Quantitative analysis of cellular networks: cell cycle entry

by

Tae Jun Lee

Department of Biomedical Engineering
Duke University

Date:_____
Approved:

_____
Dr. Lingchong You, Supervisor

_____
Dr. Craig S. Henriquez

_____
Dr. Fan Yuan

_____
Dr. Jen Tsan Chi

_____
Dr. Joseph R. Nevins

Dissertation submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in the Department of
Biomedical Engineering in the
Graduate School of Duke University

2010

ABSTRACT

Quantitative analysis of cellular networks: cell cycle entry

by

Tae Jun Lee

Department of Biomedical Engineering
Duke University

Date:_____
Approved:

_____
Dr. Lingchong You, Supervisor

_____
Dr. Craig S. Henriquez

_____
Dr. Fan Yuan

_____
Dr. Jen Tsan Chi

_____
Dr. Joseph R. Nevins

An abstract of a dissertation submitted in partial fulfillment
of the requirements for the degree of Doctor of Philosophy
in the Department of Biomedical Engineering
in the Graduate School
of Duke University

2010

Abstract

Cellular dynamics arise from intricate interactions among diverse components, such as metabolites, RNAs, and proteins. An in-depth understanding of these interactions requires an integrated approach to the investigation of biological systems. This task can benefit from a combination of mathematical modeling and experimental validations, which is becoming increasingly indispensable for basic and applied biological research.

Utilizing a combination of modeling and experimentation, we investigate mammalian cell cycle entry. We begin our investigation by making predictions with a mathematical model, which is constructed based on the current knowledge of biology. To test these predictions, we develop experimental platforms for validations, which in turn can be used to further refine the model. Such iteration of model predictions and experimental validations has allowed us to gain an in-depth understanding of the cell cycle entry dynamics.

In this dissertation, we have focused on the Myc-Rb-E2F signaling pathway and its associated pathways, dysregulation of which is associated with virtually all cancers. Our analyses of these signaling pathways provide insights into three questions in biology: 1) regulation of the restriction point (R-point) in cell cycle entry, 2) regulation of the temporal dynamics in cell cycle entry, and 3) post-translational regulation of Myc by its upstream signaling pathways. The well-studied pathways can serve as a foundation for

perturbations and tight control of cell cycle entry dynamics, which may be useful in

developing cancer therapeutics.

We conclude by demonstrating how a combination of mathematical modeling and

experimental validations provide mechanistic insights into the regulatory networks in

cell cycle entry.

To mom, dad, my sister and her family

In loving memory of my brother in heaven…

Table of Contents

List of Tables

List of Figures

List of Abbreviations

BGS: bovine growth serum

Cdks: cyclin-dependent kinases

CKI: cyclin-dependent kinase inhibitors

CME: chemical master equation

CycA: cyclin A

CycE: cyclin E

CycD: cyclin D

DMEM: Dulbecco's modified eagle medium

EGF: epidermal growth factor

Erk: extracellular signal-regulated kinase

FAK: focal adhesion kinase

FBS: fetal bovine serum

GC: growth-controlled

GFP: green fluorescence protein

GSK3β: glycogen synthase kinase 3β

MAPK: mitogen-activated protein kinase

MOI: multiplicity of infection

NGF: nerve growth factor

ODE: ordinary differential equation

PDGF: platelet-derived growth factor

PI3K: phosphatidylinositol 3–kinase

RNAi: RNA interference

R-point: restriction point

SDE: stochastic differential equation

Ser62: serine 62

siRNA: small interfering RNA

Thr58: threonine 58

TP: transition probability

tQSSA: total quasi-steady state assumption

Acknowledgements

A number of people have been instrumental in my pursuit and completion of my Ph.D. training at Duke University. First and most importantly, I would like to thank my advisor Lingchong You for his patience, encouragement and guidance throughout my time at Duke. His true passion and enthusiasm for science has most definitely motivated me throughout my sometimes arduous graduate school career. I would also like to thank my collaborator Guang Yao for guiding me through the experimental part of my graduate studies. Many techniques in molecular biology that I have acquired are attributed to Guang. Many thanks goes to Jeffery Wong, who is the only other member in the You lab that works on the mammalian system, for sharing his experiences and insights in my projects.

I would also like to thank all other colleagues in the You lab for making my time at Duke more enjoyable. The You lab is truly a melting pot of diverse backgrounds, and has been a place not only for scientific discussions but also for cultural and philosophical discussions. Cheemeng Tan and Yu Tanouchi represent East Asia (also the most critical in the You lab along with Philippe Marguet), Anand Pai alone represents South Asia (who is often bitter at his own culture), and Dr. Jeffery Wong and Dr. Rob Smith

xxx

represent Canada, whose passion for hockey often led to a clash with a hockey fanatic Stephen Payne representing the U.S.A.

I am indebted to my thesis committee members: Craig S. Henriquez, Fan Yuan, Joseph R. Nevins, and Ashley (Jen Tsan) Chi. Their insightful comments, questions, and criticisms have helped me grow intellectually. I deeply appreciate all the time, patience, and interest they have invested in my studies.

Finally, I would like to thank my family. I can never pay back my parents for the unconditional love and their prayers, and my sister and her husband Roger for their support at all times. Last but not least, I am extremely grateful for my most adorable nieces Leila and Ella, for they truly bring joy to all of us in my family.

Chapter 1: Introduction – A primer to modeling cellular networks

In Chapter 1, we present methodologies on modeling, simulation and analysis of cellular networks. We note that different types of mathematical models are widely used. Here we limit our scope to kinetic models, which represent systems of interest as coupled chemical reactions. By so doing, we steer away from discussing other widely used mathematical models, such as Boolean models, and those focusing on spatial dynamics. This chapter is reproduced by permission from the 6th chapter of '*Systems Bioinformatics: An Engineering Case-Based Approach*, Norwood, MA: Artech House, Inc., 2007'[1].

1.1 Introduction

Systems-level understanding of cellular dynamics is important for identifying biological principles and may serve as a critical foundation for developing therapeutic strategies. To date, numerous developments of therapeutics have been based on identification and comprehensive analysis of cellular dynamics, especially in the involved pathways. In cancer therapy, for instance, many researchers have focused on oncogenic pathways such as the Rb pathway, whose in-depth understanding of the pathway dynamics promises effective therapeutics [1,2,3,4,5,6,7]. The effectiveness of this approach in the development of cancer therapeutics has been illustrated in *in-vivo*

pre-clinical tests of the engineered adenovirus ONYX-015 and ONYX-411. These adenoviruses, engineered to target mutations in the Rb or p53 pathway for killing, have demonstrated high selectivity and efficiency in viral replication in tumor cells for cell killing [2,3]. However, clinical application of these methods is hindered by lack of ability to precisely predict and regulate cellular responses. This ability is essential in minimizing complications and side-effects. Especially, a large amount of biology data on these pathways generated by rapid advancements in biotechnologies and molecular biology renders integrated understanding of the pathway dynamics impossible by intuition alone. Therefore, a more systematic approach allowing incorporation of the multitude of information is necessary to improve prediction and regulation of cellular responses.

To this end, mathematical modeling is becoming increasingly indispensable for basic and applied biological research. Essentially, a mathematical model is a systematic representation of biological systems, whose analysis can confer quantitative predicting power. In recent years, advanced computing power combined with improved numerical methods has made it possible to simulate and analyze dynamics of complex cellular networks [4,5,6,7,8,9,10,11,12].

Mathematical modeling is useful in a number of ways. One of the common applications of mathematical modeling is to analyze cellular networks systematically. For example, although the mitogen-activated protein kinase (MAPK) was known to

control multiple cellular responses such as cell growth, survival, or differentiation, the molecular mechanisms for these divergent behaviors were not fully elucidated. Consequently, several models on the MAPK pathway have been developed that differentiates activation patterns in response to epidermal growth factors and neural growth factors [13], characterizes the signal-response relationship [14,15], and suggests the significance of feedback control in complete signal adaptation [16]. A more extensive modeling work investigates the emergent properties that arise from multiple signaling pathways [17]. These works illustrate the utility of mathematical modeling in understanding complex biological systems that intuition alone cannot handle.

Another use of mathematical modeling has been demonstrated in devising strategies to control cellular dynamics. The concentrations of MAPK phosphatase have been shown to play a key role in whether the MAPK pathway demonstrates monostable or bistable states [15]. Sasagawa and colleagues use their MAPK model to identify 'critical nodes', to which perturbations result in dramatic changes in system behaviors [13],. A number of critical nodes that are responsible for diverse cellular actions have also been suggested in the insulin-signaling pathways based on biochemical and computational data [18]. Such characterization of input-output response or identification of critical nodes can be utilized to effectively modulate cellular dynamics.

Furthermore, modeling can form a basis for the development of therapeutics for medical applications. Various pathway models including the MAPK models described

above can be useful in designing, or evaluating the effectiveness of, therapeutic drugs *in-silico* [13,14,15,16,17]. The predictive power and therapeutics design principles that these models offer can facilitate development of therapeutics [19,20,21]. Stemming from these studies on the MAPK signaling pathways, Kitano and colleagues have developed an EGFR Pathway Map in a software that is shared and compatible with other simulation and analysis packages [22]. Such efforts to avail and share information on biological pathways among researchers exemplify inclination towards understanding of biology via mathematical modeling.

Despite advantages of mathematical modeling for basic and applied biological research, there remain many challenges in constructing and analyzing models. Modeling of biological systems is always accompanied by assumptions, which are predicated on the modeler's goals. Therefore, a successful modeling work requires clear justification of these assumptions. Even with clear, justified goals, a modeler is faced with another challenge, lack of detailed, quantitative biological information. While biotechnologies continue to advance our knowledge on the building blocks of biological systems, parameters for the kinetics of interactions among them are often unknown. Various methodologies for inferring reaction mechanisms and parameters have been proposed [23,24,25,26,27,28,29]. Yet high-throughput biological data, generated by microarray experiments or protein expression profiling, are often not of sufficiently high resolution for using these techniques. To address these issues, a combination of mathematical

modeling and experimental validations are required. Iterations of model construction, system analysis, and experimental validation improve accuracy of the model, and lead to increased predictive power. In particular, the power to quantify gene expression with high temporal resolution at the population level or single cell level will likely complement high-throughput technologies in facilitating inference of reaction mechanisms and parameter [30,31,32,33,34,35,36].

1.2   Construction and Analysis of Kinetic Models

Construction of a kinetic model can be a daunting task for a system consisting of a large number of components with complex interactions. To build an experimentally tractable model, it is important to define the scope of abstraction. Once the scope is defined, a conventional approach begins with a minimal diagram that includes key components and interactions among them. Identification of the key components and interactions is based on the current knowledge of biology and frequently on intuition and experience. Depending on the focus of study, a modeler may choose to emphasize certain signaling pathways while de-emphasizing less relevant ones. These processes often accompany 'lumping' or deletion of molecular interactions or components. Once the diagram is completed, a minimal mathematical model is constructed from the information embedded in the diagram and is further refined or extended to reflect new

hypotheses or experimental measurements. Simulation of the final model reveals the network dynamics, which in turn gives insights into the intrinsic design principles.

1.2.1 Parameter Estimation and Modeling Resources

A major challenge in model formulation is determination of reaction mechanisms and estimation of parameters. In some systems, the network behaviors are defined mostly by the architecture of the system. These systems are highly robust to a wide range of parameters. In others, system dynamics are determined not only by the architecture, but also the parameters, which are often poorly understood. Therefore, construction of a meaningful mathematical model of a biological pathway requires two critical elements: interactions between molecular species and kinetics of the interactions

As the first step, we need to know the interactions between the molecular species in the model. Several pathway databases are available for this purpose: EcoCyc [37], Kegg [38], ERGO [39], aMAZE [40], ExPASy [41], www.sbml.org, STKE (Sci) and Nature: Signaling update. The pathways included in these databases are retrieved and constructed from specialized databases such as GenBank, PDB and EMBL. These pathway databases often provide detailed information on the molecular species. Next, we need to determine the kinetics of the interactions. In most cases, kinetic parameters are obtained from the literature data. Alternatively, we can also use kinetic parameters

from typical values, which can be based on values inferred from related process or even the experience of the modeler.

For every biological system, model construction usually goes through common process of iterations of model construction, experimental validation, and model refinement (in terms of reaction mechanisms or parameter values) (Figure 1. 1). These steps will be repeated until the mathematical model matches the experimental data to a satisfactory degree. This process can be considered as a special case of "reverse engineering" biological pathways. Additional methods, such as Bayesian [42], maximum likelihood [29], and genetic algorithms [26] can be used to infer more qualitative connectivity of biological networks from high-throughput experimental data.
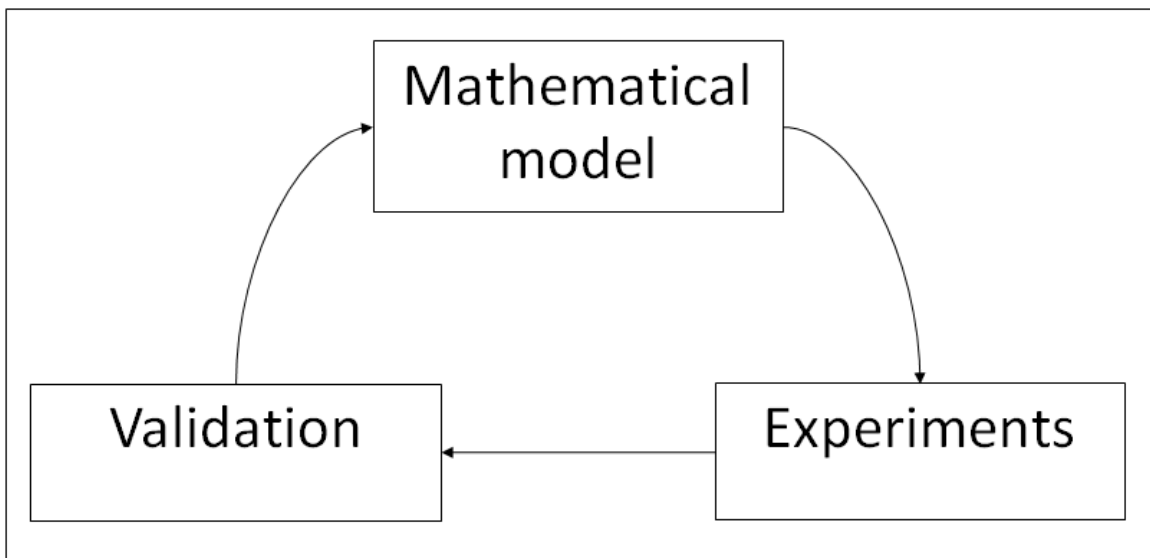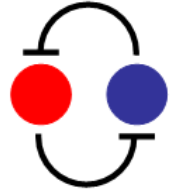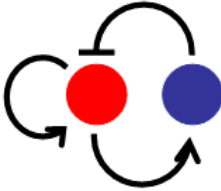
Figure 1. 1: Refining models of biological networks. Iteration of model construction and experiments enable parameter and kinetic estimation and model refinement. The experimental data can be matched to the model with various computation methods.

1.2.2 A Modular Approach to Model Formulation

Modeling and analysis of complex biological systems may benefit from a modular approach, in which a biological system is considered as a combination of smaller subnetworks with well-recognizable functions, termed as motifs and modules [43,44]. The distinction between motifs and modules is often based on the size difference but is not always clear-cut. We here use the two terms interchangeably. That is, we consider all small, repeated, and conserved regulatory subnetworks as 'modules', classifiable on the basis of function, architecture, dynamics, and biochemical process. Such conceptualization may provide insight into the qualitative network dynamics at the systems-level, and it helps clarify the modeling objective and generation of qualitative hypotheses. In addition, it also forms the basis for incorporating mathematical equations, with which more quantitative understanding can be attempted.

Table 1.1: Well-defined feedback modules involving negative, positive, or both types of regulation.

The number of variables	Negative regulation only	Negative and Positive regulation
1	Mono-stable	Bi-stable
2	Bi-stable	Oscillation
3	Oscillation	Oscillation ⋮ Chaos

The dynamics of a module are governed by both network connectivity and associated parameter values. In general, increasing complexity in either variables or connectivity will result in more complex dynamics, and modules with feedback control may show properties that are difficult to grasp by intuition alone. Structures and properties of some well-defined feedback control modules are tabulated in Table 1.1, where we have summarized their key properties. For example, a module with one variable demonstrates either mono-stable or bistable properties with negative or positive feedback control, respectively, but it is impossible to generate oscillations with a single variable in the absence of time delay. Mono-stable, bi-stable, or oscillatory behaviors can be generated with a two-variable module, but not chaos. Modules with higher number of variables can demonstrate much richer dynamics such as chaos (Table 1.1) [43,45,46].

Various feedback control mechanisms confer properties useful for different biological functions. For example, negative feedback control is essential in homeostasis, a process of maintaining the system's internal environment in a steady state. Without feedback control, sudden external changes such as those in temperature or salinity may induce significant internal damages that can be fatal to a cell. Negative feedback control can buffer the impact of such changes and facilitate homeostasis [47]. In attempts to engineer gene circuits, this property has been used to reduce variations in gene expression [48]. In addition, negative feedback may increase the response speed of simple gene cascades [49].

10

Positive feedback can create bistable behaviors. The synthetic biology approach has been used to develop bistable switches whose overall molecular mechanism is based on autocatalysis of a single gene [50,51]. These networks may be considered as synthetic models of their natural counterparts, such as signaling network controlling cell cycle regulation [52,53] and regulation of the lac operon [54]. Bistable switches can also be realized in a positive feedback system by combining negative regulations. A recent study on a synthetic 'toggle' switch, a two-component module in which two transcriptional repressors negatively regulate each other, is shown to achieve bi-stable switching behaviors [55]. A combination of negative or positive regulation between two or more components can give rise to oscillations. This was theoretically or experimentally characterized in *Escherichia coli* [56,57,58,59].

In addition to mono-stable, bi-stable, or oscillatory modules, network architectures with other connectivity have also been identified, and their properties and biological significance have been characterized [43,46,60,61]. Importantly, these modules often maintain similar function across different species. For example, oscillator modules are the molecular mechanisms that underlie molecular, physiological and behavioral rhythms [62,63] or pattern formations [64], and bistability modules may govern the cell's entry into the cell cycle and be responsible for controlling cell differentiation [52,65,66,67,68]. Thus, thorough analysis of a module in one context can provide insight into its functional roles under a wide spectrum of conditions.

11

1.2.3 Basic Kinetics

In kinetic modeling, a biological system is considered as a series of chemical reactions, whose kinetics can be described by rate expressions. The system is often composed of multiple reactions, which occur through direct interactions among reactants. If these interactions are elementary reactions, their rates can be modeled following the mass action law. That is, the reaction rate is proportional to the product of reactant concentrations. However, most biological models are frequently formulated as consisting of more complex reaction mechanisms. One important class is enzyme catalyzed reactions, which are critical for live systems where virtually all reactions are too slow to support life without enzymes. The enzymes provide a way to regulate reactions at appropriate rates and conditions.

A commonly used reaction model for enzymatic reactions is the Michaelis-Menten equation. In this reaction mechanism, one assumes that the enzyme is not consumed and the total concentration of enzyme stays constant. It only interacts directly with the substrate to form an enzyme-substrate complex, which leads to the synthesis of the product:

$$E + S \underset{k_r}{\overset{k_f}{\rightleftharpoons}} ES \xrightarrow{k2} E + P$$
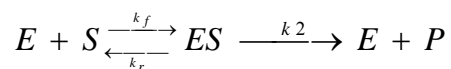
Assuming that the intermediate (ES) is at the quasi-steady-state and the substrate is in excess, we can derive the Michaelis-Menten equation:

$$\frac{dP}{dt} = \frac{V_{max}[S]}{K_M + [S]}$$

where V_{max} is the maximal reaction rate ($k_2[E_{Total}]$) and K_M is the Michaelis-Menten constant, ($\frac{(k_r + k_2)}{k_f}$).

Another recurring scheme in modeling cellular networks is the representation of gene expression. Expression of a single gene involves two basic steps: transcription and translation. This simplistic view of gene regulation starts with transcription, where the RNA polymerase binds the promoter of a gene to result in mRNA synthesis. The mRNA that carries coded information binds with ribosome, and the coded information is translated into protein (Figure 1.2A).
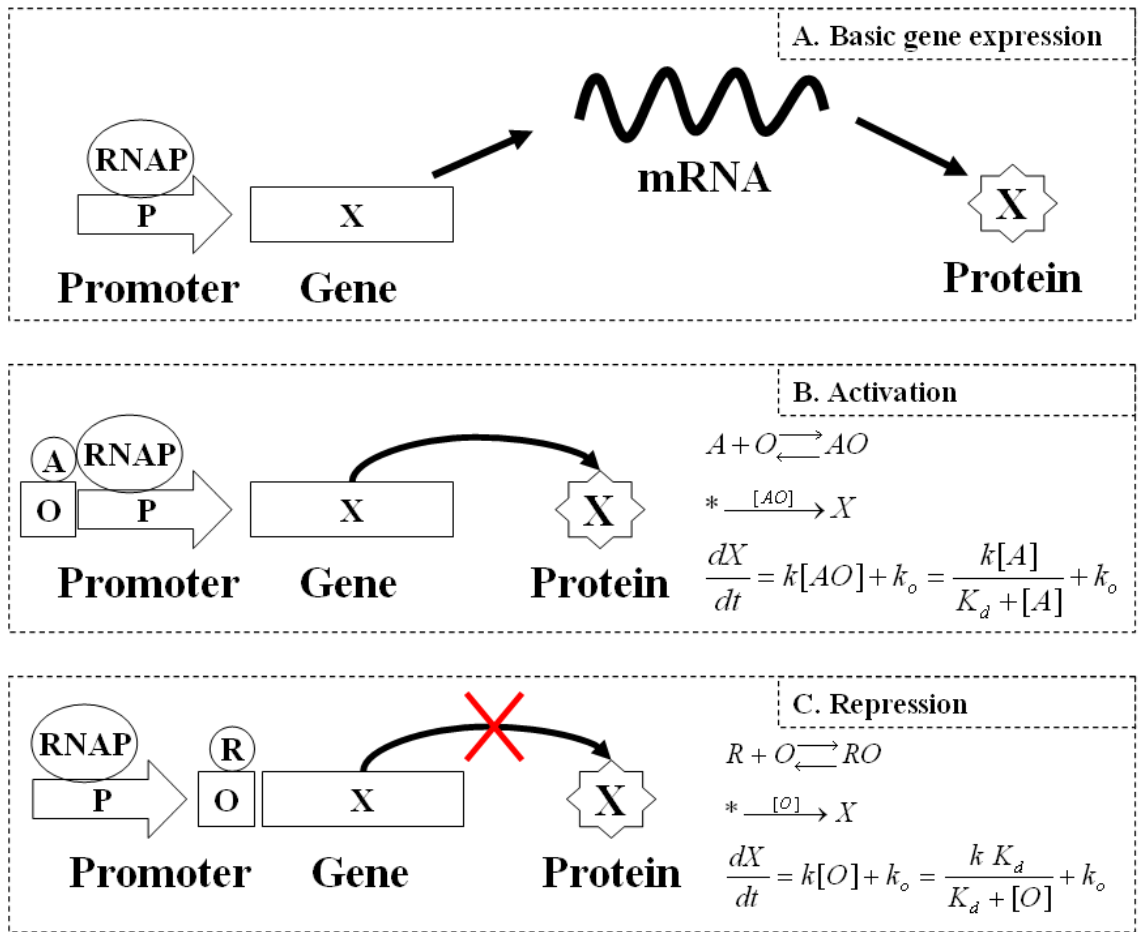
Figure 1.2: Modeling gene regulation. A simplified view of gene regulation is shown in (A). Initiation of the mRNA synthesis can be triggered by either transcription activation (B) or transcription repression (C), which can be mathematically represented with Michaelis-Menten type of kinetics.

In real systems, gene expression can be regulated at multiple layers involving interactions among inducers, repressors, and operator sites. The interactions of these components lead to two general categories of transcriptional regulations: activation and repression. When an activator binds to the operator site, this complex leads to recruitment of RNA polymerase (RNAP) and synthesis of mRNA (Figure 1.2B). In contrast, binding of a repressor will prevent initiation of transcription by blocking the RNAP. In the absence of cooperative interactions, such as dimerization and synergistic binding of transcription regulators to promoters, both types of regulation can be mathematically described by using Michaelis-Menten type of kinetics (Figure 1.2B & C).

If the transcription regulator acts as a dimer or multimer, and/or if it binds synergistically to multiple operator sites, transcription regulation can by modeled by higher-order expressions, such as the commonly used Hill kinetics:

$$\frac{dP}{dt} = \frac{V_{max}[S]^n}{K_M^n + [S]^n}$$

where n is called the Hill coefficient. For n=1, Hill kinetics is the same as the Michaelis-Menten kinetics. However, for the response curve that has a different slope from what is predicted by Michaelis-Menten kinetics, n can be adjusted to fit the Hill kinetics curve. Detailed treatment of this can be found in [69]

1.2.4 Deterministic Models

By treating each interaction as a chemical reaction, one can account for the

production and depletion of each species by using an ordinary differential equation. A

coupled system of ODEs that describes the dynamics of all elements in the network

constitutes an integrated kinetic model. The general form of ODE systems can be written

as:

$$\frac{dx_1}{dt} = f_1(x_1, x_2, x_3, \dots, x_n).$$

$$\cdot$$
$$\cdot$$
$$\cdot$$

$$\frac{dx_n}{dt} = f_n(x_1, x_2, x_3, \dots, x_n)$$

where $x_1, x_2, \dots x_n$ represent levels of different interacting species, and $f_1, f_2, \dots f_n$

represent their corresponding rate expressions.

This representation often implies that the system dynamics occur in a well-

stirred reactor in which bulk concentrations of the components are considered. Except

for simple systems, an ODE-based kinetic model is often solved numerically using

established methods [70,71,72,73]. Given the same parameter values, initial conditions,

and simulation settings (e.g. error tolerance), different rounds of simulations will

generate exactly the same temporal dynamics for each individual component. As such,

an ODE model is also called a 'deterministic' model. To assist in computational

16

modeling research, a wide range of computational methods and tools have been

developed for ODE systems [5,74,75,76,77,78,79]

1.2.5 Cellular Noise and Stochastic Methods

ODE-based models are widely used to model dynamics of both natural and

synthetic biological networks. For example, deterministic simulations predicted that a

synthetic gene circuit of transcriptional repressors would cause sustained oscillations in

gene expression [57]. Aspects of these predictions were verified in experiments where

individual cells carrying the circuit displayed oscillatory behavior. However, the real

system dynamics were quite stochastic when compared to simulation results from an

ODE model. Specifically, oscillations in the repressilator occurred in only ~40% of

individual cell lineages and were often out of phase with each other.

Such stochastic behavior is partially due to the intrinsically stochastic

biochemical reactions among small numbers of molecules. These fluctuations in gene

expression are often termed "noise" [80,81]. In general, sources of noise include

fluctuations in cellular components [33], transmitted noise from upstream genes [35],

and other cellular processes un-accounted for by the model. Recently, the origin and

propagation of noise in gene expression have been of central interest in many

experimental studies [32,36,82,83,84,85,86].

Presence of cellular noise presents both a challenge and an opportunity for

cellular function. On one hand, reliable function in the presence of noise requires

strategies that reduce impact of noise [87,88]. For instance, one such mechanism that

regulates noise is negative feedback where the output from a system reduces its own

output. In a biological context, this occurs when a protein inhibits its own expression by

binding to its promoter. This mechanism has been shown to reduce noise in gene

expression [48,89]. On the other, noise may be beneficial by serving as a source for

generating phenotypic diversity [88,90], which can facilitate adaptation to changing

environments or trigger cell differentiation [91].

Because of the important implications of noise for both natural and synthetic

cellular networks, it is often useful to model stochastic dynamics. For a well-stirred,

spatially homogeneous system, its stochastic temporal dynamics can be captured by a

generic chemical master equation (CME) [92]:

$$\frac{\partial P(\widetilde{x},t \mid \widetilde{x}_0,t_0)}{\partial t} = \sum_{j=1}^{M} a_j(\widetilde{x}-v_j) \times P(\widetilde{x}-v_j,t \mid \widetilde{x}_0,t_0) - a_j(\widetilde{x}) \times P(\widetilde{x},t \mid \widetilde{x}_0,t_0)$$

The first term of the equation describes probability of a species reacting at time t,

while the second term describes the probability of a species remaining in its current

state. x is a vector containing number of molecules for each species. $P(x,t)$ gives the

probability of the system in state x at time t. a_j is the propensity value of reaction j. v_j is a

18

vector containing the changes in state x caused by reaction j. x_0 and t_0 is the initial state and time respectively.

One can solve the CME analytically only for very simple systems. As the system size increases beyond a few reactions, the analytical solution of the CME becomes intractable. When the number of reactions and the number of molecules increases, the number of possible paths increases exponentially. Gillespie proposed a simple algorithm to solve the CME numerically using a Monte-Carlo method [93]. In this formulation, each reaction is assumed to be an elementary reaction, where collisions between reactant molecules directly lead to formation of products. The probability that a reaction happens is dependent on its reaction propensity (Table 1.2), which is analogous to a rate expression in ODE-based models.

Table 1.2: Reaction propensity for stochastic methods

Reaction	Propensity
A→B	$c*x_A$
A+B→C	$c*x_A*x_B$
2A→B	$c*x_A*(x_A-1)/2$

The reaction propensity describes the probability of one molecule colliding with another molecule, which leads to firing of a chemical reaction. Note that the reaction

propensity for dimerization reactions is equal to c*x_A*(x_A-1)/2 rather than c*x_A*x_A because a molecule cannot react with itself. This presents a consistent interpretation of stochastic rate constants. Stochastic rate constants are normally calculated from conventional rate constants [93]. Given the reaction propensity, we can now define the state of a species at time t. In order to follow the evolution of the states through time, we have to calculate which reaction (μ) is firing at time t and how much time (τ) the reaction requires. The probability of the firing event is shown in Equation (6.6) and it can be calculated by using the schemes illustrated in Table 1.3.

$$P(\tau, \mu) = a_j \times \exp(-a_0 \times \tau)$$

Table 1.3: Pseducode for Gillespie Algorithm adapted from Gillespie [93]

1.	Calculate $a_0 = \sum_{j}^{M} a_j$
2.	Generate two random numbers r_1 and r_2 from the uniform distribution (0, 1).
3.	Compute $\tau = \dfrac{1}{a_0} \ln(\dfrac{1}{r_1})$
4.	Compute μ that satisfies $\sum_{j}^{\mu} a_j \geq r_2 \times a_0 \geq \sum_{j}^{\mu-1} a_j$
5.	Execute reaction μ and advance time t by τ

Despite its simplicity, the computational cost of the Gillespie algorithm increases drastically with the number of reaction and the number of molecules in a system. The increment in computational cost is primarily due to the generation of random numbers (Step 2 in Table 1.3) and the enumeration of reactions to determine the next reaction (Step 4 in Table 1.3). For example, when the number of molecules is equal to 1e6, τ will become excessively small (order of 1e-6) which then increases the number of time-steps.

In order to simulate large scale stochastic model, Gibson [94] proposed the Next Reaction method to improve computational efficiency of the Gillespie algorithms. The first improvement involves implementation of a tree data structure to store the reaction time of each reaction, which minimizes enumeration of the reactions at every time-step. The second improvement uses a map data structure to minimize recalculation of the reaction propensity at every time step. The Gibson algorithm is significantly faster than the Gillespie algorithms for systems consisting of many reactions and many reacting species. It is also an exact algorithm in the sense that it satisfies the same basic assumptions as required by the Gillespie algorithm.

Several other algorithms were also proposed to improve computational speed of stochastic simulations. These algorithms are not exact and require users to pre-determine an extra parameter that affects accuracy of the numerical solutions. Tau-leap algorithms [95] predict multiple firing of fast reactions and hence, reduce the total

number of time-steps. Another class of algorithms will be the hybrid algorithms [88,96] which model fast reactions subsets using either ODEs or Langevin equations (see below) while treating the slow reaction subset with the stochastic algorithms.

An alternative, widely used stochastic method, remains in the framework of differential equations by adding an appropriate noise term to each of the ODEs that describe the biological network. The resulting stochastic differential equations (SDEs) can then be solved numerically. Different formulations of SDEs can be established for different types of simulation applications. With appropriate assumptions, one can obtain a special type of SDE, the chemical Langevin equation [97], which has been used to model a variety of cellular networks.

$$\frac{dX_i(t)}{dt} = \sum_{j=1}^{M} v_{ji} a_j(X(t)) + \sum_{j=1}^{M} v_{ji} a_j^{1/2}(X(t)) \Gamma_j(t)$$

where $X_i(t)$ is the number of molecules of a molecular species in the system at time t and i refers to the specific molecular species $(i = 1,...,N)$. $X(t) \equiv (X_1(t),...,X_N(t))$ is the state of the entire system at time t, $a_j(X(t))$ is the rate for a specific reaction or molecular interaction $(j = 1,...,M)$. v_{ji} is a matrix describing the change in the number of molecules as a result of one molecular interaction. In other words, interactions that result in the synthesis of $X_i(t)$ are added and interactions that result in the degradation

of $X_i(t)$ are subtracted. $\Gamma_j(t)$ are temporally uncorrelated, statistically independent Gaussian white noises.

SDEs are attractive in that they are computationally more efficient that the Gillespie algorithm and its derivatives. Also, by remaining in the framework of differential equations, they can facilitate in-depth analysis of system dynamics without always resorting to numerical simulations [98].

Regardless of the exact formulation of a stochastic algorithm, repeated rounds of stochastic simulations will generate different temporal dynamics for each individual species. One often uses an ensemble of simulated time course to gain insight into noise characteristics, as well as how they are impacted by regulatory mechanisms. One way of quantifying noise in gene expression is to normalize the standard deviation of the protein level with respect to the average protein level ($\gamma = \frac{\sigma}{\mu}$), where σ is standard deviation of protein level and μ is the mean of protein level [33]. While this metric is direct and intuitive, some noise characteristics may be obscured by the more dominant small-number effects [87]. This may make it difficult to compare the noise of proteins that are being expressed at different levels. In this case, a more advantageous metric of quantifying noise is noise strength, or the variance of the protein level normalized with respect to the average protein level, $\zeta = \frac{\sigma^2}{\mu}$. Since gene expression is often controlled through transcription factors, noise levels can be compared among different genes regardless of their expression levels. This metric was recently used to analyze relative

23

contribution of transcription rates and translation rates to the noise characteristics of

final protein products [98].

1.2.6 System Analysis Techniques

Given an integrated model, one can characterize the system behavior using

various analysis techniques, such as parametric sensitivity analysis and bifurcation

analysis. These techniques allow for exploration of potential system dynamics and

provide quantitative insights into emergent system behavior, such as robustness. Such

information is useful for revealing "design principles" of natural biological systems or

guiding design and implementation of synthetic gene circuits

1.2.6.1 Parametric Sensitivity Analysis

Sensitivity analysis is used to quantify changes in system behaviors in response to

parameter changes. Different parameter may have varying impact on the system

dynamics and the degree of the impact can be quantified by a sensitivity value. A

general method for computing the sensitivity value for an ODE system is:

$$s(I; \phi_j) = \frac{\partial I}{\partial \phi_j} = \lim_{\Delta \phi_j \to 0} \frac{I(\phi_j + \Delta \phi_j) - I(\phi_j)}{\Delta \phi_j}$$

where the sensitivity value is the ratio of change in the objective function of interest (I) to change in a parameter (ϕ_j).

Alternatively, the normalized form of sensitivity can be defined:

$$S(I;\phi_j) = \frac{\phi_j}{I} \bullet \frac{\partial I}{\partial \phi_j} = \frac{\partial \ln I}{\partial \ln \phi_j} = \frac{\phi_j}{I} \bullet s(I;\phi_j)$$

This is also called logarithmic sensitivity. It is commonly used in metabolic control analysis [99], and has the feature of being dimensionless.

The objective function of interest is determined by the goals of the analysis. In the enzymatic synthesis of product that follows Michaelis-Menten kinetics, one may be interested in the change in the synthesis rate or in the steady-state product concentration as the Michaelis-Menten constant is varied. Therefore, there may be more than one sensitivity value for a given parameter. For an extensive treatment of sensitivity analysis, refer to [100].

Sensitivity analysis has been widely used in quantifying robustness of complex biological systems with respect to parametric perturbations [101,102,103,104,105]. In a complex system with a large number of parameters, the system behaviors may be robust to changes in various parameters. Especially, feedback controls and backup or compensation mechanisms in biological systems confer additional layers of robustness [9,106,107,108]. Accurate identification of the underlying mechanisms for such

robustness is challenging, since the system behaviors result from both parameters and system architecture. By distinguishing the impact of parameters from that of the architecture, sensitivity analysis provides a way to characterize system robustness. Such mathematical exploration of various system behaviors may serve as a guide in realizing system behaviors as desired experimentally. Specifically, if the parameters with high sensitivity values can be controlled, higher efficiency in biologically feasible experiment designs and data analysis can be achieved.

1.2.6.2 Bifurcation Analysis

While sensitivity analysis provides a quantitative measure on the dependence of system dynamics on parameters, bifurcation analysis focuses on a qualitative understanding of the system dynamics. Similar to sensitivity analysis, bifurcation analysis monitors change in system behaviors in response to parameter change also, except the goal is to explore qualitative changes in the steady-state solution due to parameter changes. Bifurcation analysis is performed by varying a parameter until a qualitative change in dynamics is observed. The value at which this occurs is called the bifurcation point.

A quantitative measure of the stability can be achieved by a simple analytical method called linear stability analysis. This method provides a numerical value for the rate of decay to the stable steady-state solution from a small perturbation. Let us

consider a model consisting of only one species, $\dfrac{dx}{dt} = f(x)$. Linear stability analysis

begins with steady-state solutions (x_s), which can be found by equating the right-hand

side of the ODE expression to 0 and solving for the species concentration of interest.

Adding a small perturbation, $x = x_s + \delta(t)$, the right-hand side becomes:

$$f(x) = f(x_s + \delta(t)) \xrightarrow{\;Taylor\text{'s expan}\;sion\;} f(x_s) + \delta(t)f'(x_s) + O(\delta(t))^2$$

Assuming that the higher order terms ($O(\delta(t))^2$) are negligible and since $f(x_s)$ is 0, the

system at steady-state responds to small perturbations as $f(x) \approx \delta(t)f'(x_s)$. Since $\frac{dx}{xt}$,

the left-hand side of the ODE equation, is equal to $\frac{d(xs+\delta(t))}{dt} = \frac{d\delta(t)}{dt}$, the growth rate of

perturbations is:

$$\frac{d\delta(t)}{dt} = \delta(t)f'(x_s)$$

Therefore, the perturbation will grow exponentially if $f'(x_s)$ is positive, and decays

exponentially if $f'(x_s)$ is negative.

A bivariate system can be treated in a similar manner. For example, consider:

$$\dot{x} = f(x, y), \qquad x = x - \delta_x(t)$$
$$\dot{y} = g(x, y), \qquad v = y - \delta_y(t)$$

where $\delta_x(t)$ and $\delta_y(t)$ denote a small disturbance from the steady-state solutions. Using

the Taylor's expansion similar to the first order system, we can approximate the growth

rate of perturbations to be:

$$
\begin{pmatrix} \dot{\delta}_x \\ \dot{\delta}_Y \end{pmatrix} = A \begin{pmatrix} \delta_x \\ \delta_y \end{pmatrix}, where \ A = \begin{pmatrix} \dfrac{\partial f}{\partial x} & \dfrac{\partial f}{\partial y} \\ \dfrac{\partial g}{\partial x} & \dfrac{\partial g}{\partial y} \end{pmatrix}_{(x_s,y_s)} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}
$$

A is the Jacobian matrix at a steady state. The exponents of the growth rate are

determined by eigenvalues λ of the matrix A, given by the characteristic equation det($A-$

λI), where I is the identity matrix. Defining $\tau = trace(A) = a+d$ and $\Delta = det(A) = ad-bc$, the

eigenvalues are:

$$
\lambda_1 = \frac{\tau + \sqrt{\tau^2 - 4\Delta}}{2}, \quad \lambda_2 = \frac{\tau - \sqrt{\tau^2 - 4\Delta}}{2}
$$

Since the real part of an eigenvalue determines the rate at which the perturbation grows,

the real part of both eigenvalues must be negative for the steady-state solutions to be

stable. General analysis for yet more complex biological systems can be found in [109].

28

Varying the parameter of interest can create or destroy steady-state solutions, and the properties of these solutions can change. At bifurcation points where network behaviors undergo a qualitative change, a stable steady-state solution may become unstable or vice versa. Also, a stable steady-state solution may diverge to two or no steady-states. For an extensive treatment of bifurcation analysis, refer to [110].

1.3   Case Studies

To illustrate the basic concepts and techniques outlined above, we here provide examples of kinetic modeling and analysis using three simple biological systems: expression of a single gene, a phosphorylation-dephosphorylation cycle composed of enzymatic reactions, and a synthetic population control circuit.

1.3.1 Expression of a Single Gene

Although gene expression is a complicated process that involves a number of components, we use the simplistic view as shown in Figure 1.3A. Key assumptions in this view are that transcription of mRNA is constitutive with rate k, and that translation of protein depends on the concentration of mRNA. Although the choice of parameters depends on many factors such as the gene of interest and the internal and external environment of gene expression, commonly accepted estimation of parameters is sufficient for our gene expression model. Based on simplification and estimated

parameters, mathematical models are constructed using ODE, SDE, and stochastic

methods as shown in Table 1.4. These models are implemented and simulated in a

graphic-based simulator *Dynetica* [111] (Figure 1.3B). Also see

(http://labs.genome.duke.edu/YouLab/software/dynetica/index.php).

Table 1.4: Comparison between mathematical representation schemes for the gene expression.

ODE	SDE	Gillespie
Reactions $$* \xrightarrow{k} mRNA$$ $$* \xrightarrow{k[mRNA]} protein$$	**Reactions** $$* \xrightarrow{k} mRNA$$ $$* \xrightarrow{k[mRNA]} protein$$	**Reaction probability intensities** $$* \longrightarrow mRNA \qquad k_R * N * V$$ $$mRNA \longrightarrow * \qquad d_R * N_R$$ $$* \longrightarrow protein \qquad c_P * N_R$$ $$protein \longrightarrow * \qquad d_P * N_P$$
Ordinary Differential Equations $$\frac{d[mRNA]}{dt} = k_R - d_R[mRNA]$$ $$\frac{d[protein]}{dt} = k_p[mRNA] - d_p[protein]$$	**Stochastic Differential Equations** $$\frac{d[mRNA]}{dt} = k_1 - k_3[mRNA] + \sqrt{k_1}\Gamma_1(t)$$ $$\qquad - \sqrt{k_3[mRNA]}\,\Gamma_2(t)$$ $$\frac{d[protein]}{dt} = k_2 - k_4[protein] + \sqrt{k_2}\Gamma_3(t)$$ $$\qquad - \sqrt{k_4[protein]}\,\Gamma_4(t)$$	**Gillispie Algorithm** • Calculate $a_0 = \sum_j^M a_j$ • Generate two random numbers r1 and r2. • Compute $\tau = \frac{1}{a_0}\ln(\frac{1}{r_1})$ • Compute μ that satisfies $\sum_j^{\mu} a_j \ge r_2 \times a_0 \ge \sum_j^{\mu-1} a_j$ • Execute reaction μ and advance time t by τ

30

A) Mathematical Model

$$\frac{d[P]}{dt} = k_p * [M] - \gamma_P * [P]$$

$$\frac{d[M]}{dt} = k_R - \gamma_R * [M]$$

B) Implementation in Dynetica
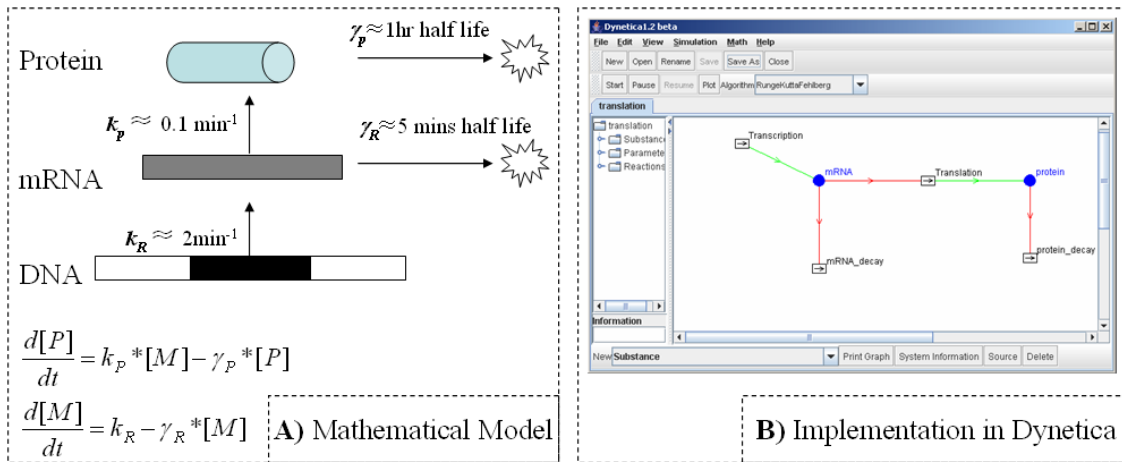
Figure 1.3: A mathematical model is constructed based on our knowledge of the single gene expression and typical reaction parameters (A). The model is implemented in a simulation and analysis software, *Dynetica*
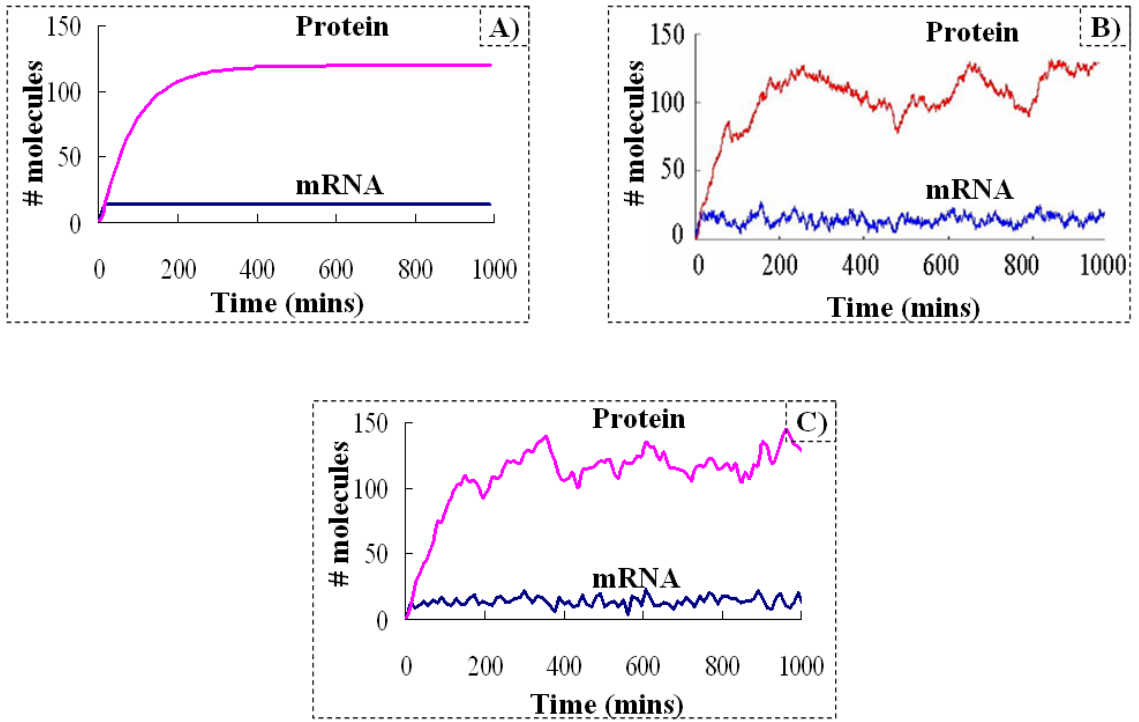
Figure 1.4: Simulation results of the model shown in Figure 1.3 by deterministic (A), SDE (B), and stochastic(C) formulations.

As shown by simulation results in Figure 1.4, the stochastic simulations generated

dynamics overall similar to that from a deterministic simulation, but their dynamics are

noisy. The deterministic simulation also reveals that mRNA synthesis reaches steady-

state faster than protein production. Assuming steady-state for mRNA synthesis, we can

carry out stability analysis of a steady state for gene expression. Equating the right-hand

side of the ODE expression for mRNA in Table 1.4 to 0, we find the mRNA level at the

steady-state to be $\frac{k_R}{d_R}$. Then, the protein expression at the steady-state concentration of

mRNA can be re-written as $\frac{d[protein]}{dt} = \frac{k_P * k_R}{d_R} - d_P[protein]$. When the decay rate

($d_P[protein]$) matches the synthesis rate ($\frac{k_P * k_R}{d_R}$) at steady-state, the system is stable.

From equation (6.10), we can calculate the exponent for the growth rate of

perturbation $\left(\frac{\partial f}{\partial P}\right)_{P=P_{SS}} = -d_P < 0$, where f is the right hand side of the rate equation at

the steady-state, and P is the protein level and P_{ss} is the steady-state protein level. Since

$f'(P_{ss})$ is negative, any perturbation around the steady-state will decay at the rate of d_P,

indicating that the steady state is globally stable.

1.3.2. A Phosphorylation-Dephosphorylation Cycle

Increasing in complexity, we analyze transient and steady-state behaviors of an

enzyme-mediated phosphorylation cycle, which has been shown to demonstrate ultra-

sensitivity when the enzymes operate outside the region of first-order kinetics [112]. To

construct a mathematical model, we begin with the conventional enzyme catalysis

scheme where a protein switches between its phosphorylated and dephosphorylated

forms (Figure 1.5). Assuming the enzymatic reactions follow the Michaelis-Menten

kinetics and the total protein concentration is constant, we develop two ODE equations

which are implemented and simulated in *Dynetica*. Since the goal of modeling here is to

identify general system behaviors of a phosphorylation cycle, we employ a set of

biologically feasible parameters.
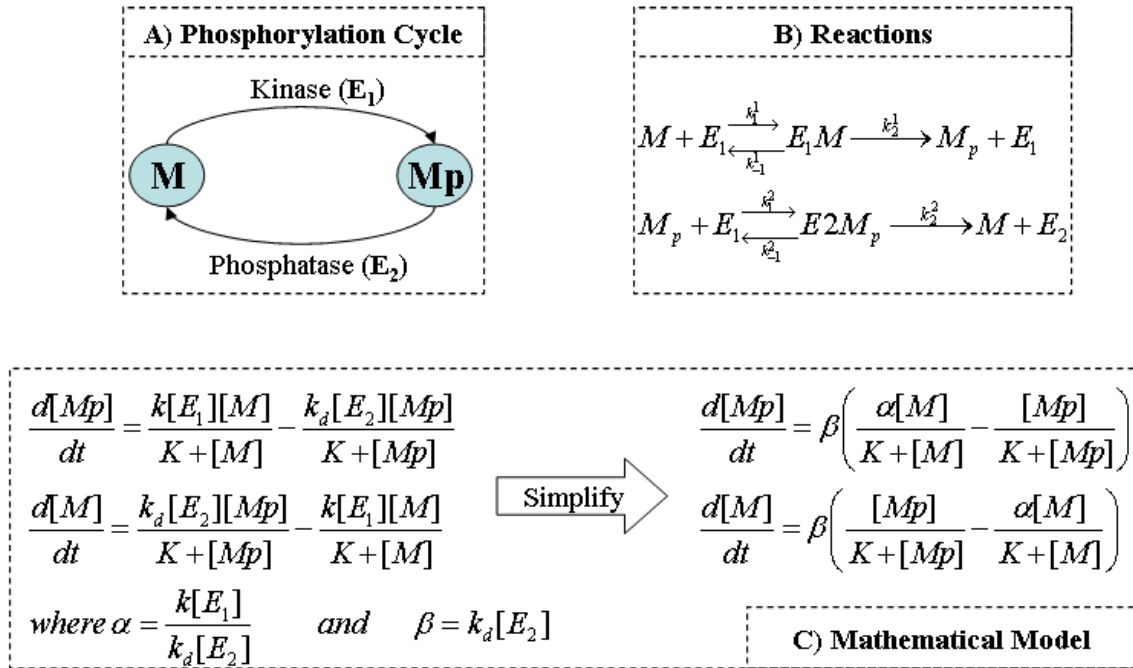
Figure 1.5: Modeling a phosphorylation-dephosphorylation cycle. Enzymatic modification cycle (A) of a protein between an unphosphorylated state and the phosphorylated state is mathematically modeled. Reaction schemes in (B) are converted to a set of ODEs based on two assumptions: 1) Michaelis-Menten kinetics for the enzymatic reactions and 2) constant total level of the protein.

Assuming that the system starts with all protein in the unphosphorylated state, the protein will switches from the unphosphorylated state to the phosphorylated state over time, leading to a steady-state distribution of the protein in the two forms (Figure 1.6A). This process is sensitive to α, a ratio between phosphorylation and dephosphorylation rates. When α is small, the amount of phosphorylated protein at the steady-state is insignificant. However, more protein is converted as α becomes large. With very large α, the phosphorylation cycle becomes virtually irreversible, favoring the phosphorylated state.
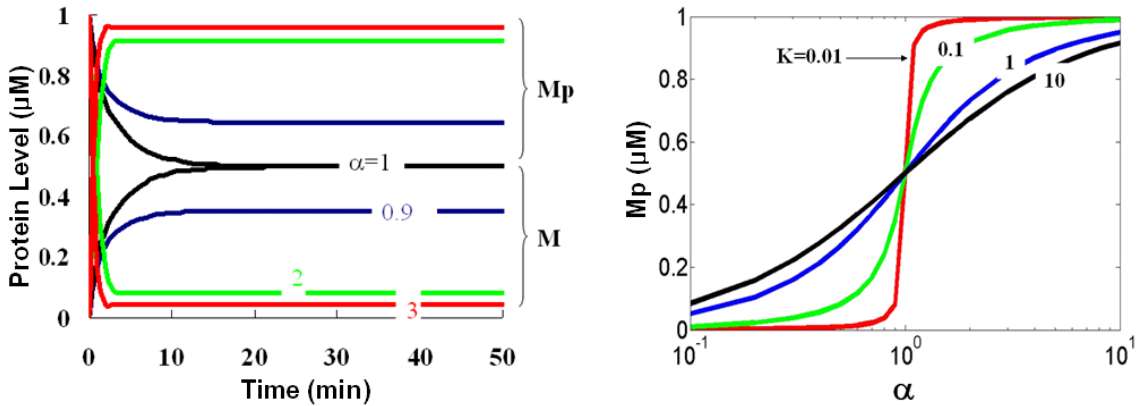
Figure 1.6: Simulation results for the model in Figure 1.5. Time-course results at varying α values show the dependence of conversion on the rate of phosphorylation and dephosphorylation (A). Protein conversion becomes ultra-sensitive for a sufficiently small Michaelis-Menten constant near $\alpha=1$, while the sensitivity becomes weaker as K is increased (B).

The sensitivity analysis in Figure 1.6B shows the dependence of conversion on α. As K, a ratio of Michaelis-Menten constant to the total protein concentration, approaches 0,

the dependence of conversion is ultrasensitive near α equal to 1. Then, the rate equation

for the protein phosphorylation becomes $\frac{dMp}{dt} = \beta(\alpha - 1)$, a zero order rate expression that

does not depend on any concentration. This dynamics is thus called zero-order

ultrasensitivity. When the Michaelis-Menten constants are comparable to the total

protein concentration (large K), the rate expression is first order and the ultrasensitivity

at $\alpha = 1$ becomes weaker. The time-courses and sensitivity analysis in Figure 1.6 reveal

two critical conditions to achieve ultrasensitivity: 1) α has to be near 1 and 2) the total

protein concentration must be much greater than the Michaelis-Menten constants. That

is, both kinase and phosphatase operate near saturation so that the overall reaction rate

does not have a linear dependence on protein concentration.

We note that modeling can facilitate discovering design principles in biological

systems. For example, ultrasensitivity is utilized in biological systems when sharp

switching behavior is desired. A study of the mitogen-activated protein kinase (MAPK)

pathway that combined both simulations and experiments has demonstrated

ultrasensitivity. Their work illustrates that the phosphorylation cycle mechanism under

the two conditions is sufficient to generate a sharp switching behavior, whose the Hill

coefficient is estimated to be 5 [113]. At least two ways by which biological systems take

advantage of ultrasensitivity can be speculated. In one scenario, a minor change in input

will result in significant output when the system is operating near $\alpha = 1$. In the other

scenario, a significant change in the input will have little impact on the output when α is

much smaller or larger than 1. This may be useful in dealing with noisy signals, allowing the system to filter out noise [114].

1.3.3. A Synthetic Population Control Circuit

In addition to revealing dynamics of natural systems, modeling has become an indispensable tool for designing synthetic circuits [91,115,116,117,118,119,120]. To illustrate this, we take as an example the synthetic population control circuit that we recently engineered [6,121]. This circuit is based on a combination of two well-characterized modules: a quorum-sensing module and a killing module. Generally, we can develop an intuition on the circuit behavior, as the design is based on a combination of previously characterized modules. For example, the quorum-sensing module allows for cell-cell communication, where the cell density is broadcasted and detected by elements in the module. When the quorum-sensing module is coupled with a killing module, detection of high cell density by the quorum-sensing module activates killing of the cells. More specifically, the signal that diffuses across cell membranes to mediate communication is a small acyl-homoserine lactorne (AHL) molecule synthesized by the LuxI protein. At high cell density, the AHL accumulates inside the cells and in the extracellular medium. At sufficiently high concentrations, it activates the LuxR transcriptional regulator, which in turn activates expression of the killer gene (E) under the control of a luxI promoter (pluxI). Accumulation of the killer protein causes cell

37

death. Based on this qualitative understanding of the programmed population control

circuit, a set of ODE equations are formulated (Figure 1.7). The model is implemented,

simulated and analyzed in *XPP-AUT* [122].
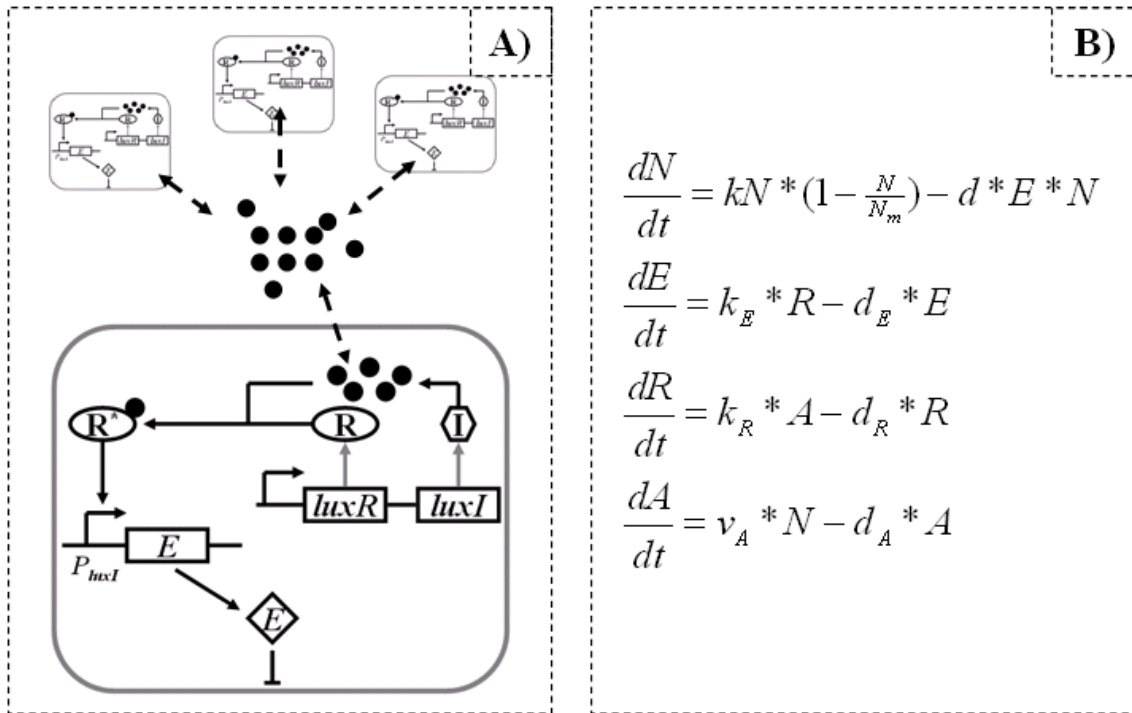
Figure 1.7: Modeling a synthetic population control circuit. A) A qualitative understanding of the circuit dynamics is illustrated in a graphical form. B) Based on the qualitative formulation, a quantitative representation of the programmed population control circuit is derived with several assumptions. More details on the assumptions and the parameters can be found in [6,121]
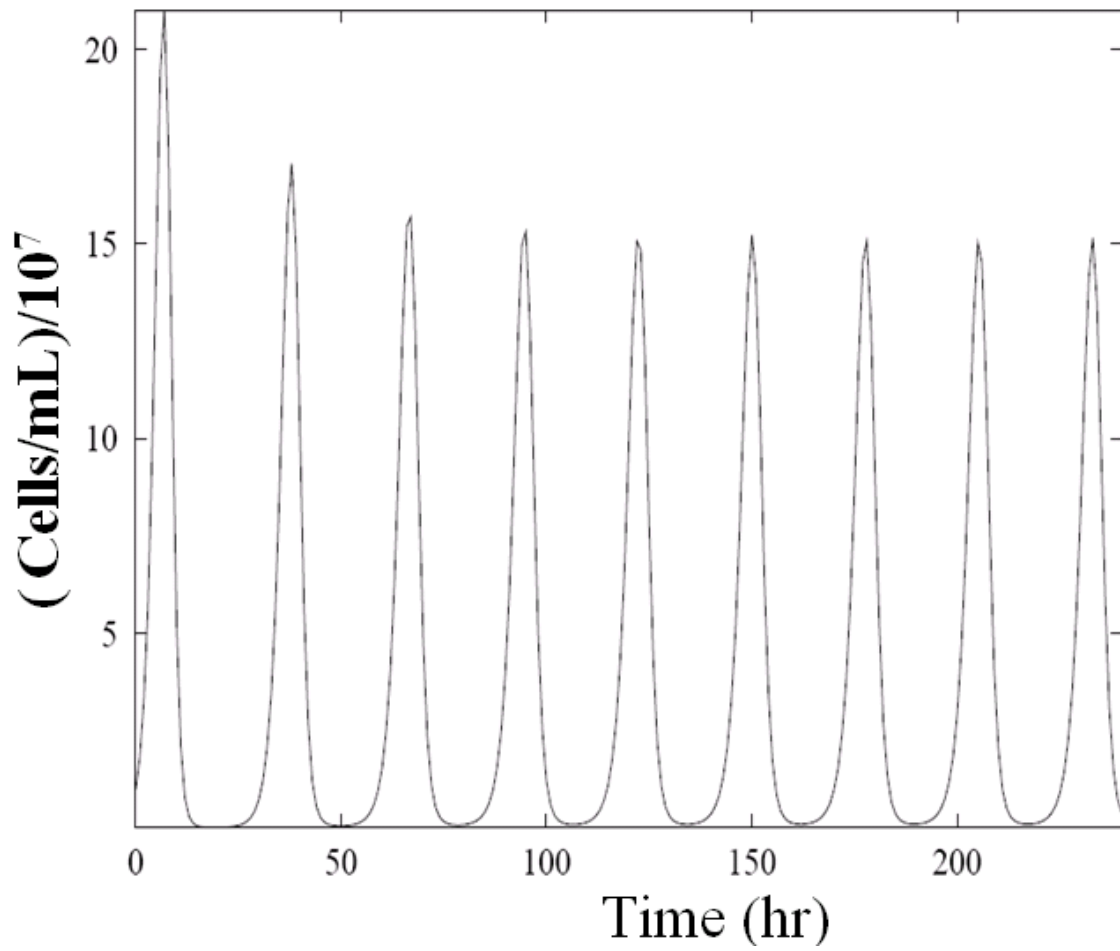
Figure 1.8: Oscillation in the cell density over time for appropriate parameter values.

This system has four components and has a negative feedback control on the cell density by the killer protein. Such structure has a potential to generate complex dynamics including oscillations. For certain biologically feasible parameters, our analysis shows that the model can indeed generate sustained oscillations over time. This prediction is consistent with experimental observations [121]. Further system stability analysis indicates that for $N \ll N_m$, there are two steady-state solutions. While the trivial steady-state is always unstable, the non-trivial steady state is stable if degradation rates of LuxR, the killer protein, and the AHL signal, and the microchemostat dilution rates are sufficiently large. However, decreases in these parameters destabilize the non-trivial steady-state, leading to oscillations. This trend is captured in Figure 1.8. For each of these parameters, bifurcation analysis is carried out using *XPP-AUT* (Figure 1.9A). In Figure 1.9B oscillations are observed for d_A less than 0.35 and the amplitude of the oscillations is the difference between the top and the bottom curves. High values of d_A (> 0.35) stabilize the system and the magnitude of oscillations decreases until damped oscillations occur (Figure 1.9C). Further increases in d_A lead to stronger dampening of the oscillations that eventually eliminate oscillations (Figure 1.9D). Similar stability analysis is carried out for the other parameters and similar behaviors of the non-trivial steady-state solution are observed (Figure 1.10).
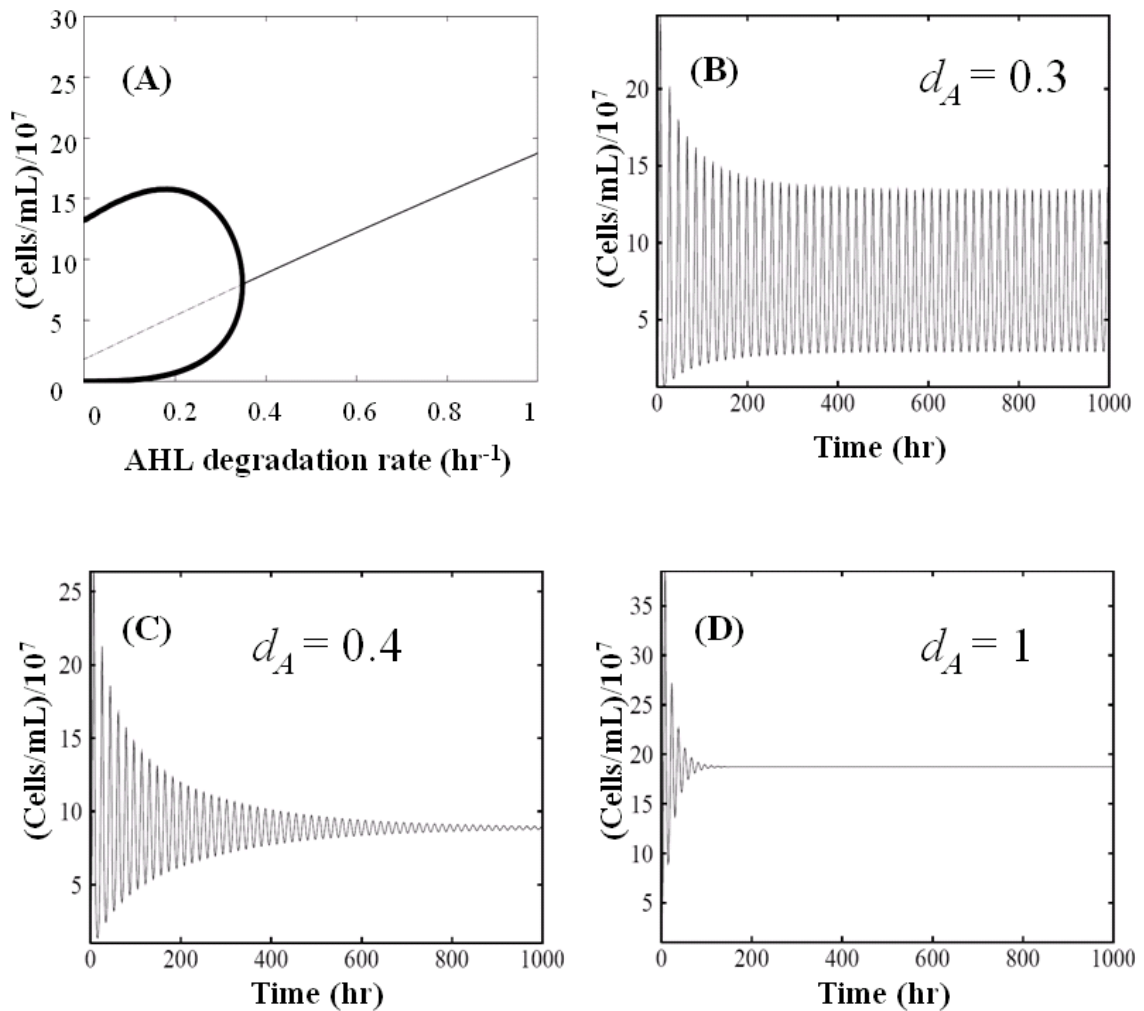
Figure 1.9: Bifurcation analysis. (A) Qualitative changes are observed in the dynamics as the AHL degradation rate (d_A) is varied. When d_A is sufficiently small (≤ 0.35), oscillation in cell density is observed (B). The population undergoes damped oscillation in cell density for increased d_A (C), and further increase in d_A stabilizes cell density at a faster rate (D).

41
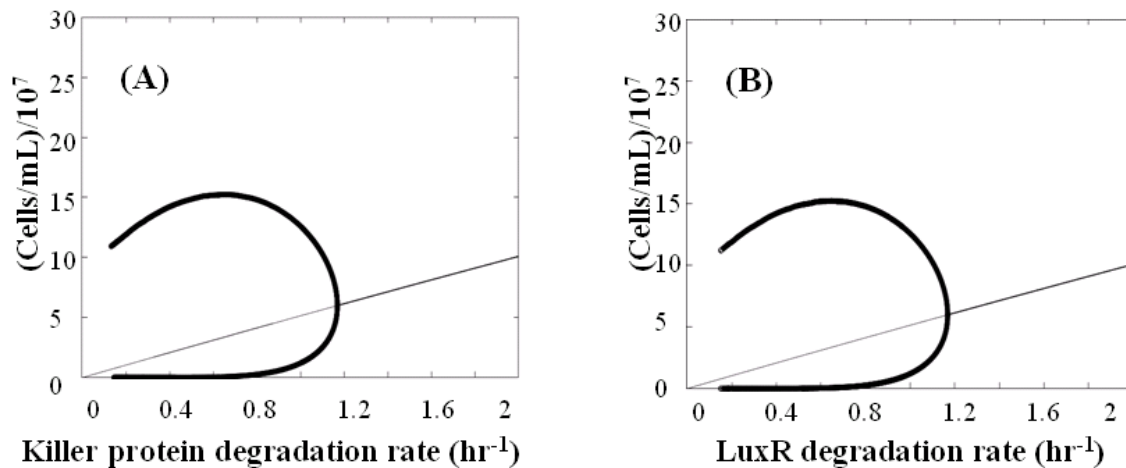
Figure 1.10: Further bifurcation analysis with rates for killer protein degradation (A) and LuxR degradation (B) is performed. Oscillations at sufficiently smaller rates diminish as the rates increase.

1.4   Discussion

We have used relatively simple, well-characterized systems to illustrate

construction and analysis of kinetic models. In these simple examples, we have

demonstrated the significance of kinetic modeling not only for improved understanding

of biological systems but also for improved predictions on cellular response to

perturbations. We note that mathematical modeling is not limited to simple systems

only, but has been used in more complex systems also. Successful application of

modeling has been demonstrated by numerous studies. The increase in the complexity

of modeled systems suggests wider applicability of mathematical modeling. Integrated

understanding of complex systems, whose dynamics cannot be conceptualized by

intuition alone, can be achieved in a quantitative manner. Also, improved predictive

power is particularly promising in the developments of therapeutics, where system-level

understanding is essential to minimize side-effects and to precisely predict drug effects.

Finally, modeling of cellular networks has become an integral part of the nascent field of

synthetic biology. The combination of design, modeling, experimental implementation,

and characterization of synthetic circuits or modules can provide substantial insights

into the design principles of more complex natural biological systems and assist in the

creation of artificial systems for practical applications.

Chapter 2: Introduction

To demonstrate use of the mathematical techniques described in Chapter 1, we focus on mammalian cell's entry into the cell cycle in the subsequent chapters. The overarching goal is to gain an integrated understanding of the signaling transduction networks that underlie mammalian cell cycle entry. This can serve as a foundation to a more in-depth view of cell signaling pathways. In addition, an integrated understanding of cell cycle entry dynamics may shed light on novel strategies for cancer therapy. Before an in-depth analyses of signaling transduction networks, we provide general introduction to regulation of the mammalian cell cycle in Chapter 2.

2.1 The mammalian cell cycle

The mammalian cell cycle consists largely of two stages: DNA replication (interphase) and segregation of the replicated chromosome into two separate daughter cells (mitosis or M phase). In the beginning of the interphase, the cell undergoes a gap phase (G1), where it grows and prepares for DNA replication. Once the cell is ready, it synthesizes a fresh copy of the DNA. This specific stage is called S phase. The S phase is followed by another gap phase (G2), where the cell prepares for mitosis. Often times, the cell can exit out of this cycle and enter into a resting, quiescent stage (G0). It can be induced to reenter into the cell cycle upon growth stimulation.
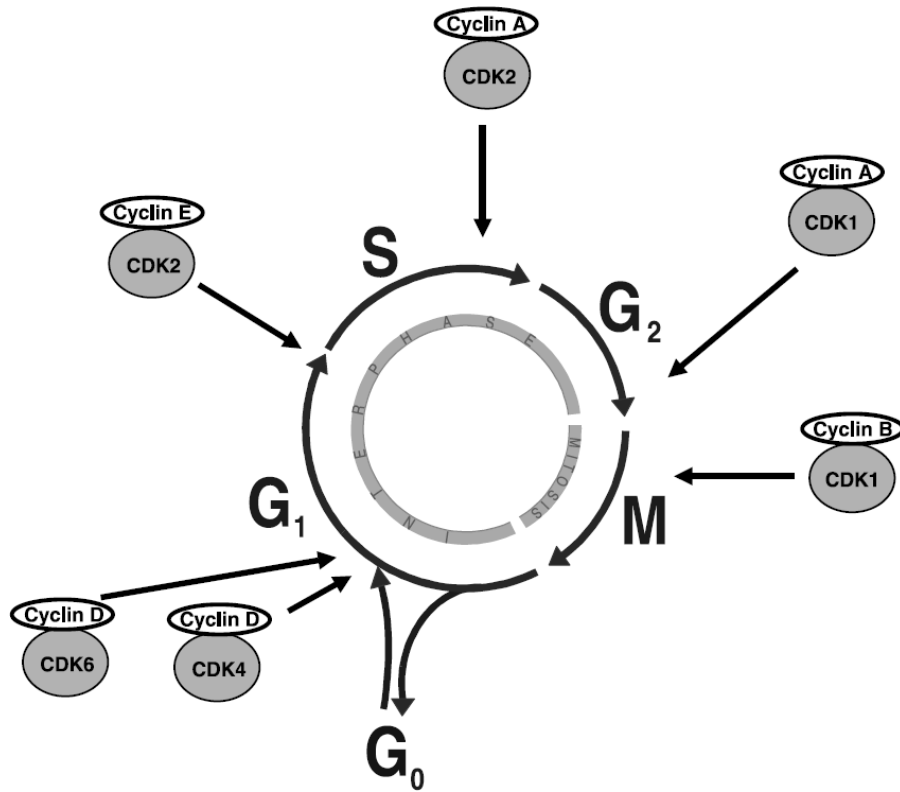
Figure 2.1: The mammalian cell cycle. The cell cycle is largely divided into two stages: mitosis (M) and interphase (G1/S/G2). The cell can exit the cell cycle and enter into the quiescent G0 phase. Appropriate stimulation may induce the cell to reenter the cell cycle (Adapted from [123]).

2.1.1 Control of the mammalian cell cycle by cyclin-dependent kinases (Cdks)

The mammalian cell cycle is intricately orchestrated by cyclin-dependent kinases (Cdks) and their binding partners, cyclins [124]. To date, a number of cyclins and Cdks have been identified. In yeast, only one major Cdk is expressed and its activity oscillates during cell-cycle progression. This oscillation in the Cdk activity is not due to oscillating Cdk level, but due to oscillating cyclin levels. By interacting with different cell-cycle stage-specific cyclins, the oscillating Cdk activity is believed to drive diverse cell-cycle transitions in the interphase (G1/S, S, and G2/M phases). In higher eukaryotic cells, many homologues of yeast Cdk have been identified, suggesting that regulation of various cell-cycle transitions is much more complex (Figure 2.1 and Table 2.1). In fact, human cells contain multiple loci encoding Cdks and cyclins (13 and 25, respectively).

Table 2.1: Cyclin-Cdk complexes at various stages of the cell cycle

Cell cycle phase	Cyclin	Cdk
G1 phase	CycD1, CycD2, CycD3	Cdk4, Cdk6
G1/S transition	CycE	Cdk2
S phase	CycA	Cdk2
G2/M transition	CycA	Cdk1
M phase	CycB	Cdk1

Despite many cyclins and Cdks, only a small number of cyclin/Cdk complexes have been shown to be directly involved in driving the mammalian cell cycle progression [125,126]. In early G1 phase, mitogenic signals are sensed by D-type cyclins (CycD) that bind and activate Cdk4 and Cdk6. Activation of these complexes initiates phosphorylation of the retinoblastoma protein (Rb) family including Rb, p107, and p130 [127]. This results in release of a transcription factor E2F, which is initially bound by the Rb protein. The released E2F activates and transcribes E2F-responsive genes required for cell cycle progression [128], including E-type cyclin (CycE) and A-type cyclin (CycA) [129]. In the late G1 phase, CycE forms a complex with Cdk2, and this complex further phosphorylates the Rb protein, leading to additional activation of E2F mediated transcription. In addition to phosphorylating Rb proteins, CycE-Cdk2 is thought to be essential for initiating DNA replication by facilitating loading of the MCM chromosome maintenance proteins onto origins of replication. These coordinated activity of Cdk2, 4, and 6 with various cyclins drives the cell to traverse through the G1/S transition into the S phase.

CycE availability is tightly controlled and is limited to the early stages of DNA synthesis to avoid re-replication of DNA [130]. During the late S-phase, Cdk2 is activated by CycA and the CycA/Cdk2 complex drives the transition from S phase to G2 phase. During G2, CycA is thought to be activated by Cdk1 and the CycA/Cdk1 complex drives the G2/M transition [131]. Near the completion of G2, CycA is degraded by

ubiquitin-mediated proteolysis, while cyclin B (CycB) is actively synthesized.

CycB/Cdk1 complex actively participates in and completes mitosis [132].

2.1.2 Cyclin-dependent kinase inhibitors (CKIs)

The activity of the Cdk4/6 and Cdk2 kinases is essential for progression through

G1 and entry into S-phase, and their activity can be modulated at multiple levels,

including cyclin accumulation and cyclin-Cdk complex formation [133]. In addition,

Cdks can be effected by their association with cyclin-dependent kinase inhibitors (CKI)

that can either physically block activation or block substrate access [134].

The CKIs can be grouped into at least two families: an Ink4 family that inhibits

CycD-associated kinase activity and a Cip/Kip family that inhibits Cdk2- and Cdk4-

containing complexes. The Ink4 family, consisting of p16Ink4a, p15Ink4b, p18Ink4c, and

p19Ink4d, inhibits Cdk activity by competing with CycD for binding to the Cdk subunit.

The Cip/Kip family, consisting of p21Cip1, p27Kip1 and p57Kip2, shares an inhibitor

domain that can bind and inhibit Cdk2- and Cdk4-containing complexes.

2.1.3 Cell cycle, cancer, and anti-cancer strategies

Given their role in the G1/S, S, or S/G2 transitions, the regulatory components

involved in the cell cycle progression must be highly orchestrated for normal cell

physiology. Alterations in the cell cycle machineries often lead to unrestrained cell

proliferation, resulting in cancer. In cancer, a number of mutations have been reported that disrupt Cdk functions. First, Cdks themselves may be mutated, although this occurs with low frequency. The mutations lead to over-expression of Cdks (Cdk4 [135], Cdk1, or Cdk2 [136,137]) or to loss of CKI binding in Cdk4 and Cdk6 [138]. Also, cyclins that bind Cdks to active them may be mutated. In particular, CycD, which senses growth signals, is often associated with a number of cancers [139]. Other cyclins including CycA and CycE have been shown to be over-expressed in lung carcinoma [140]. Finally, mutations in CKIs are observed in a high percentage of human tumors and can be inactivated by a variety of mechanisms including deletion, point mutations, and hypermethylation [141]. Cells with intact p16Ink4a, a specific inhibitor of CycD/Cdk4,6, inhibits phosphorylation of Rb, but mutated p16Ink4a removes such inhibition and cells are unrestrained to proliferate. In addition to disrupted Cdk functions, mutations in the Cdk substrate can lead to cancer development. One of the most important substrates is Rb, which binds E2F to inhibit proliferation. Absence or loss of Rb function is often observed in various types of cancer, and is associated with unrestrained cell cycle progression [139,142]. These mutations are often accompanied with disrupted Cdk functions.

Based on the current understanding of cancer development, a number of strategies have been designed to restore control of the cell cycle. Cdk inhibitors have been considered as a reasonable strategy, but the first generation of the Cdk inhibitors

has raised several concerns [126]: 1) off-target effects that limit therapeutic drug

concentrations, 2) specificity of the drugs for various mutations, and 3) toxicities. These

concerns may be addressed by gaining an in-depth understanding of cell cycle

regulation in normal and in cancers.

2.2 Checkpoints in the mammalian cell cycle

During the cell cycle, it is critical that the parental DNA is replicated faithfully

and is distributed to the daughter cells symmetrically. However, the DNA in

mammalian cells is under constant attack by agents that directly target its bases or break

the phosphodiester backbone on which the bases reside. For example, the

phosphodiester bonds in the DNA backbone can be damaged by energy released from

free oxygen radicals, which are generated either by normal metabolic processes or by

exposure to an external source of ionizing radiation [143].

To ensure survival and faithful propagation of their genetic contents, the

eukaryotic cells are equipped with elegant mechanisms to repair DNA damages. Such

repair mechanisms are critical in normal physiology and maintenance of eukaryotic cells,

as damage to cellular DNA causes cancer [144]. Highlighting their importance further,

most cancers arise from mutations in the genes involved in DNA-damage responses.

Interestingly, these repair mechanisms are intricately coordinated with the

processes of cell-cycle arrest or apoptosis. It has been shown that cells are equipped with

a number of different repair mechanisms, which are specific for cell cycle stage or the

nature of the DNA damage. These mechanisms allow for modulation of cell cycle

progression in the face of DNA damage and other stresses that affect DNA replication.

Such modulation may be critical in optimal DNA repair, as halting or slowing DNA

replication may be beneficial during the process of DNA repair.

2.2.1 Signal initiation upon stress

Upon DNA damage, a number of signals instantaneously become activated.

These include AMT (ataxia telangiectasia mutated) and ATR (AMT- and Rad3-related),

which inhibit cell cycle progression by phosphorylating their substrates. In unstressed

cells, ATM kinase is a homodimer and its kinase domain is physically blocked. Its

activity is minimal and it functions to help cells deal with cellular stresses that affect

DNA or chromatin structure [143]. Upon DNA damage, ATM undergoes a

conformational change, resulting in dissociation of the homodimer. The activated ATM

monomers get directed by a helper MRE11 to move to its substrates, some of which

localize to the sites of DNA damage. Unlike ATM, ATR exists in a complex with ATR-

interacting protein (ATRIP) and is constitutively activated. Upon DNA damage, ATRIP

binds to replication protein A (RPA), which is involved in DNA replication [145].

Accumulation of RPA, therefore, would lead to recruitment of ATR, which can

phosphorylate critical substrates in DNA repair. Once ATM and ATR reach their

appropriate destinations, they phosphorylate a number of proteins including their

effector kinases (or transducer kinases), CHK1 and CHK2 [146]. These are extremely

mobile kinases that facilitate modulation of cell-cycle progression at various stages of

the mammalian cell cycle.

2.2.2 G1/S checkpoints

2.2.2.1 The restriction point (R-point)

In unperturbed cell cycle progression, mammalian cells can only be blocked from

proliferation by withdrawing growth signals or applying growth inhibitory signals in

early-to-mid G1 phase. In this stage, cells survey a number of metabolic and

environmental signals and they develop an integrated interpretation of these signals.

Based on their interpretation, cells decide whether to enter S-phase or pause. The point

at which this decision is made is called a restriction point (R-point). When the growth

signals are weak, cells remain at its non-proliferating state. Once the growth signals are

sufficiently strong, cells proceed with proliferation and become independent of growth

signals. That is, removal of the growth signals does not prevent cells from completing

the cell cycle after the R-point. However, withdrawal of the growth signals before the R-

point will prevent cells from proliferation.

Underlying the R-point is the Rb/E2F signaling pathway, which is driven by

several cyclin/Cdk complexes. In this pathway, the binary decision at the R-point results

from the balance between growth stimulation interpreted by cyclin/Cdk complexes and opposing inhibition by Rb and CKIs. Rb represents a family of tumor-suppressor protein, consisting of retinoblastoma (Rb), p130, and p10, and is a substrate of the CycD/Cdk4,6 and CycE/Cdk2 complexes. Rb interacts with the E2F to incapacitate their transcriptional functions. E2F is a family of transcriptional factors with many members identified to date. E2F1, E2F2, and E2F3a are transcriptional activators, responsible for a large number of genes essential for DNA replication and downstream cell cycle progression. E2F3b, E2F4, and E2F5 complexed with Rb and p130 are found in quiescent cells as transcriptional repressors of S phase genes as wells as genes encoding E2F1, E2F2, and E2F3a proteins [147]. Once the growth signals overcome inhibition, cells proceed with proliferation.
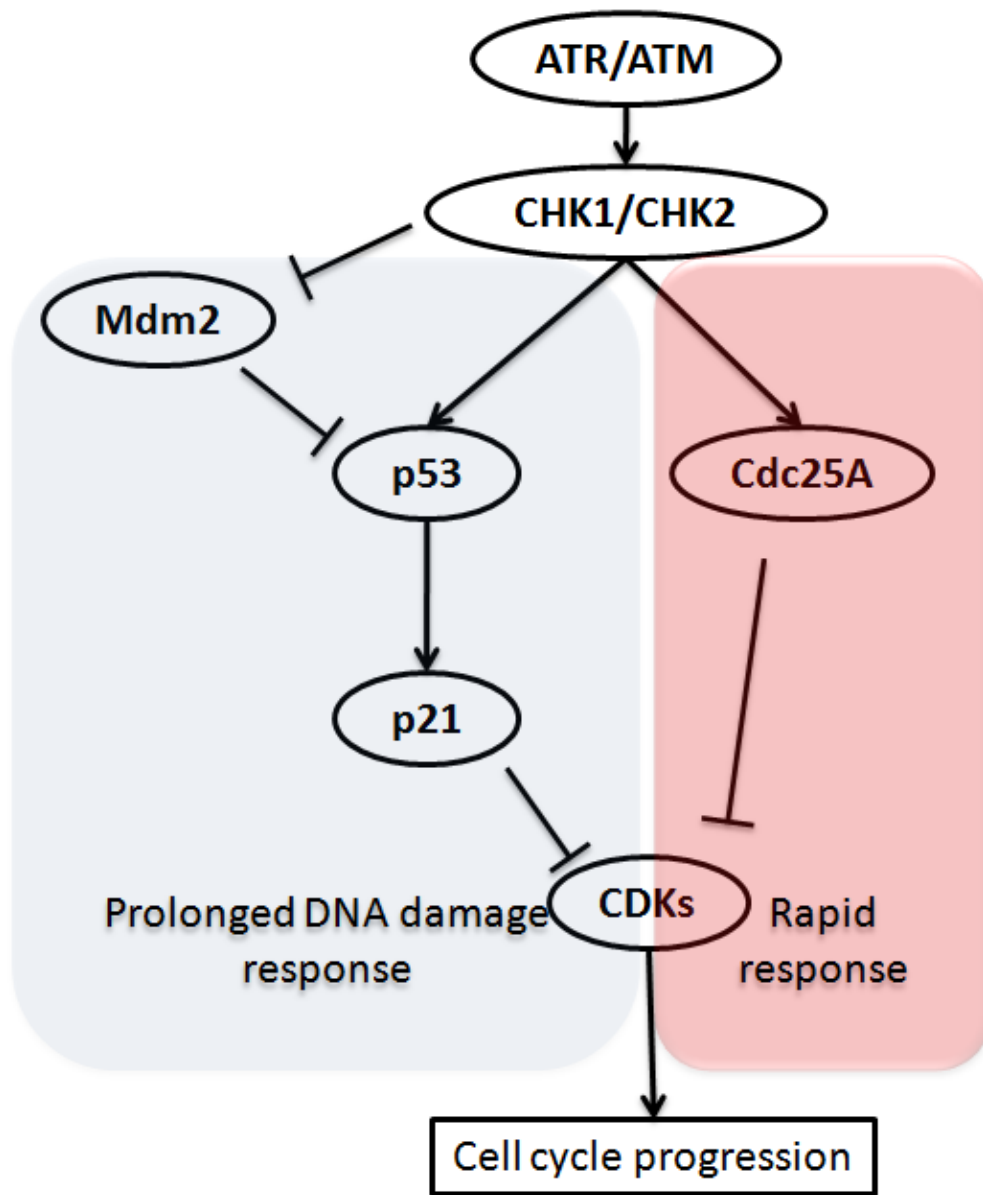
Figure 2.2: DNA damage response. Upon DNA damage, ATR/ATM with their transducers CHK1/CHK2 initiates both transient and prolonged DNA damage response, leading to inhibition of Cdks activity, which in turn results in the arrest of cell cycle progression [143].

2.2.2.2 G1/S checkpoints by ATR/AMT and CHK1/2

Under stressful condition, cells may modulate cell cycle progression to avoid replication and propagation of the damaged DNA. The dominant response to DNA damage in mammalian cells traversing through G1 is achieved by ATM/ATR along with the transducer kinases CHK2/CHK1, acting on two critical effectors: Cdc25A and p53. These effectors link the apical checkpoint kinases with the core Rb/E2F cell cycle machinery via modulation of the CycE/Cdk2 activity, as shown in Figure 2.2.

Cdc25A is a phosphatase that activates Cdk2. Cdk2 forms a complex with CycE, which is responsible for DNA replication. Therefore, the Cdc25A phosphatase activity is important for DNA replication. Upon DNA damage, CHK1 and CHK2 phosphorylates Cdc25A on multiple serine sites and target Cdc25A for degradation by enhanced ubiquitination and proteasome-mediated degradation. This reduces Cdk2 activation, thereby preventing initiation of DNA synthesis. This response by Cdc25A is rapid but transient, capable of delaying cell cycle progression only for several hours [148,149].

On the other hand, another checkpoint mediated by p53 facilitates a prolonged G1 cell cycle arrest in response to DNA damage [143,150]. p53 is a transcription factor, playing a key role in cellular decisions to either arrest the cell cycle, allowing the repair of damaged DNA, or to commit to cell death [151]. p53 accumulation is negatively regulated by Mdm2, which targets it for ubiquitin-mediated proteasome degradation. One of the key effectors of p53-mediated G1 arrest is the p21 inhibitor of Cdks. When

p21 is elevated, it effectively blocks the CycE/Cdk2 activity, arresting the cell cycle progression.

Upon DNA damage, p53 is phosphorylated by ATR/ATM and CHK1/CHK2 at various residues and is activated. In addition, its negative regulator Mdm2 is also targeted by ATR/ATM. These together contribute to stabilization and increased activity of the transcription factor p53. This results in accumulation of p21, which in turn can lead to cell cycle arrest. While this mechanism may require up to several hours, it complements and eventually replaces the transient cell cycle arrest by Cdc25A phosphorylation [150].

2.2.3 S-phase checkpoint

Once cells pass the R-point and begin their DNA replication at the S-phase, DNA damage causes cells to slow down DNA replication by inhibiting the firing from those origins of DNA replication that have not been initiated, but does not completely arrest the cell cycle progression. This is achieved by at least two parallel processes mediated by the ATR/ATM signaling machinery. Upon DNA damage, Cdk2 activity becomes inhibited by Cdc25A phosphorylation by CHK1 and CHK2, as mentioned above. This blocks the loading of Cdc45 onto chromatin, a protein required for the recruitment of DNA polymerase α into assembled pre-replication complexes. Therefore, inhibition of Cdk2 prevents firing of the new origins [148,149]. In addition, ATR/ATM can

phosphorylate other proteins that are involved in DNA repair including NBS1 and SMC1[152,153].

2.2.4 G2/M checkpoint

After completion of DNA replication, cells ensure the integrity of the DNA to prevent mitosis with damaged or unrepaired DNA at the G2/M checkpoint. Similar to the G1/S checkpoint, the G2/M checkpoint consists of rapid response mediated by Cdc25A phosphorylation and a prolonged response mediated by both p53-dependent and p53-independent mechanism.

One of the major targets in the G2/M checkpoint is Cdk1 complexed with CycB. In unperturbed cells, Cdc25A activates Cdk1, and the CycB/Cdk1 complex drives the G2/M transition. Under the condition of DNA damage or incorrect DNA replication, Cdc25A becomes phosphorylated and targeted for degradation, resulting in inactivation of Cdk1. This constitutes transient DNA repair mechanism at the G2/M transition. In addition, phosphorylation of p53 silences the Cdk1 activity via p21, resulting in cell cycle arrest.

2.3 The mammalian cell cycle and apoptosis

Under severe stress to the DNA, the integrity of the genome may be severely damaged and the elegant DNA repair mechanisms may not warrant restoration of the

genome integrity. This may pose oncogenic threat to the cell. In this case, the optimal

decision for the cell may be to undergo a programmed cell-death, or apoptosis.

Apoptotic cells undergo cell blebbing, exposure of phosphatidylserine at the cell surface,

reduction of cell size, shrinkage of the cell core, DNA condensation and the formation of

apoptotic bodies [154]. Apoptosis is achieved by the tumor suppressor protein p53,

which has also shown its ability to elicit cell cycle arrest. How p53 mediates these

drastically differential responses has been the focus of intensive research [155].

2.3.1 p53-dependent cell cycle arrest vs. apoptosis

The transcription factor p53 mediates both responses by turning on or off

appropriate genes to achieve the desirable outcome [156,157]. One strategy used by p53

to elicit different responses is by differential binding affinities. Transactivation of the p53

target genes requires sequence-specific binding at p53-response elements (P53REs), but

it has been shown that not all target genes bind to p53 with equal affinity. The binding

affinity of the genes associated with cell cycle arrest tends to be greater than that of the

apoptotic genes [158]. Therefore, the cell fate mediated by p53 depends strongly on the

its level: low level of p53 tends to favor growth arrest, while high p53 level overrides

this pathway and trigger apoptosis [159].

Apoptosis is facilitated by either an extrinsic pathway via TNF-family ligands or

an intrinsic, mitochondrial pathway induced by UV radiation, chemotherapeutics, free

radicals or DNA damage [154,160,161]. The death-receptor pathway is activated upon

binding of TNF-family ligands to their corresponding receptors, including CD95/Apo-

1/Fas or TRAIL. These receptors contain different death effector domains (DEDs) and

aggregate to form membrane-bound signaling complexes. The signaling from these

complexes ultimately leads to the formation of the death-induced signaling complex

(DISC), activation of caspase-8 (cysteinyl aspartate-specific protease), and triggering the

proteolytic caspase cascade. Active caspase-8 can cleave BID (BH3-interacting-domain

death agonist), which promotes the mitochondrial permeability transition [162]. The

intrinsic, mitochondrial pathway can be triggered when pro-apoptotic proteins such as

Bcl2-associated X protein (Bax) are activated or when anti-apoptotic proteins Bcl2 family

members are inactivated [163]. As a result, inner membrane potential becomes

dissipated and the outer mitochondrial membrane becomes permeabilized. This induces

the release of various apoptotic proteins including cytochrome c, Smac, or apoptosis

inducing factor (AIF) [162,164]. Many of the proteins involved in the mitochondrial

pathway are encoded by the target genes of p53, including the pro-apoptotic Bax protein

that generates pores sufficient for release of the apoptotic factors [165,166].

2.3.2 The cell cycle machinery linked with apoptosis

The apoptotic pathways are intricately linked with the cell cycle machinery, as

many of the cell cycle components are directly involved in inducing apoptosis. For

example, E2F1, whose ectopic expression can drive the quiescent cells through the G1/S

transition, can trigger apoptosis when continually expressed through the S-phase. In

unperturbed cells, E2F1 can induce the expression of p19[Arf], which negatively regulates

Mdm2, which in turn negatively regulates p53. Upon DNA damage, E2F1 is highly

induced by its stabilization. Similar to stabilization of p53 by phosphorylation by

ATR/ATM kinases, E2F1 is phosphorylated by these kinases and is prevented from

degradation. Therefore, the E2F accumulation upon DNA damage can further stabilize

and accumulate p53, which can facilitate p53-dependent apoptosis [129,147]. In addition,

the E2F antagonist Rb may act as a general suppressor of apoptosis, as it can suppress

apoptosis induced by IFN, TGFβ, and p53 [167]. Consistent with this notion, TNF-

induced apoptosis requires Rb cleavage by caspases [168].

Chapter 3: Modeling and experimental analysis of a bistable Myc-Rb-E2F switch

3.1 Introduction

In Chapter 3, we investigate the key signaling pathways that underlie cell cycle entry dynamics. More specifically, we have developed a mathematical model for the Myc-Rb-E2F signaling pathway, known to be involved in cell cycle entry. With this model, we predicted all-or-none switching dynamics in cell cycle entry, which were experimentally validated. Such all-or-none switching dynamics is characteristic of the restriction point (R-point). This work was achieved in collaboration between Dr. Guang Yao, who focused on experimental validations, and Tae J. Lee, who worked on mathematical modeling. The work presented in Chapter 2 appears in Nature Cell Biology 2008 Apr; 10(4): 476-482 [53].

3.2 Background

3.2.1 The Restriction Point

The restriction point (R-point) refers to a position in cell cycle where cells commit themselves to proliferation given enough growth factors [169,170,171]. Removal of growth signals before the R-point prevents cells from engaging in the cell cycle. After the R-point, however, the cell cycle is completed regardless of the growth signal condition [172]. Despite its importance, the molecular nature of the R-point remains

elusive [173] and even controversial [174]. Several modeling studies have been conducted to gain insight into the R-point from dynamics of underlying signaling networks [175,176,177,178,179,180,181,182]. An emerging consensus from these studies is that the entry of mammalian cell cycle may be governed by bistable activation of some regulatory proteins. However, these studies focused on different processes or components, leading to somewhat confusing characterization of the R-point and/or G1/S transition.

Thron proposed a highly simplified conceptual framework to account for origin of bistability during cell cycle control [52]. Aguda and colleagues built their analysis primarily around two positive feedback loops involving Cyclin E (CycE) [177,178]. Their model generated a bistable switching behavior leading to CycE activation, which they used to account for the R-point [177]. Similarly, Hatzimanikatis and colleagues predicted bistability in the CycE/Rb-E2F pathway, which they argued to be critical for G1/S transition [176]. Qu and colleagues also focused on CycE regulation as the basis of G1/S transition [183]. Yet CycE is unlikely to play a central role, as experiments indicated that it is not required for passage of the R-point [184,185]. Novak and Tyson [175] built a mammalian cell cycle model following the framework of their well-established yeast cell cycle model [186,187,188]. They characterized the R-point as activation of E2F and degradation of Cdk inhibitors. Based on their analysis, they argued that the R-point, which is independent of cell size, is dissociated from the G1/S transition driven by

accumulation of CycA, which is dependent on cell growth [175]. A limitation of these modeling studies is that their predictions have not been put to direct experimental test.

3.2.2 The Myc-Rb-E2F signaling pathway

Extensive evidences suggest that the Myc-Rb-E2F pathway plays a central role in regulating the cell cycle entry [189,190,191,192,193,194,195,196,197], but its connection with the R-point remained unclear. In our mathematical model, we considered a positive auto-regulation of E2F, which has been neglected [175,176,179] in the past or considered secondary [177]. We hypothesized that this positive feedback regulation may be sufficient to generate bistable switching behaviors, which may account for the R-point dynamics.

In the Myc-Rb-E2F pathway, the interactions between Rb and E2F define the threshold for the system activation. In quiescent cells, Rb protein inhibits function of E2F transcription activators (E2F1, E2F2, and E2F3a) by forming an E2F-Rb complex. Addition of serum (containing growth signals) leads to production of transcription factor Myc via a complex Ras signaling pathway. In turn, Myc up-regulates synthesis of cyclin D (CycD), and inactivates cyclin-dependent kinase inhibitors (CKIs). Subsequently, an active complex is formed between CycD and cyclin-dependent kinases (cdks), which phosphorylates Rb and releases its inhibition on E2F. Synergistic with Myc, active E2F can in turn promote its own synthesis, forming a positive feedback. In

addition, E2F drives the expression of cyclin E (CycE), which forms a complex with

cdks. The CycE-cdk complex further phosphorylates the Rb complex to relieve its

repression of E2F, establishing a second positive feedback.
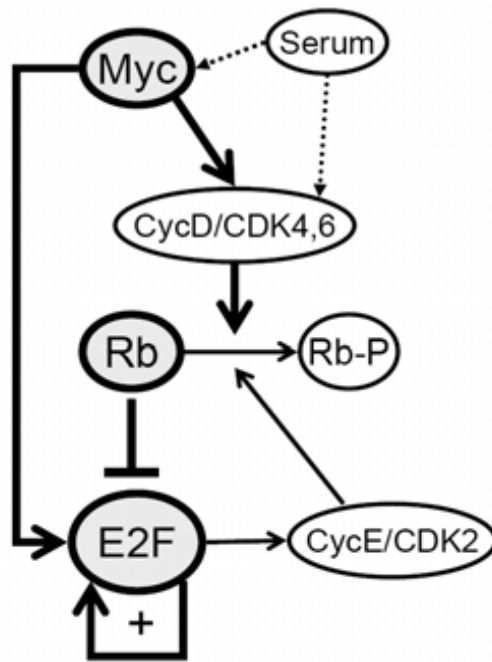
Figure 3.1: The Myc-Rb-E2F pathway (adapted from [147]). In quiescent cells E2F is
bound by Rb and its transcriptional activities are repressed. Growth stimulation
removes Rb repression by upregulating cyclin D (CycD), which, in complex with Cdk4,6,
phosphorylates Rb to release E2F. In addition, growth stimulation induces a
transcription factor Myc that upregulates CycD. The free form of E2F synergizes with
Myc to induce its own transcription, forming feed-forward and positive feedback loops.
Subsequently, E2F activates the transcription of Cyclin E (CycE), which forms a complex
with Cdk2 to further remove Rb repression by phosphorylation, constituting another
positive feedback loop.

3.3 *Quantitative analysis of the Myc-Rb-E2F signaling pathway*

3.3.1 Construction of the mathematical model

We have developed a simplified mathematical model to account for the key interactions outlined in Figure 3.1. The model consists of a set of ordinary differential equations (ODEs) as listed in Table 3.1. These ODEs were based on kinetics and parameters as shown in Table 3.2 and Table 3.3. Great care has been taken to estimate basic modeling parameters. Some parameters (e.g., Michaelis-Menten constants and decay rate constants) were obtained from the literature. Some other parameters (e.g, maximum phosphorylation and dephosphorylation rate constants) were estimated based on typical values for related or similar kinetic processes. The remaining free parameters (synthesis rates of various components) were constrained with experimental measurements. For preliminary parameter adjustments and time-course simulations, We used *Dynetica* (Figure 3.2), a graphics-based, integrated simulation platform [76]. We then implemented the model in '*XPP-AUTO*' to perform bifurcation analysis [198,199].

In developing the mathematical model, we focused on the role of Myc-Rb-E2F circuit in governing cell cycle entry. As such, the model accounted for cellular dynamics prior to the G1/S transition and neglected downstream gene regulations. Furthermore, we used E2F to generalize all E2F activators (E2F1, E2F2, and E2F3a), and used Rb to represent all pocket proteins (Rb, p107, and p130). In addition, the inhibitory activities of CKIs were lumped into the phosphorylation rate constants of Cyclin/cdk complexes.

These simplifications reduced the intricate regulatory network of cell cycle entry to an

experimentally tractable model.
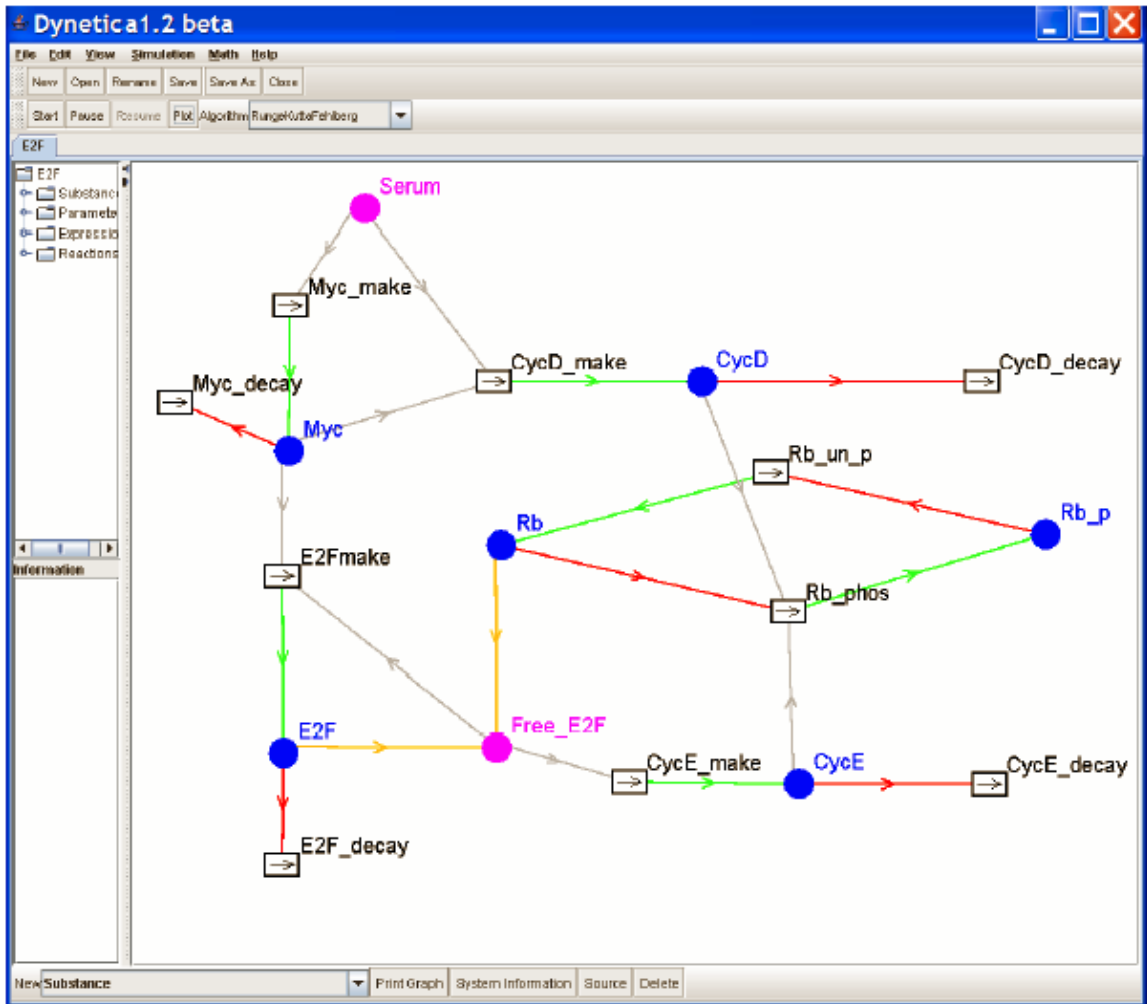
Figure 3.2: Implementation of the Myc-Rb-E2F pathway in Dynetica. Filled circles
represent interacting species. Boxes represent interactions. A red arrow pointing away
from a substance to a reaction represents degradation or consumption by the reaction. A
green arrow pointing toward a substance represents production. A gray arrow pointing
from a species to a reaction indicates that the species is modulating the rate of the
corresponding reaction.

Table 3.1: The mathematical model for the Myc-Rb-E2F pathway

$$\frac{d[M]}{dt} = \frac{k_M[S]}{[S] + K_S} - d_M[M]$$

$$\frac{d[E]}{dt} = k_E \left(\frac{[M]}{K_M + [M]} \right) \left(\frac{[E]}{K_E + [E]} \right) + \frac{k_b[M]}{K_M + [M]} + \frac{k_{P3}[CD][RE]}{K_{CD} + [RE]} + \frac{k_{P4}[CE][RE]}{K_{CE} + [RE]}$$

$$- d_E[E] - k_{RE}[R][E]$$

$$\frac{d[CD]}{dt} = \frac{k_{CD}[M]}{[M] + K_{CD}} + \frac{k_{CDS}[S]}{[S] + K_S} - d_{CD}[CD]$$

$$\frac{d[CE]}{dt} = \frac{k_{CE}[E]}{K_E + [E]} - d_{CE}[CE]$$

$$\frac{d[R]}{dt} = k_R + \frac{k_{DP}[RP]}{K_{RP} + [RP]} - k_{RE}[R][E] - \frac{k_{P1}[CD][R]}{K_{CD} + [R]} - \frac{k_{P2}[CE][R]}{K_{CE} + [R]} - d_R[R]$$

$$\frac{d[RP]}{dt} = \frac{k_{P1}[CD][R]}{K_{CD} + [R]} + \frac{k_{P2}[CE][R]}{K_{CE} + [R]} + \frac{k_{P3}[CD][RE]}{K_{CD} + [RE]} + \frac{k_{P4}[CE][RE]}{K_{CE} + [RE]} - \frac{k_{DP}[RP]}{K_{RP} + [RP]} - d_{RP}[RP]$$

$$\frac{d[RE]}{dt} = k_{RE}[R][E] - \frac{k_{P3}[CD][RE]}{K_{CD} + [RE]} - \frac{k_{P4}[CE][RE]}{K_{CE} + [RE]} - d_{RE}[RE]$$

Table 3.2: Reaction kinetics for the model

Reaction	Kinetics	Description and notes
$* \xrightarrow{S} M$	$\dfrac{k_M[S]}{[S]+K_S}$	Myc synthesis driven by growth signals (S)
$* \xrightarrow{S} CD$	$\dfrac{k_{CDS}[S]}{[S]+K_S}$	CycD synthesis driven by growth signals
$* \xrightarrow{M,E} E$	$k_E\left(\dfrac{[M]}{K_M+[M]}\right)\left(\dfrac{[E]}{K_E+[E]}\right)$ $+\dfrac{k_b[M]}{K_M+[M]}$	E2F synthesis by a synergy between Myc and E2F autocatalysis. Since neither Myc nor E2F forms a homodimer, we assumed no cooperativity in gene activation mediated by these factors, and used the Hill coefficient of 1.0. Using Hill coefficient greater than 1.0 will not change the qualitative behavior of system dynamics
$* \xrightarrow{E} CE$	$\dfrac{k_{CE}[E]}{K_E+[E]}$	CycE synthesis driven by E2F
$* \xrightarrow{M} CD$	$\dfrac{k_{CD}[M]}{K_M+[M]}$	CycD synthesis driven by Myc
$* \longrightarrow R$	k_R	Constitutive Rb synthesis
$RE \xrightarrow{CD,CE} E+RP$	$\dfrac{k_{P3}[CD][RE]}{K_{CD}+[RE]}+\dfrac{k_{P4}[CE][RE]}{K_{CE}+[RE]}$	E2F dissociation from Rb-E2F complex by CycD- and CycE-mediated phosphorylation
$E+R \longrightarrow RE$	$k_{RE}[R][E]$	E2F titration by Rb via E2F-Rb complex formation
$R \xrightarrow{CD,CE} RP$	$\dfrac{k_{P1}[CD][R]}{K_{CD}+[R]}+\dfrac{k_{P2}[CE][R]}{K_{CE}+[R]}$	Rb phosphorylation by CycD and CycE
$RP \longrightarrow R$	$\dfrac{k_{DP}[RP]}{K_{RP}+[RP]}$	Rb dephosphorylation
$M \longrightarrow *$	$d_M[M]$	Myc decay
$E \longrightarrow *$	$d_E[E]$	E2F decay
$CE \longrightarrow *$	$d_{CE}[CE]$	CycE decay

$CD \longrightarrow *$	$d_{CD}[CD]$	CycD decay
$R \longrightarrow *$	$d_{R}[R]$	Rb decay
$RP \longrightarrow *$	$d_{RP}[RP]$	Phosphorylated Rb decay
$RE \longrightarrow *$	$d_{RE}[RE]$	Rb-E2F complex decay

Table 3.3: Parameters for the model

Rate constants		Parameter values, sources, and notes
k_{EF}	0.4 μM/hr	These values were adjusted together so that:
k_{CD}	0.03 μM/hr	(1) The maximum E2F level is about 5-6 fold higher than the maximum CycD level (based on our experimental
k_{CDS}	0.45 μM/hr	observations).
k_{RB}	0.18 μM/hr	(2) E2F activation threshold is about 1% serum
k_b	0.003 μM/hr	(3) The simulated E2F level will be around the corresponding Michaelis-Menten parameter (K_E)
k_{CE}	0.35 μM/hr	Assumed to be the same as k_{EF}
d_M	0.7/hr	Myc half-life = 60 min [200,201,202]
d_E	0.25/hr	E2F half-life = 2~3 hr [203]
d_{CD}	1.5/hr	CycD half-life = 25~30 min [127,204]
d_{CE}	1.5/hr	CycE half-life = 30 min [205,206]
d_R	0.06/hr	Rb half-life = 12 hours [207]
d_{RP}	0.06/hr	Assumed to be the same as d_R
d_{RE}	0.03/hr	Rb-E2F half-life = 6 hours: The Rb-E2F complex assumed to be more stable than Rb alone [208]
$k_{P1}{}^*, k_{P2}{}^*$	18/hr	Typical value phosphorylation rate constant [209] is 3600/hr
$k_{P3}{}^*, k_{P4}{}^*$	18/hr	Typical value phosphorylation rate constant [209] is 3600/hr
$k_{DP}{}^*$	3.6 μM/hr	Typical value for dephosphorylation rate assuming a constant phosphatase concentration [209] is 720 μM/hr
K_M	0.15 μM	Estimated based on measured Myc/Max –DNA dissociation constant [210]
K_E	0.15 μM	Assumed to be the same as K_M
K_{CD}	0.92 μM	Experimentally measured [211,212]
K_{CE}	0.92 μM	Assumed to be the same as CycD
K_{RP}	0.01 μM	Typical value for Michaelis-Menten parameter for dephosphorylation [209]

Table 3.3: Continued

* Typical values of phosphorylation and dephosphorylation rate constants results in a stiff model, which drastically slows down the calculation. We have found that the overall dynamics is insensitive to the overall rates of the phosphorylation and dephosphorylation reactions, as long as they are balanced. Thus we have reduced the corresponding rate constants by 200 fold to speed up calculation. Proportionally increasing these parameters has no significant impact on the overall system dynamics.

[RE] = 0.82 μM; [Rb]=[Rbp] = [E2F] = [Myc] = [CycD] = [CycE] = 0 μM

3.3.2 Utilities and limitations of the Myc/Rb/E2F model

The complete Myc/Rb/E2F pathway is far more complex than what is shown in Figure 3.1, and understanding of the network dynamics may be challenging by intuition alone. To gain a system-level understanding of the network dynamics, we built a mathematical model that is simple enough to make experimentally tractable predictions without losing key features of the network. The ultimate goal of the model is not only to gain an in-depth, system-level understanding of the network and make predictions, but also to guide experiments that will validate model predictions.

To achieve simplicity of the model without compromising its dynamical features, we have made a number of simplifications. Some of the components of the Myc-Rb-E2F pathway not shown in the model were lumped into the model parameters by their functional similarity. For example, the activity of E2F transcriptional repressors (E2F3a, E2F4, and E2F5), which shares similar functions as the Rb family, was assumed in the Rb synthesis rate. The Rb family consists of multiple members, which were also lumped together into the Rb synthesis rate. In addition, other components that functionally oppose the model components were lumped together in the parameters that describe the model components. For instance, the activity of cyclin-dependent kinase inhibitors (CKIs) was reflected in the parameters that represent the activity of CycD/Cdk4,6 or CycE/Cdk2. Finally, many of other components in the Myc-Rb-E2F pathway were not

72

considered in our model because they are not directly involved in cell cycle entry, but

rather in the maintenance of the cell cycle progression (i.e. CycA or CycB).

The dynamics of a mathematical model is defined not only by its architecture,

but also its parameters. While some of the parameters were obtained from the literature,

others were not readily available and were estimated from our own experimental data.

Often times, these parameters would cover a wide range of values. Even the parameters

obtained from the literature may vary significantly (to several orders of magnitude)

depending on the experimental conditions and context. In addition, the parameter

values may change over time due to various mechanisms. For example, the degradation

rate of Myc is tightly controlled by its upstream signals via post-translational

modification. Before the protein modification, Myc is highly unstable. Upon protein

modifications by phosphorylation via its upstream signals, Myc becomes stabilized.

Such stabilization in Myc will depend on the timing of the upstream signals and will

change the parameter describing degradation by several folds. These sources of

uncertainty in the model parameters call for a strategic way to justify the parameter

values used in the model.

To define a permissible, biological feasible range of parameters, several criteria

can be imposed on the model parameters. One of the criteria is that the set of parameters

should give rise to the expected dynamics. In the development of the Rb-E2F model, we

expected E2F dynamics to be bistable. This is based on the previous observation that the

73

R-point control is history-dependent and irreversible, consistent with the concept of bistability. To achieve bistability, we began with the initial set of parameters obtained from the literature and our own experimental results, which were sufficient to generate bistability in our model. In case the initial set of parameters and the network architecture did not generate the expected dynamics, more systematic exploration of the parameter sets would be required. If this failed to generate the expected dynamics, modification of the network architecture would be required and another exploration of appropriate parameters would be conducted. This constitutes the first layer of model calibration.

Once a base model is established, it can be tested for sensitivity of the system dynamics to various parameters. This is usually done at one-dimension by evaluating the change in the objective function to the change in the parameter of interest. For example, the objective function in the Rb-E2F model was the temporal dynamics of E2F and the parameter of interest was the Cdk2 kinase activity. Using the model, we could predict how sensitive the temporal dynamics of E2F activation is to the phosphorylation efficiency of Cdk2. Such prediction can guide future experiments and further calibration of the model. Multi-dimensional approaches may also be taken, although experimental validation for such sensitivity analysis would be difficult with a large number of dimensions.

Such sensitivity analysis can serve as a foundation for another criterion for the model: an agreement between the predicted dynamics and experimental observation

under perturbations. Perturbing the network at different nodes introduces a change in

the parameters, and the corresponding change in network dynamics can be directly

evaluated by experiments. A good agreement between the modeling and experimental

results would demonstrate the goodness of the model, and such agreements under

various perturbations would enhance the predictive power of the model. Lack of

agreement would suggest a need for modification of the parameters set or network

architecture. To this end, we have perturbed the model at CycE/Cdk2 only, but future

studies targeting other nodes would further calibrate the model.

3.3.3 Model predictions: bistable switching behaviors

Figure 3.3 shows exemplary E2F time courses in response to different serum

stimulation schemes that demonstrate bistability of the pathway. For a low serum level

(Figure 3.3A), E2F never gets activated, which corresponds to the G0 phase or arrest in

G1 phase (Figure 3.3B). However, if the serum level starts high (Figure 3.3D, Time<5),

E2F will be activated, and it will remain at the high state (corresponding to cell cycle)

even if the serum is subsequently reduced to a low level (Figure 3.3E). That is, the

pathway has "memory": its final outcome depends on not only network parameters, but

also the prior history of network dynamics. This bistable response of E2F signaling is

distinct from expression of other genes in the Myc-Rb-E2F pathway such as CycD, which

responds to serum stimulation in a monostable fashion – that is, expression levels of these genes are independent of history of the system dynamics (Figure 3.3C,F).
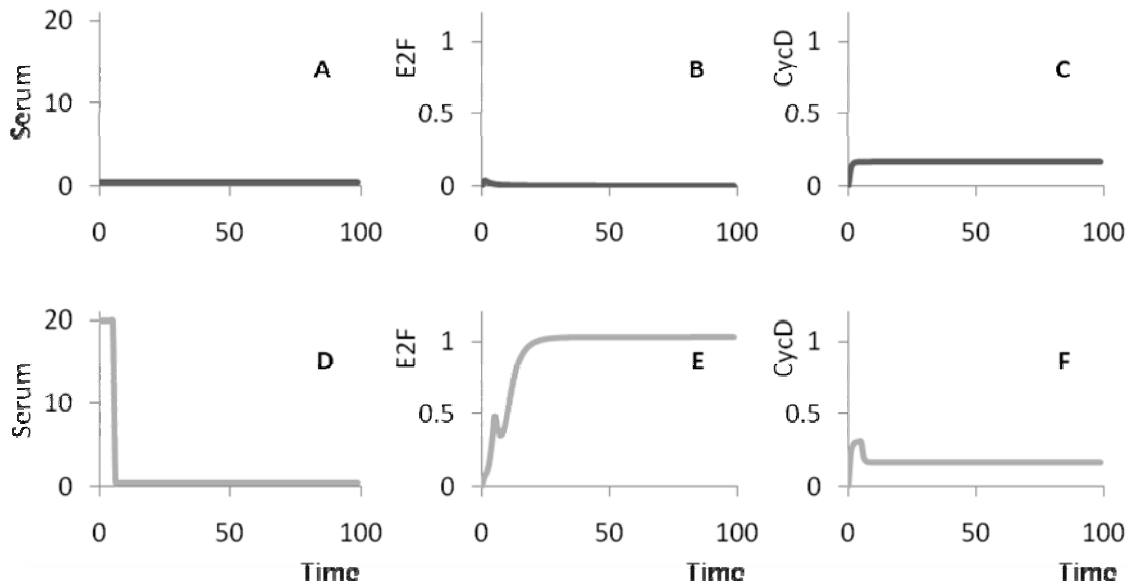
Figure 3.3: Response of E2F to serum stimulation. If serum level is kept low (A), E2F will not be activated (B). However, for a sufficiently high serum level (D, Time<100), E2F will be activated to a high level (E), and will not drop to zero even if serum level is reduced to a low level (D, Time > 5). Note that the final states of the serum level are the same for (A) and (D), but the corresponding steady-state levels of E2F are drastically different (B & E). In contrast to E2F, CycD increases or decreases in proportion to the serum level (C & F). Essentially, the circuit has a "memory": the outcome of the circuit depends on the history of circuit dynamics. Time, serum levels, and E2F and CycD levels are in arbitrary units. Similar dynamics was observed if Myc protein level, instead of serum, was varied.

The history-dependent behaviors at the steady-state are represented in the bifurcation analyses in Figure 3.4. For low serum levels (<0.2%), the system is monostable: regardless of its history, it always stays at the low state of E2F, which

corresponds to G0 phase or arrest in G1 phase. For sufficiently high serum levels (>1.0%), the system is also monostable: it always stays at the high state, corresponding to cell cycle progression. For intermediate serum levels, however, the system can be either at the low or high state, depending on its history, as demonstrated in Figure 3.4.
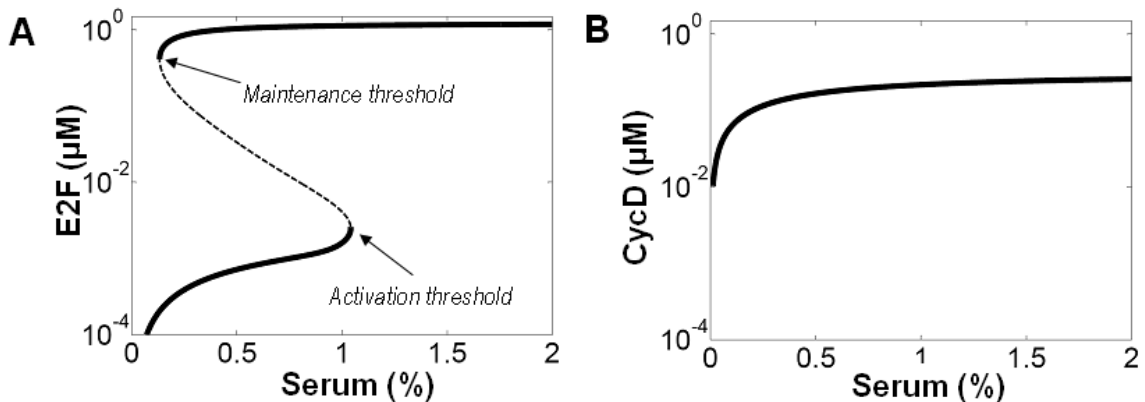
Figure 3.4: Hysteresis in E2F activation. (A) The modeling results demonstrate hysteresis in E2F activation in response to serum concentrations. At low serum concentration, E2F remains at its low state (OFF). At high concentration, E2F elevates to its high state (ON). At intermediate serum concentration, E2F can be either at the high state or low state depending on its history. E2F follows two different trajectories when switching from OFF-ON and ON-OFF. We define the minimum serum concentration at which E2F activates as 'activation threshold', and the minimum serum concentration at which E2F deactivates as 'maintenance threshold'. The region between the thresholds is a 'bistable region' where E2F can be either at ON of OFF state. (B) CycD follows an identical trajectory for both switching OFF-ON and ON-OFF.

3.3.4 Experimental validation of the model predictions (by Dr. Guang Yao)

To validate bistable switching behaviors in the Rb-E2F pathway, we monitored transcriptional activities of E2F using a destabilized green fluorescent protein (d2GFP). This transcriptional reporter, under the control of a cloned E2F promoter, was integrated

into the chromosome of a rat embryonic fibroblast cell line (REF52) by retroviral transfection. As a negative control, we constructed a CycD-d2GFP reporter system to show monostable behaviors in comparison. In either system, it was critical to establish single cell clones of relatively homogenous backgrounds (i.e. similar copy number), since we expected that stochastic cellular noise may interfere with capturing ultrasensitive dynamics of E2F [213]. Both d2GFP reporter systems exhibited strong serum responses comparable to well-characterized endogenous activities, with the half-activation time of CycD induction appeared over 7 hours earlier than that of E2F.

Consistent with our model predictions, we found that E2F-d2GFP, measured by flow cytometry, exhibited bistable behaviors with two distinct peaks at low and high serum concentrations (Figure 3.5A). When serum concentrations were sufficiently low ($\leq$ 0.1%) or adequately high ($\geq$ 3%), the E2F-d2GFP levels were unimodally OFF or ON, respectively. However, although the exact distributions were contingent on individual cell clones and experimental conditions, E2F-d2GFP distributions were clearly bimodal at intermediate concentrations (0.5% ~ 2%). Consistent with our model predictions in Figure 3.3, with serum concentration increasing, we observed gradual increase of E2F-d2GFP level at both its OFF state (left peak) and ON state (right peak); the OFF and ON peaks, however, were always well-separated. This bimodal separation of E2F-d2GFP levels indicated the "all-or-none" characteristic in E2F activation. That is, the Rb-E2F switch fully turned ON in some cells or remained OFF in others due to cellular noise

[87,88,214], but it hardly stayed at intermediate levels because of its ultrasensitive response. In contrast, the expression of CycD-d2GFP, which is not directly regulated by positive feedback, was always unimodal. Similar to E2F-d2GFP, it increased gradually as a function of increasing serum concentrations, but no clear separation between ON and OFF states was observed (Figure 3.5B).
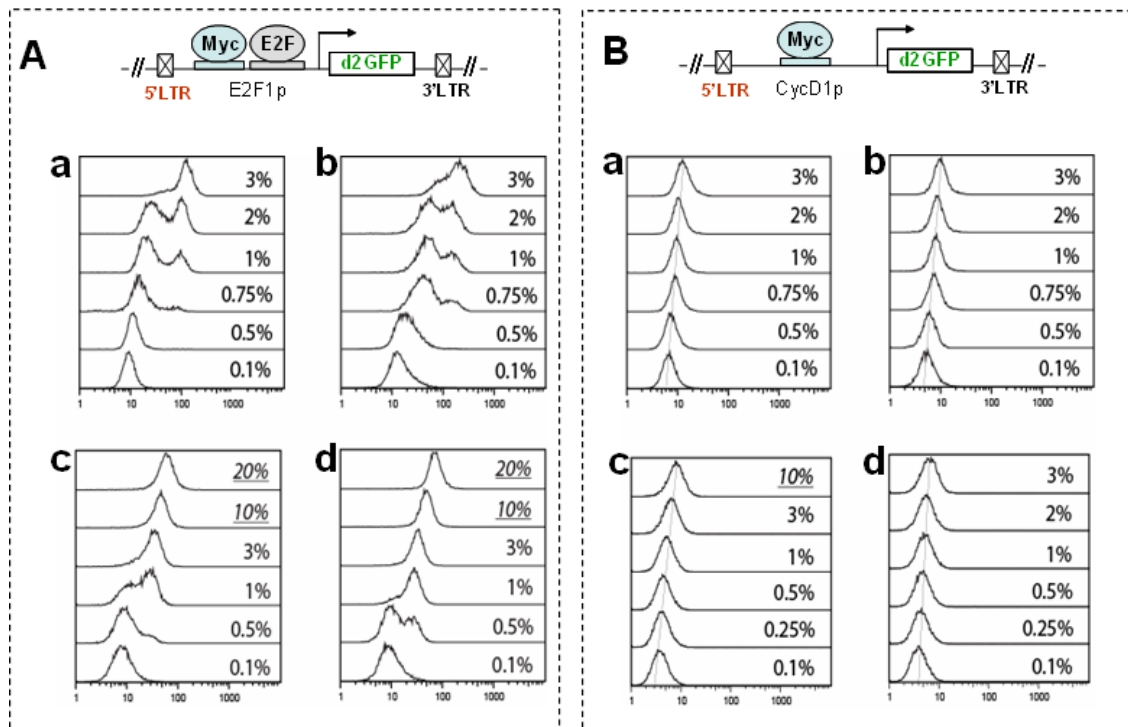
Figure 3.5: Bimodal E2F expression in cultured mammalian cells. E2F-d2GFP construct. (Upper) E2F dose responses. (Lower) (A) Cells were serum starved at 0.02% BGS (a & b) or 0.05% FBS (c & d) for 1 day and subsequently stimulated with BGS (a & b) or FBS (c & d) at indicated concentrations for 24 (a & b) or 20 (c & d) hours. (B) CycD-d2GFP construct. (Upper) CycD dose responses. (Lower) Cells were serum starved at 0.02% BGS for 1 day (a & b) or 1.5 days (c & d) and subsequently stimulated with BGS at indicated concentrations for 24 (a & b), 21 (c), or 25 (d) hours. For both A and B, Each histogram represents the distribution from ~ 10,000 cells. Data from four independent single cell clones are shown.
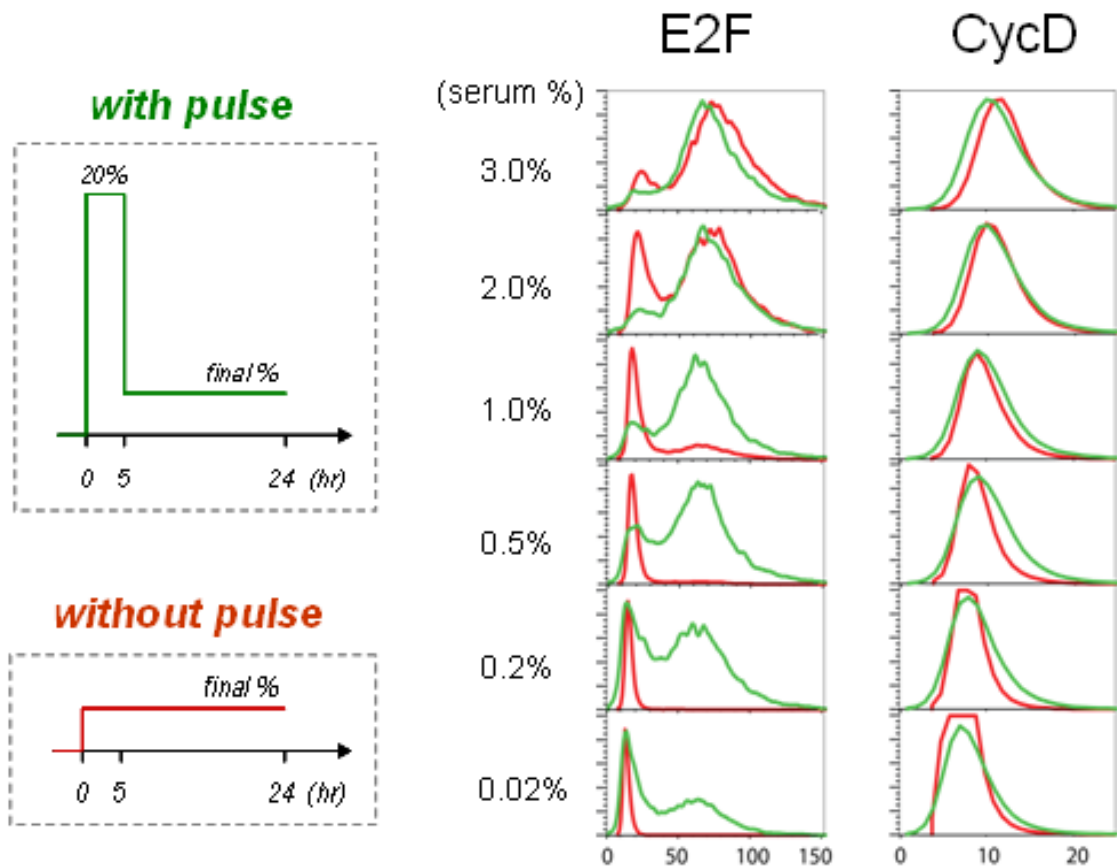
Figure 3.6: Bistable E2F activation. (Left) Serum stimulation protocols. Pulsatile input was generated by stimulating serum-starved quiescent cells at 20% serum for 5 hours, then reducing the serum level to varying concentrations. (Right) E2F-d2GFP and CycD-d2GFP dose responses (red curves, without pulse; green curves, with pulse). Hydroxyurea (HU) was applied to synchronize cells at the G1/S phase to avoid involving downstream cell cycle oscillation. Each histogram represents the distribution of GFP signals from ~ 10,000 cells after baseline correction.

The history dependence predicted in our modeling (Figure 3.4) was also experimentally observed, as shown in Figure 3.6. Similar to experimental observations in Figure 3.4, the serum threshold for E2F activation was approximately 1%. Most of the serum-starved quiescent cells stimulated with serum concentration below the threshold remained OFF, while growth stimulation with higher serum concentrations resulted in increasing number of cells switching ON. However, when the quiescent cells were stimulated with strong serum pulse for sufficiently long duration (20% for 5 hours), followed by reduction of serum to a low level, we observed that cells remained at its high state even at serum concentrations below the threshold (1%). This clearly showed that E2F expression depended not only on the input level, but also on its history. In contrast, CycD-d2GFP expression depended only on the input level: it followed the same dose response trajectory, with or without the initial pulse stimulation (Figure 3.6).

3.4 Discussion

Control of the R-point at the G1–S transition of the mammalian cell cycle is a complex process, likely to involve many other regulatory activities not included in our model. Our work here, together with other studies, nevertheless suggests that the Rb–E2F bistable switch, by its unique properties, is fundamental in establishing the R-point. First, it is 'indispensable': the R-point control was lost completely in cells with knockout

of all Rb family members [215,216]. In addition, the ability of a cell to pass the R-point to enter the cell cycle was completely blocked when all E2F activators were knocked out [217]. Second, it is 'sufficient': introducing E2F activity alone can induce quiescent cells to enter S-phase [218]. In addition, cells with E2F turned ON were the very ones that were able to pass the R-point, despite varying strengths and duration of growth stimulation [53]. To our knowledge, the Rb–E2F network is the only system identified in the mammalian cell-cycle machinery that is both indispensable and sufficient for the R-point control.

Our results indicate that the Rb–E2F pathway functions as a bistable switch in response to growth stimulation. The all-or-none and, especially, history-dependent switching behavior of this gene network (in terms of E2F accumulation in each individual cell) provides a mechanistic explanation for the R-point concept. Our work provides a foundation for future studies that will carefully analyze the contribution of individual components in the Rb–E2F gene network in achieving E2F bistability. Such studies will provide further detailed understanding how precisely the R-point is controlled in mammalian cells.

Chapter 4: Transition probability in cell cycle entry

4.1 Introduction

In Chapter 4, we extend our work in Chapter 3 to the study of temporal dynamics of E2F activation in cell cycle entry. Our analyses in this work are significant in three aspects: First, it defines a well-calibrated mathematical model that predicts the stochastic dynamics associated with the cell's transition from quiescence to proliferation. Second, it defines a conceptual framework that reconciles the seemingly conflicting views as represented by the existing phenomenological models: transition probability (TP) model and growth-controlled (GC) model. Third, our work suggests that these phenomenological models can define concise, quantitative phenotypes of the underlying cell physiology (as characterized by their parameters). This aspect has implications for classifying cell states or cell types.

4.2 Background

4.2.1 Temporal variability in the cell cycle

Cell-to-cell variability in the timing of cell-fate commitment is widely observed in biological settings [219,220,221,222]. In particular, the variable transition timing from the quiescent state to the proliferative state is a well-documented phenomenon [53,223,224,225]. In a population of proliferating cells, such variability is reflected in the partitioning of the population into subpopulations at different phases of the cell cycle.

This phenomenon is observed over time even in a population of isogenic cells that have been synchronized by serum starvation. Upon growth stimulation, cells reenter the cell cycle from quiescence and undergo the G1/S transition, but not all cells in the population proceed at the same rate. This rate also differs among different cell types [171,172], and can be modulated by external conditions [226].

4.2.2 Existing models for the temporal variability

To account for the variable transition timing in cell cycle progression, two major types of models have been proposed: transition probability (TP) model [226,227,228,229,230] and growth-controlled (GC) model [174,231,232]. The TP models attributed the temporal variability to random state transitions through different phases of the cell cycle. One of the earliest TP models was proposed to account for the inter-mitotic variability by assuming a single random transition from a non-proliferative (A-state) to a proliferative state (B-phase) [230]. It was subsequently extended to account for the timing variability in cell cycle reentry starting from quiescent (G0) cells [226,229]. In this case, the exponential drop in the fraction of G0 cells over time was suggested to indicate probabilistic nature of the transition. The original model and its subsequent variants have provided excellent fits to various types of experimental data [226,227,228,229,230]. However, a major criticism against the TP model is that the transition probability from the A-phase was assumed to be time-invariant, despite

uneven cell division at mitosis and obvious cell growth or metabolism through the cell cycle [233]. As an alternative, the GC models proposed that the observed temporal variability arises from growth rate heterogeneity within a cell population, rather than random state transitions. Remarkably, this line of models has been able to provide equally good fits to various experimental data [231,234]. Integrating these two lines of thinking, hybrid models proposed cell-size control and random transitions as regulatory elements for progression to cell division [235,236]. However, understanding of the underlying mechanisms for cell-size control and random transitions was limited at the time. Consequently, these models remain descriptive to date, although they provided excellent fits to experimental data.

There has been an active debate between these two lines of thinking since initial proposition of the TP model. While never fully resolved, the debate gradually faded after the concept of the restriction point (R-point) was proposed [169], which we have shown to be controlled by a bistable Rb-E2F switch [53]. We showed that activation of this switch is correlated with the cell's reentry from quiescence into the cell cycle. Interestingly, cell cycle reentry was explored by both the TP and GC models, which were originally developed to describe actively growing cells. For example, the TP models ascribe the quiescence and proliferation to low and high transition probabilities, respectively [226,229]. In addition, the GC model has recently been proposed as an alternative explanation for the "R-point" [174].

4.2.3 Mapping between the stochastic Rb-E2F bistable switch and phenomenological models

The temporal variability described by the GC and TP models is based on the distribution of inter-mitotic times and may differ from temporal variability in E2F activation from quiescence. However, we suggest that the stochastic Rb-E2F model embodies the concepts assumed in the TP model (stochastic) and the GC model (deterministic). Our model predictions and experiments suggest that stochastic activation of E2F can account for temporal variability in cell cycle entry, and the degree of such variability is determined by environmental cues and the regulatory network parameters. These results suggest that the TP and GC models are not mutually exclusive but rather reflect different aspects of the same temporal dynamics in cell cycle entry, as has been speculated [235,237]. In addition, we show that stochastic activation of the Rb-E2F bistable switch under various environmental conditions can be readily mapped into both TP and GC models with a small number of parameters (Figure 4.1). We propose that these parameters can potentially serve as concise, quantitative phenotypes of the cell state.
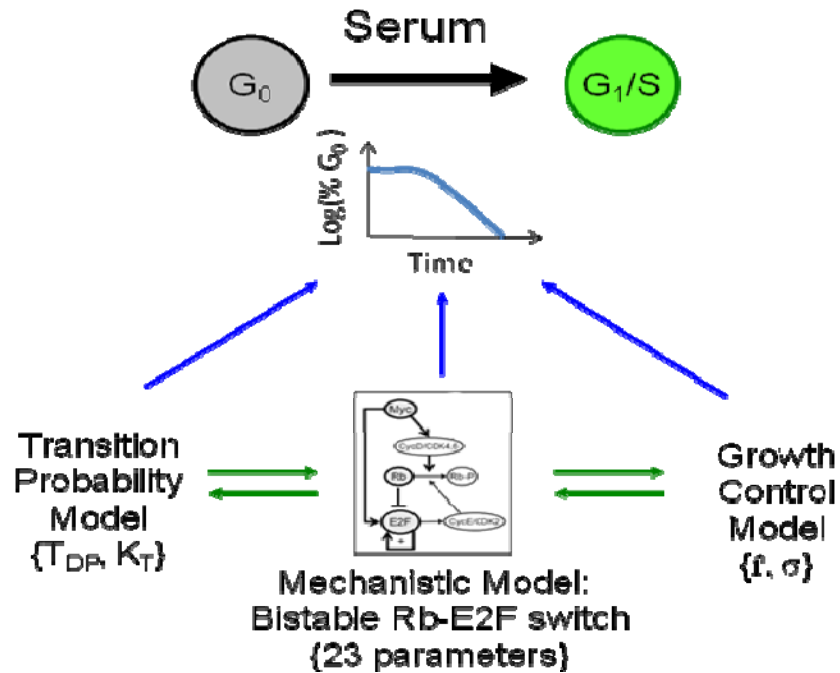
Figure 4.1: Temporal variability in cell cycle reentry. A population of quiescent cells can undergo the G1/S transition with serum stimulation. The timing of cell cycle entry is highly variable in a cell population, characterized by an exponential drop in the percentage of G0 cells over time (G0 exit curve). To account for such temporal variability, two groups of phenomenological models have been previously proposed: transition probability (TP) model, which describes the dynamics of cell cycle entry with transition rate constant (K_T) and a time delay of the cell population (T_{DP}), and growth-controlled (GC) model with the mean G1 growth rate ($\hat{f}$, defined as the reciprocal of G1 time) and the standard deviation of the G1growth rate (σ). We recently demonstrated that the G1/S transition dynamics is governed by a bistable Rb-E2F switch, whose stochastic activation may also account for the G0 exit curve. Here, we propose that the two phenomenological models in essence reflect different aspects of the cell cycle reentry dynamics, and can be recast into the framework of the mechanistic model.

4.3 Mathematical model and experimental analyses

4.3.1 Construction of the mathematical model

The mathematical framework established in our previous work [53] consists of 7 ordinary differential equations (ODEs) and 23 parameters, as shown in Table 3.1, Table 3.2, and Table 3.3. 15 of these parameters are based on the published literature and the rest of the parameters (free parameters) are adjusted to satisfy a number of stringent criteria (see supplementary materials in [53]). In our stochastic model, further adjustments are made to the free parameters, initial conditions, and extrinsic noise amplitude based on additional criteria. These include: 1) bimodal distribution at 0.5% serum stimulation and activation of the majority of cell population at 1% serum stimulation, 2) irreversible E2F activation, 3) noise terms modeled as either reaction-dependent (intrinsic) or –independent (extrinsic) Gaussian noise 4) extrinsic noise as the predominant source to the total noise in the cell population [85], 4) the time delay at 1% serum is approximately 20 hours and is halved at 10% serum and 5) moderate molecule numbers (~1000). These assumptions are consistent with our experimental observations.

Table 4.1: Revised parameter values and initial conditions:

Rate constants		Parameter values, sources, and notes
k_M	1 µM/hr	
k_E	0.15 µM/hr	These values were adjusted together so that:
k_{CD}	0.2 µM/hr	(1) The maximum E2F level is about 5-6 fold higher than the maximum CycD level (based on our experimental
k_{CDS}	0.2 µM/hr	observations).
k_R	0.35 µM/hr	(2) E2F activation threshold is about 1% serum.
k_b	0.001 µM/hr	(3) The simulated E2F level will be around the
k_{CE}	0.5 µM/hr	corresponding Michaelis-Menten parameter (K_E).
K_S	0.3 µM	(4) The time delay at 1% serum is approximately 20 hours
k_{RE}	30 / (µM * hr)	and is halved at 10% serum.

Initial conditions: [Rb]=0.4 µM, [RE]=0.25 µM, [M]=[E]=[CD]=[CE]=[RP]=0µM.

To capture stochastic aspects of the Rb-E2F signaling pathway, we adopt the

chemical Langevin formulation[1,97].

$$\frac{dX_i(t)}{dt} = \sum_{j=1}^{M} v_{ji} a_j [X(t)] + \sum_{j=1}^{M} v_{ji} a^{1/2}_j [X(t)] \Gamma_j(t) + \omega_j(t)$$

where $X_i(t)$ represents the number of molecules of a molecular species i (i=1, …, N) at

time t, and $X(t) \equiv (X_1(t),...,X_N(t))$ is the state of the entire system at time t. $X(t)$

evolves over time at the rate of $a_j[X(t)]$ $(j=1,...,M)$, and the corresponding change in

the number of individual molecules are described in vji. $\Gamma_j(t)$ and $\omega_j(t)$ are temporally

uncorrelated, statistically independent Gaussian noises. This formulation retains the

89

deterministic framework (the first term), and reaction-dependent and reaction-independent noise. The concentration units in the deterministic model were converted to molecule numbers by assuming a certain volume. As a result, the molecule number of each species is 3300x greater than its concentration value. The resulting parameters and initial conditions are shown in Table 4.1.

Since the noisy behavior of the stochastic model depends on the amplitude of the extrinsic noise (ω), we carefully select the optimal value for ω to represent stochasticity in the Rb-E2F signaling pathway. As the high limit, we use criterion 2) to constrain the noise amplitude. We assume that once the system is activated and E2F is high, the random fluctuation in the system does not drive the system to the low state. As the low limit, we use criterion 3) to ensure that the effective extrinsic noise (η_{ext}) contributes more to the overall noise (η_{tot}) in the population than the effective intrinsic noise (η_{int}). η_{tot} is calculated by dividing the standard deviation of a cell population by its mean. η_{int} can be obtained by assuming ω to be zero and calculating the population noise. Then, we can obtain η_{int} by using the relationship, $\eta_{int}^2 + \eta_{ext}^2 = \eta_{tot}^2$. This is shown in Table 4.2.

Table 4.2: Intrinsic and extrinsic noise

ω	η_{tot}	η_{int}	η_{ext}
50	0.12	.052	.11

Here, noise is defined as standard deviation divided by mean.

4.3.2 Experimental setup

4.3.2.1 Cell line, culture, synchronization, serum stimulation, and drug treatment.

The previously established REF52-d2GFP cells [53] carry an expression cassette encoding destabilized EGFP with a half-life of approximately 2 hours (d2GFP; Clontech) under E2F promoters (see supplementary information methods in [53]). These cells were regularly passed on a plate in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% of bovine growth serum (BGS, from Hyclone). To select for resistance in the REF52-d2GFP cells, puromycin (2.5 μg ml^{-1}) was added to the growth media. For all our experiments, cells were synchronized by serum starvation at the quiescent phase. For synchronization, actively growing cells were washed with PBS, trypsinized for cell detachment, and plated at a density of 10^5 cells/well in a six-well culture plate. These cells were incubated in DMEM supplemented with 0.02 % of BGS for approximately 24 hours. For serum stimulation, the starved cells were washed with PBS and were incubated in fresh DMEM with varying concentrations of BGS for the indicated durations. For drug treatment targeting the CycE/Cdk2 complex, Cdk2 inhibitor III (from Calbiochem: Cat #238803, dissolved in DMSO) was added to the fresh media at the indicated concentrations. Hydroxyurea (2mM) was added in the fresh culture media to synchronize cells at the G1-S transition.

4.3.2.2 Flow cytometry

Cells were collected after the indicated durations by trypsinization. Cells were washed with PBS, trypsinized for approximately 3 minutes for detachment, and supplemented with 10% BGS. Subsequently, cells were centrifuged at 960g for 2 minutes, and their supernatant was aspirated. Finally, they were fixed by resuspending the cells in 1% formaldehyde. For each sample, approximately 10,000 cells were assayed for their GFP signals with a flow cytometry system (BD FACSCanto™II Flow Cytometry System), and their GPF signals were analyzed with Matlab.

4.3.2.3 Western blots

REF52-d2GFP cells were serum-starved (BGS=0.02%) for 24 hours before they were treated with varying concentration of Cdk2 inhibitor III and serum. After 24 hours of serum/inhibitor drug treatment, cell lysates were collected and Western blotting was conducted with primary antibodies recognizing Rb phosphorylation at Cdk4-specific serine 780 (Santa Cruz, #sc-12901-R) and at Cdk2-specific threonine 821 (Santa Cruz, #sc-16669-R). These were conjugated with anti-rabbit secondary antibodies (GE Healthcare, #NA934) for detection. As a loading control, actin was measured with actin-recognizing primary antibodies (Santa Cruz, #sc-8432) conjugated with anti-mouse secondary antibodies (GE Healthcare, #NA9310). The effects of the inhibitor drug are shown in Figure 4.2.
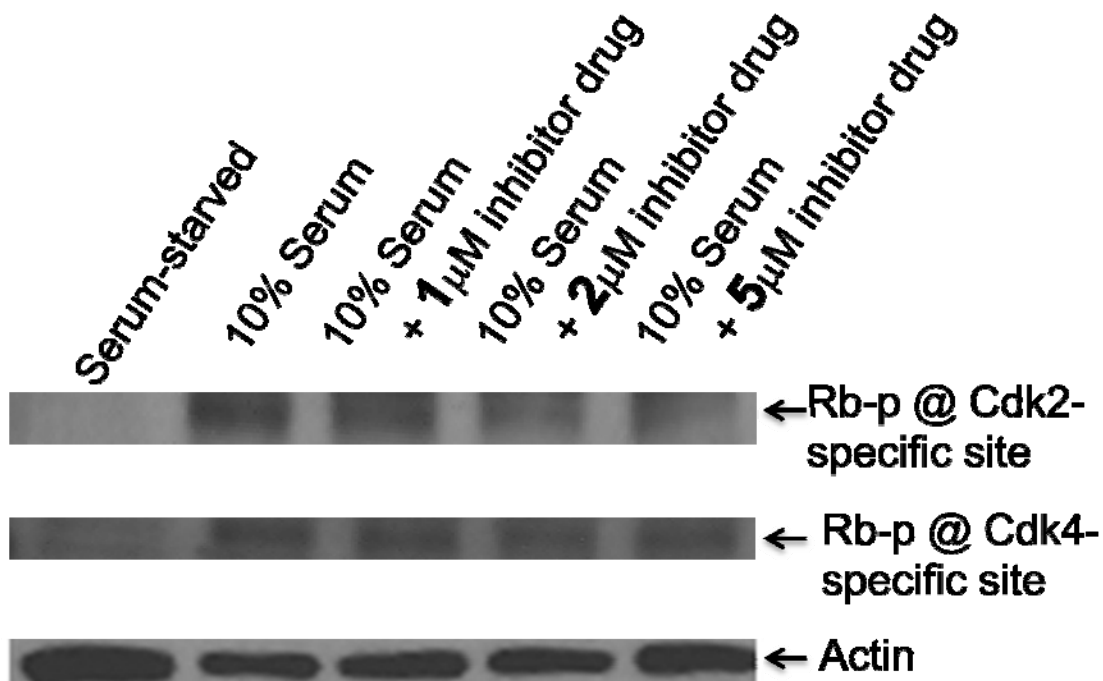
Figure 4.2: Specificity of Cdk2 inhibitor III. To demonstrate the effect of the inhibitor drug on Cdk2 kinase activity, we measured Rb phosphorylation at the Cdk2-specific and Cdk4-specific residues for varying inhibitor drug concentrations. An isogenic population of serum-starved REF52-d2GFP cells was used for Western Blotting. In serum-starvation condition (serum=0.02%), Rb phosphorylation at either residue was negligible. With serum stimulation (serum=10%), a significant increase in Rb phosphorylation at both residues was observed. For increasing inhibitor drug concentration, Rb phosphorylation efficiency decreased at the Cdk2-specific residue, but no significant change was observed at Cdk4-specific residue.

4.3.3 Simulation and experimental results

Given the bistable switching property of the Myc-Rb-E2F network, we

hypothesized that this network, when subjected to noise, might demonstrate variable

timing in E2F activation, which in turn might account for the temporal variability

observed in the cell cycle entry. This hypothesis is based on the strong correlation we

previously observed between E2F activation and DNA synthesis [53]. Using our

stochastic Rb-E2F model, we predicted temporal dynamics of E2F activation and

validated these predictions experimentally.

4.3.3.1 Modulation of E2F activation by serum stimulation: Simulation results

The fluctuations in the bistable switch result in significant discrepancies between

stochastic and deterministic simulations and the average of the stochastic simulations

significantly differs from the deterministic simulations [87,238,239,240]. Given a set of

initial conditions and parameters in the Myc-Rb-E2F network, the simulated time

courses from a deterministic model are fixed (black line in Figure 4.3), but those from a

stochastic model show drastically variable trajectories (gray lines in Figure 4.3). For

example, the stochastic Rb-E2F model can generate two modes of E2F at the steady-state

when stimulated with weak input as shown in Figure 4.3. We define a switching

threshold that distinguishes the low mode, which corresponds to non-activated

subpopulation of cells, from the high mode, which represents activated populations.

This threshold (horizontal red line in Figure 4.3) can be used to calculate the percentage of activated cells over time (or G0 exit curve). The minimum time required for E2F to reach the switching threshold is defined as a switching time (vertical red line in Figure 4.3). Similarly, the deterministic time-courses are fixed and stochastic time-courses show variable trajectories (data not shown) for strong input, but the distribution E2F activities exhibits a single mode at the high E2F level, rather than two modes.
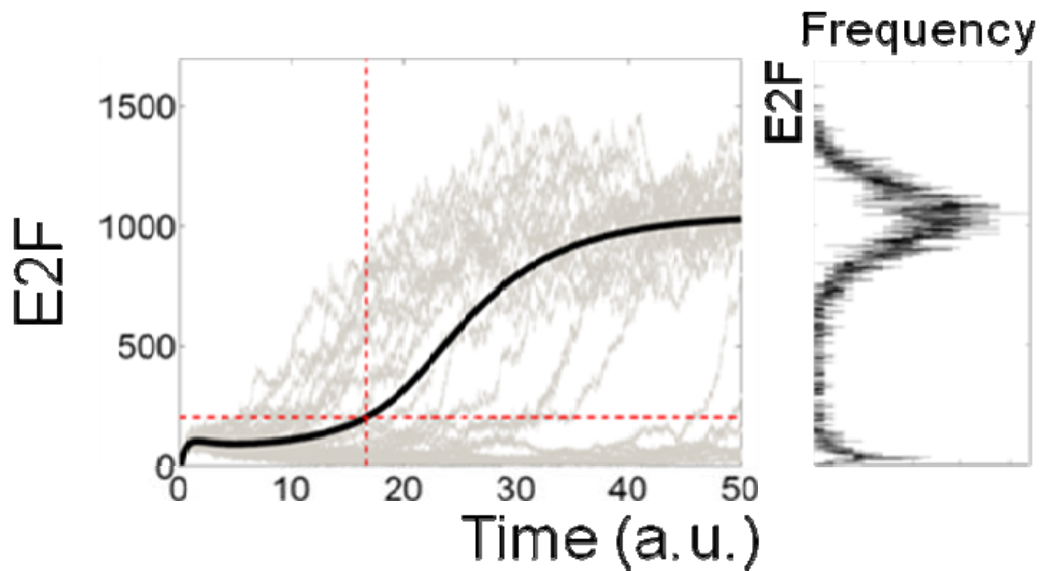
Figure 4.3: Stochastic vs. deterministic Rb-E2F model. The noise in Rb-E2F switch drives a population of cells (5000 simulations) to two modes (low and high), separated by a switching threshold (horizontal red line). The minimum time required to reach this threshold is defined as the switching time (vertical red line). The time evolution is the same for a given set of parameters in the deterministic model (black line). However, 25 stochastic simulations exhibit variable time delays (gray lines).

Based on our simulations and definitions in Figure 4.3, we obtained G0 exit curves for weak and strong input conditions as shown in Figure 4.4. These G0 exit

curves are analogous to the α-curve in the TP model, which represents frequency

distribution of inter-mitotic times [230]. Both G0 exit curve and α-curve can be fitted

with an exponential curve with two parameters (black dotted curve in Figure 4.4):

transition rate (K_T) and time delay (T_{DP}). To estimate T_{DP} and K_T of a cell population, we fit

the G0 exit curve with an exponential function: $N(t) = N_0 e^{-K_T*(t-T_{DP})}$; $N(t \leq T_{DP}) = N_0$, where

N_0 (=100%) is the initial percentage of cells in G0. K_T and T_{DP} were defined by minimizing the

root mean square deviation (RMSD) between our data and the exponential function.

   The fitting between the G0 exit curve and α-curve is possible because both exhibit

an initial time delay followed by an exponential drop [226,229,230]. The transition rate of

the G0 exit curve is inversely proportional to the temporal variability of the cell

population. For example, a population of cells with more synchronous E2F activation

E2F would have a higher transition rate than that of a population with less-synchronous

E2F activation. If the cells were fully synchronized, the G0 exit curve would have an

infinite transition rate.
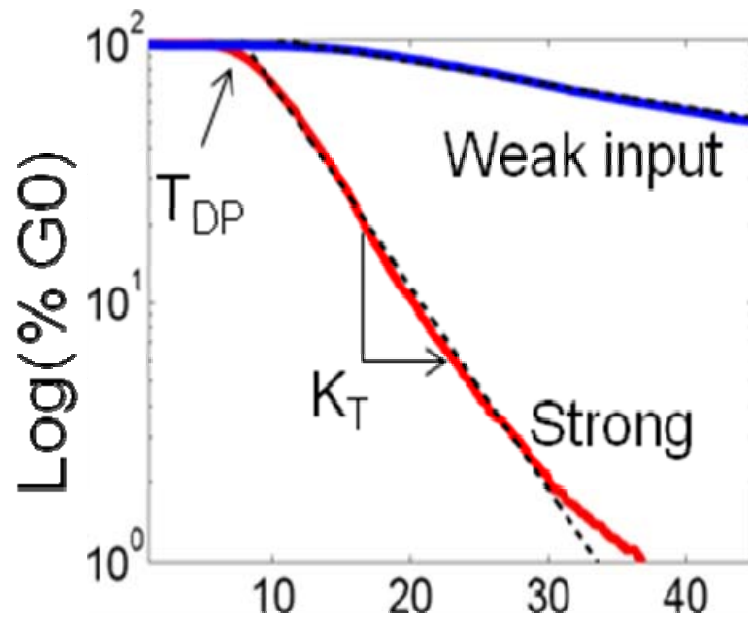
Figure 4.4: G0 exit curve. The percentage of G0 cells over time (G0 exit curve) is plotted for a population of 5,000 simulated cells stimulated at strong (red line, S = 5) and weak (blue line, S = 0.5) input concentrations. The G0 exit curve for the strong input is fitted with an exponential function (black dotted line), $N(t) = N_0 e^{-K_T*(t-T_{DP})}$; $N(t \leq T_{DP}) = N_0$, where N_0 (=100%) is the initial percentage of cells in G0, K_T is the transition rate, and T_{DP} is the population time delay. The standard error of these estimated parameters can be approximated with Monte-Carlo standard deviatioin. For increasing input strength, the transition rate was predicted to increase ($K_T = 0.018\pm0.0015$ hr^{-1} for weak input and K_T= 0.19 ±0.018 hr^{-1} for strong input) and the time delay was predicted to decrease (T_{DP} = 11.0±1.2 hours for weak and T_{DP} =8.3±0.48hrs for strong input).

Our simulated E2F activation dynamics in Figure 4.4 predict serum-dependence of transition rate and time delay. For a weak input (K_T = 0.018±0.0015 hr^{-1}and T_{DP} = 11.0±1.2 hrs, blue line in Figure 4.4), most cells were expected to remain inactivated and the percentage of G0 cells would decrease slowly. This is because the impact of noise acting on the Rb-E2F bistable switch was significant enough to activate E2F in some cells, but not in other cells. This would lead to a bimodal distribution of E2F activity (Figure 4.5), which is consistent with previous experimental observations in mouse fibroblasts [228,241,242].
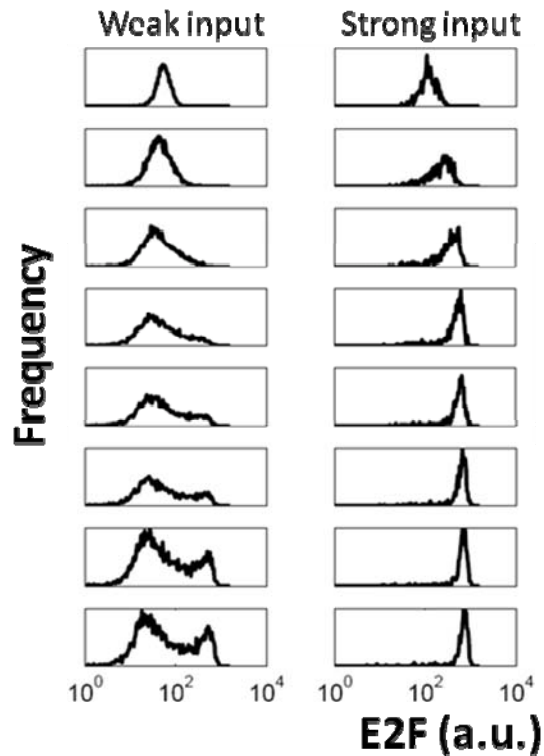
Figure 4.5: Simulated temporal dynamics of E2F activation. The Rb-E2F bistable switch was stimulated with weak (S=0.5) and strong (S=5) input strengths. E2F distributions from 5,000 simulations were sampled at various time points for both conditions. For weak input strength, bimodality was predicted to emerge at around the16th hour. At strong input strength, however, bimodaity was expected to be less clear.

In contrast, the impact of noise was negligible with strong input and all cells were predicted to be activated at high transition rate (K_T= 0.19 ±0.018 hr^{-1} and T_{DP} =8.3±0.48 hrs, red curve in Figure 4.4). The selective response of the Rb-E2F bistable switch to noise would cause an increase in K_T with increasing input strength as the population moves from bimodal distribution to monomodal distribution at the high

mode (Figure 4.6A). At sufficiently high input strength, further increase in input

strength may have negligible effect on K_T. In contrast, T_{DP} may decrease with increasing

input strength and reach a plateau at sufficiently high input strength (Figure 4.6B).
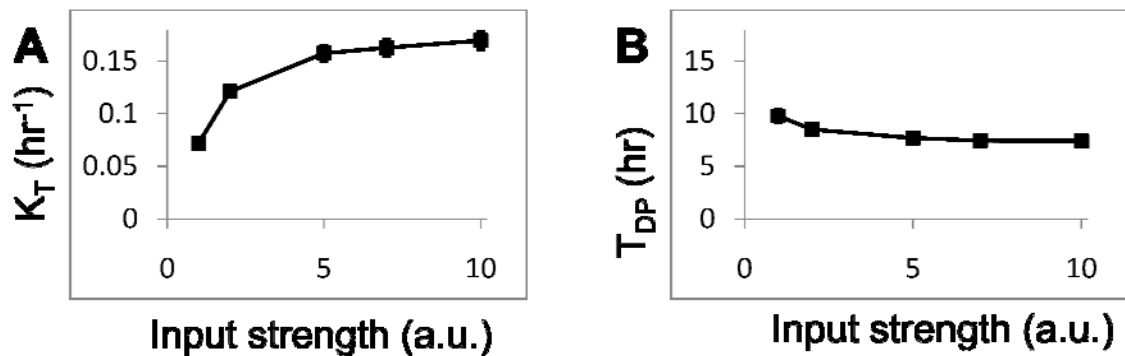
Figure 4.6: Dependence of K_T and T_{DP} on input strength. A) K_T was predicted to
increase with increasing input strength and reach a plateau at sufficiently strong input.
B) T_{DP} was predicted to decrease with increasing input strength. Here, the error-bars
represent the standard error of the estimated parameters (K_T and T_{DP}) approximated
with the Monte-Carlo standard deviation.

The dependence of K_T and T_{DP} on input strength can be recapitulated with a

minimal bistable model (Figure 4.7). Similar to our simulation results with the full Rb-

E2F model, our minimal model predicted increase in K_T and decreasing T_{DP} with

increasing input strength. This suggests that the dependence of K_T and T_{DP} may be an

intrinsic property of the bistable switch.
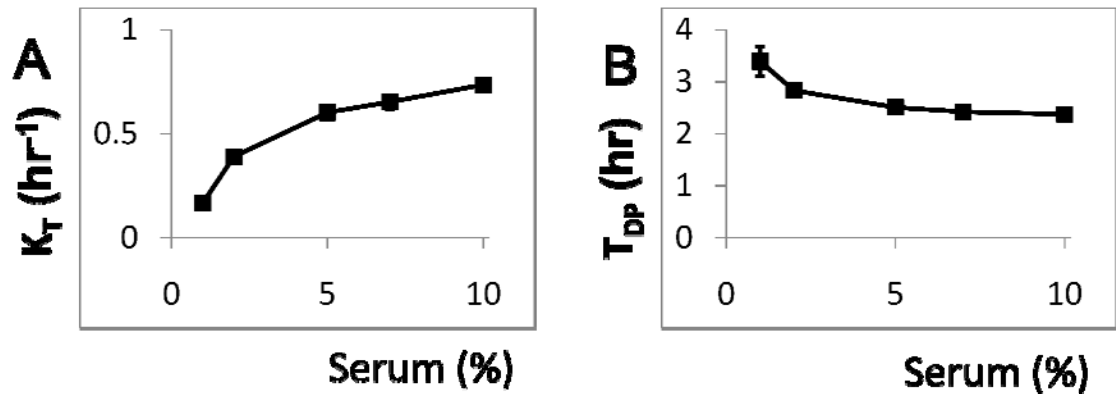
Figure 4.7: Dependence of K_T and T_{DP} on input strength in a minimal model. A) A minimal model was used to recapitulate the temporal dynamics of the bistable Rb-E2F switch. The model describes activity of a molecule X: $\dfrac{dX}{d\tau} = \left(\dfrac{[S]}{1+[S]} \right) \left(\dfrac{k_a[X]^2}{[X]^2 + K^2} \right) - [X]$,

where S is the input strength, k_a (=1) is the lumped rate term for synthesis and feedback strength, and K (=0.15) is a dimensionless Michaelis-Menten constant. This minimal model was converted to a stochastic model using the chemical Langevin formulation. The transition rates were calculated for cell populations stimulated at various input strengths. The transition rate increased with input strength and reached a plateau at sufficiently high input strength. B) In the minimal bistable model, the time delay decreased with increasing input strength and reached a plateau at sufficiently high input strength.

4.3.3.2 Modulation of E2F activation by serum stimulation: Experimental validations

To validate our model predictions, we measured E2F activation using E2F-d2GFP cells [53] stimulated with various serum conditions. Prior to serum stimulation, the REF52-d2GFP cells were synchronized at quiescence by serum-starvation (0.02% bovine growth serum, BGS) with basal E2F-GFP expression (Figure 4.8). Upon weak serum stimulation (0.3% BGS), E2F activation was rather stochastic and only a subpopulation of the cells switched to the high E2F mode over time. At earlier time

101

points ($0\sim15$th hour), the difference in E2F level between the non-activated and activated

cells was small. The difference between the two modes became increasingly clear,

resulting in distinctive bimodality starting at 18th hour. In contrast, upon strong serum

stimulation (5% BGS), E2F activation was more synchronous. The cell population

gradually switched to the high mode with greater temporal synchrony without

demonstrating detectable bimodality at any time point (Figure 4.8). It is possible that

noise may partition the cell population into two subsets (active and inactive towards

proliferation) temporarily even at high serum stimulation. However, simulations

suggest that accumulation of E2F in the activated cells at earlier time points may not be

significant enough to result in detectable difference between the two subsets (Figure 4.5).
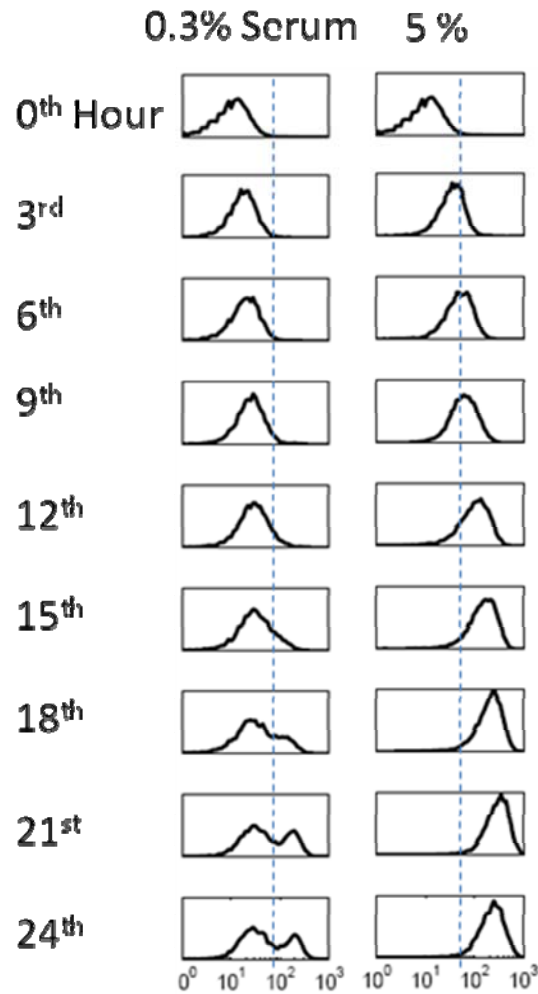
Figure 4.8: Experimental E2F distributions at varying MOIs over time. The temporal dynamics of a cell population depends on serum concentration. At 0th hour REF52-d2GFP cells were synchronized in quiescence by serum-starvation (24 hours at 0.02% bovine growth serum, BGS). These cells were then stimulated with either 0.3% or 5% serum, and corresponding E2F levels were determined by flow cytometry. The cell population treated with 0.3% serum exhibited bimodal distribution of E2F at the steady-state. In contrast, monomodal distribution was observed at the steady-state in the cell population treated with 5% serum.

Based on the distribution of E2F in Figure 4.8, we calculated the percentage of non-activated cells and obtained a G0 exit curve for each serum condition (Figure 4.9A). Consistent with predictions in Figure 4.4, we observed an increase in K_T and decrease in T_{DP} for increasing serum concentration ($K_T = 0.021\pm0.0063$ hr^{-1} and $T_{DP} = 5.5\pm1.2$ hours at 0.3% serum and $K_T = 0.16\pm0.018$ hr^{-1} and $T_{DP} = 3.9\pm0.72$ hours at 5% serum), reminiscent of modulation of the α-curve by serum [226,230]. An independent experiment under the same conditions on a different day exhibited similar dependence of K_T and T_{DP} on serum (Figure 4.9B). Consistent with our predictions in Figure 4.6, we observed initial increase in K_T, followed by a plateau, and decreasing T_{DP} for increasing serum concentrations (Figure 4.10).
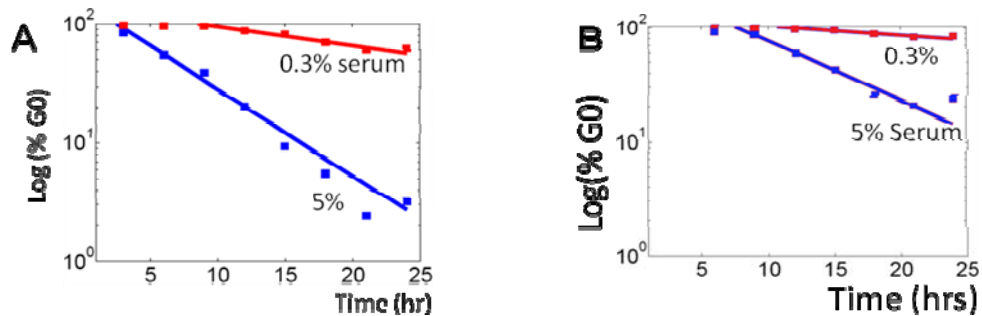
Figure 4.9: Experimental G0 exit curves. The thresholds shown as dotted blue lines in **Figure 4.8** were used to calculate the percentage of cells in the low mode of E2F. The two G0 exit curves showed that transition rate increased ($K_T = 0.021\pm0.0063$ hr^{-1} at 0.3% serum and 0.16 ± 0.018 hr^{-1} at 5% serum) and time delay decreased ($T_{DP} = 5.5\pm1.2$ hours at 0.3% and 3.9 ± 0.72 hours at 5% serum) with serum concentration. B) G0 exit curves for an independent set of experiments
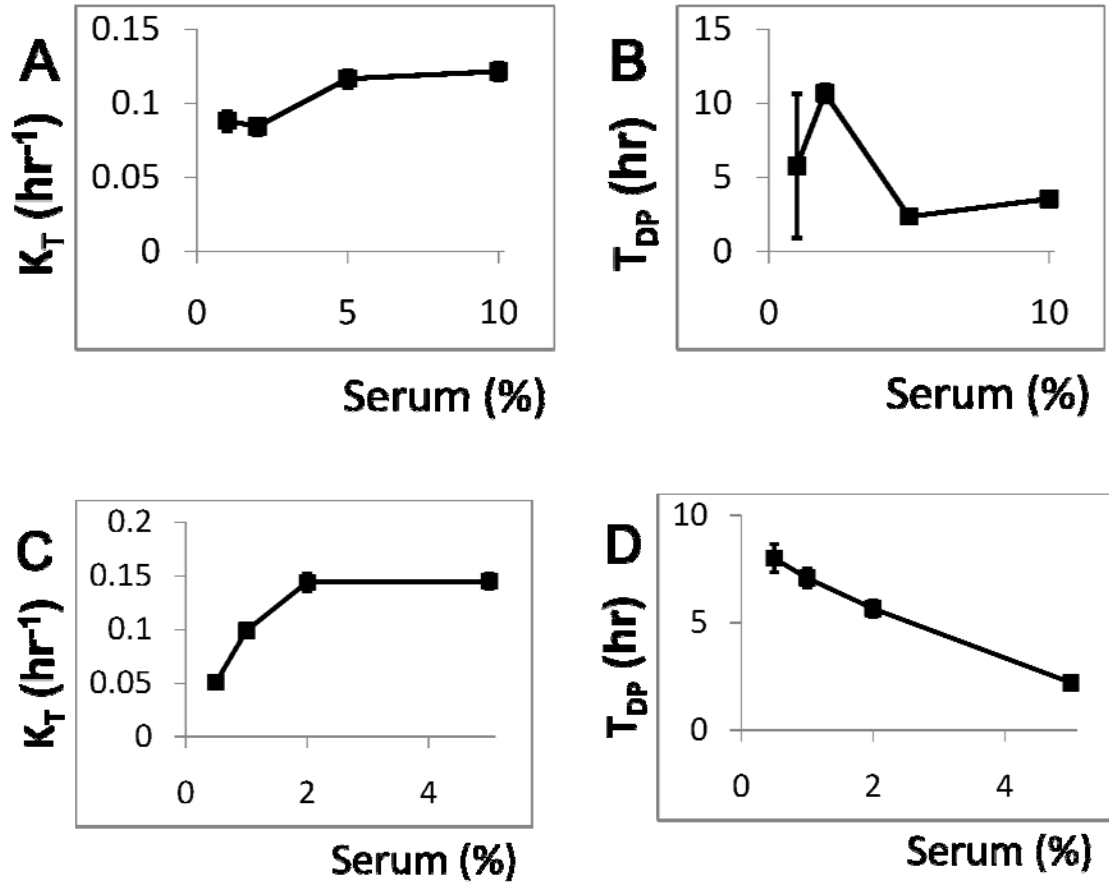
Figure 4.10: Dependence of K_T and T_{DP} on serum concentrations. A) The transition rate increased with serum concentration. B) The time delay decreased with serum concentrations. C-D) the transition rate and time delay for increasing serum from an independent set of experiments

4.3.3.3 Modulation of stochastic E2F activation by strength of CycE-mediated feedback: Simulations

The temporal dynamics of biological systems often strongly depend on network

parameters [243], and consequently, the transition rate of cell cycle entry may be

modulated by nodal perturbations. This is exemplified in a recent study on the yeast cell

cycle [244], which demonstrated that a positive feedback by G1 cyclins is responsible for temporal coherence in gene expression and proper division timing of the yeast cells. Loss of this feedback control in the cell cycle machineries was shown to promote incoherent gene expression and abnormal duration for yeast budding. Interestingly, a similar feedback module by G1 cyclin (CycE) can be found in the Myc-Rb-E2F network also, suggesting its possible role in the control of temporal dynamics.

To investigate modulation of the transition rate by nodal perturbations in the Myc-Rb-E2F network, we introduced *in-silico* perturbations on one particular node: the CycE/Cdk2 complex, which forms a positive feedback loop. Our bifurcation analyses predict that weakening of the CycE-mediated positive feedback loop will desensitize the Rb-E2F bistable switch to serum stimulation, requiring a higher critical serum concentration (Figure 4.11) for E2F activation. Similarly, we predict desensitization to serum when CycD is down-regulated or when Rb is up-regulated (data not shown).
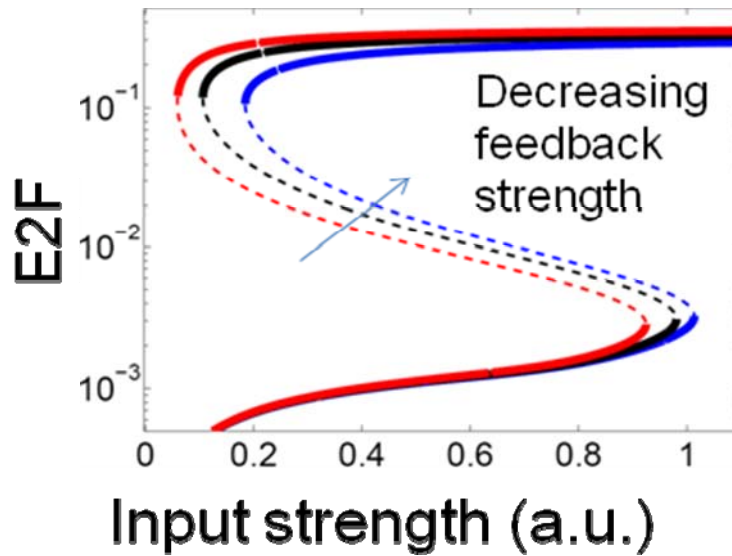
Figure 4.11: Bifurcation analyses of the Rb-E2F bistable switch for varying strengths of the CycE-mediated positive feedback. The strength of the CycE-mediated positive feedback determines the sensitivity of the system to serum stimulation. Bifurcation analyses of the Myc-Rb-E2F network with weak (Rb phosphorylation rate constant k_{P4} = 9, blue), intermediate (k_{P4} = 14.4, black), and strong strength (k_{P4} = 18, red) of the positive feedback were performed. For decreasing strength of the positive feedback, the system became less sensitive to the input strength, requiring greater critical input strength for E2F activation.

Such desensitization of the network to serum is expected to modulate the temporal dynamics of E2F activation. When the positive-feedback strength by CycE is weakened, our simulations in Figure 4.12 (corresponding simulated distributions in Figure 4.13) predicted increase in the time delay and decrease in the transition rate. For strong feedback strength, K_T was estimated to be 0.35 hr^{-1}. This value was reduced to 0.19 hr^{-1} and 0.14 hr^{-1} for intermediate and weak feedback strength, respectively. In contrast, T_{DP} for strong feedback input (=4.4 hours) was predicted to increase to 7.6

hours for intermediate feedback strength, and further extend to 11.4 hours for weak

feedback strength. Similar dependence of K_T and T_{DP} on the feedback strength was

predicted for all serum concentrations (Figure 4.12 B and C).
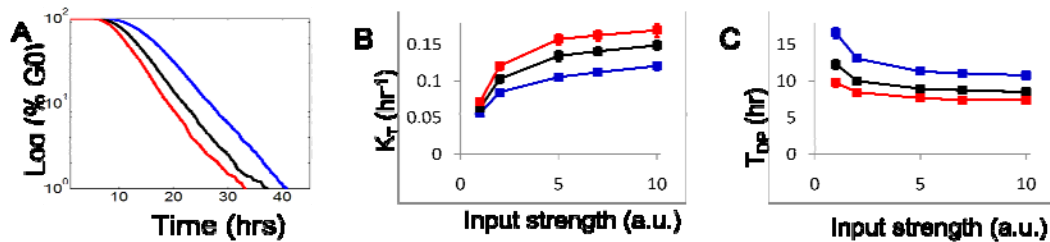
Figure 4.12: Predicting temporal dynamics of E2F activation under varying feedback strengths. A) The temporal dynamics can be modulated by adjusting the feedback strength. At the saturating input level (S = 10), the Rb-E2F switch was subjected to varying degrees of feedback strength mediated by CycE. G0 exit curves from 5,000 simulations were constructed for strong (red line, k_{P4} = 18), intermediate (black line, k_{P4} = 14.4), and weak (blue line, k_{P4} = 9) feedback strengths. For decreasing strength of the positive feedback, our simulation predicted decrease in the transition rate (K_T = 0.20±0.023 hr^{-1} for strong, 0.17±0.016 hr^{-1} for intermediate, and 00.14±0.0084 hr^{-1} for weak feedback strength), and increasing the time delay (T_{DP} = 7.8±0.52 hours for strong feedback, 9.0±0.51 and 11.6±0.40 hours for intermediate and weak feedback strength, respectively). B) Increase in K_T for increasing strength of the positive-feedback was predicted for all input strengths. C) Decrease in T_{DP} for increasing strength of the positive feedback was predicted at all input strengths.
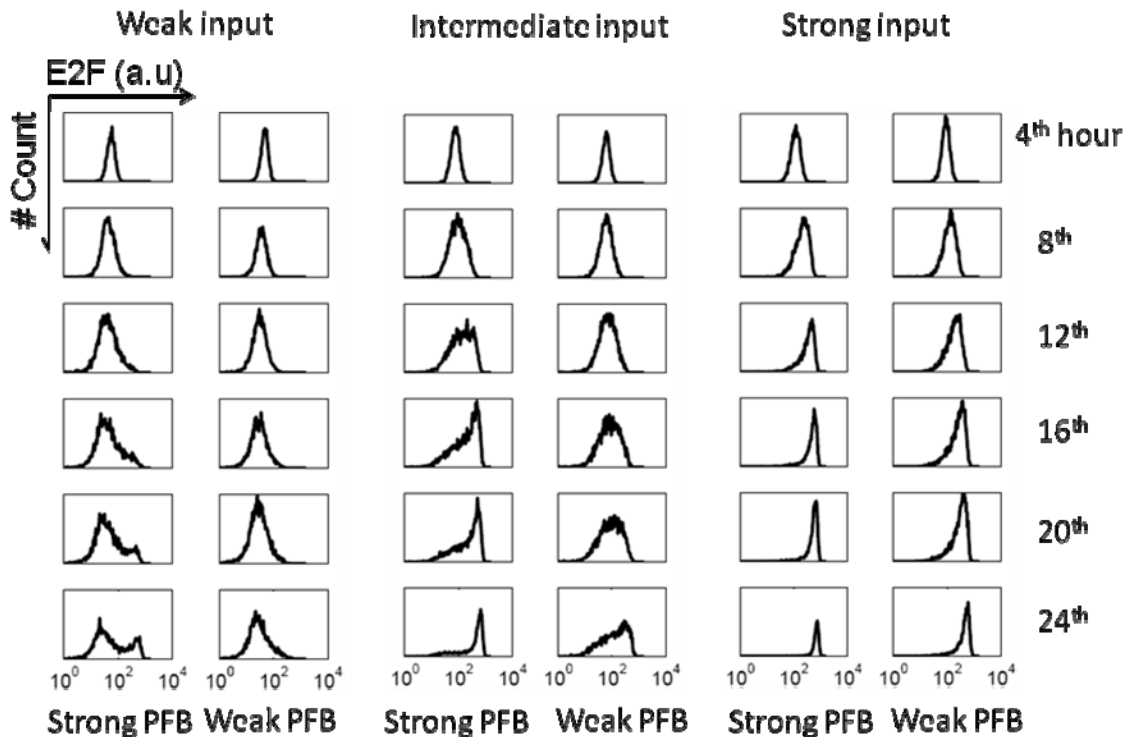
Figure 4.13: Predicted modulation of the temporal dynamics of E2F activation.
Temporal dynamics of E2F activation were simulated at varying input strengths (weak
→ S=0.5, intermediate → S =1, and strong → S = 5) and varying CycE-mediated positive
feedback strengths (strong→ k_{P4} = 18 and weak → k_{P4}=9). With strong positive feedback
(PFB), bimodality was predicted for weak input while monomodality (E2F ON) was
predicted for intermediate and strong stimulations. With weak positive feedback,
however, the percentage of E2F activation was predicted to decrease for weak and
intermediate input strengths. For strong input, however, the effect of the positive
feedback strength was minor.

4.3.3.4 Modulation of stochastic E2F activation by strength of CycE-mediated feedback: Experimental validations

To test these predictions experimentally, we perturbed the Myc-Rb-E2F network by applying varying concentrations of a cyclin-dependent kinase inhibitor (Cdk2 inhibitor III). In the context of the current study, which focuses on the dynamics leading to E2F activation, the impact of the Cdk2 inhibitor is primarily the inhibition of the CycE/cdk2 complex. However, we note that the inhibitor would also affect other components of the cell cycle regulation that were not considered in the model (i.e. CycA/cdk2 complex). When the CycE node was perturbed experimentally, we observed inhibitor dose-dependent changes in the steady-state dynamics of E2F activation. As shown in Figure 4.14A, increasing dose of the inhibitor drug reduced the percentage of cells at high E2F mode at the steady-state (24th hr), consistent with our bifurcation analysis in Figure 4.11A. For example, without the Cdk2 inhibitor, 1% serum was required for E2F activation in half of the cell population. At 2μM of the drug, Cdk2 inhibitor, 2% serum was required to achieve the same fraction of E2F activation. Such desensitization to serum stimulation was true for all drug concentrations tested (Figure 4.14).
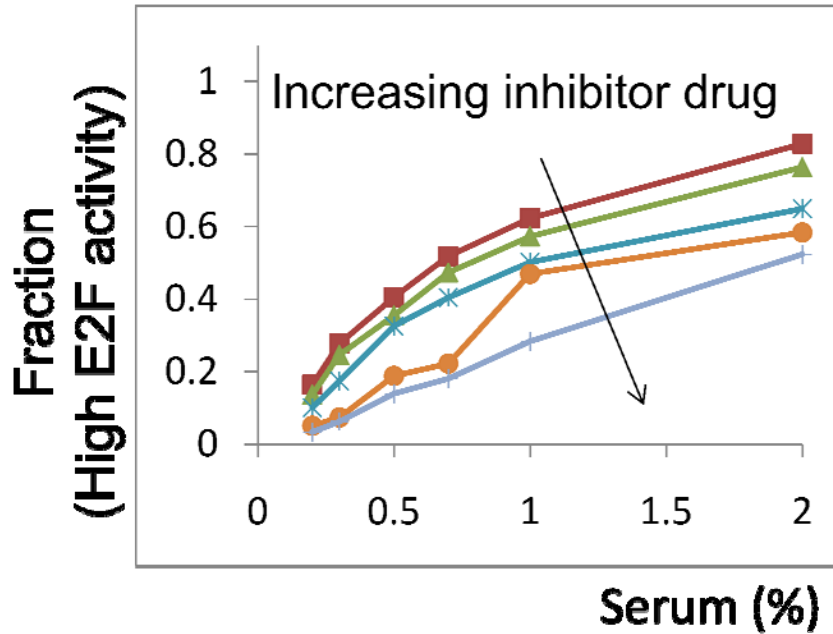
Figure 4.14: Fraction of activated E2F in a cell population under varying inhibitor drug strength and serum concentrations. E2F activity was measured at the steady-state in response to varying concentrations of the Cdk2 inhibitor and serum. The fluorescence level was quantified by flow cytometry. For each serum and inhibitor drug condition, the fraction of cells at high E2F mode was determined at 24th hour. For each serum concentration, increasing drug dose led to a decreasing fraction of cells at high E2F mode.

Next, we tested modulation of temporal dynamics by the Cdk2 inhibitor. At 2% serum, we applied Cdk2 inhibitor III (2 μM) to monitor its effect on E2F activation over time. Our results in Figure 4.15A show that the transition rate of the cell population decreased (from K_T = 0.066±0.013 to 0.022±0.0091 hr^{-1}) and time delay increased (from T_{DP} = 8.7±1.2 hrs to T_{DP} = 4.7±2.8 hrs) with addition of the inhibitor drug. Such decrease in K_T with the inhibitor drug is consistent with our model predictions in Figure 4.12, and was

111

observed for all serum concentrations, as shown in Figure 4.15B (distributions of E2F in Figure 4.16). As predicted, time delay generally decreased with serum concentrations and it increased in the presence of the inhibitor drug (Figure 4.15C).
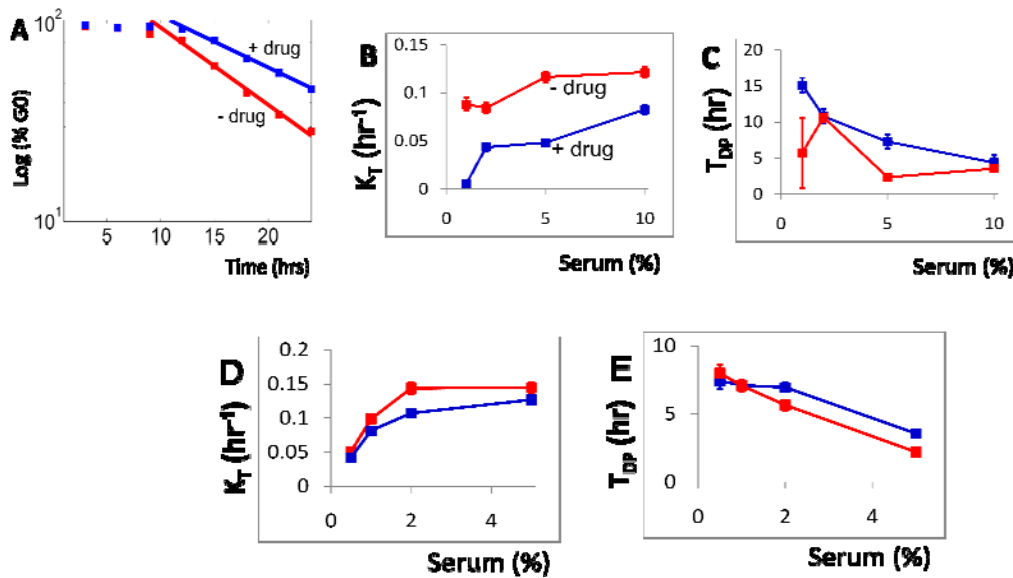
Figure 4.15: Temporal dynamics of E2F activation under drug treatment. A) The temporal dynamics of E2F activation is altered when CycE-mediated positive feedback is weakened. At 2% serum, we applied Cdk2 inhibitor III at 2μM (blue curve) and monitored the effect on E2F activation over time by flow cytometry. Compared to the case without drug (red curve), the transition rate decreased from 0.06±0.013 to 0.022±0.0091 hr^{-1} and the time delay increased from 4.7±2.8 to 8.7±1.2 hours. B) Targeting the CycE-mediated positive feedback modulates the transition rate. For a given drug dose, time-courses of cell populations treated with varying serum concentrations were obtained and the transition rate was calculated for each serum condition. The transition rate increased with serum concentration in the presence or absence of the inhibitor drug. With the addition of the inhibitor drug, the transition rate decreased for all serum concentrations. C) Time delay decreases with increasing concentration in the presence or absence of the drug. D-E) transition rate and time delay for an independent set of experiments in the presence (blue line) and absence (red line) of the Ckd2 inhibitor drug.
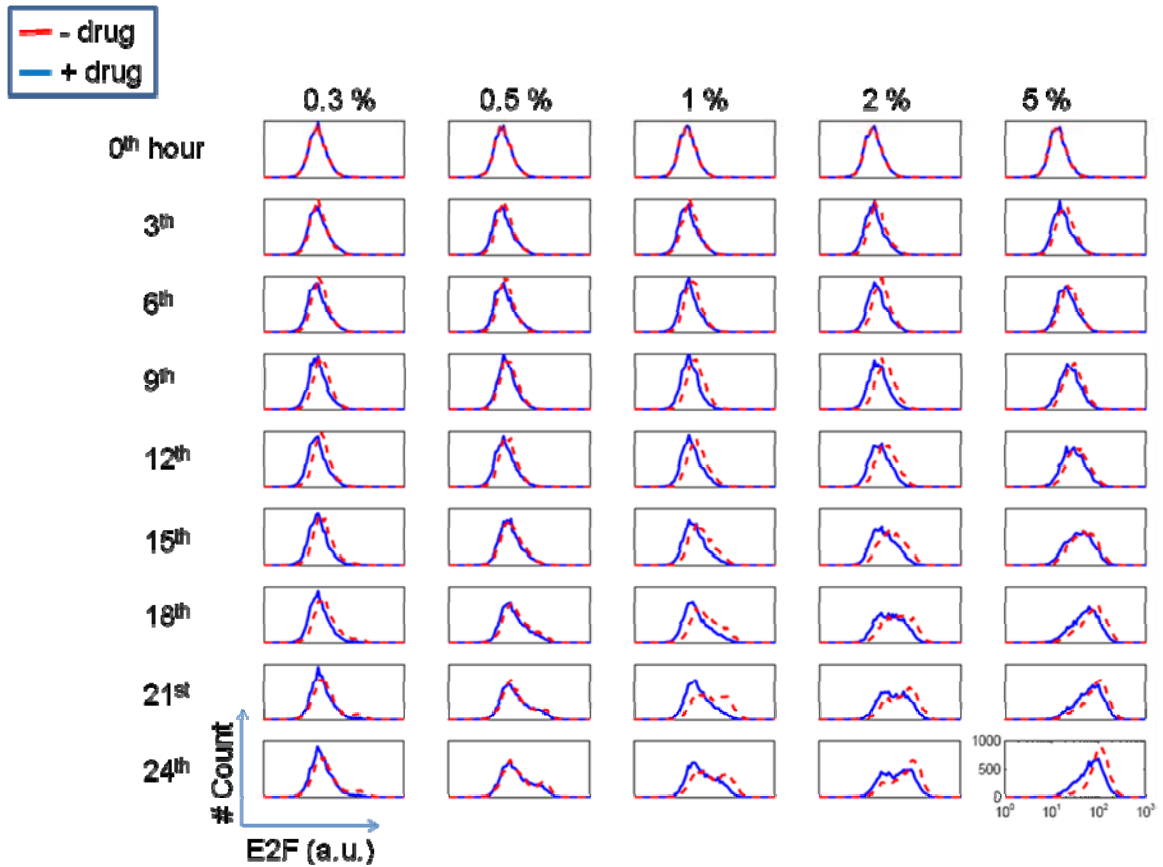
Figure 4.16: Experimentally measured E2F time courses for varying serum concentrations, in the absence or presence of Cdk2 inhibitor (at 3μM). At 0th hour REF52-d2GFP cells were synchronized in quiescence by serum-starvation (24 hours at 0.02% serum), stimulated with varying serum concentrations (with or without Cdk2 inhibitor), and measured for GFP by flow cytometry at the indicated time points.

4.3.3.5 Mapping simulated stochastic E2F activation into TP and GC model

Throughout this study, we have analyzed the temporal dynamics of E2F

activation by extracting a set of parameters defining the TP model (transition rate and

time delay). This parameter extraction establishes a connection with the mechanistic Rb-

113

E2F model. Similarly, the GC model parameters (growth rate $\bar{r}$ and its variance) can be

extracted from the stochastic dynamics of E2F activation, and a connection between the

GC model and the mechanistic Rb-E2F model can be established, as shown in Figure

4.17. Our results predicted increasing growth rate (Figure 4.17A) and decreasing

variance of the growth rate (Figure 4.17B) for increasing input strength. However,

decreasing the strength of the positive feedback mediated by CycE was predicted to

reduce growth rate without affecting its variance. Such parameter extraction of

parameters defining the phenomenological models provides a quantitative mapping

between the phenomenological models and the mechanistic Rb-E2F model.
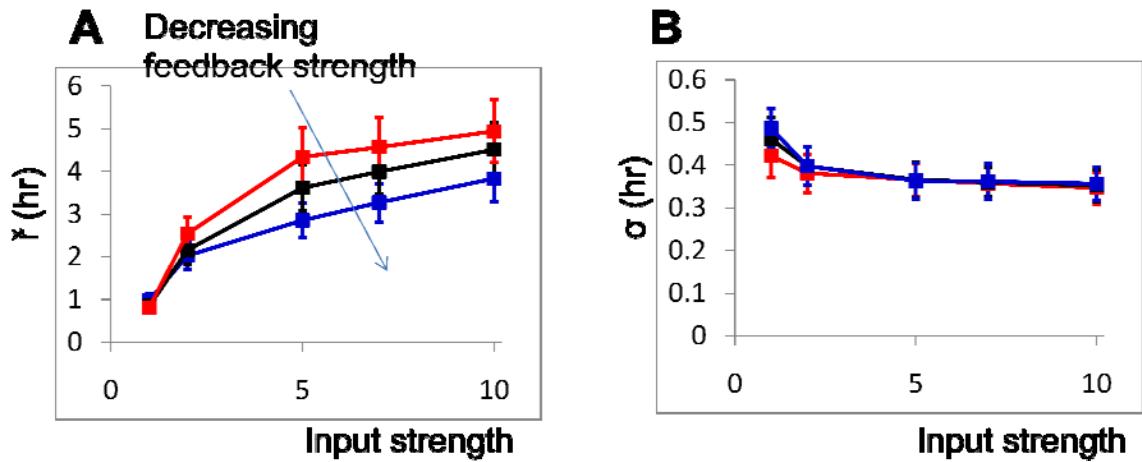
Figure 4.17: Mapping the stochastic dynamics of E2F activation with the GC model.
Simulation results from the stochastic Rb-E2F model are fitted with the GC model with
two parameters (adapted from the G-rate model [231]), which is defined as

$T = 1/X$, where $X \sim N(\bar{r}, \sigma)$. $\bar{r}$ is the mean growth rate over the entire cell population
(A) and σ is the standard deviation of the growth rate (B). The two parameters of the GC
model were estimated by assuming that X is a random variable from an inverse-gamma
distribution. The standard errors of these estimated parameters were approximated with
Monte-Carlo standard deviation. Our simulations predicted increasing growth rate for
increasing input strengths and positive feedback strengths (k_{P4}=9, 14, and 18 for blue,
black, and red lines, respectively). No significant change in the variance was predicted.

4.4 Discussion

Focusing on E2F activation, we show that the temporal variability in cell cycle

entry from quiescence can be quantitatively modeled by the stochastic activation of the

bistable Rb-E2F switch [230]. In addition, we show that the degree of such variability can

be modulated by varying the input strength or by perturbing the network parameters.

Our model predictions are overall consistent with experimental measurements.

In particular, our analysis indicates that serum and the Cdk2 inhibitor drug exerts

opposite influences on the temporal dynamics of E2F activation: transition rate increases

and time delay decreases with increasing serum, but transition rate decreases and time

delay increases with increasing Cdk2 inhibitor concentrations. We suggest that such a

well calibrated stochastic Rb-E2F model may guide further experimental analysis to gain

insights into the systems dynamics underlying cell cycle entry. For example, our model

predicts that reducing the CycD/Cdk4,6 activity may have similar effects as the Cdk2

inhibitor, while knocking down Rb may increase transition rate (data not shown). In

addition, we can predict population dynamics of E2F activation under combinatorial

perturbations including growth factors, inhibitor drugs targeting the Myc-Rb-E2F

network, or mutations within this network.

Equally importantly, we further show that these predicted stochastic dynamics

of the Rb-E2F model can be quantitatively mapped into two lines of phenomenological

models reflecting seemingly conflicting views: the TP model and the GC model. For a

given set of parameters defining the stochastic model, the simulated stochastic E2F

activation at the population level can be uniquely described by a set of parameters

defining the TP model or the GC model (Figure 4.12 and Figure 4.17). Furthermore,

different sets of parameters in the stochastic model would lead to different parameters

in the TP or the GC models. We propose that this mapping provides a simple conceptual

framework that reconciles the different views reflected in the TP and GC models, which

have been a source for an unresolved debate over the last several decades. In other

words, the stochastic model can be considered as a common mechanistic base for the two seemingly different models.

Throughout this study, we have focused on a single transition during the cell cycle progression (quiescence to proliferation) due to its experimental and computational tractability. To further simplify analysis, we have chosen not to model cell division or growth explicitly. Instead, the variability associated with these processes is lumped into the extrinsic noise terms in our SDE model. More explicit mechanisms to account for such variability may further improve the quantitative agreement between the modeling and the experiment. For example, our simulation results suggest that E2F activation depends on both stochasticity in the chemical reactions and variability in the initial conditions. This is evident when E2F activation dynamics were compared under two conditions: varying initial conditions (Var) and varying amplitude of the extrinsic noise (ω) in the stochastic model. At a fixed amplitude value of the extrinsic noise, increasing variability in the initial conditions (modeled as a Gaussian distribution of the mean being the base initial conditions and varying level of standard deviation) is predicted to decrease transition rate and time delay (Figure 4.18A-B). Similarly, increasing ω without any variability in the initial conditions (Var=0) is predicted to decrease transition rate and time delay (Figure 4.18C-D). Therefore, the impact of the variable initial conditions and extrinsic noise on variable cell cycle entry timing is similar. Interestingly, our simulation results predict greater change in the temporal

117

dynamics in E2F activation by extrinsic noise than by variability in the initial conditions. These suggest that the major source of temporal variability in cell cycle entry is extrinsic noise, but variability in the initial conditions also can lead to discernable change in the temporal dynamics of E2F activation. Such dependence on initial conditions may explain reduced time delay in actively growing cells compared to that in quiescent cells [229].
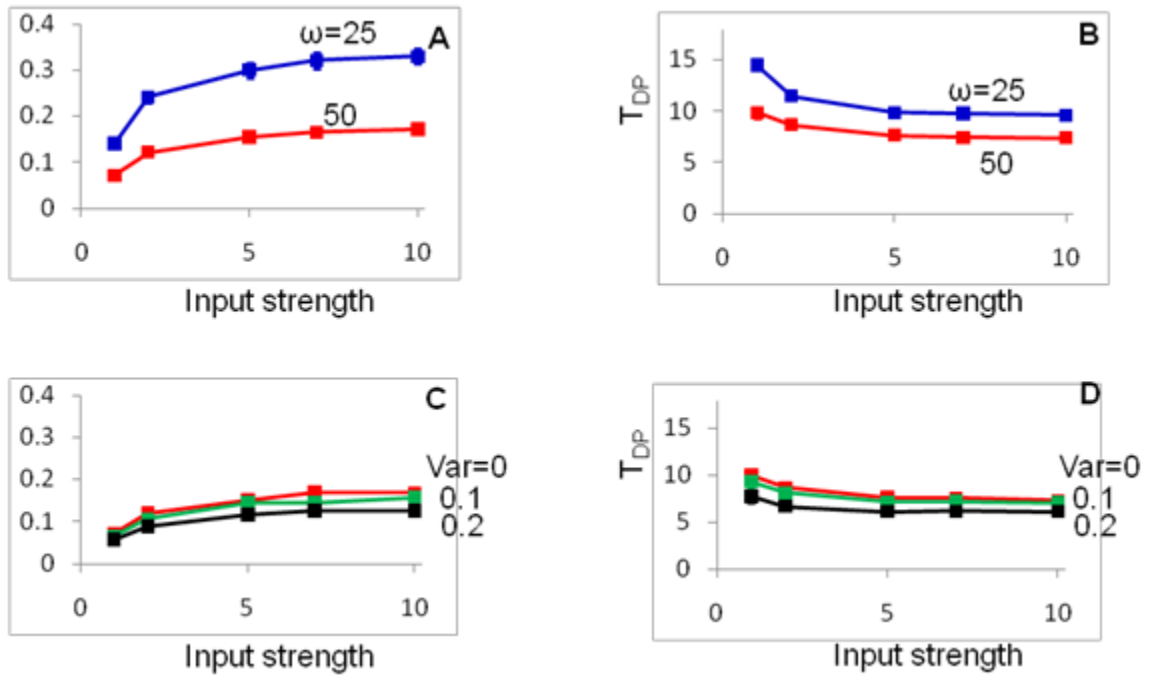
Figure 4.18: Variability in the initial conditions vs. in the rates of the chemical reactions. The effects of variability in the initial conditions and in the rates of the chemical reactions were evaluated on the temporal dynamics of E2F activation. With all else the same, our simulation results predicted that transition rate (A) and time delay (B) would decrease significantly with increasing ω. To describe variability in the initial conditions, we assumed that the initial concentrations were Gaussian-distributed with the mean being their base value and varying variance values. At a fixed extrinsic noise amplitude, our simulation results predicted that transition rate (C) and time delay (D) would decrease slightly with increasing variance of the initial conditions. Overall, the activation dynamics of E2F is much more sensitive to changes in extrinsic variability than those in the initial condition.

During the mapping from our stochastic model to the TP or GC models, details associated with individual signaling reactions are necessarily lost in the resulting TP or GC models, pointing to their limitations in offering mechanistic insights. However, this mapping suggests a potential, unappreciated utility of the TP or GC models. On one hand, these phenomenological models are simple and are able to provide quantitative description of the population-level dynamics associated with variable cell cycle entry. On the other, specific changes in the underlying reaction networks can be manifested in changes in the parameters in these simple models. As such, together with a mechanistically based model, the TP and GC models can serve as a concise platform to define quantitative phenotypes that facilitate classification of cell types or cell states.

This utility may be particularly useful for cancer diagnostic, since most cancers have defects in the Myc-Rb-E2F signaling pathway [173,245]. Recent approaches for cancer classification involve microarray-based gene expression profiling to develop cancer signatures [246], which have been used to reveal activation status of oncogenic signaling pathways [247]. Here we suggest that oncogenic phenotypes resulting from deregulation in these pathways may also serve as cancer signatures. Using the mapping technique defined in this work, a correlation may be established between the temporal dynamics of a cancer cell under various conditions and the library of predicted phenotypes based on the Myc-Rb-E2F network. This correlation may be used to reveal activation status of the Myc-Rb-E2F network. With sufficient number of testing

120

conditions, such correlation may be used to identify the underlying mechanism for the

observed patterns (i.e. mutations in this network).

Chapter 5: Sensing and integration of Erk and PI3K signals by Myc

In Chapter 5, we study post-translation regulation of Myc , a important transcriptional factor in multiple cell-fate decisions, including proliferation, growth and apoptosis. To maintain a normal cell physiology, it is critical that the control of Myc dynamics is precisely orchestrated. Recent studies suggest that such control of Myc can be achieved at the post-translational level via protein stability modulation. Myc is regulated by two Ras effector pathways: the extracellular signal-regulated kinase (Erk) and phosphatidylinositol 3-kinase (PI3K) pathways. To gain quantitative insight into Myc dynamics, we have developed a mathematical model to analyze post-translational regulation of Myc via sequential phosphorylation by Erk and PI3K. Our results suggest that Myc integrates Erk and PI3K signals to result in various cellular responses by differential stability control of Myc protein isoforms. Such signal integration confers a flexible dynamic range for the system output, governed by stability change. In addition, signal integration may require saturation of the input signals, leading to sensitive signal integration to the temporal features of the input signals, insensitive response to their amplitudes, and resistance to input fluctuations. We further propose that these characteristics of the protein stability control module in Myc may be commonly utilized in various cell types and classes of proteins. The work in this chapter appears in *PLoS Computational Biology*, 2008 Feb; 4(2): e1000013 [53].

5.1 Introduction

The proto-oncogene Myc is a transcription factor that regulated numerous signaling pathways involved in cell-fate decisions [248,249,250,251,252]. In fact, Myc is known to coordinate two coupled but distinct processes [253] (Figure 5.1): cell proliferation, a discrete process initiated by DNA synthesis as outlined in Chapter 3 and cell growth, a continuous process characterized by accumulation of mass. Disengagement of this coupling by disrupting Myc function or by interfering with either pathway will cause abnormal phenotype in cell cycle process [254,255]. In addition, excessive accumulation of Myc can induce apoptosis [256,257] when cell are under stress or deprived of growth factors. In this chapter, we discuss how Myc is tightly controlled in response to different environmental cues, which may provide insights into coordination of multiple cell-fate decisions.
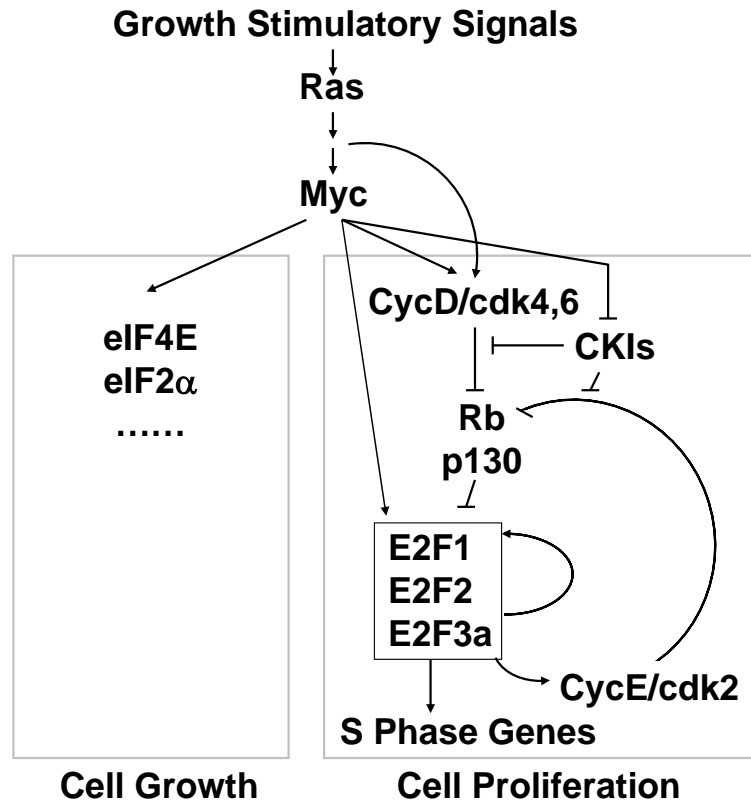
Figure 5.1: Myc's role in coordination of cell growth and proliferation by relaying signals from Ras (adapted from [147]).

5.2 Background:

5.2.1 Modulation of the Myc-Rb-E2F signaling pathway by Ras

Dynamics of the Myc-Rb-E2F pathway is known to be influenced by many signaling components, one of which is a proto-oncogene Ras. Activation of Ras signaling pathways is essential for cells to leave a quiescent state and to progress through G1 phase of the cell cycle. Based on experiments in cells expressing wild-type or mutant Rb, a primary role for Ras in G1 progression is to inactivate Rb through the activation of G1

Cdks [258,259]. This has been shown to occur through the stimulation of Cyclin D transcription as well as increases in the Cyclin D/Cdk4 kinase activity [260,261]. Three Ras effector pathways, the Raf-MEK-ERK cascade, PI3-K signaling, and Ral activation are all involved in stimulating Cyclin D gene transcription, with maximal stimulation requiring the co-operative action of several pathways [262]. In addition, PI3-K/AKT signaling, via inhibition of glycogen synthase kinase (GSK-3), increases the stability of Cyclin D [263].

Ras activation has been shown to down-regulate the cyclin-dependent kinase inhibitor p27kip1, resulting in the activation of Cyclin E/Cdk2 [264]. The down-regulation of p27 involves both ERK and PI3-K effector signaling pathways, and it is associated with a decrease in the rate of p27 translation, stability, and association with Cyclin E/Cdk2. This effect is essential for Ras-mediated entry into S phase [265]. Ras, via Raf, has also been reported to activate the Cdc25A phosphatase that removes inhibitory phosphates from Cdk2 and Cdk4 contributing to their activation [266]. Adding to the complexity, Ras also stimulates transcription of the cyclin-dependant kinase inhibitor p21 and p16INK4a, which may underlie the ability of Ras to induce cellular senescence [267].

5.2.2 Myc modulation by Ras via the Erk and PI3K signaling pathways

Ras activation is essential in collaborating with Myc to mediate serum stimulated

cell proliferation, as well as oncogenesis. On one hand, Ras signaling stabilizes the Myc

protein and increases the accumulation of functional Myc transcriptional factor in the

cell [200,201,202]. It has been shown that serum-induced increase in Myc protein half-life

depends on Ras activation. On the other hand, the activation of PI3-K/AKT signaling in

Ras pathway provides a critical survival signal, which counteracts the Myc-induced

apoptosis effect that prevents the outgrowth of a cell population [268]. Taken together,

these data suggest a central role of Ras in modulating the dynamics of Myc-Rb-E2F

pathway.

Interestingly, recent discoveries indicate that Myc is dynamically regulated at

the protein level by the Ras effector pathways [200,201,202,269]. These discoveries

suggest that Myc protein undergoes a series of modifications that are sequential and

irreversible [147,200,201,202,270]. More specifically, when Myc is newly synthesized, it is

highly unstable and quickly undergoes ubiquitination and degradation [271]. It can be

substantially stabilized when phosphorylated at serine 62 (Ser62) by Ras-activated Erk

activity (Figure 5.2). Subsequent phosphorylation of Myc at threonine 58 (Thr58) by

Gsk3β, however, initiates a destabilization process in a sequential manner. This is

achieved by a dephosphorylation mechanism by a prolyl isomerase Pin1 and a protein

phosphatase PP2A. Once Myc is phosphorylated at Thr58 (Myc[Ser62-Thr58]), Pin1 induces it

to undergo conformation changes, which are required for PP2A to dephosphorylate the

Ser62 residue (Myc[Thr58]) [200]. To date, this is the only dephosphorylation mechanism

identified in the Myc stabilization processes. Destabilization of Myc by Gsk3β can be

blocked by the Ras-activated PI3K pathway (Figure 5.2).
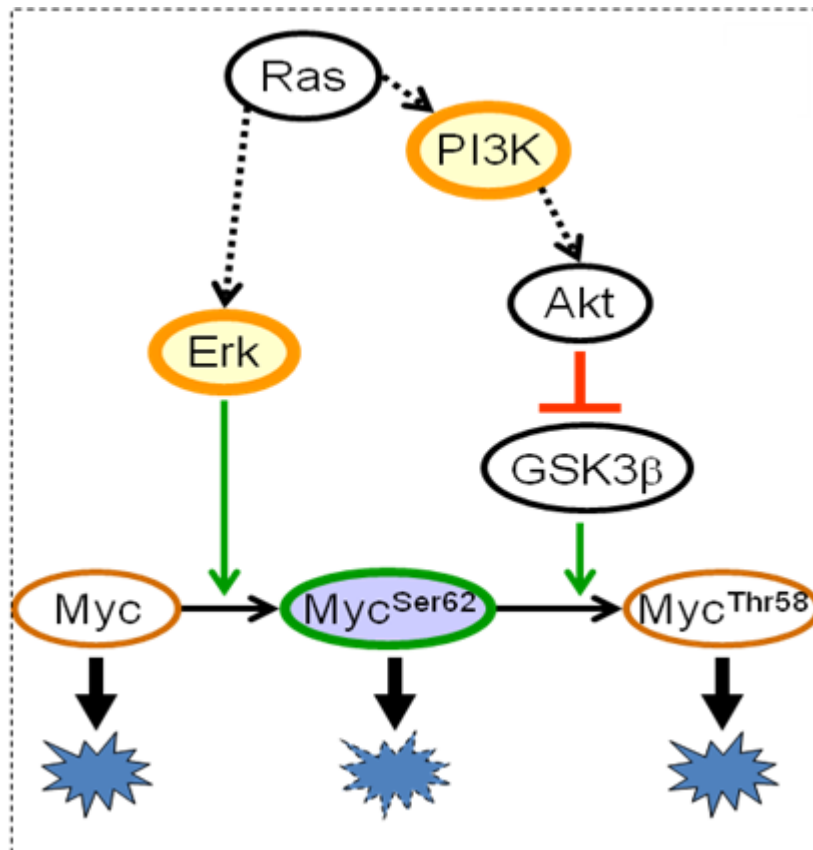
Figure 5.2: Post-translational of Myc by the two arms of Ras signaling. Stimulation with growth factors (GF) leads to activation of Ras and Myc synthesis. Active Ras induces activation of its downstream effector pathways: the MAPK and PI3K pathways. While the synthesized Myc is unstable with short half-life, its stability can be significantly increased via the Ras effector pathways. Active Ras induces Erk that stabilizes Myc by phosphorylation at Ser62. PI3K activation blocks Myc degradation by inhibiting phosphorylation at Thr58 by Gsk3β. As Ras activity declines, Gsk3β initiates phosphorylation of Myc at Thr58 and triggers degradation. Phosphorylation at Thr58 requires prior phosphorylation at Ser62, and phosphorylation at Thr58 induces dephosphorylation at Ser62.

The unique control of Myc dynamics by sequential phosphorylation allows Myc to integrate upstream signals from Erk and PI3K, which play critical roles in controlling diverse cell fates [200,272,273]. Erk often exhibits an early, transient peak of activation upon growth stimulation (Table 5.1). The peak is followed by varying residual activities, which depend on cell lines and growth factors. This residual Erk is critical in downstream signal encoding. For example, in PC12 cells, a small residual Erk activity, as a result of epidermal growth factor (EGF) stimulation, leads to proliferation. In contrast, a high residual Erk activity as a result of nerve growth factor (NGF) stimulation in the same cell line leads to differentiation [13,274] The residual Erk level has also been observed to be critical in regulating c-Fos level in fibroblasts [275].

Table 5.1: Erk signal pattern

(Transient strong activation followed by varying residual levels**)**

Cell Line	Growth Factor	Cell outcome	References
PC12	EGF	Proliferation	[13,274]
PC12	NGF	Differentiation	[13,274,276]
CCL39	Thrombin	Proliferation	[277]
Swiss 3T3	PDGF	Proliferation	[275]
Swiss 3T3	EGF	No proliferation	[275]

The PI3K activation pattern depends on cell lines and stimulants, as detailed in Table 5.2. It is bimodal (having two peaks) in various cell lines including WI38, NIH 3T3, or HepG2 when stimulated by platelet-derived growth factors (PDGF) or fetal bovine

serum (FBS) [270,278,279]. In contrast, PI3K appears to have only an early, transient

single peak in the U-2OS or PVSM cell lines stimulated with other growth stimulants

[278,280]. The bimodal activation of PI3K has been shown to be important for cell cycle

regulation [279,281,282]. In particular, the second peak has been found sufficient and

critical to drive the G1/S transition during cell cycle [270,281].

Table 5.2: PI3K signal pattern

	Cell line	Growth Factor	Cell outcome	Inter-peak delay (hrs)	References
Single peak	PVSM	PDGF	Proliferation or migration	N/A	[280]*
	PVSM	TGFα	Proliferation	N/A	[280]*
	PVSM	Thrombin	Proliferation	N/A	[280]*
	U-2OS	FCS	Proliferation	N/A	[278]
Double peaks	NIH 3T3	FBS	Proliferation	8	[270]
	HepG2	PDGF	Proliferation	5	[279]
	WI38	PDGF	Proliferation	3	[278]

* Experimental observations were made only for the first three hours after stimulation. It is possible that cells exhibit a second round of PI3K activity when measured for a longer time under these conditions.

The temporal pattern of Myc activation closely correlates with those of Erk and

PI3K (Table 5.3). Myc protein reaches its peak at ~2 hours after growth stimulation and

decreases to and remains at an intermediate value, or hump, for over ~6 hours before

reducing to its basal level [270,283]. The peak and the hump of Myc coincide with the

Erk peak (also the 1st PI3K pulse) and the 2nd PI3K pulse, respectively. These

observations suggest that Myc may sense and integrate signals from its two regulators (Erk and PI3K).

Table 5.3: Myc signal pattern

	Cell Line	Growth factor	Cell outcome	Inter-peak delay (hrs)	References
Two peaks	NIH 3T3	FBS	Proliferation	8	[270]
Overlapping peaks	Primary Human Fibroblast	FBS	Proliferation	N/A	[283]

5.3 Mathematical Analysis of post-translational regulation of Myc by Erk and PI3K

To gain insight into this control mechanism, we have constructed a mathematical model to analyze dynamics of Myc accumulation, controlled by sequential phosphorylation. Using this model, we aimed to investigate how signaling patterns of Erk and PI3K regulate Myc dynamics at the post-translational level. Also, how robust is Myc dynamics with respect to network parameters, such as phosphorylation and dephosphorylation rate constants? What is unique about this strategy of controlling Myc accumulation by sequentially modulating protein stability? Is this a common strategy by which cells achieve reliable temporal control of key regulatory proteins? By exploring these questions, we provided insights into design features of cell signaling networks and guidance for experimental intervention. Conceptually, our model defines a unique

module that connects with other models that deal with upstream signaling dynamics leading to the activation of Erk [13] or PI3K [284,285], as well as downstream dynamics leading to mammalian cell fate decisions [175,177,286], including cell proliferation governed by the Myc-Rb-E2F signaling pathway. We further propose that post-translation regulation of Myc represents an example of a generic dual-kinase motif. With appropriate parameters, this motif will enable precise temporal sensing of input signals.

5.3.1 Construction of the mathematical model: the post-translational control of Myc accumulation

Based on the reaction network outlined in Figure 5.3, we developed a kinetic mathematical model in *Dynetica*, a graphics-based, integrated simulation platform [76]. Most of reaction mechanisms and base model parameters were derived from experimental results in mouse or human cells. Others were carefully obtained or estimated from previous theoretical studies [14,209,287]. We simplified the model by lumping the sequential destabilization processes together into the phosphorylation rate constant for Myc^{Thr58}. This simplification was based on the observation that Pin1 and PP2A activation is not rate-limiting during cell cycle entry [288,289]. Explicitly accounting for the sequential events by Pin1 and PP2A did not have significant impact on Myc accumulation. In addition, we assumed that the change in transactivation capacity due to stability control [200] is not significant, since Myc^{Ser62} is much more predominant than Myc^{Thr58}. Based on experimental data, we assumed phosphorylation of

Myc at Ser62 or Thr58 to be much more significant than dephosphorylation. This results

in sequential, irreversible Myc stabilization. However, we accounted specifically for the

differential degradation dynamics of protein isoforms in this model.
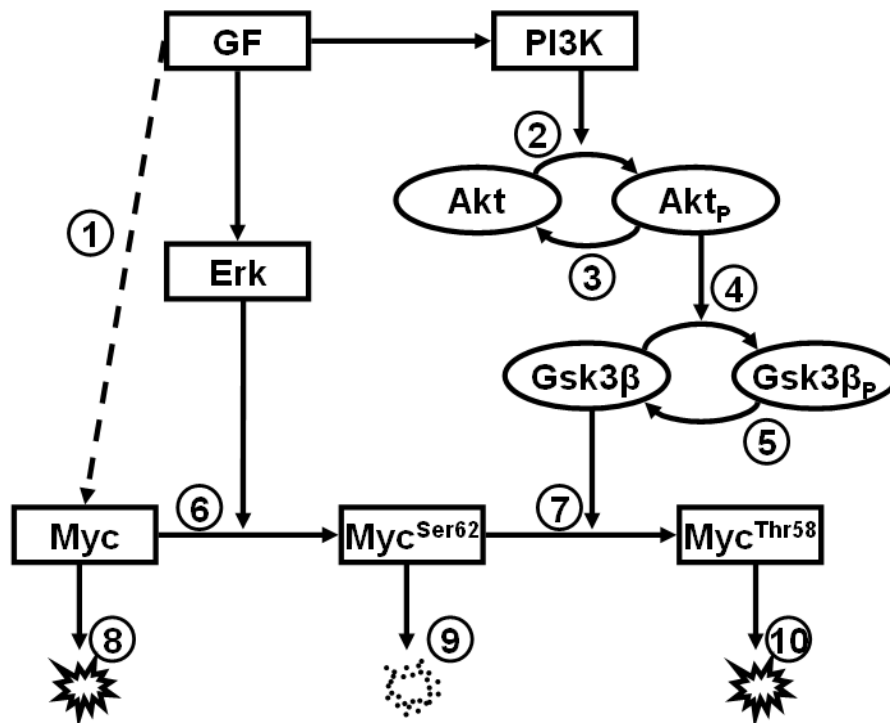
Figure 5.3: Detailed reaction diagram for Myc protein stabilization.

To establish a framework that facilitates investigation of Myc modulation by its

upstream signals, Erk and PI3K, we built the model with Erk and PI3K as the inputs and

Myc as the output. Despite the extensive interactions between the MAPK and PI3K

pathways, we decoupled Erk and PI3K signals and simplified them as a single or double

rectangular pulses, respectively. Such decoupling of these signals was driven by the

objective of our study: to characterize Myc's response to various input signal patterns.

Since activation of Erk and PI3K is specific to cell lines and stimulants (as shown in

Table 5.1and Table 5.2), and is mediated by multiple signaling pathways including Ras,

Rac, or Rap [290,291,292,293], it is not clear to what extent these signaling pathways

contribute to Erk and PI3K activation patterns. Furthermore, the Erk and PI3K pathways

that control Myc protein turnover are a conserved motif found in both mammalian and

yeast systems and such control motif has been speculated in many other protein

stabilization processes [269]. In many of these processes, input signals are not triggered

by a single protein. To set up a framework for a more general Myc stabilization process,

we decoupled the two signals from each other and from their upstream network

regulation and assumed their effects on Erk and PI3K in the decoupled inputs. This

allowed analyzing Myc's response to variations in Erk and PI3K independently.

Based on experimental observations, we approximated input signals as

rectangular pulses with three parameters: duration, the maximum level, and the residual

level. To describe two-peak PI3K activation, we introduced another parameter, inter-

peak delay. More sophisticated representations (for example, sinusoidal pulses) give

similar results (data not shown). Although we focused on the two-peak activation of

PI3K in the base model, the modeling framework can be extended to study other

patterns of PI3K signals (such as a single peak pattern) by varying duration, steady-state

values, or inter-peak delay.

Table 5.4: Reaction Kinetics

Index	Reaction	Kinetics	Description
1	$* \xrightarrow{GF} Myc$	$k_M[GF]$	Synthesis of Myc
2	$Akt \xrightarrow{PI3K} Akt_p$	$\dfrac{k_{AP}[PI3K][Akt]}{K_{AP}+[Akt]}$	Akt phosphorylation
3	$Akt_p \longrightarrow Akt$	$\dfrac{k_{AD}[Akt_p]}{K_{AD}+[Akt_p]}$	Akt dephosphorylation
4	$Gsk3\beta \xrightarrow{Akt_p} Gsk3\beta_p$	$\dfrac{k_{GP}[Akt_p][Gsk3\beta]}{K_{GP}+[Gsk3\beta]}$	Gsk3β phosphorylation
5	$Gsk3\beta_p \longrightarrow Gsk3\beta$	$\dfrac{k_{GD}[Gsk3\beta_p]}{K_{GD}+[Gsk3\beta_p]}$	Gsk3β dephosphorylation
6	$Myc \xrightarrow{Erk} Myc^{Ser62}$	$\dfrac{k_{MS}[Erk][Myc]}{K_{MS}+[Myc]}$	Myc phosphorylation at Ser62
7	$Myc^{Ser62} \xrightarrow{Gsk3\beta} Myc^{Thr58}$	$\dfrac{k_{MT}[Gsk3\beta][Myc^{Ser62}]}{K_{MT}+[Myc^{Ser62}]}$	Myc phosphorylation at Thr58
8	$Myc \longrightarrow *$	$d_M[Myc]$	Unstable Myc Degradation
9	$Myc^{Ser62} \longrightarrow *$	$d_{MS}[Myc^{Ser62}]$	MycSer62 degradation
10	$Myc^{Thr58} \longrightarrow *$	$d_{MT}[Myc^{Thr58}]$	MycThr58 degradation

Table 5.5: Base model parameters and notes

Rate Constants	Values	Source
k_M	1 /hr	Constrained by [283]
$Erk_{Max}{}^*$	1 /hr	Constrained by [13]
$PI3K_{Max}{}^{**}$	1 /hr	Constrained by [279]
k_{AP}	360 / hr	Typical value phosphorylation rate constant is 3600/hr [209]
K_{AP}	0.01 μM	Typical value for Michaelis-Menten (MM) parameter for phosphorylation is 0.01 μM [209]
k_{AD}	72 μM / hr	Typical value dephosphorylation rate constant assuming a constant phosphatase concentration is 720 μM/hr [209]
K_{AD}	0.01 μM	Typical value for MM parameter for dephosphorylation [209]
k_{GP}	360 / hr	Typical value phosphorylation rate constant is 3600/hr [209]
K_{GP}	0.01 μM	Typical value for MM parameter for phosphorylation [209]
k_{GD}	72 μM / hr	Typical value dephosphorylation rate constant assuming a constant phosphatase concentration is 720 μM/hr [209]
K_{GD}	0.01 μM	Typical value for MM parameter for dephosphorylation [209]
$k_{MS}{}^*$	2.3 /hr	These values are adjusted together so that
K_{MS}	0.01 μM / hr	(1) The peak level is at least 3 times higher than the basal level.
k_{MT}	0.4 /hr	
K_{MT}	0.01 μM / hr	(2) The level of the second hump is about 60 % of the peak level [283]
d_M	2.08 / hr	[202,294]
d_{MS}	0.35 / hr	Myc's half-life is increased by 6fold upon phosphorylation at Ser62 [202,294]
d_{MT}	2.08 / hr	Assumed to be the same as d_M

Table 5.5: Continued

Initial Akt = 0.6 μM and initial Gsk3β = 0.6 μM.

Erk_R: Residual Erk level. Adjusted to 10 percent of the maximal Erk level. To reflect experimental observations in our model, residual Erk level is required. Without Erk_R, our model cannot generate the second peak of Myc activity.

$PI3K_R$: Residual PI3K level. Adjusted to 10 percent of the maximal PI3K level.

* Erk_{Max} and k_{MS} together represent Erk's phosphorylation efficiency of Myc at Ser62. If $k_{MS}[Erk_{Max}]$ is sufficiently large, our model can generate the initial peak of Myc activity as observed in [283]. Further increase in $k_{MS}[Erk_{Max}]$ does not significantly change the initial peak of Myc activity, suggesting that Erk's phosphorylation efficiency is operating at saturation.

** Prevention of Myc destabilization by Gsk3β depends on PI3K's ability to inhibit Gsk3β activity. A sufficiently strong PI3K pulse is needed to prevent Myc from destabilization and to generate the second round of the Myc accumulation.

Due to lack of detailed quantitative analysis of the PI3K pathway, many rate parameters for our modeled reactions are unavailable. To address this limitation, we have adopted base parameter values that fall into the range of values typically used for other phosphorylation or dephosphorylation reactions (Table 5.5). In adopting these parameter values for our system, we made an implicit assumption that different types of phosphorylation-dephosphorylation (Ph-dePh) cycles share similarity, not only in terms of the mechanism, but also in terms of parameter values. It is possible that the adopted base parameter values may differ significantly from the true values. To address this caveat, we carried out extensive sensitivity analysis to test the impact of the uncertainty in these parameters. Our results suggest that overall, our conclusions will hold despite significant changes in these parameters (up to 100fold change for each parameter, while holding the others constant). Further tests of the modeling predictions against experiments will help constraining these parameters.

We have used Michaelis-Menten kinetics to model phosphorylation or dephosphorylation reactions, which is a well-established practice in modeling signaling pathways [209,287,295,296]. Under certain circumstances, the Michaelis-Menten kinetics may deviate from true dynamics. Overall, our model predictions on the response of Myc to varying input signals (Erk and PI3K) have been qualitatively consistent with existing experimental observations. Due to the lack of quantification of signaling dynamics leading to Myc activity (especially at the individual reaction level), it is unclear how, at

138

the quantitative level, the Michaelis-Menten kinetics may faithfully reflect the underlying kinetics. If our model's predictions deviate significantly from more quantitative experimental observations, it may be necessary to formulate the model using alternative methods, such as the "total quasi-steady state assumption" (tQSSA) [297].

5.3.2 Base simulation

The Myc temporal dynamics, simulated with reaction kinetics and base parameter values in Table 5.4 and Table 5.5, was overall consistent with experimental observations in Figure 5.4. [281,283]. To achieve this consistency, however, we found that the input signals Erk and PI3K needed to operate at or close to saturation, and there needed to be sufficient residual Erk (Erk_R) before the second PI3K pulse. In the base simulation (Figure 5.4), the total Myc (black line) consisted of unmodified, unstable Myc (blue line), stable Myc^{Ser62} (red line), and unstable Myc^{Thr58} (green line). Although phosphorylation state affects transactivation capacity [200], the contribution from Myc^{Ser62} to total Myc was much more significant than that from Myc^{Thr58}. Therefore, we assumed that the overall transactivation capacity of Myc does not change significantly during Myc modulation. The modification of Myc from its unstable to stable form, then back to unstable form closely followed Erk and PI3K signals. The first peak of Myc coincided with the Erk pulse and the first PI3K pulse. After these initial pulses,

139

unmodified Myc and Myc^{Ser62} recovered to new steady state levels, which depended on

the rate constants of Myc synthesis, phosphorylation, and degradation. Before Myc^{Ser62}

reached its new steady state, however, PI3K pulse became activated for the second time

and prevented Myc^{Ser62} from converting to the unstable form, sustaining total amount of

Myc at high level. Once the second PI3K pulse subsided, Myc^{Ser62} was turned to Myc^{Thr58}

via phosphorylation by Gsk3β. In other words, while Erk and the first peak of PI3K

determine initial Myc accumulation, the second peak of PI3K prevents Myc from

receding to a lower level, thus fine-tuning the Myc level.
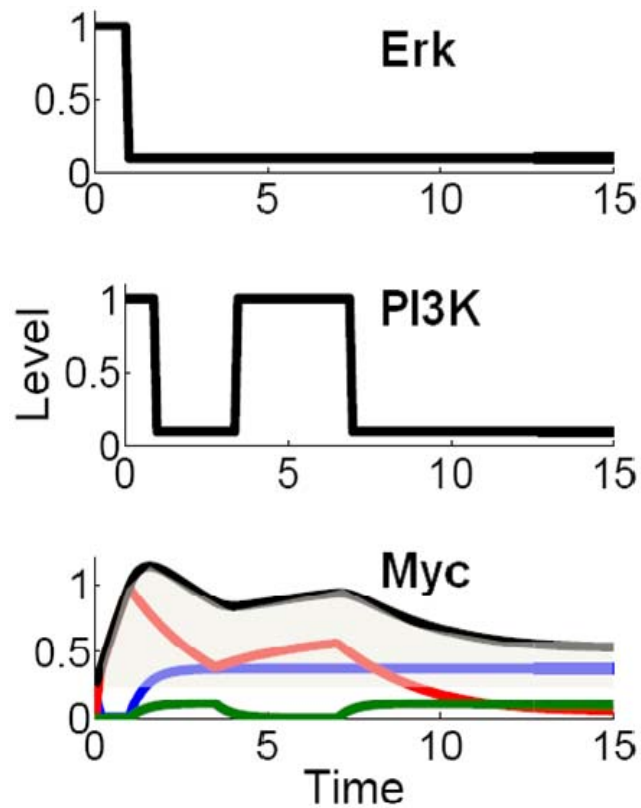
Figure 5.4: Myc protein stabilization. Activation patterns of Erk and PI3K determine Myc stability pattern. The three forms of Myc are plotted independently. The unmodified Myc (blue line) and Myc^{Thr58} (green line) accumulate only to a limited level, but stabilized Myc^{Ser62} level increases via phosphorylation (red line). The total Myc level is the sum of the three forms of Myc (black line) and its dynamics are highly correlated with input signals, Erk and PI3K. We define the shaded area under the Myc curve as 'potency', a measure of Myc accumulation.

5.3.3 Sensitivity Analysis

As Myc accumulation was determined by conversion between its unstable forms

and stable form, we expected Myc accumulation to depend on the degradation rate

constant of each form. As a quantitative estimate for Myc accumulation, we used Myc

potency [298], the shaded area in Figure 5.4. If Myc became stabilized quickly and

141

remained stabilized for an extended period of time, Myc potency would be high. In

contrast, slow stabilization and quick destabilization would yield small potency. To test

these notions, we carried out sensitivity of Myc potency to various parameters. Here,

sensitivity is the amount of change in Myc potency to a 100fold change in each

parameter. To be more specific, we calculated Myc potency for 10fold increase and

decrease in each parameter. Then we took the logarithmic ratio between the largest Myc

potency and the smallest Myc potency. Therefore, the positive values in Table 5.6

indicate logarithmic fold-increase in Myc potency over 100fold parameter change. The

negative values indicate logarithmic fold decrease. Sensitivity of '0' represents no

change, and '1' represents 10 fold changes in Myc potency.

5.3.4 Zero-order ultrasensitivity

Our sensitivity analyses in Table 5.6 indicated that Myc potency was highly

sensitive to parameters involved in stabilization of Myc and maintenance of the stable

form. In comparison, other parameters governing the signal transduction in the PI3K

pathway had little impact on Myc potency. In particular, many parameters had little

impact on Myc potency when the input Erk and PI3K signals were strong. For a 100fold

change in each of these parameters, the corresponding change in the potency was less

than 10fold. Many of these are involved in the PI3K signaling cascade. These include the

Michaelis-Menten (MM) constants for Akt and Gsk3β phosphorylation and

dephosphorylation.

Table 5.6: Parametric Sensitivity

Parameters	Sensitivity = $\log_{10}\left(\dfrac{Potency_{Increase}}{Potency_{Decrease}}\right)$
Myc synthesis rate constant (k_M)	4.57
Myc degradation rate constant (d_M)	-1.27
Myc^{Ser62} degradation rate constant (d_{MS})	-0.88
Rate constant for Myc phosphorylation at Ser62 (k_{MS})	0.88
Rate constant for Myc phosphorylation at Thr58 (k_{MT})	-0.38
Akt phosphorylation rate constant (k_{AP})	0.32
Akt dephosphorylation rate constant (k_{AD})	-0.32
Myc^{Thr58} degradation rate constant (d_{MT})	-0.31
Gsk3β dephosphorylation rate constant (k_{GD})	-0.14
Gsk3β phosphorylation rate constant (k_{GP})	0.14

* Sensitivity to other parameters was equal to or smaller than 0.1. These include 'MM constants for Myc phosphorylation at Ser62 (K_{MS}) and Thr58 (K_{MT}) , Gsk3β phosphorylation (K_{GP}) / dephosphorylation (K_{GD}), and Akt phosphorylation (K_{AP}) / dephosphorylation (K_{AD}).

The insensitive response of the Myc potency may result from the signaling

transduction in the PI3K pathway operating with zero-order ultrasensitivity [112]

around the base parameter setting, which may confer robustness to random

perturbation in a signaling cascade [299]. That is, each phosphorylation-

dephosphorylation cycle acts as a digital switch in response to changes in kinase and

phosphatase concentrations (Figure 5.5). This switching behavior occurs under the

following conditions: 1) the phosphorylation and dephosphorylation rates are close to each other, and 2) the protein concentration is much greater than the MM constants. As long as parameter changes did not cause loss of ultrasensitivity, the switching behavior of each stage was not significantly changed (Figure 5.5).

A) Phosphorylation Cycle

Kinase (Akt$_P$)

Gsk3β   Gsk3β$_P$

Phosphatase (PT)

C) Sensitivity Analysis

$K_{GD}=K_{GP}=0.1$

0.001

B) Mathematical Model

$$\frac{d[Gsk3\beta]}{dt} = \frac{k[PT][Gsk3\beta_P]}{K+[Gsk3\beta_P]} - \frac{k_{GP}[Akt_P][Gsk3\beta]}{K+[Gsk3\beta]}$$

$$\frac{d[Gsk3\beta_P]}{dt} = \frac{k_{GP}[Akt_P][Gsk3\beta]}{K+[Gsk3\beta]} - \frac{k[PT][Gsk3\beta_P]}{K+[Gsk3\beta_P]}$$

$$\frac{d[G]}{dt} = \beta\left(\frac{[G_P]}{K+[G_P]} - \frac{\gamma[G]}{K+[G]}\right)$$

$$\frac{d[G_P]}{dt} = \beta\left(\frac{\gamma[G]}{K+[G]} - \frac{[G_P]}{K+[G_P]}\right)$$

where $\gamma = \dfrac{k_{GP}[Akt_P]}{k[PT]}$, $\beta = k[PT]$,

$Gsk3\beta = G$, and $Gsk3\beta_P = G_P$

$K_{GD} = K_{GP} = K$
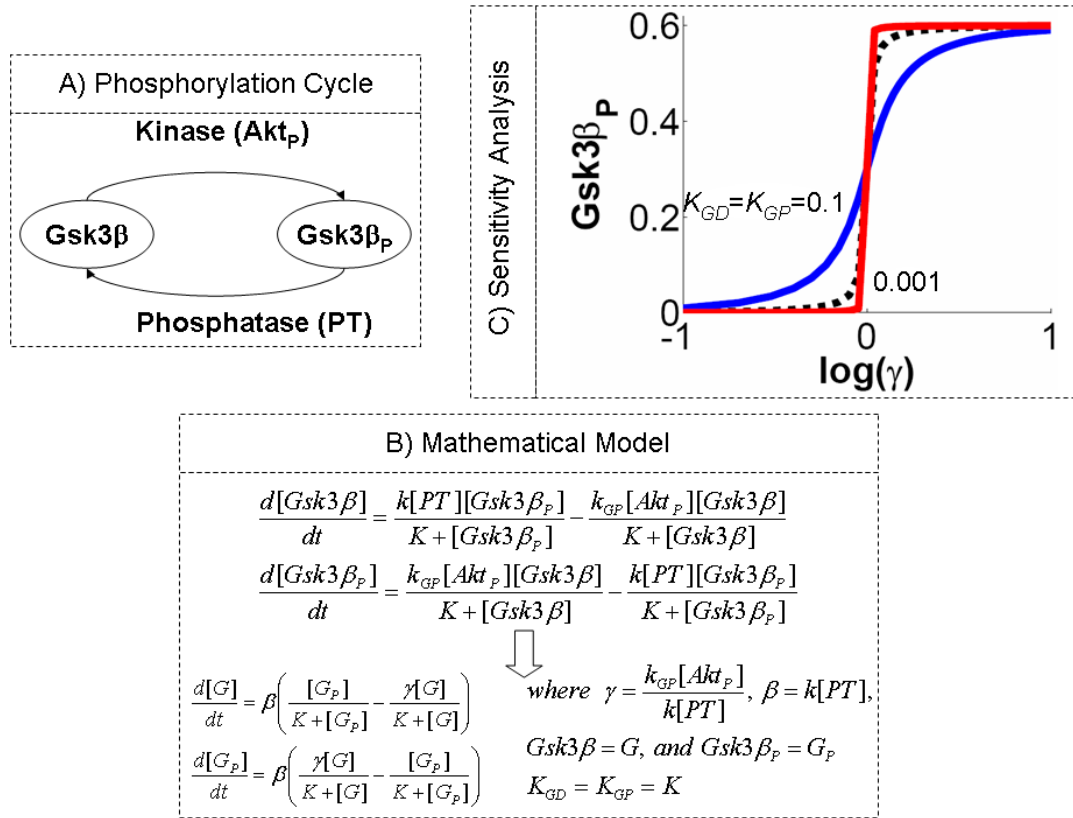
Figure 5.5: Modeling a phosphorylation-dephosphorylation cycle. An enzymatic modification cycle of Gsk3β between phosphorylated and dephosphorylated states (A) is mathematically modeled (B). k and k_{GP} are rate constants for phosphorylation and dephosphorylation, and K is the Michaelis-Menten constant. Protein conversion is ultrasensitive near $\gamma = 1$, for a sufficiently small Michaelis-Menten constant. The sensitivity becomes weaker as K is increased. Time-course simulation results at varying α values show the dependence of conversion on the rates of phosphorylation and dephosphorylation (C). Protein conversion becomes ultrasensitive near $\alpha=1$ for a sufficiently small Michaelis-Menten constant, while the sensitivity becomes weaker as K is increased.

145

To investigate ultrasensitivity in the PI3K pathway further, we evaluate

contribution of each phosphorylation-dephosphorylation (Ph-dePh) cycle of the cascade

towards overall ultrasensitivity. In particular, we characterize parameter dependence of

ultrasensitivity in each Ph-dePh cycle, and how this affects the cascade's overall

input/output response. We first start with our base model, where we assume

ultrasensitivity in the Akt ($K_{AD}=K_{AP}=K_A=0.01$) and Gsk3β ($K_{GD}=K_{GP}=K_G=0.01$) Ph-dePh

steps (blue line in Figure 5.6A,B). Consequently, the overall cascade response is

ultrasensitive. When the ultrasensitivity in Akt Ph-dePh step is lost ($K_A=1$, red line in

Figure 5.6A), the overall cascade response is not significantly affected as long as Gsk3β

Ph-dePh cycle ($K_G=0.01$) remains ultrasensitive (Figure 5.6B). When the ultrasensitivity

in Gsk3β Ph-dePh cycle is lost ($K_G=1$, Figure 5.6C), the overall ultrasensitive response can

be maintained if Akt Ph-dePh cycle is ultrasensitive (blue line in Figure 5.6C).

Otherwise, the overall system response loses ultrasensitivity (red line in Figure 5.6C).

Here we limit our analysis to two extreme cases where ultrasensitivity is present or

absent.
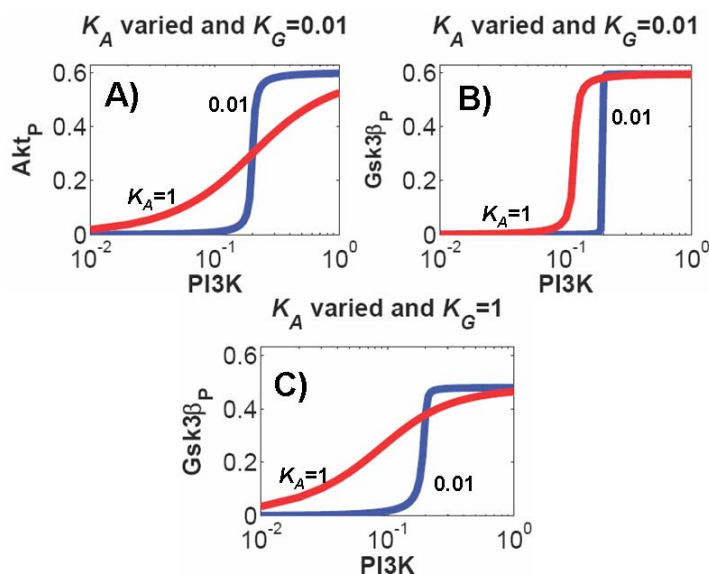
Figure 5.6: The overall ultrasensitivity arises from the input/output response in each level and across different levels down the cascade. (A) The Akt Ph-dePh cycle (in response to PI3K) can be either graded (red line) or ultrasensitive (blue line) depending on the Michaelis-Menten constants. (B) Both types of PI3K-Akt responses can lead to ultrasensitive PI3K-Gsk3β responses (both red and blue), if the Akt-Gsk3β response remains ultrasensitive. (C) If Akt-Gsk3β response is not ultrasensitive, the overall PI3K-Gsk3β remains ultrasensitive if PI3K-Akt response is ultrasensitive, but may lose ultrasensitivity if PI3K-Akt response is not ultrasensitive. Note that here we have assumed that the output from the first step (Akt_P) has an appropriate dynamic range that "matches" the input of the second step. The dependence of the overall sensitivity of the PI3K-Gsk3β response will likely be much more complex if this matching condition is not satisfied.

We note that overall input/output response is a convolution of the input/output responses between PI3K→Akt and Akt→Gsk3β. Ultrasensitivity may arise if the 'gating' between these responses matches. In this case, it has been shown that even without ultrasensitivity in each level of the cascade, ultrasensitivity may arise [300].

147

Although the ultrasensitivity in the PI3K pathway may facilitate noise-resistance, it is not absolutely required for insensitivity to parameter changes. The noise-resistant feature of the model also arises from the inputs operating at the saturation level. If the input level is sufficiently large, minor increase or decrease in the input level may not have much impact on the output response. This is clear in Table 5.7, where we repeat sensitivity analysis assuming absence of ultrasensitivity in Akt (K_A=1) and Gsk3β (K_G=1) Ph-dePh cycles. The rate constants for Myc phosphorylation at Ser62 (k_{MS}=3.4/hr) and Thr58 (k_{MT}=1.08/hr) are adjusted to match Myc's steady-state level with our base model. Although a minor increase is observed in some parameters (k_{MT}, d_{MT}, k_{GP}, and k_{GD}) the results in Table 5.7 do not significantly deviate from those in Table 5.6 as long as the inputs are sufficiently strong. For example, Myc potency is sensitive to parameters directly involved in Myc protein modification or degradation. However, many parameters involved in the PI3K signaling cascade have little impact on Myc potency. In contrast, when the input signals are weak, the sensitivity is overall increased. Ultrasensitivity with increasing effective Hill coefficient may 'shrink' the sensitive region in input/output response curve, allowing wider range of input values for noise-resistance. However, Michaelis-Menten response is sufficient to confer resistance to minor fluctuations when the input value is large. This aspect is also evident in the analysis of the simplified dual-kinase module, which assumes no cooperativity in phosphorylation of Myc isoforms.

Table 5.7: Parametric sensitivity without ultrasensitivity*

Parameters	Sensitivity = $\log_{10}\left(\dfrac{Potency_{Increase}}{Potency_{Decrease}}\right)$
Myc synthesis rate constant (k_M)	4.6
Myc degradation rate constant (d_M)	-1.3
Rate constant for Myc phosphorylation at Ser62 (k_{MS})	0.81
Myc^{Ser62} degradation rate constant (d_{MS})	-0.79
Rate constant for Myc phosphorylation at Thr58 (k_{MT})	-0.64
Myc^{Thr58} degradation rate constant (d_{MT})	-0.61
Gsk3β phosphorylation rate constant (k_{GP})	0.56
Gsk3β dephosphorylation rate constant (k_{GD})	-0.56
Akt dephosphorylation rate constant (k_{AD})	-0.41
Akt phosphorylation rate constant (k_{AP})	0.41
MM constant for Myc^{Ser62} phosphorylation (K_{MS})	-0.12
MM constant for Myc^{Thr58} phosphorylation (K_{MT})	0.10

*Sensitivity to other parameters was equal to or smaller than 0.1. These include Michaelis-Menten constants for Myc^{Thr58} phosphorylation (K_{MT}), Gsk3β phosphorylation (K_{GP}) / dephosphorylation (K_{GD}), and Akt phosphorylation (K_{AP}) / dephosphorylation (K_{AD}).

5.3.5 Effects of Erk and PI3K signal patterns on Myc accumulation

Erk and PI3K activation patterns, which determine the temporal dynamics of Myc, may vary significantly under different growth conditions and in different cell lines. Here we investigated how Myc potency responds to varying patterns of Erk and PI3K signals. Whenever possible, model predictions were compared with existing experimental observations. When the latter are unavailable, our model predictions may serve as testable hypothesis for future experiments, which in turn can further constrain our model. As shown in Figure 5.7, we quantitatively represented input signals of Erk

and PI3K with the following parameters: duration (Dur_E and Dur_P), maximal amplitude

(Erk_{Max} and $PI3K_{Max}$), and residual level (Erk_R and $PI3K_R$). For PI3K, we used an

additional parameter to describe the time interval between the two peaks (IP_P).

A

Dur$_E$

Erk$_{Max}$

Erk$_R$

B

IP$_P$

Dur$_P$

PI3K$_{Max}$

PI3K$_R$

Time
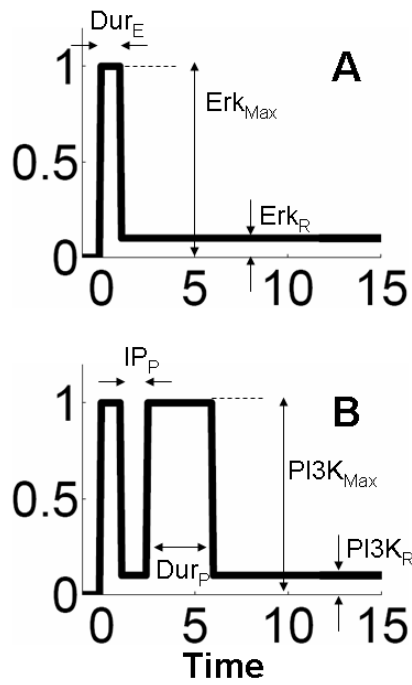
Varying Erk signal pattern

C

D

E

Myc
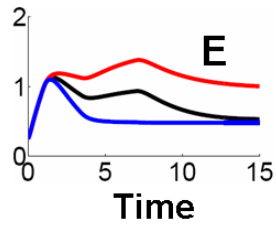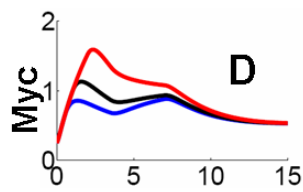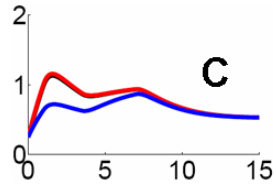
Varying PI3K signal pattern

F

G

H

Myc

Time
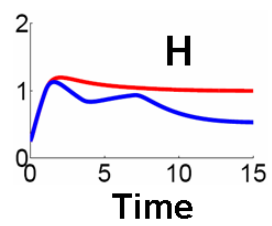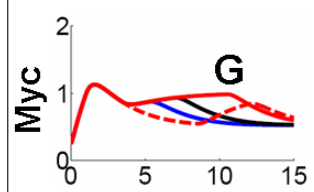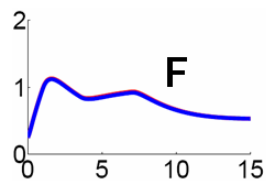
Figure 5.7: Erk and PI3K signal patterns determine Myc temporal behaviors. For all analyses, black lines represent the base case. A) The Erk signal was represented with the following parameters: duration (Dur_E), maximal Erk amplitude (Erk_{Max}), and residual Erk level (Erk_R). B) The PI3K signal was represented with the following parameters: duration (Dur_P), maximal PI3K amplitude ($PI3K_{Max}$), residual PI3K level ($PI3K_R$), and the time interval between the two peaks of PI3K (IP_P). The first peak of the PI3K was not considered, since its variations did not have a big impact. C) Myc accumulation was insensitive to Erk_{Max}. Fivefold increase in Erk_{Max} resulted in little change in Myc (red line) in comparison to the base case (black line), whereas fivefold decrease in Erk_{Max} resulted in light reduction in the main peak of Myc (blue line). D) Doubling (red line) or halving (blue line) Dur_E leads to significant change in the initial peak of Myc accumulation. E) Myc was sensitive to Erk_R. The base value of Erk_R was 10 percent of Erk_{Max} (black line). A small increase in Erk_R (20% of Erk_{Max}) resulted in excessive Myc accumulation (red line). When Erk was completely removed ($Erk_R=0$), Myc responded only to the initial, transient Erk pulse and became unresponsive to the PI3K signal (blue line). F) Myc accumulation was insensitive to $PI3K_{Max}$. Fivefold increase (red line) or decrease (blue line) in $PI3K_{Max}$ resulted in little change in Myc accumulation. G) The 2nd PI3K peak determined generation and maintenance of Myc hump. Doubling (red line) or halving (blue line) the duration of the second PI3K peak led to approximately twofold change in the Myc hump duration. Increasing IP_P from 3 hours to 8 hours delayed the timing of the second rise in Myc accumulation (red dotted line). H) A slight increase (20% of $PI3K_{Max}$) in $PI3K_R$ from the base value (10% of $PI3K_{Max}$) resulted in excessive Myc accumulation (red line). However, complete removal of $PI3K_R$ did not change Myc accumulation significantly (blue line overlapping with black line).

Our analysis predicted Myc accumulation to be insensitive to further increase in Erk amplitude. A fivefold increase in Erk_{Max} caused little change in Myc accumulation (Figure 5.7C). A fivefold decrease in Erk_{Max}, however, predicted a slight but discernable decrease in Myc accumulation. These results indicated that the base case of Erk was operating at saturation. As a result, this behavior enabled the system to be insensitive to minor changes in Erk amplitude, unless the Erk amplitude became sufficiently small. In comparison, the Myc potency was much more sensitive to the duration of Erk pulse: excessive accumulation of Myc was also observed when the duration of Erk was doubled (red line in Figure 5.7D). Halving Erk duration resulted in significant reduction in the initial peak of Myc.

Myc potency was sensitive to the residual Erk level (Erk_R). Without it ($Erk_R=0$), the total Myc level quickly reduced to a low level following the Erk pulse (blue line in Figure 5.7E). Conversely, a mere twofold increase in Erk_R from the base value (=10% of Erk_{Max}) led to excessive Myc accumulation (red line in Figure 5.7E). These results highlighted the importance of Erk_R in fine-tuning total Myc accumulation. In particular, Erk_R was important for maintaining sufficient Myc level before the arrival of the second PI3K pulse, by providing a moderate rate of Myc stabilization. In a more extreme case where the Erk signal was completely removed, no Myc accumulation was observed (data not shown). These results may provide a mechanistic explanation for differential phenotypic responses to varying residual level of Erk [13]. In PC12 cells proliferation

153

was correlated with low residual level of Erk, while high residual level of Erk was observed for differentiation. Based on our simulations, we suggest that differential regulation of Myc accumulation may be involved in determining these diverging phenotypic behaviors of these cells. This prediction can be tested by further experiments.

Similarly, Myc accumulation was insensitive to the maximum amplitude of PI3K ($PI3K_{Max}$), but much more sensitive to its residual level ($PI3K_R$) and temporal features, including duration of the 2nd peak (Dur_P) and time interval between the two peaks (IP_P). Five-fold increase or decrease in $PI3K_{Max}$ resulted in little change in Myc accumulation (Figure 5.7F). However, doubling or halving the duration of the 2nd PI3K peak caused an approximately two fold change in the duration of the Myc hump (G). Complete removal of the 2nd PI3K peak eliminated the Myc hump (Figure 5.8A). This indicates that the 2nd PI3K peak was primarily responsible for generating and maintaining the hump in Myc activation. These results are consistent with recent experimental data: removal of the 2nd PI3K peak by using a PI3K inhibitor [270] or by acid washing [281] drastically reduced total Myc accumulation. Given this role of the 2nd PI3K peak, the time interval between the two peaks of PI3K was critical for determining Myc accumulation pattern (dotted red line in Figure 5.7G). This is highlighted by a variable time interval across different cell lines or growth conditions. For example, the PI3K inter-peak delay is 3~4 hours in HepG2 cells [279] but approximately 8 hours in

154

NIH 3T3 cells [270,279]. Our model was able to account for Myc accumulation pattern in both conditions by varying only the time-interval (either 3 hrs or 8 hours) between the two peaks of PI3K (Figure 5.8B).
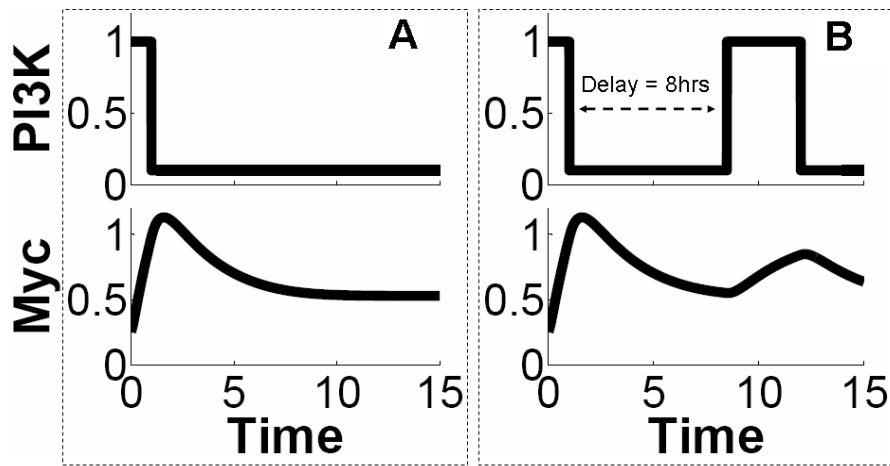
Figure 5.8: Impact of varying PI3K inputs on Myc accumulation. (A) A single peak of Myc is predicted if the second round of PI3K activity is removed. These results in reduced Myc accumulation compared to the wild-type. (B) Increased inter-peak time delay of PI3K (from 3 to 8 hours) results in wider separation between the two peaks of Myc, and the resulting Myc accumulation is less than the wild-type.

Another sensitive parameter of PI3K was its residual level. A mere two-fold increase in the residual level from the base case (10% of $PI3K_{Max}$), resulted in excessive increase in Myc level (red line in Figure 5.7H), consistent with an experimental study where exogenous Akt expression induced significantly increased Myc protein levels

155

[301]. Interestingly, however, PI3K$_R$ below a certain threshold level did not have much

impact on Myc accumulation (blue line overlapping with black line in Figure 5.7H).

Such threshold effect is due to the ultrasensitivity in the PI3K signaling cascade (Figure

5.5 and Figure 5.6). If the change in the PI3K residual level triggers a digital switching

behavior, it can cause a large change in the output (black to red lines in Figure 5.7H).

Any change in the residual level outside the ultrasensitive region will not cause any

significant output change.

The results in Figure 5.7 suggest that Myc accumulation was insensitive to

changes in the maximum amplitude of Erk and PI3K signals (Figure 5.7C and F), but

much more sensitive to their temporal features such as duration and inter-peak time

delay, and their residual values (Figure 5.7D, E, G, and H). This occurred because the

maximum amplitudes of Erk and PI3K pulses were at their saturation level. That is,

when the Erk pulse is sufficiently strong, Myc is almost completely converted into

Myc^{Ser62}; strong PI3K pulses block further phosphorylation of Myc^{Ser62} to Myc^{Thr58}. If so,

this mechanism will allow cells to resist further changes in Erk and PI3K amplitudes.

Such resistance (or insensitivity) to amplitude changes (or fluctuations) of Erk and PI3K

may underlie precise control of signal transduction by Myc, given its role as a key

regulator of downstream cellular events. That is, dysregulation of Myc activities, which

is a signature of various cancers, may have detrimental consequences [302]. To prevent

such dysregulation, activation or deactivation cues must be transmitted and integrated

156

precisely to regulate Myc accumulation. We note that this noise-resistance, which we

define as insensitivity to the changes in Erk or PI3K level in individual cells, requires the

maximum amplitudes of Erk and PI3K to be sufficiently large. If their amplitudes and

residual values are set 10 fold lower, the Myc accumulation becomes much more

sensitive to perturbations around the new base values.

5.4 Mathematical analysis of a generic signal integrator

The Erk and PI3K pathways that control Myc protein turnover are conserved in

yeast [269], and may represent a general post-translational strategy in natural signaling

pathways [303,304,305,306,307,308,309]. For instance, β-catenin stability is regulated by

casein kinase Iα (CKIα) and Gsk3β [307,309]. Similarly, an unknown kinase and Gsk3β

coordinate to modulate microtubule (MT) stabilizing activity [308]. These examples

consist of a dual-kinase motif that integrates two independent input signals (Figure

5.9A). In this motif, X represents the unphosphorylated effector protein, which is

unstable. It can be stabilized by kinase S_1 through phosphorylation (becoming Xp) and

subsequently destabilized by kinase S_2 through additional phosphorylation (becoming

Xpp), as shown in Figure 5.9B.

157
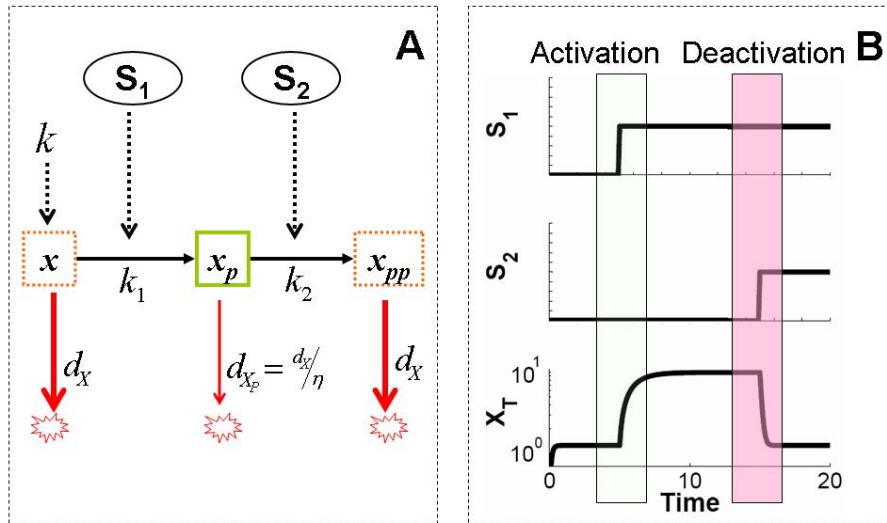
Figure 5.9: Dual-kinase module as a signal integrator. A) The dual-kinase mechanism. S_1 and S_2 determine gain and loss of X stability by sequential phosphorylation, which in turn control the total amount of the target protein ($x_T = x + x_P + x_{PP}$). k_1 and k_2 are the rate constants for phosphorylation by S1 and S2, respectively. d_X and d_{X_P} are degradation rate constants of the unstable (X or X_{pp}) and stable (X_P) forms of X. B) Given sufficiently strong input signals S_1 and S_2, the dual kinase mechanism integrates upstream activating signal S_1 to turn on, and deactivating signal S_2 to turn off. The time delay between the two signals controls the duration of activation.

158

The wide presence of this motif suggests its potential advantages for cellular signal processing. To gain insights into this issue, we developed a simplified model to analyze dynamics of the dual-kinase motif (see *Model and Methods*). In the model, we treated the two inputs of the system (Figure 5.9A) as independent, decoupled upstream signals, since most of the dual-kinase motifs found in nature often integrate independent upstream signals. Using this model, we aimed to explore what properties of this basic motif may underlie the dynamics observed for Myc regulation, and what advantages these properties may confer in cellular signaling.

5.4.1 Construction of a mathematical model: the dual-kinase motif as a generic signal integrator

Based on the connectivity in Figure 5.9A, we modeled the dual kinase motif using three highly simplified ordinary differential equations (ODEs), as presented in a dimensionless form:

$$\frac{dx}{d\tau} = \kappa - \alpha \frac{x}{1+x} - \eta x, \qquad \frac{dx_P}{d\tau} = \alpha \frac{x}{1+x} - \beta \frac{x_P}{1+x_P} - x_P, \qquad \frac{dx_{PP}}{d\tau} = \beta \frac{x_P}{1+x_P} - \eta x_{PP}$$

where x, x_P, and x_{PP} are the concentrations of the three forms of a molecule X; τ is an independent variable, time; κ describes the synthesis of X ($\frac{k}{K_d x_P}$); α is the activation efficiency by S_1 ($\frac{k_1 S_1}{K_d x_P}$); β is the deactivation efficiency by S_2 ($\frac{k_2 S_2}{K_d x_P}$); and η is the ratio of unstable protein to stable protein ($\frac{d_x}{d_{x_P}}$), or *stabilization efficiency*. Without loss of generality and for simplification, we assumed that x and x_{pp} have the same stability.

159

Similar to Myc regulation, we used the total effector concentration X (= $x + x_P + x_{PP}$) to represent the system output. As shown by a typical simulation (Figure 5.9B), this module enables integration of two signals by the effector module. Drawing analogy to electric signal processing, the output can be considered a combination of 'NOT' and 'AND' operators, which defines a pulse of output.

5.4.2 Characterization of the dual-kinase motif

To characterize the dual-kinase motif, we first examined dose response of the system with respect to the two inputs S_1 and S_2. Our results indicated that system activation (through phosphorylation by S_1) was sensitive to input variations at an intermediate α value (Figure 5.10A). In contrast, it was insensitive to input variations at either high or low α values (green curve). Similar insensitivity (or noise-resistance) at either strong or weak input signals was observed for system deactivation (Figure 5.10B) for varying β. These results provide a mechanistic explanation for the insensitivity of Myc potency to input signal strengths (Figure 5.7C and F).
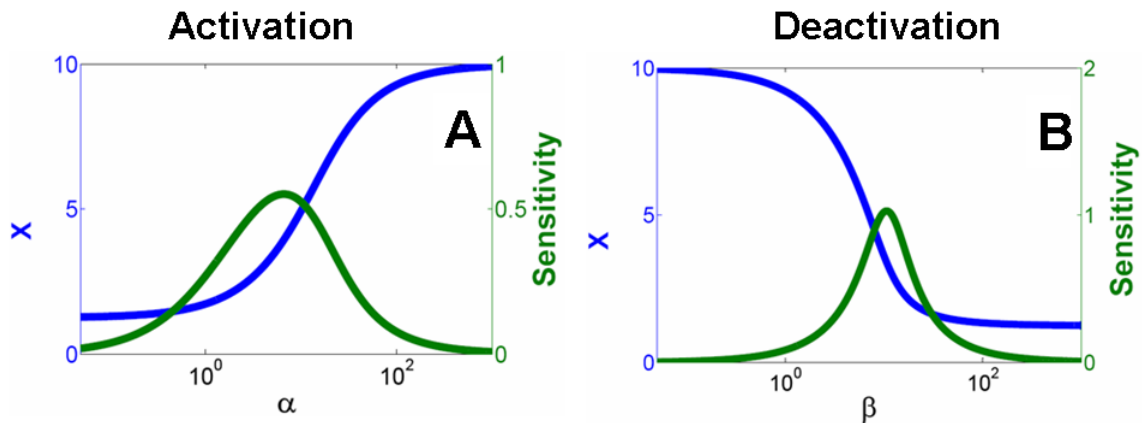
Figure 5.10: System sensitivity to input signal perturbations. A) At a given synthesis rate constant (κ=10), the maximal activated level of X at the steady-state (X_{ss}) can be modulated by α. For small or large α, sensitivity (defined as $\ln X/\ln\alpha$) was minimal, while it was the greatest at intermediate α values. We assumed 0 for β to allow decoupling of activation from deactivation. B) Deactivation from the high state depended on β at a given κ. The system was initially driven to its high state by assuming a large α (10000). Sensitivity was minimal for small or large β, and was the greatest at intermediate β values.

Another salient feature of the dual kinase motif was the stabilization of X, which could be captured by the stabilization efficiency (η), or the ratio between the degradation rate constant of the unstable form and that of the stable form. Our analysis indicated that the stabilization efficiency determines the dynamic range of the output X. In response to S_1, the upper bound of output was set by the synthesis rate of X (κ) and was asymptotically approached as the signal strength (α) increased (Figure 5.11A). The lower bound of the output, however, was set by κ/η, which corresponded to the basal level of X in the absence of S_1. Thus, η directly set the dynamic range of the output ($\kappa/\eta \sim \kappa$). Similar dependence was also applied to deactivation of X by S_2 (Figure 5.11B). Given

161

saturating activation by S_1, the dynamic range for deactivation by S_2 increased with

increasing η, allowing the dynamic range to be flexible. The system approached the

basal output level (κ/η) with an increasing strength of S_2 ($\beta \rightarrow \infty$).
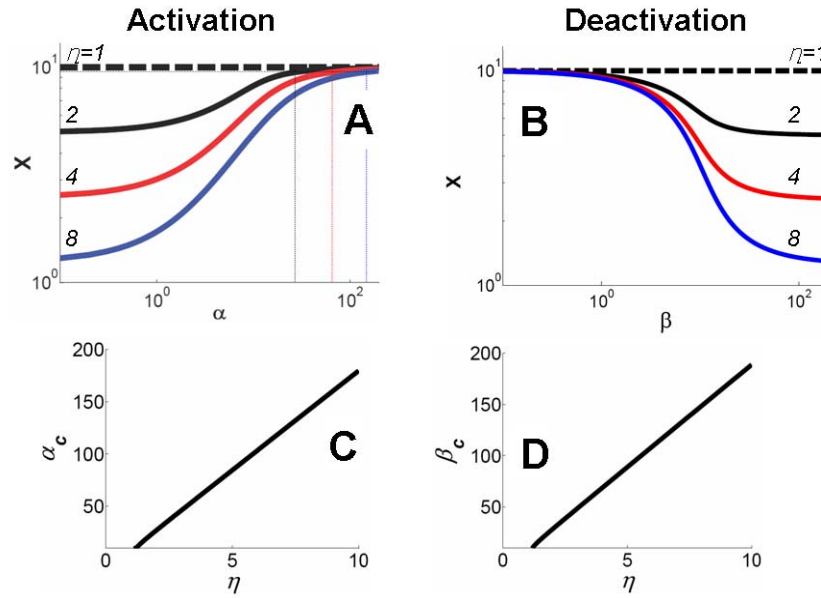
Figure 5.11: Dynamic range of output activation and deactivation. A) The dynamic range for activation was η because: $X_{ss} \approx \kappa/\eta$ as $\alpha \rightarrow 0$; $X_{ss} \approx \kappa$ as $\alpha \rightarrow \infty$. For a given η, we define a critical value α_C that corresponds to an $X_{ss} = 95\%$ of the maximal value. If $\alpha > \alpha_C$, fluctuations in X_{ss} due to fluctuations in α would be smaller than 5%. Here we consider system activation in this parameter range as effectively noise-resistant. Similar to analyses in Figure 5.10, we assumed 0 for β and a large value (10000) for α, which allowed analyzing dynamic range for activation and deactivation independently. B) Given a sufficiently large α, the dynamic range for system deactivation was also η because: $X_{ss} \approx \kappa$ as $\beta \rightarrow 0$; $X_{ss} \approx \kappa/\eta$ as $\beta \rightarrow \infty$. For a given η, we define a critical value β_C that corresponds to X_{ss} within 5% of its minimal value. Similar to (A), we consider system deactivation to be effectively noise-resistant for $\beta > \beta_C$. C) α_C increased with η almost linearly. D) β_C increased with η almost linearly.

These results highlight two appealing features of the dual kinase motif. First, differential stability control on effector protein isoforms enables flexible modulation of the output dynamic range. This dynamic range can be fully exploited if the signal strengths are sufficiently large. Second, sufficiently strong signals will also result in desensitization of the system output to minor fluctuations in the levels of these signals.

While advantageous, however, increase in noise-resistance and dynamic range comes with increasing metabolic cost. On one hand, increasing destabilization of X or Xpp is associated with increasing metabolic cost. On the other, this will also require stronger input signals to fully exploit the increased dynamic range and to achieve noise-resistance, creating another metabolic burden as characterized by α and β. To quantify this effect, we define a critical α value (α_C), which corresponds to a steady-state X (Xss) value at 95% of the maximum X (for $\alpha \rightarrow \infty$). If the input signal would fluctuate in the range of $\alpha > \alpha_C$, the resulting output fluctuation would never exceed 5% (regardless of the magnitude of input signal fluctuation). Here we can consider system activation as noise-resistant in this parameter range. With similar reasoning, we define a critical β_C, which corresponds to an Xss value within 5% of the minimum X (for $\beta \rightarrow \infty$). α_C and β_C thus determine the minimal signal strengths required to achieve noise-resistance in system activation or deactivation. As shown Figure 5.11C and D, the greater the stabilization efficiency was (larger η), the heavier would be the corresponding metabolic burden (larger α_C or β_C) required to achieve noise-resistance. Insufficient input signal

163

strength would either fail to generate response or fall into the sensitive range of the dose response curve (Figure 5.10).

5.5 Discussion

Here we demonstrate that modulation of Myc stability by sequential phosphorylation enables Myc to precisely sense and integrate upstream Erk and PI3K signals. Such regulation is likely critical to cell fate decisions. Our analysis indicates that, when operating with appropriate parameters, this mechanism enables the temporal features, instead of maximum amplitudes, of the upstream signals to precisely modulate Myc accumulation. Supporting this notion, dynamics of a minimal dual-kinase motif provide direct, intuitive explanation for the key sensitivity properties of Myc output in the full model. In this work, we have limited our study to the well-defined post-translational control of Myc. It is possible that robust control of Myc accumulation is facilitated by additional mechanisms, including Myc stabilization by a signal in the carboxy-terminus of Myc [310] and Myc sequestration for degradation [311,312]. Myc modulation is also tuned by regulations at other levels including post-transcription [313] and translation [314], along with feedback control [315]. Furthermore, the activities of Pin1 and PP2A, which we assumed to be abundant and not rate-limiting, may further contribute to more complex Myc dynamics, as seen in various cancers [49-52].

As Myc is often deregulated in cancers, quantitative understanding of the

mechanisms for Myc regulation may be helpful for developing novel strategies for

cancer treatment. Myc stabilization processes consist of two temporally coordinated

events: Myc stabilization by Erk and prevention of Myc degradation by PI3K. While the

significance of Myc degradation by the second PI3K activity has been suggested in cell

proliferation [270], the extent to which the initial Myc stabilization by Erk contributes to

cell proliferation remains unknown. Our model predicts that, for the second round of

Myc accumulation, Myc needs to be sufficiently accumulated by Erk_R prior to the second

PI3K activity (black line in Figure 5.12). With the PI3K signal fixed, a small increase in

Erk_R is predicted to result in a significant increase in Myc accumulation pattern (red line

in Figure 5.12). In contrast, removal of Erk_R renders Myc unresponsive to the PI3K signal

(blue line in Figure 5.12). This is due to the sequential nature of the Myc stabilization

processes, where Erk activity must precede PI3K activity. In other words, while Erk

'primes' Myc activity, PI3K 'fine-tunes' Myc accumulation.
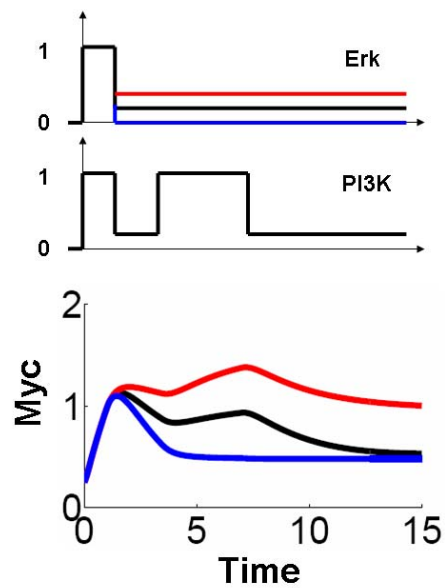
Figure 5.12: Erk 'primes' Myc activity, and PI3K 'fine-tunes' Myc accumulation level.
With the PI3K signal fixed, different residual Erk level leads to differential Myc
accumulation by the second PI3K activity. The base value of the residual Erk level
(ErkR) was 10 percent of maximal Erk level (black line). For increased level of ErkR
(20%), the second PI3K activity increased Myc accumulation level significantly (red line).
When ErkR was completely removed, Myc became unresponsive to the PI3K signal
(blue line).

The priming ability of Erk for Myc modulation may play a critical role in distinct responses to different stimulations. Studies have shown that PC12 cells can be induced to undergo differentiation or proliferation in response to NGF or EGF [316,317], and the residual Erk level may be responsible for these differential cellular responses [13]. Our analysis suggests that the ability of Erk to modulate these cellular responses is through modulation of Myc accumulation. For EGF stimulation, the low residual Erk level may induce proliferation by weakly priming Myc (black line in Figure 5.12). In contrast, the high residual Erk level upon NGF stimulation may lead to a significant increase in Myc, inducing differentiation (red line in Figure 5.12). This notion can be experimentally tested by simultaneous time-course measurements of the input signals Erk and PI3K, and the output Myc protein. Also, the input signals can be independently controlled by inhibitor drugs [318], inducible systems, or siRNA molecules targeting the MAPK or PI3K pathways [319].

The assumed saturation of the input signals in the base model can also be experimentally tested. Our simulations indicate that the assumed saturation is a necessary condition for the overall robustness of Myc to parameters. This serves as an interesting question to explore experimentally. Also, as detailed in Figure 5.8, some constituent reactions in the PI3K pathway (e.g. the Ph-dePh cycles) have not been well-characterized at the quantitative level. Our additional model predictions on how the

167

overall response of Gsk3β to PI3K depends on sensitivity characteristics of individual stages can serve as further targets for experimental tests.

The analysis of the Myc stabilization mechanism reveals a regulatory network motif that may be ubiquitously used in nature. Network motifs are small, recurring cellular regulatory networks, identified and characterized by their shared architectures and functions among diverse organisms. Well-known examples include feedback regulations, feed-forward loops, and their derivatives (see [1,43,46,320] for review). Here we suggest that the dual-kinase motif represents another example with distinctive features.

The dual-kinase motif is similar to a well-studied phosphorylation-dephosphorylation enzymatic motif of protein modification. In both motifs, protein modification events occur sequentially, and the current state of the protein hinges upon its previous state. Given appropriate input signals and parameters, the sensitivity and amplitude of the output response can be precisely controlled [321]. The dual-kinase motif differs from the phosphorylation-dephosphorylation one, however, in that protein modification process is irreversible. Once phosphorylated, the stabilized protein cannot return to its initial state, but is targeted for degradation upon further modification. This distinctive characteristic contributes to additional features of the dual-kinase motif: sequential signal integration of multiple inputs and, correspondingly, flexible dynamic range for the output governed by protein stability modulation.

168

Chapter 6: Conclusion and future directions

In summary, our works here demonstrate that a combination of mathematical modeling and experimental validations is an effective approach for understanding complex dynamics that arise from highly connected signaling pathways. With this approach we have investigated an important biological phenomenon: cell cycle regulation. In particular, we characterized the Myc-Rb-E2F signaling pathway and its upstream signaling pathways. Such characterizations provide insights into three important questions in the cell cycle regulation: 1) regulation of the restriction point (R-point) in mammalian cell cycle entry, 2) regulation of temporal variability in cell cycle entry, and 3) regulation of Myc by its upstream signals.

In Chapter 3, we demonstrate that the Myc-Rb-E2F signaling pathway underlies the R-point dynamics. First, we built a mathematical model for the Myc-Rb-E2F signaling pathway based on the current known biology. With this model we predicted bimodality in the activation of a transcriptional factor E2F, which is responsible for DNA synthesis. In addition, we predicted history-dependent activation of E2F (hysteresis). This is, once E2F is activated upon growth stimulation, it stays at its activated state even after removal of growth stimulation. However, E2F remains at its inactivated state without growth stimulation. These predictions were experimentally validated in collaboration with a molecular biologist Guang Yao. Our modeling and experimental results together showed that E2F activation dynamics in the G1-S transition demonstrate

169

bimodality and hysteresis, which are characteristic of a bistable switch. These results

suggest that the Myc-Rb-E2F signaling pathway underlie the R-point dynamics.

Our analyses of the R-point dynamics focused on the steady-state dynamics of

E2F activities, but the temporal dynamics of E2F activities in cell cycle entry is largely

unknown. In Chapter 4, we investigated the temporal dynamics of E2F activation by

extending our work in Chapter 3. In particular, we focused on the variable transition of

the mammalian cell from the quiescence to proliferation. Over the last four decades, two

lines of apparently contradictory, phenomenological models have been proposed to

account for such temporal variability. These include various forms of the transition

probability (TP) model and the growth control (GC) model. The GC model was further

proposed as an alternative explanation for the concept of restriction point, which we

demonstrated as being controlled by a bistable Rb-E2F switch in Chapter 2. Through a

combination of modeling and experiments, we show that these different lines of models

in essence reflect different aspects of the stochastic dynamics in cell cycle entry. In

particular, we show that the variable activation of E2F can be described by stochastic

activation of the bistable Rb-E2F switch. This activation dynamics can be recast into the

framework of the TP model and that of the GC model. While the phenomenological

models lack direct mechanistic insights into the underlying dynamics, we show that

there is a quantitative mapping between these models and the mechanistic model. As

such they may be useful for defining concise, quantitative phenotypes of the cell physiology.

In Chapter 5, we investigate the post-translation regulation of Myc, an input to the Myc-Rb-E2F signaling pathway. This is via protein stability modulation by its upstream Ras effector pathways: the extracellular signal-regulated kinase (Erk) and phosphatidylinositol 3-kinase (PI3K) pathways. Precise control of Myc expression is critical in regulating diverse cell-fate decisions, including growth, proliferation, and programmed cell death. Underscoring its importance, Myc expression is often found to be deregulated in cancers. However, the dynamic mechanism for Myc modulation remained unclear. To address this issue, we analyzed a well-defined signaling module for Myc regulation using a kinetic model constrained by experimental data and observations. In this module, Myc acts as an integrator of its upstream signals that differentially regulate its stability. We showed that this module can enable highly sensitive Myc response to the temporal features of the input signals, but not to their maximum amplitudes. We further suggest that this module represents a generic post-translational mechanism for signal sensing and integration in diverse signaling networks. Our work offers insight into the 'design' of natural biological networks and makes predictions that can guide further experimental studies on Myc regulation. Moreover, it defines a simple signal processing unit that may be useful for engineering synthetic gene circuits to carry our cell-based computations.

171

In Chapter 3 and Chapter 4, we have taken great care to develop a mathematical model that is consistent with experimental observations. Such well-calibrated model can be used for further predictions that are experimentally tractable. For example, *in-silico* perturbations to the CycD/Cdk4,6 node may show significant modulation of the activation threshold for E2F elevation. In fact, our simulation results predict that reducing the activity of CycD/Cdk4,6 may increase activation threshold for E2F elevation and increase temporal variability in cell cycle entry. Such perturbations can be realized experimentally by adding various Cdk4,6 inhibitors drugs (Calbiochem, Cat. No. 219476, 219478, 219477, and 219492) to the REF52-d2GFP cells. The effects of these drugs on the steady-state and temporal dynamics in the cell cycle entry can be quantitatively measured by reading GPF signals with flow cytometry. In addition, the effects of varying Rb synthesis can be simulated. Our simulation results predict that the Rb synthesis is directly correlated with activation threshold of growth stimulation and temporal variability. For higher Rb synthesis, stronger growth stimulation may be required for E2F elevation, and increased temporal variability is expected. Targetting of the Rb synthesis may be achieved by small interfering RNA (siRNA) or short-hairpin RNA (shRNA). The effects of modulating Rb synthesis can also be quantified by GFP reading with flow cytometry. In addition, we can predict population dynamics of E2F activation under combinatorial perturbations including growth factors, inhibitor drugs

or siRNA targeting the Myc-Rb-E2F network, or mutations within this network. These results may be used to further calibrate our Myc-Rb-E2F model.

We discussed the mapping between the phenomenological models and the stochastic Rb-E2F model in Chapter 3, which provides a simple conceptual framework that reconciles the different views reflected in the TP and GC models. In addition, this mapping provides quantitative description of the population-level dynamics associated with variable cell cycle entry (by phenomenological models) as well as the state of the underlying reaction networks (by stochastic Rb-E2F model).

Based on this mapping, we can develop a library of predictions that can be uniquely described by a set of parameters defining the TP model or the GC model. This library would consist of different sets of parameters in the stochastic model, leading to different parameters in the TP or the GC model. Based on this library, experimentally obtained parameters for the phenomenological models can be used to reveal the state of the underlying reaction networks. As such, the TP and GC models can serve as a concise platform to define quantitative phenotypes that facilitate classification of cell types or cell states.

This idea may be experimentally demonstrated with REF52-d2GFP cells. These cell lines may be subjected to varying testing conditions, and the unique parameters in the TP or GC model describing the population-level dynamics can be obtained. Then, a correlation may be established between the temporal dynamics of these cells and the

library of predicted phenotypes based on the Myc-Rb-E2F network. With sufficient number of testing conditions, such correlation may be used to reveal the activation status of the Myc-Rb-E2F network. Using this technique, different cell lines carrying the same expression cassette encoding destabilized EGFP may be tested to identify the underlying mechanism for the observed patterns (i.e. mutations in the Myc-Rb-E2F network).

In the study of temporal variability in cell cycle entry in Chapter 4, we focused on a single transition from quiescence to proliferation during the cell cycle progression due to its computational and experimental tractability. However, the temporal variability at the G1-S transition in actively growing cells may differ significantly from that in quiescent cells [229]. To compare the temporal variability in the quiescent and actively growing REF52-d2GFP cells, these cells may be grown over multiple cell cycles and observed under the microscope. The difference in the temporal dynamics between the two conditions may be recapitulated with our mathematical model by assuming different initial conditions.

Appendix A: Bimodality in bacterial infection

In Appendix A, we apply mathematical tools described in Chapter 1 to the understanding of bacterial infection dynamics. In particular, we focus our analyses on the bacterial attachment step, which has been the topic of interest in studying infection dynamics and developing anti-bacterial or bacteria-mediated therapeutics. Complicating these efforts is the inherent variability in bacterial uptake, but the underlying mechanism remains largely speculative. To gain more insights into such variability in bacterial uptake, we have developed an experimental model system by engineering non-pathogenic bacteria to express invasins, which interact with mammalian receptors $\beta 1$-integrins. Using this model system, we show that uptake of the engineered bacteria is variable, leading to bimodal infection kinetics. This can be accounted for by a simple mathematical model that incorporates the 'zipper mechanism', together with a positive feedback mediating the invasin-integrin interactions. Taking advantage of this simple model, we make predictions on correlation between the average receptor concentration and the percentage of infection. Our simulations predict robust correlation to variations in model parameters, which were experimentally represented by various mammalian host cell lines that span a wide range of $\beta 1$-integrin levels. Our experimental correlation was consistent with the model prediction, suggesting the model's utility in predicting bacterial infection dynamics. Our results together suggest that our work may serve as a

useful guide for studying infection dynamics or developing therapeutics. We are in the final stage of preparing a manuscript for submission.

A.1 Introduction

The infection cycle of a bacterial pathogen typically consists of multiple steps, including adherence to the host, internalization, survival and replication, and ultimate intoxication or death of the mammalian host. Although the molecular details of bacterial infection vary among pathogen species, many pathogens share common strategies in the infection processes. For example, bacterial adherence to the host cell is often mediated by adhesins, which are molecules or molecular structures that facilitate fimbrial or afimbrial binding of bacteria to their hosts [322]. In addition, bacterial internalization by the hosts is achieved by largely two common mechanisms: 'trigger' mechanism, where signaling triggers membrane ruffling and cytoskeletal rearrangement for bacterial entry, and 'zipper' mechanism, where sequential interactions between the ligands and the receptors result in encapsulation of the bacterium [323].

A.2 Background

A population of isogenic host cells subjected to isogenic pathogens often shows a drastic cell-to-cell variability in bacterial uptake [324]. Some hosts are highly infected,

while others remain free of bacteria. Even in the infected subpopulation, bacterial uptake is variable ranging from a single bacterium to hundreds of bacteria per host cell. This variability may result from heterogeneity in both pathogens and mammalian host cells. In isogenic bacterial pathogens, it has been shown that invasion efficiency may differ due to stochastic gene expression of virulence factors (i.e. phase variation for virulence gene expression [325,326,327]), external factors including bacterial growth conditions (growth media, temperature, or shaking speed), or growth phase (logarithmic, late-log, or stationary). In mammalian host cells, bacterial uptake may depend on the activation state of the involved signaling transduction pathways [328,329] or receptor distribution [330]. These individual differences in the pathogens or in the host cells together may result in variable bacterial uptake. However, the variable bacterial uptake has been largely overlooked in the study of host-pathogen interactions, and the underlying mechanism for such variability remains poorly understood.

To gain insight into variable bacterial uptake, here we analyze the infection dynamics by an engineered bacterium that expresses invasins, which interact with mammalian receptors β1-integrins. Our experimental results show that uptake of the engineered bacteria is highly variable in mammalian host cells, leading to bimodal infection kinetics. This can be accounted for by a simple mathematical model that incorporates the 'zipper mechanism', together with a positive feedback mediating the invasin-integrin interactions. With this model, we predict that the correlation between

177

the average receptor concentration and the percentage of infected cells would be robust

against variations in model parameters. This prediction was validated by further

experiments carried on cell lines with varying concentrations of β1-integrins. These

results together suggest that our work may serve as a quantitative framework for

guiding the modulation of bacterial infection kinetics.

A.3 Mathematical model and experimental analyses

A.3.1 Construction of the mathematical model

We have developed a mathematical model based on the interactions between

invasin-expressing bacteria and mammalian *β1-integrin* receptors. These interactions

begin with a bacterium with membrane-bound *invasins* initially making contact with *β1-*

*integrin*s. This initial contact facilitates sequential interactions of *invasin-integrin*

interactions, resulting in encapsulation of the bacterium [328,331]. A mathematical

model for such serial interactions between a multi-valent ligand and receptors have been

previously developed by Perelson [332], but our simulations of the Perelson's model

could not sufficiently account for the experimentally observed threshold effects . In this

work, we develop a highly simplified model based on the Perelson's model by making

additional assumptions. These assumptions include: 1) *β1-integrins* exist in either

available (R_a) or unavailable (R_u) pools for binding with bacteria, and the their total

concentration is constant, 2) once a bacterium in suspension (determined by MOI) binds

the host via the initial invasin-integrin interaction, further interactions are assumed to be

instantaneous, and 3) the downstream events after adherence are not considered in the

current model. This is based on the experimental observation that bacterial processing

by the mammalian host is minor (Figure A.1). The model consists of a set of ordinary

differential equations (ODEs), and its parameters are obtained either from the literature

or fitted from experimental data.

Table A.1: The mathematical model for bacterial adherence

$$\frac{dB_1}{dt} = -k_f[B_1][R_a]^n + k_r[B_n] + B_i[R_a] - k_r[B_1]$$

$$\frac{dB_n}{dt} = k_f[B_1][R_a]^n - k_r[B_n]$$

$$\frac{dR_i}{dt} = k_i[R_a] - \frac{k_a[R_i][B_1]}{B_1 + K_B} - k_b$$

$$\frac{dR_a}{dt} = -k_i[R_a] - n*k_f[B_1][R_a]^n - B_i[R_a] + n*k_r[B_n] + k_r[B_1] + \frac{k_a[R_i][B_1]}{B_1 + K_B} + k_b$$

Variables are:
B_1: The amount of bacteria bound with one receptor
B_n: The amount of bacteria with 'n' number of receptors
R_a: available integrin receptors available for bacterial binding
R_i: unavailable receptors

Initial conditions:
Total receptor concentration ($R_T = R_i + R_a$) = 10nM [333,334]
R_i = 10nM, R_a=0nM

179

Table A.2: Reaction kinetics for the model

Reaction	Kinetics	Description and notes
$[R_i] \xrightarrow{\frac{k_a[B_1]}{B_1+K_B}} [R_a]$	$\dfrac{k_a[R_i][B_1]}{B_1+K_B}$	Receptor activation by positive feedback
$[R_i] \xrightarrow{k_b} [R_a]$	$k_b[R_i]$	Basal receptor activation
$[R_a] \xrightarrow{k_i} [R_i]$	$k_i[R_a]$	Receptor inactivation
$[R_a] \xrightarrow{k_f B_i,} [B_1]$	$k_f[B_i][R_a]$	Initial invasin-integrin interaction
$[B_1] \xrightarrow{k_r} [R_a]$	$k_r[B_1]$	Unbinding of the initial interaction
$[B_1]+n[R_a] \xrightarrow{k_f} [B_n]$	$k_f[B_1][R_a]^n$	Subsequent invasin-integrin interactions
$[B_n] \xrightarrow{k_r} n[R_a]+[B_1]$	$k_r[B_n]$	Unbinding of the subsequent interactions

Table A.3: Parameters for the model

Rate constants		Parameter values, sources, and notes
n	10	These free parameters were adjusted to satisfy the following criteria :
k_a	0.1 nM^{-1}h^{-1}	
k_i	0.3 h^{-1}	1) Bimodal distribution
k_b	0.001h^{-1}	
k_r	0.036 h^{-1}	Rate constant for unbinding [334]
k_f	0.0072 nM^{-1} h^{-1}	Rate constant for invasin-integrin interaction (calculated from the known dissociation constant [335] and k_r)
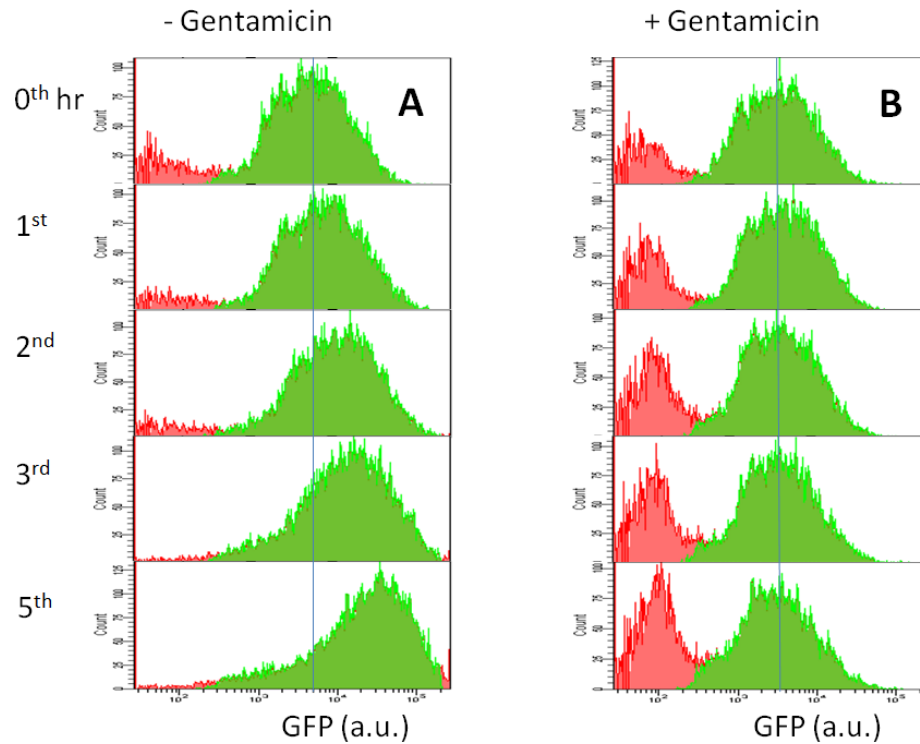
Figure A.1: Temporal dynamics of bacterial adherence/internalization after a pulsatile input of bacteria. HeLa cells were co-incubated with bacteria (grown overnight and diluted in DMEM) at high MOI (=2000) in the absence (A) and presence (B) of gentamicin. The mixture was co-incubated in DMEM with 10 % bovine growth serum (BGS) for 1 hour, and we removed bacteria in suspension by washing with PBS 3 times. The cells were replenished with fresh DMEM with 10% BGS and the antibiotic treatment condition was maintained as before. At various time points after washing, we assayed the cells for their bacterial uptake with flow cytometry. In the absence of gentamicin, the mean GFP signals became stronger over time, suggesting bacterial growth in the host cells either on the surface or inside of the host cells. In contrast, we observed minor decrease in the GFP signals strength in the presence of gentamicin, suggesting degradation of the bacterium by the host cells or by the antibiotics.

A.3.2 Experimental setup

A.3.2.1 Plasmid construction

The invasin-coding plasmid *placInv* was constructed by fusing a PCR-derived fragment of the *inv* gene from *pGB2 Ωinv-hly* [336] into *pPROLar.A*. Similarly, another invasin-coding plasmid *pSCT7Inv* was constructed by fusing *inv* gene into xxx. The reporter plasmid *ptetGFP* was created by fusing a PCR-derived fragment of the *gfp* gene into *pPROTet.E*.

A.3.2.2 Bacterial transformation and culture

Bacterial strains used in this study were Top10F' and BL21DE3. Both cell strains were transformed with a reporter plasmid *ptetGFP*, and Top10F' was transformed with *placInv* and BL21DE3 was transformed with *pSCT7Inv*. They were grown overnight at 37°C in Luria-Bertani medium, and were diluted with D-MEM until their absorbance (Abs_{600}) reading by a plate reader ( ) was approximately 1.0. At this absorbance, the number of bacteria in 1 μL of the bacterial culture was approximately 2.0×10^6.

A.3.2.3 Mammalian cell culture and media

All mammalian cells (graciously given by the Dr. Nevins lab) were grown in D-MEM (GIBCO® Cat. No. 31053-036), supplemented with L-glutaminine (Cat. No. 25030-164), sodium pyruvate (Cat. No. 11360-070), and 10 % bovine growth serum (BGS).

These cells were detached by trypsinization and seeded on 6-well plates at

approximately 2.0 x 10^5/well, and were allowed for recovery overnight in D-MEM in

10% BGS overnight.

A.3.2.4 Co-incubation and GFP assay

MOI in all the experiments in this study was determined by the number of

bacterial cell number (2.0 x 10^6) added to each well of host cells in the 6-well plate (2.0 x

10^5). The bacteria in D-MEM were co-incubated with mammalian host cells on the 6-well

plates in the presence of gentamicin at varying MOIs. After the desired co-incubation

duration, the 6-well plates were washed 3 times with PBS to wash off extracellular

bacteria. These cells were assayed for their GFP signals under the microscope (Leica) or

flow cytometry (Canto 2).

A.3.2.5 Quantification of β1-integrin concentration

For quantification of *β1-integrin* concentration, we use a primary antibody

(CBL497, Millipore) directed against the *α*-chain of the *α5β1*-integrin receptors. The

primary antibody can be conjugated to the immunofluorescent secondary antibody

(phycoerythrin, PE), allowing for quantification of the receptor concentration by flow

cytometry.

A.3.3 Simulation and experimental results

Our experimental system consists of mammalian host cells (HeLa) and engineered bacteria (BL21DE3) that carries two plasmids: *pT7Inv* and *ptetGFP*. *Invasins* facilitate binding of the bacteria with its mammalian hosts and subsequent internalization, while the *GFP* signals are used as a means for quantifying bacterial uptake. For all experiments, we allowed co-incubation of bacteria with host cells for varying durations and MOIs. During co-incubation, bacterial growth was inhibited by gentamicin treatment and heterogeneity in bacterial uptake in the host population was reduced by constant shaking (at 30 rpm). After the co-incubation, we removed bacteria in suspension by washing, and the host cells were assayed for the GFP signals.

A.3.3.1 Quantification of bacterial uptake in the population and single-cell levels

We first visualized binding of the engineered bacteria with HeLa cells under the fluorescence microscope, as shown in Figure A.2A. A population of HeLa cells co-incubated with the engineered *E.Coli* showed significant variability in bacterial uptake. In Figure A.2A, some HeLa cells were bound by a large number of bacteria (~ 20 bacteria), while others remained free of bacteria. Such variability was also reflected in our flow cytometry measurements. Our results in Figure A.2B showed MOI-dependent bimodality in GFP signals. Based on the intensity of the GFP signals, we separated the cell population into the low mode and high mode of GFP signals and observed them

184

under the microscope (data not shown). The mammalian cells with low mode of GFP signals were free of bacteria, suggesting that the GFP signals in a single bacterium may be sufficient to separate the two modes in the cell population. Those with strong GFP signals carried one or more bacteria as in Figure A.2A. A control study showed that HeLa cells co-incubated with BL21DE3 cells carrying only *ptetGFP* remained free of bacteria (data not shown). These results demonstrated that the engineered bacteria can interact with the host cells with specificity, and their interaction efficiency can be quantified.
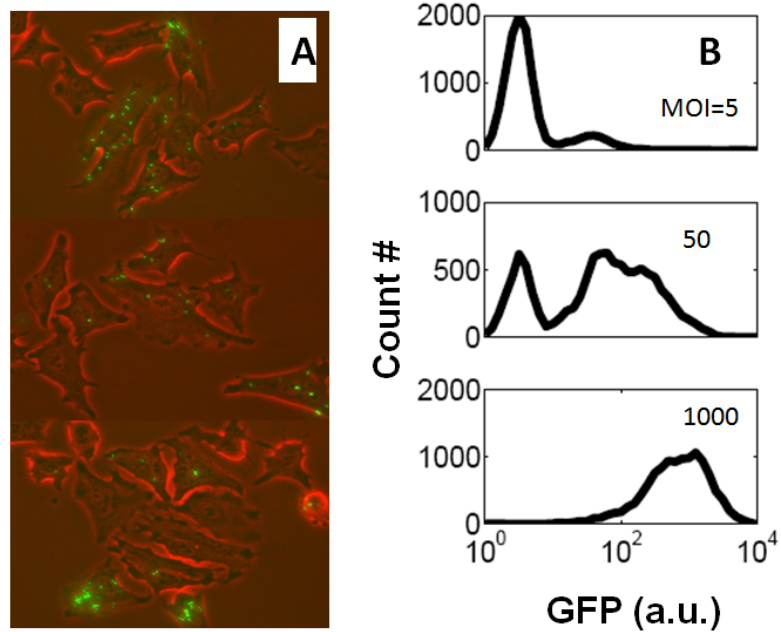
Figure A.2: Bimodal bacterial attachment. A) A microscope image of engineered bacteria attaching to HeLa cells. Bacteria were co-incubated with HeLa cells in a 6-well plate in the presence of gentamicin at 50 multiplicities of infection (MOI). After 2 hours of co-incubation, extracellular bacteria were removed by washing the wells with PBS three times. The uptaken bacteria were visualized with a inverted microscope. B) Bimodal distribution of bacterial attachment in a cell population. The infected HeLa cells on the plate were trypsized and their *gfp* signals were analyzed with flow cytometry. At sufficiently low MOI (=5), the cell population demonstrated bimodal distribution of gfp signals. For increasing MOI (=50), the low mode shifts to the high mode, until the distribution becomes monomodal at the high mode (MOI=1000). The low mode represents absence of bacterial attachment or internalization, and the high mode represents a subpopulation of cells with at least one or more bacterial uptake.

A.3.3.2 Dependence of bacterial uptake efficiency on MOI and co-incubation duration: threshold effects

It has been shown that bacterial uptake efficiency is positively correlated with increasing time and multiplicity of infection (MOI) [337]. To recapitulate such correlation, we co-incubated the engineered bacteria with a population of HeLa cells at various co-incubation durations and MOIs. Intuitively, the bacterial uptake would increase monotonically with increasing durations and MOIs, followed by a plateau at saturation. Consistent with our intuition, our experimental results showed positive dependence of bacterial uptake on co-incubation durations and MOIs (raw data shown in Figure A.3)
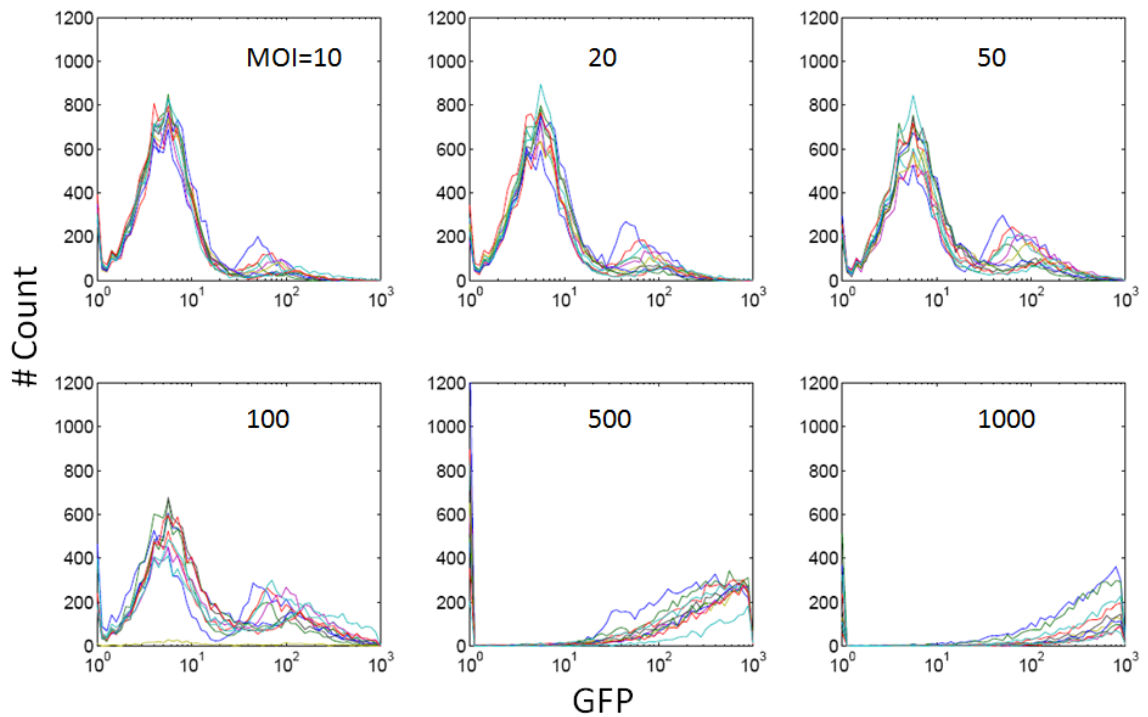
Figure A.3: Raw data for bacterial infection at various co-incubation durations and MOIs. HeLa cells grown in DMEM supplemented with 10% BGS were infected with BL21DES bacteria harboring *ptetGFP* and *pSCT7Inv* in the presence of gentamicin at varying multiplicity of infections (MOIs = 10, 20, 50, 100, 500, and 1000). After varying co-incubation durations (30, 60, 90, 120, 150, 180, 210, 240, 270, 300, and 330 mins), the extracellular bacteria were washed off with PBS, and HeLa cells were collected by trypsinization. These cells were assayed for their GFP signals under flow cytometry.
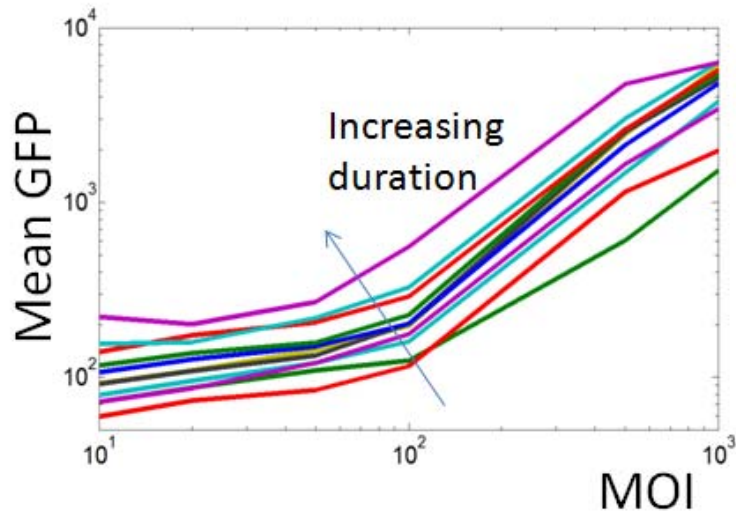
Figure A.4: The threshold effects in bacterial infection. The average GFP signal was calculated from a GFP distribution in **Figure A.3** and plotted over co-incubation durations. Below a threshold MOI (=100), the increase in bacterial uptake is insignificant. Above this threshold, however, gfp signal strength increased with increasing MOI.

A closer inspection of our results in Figure A.3 suggested two stages of infection separated by a threshold MOI ($\approx$ 100), as shown in Figure A.4. Below the threshold, the dependence of bacterial uptake on increasing MOI is small. Above the threshold, however, the dependence increases significantly with increasing MOI, demonstrating a 'sharp' transition from insensitive to sensitive infection dynamics to MOI. This sharp transition in bacterial uptake was also observed at the single-cell level (Figure A.5). At sufficiently high MOI (=100), our microscope images frequently showed a large number of bacteria on a host cell. At low MOI (=50), however, the frequency of such large number on a host cell reduced significantly. Our dose response curves at various co-

189

incubation durations suggested that the sharp transition in bacterial uptake to increasing

MOI was maintained for all co-incubation durations.
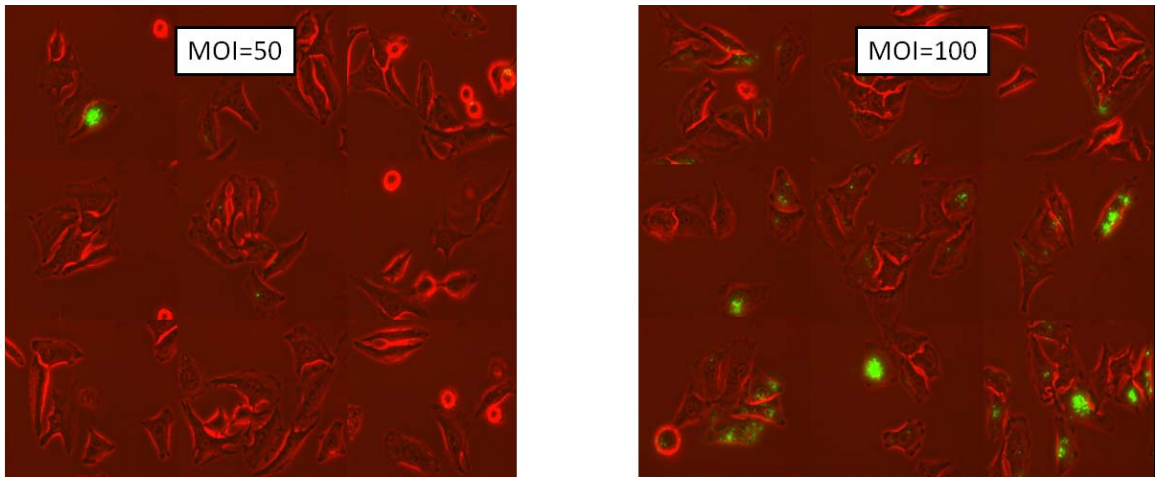
Figure A.5: Bacterial infection at MOI=50 and MOI=100. Multiple microscope images (9 images) were juxtaposed to show frequency of localization of bacterial infection. At sub-threshold MOI (=50), bacterial infection was approximately 10 fold lower than at MOI near the threshold (= 100).
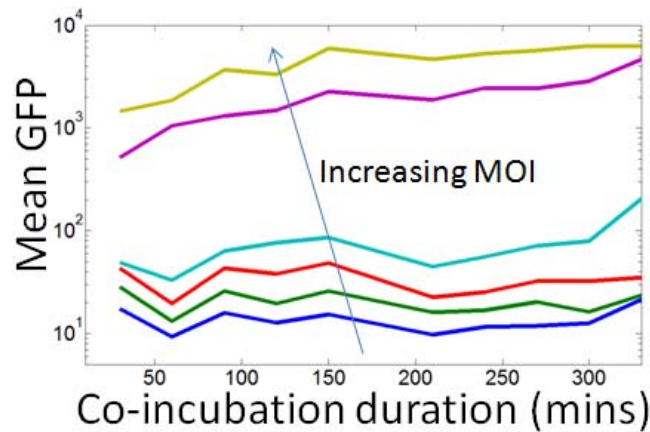
Figure A.6: Time-course plots. At sufficiently low MOIs (=10, 20, and 50), no significant increase in GFP is observed over time. At MOI=100, however, GFP signal became increasingly stronger over time after some time delay. At higher MOIs, GFP signal increased without time delay.

To investigate the cause of such sharp response to MOI, we observed the temporal dynamics of bacterial infection. Consistent with our dose response curves, our results in Figure A.6 also indicated MOI-dependent increase in bacterial uptake. At low MOI (≤ 50), the GFP signal was maintained at a low value over 5.5 hours. For an intermediate MOI (= 100), however, the GFP signal became stronger over time after some time delay. Interestingly, despite the increase in the GFP signal, we did not observe significant change in the percentage of infected host cells at the intermediate MOI (Figure A.7A). These results suggest that the increase in GFP signal is not because of more mammalian host cells being infected, but because of increasing GFP signal strength in the infected host cells. This increase in GFP was probably not because of bacterial growth or increasing GFP signal strength in bacteria due to the gentamicin

treatment. At the working concentration of this antibiotic, our plate reader

measurements showed that the drug treatment inhibited bacterial growth and reduced

GFP signal strength (Figure A.8). Moreover, our raw data in Figure A.7B showed that

the high mode of GFP at 90 minutes (red line in Figure A.7B) shifted to the right after

some time delay (330 mins, blue line in Figure A.7B), while the low mode of GFP did not

change significantly. These results together suggest that the infected host cells become

more prone to further bacterial uptake over time, while those without bacterial uptake

remains free of bacteria.
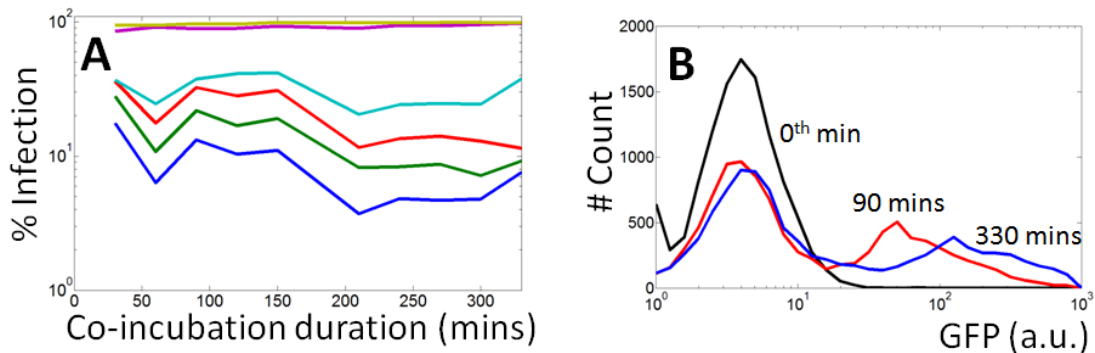
Figure A.7: Sustained percentage of infection. A) The percentage of the host cells infected with bacteria is plotted over co-incubation duration for varying MOIs. The GFP signal in a single bacterium allows for differentiation of the infected host cells from the non-infected host cells. B) GFP distributions at varying co-incubation durations for MOI=100.
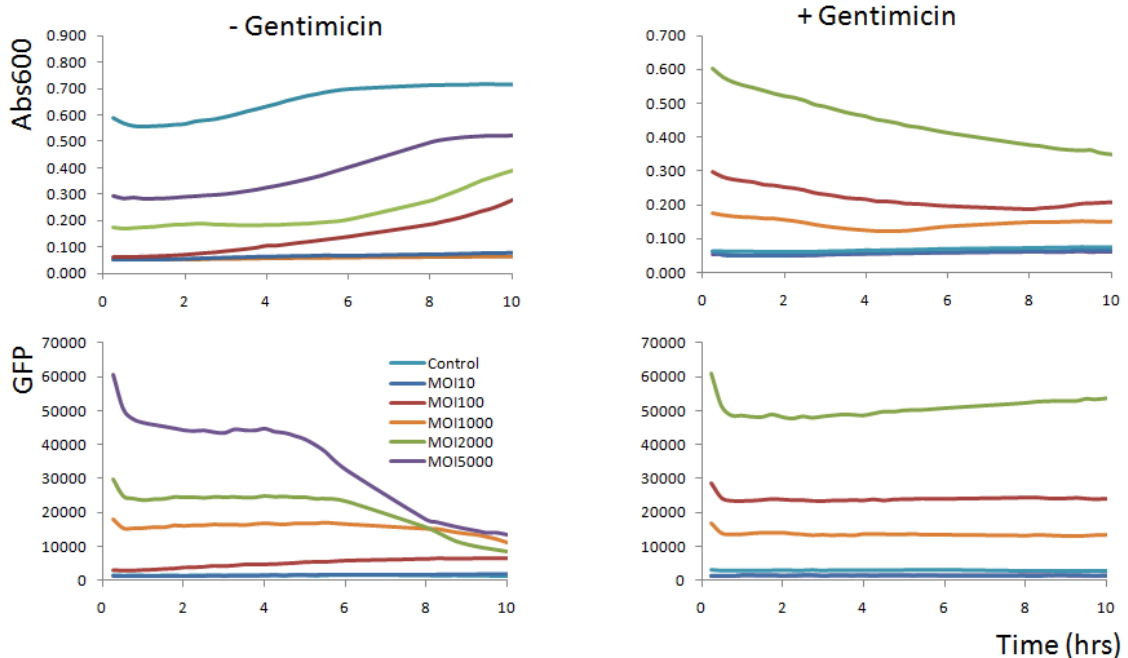
Figure A.8: The effects of gentimicin on bacterial growth. The engineered bacteria were grown overnight and diluted to varying concentrations in DMEM supplemented with 10% BGS in the presence or absence of gentimicin in a 96-well plate. The amount of bacteria in each well is similar to the infection conditions at varying MOIs.

A.3.3.3 The threshold effects observed across multiple cell lines

We hypothesized that the characteristic threshold effect may be an intrinsic

feature of the 'zipper mechanism' mediated by invasin-integrin interactions, and it

should be observed regardless of the host cell line or bacterial strain. To test our

hypothesis, we developed another bacterial strain that carries an invasin-encoding

plasmid *pSCT7Inv* and a reporter plasmid *ptetGFP* for quantification of bacterial uptake.

Under the same experimental conditions, we quantified bacterial uptake in multiple cell

lines (total of 8) at various MOIs. Consistent with our hypothesis, our results in Figure

A.9 showed threshold effects in most of the cell lines tested, suggesting that the positive

feedback regulation may be a ubiquitous mechanism in the bacterial uptake process.
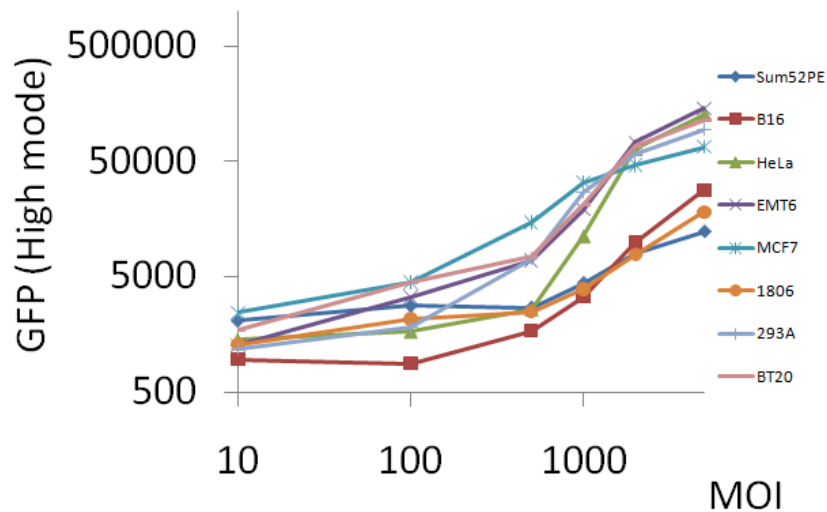
Figure A.9: Bacterial uptake for multiple host cells. Bacterial uptake was measured in multiple host cells for increasing MOIs. These host cells were subjected to varying MOIs under the same experimental conditions. After 2 hours of co-incubation, these cells were collected, fixed with formaldehyde, and assayed for their GFP signals.

A.3.3.4 A positive feedback regulation for multi-stage infection

The threshold effects observed in these experiments are characteristic of a

positive feedback regulation [46,338]. In fact, at least two positive feedback mechanisms

have been speculated in the recruitment of *β1-integrin* receptors at the bacterial binding

site: via receptor clustering and focal adhesion kinase (FAK) signaling. During bacterial

uptake, the initial bindings of invasins with *β1-integrins* facilitate receptor clustering

[339], constituting a positive feedback that increases the efficiency of bacteria binding. In

addition, these receptors directly interact with FAK signals. These FAK signals have been speculated to liberate *β1-integrins* from the focal adhesion sites [6,7], where *β1-integrins* serve as cell-adhesin molecules for anchoring the cell to the extracellular matrix. The liberated receptors become available for further binding of bacteria, constituting another positive feedback that increases receptor availability. Such positive feedback regulations may account for the two stages of bacterial uptake. The initial stage of bacterial uptake may correspond to when the positive feedback is not yet activated and the number of bound bacteria is small. The second stage may represent activated state of the positive feedback regulation when MOI is greater than the threshold, and the infected host cells are prone to further infection.

To test this possibility, we considered our mathematical model under two conditions: with and without a positive feedback. In either case, the positive dependence of bacterial uptake on MOI and co-incubation duration could be recapitulated with the model. However, our model could not generate the threshold effects without the positive feedback loop (data not shown). When the positive feedback loop was considered in the model, our model could generate the threshold effect, as shown in time plots (Figure A.10A) and dose responses (Figure A.10B).
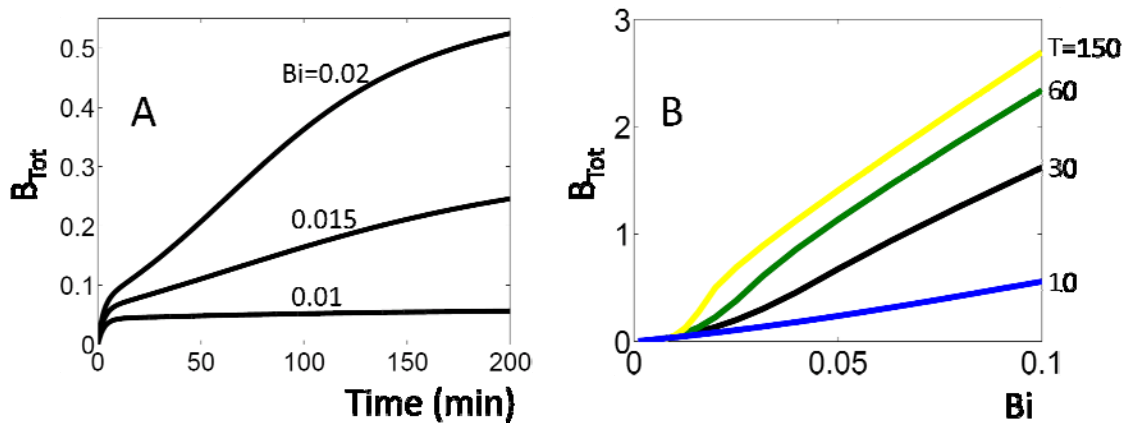
Figure A.10: Simulation results. A) Time-course simulations. Below a threshold level of B_i (≈ 0.01), the change in B_{Tot} (sum of initially bound bacteria B_1 and fully bound bacteria B_n) is expected to be small over time. Above the threshold, however, B_{Tot} is predicted to increase significantly over time. B) Threshold effects in the dose response. Simulations at various time points suggest a threshold of input B_i, below which bacterial uptake is insignificant and is independent of co-incubation durations.

A.3.3.5 Dependence of bacterial uptake on receptor concentration

Intuitively, bacterial binding would directly correlate with receptor concentrations. Such correlation has been demonstrated in a cancer cell line HEp-2 expressing different levels of *β-integrin* concentrations [340], and is consistent with our model predictions (Figure A.11). However, it is not clear whether this dependence would be maintained across different host cell lines, since different cell lines may have varying β1-integrin receptors and integrin binding affinities [341].
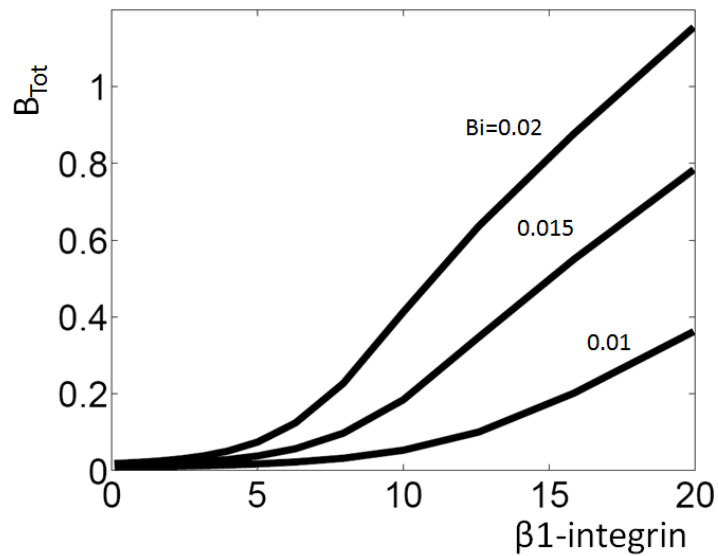
Figure A.11: Dependence of bacterial infection on receptor concentration.

Using our mathematical model, we can predict how variations in the model parameters may govern the correlation between the receptor concentration and the percentage of infection. To emulate variations in the model parameters, we assumed that a parameter is a Gaussian random variable with a variance of 1 and a mean being the basal value for the parameter. Our simulation results Figure A.12A suggest that the correlation between the receptor concentration and the percentage of infection is robust to variations in parameters.
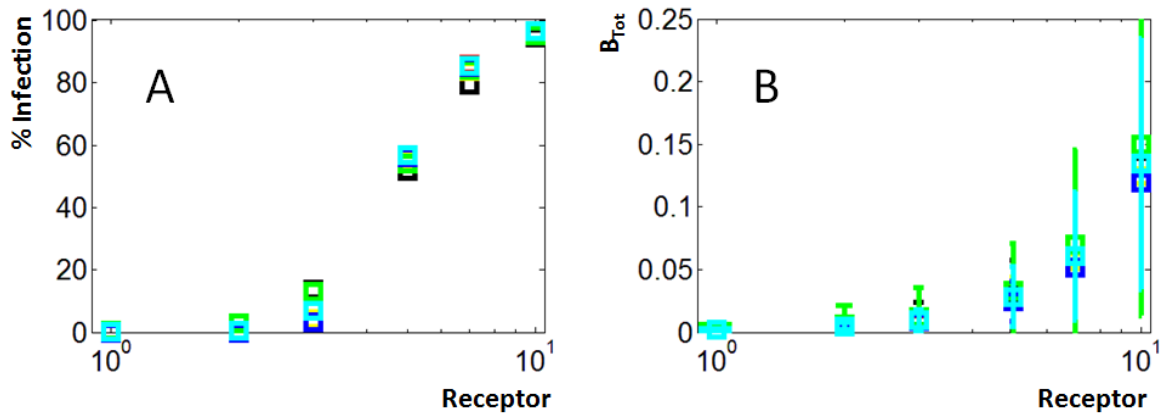
Figure A.12: Predicted correlations between *β1-integrin* level and bacterial uptake, and with the percentage of infection. A) Predicted correlation between *β1-integrin* level and the percentage of infection. B) Predicted correlation between *β1-integrin* level and bacterial uptake.

However, simulations in Figure A.12B suggest that bacterial uptake efficiency is highly sensitive to variations in some parameters. In particular, our model predicts the greatest sensitivity in the parameters that define receptor availability (k_{act} and k_{ina}) and receptor's binding affinity with invasins (k_r). In fact, it has been experimentally shown that receptors in different cell lines may have different binding affinities [340], which may lead to variable bacterial uptake. These simulation results suggest the sensitivity of bacterial uptake efficiency to various model parameters may mask the dependence of bacterial uptake efficiency on receptor concentration.
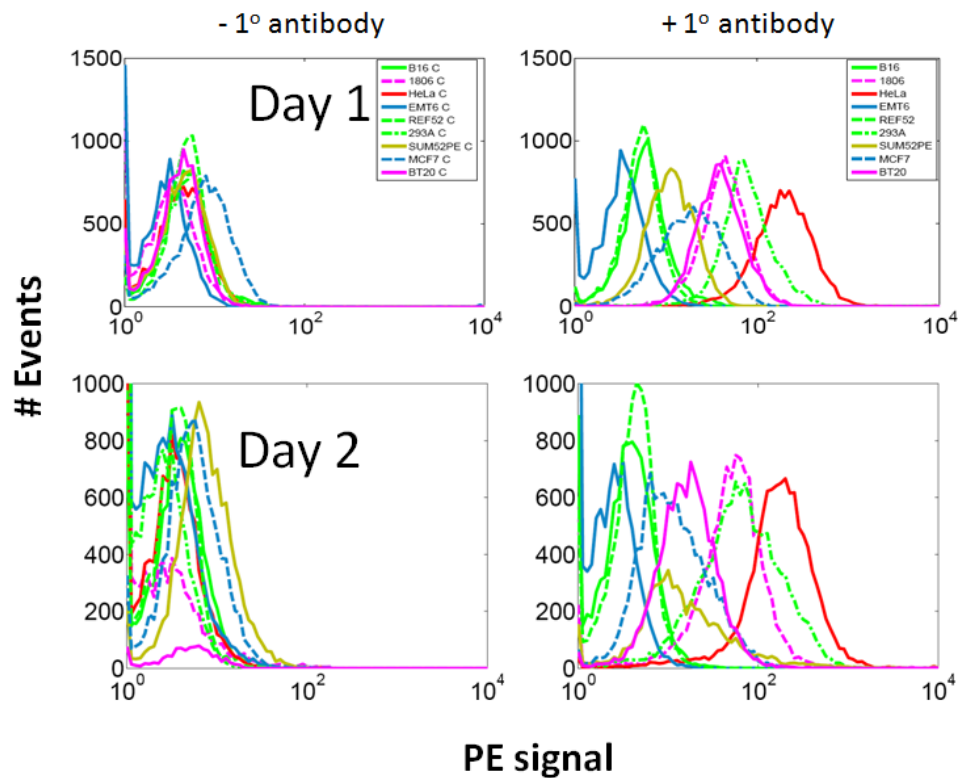
Figure A.13: Quantification of *β1-integrin* in multiple cell lines. *β1-integrin* was quantified with phycoerythrin (PE), immunofluorescent signals conjugated with the primary antibodies (Milipore, CBL479) for *β-integrin* receptors. Two independent measurements were performed in two different days.
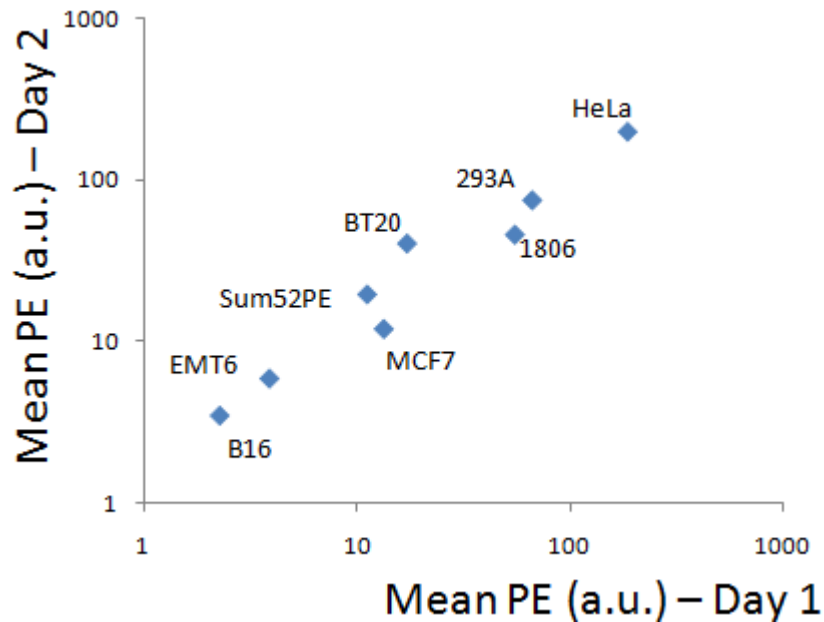
Figure A.14: Ordering of receptor levels in different cell lines. Two independent measurements of PE in various cell lines (each data point representing a cell line) were plotted against each other. Despite day-to-day variability in PE values, the order of the receptor levels in different cell lines did not change.

To test these predictions, we quantified *β-integrin* receptor concentrations in a set of mammalian cell lines that span a wide range of receptor concentrations (Figure A.13). These different cell lines represent variations in the model parameters. We used primary antibody against the *α5β1*-integrin receptors and immunofluorescent secondary antibody (phycoerythrin, PE), conjugated to the primary antibody. Under the same experimental conditions, we observed day-to-day variability in the level of PE signals, but the order of the receptor levels in different cell lines did not change significantly (Figure A.14). To check for correlation between percentage of infection and receptor concentration, we plotted the bacterial uptake efficiency in Figure A.9 against our β1-

integrin measurements in Figure A.14. Our results in Figure A.15A showed that the

correlation was not significant, suggesting that factors other than receptor concentration

may also play a role in bacterial uptake.
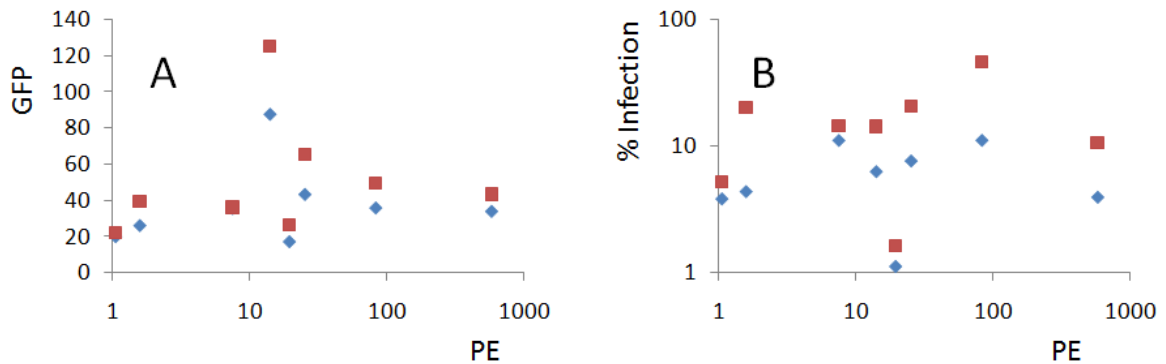
Figure A.15: Experimental correlations between receptor concentrations and bacterial uptake, and with percentage of infections. A) A correlation between bacterial uptake (represented by GFP) and β1-integrin (represented by PE) level in multiple cell lines. B) The percentage of host cells infected with bacteria in these host cell lines with different β1-integrin levels.

Despite the insignificant correlation between bacterial uptake efficiency and

receptor concentration, the percentage of infection was generally correlated with

receptor concentration, as shown in Figure A.15B.At a sub-saturating MOI (≤100), we

observed a group of cell lines that showed increasing percentage of bacterial uptake

with increasing receptor concentration (along an arbitrarily drawn black line in (Figure

A.15B). However, we also observed deviation from this black line in 1806 and HeLa cells,

suggesting that receptor concentration may not be the only determinant in governing

the percentage of infection. In fact, 1806 cells were more strongly attached to the plate

than others and required prolonged trypsinization, which might have removed attached

bacteria by cleaving the invasin-integrin interactions [342]. At sufficiently high MOI, no

significant variability was observed as most of the cells for a given cell line were infected

(data not shown). To test for reproducibility, we carried out another set of experiments

under the same experimental conditions on a different day. Our results in Figure A.16

showed similar correlation between the percentage of infection and receptor

concentration, but the correlation between bacterial uptake efficiency and receptor

concentration was more pronounced in this data set. We note that this is not inconsistent

with our model predictions in Figure A.12B, as we predicted significant variability in

bacterial uptake efficiency for variable model parameters. These modeling and

experimental results together suggest that our model can be useful in predicting further

bacterial infection dynamics.
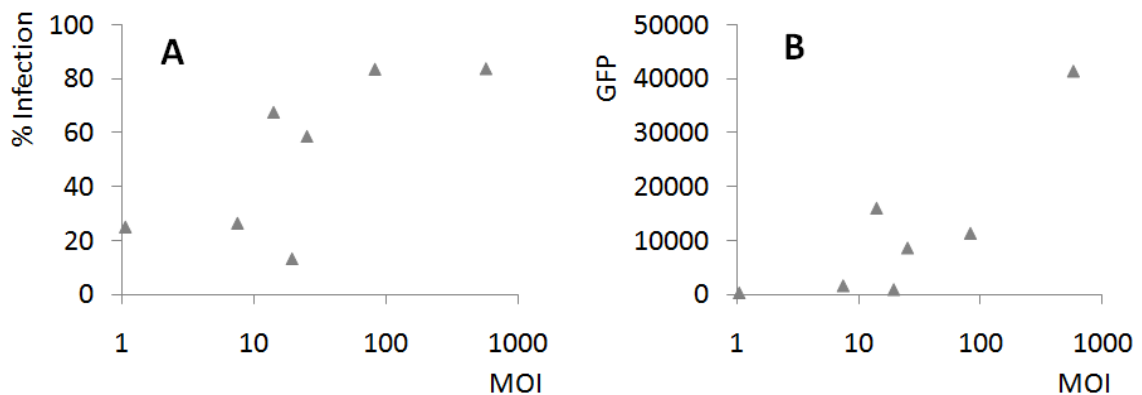
Figure A.16: Replicate experiments for correlations between receptor concentrations and bacterial uptake, and with percentage of infections

A.4 Discussion

In this work, we provide experimental and mathematical evidences for a positive feedback regulation, which may account for the variable bacterial uptake that leads to bimodal infection kinetics. In addition, we show that the positive feedback may be a ubiquitous regulation associated with the 'zipper mechanism' for bacterial uptake, as demonstrated by the dose responses in multiple host cell lines. Finally, we demonstrated our model's utility in making experimentally tractable predictions that were experimentally validated: robust correlation between receptor concentration and the percentage of bacterial infection to variable model parameters, but insignificant correlation between receptor concentration and bacterial uptake efficiency.

A mathematical framework may be beneficial in various applications that require a quantitative understanding of host-pathogen interactions. In fact, a number of predictive mathematical models for bacterial adherence have been developed [343,344,345,346,347], but the predicted dynamics have not been validated experimentally. Our quantitative framework in this work, consisting of a mathematical model and an experimental platform for model validation, may serve as a useful guide for modulating infection dynamics.

For example, our model predictions may provide insights for optimal intervention strategies in the development of anti-bacterial therapeutics. The traditional treatment of bacterial diseases has focused on eradication of the pathogens by antibiotics.

However, many pathogens have developed resistance to major antibiotics [348], necessitating alternative strategies for the resistant pathogens. One of these efforts is to reduce or eliminate host-pathogen interactions without incurring significant pressure for the emergence of resistant strains [349]. In this mode of treatment, our model may guide effective therapeutic strategies. For example, our model predictions in Figure A.12B suggest that the optimal strategy may be to limit receptor availability or to interfere with receptor's binding affinity with invasins.

In addition, the development of bacteria-mediated therapeutics may also benefit from a quantitative framework for bacterial infection. For the past two decades, increasing characterization of virulence factors has inspired emergence of genetically engineered bacteria as a delivery method [337], where the goal is to maximize delivery efficiency with minimal toxicity to the host cells. Our model predictions in Figure A.12B suggest that increasing receptor availability or increasing receptor's binding affinity may improve bacterial uptake efficiency.

References

1. Lee TJ, Tan CM, Tu D, You LC (2007) Modeling cellular networks. Bioinformatics: An Engineering Case-Based Approach Chapter 6.

2. Johnson L, Shen A, Boyle L, Kunich J, Pandey K, et al. (2002) Selectively replicating adenoviruses targeting deregulated E2F activity are potent, systemic antitumor agents. Cancer Cell 1: 325-337.

3. Khuri FR, Nemunaitis J, Ganly I, Arseneau J, Tannock IF, et al. (2000) a controlled trial of intratumoral ONYX-015, a selectively-replicating adenovirus, in combination with cisplatin and 5-fluorouracil in patients with recurrent head and neck cancer.[see comment]. Nature Medicine 6: 879-885.

4. Ideker T, Winslow L, Lauffenburger A (2006) Bioengineering and Systems Biology. Annals of Biomedical Engineering 34: 257.

5. Alves R, Antunes F, Salvador A (2006) Tools for kinetic modeling of biochemical networks. Nat Biotech 24: 667.

6. You L, Cox RS, 3rd, Weiss R, Arnold FH (2004) Programmed population control by cell-cell communication and regulated killing. Nature 428: 868-871.

7. Weston AD (2004) Systems biology, proteomics, and the future of health care: Toward predictive, preventative, and personalized medicine. Journal of proteome research 3: 179.

8. Neves SR (2002) Modeling of signaling networks. BioEssays 24: 1110.

9. Kitano H (2002) Computational systems biology. Nature 420: 206.

10. Rao CV, Arkin AP (2001) Control motifs for intracellular regulatory networks. Annual Review of Biomedical Engineering 3: 391-419.

11. Hasty J, McMillen D, Isaacs F, Collins JJ (2001) Computational studies of gene regulatory networks: in numero molecular biology. Nature Reviews Genetics 2: 268-279.

12. Endy D, Brent R (2001) Modelling cellular behaviour. Nature 409: 391-395.

13. Sasagawa S, Ozaki Y, Fujita K, Kuroda S (2005) Prediction and validation of the distinct dynamics of transient and sustained ERK activation. Nature Cell Biology 7: 365-U331.

14. Schoeberl B, Eichler-Jonsson C, Gilles ED, Muller G (2002) Computational modeling of the dynamics of the MAP kinase cascade activated by surface and internalized EGF receptors.[see comment]. Nature Biotechnology 20: 370-375.

15. Bhalla US, Ram PT, Iyengar R (2002) MAP kinase phosphatase as a locus of flexibility in a mitogen-activated protein kinase signaling network.[see comment]. Science 297: 1018-1023.

16. Asthagiri AR, Lauffenburger DA (2001) A computational study of feedback effects on signal dynamics in a mitogen-activated protein kinase (MAPK) pathway model. Biotechnology Progress 17: 227-239.

17. Bhalla US, Iyengar R (1999) Emergent properties of networks of biological signaling pathways.[see comment]. Science 283: 381-387.

18. Taniguchi CM, Emanuelli B, Kahn CR (2006) Critical nodes in signalling pathways: insights into insulin action. Nature Reviews Molecular Cell Biology 7: 85-96.

19. Somogyi R, Greller LD (2001) The dynamics of molecular networks: applications to therapeutic discovery. Drug Discovery Today 6: 1267.

20. Jackson TL, Byrne HM (2000) A mathematical model to study the effects of drug resistance and vasculature on the response of solid tumors to chemotherapy. Mathematical Biosciences 164: 17-38.

21. Butcher EC, Berg EL, Kunkel EJ (2004) Systems biology in drug discovery. Nature Biotechnology 22: 1253-1259.

22. Oda K, Matsuoka Y, Funahashi A, Kitano H (2005) A comprehensive pathway map of epidermal growth factor receptor signaling. Mol Syst Biol 1: msb4100014-E4100011.

23. Gardner TS, di Bernardo D, Lorenz D, Collins JJ (2003) Inferring genetic networks and identifying compound mode of action via expression profiling. Science 301: 102-105.

24. Butte AJ, Kohane IS (2000) Mutual information relevance networks: functional genomic clustering using pairwise entropy measurements. Pacific Symposium on Biocomputing: 418-429.

25. Friedman N, Linial M, Nachman I, Pe'er D (2000) Using Bayesian networks to analyze expression data. Journal of Computational Biology 7: 601-620.

26. Moles CG, Mendes P, Banga JR (2003) Parameter estimation in biochemical pathways: A comparison of global optimization methods. Genome Research 13: 2467-2474.

27. Arkin A, Shen P, Ross J (1997) A Test Case of Correlation Metric Construction of a Reaction Pathway from Measurements. Science 277: 1275-1279.

28. You L, Yin J (2000) Patterns of regulation from mRNA and protein time series. Metabolic Engineering 2: 210-217.

29. Ideker T, Thorsson V, Siegel AF, Hood LE (2000) Testing for differentially-expressed genes by maximum-likelihood analysis of microarray data. Journal of Computational Biology 7: 805-817.

30. Ronen M, Rosenberg R, Shraiman BI, Alon U (2002) Assigning numbers to the arrows: parameterizing a gene regulation network by using accurate expression kinetics. Proceedings of the National Academy of Sciences of the United States of America 99: 10555-10560.

31. Guido NJ, Wang X, Adalsteinsson D, McMillen D, Hasty J, et al. (2006) A bottom-up approach to gene regulation. Nature 439: 856-860.

32. Austin DW, Allen MS, McCollum JM, Dar RD, Wilgus JR, et al. (2006) Gene network shaping of inherent noise spectra. Nature 439: 608-611.

33. Elowitz MB, Levine AJ, Siggia ED, Swain PS (2002) Stochastic gene expression in a single cell. Science 297: 1183-1186.

34. Becskei A, Kaufmann BB, van Oudenaarden A (2005) Contributions of low molecule number and chromosomal positioning to stochastic gene expression.[see comment]. Nature Genetics 37: 937-944.

35. Pedraza JM, van Oudenaarden A (2005) Noise propagation in gene networks.[see comment]. Science 307: 1965-1969.

36. Rosenfeld N, Young JW, Alon U, Swain PS, Elowitz MB (2005) Gene regulation at the single-cell level. Science 307: 1962-1965.

37. Keseler IM, Collado-Vides J, Gama-Castro S, Ingraham J, Paley S, et al. (2005) EcoCyc: a comprehensive database resource for Escherichia coli. Nucleic Acids Res 33: D334-337.

38. Kanehisa M, Goto S, Hattori M, Aoki-Kinoshita KF, Itoh M, et al. (2006) From genomics to chemical genomics: new developments in KEGG. Nucleic Acids Res 34: D354-357.

39. Overbeek R, Larsen N, Walunas T, D'Souza M, Pusch G, et al. (2003) The ERGO genome analysis and discovery system. Nucleic Acids Res 31: 164-171.

40. Lemer C, Antezana E, Couche F, Fays F, Santolaria X, et al. (2004) The aMAZE LightBench: a web interface to a relational database of cellular processes. Nucleic Acids Res 32: D443-448.

41. Gasteiger E, Gattiker A, Hoogland C, Ivanyi I, Appel RD, et al. (2003) ExPASy: The proteomics server for in-depth protein knowledge and analysis. Nucleic Acids Res 31: 3784-3788.

42. Friedman N, Vardi S, Ronen M, Alon U, Stavans J (2005) Precise temporal modulation in the response of the SOS DNA repair network in individual bacteria. PLoS Biol 3: e238.

43. Wolf DM, Arkin AP (2003) Motifs, modules and games in bacteria. Current Opinion in Microbiology 6: 125-134.

44. Hartwell LH, Hopfield JJ, Leibler S, Murray AW (1999) From molecular to modular cell biology. Nature 402: C47-52.

45. Romond P-C, Rustici M, Gonze D, Goldbeter A (1999) Alternating Oscillations and Chaos in a Model of Two Coupled Biochemical Oscillators Driving Successive Phases of the Cell Cycle. Annals of the New York Academy of Sciences 879: 180-193.

46. Tyson JJ, Chen KC, Novak B (2003) Sniffers, buzzers, toggles and blinkers: dynamics of regulatory and signaling pathways in the cell. Current Opinion in Cell Biology 15: 221-231.

47. Batchelor E, Silhavy TJ, Goulian M (2004) Continuous control in bacterial regulatory circuits. Journal of Bacteriology 186: 7618-7625.

48. Becskei A, Serrano L (2000) Engineering stability in gene networks by autoregulation. Nature 405: 590-593.

49. Rosenfeld N, Elowitz MB, Alon U (2002) Negative autoregulation speeds the response times of transcription networks. Journal of Molecular Biology 323: 785-793.

50. Becskei A, Seraphin B, Serrano L (2001) Positive feedback in eukaryotic gene networks: cell differentiation by graded to binary response conversion. EMBO Journal 20: 2528-2535.

51. Kramer BP, Fussenegger M (2005) Hysteresis in a synthetic mammalian gene network. Proceedings of the National Academy of Sciences of the United States of America 102: 9517-9522.

52. Thron CD (1997) Bistable biochemical switching and the control of the events of the cell cycle. Oncogene 15: 317-325.

53. Yao G, Lee TJ, Mori S, Nevins JR, You LC (2008) A bistable Rb-E2F switch underlies the restriction point. Nature Cell Biology 10: 476-U255.

54. Acar M, Becskei A, van Oudenaarden A (2005) Enhancement of cellular memory by reducing stochastic transitions. Nature 435: 228-232.

55. Gardner TS, Cantor CR, Collins JJ (2000) Construction of a genetic toggle switch in Escherichia coli. Nature 403: 339-342.

56. Atkinson MR, Savageau MA, Myers JT, Ninfa AJ (2003) Development of genetic circuitry exhibiting toggle switch or oscillatory behavior in Escherichia coli. Cell 113: 597-607.

57. Elowitz MB, Leibler S (2000) A synthetic oscillatory network of transcriptional regulators. Nature 403: 335-338.

58. Fung E, Wong WW, Suen JK, Bulter T, Lee SG, et al. (2005) A synthetic gene-metabolic oscillator. Nature 435: 118-122.

59. Guantes R, Poyatos JF (2006) Dynamical Principles of Two-Component Genetic Oscillators. PLoS Computational Biology 2: e30.

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

61. Shen-Orr SS, Milo R, Mangan S, Alon U (2002) Network motifs in the transcriptional regulation network of Escherichia coli. Nature Genetics 31: 64-68.

62. Stoleru D, Peng Y, Nawathean P, Rosbash M (2005) A resetting signal between Drosophila pacemakers synchronizes morning and evening activity. Nature 438: 238.

63. Levine JD, Funes P, Dowse HB, Hall JC (2002) Signal analysis of behavioral and molecular cycles. BMC Neuroscience 3: 1.

64. Meinhardt H (1992) Pattern formation in biology: a comparison of models and experiments. Reports on Progress in Physics 55: 797.

65. Xiong W, Ferrell JE, Jr. (2003) A positive-feedback-based bistable 'memory module' that governs a cell fate decision.[see comment]. Nature 426: 460-465.

66. Tyson JJ, Novak B (2001) Regulation of the eukaryotic cell cycle: molecular antagonism, hysteresis, and irreversible transitions. Journal of Theoretical Biology 210: 249-263.

67. Tyson JJ, Novak B, Chen K, Val J (1995) Checkpoints in the cell cycle from a modeler's perspective. Progress in Cell Cycle Research 1: 1-8.

68. Pomerening JR, Sontag ED, Ferrell JE, Jr. (2003) Building a cell cycle oscillator: hysteresis and bistability in the activation of Cdc2. Nature Cell Biology 5: 346-351.

69. Bintu L, Buchler NE, Garcia HG, Gerland U, Hwa T, et al. (2005) Transcriptional regulation by the numbers: models. Current Opinion in Genetics & Development 15: 116-124.

70. Mathews JH, Fink KD (2004) Numerical methods using MATLAB. Upper Saddle River, N.J.: Pearson. ix, 680 p. p.

71. Atkinson KE (1989) An introduction to numerical analysis. New York: Wiley. xvi, 693 p. p.

72. Quarteroni A, Sacco R, Saleri F (2000) Numerical mathematics. New York: Springer. xx, 654 p. p.

73. Epperson JF (2002) An introduction to numerical methods and analysis. New York: J. Wiley. xv, 556 p. p.

74. Slepchenko BM, Schaff JC, Macara I, Loew LM (2003) Quantitative cell biology with the Virtual Cell. Trends in Cell Biology 13: 570-576.

75. Ramsey S, Orrell D, Bolouri H (2005) Dizzy: stochastic simulation of large-scale genetic regulatory networks. Journal of Bioinformatics & Computational Biology 3: 415-436.

76. You LC, Hoonlor A, Yin J (2003) Modeling biological systems using Dynetica - a simulator of dynamic networks. Bioinformatics 19: 435-436.

77. Dhar P, Meng TC, Somani S, Ye L, Sairam A, et al. (2004) Cellware--a multi-algorithmic software for computational systems biology. Bioinformatics 20: 1319-1321.

78. Hucka M, Finney A, Sauro HM, Bolouri H, Doyle JC, et al. (2003) The systems biology markup language (SBML): a medium for representation and exchange of biochemical network models. Bioinformatics 19: 524-531.

79. Mendes P (1997) Biochemistry by numbers: simulation of biochemical pathways with Gepasi 3. Trends in Biochemical Sciences 22: 361-363.

80. McAdams HH, Arkin A (1997) Stochastic mechanisms in gene expression. Proc Natl Acad Sci U S A 94: 814-819.

81. Arkin A, Ross J, McAdams HH (1998) Stochastic kinetic analysis of developmental pathway bifurcation in phage lambda-infected Escherichia coli cells. Genetics 149: 1633-1648.

82. Bar-Even A, Paulsson J, Maheshri N, Carmi M, O'Shea E, et al. (2006) Noise in protein expression scales with natural protein abundance. Nature Genetics 38: 636-643.

83. Hooshangi S, Thiberge S, Weiss R (2005) Ultrasensitivity and noise propagation in a synthetic transcriptional cascade. PNAS 102: 3581-3586.

84. Kaern M, Elston TC, Blake WJ, Collins JJ (2005) Stochasticity in gene expression: from theories to phenotypes. Nature Reviews Genetics 6: 451-464.

85. Raser JM, O'Shea EK (2004) Control of stochasticity in eukaryotic gene expression. Science 304: 1811-1814.

86. Blake WJ, Kaern M, Cantor CR, Collins JJ (2003) Noise in eukaryotic gene expression. Nature 422: 633-637.

87. Kaern M, Elston TC, Blake WJ, Collins JJ (2005) Stochasticity in gene expression: From theories to phenotypes. Nature Reviews Genetics 6: 451-464.

88. Rao CV, Wolf DM, Arkin AP (2002) Control, exploitation and tolerance of intracellular noise. Nature 420: 231-237.

89. Savageau MA (1974) Comparison of classical and autogenous systems of regulation in inducible operons. Nature 252: 546-549.

213

90. Weinberger LS, Burnett JC, Toettcher JE, Arkin AP, Schaffer DV (2005) Stochastic gene expression in a lentiviral positive-feedback loop: HIV-1 Tat fluctuations drive phenotypic diversity. Cell 122: 169-182.

91. Suel GM, Garcia-Ojalvo J, Liberman LM, Elowitz MB (2006) An excitable gene regulatory circuit induces transient cellular differentiation. Nature 440: 545-550.

92. Gillespie DT (1992) A rigorous derivation of the chemical master equation. Physica A: Statistical and Theoretical Physics 188: 404.

93. Gillespie DT (1977) Exact stochastic simulation of coupled chemical reactions. The Journal of physical chemistry 81: 2340.

94. Gibson MA, Bruck J (2000) Efficient exact stochastic simulation of chemical systems with many species and many reactions. Journal of Chemical Physics 104: 1876-1889.

95. Cao Y, Gillespie DT, Petzold LR (2006) Efficient step size selection for the tau-leaping simulation method. J Chem Phys 124: 044109.

96. Haseltine EL, Rawlings JB (2002) Approximate simulation of coupled fast and slow reactions stochastic chemical kinetics. Journal of Chemical Physics 117: 6959-6969.

97. Gillespie DT (2000) The Chemical Langevin Equation. Journal of Chemical Physics 113: 297-306.

98. Ozbudak EM, Thattai M, Kurtser I, Grossman AD, van Oudenaarden A (2002) Regulation of noise in the expression of a single gene. Nat Genet 31: 69-73.

99. Fell DA (1992) Metabolic control analysis: a survey of its theoretical and experimental development. Biochemical Journal 286: 313-330.

100. Varma A, Morbidelli M, Wu H (1999) Parametric sensitivity in chemical systems. Cambridge, U.K. ; New York, NY: Cambridge University Press. xvi, 342 p. p.

101. Morohashi M, Winn AE, Borisuk MT, Bolouri H, Doyle J, et al. (2002) Robustness as a measure of plausibility in models of biochemical networks. Journal of Theoretical Biology 216: 19-30.

102. Barkai N, Leibler S (1997) Robustness in simple biochemical networks.[see comment]. Nature 387: 913-917.

103. You L, Yin J (2002) Dependence of epistasis on environment and mutation severity as revealed by in silico mutagenesis of phage t7. Genetics 160: 1273-1281.

104. You L, Yin J (2006) Evolutionary design on a budget: robustness and optimality of bacteriophage T7. IEE Proceedings - Systems Biology 153: 46-52.

105. Alon U, Surette MG, Barkai N, Leibler S (1999) Robustness in bacterial chemotaxis.[see comment]. Nature 397: 168-171.

106. Freeman M (2000) Feedback control of intercellular signalling in development. Nature 408: 313.

107. Csete ME, Doyle JC (2002) Reverse Engineering of Biological Complexity. Science 295: 1664-1669.

108. Carlson JM, Doyle J (2002) Complexity and robustness. PNAS 99: 2538-2545.

109. Murray JD (1993) Mathematical biology. Berlin ; New York: Springer-Verlag. xiv, 767 p. p.

110. Strogatz SH (1994) Nonlinear dynamics and Chaos : with applications to physics, biology, chemistry, and engineering. Reading, Mass.: Addison-Wesley Pub. xi, 498 p. p.

111. You L, Hoonlor A, Yin J (2003) Modeling biological systems using Dynetica--a simulator of dynamic networks. Bioinformatics 19: 435-436.

112. Goldbeter A, Koshland DE, Jr. (1981) An amplified sensitivity arising from covalent modification in biological systems. Proceedings of the National Academy of Sciences of the United States of America 78: 6840-6844.

113. Huang CY, Ferrell JE, Jr. (1996) Ultrasensitivity in the mitogen-activated protein kinase cascade. Proceedings of the National Academy of Sciences of the United States of America 93: 10078-10083.

114. Ferrell JJE (1996) Tripping the switch fantastic: how a protein kinase cascade can convert graded inputs into switch-like outputs. Trends in Biochemical Sciences 21: 460.

115. Tu D, Lee J, Ozdere T, Lee TJ, You L (2006) Engineering Gene Circuits: Foundations and applications. Nanotechnology in Biotechnology and Medicine.

116. Church GM (2005) From systems biology to synthetic biology. Mol Syst Biol 1: msb4100007-E4100001.

117. Hasty J, McMillen D, Collins JJ (2002) Engineered gene circuits. Nature 420: 224-230.

118. Andrianantoandro E, Basu S, Karig DK, Weiss R (2006) Synthetic biology: new engineering rules for an emerging discipline. Mol Syst Biol 2: E1.

119. Colman-Lerner A, Gordon A, Serra E, Chin T, Resnekov O, et al. (2005) Regulated cell-to-cell variation in a cell-fate decision system. Nature 437: 699.

120. Endy D (2005) Foundations for engineering biology. Nature 438: 449.

121. Balagadde FK, You L, Hansen CL, Arnold FH, Quake SR (2005) Long-term monitoring of bacteria undergoing programmed population control in a microchemostat. Science 309: 137-140.

122. Doedel EJ (1983) AUTO: A program for the automatic bifurcation analysis of autonomous systems. Dynamics 38: 1493.

123. Vermeulen K, Van Bockstaele DR, Berneman ZN (2003) The cell cycle: a review of regulation, deregulation and therapeutic targets in cancer. Cell Proliferation 36: 131-149.

124. Morgan DO (1997) Cyclin-dependent kinases: engines, clocks, and microprocessors. Annu Rev Cell Dev Biol 13: 261-291.

125. Satyanarayana A, Kaldis P (2009) Mammalian cell-cycle regulation: several Cdks, numerous cyclins and diverse compensatory mechanisms. Oncogene 28: 2925-2939.

126. Malumbres M, Barbacid M (2009) Cell cycle, Cdks and cancer: a changing paradigm. Nature Reviews Cancer 9: 153-166.

127. Sherr CJ, Roberts JM (1999) Cdk inhibitors: positive and negative regulators of G1-phase progression. Genes & Development 13: 1501-1512.

128. Weinberg RA (1995) The Retinoblastoma Protein and Cell-Cycle Control. Cell 81: 323-330.

129. Lundberg AS, Weinberg RA (1998) Functional inactivation of the retinoblastoma protein requires sequential modification by at least two distinct cyclin-cdk complexes. Molecular and Cellular Biology 18: 753-761.

130. Hwang HC, Clurman BE (2005) Cyclin E in normal and neoplastic cell cycles. Oncogene 24: 2776-2786.

131. Furuno N, den Elzen N, Pines J (1999) Human cyclin A is required for mitosis until mid prophase. Journal of Cell Biology 147: 295-306.

132. Riabowol K, Draetta G, Brizuela L, Vandre D, Beach D (1989) The Cdc2 Kinase Is a Nuclear-Protein That Is Essential for Mitosis in Mammalian-Cells. Cell 57: 393-401.

133. Vidal A, Koff A (2000) Cell-cycle inhibitors: three families united by a common cause. Gene 247: 1-15.

134. Pavletich NP (1999) Mechanisms of cyclin-dependent kinase regulation: structures of Cdks, their cyclin activators, and Cip and INK4 inhibitors. J Mol Biol 287: 821-828.

135. Wolfel T, Hauer M, Schneider J, Serrano M, Wolfel C, et al. (1995) A p16INK4a-insensitive Cdk4 mutant targeted by cytolytic T lymphocytes in a human melanoma. Science 269: 1281-1284.

136. Kim JH, Kang MJ, Park CU, Kwak HJ, Hwang Y, et al. (1999) Amplified Cdk2 and cdc2 activities in primary colorectal carcinoma. Cancer 85: 546-553.

137. Yamamoto H, Monden T, Miyoshi H, Izawa H, Ikeda K, et al. (1998) Cdk2/cdc2 expression in colon carcinogenesis and effects of cdk2/cdc2 inhibitor in colon cancer cells. Int J Oncol 13: 233-239.

138. Easton J, Wei T, Lahti JM, Kidd VJ (1998) Disruption of the cyclin D/cyclin-dependent kinase/INK4/retinoblastoma protein regulatory pathway in human neuroblastoma. Cancer Res 58: 2624-2632.

139. Hall M, Peters G (1996) Genetic alterations of cyclins, cyclin-dependent kinases, and Cdk inhibitors in human cancer. Adv Cancer Res 68: 67-108.

140. Dobashi Y, Shoji M, Jiang SX, Kobayashi M, Kawakubo Y, et al. (1998) Active cyclin A-Cdk2 complex, a possible critical factor for cell proliferation in human primary lung carcinomas. Am J Pathol 153: 963-972.

141. Kamb A (1998) Cyclin-dependent kinase inhibitors and human cancer. Curr Top Microbiol Immunol 227: 139-148.

142. Horsthemke B (1992) Genetics and cytogenetics of retinoblastoma. Cancer Genet Cytogenet 63: 1-7.

143. Kastan MB, Bartek J (2004) Cell-cycle checkpoints and cancer. Nature 432: 316-323.

144. Doll R, Peto R (1981) The Causes of Cancer - Quantitative Estimates of Avoidable Risks of Cancer in the United-States Today. Journal of the National Cancer Institute 66: 1191-&.

145. Zou L, Elledge SJ (2003) Sensing DNA damage through ATRIP recognition of RPA-ssDNA complexes. Science 300: 1542-1548.

146. Bartek J, Lukas J (2003) Chk1 and Chk2 kinases in checkpoint control and cancer. Cancer Cell 3: 421-429.

147. Sears RC, Nevins JR (2002) Signaling networks that link cell proliferation and cell fate. Journal of Biological Chemistry 277: 11617-11620.

148. Mailand N, Falck J, Lukas C, Syljuasen RG, Welcker M, et al. (2000) Rapid destruction of human Cdc25A in response to DNA damage. Science 288: 1425-1429.

149. Falck J, Mailand N, Syljuasen RG, Bartek J, Lukas J (2001) The ATM-Chk2-Cdc25A checkpoint pathway guards against radioresistant DNA synthesis. Nature 410: 842-847.

150. Lukas J, Lukas C, Bartek J (2004) Mammalian cell cycle checkpoints: signalling pathways and their organization in space and time. DNA Repair 3: 997-1007.

151. Vogelstein B, Lane D, Levine AJ (2000) Surfing the p53 network. Nature 408: 307-310.

152. D'Amours D, Jackson SP (2002) The Mre11 complex: At the crossroads of DNA repair and checkpoint signalling. Nature Reviews Molecular Cell Biology 3: 317-327.

153. Kim ST, Xu B, Kastan MB (2002) Involvement of the cohesin protein, Smc1, in Atm-dependent and independent responses to DNA damage. Genes & Development 16: 560-570.

154. Maddika S, Ande SR, Panigrahi S, Paranjothy T, Weglarczyk K, et al. (2007) Cell survival, cell death and cell cycle pathways are interconnected: Implications for cancer therapy. Drug Resistance Updates 10: 13-29.

155. Aylon Y, Oren M (2007) Living with p53, dying of p53. Cell 130: 597-600.

156. Das S, Raj L, Zhao B, Kimura Y, Bernstein A, et al. (2007) Hzf determines cell survival upon genotoxic stress by modulating p53 transactivation. Cell 130: 624-637.

157. Tanaka T, Ohkubo S, Tatsuno I, Prives C (2007) hCAS/CSE1L associates with chromatin and regulates expression of select p53 target genes. Cell 130: 638-650.

158. Inga A, Storici F, Darden TA, Resnick MA (2002) Differential transactivation by the p53 transcription factor is highly dependent on p53 level and promoter target sequence. Molecular and Cellular Biology 22: 8612-8625.

159. Laptenko O, Prives C (2006) Transcriptional regulation by p53: one protein, many possibilities. Cell Death and Differentiation 13: 951-961.

160. Los M, Vandecraen M, Penning LC, Schenk H, Westendorp M, et al. (1995) Requirement of an Ice/Ced-3 Protease for Fas/Apo-1-Mediated Apoptosis. Nature 375: 81-83.

161. Maddika S, Mendoza FJ, Hauff K, Zamzow CR, Paranjothy T, et al. (2006) Cancer-selective therapy of the future - Apoptin and its mechanism of action. Cancer Biology & Therapy 5: 10-19.

162. Hengartner MO (2000) The biochemistry of apoptosis. Nature 407: 770-776.

163. Finkel E (2001) The mitochondrion: Is it central to apoptosis? Science 292: 624-626.

164. Barczyk K, Kreuter M, Pryjma J, Booy EP, Maddika S, et al. (2005) Serum cytochrome c indicates in vivo apoptosis and can serve as a prognostic marker during cancer therapy. International Journal of Cancer 116: 167-173.

165. Yin CY, Knudson CM, Korsmeyer SJ, VanDyke T (1997) Bax suppresses tumorigenesis and stimulates apoptosis in vivo. Nature 385: 637-640.

166. Miyashita T, Reed JC (1995) Tumor-Suppressor P53 Is a Direct Transcriptional Activator of the Human Bax Gene. Cell 80: 293-299.

167. Haaskogan DA, Kogan SC, Levi D, Dazin P, Tang A, et al. (1995) Inhibition of Apoptosis by the Retinoblastoma Gene-Product. Embo Journal 14: 461-472.

168. Janicke RU, Walker PA, Lin XY, Porter AG (1996) Specific cleavage of the retinoblastoma protein by an ICE-like protease in apoptosis. Embo Journal 15: 6969-6978.

169. Pardee AB (1974) A restriction point for control of normal animal cell proliferation. Proceedings of the National Academy of Sciences of the United States of America 71: 1286-1290.

170. Pardee AB (1989) G1 events and regulation of cell proliferation. Science 246: 603-608.

171. Zetterberg A, Larsson O, Wiman KG (1995) What Is the Restriction Point. Current Opinion in Cell Biology 7: 835-842.

172. Zetterberg A, Larsson O (1985) Kinetic-Analysis of Regulatory Events in G1 Leading to Proliferation or Quiescence of Swiss 3t3 Cells. Proceedings of the National Academy of Sciences of the United States of America 82: 5365-5369.

221

173. Malumbres M, Barbacid M (2001) To cycle or not to cycle: a critical decision in cancer. Nature Reviews Cancer 1: 222-231.

174. Cooper S (2003) Reappraisal of serum starvation, the restriction point, G0, and G1 phase arrest points. Faseb Journal 17: 333-340.

175. Novak B, Tyson JJ (2004) A model for restriction point control of the mammalian cell cycle. Journal of Theoretical Biology 230: 563-579.

176. Hatzimanikatis V, Lee KH, Bailey JE (1999) A mathematical description of regulation of the G1-S transition of the mammalian cell cycle. Biotechnology & Bioengineering 65: 631-637.

177. Aguda BD, Tang Y (1999) The kinetic origins of the restriction point in the mammalian cell cycle. Cell Proliferation 32: 321-335.

178. Aguda BD (2001) Kick-starting the cell cycle: From growth-factor stimulation to initiation of DNA replication. Chaos 11: 269-276.

179. Cossart P, Vicente MF, Mengaud J, Baquero F, Perez-Diaz JC, et al. (1989) Listeriolysin O is essential for virulence of Listeria monocytogenes: direct evidence obtained by gene complementation. Infect Immun 57: 3629-3636.

180. Obeyesekere MN, Knudsen ES, Wang JY, Zimmerman SO (1997) A mathematical model of the regulation of the G1 phase of Rb+/+ and Rb-/- mouse embryonic fibroblasts and an osteosarcoma cell line. Cell Prolif 30: 171-194.

181. Kohn KW (1998) Functional capabilities of molecular network components controlling the mammalian G1/S cell cycle phase transition. Oncogene 16: 1065-1075.

182. Swat M, Kel A, Herzel H (2004) Bifurcation analysis of the regulatory modules of the mammalian G1/S transition. Bioinformatics 20: 1506-1511.

183. Qu ZL, MacLellan WR, Weiss JN (2003) Dynamics of the cell cycle: Checkpoints, sizers, and timers. Biophysical Journal 85: 3600-3611.

184. Ekholm SV, Zickert P, Reed SI, Zetterberg A (2001) Accumulation of cyclin E is not a prerequisite for passage through the restriction point. Mol Cell Biol 21: 3256-3265.

185. Geng Y, Yu Q, Sicinska E, Das M, Schneider JE, et al. (2003) Cyclin E ablation in the mouse. Cell 114: 431-443.

186. Chen KC, Calzone L, Csikasz-Nagy A, Cross FR, Novak B, et al. (2004) Integrative analysis of cell cycle control in budding yeast. Mol Biol Cell 15: 3841-3862.

187. Chen KC, Csikasz-Nagy A, Gyorffy B, Val J, Novak B, et al. (2000) Kinetic analysis of a molecular model of the budding yeast cell cycle. Mol Biol Cell 11: 369-391.

188. Novak B, Csikasz-Nagy A, Gyorffy B, Chen K, Tyson JJ (1998) Mathematical model of the fission yeast cell cycle with checkpoint controls at the G1/S, G2/M and metaphase/anaphase transitions. Biophys Chem 72: 185-200.

189. Kong LJ, Chang JT, Bild AH, Nevins JR (2006) Compensation and specificity of function within the E2F family. Oncogene 26: 321-327.

190. Zhu W, Giangrande PH, Nevins JR (2004) E2Fs link the control of G1/S and G2/M transcription. Embo J 23: 4615-4626.

191. Johnson DG, Schneider-Broussard R (1998) Role of E2F in cell cycle control and cancer. Front Biosci 3: d447-448.

192. Lavia P, Jansen-Durr P (1999) E2F target genes and cell-cycle checkpoint control. Bioessays 21: 221-230.

193. Stevens C, La Thangue NB (2003) E2F and cell cycle control: a double-edged sword. Arch Biochem Biophys 412: 157-169.

223

194. Bell LA, Ryan KM (2004) Life and death decisions by E2F-1. Cell Death Differ 11: 137-142.

195. Attwooll C, Lazzerini Denchi E, Helin K (2004) The E2F family: specific functions and overlapping interests. Embo J 23: 4709-4716.

196. Leone G, DeGregori J, Jakoi L, Cook JG, Nevins JR (1999) Collaborative role of E2F transcriptional activity and G1 cyclindependent kinase activity in the induction of S phase. Proc Natl Acad Sci U S A 96: 6626-6631.

197. Dyson N (1998) The regulation of E2F by pRB-family proteins. Genes & Development 12: 2245-2262.

198. Doedel EJ (1981) AUTO, a program for the automatic bifurcation analysis of autonomous systems. Cong Numer 30: 265-384.

199. Ermentrout B (2002) Simulating, Analyzing, and Animating Dynamical Systems. SIAM Press, Philadelphia.

200. Yeh E, Cunningham M, Arnold H, Chasse D, Monteith T, et al. (2004) A signalling pathway controlling c-Myc degradation that impacts oncogenic transformation of human cells. Nature Cell Biology 6: 308-318.

201. Sears R, Leone G, DeGregori J, Nevins JR (1999) Ras enhances Myc protein stability. Molecular Cell 3: 169-179.

202. Sears R, Nuckolls F, Haura E, Taya Y, Tamai K, et al. (2000) Multiple Ras-dependent phosphorylation pathways regulate Myc protein stability. Genes & Development 14: 2501-2514.

203. Helin K (1998) Regulation of cell proliferation by the E2F transcription factors. Current Opinion in Genetics & Development 8: 28-35.

204. Diehl JA, Zindy F, Sherr CJ (1997) Inhibition of cyclin D1 phosphorylation on threonine-286 prevents its rapid degradation via the ubiquitin-proteasome pathway. Genes & Development 11: 957-972.

205. Clurman BE, Sheaff RJ, Thress K, Groudine M, Roberts JM (1996) Turnover of cyclin E by the ubiquitin-proteasome pathway is regulated by cdk2 binding and cyclin phosphorylation. Genes & Development 10: 1979-1990.

206. Won KA, Reed SI (1996) Activation of cyclin E/Cdk2 is coupled to site-specific autophosphorylation and ubiquitin-dependent degradation of cyclin E. EMBO Journal 15: 4182-4193.

207. Mihara K, Cao XR, Yen A, Chandler S, Driscoll B, et al. (1989) Cell cycle-dependent regulation of phosphorylation of the human retinoblastoma gene product. Science 246: 1300-1303.

208. Buchler NE, Gerland U, Hwa T (2005) Nonlinear protein degradation and the function of genetic circuits. PNAS 102: 9559-9564.

209. Kholodenko BN (2006) Cell-signalling dynamics in time and space. Nature Reviews Molecular Cell Biology 7: 165-176.

210. Park S, Chung S, Kim K-M, Jung K-C, Park C, et al. (2004) Determination of binding constant of transcription factor myc-max/max-max and E-box DNA: the effect of inhibitors on the binding. Biochimica et Biophysica Acta (BBA) - General Subjects 1670: 217-228.

211. Grafstrom RH, Pan W, Hoess RH (1999) Defining the substrate specificity of cdk4 kinase-cyclin D1 complex. Carcinogenesis 20: 193-198.

212. Grillo M, Bott M, Khandke N, McGinnis J, Miranda M, et al. (2006) Validation of cyclin D1/Cdk4 as an anticancer drug target in MCF-7 breast cancer cells: Effect of regulated overexpression of cyclin D1 and siRNA-mediated inhibition of endogenous cyclin D1 and Cdk4 expression. Breast Cancer Research and Treatment 95: 185-194.

213. Ozbudak EM, Thattai M, Lim HN, Shraiman BI, Van Oudenaarden A (2004) Multistability in the lactose utilization network of Escherichia coli. Nature 427: 737-740.

214. Longo D, Hasty J (2006) Dynamics of single-cell gene expression. Mol Syst Biol 2: 64.

215. Dannenberg JH, van Rossum A, Schuijff L, Riele HT (2000) Ablation of the Retinoblastoma gene family deregulates G(1) control causing immortalization and increased cell turnover under growth-restricting conditions. Genes & Development 14: 3051-3064.

216. Sage J, Mulligan GJ, Attardi LD, Miller A, Chen SQ, et al. (2000) Targeted disruption of the three Rb-related genes leads to loss of G(1) control and immortalization. Genes & Development 14: 3037-3050.

217. Wu LZ, Timmers C, Maiti B, Saavedra HI, Sang L, et al. (2001) The E2F1-3 transcription factors are essential for cellular proliferation. Nature 414: 457-462.

218. Johnson DG, Schwarz JK, Cress WD, Nevins JR (1993) Expression of Transcription Factor E2f1 Induces Quiescent Cells to Enter S-Phase. Nature 365: 349-352.

219. Nachman I, Regev A, Ramanathan S (2007) Dissecting timing variability in yeast meiosis. Cell 131: 544-556.

220. Spencer SL, Gaudet S, Albeck JG, Burke JM, Sorger PK (2009) Non-genetic origins of cell-to-cell variability in TRAIL-induced apoptosis. Nature 459: 428-432.

221. Bean JM, Siggia ED, Cross FR (2006) Coherence and timing of cell cycle start examined at single-cell resolution. Mol Cell 21: 3-14.

222. Colman-Lerner A, Gordon A, Serra E, Chin T, Resnekov O, et al. (2005) Regulated cell-to-cell variation in a cell-fate decision system. Nature 437: 699-706.

223. Brooks RF (1979) Cytoplasmic Origin of Variability in the Timing of S-Phase in Mammalian-Cells. Cell Biology International Reports 3: 707-716.

224. Brooks RF (1976) Regulation of fibroblast cell cycle by serum. Nature 260: 248-250.

225. Smith JA (1982) The Cell-Cycle and Related Concepts in Cell-Proliferation. Journal of Pathology 136: 149-166.

226. Shields R, Smith JA (1977) Cells Regulate Their Proliferation through Alterations in Transition-Probability. Journal of Cellular Physiology 91: 345-355.

227. Shields R (1978) Further Evidence for a Random Transition in Cell-Cycle. Nature 273: 755-758.

228. Shields R (1977) Transition-Probability and Origin of Variation in Cell-Cycle. Nature 267: 704-707.

229. Brooks RF, Bennett DC, Smith JA (1980) Mammalian-Cell Cycles Need 2 Random Transitions. Cell 19: 493-504.

230. Smith JA, Martin L (1973) Do Cells Cycle. Proceedings of the National Academy of Sciences of the United States of America 70: 1263-1267.

231. Castor LN (1980) A G1 Rate Model Accounts for Cell-Cycle Kinetics Attributed to Transition-Probability. Nature 287: 857-859.

232. Cooper S (1987) On G0 and Cell-Cycle Controls. Bioessays 7: 220-222.

233. Koch AL (1999) The re-incarnation, re-interpretation and re-demise of the transition probability model. Journal of Biotechnology 71: 143-156.

234. Koch AL (1980) Does the variability of the cell cycle result from one or many chance events? Nature 286: 80-82.

235. Tyson JJ, Hannsgen KB (1986) Cell growth and division: a deterministic/probabilistic model of the cell cycle. J Math Biol 23: 231-246.

236. Tyson JJ, Hannsgen KB (1985) Global asymptotic stability of the size distribution in probabilistic models of the cell cycle. J Math Biol 22: 61-68.

237. Nurse P (1980) Cell-Cycle Control - Both Deterministic and Probabilistic. Nature 286: 9-10.

238. Srivastava R, You L, Summers J, Yin J (2002) Stochastic vs. Deterministic Modeling of Intracellular Viral Kinetics. Journal of Theoretical Biology 218: 309-321.

239. Vellela M, Qian H (2009) Stochastic dynamics and non-equilibrium thermodynamics of a bistable chemical system: the Schlogl model revisited. Journal of the Royal Society Interface 6: 925-940.

240. Wang X, Hao N, Dohlman HG, Elston TC (2006) Bistability, stochasticity, and oscillations in the mitogen-activated protein kinase cascade. Biophysical Journal 90: 1961-1978.

241. Brooks RF, Richmond FN, Riddle PN, Richmond KMV (1984) Apparent Heterogeneity in the Response of Quiescent Swiss 3t3 Cells to Serum Growth-Factors - Implications for the Transition-Probability Model and Parallels with Cellular Senescence and Competence. Journal of Cellular Physiology 121: 341-350.

242. Brooks RF, Riddle PN (1988) The 3t3 Cell-Cycle at Low Proliferation Rates. Journal of Cell Science 90: 601-612.

243. Hoffmann A, Levchenko A, Scott ML, Baltimore D (2002) The I kappa B-NF-kappa B signaling module: Temporal control and selective gene activation. Science 298: 1241-1245.

244. Skotheim JM, Di Talia S, Siggia ED, Cross FR (2008) Positive feedback of G1 cyclins ensures coherent cell cycle entry. Nature 454: 291-U212.

245. Nevins JR (2001) The Rb/E2F pathway and cancer. Human Molecular Genetics 10: 699-703.

246. Golub TR, Slonim DK, Tamayo P, Huard C, Gaasenbeek M, et al. (1999) Molecular classification of cancer: Class discovery and class prediction by gene expression monitoring. Science 286: 531-537.

247. Bild AH, Yao G, Chang JT, Wang QL, Potti A, et al. (2006) Oncogenic pathway signatures in human cancers as a guide to targeted therapies. Nature 439: 353-357.

248. Pelengaris S, Khan M, Evan G (2002) c-MYC: More than just a matter of life and death. Nature Reviews Cancer 2: 764-776.

249. Adhikary S, Eilers M (2005) Transcriptional regulation and transformation by MYC proteins. Nature Reviews Molecular Cell Biology 6: 635-645.

250. Elend M, Eilers M (1999) Cell growth: downstream of Myc - to grow or to cycle? Current Biology 9: R936-938.

251. Secombe J, Pierce SB, Eisenman RN (2004) Myc: a weapon of mass destruction. Cell 117: 153-156.

252. Schmidt EV (1999) The role of c-myc in cellular growth control. Oncogene 18: 2988-2996.

253. Schorl C, Sedivy JM (2003) Loss of protooncogene c-Myc function impedes G1 phase progression both before and after the restriction point. Mol Biol Cell 14: 823-835.

254. Classon M, Salama S, Gorka C, Mulloy R, Braun P, et al. (2000) Combinatorial roles for pRB, p107, and p130 in E2F-mediated cell cycle control. Proc Natl Acad Sci U S A 97: 10820-10825.

255. Herrera RE, Sah VP, Williams BO, Makela TP, Weinberg RA, et al. (1996) Altered cell cycle kinetics, gene expression, and G1 restriction point regulation in Rb-deficient fibroblasts. Mol Cell Biol 16: 2402-2407.

256. Evan GI, Wyllie AH, Gilbert CS, Littlewood TD, Land H, et al. (1992) Induction of Apoptosis in Fibroblasts by C-Myc Protein. Cell 69: 119-128.

257. Prendergast GC (1999) Mechanisms of apoptosis by c-Myc. Oncogene 18: 2967-2987.

258. Leone G, DeGregori J, Sears R, Jakoi L, Nevins JR (1997) Myc and Ras collaborate in inducing accumulation of active cyclin E/Cdk2 and E2F. Nature 387: 422-426.

259. Peeper DS, Upton TM, Ladha MH, Neuman E, Zalvide J, et al. (1997) Ras signalling linked to the cell-cycle machinery by the retinoblastoma protein. Nature 386: 177-181.

260. Liu JJ, Chao JR, Jiang MC, Ng SY, Yen JJ, et al. (1995) Ras transformation results in an elevated level of cyclin D1 and acceleration of G1 progression in NIH 3T3 cells. Mol Cell Biol 15: 3654-3663.

261. Robles AI, Rodriguez-Puebla ML, Glick AB, Trempus C, Hansen L, et al. (1998) Reduced skin tumor development in cyclin D1-deficient mice highlights the oncogenic ras pathway in vivo. Genes Dev 12: 2469-2474.

262. Marshall C (1999) How do small GTPase signal transduction pathways regulate cell cycle entry? Curr Opin Cell Biol 11: 732-736.

263. Diehl JA, Cheng M, Roussel MF, Sherr CJ (1998) Glycogen synthase kinase-3beta regulates cyclin D1 proteolysis and subcellular localization. Genes Dev 12: 3499-3511.

264. Aktas H, Cai H, Cooper GM (1997) Ras links growth factor signaling to the cell cycle machinery via regulation of cyclin D1 and the Cdk inhibitor p27KIP1. Mol Cell Biol 17: 3850-3857.

265. Takuwa N, Takuwa Y (1997) Ras activity late in G1 phase required for p27kip1 downregulation, passage through the restriction point, and entry into S phase in growth factor-stimulated NIH 3T3 fibroblasts. Mol Cell Biol 17: 5348-5358.

266. Galaktionov K, Jessus C, Beach D (1995) Raf1 interaction with Cdc25 phosphatase ties mitogenic signal transduction to cell cycle activation. Genes Dev 9: 1046-1058.

267. Olson MF, Paterson HF, Marshall CJ (1998) Signals from Ras and Rho GTPases interact to regulate expression of p21Waf1/Cip1. Nature 394: 295-299.

268. Kauffmann-Zeh A, Rodriguez-Viciana P, Ulrich E, Gilbert C, Coffer P, et al. (1997) Suppression of c-Myc-induced apoptosis by Ras signalling through PI(3)K and PKB. Nature 385: 544-548.

269. Escamilla-Powers JR, Sears RC (2007) A conserved pathway that controls c-Myc protein stability through opposing phosphorylation events occurs in yeast. Journal of Biological Chemistry 282: 5432-5442.

270. Kumar A, Marques M, Carrera AC (2006) Phosphoinositide 3-Kinase Activation in Late G1 Is Required for c-Myc Stabilization and S Phase Entry. Mol Cell Biol 26: 9116-9125.

271. Gregory MA, Hann SR (2000) c-Myc proteolysis by the ubiquitin-proteasome pathway: stabilization of c-Myc in Burkitt's lymphoma cells. Molecular & Cellular Biology 20: 2423-2435.

272. Vojtek AB, Der CJ (1998) Increasing complexity of the Ras signaling pathway. Journal of Biological Chemistry 273: 19925-19928.

273. Ebisuya M, Kondoh K, Nishida E (2005) The duration, magnitude and compartmentalization of ERK MAP kinase activity: mechanisms for providing signaling specificity. Journal of Cell Science 118: 2997-3002.

274. Traverse S, Seedorf K, Paterson H, Marshall CJ, Cohen P, et al. (1994) EGF triggers neuronal differentiation of PC12 cells that overexpress the EGF receptor. Current Biology 4: 694-701.

275. Murphy LO, Smith S, Chen RH, Fingar DC, Blenis J (2002) Molecular interpretation of ERK signal duration by immediate early gene products.[see comment]. Nature Cell Biology 4: 556-564.

276. Yaka R, Gamliel A, Gurwitz D, Stein R (1998) NGF induces transient but not sustained activation of ERK in PC12 mutant cells incapable of differentiating. Journal of Cellular Biochemistry 70: 425-432.

277. Balmanno K, Cook SJ (1999) Sustained MAP kinase activation is required for the expression of cyclin D1, p21Cip1 and a subset of AP-1 proteins in CCL39 cells. Oncogene 18: 3085-3097.

278. Chaussepied M, Ginsberg D (2004) Transcriptional regulation of AKT activation by E2F. Molecular Cell 16: 831-837.

279. Jones SM, Klinghoffer R, Prestwich GD, Toker A, Kazlauskas A (1999) PDGF induces an early and a late wave of PI 3-kinase activity, and only the late wave is required for progression through G1. Current Biology 9: 512-521.

280. Goncharova EA, Ammit AJ, Irani C, Carroll RG, Eszterhas AJ, et al. (2002) PI3K is required for proliferation and migration of human pulmonary vascular smooth muscle cells. American Journal of Physiology - Lung Cellular & Molecular Physiology 283: L354-363.

281. Jones SM, Kazlauskas A (2001) Growth-factor-dependent mitogenesis requires two distinct phases of signalling. Nature Cell Biology 3: 165-172.

282. Garcia Z, Kumar A, Marques M, Cortes I, Carrera AC (2006) Phosphoinositide 3-kinase controls early and late events in mammalian cell division. EMBO Journal 25: 655-661.

283. O'Donnell KA, Wentzel EA, Zeller KI, Dang CV, Mendell JT (2005) c-Myc-regulated microRNAs modulate E2F1 expression. Nature 435: 839-843.

284. Park CS, Schneider IC, Haugh JM (2003) Kinetic Analysis of Platelet-derived Growth Factor Receptor/Phosphoinositide 3-Kinase/Akt Signaling in Fibroblasts. J Biol Chem 278: 37064-37072.

285. Kaur H, Park CS, Lewis JM, Haugh JM (2006) Quantitative model of Ras phosphoinositide 3-kinase signalling cross-talk based on co-operative molecular assembly. Biochem J 393: 235-243.

286. Qu Z, MacLellan WR, Weiss JN (2003) Dynamics of the cell cycle: checkpoints, sizers, and timers. Biophysical Journal 85: 3600-3611.

287. Hatakeyama M, Kimura S, Naka T, Kawasaki T, Yumoto N, et al. (2003) A computational model on the modulation of mitogen-activated protein kinase (MAPK) and Akt pathways in heregulin-induced ErbB signalling. Biochem J 373: 451-463.

288. Shen M, Stukenberg PT, Kirschner MW, Lu KP (1998) The essential mitotic peptidyl-prolyl isomerase Pin1 binds and regulates mitosis-specific phosphoproteins. Genes & Development 12: 706-720.

289. Baharians Z, Schonthal AH (1998) Autoregulation of protein phosphatase type 2A expression. Journal of Biological Chemistry 273: 19019-19024.

290. Chang F, Lee JT, Navolanic PM, Steelman LS, Shelton JG, et al. (2003) Involvement of PI3K/Akt pathway in cell cycle progression, apoptosis, and neoplastic transformation: a target for cancer chemotherapy. Leukemia 17: 590-603.

291. Luo J, Manning BD, Cantley LC (2003) Targeting the PI3K-Akt pathway in human cancer: Rationale and promise. Cancer Cell 4: 257-262.

292. Stork PJS, Schmitt JM (2002) Crosstalk between cAMP and MAP kinase signaling in the regulation of cell proliferation. Trends in Cell Biology 12: 258-266.

293. Shen YH, Godlewski J, Zhu J, Sathyanarayana P, Leaner V, et al. (2003) Cross talk between JNK/SAPK and ERK/MAPK pathways: Sustained activation of JNK blocks ERK activation by mitogenic factors. J Biol Chem: M303264200.

294. Lepique AP, Moraes MS, Rocha KM, Eichler CB, Hajj GNM, et al. (2004) c-Myc protein is stabilized by fibroblast growth factor 2 and destabilized by ACTH to control cell cycle in mouse Y1 adrenocortical cells. J Mol Endocrinol 33: 623-638.

295. Kholodenko BN (2000) Negative feedback and ultrasensitivity can bring about oscillations in the mitogen-activated protein kinase cascades. European Journal of Biochemistry 267: 1583-1588.

296. Igoshin OA, Goldbeter A, Kaiser D, Oster G (2004) A biochemical oscillator explains several aspects of Myxococcus xanthus behavior during development. Proceedings of the National Academy of Sciences 101: 15760-15765.

297. Ciliberto A, Capuani F, Tyson JJ (2007) Modeling Networks of Coupled Enzymatic Reactions Using the Total Quasi-Steady State Approximation. PLoS Computational Biology 3: e45.

298. Heinrich R, Neel BG, Rapoport TA (2002) Mathematical models of protein kinase signal transduction. Molecular Cell 9: 957-970.

299. Wang J, Huang B, Xia X, Sun Z (2006) Funneled Landscape Leads to Robustness of Cellular Networks: MAPK Signal Transduction. Biophys J: biophysj.106.086777.

300. Huang C-YF, Ferrell JE, Jr. (1996) Ultrasensitivity in the mitogen-activated protein kinase cascade. PNAS 93: 10078-10083.

301. Segrelles C, Moral M, Lara MF, Ruiz S, Santos M, et al. (2006) Molecular determinants of Akt-induced keratinocyte transformation. Oncogene 25: 1174-1185.

302. Nilsson JA, Cleveland JL (2003) Myc pathways provoking cell suicide and cancer. Oncogene 22: 9007-9021.

303. Assoian RK (2002) Common sense signalling. Nature Cell Biology 4: E187-188.

304. Musti AM, Treier M, Bohmann D (1997) Reduced ubiquitin-dependent degradation of c-Jun after phosphorylation by MAP kinases. Science 275: 400-402.

305. Zeng X, Tamai K, Doble B, Li ST, Huang H, et al. (2005) A dual-kinase mechanism for Wnt co-receptor phosphorylation and activation. Nature 438: 873-877.

306. Mailand N, Podtelejnikov AV, Groth A, Mann M, Bartek J, et al. (2002) Regulation of G(2)/M events by Cdc25A through phosphorylation-dependent modulation of its stability. EMBO Journal 21: 5911-5920.

307. Amit S, Hatzubai A, Birman Y, Andersen JS, Ben-Shushan E, et al. (2002) Axin-mediated CKI phosphorylation of beta-catenin at Ser 45: a molecular switch for the Wnt pathway. Genes & Development 16: 1066-1076.

308. Hergovich A, Lisztwan J, Thoma CR, Wirbelauer C, Barry RE, et al. (2006) Priming-dependent phosphorylation and regulation of the tumor suppressor pVHL by glycogen synthase kinase 3. Molecular & Cellular Biology 26: 5784-5796.

309. Liu CM, Li YM, Semenov M, Han C, Baeg GH, et al. (2002) Control of beta-catenin phosphorylation/degradation by a dual-kinase mechanism. Cell 108: 837-847.

310. Tworkowski KA, Salghetti SE, Tansey WP (2002) Stable and unstable pools of Myc protein exist in human cells. Oncogene 21: 8515-8520.

311. Sanders JA, Gruppuso PA (2005) Nucleolar localization of hepatic c-Myc: a potential mechanism for c-Myc regulation. Biochimica et Biophysica Acta 1743: 141-150.

312. Arabi A, Rustum C, Hallberg E, Wright APH (2003) Accumulation of c-Myc and proteasomes at the nucleoli of cells containing elevated c-Myc protein levels. J Cell Sci 116: 1707-1717.

313. Endo T, Nadal-Ginard B (1986) Transcriptional and posttranscriptional control of c-myc during myogenesis: its mRNA remains inducible in differentiated cells and does not suppress the differentiated phenotype. Molecular & Cellular Biology 6: 1412-1421.

314. Jopling CL, Willis AE (2001) N-myc translation is initiated via an internal ribosome entry segment that displays enhanced activity in neuronal cells. Oncogene 20: 2664-2670.

315. Penn LJ, Brooks MW, Laufer EM, Land H (1990) Negative autoregulation of c-myc transcription. EMBO Journal 9: 1113-1121.

316. Peunova N, Enikolopov G (1995) Nitric oxide triggers a switch to growth arrest during differentiation of neuronal cells. Nature 375: 68-73.

317. Greene LA, Tischler AS (1976) Establishment of a noradrenergic clonal line of rat adrenal pheochromocytoma cells which respond to nerve growth factor. Proceedings of the National Academy of Sciences of the United States of America 73: 2424-2428.

318. Hennessy BT, Smith DL, Ram PT, Lu Y, Mills GB (2005) Exploiting the PI3K/AKT Pathway for Cancer Drug Discovery. Nat Rev Drug Discov 4: 988-1004.

319. Santos SDM, Verveer PJ, Bastiaens PIH (2007) Growth factor-induced MAPK network topology shapes Erk response determining PC-12 cell fate. Nat Cell Biol 9: 324-330.

320. Alon U (2007) Network motifs: theory and experimental approaches. Nat Rev Genet 8: 450-461.

321. Goldbeter A, Koshland DE, Jr. (1987) Energy expenditure in the control of biochemical systems by covalent modification. Journal of Biological Chemistry 262: 4460-4471.

322. Pizarro-Cerda J, Cossart P (2006) Bacterial adhesion and entry into host cells. Cell 124: 715-727.

323. Cossart P, Sansonetti PJ (2004) Bacterial invasion: The paradigms of enteroinvasive pathogens. Science 304: 242-248.

324. Garcia-Del Portillo F (2008) Heterogeneity in tissue culture infection models: a source of novel host-pathogen interactions? Microbes and Infection 10: 1063-1066.

325. Hallet B (2001) Playing Dr Jekyll and Mr Hyde: combined mechanisms of phase variation in bacteria. Current Opinion in Microbiology 4: 570-581.

326. Turner KH, Vallet-Gely I, Dove SL (2009) Epigenetic Control of Virulence Gene Expression in Pseudomonas aeruginosa by a LysR-Type Transcription Regulator. Plos Genetics 5: -.

327. van der Woude MW, Baumler AJ (2004) Phase and antigenic variation in bacteria. Clinical Microbiology Reviews 17: 581-+.

328. Isberg RR, Hamburger Z, Dersch P (2000) Signaling and invasin-promoted uptake via integrin receptors. Microbes and Infection 2: 793-801.

329. Wong KW, Isberg RR (2005) Emerging views on integrin signaling via Rac1 during invasin-promoted bacterial uptake. Current Opinion in Microbiology 8: 4-9.

330. Gerceker AA, Zaidi T, Marks P, Golan DE, Pier GB (2000) Impact of heterogeneity within cultured cells on bacterial invasion: Analysis of Pseudomonas aeruginosa and Salmonella enterica serovar typhi entry into MDCK cells by using a green fluorescent protein-induced cystic fibrosis transmembrane conductance regulator receptor. Infection and Immunity 68: 861-870.

331. Isberg RR, Barnes P (2001) Subversion of integrins by enteropathogenic Yersinia. Journal of Cell Science 114: 21-28.

332. Perelson AS (1981) Receptor Clustering on a Cell-Surface .3. Theory of Receptor Cross-Linking by Multivalent Ligands - Description by Ligand States. Mathematical Biosciences 53: 1-39.

333. Goennenwein S, Tanaka M, Hu B, Moroder L, Sackmann E (2003) Functional incorporation of integrins into solid supported membranes on ultrathin films of cellulose: Impact on adhesion. Biophysical Journal 85: 646-655.

334. Palecek SP, Horwitz AF, Lauffenburger DA (1999) Kinetic model for integrin-mediated adhesion release during cell migration. Annals of Biomedical Engineering 27: 219-235.

335. Van Nhieu GT, Isberg RR (1991) The Yersinia pseudotuberculosis invasin protein and human fibronectin bind to mutually exclusive sites on the alpha 5 beta 1 integrin receptor. Journal of Biological Chemistry 266: 24367-24375.

336. Grillot-Courvalin C, Goussard S, Huetz F, Ojcius DM, Courvalin P (1998) Functional gene transfer from intracellular bacteria to mammalian cells. Nature Biotechnology 16: 862-866.

337. Xiang SL, Fruehauf J, Li CJ (2006) Short hairpin RNA-expressing bacteria elicit RNA interference in mammals. Nature Biotechnology 24: 697-702.

338. Ferrell JE (2002) Self-perpetuating states in signal transduction: positive feedback, double-negative feedback and bistability. Current Opinion in Cell Biology 14: 140-148.

339. Knodler LA, Celli J, Finlay BB (2001) Pathogenic trickery: Deception of host cell processes. Nature Reviews Molecular Cell Biology 2: 578-588.

340. Vannhieu GT, Isberg RR (1993) Bacterial Internalization Mediated by Beta-1 Chain Integrins Is Determined by Ligand Affinity and Receptor Density. Embo Journal 12: 1887-1895.

341. Isberg RR, Leong JM (1990) Multiple Beta-1-Chain Integrins Are Receptors for Invasin, a Protein That Promotes Bacterial Penetration into Mammalian-Cells. Cell 60: 861-871.

342. Isberg RR, Voorhis DL, Falkow S (1987) Identification of Invasin - a Protein That Allows Enteric Bacteria to Penetrate Cultured-Mammalian-Cells. Cell 50: 769-778.

343. Hsu JP (1987) Stochastic Modeling of Bacterial Adhesion - a 2-Step Mechanism with Linear Adhesion Rate. Journal of Theoretical Biology 124: 495-504.

344. Hsu JP, Wang HH (1987) Kinetics of Bacterial Adhesion - a Stochastic-Analysis. Journal of Theoretical Biology 124: 405-413.

345. Hsu JP, Wang HH (1986) A Stochastic-Analysis of Bacterial Adhesion. Journal of Theoretical Biology 119: 435-444.

346. Galvez J, Lajarin F, GarciaPenarrubia P (1997) Penetration of host cell lines by bacteria. Characteristics of the process of intracellular bacterial infection. Bulletin of Mathematical Biology 59: 857-879.

347. Galvez J, Lajarin F, GarciaPenarrubia P (1997) Mathematical modeling of adhesion of bacteria to host cell lines. Bulletin of Mathematical Biology 59: 833-856.

348. Levy SB, Marshall B (2004) Antibacterial resistance worldwide: causes, challenges and responses. Nature Medicine 10: S122-S129.

349. Alksne LE, Projan SJ (2000) Bacterial virulence as a target for antimicrobial chemotherapy. Current Opinion in Biotechnology 11: 625-636.

Biography

Name: Tae Jun Lee

Date of Birth: 02.07.1979

Place of Birth: Kwang-Ju, South Korea

Education:

08/03 – 05/10   Duke University, Durham NC

                    Ph.D. Department of Biomedical Engineering 5/2010

08/99 – 05/03   Duke University, Durham NC

                    B.S.E Electrical Engineering / Biomedical Engineering 05/2003

Publications:

- **Peer-reviewed articles:**

  - <u>Lee T</u>, Yao G, Nevins J, You L (2008) Sensing and integration of Erk and PI3K signals by Myc. PLoS Computational Biology 4:e10000013
  - Yao G, <u>Lee TJ</u>, Mori S, Nevins JR, You LC (2008) A Bistable Rb-E2F switch underlies the restriction point. Nature Cell Biology 10:476-U255
  - Bobkov, K.V. and Bowsher, J.E. and Greer, K.L. and Jaszczak, R.J. and Braun, K.E. and <u>Lee, T.J.,</u> *Phantom assessment of new acquisition geometries for breast pinhole SPECT imaging*, IEEE Trans. Nucl. Sci. (USA), vol. 53 no. 3 (2006), pp. 1162 - 7

- **Book chapters:**

  - <u>T. Lee,</u> D. Tu, C.M. Tan, and L. You (2007) Modeling cellular networks. In "***Systems Bioinformatics: An Engineering Case-Based Approach,***" Eds. G. Alterovitz & M. Ramoni. Chapter 6. Page 151-178 (Peer reviewed)

o D. Tu, J. Lee, T. Ozdere, <u>T. Lee</u>, and L. You. (2007) Engineering gene circuits: foundations and applications. In "***Nanotechnology in Biology and Medicine: Methods, Devices and Applications,***" Ed. T. Vo-Dinh. CRC Press. Chapter 20. Page 363-380