Chapter 6
Kinetic Models of Biochemical Signaling Networks

Mehdi Bouhaddou and Marc R. Birtwistle

Abstract Kinetic models of biochemical signaling networks are a mechanistic description of pharmacodynamics, and thus are potentially well-poised to fill gaps in the drug development pipeline by: (i) allowing putative drugs to be tested via simulations for efficacy and safety before expensive experiments and failed clinical trials; (ii) providing a framework for personalized and precision medicine that incorporates genomic information into a prediction of drug action in an individual; and (iii) interfacing with traditional pharmacokinetic models to yield computable yet mechanistic simulations that can inform drug dosing and frequency. However, biochemical signaling networks are currently incompletely understood on a basic level and are extremely complex compared to traditional applications of kinetic modeling. Herein, we describe current methods used to build such models and highlight strengths and weaknesses of the various approaches, as well as identify areas that need more research to drive the field towards influencing these important potential applications.

Keywords Crosstalk · Input signal · Feedforward · Feedback · Perturbations · Enhanced pharmacodynamic (ePD) models · Occam's razor · Reaction rate laws · Michaelis-Menten equations · Microdomains · Parameter values · Ordinary differential equation (ODE) · Stoichiometries

Electronic supplementary material The online version of this article (doi:10.1007/978-3-319-44534-2_6) contains supplementary material, which is available to authorized users.

M. Bouhaddou · M.R. Birtwistle (✉)
Department of Pharmacology and Systems Therapeutics, Icahn School of Medicine at Mount Sinai, One Gustave L. Levy Place, Box 1603, New York, NY 10029, USA
e-mail: marc.birtwistle@mssm.edu

D.E. Mager and H.H.C. Kimko (eds.), *Systems Pharmacology and Pharmacodynamics*, AAPS Advances in the Pharmaceutical Sciences Series 23,
DOI 10.1007/978-3-319-44534-2_6

6.1 Introduction

Kinetic models are those that represent the dynamics of a system in response to perturbation, and are almost ubiquitously quantitative. This chapter focuses on how one builds a quantitative kinetic model to describe the coupled chemical reactions that together dictate how cells respond dynamically to a perturbation, such as treatment with a drug. These collections of coupled chemical reactions are often called biochemical signaling networks, or signal transduction networks. The word network should be stressed, because although the historic notion of linear signaling pathways has allowed us to understand basic routes of biochemical signaling, it is becoming clear that signaling pathways in cells rarely operate linearly in isolation, but rather are highly interconnected with feedforward and feedback loops and exhibit significant crosstalk (Kholodenko et al. 2010).

Why should one care about kinetic models of biochemical signaling networks? They are useful, if not essential tools in understanding and predicting how perturbations to biochemical signaling networks influence cell behavior. The validity of this assertion is not clear if one still views signaling in terms of linear pathways without feedforward and feedback loops or crosstalk. However, when one embraces these ubiquitous features of signaling systems, the importance of kinetic models as a tool to predict quantitative signaling behavior becomes lucent. As a simple example, consider the simple biochemical network shown in Fig. 6.1, where an input signal S is controllable and activation of D leads to a cell fate. In this simple yet common "incoherent feedforward loop" network, the input signal S leads to activation of A and C, but A activates D while C represses D. What will the cell fate be if S is increased? If the levels of C were high, then increasing S would decrease D, and if they were low, the opposite. Further compounding this seemingly simple question are the spatiotemporal dynamics of A and C activation. If A is localized with D but C is not, then D would go up, but if C is localized with D and A is not, the opposite. Moreover, if C affects D more quickly than does A, D would go down then up, and if C affects D more slowly than does A, vice versa. Thus, even in this idealized example, *quantitative knowledge of network spatiotemporal dynamics is needed to predict cell fate*. Qualitative knowledge for maps such as those in Fig. 6.1 is currently abundant, but the quantitative knowledge needed to predict behavior in response to perturbations is scarce. This quantitative predictability problem is amplified by the overwhelming complexity of real biochemical networks, in which a large number of species are interconnected by a multitude of feedforward and feedback regulatory motifs. Kinetic models of biochemical signaling networks have properties that are suitable for dealing with these types of problems.

This kind of quantitative understanding of signaling dynamics is at the heart of much current signal transduction research. If one understands the relationship between biochemical network behavior and cell fate, it becomes possible to answer a related but critically important question: how can one manipulate a biochemical network to control cell fate? The ability to answer this second question has seemingly countless potential applications. For instance, understanding how to

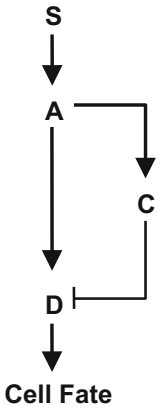

Fig. 6.1 A simple incoherent feedforward motif

make stem cells migrate, differentiate, and proliferate as needed could potentially lead to cures for several hitherto untreatable diseases, and understanding how to perturb deregulated signaling networks in cancer cells such that they cease proliferation and migration could lead to novel cancer treatments. Improving the efficiency of the drug development pipeline is critical as the cost of drug development rises while success rates fall (Birtwistle et al. 2013; Returns on R&D investments continue to fall 2014), and the ability of kinetic models to predict the effects of putative new drugs or drug combinations, both for efficacy and toxicity, can place a meaningful computational layer into the pharmaceutical drug development pipeline that can help prevent costly failed clinical trials (Schoeberl et al. 2009). Moreover, if the model is grounded in biochemical mechanisms, as we advocate for in this chapter, then adapting the model to different patients based on their particular genomic characteristics may be straightforward. In this regard, kinetic models of biochemical signaling networks have been referred to as enhanced pharmacodynamic (ePD) models (Iyengar et al. 2012). The ability of ePD models to cleanly interface with traditional pharmacokinetic models can also help inform drug dosing and scheduling (Zhang et al. 2014). These potential applications poise kinetic modeling of biochemical signaling pathways to have significant impact in the developing disciplines of systems and personalized pharmacology.

6.2 Building Kinetic Models of Biochemical Signaling Networks

Figure 6.2 illustrates the general process that one would follow to develop a kinetic model of a biochemical signaling network. It begins with a question or problem-of-interest, and ends with a validated model that describes the relevant system

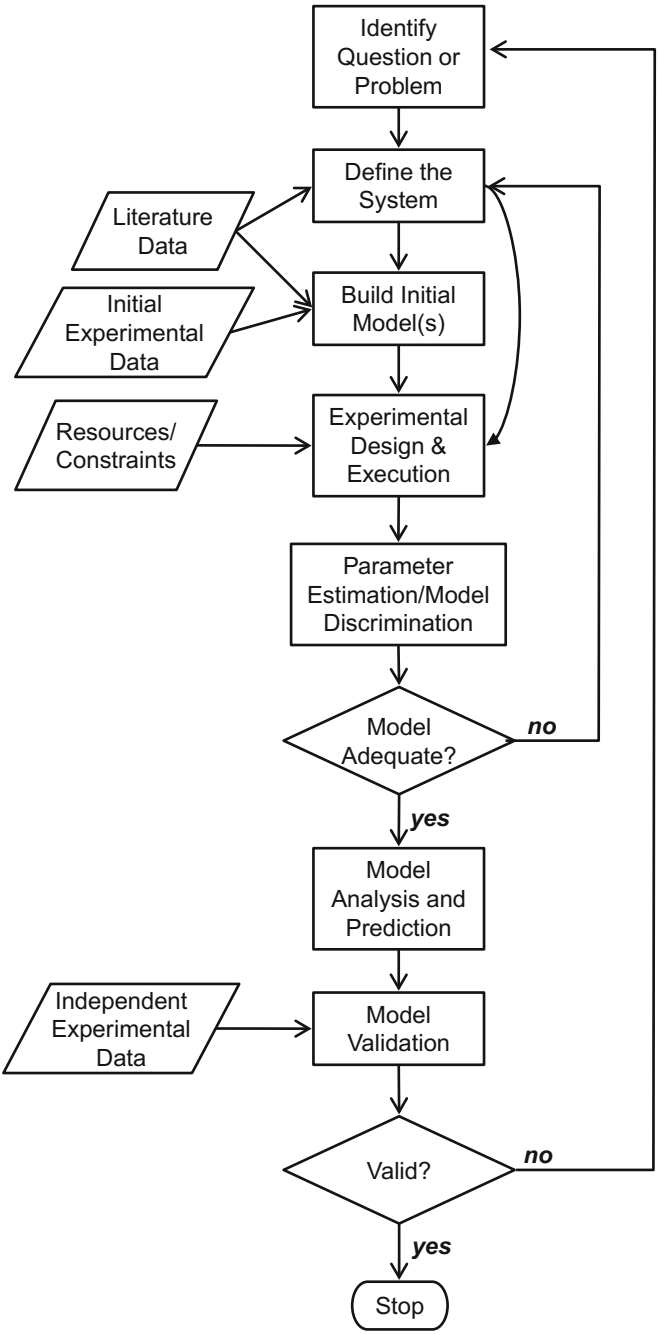

Fig. 6.2 An overall workflow for building kinetic models of biochemical signaling networks. The various aspects of the process are described throughout this chapter in sequence. *Box*—process to be carried out, *trapezoid*—input information or data, *diamond*—decision to be made

behavior and answers the initial question in a meaningful way. This process is not unique to the models we consider in this chapter but is instead rather general to many types of modeling. Notably, it depicts an ideal situation; in practice typically many of these steps are either not performed or do not follow rigorous protocols. Models that are not developed via this ideal process nevertheless are often quite useful and have had impact in systems biology research. Yet, applying all the steps of this pipeline would certainly improve the final developed model and would help to standardize the currently widely-varying protocols employed to develop these models. Such protocol standardization would help to boost confidence in the fidelity of this class of models, as well as increase model sharing and reuse. This is of particular importance if the goal of one's modeling efforts is to inform preclinical or clinical pharmaceutical development in an industrial or clinical setting.

In what follows, we expand on each step of this process, surveying current methods and approaches as well as illustrating with real examples when possible. Also, although we give an inherent bias towards mammalian systems and methods, as the focus of this chapter is pharmacology, many of the presented methods are general and can be applied to a multitude of biological systems.

6.2.1 Identify the Question or Problem

The process begins by identifying the question or problem-of-interest. This step is absolutely critical to embarking on a well-posed modeling exercise that is likely to yield meaningful research results, and is analogous to starting a research project with a well-founded and experimentally-testable hypothesis. The question should (i) rely on or be enhanced by kinetic modeling for providing an answer and (ii) provide focus for the scope and granularity of the model when choosing from among the many alternatives to be discussed below.

Throughout this Chapter, we will illustrate some (but not all) of the methods with a case study focused on the role of Kinase Suppressor of Ras 2 (KSR2) in extracellularly-regulated kinase 1 (ERK1) and ERK2 signaling mediated by the B-isoform of rapidly accelerated fibrosarcoma (B-Raf) (Brennan et al. 2011). The ERK1/2 signaling pathway plays a central role in a variety of cellular processes, including migration, differentiation, and death (Yoon and Seger 2006). B-Raf is commonly mutated in a variety of human cancers. A well-documented B-Raf mutation is V600E, which is frequent in melanoma and constitutively activates the kinase activity of B-Raf, leading to inappropriate ERK1/2 activation and cell proliferation, survival, and migration (Dankort et al. 2009). A targeted small molecule inhibitor of only B-Raf V600E, but not wild-type B-Raf or C-Raf (vemurafenib) was recently developed with structure-based approaches, and has

showed a remarkable 80 % response rate in melanoma patients with the V600E mutation (although the tumor regression only lasted 2–18 months presumably due to adaptive resistance) (Sondergaard et al. 2010; Das Thakur et al. 2013; Bollag et al. 2010). However, not all Raf inhibitor drugs have worked out so well. Pan small molecule inhibitors of Raf family proteins, such as GDC-0879 and PLX4720 that target mutated and non-mutated B-Raf as well as the C-Raf isoform, have been developed but have seen relatively limited clinical success (Hatzivassiliou et al. 2010; Poulikakos et al. 2010). This may be due to the fact that, in some cases, these inhibitors can actually have paradoxical activating effects on Raf-family proteins by inducing their dimerization (Poulikakos et al. 2010) and depend on the mutational status of the upstream Ras family proteins (Hatzivassiliou et al. 2010). A new and potentially critical mechanism has been shown for how the scaffold protein KSR2 operates in the context of mediating B-Raf activation and signaling to ERK1/2 (Brennan et al. 2011), and it may be feasible to design drugs that disrupt KSR2's ability to activate B-Raf-mediated ERK1/2 signaling. What oncogenic mutations might make a particular tumor sensitive and/or resistant to such a KSR2 drug? What other drugs, such as Raf inhibitors, may synergize with a KSR2 drug? What affinity must a KSR2 drug have to exert ERK1/2 pathway inhibition, and what are the most effective mechanisms to target? A kinetic model of the ERK1/2 signaling pathway incorporating these new KSR2 mechanisms may help answer such questions.

6.2.2 Define the System

Defining the system forms the foundation for the entire modeling process, yet it is very difficult to systematize and embodies what many refer to as "the art of modeling". Many choices and assumptions must be made that are based more on intuition rather than rigor and mathematics. Yet, one guiding principal that almost universally holds true is that of Occam's Razor, which suggests that when confronted with many possibilities, one should choose the simplest that describes all desired features. Thus, we prefer to start simple with relatively restrictive models, and only expand the model to be more complex when warranted by inability to describe experimental data relevant to the question-of-interest. At this stage it is prudent to consider the experimental system as well (e.g., cell lines vs. animal models; which lines or animals, etc.), although this could also be considered at the experimental design phase.

6.2.2.1 Inputs and Outputs

The question-of-interest should inform a clear choice for the input(s) and output(s) of the model. The input should be experimentally perturbable, the output should be

experimentally observable, and ideally vice versa. For our example, we choose Ras as the input and ERK1/2 as the output. Ras is a small G-protein that sits at the top of the B-Raf-ERK1/2 signaling cascade. Activation of Ras leads to B-Raf and ERK1/2 activation. ERK1/2 activation is a key regulator of proliferation, and its aberrant activation in cancer is often a driver of unregulated growth.

6.2.2.2 Connecting the Inputs and Outputs with a Kinetic Scheme

The kinetic scheme is a precise pictorial representation of the chemical transformations that are assumed to occur between the model inputs and outputs. It is the foundation and a central part of the model. There have been some formalisms proposed to depict kinetic schemes, such as those suggested by CellDesigner (Kitano et al. 2005), but they have yet to be widely adopted. Nevertheless, there are a few basic features that are widely understood (see Fig. 6.3 for examples): (i) binding of one species to another to form a complex is depicted by two arrows smoothly coming together into one arrow, and dissociation as the reverse; (ii) enzymatic transformation is depicted by a curved unidirectional arrow from substrate to product, with a straight unidirectional arrow from the enzyme to the nadir of the curved arrow; (iii) degradation is depicted by chemical transformation to the empty set $\varnothing$; (iv) synthesis is depicted by chemical transformation from nothing. Of course, there are many variations on this theme, but these generalities will help one to understand many kinetic schemes.

Deciding on reactions to connect the inputs and outputs is again part of the art of modeling. Moving backwards from the output to the input is usually a reasonable method. How does my output get produced? What biochemical entities are needed to do that? How does this mechanism move me one step closer to my input? By repeatedly answering these questions, one can arrive at a list of species and reactions that are likely important to connect the model inputs and outputs. It will also suggest when "dead-ends" are present and thus suggest species and reactions that may be culled.

We opt to build "mechanistic" models when possible, meaning that if an elementary reaction mechanism is known, then we prefer to include it in the model explicitly. This is desirable as the model is grounded by its clear connection to real biochemical entities and mechanisms, and as such is usually easier to compare to experimental data than a model that is not mechanistic. Most importantly, a mechanistic model can be quite adaptable to different cell systems or contexts (Bouhaddou and Birtwistle 2014) because its parameters and species have explicit biochemical meaning and can be changed by measurements in the new system. This is in contrast to empirical models that may be very good at describing the behavior of a particular system, but are not easily adaptable to other systems because the parameters do not have biochemical meaning. Yet, more often than not in biochemical signaling networks, uncertainty is abundant, and mechanisms are not known. This necessitates semi-mechanistic or empirical approaches, at least for connecting those species where mechanisms are not known. Approaches based on

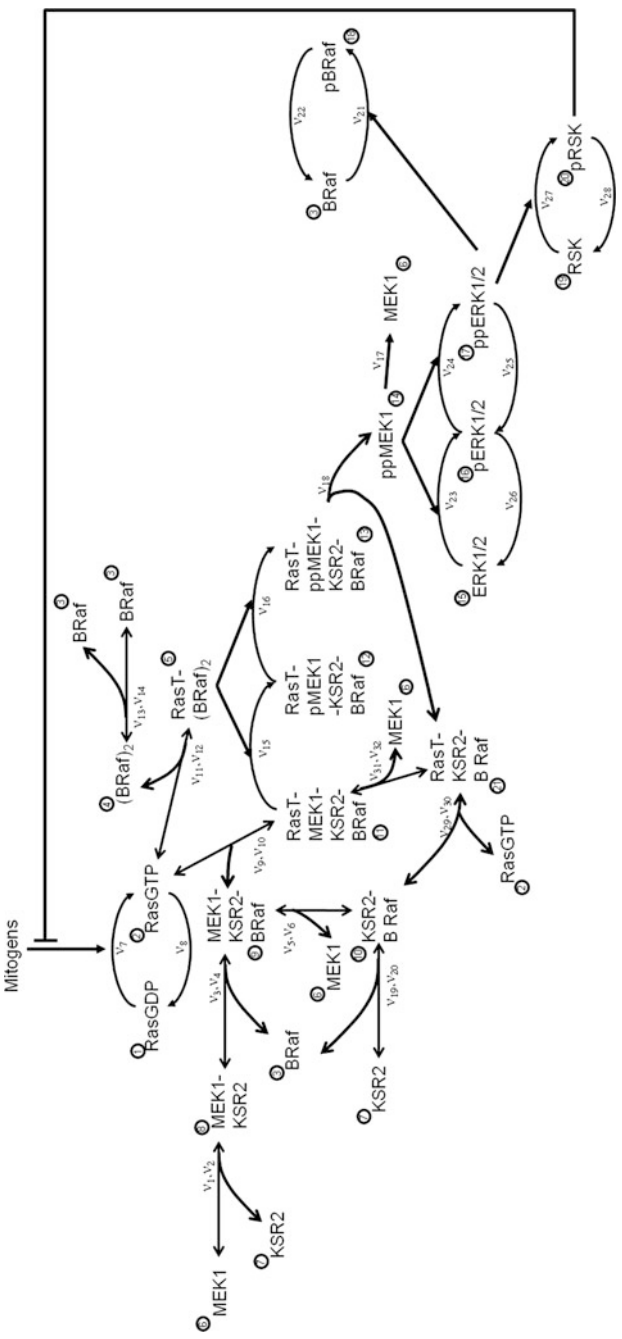

Fig. 6.3 Kinetic scheme for our B-Raf/KSR2 model example. Reactions are chosen as described throughout the text, and the equations underlying this model are given in the supplementary material in the MATLAB code. *Circled numbers* correspond to species indices, and reaction rates are *labeled* next to their *arrows*

fuzzy logic, Bayesian inference, and quantitative logic gates have had success in such endeavors (Morris et al. 2011; Kirouac and Onsum 2013; Sachs et al. 2005), as well as those based on modular response analysis (Klinger et al. 2013) or similar perturbation-centric methods (Molinelli et al. 2013). Moreover, the mechanistic links between signaling kinetics and cell fate, such as apoptosis, are sometimes not clear; in such cases regression based methods, such as partial least squares regression, have been shown to work well (Miller-Jensen et al. 2007; Janes et al. 2005). Useful models will likely be a mix of mechanistic with empirical functions when necessary.

Sometimes, explicitly modeling all known biochemical mechanisms is overly burdensome or not relevant to the question of interest. For example, often times in models of kinase signaling pathways, ATP is not explicitly modeled. Of course, ATP is important for kinase signaling; yet, its concentration almost never changes in the cell and its binding to enzyme is typically very fast and does not appreciably influence the overall reaction kinetics. Thus, one must exercise some restraint to balance the costs vs. the benefits of incorporating mechanistic detail. Without this restraint, one ends up with an extremely large model that becomes computationally impractical to apply the downstream modeling process steps. Thus, we strongly prefer parsimony as a first assumption over complexity.

The initial kinetic scheme we have chosen to depict our KSR2/B-Raf model is shown in Fig. 6.3. Note that we have also indicated reaction rate numbers and species indices; this greatly facilitates model development and its understanding by others. As stated above, we describe scheme development starting with the output and work upstream. ERK1 and ERK2 are doubly phosphorylated and activated by active MEK1 or MEK2; however, only MEK1 has been sufficiently studied to be included in the model (Brennan et al. 2011). MEK1 is activated by double phosphorylation as well, and it is claimed that only MEK1, which is bound to the KSR2-B-Raf complex, can be doubly phosphorylated by a different B-Raf molecule (Brennan et al. 2011). We assume that active GTP-bound Ras serves to recruit these MEK1-KSR2-B-Raf complexes (presumably through the Ras binding domain of B-Raf) and catalytically active B-Raf dimers to the plasma membrane so that the B-Raf dimers can phosphorylate and activate MEK1. We assume that KSR2, MEK1, and B-Raf can bind in any order to one another, but that B-Raf is needed for binding to RasGTP. As mentioned above, we choose Ras as our model input, which can be in two states, the inactive GDP or active GTP bound state. The conversion of RasGDP to RasGTP is mediated by guanine exchange factors, whose activity are typically regulated by external factors such as mitogens. It is well established that active, doubly phosphorylated ERK1/2 (ppERK1/2) enacts strong negative feedback upstream by directly phosphorylating and inactivating C-Raf, and possibly to a lesser extent B-Raf (Sturm et al. 2010; Fritsche-Guenther et al. 2011; Pratilas et al. 2009). Another well-established mechanism of ERK1/2-mediated feedback, which is thought to travel through p90 ribosomal S6 kinase (RSK), is dampening of Ras

activation by inhibiting the GEFs (von Kriegsheim et al. 2009). Negative feedback is important for dictating input/output responses and drug sensitivity/resistance, and therefore we include them since one of the modeling goals is to investigate drug sensitivity/resistance.

Inspection of all the various association and dissociation reactions presented in Fig. 6.3 highlights that it is often difficult to allow for all combinations of biochemical transformations in a rigorous manner (i.e., it is easy to leave some out by mistake). Furthermore, consider how difficult it would be to incorporate C-Raf into this model, which can heterodimerize with B-Raf and can also bind many of the same targets as B-Raf. Yet, incorporating C-Raf would likely be important to our question-of-interest, since it is the target of many current Raf drugs and influences their effectiveness (Sullivan and Flaherty 2013; Heidorn et al. 2010). This "combinatorial complexity" arises when model species have multiple sites that can each be in many states, and arises not only from complex binding scenarios (as described here), but more commonly from multisite modification such as phosphorylation. For example, the epidermal growth factor receptor has approximately 10 tyrosines that can each be unphosphorylated, phosphorylated, or bound to various downstream adaptor proteins; accounting for all these combinations yields 3^{10} possible species. This is clearly infeasible to draw on a kinetic scheme (and code) (Birtwistle 2014). An innovative solution to building models that account for such combinatorial complexity is "rule-based modeling" (Chylek et al. 2013; Sorokina et al. 2013). BioNetGen and its derivatives, such as NFsim, are the more commonly used variants (Sneddon et al. 2011). Instead of specifying a detailed kinetic scheme of all unique chemical species, one only lists the domains and states of each molecule and a handful of reaction rules that describe interactions between these domains and states, which the software then iterates over to generate all potential chemical species. Although attractive-in-principle, rule-based approaches tend to generate extremely large models (sometimes even of infinite size if there are ring-forming or polymerization reactions), which practically limits their applicability and sometimes necessitates assumptions to reduce model size so that they are computationally feasible. As computational power grows and rule-based algorithms improve, such approaches will become more attractive. Nevertheless, whether the kinetic scheme is depicted with the traditional enumeration of unique chemical species or by reaction rules, the modeling steps illustrated in Fig. 6.2 and discussed below still generally apply.

6.2.2.3 System Properties

Whereas the kinetic scheme is a quite comprehensive summary of all the biochemical species and reactions being considered, there are still several key

properties of the system to specify beyond that. Below we list these properties and describe rationale for their selection:

1. Single Cell or Cell Population Average: Are the experimental data to be gathered on a single cell level, or from millions of cells combined into a single measurement? Are the questions-of-interest relevant to behavior of individual cells or a population of cells?
2. Spatial or Compartmental: Do possible diffusional effects or spatial gradients of biochemical species play a role? Can various subcellular compartments be viewed as effectively well-mixed? How many cellular compartments are needed?
3. Stochastic or Deterministic: Are any molecular species present in low copy number (~ 100), making stochastic effects potentially important? Is one interested in cell heterogeneity? Does gene expression noise affect the question-of-interest?
4. Dynamic or Steady-State: Are cellular dynamics important for your question-of-interest? Will the experimental data consist of time courses?

Determining these properties gives well-defined tasks and methods for later in the model development process.

For our example, we choose (1) cell population average, as drug sensitivity and resistance are typically evaluated in cell populations, (2) compartmental with one whole cell compartment, as the kinetic scheme does not involve any transport or subcellular organelles, (3) deterministic, as these components are not very lowly expressed (see below Build Initial Model section), and we do not wish to explain any cellular heterogeneity (although drug response heterogeneity may be interesting), and (4) dynamic, because often ERK1/2 and/or drug dynamics can be important for mediating biological effects, and there are dynamic data available to help constrain the model behavior.

Unfortunately, there is not enough space in this chapter to describe how to build spatial models of biochemical signaling networks; however, there are a number of excellent reviews on the topic (Kholodenko 2006; Neves and Iyengar 2002), and many pieces of the model building process do not depend on whether the model is spatial or compartmental.

6.2.3 Build Initial Model(s)

Once the system is defined and the kinetic scheme is in hand, one can specify the model equations and then simulate and adjust the model, if needed, based on literature or pre-existing experimental data.

6.2.3.1 Choosing Reaction Rate Laws

Each reaction in the kinetic scheme has a corresponding rate law, which describes the rate at which that reaction proceeds in units of concentration per time. The type of reaction will drive the choices for the mathematical form of the rate law. These rate laws can be used directly for deterministic simulations, and with minimal effort for stochastic simulations (by converting units to reaction propensities and molecule numbers).

Association and dissociation reactions are often assumed to follow mass-action kinetics, in which the reaction rate is proportional to the concentration of the reactants. For example, v_1 is an association reaction between MEK1 and KSR2 and we assume the rate law to be $v_1 = k_1$[MEK1][KSR2], where k_1 is a rate constant (1/time/concentration) and [] denotes concentration. The corresponding dissociation reaction is $v_2 = k_2$[MEK1-KSR2], where k_2 is a rate constant (1/time). For simplicity and organizational purposes, we prefer to match rate constant subscript indices with those of the reaction.

Enzymatic reactions can be broken down into elementary steps of association, catalysis, and dissociation, or a lumped rate law can be assumed. Michaelis-Menten lumped rate laws are common and capture the saturating nature of enzymatic reactions even though the basic assumptions that go into such rate laws may not firmly hold in some scenarios to which they are applied (such assumptions are often difficult to validate in vivo) (Chen et al. 2010). If competition or sequestration is thought to be important, then full elementary reactions should be used, in which the binding of enzyme and substrate are considered explicitly.

Many times we further reduce the Michaelis-Menten equations to effective linear rate laws that on their surface look like mass action laws. This requires assumptions of Michaelis constants being much greater than the substrate concentration, giving an effective first-order rate constant of k_{cat}/K_m, where k_{cat} is the enzyme catalytic constant and K_m is the Michaelis constant. However, often due to parameter identifiability issues (see below), one usually cannot determine whether such assumptions are valid or not, and therefore we opt to use Occam's Razor and go with the simpler linear rate laws unless there is experimental evidence that they are too simplistic.

Other potential rate laws are many, depending on specific assumptions the researcher would like to make about each reaction. They include, for example, Hill-type equations for multi-step or unknown mechanisms, and also for transcriptional regulation. Our choices for all the rate laws in our example cannot be described here due to space constraints, but can be found in the MATLAB code for the model contained in the Supplementary Material (SimulateKSRv1.m and SimulateKSRv2.m).

One important consideration is that all of these reaction rate laws are predicated upon the hypothesis that cellular compartments are "well-mixed", meaning that the species are uniformly distributed throughout a compartment. Clearly, in many biological situations such an assumption is suspect. For example, the plasma membrane is not a homogeneous compartment but rather contains microdomains

where receptors are located and downstream signaling occurs. Yet, models based on this well-mixed assumption over the past 40 years, from metabolism to signaling, have been very successful at describing the biological behaviors of interest and providing insight into their mechanisms and function. Nevertheless, one must be aware of this underlying assumption of traditional reaction kinetics approaches, and if there is reason to believe it is causing model-experiment mismatch, one can use more complex agent-based or spatial kinetic Monte Carlo approaches (Collins et al. 2010; Costa et al. 2009). Alternatively, one can instead opt to use empirical approaches in such situations, as described above.

6.2.3.2 Initial Parameter Value Choices

Although parameter estimation is an explicit and significant step later in the modeling process, much can be done here at the initial model development stage to determine parameter values. There is a rich literature describing in vitro studies on enzyme kinetics for many enzymes of interest, as well as affinities of protein–protein interactions. These should serve as initial parameter value estimates. Even if a particular protein–protein interaction or enzyme was not explicitly studied, one can infer a feasible range for the parameter based on experimental studies of homologs or similar processes. Such primary experimental studies should be a first resource for determining initial parameter values. However, it should be noted that there is no guarantee that parameter values measured in an in vitro setting correspond to a live cell situation; thus, caution must be taken to remember that these values are only initial estimates, and may need to be refined.

Another source of parameter values is previously developed kinetic models. These are not as trustworthy as values that have been directly determined experimentally, because parametric identifiability of these types of models is largely not guaranteed (see Parameter Estimation section below). Nevertheless, such values are reasonable starting points when no other information is available.

Because many of the reactions are founded in mechanism, reasonable lower and upper bounds can be posited based on thermodynamic and physical principles, such as detailed balance and diffusion limit. Detailed balance specifies that the product of equilibrium constants for a circular cycle of binding reactions that do not produce or consume energy must equal one, and the diffusion limit specifies that no association reaction can proceed faster than the two reactants can find each other. If one assumes that a particular protein–protein association has the same equilibrium constant no matter what the states of the proteins (e.g., bound to other proteins, phosphorylated, etc.), then detailed balance will be satisfied. There is some spread in the literature as to what the diffusion limit is, perhaps because of uncertainty in how diffusion proceeds in the complex cellular environment as compared to a pure solution, but it is typical to limit on rate constants at approximately 0.1 s^{-1} nM^{-1}. In our experience, association rate values are well described by values a few orders of magnitude less than this ($\sim$0.001 s^{-1} nM^{-1}). One will often find values of dissociation constants or affinities in the literature; we prefer to hold the association

constant at 0.001 s^{-1} nM^{-1} and then calculate the off rate constant, based on the reasoning that affinity will more likely affect the lifetime of the complex (k_{off}), rather than the association rate constant (often determined by encounter rate set by diffusion). If one finds that the association rate constant must be increased far past the diffusion limit to describe the data well, a likely explanation is that the two components are co-localized in a volume smaller than the compartment being considered. For example, the association of two signaling proteins recruited to the plasma membrane will have a much higher apparent rate constant due to their co-localization. Thus, such constraint-violating situations can be resolved with new mechanistic hypotheses, and then the rate law can be adjusted by assuming a volume of this new subcompartment and scaling the two reactant concentrations accordingly.

Another major class of parameters to determine is initial species concentrations. Cellular protein concentrations range from about 0.1 nM ($\sim$100 molecules/cell) on the low end to about 1 μM ($\sim 10^6$ molecules/cell) at the high end (based on a 2000 μm^3 cell—8.3×10^{-4} nM × cell/molecules). A recent study has obtained absolute quantification of many cellular protein concentrations (Schwanhausser et al. 2011), and we routinely use this resource to obtain species concentration estimates. We also used this resource to estimate the overall levels of the species in Fig. 6.3 (see Supplementary Material), but KSR2 levels were not available so this is something we must estimate by analyzing its effects on model behavior. However, it is important to note that these estimates should be refined later by parameter estimation or through direct experimentation, as protein concentrations vary widely across cell types.

6.2.3.3 Deriving the Differential Equations

For a deterministic model, each species will have its own ordinary differential equation (ODE) (or a partial differential equation if a spatial model is being considered; stochastic models do not have differential equations but rather reactions are fired probabilistically using rate laws derived from the deterministic case). A large majority of kinetic models of biochemical signaling networks are based on ODEs. Thus we describe in detail here the derivation of the ODEs (there are numerous excellent reviews on these other topics for interested readers).

The ODE for each species is simply the sum over all reactions that produce or consume that species, with all reactions multiplied by the stoichiometry of the species in that reaction. Consuming stoichiometries are negative, whereas producing stoichiometries are positive. The entire system of ODEs is succinctly represented in matrix-vector notation using the stoichiometric matrix **S**

$$\frac{d\mathbf{x}}{dt} = \mathbf{S}\mathbf{v} \tag{6.1}$$

here, **x** is an n-by-1 vector of species concentrations, $\boldsymbol{v}$ is an m-by-1 vector of reaction rates, and **S** is an n-by-m matrix containing the stoichiometric coefficients, with species corresponding to rows and reactions corresponding to columns. In principal, one does not need to specify the stoichiometric matrix to derive the differential equations, and can instead write out sums of the reaction rates for each species. However, we strongly prefer to use this stoichiometric matrix approach to deriving the differential equations for several reasons: (i) reaction stoichiometries are stored in one place and it is much easier to ensure their correctness; (ii) calculating the differential equations requires only one line of code (with linear algebra libraries); (iii) any errors will be confined to the stoichiometric matrix itself, which is straightforward to troubleshoot, rather than possibly being contained in hard coded sums of reaction rates in multiple places; (iv) there are a host of analyses that provide model information based on analysis of the stoichiometric matrix alone (e.g., conserved moieties (Vallabhajosyula et al. 2006)); (v) providing a stoichiometric matrix facilities combining, modifying, and sharing models. An example stoichiometric matrix for our model is provided in the Supplementary Material (KSRModelv2Stoich.csv).

When reactions transport species across compartments, or when two reactants are localized to a compartment but the reactant concentrations are defined with respect to different volumes, special care must be taken to ensure the ODEs are correct. In the case of transport, the product stoichiometry should be the volume ratio between the compartments. For example, if a species A is transported from the cytoplasm to the nucleus, with rate $v = k_t[A]_{cyt}$, then the stoichiometric matrix entry for this reaction for cytoplasmic A is -1, but the entry for nuclear A is the volume ratio V_{cyt}/V_{nuc}, with the subscripts cyt and nuc referring to the cytoplasm and nucleus. Alternatively, one can multiply the vector of reaction rate laws by their respective volumes, to obtain a left-hand side that is in terms of molecules (or moles) per time, rather than concentration. In the case of species localization, both species concentrations must be rescaled to the volume of the reaction compartment. A typical example is ligand-receptor binding. Ligand-receptor association occurs in the extracellular space, but ligand concentration is defined in the extracellular compartment and receptor concentration in the cellular compartment. In this case, the receptor concentration in the ligand-receptor association rate law should be multiplied by the factor V_c/V_{ec}, with the subscripts ec and c denoting extracellular and cellular compartments. This rescales the receptor concentration to the extracellular compartment for this particular reaction (but not in the entire model).

6.2.3.4 Simulating the Model-Deterministic

At this point one will have an ODE model that describes the mechanistic relationships between the chosen inputs and outputs. To simulate this model, one first needs to specify the initial conditions. In most scenarios this is a two-step process. First, one sets all of the unmodified, unbound species equal to the total concentrations determined above, and all other concentrations to zero. Then, with the input

level set to that corresponding to the system prior to experimental perturbation, the model is integrated to its natural steady-state. This step is called "equilibration." Integration should be done with an algorithm designed for stiff systems, such as `ode15s` in MATLAB. The values for all states after equilibration become the initial conditions for a relevant simulation in response to a perturbation. Then, the response of a system to the perturbation is simulated by integrating over the desired time interval, starting from the equilibrated initial conditions.

Simulations with the initial model should be used to verify that the model does not have errors and that it displays expected behavior. No species should have negative concentrations and moiety conservation should be confirmed. It is typical to do a preliminary study of the dynamics and dose response of the system, and compare it to what might be expected. Almost always the model must be altered. We did this for our KSR2 model with the following characteristics in mind: (i) ERK1/2 activation in response to mitogens usually occurs over $\sim$5 min., and in the presence of strong negative feedback with decline by $\sim$30 min.; (ii) there should be a smooth dose response of RasGTP, active B-Raf (RasGTP bound), ppMEK1, and ppERK1/2 to mitogen levels 60 min. post mitogen stimulus; (iii) most MEK1 and ERK1/2 should be in the doubly phosphorylated form at high mitogen doses; (iv) there should be an optimal KSR2 concentration for causing maximal ERK1/2 activation 60 min post mitogen stimulus; (v) increasing negative feedback strength should smoothen the dose response to make ppMEK1 and ppERK1/2 (at 60 min.) increase more gradually in response to increased mitogen levels. Unfortunately, we cannot expand in detail how we made specific modifications to meet these criteria here, but the model code with comments before (file name with suffix '*v*1') and after ('*v*2') modification is given in the Supplementary Material (SimulateKSRv1.m and SimulateKSRv2.m). Interested readers are encouraged to analyze the differences between the two, and contact us with questions.

6.2.3.5 Simulating the Model-Stochastic

If the model is stochastic, then there are no ODEs to integrate, and reactions are fired through random sampling approaches. If one is unsure whether the molecule numbers of various species are low enough to warrant stochastic simulation, it is prudent to compare deterministic to stochastic simulation results and analyze if there are significant differences. Common stochastic simulation methods reviewed (Golightly and Gillespie 2013) are predominantly based on the Gillespie algorithm or variants thereof. A preferred exact algorithm is that of Gibson and Bruck (2000), but this is usually computationally intensive and cannot be widely applied to larger models, particularly for stiff systems as is common for biochemical signaling networks. The implicit tau-leaping algorithm (Cao et al. 2007) is the numerical equivalent to ODE integrators for stiff systems, and although it is not exact, it usually gives acceptable results by time-averaging the behavior of fast reactions. There are also hybrid methods that divide the model into stochastic and deterministic portions (Salis et al. 2006). Many software packages make many such

algorithms readily available in standard code libraries or with graphical user interfaces, such as StochKit and Cain.

Sometimes, even though system behavior on the single cell level is stochastic, all the molecule numbers are large resulting in little variability in reaction rates with traditional stochastic simulations. Such noise may be largely due to cell-to-cell protein expression variability, which arises from the stochastic nature of gene expression. One can take such noise into account without direct stochastic simulations, or rather by sampling total protein concentrations from known distributions, and then integrating the ODE model for many sets of these initial conditions (~ 1000) (Birtwistle et al. 2012a; Gaudet et al. 2012). Although some have used the log-normal distribution for such sampling, it predicts a scaling behavior between mean protein concentrations and the variance of protein concentrations that is not consistent with experimental observations (Birtwistle et al. 2012b). The gamma distribution is a more appropriate choice that captures cell-to-cell variability in protein expression over a wider range of conditions (Birtwistle et al. 2012b; Shahrezaei and Swain 2008). Alternatively, knowledge of the transcription and translation rates, as well as mRNA and protein degradation rates for each species in a model, permits the simulation of burst-like transcription and translation events de novo using algorithms that simulate stochastic processes such as the Gillespie algorithm.

6.2.3.6 Annotating the Model

For others to understand and reuse the model, it is essential to provide carefully documented code, record the assumptions and parameter values that were used, and then upload them to model-sharing resources such as BioModels or convert them into a universal format such as SBML as well as provide source code (e.g., MATLAB or C++ code) (Waltemath et al. 2011). There is much literature devoted to these topics and we do not discuss them in detail, but it is nevertheless important to emphasize this step. We mention it here, rather than at the end, because it is typically easier to annotate while the model is being initially developed, rather than at the end when many versions of the model have been created, and many of the assumptions were made long ago (sometimes several years).

6.3 Experimental Design and Execution of Experiments

It can often take months, if not longer, to gather reliable experimental data for the initial development of a kinetic model. Thus, one should aim to plan such experiments at a very early stage in the model development process, perhaps as early as the kinetic scheme is produced (see curved arrow in Fig. 6.2). Although some experimental design methods depend on an initial model, some experiments can be planned and completed prior to completion of an initial model.

Table 6.1 Selected experimental methods and their properties relevant to kinetic modeling

Method	What it measures	Single cell or population measure	Live, fixed, or lysed cells	Ability to multiplex	Ability to generate time-course data	Absolute or relative quantification
qRT-PCR	RNA	Both	Lysed	Medium	Medium	Relative
RNA-seq	RNA	Both	Lysed	High	Low	Absolute or relative
FISH	RNA or DNA	Single cell	Live or fixed	Medium	Low	Relative
Protein tagging	Protein	Both	Live or fixed	Low	High	Relative
FRET	Protein interactions	Single cell	Live or fixed	Low	High	Relative
Western blot	Protein	Population	Lysed	Low-medium	Medium	Absolute or relative
Flow cytometry	Protein	Both	Live or fixed	Low-medium	Medium	Relative
Luminex	Protein	Population	Lysed	Medium	Medium	Relative
Mass spectrometry	Protein	Population	Lysed	High	Low	Absolute or relative
Mass cytometry (CyTOF)	Protein	Both	Fixed	High	Low	Relative

Experimental design for kinetic models of biochemical signaling networks consists of answering two questions. What types of perturbations should I apply? What should I measure, and when should I measure it for each perturbation? If a biological system has not yet been chosen, this also needs to be addressed here based on the question-of-interest and system definition above. Are cell lines sufficient or do we need animal models? What lines are appropriate? We focus on cell lines. Here, we briefly describe the types of perturbation and measurement methods that are commonly used, with the goal of providing enough information for readers to make an informed choice, although it is impossible to comprehensively list all methods here. As an aid, we provide a summary in Table 6.1.

We note that although it is seldom done, we strongly advocate for directly measuring the total absolute abundances of all model proteins in the experimental system of interest. Without these measurements to constrain the magnitude of the various species in the model and the initial conditions, it is generally difficult to make relevant predictions of the system behavior.

6.3.1 Types of Perturbations

The most straightforward way to perturb the system is to apply a defined and time-invariant concentration (dose) of a compound. Such compounds can be pharmacological (e.g., a small molecule kinase inhibitor), or biological (e.g., a growth factor). While biological compounds are usually applied at the moment defined as time point 0, pharmacological inhibitors are often applied to the system prior to biological compound treatment, to ensure its uniform distribution in cells. Standard pipette-based approaches are limited to simple perturbation time courses, such as the time 0 step response or a pulse-chase where the original dose is washed out and then replaced with a new dose. However, sophisticated pump-based perfusion chamber or microfluidics methods (Mettetal et al. 2008) and optogenetic-based methods (Toettcher et al. 2011, 2013) can allow for complex time-dependent perturbations, such as sine waves or increasing ramps.

The system can also be perturbed at the transcript level. One can transfect (liposome-based methods) or infect (virus-based methods) genetic material into cells. The genetic material (usually a DNA plasmid) can express a gene ectopically, overexpress an endogenous gene, or downregulate a gene via RNA interference (e.g., shRNA) in a transient setting. A negative selection mechanism is needed to obtain cells that permanently exhibit the alteration. To do this, the plasmid also usually encodes resistance to a particular drug, such as puromycin. However, genome-editing methods, such as those involving CRISPR and TALEN (Ran et al. 2013; Reyon et al. 2012), are becoming more realistic.

Once a gene has been permanently integrated, it can be controlled using drugs such as tetracycline (or more commonly its higher affinity analog doxycycline) so long as the gene is engineered to be Tet responsive. Expression can be modulated in a dose-dependent manner, permitting a level of control similar to that describable by a kinetic model. Aside from such transcription rate control, one can use the small molecule Shield1 to increase the half-life of proteins that contain a DD domain (Banaszynski et al. 2006).

6.3.2 Types of Measurements-Transcripts

To obtain a quantitative measure of the mRNA expression levels, one can perform a qRT-PCR, or quantitative reverse-transcription polymerase chain reaction. This technique is useful if one is interested in a few transcripts in a cell population setting, but sometimes is done on single cells (although it can be technically challenging). The alternative to enumerating individual transcripts a priori is to measure transcripts globally, using a technique such as RNA-seq. RNA-seq data can provide absolute mRNA copy number information as long as internal controls are included in the analysis. Although RNA-seq is most commonly performed on cell populations, the ability to perform RNA-seq in single-cells is rapidly emerging

(Islam et al. 2013). If one is unable obtain RNA-seq data, microarray is a comparable methodology though it possesses several drawbacks such as high background noise and the limitation of having to define target sequences of interest a priori (Xu et al. 2013). Fluorescence in situ hybridization (FISH) is a technique one can use to visualize a particular mRNA transcript via a fluorescently labeled nucleic acid probe. FISH can lend insight into differences in gene expression across a population of cells and has been used to study stochastic gene expression in mammalian cells with success (Raj et al. 2006). Generally, one fixes and permeabilizes the cells prior to FISH, thus killing them; however, it is possible to perform FISH in live cells with minimal perturbation (Simon et al. 2010). Highly multiplexed FISH methods are also becoming available (Lubeck and Cai 2012).

6.3.3 Types of Measurements-Proteins

One way to measure the levels (and/or spatial location) of a protein over time is to "tag" it with a fluorescent protein, by cloning them together such that they are transcribed as a fusion protein. This method could serve to monitor protein expression levels in response to a stimulus, such as cFos levels following growth factor-mediated ERK activation. Importantly, protein tagging can be used to measure protein levels and/or localization in live cells over time. Another live-cell technique involves genetically encoded probes based on Forster Resonance Energy Transfer (FRET) (Miyawaki 2011). Such probes allow one to measure the spatio-temporal dynamics of biochemical signaling activities. For example, the EKAR-EV FRET probe responds to ERK1/2 activity by changing its FRET (Komatsu et al. 2011), and the past decade has provided a wealth of these probes for various biochemical activities one may be interested in.

To measure protein levels in a cell population, the western blot is the gold standard. In this technique, an antibody is used to bind to and measure the amount of a protein or protein state in whole cell lysate that has been separated by molecular weight. As an example, one could measure the level of a phosphorylated protein in response to increasing mitogen dose. Further, the technique can be combined with immunoprecipitation to quantify levels of protein–protein interactions. Although enhanced chemiluminescence is often used to quantify western blot signals, there can be non-linearity in such measurements; therefore for quantitative kinetic models, systems such as LI-COR are preferred which provide a linear signal-response. Absolute quantification is seldom done but possible by including known protein concentrations as internal controls. It should be noted that the western blot has been miniaturized into a so-called "microwestern" which is potentially useful for kinetic modeling applications, since it allows probing with many antibodies over many perturbation conditions (dose/time points). A Luminex assay allows multiplexing to a similar extent as microwesterns (Ciaccio et al. 2010).

Flow cytometry allows one to measure the relative level of proteins in a large number of single cells at a single time point, capturing the distribution of protein

levels found in a population of cells using fluorescently labeled antibodies. In principal, modern flow cytometers can quantify up to 16 analytes, but in practice this is quite difficult and 3–4 color imaging is more typical. A flow cytometer coupled to a mass spec (CyTOF) potentially allows quantification of ~40 analytes in single cells (Bendall et al. 2011), but requires specialized equipment and antibodies.

As previously stated, one must determine initial protein concentrations in order to construct a kinetic model that is biologically meaningful. One way to gather initial protein concentrations in the cells of interest is via a proteomics approach using mass spectrometry (MS), which can globally quantify protein levels in a population of cells. Advances in mass spectrometry have enabled the quantification of proteins in terms of absolute copy numbers, easily convertible to units of molecules/cell (Schwanhausser et al. 2011).

6.3.4 Formal Statistical Design of Experiments

With an initial model that is based on and reproduces to an acceptable level available literature or preliminary experimental data, it is possible to implement a formal statistical design of experiments (DoE) tailored to the goal of the modeling exercise. DoE is a mature field, and many have demonstrated how it might be applied to the types of biochemical signaling models we describe here (Bandara and Meyer 2012; Banga and Balsa-Canto 2008). The specific approaches that have been described depend on whether the goal is parameter estimation for one model, or discrimination between many candidate models. Although such established DoE methods have historically been successful in other fields, their original development is largely grounded in application to relatively low dimensional linear models, with only a handful of states and experimental decision variables. Kinetic models of biochemical signaling networks, however, have many properties that in our opinion preclude meaningful application of such approaches with current computational technology:

1. They are almost ubiquitously highly non-linear. Thus, local linear approximations of the model are used to apply these traditional DoE methods, the validity of which is often unclear.
2. They are typically high-dimensional (>10 states). Thus, the potential number of states to measure that the DoE algorithm must choose from is often overwhelmingly large.
3. There are many potential experimental perturbations. This causes a combinatorial explosion of possibilities for these DoE methods that optimize over the experimental decision variable space.
4. It is generally not yet known how to guarantee parametric identifiability (see Parameter Estimation below). Thus, it is unlikely that these DoE methods would produce a design that significantly enhances our ability to identify model

parameters or discriminate between two models, as compared to a more pragmatic approach.
5. The validity of the initial model structure and parameter values is unknown at this point. Thus, it makes little practical sense to spend a significant amount of experimental resources to implement an optimal experimental design based on an initial model that is likely inadequate.

These major reasons are most likely why in practice, formal DoE methods are typically not employed to develop kinetic models of biochemical signaling networks. This is not to say that such methods lack importance, but rather that much research is needed to develop new DoE methods that are suited to the properties of this class of models and fill the current void in the model development process.

6.3.5 *Practical Design of Experiments*

The more pragmatic yet common approach to experimental design is to use the insight of expert biologists to answer the two basic questions needed for experimental design. This can often be done only with a kinetic scheme and thus can be done before the initial model is completed. At this initial stage of model development, a broad experimental design that perturbs the system using the model inputs and measures across "important" states is preferred. Typically, perturbations consist of applying extracellular agonists or antagonists (e.g., a growth factor) to cells that have been serum-starved overnight (to minimize confounding variables), in the presence or absence of pharmacological or small molecule inhibitors of "key pathways" in the model. What states are "important" and what pathways are "key" is best informed by expert opinion or initial experimental data. Dose responses for the extracellular agonists or antagonists give important information to constrain model behavior and should be done if the resources are available. Logarithmic dose spacing (e.g., base 10) between saturating and limit of detection levels is often most informative. Which species to measure is also best informed by an experimental biologist with expertise in that system, and of course is limited by available technologies and resources. The time point selection should also be informed by expert opinion, and depends on how one defined the system, the question(s) of interest, the limits of the chosen technology, and available resources.

Biochemical signaling models have been termed "sloppy" (Gutenkunst et al. 2007), which refers to the fact that many key system outputs are quite robust to variations in many model parameters (discussed more in Parameter Estimation below). This property may be why such a pragmatic approach to experimental design often results in a successful modeling exercise, because sloppiness dictates that many of the choices simply do not matter for the behavior of key system outputs.

6.3.6 *Comparison of Experimental Data to Model Simulations*

In general an experimental measurement does not directly correspond to a particular model species. For example, if one is measuring a kinase activity in live cells by FRET, the resulting FRET measurement cannot be directly related to the kinase activity in the model, because there is not a linear relationship between the two (Birtwistle et al. 2011). As another example, if one measures the total amount of cellular RasGTP by pull down and western blot, this would correspond to a sum over several model states in Fig. 6.3. Moreover, the way in which the western blot data are normalized can have a significant impact on the quality of the normalized data (Degasperi et al. 2014). Thus, great care should be taken to ensure that the best comparison between experimental data and model simulations is being employed, and often requires mapping the model variables onto so-called "observable" variables with defined functions. This requires a thorough understanding of both the computational model and the experimental data; thus close collaboration and effective communication between wet and dry lab researchers is essential.

6.4 Parameter Estimation

With an initial model and experimental data in hand, the next task is to determine whether the model is capable of describing the experimental data, and what range of parameter values give good fit. This exercise is called parameter estimation or "training." Parameters include total protein abundances (if not directly measured) and kinetic rate parameters in each rate equation. Although one will have reasonable initial values for all these quantities, it is highly unlikely that the model will be able to reproduce the new experimental data without modifying the parameter values. This is expected since many of the initial parameter values will have come from in vitro studies or from data collected in a different biological system.

Parameter estimation for kinetic models of biochemical signaling pathways is an extremely challenging exercise for two main reasons. First, the model is high-dimensional and nonlinear. Thus it is computationally expensive to explore the parameter space extensively when searching for good-fitting parameter sets. Second, it is not understood how to guarantee parametric identifiability for these models, and even this general class of nonlinear chemical kinetic models. An identifiable parameter is one whose value is well-constrained by the experimental data, such that it is known with acceptable precision. A typical kinetic model of a biochemical signaling pathway will not have identifiable parameters. This is quite shocking and perhaps even disturbing to modelers from other disciplines, such as pharmacokinetics and pharmacodynamics described in other chapters in this book. Despite this ubiquitous parametric uncertainty, an emerging theme in this type of modeling is that key temporal outputs are typically robust to large changes in most

parameter values. This property seems to be general for this class of systems biology models and is referred to as "sloppiness" (Gutenkunst et al. 2007). From a biological and evolutionary perspective, this makes sense, because key dynamic behavior should not be affected by common noise sources. From a modeling perspective, this to some extent mitigates the problem that unidentifiable parameters cause, with respect to reproducing biologically relevant behavior. However, one should still strive to ensure all rate constants and concentrations are within biophysically feasible ranges, such as not exceeding the diffusion limit, and are justified to do so because model parameters typically have a biophysical interpretation. Nonetheless, we are still left with the problem that we cannot be certain that unidentifiable parameters do not affect our conclusion. That is why downstream model analyses must account for how parametric uncertainty affects predictions, such as global sensitivity analysis methods (see Model Analysis).

The first step in parameter estimation is to define lower and upper bounds for the unknown parameter values. As described above, these can typically be set through hard biophysical or thermodynamic limits. Next, one must define an objective function that represents goodness-of-fit. There are many options, including log-likelihood and sum-of-squared errors between simulations and data, and the particular choice depends on assumptions for the expected errors in the experimental data (Raue et al. 2013). Regardless, it is essential to scale each quantity such that error does not depend on units (variance scaling is common and often statistically valid). Then, one must choose an algorithm that will vary the parameters over the bounds to optimize the objective function. Local, deterministic gradient-based optimization is inappropriate for this class of models as they are nonlinear and of high dimension. Global optimization methods are a necessary component of any choice. One simple global method is to repeatedly employ local methods but from different initial parameter values judiciously chosen from across the parameter space (e.g., with latin hypercube sampling), and in fact such methods may be both accurate and efficient (Raue et al. 2013). However, the majority of studies have had success using either the genetic algorithm (Nakakuki et al. 2010; Schoeberl et al. 2002) or simulated annealing (Wang et al. 2009). Bayesian methods have been applied in a few cases with success (Vyshemirsky and Girolami 2008; Eydgahi et al. 2013), and such methods are very attractive since they rigorously account for multi-dimensional parametric uncertainty, although at much higher computational cost than other global methods. Lastly, the chosen algorithm should be run many times over, due to the inherent sloppiness of these models and therefore parameter uncertainty. This allows one to estimate the range of parameter values that give rise to models with "acceptable" fit. We suggest obtaining at least 10 good fitting parameter sets; $\sim$100 would give a much better indicator of parametric uncertainty but even 10 is sometimes difficult to obtain due to the computational burden.

Most of these parameter estimation algorithms are well-suited to parallelization and should be implemented on high performance computing resources. One potentially promising new technology is graphical processing unit (GPU)-based computation. A single GPU card can contain $\sim$3000 processors that run the same

program (i.e., model) given different input values (i.e., parameter sets), which is ideal for this parameter estimation task. A desktop workstation can house up to 4 such cards, giving ~12,000 GPU processors in a single machine. However, robust ODE solvers that operate in the specialized GPU environment (e.g., NVIDIA CUDA language) must first be developed for such an approach to be implemented. Some attempts exist, such as `cudasim` which can take an SBML model input and use the GPU to simulate it both stochastically or deterministically (Zhou et al. 2011), or code libraries such as `odeint`. Any GPU-based solver must be able to implement implicit solver methods that can tackle the stiffness that is present in these types of models.

After parameter estimation, one must decide whether the model has acceptable fit or not. This is commonly done by simply plotting the model simulations against the experimental data, and looking for close match between the two. In addition, one can analyze the distribution of residual errors (differences between simulations and data) for evidence of bias (non-zero mean). If there is bias, then that suggests that the model structure and/or parameter bounds must be changed. A clear indicator that parameter bounds should be changed is if estimated parameter values are constantly on or near the bound. It is desired to first try to expand parameter bounds if it is likely to improve fits, before altering the model structure. How to alter the model structure is highly dependent on the nature of the model-experiment mismatch and needs to be analyzed on a case-by-case basis. Regardless, if the structure needs refinement, one must return to defining the system to come up with new hypotheses.

6.5 Model Discrimination

Model discrimination refers to determining which model among a set is most appropriate given experimental observations. Usually, parameter estimation must be done before model discrimination. Some have investigated model discrimination in a formal way, using Bayesian methods to compute Bayes factors for each model (Xu et al. 2010). More straightforward and computationally inexpensive methods are simply considering the sum of squared residual errors for each model and weighting it by the number of free parameters with the Akaike or Bayesian Information Criterion. As mentioned briefly above, there are some statistical design of experiments methods that are focused on model discrimination, but largely any experiments that are tailored to determining which model among many is most appropriate are designed in a pragmatic manner.

6.6 Model Analysis and Prediction

After successfully completing parameter estimation, one is ready to analyze the model to provide potential answers for the question-of-interest in the form of experimentally testable hypotheses. The type of model analysis will differ depending on the question-of-interest. For example, if nonlinear dynamical phenomena such as bistability or oscillations are of interest (e.g., cell cycle or circadian clock), then traditional bifurcation analysis techniques can be applied (Chickarmane et al. 2005), although this can be difficult with the high dimensional models typical of biochemical signaling pathways and need expert knowledge to reduce the number of parameters one is considering.

One general type of model analysis that is almost universally useful is parameter sensitivity analysis. This consists of varying parameter values and observing the effect on outputs-of-interest. Sensitivity analysis, like parameter estimation methods, can be local or global. Local methods consider only a particular region of parameter space, and are typically inappropriate because (i) the models are nonlinear and (ii) parameters are not identifiable and therefore their values are not known precisely. Unfortunately, local methods include metabolic control analysis which has been widely applied to understand steady-state phenomena in metabolic networks (Kholodenko et al. 1994). Global methods consider an entire multi-dimensional region of parameter space, and therefore can account for the inherent parametric uncertainty present in these models. There are many global sensitivity analysis methods available (reviewed in (Saltelli 2008)), and it is not yet clear which may be best for these types of models. We have previously used a rigorous yet straightforward global method called Sobol sensitivity analysis, which quantitatively decomposes the total variance in outputs-of-interest into the contributions by individual parameters and the interactions between parameters (Sobol 2001). A larger variance indicates a more important parameter and therefore important mechanism. The method functions by evaluating model outputs for a large number of different parameter sets, and, importantly, is capable of providing error estimates on the sensitivity coefficients. Although a very large number of model evaluations are needed to produce statistically significant results, the algorithm is easily parallelizable. We were able to perform Sobol sensitivity analysis on a model of the VEGFR pathway containing 77 parameters with relative computational ease (Zhang et al. 2014).

6.7 Model Validation

The model analysis stage will produce many predictions, and these predictions must be sorted into those that can be experimentally tested and those that cannot, which, like experimental design, requires close contact between the wet and dry labs. Among those predictions that can be experimentally tested, typically the

counter-intuitive or unexpected ones are the best to explore experimentally, in addition of course to those that directly address the question-of-interest. Only a small subset of all predictions can be experimentally addressed, so it is important to carefully select those to further consider. Importantly, any experiments for model validation must be independent from those used to develop and train the model.

After the new experiments are performed (or mined from the literature), one must compare simulation predictions to the new data, and then interpret what it means for the question-of-interest. Currently, a model is considered valid if it is able to reproduce independent experimental data outside the scope of the original training data set. If the model is not valid, typically it can still yield insight into the question-of-interest, and may still be valuable in that regard. Such disagreement prompts a new hypothesis and iteration back to the first step of the modeling process. Yet, even if a model is validated in this way, it is not certainly universally valid, and assuming that the model can predict many other quantities outside of its training set would be grossly premature. Much more research must be done to elucidate how a more unbiased approach to model validation can be designed, so that confidence in model predictions can be quantified in a rigorous manner.

6.8 Conclusions

Building a kinetic model of a biochemical signaling network is a significant investment of time and effort, and therefore one should have very clear goals and expectations for what the eventual model will accomplish for the research question-of-interest. Such kinetic models have many properties that can potentially fill a significant gap in the drug development pipeline and inform personalized medicine approaches. However, to reach this potential, much theoretical work must be done to improve and standardize each step of the model building process shown in Fig. 6.2. Any new methods must take special care to accommodate the properties of biochemical signaling networks that hamper current methods, namely, the complexity of these networks, their large scale, and inherent uncertainty.

Acknowledgments We thank Arvin Dar for helpful discussions about the KSR2/B-Raf example. We gratefully acknowledge funding from the Icahn School of Medicine at Mount Sinai, an IBM Faculty Award, and the NIH Grants P50 GM071558 (Systems Biology Center New York), R01GM104184, and U54HG008098 (LINCS Center). MB is supported by a NIGMS-funded Integrated Pharmacological Sciences Training Program grant (T32 GM062754). We regret not being able to cite many good papers in the literature on these diverse topics due to space constraints.

References

Banaszynski LA, Chen LC, Maynard-Smith LA, Ooi AG, Wandless TJ (2006) A rapid, reversible, and tunable method to regulate protein function in living cells using synthetic small molecules. Cell 126(5):995–1004

Bandara S, Meyer T (2012) Design of experiments to investigate dynamic cell signaling models. Methods Mol Biol 880:109–118

Banga JR, Balsa-Canto E (2008) Parameter estimation and optimal experimental design. Essays Biochem 45:195–209

Bendall SC, Simonds EF, Qiu P, el Amir AD, Krutzik PO, Finck R, Bruggner RV, Melamed R, Trejo A, Ornatsky OI, Balderas RS, Plevritis SK, Sachs K, Pe'er D, Tanner SD, Nolan GP (2011) Single-cell mass cytometry of differential immune and drug responses across a human hematopoietic continuum. Science 332(6030):687–696

Birtwistle MR (2014) Analytical reduction of combinatorial complexity arising from multiple protein modification sites. J R Soc Interface 12(103)

Birtwistle MR, von Kriegsheim A, Kida K, Schwarz JP, Anderson KI, Kolch W (2011) Linear approaches to intramolecular Forster resonance energy transfer probe measurements for quantitative modeling. PLoS ONE 6(11):e27823

Birtwistle MR, Rauch J, Kiyatkin A, Aksamitiene E, Dobrzynski M, Hoek JB, Kolch W, Ogunnaike BA, Kholodenko BN (2012a) Emergence of bimodal cell population responses from the interplay between analog single-cell signaling and protein expression noise. BMC Syst Biol 6:109

Birtwistle MR, von Kriegsheim A, Dobrzynski M, Kholodenko BN, Kolch W (2012b) Mammalian protein expression noise: scaling principles and the implications for knockdown experiments. Mol BioSyst 8(11):3068–3076

Birtwistle MR, Mager DE, Gallo JM (2013) Mechanistic vs. empirical network models of drug action. CPT Pharmacomet Syst Pharmacol 2:e72

Bollag G, Hirth P, Tsai J, Zhang J, Ibrahim PN, Cho H, Spevak W, Zhang C, Zhang Y, Habets G, Burton EA, Wong B, Tsang G, West BL, Powell B, Shellooe R, Marimuthu A, Nguyen H, Zhang KY, Artis DR, Schlessinger J, Su F, Higgins B, Iyer R, D'Andrea K, Koehler A, Stumm M, Lin PS, Lee RJ, Grippo J, Puzanov I, Kim KB, Ribas A, McArthur GA, Sosman JA, Chapman PB, Flaherty KT, Xu X, Nathanson KL, Nolop K (2010) Clinical efficacy of a RAF inhibitor needs broad target blockade in BRAF-mutant melanoma. Nature 467(7315):596–599

Bouhaddou M, Birtwistle MR (2014) Dimerization-based control of cooperativity. Mol Bio Syst 10:1824–1832

Brennan DF, Dar AC, Hertz NT, Chao WC, Burlingame AL, Shokat KM, Barford D (2011) A Raf-induced allosteric transition of KSR stimulates phosphorylation of MEK. Nature 472 (7343):366–369

Cao Y, Gillespie DT, Petzold LR (2007) Adaptive explicit-implicit tau-leaping method with automated tau selection. J Chem Phys 126

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

Chickarmane V, Paladugu SR, Bergmann F, Sauro HM (2005) Bifurcation discovery tool. Bioinformatics 21(18):3688–3690

Chylek LA, Harris LA, Tung CS, Faeder JR, Lopez CF, Hlavacek WS (2013) Rule-based modeling: a computational approach for studying biomolecular site dynamics in cell signaling systems. Wiley Interdiscip Rev Syst Biol Med

Ciaccio MF, Wagner JP, Chuu CP, Lauffenburger DA, Jones RB (2010) Systems analysis of EGF receptor signaling dynamics with microwestern arrays. Nat Methods 7(2):148–155

Collins S, Stamatakis M, Vlachos DG (2010) Adaptive coarse-grained Monte Carlo simulation of reaction and diffusion dynamics in heterogeneous plasma membranes. BMC Bioinform 11:218

Costa MN, Radhakrishnan K, Wilson BS, Vlachos DG, Edwards JS (2009) Coupled stochastic spatial and non-spatial simulations of ErbB1 signaling pathways demonstrate the importance of spatial organization in signal transduction. PLoS ONE 4(7):e6316

Dankort D, Curley DP, Cartlidge RA, Nelson B, Karnezis AN, Damsky WE Jr, You MJ, DePinho RA, McMahon M, Bosenberg M (2009) Braf(V600E) cooperates with Pten loss to induce metastatic melanoma. Nat Genet 41(5):544–552

Das Thakur M, Salangsang F, Landman AS, Sellers WR, Pryer NK, Levesque MP, Dummer R, McMahon M, Stuart DD (2013) Modelling vemurafenib resistance in melanoma reveals a strategy to forestall drug resistance. Nature 494(7436):251–255

Degasperi AB, MR; Volinsky, N; Rauch, J; Kolch, W; Kholodenko, BN (2014) Evaluating strategies to normalise biological replicates of Western Blot Data. PLoS One (in press)

Eydgahi H, Chen WW, Muhlich JL, Vitkup D, Tsitsiklis JN, Sorger PK (2013) Properties of cell death models calibrated and compared using Bayesian approaches. Mol Syst Biol 9:644

Fritsche-Guenther R, Witzel F, Sieber A, Herr R, Schmidt N, Braun S, Brummer T, Sers C, Bluthgen N (2011) Strong negative feedback from Erk to Raf confers robustness to MAPK signalling. Mol Syst Biol 7:489

Gaudet S, Spencer SL, Chen WW, Sorger PK (2012) Exploring the contextual sensitivity of factors that determine cell-to-cell variability in receptor-mediated apoptosis. PLoS Comput Biol 8(4): e1002482

Gibson MA, Bruck J (2000) Efficient exact stochastic simulation of chemical systems with many species and many channels. J Phys Chem 104(9)

Golightly A, Gillespie CS (2013) Simulation of stochastic kinetic models. Methods Mol Biol 1021:169–187

Gutenkunst RN, Waterfall JJ, Casey FP, Brown KS, Myers CR, Sethna JP (2007) Universally sloppy parameter sensitivities in systems biology models. PLoS Comput Biol 3(10):1871–1878

Hatzivassiliou G, Song K, Yen I, Brandhuber BJ, Anderson DJ, Alvarado R, Ludlam MJ, Stokoe D, Gloor SL, Vigers G, Morales T, Aliagas I, Liu B, Sideris S, Hoeflich KP, Jaiswal BS, Seshagiri S, Koeppen H, Belvin M, Friedman LS, Malek S (2010) RAF inhibitors prime wild-type RAF to activate the MAPK pathway and enhance growth. Nature 464 (7287):431–435

Heidorn SJ, Milagre C, Whittaker S, Nourry A, Niculescu-Duvas I, Dhomen N, Hussain J, Reis-Filho JS, Springer CJ, Pritchard C, Marais R (2010) Kinase-dead BRAF and oncogenic RAS cooperate to drive tumor progression through CRAF. Cell 140(2):209–221

Islam S, Zeisel A, Joost S, La Manno G, Zajac P, Kasper M, Lonnerberg P, Linnarsson S (2013) Quantitative single-cell RNA-seq with unique molecular identifiers. Nat Methods

Iyengar R, Zhao S, Chung SW, Mager DE, Gallo JM (2012) Merging systems biology with pharmacodynamics. Sci Transl Med 4(126):126–127

Janes KA, Albeck JG, Gaudet S, Sorger PK, Lauffenburger DA, Yaffe MB (2005) A systems model of signaling identifies a molecular basis set for cytokine-induced apoptosis. Science 310 (5754):1646–1653

Kholodenko BN (2006) Cell-signalling dynamics in time and space. Nat Rev Mol Cell Biol 7 (3):165–176

Kholodenko BN, Sauro HM, Westerhoff HV (1994) Control by enzymes, coenzymes and conserved moieties. A generalisation of the connectivity theorem of metabolic control analysis. Eur J Biochem 225(1):179–186

Kholodenko BN, Hancock JF, Kolch W (2010) Signalling ballet in space and time. Nat Rev Mol Cell Biol 11(6):414–426

Kirouac DC, Onsum MD (2013) Using network biology to bridge pharmacokinetics and pharmacodynamics in oncology. CPT Pharmacomet Syst Pharmacol 2:e71

Kitano H, Funahashi A, Matsuoka Y, Oda K (2005) Using process diagrams for the graphical representation of biological networks. Nat Biotechnol 23(8):961–966

Klinger B, Sieber A, Fritsche-Guenther R, Witzel F, Berry L, Schumacher D, Yan Y, Durek P, Merchant M, Schafer R, Sers C, Bluthgen N (2013) Network quantification of EGFR signaling unveils potential for targeted combination therapy. Mol Syst Biol 9:673

Komatsu N, Aoki K, Yamada M, Yukinaga H, Fujita Y, Kamioka Y, Matsuda M (2011) Development of an optimized backbone of FRET biosensors for kinases and GTPases. Mol Biol Cell 22(23):4647–4656

Lubeck E, Cai L (2012) Single-cell systems biology by super-resolution imaging and combinatorial labeling. Nat Methods 9(7):743–748

Mettetal JT, Muzzey D, Gomez-Uribe C, van Oudenaarden A (2008) The frequency dependence of osmo-adaptation in Saccharomyces cerevisiae. Science 319(5862):482–484

Miller-Jensen K, Janes KA, Brugge JS, Lauffenburger DA (2007) Common effector processing mediates cell-specific responses to stimuli. Nature 448(7153):604–608

Miyawaki A (2011) Development of probes for cellular functions using fluorescent proteins and fluorescence resonance energy transfer. Annu Rev Biochem 80:357–373

Molinelli EJ, Korkut A, Wang W, Miller ML, Gauthier NP, Jing X, Kaushik P, He Q, Mills G, Solit DB, Pratilas CA, Weigt M, Braunstein A, Pagnani A, Zecchina R, Sander C (2013) Perturbation biology: inferring signaling networks in cellular systems. PLoS Comput Biol 9 (12):e1003290

Morris MK, Saez-Rodriguez J, Clarke DC, Sorger PK, Lauffenburger DA (2011) Training signaling pathway maps to biochemical data with constrained fuzzy logic: quantitative analysis of liver cell responses to inflammatory stimuli. PLoS Comput Biol 7(3):e1001099

Nakakuki T, Birtwistle MR, Saeki Y, Yumoto N, Ide K, Nagashima T, Brusch L, Ogunnaike BA, Okada-Hatakeyama M, Kholodenko BN (2010) Ligand-specific c-Fos expression emerges from the spatiotemporal control of ErbB network dynamics. Cell 141(5):884–896

Neves SR, Iyengar R (2002) Modeling of signaling networks. BioEssays 24(12):1110–1117

Poulikakos PI, Zhang C, Bollag G, Shokat KM, Rosen N (2010) RAF inhibitors transactivate RAF dimers and ERK signalling in cells with wild-type BRAF. Nature 464(7287):427–430

Pratilas CA, Taylor BS, Ye Q, Viale A, Sander C, Solit DB, Rosen N (2009) (V600E)BRAF is associated with disabled feedback inhibition of RAF-MEK signaling and elevated transcriptional output of the pathway. Proc Natl Acad Sci USA 106(11):4519–4524

Raj A, Peskin CS, Tranchina D, Vargas DY, Tyagi S (2006) Stochastic mRNA synthesis in mammalian cells. PLoS Biol 4(10):e309

Ran FA, Hsu PD, Wright J, Agarwala V, Scott DA, Zhang F (2013) Genome engineering using the CRISPR-Cas9 system. Nat Protoc 8(11):2281–2308

Raue A, Schilling M, Bachmann J, Matteson A, Schelke M, Kaschek D, Hug S, Kreutz C, Harms BD, Theis FJ, Klingmuller U, Timmer J (2013) Lessons learned from quantitative dynamical modeling in systems biology. PLoS ONE 8(9):e74335

Returns on R&D investments continue to fall (2014). Nat Rev Drug Discov 13 (1):9 doi:10.1038/nrd4224

Reyon D, Tsai SQ, Khayter C, Foden JA, Sander JD, Joung JK (2012) FLASH assembly of TALENs for high-throughput genome editing. Nat Biotechnol 30(5):460–465

Sachs K, Perez O, Pe'er D, Lauffenburger DA, Nolan GP (2005) Causal protein-signaling networks derived from multiparameter single-cell data. Science 308(5721):523–529

Salis H, Sotiropoulos V, Kaznessis YN (2006) Multiscale Hy3S: hybrid stochastic simulation for supercomputers. BMC Bioinform 7:93

Saltelli A (2008) Global sensitivity analysis: the primer. Wiley, Chichester

Schoeberl B, Eichler-Jonsson C, Gilles ED, Muller G (2002) Computational modeling of the dynamics of the MAP kinase cascade activated by surface and internalized EGF receptors. Nat Biotechnol 20(4):370–375

Schoeberl B, Pace EA, Fitzgerald JB, Harms BD, Xu L, Nie L, Linggi B, Kalra A, Paragas V, Bukhalid R, Grantcharova V, Kohli N, West KA, Leszczyniecka M, Feldhaus MJ, Kudla AJ, Nielsen UB (2009) Therapeutically targeting ErbB3: a key node in ligand-induced activation of the ErbB receptor-PI3K axis. Sci Signal 2(77):ra31

Schwanhausser B, Busse D, Li N, Dittmar G, Schuchhardt J, Wolf J, Chen W, Selbach M (2011) Global quantification of mammalian gene expression control. Nature 473(7347):337–342

Shahrezaei V, Swain PS (2008) Analytical distributions for stochastic gene expression. Proc Natl Acad Sci USA 105(45):17256–17261

Simon B, Sandhu M, Myhr KL (2010) Live FISH: imaging mRNA in living neurons. J Neurosci Res 88(1):55–63

Sneddon MW, Faeder JR, Emonet T (2011) Efficient modeling, simulation and coarse-graining of biological complexity with NFsim. Nat Methods 8(2):177–183

Sobol IM (2001) Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates. Math Comput Simulat 55(1–3):271–280

Sondergaard JN, Nazarian R, Wang Q, Guo D, Hsueh T, Mok S, Sazegar H, MacConaill LE, Barretina JG, Kehoe SM, Attar N, von Euw E, Zuckerman JE, Chmielowski B, Comin-Anduix B, Koya RC, Mischel PS, Lo RS, Ribas A (2010) Differential sensitivity of melanoma cell lines with BRAFV600E mutation to the specific Raf inhibitor PLX4032. J Transl Med 8:39

Sorokina O, Sorokin A, Douglas Armstrong J, Danos V (2013) A simulator for spatially extended kappa models. Bioinformatics 29(23):3105–3106

Sturm OE, Orton R, Grindlay J, Birtwistle M, Vyshemirsky V, Gilbert D, Calder M, Pitt A, Kholodenko B, Kolch W (2010) The mammalian MAPK/ERK pathway exhibits properties of a negative feedback amplifier. Sci Signal 3(153):ra90

Sullivan RJ, Flaherty K (2013) MAP kinase signaling and inhibition in melanoma. Oncogene 32 (19):2373–2379

Toettcher JE, Gong D, Lim WA, Weiner OD (2011) Light-based feedback for controlling intracellular signaling dynamics. Nat Meth 8(10):837–839

Toettcher JE, Weiner OD, Lim WA (2013) Using optogenetics to interrogate the dynamic control of signal transmission by the ras/erk module. Cell 155(6):1422–1434

Vallabhajosyula RR, Chickarmane V, Sauro HM (2006) Conservation analysis of large biochemical networks. Bioinformatics 22(3):346–353

von Kriegsheim A, Baiocchi D, Birtwistle M, Sumpton D, Bienvenut W, Morrice N, Yamada K, Lamond A, Kalna G, Orton R, Gilbert D, Kolch W (2009) Cell fate decisions are specified by the dynamic ERK interactome. Nat Cell Biol 11(12):1458–1464

Vyshemirsky V, Girolami M (2008) BioBayes: a software package for Bayesian inference in systems biology. Bioinformatics 24(17):1933–1934

Waltemath D, Adams R, Beard DA, Bergmann FT, Bhalla US, Britten R, Chelliah V, Cooling MT, Cooper J, Crampin EJ, Garny A, Hoops S, Hucka M, Hunter P, Klipp E, Laibe C, Miller AK, Moraru I, Nickerson D, Nielsen P, Nikolski M, Sahle S, Sauro HM, Schmidt H, Snoep JL, Tolle D, Wolkenhauer O, Le Novere N (2011) Minimum information about a simulation experiment (MIASE). PLoS Comput Biol 7(4):e1001122

Wang CC, Cirit M, Haugh JM (2009) PI3K-dependent cross-talk interactions converge with Ras as quantifiable inputs integrated by Erk. Mol Syst Biol 5:246

Xu TR, Vyshemirsky V, Gormand A, von Kriegsheim A, Girolami M, Baillie GS, Ketley D, Dunlop AJ, Milligan G, Houslay MD, Kolch W (2010) Inferring signaling pathway topologies from multiple perturbation measurements of specific biochemical species. Sci Signal 3(134):ra20

Xu X, Zhang Y, Williams J, Antoniou E, McCombie WR, Wu S, Zhu W, Davidson NO, Denoya P, Li E (2013) Parallel comparison of Illumina RNA-Seq and Affymetrix microarray platforms on transcriptomic profiles generated from 5-aza-deoxy-cytidine treated HT-29 colon cancer cells and simulated datasets. BMC Bioinf 14(Suppl 9):S1

Yoon S, Seger R (2006) The extracellular signal-regulated kinase: multiple substrates regulate diverse cellular functions. Growth Factors 24(1):21–44

Zhang XY, Birtwistle MR, Gallo JM (2014) A general network pharmacodynamic model-based design pipeline for customized cancer therapy applied to the VEGFR pathway. CPT Pharmacomet Syst Pharmacol 3:e92

Zhou Y, Liepe J, Sheng X, Stumpf MP, Barnes C (2011) GPU accelerated biochemical network simulation. Bioinformatics 27(6):874–876