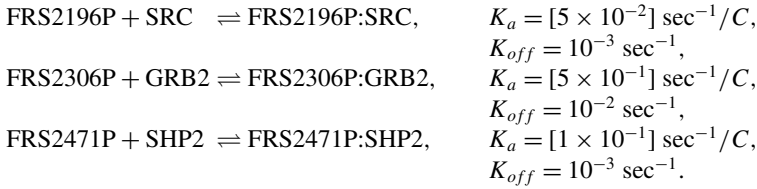


Phosphorylation does occur at other FRS2 sites, but only the above sites have downstream consequences in this model.

Reaction 6. Dephosphorylation of FRS2 when SHP2 bound to FRS2471P



Reaction 7. FRS2 effectors bind phosphorylated FRS2

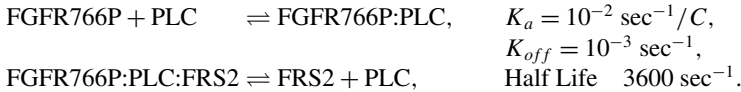


Reaction 8. FRS2196P:SRC causes the whole molecule to be relocated out



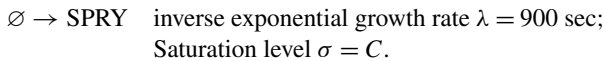
Note the symbol \emptyset refers to the fact that this relocation is, in terms of the modeling equations, equivalent to decay with no product.

Reaction 9.1. PLC Dynamics



The FRS2 produced by the latter decay (and any molecules attached to it) can subsequently influence the system if there is free phosphorylated FGFR with which the FRS2 can bind.⁵

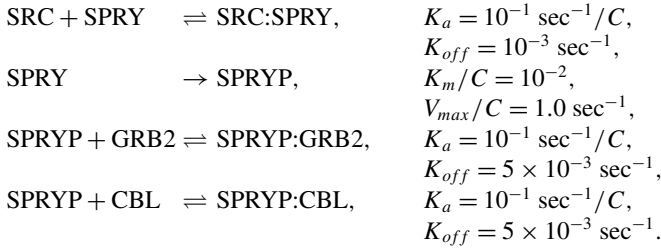
Reaction 9.2. Spry appears in time dependent manner



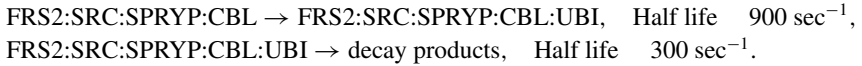
⁵To facilitate the numerical solution of the equations in the full model (by avoiding extreme stiffness once the initial ultra-fast dynamics has completed), the binding of any released FRS2 decay products to any free phosphorylated FGFR does not happen quite as fast as in reaction 4. The timescale used is nonetheless asymptotically fast compared to any other timescale present in the model by the time such decays are significant and thus this still is a very accurate approximation in the numerical scheme. Analogous comments apply for the decay products of reaction 11.

Note the “ \emptyset ” refers to the fact that this appearance of SPRY is equivalent, in terms of the modeling equations, to production with no reactant. The appearance of SPRY is modeled as $\sigma(1 - \exp[-t/\lambda])$.

Reactions 10. Spry binds SRC. Phosphorylation of SPRY. SPRYP binds other effectors.

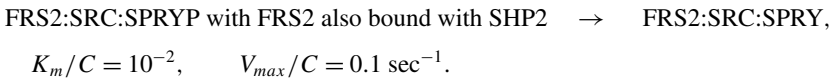


Reactions 11. SRC:SPRYP:CBL, when bound to FRS2, degrades FRS2 via ubiquitination



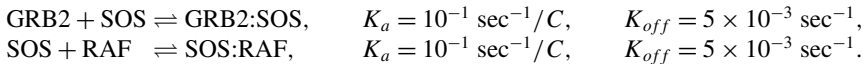
The decay products are the released chains of molecules when FRS2 is degraded. Given FRS2 is bound to phosphorylated FGFR, a phosphorylated FGFR molecule is an example of a decay product; it can subsequently bind FRS2 if the latter is available. Similarly, for a SHP2 molecule, or any chain of molecules headed by SHP2, which are bound at the phosphorylated site FRS2471P. Analogous comments apply for the binding of GRB2 to site FRS2306P and SRC at FRS2196P.

Reactions 12. SPRYP, bound to FRS2, is dephosphorylated by SHP2



The SHP2 and SPRY molecules stay bound to the FRS2 complex, while any effector which binds to SPRYP is released back into the reacting compartment on dephosphorylation.

Reactions 13. GRB2 binds SOS which binds RAF



However, in this work, we often do not explicitly include SOS or RAF to reduce the combinatorial explosion of possible molecular states.

The initial conditions

In our reference model, without knockouts, the initial conditions are

$$\text{FGF} = \text{FGFR} = \text{FRS2} = \text{PLC} = \text{SHP2} = \text{GRB2} = \text{CBL} = \text{SRC} = C,$$

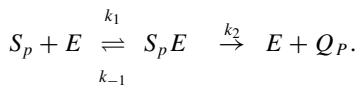
with all other concentrations zero with one possible exception: if SOS and RAF are present, as in the results presented in Fig. 10, their initial concentrations are also C .

These initial conditions reflect an implicit assumption that the key molecular players are present in approximately equal quantities. While this is a restriction, in Section 4, we motivate the observation that altering the levels of SHP2 is the key change that will have a complex and counter-intuitive impact on the system dynamics and this is investigated in detail; see Fig. 13. We also consider the effects of dramatically reducing SHP2, SRC, and CBL in the modeling of knockouts; see Figs. 6, 7, and 8, for example. Even under such extreme initial condition alterations, the model simplification procedure remains valid, clearly demonstrating that it does not require unreasonable restrictions on the initial conditions. Similarly, the qualitative behavior of the levels of GRB2 binding in the presence of knockouts is still that of an initial fast rise with a slower decay and much higher levels of GRB2 binding at site 3 rather than site 2. All other simulations performed with differing initial conditions, given significant initial levels of GRB2, FGF, FGFR, and FRS2 and at least one signal down-regulator, constituting a necessary pre-requisite for transient FGF-induced GRB2 binding, have been entirely consistent with these statements. Consequently, the above constraints on the initial conditions do not constitute overly restrictive assumptions in terms of our purpose of understanding and simplifying this system plus extracting its qualitative trends.

Appendix B: Michaelis–Menten kinetics

It is assumed that the phosphorylation and dephosphorylation reactions are driven by Michaelis–Menten kinetics, which is the large time asymptotic approximation of an enzyme-catalyzed biochemical reaction (Murray, 1989). Here we assume, for example, the phosphorylation of the PLC binding site is such that all molecules with a unphosphorylated PLC binding site are in competition for the same enzyme. We thus need to derive the appropriate large time asymptotic approximation.

Let the substrates be denoted by S_p , where p indexes substrate number. The enzyme is denoted by E , while the complex formed by substrate S_p and the enzyme is denoted by S_pE and finally the product of the enzyme action on the p^{th} substrate is denoted by Q_p . The concentration of S_pE is denoted by c_p while the concentrations of S_p , E , Q_p are s_p , e , q_p , respectively. From the standard enzyme reaction scheme for the S_p , we have⁶



Thus, for each p , we have the law of mass action ordinary differential equations

$$\frac{ds_p}{dt} = -k_1es_p + k_{-1}c_p, \quad \frac{dc_p}{dt} = k_1es_p - (k_{-1} + k_2)c_p, \quad \frac{dq_p}{dt} = k_2c_p,$$

⁶The fact the reaction rate constants are independent of p reflects an implicit assumption that the phosphorylation and dephosphorylation rates of sites are not affected by bindings elsewhere. However, the model can, and does, incorporate the fact binding molecules at one site can alter phosphorylation or dephosphorylation at other binding sites, as with the dynamics of SHP, for example. The effects of such binding molecules is taken to supplement base levels of phosphorylation or dephosphorylation. It is these base levels that are taken to be independent of other bindings and are the focus of this Appendix.

with

$$\frac{de}{dt} = \sum_p [-k_1 e s_p + (k_{-1} + k_2) c_p].$$

With the initial condition $c_p(t=0) = 0$ for all p and $e(t=0) = e_0$, we have $e = e_0 - \sum_p c_p$. In addition, one can readily find that

$$\frac{d}{dt} \left(\sum_p c_p \right) = k_1 \left(e_0 - \sum_p c_p \right) \sum_p s_p - (k_{-1} + k_2) \sum_p c_p.$$

With the usual Michaelis–Menten approximation that the complexes are in quasi-steady equilibrium (see Murray, 1989 for further details and justification), we have that

$$\sum_p c_p = \frac{e_0 \sum_p s_p}{K_M + \sum_p s_p}, \quad e = \frac{e_0 K_M}{K_M + \sum_p s_p}, \quad K_M \stackrel{\text{def}}{=} \frac{k_{-1} + k_2}{k_1},$$

and thus

$$\frac{ds_p}{dt} = -k_2 e_0 \frac{s_p}{K_M + \sum_q s_q}, \quad \frac{dq_p}{dt} = k_2 e_0 \frac{s_p}{K_M + \sum_q s_q}.$$

Hence, we see subtleties in the term arising from the Michaelis–Menten approximation when there are multiple molecular states, but each being phosphorylated by the same enzyme. In particular, the denominator must contain the sum $\sum_q s_q$. These expressions are used for all phosphorylations and dephosphorylations in the model presented in this paper. Finally, note when there is only one molecular state undergoing a reaction the indices p, q are limited to a single value, without loss of generality $q = p = 1$. Then $\sum_q s_q = s_1 = s_p$ and the above expressions reduce to the standard Michaelis–Menten approximation.

Appendix C: The ultra-fast initial dynamics of the model

In the following, we investigate the model's ultra-fast dynamics. For notational convenience, let the reactants concentrations $R_1 \dots R_{12}$ be defined by

- R_1 FGF
- R_2 FGFR, unphosphorylated
- R_3 FGF:FGFR, unphosphorylated
- R_4 FGF:FGFR, 653P
- R_5 FGF:FGFR, 654P
- R_6 FGF:FGFR, 653P654P, site 766 unphosphorylated
- R_7 FGF:FGFR, 653P654P, site 766 phosphorylated
- R_8 FGF:FGFR, 653P654P, site 766 phosphorylated, PLC bound
- R_9 FRS2
- R_{10} FRS2:FGF:FGFR, 653P654P, site 766 unphosphorylated
- R_{11} FRS2:FGF:FGFR, 653P654P, site 766 phosphorylated

R_{12} FRS2:FGF:FGFR, 653P654P, site 766 phosphorylated, PLC bound

with R_q , $q > 12$, systematically counting through all remaining molecular states. The (non-dimensionalized) equations for R_{10} , R_{11} , R_{12} are of the form

$$\begin{aligned}\dot{R}_{10} &= \frac{1}{\epsilon_1} R_6 R_9 + f_{10}(R_{10}, \dots), & \dot{R}_{11} &= \frac{1}{\epsilon_1} R_7 R_9 + f_{11}(R_{10}, \dots), \\ \dot{R}_{12} &= \frac{1}{\epsilon_1} R_8 R_9 + f_{12}(R_{10}, \dots),\end{aligned}$$

where f_{10} , f_{11} , f_{12} are functions representing all but the very fastest dynamics and $\epsilon_1 \ll 1$. The asymptotic smallness of ϵ_1 is due to the fact the timescale associated with the influence of $R_6 R_9$, $R_7 R_9$ and $R_8 R_9$ on R_{10} , R_{11} and R_{12} is much faster than any other timescale. A matched asymptotic expansion (e.g. Chapter 7 (Bender and Orszag, 1991)) reveals that the composite solutions at leading order in ϵ_1 are

$$\begin{aligned}R_{10+v} &= R_{10+v}^{outer} + \frac{1}{\epsilon_1} \int_t^\infty ds R_{6+v} R_9, \\ \text{where } \dot{R}_{10+v}^{outer} &= f_{10+v}(R_{10}, \dots), \lim_{t \rightarrow 0} R_{10+v}^{outer} = \frac{1}{\epsilon_1} \int_0^\infty ds R_{6+v} R_9, \quad v \in \{0, 1, 2\}\end{aligned}$$

and

$$R_w = R_w^{outer}, \quad \text{where } \dot{R}_w^{outer} = f_w(R_{10}, \dots), \lim_{t \rightarrow 0} R_w^{outer} = R_w^0, \quad w > 12.$$

The quantity R_w^0 is the initial condition for R_w in the full model. If we are only interested in downstream quantities, such as levels of GRB2 binding, then the only quantities from the first stage of the reaction scheme which could be important are

$$\int_0^\infty ds R_6 R_9, \quad \int_0^\infty ds R_7 R_9, \quad \int_0^\infty ds R_8 R_9,$$

as these provide initial conditions in the outer equations for R_{10} , R_{11} , R_{12} . Also, the above outer solutions are highly accurate for a sufficiently large time compared to the timescale of the ultra-fast dynamics as then

$$\int_t^\infty ds R_6 R_9, \quad \int_t^\infty ds R_7 R_9, \quad \int_t^\infty ds R_8 R_9,$$

are small. Thus, we investigate these integrals.

As we will conclude below, R_8 is effectively negligible in the ultra-fast dynamics, and thus we do not consider its dynamics explicitly in the following. The non-dimensionalized model equations for R_1 to R_7 , R_9 are as follows:

$$\begin{aligned}\dot{R}_1 &= -q_1 R_1 R_2, & \dot{R}_2 &= -q_1 R_1 R_2, \\ \dot{R}_3 &= q_1 R_1 R_2 - \frac{q_2 R_3}{K + R_3 + R_4} - \frac{q_2 R_3}{K + R_3 + R_5},\end{aligned} \tag{C.1}$$

$$\begin{aligned}\dot{R}_4 &= \frac{q_2 R_3}{K + R_3 + R_5} - \frac{q_2 R_4}{K + R_3 + R_4}, & \dot{R}_5 &= \frac{q_2 R_3}{K + R_3 + R_5} - \frac{q_2 R_5}{K + R_3 + R_5}, \\ \dot{R}_6 &= -q_3 R_6 R_9 + \frac{q_2 R_4}{K + R_3 + R_4} + \frac{q_2 R_5}{K + R_3 + R_5} - O\left(\frac{q_4}{q_3}\right), \\ \dot{R}_7 &= -q_3 R_7 R_9 + \frac{q_4 R_6}{K + R_6 + \dots} + O\left(\frac{q_2}{q_3}, \frac{q_4}{q_3}\right), & \dot{R}_9 &= -q_3(R_6 + R_7)R_9\end{aligned}$$

with the initial conditions

$$R_1 = R_1^0, \quad R_2 = R_2^0, \quad R_3 = R_4 = R_5 = R_6 = R_7 = 0, \quad R_9 = R_9^0.$$

The parameters q_1 , q_2 , $q_3 \stackrel{\text{def}}{=} \epsilon_1^{-1}$, q_4 are such that $q_2/q_3, q_4/q_3 \ll 1$. Typically $K \sim O(1)$. As we are only interested in $R_6 R_9$, $R_7 R_9$, we consider these equations only while $R_6 R_9$, $R_7 R_9$ are non-negligible, and thus we can neglect the $O(q_2/q_3)$, $O(q_4/q_3)$ terms in the above.

After solving for R_1 , R_2 the equations for $R_3 \dots R_6$ reduce to:

$$\begin{aligned}\dot{R}_3 &= \frac{q_1 R_1^0 R_2^0 (R_2^0 + R_1^0)^2}{(R_1^0 - R_2^0 e^{q_1(R_2^0 - R_1^0)t})(-R_2^0 + R_1^0 e^{-q_1(R_2^0 + R_1^0)t})} - \frac{2q_2 R_3}{K + R_3 + R_5}, \\ \dot{R}_5 &= \frac{q_2 R_3}{K + R_3 + R_5} - \frac{q_2 R_5}{K + R_3 + R_5}, \\ \dot{R}_6 &= -q_3 R_6 R_9 + \frac{2q_2 R_5}{K + R_3 + R_5} - O\left(\frac{q_4}{q_3}\right),\end{aligned}\tag{C.2}$$

with $R_4 = R_5$.

Solving an equation of the form

$$\begin{aligned}\frac{dx}{dt} &= -q_3 g(t)x + f(t), \quad q_3 \gg 1, \quad q_3 g(t) \gg 1, \quad g(t) > 0, \\ \frac{dg}{dt} &> 0, \quad |df/dt| \ll q_3 |dg/dt|\end{aligned}$$

and implementing Watson's lemma (e.g. Chapter 6 (Bender and Orszag, 1991)) reveals that the leading order solution, in an expansion with respect to q_3 , is given by $x = f(t)/[q_3 g(t)]$. Thus, while $q_3 R_9 \gg 1$, we have the quasi-steady state hypothesis holds at leading order in $q_3^{-1} \ll 1$ for the R_6 equation. Hence, the flux into the R_{10} compartment is simply

$$q_3 R_6 R_9 = \frac{2q_2 R_5}{K + R_3 + R_5}$$

at leading order. Note this implies that $R_5 \gg R_6$, $R_6 \ll 2q_2$, and hence $q_3 R_6 R_9 \gg q_3 R_7 R_9$ (at least for $q_3 R_9$ sufficiently large). Thus, the total flux into the R_{11} compartment is sub-leading compared to the total flux into the R_{10} compartment. Hence, the dominant influence of the early stage dynamics on the rest of the model is via

$$\int_0^\infty q_3 R_6 R_9.$$

Given the effects of R_7 are thus essentially negligible in the ultra-fast dynamics, even on neglect of the production of R_8 from R_7 , the levels of R_8 must also be negligible in the ultra-fast dynamics.

If the initial levels of R_9 , i.e. FRS2, is sufficiently small, then levels of FRS2 will limit the reaction and we have $R_9(\infty) = 0$, $R_6(\infty) > 0$. Integrating the equations in (C.2) over the positive real line and utilizing the above leading order approximations then yields

$$\int_0^\infty dt q_3 R_6 R_9 \sim R_9^0.$$

Conversely, for abundant FRS2, we have $R_9(\infty) > 0$, $R_6(\infty) = 0$. Integrating Eqs. (C.2) over the positive real line then quickly reveals that

$$\int_0^\infty dt q_3 R_6 R_9 \sim \int_0^\infty dt \frac{q_1 R_1^0 R_2^0 (R_2^0 + R_1^0)^2}{(R_1^0 - R_2^0 e^{q_1 (R_2^0 - R_1^0)t}) (-R_2^0 + R_1^0 e^{-q_1 (R_2^0 + R_1^0)t})} = \min(R_1^0, R_2^0).$$

Note these results, though contingent on leading order asymptotic approximations, respect conservation principles exactly, i.e. the molecules that enter the initial reaction phase leave the initial reaction phase.

There is sufficient R_9 to prevent $R_9(\infty) = 0$ if $R_9^0 > \min(R_1^0, R_2^0)$. Hence, the above can be summarized by

$$\int_0^\infty dt q_3 R_6 R_9 \sim \min(R_1^0, R_2^0, R_9^0).$$

Thus, we have the immediate modeling prediction that the amount of FGF:FGFR653P654P:FRS2 produced in the initial stages of the reaction scheme is simply dependant on the initial levels of FGF, FGFR, and FRS2 at leading order in $\epsilon_1 \ll 1$. Neither the details of the phosphorylation and dephosphorylation Michaelis–Menten dynamics nor how the constituents of the early reaction phase interact are important for the initial dynamics. We also have the prediction that solving for R_k^{outer} , $k \geq 10$ with the above initial conditions will give accurate predictions for sufficiently large timescales compared to the timescales of the initial, ultra-fast dynamics. Thus, the initial dynamics effectively decouple.

We have seen the robustness of the model's predictions for the formation of FGF:FGFR653P654P:FRS2, given FRS2 binds to FGF:FGFR653P654P but not FGF:FGFR653P, FGF:FGFR654P or FGF:FGFR. However, consider a model which allowed these possibilities with similar reaction rates. We would still anticipate that the amount of FGF:FGFR bound to FRS2 in the initial dynamics again simply depends on the initial levels of FGF, FGFR and FRS2. Thus, the modeling predictions would be unaltered on all but the very fastest of timescales, thus motivating our simplification of the initial dynamics.

References

- Adachi, T., Cui, C., Kanda, A., Kayaba, H., Ohta, K., Chihara, J., 2004. Activation of epidermal growth factor receptor via CCR3 in bronchial epithelial cells. *Biochem. Biophys. Res. Commun.* 320, 292–296.

- Agazie, Y.M., Hayman, M.J., 2003. Molecular mechanism for a role of SHP2 in epidermal growth factor receptor signaling. *Mol. Cell. Biol.* 23, 7875–7886.
- Aldridge, B.B., Burke, J.M., Lauffenburger, D.A., Sorger, P.K., 2006. Physicochemical modelling of cell signalling pathways. *Nat. Cell Biol.* 8, 1195–1203.
- Augoff, K., Tabola, R., Kula, J., Gosk, J., Rutowski, R., 2005. Epidermal growth factor receptor (EGF-R) in dupuytren's disease. *J. Hand Surg.: J. Br. Soc. Surg. Hand* 30, 570–573.
- Basilico, C., 2005. Editorial: Introduction to the special FGF issue. *Cytokine Growth Factor Rev.* 16, 105–106.
- Bender, C.M., Orszag, S.A., 1991. *Advanced Mathematical Methods for Scientists and Engineers: Asymptotic Methods and Perturbation Theory*. Springer, New York.
- Billingham, J., King, A.C., 2000. *Wave Motion*. Cambridge University Press, Cambridge.
- Borisov, N.M., Markevich, N.I., Hoek, J.B., Kholodenko, B.N., 2005. Signaling through receptors and scaffolds: Independent interactions reduce combinatorial complexity. *Biophys. J.* 89, 951–966.
- Borisov, N.M., Markevich, N.I., Hoek, J.B., Kholodenko, B.N., 2006. Trading the micro-world of combinatorial complexity for the macro-world of protein interaction domains. *Biosystems* 83, 152–166.
- Bublit, E.M., Yarden, Y., 2007. The EGF receptor family: Spearheading a merger of signalling and therapeutics. *Curr. Opin. Cell Biol.* 19, 124–134.
- Chen, P., Xie, H., Wells, A., 1996. Mitogenic signaling from the EGF receptor is attenuated by a phospholipase C-gamma/protein kinase C feedback mechanism. *Mol. Biol. Cell* 7, 871–881.
- Citri, A., Yarden, Y., 2006. EGF-ErbB signalling: Towards the systems level. *Nat. Rev. Mol. Cell Biol.* 7, 505–516.
- Crampin, E.J., Schnell, S., McSharry, P.E., 2004. Mathematical and computational techniques to reduce complex biochemical reaction mechanisms. *Prog. Biophys. Mol. Biol.* 86, 77–112.
- Dailey, L., Ambrosetti, D., Mansukhani, A., Basilico, C., 2005. Mechanisms underlying differential responses to FGF signaling. *Cytokine Growth Factor Rev.* 16, 233–247.
- Eswarakumar, V.P., Lax, I., Schlessinger, J., 2005. Cellular signaling by fibroblast growth factor receptors. *Cytokine Growth Factor Rev.* 16, 139–149.
- Gavutis, M., Jaks, E., Lamken, P., Piehler, J., 2006. Determination of the two-dimensional interaction rate constants of a cytokine receptor complex. *Biophys. J.* 90, 3345–3355.
- Greenman, C., Stephens, P., Smith, R., Dalgliesh, G.L., Hunter, C., Bignell, G., Davies, H., Teague, J., Butler, A., Stevens, C., et al., 2007. Patterns of somatic mutation in human cancer genomes. *Nature* 446, 153–158.
- Grose, R., Dickson, C., 2005. Fibroblast growth factor signaling in tumorigenesis. *Cytokine Growth Factor Rev.* 16, 179–186.
- Hanafusa, H., Torii, S., Yasunaga, T., Matsumoto, K., Nishida, E., 2004. Shp2, an SH2-containing protein-tyrosine phosphatase, positively regulates receptor tyrosine kinase signaling by dephosphorylating and inactivating the inhibitor sprouty. *J. Biol. Chem.* 279, 22992–22995.
- Heath, J.K., Kwiatkowska, M.Z., Norman, G., Parker, D., Tymchyshyn, O., 2006. Probabilistic model checking of complex biological pathways. In: *Proc. Computational Methods in Systems Biology CMSB'06*.
- Heath, J., Kwiatkowska, M., Norman, G., Parker, D., Tymchyshyn, O., 2007. Probabilistic model checking of complex biological pathways. *Theor. Comput. Sci. (Special Issue on Converging Sciences: Informatics and Biology)*.
- Hucka, M., Finney, A., Sauro, H.M., Bolouri, H., Doyle, J.C., Kitano, H., Arkin, A.P., Bornstein, B.J., Bray, D., et al., 2003. (SBML Forum). The systems biology markup language (SBML): a medium for representation and exchange of biochemical network models. *Bioinformatics* 19, 524–531.
- Jarvis, L.A., Toering, S.J., Simon, M.A., Krasnow, M.A., Smith-Bolton, R.K., 2006. Sprouty proteins are in vivo targets of Corkscrew/SHP-2 tyrosine phosphatases. *Development* 133, 1133–1142.
- Julien, S.G., Dube, N., Read, M., Penney, J., Paquet, M., Han, Y., Kennedy, B.P., Muller, W.J., Tremblay, M.L., 2007. Protein tyrosine phosphatase 1B deficiency or inhibition delays ErbB2-induced mammary tumorigenesis and protects from lung metastasis. *Nat. Genet.* 39, 338–346.
- Kiyatkin, A., Aksamitiene, E., Markevich, N.I., Borisov, N.M., Hoek, J.B., Kholodenko, B.N., 2006. Scaffolding protein Grb2-associated binder 1 sustains epidermal growth factor-induced mitogenic and survival signaling by multiple positive feedback loops. *J. Biol. Chem.* 281, 19925–19938.
- Kramer, S., Okabe, M., Hacohen, N., Krasnow, M.A., Hiromi, Y., 1999. Sprouty: A common antagonist of FGF and EGF signaling pathways in *Drosophila*. *Development* 126, 2515–2525.
- Krueger, G., Ellis, C.N., 2005. Psoriasis-recent advances in understanding its pathogenesis and treatment. *J. Am. Acad. Dermatol.* 53(Supplement 1), S94–S100.

- Kudryavtsev, A.B., Jameson, R.F., Linert, W., 2001. *The Law of Mass Action*, Springer Statistical Thermodynamics. Springer, New York.
- Kwiatkowska, M.Z., Norman, G., Parker, D., Tymchyshyn, O., Heath, J.K., Gaffney, E., 2006. Simulation and verification for computational modelling of signalling pathways. In: Perrone, L.F., Wieland, F.P., Liu, J., Lawson, B.G., Nicol, D.M., Fujimoto, R.M. (Eds.), *Proc. Winter Simulation Conference 2006 (WSC'06)*, ISBN: 1-4244-0501-7.
- Li, X., Brunton, V.G., Burgar, H.R., Wheldon, L.M., Heath, J.K., 2004. FRS2-dependent SRC activation is required for fibroblast growth factor receptor-induced phosphorylation of Sprouty and suppression of ERK activity. *J. Cell Sci.* 117, 6007–6017.
- Millat, T., Bullinger, E., Rohwer, J., Wolkenhauer, O., 2007. Approximations and their consequences for dynamic modelling of signal transduction pathways. *Math. Biosci.* 207, 4057.
- Mohammadi, M., Olsen, S.K., Ibrahimi, O.A., 2005. Structural basis for fibroblast growth factor receptor activation. *Cytokine Growth Factor Rev.* 16, 107–137.
- Murray, J.D., 1989. *Mathematical Biology*. Springer, New York.
- Schlessinger, J., 2000. Cell signaling by receptor tyrosine kinases. *Cell* 103, 211–225.
- Schoeberl, B., Eichler-Jonsson, C., Gilles, E.D., Muller, G., 2002. Computational modeling of the dynamics of the MAP kinase cascade activated by surface and internalized EGF receptors. *Nat. Biotechnol.* 20, 370–375.
- Sevecka, M., MacBeath, G., 2006. State-based discovery: A multidimensional screen for small-molecule modulators of EGF signalling. *Nat. Methods* 3, 825–831.
- Suckley, R., Biktashev, V.N., 2003. Comparison of asymptotics of heart and nerve excitability. *Phys. Rev. E* 68, 011902.
- Thisse, B., Thisse, C., 2005. Functions and regulations of fibroblast growth factor signaling during embryonic development. *Dev. Biol.* 287, 390–402.
- Tikhonov, A.N., 1952. Systems of differential equations containing a small parameter in the derivatives. *Mat. Sb.* 31, 575–586.
- Tyson, J.J., Chen, K., Novak, B., 2001. Network dynamics and cell physiology. *Nat. Rev. Mol. Cell Biol.* 2, 908–916.
- Vallabhajosyula, R.R., Sauro, H.M., 2006. Complexity reduction of biochemical networks. In: Perrone, L.F., Wieland, F.P., Liu, J., Lawson, B.G., Nicol, D.M., Fujimoto, R.M. (Eds.), *Proceedings of the 2006 Winter Simulation Conference*, ISBN: 1-4244-0501-7.
- Wilkie, A.O.M., 2005. Bad bones, absent smell, selfish testes: The pleiotropic consequences of human FGF receptor mutations. *Cytokine Growth Factor Rev.* 16, 187–203.
- Wolkenhauer, O., Ullah, M., Kolch, W., Cho, K., 2004. Modelling and simulation of IntraCellular dynamics: Choosing an appropriate framework. *IEEE Trans. Nanobiosci.* 3, 200–207.
- Wong, E.S., Lim, J., Low, B.C., Chen, Q., Guy, G.R., 2001. Evidence for direct interaction between sprouty and Cbl. *J. Biol. Chem.* 276, 5866–5875.
- Yamada, S., Taketomi, T., Yoshimura, A., 2004. Model analysis of difference between EGF pathway and FGF pathway. *Biochem. Biophys. Res. Commun.* 314, 1113–1120.
- Yarden, Y., Sliwkowski, M.X., 2001. Untangling the ErbB signalling network. *Nat. Rev. Mol. Cell Biol.* 2, 127–137.