Published in final edited form as:

Methods Mol Biol. 2015 ; 1244: 259–276. doi:10.1007/978-1-4939-1878-2_12.

Using Computational Modeling and Experimental Synthetic Perturbations to Probe Biological Circuits

Joshua R. Porter¹ and Eric Batchelor¹

¹Laboratory of Pathology, Center for Cancer Research, National Cancer Institute, National Institutes of Health, Building 10, Room B1B42, 10 Center Dr., MSC 1500, Bethesda, MD, 20892, USA.

Abstract

This chapter describes approaches for using computational modeling of synthetic biology perturbations to analyze endogenous biological circuits, with a particular focus on signaling and metabolic pathways. We describe a bottom-up approach in which ordinary differential equations are constructed to model the core interactions of a pathway of interest. We then discuss methods for modeling synthetic perturbations that can be used to investigate properties of the natural circuit. Keeping in mind the importance of the interplay between modeling and experimentation, we next describe experimental methods for constructing synthetic perturbations to test the computational predictions. Finally, we present a case study of the p53 tumor-suppressor pathway, illustrating the process of modeling the core network, designing informative synthetic perturbations in silico, and testing the predictions in vivo.

Keywords

Synthetic biology; Dynamical systems; Computational modeling; Experimental design; Signal transduction; Metabolism

1 Introduction

Synthetic biology provides a powerful approach for generating novel biological devices that operate on a molecular scale. These devices have most often been developed using genetically tractable microorganisms as a chassis into which biological parts orthogonal to the host organism can be engineered. Most synthetic biology efforts thus far have been focused on one of two broad tasks: creating small-scale devices that perform a desired computation or function, or engineering the metabolism of a microorganism for the synthesis of useful compounds. Predictive computational models have been instrumental in designing synthetic circuits that achieve these goals.

Based on the principle “What I cannot create, I do not understand” (Richard Feynman), synthetic biologists seek to design and experimentally implement circuits that produce a desired, predicted behavior. This, in turn, helps one understand how natural biological circuits generate such behaviors mechanistically. Several proof-of-principle devices have been constructed to perform simple tasks, including generating oscillatory outputs [1] and performing logic computations [2]. Quantitative modeling has aided the construction of even

the simplest circuits during both initial circuit design and circuit refinement. For example, modeling the repressilator [1] indicated that the individual protein components of the circuit needed to be rapidly degraded. Accordingly, the addition of protein degradation tags was necessary to implement the circuit successfully in vivo. Subsequent, more detailed modeling has led to further improvement of minimal transcriptional oscillators [3, 4].

Metabolic pathways are complex even in the simplest organisms; therefore, manipulation of such pathways for novel purposes can lead to non-intuitive, unanticipated results. Quantitative models of metabolic pathways are required not only to identify the best avenues for pathway manipulation but also to interpret the large amounts of data generated by metabolic studies. Computational approaches, including flux balance analysis, have proven invaluable for a wide range of such applications, from identifying a naturally occurring secondary metabolite and increasing its production [5] to improving methods for synthesizing heterologous compounds such as anti-malarial drugs [6].

In addition to providing methods for constructing new biological devices and engineering novel metabolites, synthetic biology provides a means of analyzing complex naturally occurring cellular networks. By making directed synthetic perturbations, one can probe the structure and function of endogenous signaling and metabolic pathways. At the simplest level, such perturbations might involve genetic knockouts, transcript knockdowns by RNA interference, overexpression of pathway components, or temperature-sensitive mutations of proteins. Such perturbations have been widely used by biologists for decades. More complex perturbations can enable a more accurate probing of natural circuit function, allowing one to plug in new network connections, bypass or eliminate existing connections, or tune interactions between circuit parts. Such approaches include mutation of transcription factor binding sequences at target promoters, construction of hybrid allosteric switches in proteins, or introduction of novel multicomponent network feedbacks.

In this chapter, we describe general methods for using synthetic approaches to probe the function of endogenous signaling and metabolic networks, focusing on the role of computational modeling in this process. We detail procedures for defining the natural cellular circuit of interest, constructing a computational model of the natural circuit, identifying possible perturbations to generate a desired novel output or probe regulatory connections, and designing experiments to test in vivo the perturbations designed in silico.

2 Methods

2.1 Modeling the Natural Circuit

Before one can reengineer a natural biological system, one must have sufficient quantitative understanding of how the natural system works. The best way to express and assess this quantitative understanding is to recreate the system in silico with a mathematical model.

Here we take a “bottom-up” modeling approach, which involves experimentally examining individual parts of a system and their interactions as modules, then combining the tractable modules to construct a larger, more complete model. This approach contrasts with “top-down” modeling techniques, in which measurements of many biological quantities under

different perturbations of the system are used to infer the underlying system structure (see ref. 7 for one review). Top-down modeling techniques assume by default that the measurements provided are sufficient to identify what matters in the structure of a system. By contrast, the bottom-up approach does not assume this, and it has been our experience that the process of bottom-up modeling identifies which important details are still unknown, enabling a more systematic and comprehensive analysis of the key components of a system. For this reason, we believe that bottom-up modeling is better for building the sort of model required to re-engineer a biological system, as such systems frequently contain redundancies that are not apparent during normal system operation but may make a difference when the system is synthetically modified.

Within bottom-up modeling, we will also restrict our discussion to systems in which all chemical species involved are present in large numbers (on the order of 10^2 – 10^3 or greater) and can be assumed to be well mixed within their containing compartment(s) [8]. Such systems can reasonably be modeled using ordinary differential equations (ODEs), which are relatively simple to construct and solve. Systems with some chemical species present in very small numbers are best described using stochastic models; systems in which diffusion through space is a concern can be modeled using partial differential equations. These more complex types of models are useful but beyond the scope of this discussion.

The first step in modeling a biological system from the bottom up is to identify the relevant known parts of the system of interest. We define a part as a distinct biochemical species located in a particular reaction compartment. We then identify the processes in the system that change the concentration of the individual parts.

Table 1 lists common processes that are important in signal transduction and metabolic pathways. The most fundamental processes of interest are two parts binding to form a complex and the complex unbinding to form its constituent parts. Binding and unbinding processes can be combined in complex ways to form higher-level processes such as production, degradation, and enzyme catalysis. Other important processes include passive transport, the diffusion of parts down a concentration gradient between compartments, and the dilution of parts due to the exponential growth of their containing compartments. As a general rule, the less detail one understands about a process, the higher level an abstraction one uses to model it.

Once parts and processes have been identified, this information can be organized by combining it into a diagram using the conventional notation of symbols connected by arrows, as shown in Table 1. A simple diagram can be a powerful tool for understanding a biological system beyond the reductionist perspective of considering one part at a time, while at the same time it provides a nonmathematical description of a system that is accessible to a nontechnical audience.

When parts and processes have been identified and organized, the next step is to translate this information into a set of differential equations. In general, for every part in the system, this will yield one differential equation of the form $\frac{d[\text{part}]}{dt} = \sum \text{process rates}$, where $[\text{part}]$ is the concentration of the part and the right side of the equation is a summation of terms for

the rate of each process acting on that part. Table 1 shows these rate terms for each of the common processes listed.

In general, the system of study will be of sufficient complexity that the set of differential equations describing it will not have an analytical solution. Instead, a numerical solution must be obtained. Several numerical computing packages, including MATLAB and Mathematica, can readily solve such systems of differential equations. Using the ODE solvers in these packages simply requires writing the differential equations in the appropriate coding format and specifying initial conditions for the concentration of each part. Several other software packages, including little b [9], BioNetGen [10], Kappa [11] (see Chapter 6), Simmune [12], and PySB [13], have been written specifically for the analysis of biochemical networks. These packages help to automate and organize the process of model-building, which is increasingly useful as more complex systems are considered. In general, the various solvers can be used to explore both steady-state and dynamical behavior of biological circuits in silico.

Biochemical systems are often *stiff*, meaning that they are composed of processes that operate on different time scales. In particular, binding and unbinding processes are generally much faster than production, degradation, catalysis, transport, etc. Stiff systems require special techniques to solve computationally in an efficient manner. One convenient solution is to perform a separation of time scales, in which one considers the faster processes to be at steady state. This simplification can reduce both the number of equations in the model and the number of parameters required, as a single dissociation constant can replace two rate constants for binding and unbinding. As a result, a numerical solution can be computed faster, and in some cases an analytical solution can be determined.

Finding realistic parameter values is often one of the most challenging aspects of model building. The best choice for a parameter is the one connected as closely as possible with the simplest, best understood physical reality. Ideally, this would be obtained from a direct biochemical measurement. However, biochemical parameters are often not available in the literature and cannot always be directly measured by the researchers performing the modeling. In such cases, one must choose parameters such that the solution of the equations matches experimental observations. This is one reason why a bottom-up approach to model construction is often more feasible than a top-down approach, which is by nature dependent on a greater number of unknown parameters. As more biochemical details become known, the model can be refined with additional complexity as needed.

One can assess the quality of a model by how well its solution matches experimental observations. It is important to remember that models are not “right” or “wrong” so much as they are more or less useful—any model is a deliberately simplified representation of reality. The chances that a model will be useful can be improved by testing it against different experimental observations, ideally including some observations that were not used to determine parameters.

2.2 Probing Possible Synthetic Changes In Silico

Having constructed a computational model of the natural circuit of interest, one can use the model to make testable predictions about circuit performance. Confirming these predictions experimentally validates one's understanding of the circuit. When designing a synthetic circuit for a specific engineering task, a model is useful for identifying necessary connections and parameter operating regimes that give rise to a desired functional output. Similarly, when studying a natural circuit, one can use the model to identify the parts or parameter values necessary for a desired biological outcome. How dependent is the yield of a metabolite on the degradation rate of a metabolic intermediate? If there were a negative feedback on the production of phosphatase A, would signaling through kinase B be attenuated at a later point? Such questions can readily be explored in silico and experimentally tested in vivo with synthetic perturbations.

We can consider synthetic perturbations as falling into the following broad categories:

Synthesis Perturbations: The synthesis rates of individual parts are modified, or novel methods for increasing the (real or effective) concentration of parts are generated.

Degradation Perturbations: The degradation rates of parts are modified, or new methods for decreasing the (real or effective) concentration of parts are generated.

Interaction Perturbations: The interactions between natural parts are modified.

Novel Regulatory Connections: New processes (possibly involving additional synthetic parts) connecting existing parts are generated. These new processes can be broadly classified as positive feedback, negative feedback, coherent feedforward, or incoherent feedforward based on which parts influence which other parts and whether that influence is activating or repressing (Fig. 1) [14].

Once a change to the system is conceived, it can be tested in the model by altering parameters or adding and removing rate terms as described earlier and simulating the new model.

During the process of model exploration, it is important to keep in mind which types of synthetic manipulations are experimentally feasible; otherwise, the model predictions will remain untestable hypotheses and little will be learned about the natural biological circuit. It is also important to keep synthetic perturbations as simple as possible. While a Rube Goldberg machine may look exciting on paper, it rarely yields informative results in the laboratory.

2.3 Implementing Changes In Vivo

Having developed a quantitative model of the biological circuit of interest and identified possible perturbations that will yield insight into the function of the circuit, one can perturb the actual circuit using a variety of experimental techniques. Several synthetic biology tools are readily available to modify a biological circuit at the level of individual genes, mRNAs, or proteins, and these perturbations can be combined to generate multi-component

regulatory loops and modules. Here we briefly mention a few such tools for perturbing signaling or metabolic pathways.

The introduction of synthetic perturbations often depends on precise manipulation of a model organism's genome. Several methods are available for deleting or inserting DNA sequences. Random insertion of genetic material using electroporation, lipid-based transfection, or viral-mediated infection procedures is relatively easy. Directed genome manipulations can present more of a challenge, and the difficulty tends to scale with the complexity of the model organism being studied. Regardless of the target, the basic techniques are similar. The most widely used methods for directed editing are the phage recombinase systems, including lambda Red, FLP-FRT, and Cre/*lox* recombination [15, 16]. Recent advances in the use of engineered nucleases, such as TALENs and zinc-finger nucleases [17], offer other powerful methods for genome editing in a variety of organisms.

It is often desirable to control the synthesis rates of biological parts when testing model predictions. This control can be performed *in vivo* by altering expression of natural or synthetic parts through the use of regulatable promoters. Several different inducible or repressible expression systems are available from prokaryotic or eukaryotic origins. The well-characterized *lac*- and *ara*-inducible systems from bacteria are regulated by the presence of sugars and sugar analogs such as isopropyl β -D-1-thiogalactopyranoside (IPTG) [18]. The *GAL1* and *CUP1* promoters, which are induced in response to galactose and copper, respectively, are examples of systems effective in yeast [19]. In mammalian cells, one of the most widely used systems is the tetracycline-responsive system [20]. This system is a particularly effective tool in that versions for induction in either the presence or the absence of tetracycline/doxycycline have been developed, and they can be used in a variety of organisms. These are but a few of the promoters available, and they all provide powerful synthetic methods for regulating both the level and the timing of gene expression.

For changing the degradation rates of endogenous mRNAs, one can use RNA interference (RNAi) technologies. Originally developed in *C. elegans* [21] and now widely used in mammalian cell culture, RNAi can be used for transient knockdown of target mRNAs in a variety of systems. For example, transfection of small interfering RNAs (siRNAs) into mammalian cells enables the knockdown of endogenous mRNAs for up to several days. Long-term and regulatable knockdown of mRNAs can be achieved by engineering small hairpin RNAs (shRNAs), expressed from either constitutive or inducible promoters, directly into the model system's genome.

The degradation rate of proteins can also be modified synthetically. The addition of degradation tags such as PEST sequences [22] can shorten protein lifetimes. One notable way to regulate protein degradation more tightly is to use the *ssrA* proteasomal machinery from *E. coli* in an orthogonal organism [23]. By altering the concentration of the degradation machinery via an inducible promoter, this system enables tunable regulation of degradation rates for specific proteins within a network of interest.

Beyond direct control of protein degradation, it is possible to synthetically alter the effective concentration of a protein. In particular, changing protein localization can be a means of

changing protein concentration in a subcellular compartment of interest. For example, mutating nuclear localization sequences or nuclear exclusion signals can alter protein transport into or out of the nucleus. This synthetic manipulation may be particularly useful for eukaryotic transcription factors, as their gene regulatory activity depends on their nuclear localization.

The next major class of synthetic perturbations one can consider is altering the interactions between endogenous cellular parts. Many mechanisms exist for making such alterations. For example, manipulating organelle localization tags and membrane tethers allows additional regulation of protein interactions by altering the effective protein concentration. Such changes can also be effected by using modified scaffold proteins or direct synthetic linkage of protein domains to increase the probability of interactions occurring. For example, Bashor et al. [24] used a scaffold modification approach to dissect the function of the yeast mating MAPK pathway in *S. cerevisiae*. Linking proteins or protein domains is also an effective way to generate novel allosteric switches for a variety of purposes.

Recent developments in optogenetics also hold great promise as tools for synthetic manipulation of biological circuits. Optogenetics was originally developed as a method to control nerve cell activation noninvasively using light-activated ion channels [25]. Recently, additional photoreceptor domains, such as the LOV domain from *A. sativa*, have been engineered to control signaling molecules, providing precise temporal and spatial control of protein activation. For example, Wu et al. [26] fused the LOV domain to forms of the GTPase Rac1, creating a hybrid protein that could modulate cell motility in a reversible and repeatable manner.

Ultimately, synthetic manipulations at the genetic, mRNA, and protein levels can be combined to engineer large-scale, multicomponent network connections. For example, new feedback and feed-forward motifs can be wired into natural signaling circuits to control the magnitude or duration of a signal. The possibilities for generating such manipulations experimentally are limitless, although they depend on the natural circuit one is interested in modifying. For illustrative purposes, we will describe an example of such rewirings in the following case study.

3 Case Study: The p53 Signaling Pathway

The p53 signaling pathway in human cells has been extensively studied, as it plays a critical role in preventing tumor formation. p53 is a transcription factor that is activated by different forms of cell stress and regulates the expression of over a hundred genes, impacting several different processes including DNA repair, apoptosis, cell cycle arrest, and senescence [27]. One gene upregulated by p53 codes for the E3 ubiquitin ligase Mdm2, which tags p53 for degradation, thereby forming a negative feedback loop (Fig. 2a). When cells are γ -irradiated to generate DNA double-strand breaks (DSBs), the concentration of p53 and Mdm2 in the nucleus changes in a series of oscillatory pulses characterized by fixed amplitude and timing [28].

Toettcher et al. [29] sought to modify the p53 pathway to alter the expression dynamics of p53 and Mdm2. The goal of these modifications was to better understand the mechanisms generating the dynamics and to gain insight into their functional consequences. Since treating cells with γ -radiation triggers a complex response involving a larger network of components, they first sought a way to activate p53 that would bypass this complexity. This bypass took the form of a cell line in which p53 fused to CFP was expressed from the zinc-inducible rat metallothionein promoter. Using this synthetic system enabled direct manipulation of *p53* transcription independent of the endogenous *p53* promoter. Treating the cells with different concentrations of zinc chloride (ZnCl_2) maintained the oscillatory dynamics of p53 and Mdm2 expression, although the oscillations were damped, in contrast to the undamped oscillations observed in the response to γ -radiation. In this way, Toettcher et al. were able to use a synthetic bypass of the full natural DSB response, effectively narrowing the scope of the biological system they needed to consider.

Having established a simplified experimental system with which to study the core process of interest (Fig. 2a), Toettcher et al. next constructed a computational model to explore additional possible synthetic perturbations. The parts of the simplified system, in addition to p53 and Mdm2, included ZnCl_2 , which promotes p53 production, and Nutlin3A, a small molecule that inhibits the interaction between Mdm2 and p53. Four key processes connected these four biochemical parts. First, p53 is produced at a constant rate plus an additional zinc-dependent rate. Second, Mdm2 catalyzes the ubiquitination of p53, tagging it for subsequent degradation; this catalysis is competitively inhibited by Nutlin3A. Third, p53 promotes the transcription of the *mdm2* gene, leading to an increased Mdm2 protein concentration with a delay due to protein translation, folding, and nuclear localization. Fourth, Mdm2 catalyzes its own ubiquitination, leading to its own degradation. Figure 2b graphically summarizes these parts and processes.

The next step was to translate this information about parts and processes into a set of differential equations:

p53 protein:

$$\frac{d[P]}{dt} = \underbrace{\xi_p(t)}_{\text{noise}} \left(\underbrace{\alpha_p + \frac{p_Z[Z]^3}{K_Z^3 + [Z]^3}}_{\text{ZnCl}_2 - \text{dependent production}} \right) - \underbrace{\frac{\delta_p[M][P]}{K_p(1 + [N]) + [P]}}_{\text{Mdm2-mediated degradation}}$$

mdm2 mRNA:

$$\frac{d[M_0]}{dt} = \gamma_{m0} \left(\underbrace{\alpha_{m0} + \beta_{m0} \frac{[P]^2}{K_{m0}^2 + [P]^2}}_{\text{p53-upregulated transcription}} - \underbrace{[M_0]}_{\text{degradation}} \right)$$

mdm2 mRNA $\rightarrow$ protein
delays:

$$\frac{d[M_i]}{dt} = \gamma_{m0}([M_{i-1}] - [M_i]), \quad i = 1 \dots 4$$

Mdm2 protein:

$$\frac{d[M]}{dt} = \underbrace{\xi_m(t)[M_4]}_{\text{noisy production}} - \underbrace{\frac{\delta_m[M]}{K_m + [M]}}_{\text{autocatalytic degradation}} - \underbrace{\gamma_m[M]}_{\text{constant degradation}}$$

Quantities described by these equations are defined in Table 2. Notably, the equations for p53 and Mdm2 protein incorporate the stochastic processes $\xi_p(t)$ and $\xi_m(t)$ to model noise in production of the p53 and Mdm2 proteins, respectively. To avoid using a stiff delay to model the time-delayed production of Mdm2, Toettcher et al. used a “boxcar” procedure of four separate equations representing intermediates in the processing and translation of *mdm2* mRNA (represented by the equations for $[M_i]$, $i = 1-4$) [29].

Toettcher et al. next found values for the model parameters. They first measured the relationship between zinc concentration and p53 transcription using a cell line in which production of CFP was driven by the same zinc-inducible promoter as that driving p53-CFP in this system. To obtain values for K_z and p_z , they fit a Hill function with $n = 3$ to the relationship between zinc and CFP fluorescence. For the remaining parameters, they used an optimization method to obtain parameter values that minimized the difference between simulated and experimental measurements of p53 and Mdm2 first pulse amplitude, p53 pulse frequency, and p53 pulse damping rate at five different concentrations of zinc. They further validated the optimal parameters by simulating the model with noise and verifying that these simulations were consistent with experimental measurements.

Having constructed the base model, Toettcher et al. then sought to identify methods by which they could perturb various characteristics of p53 oscillations. They first simulated the addition of a synthetic positive (NPF, Fig. 3a) or negative (2NF, Fig. 3b) feedback loop to the core p53-Mdm2 network. Based on their knowledge of possible experimental perturbations, they modeled these synthetic loops as transcriptional targets of p53 that could increase or decrease p53 expression. To do this, they added equations describing concentrations of mRNA ($[F_0]$) and protein ($[F]$) of a putative positive or negative feedback mediator (PFM or NFM), as well as equations modeling the delays between mRNA and protein ($[F_i]$, $i = 1 \dots 4$) (Fig. 3c, d). Moreover, since p53 production was now part of a feedback loop, they made the model of p53 appropriately more complex, adding differential equations describing concentrations of p53 mRNA ($[P_0]$) and modeling the delays between mRNA and protein ($[P_i]$, $i = 1 \dots 4$). The constant-rate degradation of Mdm2 was removed from the model, as the optimal value for γ_m in the old model was found to be 0, and the noise terms were removed as well to consider only the deterministic system. The new set of equations was nearly identical for the NPF and 2NF models, with the only difference being whether the PFM/NFM ($[F]$) activates or represses production of p53 mRNA.

p53 mRNA:

$$\frac{d[P_0]}{dt} = \begin{cases} \gamma_{p0} \left(\underbrace{\alpha_{p0} + \beta_{p0} \left(\frac{[Z]^3}{K_Z^3 + [Z]^3} \right) \left(\frac{[F]^2}{K_f^2 + [F]^2} \right)}_{\text{ZnCl}_2\text{-activated, feedback-activated transcription}} - \underbrace{[P_0]}_{\text{degradation}} \right) & \text{NPF model} \\ \gamma_{p0} \left(\underbrace{\alpha_{p0} + \beta_{p0} \left(\frac{[Z]^3}{K_Z^3 + [Z]^3} \right) \left(\frac{K_f^2}{K_f^2 + [F]^2} \right)}_{\text{ZnCl}_2\text{-activated, feedback-activated transcription}} - \underbrace{[P_0]}_{\text{degradation}} \right) & \text{2NF model} \end{cases}$$

p53 mRNA → protein delays:

$$\frac{d[P_i]}{dt} = \gamma_{p0}([P_{i-1}] - [P_i]), \quad i = 1 \dots 4$$

p53 protein:

$$\frac{d[P]}{dt} = \underbrace{\alpha_p [P_4]}_{\text{production}} - \underbrace{\frac{\delta_p [M][P]}{K_p(1 + [N]) + [P]}}_{\text{Mdm2-mediated degradation}}$$

mdm2 mRNA:

$$\frac{d[M_0]}{dt} = \gamma_{m0} \left(\underbrace{\alpha_{m0} + \beta_{m0} \frac{[P]^2}{K_{m0}^2 + [P]^2}}_{\text{p53-upregulated transcription}} - \underbrace{[M_0]}_{\text{degradation}} \right)$$

Mdm2 mRNA → protein delays:

$$\frac{d[M_i]}{dt} = \gamma_{m0}([M_{i-1}] - [M_i]), \quad i = 1 \dots 4$$

Mdm2 protein:

$$\frac{d[M]}{dt} = \underbrace{[M_4]}_{\text{production}} - \underbrace{\frac{\delta_m [M]}{K_m + [M]}}_{\text{autocatalytic degradation}}$$

Synthetic positive/negative feedback mediator mRNA:

$$\frac{d[F_0]}{dt} = \gamma_{f0} \left(\underbrace{\alpha_{f0} + \beta_{f0} \frac{[P]^2}{K_{f0}^2 + [P]^2}}_{\text{p53-upregulated transcription}} - \underbrace{[F_0]}_{\text{degradation}} \right)$$

Synthetic positive/negative feedback mediator mRNA → protein delays:

$$\frac{d[F_i]}{dt} = \gamma_{f0}([F_{i-1}] - [F_i])$$

Synthetic positive/negative feedback mediator protein:

$$\frac{d[F]}{dt} = \underbrace{[F_4]}_{\text{production}} - \underbrace{\gamma_f [F]}_{\text{degradation}}$$

All quantities described by these equations are defined in Table 3. In this modified model, Toettcher et al. kept the original model parameters governing the interactions between p53 and Mdm2, as this aspect of the model did not change. Using a range of feedback strengths and delay times, they simulated the model to probe the effects of the new feedbacks on p53 dynamics. They found that the additional feedback loops affected oscillation amplitude and

damping rate but not frequency. Further simulations revealed that changes to p53 pulse frequency could be obtained by modulating the concentration of Nutlin3A.

Toettcher et al. next established experimental systems to validate the predictions from the model of the synthetic-natural circuits. To generate a synthetic positive feedback loop, they constructed a cell line in which MTF1, the zinc-responsive transcription factor that acts directly on the zinc-inducible promoter, was expressed from a p53-responsive promoter (NPF, Fig. 3a). Thus, p53 upregulated MTF1 production, which in turn upregulated p53 production via the zinc-inducible promoter, forming the requisite synthetic positive feedback loop (Fig. 3c). In a separate design, to generate a synthetic negative feedback in addition to the core network, they used MTF1-KRAB, a dominant-negative form of MTF1, expressed from a p53-dependent promoter (2NF, Fig. 3b). In this cell line, p53 upregulated MTF1-KRAB production, and MTF1-KRAB repressed p53 production by competing with endogenous MTF1, forming a new negative feedback loop (Fig. 3d). Stimulation of these cells with ZnCl_2 confirmed that the damping rates of the oscillations, but not the frequency, were altered by the new feedback loops, as predicted by the computational model [29]. Additional experiments showed that perturbing the base experimental system with Nutlin3A altered oscillation frequency as predicted by the model [29].

4 Conclusions

While much attention in synthetic biology research has focused on engineering new biological devices, in this chapter we have described a method for using synthetic biology approaches to address questions in basic research. It is important to remember that the combined computational and experimental approach described here is iterative—experimental observations are used to build a quantitative model, which is used to generate new hypotheses to test experimentally, etc. It is hoped that through each iteration, one can better understand the network structure and the function of the natural biological circuit. For modeling and manipulating such circuits, the computational methods and experimental tools described here are by no means a complete catalog of the synthetic biology toolbox. However, they are broadly applicable for tackling a variety of problems in signal transduction and metabolic research, as illustrated in the p53 case study. As new methods and tools are developed, they can readily be applied to the general framework described here.

Acknowledgment

This work was supported by the Intramural Research Program of the Center for Cancer Research, National Cancer Institute, National Institutes of Health.

References

1. Elowitz MB, Leibler S (2000) A synthetic oscillatory network of transcriptional regulators. *Nature* 403(6767):335–338. doi:10.1038/35002125 [PubMed: 10659856]
2. Benenson Y (2012) Biomolecular computing systems: principles, progress and potential. *Nat Rev Genet* 13(7):455–468. doi:10.1038/nrg3197 [PubMed: 22688678]

3. Atkinson MR, Savageau MA, Myers JT, Ninfa AJ (2003) Development of genetic circuitry exhibiting toggle switch or oscillatory behavior in *Escherichia coli*. *Cell* 113(5):597–607. doi: 10.1016/S0092-8674(03)00346-5 [PubMed: 12787501]
4. Stricker J, Cookson S, Bennett MR, Mather WH, Tsimring LS, Hasty J (2008) A fast, robust and tunable synthetic gene oscillator. *Nature* 456(7221):516–519. doi:10.1038/nature07389 [PubMed: 18971928]
5. Nguyen QT, Merlo ME, Medema MH, Jankevics A, Breitling R, Takano E (2012) Metabolomics methods for the synthetic biology of secondary metabolism. *FEBS Lett* 586(15):2177–2183. doi: 10.1016/j.febslet.2012.02.008 [PubMed: 22710183]
6. Keasling JD (2012) Synthetic biology and the development of tools for metabolic engineering. *Metab Eng* 14(3):189–195. doi:10.1016/j.ymben.2012.01.004 [PubMed: 22314049]
7. Lee WP, Tzou WS (2009) Computational methods for discovering gene networks from expression data. *Brief Bioinform* 10(4):408–423. doi:10.1093/bib/bbp028 [PubMed: 19505889]
8. Chen WW, Niepel M, Sorger PK (2010) Classic and contemporary approaches to modeling biochemical reactions. *Genes Dev* 24(17):1861–1875. doi:10.1101/gad.1945410 [PubMed: 20810646]
9. Mallavarapu A, Thomson M, Ullian B, Gunawardena J (2009) Programming with models: modularity and abstraction provide powerful capabilities for systems biology. *J R Soc Interface* 6(32):257–270. doi:10.1098/rsif.2008.0205 [PubMed: 18647734]
10. Hlavacek WS, Faeder JR, Blinov ML, Posner RG, Hucka M, Fontana W (2006) Rules for modeling signal-transduction systems. *Sci STKE* 2006 (344):re6. doi:10.1126/stke.3442006re6 [PubMed: 16849649]
11. Danos V, Feret J, Fontana W, Harmer R, Kri-vine J (2008) Rule-based modeling, symmetries, refinements In: *Formal methods in systems biology*, vol 5054, *Lecture notes in bioinformatics* 2008. Springer, Berlin, pp 103–122
12. Meier-Schellersheim M, Xu X, Angermann B, Kunkel EJ, Jin T, Germain RN (2006) Key role of local regulation in chemosensing revealed by a new molecular interaction-based modeling method. *PLoS Comput Biol* 2(7):e82. doi:10.1371/journal.pcbi.0020082 [PubMed: 16854213]
13. Lopez CF, Muhlich JL, Bachman JA, Sorger PK (2013) Programming biological models in Python using PySB. *Mol Syst Biol* 9:646. doi:10.1038/msb.2013.1 [PubMed: 23423320]
14. Alon U (2007) Network motifs: theory and experimental approaches. *Nat Rev Genet* 8(6):450–461. doi:10.1038/nrg2102 [PubMed: 17510665]
15. Murphy KC (2012) Phage recombinases and their applications. *Adv Virus Res* 83:367–414. doi: 10.1016/B978-0-12-394438-2.00008-6 [PubMed: 22748814]
16. Turan S, Zehe C, Kuehle J, Qiao J, Bode J (2012) Recombinase-mediated cassette exchange (RMCE)—a rapidly-expanding tool-box for targeted genomic modifications. *Gene* 515(1):1–27. doi:10.1016/j.gene.2012.11.016 [PubMed: 23201421]
17. Gaj T, Gersbach CA, Barbas CF, 3rd (2013) ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering. *Trends Biotechnol* 31(7):397–405. doi:10.1016/j.tibtech.2013.04.004 [PubMed: 23664777]
18. Terpe K (2006) Overview of bacterial expression systems for heterologous protein production: from molecular and biochemical fundamentals to commercial systems. *Appl Microbiol Biotechnol* 72(2):211–222. doi:10.1007/s00253-006-0465-8 [PubMed: 16791589]
19. Maya D, Quintero MJ, de la Cruz Munoz-Centeno M, Chavez S (2008) Systems for applied gene control in *Saccharomyces cerevisiae*. *Biotechnol Lett* 30(6):979–987. doi:10.1007/s10529-008-9647-z [PubMed: 18239858]
20. Gossen M, Bujard H (1992) Tight control of gene expression in mammalian cells by tetracycline-responsive promoters. *Proc Natl Acad Sci U S A* 89(12):5547–5551 [PubMed: 1319065]
21. Fire A, Xu S, Montgomery MK, Kostas SA, Driver SE, Mello CC (1998) Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature* 391 (6669):806–811. doi:10.1038/35888 [PubMed: 9486653]
22. Rogers S, Wells R, Rechsteiner M (1986) Amino acid sequences common to rapidly degraded proteins: the PEST hypothesis. *Science* 234(4774):364–368 [PubMed: 2876518]

23. Grilly C, Stricker J, Pang WL, Bennett MR, Hasty J (2007) A synthetic gene network for tuning protein degradation in *Saccharomyces cerevisiae*. *Mol Syst Biol* 3:127. doi:10.1038/msb4100168 [PubMed: 17667949]
24. Bashor CJ, Helman NC, Yan S, Lim WA (2008) Using engineered scaffold interactions to reshape MAP kinase pathway signaling dynamics. *Science* 319(5869):1539–1543. doi:10.1126/science.1151153 [PubMed: 18339942]
25. Szobota S, McKenzie C, Janovjak H (2013) Optical control of ligand-gated ion channels. *Methods Mol Biol* 998:417–435. doi:10.1007/978-1-62703-351-0_32 [PubMed: 23529448]
26. Wu YI, Frey D, Lungu OI, Jaehrig A, Schlichting I, Kuhlman B, Hahn KM (2009) A genetically encoded photoactivatable Rac controls the motility of living cells. *Nature* 461 (7260):104–108. doi:10.1038/nature08241 [PubMed: 19693014]
27. Riley T, Sontag E, Chen P, Levine A (2008) Transcriptional control of human p53-regulated genes. *Nat Rev Mol Cell Biol* 9(5):402–412 [PubMed: 18431400]
28. Geva-Zatorsky N, Rosenfeld N, Itzkovitz S, Milo R, Sigal A, Dekel E, Yarnitzky T, Liron Y, Polak P, Lahav G, Alon U (2006) Oscillations and variability in the p53 system. *Mol Syst Biol* 2(2006):0033 [PubMed: 16773083]
29. Toettcher JE, Mock C, Batchelor E, Loewer A, Lahav G (2010) A synthetic-natural hybrid oscillator in human cells. *Proc Natl Acad Sci U S A* 107(39):17047–17052. doi:10.1073/pnas.1005615107 [PubMed: 20837528]

Fig. 1.

Basic network structures. $A \rightarrow B$ means “A activates B”; $A \nrightarrow B$ means “A inhibits B.” Each $\rightarrow$ or $\nrightarrow$ symbol represents at least one process, and possibly additional parts, by which part A influences part B or vice versa

Fig. 2.

Diagrams of the p53 signaling pathway. **(a)** A general depiction of the interactions between p53 and Mdm2 and how those interactions are modulated by external chemical signals, adapted from Toettcher et al. [29]. **(b)** A more detailed diagram, showing specific parts and processes, is often more useful for building a model

Negative and Positive Feedbacks (NPF)

Two Negative Feedbacks (2NF)

Fig. 3.

Diagrams of two p53 signaling pathways with feedback loops added to the original pathway.

(a) The NPF model contains an extra positive feedback loop: p53 promotes PFM (positive feedback mediator) production, which in turn promotes p53 production. Adapted from ref. 29.

(b) The 2NF model contains an extra negative feedback loop: p53 promotes NFM

(negative feedback mediator) production, which in turn inhibits p53 production. Adapted

from ref. 29. (c) The detailed diagram of the NPF model includes one extra part (PFM) and three extra processes: p53-upregulated PFM production, PFM degradation, and PFM-

upregulated p53 production. (d) The detailed diagram of the 2NF model is similar to that of the NPF model except that NFM inhibits rather than promotes p53 production

Table 1

Processes of interest in a biological system and their representations in a mathematical model

Process	Diagram	Rate
Binding	$X + Y \rightarrow XY$	$k_b[X][Y]$
Unbinding	$XY \rightarrow X + Y$	$k_u[XY]$
Production (constant)	$\rightarrow X$	$k_{p,X}$
Degradation	$X \rightarrow \emptyset$ (nothing)	$k_{d,X}[X]$
Catalysis	$\begin{array}{c} E \\ S \xrightarrow{\quad} P \end{array}$	$\frac{k_{\text{cat}}[E][S]}{K_M + [S]}$
Catalysis with competitive inhibition	$\begin{array}{c} E \\ S \xrightarrow{I \dashv} P \end{array}$	$\frac{k_{\text{cat}}[E][S]}{K_M \left(1 + \frac{[I]}{K_I}\right) + [S]}$
Catalysis with noncompetitive inhibition	$\begin{array}{c} E \\ S \xrightarrow{I \dashv} P \end{array}$	$\frac{k_{\text{cat}}[E][S]}{\left(K_M + [S]\right) \left(1 + \frac{[I]}{K_I}\right)}$
Passive transport	$X_A \leftrightarrow X_B$	$k_T([X_B] - [X_A])$
Dilution due to exponential growth		$k_{\text{dil}}[X], k_{\text{dil}} = \frac{\ln(2)}{t_d}$ where t_d is doubling time

Table 2

Symbols and parameters used in the base model of p53 signaling

Symbol	Description
$[P]$	Concentration of p53 protein
$[M_0]$	Concentration of <i>mdm2</i> mRNA
$[M_i], i = 1 \dots 4$	Concentrations of intermediates in Mdm2 protein production
$[M]$	Concentration of Mdm2 protein
$[Z]$	Concentration of $ZnCl_2$
$[N]$	Concentration of Nutlin3A
Parameter	Description
$\xi_p(t)$	Noise in p53 protein production
α_p	p53 protein production rate
ρ_z	$ZnCl_2$ -mediated p53 production rate
K_z	Saturation of $ZnCl_2$ -mediated p53 production
δ_p	Mdm2-mediated p53 degradation rate
K_p	Saturation of Mdm2-mediated p53 degradation
γ_{m0}	Mdm2 production delay
α_{m0}	Basal transcription rate of <i>mdm2</i> mRNA
β_{m0}	p53-mediated transcription rate of <i>mdm2</i> mRNA
K_{m0}	Saturation of p53-mediated Mdm2 production
$\xi_m(t)$	Noise in Mdm2 protein production
δ_m	Mdm2-mediated Mdm2 degradation rate
K_m	Saturation of Mdm2-mediated Mdm2 degradation
γ_m	Mdm2 degradation rate

Table 3

Symbols and parameters used in the modified model of p53 signaling

Symbol	Description
$[P_0]$	Concentration of <i>p53</i> mRNA
$[P_i], i = 1 \dots 4.$	Concentrations of intermediates in p53 protein production
$[P]$	Concentration of p53 protein
$[M_0]$	Concentration of <i>mdm2</i> mRNA
$[M_i], i = 1 \dots 4$	Concentrations of intermediates in Mdm2 protein production
$[M]$	Concentration of Mdm2 protein
$[F_0]$	Concentration of mRNA for synthetic positive/negative feedback mediator
$[F_i], i = 1 \dots 4$	Concentrations of intermediates in production of synthetic positive/negative feedback mediator
$[F]$	Concentration of synthetic positive/negative feedback mediator protein
$[Z]$	Concentration of ZnCl_2
$[N]$	Concentration of Nutlin3A
Parameter	Description
γ_{p0}	p53 production delay
α_{p0}	Basal transcription rate of <i>p53</i> mRNA
β_{p0}	Feedback-mediated transcription rate of <i>p53</i> mRNA
K_z	Saturation of ZnCl_2 -mediated <i>p53</i> transcription
K_f	Saturation of feedback-mediated <i>p53</i> transcription
γ_{p0}	p53 production delay
a_p	p53 protein production rate
δ_p	Mdm2-mediated p53 degradation rate
K_p	Saturation of Mdm2-mediated p53 degradation
γ_{m0}	Mdm2 production delay
α_{m0}	Basal transcription rate of <i>mdm2</i> mRNA
β_{m0}	p53-mediated transcription rate of <i>mdm2</i> mRNA
K_{m0}	Saturation of p53-mediated <i>mdm2</i> transcription
δ_m	Mdm2-mediated Mdm2 degradation rate
K_m	Saturation of Mdm2-mediated Mdm2 degradation
γ_{f0}	Production delay for synthetic positive/negative feedback mediator
α_{f0}	Basal transcription rate of mRNA for synthetic positive/negative feedback mediator
β_{f0}	p53-mediated transcription rate of mRNA for synthetic positive/negative feedback mediator
K_{f0}	Saturation of p53-mediated transcription of synthetic positive/negative feedback mediator
γ_f	Degradation rate of synthetic positive/negative feedback mediator