Relating topology and dynamics in cell signaling networks

by

Jared E. Toettcher

B.S. Bioengineering,
Minor Mathematics
University of California, Berkeley 2004

Submitted to the Department of Biological Engineering
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 2009

Author .....
Department of Biological Engineering
August 3, 2009

Certified by .....
Bruce Tidor
Professor of Biological Engineering and Computer Science
Thesis Supervisor

Certified by .....
Galit Lahav
Professor of Systems Biology, Harvard Medical School
Thesis Supervisor

Accepted by .....
Alan J. Grodzinsky
Professor of Electrical, Mechanical and Biological Engineering
Chair, Course XX Graduate Program Committee

Relating topology and dynamics in cell signaling networks

by

Jared E. Toettcher

Submitted to the Department of Biological Engineering
on August 3, 2009, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Abstract

Cells are constantly bombarded with stimuli that they must sense, process, and interpret to make decisions. This capability is provided by interconnected signaling pathways. Many of the components and interactions within pathways have been identified, and it is becoming clear that the precise dynamics they generate are necessary for proper system function. However, our understanding of how pathways are interconnected to drive decisions is limited. We must overcoming this limitation to develop interventions that can fine tune a cell decision by modulating specific features of its constituent pathway's dynamics.

How can we quantatively map a whole cell decision process? Answering this question requires addressing challenges at three scales: the detailed biochemistry of protein–protein interactions, the complex, interlocked feedback loops of transcriptionally regulated signaling pathways, and the multiple mechanisms of connection that link distinct pathways together into a full cell decision process. In this thesis, we address challenges at each level. We develop new computational approaches for identifying the interactions driving dynamics in protein–protein networks. Applied to the cyanobacterial clock, these approaches identify two coupled motifs that together provide independent control over oscillation phase and period. Using the p53 pathway as a model transcriptional network, we experimentally isolate and characterize dynamics from a core feedback loop in individual cells. A quantitative model of this signaling network predicts and rationalizes the distinct effects on dynamics of additional feedback loops and small molecule inhibitors. Finally, we demonstrated the feasibility of combining individual pathway models to map a whole cell decision: cell cycle arrest elicited by the mammalian DNA damage response. By coupling modeling and experiments, we used this combined perspective to uncover some new biology. We found that multiple arrest mechanisms must work together in a proper cell cycle arrest, and identified a new role for p21 in preventing G2 arrest, paradoxically through its action on G1 cyclins. This thesis demonstrates that we can quantitatively map the logic of cellular decisions, affording new insight and revealing points of control.

Thesis Supervisor: Bruce Tidor
Title: Professor of Biological Engineering and Computer Science

Thesis Supervisor: Galit Lahav
Title: Professor of Systems Biology, Harvard Medical School

Acknowledgments

Three faculty members have been instrumental in shaping the work contained here, and in my growth over the five years I have spent in Cambridge. My advisors, Bruce Tidor and Galit Lahav, have shaped my thinking about what problems to solve, and taught me how to channel my enthusiasm into a deeper form of inquiry, and Jacob White has been a close collaborator, incredible colleague, and great friend. I would also like to thank my thesis committee members, Michael Yaffe and Arup Chakraborty, for their enthusiasm and support over the course of the projects described herein.

I learned nearly everything I know about the day-to-day pursuit of research from Alex Loewer. His advice on how to read, write, and think about science has been invaluable, and I hope he understands that my graduation will not save him from my seeking it in the future.

Other collaborators provided key contributions to specific parts of this thesis. I thank Caroline Mock for her experimental contributions to the p53-Mdm2 circuit and synthetic feedback project described in Chapter 2. The work of Chapter 3 was done as an equal collaboration with Alex Loewer, who contributed both to the modeling and experiments contained there. Finally, Jacob White and Anya Castillo made critical contributions to the theory and computation used to perform sensitivity analysis in Chapters 4 and 5. I also thank my labmates, especially Bracken King, Josh Apgar and Jaydeep Bardhan for their support and stimulating conversations throughout these years.

On a personal note, I thank my brother Alex and parents for their support through these years spent so far from family, and their efforts to make that distance shorter. Finally, I would like to express my deepest gratitude to my fiancee Jacquie Bicais, without whose love, support, and insightful advice none of this would have been possible.

Chapter 1

Introduction

The purpose of computing is insight, not numbers. - Richard Hamming

A central goal of systems biology is to develop a predictive understanding of how cell decisions arise from the signaling pathways that sense and process information inside the cell. To be complete, this understanding must be end-to-end: it should quantitatively relate an input stimulus – whether the binding of an extracellular ligand, or damage to a cell's genetic information – to the cell's eventual commitment to an appropriate response – cell division, differentiation or even death – while accounting for the appropriate context – mutational status, or the presence of pharmacological inhibitors. Developing this understanding holds the promise of tuning cell decisions toward therapeutic goals, or restoring them in circumstances where they have been lost or dysregulated.

Achieving this understanding requires the combined application of experimental, computational and theoretical tools. Decisions in response to stimuli are made by individual cells, and not every cell reacts identically. Thus, we are challenged to investigate these mechanisms in individual cells, and derive insight that reflects the different choices elicited by random or probabilistic processes. Because our goal is *understanding*, it is not sufficient to statistically map the cell decisions associated with each input; rather, we must develop a quantitative rationale for how variations over time in these intermediate signals logically determine the cell's eventual course

of action. For this, we must identify the pathways and connections that are crucial for transmitting these signals, and that represent the functional units of these processes. Developing mechanistic models for the intermediate connections that process, filter and transmit these signals is a useful tool for demonstrating the sufficiency of these intermediate steps, encapsulating existing knowledge about their connectivity, and predicting the effect of perturbations; at its best, modeling can provide insight into how individual features of the transmitted signal are controlled by specific interactions.

The complexity of cellular processes prevents the full description of a signaling process at this time. By focusing on specific questions, identifying major obstacles and solving them, this thesis constitutes a step towards this goal. This work's contribution arises in case studies relating network topology to dynamics across three levels of system complexity.

The following section describes challenges faced in understanding the end-to-end function of cell signaling networks, and introduces the model systems in which we have their relationship in detail. Section 1.2 reviews current methods for tackling these challenges, and the formalisms within which our contributions arise. Finally, this introduction concludes by outlining the organization of the following thesis chapters.

1.1 Challenges in mapping signaling pathways to cell decisions

1.1.1 Signaling pathways rely on heterogeneous dynamical responses

Cell often respond to stimuli on different timescales compared to those of their input signals [1–3]. In the underlying networks, the dynamics with which signals are stored and transmitted can determine the cell's response to stimulation. In some cases the role played by dynamics in driving proper system function can be easily intuited. For instance, cells' orchestration of periodic processes (*e.g.* the cell division

cycle; changes in day/night metabolic cycles) are driven by oscillatory networks [4,5]. However, as many other signaling pathways are measured with finer temporal resolution, it is becoming clear that their activation elicits complex dynamics in the level or activity of signaling proteins, and that these dynamics can determine the cell's downstream response. A notable example is found in the decision to grow or differentiate elicited in PC12 neuroblastoma cells by EGF or NGF stimulation, respectively [6,7]. For this cell fate decision, the duration of pathway activation, but not its amplitude, determine the cell's response. Even more complex dynamics can arise in signaling pathways. In recent years, a growing catalog of mammalian signaling networks has been shown to drive oscillations in the concentration of key transcription factors, such as p53 and NF-κB [8,9].

Here, we focus on the network driving pulses of p53 as a model system in which to understand dynamics in cell signaling. p53 is a transcription factor activated in response to a variety of cellular stresses, such as DNA damage or the activation of oncogenes [10]. A series of studies characterizing the dynamics of p53 activation in detail have ensured this transcriptional network's status as one of the canonical dynamical processes in mammalian cells [9,11–13]. Nearly ten years ago, it was demonstrated that at the population level, MCF-7 cells exposed to γ-ionizing radiation (IR) undergo damped oscillation in p53 levels [11]. Subsequent studies monitoring individual cells over time further refined this perspective [9,12,13]. These studies identify three distinct dynamical features that are tightly controlled in this response:

1. the mean pulse amplitude does not depend on the IR dose;

2. the mean amplitude of successive pulses, averaged between cells is constant;

3. the timing of pulses is tightly controlled.

Further characterization of p53's oscillatory dynamics after IR has demonstrated the presence of these dynamics in a variety of cellular contexts: they arise in multiple cell lines [9,14], and have even been observed *in vivo* after γ-irradiation of transgenic mice expressing luciferase in a p53-dependent manner [15]. These observations raise

a question: as p53's importance in the IR response has been known for decades, and it is one of the most carefully studied proteins in eukaryotic cell biology, how did its dynamics remain uncharacterized for so long? The answer lies in the technical requirements for these experiments, which have been solved only recently. First, the dynamics are only revealed by monitoring a system over time with fine temporal sampling. This requires the use of minimally perturbative, and certainly nonlethal, measurement techniques such as live cell reporters and microscopy [16, 17]. Second, even in clonal populations of cells under identical stimulation, responses between individuals can quickly lose synchrony, requiring careful measurement and analysis of single cells.

This last observation underscores a challenge in understanding signaling dynamics: although individual cells are heterogeneous, some features of their responses must be tightly controlled. Some heterogeneity arises through stochastic variation due to small numbers of molecules [18–20]. Another important source of variation arises from time-varying transcription rates and changes in global cellular metabolic programs [12, 21]. Transcriptional noise is also of special significance because of its timescale. Many biochemical interactions functionally act as low pass filters [22], which can suppress fast stochastic variation to maintain synchrony of processes on timescales relevant for cell signaling [23]. As transcriptional noise varies on the timescales of signaling processes themselves, it cannot be filtered by frequency alone.

1.1.2 Cell decisions are driven by networks with complex topology

The activation of eukaryotic signaling pathways is rarely specific to a single input, or limited to initiating a single response. Rather, inputs may activate a number of parallel and serial pathways to varying extents. The effects of signals are often combined to achieve a proper response [24], and can even induce autocrine production of additional input stimuli [25]. Moreover, this complexity is not limited to connections between pathways. Individual signaling modules frequently include seemingly

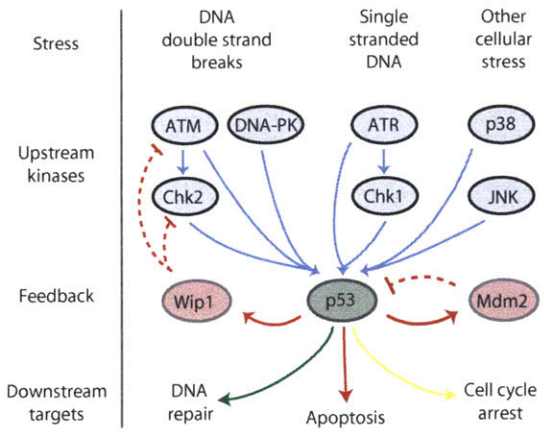


Figure 1-1: Signals and responses acting through the p53 pathway. Many stresses signal to p53 through a variety of upstream kinases. These signals are processed by a feedback network (see Figure 1-2) and lead to a downstream fates including cell cycle arrest, apoptosis, and/or repair of damaged DNA.

redundant and antagonistic connections. This phenomenon is especially prominent in feedback connected pathways, where the activation of multiple positive and negative feedback loops are initiated by the same signal [26].

The mammalian DNA damage response network is a model system for the complex topology within and between distinct pathways. Different types of damage converge to regulate shared downstream signaling processes through parallel branches of stress-activated kinases, acting through ataxia telangiectasia mutated (ATM) and checkpoint kinase 2 (Chk2), ataxia telangiectasia mutated and Rad-3 related (ATR) and checkpoint kinase 1 (Chk1), and p38 or Jun N-terminal kinase (JNK) mitogen-activated protein kinase (MAPK) cascades (Figure 1-1) [27–30]. Not all branches are equally responsive to each input. For instance, p38 is required for the response to ultraviolet light (UV) induced damage, but dispensable for initiating cell cycle arrest after IR [28]. Ionizing radiation is canonically thought to elicit double stranded DNA break that are sensed through the kinases ATM and Chk2 [27, 31, 32]. The essentiality of this kinase cascade in generating a dynamic p53 response has also been established [13]. This multiplicity of inputs is matched by a multiplicity of p53-regulated downstream processes, including apoptosis, DNA repair and cell cycle arrest (Figure 1-1) [33].

One example of the diversity of interconnections between pathways arises at the interface between the p53 signaling pathway and its effect on the cell cycle. Progression through the cell cycle is normally driven by a network regulating the sequential activation of cyclin dependent kinases, consisting of a Cdk kinase subunit and a regulatory cyclin subunit [4]. Halting progression through the cell cycle in the presence of DNA damage involves multiple interactions through a variety of biochemical processes, including the induction and repression of target cell cycle genes such as cyclins A and B [34–36]; the binding and stoichiometric repression of cyclin dependent kinases by p21 [37]; and the enzymatic inactivation by post-translational modification of key regulators of cell cycle progression such as Cdc25 [38, 39]. The p53 signaling pathway plays a direct role in this regulation to elicit both G1 and G2 cell cycle arrests [37, 40].

Interactions within the p53 signaling pathway reveals the enormous structural complexity of this pathway (Figure 1-2). p53 regulates many genes that can in turn modulate its own activation or stability, as well as the activation of its upstream kinases, forming multiple positive and negative feedback loops on its activity [33]. One of the best characterized feedback loops in the network acts through Mdm2, which can be induced by p53 even in the absence of stress and targets p53 for proteasomal degradation [41].

p53's transcriptional activity for these targets is tightly regulated through post-translational modification to prevent unwanted activation of downstream genes [42, 43], providing a mechanism by which subsets of feedback connections can be activated in a stimulus-dependent manner. Acetylation, a major class of activating modifications, is required for p53's ability to induce many downstream genes [44], and different patterns of acetylation can selectively activate individual downstream gene programs [45]. Even the induction of canonical targets with high-affinity p53 DNA binding sites, such as p21, requires the recruitment of p53 modification-dependent coactivators [45, 46]. Notably, supporting the homeostatic role of the p53-Mdm2 feedback loop, p53 acetylation is not required for *mdm2* gene activation [44, 45].

Prior work has identified some feedback loops that are required for the IR in-

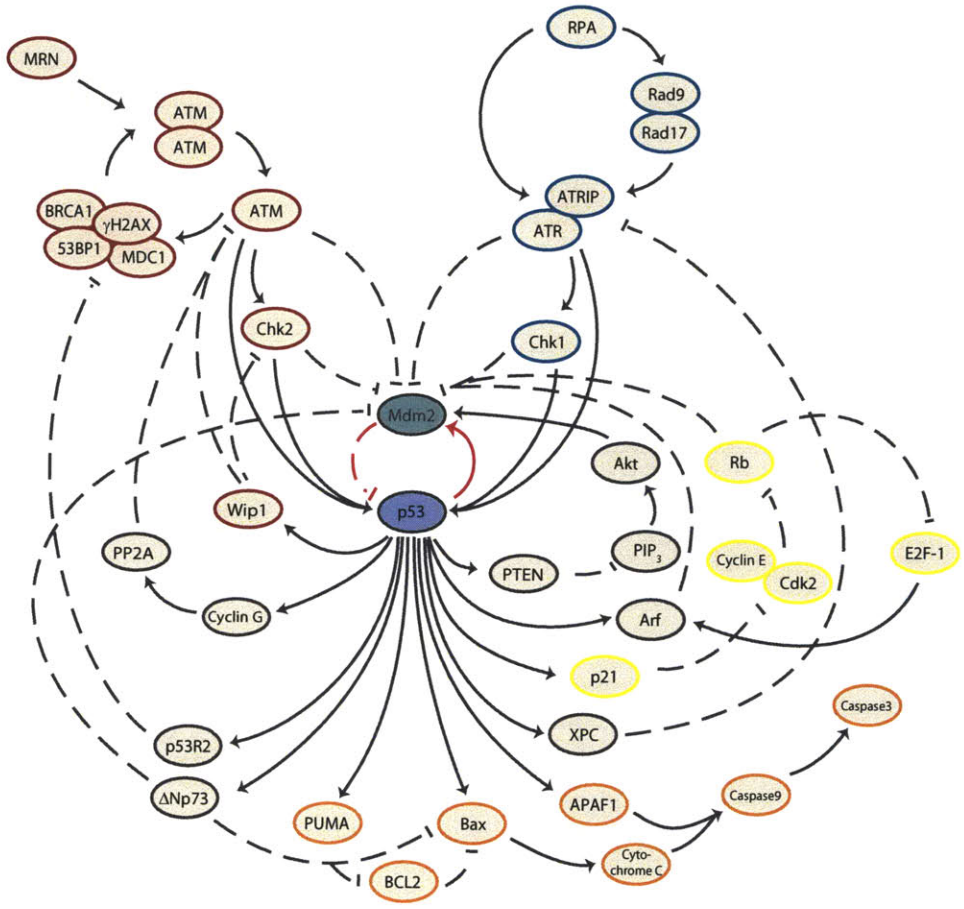


Figure 1-2: **p53 at the heart of multiple feedback loops.** In response to cellular stress, many feedback loops act on p53. The weights of individual connections vary depending on cellular context and the nature of the applied stress. Mdm2, the core negative regulator of p53, participates in a constitutively active homeostatic negative feedback loop (red arrows). Other feedback loops act through species involved in DNA damage sensing and repair (double stranded breaks, maroon-circled species; single stranded DNA, teal species), the cell cycle (yellow species), and apoptosis (orange species), as well as other pathways. Solid and dashed lines indicate activating and inhibitory connections, respectively.

duced dynamical response, and excluded others from playing a role in this process [9,12,13,47]. The oscillatory response to IR requires at least two feedback loops: the core Mdm2 loop, and one acting through the phosphatase Wip1's inhibition of upstream kinases [13]. Three other connections are dispensable for the generation of these dynamics: levels of PTEN and Cyclin G are constant and high during oscillation [13], and oscillations have been observed in cells lacking Arf [9]. A number of negative and positive feedback loops, however, may yet play additional, undefined roles.

1.1.3 Detailed biochemistry drives systems-level behavior

The preceding section describes complex interactions driving systems-level responses, both between pathways and within a single pathway. Three distinct lines of theoretical and experimental evidence suggests they may also arise at a finer scale: within the detailed biochemistry of interactions between small sets of proteins. First, a recent theoretical study by Thomson and Gunawardena shows that a protein with multiple sites of modification, coupled to a single kinase and phosphatase, provides a mechanism to generate a vast number of stable states [48]. While untested experimentally, this mechanism is plausible, as multiple modification states, multimeric complexes and protein isoforms are present in a variety of dynamically varying signaling pathways [43,49,50], often with unknown roles.

Second, theoretical studies have proposed that systems-level properties such as bistability and oscillation can arise in networks at the post-translational level, without relying on easily identifiable feedback loops [51,52]. Notably, these phenomena rely on the precise details of enzyme-substrate complex, enzyme-inhibitor complex, and multimer formation (for startling examples of the importance of these details, see [51]). Experimental studies have shown that this capability is not just theoretical; natural biological systems have utilized protein-protein interaction networks in driving complex dynamical responses [52,53].

A third mechanism of systems-level complexity in biochemical interactions arises through the formation of competitive inhibitory protein-protein complexes, a common

property of nearly all biochemical interactions [54]. It was shown theoretically that ultrasensitive transfer functions with Hill coefficients >10 could be generated through 'molecular titration,' or competition between binding partners for complex formation [55]. This effect was subsequently demonstrated experimentally for the Wee1 cell cycle kinase [56] and leucine zipper transcription factors [57]. Explicit treatment of these effects in quantitative models is rare, often substituted by Langmuir and Michaelis expressions to represent binding and enzymatic processes, respectively, as models that explicitly treat these reactions grow combinatorially in the number of modeled species. While direct approaches have been successful in some cases [50], major questions remain how to cope with the dense connectivity and large numbers of species in such models.

An elegant demonstration of the role played by detailed biochemical interactions in organizing a dynamical process arose through experimental study of the circadian clock of *Synechococcus elongatus*. This network enables *S. elongatus* to adapt its genetic and metabolic programs to daily changes in the environment and provides a daily rhythm to photosynthetic regulation [58, 59]. In its normal context in the bacterium, the circadian clock involves transcriptional regulation acting through multiple feedback loops [60]. Surprisingly, the essential characteristics of this system's dynamics – oscillation with an approximately 24 h period at a wide range of temperatures – can be reconstituted *in vitro* with only three proteins: KaiA, KaiB and KaiC [61]. This network is a model system of detailed biochemical complexity, as it is tractably small (consisting of only three proteins) but tightly regulated at the level of multimerization [62], post-translational modification [63, 64], and incorporates molecular titration by inhibitory complex formation [65]. Many models have been constructed demonstrating that this complexity can drive oscillation [49, 65–67], but in many cases, the specific network topology and interactions driving oscillations is unclear.

1.2 Gaining insight into the operation of complex networks

The prior section outlined three levels at which complex interactions within and between processes can give rise to a dynamic response. To develop a quantitative understanding of the operation of these systems, we are challenged to relate specific features of a dynamical response to the structure and parameterization of the network generating them. Such an understanding requires models representing these features as they cannot be quantitatively described by network connectivity diagrams alone. This section describes the current experimental and computational approaches aimed at mapping the relationship between network topology and dynamics, and informing and analyzing models of dynamically varying networks.

1.2.1 Network motifs map pathways to dynamical features

A growing body of work from the Alon lab and others coined the term 'network motif' to describe network structures overrepresented in transcriptional or post-translational interaction networks that elicit a predictable qualitative effect on signaling dynamics [68–70] (see [71] for a general introduction to this approach). This approach has been used to classify and describe pathways arising in bacteria [69] and yeast [70], as well as to exhaustively classify behavior from networks consisting of four or fewer nodes [72]. Some prevalent motifs, for which strong predictions of function could be made, were subsequently identified in their natural context, and their effect on signal processing was validated experimentally. For instance, the coherent feedforward motif was identified as a delayed response module with noise-filtering capabilities [69, 73]; multiple examples of this motif, each harboring these properties, were subsequently characterized in *E. coli* [74, 75].

Initial studies of motifs treated these network structures as signal transduction elements that shape a dynamic response to one or more inputs, rather than as closed, autonomous networks. Recently, efforts were made to extend the principles underly-

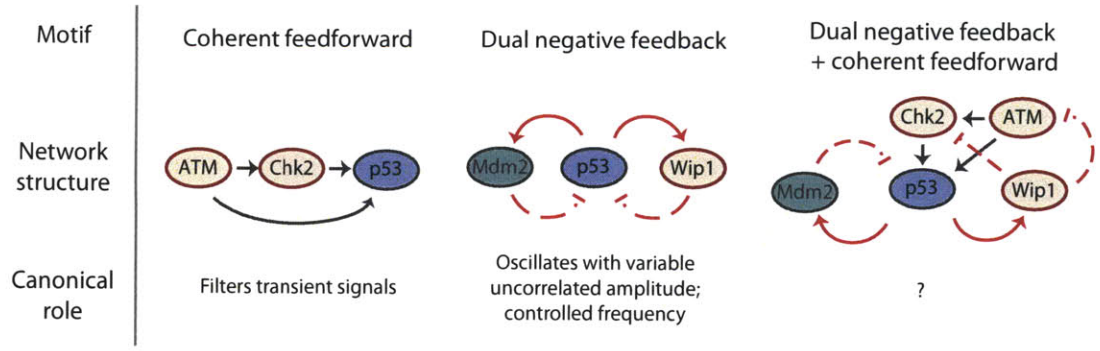


Figure 1-3: Motifs in the IR-induced p53 network. Columns show three network structures – a coherent feedforward motif, a double negative feedback oscillator and their combination – whose connections have been experimentally validated as indispensible for generating dynamics in the IR-induced p53-Mdm2 network. Rows indicate the name, interacting species, and the canonical role (where known) associated with each structure.

ing this approach to autonomous oscillatory networks, leading to the characterization of the principles underlying oscillations driven by delayed negative feedback loops, and combinations of a positive and negative feedback loop [26, 76, 77]. In these cases, the inclusion of a fast positive feedback loop was predicted to increase the size of the parameter regime in which undamped oscillation can occur, and provide a mechanism by which oscillation frequency can be tuned over a broad range without large changes in amplitude [26].

Testing these predictions has proven challenging because most natural oscillating networks comprise more complicated topologies, and are not amenable to manipulation of key parameters. For these cases, synthetic biological studies provides a platform to experimentally test the relationship between topology and dynamics. Oscillating transcriptional networks have been developed in bacteria first utilizing negative feedback loops [78] and more recently using combinations of negative and positive feedback loops [79]. Along with the recent demonstration of an oscillatory synthetic network in mammalian cells [80], these studies have helped confirm properties such as the robustness and frequency-tunability of oscillation driven by combinations of loops.

It is noteworthy that not all bacterial signaling networks comprise single, ratio-

nalizable network motifs, and far fewer eukaryotic networks are structured in this way. For example, the connections experimentally validated as crucial for generating dynamics in the IR-induced p53 network contain both a feedforward motif and two delayed negative feedback loops (Figure 1-3). Furthermore, many of the dynamical features described in Section 1.1.1, such as combination of tight frequency regulation and dose independence of amplitude, are not representative of either basic oscillatory motif yet characterized. It is an open question whether combining motifs should preserve, augment, or replace their isolated functions.

1.2.2 Optimization: coping with more complex systems

A variety of tools are available to gain intuition about the the principles underlying the operation of large networks for which quantitative models are available. One powerful class of methods is rooted in the concept of local optimization. Applied to ordinary differential equation (ODE) models of biochemical processes, optimization identifies values for parameters (*e.g.* reaction rate constants; initial protein concentrations) that provide at least a local minimum of an objective function (a scalar-valued function that can be evaluated at any parameterization of the model in the neighborhood of an initial guess at these parameter values). The power of optimization-based techniques lies in two sources of flexibility: the flexibility with which one may define an objective function to be minimized and the parameters over which to optimize. For example, an objective function that mathematically represents the error between modeled protein concentrations and their experimental measurements allows the modeler to update the model to reflect new data. Choosing parameters also provides another source of flexibility. We have shown that local optimization can be used to find time-varying inputs capable of driving a model output to match a specified temporal pattern, and that time-varying stimuli designed in this way can be advantageous for discriminating between candidate models[1] [82].

[1] As a supporting author, I contributed to the development of the mathematical and numerical methods used for the nonlinear optimal controller, and in writing the software used to perform this optimization. A full description of these methods and software appears in [81].

The utility of these methods is not limited to controlling model output or matching experimental data, and a second class of applications arise through sensitivity analysis. The sensitivity of an objective function with respect to a parameter describes the change in objective function due to an infinitesimal change of a parameter. Computing sensitivities of dynamical features of interest (*e.g.* the amplitude of a peak in concentration; the period of oscillation) permits the identification of reactions critical in setting the value of those features. Sensitivity analysis has been applied to oscillating biological networks, especially in the context of the eukaryotic circadian clock. Computing period sensitivities for a complex model of the mammalian circadian clock identified a single negative feedback loop responsible for setting this dynamical feature [83], and relationships between sensitivity profiles to multiple dynamical features have been compared using models of the *Drosophila melanogaster* and murine clock networks [84].

The computation and analysis of sensitivities of dynamical features to changes in parameters is an area of active research, particularly for features such as oscillation period and phase [85–88], or the amplitude or timing of a peak in concentration [89]. However, in many cases computing sensitivities is feasible even for models of large pathways; efficient computation of sensitivities using an adjoint formulation has been described for general ODE models [90]. Major challenges remain in interpreting these results to gain insight from complex models, and to understand them in the context of their network topology.

1.3 The present work

In this first part of this thesis, we set out to map out an end-to-end cell signaling process, DNA damage signaling and its effect on the cell cycle, from the early signaling after damage induction to cells' decision to sustain arrest or reenter the cell division cycle up to days after the original introduction of damage. This pathway is characterized by the coupling of two oscillatory networks, each consisting of combinations of feedback loops: the p53 signaling pathway, the cell cycle, and their

interconnection. Throughout this work, we used a quantitative approach to refine models with experimental data, and to query these models for a deeper understanding of how individual interactions determine the response of cell signaling pathways.

The p53 network tightly regulates multiple distinct features of the dynamical response to DNA damage. To understand how this network achieves such regulation it is necessary to understand what interactions and feedback loops are responsible for generating these dynamical features. In Chapter 2, we use an approach grounded in synthetic biology to isolate and characterize a core regulatory feedback loop of the p53 network. Through quantitative modeling coupled with single cell experiments, we find that additional feedback loops and small molecule inhibitors modulate specific features of the dynamics, predictions we confirm experimentally.

After building a quantitative description of the core regulation in the p53 pathway, we turned to the broader regulation of cell cycle arrest through the DNA damage response network. In this context, we were able to address a computational challenge in systems biology, demonstrating that it is feasible to interconnect individual pathway models to quantitatively understand the action of their combined network. Chapter 3 describes our work to interconnect our p53 model to an existing model of the cell cycle, parameterized to match data from mammalian cells. A combined computational and experimental approach reveals that individual mechanisms of cell cycle arrest contribute specific features to the overall arrest state, and that their dysregulation can lead to grave errors in cell cycle progression.

In the second part of this thesis, we turn to the detailed biochemistry of post-translational networks, a third scale at which network complexity can lead to systems-level dynamics. Through the application of sensitivity analysis tools, we identified a previously unappreciated subtlety in the computation of sensitivities for biochemical networks exhibiting mass conservation. Chapter 4 of this thesis develops this subtlety in detail, showing that for a large class of ODE systems exhibiting hidden conservation laws, additional constraints must be included in the framework for oscillator sensitivity analysis.

Finally, in Chapter 5, we set out to understand how network structure drives dy-

namics in detailed biochemical networks, using the *in vitro* cyanobacterial circadian clock as a model system. We find that oscillator sensitivity analysis identifies groups of reactions that drive dynamics as part of larger, self-consistent processes. However, we find that many highly sensitive processes are distributed throughout the network. This result starkly contrasts prior work in transcriptional oscillatory networks, where oscillator sensitivity analysis identifies single feedback loops as responsible for driving the dynamic response. By exhaustively enumerating subsets of reactions that still undergo oscillation, we identify two network motifs – a delayed negative feedback oscillator and a coupled negative-positive feedback oscillator – that each contribute to the dynamics of the full network. We suggest that this coupled oscillator combining two well-known motifs is an excellent topology with which to tune oscillation phase while preserving oscillation period, a crucial characteristic for a circadian system.

Chapter 2

Modulating dynamics in a synthetic p53 network

Many cell signaling networks sense and encode dynamical information. The cell orchestrates periodic processes such as the cell cycle and circadian metabolic processes using biochemical oscillators to ensure the proper timing of events [4, 5], and cell fate decisions can depend on the transient or sustained activation of upstream signals [7]. Specific features of these dynamics must be tightly regulated in order to ensure information is properly processed in the noisy environment of the cell. Understanding how specific motifs (such as feedback loops) and interactions within these networks allow for tight control of dynamical features is a central challenge of systems biology [91].

In recent years, synthetic biology has proven useful for engineering complex dynamical behaviors in designed signaling networks [78–80, 92], thereby demonstrating their sufficiency in generating these dynamics. Oscillating transcriptional networks form a major class of these transcriptional networks, and have been developed in bacteria first utilizing negative feedback loops [78] and more recently using combinations of negative and positive feedback loops [79]. Along with the recent demonstration of synthetic network oscillations in mammalian cells [80], these studies have helped elucidate properties such as the robustness of oscillation in cell populations, and the tunability of oscillation frequency. Native biological networks are often more complex

and can be interconnected with other pathways [14]. Here, using the p53 oscillating network as a model system, we show that synthetic biology also offers a powerful set of tools for the dissection and control of natural systems.

p53 is a mammalian transcription factor that is activated in response to a variety of cellular stresses [10]. In addition to inducing genes regulating the response to these stresses, p53 regulates many targets that provide feedback regulation on its own activity [33]. One of the best characterized feedback loops in the network acts through Mdm2, which can be induced by p53 even in the absence of stress and targets p53 for proteasomal degradation [41]. After γ-irradiation (IR), individual cells undergo a series of p53 pulses in a dynamical response that maintains tight control over three distinct dynamical features. First, the pulse amplitude does not depend on the IR dose. Second, these pulses are undamped; the amplitude of successive pulses remains constant. Finally, although the amplitude of individual pulses can be highly variable, the timing of these pulses is more tightly controlled. At least two negative feedback loops, acting through Mdm2 and Wip1 respectively, are required for the proper dynamical response after IR. [13], and IR induces post-translational modifications on both Mdm2 and p53, modulating many additional system parameters. In the context of this complex regulation, it is unknown whether specific interactions and feedback loops can control individual dynamical properties. Identifying these control points may also allow tuning of these features to assess their importance or rescue them from dysregulation.

This chapter describes the construction of a synthetic variant of the p53 network based on transcriptional stimulation of the core p53-Mdm2 negative feedback circuit in the absence of a cellular stress response. This reduced network undergoes damped oscillation, and shares a subset of the features of the IR response. Using mathematical modeling coupled with experiments, we demonstrate that addition of synthetic positive and negative feedback loops can specifically modulate the damping rate of these oscillations, and that varying the core loop's feedback strength using a small molecule inhibitor allows tuning of the oscillation frequency.

2.1 Transcriptional stimulation of the NF circuit leads to damped oscillation

We identified the core p53-Mdm2 negative feedback as a suitable reduced network because this feedback loop is active even in unstressed cells [93]. To bypass the usual mode of activation through post-translational modification, we used the transcriptional activation of p53 as a synthetic input. We used a cell line in which the expression of a p53-CFP fusion protein is driven by an zinc-inducible metallothionein promoter [94], and Mdm2-YFP is driven by its native promoter [9].

We first set out to characterize dynamics from this core p53-Mdm2 negative feedback loop. Stimulation with varying doses of $ZnCl_2$ led to the induction of p53-CFP and Mdm2-YFP observable in individual cells by time-lapse microscopy (Figure 2-1B; see Section 2.4), and subsequent quantification of individual cell trajectories revealed pulses of p53 and Mdm2 (Figure 2-1C). Figure 2-1D shows the collective dynamics of individual cells tracked over time after 50 μM zinc stimulation; cells undergo a high amplitude, tightly synchronized first p53 pulse, reaching a maximum at about 5 h (Figure 2-1D). The corresponding first Mdm2 pulse was also tightly synchronized and delayed by approximately 2 h (Figure 2-1E); synchrony between cells in both p53 and Mdm2 dynamics was lost in successive pulses.

These dynamics are reminiscent of those observed in p53 levels after γ-irradiation, in which p53 is seen to undergo a series of undamped pulses whose amplitude and frequency are independent of the radiation dose [9, 12, 13]. To assess whether the core p53-Mdm2 loop exhibits similar control mechanisms, we used an automated pulse detection algorithm to identify pulse maxima and minima from individual cell trajectories (see Appendix A); from these data we computed the mean amplitude and timing of successive pulses in each condition (Figure 2-1F-H). This quantitative characterization revealed key differences from the γ response. We found that the amplitude of each p53 pulse was no longer 'digital,' but rather was highly sensitive to the zinc concentration applied (Figure 2-1F), with a tenfold difference in first pulse amplitude between the highest and lowest zinc concentrations. In addition, we find

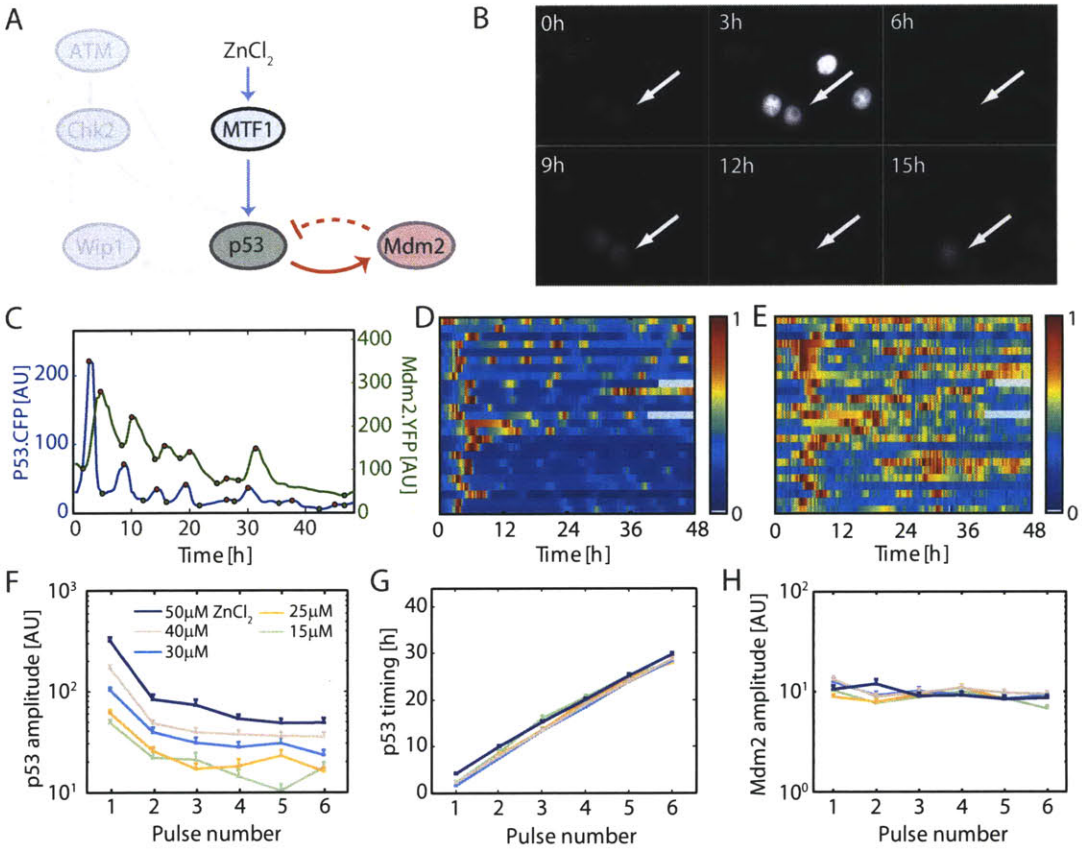


Figure 2-1: **Stimulating and measuring dynamics from the core p53-Mdm2 feedback loop.** (A) The p53-Mdm2 circuit. Zinc stimulates p53-CFP transcription from a metallothionein promoter, bypassing IR-induced activation through the ATM/Chk2 kinase cascade and Wip1 negative feedback loop (gray arrows). Induced p53 activates Mdm2-YFP transcription, which negatively regulates p53 stability. (B) Time-lapse microscopy of p53-CFP after stimulation with 50 μM $ZnCl_2$. Arrows denote the same representative cell in subsequent frames. (C) Trajectory of cell marked in (B), with pulses identified using an automated approach (see Section 2.4). Fluorescence intensities of p53-CFP (blue curve) and Mdm2-YFP (green curve) are shown on left and right axes, respectively. (D-E) Heat maps of 25 representative cells after 50 μM $ZnCl_2$ treatment. Rows represent (D) p53-CFP and (E) Mdm2-YFP levels normalized to the maximum amplitude for each cell. (F) p53-CFP fluorescence, (G) Mdm2-YFP fluorescence, and (H) p53-CFP pulse timing are shown for each pulse after stimulation (mean + s.e.m.).

that successive p53 pulses decrease in amplitude, indicating individual cells undergo damped oscillation after zinc stimulation (Figure 2-1F). Other features of the dynamics are more tightly controlled. As is the case after γ-irradiation, the timing of successive p53 pulses is also tightly controlled with approximately 5.5 h between successive pulses (Figure 2-1G).

We next set out to validate that our reduced system did not lead to the activation of a cellular stress response, another source of p53 dynamics, by monitoring cell death and division in zinc-treated cells. We did not observe any difference in the number of cells diving after 50 μM $ZnCl_2$ treatment compared to untreated cells (Figure A-1D), and fewer than 5% of cells died after treatment at any zinc dose (data not shown), suggesting that these p53 dynamics did not lead to a cellular stress response. We also found that p53 and Mdm2 dynamics persisted through cell division events, with division sometimes occuring during a pulse, consistent with observations after low levels of γ-irradiation and in other transcriptional oscillating networks [12, 95].

Surprisingly, our initial experiments revealed that Mdm2 amplitude was less variable than p53 amplitude across both changes in zinc dose and the number of pulses after treatment (Figure 2-1H). We reasoned that this might reflect the network's tight control over two processes: p53-induced Mdm2 transcription and Mdm2 protein stability. In the former case, p53 activation above a threshold might saturate Mdm2 promoter activity, leading to a controlled increase in Mdm2 level. In the latter case, the ability of Mdm2 to regulate its own level through autoubiquitination and subsequent degradation might prevent further increases in Mdm2 level. To separate these two effects, we used a cell line containing the same inducible p53-CFP construct coupled to the Mdm2 promoter driving expression of YFP. In this cell line, YFP induction should be subject only to transcriptional regulation. Stimulating these cells with zinc still led to damped p53 dynamics, as well as a slower, sustained increase in YFP levels (Figure A-2A). We compared the mean amplitude of the first p53 pulse to the amplitude of the first YFP maximum at various zinc doses, and found that YFP activation was less pronounced than p53 activation, supporting transcriptional control as one mechanism for Mdm2 regulation (s Figure A-2B-C). However, YFP

levels were still substantially more variable than Mdm2 levels observed previously, suggesting the presence of additional regulation on Mdm2 levels, possibly through autoregulation of its stability.

We have shown that transcriptional stimulation of the core p53-Mdm2 negative feedback circuit is capable of generating complex dynamics in unstressed cells, and that these dynamics share a subset of characteristics with the response after IR. To better understand how these different dynamical features may be controlled in the network, we constructed a mathematical model of the core p53-Mdm2 negative feedback loop. Our model consists of a series of ordinary differential equations (ODEs) representing p53 and Mdm2 (see Appendix A for details). It incorporates three nonlinear interactions: Mdm2-mediated ubiquitation of p53 [41], the transcriptional effect of p53 on the Mdm2 promoter [96], and autoregulation of Mdm2 on its own stability [97]. We parameterized our model to quantitatively match our experimental observations of amplitude, frequency and damping after zinc treatment (Figure 2-2).

To compare our model directly at the relevant zinc concentrations used experimentally, we determined the transfer function from zinc dosage to MTF1-induced p53 transcription (see Appendix A). We simulated the model at five zinc concentrations and compared the resulting p53 and Mdm2 first pulse amplitudes, frequencies, and damping coefficients to those obtained previously from experimental data (Figure 2-2A-C; see Appendix A for computation details). Our model captures the scaling of p53 amplitude as well as the invariance of frequency and damping over varying zinc concentrations (Figure 2-2A-C). While the model qualitatively exhibits a decreased dependence of Mdm2 amplitude on zinc dose, it predicts more variability than was observed experimentally (Figure 2-2A). In addition to these deterministic simulations, we ran our model in the presence of multiplicative transcriptional noise to compare its results to the cell population data of Figure 2-2 [12, 21]. Figure 2-2D shows a representative modeled cell with multiplicative noise modulating both the p53 and Mdm2 production rates. Using the same data processing algorithms from our experimental protocol, we tabulated the p53 amplitude and timing profiles from 500 modeled cells at each of five zinc concentrations (Figure 2-2E-F). In agreement with our experi-

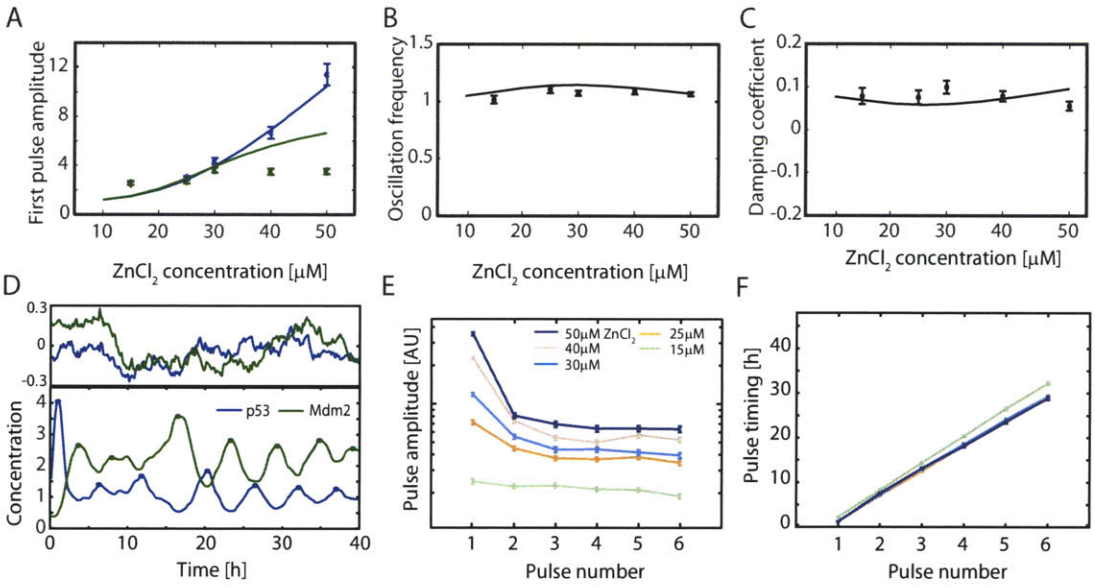


Figure 2-2: **Mathematical modeling of the core negative feedback circuit.** (A) (A) p53 (blue) and Mdm2 (green) amplitudes, (B) frequency of oscillation and (C) p53 damping coefficient at varying $ZnCl_2$ concentrations. Points represent experimental data (mean + s.e.m); curves show model results. (D) Model dynamics in the presence of multiplicative transcriptional noise. The lower panel shows modeled p53 (blue trace) and Mdm2 (green trace) levels after 25 μM $ZnCl_2$ stimulation in the presence of multiplicative noise in p53 and Mdm2 production. The upper trace shows the corresponding noise levels. (E) p53 amplitude and (F) timing for 100 modeled cells are plotted for each pulse since stimulation (mean + s.e.m.). Colors are as in Figure 2-1F-H.

mental results, we find that noise contributes to wide variation in p53 amplitudes in individual modeled cells but a more tightly controlled distribution of pulse times.

2.2 Addition of synthetic transcriptional feedback loop modulates network stability

Network motifs controlling oscillations have been the subject of much recent scrutiny, both computationally and experimentally [26, 77]. Recent work suggests that combinations of negative and positive feedback loops can lead to both robust oscillation and tunable frequency [26, 79]. We set out to determine whether additional positive and negative feedback loops can play these roles in the context of the core p53-Mdm2 negative feedback circuit. We reasoned that our input to the p53-Mdm2 negative feedback loop – transcriptional induction of p53 – provides a node at which to add new synthetic feedback connections in the network.

We first turned to our model to predict the effects of adding positive and negative feedback loops through p53's induction of either an inducer or repressor of p53 transcription. To account for the action of this new feedback connection, we augmented our model to include delay in producing p53 and the feedback protein, as well as terms representing p53's induction of the feedback protein and its subsequent effect on p53 (for detailed equations and parameters see Appendix A). We queried the effect of these additional feedback loops on dynamics by sampling the synthetic loop's delay time and feedback strength over a wide range of parameter values. We found that incorporating additional feedback loops was unable to strongly affect the amplitude or frequency of oscillation, but had a much stronger effect on the damping rate (Figure 2-3A-B; Figure A-3). For a broad range of parameter values where oscillation was preserved, we found that addition of a synthetic positive feedback destabilizes the network (lower damping rates or undamped oscillation), while a synthetic negative feedback loop has the opposite effect.

We set out to test these predictions experimentally by supplementing the core p53-

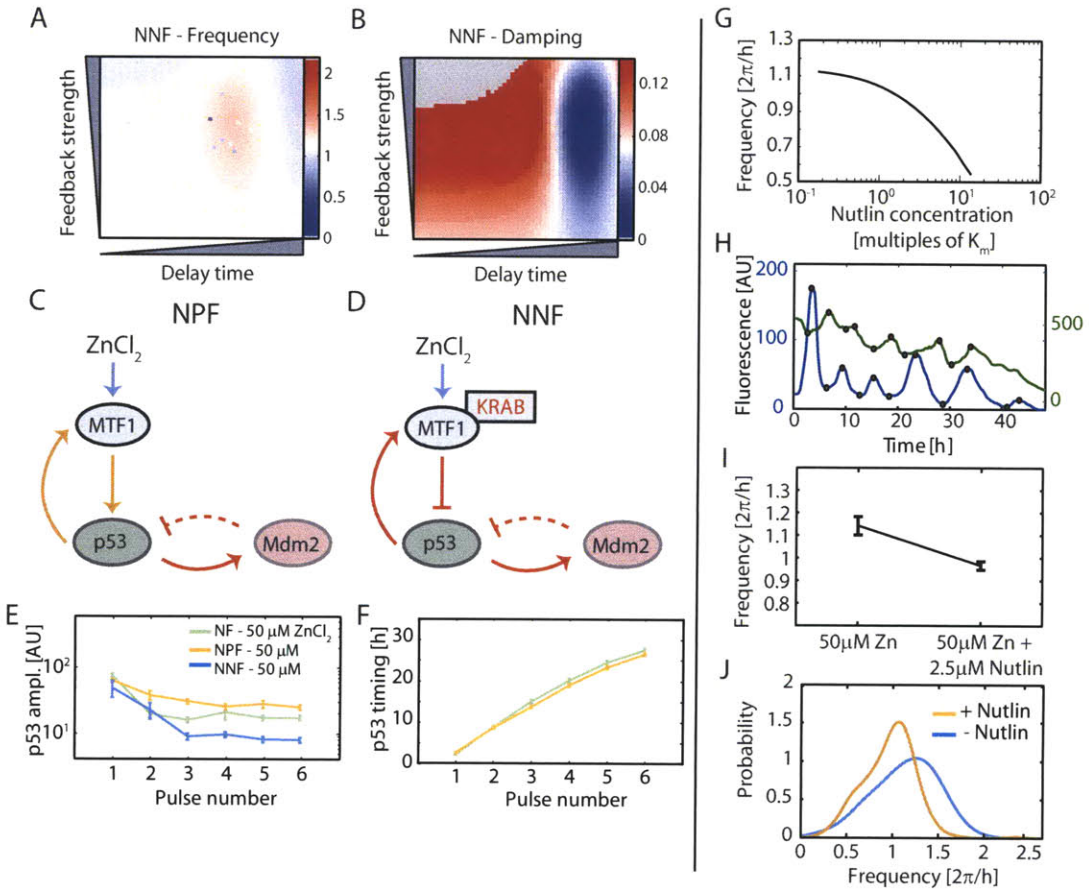


Figure 2-3: **Modulating stability and frequency of oscillation.** Model predictions for (A) frequency and (B) damping in the presence of a second negative feedback loop (NNF). x- and y-axes represent variation of two parameters, γ_{f0} and β_{p0} across two orders of magnitude, representing feedback protein production delay and its effect on p53 transcription. Color bars show frequency and damping at each point, where white indicates the value from the model without additional feedback. (C-D) Schematics of experimental synthetic feedback systems: (C) NPF, positive feedback on p53 through MTF1 and (D) NNF, negative feedback on p53 through MTF1-KRAB. (E) p53 amplitude and (F) timing measured in NF (green), NPF (orange) and NNF (blue) cells after stimulation with 50 μM $ZnCl_2$. (G) Predicted oscillation frequencies in the p53-Mdm2 circuit is plotted against Nutlin3A dose, measured in multiples of its IC50. (H) Representative single cell trajectories showing p53-CFP (blue trace) and Mdm2-YFP (green trace) after treatment with 50 μM $ZnCl_2$ in the presence of 2.5 μM Nutlin3A. (I) Oscillation frequency in nutlin-pretreated cells (H) compared to control cells. (J) Distribution of oscillation frequencies in Nutlin3A-pretreated (orange trace) and control cells (blue trace).

Mdm2 negative feedback circuit with an additional negative or positive feedback loop on p53 transcription (Figure 2-3C-D) using variants of MTF1, the zinc-responsive transcription factor that acts on the metallothionein promoter. Cells with synthetic positive feedback in addition to the core p53-Mdm2 negative feedback loop (NPF cells) were generated using a construct containing the Mdm2 promoter driving transcription of MTF1 fused to the mCherry fluorescent protein [98]; in this circuit, p53 induces MTF1-mCherry, which, in the presence of zinc, induces p53 (Figure 2-3C). Similarly, we constructed a cell line harboring a second negative feedback loop (NNF cells) by utilizing p53 transcriptional control over a MTF1-KRAB fusion protein fused to mCherry to repress p53 transcription [99] (Figure 2-3D).

Like NF cells, both NPF and NNF cells generated pulses of p53 after zinc treatment. We found that the amplitude of the first pulse was comparable across all three cell lines (Figure 2-3E). In NPF cells the damping rate was lower than in NF cells, indicated by the higher amplitude of subsequent pulses (Figure 2-3E). Conversely, NNF cells exhibit a faster damping rate than NF cells, so that after the third pulse the amplitude was too low to detect (Figure 2-3E). For NF and NPF cells, which still exhibited sustained pulsing, the timing of subsequent pulses was unaffected (Figure 2-3F). Together, these results confirm our prediction that additional transcriptional feedback loops can modulate the system stability, but do not affect the timing of pulses or the first pulse amplitude.

2.3 A small molecule inhibitor of p53-Mdm2 interaction modulates oscillation frequency

The previous section's results highlight a distinguishing feature of the p53 network: its tight regulation of pulse timing. This precise control arises both in our data from the core p53-Mdm2 circuit as well as in the full network in response to IR [9]. We next set out to identify the interactions in the core negative feedback loop that could be used to modulate oscillation frequency. We varied each parameter over two

orders of magnitude and computed the oscillation amplitude and frequency to identify parameters capable of modifying period while maintaining oscillation of reasonable amplitude (Appendix A and Figures A-4,A-5). We found that parameters controlling basal transcription rates of p53 and Mdm2 led to changes in amplitude without significantly affecting frequency. This is intuitive, as stimulation through zinc acts at the transcriptional level, and noise in protein production rates do not affect the tight control of frequency. Parameters affecting p53-Mdm2 feedback strength and the delay time of Mdm2 protein maturation were more sensitive to period. Many of the biological processes associated with these sensitive parameters are difficult to modulate experimentally. However, we identified the strength of the p53-Mdm2 interaction as a target through the use of nutlin-3A, a small molecule inhibitor of this interaction [100]. To investigate this further, we predicted the effect of nutlin addition on oscillation frequency in multiples of its IC50 (Figure 2-3G), and found that addition of nutlin at a concentration tenfold higher than the IC50 was predicted to lead to an approximately 30% decrease in period.

To test this prediction, we preincubated NF cells with 2.5 μM nutlin-3A for 24 h before stimulating oscillation with 50 μM $ZnCl_2$. A representative cell from this experiment is shown in Figure 2-3H, indicating that cells preincubated with nutlin-3A are still able to undergo sustained oscillation but exhibit wider, lower-frequency pulses. By tabulating the timing between consecutive pulses in individual cells, we observed an approximately 20% decrease in oscillation frequency in nutlin-pretreated cells (Figure 2-3I). We used this data to estimated the distribution of pulse frequencies in cells with or without nutlin pretreatment (Figure 2-3J). This difference is significant by the Kolmogorov-Smirnoff test indicates that the nutlin pretreatment frequency distribution is left-shifted with a p-value of $< 10^{-6}$; examining these distributions shows a distinct shift to lower frequencies. These results demonstrate that in this core NF loop, one of the most tightly controlled dynamical features – oscillation frequency – can be experimentally modulated.

Two major thrusts of systems biology are the better understanding of the operation of complex natural networks, and the *de novo* design of simple networks. The

current study suggests that these approaches need not be mutually exclusive, and that we can elucidate fundamental properties of complex signaling networks by isolating simpler subnetworks. In this work, we have described how individual feedback loops and interactions tune each of three distinct dynamical features. Manipulating the transcriptional rate of p53 using an inducible promoter tunes the amplitude of oscillation without affecting oscillation frequency or damping rates. Transcriptional synthetic feedback loops on p53 can modulate pulse damping with a less pronounced effect on amplitude and frequency. Finally, targeted perturbation of the p53-Mdm2 interaction leads to a modulation in oscillation frequency. Taken together, these results show that even a 'simple' oscillatory network motif – the delayed negative feedback loop – provides a platform allowing the independent modulation of three crucial dynamical features: amplitude, frequency and stability.

2.4 Methods

Cell lines and expression constructs

We used MCF7 cells stably transfected with MTp-p53-CFP and Mdm2p-MDM2-YFP as described [9]. To create the NPF and NNF plasmids, Mdm2p-MTF1-mCherry and Mdm2p-MTF1-KRAB-mCherry, we used MultiSite-Gateway recombination (Invitrogen). The human Mdm2 promoter [9], the MTF1 cDNA and MTF1-KRAB cDNA [99] and mCherry (gift from Dr. Tsien) were cloned into a modified pDEST-R4R3 vector containing the puromycin gene according to manufacturer's instructions via two sequential recombination reactions (Invitrogen). After transfection into the MCF7 cell line containing MTp-p53-CFP and Mdm2p-MDM2-YFP (FuGene6, Roche; [9]), cells were selected and clonal populations were obtained by single cell dilution.

Cell lines were grown at 37C in RPMI medium supplemented with 10% fetal bovine serum, 100U/mL penicillin, 100μg/mL streptomycin, 250ng/mL amphotericin B and appropriate selective antibiotics: G418 (0.4mg/mL), hygromycin (100μg/mL), and/or

puromycin (0.5μg/mL).

Time-lapse microscopy

Two days prior to microscopy, cells were plated onto poly-D-lysine-coated glass bottom plates (MatTek Corporation). One day prior to microscopy, the media was changed RPMI lacking ribloflavin and phenol red supplemented with only 2-5% fetal bovine serum and antibiotics to minimize autofluorescence. Nutlin-3A was added 24 hours prior to the start of microscopy (Fig. 5). Prior to the start of the movie, media was exchanged for fresh medium. Cells were viewed with two types of inverted fluorescence microscope systems named FMS1 (Figs. 1, 4, and 5) and FMS2 (Fig. 2). Each is surrounded by an enclosure to maintain constant temperature, CO_2 concentration and humidity. FMS1 consists of a Nikon Eclipse-TI-E perfect focus inverted microscope with a cooled CCD camera Hamamatsu Orca-R2. Brightfield, CFP, YFP, and mCherry images were taken every 20 min with Prior Lumen 200 metal arc lamp. FMS2 consists of a Nikon Eclipse TE2000-E inverted microscope with a cooled CCD camera Hamamatsu Orca-ER. CFP and YFP images were taken every 20 min with a mercury lamp. CFP filter set: 436nm/20nm; 455nm nm dichroic beam splitter, and 480nm/40nm emission. YFP filter set: 500 nm/20nm excitation, 515nm dichroic beam splitter, and 535nm/30nm emission (Chroma). mCherry filter set: 560nm/40nm excitation, 585nm dichroic beam splitter, and 630nm/75nm emission (Chroma). Images were acquired using MetaMorph software (Molecular Devices) for 48 hours.

Cell tracking and fluorescence quantification

Cell nuclei in the brightfield images were tracked manually in every frame using ImageJ (NIH). Mean nuclear fluorescence intensity was measured using custom written MATLAB software (Mathworks Inc.) which measured and subtracted each image's background fluorescence and excluded nucleolar regions from each tracked nucleus. Because of autofluorescence caused by the rounding up of cells near times of

cell division, fluorescence signal was masked and interpolated for 2 h before and after cell division events.

Data processing and automated pulse identification

For all data analysis, we followed the dynamics of a single daughter cell after each cell division event to avoid any bias arising from correlations between daughter cells. Trajectories were smoothened by Blaise filtering as described in [12, 13]. To identify pulse maxima and minima, trajectories were processed by reducing the depth local minima by 0.2 fluorescence units and performing the morphological opening operation with a width of 3 timepoints to exclude short, noisy fluctuations in amplitude. Maxima were identified from the processed trajectories using the watershed algorithm; minima were identified using the watershed algorithm on the negative reflection of the processed trajectory.

Computational methods

For all simulations, numerical integration was performed in MATLAB using ode15s (The Mathworks, Natick MA). Optimization was implemented using fmincon configured to use Quasi-Newton with BFGS in the MATLAB Optimization Toolbox Version 3.0.4.

Chapter 3

Distinct mechanisms act in concert to mediate sustained cell cycle arrest[1]

In the previous chapter, we developed and parameterized a model of p53 dynamics to match a wealth of experimental data. This model accurately reflects the operation of this signaling pathway but leaves open a crucial question: what role do signals from this upstream pathway in determining downstream cell decisions? This is an example of a broader challenge to systems biologists: as mathematical models of individual pathways emerge, we are challenged to interconnect them into a detailed understanding of how different pathways control the processing of information within the cell.

We chose to address this question in the context of one crucial downstream fate: the control of cell cycle progression in response to DNA damage. These networks are natural choices for such an integrative study. Each has been individually modeled successfully, and a great deal is understood about how specific interactions and regulation affect the dynamics of each network. However, in the absence of an extended model bridging these two pathways, the quantitative interaction between them remains undescribed. Here we develop a computational model of the combined net-

[1]This chapter has been previously published as Toettcher JE, Loewer A, Ostheimer GJ, Yaffe MB, Tidor B, Lahav G. *PNAS* **106**:785 (2009).

works and use it together with experimental measurements to determine the specific function of different cell cycle arrest mechanisms in response to DNA damage and their relative contribution to the proper execution of this cell decision.

During the cell cycle, mammalian cells coordinate cell growth, genome replication, and division. Two irreversible events subdivide the cell cycle into distinct phases: the onset of DNA replication defines S phase; and cell division defines M phase. Cells grow and carry out additional functions during the gap phases G1 and G2. The changing activity states of cyclin-dependent kinases (CDKs) regulate the transition between different stages of the cell cycle [101]. Cyclin D/Cdk4 and 6 and cyclin E/Cdk2 complexes drive the sequential progression from G1 to S phase, respectively. Cyclin A/Cdk2 and Cdk1 complexes become active during S and G2 phase, and cyclin B/Cdk1 complexes control the G2/M transition as well as various processes during mitosis. The cell cycle has long been a fruitful subject for mathematical modeling [102]. Models have proven useful for understanding the impact of perturbations to protein levels, network connections, and the cellular environment on cell cycle progression [103, 104].

A separate, well-studied regulatory network senses DNA double stranded breaks (DSBs) caused by ionizing radiation (IR). DSBs activate the ATM/Chk2 kinase cascade that phosphorylates p53, contributing to its stabilization and activation [27, 31, 32]. p53 transcriptionally modulates a variety of genes involved in cell cycle arrest, DNA repair, apoptosis, and in regulating p53 itself [10]. The feedback loops between p53, its upstream activating kinases ATM and Chk2, and its downstream regulators Mdm2 and Wip1 generate oscillatory dynamics in single cells [9, 12, 13]. Mathematical modeling contributed to understanding the dynamical behavior exhibited by this network as well [12, 47].

Upon DNA damage, interactions between the damage sensing and the cell cycle networks induce cell cycle arrest by modulating cyclin/Cdk activity. These interactions must fulfill three main requirements: first, to prevent alterations to the genome, they must relay the damage signal and halt the cell cycle promptly. Second, the arrest must persist as long as damage is present. Lastly, as cyclin/Cdk activity might

be changed during the arrest, cell cycle re-entry should only proceed from a state of cyclin activation that ensures the proper sequence of DNA replication and mitosis.

Multiple mechanisms that connect the DNA damage response to the cell cycle have been identified [105], and there is evidence for cooperation between some of them [106]. However, little is known about their relative contribution in the context of the full signaling networks. Furthermore, it is unclear whether individual mechanisms are sufficient to fulfill all of the above criteria, or if combinations of mechanism confer specific characteristics to a proper cell cycle arrest.

We address these questions systematically by combining experimental measurements of cell cycle distributions and cyclin levels together with the development of an integrated model of the DNA damage response and cell cycle networks. We find that individual arrest mechanisms act in concert to specifically establish immediate and sustained arrest after damage, as well as to prevent improper cell cycle re-entry.

3.1 A model of the DNA damage and cell cycle networks

We constructed an integrated model of the DNA damage response network and the cell cycle (Figure 3-1A). The model includes interactions previously studied in the context of the p53 network (shown in blue), and the cell cycle (shown in black). The interactions between the two networks represent the effect of DNA damage on the cell cycle (shown in red). The topology of the DNA damage model was derived from the model of Batchelor et al., in which oscillations are driven by a combination of two negative feedback loops: the core p53-Mdm2 loop and a loop in which the upstream checkpoint kinases are inhibited by a p53-inducible gene product, the phosphatase Wip1 [13]. To provide an extensible framework for future modeling of the DNA damage network, we incorporate additional feedback loops [33] in our model (Figure B-1A). With the current parameterization, however, these loops do not significantly affect the network's dynamics.

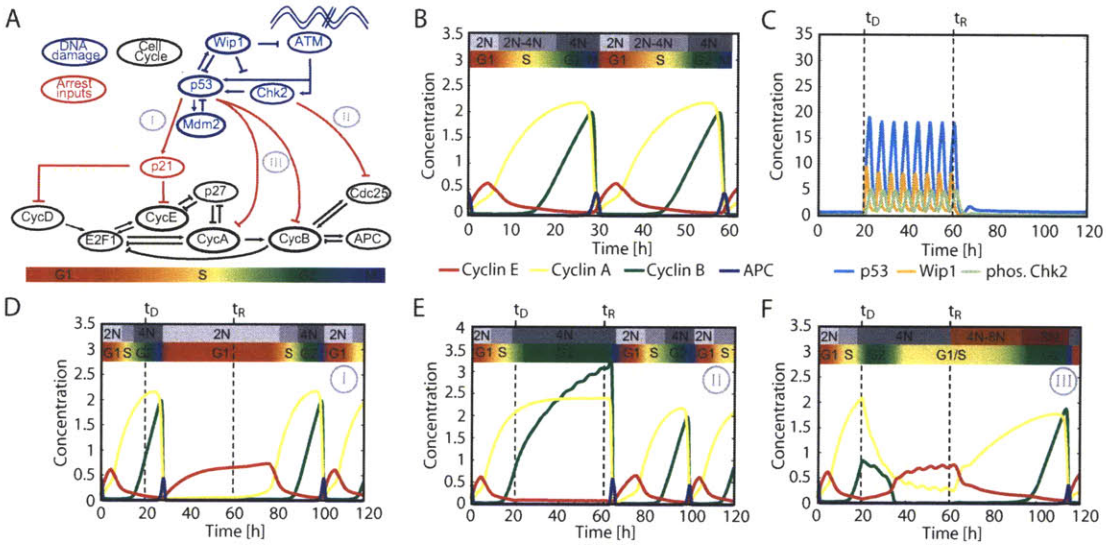


Figure 3-1: **Cell cycle and DNA damage models.** (A) Diagram of key species in the integrated model of DNA damage signaling (blue) and cell cycle arrest (black). Bridging connections consist of species modulating cell cycle arrest (red). The approximate cell cycle phases are shown below the diagram. Three classes of arrest mechanisms are indicated by numerals; (I) G1 arrest by p21 induction, (II) G2 arrest by G2 cyclin inactivation, and (III) G2 arrest by G2 cyclin transcriptional repression. (B) Cell cycle model simulation showing cyclins E, A, and B, and phosphorylated anaphase promoting complex (APC_P). Progression through cell cycle phases and changes in DNA content are indicated above the simulation. (C) Simulation of the DNA damage network after onset of damage at times t_D, until the repair time t_R. Nuclear p53, phospho-Chk2, and Wip1 species are shown. (D-F) Simulation of arrest mechanisms (I-III). Dynamics are influenced by the p53 and Chk2 activity from the DNA damage module shown in (C).

Our cell cycle model is based on the recently published cell cycle model of Tyson and colleagues [107]. This comprehensive model is comprised of generic network modules that have been parameterized to match data from yeast to mammals. To adapt the model as a platform to study cell cycle arrest in human cells, it was necessary to modify it in both parameterization and topology, while ensuring that it remains capable of recapitulating known experimental results.

Three classes of changes are introduced in the present study: (1) species previously treated at quasi-steady state with algebraic expressions were expanded to dynamic differential equations, (2) protein synthesis and degradation terms were added for each species in the model, and (3) the intracellular signal resulting from extracellular growth factor present in the medium, M, replaced the dependence between cell size and progression through the cell cycle [108] (see B.1).

Simulation of the freely cycling model shows qualitative similarity to trajectories obtained previously [107], with sequential peaks of cyclins E, A, and B defining G1, S, and G2 phase (Figure 3-1B). These cell cycle phases are associated with the transition from a 2N DNA content to 4N and the subsequent distribution of chromosomes to daughter cells during mitosis, which is represented by a peak in Anaphase Promoting Complex (APC) activity (Figure 3-1B) [101]. The model also matches a variety of experimental results from the literature including (i) G1 synchronization by serum starvation or cycloheximide treatment [109] (Figure B-1B,C), (ii) free cycling without cyclin E [110] (Figure B-1D), and (iii) G1 arrest at normal mitogen levels but continued cycling at high mitogen levels for the cyclin D knockout model (Figure B-1E) [111]. Taken together these results demonstrate that our cell cycle model recapitulates a wide range of experimental observations.

The two models were initially joined by incorporating well-described interactions that represent larger classes of G1 and G2 arrest mechanisms (Figure 3-1A, section B.1). For simplicity, we divided these mechanisms into three classes and analyzed one representative mechanism from each class (Figure 3-1A): (I) G1 arrest represented by p53-dependent inhibition of cyclin E/Cdk and cyclin D/Cdk complexes by p21 [112, 113], (II) p53-independent G2 arrest represented by posttranslational

inactivation of cyclin A/Cdk and cyclin B/Cdk complexes [38, 39], and (III) p53-dependent G2 arrest represented by transcriptional repression of cyclin A, cyclin B, and Cdk1 [34–36, 114].

3.2 Computational analysis of different arrest mechanisms

To assess the relative contribution of different arrest mechanisms, we implemented each mechanism individually and tested the resulting network behavior (Figure 3-1C-F). Cell cycle arrest was simulated by activating DNA damage between the time of damage (t_D) and recovery (t_R) (see B.1). The damage stimulus activated the p53 network, leading to oscillations of p53 and active Chk2 with a period of about 5.5 h (Figure 3-1C) [13]. Each arrest mechanism was capable of halting the cell cycle on its own, but there were distinct differences in the state of the network during the arrest (Figure 3-1D-F).

When damage was applied during G2 phase, mechanism I (implementing p21-dependent inhibition of Cyclin E/Cdk2) led to a stable arrest in G1 (2N DNA content) after one cell division (Figure 3-1D). During the arrest, p21 and cyclin E reached high levels. After removal of the damage signal, re-entry into S phase was delayed approximately 20 h. Mechanism II (implementing Chk2-mediated G2 cyclin inactivation) induced the arrest of 4N cells with high levels of cyclins A and B; upon re-entry, cells immediately entered mitosis (Figure 3-1E). In contrast, p53-dependent G2 cyclin downregulation implemented by mechanism III led to the arrest of 4N cells in which cyclin A and B levels progressively decreased and cyclin E levels were elevated (Figure 3-1F). Thus, in cells arrested by mechanism III, the cyclin network uncouples from the status of DNA replication: cyclins switch from a G2 to a G1/S-like state, even though cells do not divide. Upon deactivation of the damage stimulus, these cells increase cyclin A and B levels, thereby progressing twice through the cell cycle without an intervening mitosis. This implies a danger inherent to arrest mechanism

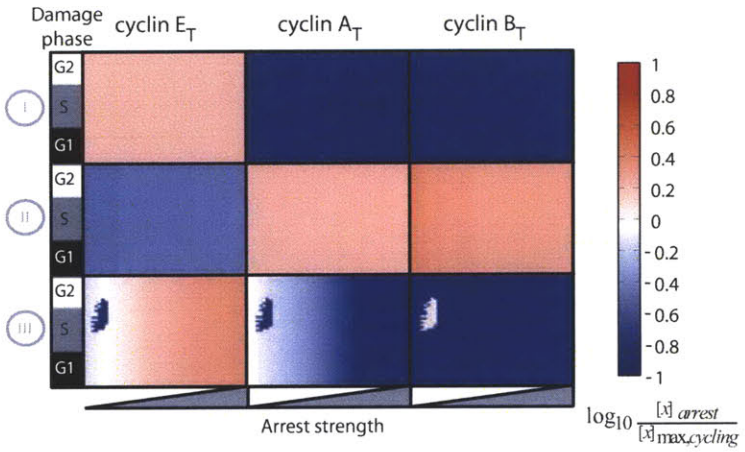


Figure 3-2: **Steady state cyclin levels during simulated arrest.** Rows indicate the arrest mechanism; columns indicate the cyclin measured. In each square both strength and time of arrest application are varied. Arrest strength was varied by scaling the value of the parameters implementing each arrest mechanism over two orders of magnitude from the minimum value required to generate arrest (see B.1). Colors indicate the ratio of steady state cyclin levels to their maximum level during normal cycling.

III: downregulation of the G2 cyclins might lead to loss of information about the cell cycle phase prior to damage and endoreduplication.

The dynamical behaviors of Figure 3-1 were obtained by applying damage at a specific time, with fixed parameters controlling p53's and Chk2's activation of each arrest mechanism (see B.1). To expand this analysis, we simulated cell cycle arrest varying the time of damage induction and the arrest strength. From the resulting trajectories, we determined the steady state levels of cyclins E, A, and B (see B.1; Figure 3-2). Mechanism I led to a moderate increase of cyclin E levels with low levels of the G2 cyclins A and B, consistent with arrest in G1. Notably, mechanisms II and III, which are both known to act in G2 cells with 4N DNA content, resulted in distinct cyclin profiles during the arrest: mechanism II led to increased levels of cyclins A and B and low levels of cyclin E, while mechanism III showed the reverse pattern in most simulations. At weak arrest strengths this mechanism led to bimodality in cyclin B and E profiles.

3.3 Experimental measurements of cell cycle arrest

Our modeling results indicate that the dominant arrest mechanism can be uniquely identified by measuring DNA content and cyclin levels in arrested cells. Within 8 h after irradiation, an asynchronous population of wild-type HCT116 cells arrested with approximately 30% of the cells in G1 (2N) and 70% in G2 (4N). The cells remained arrested for at least 72 h (Figure 3-3A). Notably, we observe only low levels of apoptosis during the time period studied (Figure 2F, [115]). Finer temporal sampling showed that the G2 arrest was immediate, while the G1 population initially decreased and was stabilized only after 4 h (Figure B-3A). This is consistent with the dependence of the G1 arrest on p53-mediated p21 expression (mechanism I, Figure B-3B). The ratio of cells arrested in G1 versus G2 thus depends on the initial cell cycle distribution and the kinetics of p21 expression. To determine the contribution of mechanisms II and III for the G2 arrested cells, we measured cyclin levels in the G2 population (4N). In freely cycling cells (Figure 3-3A; 0 h time point), most cells with 4N DNA content had high cyclin B1 and low cyclin E levels. Only a small population had low levels of cyclin B1, which likely represent post-metaphase mitotic cells. At 8 h after IR, most of the now-arrested G2 cells maintained high cyclin B1 and low cyclin E levels, suggesting that mechanism II was dominant during this time. At later times, the population of cells with low cyclin B1 increased, with nearly all cells having low cyclin B levels by 48 h. We observed corresponding decreases in cyclin A and Cdk1 levels during these times by western blot (Figure B-3C). Conversely, the cyclin E distribution increased dramatically, reaching levels that exceed those observed during G1 in freely cycling cells (Figure B-5A). These results argue that in cells initially arrested in G2, mechanism III gradually becomes dominant over mechanism II during the course of arrest, uncoupling the cyclin state (now G1/S like) from DNA content (4N). Similar results were obtained in non-transformed RPE-hTERT cells (Figure B-4A-B).

These observations led us to ask whether mechanism II remains active at late

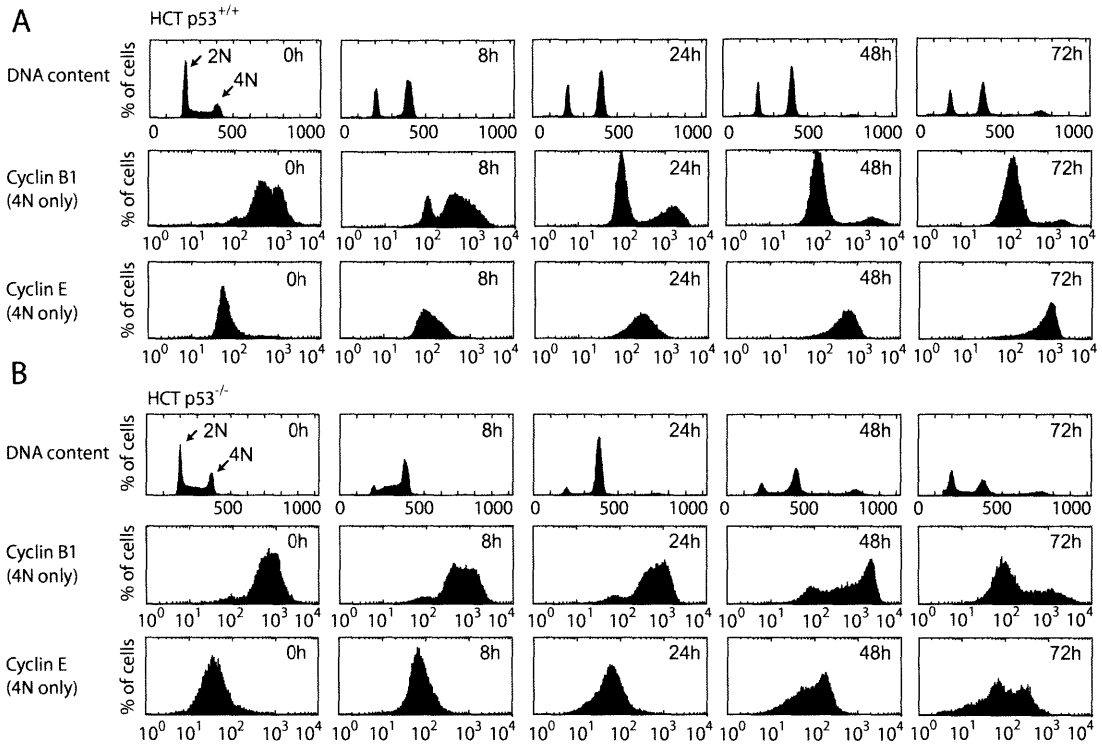


Figure 3-3: **Cell cycle progression and cyclin levels during arrest.** Flow cytometry histograms of DNA content and cyclin levels in HCT $p53^{+/+}$ (A) cells and $p53^{-/-}$ (B) after IR. Cells were irradiated and stained for DNA content and cyclin levels. Histograms of cyclin levels are gated from the 4N population only. Quantification of apoptotic cells is shown in Figure B-4C.

times after IR, or if it turns off after mechanism III is initiated. In the first case, both pathways play redundant roles, while in the second case, each mechanism is utilized at different times during arrest. To distinguish between these cases, we examined DNA content and cyclin levels in HCT116 cells lacking p53, which are restricted to using mechanism II (Figure 3-3B, Figure B-3A). These cells arrested in G2 (4N) by 16-24 h, with most cells retaining high cyclin B1 and low cyclin E levels. This is in agreement with a G2 arrest solely mediated by the p53-independent mechanism II. Most noticeably, the G2 arrest was transient: 24 h post irradiation, cells re-enter G1 (2N) and S, which is reflected by changes in cyclin levels. These results show that mechanism II is sufficient to induce an immediate G2 arrest (Figure 3-3B and Figure B-3A). However, sustained G2 arrest depends on mechanism III, arguing against redundancy of these mechanisms. Instead, we suggest that mechanisms II and III complement each other, with mechanism II operating on a fast and mechanism III on a slow time scale.

3.4 Merging model and measurements

Our original model addressed individual arrest mechanisms in the context of the generic mammalian cell cycle. After acquiring quantitative measurements of the combination of arrest mechanisms and their relative timings, we next set out to parameterize our model to reflect this data. Our fitting procedure accounts for the steady state concentrations of molecular species during arrest as well as the relative timing of their induction during cell cycle progression (see B.3). The sensitivities of both were calculated efficiently using an adjoint method. The fitted model recapitulated the amount of time spent in G1, S, and G2/M (Figure 3-4A) and matched the cyclin levels reached during arrest (Figure 3-4B). Cycloheximide treatment, serum starvation, and cyclin knock-outs still elicited the appropriate phenotypes (data not shown).

After fitting, two features of the arrest dynamics remained undetermined: the activation and deactivation time of each arrest mechanism. Based on the arrest profile of $p53^{-/-}$ cells (Figure B-3A), we assume that mechanism II is initiated immediately

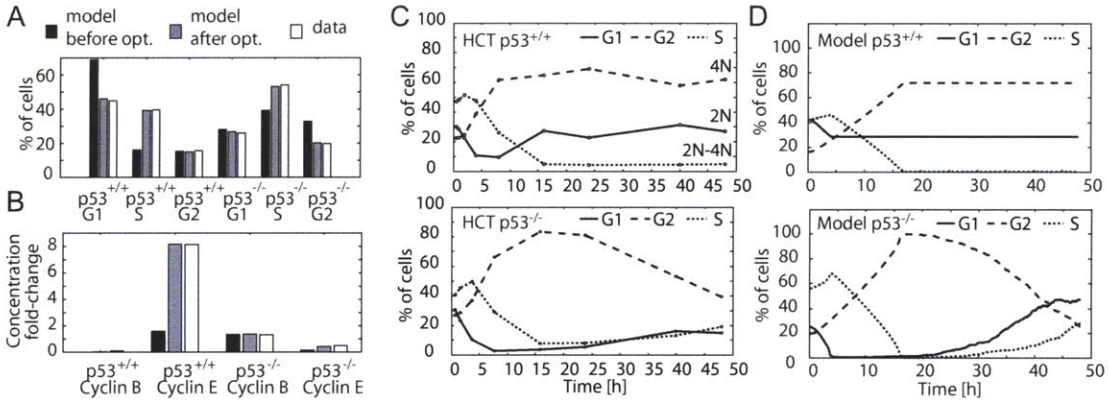


Figure 3-4: **Cell cycle model training and prediction.** (A) Fractions of G1, S, and G2 cells in freely cycling HCT $p53^{+/+}$ and HCT $p53^{-/-}$ populations are compared to the amount of time spent by the initial and the fitted model in G1, S, and G2. (B) The ratios of cyclins E and B during IR-induced arrest to their maximum level during normal cycling in both HCT $p53^{+/+}$ and $p53^{-/-}$ cells, compared to the ratios calculated from model trajectories. (C) Time courses of DNA content after treatment with 10 Gy IR. HCT $p53^{+/+}$ and $p53^{-/-}$ cells were irradiated, and the fractions of cells with G1, S, and G2 DNA contents were measured by FACS. For apoptotic fraction, see Figure B-4C. (D) Model-generated cell cycle distribution time courses. 500 individual model trajectories were simulated from initial conditions distributed through the cell cycle (see B.1).

after damage. It is gradually deactivated between 16-48 h, as indicated by the slow decrease of the G2-arrested population of $p53^{-/-}$ cells (Figure 3-4C). We modeled mechanism II deactivation times with a mean of 35 h and a standard deviation of 10 h, kinetics that are consistent with the stochastic repair of DSBs using repair rates found in literature (see B.1; Figure B-6A-B) [47]. For the p53-dependent mechanisms I and III, we define the activation time to be 4 h after damage, as p53 was induced at that time (Figure B-3B). In wild-type cells, the arrest was sustained through the time course (Figure 3-3A), during which p53 levels remained high (Figure B-3C). We therefore assume that mechanisms I and III are activated permanently on the time scale of our simulations.

Our final model implements the combined action of three arrest mechanisms, as well as their temporal organization, allowing direct comparison to experimentally measured distributions of arrested cells over time. Indeed, simulation of an asynchronous population of cells after arrest largely recapitulated the arrest dynamics and cyclin profiles of $p53^{+/+}$ and $p53^{-/-}$ cells (compare Figure 3-4C-D and Figure B-7A-B). Notably, in contrast to the modeling results, the HCT116 $p53^{+/+}$ G1 population increased 8 h after irradiation, although damage was still present, while RPE-hTERT cells do not show this behavior (Figure B-4B,D). This suggests that at early arrest times some cells escaped G2 arrest and entered G1, which may be the result of checkpoint defects in the cancer cell line used, phenomena that are not considered in our model.

3.5 Model validation and predictions

To validate the quality of our fitted model, we tested its ability to predict the mean protein levels in irradiated populations of $p53^{+/+}$ and $p53^{-/-}$ cells, and the behavior of freely cycling and arrested cells after silencing individual cell cycle genes (see B.4; Fig. S5C; Table B.2). These results demonstrate that the fitted model accurately reflects features of arrested mammalian cells.

One additional model prediction captured our attention. Simulating mechanism

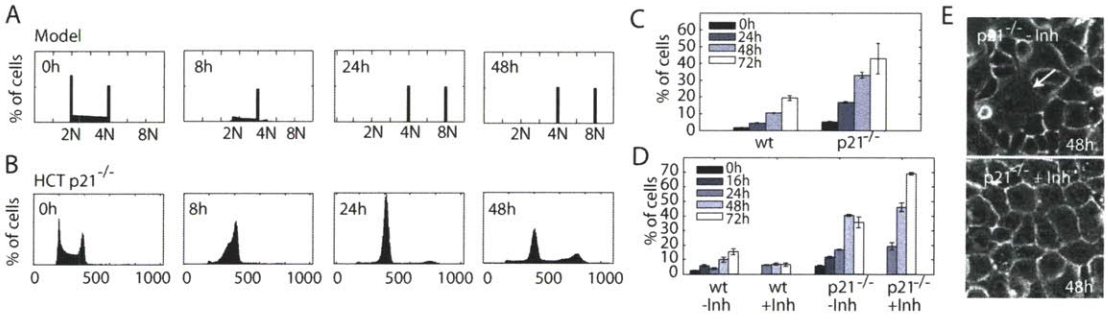


Figure 3-5: p21 function in preventing endoreduplication. (A) Model prediction of HCT $p21^{-/-}$ endoreduplication. $5x10^2$ individual model trajectories were simulated from initial conditions distributed through the cell cycle. (B) DNA content profiles for HCT $p21^{-/-}$ cells after irradiation with 10Gy IR. For apoptotic fraction, see Figure B-4C. (C) Fractions of cells with >G2 DNA content in HCT wt and $p21^{-/-}$ cell lines (mean ± SE) as a function of time after irradiation. (D) Fractions of cells with >G2 DNA content after irradiation followed by addition of Cdk1 inhibitor RO-3306 at 16 h post-IR (mean ± SE). The increased endoreduplication of $p21^{-/-}$ cells after inhibitor treatment can be explained by a decrease in cells entering mitosis prematurely, creating a larger pool of cells able to undergo a second S-phase. (E) Phase contrast images of $p21^{-/-}$ cells 72 h after irradiation. Without added inhibitor, cells are multinucleated (arrow), indicative of failures during mitosis. Cells treated with inhibitor are mononucleate but have also undergone endoreduplication.

III individually predicted that downregulation of G2 cyclins would prime cells for endoreduplication (Figure 3-1F). However, this phenomenon was not observed in the model combining all arrest mechanisms. We next interrogated the model to determine what feature of the combined arrest prevented endoreduplication. Simulations of mechanism III alone and of the final model resulted in high total cyclin E levels and low cyclin A and B1 levels during arrest (Figure 3-1F and Figure B-9A). However, in the final model, cyclin E was inhibited by p21 through mechanism I, suggesting that p21 prevented initiation of DNA replication of G2-arrested cells (Figure B-9B). Indeed, eliminating p21 from the final model predicted that approximately 50% of the cells will initiate a second S phase and endoreduplicate (Figure 3-5A and Figure B-9B).

To test this prediction experimentally, we measured the cell cycle distribution of a HCT116 cell line lacking p21 [112] (Figure 3-5B). Consistent with our modeling results and previous reports [115], approximately half of the $p21^{-/-}$ population had DNA content greater than 4N (Figure 3-5C). Surprisingly, about 20% of wild-type cells endoreduplicated as well. We suggest that this is caused by cells attempting to re-enter the cell cycle prematurely with remaining damage, leading to failures in mitosis [37].

Simulation of our model suggests that p21 prevents endoreduplication by inhibiting cyclin E/Cdk2 activity in G2-arrested cells. Alternatively, p21 has been implicated in inhibiting Cdk1 activity [37, 116]. In this scenario, elimination of p21 may cause cells to prematurely enter mitosis, fail, and ultimately endoreduplicate. Morphologically, a fraction of $p21^{-/-}$ cells possessed fragmented nuclei, supporting mitotic failure as a cause of endoreduplication (Figure 3-5E). To determine whether this is the only source of endoreduplication, we used the Cdk1 inhibitor RO-3306 [117]. This drug mimics the effect of p21 on Cdk1 and prevents cells from attempting mitosis. Under these conditions, G2-arrested $p21^{-/-}$ cells still re-replicated their DNA while remaining mononucleate, suggesting that cyclin E/Cdk2 inhibition by p21 is necessary as well to prevent endoreduplication (Figure 3-5D-E). Taken together, these results suggest that p21 plays an important role in maintaining the sustained G2 arrest by preventing

endoreduplication after downregulation of G2 cyclins.

3.6 Conclusions

An intricate network of protein interactions mediates cellular signaling. To facilitate our understanding, this network is often subdivided into individual units. However, these units do not act in isolation, as they influence each other through common interactions and complex feedbacks. Here we present the integration of two models of subnetworks by implementing specific, experimentally verified connections supplemented by a thorough investigation of the space of possible arrest mechanisms. We found that a variety of interactions lead to similar arrest profiles, and that the specific connections implemented are representative of these larger classes of arrests. One benefit of such an approach lies in the ability to individually study these mechanisms and their effect on the behavior of the integrated network. Furthermore, by fitting to experimental data, the model can be used to analyze the combined action of multiple mechanisms and their relative contribution to the signal processing.

Upon DNA damage, cells must activate arrest immediately, maintain it as long as the insult persists, and be prevented from re-entering into inappropriate cell cycle phases [105]. Our analysis shows that a combination of different arrest mechanisms contributes to fulfilling these requirements. However, the requirements seem to pose a paradox for G2-arrested cells: cells undergoing sustained arrest lower their G2 cyclin levels, whereas appropriate cell cycle re-entry depends on these cyclins to convey information about the pre-arrest state. To resolve this paradox, we propose that in response to high levels of DNA damage, cells that arrest by cyclin downregulation must do so permanently. Downregulation of cyclins by p53 may therefore be the first step in establishing senescence, a terminal cell fate characterized by the irreversible exit from the cell cycle [118]. Our model can now be used to generate testable predictions for thresholds in time, damage levels, and cyclin concentration that define the decision between cell-cycle re-entry and senescence.

Our integrated model also revealed a function for p21 in sustaining G2 cell cycle

arrest. p21's involvement in G2 and endoreduplication arrest has been reported before [115, 116], but the exact mechanism remains less well defined. p21 was previously suggested to inhibit Cdk1-activating kinases [119] or alter the subcellular locations of the Cdk1 complex [120]. In addition, we now propose that p21 contributes to a sustained G2 arrest by inhibiting G1 cyclins. This function is crucial to prevent DNA replication after downregulation of G2 cyclins and may explain previously observed endoreduplication after recovery from transient p21 overexpression [121] or mitotic spindle disruption [122] in cells lacking p21.

In the present study, we have abstracted certain processes involved in the DNA damage and cell cycle networks. For example, we do not address all of the details of DNA repair, but rather rely on a simple stochastic representation of this process. Additionally, we use the activation of APC as a surrogate for the complex process of mitosis. While this abstracted model was sufficient to characterize the interactions transmitting the DNA damage signal to the cyclin network, a detailed treatment of these processes would allow us to address further questions. For example, we show that p53 activation is sustained for at least 96 h (Figure B-3C), and that cells lacking p53 re-enter the cell cycle after 24 h (Figure 3-4C). This indicates either that DNA damage is repaired by 24 h and p53 activity is sustained after DNA repair is complete, or that DNA damage persists and cells lacking p53 adapt to the damage checkpoint. Including the details of mitotic progression and the possibility of mitotic failure would provide the framework necessary to better understand checkpoint adaptation. Accounting for the details of DNA damage and its repair may reveal whether sustained p53 activation is mediated by persistent damage or by the activation of additional network interactions.

3.7 Materials and Methods

Cell culture

HCT116 $p53^{+/+}$, $p53^{-/-}$, and $p21^{-/-}$ cells were grown in McCoy's media including 10% FBS under standard conditions. $5x10^5$ (6-cm dish) or $1.5x10^5$ (10-cm dish) cells were plated and irradiated two days later with 10 Gy using a Co60 source.

Immunoblots

Western blots were performed as described before [13]. Antibodies used were αp53 DO-1, αCyclinB1 (H433), αCyclinA (C-19), αCdk1 (all Santa Cruz Biotechnology), αp21 (Calbiochem), and $\alpha\beta$-tubulin (E7, Developmental Studies Hybridoma Bank).

Flow Cytometry

Cells were trypsinized and fixed in 70% ethanol at -20C. For DNA content analysis, cells were washed in PBS, incubated with 25 μg/ml propidium iodide (PI), 0.1% Triton, and 0.2 μg/ml RNase and analyzed on a FACSCalibur flow cytometer (BD Biosciences). For cyclin A, E, and B labeling, fixed cells were washed, permeabilized in 0.25% Triton, and blocked in 0.5% BSA. 1x106 cells were incubated with 1 μg primary antibodies, washed and incubated with Alexa488-coupled secondary antibody. Cells were stained with PI and analyzed as above (CyclinE (HE12), Santa Cruz Biotechnology).

Only cell singlets were analyzed, based on the pulse width versus height ratio. To obtain the percentages of G1, S, G2/M, and endoreduplicated cells, we computationally fit the DNA content distributions using a modification of the Dean-Jett model, augmented to include the 8N and second S phase population [123].

Computational methods

For all simulations, numerical integration was performed in MATLAB using ode15s (The Mathworks, Natick MA). Optimization was implemented using fmincon config-

ured to use Quasi-Newton with BFGS in the MATLAB Optimization Toolbox Version 3.0.4.

Chapter 4

Oscillator sensitivity analysis in the presence of hidden conservation constraints

4.1 Introduction

The numerical techniques for analyzing oscillators whose trajectories satisfy systems of autonomous ordinary differential equations (ODEs), such as the systems generated by lumped-element models of circuits or mass-action kinetics models of biochemical networks, have been investigated from so many perspectives it is hard to imagine that there anything new to report. The basic finite-difference, spectral, and shooting-Newton methods have been known for decades [124, 125], the issues associated with computing the impact of noise in oscillators has been examined from a variety of perspectives [126–128], and a number of research groups in different application domains have investigated approaches to computing parameteric sensitivities [83, 129, 130].

Nevertheless there does seem to be an unknown subtlety, at least that the authors have yet to find in the literature, associated with what we refer to as hidden conservation constraints. The subtlety shows itself in the simplest of experiments. Consider

the following simple differencing procedure for computing oscillator parameteric sensitivity:

1. Starting from an initial condition, x_o, integrate the differential equation system until the computed trajectories form a periodic orbit with period T;

2. Perturb some system parameter of interest;

3. Integrate the perturbed differential equation system, again starting from the initial condition x_o, until the computed trajectories form a new periodic orbit with period $\hat{T}$.

The surprising result, and the key result of this chapter, is that standard approach to oscillator parameteric sensitivity analysis *generates results that do not match the results from the above differencing algorithm.* Even the period sensitivity is inconsistently predicted.

It should be noted that hidden conservation constraints are quite common in mass-action kinetics models of biochemical systems, weighted sums of species are often conserved by sequences of reversible reactions. Hidden conservation constraints are less common in circuits, but can occur if there are cut-sets of capacitors in the circuit, as sometimes happens in circuits with MOS devices. If there is a cut-set of capacitors, then there is usually a hidden conservation constraint, a weighted sum of node voltages that is conserved.

We begin in the next section by presenting a brief review of the state-transition function approach to computing oscillator periodic steady-states, and the techniques for computing sensitivities for oscillating systems. In Section 4.3 we are more precise about conservation constraints, and describe their impact on oscillator analysis. In Section 4.4 it is shown how neglecting these conservation constraints can impact the solution of the oscillator sensitivities for two examples, a classical circuit oscillator and a model of a canonical biochemical oscillator, the cyanobacterial circadian clock.

4.2 Background

4.2.1 Oscillators described by ODEs

For simplicity, we will consider systems of differential equations in normal form, though most of the results generalize directly to the nonlinear descriptor systems typically generated by modified nodal analysis of circuits. Specifically, consider the ordinary differential equation system

$$\dot{\mathbf{x}}(t,\mathbf{p}) = \mathbf{f}(\mathbf{x}(t,\mathbf{p}),\mathbf{p}) \tag{4.1}$$

$$\mathbf{x}(0,\mathbf{p}) = \mathbf{x}_0 \tag{4.2}$$

where x is the vector of state variables of size n and p is a vector of parameters of size m, and $\mathbf{f}(\mathbf{x},\mathbf{p})$ describes the time evolution of the trajectory $\mathbf{x}(t,\mathbf{p})$. Note that the system is autonomous, the function $\mathbf{f}$ does not depend explicitly on time. Such a system oscillates for a particular choice of $\mathbf{x}(t,\mathbf{p})$ if there exists a $T(\mathbf{p}) > 0$ such that

$$\mathbf{x}(T(\mathbf{p}),\mathbf{p}) = \mathbf{x}(0,\mathbf{p}) \tag{4.3}$$

is satisfied. Also, if $\mathbf{f}$ is well-behaved (Lipschitz continuous in $\mathbf{x}$ and continuous in t), uniqueness guarantees that satisfying Equation 2 implies

$$\mathbf{x}(t+T(\mathbf{p}),\mathbf{p}) = \mathbf{x}(t,\mathbf{p}) \quad \forall t. \tag{4.4}$$

When analyzing oscillating ODE systems, it is convenient to consider the state transition function defined by solving (4.2),

$$\mathbf{x}(\tau,\mathbf{p}) = \mathbf{F}(\tau,\mathbf{x}_0,\mathbf{p}). \tag{4.5}$$

The state transition function is implicitly defined, and describes how (4.2) maps an initial condition, x_0, and an interval τ to x at time τ.

The state transition function can be differentiated with respect to x_0, yeilding

$$\frac{\partial \mathbf{x}(\tau, \mathbf{p})}{\partial \mathbf{x}_0} = \frac{\partial \mathbf{F}(\tau, \mathbf{x}_0, \mathbf{p})}{\partial \mathbf{x}_0} \tag{4.6}$$

where $\frac{\partial \mathbf{F}}{\partial \mathbf{x}}$ is the Jacobian with respect to x of the state transition function, and usually referred to as the sensitivity matrix.

In the special case where $x(t, \mathbf{p})$ is a periodic solution of (4.2) with period $T(\mathbf{p})$, the sensitivity matrix has special name and special properties. To see this, consider that if $x(t, \mathbf{p})$ satisfies (4.4), then differentiating (4.5) with respect to time yields

$$\begin{aligned} \dot{\mathbf{x}}(t + T(\mathbf{p}), \mathbf{p}) &= \dot{\mathbf{x}}(t, \mathbf{p}) && (4.7) \\ &= \frac{\partial \mathbf{F}}{\partial \mathbf{x}}(T(\mathbf{p}), \mathbf{x}(t, \mathbf{p}))\, \dot{\mathbf{x}}(t, \mathbf{p}) && (4.8) \end{aligned}$$

where the sensitivity matrix, $\frac{\partial \mathbf{F}}{\partial \mathbf{x}}$, for this periodic special case is referred to as the Monodromy matrix. The Monodromy matrix has the important property that one its eigenvalues is one, with a corresponding eigenvector equal to $\dot{\mathbf{x}}(t)$.

4.2.2 Computing the periodic steady state

Given a system (4.2), we can solve for a periodic solution $(\mathbf{x}_0(\mathbf{p}), T(\mathbf{p}))$ by numerically solving the system of equations

$$\mathbf{x}_0(\mathbf{p}) - \mathbf{F}(T(\mathbf{p}), \mathbf{x}_0(\mathbf{p})) = 0. \tag{4.9}$$

Though not guaranteed to converge, the iterative Newton's method is often used to solve (4.9). Each Newton iteration involves solving the system

$$\begin{bmatrix} \mathbf{I}_n - \frac{\partial \mathbf{F}}{\partial \mathbf{x}}(\mathbf{x}_0^k, T^k) & \frac{\partial \mathbf{F}}{\partial \mathbf{x}} \end{bmatrix} \begin{bmatrix} \Delta \mathbf{x}_0^{k+1} \\ \Delta T^{k+1} \end{bmatrix} = \left[\mathbf{x}_0^k - \mathbf{F}(x_0^k, T^k)\right]. \tag{4.10}$$

where $\mathbf{I}_n$ denotes the $n \times n$ identity matrix, and the superscript k refers to the k^{th} Newton's method iteration. The system solutions are used to update guesses of the

oscillation period and a point on the periodic orbit, as in

$$(\mathbf{x}_0^{k+1}, T^{k+1}) = (\mathbf{x}_0^k, T^k) + (\Delta\mathbf{x}_0^{k+1}, \Delta T^{k+1}), \tag{4.11}$$

and the procedure is repeated until convergence is achieved.

It can be readily seen that the system described in (4.10) is underdetermined, there are n equations but $n+1$ unknowns. This indeterminacy stems from the fact that (4.9) has a continuum of solutions, associated with any point on a periodic orbit. In particular, once a periodic solution is found, such a solution can always be perturbed in the direction $\dot{\mathbf{x}}(0, \mathbf{p})$ without changing $\frac{\partial T}{\partial \mathbf{p}}$ because (4.8) guarantees that such a perturbation will always lie in the null space of $\mathbf{I}_n - \frac{\partial \mathbf{F}}{\partial \mathbf{x}}(\mathbf{x}_0^k, T^k)$.

This indeterminacy, often referred to as an indeterminacy in phase, can be resolved by adding an additional constraint, known as a phase locking condition, to select a particular point in time along the periodic orbit of interest. Common phase locking conditions include fixing the value some state variable or requiring that a particular state variable reach an extremum at $t = 0$. These conditions can be represented by ϕ_1 and ϕ_2 such that

$$\phi_1 : \mathbf{x} \rightarrow x_i \quad \phi_1(\mathbf{x}) = x_{i0} \tag{4.12}$$

$$\phi_2 : \mathbf{x} \rightarrow f_i(\mathbf{x}, \mathbf{p}) \quad \phi_2(\mathbf{x}) = 0 \tag{4.13}$$

Such an equation can be provided as an additional constraint in the Newton's method formulation (4.10), leading to the augmented system

$$\begin{bmatrix} \mathbf{I}_n - \frac{\partial \mathbf{F}}{\partial \mathbf{x}} & \frac{\partial \mathbf{F}}{\partial \mathbf{x}} \\ \frac{\partial \phi}{\partial \mathbf{x}} & 0 \end{bmatrix} \begin{bmatrix} \Delta\mathbf{x}_0^k \\ \Delta T^k \end{bmatrix} = \begin{bmatrix} \mathbf{x}_0^k - \mathbf{F}(x_0^k, T^k) \\ \phi(\mathbf{x}_0^k) - \phi_0 \end{bmatrix}. \tag{4.14}$$

In general, such a phase locking condition will be sufficient to isolate a solution

along an oscillating trajectory if

$$\frac{\partial \phi}{\partial \mathbf{x}} \dot{\mathbf{x}}(0, \mathbf{p}) \neq 0 \tag{4.15}$$

thereby removing $\dot{\mathbf{x}}(0, \mathbf{p})$ from the left hand matrix's null space.

4.2.3 Computing parametric sensitivities

In a variety of applications, such as oscillator circuit design [130] or determining dominant mechanisms in circadian rhythms [83], it is important to understand how properties of the periodic solution (such as $\mathbf{x}_0$ and T) vary with changes in $\mathbf{p}$. As mentioned in the introduction, it is possible to compute parameteric sensitivities as follows: for some choice of $\mathbf{p}$, integrate the system for a long time to obtain a periodic orbit. Then, after a small perturbation to the parameters $\mathbf{p}+\Delta\mathbf{p}$, integrate the system from the same initial conditions a second time and use the differences between the periodic solutions to compute approximate sensitivities. Any alternative method for computing sensitivies should match this finite differencing technique, at least in the limit of small $\Delta\mathbf{p}$.

One technique for obtaining the parametric sensitivity pair $\left(\frac{\partial \mathbf{x}_0}{\partial \mathbf{p}}, \frac{\partial T}{\partial \mathbf{p}}\right)$ involves solving a two-point boundary value problem that is similar to that described in Section 4.2.2, once a trajectory satisfying (4.9) is known. Differentiating (4.9) and (4.13) with respect to $\mathbf{p}$ results in a system that can be used to compute $\left(\frac{\partial \mathbf{x}_0}{\partial \mathbf{p}}, \frac{\partial T}{\partial \mathbf{p}}\right)$:

$$\begin{bmatrix} \mathbf{I}_n - \frac{\partial \mathbf{F}}{\partial \mathbf{x}} & \frac{\partial \mathbf{F}}{\partial \mathbf{x}} \\ \frac{\partial \phi}{\partial \mathbf{x}} & 0 \end{bmatrix} \begin{bmatrix} \frac{\partial \mathbf{x}_0}{\partial \mathbf{p}} \\ \frac{\partial T}{\partial \mathbf{p}} \end{bmatrix} = \begin{bmatrix} \frac{\partial \mathbf{F}}{\partial \mathbf{p}}(x_0, T, p) \\ 0 \end{bmatrix}, \tag{4.16}$$

where terms have the same meaning as in (4.14).

4.3 Hidden Conservation Constraints

In this section, we consider how the analyses above change when our ODE system contains hidden conservation constraints, ones that are not explicitly represented in

the ODE description. If weighted sum of certain state variables are invariant during time-evolution of the system, such constraints take the form

$$\mathbf{C}^T\mathbf{x}(t,\mathbf{p}) = \mathbf{C}^T\mathbf{x}(0,\mathbf{p}) \qquad \forall t, \mathbf{x}(0,\mathbf{p}) \tag{4.17}$$

where $\mathbf{C}$ is a $n \times k$ matrix containing the k conservation constraints of the system. Such constraints can arise from a variety of modeled phenomena. For example, in mass action chemical kinetics models, subsets of chemical species that interact through reversible reactions may have concentrations that always sum to a constant, or in lumped-element circuit analysis a subset of voltages in a circuit may always sum to a constant to maintain a certain net charge. Furthermore, such conserved quantities may not be obvious by inspection of either a network diagram or the associated ODE system. In this section, we show that these conservation laws play a crucial role in the proper computation of the periodic orbit and sensitivities using the boundary value equations (4.14) and (4.16).

First, note that if there are conservation conditions, they are always left eigenvectors of the sensitivity matrix with eigenvalues associated with one.

Theorem 1 *If a solution to (4.2) with initial condition x_0 always satisfies*

$$C^T\mathbf{x}(t,\mathbf{p}) = C^T\mathbf{x}_0 \tag{4.18}$$

then

$$\left(\frac{\partial \mathbf{F}}{\partial \mathbf{x}}\right)^T C = C \tag{4.19}$$

Proof Using the definition of the state transition function and multiplying by C^T yields

$$C^T\mathbf{x}(t,\mathbf{p}) = C^T\mathbf{F}(t,\mathbf{x}_0,\mathbf{p}). \tag{4.20}$$

Noting that $C^T\mathbf{x}(t,\mathbf{p}) = C^T\mathbf{x}_0$, (4.20) becomes

$$C^T\mathbf{x}_0 = C^T\mathbf{F}(t,\mathbf{x}_0,\mathbf{p}). \tag{4.21}$$

and differentiating with respect to $\mathbf{x}_0$ results in

$$C^T = C^T \frac{\partial \mathbf{F}(t, \mathbf{x}_0, \mathbf{p})}{\partial \mathbf{x}_0}. \tag{4.22}$$

Transposing both sides of (4.22) proves the theorem. ∎

Since the hidden conservation conditions are left eigenvectors of any sensitivity matrix, they can be extracted numerically with reasonable reliability.

Theorem 2 *Consider the system defined by (4.2), (4.3) with k hidden conservation constraints (4.17). The oscillating sensitivity system (4.23) is not overdetermined.*

$$\begin{bmatrix} \mathbf{I}_n - \frac{\partial \mathbf{F}}{\partial \mathbf{x}} & \frac{\partial \mathbf{F}}{\partial \mathbf{x}} \\ \mathbf{C}^T & 0 \\ \frac{\partial \phi}{\partial \mathbf{x}} & 0 \end{bmatrix} \begin{bmatrix} \frac{\partial \mathbf{x}_0}{\partial \mathbf{p}} \\ \frac{\partial T}{\partial \mathbf{p}} \end{bmatrix} = \begin{bmatrix} \frac{\partial \mathbf{F}}{\partial \mathbf{p}}(x_0, T) \\ 0 \\ 0 \end{bmatrix}, \tag{4.23}$$

Proof First, we show that systems (4.16) can always be augmented with k additional equations representing these conservation constraints. By premultiplying (4.5) by $\mathbf{C}^T$ and differentiating with respect to $\mathbf{x}_0$, we find that

$$\mathbf{C}^T \left(\mathbf{I}_n - \frac{\partial \mathbf{F}}{\partial \mathbf{x}}(\mathbf{x}_0, t) \right) = 0 \qquad \forall t, \mathbf{x}_0. \tag{4.24}$$

By the biorthogonality of left and right eigenvectors for the matrix $\frac{\partial \mathbf{F}}{\partial \mathbf{x}}$ there exist at least k right eigenvectors $\mathbf{v}_i$ satisfying

$$\left(\mathbf{I}_n - \frac{\partial \mathbf{F}}{\partial \mathbf{x}}(\mathbf{x}_0, t) \right) \mathbf{v}_i = 0 \qquad i \in 1 \ldots k \tag{4.25}$$

and for which $\mathbf{C}^T \mathbf{v}_i \neq 0$. Furthermore, by differentiating (4.17) with respect to t we find that

$$\mathbf{C}^T \dot{\mathbf{x}}(t, \mathbf{p}) = 0 \qquad \forall t, \mathbf{p}. \tag{4.26}$$

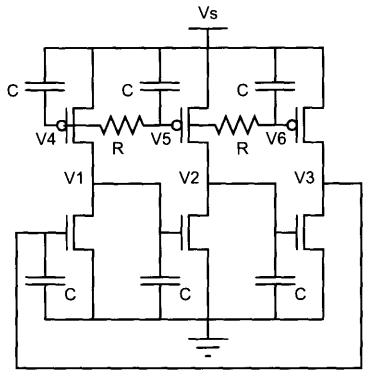


Figure 4-1: The ring oscillator circuit.

Taken together with (4.8), (4.26) shows that $\dot{\mathbf{x}}$ is still in the null space of the first $n + k$ rows of the matrix in (4.23). Thus, a phase locking condition satisfying (4.17) can be included as an additional row in the augmented system (4.23). An analogous argument justifies including these conservation constraints when using Newton's method to compute the periodic steady state (4.14). ∎

4.4 Examples

We have shown that inclusion of conservation constraints is always feasible in the context of the two-point boundary value formulation of oscillator sensitivity analysis. In this section, we will demonstrate that inclusion of these constraints is not only feasible but *necessary* to ensure that the sensitivities will accurately match the direct finite difference calculation,. We demonstrate this necessity in the following section through considering two examples of oscillators with hidden conservation constraints, a simple ring oscillator circuit and a minimal biochemical circadian oscillator from *Synechococcus elongatus.*

4.4.1 A simple ring oscillator circuit

As a first numerical example, we consider the simple oscillator circuit shown in Figure 4-1. For simplicity, we eliminate the capacitance variable from the resulting equations by setting the value of $C = 1$. A system of equations governing the

dynamics of this circuit can be derived of the form

$$\begin{aligned}
\dot{V}_1 &= i_{DS}^P(V_S - V_4 - V_T, V_S - V_1) - i_{DS}^N(V_3 - V_T, V_1) \\
\dot{V}_2 &= i_{DS}^P(V_S - V_5 - V_T, V_S - V_2) - i_{DS}^N(V_1 - V_T, V_2) \\
\dot{V}_3 &= i_{DS}^P(V_S - V_6 - V_T, V_S - V_3) - i_{DS}^N(V_2 - V_T, V_3) \\
\dot{V}_4 &= \frac{1}{R}(V_5 - V_4) \\
\dot{V}_5 &= \frac{1}{R}(V_4 + V_6 - 2V_5) \\
\dot{V}_6 &= \frac{1}{R}(V_5 - V_6)
\end{aligned}$$

where V_i are the voltages at the six nodes depicted in Figure 4-1, and the currents through each MOSFET i_{DS} are each given by the simple transistor model

$$i_{DS}(V_{OV}, V_{DS}) = \begin{cases} k\left(V_{OV} \cdot V_{DS} - 2V_{DS}^2\right) & V_{DS} < V_{OV} \\ kV_{OV}^2 & V_{DS} \geq V_{OV}. \end{cases} \tag{4.27}$$

In this model, $V_{OV} = V_{GS} - V_T$ is the overdrive voltage, and voltages are taken with the appropriate sign conventions for N- and P-type devices. In this circuit example, the system contains a single conservation law for the voltages at the gates of the PMOS devices, given by

$$V_4 + V_5 + V_6 = V_{Tot}. \tag{4.28}$$

We solved for the parameter sensitivities using multiple methods, either incorporating or neglecting the conservation constraints on this system. Table 4.1 shows the sensitivities to the NMOS and PMOS transistor parameters k_N and k_P and the source voltage V_S. These sensitivities were computed three ways: without the conservation equation 4.28 in the period sensitivity calculations ('raw'); with conservation ('cons'); and using a finite-difference approximation to the full sensitivities using the differencing procedure outlined in Section 4.1 ('FD'). The overbars indicate that each

	$\frac{\partial \bar{T}}{\partial \bar{\mathbf{p}}_{raw}}$	$\frac{\partial \bar{T}}{\partial \bar{\mathbf{p}}_{cons}}$	$\frac{\partial \bar{T}}{\partial \bar{\mathbf{p}}_{FD}}$
k_N	-3.1065×10^{-5}	-2.6506×10^{-1}	-2.6476×10^{-1}
k_P	-4.1779×10^{-5}	-7.3494×10^{-1}	-7.3427×10^{-1}
V_S	-2.3996×10^{-4}	-4.4972	-4.4948

Table 4-1: Normalized period sensitivities, ring oscillator.

sensitivity shown has been normalized to the total period and parameter value (i.e. $\partial \bar{x} = \partial log x = \frac{1}{x} \partial x$). We find close agreement between sensitivities computed with conservation and the differencing scheme, and a wide discrepancy between these results and the naive boundary value scheme. Each of these sensitivity calculations was initiated from the same point on the system's periodic orbit, thereby showing that even when a proper periodic orbit has been previously obtained, neglecting conservation can lead to an incorrectly computed period sensitivity.

4.4.2 A mass-action biochemical model

Conservation constraints arise naturally in many biochemical networks as a result of the conservation of mass of chemical species. Due to the size and complexity of the resulting networks, however, these conservation laws are often difficult to determine by direct examination of the underyling rate equations. In this section, we consider an example of an oscillating biochemical network harboring such conservation relations, a mass action model of the *in vitro* cyanobacterial circadian clock [65].

This network consists of three proteins – KaiA, KaiB, and KaiC – that associate with one another and undergo modification. Surprisingly, the canonical behavior of a circadian clock, persistent oscillations with a period of about 24 h, can be experimentally reconstituted in a minimal solution containing only the three purified proteins and their cofactors [61].

Our model of this network, based on that of [65], tracks every complex formed by the association and dissociation of the three Kai proteins, and consists of $n_x = 73$ state variables, $n_r = 277$ distinct chemical reactions, and $n_p = 22$ parameters. However, the combinatorial complexity of these reactions can be abstracted to yield a simple diagram containing 12 species representing the four phosphorylation states of

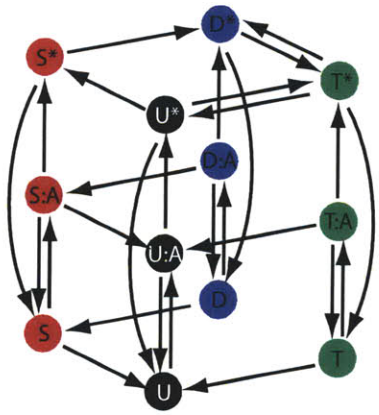


Figure 4-2: The cyanobacterial *in vitro* circadian clock. The twelve major states of KaiC are pictured, with arrows representing reactions that modulate transfer between states.

KaiC (U - unphosphorylated; S - phosphorylated at Ser431; T - phosphorylated at Thr432; D - doubly phosphorylated) in three KaiA-related states (lower plane - free KaiC; middle plane - KaiC:KaiA complexes; upper plane - KaiC* after modification by KaiA) (Figure 4-2).

To identify mass conservation reactions in this example, we implemented a general solution that relies on the structure of mass-action biochemical models. Our model of the *in vitro* cyanobacterial circadian clock is of the form

$$\dot{\mathbf{x}} = \mathbf{S} \cdot \mathbf{v}(\mathbf{x}, \mathbf{p}) \tag{4.29}$$

where $\mathbf{S}$ is the $n_x \times n_r$ stoichiometry matrix, indicating the number of molecules of each species that participates in each reaction, and $\mathbf{v}(\mathbf{x}, \mathbf{p})$ is the $n_r \times 1$ vector of the reaction rates. The conservation constraints exhibited by such systems can be identified automatically from the structure of such models since they comprise a basis for the left null space of the stoichiometry matrix $\mathbf{S}$ [131]. Computing this null space for our model identifies three such conservation constraints, one for the total

Method	No conservation	Conservation
T	20.000 h	20.468 h

Table 4-2: Oscillation period, circadian clock model.

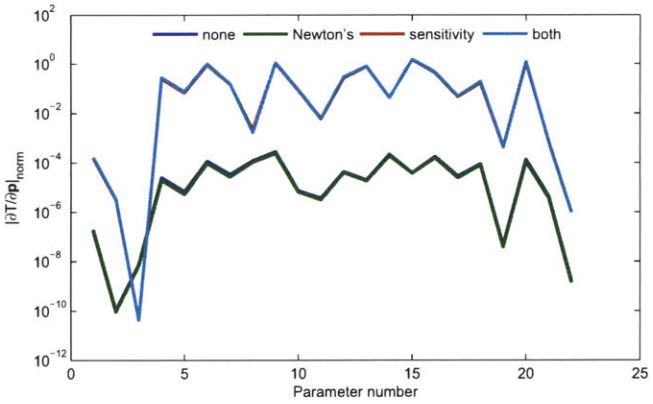


Figure 4-3: Normalized sensitivities, circadian clock model.

concentration of each of the three Kai proteins.

With these conservation relations in hand, we set out to identify whether or not their inclusion significantly alters the periodic initial conditions identified by Newton's method or the sensitivity of the period to parameters computed from this periodic trajectory. We computed conditions on the limit cycle using (4.14) or by augmenting this system to also include the three mass conservation constraints, which led to different periodic initial conditions and periods of oscillation (Table 4.2).

We next asked whether the differences in period sensitivity in the presence or absence of such constraints were dominated by differences in the periodic orbit obtained, or by the subsequent sensitivity calculation. To test this, we performed the sensitivity analysis with and without conservation constraints, from either of the two periodic orbits identified in Table 4.2. Figure 5-3 shows the results of this calculation, where the color of each curve indicates at which stage mass conservation was incorporated. Two pairs of curves are evident, separated by whether or not conservation is included in the sensitivity calculation. Our results indicate that including constraints in the sensitivity calculation is cruicial to their accurate computation, and that (at least for this example), this source of error is dominant.

4.5 Conclusions

We showed that in order for the standard approach to parameteric sensitivity analysis of oscillators to yield correct results, hidden conservation conditions must be explicitly included. We also showed that the conservation conditions can be extracted by computing left eigenvectors of the the sensitivity matrix at any timepoint, so multiple timepoints can be used to extract these hidden conservation conditions reliably. We also showed that adding the conservation conditions does not make the matrix needed to compute sensitivities over-determined. Finally, we demonstrated that neglecting conservation constraints results serious errors for the period and orbit sensitivities for a simple ring oscillator circuit and a model of a canonical biochemical oscillator, the cyanobacterial circadian clock. In the next chapter, we show how these tools can be used to gain a new understanding of the mechanisms underlying the operation of a mass conserved biochemical oscillator, the cyanobacterial circadian clock.

Chapter 5

Coupled oscillators drive dynamics in the cyanobacterial circadian network

5.1 Introduction

Cells use signaling networks to determine how and when to divide, communicate with one another, and orchestrate responses to environmental cues. A crucial function of many of these signaling networks is the generation of complex time-varying signals, such as oscillation or a transient pulse of activity [4, 5, 7, 9]. Understanding and controlling how signaling networks reliably generate these complex dynamical signals is a central goal of systems biology. Perturbation to the appropriate dynamics of biological systems plays a role in many disease states, such as such as cancer [13, 132] and familial advanced sleep phase syndrome (FASPS) [133]. Similarly, dosing schedules can be adjusted to a system's dynamical response to increase efficacy or decrease undesirable side effects [134].

Mathematical modeling is a particularly useful tool for characterizing dynamically varying systems because these dynamics cannot be understood *via* static diagrams of network connectivity. In recent years, models of individual network motifs have helped

to characterize the relationship between topology and dynamics for feedforward and feedback loops [91,135]. Oscillating networks have been the subject of special scrutiny. Theoretical and experimental studies of two major oscillating motifs—delay oscillators (comprising a delayed negative feedback loop) and relaxation or hysteresis oscillators (comprising a fast positive and slower negative feedback loop)—have helped elucidate their properties [26, 77–79, 136].

As biological signaling networks are understood in greater detail, it is becoming clear that they rarely consist of isolated motifs. Rather, networks are often more complex, with seemingly redundant or antagonistic connections [5, 33]. This complexity may arise from two sources. First, newly identified species and interactions often result in additional feedback loops. Detailed models utilizing many combinations of these motifs have been constructed for larger systems [13, 107, 137]. Fortunately, these feedback loops are typically biochemically distinct, with few points of interaction between loops. This separability can be exploited by computationally or experimentally perturbing species involved in individual feedback loops to probe the roles of these loops in system operation [83, 138, 139].

A second source of complexity arises from a growing understanding of the detailed biochemical interactions between previously identified system components. Biological systems consisting of even very few proteins contain a large number of conformational, post-translational modification and multimeric states, all of which may affect reaction rates. In systems regulated at this level, the numbers of distinct chemical species and reactions between them increase combinatorially with the number of states. Multiple modification states, multimeric complexes and protein isoforms are present in a variety of signaling networks [43, 49, 50], but the roles are often unknown. Because of combinatorial interactions between species, such networks can be densely connected, making even the identification of network motifs a challenging task. Furthermore, this detailed biochemical complexity has been shown to have important consequences for system dynamics in theoretical [51] and experimental studies [53]. Many open questions remain in understanding such systems. What are the processes governing dynamics in densely connected biochemical networks? What motifs or combinations

of motifs comprise these networks? What tools can be used to identify these processes?

We address these questions in the context of a model system of detailed biochemical complexity: the *in vitro* circadian clock extracted from *Synechococcus elongatus*. The cyanobacterial circadian clock enables *S. elongatus* to adapt its genetic and metabolic programs to daily changes in the environment and provides a daily rhythm to photosynthetic regulation [58, 59]. In its normal context in the bacterium, the circadian clock involves transcriptional regulation acting through multiple feedback loops [60]. Strikingly, Kondo *et al.* have shown that oscillation with a period of approximately 24 hours can be reconstituted *in vitro* with only three proteins: KaiA, KaiB, and KaiC [61]. As this *in vitro* reaction lacks other proteins and nucleic acids, no additional transcriptional or post-translational pathways can be implicated in generating these dynamics. During oscillation, the three Kai proteins traverse a variety of modification states and multimeric complexes, making this system ideal to study how these biochemical details contribute to complex system dynamics [49, 65, 140].

Several mathematical models have been proposed for this network [65–67]. Some of these models derive oscillation from monomer exchange between KaiC hexamers [66, 67], a phenomenon still under experimental investigation [49, 67], while others rely on the dynamics of KaiC phosphorylation [65]. We used the model of Rust *et al.* as the basis for our work [65]. This model, parameterized to experimental KaiC phosphorylation and dephosphorylation data under a variety of conditions, concisely treats the dynamics of the system using four ordinary differential equations (ODEs), one algebraic equation, and thirteen rate parameters. It quantitatively describes the dynamics of two phosphorylation sites of KaiC, serine 431 (S431) and threonine 432 (T432). The rate of transition between four phosphorylation states (U-KaiC, unphosphorylated KaiC; S-KaiC, S431-phosphorylated KaiC; T-KaiC, T432-phosphorylated KaiC; and D-KaiC, doubly S431- and T432-phosphorylated KaiC) is dependent on interaction with KaiA, which is in turn regulated by S-KaiC and KaiB, providing the basis for feedback in the model. By examining the relative magnitude of the fitted model's reaction rates, Rust *et al.* conjectured that a subset of the reactions of the

network—the phosphorylation of KaiC in the presence of free KaiA first at T432 followed by S431, followed by the dephosphorylation of KaiC first at T432 and then at S431—constitute the network's core oscillatory pathway.

In this work, we sought to better understand how progression through different phosphorylation states controls dynamics in the oscillating network. Sensitivity analysis provides a powerful technique for defining how individual reactions contribute to specific features of a dynamical trajectory. For oscillatory systems, sensitivity analysis probes how small variations in reaction rate constants affect features of the periodic orbit, such as oscillation period, amplitude for individual species, and the difference in oscillation phase between two species' extrema. Oscillator sensitivity analysis has been successfully applied to other biological models [83, 138, 141, 142], where it has been observed that reactions of high sensitivity can help to identify particularly crucial pathway steps. Applying oscillator sensitivity analysis necessitates the careful consideration of mathematical details. ODE models of biological oscillators are typically limit cycles, where the system will settle to the same periodic orbit regardless of the initial conditions from which simulation is initiated. This is not the case for a biochemical oscillator of the type described here, where certain combinations of species (in this case, the total concentration of KaiA, KaiB, and KaiC) cannot be changed through reactions within the system. Such a system can be characterized as an intermediate oscillator, in which initiating simulation with different total protein concentrations will lead to distinct periodic orbits, but varying initial conditions such that total protein concentrations are kept constant will always converge to the same orbit [88]. We use a recently developed approach to oscillator sensitivity analysis that extracts the appropriate sensitivities in the subspace preserving these conservation laws [88].

Here, we recast the Rust *et al.* circadian clock model into a purely mass action description and applied oscillator sensitivity analysis to it. We find the sensitivity profiles obtained through these methods identify self-consistent processes in the network. By comparing these profiles, we found that certain features of the dynamics co-vary regardless of which parameters are tuned. To further characterize the feed-

back connections that control dynamics, we exhaustively enumerated all subnetworks and queried them for the ability to undergo oscillation. Using this approach, we recovered two feedback motifs, a delayed negative feedback loop (similar to the topology identified by Rust *et al.*) and a coupled negative and positive feedback loop that are independently capable of oscillation with similar periods. Although the core reactions in the second motif have smaller kinetic parameters than those in the first, these loops are coupled and each contributes to the dynamics of the full network. Finally, we show that the full network combining the two motifs demonstrates the ability to tune phase but preserve oscillation period, a crucial characteristic for a circadian system.

5.2 A mass action model of the circadian clock

In the *in vitro* circadian oscillator, the biochemical reactions between the three Kai proteins represent processes governed by mass action chemical kinetics. This formalism treats the bimolecular association of chemical species into complexes (such as KaiA–KaiC or KaiB–KaiC binding), and the unimolecular dissociation and modification of both individual species and their complexes. We converted the model of Rust *et al.* to a mass action kinetic model representing the association of KaiA, KaiB, and KaiC as well as the phosphorylation of KaiC, allowing us to use modeling and analysis tools that we had previously developed for the efficient analysis of mass action models.

The original model of Rust *et al.* incorporates two types of nonlinear reactions: a saturating expression for the fraction of KaiC in a phosphorylation-competent state, and an algebraic relation for the concentration of available KaiA. The precise mathematical form of these nonlinearities can be realized by mass action processes under the assumption that KaiA–KaiC interaction occurs fast relative to KaiC phosphoconversion. We implemented a mass action reaction mechanism and derived the limits in which it reduces to the original model's nonlinear relations to ensure that our mass action model's dynamics were identical to those of the original model (see Appendix C.1). We found that the simplest mass action model of KaiC activation, in

which the KaiA-bound form of KaiC undergoes phosphorylation at an enhanced rate, cannot oscillate because the concentration of free KaiA is depleted through binding to KaiC. Instead, we implemented an alternative model, in which KaiA binds to and modifies KaiC in a manner that persists after KaiA dissociation. In this model, only the modified form of KaiC (denoted KaiC*) is capable of autophosphorylation. This modified KaiC* state could represent a conformational change of KaiC induced by KaiA binding, or the *trans* activation of other KaiC monomers in the experimentally observed but not directly modeled hexameric KaiC complex upon KaiA binding.

The dynamics of our mass action model are identical to those of the original model in the limit of infinitely fast reactions. To ensure that our mass action model accurately reflected the original model's dynamics for finite rate constant values, we fit our model to the output of the original model simultaneously across all conditions under which it had been fitted [65] (for fitting procedure, see Methods). Simulation of the fitted model generates trajectories indistinguishable from the original model (Figure 5-1A), and all of the parameters shared between the original and fitted models matched to within 1% (Table C.1). Additionally, the parameters governing KaiA–KaiC binding and modification led to an effective $K_{1/2}$ value close to that of the original model (Table C.1). The simulations oscillate regularly with a period of 21.5 hours. The first period shows somewhat different dynamics than the remainder as the simulation relaxes onto a periodic orbit from the initial conditions of Rust *et al.*

Our mass action model consists of 75 state variables representing the three Kai proteins, their modification states, and their complexes. Simulations of this signaling network demonstrate that many of these states only attain very low concentrations during the oscillatory cycle. This reduction arises primarily from the rules assumed to govern KaiC–KaiA–KaiB association: S-KaiC associates with KaiB and inactivated KaiA dimers tightly and quickly, and all other KaiC phosphoforms undergo fast dissociation from inactivated KaiA and KaiB. Illustrating this point, the concentration of S-KaiC–KaiB complexes is $> 10^6$ higher than the concentration free S-KaiC under oscillating conditions (where KaiB is present in excess). Also, S-KaiC–KaiB complexes are always bound to inactive dimers of KaiA (S_i, Figure 5-1C), except

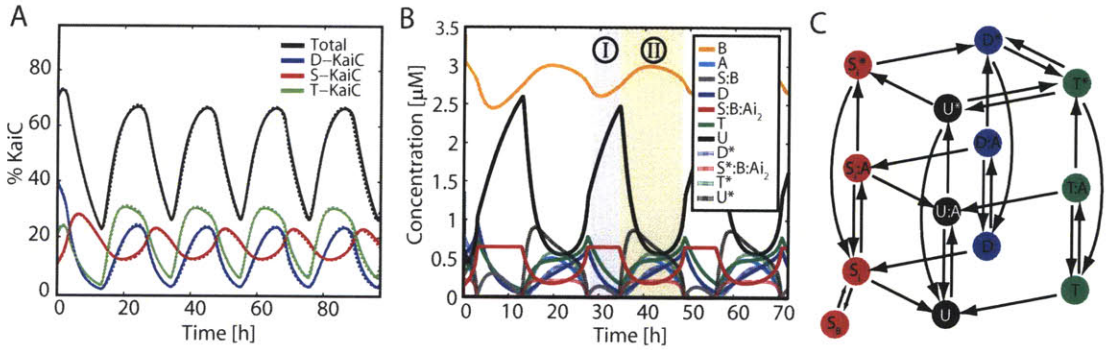


Figure 5-1: **Identifying processes in a model of the *in vitro* circadian clock.** (A) Trajectories from a mass action circadian clock model, fit to the dynamics of the model of Rust *et al.* [65]. Total concentrations of D-, S-, T- and total phosphorylated KaiC are shown from the original model (dashed lines) and fitted mass-action model (solid lines). (B) Simulation of all species in the model that accumulate to ≥ 0.01 μM. Major isoforms of U-, S-, T- and D-KaiC are represented in black, red, green and blue respectively. KaiC, KaiC–KaiA complexes and KaiC* are shown in solid, dashed, and dotted lines respectively. Other species shown are free KaiB (orange) and KaiA (teal), as well as S-KaiC–KaiB complexes (gray). Shaded regions indicate partitioning of this trajectory into a dephosphorylation-dominated regime lacking free KaiA (I) and a phosphorylation-dominated regime with KaiA present(II). (C) Diagram of the reactions in the mass-action model. KaiC phosphoconversion is depicted by reactions within horizontal planes, while KaiA binding, dissociation and catalytic activation of each KaiC phosphoform is depicted by the vertical arrows. S_i denotes S-KaiC–KaiB–$KaiA_2$ complexes, while S_B denotes S-KaiC–KaiB complexes.

when all free KaiA is depleted by this binding and excess S-KaiC–KaiB complexes are present (S_B, Figure 5-1C). During simulation, only eleven species accumulate to appreciable concentration: free KaiA, free KaiB, and nine phospho-KaiC states and complexes (Figure 5-1C). The detailed dynamics of these eleven species are shown in Figure 5-1B. The model exhibits a biphasic response consisting of a dephosphorylation regime (Regime I marked in the figure) and a phosphorylation regime (II). Regime I is characterized by the sequestration of KaiA to S-KaiC (S:B:Ai2, where Ai2 refers to the two sequestered KaiA molecules; red trace), the accumulation of excess S-KaiC–KaiB complexes (S:B; gray curve), and the gradual conversion of D-KaiC and T-KaiC to U-KaiC. In Regime II, free KaiA (A; teal curve) generates $KaiC^*$ species, which leads to the accumulation of phosphorylated KaiC.

By focusing on the subset of species that accumulate during cycling, we developed a simplified visual representation for the reactions between these species (Figure 5-1C). This visual representation includes transitions between the nine populated KaiC and $KaiC^*$ states as well as their respective KaiC–KaiA intermediates, and provides a useful framework for visualizing the effects of individual reactions on dynamics in what follows. The diagram consists of three planes: free KaiC (lower plane), KaiC–KaiA complexes (middle plane) and modified $KaiC^*$ (upper plane). Reactions in the upper plane proceed with kinetics that favor phosphorylation, while reactions in the lower two planes proceed by dephosphorylation kinetics governed by identical rate constants. Reactions between planes model the association and dissociation of KaiA, and the KaiA-mediated modification of KaiC. Finally, in addition to being activated by KaiA, S-KaiC can also tightly bind and sequester two KaiA molecules, and is present in S-KaiC–KaiB–$KaiA_2$ complexes (indicated as S_i in Figure 5-1C). During the dephosphorylation phase, S-KaiC–KaiB complexes (indicated as S_b in Figure 5-1C) accumulate when no additional KaiA is present to be sequestered. We verified that this representation is valid because for each of the subsequent analyses performed by noting that the largest effect of reactions excluded from this diagram is at least 5 orders of magnitude lower than the largest effect of those included.

Thus, we demonstrated that our mass action model reproduces the dynamical re-

sponse of the original model under all conditions used in the original model's fit, and that although it introduces a large number of species, it can be compactly represented as transitions between only eleven distinct biochemical species. Taken together, our initial results indicate the utility of further investigation to elucidate the mechanisms generating dynamics using our mass action model.

5.3 Diverse reactions modulate dynamical features in a concerted fashion

We next set out to probe the importance of individual reactions in setting the dynamical response of the network. To accomplish this goal, we performed oscillator sensitivity analysis by computing the variation in the oscillation period, angular phase, and amplitude due to changes in the value of each of the model's "unlumped" reaction rate constants, essentially using each rate constant as a handle to locally perturb each reaction individually (see Methods). This type of analysis can provide insight into the mechanisms responsible for determining specific features of the dynamics. A single important reaction can indicate a rate limiting step, or multiple sensitive reactions can be clustered close to one another, indicating important multiple contributions to a single, larger scale process [83]. To visualize the resulting sensitivity profiles, we mapped the sign and magnitude of normalized period, angular phase, and amplitude sensitivities of each reaction onto the arrows connecting pairs of species (Figure 5-2A–C).

Figure 5-2A shows the normalized period sensitivities obtained by this method. The individual normalized sensitivities vary by over five orders of magnitude, indicating that some reactions are significantly more important than others in setting the period of oscillation in the network. By inspection of the sensitivity magnitudes (arrow weights) on this diagram, we find that there are multiple reactions of comparable sensitivity, and that these reactions are distributed throughout the network; control over the period of oscillation appears distributed across reactions involving all four

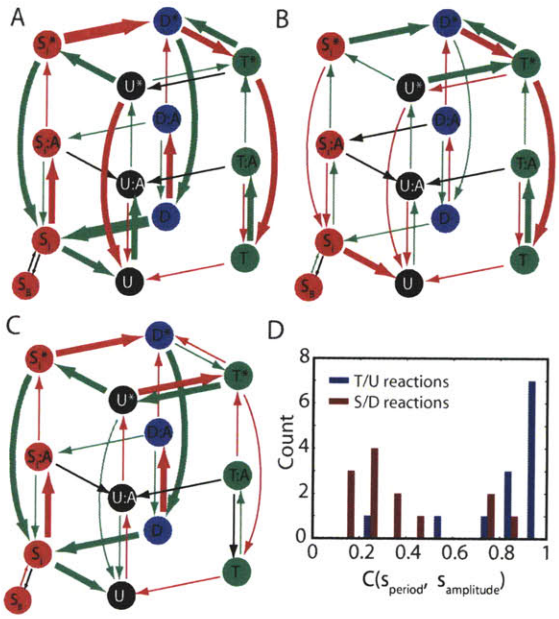


Figure 5-2: **Oscillator sensitivity analysis for the mass action model.** (A–C) Diagrams showing normalized sensitivities of (A) oscillation period, (B) angular phase of Regime II, taken as the phase difference between U-KaiC and T-KaiC maxima, and (C) relative amplitude of total U-KaiC. Arrow thicknesses correspond to $0.01 \leq |\bar{s}| < 0.1$, $0.1 \leq |\bar{s}| < 1.0$, and $|\bar{s}| \geq 1.0$, where $|\bar{s}|$ is the absolute value of the appropriate normalized sensitivity. Arrows are red and green for positive and negative sensitivities, respectively. Black arrows represent sensitivities $|\bar{s}| < 0.01$. (D) Histogram of the magnitude of correlation between period and amplitude sensitivity profiles, considering reactions of S-KaiC and D-KaiC phosphoforms (back vertical plane of Figure 5-1C; brown bars) or reactions of T-KaiC and U-KaiC phosphoforms (front vertical plane of Figure 5-1C; blue bars). Each count represents correlation with the relative amplitude sensitivity of a single species.

KaiC phosphoforms. This distributed sensitivity profile stands in contrast to prior studies in which similar analyses have been performed [83], and arise regardless of which dynamical feature is considered (see Figures C-1 and C-2). We suggest that this distributed sensitivity may be due to the mass conserved nature of this network. In this system any increase in the concentration of one species necessitates depletion of others because total protein concentrations are fixed; in a transcriptional network, an increase in protein level does not deplete its transcript, potentially insulating these networks from such effects. Indeed, the contrasting prior studies did involve transcriptional feedback [83], consistent with this suggestion. Nevertheless, other differences between networks may also contribute, including the greater degree of connectivity in the current network.

An important feature of the sensitivity analysis profile of Figure 5-2A is the grouping of sensitivities into self-consistent processes. Here, we define a 'process' as a set of reactions exhibiting a consistent relationship between species concentration and dynamics. For instance, increasing rates leading to generation of S-KaiC* from S-KaiC—whether in the association of S-KaiC with KaiA or the release of S-KaiC*—always increase the oscillation period, whereas increasing rates towards S-KaiC production—whether in the dissociation of S-KaiC–KaiA or in the direct S-KaiC*→S-KaiC reactivation reaction—have the opposite effect. This consistency finding generalizes to all KaiC↔KaiC* reactions, although the sign is flipped for the U-KaiC and T-KaiC branches. From this framework, two larger self-consistent processes emerge: all reactions from S-KaiC to D-KaiC* in the back vertical plane are of coherent sign, as are all reactions in the front vertical plane leading from U-KaiC to T-KaiC*. Notably, although these processes both lead to an increase in T-432 phosphorylation state, increasing rates from S-KaiC→D-KaiC* lead to a longer oscillation period, while biasing the network towards U-KaiC→T-KaiC* shortens the period.

Circadian oscillators might exhibit differential control over various features of their oscillating dynamics. For instance, modulating oscillation period should have a deleterious effect due to the constancy of the 24-hour day, whereas the fraction of the cycle (*i.e.*, phase) representing day and night might be more tunable to account for

evolution of adjustable strategies and seasonal variation in daylight. We set out to determine whether the mass action *in vitro* circadian clock model exhibited differential control over different dynamical features by comparing the period sensitivity profile obtained above to those of individual species' amplitudes and phases. Figure 5-2B shows the angular phase sensitivity profile for the fraction of time spent in the phosphorylation regime (Regime II). Such an angular phase sensitivity only computes the fraction of the total cycle spent in the phosphorylation regime; thus, a low sensitivity value indicates that the time spent in the phosphorylation regime scale proportionally with the total period. Surprisingly, many reactions important for setting the period (such as the D-KaiC→S-KaiC arrow representing the largest magnitude period sensitivity) have a minimal effect on the relative duration of the phosphorylation phase. We find that this analysis identifies a clustered set of high-sensitivity reactions governing phosphorylation through T-KaiC*, suggesting that this phosphoform represents a point of differential control over period and phase.

We next investigated whether these conclusions generalized to other types of dynamical sensitivities [84]. Figure 5-2C shows an example of a typical amplitude sensitivity profile for the total concentration of U-KaiC. Compared to the period sensitivity profile (Figure 5-2A), we observed a high degree of similarity between reaction sensitivities in the S-KaiC/D-KaiC plane (back vertical plane of Figure 5-2C), whereas reactions comprising the U-KaiC/T-KaiC plane (front vertical plane) varied more widely. This relationship was preserved when the amplitude of other phosphoforms was considered (Figure C-1). To quantify this difference, we computed the correlation across all parameters between amplitude and period sensitivity profiles for those reactions in the front vertical plane and back vertical plane (Figure 5-2D). A high correlation indicates that changes made to one dynamical feature leads to proportional changes in another feature; such features would be proportionally controlled by these parameters. We find that for those reactions in the back vertical plane, amplitude and period are modulated together ($\rho > 0.9$ for 7/13 species), whereas reactions in the front vertical plane modulate these dynamical features differentially ($\rho < 0.3$ for 7/13 species). Taken together, our sensitivity results indicate that reactions modulating

S432-phosphorylated KaiC levels (S-KaiC and D-KaiC phosphoforms) broadly control shared features of oscillation, whereas reactions modulating T431-phosphorylated KaiC can fine-tune specific features of the dynamical response.

The *in vitro* circadian clock exhibits a biphasic response characterized by a dephosphorylation (KaiA-inhibited) regime and a phosphorylation (KaiA-activated) regime (Figure 5-1A). We reasoned that while sensitivity analysis identifies important biochemical reactions for modulating dynamics over a full trajectory, the contribution from specific reactions during portions of that trajectory might vary. To separate the effects of parameters in the phosphorylation and dephosphorylation regimes, we computationally added a fictitious enzyme cascade, $E \rightarrow X$, that is activated by KaiA and catalyzes the conversion of all KaiC species to a corresponding modified form that behaves with identical kinetics, but distinct parameters. We verified that this procedure for isolating sensitivities in each regime does not lead to any changes in the overall model trajectories (Figure 5-3A), and that the sum of the sensitivities in the two regimes are equal to the overall period sensitivity (data not shown).

The resulting period sensitivities are shown in Figure 5-3B–C. In the dephosphorylation regime (Figure 5-3B), only a single reaction shows a sensitivity comparable to the high magnitude sensitivities in the full model: the conversion of S-KaiC to U-KaiC. The sign of this sensitivity indicates that increasing the dephosphorylation rate of S-KaiC leads to a decrease in the oscillation period, as would be expected from the rationale that phosphorylation resumes when S-KaiC levels fall below the concentration required to fully inhibit KaiA. That this is the only reaction with appreciable sensitivity suggests the dephosphorylation of S-KaiC functions as a classical rate-limiting step governing transition from dephosphorylation to phosphorylation. This interpretation is supported by its threefold lower magnitude than the preceding D-KaiC*$\rightarrow$S-KaiC* reaction.

The dephosphorylation regime's sensitivity profile contrasts starkly with that obtained from the phosphorylation regime (Figure 5-3C), which much more closely resembles the full model's sensitivity profile. In this decomposition, an additional feature emerges. All reactions leading towards the production of S-KaiC in the phos-

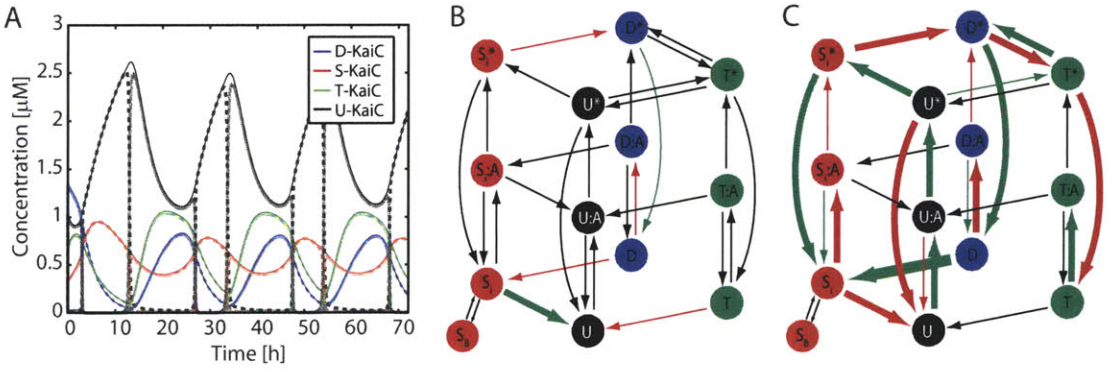


Figure 5-3: Partitioning trajectories into phosphorylation and dephosphorylation regimes. (A) Computational partitioning of the full model through introduction of a fictional enzyme whose activity depends on the presence of KaiA. Species active during Regime I and II are shown in dotted and dashed lines, respectively, with total U-, S-, T- and D-KaiC trajectories shown in solid lines. (B–C) Normalized period sensitivities for each parameter's contribution during (B) regime I (dephosphorylation) and (C) regime II (phosphorylation) only. Arrows are as described for Figure 5-1A.

phorylation regime have negative sensitivities (green arrows) and are opposed by reactions with positive sensitivities (red arrows), indicating that increasing conversion to S-KaiC decreases the period of oscillation. This suggests that progression through the phosphorylation regime is dominated by the delay in producing S-KaiC. We note that many of these sensitivities have comparably high magnitude, indicating that S-KaiC formation may rely on multiple reactions rather than a single rate-limiting step. The S-KaiC $\rightarrow$ U-KaiC conversion is high in magnitude but opposite in sign in the two regimes, supporting the hypothesis that reactions govern specific processes at different times during a trajectory.

5.4 Minimal oscillating subnetworks identify feedback loops required for oscillation

Oscillator sensitivity analysis applied to the *in vitro* circadian clock identifies modes along which dynamics co-vary, and suggests sets of reactions representing processes with self-consistent effects on system dynamics. While sensitivity analysis has been useful in other studies to identify feedback loops governing dynamics, the high magnitude sensitivities in the circadian clock are distributed throughout the network (Figure 5-2A–C). We reasoned that enumerating subsets of reactions that still undergo oscillation would be a stringent test to identify reactions important for generating this dynamical response.

Because the transformations between the dominant KaiC isoforms are characterized by twelve phosphoconversion reactions and eight KaiC$\leftrightarrow$KaiC* conversion processes, we were able to exhaustively enumerate all models in which combinations of these reactions were removed. (For larger networks in which enumeration would not be feasible, other procedures such as a greedy search for reactions may be useful). In this procedure, reactions were removed by setting their corresponding rate constants to zero; other parameter values were not allowed to vary. Figure 5-4A shows the incidence of oscillating subnetworks as all combinations of reactions are removed. Rows

represent reactions between planes, and columns represent reactions within planes connecting the dominant KaiC isoforms (Figure 5-1C). We found that oscillation is preserved in $9,454/1,048,576$ subnetworks, 0.9% of the possible configurations. The horizontal and vertical banding pattern of Figure 5-4A indicates that the inclusion of certain reactions can give rise to oscillation in a variety of configurations, indicating that many of the oscillating subnetworks identified by this approach share components of core oscillator modules.

We further probed this observation using a computational procedure to identify these oscillating modules, which we termed minimal oscillating subnetworks (MOSs). This was performed iteratively by (1) finding the oscillating subnetwork with the fewest number of reactions and (2) removing all networks containing this set of reactions from the remaining pool of oscillating subnetworks. This procedure identified two sets of reactions giving rise to oscillation (Figure 5-4B–C). These two MOSs do not share any closed loops between the different KaiC isoforms, yet simulation of both subnetworks generates oscillating trajectories with the same ordering of S-, U- and D-KaiC peaks, suggesting that at least two independent sets of feedback loops are capable of driving oscillation in this network.

Although the dynamics generated by these two MOSs are similar, the subnetworks represent very different mechanisms of oscillation. The first subnetwork consists of a closed loop coupling dephosphorylation from D-KaiC through S-KaiC to U-KaiC (bottom plane, Figure 5-4B) with phosphorylation progressing through the T-KaiC* intermediate (top plane, Figure 5-4B), the kinetically favored intermediate because the U-KaiC*→T-KaiC* reaction is an order of magnitude faster than the U-KaiC*→S-KaiC* reaction. This reaction scheme, comprising a chain of interconversion steps arranged in an a cycle, is very similar to that described by Rust *et al.* as a likely mechanism for oscillation in the full network [65]. This subnetwork does not include either positive feedback loop between KaiA and S-KaiC, and only includes one of the full network's two negative feedback loops. By inspection of its topology, we reasoned that the mechanism used by this subnetwork to generate oscillations is a single negative feedback loop where phosphoconversion reactions represent multiple

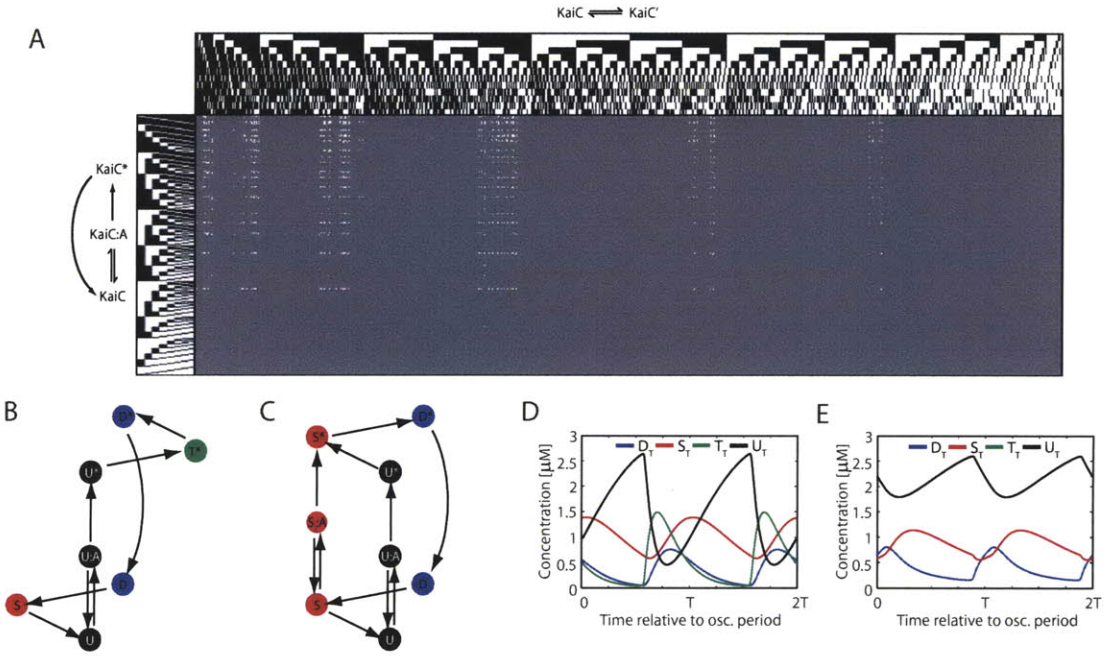


Figure 5-4: **Identification of oscillating subnetworks.** (A) Exhaustive enumeration of oscillation status for knockout models. Rows represent different combinations of the eight reactions, while columns represent the twelve KaiC phosphoconversion reactions for KaiC, KaiC–KaiA and KaiC* isoforms. Presence or absence of a reaction is indicated by black or white, respectively, in the margin area. Oscillating subnetworks are shown in white. (B–C) Diagrams indicating the minimal oscillating subnetworks (MOSs) of the mass action model. (D–E) Simulations over two full periods of oscillation of representative MOSs from (B–C). The periods of oscillation in the two MOSs are 25.7 and 30.7 h, respectively.

intermediate delay steps.

The second minimal subnetwork utilizes a strikingly different topology to generate oscillation (Figure 5-4C). Like the first subnetwork, it includes a delayed negative feedback loop operating through phases of KaiC dephosphorylation and phosphorylation gated by the interaction between S-KaiC and KaiA. Unlike the first subnetwork, the phosphorylation phase progresses through S-KaiC* rather than T-KaiC*. In addition to this negative feedback loop, the second MOS requires KaiA-mediated activation of S-KaiC, forming a strong positive feedback loop. The action of this positive feedback loop can be intuited by considering the beginning of the phosphorylation phase, when levels of S-KaiC decrease enough for a small amount of KaiA to escape inhibition. Some of this free KaiA activates S-KaiC to S-KaiC*, which reacts quickly to form D-KaiC*, freeing additional molecules of KaiA and closing the positive feedback loop. This positive feedback loop acts quickly, as modification by KaiA is assumed to be nearly instantaneous, and interconversion between S-KaiC and D-KaiC are among the fastest phosphoconversion reactions in the system (Table C.1). Oscillators utilizing a fast positive and slow negative feedback loop are often termed relaxation or hysteresis oscillators [26], and many oscillating phenomena are thought to be driven by such a topology [77].

We identified both a hysteresis and delay oscillator in the *in vitro* circadian clock network. However, in the full biological system, both oscillators are present together, and they share many species and reactions. This mode of organization is not specific to this system; rather, many biological networks include seemingly redundant connections [14]. We set out to identify the relative contribution from both minimal subnetworks to the overall system dynamics. For metabolic networks and other mass conserved systems, analysis of flux through different reaction pathways has long been appreciated as a useful tool for gauging the relative utilization of these pathways [143, 144]. We applied a flux analysis scheme to our full model to assess the contribution of the two oscillator motifs described above. We tracked the total concentration flux over time through each of the twelve KaiC interconversion connections, with the sign convention that positive flux flows towards increasing phosphorylation

and modification state (Figure 5-5A). The resulting fluxes for all twelve reactions are shown in Figure 5-5C–E. The sum of all fluxes into each node leads to the evolution of total protein concentrations shown in Figure 5-5B.

This procedure provides some insight into the relative importance of different reactions along the periodic orbit. Initially the flux is dominated by dephosphorylation, leading to low concentrations of active KaiC* species (Figure 5-5C). During this phase, contributions from the two oscillating subnetworks are indistinguishable, as these subnetworks share the reactions through which dephosphorylation proceeds. These subnetworks diverge in the subsequent phosphorylation phase, allowing their influence to be tabulated separately. To measure fairly the contribution through each subnetwork, we use the minimum flux along a pathway as a measure of the overall flux through that pathway:

$$J_{T^*} = \min\{J_{U^*\leftrightarrow T^*}, J_{T^*\leftrightarrow D^*}\} \tag{5.1}$$

$$J_{S^*} = \min\{J_{U^*\leftrightarrow S^*}, J_{S^*\leftrightarrow D^*}\} \tag{5.2}$$

where J_{T^*} and J_{S^*} are the overall fluxes through each of the two subnetworks, and $J_{C\leftrightarrow C'}$ is the flux through a specific KaiC conversion reaction. This "flux-limiting step" approach accounts for eddies within sets of reversible reactions that would otherwise overestimate the overall flux passing through these reactions.

Figure 5-5C shows that while the U-KaiC$^*\leftrightarrow$T-KaiC* conversion is fast (green trace; approximately 2.5 μM total KaiC flows through this branch), far less subsequently flows from T-KaiC* to D-KaiC* (teal trace), indicating that the overall flux through the first subnetwork is strikingly lower than would be assumed from considering the kinetically favored T-KaiC$^*\rightarrow$D-KaiC* reaction alone. This difference can be understood by noting the strength of the opposing dephosphorylation reactions that limit the amount of D-KaiC* formed by this branch. Notably, the total flux through the T-KaiC$^*\leftrightarrow$D-KaiC* pathway follows a trajectory very close to that of U-KaiC$^*\rightarrow$S-KaiC* (blue trace), the first reaction along the negative feedback loop of the second subnetwork. (The subsequent S-KaiC$^*\rightarrow$D-KaiC* flux includes this sub-

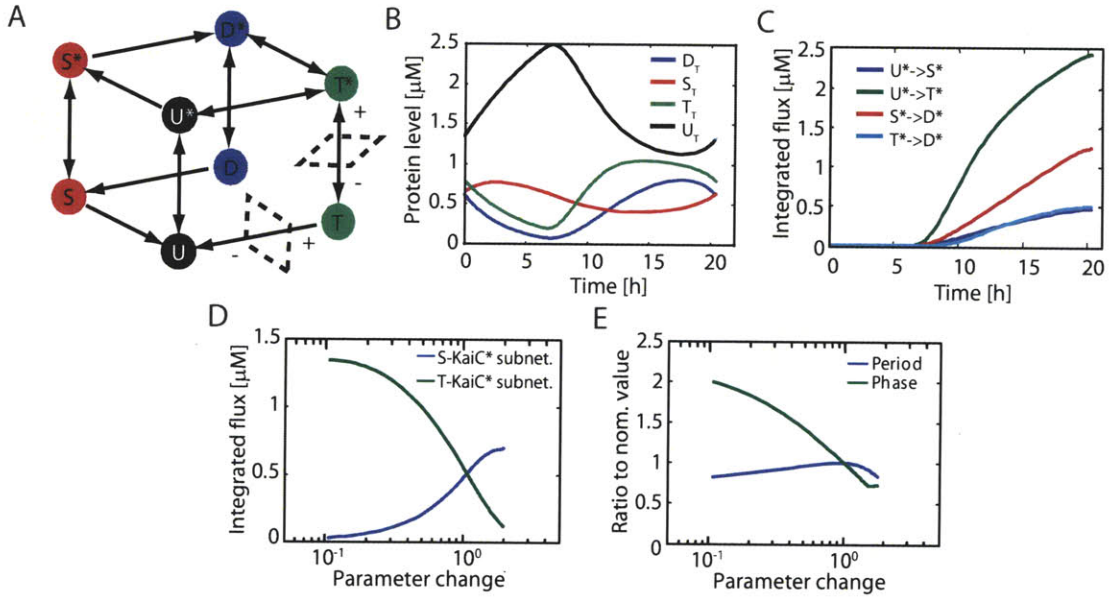


Figure 5-5: Analysis of fluxes in the full network. (A) Diagram of flux analysis procedure. The flux passing between nodes along each branch was tracked during one period of oscillation from the full model. Sign conventions are taken such that the flux is considered positive when flowing towards increasing phosphorylation state within each horizontal plane, or towards autokinase-enhanced KaiC between planes. (B) Simulation of the full model, beginning at the onset of Regime I, from which all fluxes are computed. The concentrations of total D-, S-, T- and U-KaiC are plotted over time. (C) Integrated flux as a function of time between all KaiC* phosphoforms. As in (B), the x-axis shows time elapsed since the beginning of Regime I. See Figure C-4 for fluxes through all other branches. (D) Integrated flux after one full period through the T* (blue trace) and S* (green trace) subnetworks as k^*_{usU} and k^*_{tdT}, the reaction rates associated with these subnetworks, are simultaneously varied. (E) Variation in oscillation period (blue trace) and the length of the dephosphorylation phase (green trace) relative to the initial parameterization, as the reaction rates associated with each subnetwork are varied as in (D).

network's fast positive feedback loop, leading to its larger magnitude). From these data we reason that the flux-limiting reactions through each pathway are balanced in magnitude throughout the phosphorylation phase; thus, the full network utilizes each of its two oscillating submodules equally.

What benefit might be provided by driving dynamics using two balanced, coupled oscillating motifs? Each of the two motifs we identified shares a similar oscillation period, but varies in the timing of activation of the KaiC phosphoforms driving oscillation. We hypothesized that modulating the flux through each motif might provide a key functionality to the circadian clock: a mechanism to tune oscillation phase without a strong effect on period. The cyanobacterial circadian clock has been shown to align KaiC phosphorylation with the day/night cycle: KaiC is hypophosphorylated during daylight hours and hyperphosphorylated at night [145]. To test this mechanism, we varied the reaction rates corresponding to the flux-limiting steps for each minimal network (U-KaiC*→S-KaiC* and T-KaiC*→D-KaiC*), and monitored the fluxes through each subnetwork, J_{T^*} and J_{S^*}. Figure 5-5F shows the total integrated flux through each subnetwork; the full time-dependent fluxes are shown in Figure C-4. The full network continued to support oscillation over an approximately 20-fold variation in these rate constants, over which the flux through each subnetwork could be tuned in a graded fashion. For each resulting parameterization, we measured the period of oscillation and the length of the dephosphorylation phase (Figure 5-5G). We found that varying the utilization of each subnetwork supported phases from 60% to 200% of the initial model, with less than a 20% variation in period.

5.5 Conclusions

In this chapter, we have identified the principles underlying the operation of a complex network, a detailed model of the cyanobacterial circadian clock. Using local sensitivity analysis, we found that the sensitive parameters represent self-consistent processes controlling oscillation period, amplitude and angular phase. We identified portions of the network utilizing a classical 'limiting step' (e.g. progression through

the dephosphorylation-dominated oscillation regime) as well as those exhibiting more distributed control. Large scale perturbations to network structure revealed that two motifs—a delayed negative feedback oscillator and a hysteretic negative-positive feedback oscillator—are contained in the full network. By tabulating fluxes through these two motifs we found that they are each utilized to drive dynamics. Finally, we suggest a mechanism by which these two motifs can provide control over day/night phase, with only a small effect on oscillation period.

To accomplish these goals, we developed two general approaches for mapping dynamical features to the biochemical reactions that control them. The first approach relies on sensitivity analysis to features of oscillation. Sensitivity analysis is typically used to identify individual reactions to which an system output is sensitive. We take a different approach, using sensitivities to identify clusters of important reactions and comparing sensitivity profiles between multiple dynamical features. While sensitivity analysis is an efficient technique for probing the effect of parameter changes on dynamics, it is a local method applied to individual parameters. Large-scale, concerted changes in multiple parameters might have effects different than those expected from this local characterization. Our work does demonstrate that in one case, large-scale perturbation to parameters in both motifs modulate phase and period consistently with the expected sensitivity result (Figure 5-5F; compare to Figure 5-2A–B). This agreement may be system-dependent, and should be verified for other networks to which such methods are applied.

In the second section of this work, we exhaustively enumerated subnetworks capable of generating an oscillatory response. This approach is especially well suited for extracting intuition from densely connected models such as the system studied here, but may also be useful to identify functional motifs in gene regulatory networks. For this procedure, removal of some reactions (*e.g.*, opposing forward and reverse reactions) might be expected to occur in a concerted fashion. To ensure that the MOSs we identify represent biologically meaningful combinations of reactions, we repeated their identification while requiring the simultaneous removal of KaiC→KaiC* and KaiC*→KaiC reactions and of free KaiC and complexed KaiC–KaiA dephosphoryla-

tion reactions. This procedure yielded two MOSs similar to those obtained from the large-scale approach (Figure C-3A–B). Notably, the two subnetworks undergo oscillation with a period of 21.5 h and 21.2 h respectively (Figure C-3C-D), suggesting that their coupling productively contributes to the dynamics of the full network (whose period is 21.5 h).

The approaches described here allow identification of reactions and motifs that control dynamical responses of interest. Here, we chose to focus on the oscillation period, amplitude and phase, characteristics of intuitive importance for a circadian system. However, additional features might be important in other modes of operation, such as this oscillating network's coupling to other transcriptionally-driven feedback loops, or its operation under variation in light and temperature. When applied to these other features, the methods described here might identify further subsets of reactions, and additional network motifs.

As part of this study, we constructed a mass action model with identical dynamics to those of the abstracted model, but that explicitly represents each protein-protein complex and modification state. This 'unpacking' procedure proved crucial for identifying the motifs contributing to oscillation. In the original model, each KaiC phosphoform was represented as a single dynamical variable. In this treatment, even the sign of KaiA's effect on S-KaiC levels—and therefore, the identification of positive or negative feedback loops between S-KaiC and KaiA—will depend on the concentrations of these and other species. Thus, identifying specific positive and negative feedback loops is challenging, as the number and sign of these loops will depend on the network's operating point. Mass action biochemical models typically implement network connections that are monotonic and consistent in sign, which resolves this challenge and may allow for the application of other analytical tools [51, 52].

5.6 Methods

5.6.1 Model construction

We constructed a model based on that described by Rust *et al.* (henceforth referred to as the original model). The model was implemented in MATLAB R2008b and is available upon request (from BT). This model uses mass action chemical kinetics, implementing first- and second-order reactions that represent the association, dissociation, and modification of the three Kai proteins. To facilitate matching the dynamics of the original model, we ensured that in the limit of fast association and dissociation, our mass action model simplified to the exact mathematical form of the original model.

Our model implements three classes of reactions: (1) binding and activation of all KaiC isoforms by KaiA, (2) the binding and sequestration of KaiA and KaiB to S-KaiC, and (3) the phosphoconversion of KaiC. For the third class, we found that the simplest mass action representation of the reactions described by Rust *et al.*, in which KaiC and KaiA reach a fast binding equilibrium and the KaiA-bound form of KaiC undergoes enhanced autophosphorylation, did not produce oscillation. Furthermore, using optimization to vary the model's rate constants to match the oscillatory, phosphorylation, and dephosphorylation data of ref. 65 did not lead to a close fit of one model to the other (unpublished data). We found that this discrepancy was due to the low total concentration of KaiA (1.3 μM) relative to that of KaiC (3.4 μM), implying that the concentration of free KaiA would be depleted upon binding, an effect not treated in the original model's binding equilibrium expression. We instead implemented a three-state model in which KaiA catalyzes the modification of KaiC after binding, and only this modified KaiC (denoted KaiC*) has an enhanced rate of phosphorylation. The KaiC* state could approximate a persistent conformational change undergone by KaiC after KaiA binding [145], or the effect of KaiA binding on neighboring, unbound subunits of a KaiC hexamer [62]. Our three-state model assumes that the phosphoconversion rates for unbound KaiC and KaiA–KaiC complexes are identical, and that KaiC* becomes unmodified with first-order kinetics.

With these assumptions, the model precisely recovers the quasi-steady state limit modeled in ref. 65 (see Appendix C.1).

Our mass action model consists of 75 state variables (concentrations), 349 reactions, 18 distinct reaction rates (which when unlumped correspond to 349 rate constants) [83], and six non-zero initial concentrations. The number of parameters is significantly smaller than the number of reactions because many species (such as free KaiC and KaiC–KaiA complexes) are assumed to undergo certain reactions with identical kinetics. To ensure that the mass action model matched the dynamics of the original model, we fitted the mass action model to the U-, S-, T- and D-KaiC trajectories generated by the original model in response to the three experiments used to obtain their fit, as well as its full oscillatory dynamics. For this fitting procedure, we performed local optimization using FMINCON configured to use Quasi-Newton [146] with BFGS [147, 148] in the MATLAB Optimization Toolbox Version 4.2. We used parameters obtained by Rust *et al.* as initial guesses. For parameters not present in the original model, representing the association, dissociation, and catalysis reactions between KaiB, KaiA, and KaiC species, we chose values fast compared to those of KaiC phosphoconversion. We assessed goodness of fit using the integrated squared error objective function

$$G(\mathbf{p}) = \int_0^{t_F} ||\mathbf{y}_{\text{model}}(t, \mathbf{p}) - \mathbf{y}_{\text{data}}(t)||^2 \, dt \tag{5.3}$$

where $\mathbf{y}_{\text{model}}(t, \mathbf{p})$ is a vector of the mass action model's concentration of the total quantities of U-, S-, T- and D-KaiC, $\mathbf{y}_{\text{data}}(t)$ is the corresponding concentration of these KaiC isoforms in the original model, and t_F is taken to be 72 h, roughly 3 oscillation cycles. The fitted model matched the original model very closely (Figure 5-1A); all fitted rate constants were within 1% of those in the original model (Table C.1).

Computing oscillator sensitivities for models with mass conservation

We computed the sensitivities with respect to each parameter of the period, the amplitude of oscillation of each species, and the relative phase of the maximum concentration of each species. In order to compute these sensitivities, an initial condition on the periodic orbit and the period of oscillation must first be obtained. We developed an approach for models of the form

$$\dot{\mathbf{x}}(t,\mathbf{p}) = \mathbf{f}(\mathbf{x}(t,\mathbf{p}),\mathbf{p}) \tag{5.4}$$

$$\mathbf{x}(0,\mathbf{p}) = \mathbf{x}_0, \tag{5.5}$$

where $\mathbf{x}$ is the $n \times 1$ vector of state variables, $\mathbf{p}$ is the vector of parameters and $\mathbf{f}(\mathbf{x},\mathbf{p})$ describes the time evolution of the trajectory $\mathbf{x}(t,\mathbf{p})$. For such systems, the state transition function $\mathbf{F}(\tau,\mathbf{x}_0,\mathbf{p})$ is defined as

$$\mathbf{x}(\tau,\mathbf{p}) = \mathbf{F}(\tau,\mathbf{x}_0,\mathbf{p}) \tag{5.6}$$

and describes how Eqs. (5.4) and (5.5) map an initial condition $\mathbf{x}_0$ and a time interval τ to $\mathbf{x}$ at time τ. Because our system is periodic, we implemented a fixed-timestep trapezoidal rule integrator to integrate over one period [124]. Due to the cancellation of errors in Fourier terms, this method has the advantage of spectral convergence when applied over one period. This formulation also allows the efficient integration of the corresponding full sensitivity system,

$$\frac{\partial \dot{\mathbf{x}}}{\partial \mathbf{p}}(t,\mathbf{p}) = \frac{\partial \mathbf{f}}{\partial \mathbf{x}}\frac{\partial \mathbf{x}}{\partial \mathbf{p}} + \frac{\partial \mathbf{f}}{\partial \mathbf{p}} \tag{5.7}$$

$$\frac{\partial \mathbf{x}}{\partial \mathbf{p}}(0,\mathbf{p}) = \mathbf{0}, \tag{5.8}$$

at the same points used in the integration of the forward system. In Eqs. (5.7) and (5.8), $\frac{\partial \mathbf{x}}{\partial \mathbf{p}}(t,\mathbf{p})$ denotes the sensitivity of the concentration of state variables with respect to all parameters at time t.

For models of this type, one can use Newton's method iteratively solving for T and $\mathbf{x}_0$ to satisfy the equation

$$\mathbf{x}_0(\mathbf{p}) - \mathbf{F}(T(\mathbf{p}), \mathbf{x}_0(\mathbf{p}), \mathbf{p}) = 0, \tag{5.9}$$

where T is the period of oscillation, so Eq. (5.9) expresses the periodicity constraint on the trajectory of all species relative to an initial condition. The initial condition $\mathbf{x}_0$ becomes a function of parameters $\mathbf{p}$ because it must be on the limit cycle to initiate pure oscillation and the location and shape of the limit cycle in phase space depends on the parameters. For any oscillating system there are infinitely many initial conditions that satisfy Eq. (5.9) because any point on the periodic orbit will satisfy this equation; one must add an additional constraint (known as a phase locking condition) to pick out an isolated solution. We choose to pin the solution to a maximum of the concentration of the i^{th} species by defining a phase locking function $\phi(\mathbf{x})$ and fixing its value to pick an isolated extremum (either a maximum or minimum) on the periodic orbit:

$$\phi : \mathbf{x} \to f_i(\mathbf{x}, \mathbf{p}) \quad \phi(\mathbf{x}) = 0. \tag{5.10}$$

For the mass action model considered here, a second indeterminacy still leads to a continuum of solutions to this problem. The oscillation period and initial condition depend continuously not only on the parameters but also on the total concentrations of KaiA, KaiB, and KaiC with which the system is initialized; these conservation constraints must be included to obtain the proper solution to Eq. (5.9) [149]. We formulate these conservation constraints as

$$\mathbf{C}^T\mathbf{x}(t, \mathbf{p}) = \mathbf{C}^T\mathbf{x}_0(t, \mathbf{p}) \tag{5.11}$$

where $\mathbf{C}$ is the $n \times 3$ matrix that sums model species to obtain the three conserved total protein concentrations. The constraint equations (5.10) and (5.11) must be satisfied directly by the solution $(\mathbf{x}_0, T)$; to apply them in solving for the sensitivities in Eq. (5.12), their derivatives with respect to parameters must be included.

To compute the period sensitivity $\frac{\partial T}{\partial \mathbf{p}}$ from the point $(\mathbf{x}_0, T)$ on the limit cycle, we followed a boundary value problem (BVP) approach [83], which solves the system

$$\begin{bmatrix} \mathbf{I}_n - \frac{\partial \mathbf{F}}{\partial \mathbf{x}} & \frac{\partial \mathbf{F}}{\partial \mathbf{x}} \\ \mathbf{C}^T & 0 \\ \frac{\partial \phi}{\partial \mathbf{x}} & 0 \end{bmatrix} \begin{bmatrix} \frac{\partial \mathbf{x}_0}{\partial \mathbf{p}} \\ \frac{\partial T}{\partial \mathbf{p}} \end{bmatrix} = \begin{bmatrix} \frac{\partial \mathbf{F}}{\partial \mathbf{p}}(\mathbf{x}_0, T) \\ 0 \\ 0 \end{bmatrix}, \tag{5.12}$$

We define amplitude sensitivity for a species x_i as the sensitivity of difference between the maximum and minimum concentration achieved during the cycle; this can be calculated by evaluating the sensitivity system at the times of these maxima and minima:

$$\frac{\partial a_i}{\partial \mathbf{p}} = \frac{\partial x_i}{\partial \mathbf{p}}(t_{i,\max}, \mathbf{p}) - \frac{\partial x_i}{\partial \mathbf{p}}(t_{i,\min}, \mathbf{p}). \tag{5.13}$$

Finally, we define the angular phase sensitivity θ_{ij} as the sensitivity to parameters of the timing of the maxima of species x_i and x_j relative to the oscillation period. Intuitively, the angular phase sensitivity measures the fraction of the period orbit spent between these peaks. It can be computed using Eqs. (5.14) and (5.15), where $\frac{\partial t_i^*}{\partial \mathbf{p}}$ is the sensitivity of the peak time of species i to parameters, and $\frac{\partial \mathbf{f}}{\partial \mathbf{x}}$ and $\frac{\partial \mathbf{f}}{\partial \mathbf{p}}$ are the Jacobians of the ODE system with respect to state variables and parameters, respectively.

$$\frac{\partial \theta_{ij}}{\partial \mathbf{p}} = \frac{1}{T}\left(\frac{\partial t_i^*}{\partial \mathbf{p}} - \frac{\partial t_j^*}{\partial \mathbf{p}}\right) - \frac{t_i^* - t_j^*}{T^2}\frac{\partial T}{\partial \mathbf{p}} \tag{5.14}$$

$$\frac{\partial t_i^*}{\partial \mathbf{p}} = \frac{-1}{\frac{\partial f_i}{\partial \mathbf{x}} \cdot \mathbf{f}} \cdot \left(\frac{\partial f_i}{\partial \mathbf{x}}\frac{\partial \mathbf{x}}{\partial \mathbf{p}} - \frac{\partial f_i}{\partial \mathbf{p}}\right) \tag{5.15}$$

For all analysis in this work, we converted the raw sensitivities described above to their corresponding *relative* sensitivities,

$$\frac{\partial \bar{s}_i}{\partial \bar{p}_j} = \frac{\partial s_i}{\partial p_j}\frac{p_j}{s_i}, \tag{5.16}$$

defined by normalizing each raw sensitivity $\frac{\partial s_i}{\partial p_j}$ to both to the value of the parameter p_j and feature s_i.

As stated previously, many corresponding reactions in the model are assumed to occur with identical reaction rates. In order to separate the effects on period, amplitude, and phase of these individual reactions, we assigned a unique parameter to each reaction and computed sensitivities with respect to these 349 'unlumped' parameters [83]. This procedure did not change the simulated dynamics of the model, as the 349 parameters were assigned their original 18 distinct values. We note that this procedure is a useful computational tool, and its application is not meant to imply that these reactions can be modulated independently in the biological system.

Chapter 6

Conclusions and future directions

The introduction to this thesis posed a question: can we develop a predictive, end-to-end understanding of a cell decision process? In the subsequent chapters, we have demonstrated that interconnecting pathways to such an end-to-end understanding is feasible. Along the way, we developed a new understand of pathways at three levels of complexity, drawing on techniques from cell biology, biochemistry, mathematics and computer science.

The p53 protein is a crucial node regulating the cell's response to stress and conferring protection against tumorigenesis. Its importance is evidenced by its inactivation in cancer: approximately 40% of tumors harbor p53 mutations preventing its induction of downstream genes. Alongside its clinical significance, p53 is a signaling hub, serving as an integration point for signals arising from a variety of inputs, as well as an activator of multiple distinct output pathways [33]. Understanding signaling through such hubs provides a challenge to the biologist. How can distinct downstream fates be chosen by a single protein when it is activated by different signals, in specific contexts? In Chapters 2 and 3, we quantitatively addressed two potential answers to this question: the dynamics of activation may select certain downstream programs, and the systems-level properties of the connections between upstream and downstream pathways can provide higher-level logical control.

Dynamics offer a resolution the dilemma of hub signaling. Different patterns of activation can select individual downstream fates, or provide a mechanism to sense

the time or duration of an input signal. However, while such mechanisms have been proposed and tested theoretically, experimental evidence for them has been largely lacking. To demonstrate the causal relationship between dynamics and downstream processes such as target gene activation, features of these dynamics must be perturbed and the resulting effect measured. This procedure is reminiscent of the classical approach for determining the function of individual network components: silencing, knockout, and overexpression experiments can demonstrate the effect of individual genes on system function.

In Chapter 2, we provide a framework for the application of these approaches to probe mammalian signaling dynamics. Through quantitative construction and investigation of a reduced p53 circuit, we identifying three perturbations that allow *dynamical features*, rather than genes, to be modulated individually. Through a synthetic transcriptional input, we can tune p53 pulse amplitude without affecting frequency or the rate of damping. Model prediction, coupled with experimental validation, demonstrates that additional synthetic feedback loops tune damping without affecting amplitude, and that frequency can be modulated through parameters controlling p53-Mdm2 feedback. These results provide a platform permitting a new line of investigation. By treating individual dynamical features as entities that can be 'knocked down,' their contribution to system function can be assessed.

In Chapter 3, we have described a series of complex interactions linking DNA damage induced by ionizing radiation to cell cycle arrest. This work demonstrates, to our knowledge, the first attempt at integrating large models of mammalian pathways to understand their combined function. The full model comprising both pathways revealed new systems-level organization, showing that multiple mechanisms of cell cycle arrest can each contribute individual features to the cell's response. It was also able to match the data of combinations of mechanisms observed experimentally, showing that multiple mechanisms interact with one another to properly maintain arrest.

One of the major sources of complexity that arises in the cell cycle is the multiplicity of closely related proteins participating in its regulation. Each of the cyclins E, A and B represent *families* of related proteins, rather than single entities. Not one,

but three protein phosphatases—Cdc25A, Cdc25B, and Cdc25C—can participate in activating these cyclins. A third example of this multiplicity arises in the cyclin-dependent kinase inhibitors, p21 and p27, which play overlapping roles in cyclin/Cdk inactivation. This multiplicity can have surprising consequences. For instance, a number of cell cycle regulators can be knocked out without affecting the cell's ability to cycle, including cyclin D, cyclin E, or the Cdc25 isoforms [110, 111]. Notably, our model is able to recapitulate many of these results (B.2).

However, despite this robustness of function, some seemingly-redundant species confer specific and irreplaceable functions. In this thesis, we identify a role for p21 in conferring protection against endoreduplication after DNA damage. This function cannot be provided by the related cyclin-dependent kinase inhibitor p27, as it requires the induction of p21 by p53 in response to damage. This specific role suggests an explanation to a prior clinical observation: although p53 is so frequently mutated in response to cancer, inactivation of p53 downstream genes such as p21 is much rarer. Our work shows that maintaining p53 activity in the absence of p21 reveals systems-level problems in the maintenance of cell cycle arrest, leading to reentry into the wrong cycle phase.

This work demonstrates the possibility of deriving new insights from interconnecting models of mammalian signaling pathways. Applying this approach to additional pathways may yield immediate insights into other cell decisions. Many DNA damaging agents and other stresses signal through the p53 pathway, which in turn can regulate apoptosis and DNA repair in addition to cell cycle arrest. It will be critical to extend these efforts through characterizing additional pathways experimentally and computationally, and work to integrate them into a quantitative picture of the larger network. Some important pathways that await integration are:

1. signaling through ATR/Chk1 and p38 in response to UV and DNA alkylating agents,

2. regulation through ARF in response to oncogenic stress,

3. apoptosis and its regulation by p53 its upstream kinases, and

4. a detailed treatment of mitosis and mitotic failure.

Incorporating these pathways into combined, mechanistic models can offer complementary insights to those obtained through statistical modeling, such as PLSR and Bayesian network analysis. In particular, accounting for the specific temporal profiles of signals poses a challenge to statistical models but can be naturally captured in an ODE framework. Together these approaches hold great promise in overcoming the challenges faced by systems biology to quantitatively predict and provide intuition into biological processes.

Some signaling interactions are conferred by well-understood biochemical machinery, such as the enzymatic action of a kinase on its substrate, or binding between a protein and its ligand. As an increasing number of signaling processes are understood in detail, however, it is becoming clear that such simple binding-equilibrium and enzymatic processes are not ubiquitous, and might even be in the minority. ATM activation after ionizing radiation proceeds through the cross-phosphorylation of inactive dimers, and their subsequent dissociation. Scaffolding proteins can provide both spatial and kinetic selectivity of pathway activation. While theoretical studies have demonstrated the potential power conferred by specific biochemical mechanisms, we lack practical techniques to gain intuition for the operation of natural systems exhibiting this complexity.

We addressed these challenges by developing mathematical and computational techniques that can be applied to a biochemical systems governed by mass action kinetics. Using two complementary approaches, we identified the mechanisms underlying the operation of a detailed biochemical model of the cyanobacterial circadian clock. By making infinitesimal perturbations to the full system, we showed that a distributed set of reactions control dynamics in the full network, but that these reactions cluster into groups of self-consistent processes. These processes provide a principled approach towards abstracting the detail of individual reactions to more general pathway components with a defined effect on dynamics. We found that these processes also controlled multiple dynamical features in concert, demonstrating their generality. Notably, the processes we identified were not clustered into intuitive feedback loops,

as had been observed in transcriptional networks, systems with a much lower degree of connectivity between components.

To obtain an intuitive picture of the network topology driving dynamics in such densely connected networks, we took an approach utilizing larger scale perturbations. By exhaustively enumerating the subnetworks still capable of oscillation, we were able to identify two motifs driving dynamics in the circadian system. Each oscillates with a period of approximately 1 day, and the two motifs coupled with roughly equal weight in the full system. Motifs, especially those driving oscillating dynamics, have been the subject of much scrutiny in recent years, and identifying the design principles of biological oscillators is an open area of research. Here, we reveal a novel oscillator topology, consisting of coupled delay and hysteretic motifs, in a network consisting of only three proteins. This topology generates robust oscillation, as is apparent from the success of this model in explaining observed experimental data, even when oscillatory data was not part of the fitting procedure. Furthermore, it provides a specific advantage in a circadian system: the ability to tune the phase of oscillation without affecting the period. This is useful because it provides a mechanism for maintaining a 24 hour cycle while accounting for seasonal variation in the number of daylight hours. Indeed, cyanobacteria are found at a variety of latitudes, suggesting that this tunability could provide a selective advantage. It should be noted that the techniques we employed to identify these motifs are general, suggesting that in the future, more complex motifs can be similarly identified and characterized.

Our results demonstrate that densely interconnected biochemical networks can be both described in mechanistic detail, and analyzed to offer intuition for the function of the full system. Future work must translate this initial success to considering larger networks in more complex contexts. For instance, different ErbB heterdimereric complexes and modification states make the EGF network an ideal candidate for these methods' application in a mammalian system. While our sensitivity analysis based approaches can translate immediately to more complex networks, exhaustive enumeration of subnetworks is limited to smaller systems by the complexity of this approach, scaling combinatorially with the number of reactions. However, there is considerable

room for improvement in efficiency through the use of guaranteed search procedures, especially if enumeration of only the minimal subnetworks are desired.

Through these studies, we have demonstrated a paradigm for coupling the analysis of detailed, mechanistic pathway models to the intuition gained from biologicial experiments. In this paradigm, a model is quantitatively matched to a broad variety data from a well-studied pathway, and the addition or modulation of specific components are tested exhaustively across the parameters controlling their action, leading to predictions that can qualitatively discriminate behavior across a range of parameter values. Through application to the mammalian DNA damage response, we used this approach to comprehensively demonstrate how synthetic feedback loops and reaction rate constants affects p53 dynamics a reduced pathway, and to identify broad classes of arrest states in the cell cycle network through the action of stoichiometric and enzymatic inhibition, as well as transcriptional repression. Successive rounds of closely coupled modeling and experimentation are an absolute requirement for this approach, as exhaustive sampling of the behavior resulting from new connections is only feasible if it proceeds from an existing quantitative understanding, and can be subsequently verified experimentally.

Appendix A

Supporting materials and methods for Chapter 2

A.1 Estimating pulse statistics from trajectories

We used frequency, amplitude and rate of damping to quantitatively characterize oscillatory trajectories. For both experimental and computational data, we estimated these three characteristics from measurements of the steady state concentration (x_{ss}) and the times $t_{max}^{(i)}$ and amplitudes $x_{max}^{(i)}$ of the i$^{\text{th}}$ maxima during oscillation. For experimental data, where autofluorescence near the time of cell division often obscured dynamics, we split trajectories before and after cell division events and computed frequency and damping independently for each.

We estimated the frequency of oscillation by computing the differences in timing between successive peaks using the equation

$$\omega = \left\langle \frac{2\pi}{t_{max}^{(i+1)} - t_{max}^{(i)}} \right\rangle, \tag{A.1}$$

where the average is taken over all pairs of pulses throughout the population of cells. We ensured that this approach yielded similar results to those obtained by pitch analysis as in [12]. To estimate the amplitude of the first pulse, we computed the ratio of the steady state concentration to the peak concentration. This measurement

represents the fold upregulation of p53 from its basal level.

$$a_1 = x_{max}^{(1)}/x_{ss}. \tag{A.2}$$

To estimate the rate of oscillatory damping, one must take into account both the relative amplitude of subsequent pulses as well as the time elapsed between them. We assumed that damping between successive extrema occurred *via* an exponential decay, and that this decay is offset from the origin by the steady state level of protein:

$$x(t_{max}) - x_{ss} = (x_0 - x_{ss})e^{-rt_{max}}. \tag{A.3}$$

We verified that these assumptions accurately fit the behavior of successive peaks obtained from simulation. From measurement of maxima of successive pulses, we could compute the the decay constant of this exponential decay for each trajectory according to the formula

$$r = \left\langle \frac{log\left(x_{max}^{(i)} - x_{ss}\right) - log\left(x_{max}^{(i+1)} - x_{ss}\right)}{t_{max}^{(i+1)} - t_{max}^{(i)}} \right\rangle, \tag{A.4}$$

where the average is again taken over all pairs of pulses throughout the population of cells.

Our live cell microscopy experiments measured levels of exogenous p53-CFP fusion protein in individual cells; we do not expect the endogenous p53 copy to be affected proportionally as it is not under the control of a zinc-inducible promoter. In order to correctly account for the endogenous protein level in our calculations of p53 pulse amplitude and damping, we used Western blotting to measure the induction of total protein levels during the first p53 pulse after 25 and 50 μM $ZnCl_2$ stimulation (Figure A-1B). We observed some change in endogenous p53 levels, although these differences were much smaller than the increase in exogenous protein levels. This change may be due to competition for Mdm2 binding between exogenous and endogenous p53, leading to indirect stabilization of the endogenous protein.

By summing the total protein in both the endogenous and exogenous bands, we found that zinc induction leads to an approximately 2-fold (11-fold) activation after zinc treatment at 25 (50) μM. We used these points to compute the equivalent fluorescence intensity expected from endogenous p53 using the equation

$$\frac{(p53_{max}^{exog} + p53^{endog})}{(p53_{0}^{exog} + p53^{endog})} = f \tag{A.5}$$

where f denotes the total p53 pulse amplitude after zinc treatment (as measured by Western blot), $p53_{max}^{exog}$ is the mean first-amplitude fluorescence observed by live-cell microscopy, $p53^{endog}$ is the unknown equivalent fluorescence of endogenous p53, and $p53_{0}^{exog}$ is mean initial level of exogenous p53 observed by live-cell microscopy. For our initial experiment of zinc stimulation, we obtained a least-squares solution using our data at 25 and 50 μM to solve for $p53^{endog}$, and computed the p53 amplitudes at each zinc concentration using this value and the corresponding $p53^{exog}$ values determined by microscopy. In every subsequent experiment, we included a 50μM zinc treatment control to compute the corresponding value of $p53^{endog}$ and obtain an estimate for the true amplitudes associated with p53 dynamics.

A.2 Model construction and parameterization

Our model of the core p53-Mdm2 negative feedback loop was based on the topology of Model IV from Ref. 12. Rather than a stiff delay, we implemented a boxcar procedure of 3 linear ODEs representing the combined delay of translation, protein folding, and nuclear translocation (*i.e.* $1 \leq i \leq 3$ in Eq. (A.6)). The perfectly stiff delay is recovered in the limit of an infinite number of boxcar steps, and as we assume equally distributed boxcar steps, our approach does not add any additional parameters. The full model is shown in Eq. (A.6) and all results were generated using

the parameters listed in A.1.

$$\begin{aligned}
\dot{p} &= \alpha_x(1 + p_z Z(t)) - \frac{\delta_p}{1 + N(t)} mp \\
\dot{m}_0 &= \gamma_0(\alpha_0 + \beta_0 \frac{Kp^2}{K_0^2 + p^2} - m_0) \\
\dot{m}_i &= \gamma_0(m_{i-1} - m_i) \\
\dot{m} &= \alpha_m m_4 - \delta_m \frac{m}{K_m + m} - \gamma_m m
\end{aligned} \tag{A.6}$$

To compare the negative feedback model with data after zinc induction, we estimated the transfer function from zinc stimulation to p53 induction through the action of MTF1. From initial model simulations, we found that the the p53 first pulse amplitude scaled linearly with the transcriptional stimulus applied to the model (data not shown). However, we observed experimentally that zinc dose scaled superlinearly with amplitude. We fit a Hill-model saturation curve (Eq. (A.7)) with $n = 3$ representing this transfer function to the experimentally observed p53 increase in first pulse concentration at five zinc concentrations (Figure A-1D), and used this transfer function to map specific zinc doses to the model input $Z(t)$ for all subsequent analyses. The remaining parameter values matching experimentally determined pulse amplitudes, frequencies and timings were determined by hand and are listed in Table A.1; the qualitative features of these trajectories were observed over a broad range of parameter values.

$$TF(z) = a \frac{z^n}{K_z^n + z^n} \tag{A.7}$$

A.3 Noise simulations

Noise in protein synthesis rates constitutes a major source of variability in mammalian cells, and varies with a correlation time on the order of hours [12, 21]. We modeled variability between cells using our core NF model by implementing multiplicative transcriptional noise on the parameters α_p and $\alpha_m 0$. To incorporate a

nonzero correlation time without driving the system with a strongly periodic signal, we implemented a simple iterative algorithm to generate Gaussian colored noise as described in Ref. 150. In applying this algorithm, we chose a correlation time τ of 5 hours and a standard deviation $\sigma = 0.1$, and a sampling time of 10 minutes. To numerically integrate the ODE system in the presence of this noise function, we assumed that the value of the noise term was constant between sampling times. Similar results were obtained using a sampling time of 1 minute (data not shown), validating this assumption.

A.4 Modeling synthetic feedback

For predictions and analyses of additional synthetic negative and positive feedback loops, we constructed an augmented model in which we also included boxcar equations for p53 and MTF1 (or MTF1-KRAB) protein production (the full models are shown in Eq. (A.8), where $1 \leq i \leq 4$). Because these models add additional positive and negative feedback loops to the original negative feedback (NF) circuit, we refer to them as NPF and NNF systems, respectively. No parameter values of the p53-Mdm2 interaction were modified in construction of these models. Furthermore, because the Mdm2 promoter was used to drive MTF1 expression, the same parameter values were used to represent p53 induction of MTF1 as used for Mdm2. We ensured that in the absence of feedback, these augmented models led to identical p53 and Mdm2 dynamics as the NF model. Table A.2 list the parameter values used for these models.

$$
\begin{aligned}
\dot{p}_0 &= \begin{cases} \gamma_{p0}(\alpha_{p0} + \beta_{p0} Z(t) \frac{f^2}{K_f^2 + f^2} - p_0) & \text{(NPF model)} \\ \gamma_{p0}(\alpha_{p0} + \beta_{p0} Z(t) \frac{K_f^2}{K_f^2 + f^2} - p_0) & \text{(NNF model)} \end{cases} \\
\dot{p}_i &= \gamma_{p0}(p_{i-1} - p_i) \\
\dot{p} &= \alpha_p p_4 - \frac{\delta_p}{1 + N(t)} mp \\
\dot{m}_0 &= \gamma_{m0}(\alpha_{m0} + \beta_{m0} \frac{p^2}{K_0^2 + p^2} - m_0) \\
\dot{m}_i &= \gamma_{m0}(m_{i-1} - m_i) \\
\dot{m} &= \alpha_m m_4 - \delta_m \frac{m}{K_m + m} - \gamma_m m \\
\dot{f}_0 &= \gamma_{f0}(\alpha_{f0} + \beta_{f0} \frac{p^2}{K_0^2 + p^2} - f_0) \\
\dot{f}_i &= \gamma_{f0}(f_{i-1} - f_i) \\
\dot{f} &= \alpha_f f_4 - \gamma_f f
\end{aligned}
\tag{A.8}
$$

After this procedure, two sources of parameters remained unconstrained: the delay through the synthetic feedback loop represented by p53 and MTF1 protein synthesis, and the feedback strength represented by the effect of variable MTF1 levels on p53 production. We did not know *a priori* which values these processes would take, and chose to query oscillatory characteristics over a wide range of parameter values. For this analysis, we sampled the parameters γ_{f0} and β_{p0} (to change delay time and feedback strength, respectively) at 51 points one order of magnitude above and below their nominal parameterization, and computed the resulting p53 amplitudes, frequencies and damping coefficients; the full results are shown in Figure A-3. We briefly summarize and discuss these results here.

For the synthetic negative feedback loop, we observed damped oscillation with a first pulse amplitude and frequency that remained consistent across delay times and feedback strengths (Figure A-3A-C). The damping results varied widely, and trajectories were both stabilized and destabilized for some values of the parameters being

varied. Both long and short delay times stabilized the system by increasing the oscillation damping rate. Examination of the corresponding trajectories showed that for many of these, a single first pulse dominates the dynamical response. Notably, however, we found the addition of the synthetic negative feedback loop could destabilize oscillation if its delay time is tuned near that of the corresponding parameter in the p53-Mdm2 loop (blue region, Figure A-3A).

As we had observed for the synthetic negative feedback, addition of a positive feedback loop did not have a pronounced effect on frequency and amplitude except at strong feedback strengths (Figure A-3D-F). Examination of the corresponding trajectories in this regime indicates a single high-amplitude pulse followed by low amplitude pulses. The transition to this regime was sudden, occurring over a small range of feedback strengths (see Figure A-3E, upper region); at lower feedback strengths, only a small dependency of frequency and amplitude on the presence of feedback was observed. Like the NNF system, the NPF system is predicted to have a more pronounced effect on damping, but with opposite effects (Figures A-3A,D). For a broad range of parameter values, the synthetic positive feedback loop is predicted to destabilize oscillation. Furthermore, for some values of feedback strengths and delay times, the model predicts a transition to undamped oscillation (Figure A-3D).

A.5 Parameter perturbation analysis

To map the effect of parameter variation on oscillation amplitude and frequency, we simulated our model after individually varying each parameter at 50 logarthmically spaced points from one order of magnitude below and above its nominal value, and computed the oscillation frequency and amplitude at each point. Simulations were run for until the 5000 h elapsed or five maxima were observed; simulations that were not seen to oscillate by undergoing at least three pulses were discarded. The resulting trajectories are shown in Figure A-4, and are colored lighter for higher parameter values. The corresponding points in amplitude–frequency space are shown in Figure A-5. Broadly, we found that parameters affecting the negative feedback

loop's delay (*e.g.* γ_0, the delay in Mdm2 protein maturation and δ_p, the rate of Mdm2-mediated p53 degradation; see Figure A-5C,D,F) had the largest effect on frequency, while basal protein production rates (*e.g.* α_p and α_0 for p53 and Mdm2, respectively; see Figure A-5A,B,E) had a smaller effect on frequency.

A.6 Supplementary experimental methods

A.6.1 Immunoblots

Western blots were performed as described previously [13]. Antibodies used were anti-p53 (DO-1, Santa Cruz Biotechnology) and anti-β-tubulin (E7, Developmental Studies Hybridoma Bank).

A.6.2 Cell lines and expression constructs

Mdm2p-YFP was constructed previously as described in Ref. 9. MCF7 cells were transfected (FuGene6, Roche) and stable clones were selected by flow cytometry for expression of YFP.

A.6.3 Time lapse microscopy

MCF7 Mdm2p-YFP cell line was imaged using FMS2 for 48 hours every 20 min after addition of differing concentrations of $ZnCl_2$. Cells were analyzed as described in Materials and Methods.

Parameter name	Parameter value	Description	Units
α_p	0.040	p53 protein production rate	[p53/h]
p_z	1.529	Zinc-mediated increase in p53 production rate	[1/Zn]
δ_p	0.329	Mdm2-mediated p53 degradation rate	[1/h]
p_n	1.738	Nutlin-associated decrease in p53-Mdm2 assocation	[1/EC50]
α_{m0}	0.052	Mdm2 mRNA basal transcription rate	[Mdm2 mRNA]
β_{m0}	0.554	p53-mediated Mdm2 mRNA transcription rate	[Mdm2 mRNA]
K_{m0}	0.744	Saturation of p53-mediated Mdm2 production	[p53]
γ_{m0}	0.239	Mdm2 production delay	[1/h]
α_m	1	Mdm2 protein production rate	[Mdm2/h]
γ_m	0.016	Mdm2 degradation rate	[1/h]
δ_m	0.748	Mdm2-mediated Mdm2 degradation rate	[1/h]
K_m	2.169	Saturation of Mdm2-mediated Mdm2 degradation	[Mdm2]

Table A-0: Parameters for the core p53-Mdm2 negative feedback (NF) model. The table indicates each parameter's name, its nominal value, a description of its effect, and the units in which it is measured.

Parameter name	Parameter value	Description	Units
α_{p0}	0.040	p53 mRNA basal transcription (txn)	[p53 mRNA]
β_{p0}^{NPF}	14.82	NPF p53-mediated Mdm2 mRNA txn	[Mdm2 mRNA]
β_{p0}^{NNF}	0.317	NNF p53-mediated Mdm2 mRNA txn	[Mdm2 mRNA]
K_{p0}	2	Saturation of p53-mediated Mdm2 prod.	[p53]
γ_{p0}	3.293	Mdm2 production delay	[1/h]
α_p	1	Mdm2-mediated p53 degradation	[p53/h]
δ_p	0.329	Mdm2-mediated p53 degradation	[1/h]
p_n	1.738	Nutlin-associated decrease in p53-Mdm2 assoc.	[1/EC50]
α_{m0}	0.052	Mdm2 mRNA basal txn	[Mdm2 mRNA]
β_{m0}	0.554	p53-mediated Mdm2 mRNA txn	[Mdm2 mRNA]
K_{m0}	0.744	Saturation of p53-mediated Mdm2 prod.	[p53]
γ_{m0}	0.239	Mdm2 production delay	[1/h]
α_m	1	Mdm2 protein production	[Mdm2/h]
γ_m	0.016	Mdm2 degradation	[1/h]
δ_m	0.748	Mdm2-mediated Mdm2 degradation	[1/h]
K_m	2.169	Saturation of Mdm2-mediated Mdm2 deg.	[Mdm2]
α_{f0}	0.040	feedback species mRNA basal txn	[feedback mRNA]
β_{f0}^{NPF}	0.396	NPF p53-mediated Mdm2 mRNA txn	[feedback mRNA]
β_{f0}^{NNF}	31.05	NNF p53-mediated Mdm2 mRNA txn	[feedback mRNA]
K_{f0}	0.744	Saturation of p53-mediated feedback prod.	[p53]
γ_{f0}	1	Feedback protein production delay	[1/h]
α_f	1	Feedback protein production	[feedback protein/h]
γ_f	0.853	Feedback protein degradation	[1/h]

Table A-0: Parameters for the synthetic p53-Mdm2 models. The table indicates each parameter's name, its nominal value, a description of its effect, and the units in which it is measured. Parameters changed between models incorporating synthetic positive (NPF) and negative (NNF) feedback loops are indicated with superscripts. Dynamics of the core NF loop in the absence of feedback (obtained by setting β_{f0} to zero) are identical to those obtained from the original NF model.

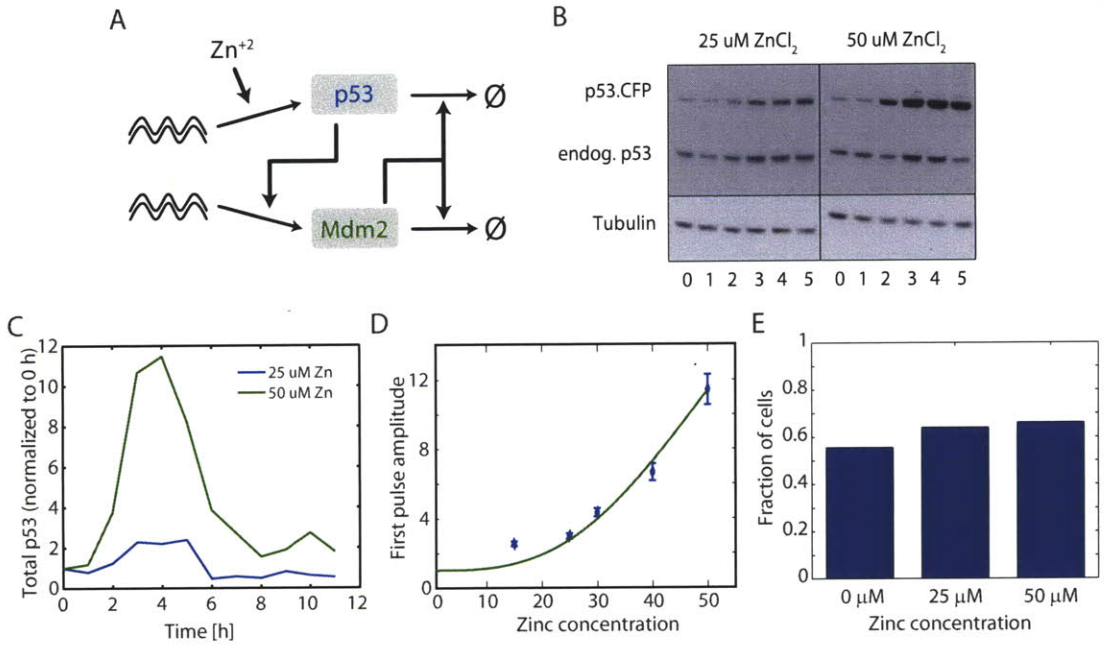


Figure A-1: **Cell population data of the first p53 pulse after zinc stimulation.** (A) Detailed schematic diagram of the negative p53-Mdm2 feedback circuit, showing the three nonlinear reactions (bold arrows in center) representing the action of p53 on the Mdm2 promoter, the ubiquitination of p53 by Mdm2, and Mdm2 autoubiquitination. (B) Timecourse of endogenous p53 and exogenous p53-CFP levels in MCF7 cells after addition of 25 or 50 μM $ZnCl_2$. Cells were collected every hour and analyzed by Western blot. (C) Timecourse of total p53 induction, quantified from the Western blot in (A). All samples are normalized to tubulin and are shown relative to the value at $t = 0$. (D) The p53 first pulse amplitudes at five zinc concentrations (points represent as mean + s.e.m.) are shown with the best-fit Hill equation (Eq. (A.7)) with parameters $a = 62.77$, $K_{zn} = 30.47$ and $n = 3$. (E) Percentage of cells that divide within 24 h after zinc treatment. 25 and 50 μM data shows the mean of four independent experiments; 0 μM (untreated) results were tabulated from a single experiment.

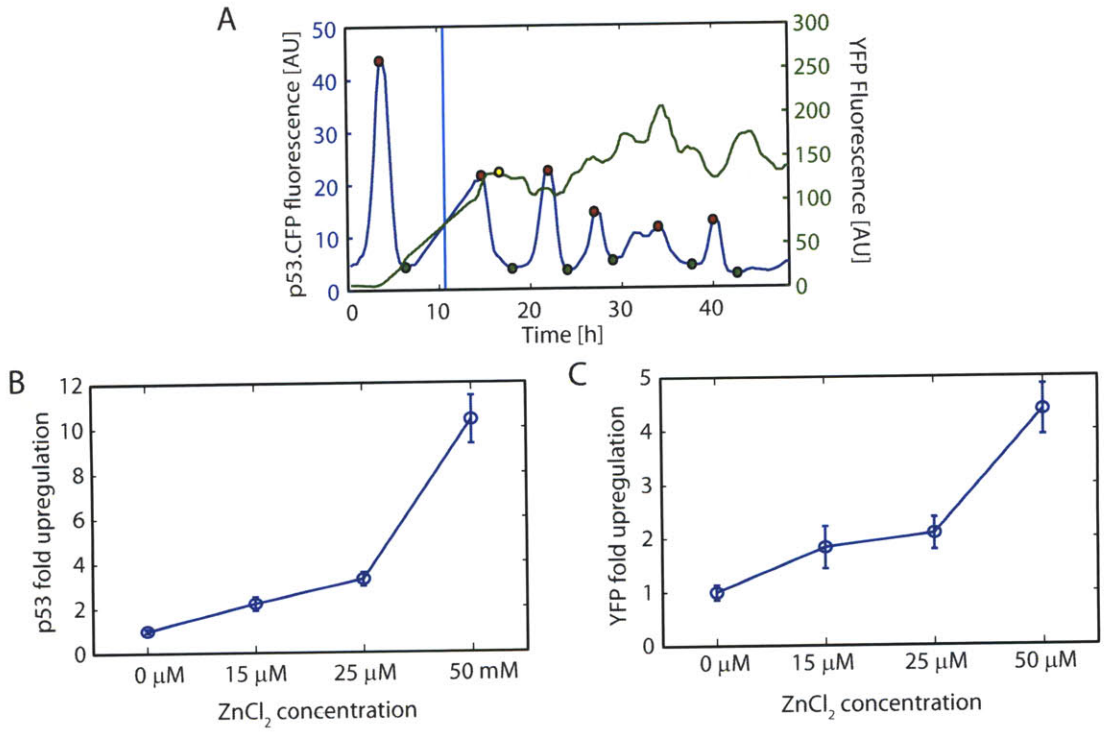


Figure A-2: **p53-responsive YFP induction in response to varying inducer levels.** (A) Representative single-cell trajectory showing p53-CFP (blue curve) and YFP fluorescence (green curve) intensities over time. Automatically identified maxima and minima are shown as points on each trajectory; vertical cyan bar represents time of cell division. (B-C) First pulse amplitude (mean + s.e.m) for (B) p53 and (C) YFP curves at three zinc concentrations. Data points are normalized to 0 μM condition.

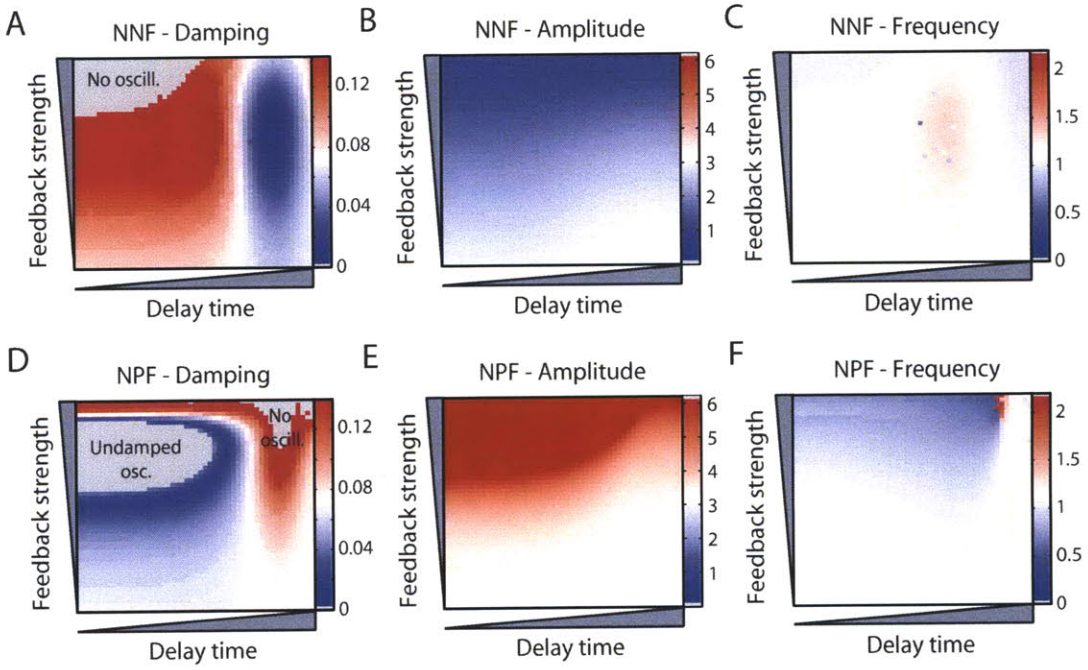


Figure A-3: **Predictions for addition of synthetic feedback loops.** (A-C) Prediction of (A) p53 damping, (B), p53 amplitude and (C) p53 frequency coefficient for the model incorporating a second negative feedback in addition to the core p53-Mdm2 loop (NNF model). X and y axes represent 51 logarithmically spaced values from 0.1x to 10x the nominal parameter value for the MTF1 protein production delay γ_{f0} and synthetic feedback strength β_{p0}, respectively. Colors map values as indicated on each panel. (D-F) Predictions for the model containing a synthetic positive feedback loop (NPF model). All plots were generated as described in (A-C).

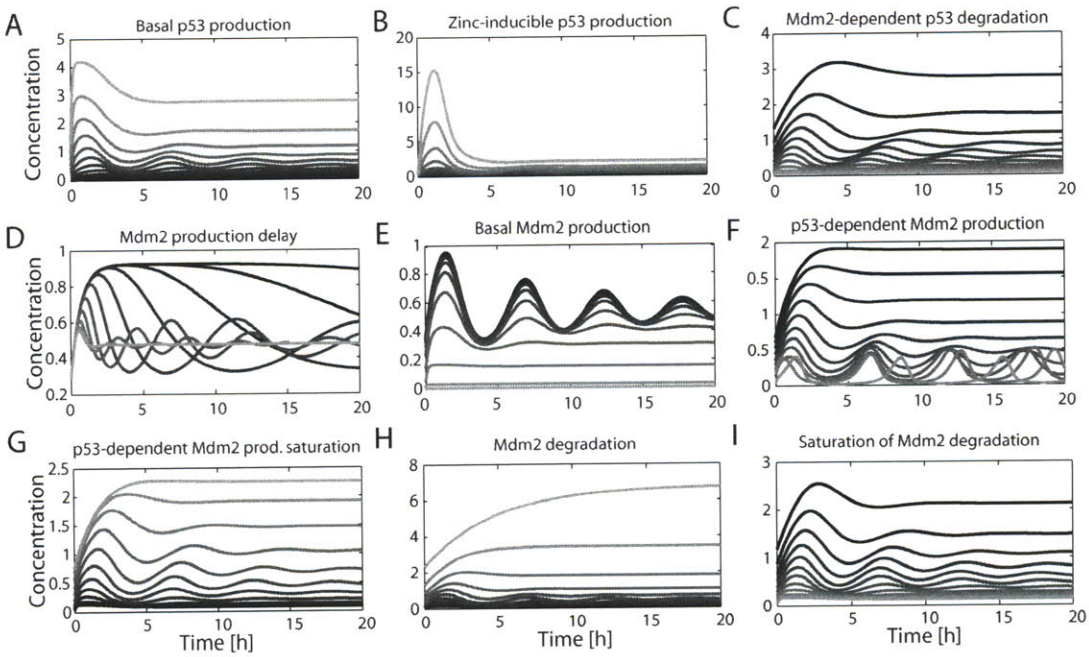


Figure A-4: Effect of parameter perturbations on oscillatory trajectories. Trajectories from the core negative feedback model are shown for a range of values of key model parameters. In each panel, p53 concentration is plotted over time for 21 logarithmically spaced values of the variable parameter, ranging from 1.5 orders of magnitude below to 1.5 orders of magnitude above its nominal value. Parameters shown are (A) α_p, (B) p_z, (C) δ_p, (D) γ_0, (E) α_0, (F) β_0, (G) K_0, (H) δ_y and (I) K_y.

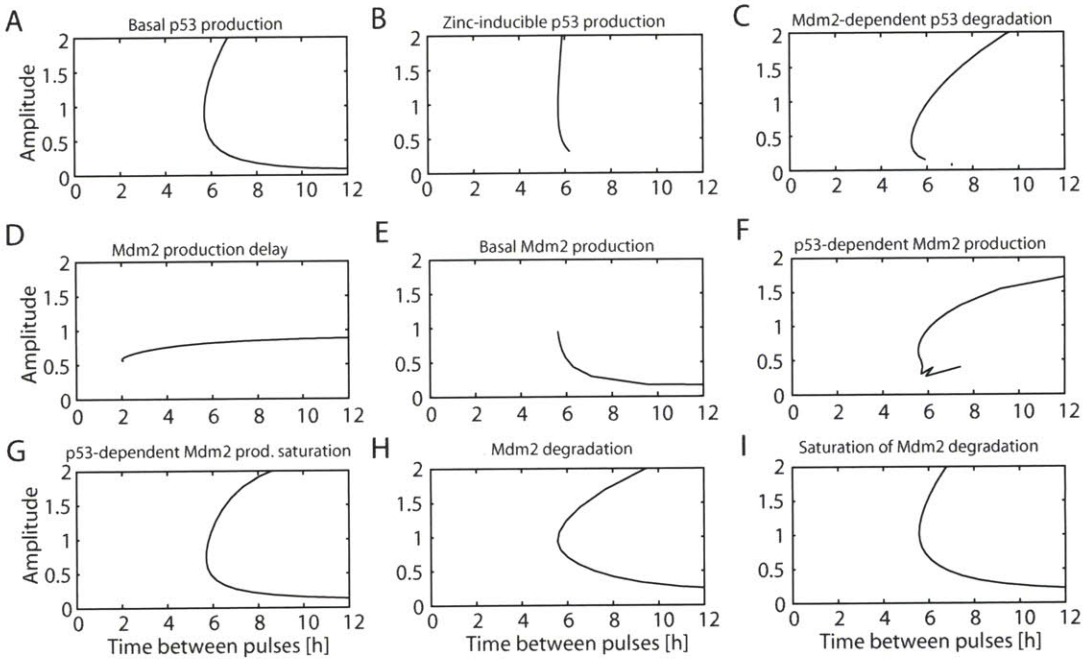


Figure A-5: Effect of parameter perturbations on oscillation amplitude and frequency. The variation in oscillation features are shown for a range of values of key model parameters. In each panel, curves trace the oscillation frequency and first pulse amplitude at different logarithmically spaced values of an individual model parameter, ranging from 1.5 orders of magnitude below to 1.5 orders of magnitude above its nominal value. Parameters shown are (A) α_p, (B) p_z, (C) δ_p, (D) γ_0, (E) α_0, (F) β_0, (G) K_0, (H) δ_y and (I) K_y.

Appendix B

Supporting materials and methods for Chapter 3

B.1 Computational model of the integrated cell cycle and DNA damage network

Model construction

The integrated model is comprised of sub-models for the cell cycle and the DNA damage response. The DNA damage signaling model consists of 72 state variables and 135 nonzero parameters, while the cell cycle model and its mechanisms of arrest consists of 31 state variables and 85 nonzero parameters. Both are ordinary differential equation models that describe rates of change of concentrations of proteins and other species resulting from chemical reaction processes.

The cell cycle model is based on a recently published model by Tyson and colleagues (1). In the original model, it was assumed that a variety of molecular species react quickly and are at quasi-steady state. Such species are represented by algebraic rather than differential equations. In order to incorporate interactions between these species and the DNA damage model, we expanded the algebraic relations into dynamical equations. In addition, we added protein synthesis and degradation reactions

for each species in the model to simulate the decrease in protein synthesis caused by cycloheximide treatment, an experiment that helped define the restriction point in mammalian cells (2). These changes affect the equations governing the dynamics of the model species `Wee1`, `Cdc25C`, `TFE` (E2F1), `TFB` (Mcm), and `APC` (model species are named according to the Tyson model). We ensured that the model's freely cycling dynamics were unchanged by incorporating synthesis and degradation of each cell cycle protein, and by treating variables dynamically rather than at steady state. These changes were essential for simulating arrest, as this process requires dynamical interaction with cell cycle proteins, and can connect to the cell cycle by reducing protein synthesis rates.

The Tyson model links cell cycle progression to cell size by setting cyclin synthesis rates proportional to an exponentially increasing cell mass (1). While size dependent progression through the cell cycle has been shown in yeast, the existence of such a mechanism in mammalian cells is not well understood (3). In addition, the regulation of cell growth during arrest is poorly understood, and initial simulations of cell cycle arrest without incorporating such regulation resulted in growth to unphysical cell sizes (data not shown). We therefore define a mitogen parameter, M, representing the intracellular signal resulting from extracellular growth factors present in the medium. We model this mitogen parameter's effect as increasing the synthesis rates of CycD and CycE transcripts, replacing their dependence on cell size. For cell cycle simulations we set $M = 1.8$ unless otherwise indicated, a value that leads to cyclin synthesis rates comparable to those attained with the original model.

The topology of the DNA damage model was derived from the model of Batchelor et al., in which oscillations are driven by a combination of two negative feedback loops: the core p53-Mdm2 loop and a loop in which the upstream checkpoint kinases are inhibited by a p53-inducible gene product, the phosphatase Wip1 [11]. To provide an extensible framework for future modeling of the DNA damage network, we incorporate additional feedback loops involving p53 [14] in our model (FigureB-1A shows a detailed network diagram). With the current parameterization, however, these loops do not significantly affect the network's dynamics.

Modeling cyclin knockouts and the restriction point

To test the model's ability to match known experimental results for mammalian cells, we simulated serum starvation and cycloheximide treatment. To do the former, we modeled an immediate, step-like decrease in M from 1.8 to 1 at different times during a normal cell cycle (FigureB-1B,B'). These simulations recapitulate the existence of a restriction point: if M is decreased within 3 hours of cell division, cells arrest immediately in G1, while for later serum removal times, a full cell cycle is completed before the cell arrests. Cycloheximide treatment is simulated by a decrease in all protein synthesis rates by 20% at various times, leading to similar results (FigureB-1C,C'). For cyclin D and E knockout simulations, the synthesis rates and initial conditions of these species were set to zero. For $M = 2$, the cyclin $E^{-/-}$ model still cycles, albeit more slowly (FigureB-1D). At this mitogen level, however, the cyclin $D^{-/-}$ model undergoes G1 arrest. The cyclin $D^{-/-}$ model only cycles for values of M ¿ 10, consistent with the qualitative observation that cyclin $D^{-/-}$ cells have an increased mitogen requirement for proliferation (FigureB-1E).

Modeling cell cycle arrest

In addition to the cell cycle arrest mechanisms described in detail in the main text, other interactions leading to arrest have been described, for example (4, 5). To ensure that the interactions we consider are representative of the space of possible G1 and G2 arrest mechanisms, we systematically applied arrest by stoichiometric inhibition, enzymatic inactivation or transcriptional repression to all species of the cell cycle model.

Modeling arrest mechanism biochemistry

We modeled cell cycle arrest using three distinct biochemical interactions: protein inactivation by inhibitor binding, protein inactivation by enzymatic phosphorylation, and a decrease in protein level by repression. In the first of these, species x_i is inhibited by the inhibitor I through reactions modeling inhibitor binding formation

of the complex c, as well as the degradation of each constituent species (Equation B.1). Any additional species-specific terms governing the dynamics of are kept, and the protein's specific degradation term, denoted by $V_{D,i}$, is assumed to act on both the bound and unbound species. We assume the inhibitor is upregulated in proportion to the level of tetrameric nuclear p53.

$$\begin{aligned} \dot{x}_i &= (\text{other terms}) - k_f x_i I + k_r c - V_{D,i} x_i + k_d I \\ \dot{I} &= k_s p53_{4np}(t) - k_d I + V_{D,i} x_i \\ c &= k_f x_i I - k_r c - V_{D,i} c - k_d I \end{aligned} \tag{B.1}$$

We model enzymatic phosphorylation by a Michaelis-Menten rate law in which species x_i is modified to an inactive phosphorylated form x_i^P by phospho-Chk2 (Equation B.2). The phosphorylated form is assumed to degrade at the same rate as the unmodified form.

$$\begin{aligned} \dot{x}_i &= (\text{other terms}) - k_{cat} Chk2_p(t) \tfrac{x_i}{K_m + x_i} - V_{D,i} x_i \\ \dot{x}_i^P &= k_{cat} Chk2_p(t) \tfrac{x_i}{K_m + x_i} - V_{D,i} x_i^P \end{aligned} \tag{B.2}$$

Finally, protein synthesis inhibition was modeled through a repressive saturation term dependent on tetrameric nuclear p53 (Equation 3). This term modulates the protein's usual production rate, denoted by k_{si}.

$$\dot{x}_i = k_{si} \frac{K_D}{K_D + p53_{4np}(t)} + (\text{other terms}) \tag{B.3}$$

To vary arrest strength as shown in Figs. 2 and S3, we varied the values of k_s, k_{cat}, and K_D in the equations for the three interactions. The initial value of each was set at the limit required for cell cycle arrest, and varied to two orders of magnitude in the direction of stronger arrest.

Analysis of individual arrest mechanisms

To obtain a comprehensive picture of how cell cycle arrest could be achieved by different biochemical implementations and cell cycle targets, we inhibited each cell cycle protein individually by each binding, phosphorylation and transcriptional repression. The results of this analysis are shown in FigureB-2. We find that cyclin profiles can be grouped into different types of arrest: mechanism I-like G1 arrest with high G1 cyclins and low G2 cylins; mechanism II-like G2 arrest with high G2 cyclins and low G1 cyclins; mechanism III-like G2 arrest with low G2 cyclins and high G1 cyclins; M-phase arrest characterized by high cyclin B but low cyclins A and E (6); and finally S-phase arrest characterized by high cyclin A but low cyclins E and B. All of the resulting arrests fall into the three classes of G1 and G2 arrest described in the main text, with the exception of interactions leading to arrest in M- or S-phase, which are less prominent in response to DSBs. Inhibiting some species by any of the three mechanisms did not lead to arrest; these species are not pictured. We find that different choices of biochemical implementation frequently lead to qualitatively similar arrest phenotypes, and that inhibiting proteins at similar points in the cell cycle (e.g. cyclins D and E; cyclin A and TFE; APC and Cdc20) also leads to similar arrests profiles. Thus, the cell cycle arrest mechanisms discussed in the main text, implemented via canonical targets and biochemical interactions, are representative of larger subsets of targets and interactions that lead to similar arrest states (e.g. (4, 5)).

Model simulation details

Computing cell cycle transitions

For all freely cycling simulations (Fig. 1B and during each optimization run), the G1/S, S/G2, and M/G1 transitions were identified as the first peak times of cyclin E (CycE), phosphorylated cyclin B (pB) and APCCdc20A (Cdc20A). However, during cell cycle arrest simulations, cyclin trajectories might not peak, or peaks might be significantly lower than during the corresponding free cycling simulation. For arrest

simulations, the G1/S transition was defined as the time at which CycE reached 90% of its maximal free cycling value. Similarly M/G1 transitions were defined as times at which Cdc20A reached 80% of its maximal value. For arrest simulations, we assume that S phase is of fixed duration, equal to the duration in freely cycling cells.

Simulating DNA damage and measuring arrest steady states

We modeled the induction of DNA damage as the presence of an input S, representing DSBs as well as accessory proteins such as the MRN complex assembled at the site of damage. This input was modeled as capable of binding inactive ATM2 dimers and catalyzing their phosphorylation and disassociation. At the time of damage t_D, S was set to a value of 1, while at the time of recovery from arrest t_R, S was returned to a value of 0. The dynamics of p53, Wip1, and Chk2 cause cyclin levels to oscillate with low amplitude during arrest (Fig. 1C-F). In generating Fig. 2, we computed an approximate steady state by averaging cyclin levels from 180-200 h after constant application of the damage stimulus that causes p53 oscillations. We confirmed that the averages were unchanged compared to the preceding interval from 160-180 h, indicating that steady state has been reached.

Simulating populations of cells (Fig. 4, S4 and S5)

The distribution of cell ages in a free cycling population is expected to follow the distribution given by Equation B.8. Thus, for Figs. 4C,D 5A and S8D, 500 cells were simulated with initial ages (and corresponding initial conditions) sampled from this distribution. Damage was induced by applying the damage stimulus to all cells at $t = 0$. To account for damage repair, we implemented a simple model of the stochastic repair of DNA damage (7), parameterized using values found in literature, and measured the distribution of repair times (FigureB-6A). We modeled the initial damage distribution in each cell to be Poisson distributed with a mean of 25 DSBs per Gy (8) and a distribution of 80% easy and 20% difficult breaks. Simple breaks were repaired with a half-life of 15 min, while difficult breaks were repaired with a half-life

of 10 h (9, 10). We assumed cell cycle reentry occurs when fewer than three breaks remain (7). FigureB-6B shows the distribution of repair times from 500 simulations of this stochastic process as computed by independent runs of the Gillespie algorithm (11), compared to a Gaussian distribution with mean 35 h and standard deviation 10 h. Despite presumed completion of DNA repair, $p53^{+/+}$ cells remain arrested for the duration of our experiments. We model this process as sustained activation of p53-dependent arrest mechanisms.

B.2 Analyzing experimental data for fitting

Quantifying the fold-change in cyclin levels during arrest

Two types of flow cytometry measurements were used to train the model to HCT $p53^{+/+}$ and $p53^{-/-}$ data: (i) the fraction of cells in G1, S, and G2, and (ii) the distribution of cyclin B and E levels during G2 arrest. To extract typical cyclin E and B levels reached during arrest, we fit log-normal distributions to the data of Fig. 3 at each time point (FigureB-5). A single log-normal distribution was fit to the unimodal G2 cyclin E levels at each time point. G1 cyclin E levels were bimodal, possibly indicating that early-G1 and late-G1 subpopulations vary in cyclin E levels; for these data, a sum of two log-normal distributions was used. The G2 cyclin B1 data was fit by a sum of three log-normal distributions.

We measured the fold change of cyclin E levels after G2 arrest by computing the ratio of the mean of the log-normal cyclin E distribution in arrested G2 cells to the mean of the high cyclin E distribution in freely cycling G1 cells. Similarly, we computed the fold change in cyclin B1 levels from the means of the most populated G2 cyclin B1 peak during arrest to the mean of the highest peak in freely cycling G2 cells. In each case, the mean intensity of the appropriate isotype control was subtracted from each measurement to remove background fluorescence. To capture the levels attained during each cell line's arrest, we performed this analysis at the 48 h time point for HCT $p53^{+/+}$ cells, and the 24 h time point for HCT $p53^{-/-}$ cells (as

many $p53^{-/-}$ cells have re-entered the cell cycle at later times).

Quantifying the time spent by HCT cells in G1, S, and G2/M

To measure the amount of time spent by HCT $p53^{+/+}$ and $p53^{-/-}$ cells in G1, S, and G2/M, we assumed that unirradiated cell populations at 0 h are completely asynchronous, and that they maintain a stationary distribution across cell cycle phases. However, because cells undergoing mitosis divide into two daughter cells, the observed distribution does not directly reflect the amount of time spent by an individual cell in each cell cycle phase. Instead, young cells will be over-represented relative to older cells. To properly account for this bias towards observing young cells during proliferation, we derived an analytical distribution of cell age in a stationary population (similar derivations can be found in (12)).

We considered a model in which cells progress through N states before dividing, and all transitions between states occur at the rate a. We represented this process as the linear system of ordinary differential equations (ODEs) shown in Equation B.4, where the factor of 2 arises from a parent cell dividing into two daughters.

$$\begin{aligned} \dot{X}_1 &= 2aX_N - aX_1 \\ &\vdots \\ \dot{X}_i &= aX_{i-1} - aX_i \\ &\vdots \\ \dot{X}_N &= aX_{N-1} - aX_N \end{aligned} \tag{B.4}$$

While the total number of cells in such a model diverge, the fractions of cells in each phase, computed by normalizing to the total population $X_T = \sum_{i=1}^{N} X_i$, approach a steady state. We denoted these fractions as state variables $x_i = X_i/X_T$. Taking

derivatives of x_i terms, we defined a new system of nonlinear ODEs (Equation B.5).

$$\begin{aligned}
\dot{x}_1 &= 2ax_N - ax_1 - ax_1x_N \\
&\vdots \\
\dot{x}_i &= ax_{i-1} - ax_i - ax_ix_N \\
&\vdots \\
\dot{x}_N &= ax_{N-1} - ax_N - ax_N^2
\end{aligned} \tag{B.5}$$

This system was analytically solved at steady state to obtain expressions for each species x_i, and resulted in an exponential distribution (Equation B.6).

$$\begin{aligned}
x_i &= x_1 \cdot 2^{-\frac{i-1}{N}} \\
x_1 &= 2\left(1 - 2^{-\frac{1}{N}}\right)
\end{aligned} \tag{B.6}$$

For the linear transitions in the model of Equation B.4, the mean time spent in each state X_i is given by $\tau_i = \frac{1}{a}$, and the total mean cell cycle period is $T = \frac{N}{a}$. From these results we defined the probability distribution $p_N[t]$ of mean cell age in terms of the cell population x_i.

$$\begin{aligned}
p_N\left[t = \tfrac{i}{a}\right] &= x_i \\
p_N[t] &= p_0 \cdot 2^{\frac{1-at}{N}}
\end{aligned} \tag{B.7}$$

In the limit of large N, the distributions of ages in each stage become arbitrarily narrow around each state's mean lifetime, representing a deterministic cell spending a fixed amount of time in each stage. In this limit, the distribution of cell ages becomes continuous in time. We non-dimensionalized this distribution by normalizing the age variable to the total period (i.e. defining $\tau = \frac{a}{N}t$), so that the probability distribution of cell ages in the free cycling population can be written as shown in Equation B.8.

$$\begin{aligned}
p_\infty(\tau) &= 2\log 2 \cdot 2^{-\tau} \\
\tau &\in [0, 1]
\end{aligned} \tag{B.8}$$

From this distribution and the relationships of Equation B.8, we solved for the time spent in G1, S and G2/M by an individual cell given any observed distribution of

cells in these phases.

$$\begin{aligned} \%\mathrm{G1} &= \int_0^{\tau_1} p(\tau)\, d\tau = 2\left(1 - 2^{-\tau_1}\right) \\ \%\mathrm{G2/M} &= \int_{1-\tau_2}^{1} p(\tau)\, d\tau = 2^{\tau_2} - 1 \end{aligned} \tag{B.9}$$

We obtained the fraction of cells in each cell cycle phase from the measured DNA profiles by fitting distributions to a modified Dean-Jett model (13) (see Methods), and computing the number of cells in each fit population. The measured population data, and its conversion to the time spent in each phase, is reproduced in Figure S3C.

B.3 Fitting the model to data

Mathematical notation for fitting and sensitivities

Our model takes the form of a system of coupled ODEs that depends on initial conditions $\mathbf{x_0}$, the time t, the current value of all state variables $\mathbf{x}$, parameters $\mathbf{p}$, and inputs $\mathbf{u}$.

$$\begin{aligned} \dot{\mathbf{x}} &= \mathbf{f}(t, \mathbf{x}, \mathbf{p}, \mathbf{u}) \\ \mathbf{x}(0) &= \mathbf{x}_0 \end{aligned} \tag{B.10}$$

We define a model output of interest y (i.e., the sum of all cyclin E or B species, active cyclin E or B, or Cdc20A) as some linear combination of state variables, formed by multiplication of the row vector $\mathbf{c_y^T}$ with the state vector $\mathbf{x}$. These outputs consist of total Cyclin E (`CycE` + `TriE` + `TriE21`) to compare to flow cytometry data; total Cyclin B (`CycB` + `pB` + `BCKI` + `pBCKI`) to compare to flow cytometry data; active Cyclin E (`CycE`) to identify the G1/S transition; phosphorylated Cyclin B (`pB`) to identify the S/G2 transition; and APC$^{\text{(}}$Cdc20A) (`Cdc20A`) to identify mitosis.

$$y = \mathbf{c}_y^T \mathbf{x} \tag{B.11}$$

Fitting procedure details

We performed a local optimization procedure on the cell cycle arrest model. The aim of this procedure was to find a set of parameters that would allow our fitted model to simultaneously match: (i) the percentage of time spent in each cell cycle phase by undamaged, freely cycling HCT $p53^{+/+}$ and $p53^{-/-}$ cells, (ii) the ratio of steady state cyclin E and cyclin B1 arrest levels to their maximal freely cycling levels, (iii) the requirement for cyclin $E^{-/-}$ and cyclin $D^{-/-}$ cells to still cycle (14, 15), and (iv) the existence of a restriction point after serum starvation and cycloheximide treatment (2). In our fitting procedure, all parameters in the cell cycle model and arrest mechanisms I-III were allowed to vary.

For fitting, we treated p53 and Chk2 as constant-level inputs to the cell cycle model. $p53^{+/+}$ cells were assumed to arrest with a combination of arrest mechanisms I, IIa, IIIa, and IIIb, while $p53^{-/-}$ cells were assumed to arrest by mechanism IIa alone. For the cell cycle model to account for differences in the length of cell cycle phases between cell lines, we assumed that a basal level of p53 activation is responsible for some p21 activation in unirradiated $p53^{+/+}$ cells. This might be due to low levels of damage incurred during the normal cell cycle, or to p53's involvement in cell cycle checkpoints in undamaged, freely cycling cells. All other parameters were shared by both cell lines. Because the total time scale of all simulations was left unconstrained by data, this degree of freedom in parameter space was removed by assuming a total cell cycle period of 30 h for wild-type ($p53^{+/+}$) cells. This constraint was applied by scaling all parameters that include units of time, and otherwise left all trajectories unchanged.

At each parameterization during optimization, we ran eight independent simulations: 3 divisions cycles of the $p53^{-/-}$ model; a 200 h simulation of cell cycle arrest in the $p53^{-/-}$ model; 3 division cycles of the wild-type model; a 200 h simulation of arrest in the wild-type model; three division cycles of the cyclin $D^{-/-}$ model; three division cycles of the cyclin $E^{-/-}$ model; 200 h of arrest after cycloheximide treatment at 3 h after the first division; and 200 h of arrest after cycloheximide treatment at 20 h after

the first division. These simulations allow comparison to restriction point and cyclin knockout results, as well as the free cycling and steady state data collected for HCT p53$^{+/+}$ and p53$^{-/-}$ cells. We computed various quantities from these simulations. $\tau_{G1}^{(i)}$ and $\tau_{G2}^{(i)}$ represent the fraction of the cell cycle spent in G1 ($2N$ DNA content) and G2/M ($4N$ DNA content). Because a normal cell cycle is split only between G1, S, and G2/M, the fraction of time spent in S is completely specified by these two times. The $r_y^{(i,j)}$ terms represent the ratios of model outputs during arrest to those attained during normal cell cycling, and is computed as shown in Equation B.12.

$$r_y^{(i,j)} = \frac{\mathbf{c}_y^T \cdot \mathbf{x}^{(i)}(t_F)}{\mathbf{c}_y^T \cdot \mathbf{x}^{(j)}(t_y^*)} \tag{B.12}$$

$T^{(i)}$ is the total cell cycle period, and is computed by measuring the time between mitoses. Finally, the degree to which steady state is achieved by the final simulation time is measured by the norm of the time derivative at this time, $\left\|\dot{\mathbf{x}}^{(i)}(t_F)\right\|^2$. To measure the deviation of these quantities from their experimental values, we evaluated an objective function, shown in Equation B.13. Each line of the objective function represents terms that are computed after the corresponding simulation.

$$\begin{aligned}
O(\mathbf{p}) =& 100 \cdot \left(\left(\tau_{G1}^{(1)} - \bar{\tau}_{G1}^{(1)} \right)^2 + \left(\tau_{G2}^{(1)} - \bar{\tau}_{G2}^{(1)} \right)^2 \right) \\
&+ \left(r_{CycE_T}^{(2,1)} - \bar{r}_{CycE_T}^{(2,1)} \right)^2 + \left(r_{CycB_T}^{(2,1)} - \bar{r}_{CycB_T}^{(2,1)} \right)^2 + 10^3 \left\| \dot{\mathbf{x}}_{ss}^{(2)} \right\|^2 \\
&+ 100 \cdot \left(\left(\tau_{G1}^{(3)} - \bar{\tau}_{G1}^{(3)} \right)^2 + \left(\tau_{G2}^{(3)} - \bar{\tau}_{G2}^{(3)} \right)^2 \right) \\
&+ \left(r_{CycE_T}^{(4,3)} - \bar{r}_{CycE_T}^{(4,3)} \right)^2 + \left(r_{CycB_T}^{(4,3)} - \bar{r}_{CycB_T}^{(4,3)} \right)^2 + 10^3 \left\| \dot{\mathbf{x}}_{ss}^{(4)} \right\|^2 \\
&+ 5 \cdot \theta \left(T^{(5)}, T^{(1)} \right) \\
&+ 5 \cdot \theta \left(T^{(6)}, T^{(1)} \right) \\
&+ \theta \left(0.5, r_{CycE_T}^{(7,3)} \right) + 10^3 \left\| \dot{\mathbf{x}}_{ss}^{(7)} \right\|^2 \\
&+ \theta \left(0.5, r_{CycE_T}^{(8,3)} \right) + \theta \left(r_{CycE}^{(8,3)}, 0.5 \right) + \theta \left(0.9, r_{Cdc20A}^{(8,3)} \right) + 10^3 \left\| \dot{\mathbf{x}}_{ss}^{(8)} \right\|^2
\end{aligned} \tag{B.13}$$

This objective takes the form of a weighted sum of squared errors, where individual terms are scaled so that their contributions are of comparable magnitude. Terms containing the function $\theta(a, b)$, defined in Equation B.14, only penalize for deviation in one direction.

$$\theta(a, b) = \begin{cases} 0a < b \\ (a - b)^2 a \geq b \end{cases} \tag{B.14}$$

For instance, the term $5 \cdot \theta\left(T^{(5)}, T^{(1)}\right)$ only contributes to the objective function if the period of the cyclin $D^{-/-}$ simulation is longer than that of the normal cell cycle. If computing this objective function was impossible at any parameter set during optimization (for example, if arrested cells failed to arrest at all), the objective function value was set to an arbitrary high number. After fitting, model parameters had changed from the initial parameterization by no more than roughly an order of magnitude above and below between the initial model parameterization, with no parameters reaching upper or lower bounds (FigureB-8A).

Efficient timing sensitivities for optimization

An adjoint method was used to compute the objective function's parameter sensitivities in a computationally efficient manner (16). For this method to be applied, the partial derivatives $\frac{\partial \theta}{\partial \mathbf{x}}(t)$ and $\frac{\partial \theta}{\partial \mathbf{p}}(t)$ must be computed. Because many terms of the objective function involve the times at which the sum of certain species is maximal, we solved for the sensitivities to parameters of these times. We define the times of maxima, t_y^*, as times at which the slope of output y is zero, and its second derivative is negative. Mathematically, this is written by the two conditions of Equation B.15.

$$\begin{aligned} \mathbf{c}_y^T \cdot \mathbf{f}\left(t_y^*, \mathbf{x}, \mathbf{p}\right) &= 0 \\ \mathbf{c}_y^T \cdot \nabla_{\mathbf{x}} \mathbf{f}\left(t_y^*, \mathbf{x}, \mathbf{p}\right) \cdot \mathbf{c}_y &< 0 \end{aligned} \tag{B.15}$$

Differentiating Equation B.15 and using the chain rule, we derived an expression for the sensitivity of the timing of the i^{th} species' extremum (Equation B.16; a similar

result is found in Equation B.18 of Rand *et al.* (17)).

$$\frac{dt_y^*}{d\mathbf{p}} = -\frac{\mathbf{c}_y^T \cdot \nabla_{\mathbf{x}} \mathbf{f}\left(t_y^*, \mathbf{x}, \mathbf{p}\right)}{\left(\mathbf{c}_y^T \cdot \nabla_{\mathbf{x}} \mathbf{f}\left(t_y^*, \mathbf{x}, \mathbf{p}\right)\right) \cdot \mathbf{f}\left(t_y^*, \mathbf{x}, \mathbf{p}\right)} \frac{\partial \mathbf{x}}{\partial \mathbf{p}} - \frac{\mathbf{c}_y^T \cdot \nabla_{\mathbf{p}} \mathbf{f}\left(t_y^*, \mathbf{x}, \mathbf{p}\right)}{\left(\mathbf{c}_y^T \cdot \nabla_{\mathbf{x}} \mathbf{f}\left(t_y^*, \mathbf{x}, \mathbf{p}\right)\right) \cdot \mathbf{f}\left(t_y^*, \mathbf{x}, \mathbf{p}\right)} \tag{B.16}$$

Notably, terms in this equation can be grouped into the form of Equation B.17. This form permits the application of the adjoint sensitivity method without ever computing $\frac{\partial \mathbf{x}}{\partial \mathbf{p}}$, by providing the partial derivatives $\frac{\partial t_i^*}{\partial \mathbf{x}}$ and $\frac{\partial t_i^*}{\partial \mathbf{p}}$.

$$\frac{dt_y^*}{d\mathbf{p}} = \frac{\partial t_y^*}{\partial \mathbf{x}} \frac{\partial \mathbf{x}}{\partial \mathbf{p}} + \frac{\partial t_y^*}{\partial \mathbf{p}} \tag{B.17}$$

B.4 Fitted model validation

We varied individual parameters in our fitted model to identify parameters that, when varied individually, remained poorly constrained by our data and optimization procedure. We also investigated the fitted model's ability to predict experimental results, including the population-averaged levels of cell cycle species after irradiation and the results of various knockout experiments in freely cycling and arrested scenarios.

Sensitivity analysis of fitted model

To address whether or not the fitting procedure described above is able to constrain the value of individual parameters, we computed the second derivatives of the objective function with respect to each parameter by taking finite differences of the adjoint sensitivity (described above) and used these values to estimate the curvature of the local minimum obtained by optimization. This procedure does not account for any additional parameter identification provided by constraints satisfied at the fit (e.g. the requirement of cyclin knockout cells to cycle). To map these entries to actual changes in parameters, we used the curvature to estimate the change in each parameter required to increase the objective function by a value of 0.1 from the

minimum (Equation B.18).

$$O(p) \approx O(p^*) + \frac{1}{2} \frac{d^2O}{d(\log p)^2}\bigg|_{p^*} \left(\log \frac{p}{p^*}\right)^2$$

$$\frac{p}{p^*} = \exp\left(\sqrt{2\,(0.1)\left(\frac{d^2O}{d(\log p)^2}\right)^{-1}}\right) \tag{B.18}$$

The results of this analysis are shown in FigureB-8B. We find that many parameters, when varied individually, strongly affect the objective function value. Eleven parameters are sufficiently unconstrained that they have to vary at least twofold to achieve a change in objective function value of 0.1 (Table B.1). Many of these parameters are association rates between cyclins and either p21 or p27; binding of these species might be sufficiently tight such that the precise value of this association rate is unconstrained. In addition to the parameters pictured in FigureB-8B, this analysis correctly identified five parameters (`kifi_p`, `kifib_pp`, `Jafi`, `kafi`, `Jifi`) that are infinitely unconstrained. These are all parameters affecting `TFI`, a species that is disconnected from the rest of the cell cycle network in Tyson's mammalian cell parameterization.

Western blot predictions

To further validate the model fitted to the cell cycle timing and steady state arrest data, we simulated populations of $p53^{+/+}$ and $p53^{-/-}$ models initially distributed as a freely cycling population (see previous section) and computationally predicted the levels of different cell cycle proteins at times after the application of damage (FigureB-8C). These results were compared with Western blots of total protein levels from HCT $p53^{+/+}$ and $p53^{-/-}$ cell lines (FigureB-6C). We find similar qualitative trends between prediction and experiment, although some changes in protein levels are predicted to occur earlier than seen experimentally (e.g. Cyclin A and B dynamics in HCT $p53^{-/-}$ cells; Cdc20 dynamics in HCT $p53^{+/+}$ cells).

Knockout cell line predictions

We also validated the model by predicting the results of experiments in which various cell cycle proteins might be targeted for silencing or knockout. As described earlier, the ability for cyclin D and E knockout cell lines to continue cycling was considered in Tyson's original model and used as part of the fitting process. To test the consequences of other deletion experiments, we computationally eliminated each cell cycle protein individually for low (M = 2) and high (M = 16) mitogen levels, and compared the simulation results to published experimental results obtained in mammalian systems. The results are shown in Table B.2. We find that the model accurately reproduces most known phenotypes, with only two notable exceptions. Cyclin $A^{-/-}$ cells are predicted to cycle at both low and high mitogen levels, but experimentally cyclin A is found to be essential. This might be due to unmodeled functions of cyclin A in DNA replication, such as a basic requirement for initiating replication. The model predicts that $Wee1^{-/-}$ cells can still undergo cycling, although Wee1 deletion is embryonic lethal. However, this lethality is due to premature mitosis and subsequent cell death, indicating that cells lacking Wee1 continue cycling, albeit too rapidly. We also tested whether the model predicts normal arrest for knockout cell lines that are still able to cycle. All knockout cell lines are predicted to maintain normal arrest regulation except for the $Cdh1^{-/-}$ model, which arrests transiently by mechanism II but resumes cycling after repair, even in the continued presence of mechanisms I and III. This prediction is consistent with evidence that $Cdh1^{-/-}$ cells are able to cycle in the absence of damage, and are defective in maintaining cell cycle arrest after IR (18).

Parameter name	Biochemical role	Parameter change
kass21e	p21-cyclin E association	$> 10^{10}$
kasse	p27-cyclin E association	$> 10^{10}$
kass21d	p21-cyclin D association	$> 10^{10}$
kassa	p27-cyclin A association	$> 10^{10}$
kiwee_pp	cyclin B-mediated Wee1 inactivation	9.93×10^2
kdie_p	cyclin E-mediated p27 degradation	4.49×10^2
kdisse	p21-cyclin E dissociation	4.53
kdib_pp	cyclin B-mediated p27 degradation	3.02
kdia_pp	cyclin A-mediated p27 degradation	2.92
Kp53a	p53 repression of cyclin A	2.80
J20	Cyclin B-mediated Cdc20 phosphorylation	2.06

Table B-9: Model parameters expected to vary more than twofold for an objective function change of 0.1 (using Equation B.18). The parameter name, its biochemical role, and the parameter variation calculated to change the objective function by 0.1 are shown.

Condition	Cycles (predicted)	Cycles (experiment)	Arrests
$CycD^{-/-}$	Yes	Yes [106]	Arrests normally
$CycE^{-/-}$	Yes	Yes [33]	Arrests normally
$CycA^{-/-}$	Yes	No [111]	Arrests normally
$CycB^{-/-}$	No	No [112]	–
$CKI^{-/-}$	Yes	Yes [113]	Arrests normally
$Cdc20^{-/-}$	No	No [38]	–
$Cdh1^{-/-}$	Yes	Yes [109]	Continues cycling [109]
$Wee1^{-/-}$	Yes	No* [39]	Arrests normally
$Cdc25^{-/-}$	No	No [34]	–
$TFE^{-/-}$	No	No [35]	–
$TFB^{-/-}$	No	–	–
$APC^{-/-}$	No	No [36]	–

Table B-9: The predicted effect of knocking out any modeled cell cycle proteins, as well as the predicted effect under arresting conditions for those knockouts expected to cycle. For knockouts with a reported literature phenotype, the experimental result and reference are provided. Dashed entries indicate that non-cycling mutants cannot be further tested under arresting conditions.

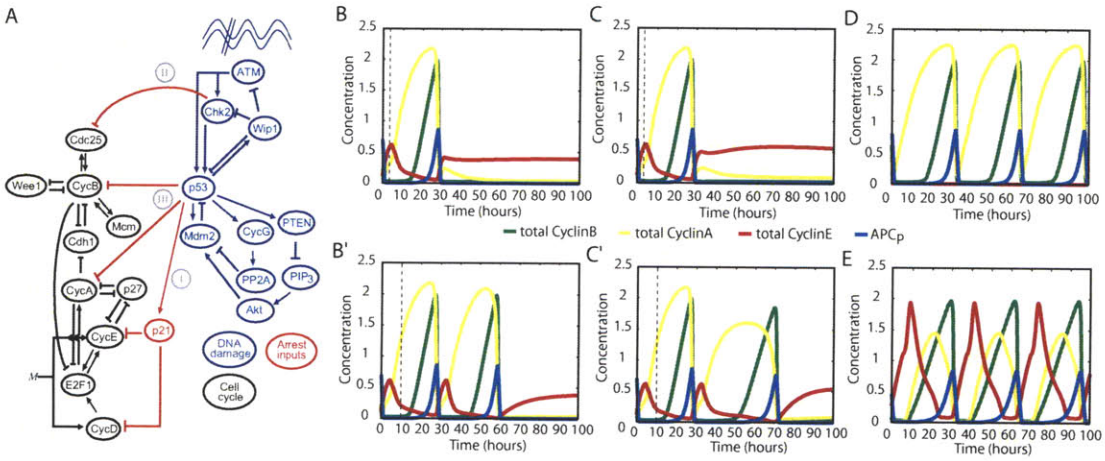


Figure B-1: (A) Diagram of the cell cycle and DNA damage models. Arrest mechanisms are labeled in gray. Mechanism I indicates the p53-dependent G1 arrest induced by p21. Mechanism II indicates Cdc25 phosphorylation and inactivation by Chk2 and its subsequent inhibition by 14-3-3 proteins. Mechanism III models the p53-dependent transcriptional repression of cyclin B and cyclin A. Additional feedback loops involving p53 are shown, but do not significantly affect the network's dynamics in their current parameterization. (B-E) Matching experiments form the literature. Trajectories from the initial model are plotted for all simulations, showing levels of total cyclins B, A, and E and phosphorylated anaphase promoting complex (APC_P) over 100 h. (B) Serum starvation applied 3 h (dotted line) after cell division leads to immediate G1 arrest. (B') Serum starvation applied 10 h (dotted line) after cell division leads to a G1 arrest after the completion of one full cell division. (C,C') Cycloheximide treatment applied 3 and 10 h after cell division have phenotypes similar to those achieved by serum starvation. (D) The cyclin $E^{-/-}$ model, simulated with a mitogen level $M = 2$, is able to cycle. (E) Similarly, the cyclin $D^{-/-}$ model cycles with $M = 10$.

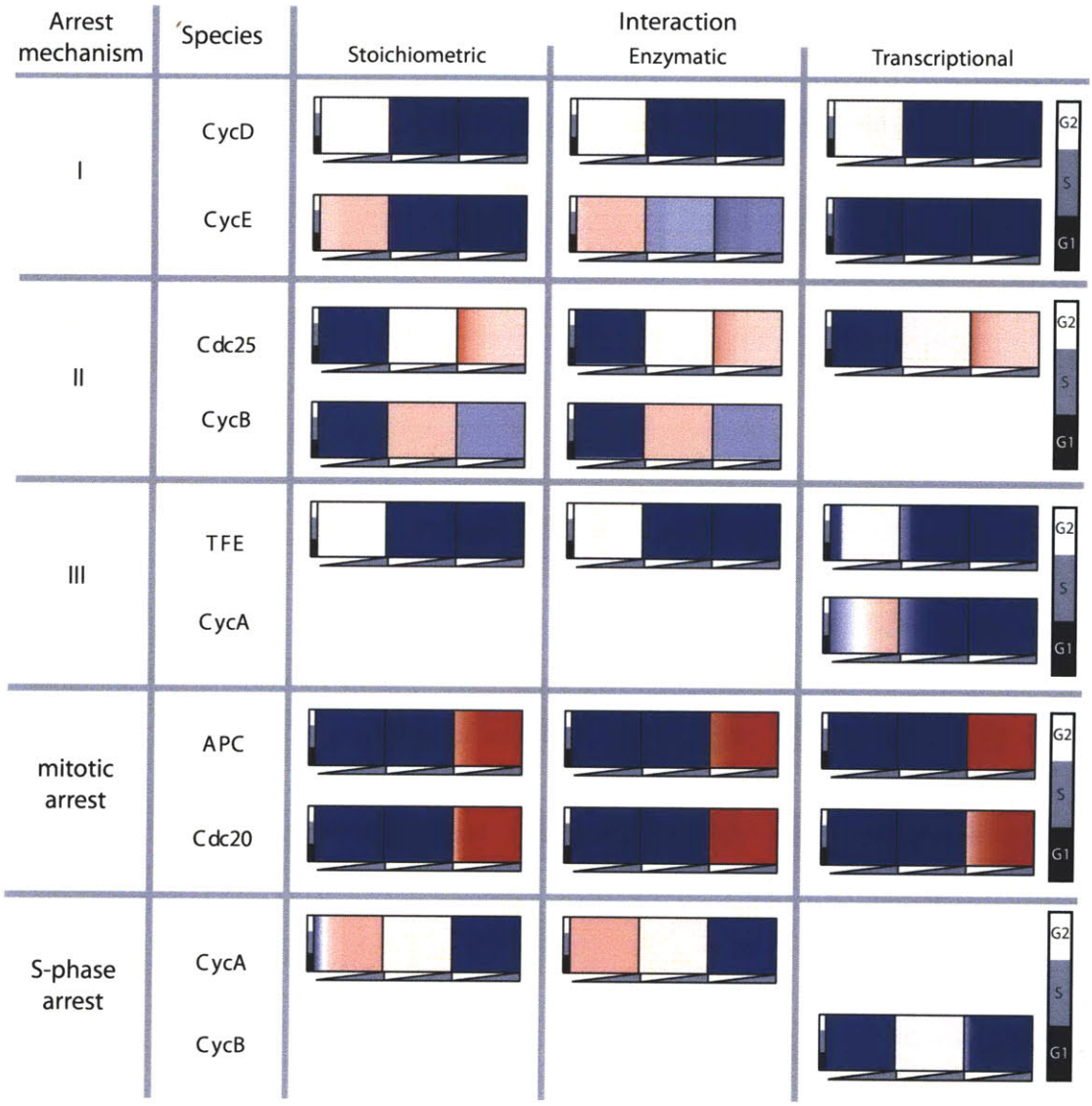


Figure B-2: Each group of three boxes is colored as described for Figure 3-2. Inhibition of each cell cycle protein is applied by binding, enzymatic modification, or transcriptional repression by an inhibitor. The resulting arrest states are grouped into similar mechanisms. (i) Mechanism I-like arrests result in a G1 DNA content state with high levels of G1 cyclins and low levels of G2 cyclins. (ii) Mechanism II-like G2 arrests lead to a high G2 cyclin and low G1 cyclin state. (iii) Mechanism III-like G2 arrests result in a high G1 cyclin and low G2 cyclin state. (iv) The model can also arrest in a mitotic arrest state by inhibition of APC or Cdc20, in which high active cyclin B levels maintain a mitotic state. (v) An S-phase arrest state consisting of high cyclin A levels but low cyclin B levels can arise from cyclin B repression or cyclin A inhibition.

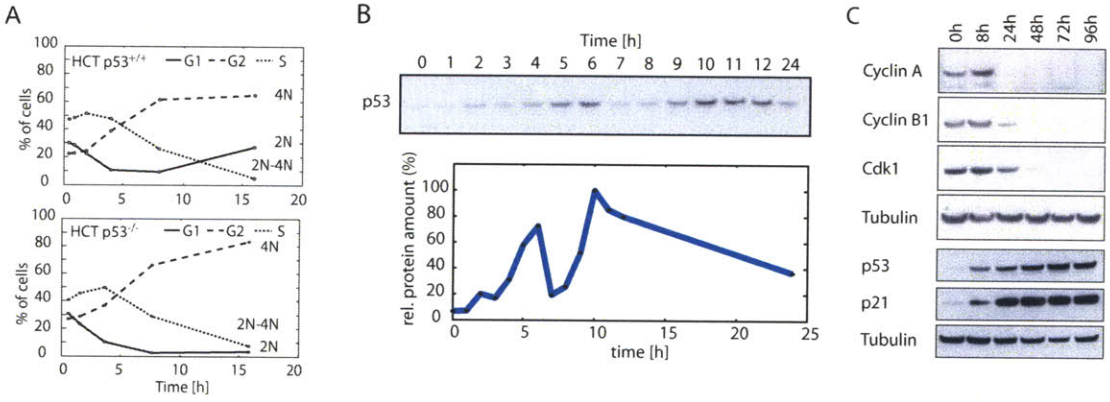


Figure B-3: (A) Time courses of DNA content after treatment with 10 Gy IR. HCT $p53^{+/+}$ and HCT $p53^{-/-}$ cells were irradiated at 0 h, and the fractions of cells with G1, S, and G2 DNA contents were measured by FACS at the indicated time points. (B) Timecourse of p53 levels after DNA damage. HCT $p53^{+/+}$ cells were irradiated with 10 Gy at 0 h. Samples were taken every hour and analyzed by western blot. Densitometric quantification of the western blot, normalized to the maximum p53 level, is shown in the graph below. p53 begins to accumulate after 2 h, and shows two peaks (6 h and 10 h) in a 12 h period. Similar results were obtained in (19) (C) Protein levels of cell cycle regulators during the arrest. HCT $p53^{+/+}$ cells were irradiated with 10 Gy and samples were taken at the indicated time points. Protein levels were analyzed by western blot, with tubulin as a loading control. Cyclin A, Cyclin B1, and Cdk1 are downregulated within the first 24 h of arrest. Concomitantly, p53 and p21 levels increase and stay high until at least 96 h.

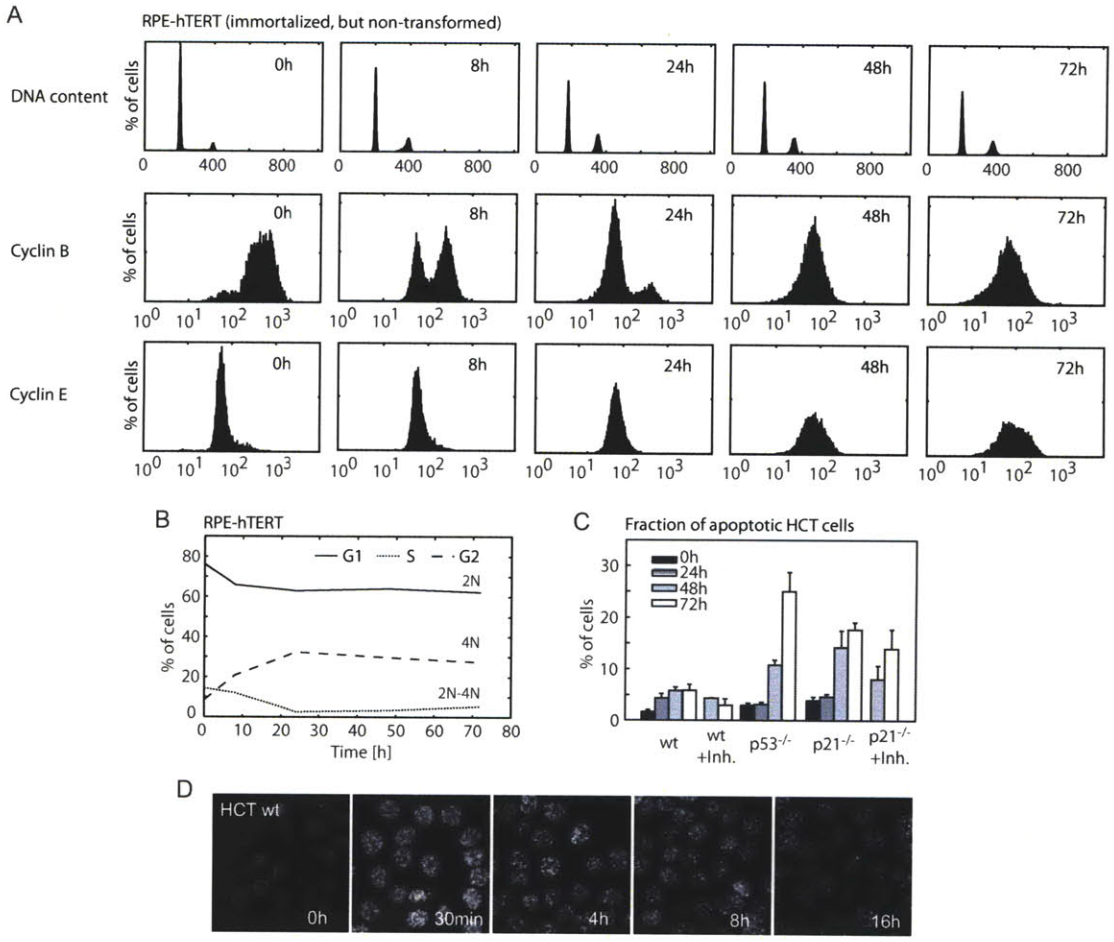


Figure B-4: (A) Flow cytometry histograms of DNA content and cyclin levels in untransformed RPE-hTERT cells after IR (10 Gy). Cells were irradiated at 0 h and stained for DNA content and cyclin levels for 72 h post-irradiation. Histograms of cyclin levels are gated from the $4N$ population only. Cyclin B downregulation was similar to that observed in HCT116 G2 arrested cells. Cyclin E was upregulated by 2- to 3-fold in G2 cells between 0 h and 72 h, a range similar to the 2-fold increase in cyclin E levels observed in freely cycling G1 RPE-hTERT cells (data not shown). (B) Quantification of the fractions of RPE-hTERT cells with G1 ($2N$), S ($2N$-$4N$), and G2 ($4N$) DNA contents as measured by FACS. These cells showed slower cell cycle and arrest dynamics and a changed ratio of G1 versus G2 arrested cells compared to HCT116. (C) Fraction of apoptotic subG1 cells in HCT wt, $p53^{-/-}$ and $p21^{-/-}$ cell lines (mean SE) over time after irradiation with 10 Gy (see Fig. 3, 4 and 5). For HCT wt and p21-/-, the fraction of apoptotic cells in the presence of the Cdk1 inhibitor RO-3306 is shown as well (+Inh., see Fig. 5). (D) Immunofluorescent staining of γH2AX in HCT116 wt cells after 10 Gy γ-irradiation. The intensity of γH2AX and the number of foci representing individual DSBs is maximal 30min post damage. At later time points, DNA repair is evident by a reduction in the number of foci, but cells retain a significant amount of damage throughout the 16h time course.

Supplementary Figure S3

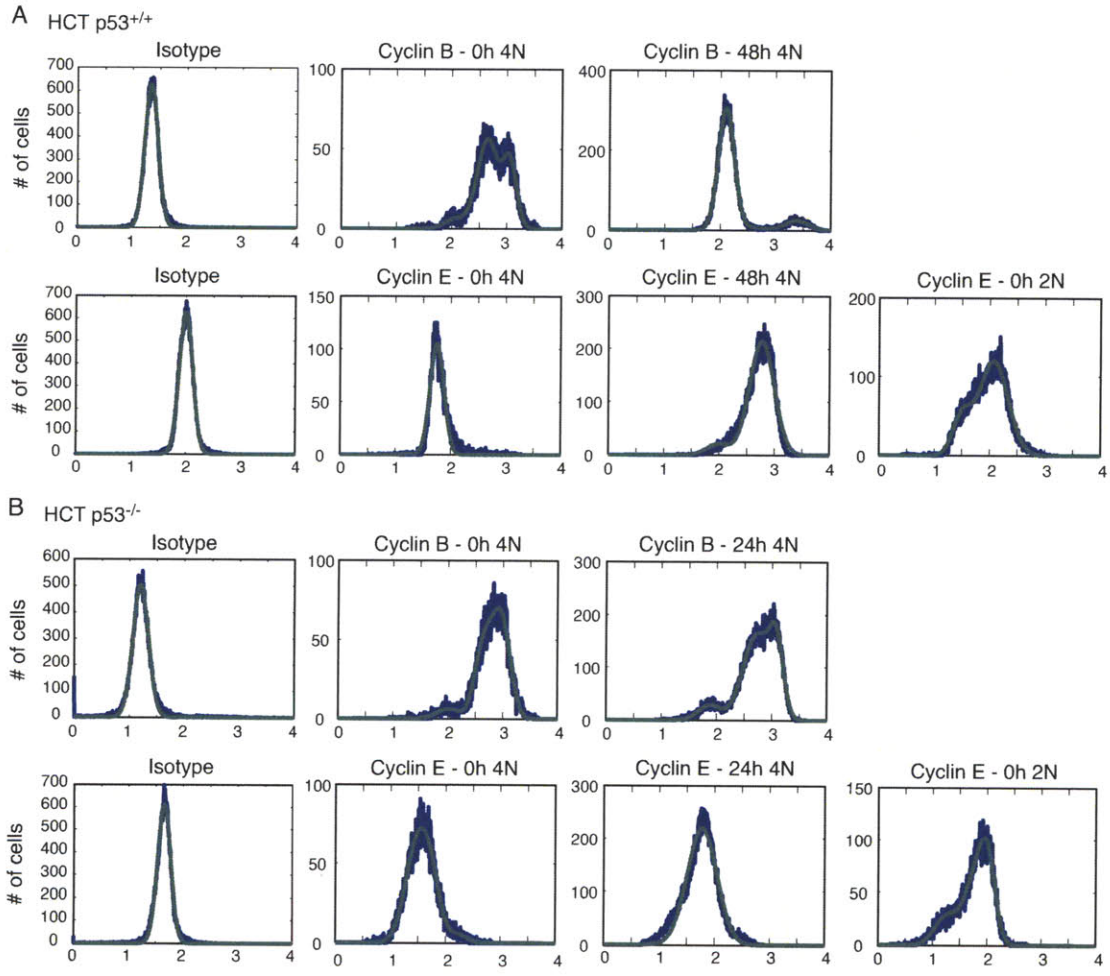
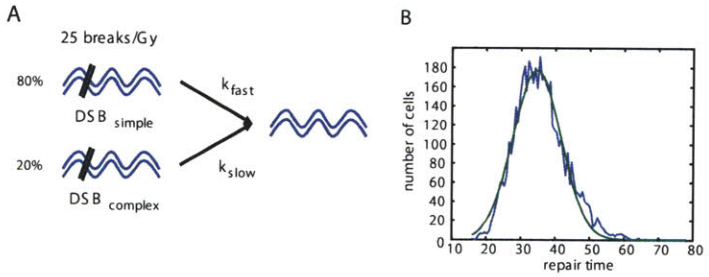


Figure B-5: (A) Distribution of cyclin B and E levels in HCT $p53^{+/+}$ cells. Cyclin B levels in cells with $4N$ DNA content showed three peaks, probably representing G2 cells (intermediate), mitotic cells (high) and cells after metaphase, but before cytokinesis (low). The sum of three Gaussian distributions (green) was fitted to the measured data (blue). Cyclin E levels in cells with $4N$ DNA content were unimodal. In cells with $2N$ DNA content, they were bimodal, probably representing early (low) and late G1 cells (high). The measured distributions were fitted with single Gaussian and the sum of two Gaussian distributions, respectively. The distribution of the respective isotype controls were fitted with single Gaussian distributions. (B) Distributions of cyclin B and cyclin E levels in HCT $p53^{-/-}$ cells. The measured data was fitted as described above.



Percent cells in cell cycle phase		
	HCT p53$^{+/+}$	HCT p53$^{-/-}$
g1	44.8%	26.1%
s	39.6%	54.0%
g2	15.6%	19.9%
Time spend in cell cycle phase		
τ_{G1}	36.6%	20.2%
τ_S	42.5%	53.7%
τ_{G2}	20.9%	26.1%

Figure B-6: (A) Diagram of the stochastic DNA double-stranded break (DSB) repair model. See Suppl. Text for details. (B) Distribution of repair times from the model (blue curve), compared to a normal distribution with mean 35 h and standard deviation 10 h (green curve). Two types of DSBs - easy and difficult to repair - were modeled, with an initial distribution of 80% and 20% of the total damage, respectively. Easy breaks were repaired with a half-life of 15 min, while difficult breaks were repaired with a half-life of 10 h. The initial number of breaks was distributed according to a Poisson distribution with a mean of 25 DSBs/Gy IR, and repair was assumed to occur when fewer than three breaks remained. (C) The difference between the fraction of cells in G1, S, and G2, and the mean time spent by each cell in these phases. Results for both wild-type and p53$^{-/-}$ cells are shown. The columns labeled $g1$, s, and $g2$ indicate the fractions of cells in each cell cycle phase as determined by flow cytometry analysis. The columns labeled τ_{G1}, τ_S, and τ_{G2} indicate the mean time spent in each cell cycle phase, as computed from Equation 9.

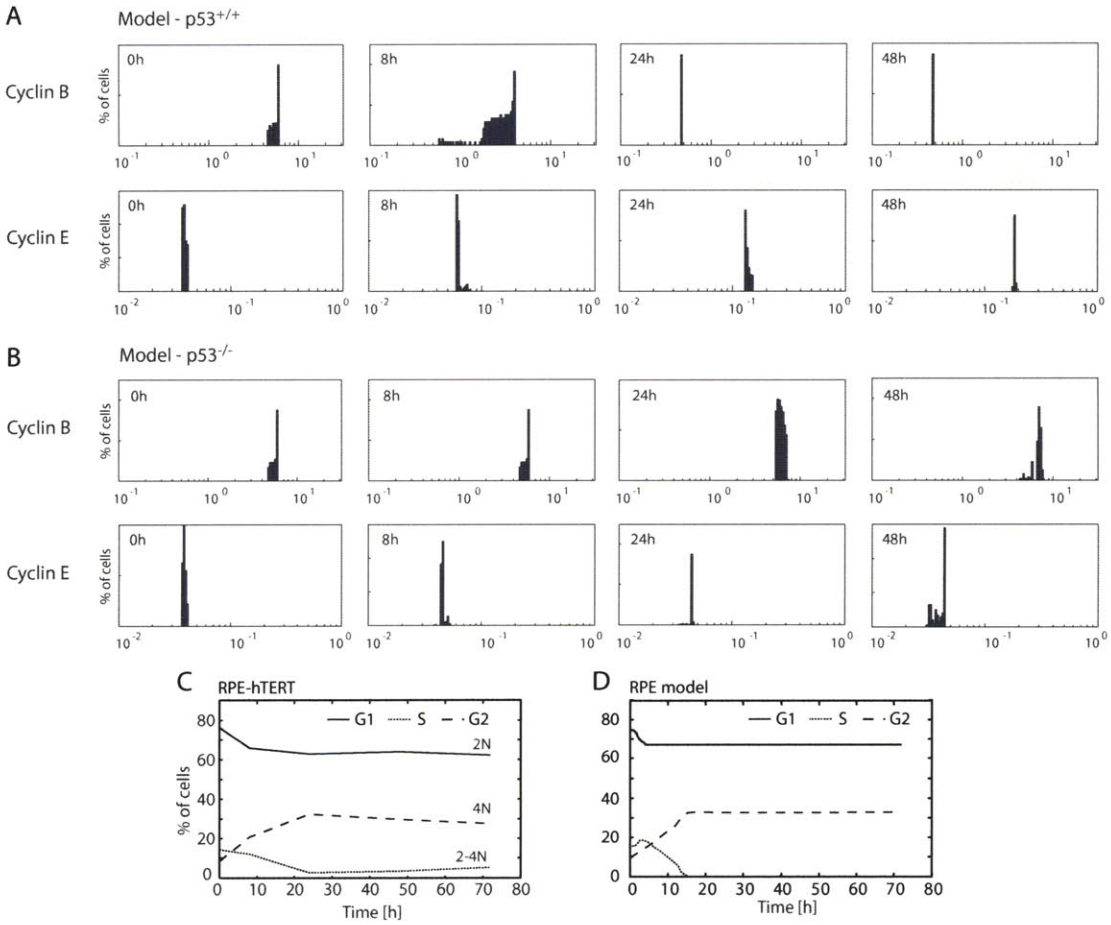


Figure B-7: (A-B) Cyclin distributions from fitted models. The same ensemble of models used to generate the population data of Figs. 4C,D was queried for cyclin levels during arrest. (A) Histograms of total cyclin B and E levels are shown at 0, 8, 24, and 48 h after IR for the $p53^{+/+}$ model. The log species concentration is plotted as the independent variable. Cyclin B levels decrease quickly, while cyclin E levels rise gradually by 48 h. (B) Corresponding histograms of total cyclin B and E levels in the $p53^{-/-}$ model. Levels of both species are largely maintained during arrest, with an increase in variability as cells re-enter the cell cycle. (C) Time courses of DNA content after treatment with 10 Gy IR. RPE-hTERT cells were irradiated at 0 h, and the fractions of cells with G1, S, and G2 DNA contents were measured by FACS. The same graph is shown in Suppl. FigureB-4B. (D) Model-generated cell cycle distribution time courses. The model was adjusted to fit the different cell cycle times of RPE-hTERT cells. $5x10^2$ individual model trajectories were simulated from initial conditions distributed through the cell cycle (see Suppl. Text). In each simulation, p53 was induced 4 h after damage was applied. The DNA repair time was selected from the distribution described in Suppl. FigureB-6B (see Supp. Text).

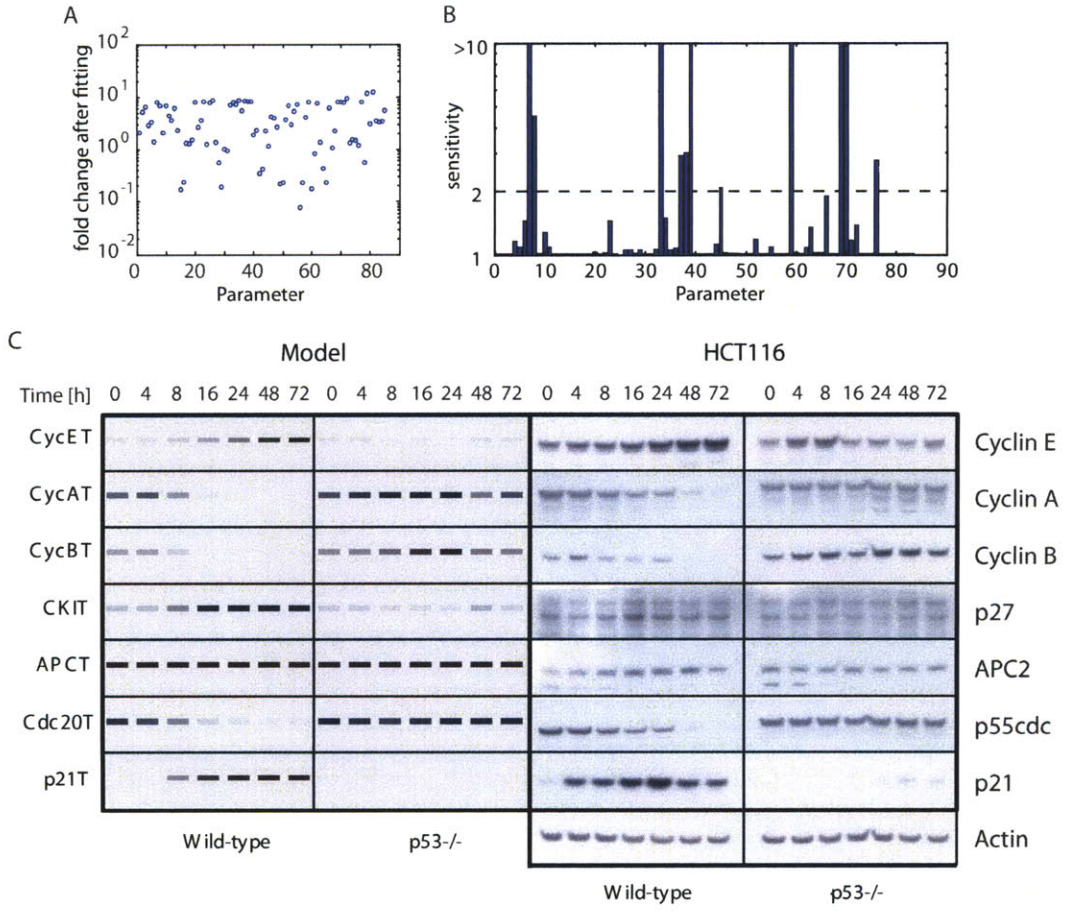


Figure B-8: (A) Influence of optimization on parameters. The indexed number of each fitted parameter in the cell cycle arrest model is plotted against the fold-change in that parameter between the initial and final parameterization. To compute this change in parameterization, both initial and final parameter sets were scaled such that the period of the $p53^{+/+}$ model was 30 h. (B) Parameters constrained by optimization to the objective function. The plot indicates the fold change in parameter value required to increase the objective function value by 0.1 for each parameter considered in optimizing the cell cycle arrest model. Bars corresponding to parameters that vary more than tenfold are cut off at this value. The horizontal dotted line indicates a twofold parameter change. Parameters that are unconstrained to more than a twofold change are listed in Table B.1. (C) Predicted protein levels in fitted $p53^{+/+}$ and $p53^{-/-}$ models. Total concentrations of cyclins A, E, and B, p27, APC, Cdc20, and p21 were measured from a simulated population of 500 cells after induction of damage. Intensity varies linearly with the concentration of each species, and both cell lines are plotted on the same scale. The same species' total protein levels were measured in HCT $p53^{+/+}$ and $p53^{-/-}$ cells by western blot.

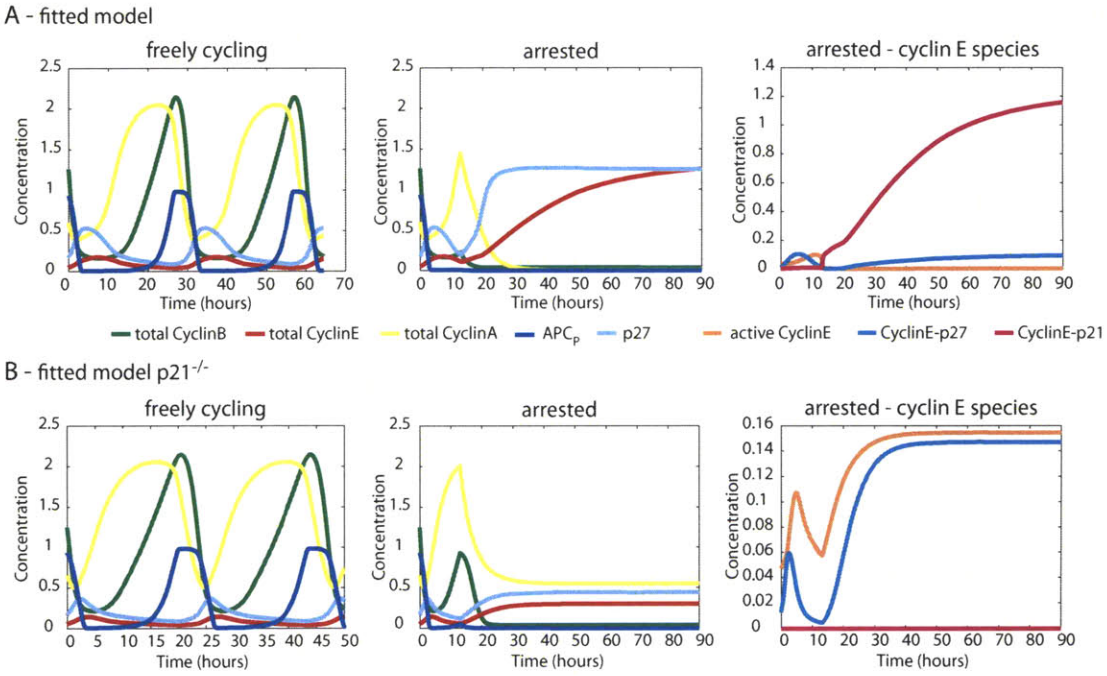


Figure B-9: (A) Cell cycle simulation from the fitted model showing cyclins E, A, and B, and phosphorylated anaphase promoting complex (APC_P). In the absence of DNA damage, the fitted model shows qualitatively similar behavior as the initial model, with consecutive peaks of cyclin E, cyclin A, cyclin B, and APC_P. When DNA damage is applied during G2 phase (13 h), the model arrests with low levels of cyclins A and B and high levels of total Cyclin E. The level of active Cyclin E is low, as it is bound by p21 (and to a lesser extent by p27). (B) Cell cycle simulation in the absence of p21. Without damage, the $p21^{-/-}$ model behaves similarly to the full model. When DNA damage is applied during G2 phase (13 h), the model arrests with low cyclin B levels, but intermediate cyclin A levels. Cyclin E levels are higher then in freely cycling cells, but lower than during arrest in the full model. In contrast to the full model, cyclin E is active and reaches levels higher than during the normal G1/S peak, as inhibition by p27 alone is insufficient. Therefore, cells are predicted to enter S phase without going through mitosis (which would be indicated by a pulse of APC_P), leading to endoreduplication.

Appendix C

Supporting materials and methods for Chapter 5

C.1 Mass action model derivation

In this work, we employed a mass action model representing the interaction between KaiC, KaiA and KaiB, and the phosphoconversion of KaiC. To ensure that this model recapitulates the findings of the abstracted model of Rust *et al.*, we set out to ensure that in the quasi-steady state limit, simulations of the mass action model are identical to those of the original model. The original model represents the concentration of free KaiA as

$$A = \max\{0, S - 2A_T\} \tag{C.1}$$

where A_T is the total concentration of KaiA, and S is the concentration of S–KaiC. This suggests a mass action mechanism in which each molecule of S–KaiC tightly and quickly binds two molecules of KaiA. We implemented this mechanism in our mass action model, with a K_D for S–KaiC–KaiA binding of 10 nM.

The phosphoconversion rate of each KaiC phosphoform is increased by the concentration of available KaiA, such that the total phosphoconversion rate from any

phosphoform C to C', $r_{C\to C'}$, is given by

$$r_{C\to C'} = k_{C\to C'}\frac{K_{1/2}}{A+K_{1/2}} + k^*_{C\to C'}\frac{A}{A+K_{1/2}} \tag{C.2}$$

where $k_{C\to C'}$ is the rate for KaiC conversion in the absence of KaiA, and $k^*_{C\to C'}$ is the conversion rate in the presence of KaiA. The most straightforward mass action implementation of this rate law, where KaiA binds each KaiC phosphoform with a $K_D = K_{1/2}$ is unable to match the oscillation and partial reaction dynamics from the original model or experimental data (unpublished data). This discrepancy arises because the concentration of KaiA (1.3 μM) is not much larger than the concentration of KaiC (3.4 μM), which is required for Equation C.2 to be valid under this model.

Instead, we propose a model in which KaiA binds KaiC transiently, and catalyzes a conformational change or post-translation modification of KaiC that persists after KaiA dissociation. This modification is subsequently lost through a first order process. These reactions can be written as

$$\dot{C} = -k_f C\cdot A_u + k_r C:A + k_d C^* \tag{C.3}$$

$$\dot{C:A} = k_f C\cdot A_u - k_r C:A - k_c C:A \tag{C.4}$$

$$\dot{C^*} = k_c C:A - k_d C^* \tag{C.5}$$

where k_f, k_r and k_c are the forward, reverse and catalytic rates for KaiA's reaction with KaiC, k_d is the rate constant for the first-order regeneration of unmodified KaiC, and $A_u = A - C:A$ is the concentration of unbound KaiA. This model can be solved at quasi steady state to obtain the concentration of C^* in terms of the Michaelis Menten constant $K_M = \frac{(k_c+k_r)}{k_f}$ and the ratio of catalysis to de-modification $\phi = \frac{k_d}{k_c}$ to obtain

$$\phi(1+\phi)C^{*2} - (\phi C_T + K_M\phi + (1+\phi)A)C^* + C_T * A = 0. \tag{C.6}$$

While this quadratic expression is always valid, we find that in the limit of low ϕ (such that $\phi \ll 1$ and $A \gg \phi C^*$) it simplifies to the relation

$$\frac{C^*}{C_T} = \frac{A}{A + K_M \phi}. \tag{C.7}$$

Thus, under the assumption that the modified KaiC species C^* reacts with the rate constant $k^*_{C \to C'}$, and that both C and $C : A$ react with rate $k_{C \to C'}$, the fraction of C^* given by Equation C.7 exactly recovers the rate law of Equation C.2, where $K_{1/2} = \phi K_M$. To verify that our fitted model parameters were in this regime, we computed the concentration of activated KaiC from the original model (using the expression $\frac{A}{A+K_{1/2}}$), and using Equation C.6; the two approaches are in excellent agreement (see Figure C-5).

Parameter name	Initial value	Fitted value	Description
k_{fSAi}	1×10^4	N/A	S–KaiC/KaiA association
k_{fSAi2}	1×10^7	N/A	S–KaiC/KaiA$_2$ association
k_{rAi}	1×10^3	N/A	S–KaiC/KaiA dissociation
k_{fSB}	1×10^5	N/A	S–KaiC/KaiB binding association
k_{rCB}	1×10^3	N/A	S–KaiC/KaiB binding dissociation
k_{suS}	0.110	0.110	S–KaiC → U–KaiC reaction
k_{tuT}	0.210	0.209	T–KaiC → U–KaiC reaction
k_{dsD}	0.310	0.311	D–KaiC → S–KaiC reaction
k^*_{usU}	0.0532	0.0534	U–KaiC*→ S–KaiC*reaction
k^*_{utU}	0.479	0.481	U–KaiC*→ T–KaiC*reaction
k^*_{sdS}	0.506	0.504	S–KaiC*→ D–KaiC*reaction
k^*_{tdT}	0.213	0.213	T–KaiC*→ D–KaiC*reaction
k^*_{tuT}	0.290	0.292	T–KaiC*→ U–KaiC*reaction
k^*_{dtD}	0.173	0.173	D–KaiC*→ T–KaiC*reaction
k_{fCA}	120	122	KaiC–KaiA association
k_{rCA}	640	640	KaiC–KaiA dissociation
k_{cCA}	2×10^4	2×10^4	KaiC → KaiC*catalytic
k_{dC}	50	49.7	KaiC*→ KaiC reaction
$K_{1/2}$	4.30	4.20	Half-maximal KaiC activation

Table C-0: The initial and fitted values of all parameters in the mass action model are shown. Parameters that were not used in the fit are indicated with a N/A entry in the 'Fitted value' column. The final parameter $K_{1/2}$ is not set independently but computed as $K_{1/2} = \frac{k_d}{k_c} \cdot \frac{(k_c+k_r)}{k_f}$, and shown for comparison to the original model.

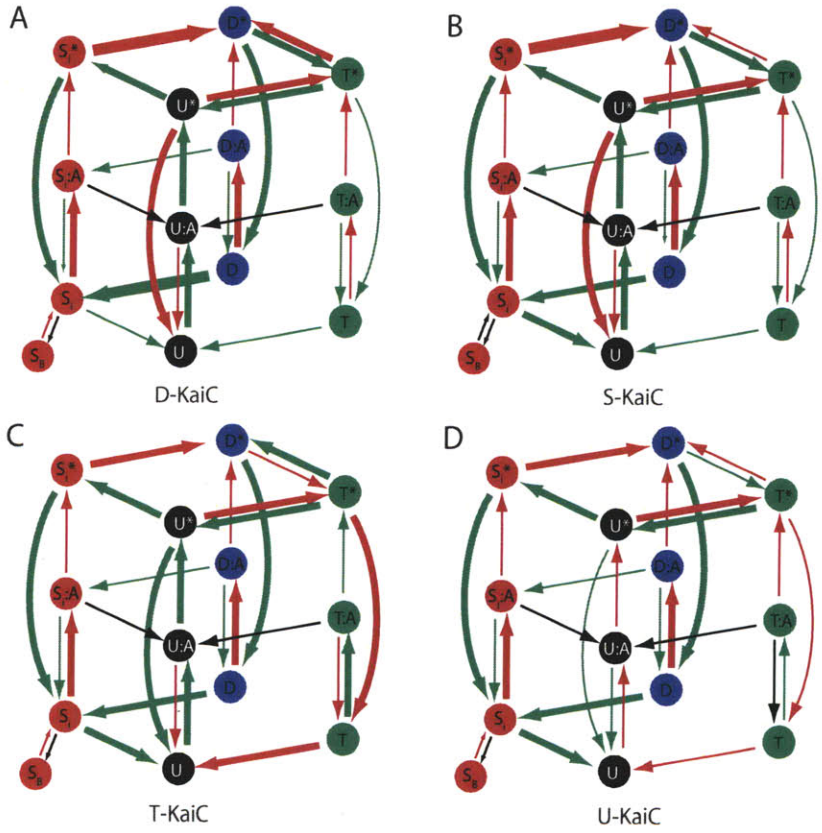


Figure C-1: The relative sensitivities to (a) D–KaiC, (b) S–KaiC, (c) T–KaiC, and (d) U–KaiC amplitudes are shown, where amplitudes are computed as the difference between maximum and minimum concentration for each species. Arrow thicknesses and colors are as described for Figure 5-2B.

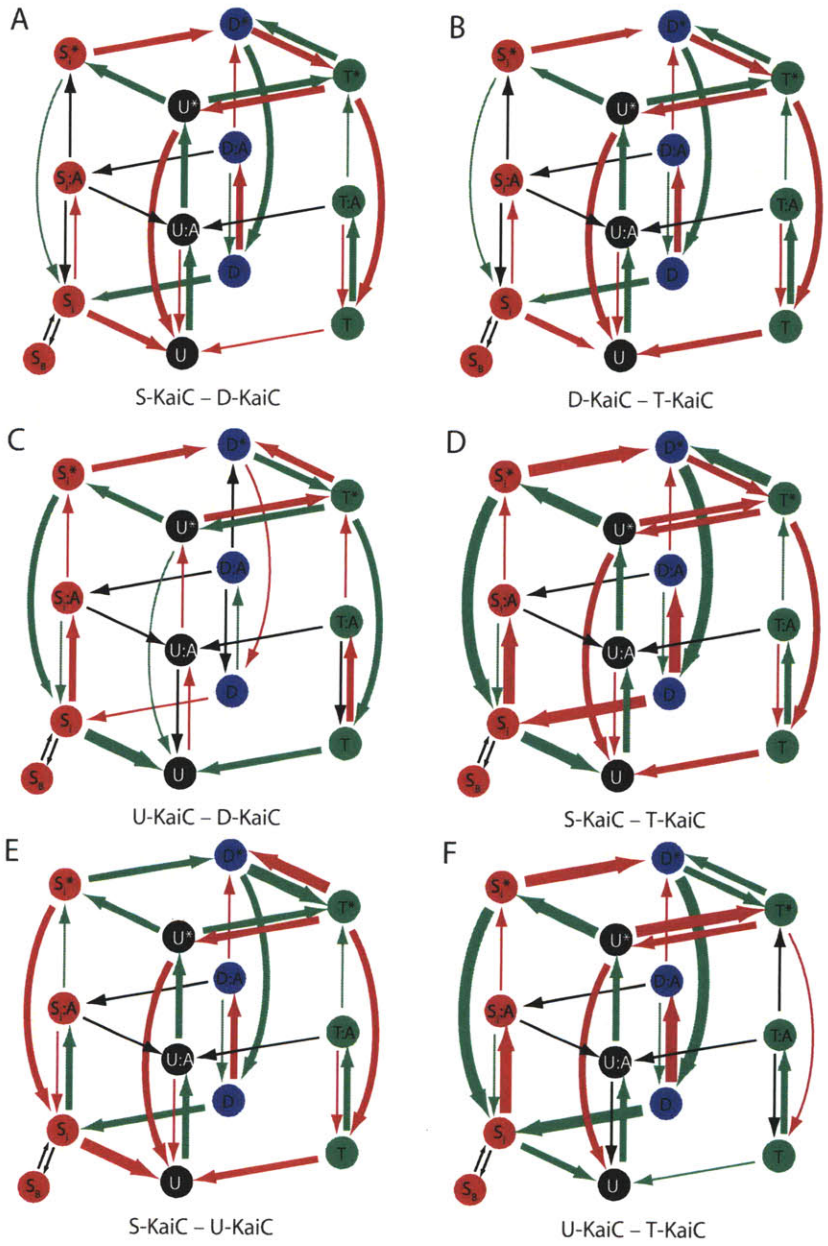


Figure C-2: The relative angular sensitivities are shown for the fraction of the period between all total KaiC phosphoform maxima. Each pair of phosphoforms is listed in the corresponding figure panel, where the first species is taken to be the earlier of the two peaks. Arrow thicknesses and colors are as described for Figure 5-2C.

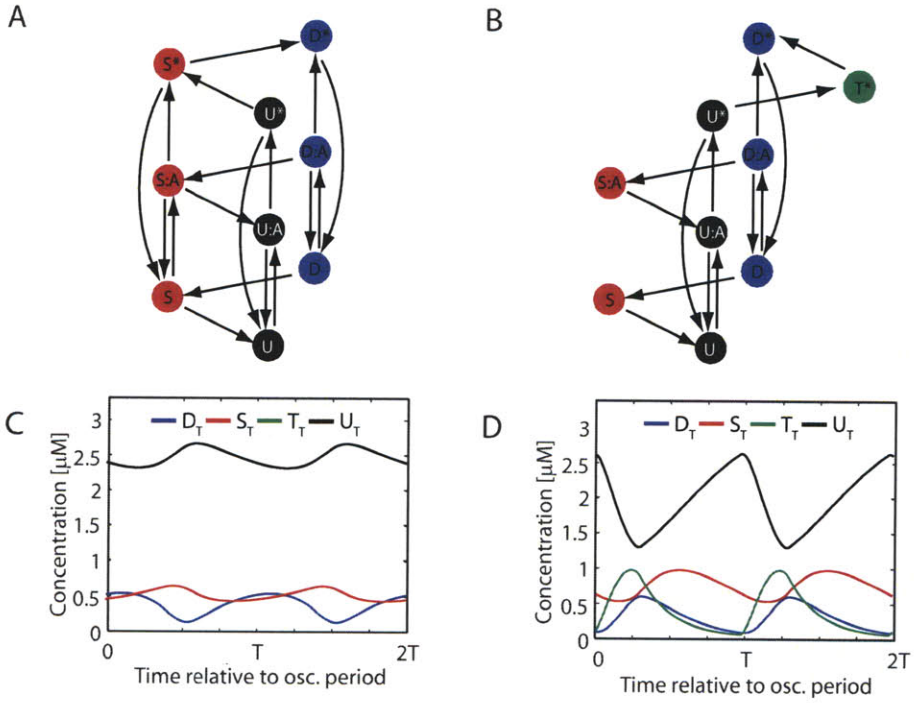


Figure C-3: (a-b) The two minimal oscillating subnetworks identified from the set of subnetworks in which KaiC $\leftrightarrow$ KaiC*arrows are simultaneously. (c-d) Total KaiC phosphoform trajectories shown over two oscillation periods corresponding to the subnetworks of (a) and (b), respectively. The concentrations of D–, S–, T– and U–KaiC are plotted over time relative to the period of oscillation.

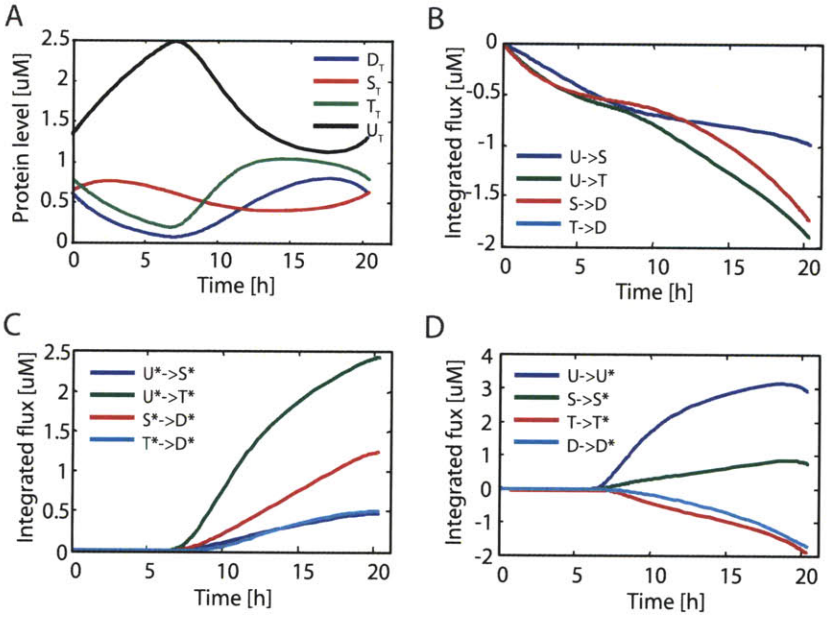


Figure C-4: (a) Simulation of the full model from which all fluxes are computed. The concentrations of total D-, S-, T- and U-KaiC are plotted over time. (b-d) Integrated flux as a function of time between (b) all unmodified KaiC phosphoforms, (c) all modified (i.e. KaiC*) phosphoforms and (d) between corresponding KaiC and KaiC*phosphoforms. For each panel, the x axis shows time elapsed since the beginning of Regime I.

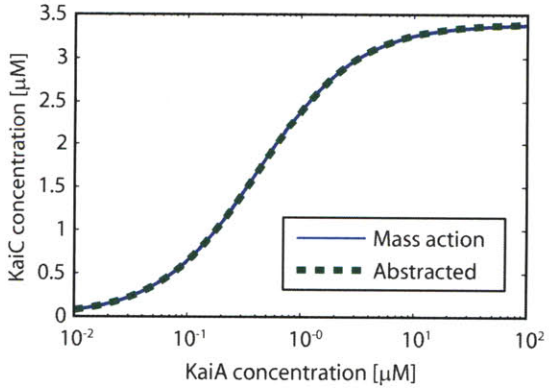


Figure C-5: The steady state concentration of KaiC*is plotted against variation in KaiA concentration across four orders of magnitude, and is computed using the expression $\frac{A}{A+K_{1/2}}$ and Equation C.6 for the abstracted and mass action models, respectively. The total KaiC concentration is 3.4 μM.

Bibliography

[1] Kholodenko B (2006) Cell-signalling dynamics in time and space. Nat Rev Mol Cell Biol 7: 165–176.

[2] Mathon N, Lloyd A (2001) Cell senescence and cancer. Nat Rev Cancer 1: 203–13.

[3] Berridge M, Bootman M, Roderick H (2003) Calcium signalling: dynamics, homeostasis and remodelling. Nat Rev Mol Cell Biol 4: 517–529.

[4] Mitchison J (1971) The biology of the cell cycle. Cambridge University Press Cambridge.

[5] Dunlap J, Loros J, Decoursey P (2004) Chronobiology: biological timekeeping. Sinauer Associates.

[6] Marshall C (1995) Specificity of receptor tyrosine kinase signaling: transient versus sustained extracellular signal-regulated kinase activation. Cell 80: 179.

[7] Santos S, Verveer P, Bastiaens P (2007) Growth factor-induced MAPK network topology shapes Erk response determining PC-12 cell fate. Nat Cell Biol 9: 324–330.

[8] Hoffmann A, Levchenko A, Scott M, Baltimore D (2002) The IκB-NF-κB signaling module: temporal control and selective gene activation. Science 298: 1241–1245.

[9] Lahav G, Rosenfeld N, Sigal A, Geva-Zatorsky N, Levine AJ, et al. (2004) Dynamics of the p53-Mdm2 feedback loop in individual cells. Nat Genet 36: 147–150.

[10] Michael D, Oren M (2003) The p53-Mdm2 module and the ubiquitin system. Semin Cancer Biol 13: 49–58.

[11] Bar-Or L, et al. (2000) Generation of oscillations by the p53-Mdm2 feedback loop: A theoretical and experimental study. Proc Natl Acad Sci 97: 11250.

[12] Geva-Zatorsky N, Rosenfeld N, Itzkovitz S, Milo R, Sigal A, et al. (2006) Oscillations and variability in the p53 system. Mol Syst Biol 2.

[13] Batchelor E, Mock C, Bhan I, Loewer A, Lahav G (2008) Recurrent initiation: a mechanism for triggering p53 pulses in response to DNA damage. Mol Cell 30: 277–289.

[14] Toettcher J, Loewer A, Ostheimer G, Yaffe M, Tidor B, et al. (2009) Distinct mechanisms act in concert to mediate cell cycle arrest. Proc Natl Acad Sci 106: 785.

[15] Hamstra D, Bhojani M, Griffin L, Laxman B, Ross B, et al. (2006) Real-time evaluation of p53 oscillatory behavior in vivo using bioluminescent imaging. Cancer Res 66: 7482.

[16] Chalfie M, Tu Y, Euskirchen G, Ward W, Prasher D (1994) Green fluorescent protein as a marker for gene expression. Science 263: 802–805.

[17] Zhang J, Campbell R, Ting A, Tsien R (2002) Creating new fluorescent probes for cell biology. Nat Rev Mol Cell Biol 3: 906–918.

[18] Thattai M, van Oudenaarden A (2001) Intrinsic noise in gene regulatory networks. Proc Natl Acad Sci 98: 8614.

[19] McAdams H, Arkin A (1997) Stochastic mechanisms in gene expression. Proc Natl Acad Sci 94: 814–819.

[20] Weinberger L, Burnett J, Toettcher J, Arkin A, Schaffer D (2005) Stochastic gene expression in a lentiviral positive-feedback loop: HIV-1 Tat fluctuations drive phenotypic diversity. Cell 122: 169–182.

[21] Rosenfeld N, Young J, Alon U, Swain P, Elowitz M (2005) Gene regulation at the single-cell level. Science 307: 1962–1965.

[22] Samoilov M, Arkin A, Ross J (2002) Signal processing by simple chemical systems. J Phys Chem A 106: 10205–10221.

[23] Barkai N, Leibler S (2000) Biological rhythms: Circadian clocks limited by noise. Nature 403: 267–268.

[24] Janes K, Albeck J, Gaudet S, Sorger P, Lauffenburger D, et al. (2005) A systems model of signaling identifies a molecular basis set for cytokine-induced apoptosis. Science 310: 1646–1653.

[25] Janes K, Gaudet S, Albeck J, Nielsen U, Lauffenburger D, et al. (2006) The response of human epithelial cells to TNF involves an inducible autocrine cascade. Cell 124: 1225–1239.

[26] Tsai T, Choi Y, Ma W, Pomerening J, Tang C, et al. (2008) Robust, tunable biological oscillations from interlinked positive and negative feedback loops. Science STKE 321: 126.

[27] Banin S, Moyal L, Shieh S, Taya Y, Anderson C, et al. (1998) Enhanced phosphorylation of p53 by ATM in response to DNA damage. Science 281: 1674.

[28] Bulavin D, Saito S, Hollander M, Sakaguchi K, Anderson C, et al. (1999) Phosphorylation of human p53 by p38 kinase coordinates N-terminal phosphorylation and apoptosis in response to UV radiation. EMBO J 18: 6845.

[29] Tibbetts R, Brumbaugh K, Williams J, Sarkaria J, Cliby W, et al. (1999) A role for ATR in the DNA damage-induced phosphorylation of p53. Genes Dev 13: 152–157.

[30] Fuchs S, Adler V, Pincus M, Ronai Z (1998) MEKK1/JNK signaling stabilizes and activates p53. Proc Natl Acad Sci 95: 10541–10546.

[31] Shieh S, Ahn J, Tamai K, Taya Y, Prives C (2000) The human homologs of checkpoint kinases Chk1 and Cds1 (Chk2) phosphorylate p53 at multiple DNA damage-inducible sites. Genes Dev 14: 289–300.

[32] Matsuoka S, Rotman G, Ogawa A, Shiloh Y, Tamai K, et al. (2000) Ataxia telangiectasia-mutated phosphorylates Chk2 in vivo and in vitro. Proc Natl Acad Sci 97: 10389–10394.

[33] Harris S, Levine A (2005) The p53 pathway: positive and negative feedback loops. Oncogene 24: 2899–2908.

[34] Taylor W, Schonthal A, Galante J, Stark G (2001) p130/E2F4 binds to and represses the cdc2 promoter in response to p53. J Biol Chem 276: 1998–2006.

[35] Spurgers K, Gold D, Coombes K, Bohnenstiehl N, Mullins B, et al. (2006) Identification of cell cycle regulatory genes as principal targets of p53-mediated transcriptional repression. J Biol Chem 281: 25134.

[36] Innocente S, Abrahamson J, Cogswell J, Lee J (1999) p53 regulates a G2 checkpoint through cyclin B1. Proc Natl Acad Sci 96: 2147–2152.

[37] Bunz F, Dutriaux A, Lengauer C, Waldman T, Zhou S, et al. (1998) Requirement for p53 and p21 to sustain G2 arrest after DNA damage. Science 282: 1497.

[38] Ahn J, Prives C (2002) Checkpoint kinase 2 (Chk2) monomers or dimers phosphorylate Cdc25C after DNA damage regardless of threonine 68 phosphorylation. J of Biol Chem 277: 48418–48426.

[39] Kaneko Y, Watanabe N, Morisaki H, Akita H, Fujimoto A, et al. (1999) Cell cycle-dependent and ATM-independent expression of human Chk 1 kinase. Oncogene 18: 3673–3681.

[40] Taylor W, Stark G (2001) Regulation of the G2/M transition by p53. Oncogene 20: 1803–1815.

[41] Haupt Y, Maya R, Kazaz A, Oren M (1997) Mdm2 promotes the rapid degradation of p53. Nature 387: 296.

[42] Kruse J, Gu W (2009) Modes of p53 Regulation. Cell 137: 609–622.

[43] Bode A, Dong Z (2004) Post-translational modification of p53 in tumorigenesis. Nat Rev Cancer 4: 793–805.

[44] Tang Y, Zhao W, Chen Y, Zhao Y, Gu W (2008) Acetylation is indispensable for p53 activation. Cell 133: 612–626.

[45] Tang Y, Luo J, Zhang W, Gu W (2006) Tip60-dependent acetylation of p53 modulates the decision between cell-cycle arrest and apoptosis. Molecular Cell 24: 827–839.

[46] Sykes S, Mellert H, Holbert M, Li K, Marmorstein R, et al. (2006) Acetylation of the p53 DNA-binding domain regulates apoptosis induction. Molecular cell 24: 841–851.

[47] Ma L, Wagner J, Rice J, Hu W, Levine A, et al. (2005) A plausible model for the digital response of p53 to DNA damage. Proc Natl Acad Sci 102: 14266–14271.

[48] Thomson M, Gunawardena J (2009) Unlimited multistability in multisite phosphorylation systems. Nature 460: 274–277.

[49] Kageyama H, Nishiwaki T, Nakajima M, Iwasaki H, Oyama T, et al. (2006) Cyanobacterial circadian pacemaker: Kai protein complex dynamics in the KaiC phosphorylation cycle in vitro. Mol Cell 23: 161–171.

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

[51] Craciun G, Tang Y, Feinberg M (2006) Understanding bistability in complex enzyme-driven reaction networks. Proc Natl Acad Sci 103: 8697.

[52] Angeli D, Ferrell J, Sontag E (2004) Detection of multistability, bifurcations, and hysteresis in a large class of biological positive-feedback systems. Proc Natl Acad Sci 101: 1822–1827.

[53] Nakajima M, Imai K, Ito H, Nishiwaki T, Murayama Y, et al. (2005) Reconstitution of circadian oscillation of cyanobacterial KaiC phosphorylation in vitro. Science 308: 414–415.

[54] Lauffenburger D, Linderman J (1996) Receptors: Models for binding, trafficking, and signaling. Oxford University Press, USA.

[55] Buchler N, Louis M (2008) Molecular titration and ultrasensitivity in regulatory networks. J Molec Biol 384: 1106–1119.

[56] Kim S, Ferrell J (2007) Substrate competition as a source of ultrasensitivity in the inactivation of Wee1. Cell 128: 1133–1145.

[57] Buchler N, Cross F (2009) Protein sequestration generates a flexible ultrasensitive response in a genetic network. Molecular Systems Biology 5.

[58] Harmer S, Panda S, Kay S (2001) Molecular bases of circadian rhythms. Ann Rev Cell Dev Biol 17: 215–253.

[59] Herrero A, Flores E (2008) The cyanobacteria: molecular biology, genomics, and evolution. Caister Academic Pr.

[60] Ishiura M, Kutsuna S, Aoki S, Iwasaki H, Andersson C, et al. (1998) Expression of a gene cluster kaiABC as a circadian feedback process in cyanobacteria. Science 281: 1519.

[61] Nakajima M, Imai K, Ito H, Nishiwaki T, Murayama Y, et al. (2005) Reconstitution of circadian oscillation of cyanobacterial KaiC phosphorylation in vitro. Science 308: 414–415.

[62] Hayashi F, Ito H, Fujita M, Iwase R, Uzumaki T, et al. (2004) Stoichiometric interactions between cyanobacterial clock proteins kaia and kaic. Biochem Biophys Res Commun 316: 195 - 202.

[63] Nishiwaki T, Satomi Y, Nakajima M, Lee C, Kiyohara R, et al. (2004) Role of KaiC phosphorylation in the circadian clock system of Synechococcus elongatus PCC 7942. Proc Natl Acad Sci 101: 13927–13932.

[64] Xu Y, Mori T, Pattanayek R, Pattanayek S, Egli M, et al. (2004) Identification of key phosphorylation sites in the circadian clock protein KaiC by crystallographic and mutagenetic analyses. Proc Natl Acad Sci 101: 13933–13938.

[65] Rust M, Markson J, Lane W, Fisher D, O'Shea E (2007) Ordered phosphorylation governs oscillation of a three-protein circadian clock. Science 318: 809.

[66] Emberly E, Wingreen N (2006) Hourglass model for a protein-based circadian oscillator. Phys Rev Lett 96: 38303.

[67] Mori T, Williams D, Byrne M, Qin X, Egli M, et al. (2007) Elucidating the ticking of an in vitro circadian clockwork. PLoS Biol 5: e93.

[68] Milo R, Shen-Orr S, Itzkovitz S, Kashtan N, Chklovskii D, et al. (2002) Network motifs: simple building blocks of complex networks. Science 298: 824–827.

[69] Shen-Orr S, Milo R, Mangan S, Alon U (2002) Network motifs in the transcriptional regulation network of Escherichia coli. Nat Genet 31: 64–68.

[70] Yeger-Lotem E, Sattath S, Kashtan N, Itzkovitz S, Milo R, et al. (2004) Network motifs in integrated cellular networks of transcription-regulation and protein-protein interaction. Proc Natl Acad Sci 101: 5934–5939.

[71] Alon U (2007) An introduction to systems biology: design principles of biological circuits. Chapman & Hall/CRC.

[72] Kashtan N, Itzkovitz S, Milo R, Alon U (2004) Topological generalizations of network motifs. Physical Review E 70: 31909.

[73] Mangan S, Alon U (2003) Structure and function of the feed-forward loop network motif. Proc Natl Acad Sci 100: 11980–11985.

[74] Mangan S, Zaslaver A, Alon U (2003) The coherent feedforward loop serves as a sign-sensitive delay element in transcription networks. J Molec Biol 334: 197–204.

[75] Kalir S, Mangan S, Alon U (2005) A coherent feed-forward loop with a SUM input function prolongs flagella expression in Escherichia coli. Mol Syst Biol 1.

[76] Tyson J, Chen K, Novak B (2003) Sniffers, buzzers, toggles and blinkers: dynamics of regulatory and signaling pathways in the cell. Curr Op Cell Biol 15: 221–231.

[77] Novák B, Tyson J (2008) Design principles of biochemical oscillators. Nat Rev Mol Cell Biol 9: 981–991.

[78] Elowitz M, Leibler S (2000) A synthetic oscillatory network of transcriptional regulators. Nature 403: 335–338.

[79] Stricker J, Cookson S, Bennett M, Mather W, Tsimring L, et al. (2008) A fast, robust and tunable synthetic gene oscillator. Nature 456: 516–519.

[80] Tigges M, Marquez-Lago TT, Stelling J, Fussenegger M (2009) A tunable synthetic mammalian oscillator. Nature 457: 309–312.

[81] Apgar J (2009) Experiment design for systems biology. Ph.D. thesis, Massachusetts Institute of Technology.

[82] Apgar J, Toettcher J, Endy D, White F, Tidor B (2008) Stimulus design for model selection and validation in cell signaling. PLoS Comput Biol 4: e30.

[83] Wilkins A, Barton P, Tidor B (2007) The Per2 negative feedback loop sets the period in the mammalian circadian clock mechanism. PLoS Comput Biol 3: e242.

[84] Bagheri N, Stelling J, Doyle F (2007) Quantitative performance metrics for robustness in circadian rhythms. Bioinformatics 23: 358-364.

[85] Rosenwasser E, Yusupov R (2000) Sensitivity of automatic control systems. CRC.

[86] Ingalls B (2004) Autonomously oscillating biochemical systems: parametric sensitivity of extrema and period. IEE Proceedings Systems Biology 1: 62–70.

[87] Gunawan R, Doyle F (2006) Isochron-based phase response analysis of circadian rhythms. Biophysical journal 91: 2131–2141.

[88] Wilkins A, Tidor B, White J, Barton P (2009) Sensitivity Analysis for Oscillating Dynamical Systems. SIAM J Sci Comput 31: 2706.

[89] Rand D, Shulgin B, Salazar D, Millar A (2004) Design principles underlying circadian clocks. J Roy Soc Interface 1: 119–130.

[90] Cao Y, Li S, Petzold L, Serban R (2003) Adjoint sensitivity analysis for differential-algebraic equations: The adjoint DAE system and its numerical solution. SIAM Journal on Scientific Computing 24: 1076.

[91] Alon U (2007) Network motifs: theory and experimental approaches. Nat Rev Genet 8: 450–461.

[92] Bashor C, Helman N, Yan S, Lim W (2008) Using engineered scaffold interactions to reshape MAP kinase pathway signaling dynamics. Science 319: 1539.

[93] Kaku S, Iwahashi Y, Kuraishi A, Albor A, Yamagishi T, et al. (2001) Binding to the naturally occurring double p53 binding site of the mdm2 promoter alleviates the requirement for p53 c-terminal activation. Nucl Acids Res 29: 1989–1993.

[94] Andersen RD, Taplitz SJ, Wong S, Bristol G, Larkin B, et al. (1987) Metal-dependent binding of a factor in vivo to the metal-responsive elements of the metallothionein 1 gene promoter. Mol Cell Biol 7: 3574–3581.

[95] Nagoshi E, Saini C, Bauer C, Laroche T, Naef F, et al. (2004) Circadian Gene Expression in Individual Fibroblasts Cell-Autonomous and Self-Sustained Oscillators Pass Time to Daughter Cells. Cell 119: 693–705.

[96] Wu X, Bayle J, Olson D, Levine A (1993) The p53-mdm-2 autoregulatory feedback loop. Genes Dev 7: 1126–1132.

[97] Fang S, Jensen J, Ludwig R, Vousden K, Weissman A (2000) Mdm2 is a RING finger-dependent ubiquitin protein ligase for itself and p53. J Biol Chem 275: 8945–8951.

[98] Shaner N, Campbell R, Steinbach P, Giepmans B, Palmer A, et al. (2004) Improved monomeric red, orange and yellow fluorescent proteins derived from Discosoma sp. red fluorescent protein. Na biotechnol 22: 1567–1572.

[99] Brugnera E, Georgiev O, Radtke F, Heuchel R, Baker E, et al. (1994) Cloning, chromosomal mapping and characterization of the human metal-regulatory transcription factor MTF-1. Nucl acids res 22: 3167.

[100] Vassilev LT, Vu BT, Graves B, Carvajal D, Podlaski F, et al. (2004) In vivo activation of the p53 pathway by Small-Molecule antagonists of MDM2. Science 303: 844–848.

[101] Murray A (2004) Recycling the Cell Cycle Cyclins Revisited. Cell 116: 221–234.

[102] Tyson J (1991) Modeling the cell division cycle: cdc2 and cyclin interactions. Proc Natl Acad Sci 88: 7328–7332.

[103] Haberichter T, Mädge B, Christopher R, Yoshioka N, Dhiman A, et al. (2007) A systems biology dynamical model of mammalian G1 cell cycle progression. Molecular Systems Biology 3.

[104] Novak B, Tyson J (2004) A model for restriction point control of the mammalian cell cycle. J Theor Biol 230: 563–579.

[105] Nyberg K, Michelson R, Putnam C, Weinert T (2002) Toward maintaining the genome: DNA damage and replication checkpoints. Annual review of genetics 36: 617.

[106] Agami R, Bernards R (2000) Distinct initiation and maintenance mechanisms cooperate to induce G1 cell cycle arrest in response to DNA damage. Cell 102: 55–66.

[107] Csikász-Nagy A, Battogtokh D, Chen K, Novák B, Tyson J (2006) Analysis of a generic model of eukaryotic cell-cycle regulation. Biophys J 90: 4361–4379.

[108] Conlon I, Raff M (2003) Differences in the way a mammalian cell and yeast cells coordinate cell growth and cell-cycle progression. J Biol 2: 7.

[109] Zetterberg A, Larsson O (1985) Kinetic analysis of regulatory events in G1 leading to proliferation or quiescence of Swiss 3T3 cells. Proc Natl Acad Sci 82: 5365–5369.

[110] Geng Y, Yu Q, Sicinska E, Das M, Schneider J, et al. (2003) Cyclin E ablation in the mouse. Cell 114: 431–443.

[111] Kozar K, Ciemerych M, Rebel V, Shigematsu H, Zagozdzon A, et al. (2004) Mouse development and cell proliferation in the absence of D-cyclins. Cell 118: 477–491.

[112] Waldman T, Kinzler K, Vogelstein B (1995) p21 is necessary for the p53-mediated G1 arrest in human cancer cells. Cancer res 55: 5187–5190.

[113] Harper J, Elledge S, Keyomarsi K, Dynlacht B, Tsai L, et al. (1995) Inhibition of cyclin-dependent kinases by p21. Mol biol cell 6: 387.

[114] Krause K, Wasner M, Reinhard W, Haugwitz U, Lange-zu Dohna C, et al. (2000) The tumour suppressor protein p53 can repress transcription of cyclin B. Nucl Acids Res 28: 4410.

[115] Waldman T, Lengauer C, Kinzler K, Vogelstein B (1996) Uncoupling of S phase and mitosis induced by anticancer agents in cells lacking p21. Nature 381: 713.

[116] Niculescu A, Chen X, Smeets M, Hengst L, Prives C, et al. (1998) Effects of p21Cip1/Waf1 at both the G1/S and the G2/M cell cycle transitions: pRb is a critical determinant in blocking DNA replication and in preventing endoreduplication. Mol cell biol 18: 629–643.

[117] Vassilev L, Tovar C, Chen S, Knezevic D, Zhao X, et al. (2006) Selective small-molecule inhibitor reveals critical mitotic functions of human CDK1. Proc Natl Acad Sci 103: 10660.

[118] di Fagagna F, Asia N (2008) Living on a break: cellular senescence as a DNA-damage response. Nat Rev Cancer 8: 512–522.

[119] Smits V, Klompmaker R, Vallenius T, Rijksen G, Makela T, et al. (2000) p21 inhibits Thr161 phosphorylation of Cdc2 to enforce the G2 DNA damage checkpoint. J Biol Chem 275: 30638–30643.

[120] Charrier-Savournin F, Chateau M, Gire V, Sedivy J, Piette J, et al. (2004) p21-Mediated nuclear retention of cyclin B1-Cdk1 in response to genotoxic stress. Mol Biol Cell 15: 3965–3976.

[121] Bates S, Ryan K, Phillips A, Vousden K (1998) Cell cycle arrest and DNA endoreduplication following p21Waf1/Cip1 expression. Oncogene 17: 1691.

[122] Stewart Z, Leach S, Pietenpol J (1999) p21Waf1/Cip1 inhibition of cyclin E/Cdk2 activity prevents endoreduplication after mitotic spindle disruption. Mol Cell Biol 19: 205–215.

[123] Dean P, Jett J (1974) Mathematical analysis of DNA distributions derived from flow cytometry. J Cell Biol 60: 523–527.

[124] Aprille T, Trick T (1972) A computer algorithm to determine the steady-state response of nonlinear oscillators. IEEE Trans Circuit Theory 19: 354–360.

[125] Kundert K, White J, Sangiovanni-Vincentelli A (1990) Steady-state methods for simulating analog and microwave circuits. Kluwer Academic Pub.

[126] Kaertner F (1990) Analysis of white and f-α noise in oscillators. International Journal of Circuit Theory and Applications 18.

[127] Lee T, Hajimiri A (2000) Oscillator phase noise: A tutorial. IEEE J Solid-State Circuits 35: 326–336.

[128] Demir A, Mehrotra A, Roychowdhury J (2000) Phase noise in oscillators: a unifying theory and numerical methodsfor characterization. IEEE Trans Circuits and Systems I: Fundamental Theory and Applications 47: 655–674.

[129] Taylor S, Gunawan R, Petzold L, Doyle F (2008) Sensitivity Measures for Oscillating Systems: Application to Mammalian Circadian Gene Network. IEEE Trans Automat Control 53: 177–188.

[130] Vytyaz I, Lee D, Hanumolu P, Moon U, Mayaram K (2008) Sensitivity analysis for oscillators. IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems 27: 1521–1534.

[131] Famili I, Palsson B (2003) The convex basis of the left null space of the stoichiometric matrix leads to the definition of metabolically meaningful pools. Biophys J 85: 16–26.

[132] Hu W, Feng Z, Ma L, Wagner J, Rice J, et al. (2007) A single nucleotide polymorphism in the MDM2 gene disrupts the oscillation of p53 and MDM2 levels in cells. Cancer Res 67: 2757.

[133] Toh K, Jones C, He Y, Eide E, Hinz W, et al. (2001) An hPer2 Phosphorylation Site Mutation in Familial Advanced Sleep Phase Syndrome. Science 291: 1040–1043.

[134] Cappuccio A, Castiglione F, Piccoli B (2007) Determination of the optimal therapeutic protocols in cancer immunotherapy. Math Biosci 209: 1–13.

[135] Cinquin O, Demongeot J (2002) Positive and negative feedback: striking a balance between necessary antagonists. J Theor Biol 216: 229–242.

[136] Kut C, Golkhou V, Bader J (2009) Analytical approximations for the amplitude and period of a relaxation oscillator. BMC Systems Biology 3: 6.

[137] Forger D, Peskin C (2003) A detailed predictive model of the mammalian circadian clock. Proc Natl Acad Sci 100: 14806–14811.

[138] Leloup J, Goldbeter A (2004) Modeling the mammalian circadian clock: Sensitivity analysis and multiplicity of oscillatory mechanisms. J Theor Biol 230: 541–562.

[139] Liou J, Kim M, Do Heo W, Jones J, Myers J, et al. (2005) STIM is a Ca2+ sensor essential for Ca2+-store-depletion-triggered Ca2+ influx. Curr Biol 15: 1235–1241.

[140] Wang J (2005) Recent cyanobacterial Kai protein structures suggest a rotary clock. Structure 13: 735–741.

[141] Stelling J, Gilles E, Doyle F (2004) Robustness properties of circadian clock architectures. Proc Natl Acad Sci 101: 13210–13215.

[142] Zak D, Stelling J, Doyle F (2005) Sensitivity analysis of oscillatory (bio) chemical systems. Comput Chem Eng 29: 663–673.

[143] Varma A, Palsson B (1994) Metabolic flux balancing: basic concepts, scientific and practical use. Nat Biotechnol 12: 994–998.

[144] Stelling J, Klamt S, Bettenbrock K, Schuster S, Gilles E (2002) Metabolic network structure determines key aspects of functionality and regulation. Nature 420: 190–193.

[145] Kim Y, Dong G, Carruthers C, Golden S, LiWang A (2008) The day/night switch in KaiC, a central oscillator component of the circadian clock of cyanobacteria. Proc Natl Acad Sci 105: 12825.

[146] Strang G (1986) Introduction to applied mathematics. Wellesley Cambridge Pr.

[147] Nocedal J, Wright S (1999) Numerical optimization. Springer.

[148] Bertsekas D, Homer M, Logan D, Patek S (1995) Nonlinear programming. Athena scientific.

[149] Toettcher J, Castillo A, White J, Tidor B Oscillator sensitivity analysis in the presence of hidden conservation constraints. Submitted .

[150] Fox RF, Gatland IR, Roy R, Vemuri G (1988) Fast, accurate algorithm for numerical simulation of exponentially correlated colored noise. Phys Rev A 38: 5938.