UNIVERSITY OF CALIFORNIA, SAN DIEGO

Steady State Control of the Cellular Response to Stress

A dissertation submitted in partial satisfaction of the
requirements for the degree
Doctor of Philosophy

in

Bioinformatics & Systems Biology

by

Paul Michael Loriaux

Committee in charge:

Professor Alexander Hoffmann, Chair
Professor Pavel Pevzner, Co-Chair
Professor Gourisankar Ghosh
Professor Bing Ren
Professor Wei Wang

2013

The dissertation of Paul Michael Loriaux is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

_____

_____

_____

_____
Co-Chair

_____
Chair

University of California, San Diego

2013

EPIGRAPH

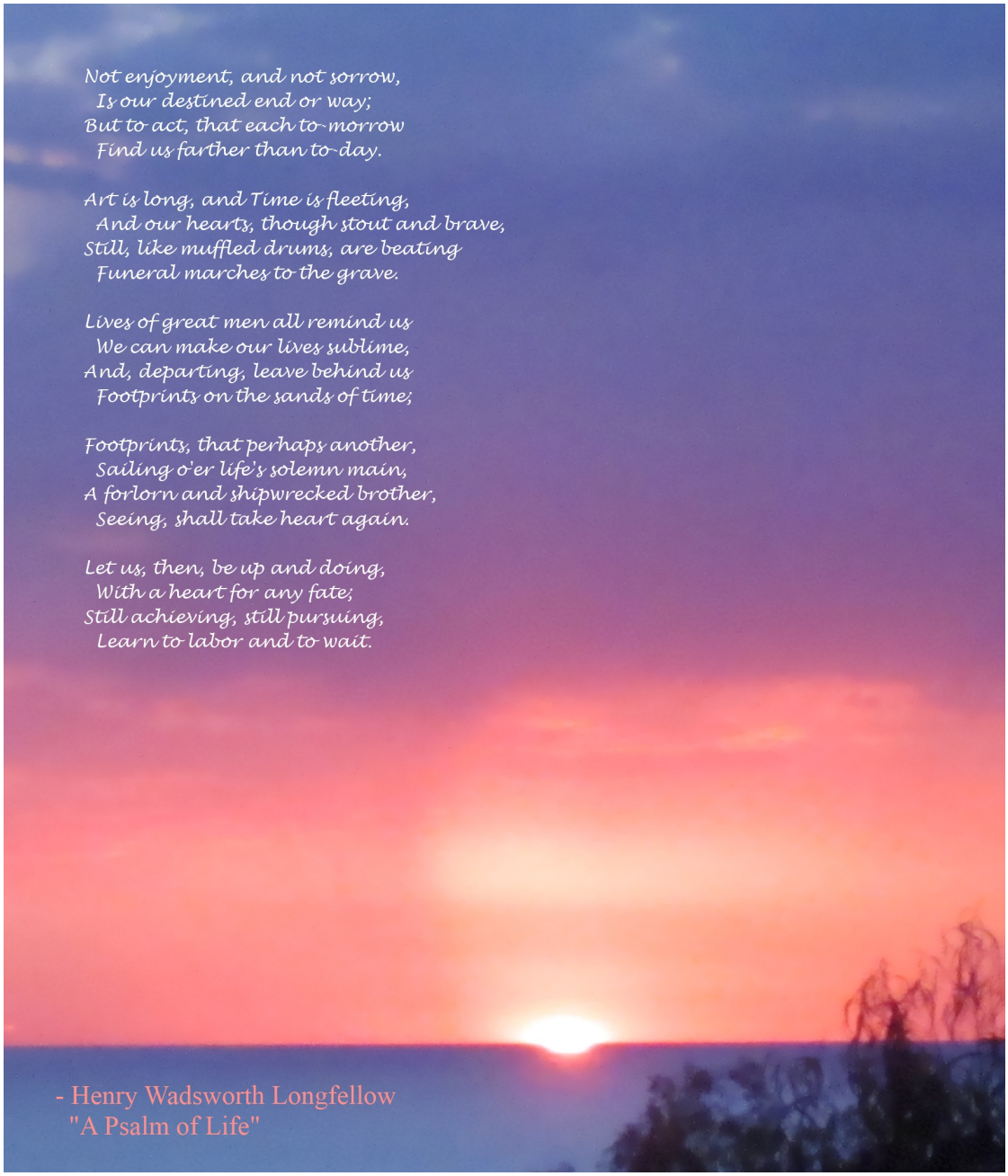

Not enjoyment, and not sorrow,
Is our destined end or way;
But to act, that each to-morrow
Find us farther than to-day.
Art is long, and Time is fleeting,
And our hearts, though stout and brave,
Still, like muffled drums, are beating
Funeral marches to the grave.
Lives of great men all remind us
We can make our lives sublime,
And, departing, leave behind us
Footprints on the sands of time;
Footprints, that perhaps another,
Sailing o'er life's solemn main,
A forlorn and shipwrecked brother,
Seeing, shall take heart again.
Let us, then, be up and doing,
With a heart for any fate;
Still achieving, still pursuing,
Learn to labor and to wait.
- Henry Wadsworth Longfellow
"A Psalm of Life"

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

GFP	Green fluorescent protein
IGF-1R	Insulin-like growth factor 1 receptor
IKK	Inhibitor of kappa B kinase
IκB	Inhibitor of kappa B
IR	Ionizing gamma radiation
LPS	Lipopolysaccharide
MAPK	Mitogen-activated protein kinase
Mdm2	Mouse double minute 2 homolog
NF-κB	Nuclear factor-kappa B
ODE	Ordinary differential equation
PEST	Proline Glutamate Serine Threonine
PLoS	Public Library of Science
TNF	Tumor necrosis factor alpha
TRAIL	TNF-related apoptosis-inducing ligand

LIST OF SYMBOLS

System elements

a_i	The i^{th} molecular species, $1 \leq i \leq d_x$.
r_j	The j^{th} biochemical reaction, $1 \leq j \leq d_k$.

Scalars

kj	The rate constant corresponding to reaction j.
k_j	The rate constant corresponding to reaction j.
$k'_{a,b}$	The transition rate constant from species a to species b, $1 \leq a, b \leq d_x$.
x_i	The concentration or abundance of a_i.
x_{tot}	The total concentration of all molecular species, $\sum_1^{d_x} x_i$.
v_j	The velocity of reaction j.
$s_{i,j}^{in}$	The stoichiometry of reactant i in reaction j.
$s_{i,j}^{out}$	The stoichiometry of product i in reaction j.
$h_{i,j}$	The order of reactant i in reaction j.

Vectors

$\mathbf{k}$	A vector of reaction rate constants.
$\mathbf{x}$	A vector of species concentrations.
$\dot{\mathbf{x}}$	The first derivative of $\mathbf{x}$ with respect to time.
$\mathbf{v}$	A vector of reaction velocities
$\overline{\mathbf{v}}$	A vector of reaction velocities satisifying $\mathbf{S}\overline{\mathbf{v}} = \mathbf{0}$.
$\mathbf{y}$	A vector of quantities on which $\dot{\mathbf{x}}$ has linear dependence.
$\overline{\mathbf{y}}$	A vector of quantities satisfying $\mathbf{C}\mathbf{y} = \mathbf{0}$.
$\mathbf{q}$	Coefficients of a linear combination of null space basis vectors; these parameterize solutions of $\psi_p(\dot{\mathbf{x}}) = \mathbf{0}$.
$\mathbf{n}$	A null space basis vector. Also, a column of $\mathbf{N}$.
$\mathbf{c}$	A column vector in $\mathbf{C}_{\texttt{rref}}$.

Matrices

$\mathbf{S}$	The matrix whose elements $(\mathbf{S})_{ij}$ give the net stoichiometry of species i in reaction j.
$\mathbf{P}$	The matrix of partial derivatives $(\mathbf{P})_{ij} = (\partial v_i / \partial y_j)$.
$\mathbf{P}_k$	The matrix of partial derivatives $(\mathbf{P}_k)_{ij} = (\partial v_i / \partial k_j)$.
$\mathbf{P}_x$	The matrix of partial derivatives $(\mathbf{P}_x)_{ij} = (\partial v_i / \partial x_j)$.
$\mathbf{C}$	The coefficient matrix, formed by the product of $\mathbf{S}$ and $\mathbf{P}$.
$\mathbf{C}_{\texttt{rref}}$	The row reduced echelon form of $\mathbf{C}$.
$\mathbf{N}$	A matrix whose columns are orthogonal and span the solution space to Equation 2.14, $\mathbf{Cy} = \mathbf{0}$, which is $\psi_p(\dot{\mathbf{x}}) = \mathbf{0}$ in a certain basis.
$\mathbf{K}$	A matrix whose elements are $(\mathbf{K})_{ij} = k'_{ij}$. See Equation 2.7.
$\mathbf{M}_i$	The i^{th} minor of $\mathbf{K}$, formed by removing the i^{th} row and column and computing its determinant.

Sets

$\mathcal{K}$	A set of reaction rate constants, $\{k_1, k_2, \ldots, k_{d_k}\}$.
$\mathcal{X}$	A set of molecular species concentrations, $\{x_1, x_2, \ldots, x_{d_x}\}$.
$\mathcal{R}$	A set of biochemical reactions, $\{r_1, r_2, \ldots, r_{d_k}\}$.
$\mathcal{Y}$	A selected subset of variables $\mathcal{K} \cup \mathcal{X}$ on which every reaction velocity has linear dependence, relabeled as $\{y_1, y_2, \ldots, y_{d_y}\}$.
$\mathcal{P}$	All remaining variables in $\mathcal{K} \cup \mathcal{X}$ not mapped to $\mathcal{Y}$, relabeled as $\{p_1, p_2, \ldots, p_{d_p}\}$. All variables on which $\dot{\mathbf{x}}$ has nonlinear dependence should be placed in $\mathcal{P}$. All variables which should be forced to be independent should be placed in $\mathcal{P}$.
$\mathcal{Q}$	Parameterization $\{q_1, q_2, \ldots, q_{d_q}\}$ of solutions of $\psi_p(\dot{\mathbf{x}}) = \mathbf{0}$.
$\mathcal{K}_{lin}, \mathcal{X}_{lin}$	User-selected subsets of $\mathcal{K}$ and $\mathcal{X}$ on which the reaction velocities have a linear relationship. These are mapped by ψ_p to elements in $\mathcal{Y}$.
$\mathcal{K}_p, \mathcal{X}_p$	All remaining variables in $\mathcal{K}$ or $\mathcal{X}$. These are mapped by ψ_p to p's.
$\mathcal{K}_q, \mathcal{X}_q$	The subsets of $\mathcal{K}_{lin}$ and $\mathcal{X}_{lin}$ determined to be free parameters.
$\mathcal{K}_y, \mathcal{X}_y$	The subsets of $\mathcal{K}_{lin}$ and $\mathcal{X}_{lin}$ determined to be dependent parameters.
$\mathcal{K}_{pq}, \mathcal{X}_{pq}$	The unions of subsets $\mathcal{K}_{pq} = \mathcal{K}_p \cup \mathcal{K}_q$, $\mathcal{X}_{pq} = \mathcal{X}_p \cup \mathcal{X}_q$.

$\mathcal{Y}_q$	The subset of $\mathcal{Y}$ giving free variables for the solution of $\psi_p(\dot{\mathbf{x}}) = \mathbf{0}$. There is a bijection between these and $\mathcal{Q} = \{q_1, q_2, \ldots, q_{d_q}\}$.
$\mathcal{Y}_q^c$	The complement of $\mathcal{Y}_q$ in $\mathcal{Y}$. The dependent variables.
$\mathcal{R}'$	A subset of $\mathcal{R}$ where one reaction has been removed.
$\mathcal{G}$	The graph $\{\mathcal{X}, \mathcal{R}\}$ formed by treating each species in $\mathcal{X}$ as a node and each reaction in $\mathcal{R}$ as an edge.
$\mathcal{S}_i$	The set of King-Altman patterns for species i.
$\mathbb{N}_0$	The set of natural numbers, including zero.
$\mathbb{R}_0$	The set of positive real numbers, including zero.

Dimensions

d_x	The number of molecular species in a system. Also, the size of $\mathcal{X}$.
d_k	The number of reactions in a system. Also, the sizes of $\mathcal{R}$ and $\mathcal{K}$.
d_y	The number of selected linear quantities for a particular mapping function ψ_p. Also, the size of $\mathcal{Y}$.
d_p	The number of remaining parameters for a particular mapping function ψ_p; that is, $d_p = d_k + d_x - d_y$. Also, the size of $\mathcal{P}$.
d_q	The number of free variables in $\mathcal{Y}$ in the solution of $\psi_p(\dot{\mathbf{x}}) = 0$. See Equation 2.17.
d_c	The number of dependent variables in $\mathcal{Y}$ in the solution of $\psi_p(\dot{\mathbf{x}}) = 0$. Also, the size of $\mathcal{Y}_q^c$.

Functions

ψ_p	A map that relabels variables $\mathcal{K} \cup \mathcal{X}$ as $\mathcal{P} \cup \mathcal{Y}$.
ψ_y	A map that imposes the relations $\psi_p(\dot{\mathbf{x}}) = \mathbf{0}$ by expressing dependent variables $\mathcal{Y}_q^c$ in terms of free variables $\mathcal{P} \cup \mathcal{Q}$. See Equation 2.20.
ψ_q	A bijection from elements in $\mathcal{Y}_q$ to elements in $\mathcal{Q}$.
ψ_q^{-1}	A map from $\mathcal{Q}$ back to $\mathcal{Y}$.
ψ_{py}	The composite map $\psi_y \circ \psi_p$.
ψ_{qp}^{-1}	The composite map $\psi_p^{-1} \circ \psi_q^{-1}$.

ψ_{ss} The composite map $\psi_{qp}^{-1} \circ \psi_{py}$. This expresses dependent variables among $\mathcal{K} \cup \mathcal{X}$ in terms of a set of free variables.

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

I would like to acknowledge Professor Alexander Hoffmann for his patience and support as the chair of my committee and for allowing me the freedom to pursue my research objectives. I also acknowledge the theoretical guidance given by my co-chair, Professor Pavel Pevzner, and my committee, Professors Gourisankar Ghosh, Bing Ren, and Wei Wang, for numerous insightful suggestions.

Two faculty members who do not serve on my committee but also deserve recognition are Professor Glenn Tesler, for his tireless revisions of our *py*-substitution manuscript (see Chapter 2), and Professor Charles Elkan, for chairing my masters project in computer science and for his generous and welcome guidance in bringing our cell death manuscript to fruition (Chapter 4).

Chapter 1, in part, is a reprint of material as it appears in Methods in Cell Biology. Loriaux, Paul M; Hoffmann, Alexander, Elsevier 2012. Chapter 1 is also, in part, a reprint of material as it appears in Public Library of Science (PLoS) Computational Biology. Loriaux, Paul M; Tesler, Glenn; Hoffmann, Alexander, Public Library of Science 2013. The dissertation author was the primary investigator and author of these papers.

Chapter 2, in part, is a reprint of material as it appears in PLoS Computational Biology. Loriaux, Paul M; Tesler, Glenn; Hoffmann, Alexander, Public Library of Science 2013. The dissertation author was the primary investigator and author of this paper.

Chapter 3, in full, is a reprint of material as it appears in PLoS Computational Biology. Loriaux, Paul M; Hoffmann, Alexander, Public Library of Science 2013. The dissertation author was the primary investigator and author of this paper.

Chapter 4, in part, is currently being prepared for submission for publication. Loriaux, Paul M; Elkan, Charles; Hoffmann, Alexander. The dissertation author was the primary investigator and author of this material.

VITA

1996-1997	Laboratory Assistant, Center for Endocrinology and Molecular Medicine, Northwestern University, Chicago, IL
2001	B. S. in Bioengineering, University of Washington, Seattle, WA
2001-2004	Research Science Engineer, Cell Systems Initiative, Seattle, WA
2004-2005	Computational Biologist, Institute for Systems Biology, Seattle, WA
2005-2013	Research Assistant, University of California, San Diego, San Diego, CA
2011	M. S. in Computer Science and Engineering, University of California, San Diego, San Diego, CA
2013	Ph. D. in Bioinformatics and Systems Biology, University of California, San Diego, San Diego, CA

PUBLICATIONS

"Of elections and cell-death decisions", *Mol Cell.* 2009 May 15;34(3):257-8.

"A framework for modeling the relationship between cellular steady-state and stimulus-responsiveness", *Methods Cell Biol.* 2012;110:81-109.

"Characterizing the relationship between steady state and response using analytical expressions for the steady states of mass action models", *PLoS Comput Biol.* 2013 Feb;9(2):e1002901.

"A protein turnover signaling motif controls the stimulus-sensitivity of stress response pathways", *PLoS Comput Biol.* 2013 Feb;9(2):e1002932

"A steady state sampling framework identifies kinetic versus static features as better predictors of TRAIL-induced cell death", *In Preparation.*

AWARDS

2005-2007	NIH Bioinformatics Training Grant No.GM00806-06
2007-2008	Whitaker International Fellowship
2008-2011	DOE Computational Sciences Graduate Fellowship

ABSTRACT OF THE DISSERTATION

Steady State Control of the Cellular Response to Stress

by

Paul Michael Loriaux

Doctor of Philosophy in Bioinformatics & Systems Biology

University of California, San Diego, 2013

Professor Alexander Hoffmann, Chair
Professor Pavel Pevzner, Co-Chair

In response to stress, intracellular signaling proteins activate gene expression programs that protect the cell, address the instigating stress, or result in programmed cell death. In many cases, information about the stimulus is encoded in the dynamics of the signal. Stress-induced signaling dynamics can therefore dictate the cellular response to stress. Recently, it was shown that these dynamics are affected by the resting state of the cell prior to stimulation. If this relationship between steady state and stimulus-induced dynamics was known, then we might predict the cellular response to a particular stimulus using steady state measurements, or engineer a stimulus to elicit a desired response. These are the foundations of diagnostic biomarkers and personalized medicine. To characterize

the relationship between steady state and the cellular response, I developed a suite of computational methods and applied them to the p53, NF-κB, and cell death pathways. First, I developed a method to derive analytical expressions for the steady states of mass action models. By applying this method to a model of cell death, I show how the steady state concentrations of different signaling proteins can affect the tolerance to the death-inducing ligand, TRAIL. Next I extended this method to examine perturbations in the steady state that don't affect the steady state protein concentrations. Applying this method to the p53 and NF-κB stress-response pathways, I show that a protein turnover signaling motif controls the stimulus-sensitivity of these two different pathways. Finally, using a Monte Carlo method, I show how sampling of the steady state prior to simulation can identify steady state predictors of the response to TRAIL. Interestingly, kinetic features, rather than steady state concentrations, figured prominently among the best predictors. If true, this has severe consequences for clinical biomarker discovery, which is based on measurements of protein abundance and not kinetic features.

Chapter 1

Introduction

1.1 The cellular response to stress

Cells must constantly respond to stress in their environment. Depending on the nature of the insult, cells can experience metabolic, ribotoxic, or oxidative stress, as well as DNA damage, heat shock, or ligand-induced aggregation of cell death receptors [1, 2]. In response to stress, transient activation of signaling molecules activate gene expression programs that function to protect the cell, arrest the cell cycle, or initiate programmed cell death. DNA damage caused by ionizing radiation or ultraviolet light, for example, activates the transcription factor p53, causing cell-cycle arrest, apoptosis, and DNA repair [3]. The inflammatory cytokine tumor necrosis factor alpha (TNF), and the pathogen-derived lipopolysaccharide (LPS), activate the transcription factor NF-κB resulting in expression of target genes that control cell survival and division [4]. In cells with compromised NF-κB activity, TNF stimulation results in activation of executioner caspases and cell death [5].

Because transient activation of signaling molecules is a hallmark of the response to stress, it is no surprise that the dynamics of activation are an important determinant of the cellular response [6, 7, 8]. In response to ionizing radiation (IR), for example, cells that experience pulses in p53 recover from the insult, while sustained p53 activation results in senescence [9]. Negative feedback attenuation of NF-κB in response to TNF, versus sustained activation in response to LPS, results

in TNF- and LPS-specific gene expression programs [10, 11, 12]. And temporal segregation of TNF-induced activation of IKK – the canonical kinase responsible for activating NF-κB – from activation of the pro-apoptotic MAP kinase (MAPK), is thought to be a major contributor to the cytoprotective effects of NF-κB [13].

Modulating the dynamics of signaling molecules in response to stimulation may therefore be a ubiquitous mechanism for encoding information about the instigating stimulus. However, it remains unclear how the resting state of the cell prior to stimulation affects these dynamics. Characterizing this relationship is important because aberration of the resting state is known to adversely affect cellular responsiveness and human health. Cells with increased abundance of p53 or its negative regulator Mdm2, for example, have compromised sensitivity to IR [14, 15, 16, 17]. Stable overexpression of the negative regulator A20 results in dampened NF-κB activation in response to TNF and inflammatory tolerance [18]. And cells that overexpress the antiapoptotic proteins Bcl-xL or Bcl-2 are insensitive to stimulation by the death-inducing TNF superfamily member, TRAIL [19, 20, 21]

If the abundance of intracellular signaling molecules affect the stimulus-induced dynamics and cellular response to stress, other studies have shown that the kinetics of protein turnover, and not their outright abundance, can affect the cell's response as well. For example, high turnover of a cell surface receptor is required to maintain a linear, non-refractory response to a broad range of stimulus concentrations [22]. And cells conditioned in an inflammatory environment experience elevated activity of IKK and turnover of IκBα– the primary inhibitor of NF-κB – resulting in hypersensitivity to ribotoxic stress [23].

The ability to systematically characterize this complex relationship between steady state and response would greatly enhance our understanding of stimulus-induced dynamics and the cellular response to stress. In what follows, I describe the development and application of computational methods to investigate the steady state control of p53, NF-κB, and cell death in response to different stress-inducing stimuli. First, in chapter 2 I describe a method, *py*-substitution, for deriving analytical expressions for the steady states of a common family of models, those that obey mass action kinetics. I then use this method to identify determinants

of sensitivity in TRAIL-induced cell death. In chapter 3 I extend *py*-substitution to characterize the effects of steady state protein turnover on the specificity of the p53 and NF-κB pathways. Finally, in chapter 4, I use Monte Carlo sampling of the steady state to identify predictors of the population response to TRAIL. The remainder of this chapter proceeds with an introduction and motivation of the methodology.

1.2 Methodology

Systematic changes in the resting states of living cells are difficult to engineer. Gene knockdowns using short interfering (si) RNA, for example, can be used to reduce the concentration of a target protein by blocking its translation [24, 25]. Changes in stimulus-responsiveness that result from siRNA knockdown, however, may be caused by a reduction in the protein's turnover kinetics, its intended reduction in concentration, or both. RNA dilution in rapidly dividing cells [26] and secondary induction of the interferon response [27] can further cloud interpretation of the data. To control a protein's turnover, tet-response elements can be introduced into its promoter [28, 29], or the protein can be tagged with an ssrA sequence for recognition by an inducible clpXP protease [30]. However, each target protein must be studied in isolation to avoid confounding effects from shared synthesis and degradation machinery [31]. Due to these inherent difficulties in effecting changes in the resting state of living cells, the relationship between resting state and stress responsiveness cannot be addressed by laboratory science alone.

As a complement to laboratory science, mathematical models have long been used in cell biology to rapidly investigate alternative hypotheses [32, 33, 34]. A model of p53 activation, for example, led to the discovery of the negative feedback regulator required for oscillations after stimulation by IR [35]. This same model was used by another group to show that heterogeneity observed in the population response to IR is due to stochasticity in protein synthesis, not in the processing of the signal itself [36]. A novel negative feedback regulator of NF-κB was also discovered with the aid of a mathematical model [37]. Indeed, canonical NF-κB

signaling may be one of the most frequently modeled systems in all of cell biology [38, 39]. Models of NF-κB activation have identified rapid degradation of unbound IκBα as being critical to TNF-induced NF-κB activity [23, 40], and the late-phase of IKK activity as required for distinguishing LPS from TNF stimulation [12]. Finally, models of cell death have revealed this pathway's capability for bistability or ultrasensitivity, as well as the principle regulators of this important switching behavior [41, 42].

The mathematical formalisms used to model biological processes are as varied as the processes themselves [43]. On one end of the spectrum are those that require only knowledge of the network of interacting molecules. Assigning a binary value to each molecule and a boolean operator to each interaction yields a boolean logic model. This type of model was proposed 45 years ago by Stuart Kauffman [44, 45], and has been used extensively to identify dependencies in network behavior [46, 47]. Recently, this approach was extended to allow variables to exist on the continuous interval $[0, 1]$ [48, 49]. These "fuzzy" logic models are a more natural representation of the probabilistic transitions between network states, and are more appropriate for reconciling large amounts of data with incomplete knowledge of an underlying network [50]. Still another formalism treats the reaction velocities as real-valued state variables, then uses the structure of the network to map these velocities onto the zero vector [51]. The solution space for this mapping is a high-dimensional polytope that defines the space of allowable steady states. Vertices on this polytope represent network optima, and their identification by linear programming is fundamental to metabolic engineering and genome-scale studies of metabolism [52, 53].

While these models are a powerful means to explore network behavior, because the concentrations of molecules are not explicit functions of time, they are inappropriate for studying signaling dynamics. To model dynamical systems, physicochemical models based on the Law of Mass Action are used instead [54]. Here, changes in molecule concentrations with respect to time are modeled as a system of coupled ordinary differential equations (ODEs) [55, 56]. Numerical integration of the ODEs from an initial condition gives the dynamic behavior of

the system in response to perturbation. Indeed, this methodology is so widespread that it now enjoys a common interoperable language [57], software library [58, 59], and database of curated models [60, 61]. In the chapters that follow I used mass action models of stress-response pathways to elucidate behaviors that are governed by the steady state. Prior to this, the basic tenets of mass action models are introduced below, then applied to a simple negative feedback system by way of an illustrative example.

1.3 Preliminaries

In this section I provide a brief mathematical foundation for models obeying mass action kinetics. First, let $\mathbb{N}_0$ be the set of non-negative natural numbers and $\mathbb{R}_0$ be the set of non-negative real numbers. Let $\mathcal{A} = \{a_1, a_2, \ldots, a_{d_x}\}$ be a set of d_x species and $\mathcal{R} = \{r_1, r_2, \ldots, r_{d_k}\}$ be a set of d_k reactions. Each reaction $r_j \in \mathcal{R}$ follows the normal definition,

$$r_j \colon s^{in}_{1,j} a_1 + s^{in}_{2,j} a_2 + \cdots + s^{in}_{d_x,j} a_{d_x} \xrightarrow{v_j} s^{out}_{1,j} a_1 + s^{out}_{2,j} a_2 + \cdots s^{out}_{d_x,j} a_{d_x}$$

where $s^{in}_{i,j} \in \mathbb{N}_0$ is the stoichiometric coefficient of the i^{th} reactant and $s^{out}_{i,j} \in \mathbb{N}_0$ is the stoichiometric coefficient of the i^{th} product [56]. We define $x_i \in \mathbb{R}_0$ to be the concentration of species a_i and $v_j \in \mathbb{R}_0$ to be the velocity at which r_j converts reactants into products. By the *Law of Mass Action*,

$$v_j = k_j \prod_{i=1}^{d_x} x_i^{h_{i,j}}. \tag{1.1}$$

The quantity $h_{i,j} \in \mathbb{R}_0$ is often, but not necessarily, equal to $s^{in}_{i,j}$. The coefficient $k_j \in \mathbb{R}_0$ is called the *rate constant*. Assuming conservation of mass, the concentration x_i changes according to

$$\dot{x}_i = \sum_{j=1}^{d_k} (s^{out}_{i,j} - s^{in}_{i,j}) v_j, \tag{1.2}$$

where $\dot{x}_i$ is the first derivative of x_i with respect to time. Any collection $\{\mathcal{A}, \mathcal{R}\}$ where the concentration x_i of $a_i \in \mathcal{A}$ obeys Equation 1.2 and the velocity v_j of $r_j \in \mathcal{R}$ obeys Equation 1.1 is called a *mass action model*. In what follows, we assume i, i_1, and i_2 are indices over the interval $1, \ldots, d_x$ and j is an index over $1, \ldots, d_k$. When $v_1, \ldots, v_{d_k}$ are such that all

$$\dot{x}_i = 0, \tag{1.3}$$

the model is said to be at *steady state*. If all $v_j = 0$ we call the steady state *trivial*. In this manuscript we are typically concerned with symbolic, non-trivial solutions to Equation 1.3. A solution is symbolic if all k_j and x_i are left as uninterpreted variables, rather than being assigned numerical values. Finally, we make the distinction between $x_i(t)$, which is a function of time, and $\bar{x}_i$, which represents the steady state concentration of a_i and is independent of time.

Let $\mathbf{x} \in \mathbb{R}_0^{d_x}$ and $\mathbf{v} \in \mathbb{R}_0^{d_k}$ be the vectors with elements $(\mathbf{x})_i = x_i$, and $(\mathbf{v})_j = v_j$. Throughout this manuscript, we use $(\mathbf{b})_i$ to denote the i^{th} element of vector $\mathbf{b}$ and $(\mathbf{A})_{ij}$ to denote the element at row i, column j of matrix $\mathbf{A}$. Let $\mathbf{S}$ be the *stoichiometric matrix*, i.e., the matrix whose elements are $(\mathbf{S})_{ij} = s_{i,j}^{out} - s_{i,j}^{in}$. Using this notation, Equation 1.2 becomes

$$\dot{\mathbf{x}} = \mathbf{S}\mathbf{v}, \tag{1.4}$$

and the steady state equation becomes

$$\mathbf{S}\overline{\mathbf{v}} = \mathbf{0}. \tag{1.5}$$

Again, by convention, we use the overline to denote vectors that satisfy steady state.

1.4 An example

To motivate the methodology used over the next three chapters, we'll consider a prototypical negative feedback model reminiscent of the p53 or NF-κB stress-response pathways (Figure 1.4A). In it, an activator X is constitutively expressed but catalytically degraded by an inhibitor, Y. The inhibitor is constitutively degraded but its synthesis requires X. Let x_1 and $x_2 \in \mathbb{R}_0$ denote the concentrations of X and Y, respectively. The set $\mathcal{R}$ of reactions for this model can then be given by

$$\begin{aligned} r_1 &: \emptyset \xrightarrow{v_1} x_1 \\ r_2 &: x_1 \xrightarrow{v_2} x_1 + x_2 \\ r_3 &: x_1 + x_2 \xrightarrow{v_3} x_2 \\ r_4 &: x_2 \xrightarrow{v_4} \emptyset \end{aligned}$$

The symbol $\emptyset$ represents a *source* or *sink* for mass and is not modeled by a time-varying species. From the set $\mathcal{R}$ we derive the stoichiometric matrix and reaction velocity vector,

$$\mathbf{S} = \begin{bmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & -1 \end{bmatrix}, \quad \mathbf{v} = \begin{bmatrix} k_1 \\ k_2 x_2 \\ k_3 x_1 x_2 \\ k_4 x_2 \end{bmatrix}.$$

By Equation 1.4 this results in the following system of equations,

$$dx_1/dt = k_1 - k_3 x_1 x_2 \tag{1.6}$$

$$dx_2/dt = k_2 x_1 - k_4 x_2 \tag{1.7}$$

Activation of X is achieved by instantaneous depletion of Y, the result of which is accumulation of X until negative feedback forces a return to steady state. The dynamics of this response can be described by two values: A, the amplitude

or maximum value of X after stimulation, and T, the time at which A is observed (Figure 1.4B). Parameters for this model were chosen such that the abundances of both X and Y are one arbitrary unit and X achieves its maximum value of $A = 10$ a.u. at time $T = 24$, where the units of time are also arbitrary.

To characterize the effects of steady state on the dynamic response, it is tempting to look for analytical expressions for A and T. If we could solve Equations 1.6 and 1.7 for x_1 and x_2 in terms of t, we could find a time $T = \arg\max_t x_1(t)$ and amplitude $A = x_1(T)$, resulting in expressions for T and A, respectively, in terms of $k_1, \ldots, k_4$, and the boundary conditions $x_1(t \leq 0) = \bar{x}_1$ and $x_2(t \leq 0) = \bar{x}_2$. Unfortunately, Equation 1.6 is nonlinear in the time-dependent variables x_1 and x_2, making an analytical solution impossible. Note that Equation 1.6 is rendered nonlinear by the monomial $k_3 x_1 x_2$. This monomial describes the mass action velocity of Y-catalyzed degradation of X. In general, any biochemical reaction involving two or more molecules results in a nonlinear monomial with respect to time. As a result, mass action models in general can not be solved by analytical methods.

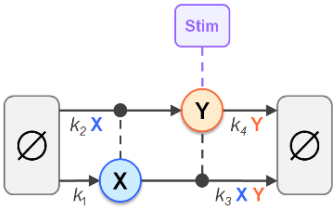
An alternative approach, however, is to derive an analytical solution for the steady state, then use numerical integration of the ODEs to investigate the impact of steady state on the dynamic response. A steady state solution for this system is any analytical relationship between $\bar{x}_1$, $\bar{x}_2$, and $k_1, \ldots, k_4$ that satisfies

$$0 = k_1 - k_3 \bar{x}_1 \bar{x}_2$$
$$0 = k_2 \bar{x}_1 - k_4 \bar{x}_2.$$

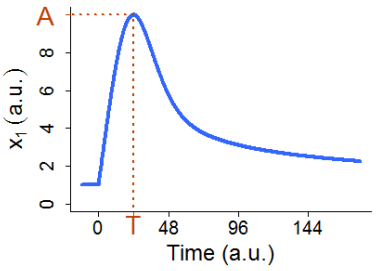
One obvious relationship is obtained by expressing the rates of synthesis, k_1 and k_2, in terms of the rates of degradation k_3 and k_4, and the steady state concentrations $\bar{x}_1$ and $\bar{x}_2$. This approach is prudent because measurements can often be found for protein concentrations and their rates of degradation [62]. Doing so gives

Figure 1.1: **A prototypical negative feedback model. A.** The activator X is constitutively expressed but catalytically degraded by an inhibitor, Y. The inhibitor is constitutively degraded but its synthesis requires X. **B.** A mass action model of the system in (A) was built and stimulated by instantaneously depleting Y. In response, X experiences a transient increase in concentration. This response can be characterized by two quantities: A and T. **C.** Mathematical representation of the perturbed steady state, indicated by the prime symbol. **D.** The effects of changes to the steady state concentrations of X and Y on stimulus-responsiveness. **E.** The effects of changes to the rates of turnover in X and Y on stimulus-responsiveness. **F.** Monte Carlo sampling of a heterogeneous steady state population and its response to stimulus.

A



B



C

$\bar{x}_1' = \phi_1 \bar{x}_1$

$\bar{x}_2' = \phi_2 \bar{x}_2$

$k_1' = k_3' \bar{x}_1' \bar{x}_2'$

$k_2' = k_3' \bar{x}_2' / \bar{x}_1'$

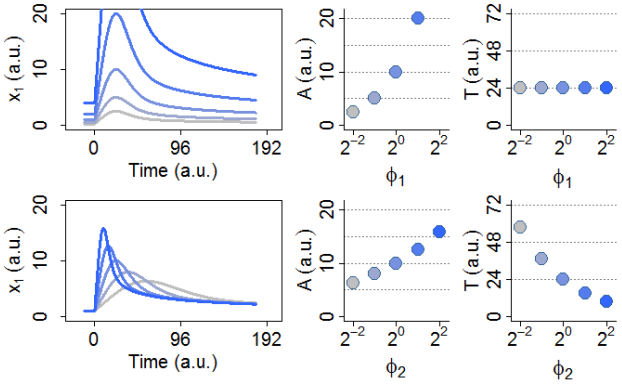
$k_3' = \theta_1 k_3$

$k_4' = \theta_2 k_4$

D

	ϕ_1	$\bar{x}_1$	$\bar{x}_2$	k_1	k_2	k_3	k_4	A	T
	0.25	0.25	1	0.14	2.52	5.67	6.31	2.5	23.9
	0.5	0.5	1	0.28	1.26	5.67	6.31	5	23.9
	1	1	1	0.57	0.63	5.67	6.31	10	23.9
	2	2	1	1.13	0.32	5.67	6.31	20	23.9
	4	4	1	2.27	0.16	5.67	6.31	40	23.9
					$\times 10^{-3}$	$\times 10^{-1}$	$\times 10^{-4}$		

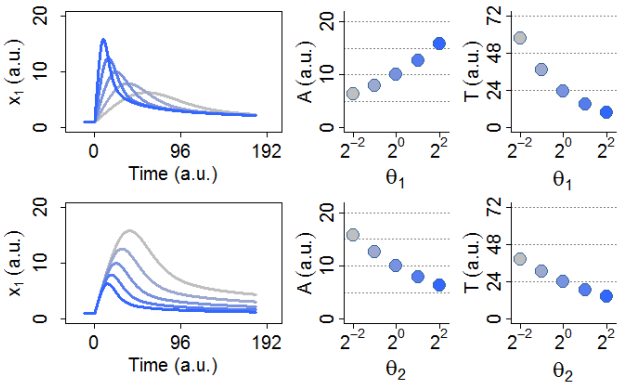
	ϕ_2	$\bar{x}_1$	$\bar{x}_2$	k_1	k_2	k_3	k_4	A	T
	0.25	1	0.25	0.14	0.16	5.67	6.31	6.34	57.7
	0.5	1	0.5	0.28	0.32	5.67	6.31	7.96	37.2
	1	1	1	0.57	0.63	5.67	6.31	10	23.9
	2	1	2	1.13	1.26	5.67	6.31	12.58	15.2
	4	1	4	2.27	2.52	5.67	6.31	15.83	9.7
					$\times 10^{-3}$	$\times 10^{-1}$	$\times 10^{-4}$		



E

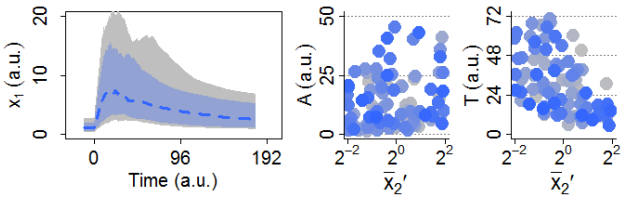
	θ_1	$\bar{x}_1$	$\bar{x}_2$	k_1	k_2	k_3	k_4	A	T
	0.25	1	1	0.14	0.63	1.42	6.31	6.34	57.7
	0.5	1	1	0.28	0.63	2.84	6.31	7.96	37.2
	1	1	1	0.57	0.63	5.67	6.31	10	23.9
	2	1	1	1.13	0.63	11.34	6.31	12.58	15.2
	4	1	1	2.27	0.63	22.69	6.31	15.83	9.7
					$\times 10^{-3}$	$\times 10^{-1}$	$\times 10^{-4}$		

	θ_2	$\bar{x}_1$	$\bar{x}_2$	k_1	k_2	k_3	k_4	A	T
	0.25	1	1	0.57	0.16	5.67	1.58	15.83	38.8
	0.5	1	1	0.57	0.32	5.67	3.15	12.58	30.5
	1	1	1	0.57	0.63	5.67	6.31	10	23.9
	2	1	1	0.57	1.26	5.67	12.61	7.96	18.6
	4	1	1	0.57	2.52	5.67	25.23	6.34	14.4
					$\times 10^{-3}$	$\times 10^{-1}$	$\times 10^{-4}$		



F

	Q	$\bar{x}_1$	$\bar{x}_2$	k_1	k_2	k_3	k_4	A	T
	0.15	0.36	0.41	0.11	0.11	2.16	2.58	3.35	12.12
	0.25	0.5	0.51	0.18	0.23	2.58	3.31	4.47	15.05
	0.5	1.17	0.81	0.54	0.63	4.85	6	9.65	27.85
	0.75	1.99	1.65	1.16	1.5	9.5	12.67	20.44	46.9
	0.85	2.68	2.36	1.84	2.59	12.08	16.48	25.91	58.89
					$\times 10^{-3}$	$\times 10^{-1}$	$\times 10^{-4}$		



$$k_1 = k_3 \bar{x}_1 \bar{x}_2 \tag{1.8}$$

$$k_2 = k_4 \frac{\bar{x}_1}{\bar{x}_2}. \tag{1.9}$$

With an analytical expression for the steady state (Figure 1.4C), the first analysis we might perform is to characterize the dependence of A and T on the steady state concentrations of X and Y. To do so we let $\bar{x}_1' = \phi_1 \bar{x}_1$ and $\bar{x}_2' = \phi_2 \bar{x}_2$ represent the perturbed steady state. Next we let ϕ_1 take values over some interval while holding $\phi_2 = 1$, then inter grate the ODEs and calculate A and T numerically (Figure 1.4D). Conversely, to investigate the dependence of A and T on the steady state concentration of Y we let ϕ_2 take values of some interval while holding $\phi_1 = 1$. The results show that A scales with the steady state concentration of X while T is unaffected. In contrast, an increased concentration of Y results in increased responsiveness, characterized by an increase in A and a reduction in T. This may seem counterintuitive, as Y is the "negative regulator" of X. However, an increase in Y with no concomitant decrease in X forces an increase in the synthesis and degradation of X, resulting in the observed increase in responsiveness.

This first type of analysis characterizes the relationship between the dynamic response and the steady state concentrations of X and Y. Given measurements for X and Y, we might conclude that we know a thing or two about how this system will respond to stimulation. However, the dynamic response can also depend on the kinetic rate constants, independently of X and Y. To characterize these dependencies, we need to find perturbations in the rate constants that do not affect $\bar{x}_1$ and $\bar{x}_2$. Since both $\bar{x}_1$ and $\bar{x}_2$ are independent parameters, this is relatively simple in our prototypical negative feedback model. Let $k_3' = \theta_1 k_3$ and $k_4' = \theta_2 k_4$ be the perturbed steady state. Notice from Equations 1.8 and 1.9 that a perturbation $\theta_1 \neq 1$ results in a change to the rates of synthesis and degradation of X such that its steady state concentration is preserved. The same holds true for θ_2 and Y. If we let θ_1 or θ_2 take values over an interval, we can numerically integrate the ODEs and calculate A and T as before (Figure 1.4). The results show that an increase in the turnover of X results in increased responsiveness, identical to

the result obtained by increasing the concentration of Y. This result suggests that steady state turnover of X is the major determinant of responsiveness. In contrast, as the turnover of Y is increased, X experiences a faster but weaker activation in response to stimulation. I will have more to say about these results in Chapter 3.

Finally, we may wish to entertain uncertainty in the parameters and see how that uncertainty propagates into the response. This effectively models a heterogeneous population of cells and can be used to identify steady state features that correlate well with the response. To do so, we let the independent parameters be described by probability density functions and use Monte Carlo sampling and numerical integration to calculate A and T. In Figure 1.4 we let

$$\phi_1 \sim \phi_2 \sim \theta_1 \sim \theta_2 \sim 2^{U(-2,2)}$$

then simulate the response of 100 randomly chosen samples. The results show that there is significant variation in A and T despite relatively little variation in the steady state itself. Interestingly, while some anticorrelation can be observed between $\bar{x}_2'$ and T, little to no correlation is observed between $\bar{x}_2'$ and A. This suggests that steady state features that are determinants of the dynamic response may nevertheless may not be good predictors of the response. All of these topics will be explore in much greater detail in the following chapters.

Chapter 1, in part, is a reprint of material as it appears in Methods in Cell Biology. Loriaux, Paul M; Hoffmann, Alexander, Elsevier 2012. Chapter 1 is also, in part, a reprint of material as it appears in Public Library of Science (PLoS) Computational Biology. Loriaux, Paul M; Tesler, Glenn; Hoffmann, Alexander, Public Library of Science 2013. The dissertation author was the primary investigator and author of these papers.

Chapter 2

Deriving analytical expressions for the steady states of mass action models

The steady states of cells affect their response to perturbation. However, no method exists to systematically characterize the relationship between steady state and response. While mathematical models are established tools for studying cellular responses, characterizing their relationship to the steady state requires that it have a parametric, or analytical, expression. For some models, this expression can be derived by the King-Altman method. However, King-Altman requires that no substrate act as an enzyme, and is therefore not applicable to most models of signal transduction. For this reason I developed *py*-substitution, a simple but general method for deriving analytical expressions for the steady states of mass action models. Where the King-Altman method is applicable, I show that *py*-substitution yields an equivalent expression, and at comparable efficiency. We use *py*-substitution to study the relationship between steady state and sensitivity to the anti-cancer drug candidate, dulanermin (recombinant human TRAIL). First, I use *py*-substitution to derive an analytical expression for the steady state of a published model of TRAIL-induced apoptosis. Next, I show that the amount of TRAIL required for cell death is sensitive to the steady state concentrations of procaspase 8 and its negative regulator, Bar, but not the other procaspase

molecules. This suggests that activation of caspase 8 is a critical point in the death decision process. Finally, I show that changes in the threshold at which TRAIL results in cell death is not always equivalent to changes in the time of death, as is commonly assumed. This work demonstrates that an analytical expression is a powerful tool for identifying steady state determinants of the cellular response to perturbation.

2.1 Introduction

Transient activation of signaling molecules is a hallmark of the cellular response to perturbation. As mentioned in Chapter 1, the dynamics of signaling molecules can encode information about the instigating stimulus [12, 63, 64], and these dynamics are affected by the steady state prior to perturbation [65, 23]. Non-genetic variation in the proteome, for example, is sufficient to explain variability in the sensitivity of HeLa cells to the pro-apoptotic ligand TRAIL [66]. Like other TNF superfamily members, TRAIL is a promising anti-cancer therapeutic [67]. Recombinant human TRAIL, or dulanermin, as well as antibodies raised against the TRAIL receptors DR4 and DR5, are currently in clinical trials [68]. To improve the efficacy of these and other drugs, understanding how sensitivity is affected by the cellular resting state is of great importance [69].

Mathematical models are powerful tools for characterizing the behavior of signaling systems in response to perturbation [70, 33, 38, 39]. Assuming conservation of mass, these models equate the change in concentration of a molecular species with the sum of reaction velocities that produce the species, minus the sum of those that consume it. The reactions themselves are often modeled by the *Law of Mass Action*. This law assumes that the velocity of a reaction is proportional to the product of the concentrations of its reactants. Since many signaling reactions are bimolecular, the resulting *mass balance* equations are non-linear in the concentrations. A system is at *steady state* if no species is consumed faster than it is produced, nor produced faster than it is consumed. By this formalism, the steady state of a signaling system is equivalent to the root of a non-linear system

of equations. Because of this, no universal method has been developed to identify the steady states of mass action models, despite their importance to basic and clinical research. As a result, even with the help of mathematical models, investigating the relationship between steady state and stimulus-responsiveness remains cumbersome.

With any model, simulating the response to perturbation often requires the system to be at steady state prior to perturbation. To achieve this, one of several techniques is currently used. The most common technique is to assume a "trivial" steady state where every reaction velocity is zero [71, 63]. While straightforward, this approach may not reflect biological reality, where tonic signaling is common [72, 73] and can strongly influence the response to perturbation [74, 75, 76]. A second technique is to approach the steady state asymptotically via numerical integration of the mass balance equations [10, 12, 39]. While this approach can yield non-trivial steady states, the number of integration steps required to reach the steady state may dominate the number of steps required to simulate the perturbation. Also, identifying the parameter values that result in a desired steady state is an inverse problem that requires non-linear optimization. For these reasons, numerical derivation of the steady state is impractical when characterizing its effect on the response to perturbation, and an analytical expression is required instead.

The best-known method for deriving analytical expressions for the steady states of mass action models was developed by King and Altman in 1956 [77]. This method assumes that all molecular species can be divided into enzymes and substrates, that no enzyme is itself a substrate, and that all substrates remain constant over the time-scale of steady state formation [78]. A number of improvements have been made to the King-Altman method over the years [79, 80, 81]. Many of these are now implemented in the Matlab application, *KAPattern* [82]. The King-Altman methodology was also recently formalized using concepts from algebraic geometry [78, 83], and extended to layered signaling cascades [84] and post-translational modification networks [85]. Despite these improvements, however, these methods do not extend to mass action models with arbitrary reaction structure, as is common in contemporary models of signaling systems. Further-

more, only the King-Altman method has been reduced to practice.

For these reasons we developed *py*-substitution, a simple, algebraic method for deriving steady state expressions for mass action models with arbitrary structure. Our method can be explained using concepts from linear algebra, and full code has been provided for all examples in this manuscript, implemented in either Matlab or Maple. A particular benefit of *py*-substitution is that it affords considerable flexibility when selecting independent quantities for the steady state expression. Often, this permits explicit derivation of kinetic rate constants from steady state concentration measurements. More generally, it allows independent quantities to be chosen that maximize incorporation of known or measured parameter values. This not only simplifies model fitting, but typically reduces the total number of parameters required as well. We compare *py*-substitution to the King-Altman method and show that, where King-Altman is applicable, the two methods yield equivalent results. Computationally, however, we find that our method is more efficient, and, because *py*-substitution does not require a particular reaction structure, more general than King-Altman.

Finally, we use *py*-substitution to derive a steady state expression for a recent model of apoptosis induced by the death-receptor ligand TRAIL [71]. We find that incorporation of a non-trivial steady state changes the qualitative behavior of the model. Specifically, tonic signaling desensitizes the system to low doses of TRAIL, while high doses of TRAIL still result in the "snap-action" signaling dynamics indicative of cell death. We then systematically alter the steady state and show that changes in steady state affect the threshold at which TRAIL results in death. We find that the threshold is highly sensitive to the steady state abundances of procaspase 8 and its negative regulator, Bar, but not the other procaspase molecules. This suggests that the activation of caspase 8 is a critical point in the cell death decision. Finally, without recourse to a model that is tolerant to low doses of TRAIL, a common practice is to approximate the sensitivity to TRAIL by the time at which death occurs. Using our tonic signaling model, we show that these two metrics are not universally equivalent. Caution should therefore be taken when equating the dynamics of cell death with the probability that

death occurs.

2.2 Methods

In this section we describe the process for deriving analytical expressions for the steady states of mass action models using *py*-substitution. First we review existing methods for deriving analytical expressions for the steady states of mass action models. Afterwards, we describe *py*-substitution using some formal concepts from algebra. In the results section we provide several examples, beginning with a version of the classical Michaelis-Menten model of enzyme action. All code for these examples, as well as detailed instructions for use and full transcripts of the output, are provided in Protocol S1 and on our website, http://signalingsystems.ucsd.edu/pg/pysub.

2.2.1 Prior Work

Recall from Section 1.3 that we are interested in solutions to the steady state equation,

$$\mathbf{S}\overline{\mathbf{v}} = \mathbf{0}, \tag{2.1}$$

where $\mathbf{S}$ is the stoichiometric matrix and $\mathbf{v}$ is a vector of reaction velocities. By convention we use the overline to denote vectors that satisfy steady state. Equation 2.1 is often seen in flux balance analysis [51, 86, 87, 88]. Here $\overline{\mathbf{v}}$ is a real-valued vector and is calculated numerically. However, prior work has shown that Equation 2.1 can also be used to calculate a vector of rate constants from a vector of steady state concentrations [89]. Let $\mathbf{k} \in \mathbb{R}_0^{d_k}$ be the vector with elements $(\mathbf{k})_j = k_j$. Let $\mathbf{P}_k$ be the diagonal matrix with elements $(\mathbf{P}_k)_{j,j} = (\partial v_j / \partial k_j)$. The vector $\mathbf{v}$ can then be expressed as

$$\mathbf{v} = \mathbf{P}_k \mathbf{k}. \tag{2.2}$$

Substituting Equation 2.2 into Equation 2.1 and solving for $\mathbf{k}$ yields the k-cone [89] — equivalently, the left null space of the matrix product $\mathbf{SP}_k$. Given a basis for this null space and a vector of steady state concentrations, a vector of rate constants can be calculated that satisfies Equation 2.1. While this approach is useful for deriving kinetic parameters from metabolomic measurements, it is less well suited to signaling systems where transient and low-abundance species confound accurate measurement of the concentrations.

If the velocity of every $r_j \in \mathcal{R}$ is homogeneous of degree 1 in $x_1, \ldots, x_{d_x}$, then an analogous approach allows $\mathbf{v}$ to be expressed in terms of $\mathbf{x}$. We call models that satisfy this condition *linear models*. An alternative, stoichiometric definition for a linear model is given by the following,

$$\forall r_j \in \mathcal{R},\ \sum_{i_1=1}^{d_k} s_{i_1,j}^{in} = \sum_{i_2=1}^{d_k} s_{i_2,j}^{out} = 1. \tag{2.3}$$

Equation 2.3 requires that every reaction defines a transition from exactly one time-varying species to another. Let $\mathbf{P}_x$ be the matrix with elements $(\mathbf{P})_{i,j} = (\partial v_i / \partial x_j)$. If $\mathbf{v}$ is a vector of linear reaction velocities, it can likewise be expressed as

$$\mathbf{v} = \mathbf{P}_x \mathbf{x}. \tag{2.4}$$

Substituting Equation 2.4 into Equation 2.1 results in the matrix product $\mathbf{SP}_x$, also called the *Jacobian matrix* [90]. Given a basis for the null space of the Jacobian, a vector of steady state concentrations can be calculated from a vector rate constants.

For linear models, an alternative, graphical method for deriving expressions for the steady state species concentrations was introduced by King and Altman in 1956 [77]. Notice that Equation 2.3 permits a two-dimensional indexing of the rate constants,

$$k'_{i_1,i_2} = \begin{cases} k_j & \text{if } \exists\ r_j \in \mathcal{R} \colon s_{i_1,j}^{in} = s_{i_2,j}^{out} = 1 \\ 0 & \text{otherwise.} \end{cases} \tag{2.5}$$

We call k'_{i_1,i_2} a *transition* rate constant since the product $k'_{i_1,i_2} x_{i_1}$ defines the rate

of transition from species x_{i_1} to x_{i_2}. Substituting Equation 2.5 into Equations 1.1 and 1.2 gives

$$\frac{dx_{i_1}}{dt} = \sum_{i_2=1}^{d_x} k'_{i_2,i_1} x_{i_2} - x_{i_1} \sum_{i_2=1}^{d_x} k'_{i_1,i_2}. \tag{2.6}$$

By defining the matrix $\mathbf{K}$ with elements

$$(\mathbf{K})_{i_1,i_2} = \begin{cases} k'_{i_2,i_1} & \text{if } i_1 \neq i_2, \\ \sum_{m \neq i_1} -k'_{i_1,m} & \text{if } i_1 = i_2, \end{cases} \tag{2.7}$$

the steady state equation becomes

$$\mathbf{K}\mathbf{x} = \mathbf{0}. \tag{2.8}$$

Note that $\mathbf{K}$ is simply the Jacobian matrix for a linear model, $\mathbf{K} = \mathbf{S}\mathbf{P}_x$. The general solution to Equation 2.8 was found in [77] to be the vector $\overline{\mathbf{x}}$ with elements

$$(\overline{\mathbf{x}})_i = \frac{\mathbf{M}_i}{\sum_{i_2=1}^{d_x} \mathbf{M}_{i_2}}, \tag{2.9}$$

where $\mathbf{M}_i$ is the i^{th} minor of $\mathbf{K}$, formed by removing its i^{th} row and column and computing its determinant. For sufficiently small systems, Equation 2.9 can be solved directly using modern mathematical computing software [91]. Prior to the advent of modern computers, King and Altman realized that the minors can also be derived by graph theoretic means. Note that for a linear model, $\mathcal{A}$ and $\mathcal{R}$ imply a directed graph,

$$\mathcal{G} = (\mathcal{A}, \mathcal{R}), \tag{2.10}$$

where each $a_i \in \mathcal{A}$ defines a vertex and each $r_j \in \mathcal{R}$ defines an edge between vertices a_{i_1} and a_{i_2} (provided i_1 and i_2 are such that $s^{in}_{i_1,j} = s^{out}_{i_2,j} = 1$). The King-Altman method enumerates for each species $a_i \in \mathcal{A}$ the set $\mathcal{S}_i$ of simple connected subgraphs

$$\mathcal{S}_i = \{\mathcal{G}' = (\mathcal{X}, \mathcal{R}') \colon \mathcal{R}' \subset \mathcal{R},\ |\mathcal{R}'| = d_x - 1\}$$

where vertex a_i has out-degree 0 and all other vertices have out-degree 1 [79, 80]. These are the directed spanning trees of G, with all edges directed towards root a_i. A subgraph $\mathcal{G}'$ is called a King-Altman *pattern*. The minor $\mathbf{M}_i$ can then be expressed as

$$\mathbf{M}_i = \sum_{\mathcal{G}' \in \mathcal{S}_i} \prod_{r_j \in \mathcal{R}'} k_j, \tag{2.11}$$

where $k_j = k'_{i_1,i_2}$ is the transition rate constant between species a_{i_1} and a_{i_2}. For a more thorough derivation of the King-Altman method, see [92].

Of course, many biochemical reactions are bimolecular. By Equation 1.1, the velocity of a bimolecular reaction is degree 2 in $x_1, \ldots, x_{d_x}$. To preserve linearity, one can assume the concentration of one reactant is so high as to be effectively constant. This concentration is incorporated into the kinetic rate constant, and the techniques described above can still be used to solve Equation 1.3. If this assumption fails, then Equation 1.2 describes a polynomial in $x_1, \ldots, x_{d_x}$ with coefficients in $\mathbb{Q}[k_1, \ldots, k_{d_k}]$. In this case the solutions to Equation 1.3 form an algebraic variety. Deriving an expression for the steady state of a non-linear model thus requires finding a parameterization of the variety [93]. One way to achieve this is to calculate a Gröbner basis for the ideal generated by $\dot{x}_1, \ldots, \dot{x}_{d_k}$ and eliminate variables [94, 95]. Alternatively, if the model displays certain structural properties, variables can be eliminated by identifying conservation relationships. The best-known example of this is when $\{\mathcal{A}, \mathcal{R}\}$ defines a cascade of post-translational modifications. In this case, enzyme-substrate intermediates can be eliminated and the variety can be parameterized by rational functions of the free enzyme concentrations with coefficients in $\mathbb{Q}(k_1, \ldots, k_{d_k})$ [78, 84]. Although these methods do not require linearity, calculating a Gröbner basis can be computationally intractable, while identifying conservation relationships can be difficult for models of arbitrary reaction structure.

2.2.2 *Py*-substitution

Py-substitution allows mass action models to be solved using simple linear algebra. We make use of the following observations: (a) $\dot{x}_i$ is always homogeneous of degree 1 in $k_1, \ldots, k_{d_k}$, and (b) $\dot{x}_i$ is often no greater than degree 2 in $x_1, \ldots, x_{d_x}$. If a subset of elements in $\mathcal{K} \cup \mathcal{X}$ can be found on which every $\dot{x}_i$ has only linear dependence, then Equation 2.1 can be solved using linear methods.

To begin, we define sets of symbolic variables $\mathcal{P} = \{p_1, \ldots, p_{d_p}\}$ and $\mathcal{Y} = \{y_1, \ldots, y_{d_y}\}$ such that $d_p + d_y = d_k + d_x$ and $d_y \geq$ rank $\mathbf{S}$. We then relabel, or map, every element in $\mathcal{K} \cup \mathcal{X}$ to a unique element in $\mathcal{P} \cup \mathcal{Y}$ so that every $\dot{x}_i$ is linear in $\mathcal{Y}$. By Equations 1.1 and 1.2 this requires that all v_j are linear in $\mathcal{Y}$. Variables that we want to remain independent, as well as variables on which $\dot{x}_i$ has non-linear dependence, should be mapped to $\mathcal{P}$. As we shall see, there is considerable flexibility in choosing this map.

Let $\mathcal{K}$ and $\mathcal{X}$ be partitioned into disjoint (but possibly empty) subsets $\mathcal{K} = \mathcal{K}_p \cup \mathcal{K}_{lin}$ and $\mathcal{X} = \mathcal{X}_p \cup \mathcal{X}_{lin}$. We define ψ_p to be a bijective map

$$\psi_p \colon \begin{cases} \mathcal{K}_p \cup \mathcal{X}_p \to \mathcal{P} \\ \mathcal{K}_{lin} \cup \mathcal{X}_{lin} \to \mathcal{Y}, \end{cases}$$

and extend it homomorphically over $\mathbb{Q}[\mathcal{K}][\mathcal{X}]$. Our linearity restriction is to consider maps of this form such that

$$\psi_p(v_j) = y_n \prod_{m=1}^{d_p} p_m^{h'_{j,m}} \tag{2.12}$$

for some $y_n \in \mathcal{Y}$. For $p_m = \psi_p(x_i)$, the exponent is $h'_{j,m} = h_{i,j}$. For $p_m = \psi_p(k_j)$, the exponent $h'_{j,m} = 1$. In words, ψ_p defines a change of variables such that $\psi_p(v_j)$ is homogeneous of degree 1 in $y_1, \ldots, y_{d_y}$. By Equation 1.2, $\psi_p(\dot{x}_i)$ becomes a homogeneous polynomial of degree 1 in $y_1, \ldots, y_{d_y}$ with coefficients in $\mathbb{Q}[p_1, \ldots, p_{d_p}]$. We can now write

$$\psi_p(\mathbf{v}) = \mathbf{P}\mathbf{y}, \tag{2.13}$$

where $\mathbf{P}$ is the $d_k \times d_y$ Jacobian matrix with elements $(\mathbf{P})_{ij} = (\partial v_i / \partial y_j)$. Here and elsewhere we use the notation $\psi(\mathbf{v}) = \mathbf{w}$ to mean that $\mathbf{w}$ is the vector formed by applying the function ψ element-wise to $\mathbf{v}$. Note that the trivial partition $\mathcal{K}_{lin} = \mathcal{K}$ and $\mathcal{X}_p = \mathcal{X}$ recovers the k-cone procedure described above. For the remainder of this section, we treat j as an index over $1, \ldots, d_y$. Substituting Equation 2.13 into Equation 2.1 gives

$$\mathbf{C}\mathbf{y} = \mathbf{0} \tag{2.14}$$

where $\mathbf{C} = \mathbf{S}\mathbf{P}$ is called the *coefficient matrix*. The solution to Equation 2.14 is precisely the null space of $\mathbf{C}$. Let $\mathbf{N}$ be a matrix whose columns form a basis for this null space. Let d_q be the number of columns in $\mathbf{N}$. By the rank-nullity theorem, we have

$$d_q = \text{ncols } \mathbf{C} - \text{rank } \mathbf{C}, \tag{2.15}$$

where ncols $\mathbf{C} = d_y$ is the number of columns in $\mathbf{C}$. Furthermore, because $\psi_p(\mathbf{v})$ is linear in $\mathbf{y}$ and ψ_p^{-1} exists, the matrix $\mathbf{P}$ must be full rank. By the properties of the rank, we can write

$$\text{rank } \mathbf{C} = \text{rank } \mathbf{S}\mathbf{P} = \text{rank } \mathbf{S}. \tag{2.16}$$

Together, Equations 2.15 and 2.16 give

$$d_q = d_y - \text{rank } \mathbf{S}, \tag{2.17}$$

thus calling for the constraint $d_y \geq$ rank $\mathbf{S}$. This, in conjunction with Equation 2.12, are the only constraints on ψ_p. If we now let $\overline{\mathbf{y}}$ be some linear combination of the basis vectors,

$$\overline{\mathbf{y}} = \mathbf{N}\mathbf{q}, \tag{2.18}$$

then $\overline{\mathbf{y}}$ satisfies Equation 2.14 and steady state is achieved. In general, Equation 2.18 is underdetermined. Equation 2.18 therefore implies a partition of $\mathcal{Y}$ into independent variables (denoted $\mathcal{Y}_q$) and dependent variables (denoted $\mathcal{Y}_q^c$). We will now describe this partition by a second mapping function, ψ_y.

Recall that a basis for the null space of $\mathbf{C}$ can be constructed from $\mathbf{C}_{\text{rref}}$, the reduced row echelon form of $\mathbf{C}$. Let $\mathbf{c}_j$ be the j^{th} column of $\mathbf{C}_{\text{rref}}$. If $\mathbf{c}_j$ contains a pivot position, then y_j is a dependent variable. If $\mathbf{c}_j$ does not contain a pivot, then y_j is free, or independent. Let

$$\begin{aligned} \mathcal{Y}_q &= \{y_j \in \mathcal{Y} \colon \text{column } \mathbf{c}_j \text{ does } \textit{not} \text{ contain a pivot}\} \\ \mathcal{Y}_q^c &= \{y_j \in \mathcal{Y} \colon \text{column } \mathbf{c}_j \text{ contains a pivot}\}. \end{aligned} \tag{2.19}$$

Let d_q be the cardinality of $\mathcal{Y}_q$. Enumerate these variables as $\mathcal{Y}_q = \{y_{j_1}, y_{j_2}, \ldots, y_{j_{d_q}}\}$, with $j_1 < \cdots < j_{d_q}$. For every $\mathbf{c}_j$ not containing a pivot, there is a basis vector $\mathbf{n}_k$ (related by $j = j_k$) whose j^{th} element equals 1 and whose elements in positions $> j$ are 0. By Equation 2.18, this gives an independent parameter, $y_j = (\mathbf{q})_k = q_k$. Equation 2.18 thus defines a function $\psi_y \colon y_j \mapsto q_k$. Let $\mathcal{Q} = \{q_1, q_2, \ldots, q_{d_q}\}$ be the set of independent parameters. If column $\mathbf{c}_j$ does contain a pivot, then y_j depends on variables in $\mathcal{P} \cup \mathcal{Q}$, giving $\psi_y \colon y_j \mapsto f_j(\mathcal{P}, \mathcal{Q}) \in \text{span}_{\mathbb{Q}(\mathcal{P})}(\mathcal{Q})$ where $f_j(\mathcal{P}, \mathcal{Q})$ is the specific function resulting from the row operations used to reduce $\mathbf{C}$ to $\mathbf{C}_{\text{rref}}$. Equation 2.18 can now be described in its entirety by the mapping function ψ_y,

$$\psi_y \colon \begin{cases} \mathcal{P} \to \mathcal{P} \ \ (\text{identity}) \\ \mathcal{Y}_q \to \mathcal{Q} \\ \mathcal{Y}_q^c \to \text{span}_{\mathbb{Q}(\mathcal{P})}(\mathcal{Q}). \end{cases} \tag{2.20}$$

The notation $\psi \colon \mathcal{P} \to \mathcal{P}$ (identity) indicates that $\psi(p) = p$ for every $p \in \mathcal{P}$. Note that we define $\text{span}_{\mathcal{F}}(\mathcal{Q})$ as the set of all linear combinations $a_1 q_1 + a_2 q_2 + \cdots$, where $a_1, a_2, \cdots \in \mathcal{F}$ and $q_1, q_2, \ldots$ are distinct elements of $\mathcal{Q}$. $\mathbb{Q}[\mathcal{P}]$ is the set of all polynomials in variables $\mathcal{P}$ with rational numbers as coefficients. $\mathbb{Q}(\mathcal{P})$ is the

field of fractions of $\mathbb{Q}[\mathcal{P}]$: any $f \in \mathbb{Q}[\mathcal{P}]$ can be expressed as $f = g_1/g_2$, where g_1, $g_2 \in \mathbb{Q}[\mathcal{P}]$.

As with ψ_p, there is some flexibility in choosing how $\mathcal{Y}$ is partitioned into free variables, $\mathcal{Y}_q$, and dependent variables, $\mathcal{Y}_q^c$. A different indexing of the variables in $\mathcal{Y}$ simultaneously permutes the vector $\mathbf{y}$ and the columns of $\mathbf{C}$. This leads to different reduced row echelon forms, with different partitions into free and dependent variables. The null space basis obtained by reducing $\mathbf{C}$ to $\mathbf{C}_{\text{rref}}$ greedily classifies low-numbered columns as dependent columns when possible, or free columns when not possible. Quantities in $\mathcal{Y}$ for which good numerical estimates exist should therefore be assigned to higher indices. These quantities are favored, but not guaranteed, to be mapped to independent parameters. Quantities for which good numerical estimates do not exist should be assigned to low indices in $\mathcal{Y}$.

Finer control over the partition of $\mathcal{Y}$ into dependent and independent parameters is possible by working directly with $\mathbf{C}_{\text{rref}}$ or $\mathbf{N}$. Let $\mathcal{Y}_q' = \{y_{j_1}, \ldots, y_{j_{d_q}}\}$ be the set of d_q elements in $\mathcal{Y}$ that we want mapped to $\mathcal{Q}$. Let $\mathbf{N}'$ be the square matrix formed by rows $j_1, \ldots, j_{d_q}$ of $\mathbf{N}$. To map $\mathcal{Y}_q'$ to $\mathcal{Q}$ requires that we find a vector $\mathbf{q}'$ such that

$$\mathbf{N}'\mathbf{q}' = \mathbf{q},$$

where $\mathbf{q}$ is the vector with elements $(\mathbf{q})_k = q_k$. Solving for $\mathbf{q}'$ gives

$$\mathbf{q}' = (\mathbf{N}')^{-1}\mathbf{q}. \tag{2.21}$$

Thus, for a given map ψ_p, not all partitions of $\mathcal{Y}$ into $\mathcal{Y}_q$ and $\mathcal{Y}_q^c$ are possible, but only those for which $\det(\mathbf{N}') \neq 0$. An example of this can be seen in the file "fum2.m" in Supporting Protocol S1, discussed below.

Next let $\mathcal{K}_q = \{k \in \mathcal{K}\colon \psi_{py}(k) \in \mathcal{Q}\}$, and $\mathcal{K}_y = \mathcal{K}_{lin} \backslash \mathcal{K}_q$. Let $\mathcal{X}_q$ and $\mathcal{X}_y$ be defined analogously. The composition $\psi_{py} = (\psi_y \circ \psi_p)$ captures the entire process of linearizing $\mathbf{v}$ with the function ψ_p, solving the linear system $\mathbf{S}\psi_p(\mathbf{v}) = \mathbf{0}$, and taking an arbitrary combination of solution space basis vectors:

$$\psi_{py}\colon \begin{cases} \mathcal{K}_p \cup \mathcal{X}_p \to \mathcal{P} \\ \mathcal{K}_q \cup \mathcal{X}_q \to \mathcal{Q} \\ \mathcal{K}_y \cup \mathcal{X}_y \to \mathrm{span}_{\mathbb{Q}(\mathcal{P})}(\mathcal{Q}). \end{cases}$$

Applying ψ_{py} to the sets $\mathcal{K}$ and $\mathcal{X}$ results in a parametric description of the steady state that is typically the most useful: every element in $\mathcal{K}$ or $\mathcal{X}$ is mapped to an element in $\mathcal{P}$ or $\mathcal{Q}$, or a function in $\mathrm{span}_{\mathbb{Q}(\mathcal{P})}(\mathcal{Q})$. Assigning numerical values to elements in $\mathcal{P}$ and $\mathcal{Q}$ results in elements in $\mathrm{span}_{\mathbb{Q}(\mathcal{P})}(\mathcal{Q})$ taking values that satisfy the steady state equation. In some cases we may wish to reverse the substitution so that functions of variables $\mathcal{P} \cup \mathcal{Q}$ are mapped back to functions of $\mathcal{K} \cup \mathcal{X}$. To do so, let $\mathcal{K}_{pq} = \mathcal{K}_p \cup \mathcal{K}_q$ and $\mathcal{X}_{pq} = \mathcal{X}_p \cup \mathcal{X}_q$. Let ψ_q^{-1} be the inverse of ψ_y restricted to the independent parameters, $\mathcal{P} \cup \mathcal{Q}$.

$$\psi_q^{-1}\colon \begin{cases} \mathcal{P} \to \mathcal{P} \quad \text{(identity)} \\ \mathcal{Q} \to \mathcal{Y}_q. \end{cases}$$

The composition of ψ_p^{-1} and ψ_q^{-1} now defines a map from the set of independent parameters to their counterparts in $\mathcal{K}$ and $\mathcal{X}$,

$$\psi_{pq}^{-1} = (\psi_p^{-1} \circ \psi_q^{-1})\colon \begin{cases} \mathcal{P} \to \mathcal{K}_p \cup \mathcal{X}_p \\ \mathcal{Q} \to \mathcal{K}_q \cup \mathcal{X}_q. \end{cases}$$

If we extend ψ_{pq}^{-1} to $f \in \mathrm{span}_{\mathbb{Q}(\mathcal{K}_p, \mathcal{X}_p)}(\mathcal{K}_q, \mathcal{X}_q)$ homomorphically, we can compose ψ_{pq}^{-1} with ψ_{py},

$$\psi_{ss} = (\psi_{pq}^{-1} \circ \psi_{py})\colon \begin{cases} \mathcal{K}_{pq} \to \mathcal{K}_{pq} \quad \text{(identity)} \\ \mathcal{X}_{pq} \to \mathcal{X}_{pq} \quad \text{(identity)} \\ \mathcal{K}_y \cup \mathcal{X}_y \to \mathrm{span}_{\mathbb{Q}(\mathcal{K}_p, \mathcal{X}_p)}(\mathcal{K}_q, \mathcal{X}_q) \end{cases}$$

The function ψ_{ss} then defines a map for which

$$\mathbf{S}\psi_{ss}(\mathbf{v}) = \mathbf{S}\overline{\mathbf{v}} = \mathbf{0},$$

where steady state velocities in $\overline{\mathbf{v}}$ are in terms of elements in $\mathcal{K}$ and $\mathcal{X}$. A visual overview of the *py*-substitution method is given in Figure 2.1.

2.3 Results

2.3.1 *Py*-substitution permits flexible derivation of a steady state solution

An important goal in developing *py*-substitution was that it be generally applicable to any model whose reaction rates obey mass action kinetics. This requires that the independent quantities be chosen freely among the species concentrations and reaction rate constants, and that non-linear rate equations do not confound the derivation of a steady state expression. To demonstrate these capabilities we consider an open-system analog of the classical Michaelis-Menten model of enzyme kinetics (OMM, see also Figure 2.2). Substrate synthesis and product degradation allow this system to achieve a non-trivial steady state $\overline{\mathbf{v}} \neq \mathbf{0}$, which we derive here using four different substitution strategies. The set $\mathcal{R}$ of reactions for this model is given by

$$\begin{aligned}
r_1 &: x_1 + x_2 \xrightarrow{v_1} x_3 \\
r_2 &: x_3 \xrightarrow{v_2} x_2 + x_1 \\
r_3 &: x_3 \xrightarrow{v_3} x_2 + x_4 \\
r_4 &: \emptyset \xrightarrow{v_4} x_1 \\
r_5 &: x_4 \xrightarrow{v_5} \emptyset
\end{aligned}$$

The symbol $\emptyset$ represents a *source* or *sink* for mass and is not modeled by a time-varying species. From the set $\mathcal{R}$ we derive the stoichiometric matrix and reaction velocity vector,

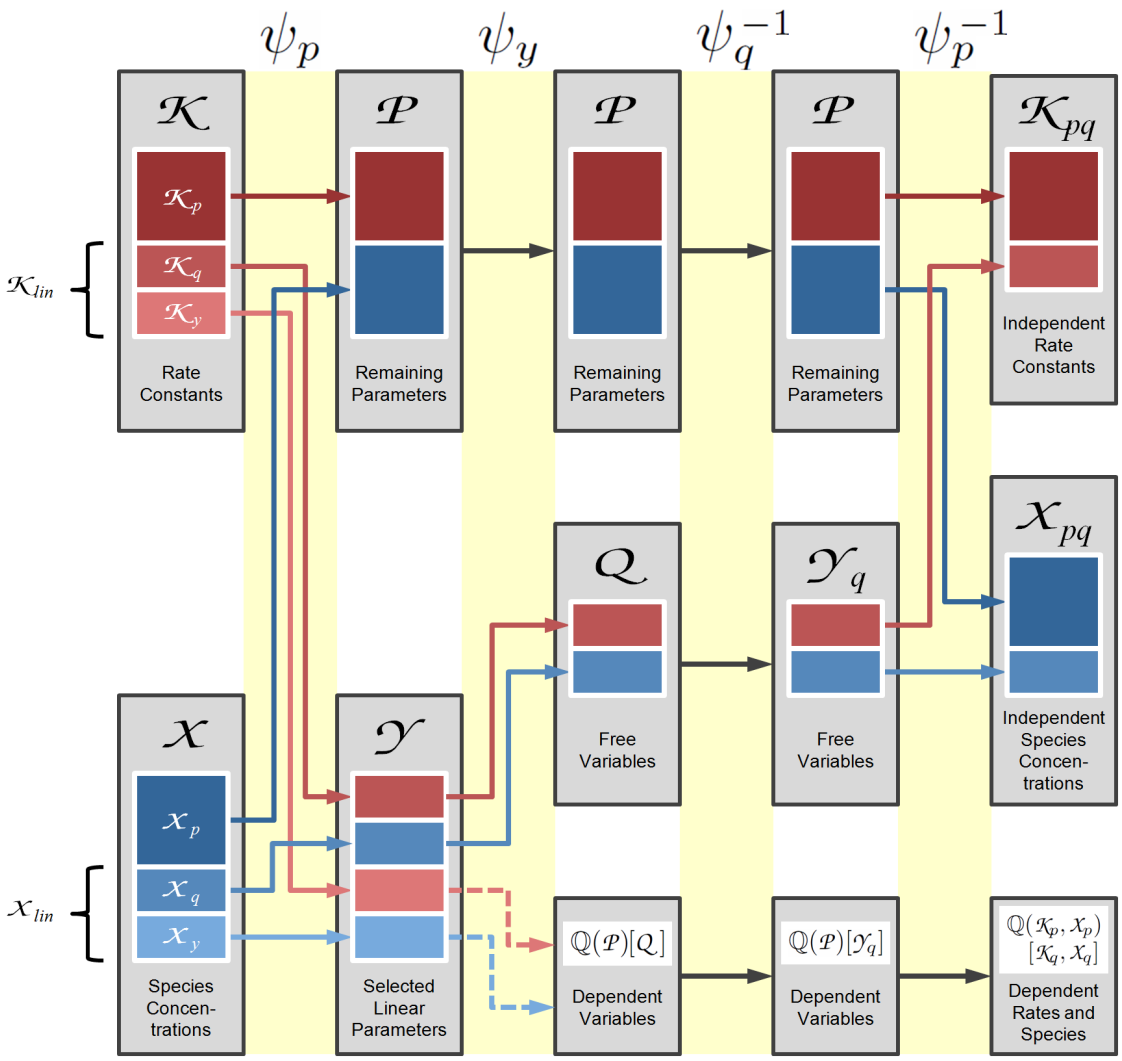


Figure 2.1: **Overview of the *py*-substitution method.** Quantities in a mass action model can be separated into kinetic rate constants (set $\mathcal{K}$, red) and species abundances or concentrations (set $\mathcal{X}$, blue). From $\mathcal{K} \cup \mathcal{X}$ a subset $\mathcal{K}_{lin} \cup \mathcal{X}_{lin}$ is selected on which all reaction velocities have only linear dependence. A function ψ_p maps these to elements in $\mathcal{Y}$ and the remaining $\mathcal{K}_p \cup \mathcal{X}_p$ to elements in $\mathcal{P}$. A second function ψ_y imposes the relations $\psi_p(\dot{\mathbf{x}}) = \mathbf{0}$ by expressing dependent variables in $\mathcal{Y}$ in terms of independent parameters $\mathcal{P} \cup \mathcal{Q}$. A third function, ψ_q^{-1}, is the inverse of ψ_y restricted to the independent parameters. The composition of ψ_p^{-1} with ψ_q^{-1} results in variables in $\mathcal{K}_y \cup \mathcal{X}_y$ being expressed in terms of variables in $\mathcal{K}_{pq} \cup \mathcal{X}_{pq}$, such that steady state is achieved. In the diagram, solid arrows are isomorphisms while dashed arrows are homomorphisms that replace dependent variables by equivalent expressions in independent parameters.

$$\mathbf{S} = \begin{bmatrix} -1 & 1 & 0 & 1 & 0 \\ -1 & 1 & 1 & 0 & 0 \\ 1 & -1 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 & -1 \end{bmatrix}, \quad \mathbf{v} = \begin{bmatrix} k_1 x_1 x_2 \\ k_2 x_3 \\ k_3 x_3 \\ k_4 \\ k_5 x_4 \end{bmatrix}.$$

By Equation 1.4 this results in the following system of equations,

$$\begin{aligned} dx_1/dt &= -k_1 x_1 x_2 + k_2 x_3 + k_4 \\ dx_2/dt &= -k_1 x_1 x_2 + k_2 x_3 + k_3 x_3 \\ dx_3/dt &= k_1 x_1 x_2 - k_2 x_3 - k_3 x_3 \\ dx_4/dt &= k_3 x_3 - k_5 x_4 \end{aligned}$$

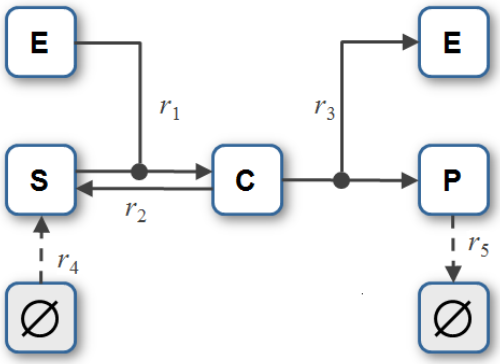
for which we now derive functions ψ_{ss} such that $\mathbf{S}\psi_{ss}(\mathbf{v}) = \mathbf{S}\overline{\mathbf{v}} = \mathbf{0}$.

Homogeneous substitution: steady state concentrations do not uniquely determine reaction rate constants

The most straightforward substitution strategy is to let $\mathcal{K}_{lin} = \mathcal{K}$ and $\mathcal{X}_p = \mathcal{X}$. The corresponding function ψ_p maps

$$\begin{aligned} k_1 &\mapsto y_1 & x_1 &\mapsto p_1 \\ k_2 &\mapsto y_2 & x_2 &\mapsto p_2 \\ k_3 &\mapsto y_3 & x_3 &\mapsto p_3 \\ k_4 &\mapsto y_4 & x_4 &\mapsto p_4 \\ k_5 &\mapsto y_5 \end{aligned}$$

See "omm1.m.trace.pdf" in Protocol S1 for details of this partition and all subsequent steps. Applying ψ_p to $\mathbf{v}$ results in a reaction velocity vector that is linear in $\mathbf{y}$, as required by Equation 2.13,



Var	Symbol	Description
x_1	E	Enzyme
x_2	S	Substrate
x_3	C	Complex
x_4	P	Product

Figure 2.2: **An open system analog of the classical Michaelis-Menten model for enzyme catalysis.** Enzyme and substrate bind to form an intermediate complex, followed by catalysis and dissociation of the product. The substrate is synthesized by a zero-order reaction, r_4, and the product is degraded by a first-order reaction, r_5. See "omm1.m" in Protocol S1 for a complete description of the model.

$$\psi_p(\mathbf{v}) = \mathbf{P}\mathbf{y} = \begin{bmatrix} p_1p_2 & 0 & 0 & 0 & 0 \\ 0 & p_3 & 0 & 0 & 0 \\ 0 & 0 & p_3 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & p_4 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \end{bmatrix}.$$

The resulting coefficient matrix is given by

$$\mathbf{C} = \mathbf{S}\mathbf{P} = \begin{bmatrix} -p_1p_2 & p_3 & p_3 & 0 & 0 \\ -p_1p_2 & p_3 & 0 & 1 & 0 \\ p_1p_2 & -p_3 & -p_3 & 0 & 0 \\ 0 & 0 & p_3 & 0 & -p_4 \end{bmatrix},$$

which row reduces to

$$\mathbf{C} \sim \mathbf{C}_{\text{rref}} = \begin{bmatrix} 1 & -p_3/(p_1p_2) & 0 & 0 & -p_4/(p_1p_2) \\ 0 & 0 & 1 & 0 & -p_4/p_3 \\ 0 & 0 & 0 & 1 & -p_4 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}. \tag{2.22}$$

From Equation 2.22, we observe that rank $\mathbf{C} = 3$. Thus, of the 9 degrees of freedom in this system (5 rate constants plus 4 species concentrations), 3 will have values that are constrained by Equation 2.1. Since our substitution strategy only identifies 4 independent parameters, 2 additional elements mapped to $\mathcal{Y}$ must in fact be independent as well. These elements can be identified by the columns in $\mathbf{C}_{\text{rref}}$ that do not contain pivots, namely columns 2 and 5. To see this, note that Equation 2.22 yields the following basis for the null space of $\mathbf{C}$,

$$\mathbf{N} = \begin{bmatrix} p_3/(p_1p_2) & p_4/(p_1p_2) \\ 1 & 0 \\ 0 & p_4/p_3 \\ 0 & p_4 \\ 0 & 1 \end{bmatrix}.$$

Letting $\mathbf{q} = [q_1, q_2]^{\mathrm{T}}$, Equation 2.18 gives

$$\overline{\mathbf{y}} = \begin{bmatrix} (q_1p_3 + q_2p_4)/(p_1p_2) \\ q_1 \\ (q_2p_4)/p_3 \\ q_2p_4 \\ q_2 \end{bmatrix}. \tag{2.23}$$

Thus by Equation 2.19, we have that $\mathcal{Y}_q = \{y_2, y_5\}$ and $\mathcal{Y}_q^c = \{y_1, y_3, y_4\}$. By Equation 2.20, Equation 2.23 can be described by the mapping function ψ_y:

$$\begin{aligned} p_1 &\mapsto p_1 & y_1 &\mapsto (q_1p_3 + q_2p_4)/(p_1p_2) \\ p_2 &\mapsto p_2 & y_2 &\mapsto q_1 \\ p_3 &\mapsto p_3 & y_3 &\mapsto (q_2p_4)/p_3 \\ p_4 &\mapsto p_4 & y_4 &\mapsto q_2p_4 \\ & & y_5 &\mapsto q_2 \end{aligned}$$

From ψ_y and ψ_p we construct the composite forward map, $\psi_{py} = (\psi_y \circ \psi_p)$:

$$
\begin{aligned}
k_1 &\mapsto (q_1 p_3 + q_2 p_4)/(p_1 p_2) & x_1 &\mapsto p_1 \\
k_2 &\mapsto q_1 & x_2 &\mapsto p_2 \\
k_3 &\mapsto (q_2 p_4)/p_3 & x_3 &\mapsto p_3 \\
k_4 &\mapsto q_2 p_4 & x_4 &\mapsto p_4 \\
k_5 &\mapsto q_2
\end{aligned}
$$

To reverse the substitution, notice from Equation 2.23 that $y_2 = q_1$ and $y_5 = q_2$, giving the following map, ψ_q^{-1}:

$$
\begin{aligned}
\mathcal{P} \to \mathcal{P} \ \text{(identity)} \qquad & & q_1 &\mapsto y_2 \\
& & q_2 &\mapsto y_5
\end{aligned}
$$

This yields a composite backward map, $\psi_{qp}^{-1} = (\psi_p^{-1} \circ \psi_q^{-1})$:

$$
\begin{aligned}
p_1 &\to x_1 & q_1 &\mapsto k_2 \\
p_2 &\to x_2 & q_2 &\mapsto k_5 \\
p_3 &\to x_3 & & \\
p_4 &\to x_4 & &
\end{aligned}
$$

The complete steady state mapping $\psi_{ss} = (\psi_{qp}^{-1} \circ \psi_{py})$ is therefore

$$
\begin{aligned}
k_1 &\mapsto (k_2 x_3 + k_5 x_4)/(x_1 x_2) & x_1 &\mapsto x_1 \\
k_2 &\mapsto k_2 & x_2 &\mapsto x_2 \\
k_3 &\mapsto k_5 x_4 / x_3 & x_3 &\mapsto x_3 \\
k_4 &\mapsto k_5 x_4 & x_4 &\mapsto x_4 \\
k_5 &\mapsto k_5
\end{aligned}
\tag{2.24}
$$

Applying this transformation to the original vector of reaction velocities yields

$$\overline{\mathbf{v}} = \psi_{ss}(\mathbf{v}) = \begin{bmatrix} k_2x_3 + k_5x_4 \\ k_2x_3 \\ k_5x_4 \\ k_5x_4 \\ k_5x_4 \end{bmatrix},$$

which one can verify satisfies Equation 2.1. An interesting implication of this trivial application of py-substitution is that, because ψ_{py} maps every species concentration to an independent parameter, we can interpret Equation 2.24 to mean that any vector of steady state concentrations will be consistent with an infinite number of reaction rate constants. In this particular case, knowing all four concentrations tells us nothing about the rates of enzyme-substrate dissociation or product degradation. As we shall see, by using different substitution strategies, we have some flexibility in choosing which rate constants are constrained by the steady state concentrations, but the structure of the OMM model makes finding a unique set of rate constants impossible. In general, a unique set of reaction rate constants requires that the coefficient matrix be full rank, or

$$\text{rank } \mathbf{C} = d_y. \tag{2.25}$$

Since complete knowledge of the species concentrations implies $d_p = d_x$ and $d_y = d_k$, by Equation 2.16, Equation 2.25 becomes

$$\text{rank } \mathbf{S} = d_k.$$

In other words, a unique set of rate constants requires that the stoichiometric matrix be full rank, which is equivalent to requiring that the corresponding reaction network have no cycles. Since even a single reversible reaction represents a cycle, we conclude that in the general case, a set of steady state species concentrations does not imply a unique set of reaction rate constants.

Heterogeneous substitution: the number of independent model parameters is constant

Often, models contain species whose concentrations are difficult to measure or reactions whose rates have been well characterized. For such models it is preferable to partition sets $\mathcal{K}$ and $\mathcal{X}$ so that species whose concentrations are difficult to measure are mapped to $\mathcal{Y}$ while well-characterized reaction rates are mapped to $\mathcal{P}$. For example, if the kinetics of the enzyme are well characterized, an attractive partitioning of the OMM model might be $\mathcal{K}_p = \{k_1, k_2, k_3, k_5\}$ and $\mathcal{X}_p = \{x_2\}$. This yields a map ψ_p:

$$\begin{aligned} k_1 &\mapsto p_2 & x_1 &\mapsto y_1 \\ k_2 &\mapsto p_3 & x_2 &\mapsto p_1 \\ k_3 &\mapsto p_4 & x_3 &\mapsto y_2 \\ k_4 &\mapsto y_4 & x_4 &\mapsto y_3 \\ k_5 &\mapsto p_5 \end{aligned}$$

Again, see "omm2.m.trace.pdf" in Protocol S1 for complete details. Notice here that we have forced the enzyme kinetic parameters k_1, k_2, and k_3 to be independent by mapping them to elements in $\mathcal{P}$. The resulting coefficient matrix and null space basis are

$$\mathbf{C} = \begin{bmatrix} -p_1p_2 & p_3+p_4 & 0 & 0 \\ -p_1p_2 & p_3 & 0 & 1 \\ p_1p_2 & -(p_3+p_4) & 0 & 0 \\ 0 & p_4 & -p_5 & 0 \end{bmatrix}, \quad \mathbf{N} = \begin{bmatrix} (p_3+p_4)/(p_1p_2p_4) \\ 1/p_4 \\ 1/p_5 \\ 1 \end{bmatrix},$$

which yield the steady state map ψ_{ss}:

$$
\begin{aligned}
k_1 &\mapsto k_1 & x_1 &\mapsto k_4(k_2+k_3)/(k_1k_3x_2)\\
k_2 &\mapsto k_2 & x_2 &\mapsto x_2\\
k_3 &\mapsto k_3 & x_3 &\mapsto k_4/k_3\\
k_4 &\mapsto k_4 & x_4 &\mapsto k_4/k_5\\
k_5 &\mapsto k_5
\end{aligned}
$$

As desired, x_2 is the only independent species concentration. Applying this transformation to the original vector of reaction velocities gives

$$
\overline{\mathbf{v}} = \psi_{ss}(\mathbf{v}) = \begin{bmatrix} k_4(k_2+k_3)/k_3 \\ k_2k_4/k_3 \\ k_4 \\ k_4 \\ k_4 \end{bmatrix}.
$$

Notice that even though the cardinality of $\mathcal{P}$ differs in this example as compared to the one above (5 *versus* 4), the cardinality of $\mathcal{Y}_q^c$ does not (3). Let d_c denote this cardinality. Obviously, $d_c = d_y - d_q$, or equivalently,

$$
d_c = \text{ncols } \mathbf{C} - \text{ncols } \mathbf{N}.
$$

This is simply the rank-nullity theorem again. By Equation 2.16, we can therefore conclude that

$$
d_c = \text{rank } \mathbf{S}.
$$

In other words, the final number of dependent elements in the steady state expression for a system is independent of the substitution strategy, and only depends on the structure of the reaction network.

Substitution with sublinear velocities: using *py*-substitution to resolve non-linearities (I)

Some reaction velocities are zero-order but well-characterized. For example, if the rate v_4 of substrate synthesis in the OMM model has been accurately measured, we may wish to partition $\mathcal{K}$ such that $k_4 \in \mathcal{K}_p$. The resulting mapping function ψ_p, however, fails to linearize $\mathbf{v}$. To compensate, we introduce a pseudospecies $\hat{x}_5 = 1$ and let $v_4 = k_4\hat{x}_5$. If we now partition $\mathcal{X}$ such that $\hat{x}_5 \in \mathcal{X}_{lin}$, the linearity of $\psi_p(\mathbf{v})$ in $\mathbf{y}$ is preserved and we may continue as before.

To illustrate this approach, we again let $\mathcal{X}_p = \{x_1, x_2, x_3, x_4\}$ and $\mathcal{K}_p = \{k_4\}$. The remaining rate constants and one pseudospecies are partitioned into sets $\mathcal{K}_{lin}$ and $\mathcal{X}_{lin}$, respectively, such that $\psi_p(\hat{x}_5) = y_5$. See "omm3.m.trace.pdf" in Protocol S1 for details. The resulting velocity vector is linear and yields a coefficient matrix whose null space is two-dimensional,

$$\mathbf{N} = \begin{bmatrix} p_3/(p_1p_2) & p_5/(p_1p_2) \\ 1 & 0 \\ 0 & p_5/p_3 \\ 0 & p_5/p_4 \\ 0 & 1 \end{bmatrix}. \tag{2.26}$$

However, one of these two dimensions is constrained by the pseudospecies. We are thus not at liberty to take a general linear combination as per Equation 2.18 but must find $\mathbf{q}$ such that

$$\psi_{py}(\hat{x}_5) = 1. \tag{2.27}$$

By our choice of ψ_p, and by Equations 2.18 and 2.26, we have $\psi_{py}(\hat{x}_5) = (\mathbf{q})_2$. Equation 2.27 is therefore satisfied when $(\mathbf{q})_2 = 1$. This gives $\mathbf{q} = [q_1, 1]^{\mathrm{T}}$ and

$$\overline{\mathbf{y}} = \begin{bmatrix} (p_3 q_1 + p_5)/(p_1 p_2) \\ q_1 \\ p_5/p_3 \\ p_5/p_4 \\ 1 \end{bmatrix}.$$

The complete steady state mapping $\psi_{ss} = (\psi_{qp}^{-1} \circ \psi_{py})$ is

$$\begin{aligned} k_1 &\mapsto (k_4 + k_2 x_3)/(x_1 x_2) & x_1 &\mapsto x_1 \\ k_2 &\mapsto k_2 & x_2 &\mapsto x_2 \\ k_3 &\mapsto k_4/x_3 & x_3 &\mapsto x_3 \\ k_4 &\mapsto k_4 & x_4 &\mapsto x_4 \\ k_5 &\mapsto k_4/x_4 & \hat{x}_5 &\mapsto 1 \end{aligned}$$

As desired, k_4 remains an independent parameter. Applying this transformation to the original vector of reaction velocities yields

$$\overline{\mathbf{v}} = \psi_{ss}(\mathbf{v}) = \begin{bmatrix} k_4 + k_2 x_3 \\ k_2 x_3 \\ k_4 \\ k_4 \\ k_4 \end{bmatrix}.$$

Substitution with superlinear velocities: using *py*-substitution to resolve non-linearities (II)

Some reaction velocities are superlinear in their reactant concentrations. If good estimates for these concentrations do not exist, we would like to partition these species into $\mathcal{X}_{lin}$. Analogous to the sublinear case above, doing so results in a velocity vector $\psi_p(\mathbf{v})$ that is non-linear in $\mathbf{y}$. Fortunately, the strategy above is useful here as well: introduce a pseudospecies for each superlinearity, calculate a basis

for the null space of the coefficient matrix, and identify basis vector coefficients that satisfy the constraints imposed by the pseudospecies.

Let us consider a version of the OMM model where the rate of product formation is proportional to the square of the enzyme-substrate complex, $v_3 = k_3 x_3^2$. Let us further assume that no estimate exists for the value of x_3. We would therefore like $\psi_p(x_3) \in \mathcal{Y}$. Since this fails to linearize the velocity, we introduce a pseudospecies $\hat{x}_5 = x_3^2$ and let $v_3 = k_3 \hat{x}_5$. We now define ψ_p such that

$$\begin{aligned} k_1 &\mapsto p_2 & x_1 &\mapsto y_4 \\ k_2 &\mapsto p_3 & x_2 &\mapsto p_1 \\ k_3 &\mapsto p_4 & x_3 &\mapsto y_1 \\ k_4 &\mapsto y_5 & x_4 &\mapsto y_3 \\ k_5 &\mapsto p_5 & \hat{x}_5 &\mapsto y_2 \end{aligned}$$

This satisfies the linearity requirement and maps x_3 and $\hat{x}_5$ to the lowest indices in $\mathcal{Y}$, thereby favoring these quantities to become dependent parameters. See "omm4.m.trace.pdf" in Protocol S1 for details. The resulting coefficient matrix has a null space that is spanned by the columns of

$$\mathbf{N} = \begin{bmatrix} p_1 p_2 / p_3 & -1/p_3 \\ 0 & 1/p_4 \\ 0 & 1/p_5 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Letting $\mathbf{q} = [q_1,\ q_2]^T$ maps k_4 and x_1 to $\mathcal{Q}$ and satisfies our requirement that $\psi_{py}(x_3)$ and $\psi_{py}(\hat{x}_5) \in \mathcal{Y}_q^c$. As in the previous section, however, one dimension of $\mathbf{N}$ is constrained by the pseudospecies. Specifically, we require that $\psi_{py}(\hat{x}_5) = \psi_{py}(x_3^2)$. by Equation 2.18, this requires that

$$(q_2 - p_1 p_2 q_1)^2 / p_3^2 = q_2 / p_4.$$

Solving for q_1 (we may just as easily have solved for q_2; in this example, whether k_4 or x_1 map to $\mathcal{Q}$ is immaterial), we are left with the following:

$$\overline{\mathbf{y}} = \begin{bmatrix} \sqrt{q_2/p_4} \\ q_2/p_4 \\ q_2/p_5 \\ q_2(p_3\sqrt{q_2})/(p_1p_2\sqrt{p_4}) \\ q_2 \end{bmatrix}.$$

The complete steady state mapping $\psi_{ss} = (\psi_{qp}^{-1} \circ \psi_{py})$ is

$$\begin{aligned} k_1 &\mapsto k_1 & x_1 &\mapsto (k_4 + k_2\sqrt{k_4/k_3})/(k_1x_2) \\ k_2 &\mapsto k_2 & x_2 &\mapsto x_2 \\ k_3 &\mapsto k_3 & x_3 &\mapsto \sqrt{k_4/k_3} \\ k_4 &\mapsto k_4 & x_4 &\mapsto k_4/k_5 \\ k_5 &\mapsto k_5 & \hat{x}_5 &\mapsto k_4/k_3 \end{aligned}$$

Applying this transformation to the original vector of reaction velocities yields

$$\overline{\mathbf{v}} = \psi_{ss}(\mathbf{v}) = \begin{bmatrix} k_4 + k_2\sqrt{k_4/k_3} \\ k_2\sqrt{k_4/k_3} \\ k_4 \\ k_4 \\ k_4 \end{bmatrix}.$$

This example illustrates that, using pseudospecies, a mapping function ψ_p can always be found such that ψ_y can be derived using linear methods. Non-linearities introduced by pseudospecies can then be resolved on a case-by-case basis, resulting in the final steady state solution.

2.3.2 *Py*-substitution is more general, but not less efficient, than King-Altman

Some chemical reaction systems are linear in the species concentration vector, or can be rendered linear by assuming that the concentrations of certain species don't change over time. The classical model for malate synthesis is an example of the latter [96]. Here, the enzyme fumarase binds reversibly to fumarate and hydrogen in either order, followed by reversible binding of hydroxyl and reversible formation of malate (Figure 2.3). The reactions for this model are

$$\begin{aligned}
r_1 &: x_1 + x_6 \xrightarrow{v_1} x_3 & \qquad r_7 &: x_3 \xrightarrow{v_7} x_1 + x_6 \\
r_2 &: x_3 + x_7 \xrightarrow{v_2} x_4 & \qquad r_8 &: x_4 \xrightarrow{v_8} x_3 + x_7 \\
r_3 &: x_1 + x_7 \xrightarrow{v_3} x_5 & \qquad r_9 &: x_5 \xrightarrow{v_9} x_1 + x_7 \\
r_4 &: x_5 + x_6 \xrightarrow{v_4} x_4 & \qquad r_{10} &: x_4 \xrightarrow{v_{10}} x_5 + x_6 \\
r_5 &: x_4 + x_8 \xrightarrow{v_5} x_2 & \qquad r_{11} &: x_2 \xrightarrow{v_{11}} x_4 + x_8 \\
r_6 &: x_2 \xrightarrow{v_6} x_1 + x_9 & \qquad r_{12} &: x_1 + x_9 \xrightarrow{v_{12}} x_2
\end{aligned}$$

By Equation 1.1, the corresponding reaction velocities are

$$\begin{aligned}
v_1 &= k_1 x_1 x_6 & \qquad v_7 &= k_7 x_3 \\
v_2 &= k_2 x_3 x_7 & \qquad v_8 &= k_8 x_4 \\
v_3 &= k_3 x_1 x_7 & \qquad v_9 &= k_9 x_5 \\
v_4 &= k_4 x_5 x_6 & \qquad v_{10} &= k_{10} x_4 \\
v_5 &= k_5 x_4 x_8 & \qquad v_{11} &= k_{11} x_2 \\
v_6 &= k_6 x_2 & \qquad v_{12} &= k_{12} x_1 x_9
\end{aligned} \tag{2.28}$$

The stoichiometric matrix is

$$\mathbf{S} = \left[\begin{array}{cccccccccccc}
-1 & 0 & -1 & 0 & 0 & 1 & 1 & 0 & 1 & 0 & 0 & -1 \\
0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 & -1 & 1 \\
1 & -1 & 0 & 0 & 0 & 0 & -1 & 1 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 1 & -1 & 0 & 0 & -1 & 0 & -1 & 1 & 0 \\
0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 & -1 & 1 & 0 & 0 \\
\hline
-1 & 0 & 0 & -1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\
0 & -1 & -1 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 1
\end{array}\right].$$

Notice that the submatrix formed by the first five rows of $\mathbf{S}$ satisfies the definition for a linear model given in Equation 2.3. Call this submatrix $\mathbf{S}_5$ and let $\mathbf{x}' = [x_1, \ldots, x_5]^{\mathrm{T}}$ be the vector of enzyme concentrations. If we assume that the substrate concentrations $x_6, \ldots, x_9$ are time-invariant, the steady state equation for this model becomes

$$\mathbf{S}_5\mathbf{v} = \mathbf{0}. \tag{2.29}$$

Because $\mathbf{S}_5$ satisfies Equation 2.3, we may define the following transition rate constants

$$\begin{aligned}
k'_{1,3} &= k_1 x_6 & k'_{3,1} &= k_7 \\
k'_{3,4} &= k_2 x_7 & k'_{4,3} &= k_8 \\
k'_{1,5} &= k_3 x_7 & k'_{5,1} &= k_9 \\
k'_{5,4} &= k_4 x_6 & k'_{3,4} &= k_{10} \\
k'_{4,2} &= k_5 x_8 & k'_{2,4} &= k_{11} \\
k'_{2,1} &= k_6 & k'_{1,2} &= k_{12} x_9
\end{aligned} \tag{2.30}$$

Substituting Equations 2.30 into 2.28 results in a velocity vector $\mathbf{v}'$ that is linear in $\mathbf{x}'$. Let $\mathbf{P}_{x'} = (\partial v'_i/\partial x'_j)$ as before, where $i = 1, \ldots, d_k$ and $j = 1, \ldots, d_{x'}$. This gives

$$\mathbf{v}' = \mathbf{P}_{x'}\mathbf{x}'. \tag{2.31}$$

If we now define a matrix

$$\mathbf{K} = \mathbf{S}_5\mathbf{P}_{x'}, \tag{2.32}$$

Equation 2.29 becomes

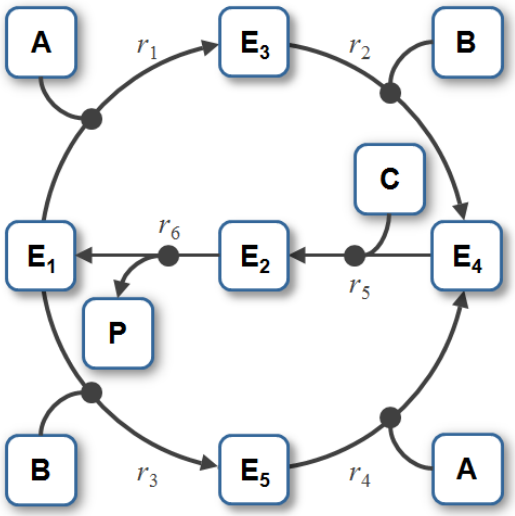
$$\mathbf{K}\mathbf{x}' = \mathbf{0}, \tag{2.33}$$

where the elements of $\mathbf{K}$ are given in Equation 2.7. The solution to Equation 2.33 is given by Equation 2.9, which we saw may be evaluated using the King-Altman method. Alternatively, we may solve Equation 2.29 directly using *py*-substitution. Given that *py*-substitution applies to a more general class of mass action models then King-Altman, we wondered whether this flexibility came at the cost of computational efficiency. Here we show that, for models that can be treated using the King-Altman method, *py*-substitution yields an equivalent result, and at no loss of efficiency.

***Py*-substitution and King-Altman yield equivalent steady state expressions**

Equation 2.33 has been solved previously using *KAPattern* [82]. The solution is reproduced here in "fum1.m.trace.pdf" in Protocol S1. For each enzyme i, $1 \leq i \leq 5$, the steady state concentration has the form

$$\bar{x}_i^{ka} = \frac{N_i^{ka}}{D^{ka}}. \tag{2.34}$$



Var	Symbol	Description
x_1	E_1	Fumarase
x_2	E_2	E_1+A+B+C
x_3	E_3	E_1+A
x_4	E_4	E_1+A+B
x_5	E_5	E_1+B
x_6	A	Fumarate
x_7	B	Hydrogen
x_8	C	Hyrdoxyl
x_9	P	Malate

Figure 2.3: **The model of malate synthesis used to compare *py*-substitution with the King-Altman method.** This mechanism for the conversion of fumarate to malate by the enzyme fumarase was proposed in [96]. Fumarase binds to fumarate and hydrogen in either order, then hydroxyl, followed by formation of the product, malate. All reactions are reversible. See "fum1.m" in Protocol S1 for a complete description of the model.

In this subsection only, we use $\bar{x}_i^{ka}$ to mean the i^{th} element of the vector $\overline{\mathbf{x}}$, made to satisfy Equation 2.33 by the King-Altman method. The element $\bar{x}_i^{py}$ is defined analogously for *py*-substitution. To solve Equation 2.29 by *py*-substitution, we partition $\mathcal{X}$ into subsets

$$\begin{aligned} \mathcal{E} &= \{x_1, \ldots, x_5\} \\ \mathcal{E}^c &= \{x_6, \ldots, x_9\}, \end{aligned} \tag{2.35}$$

and define ψ_p such that

$$\psi_p \colon \begin{cases} \mathcal{K} \cup \mathcal{E}^c \to \mathcal{P} \\ \mathcal{E} \to \mathcal{Y}. \end{cases}$$

The resulting coefficient matrix is precisely the matrix of rate constants, $\mathbf{K}$. The null space of $\mathbf{K}$ is one-dimensional and spanned by a single basis vector $\mathbf{n}$. In our solution, the basis vector is normalized to element $(\mathbf{n})_5$, which by Equations 2.18 and 2.19 yield a partition of $\mathcal{Y}$ into subsets $\mathcal{Y}_q = \{y_5\}$ and $\mathcal{Y}_q^c = \{y_1, \ldots, y_4\}$. After reversing the substitution we find that the steady state concentration of each enzyme likewise has the form

$$\bar{x}_i^{py} = \frac{N_i^{py}}{D^{py}}. \tag{2.36}$$

By inspection, Equations 2.34 and 2.36 are related by the following:

$$N_i^{py} = \bar{x}_5^{py} N_i^{ka}, \quad i = 1, 2, 3, 4 \tag{2.37}$$

$$D^{py} = N_5^{ka}. \tag{2.38}$$

In other words, the solutions given by *KAPattern* and *py*-substitution are not identical. The disparity arises from Equation 2.9, which imposes the constraint $\sum_i \bar{x}_i^{ka} = 1$. When derived by King-Altman, the steady state expression for each enzyme is therefore a ratio of the total enzyme concentration. In contrast, *py*-substitution results in $\bar{x}_1^{py}, \ldots, \bar{x}_4^{py}$ being expressed in terms of $\bar{x}_5^{py}$, the only element

$x \in \mathcal{E}$ for which $\psi_{py}(x) \in \mathcal{Q}$. Despite this disparity, Equations 2.36 to 2.38 can be combined to give

$$\bar{x}_i^{py} = \bar{x}_5^{py} N_i^{ka} / N_5^{ka}.$$

Therefore,

$$\begin{aligned}\sum_{i=1}^{5} \bar{x}_i^{py} &= \frac{\bar{x}_5^{py}}{N_5^{ka}} \sum_{i=1}^{5} N_i^{ka} \\ &= \bar{x}_5^{py}(D^{ka})/N_5^{ka} \\ &= \bar{x}_5^{py}/\bar{x}_5^{ka}.\end{aligned}$$

If we likewise impose the constraint $\sum_i \bar{x}_i^{py} = 1$, then $\bar{x}_5^{ka} = \bar{x}_5^{py}$, and for $i \neq 5$,

$$\begin{aligned}\bar{x}_i^{py} &= \bar{x}_5^{py} N_i^{py} / N_5^{ka} \\ &= (N_5^{ka}/D^{ka}) N_i^{ka} / N_5^{ka} \\ &= N_i^{ka}/D^{ka} \\ &= \bar{x}_i^{ka}.\end{aligned}$$

The two solutions are thus equivalent.

Py-substitution is not less efficient than King-Altman

We next wondered whether the King-Altman method is computationally more efficient than direct algebraic solution of the linear steady state equation (Equation 2.14). The King-Altman method requires exhaustive enumeration of valid King-Altman patterns. The number of patterns depends critically on the structure of the model. A model of strongly connected species generates $d_x^{d_x-2}$ patterns while a simple cycle generates only d_x [92]. By comparison, solving Equation 2.14 requires Gaussian elimination on the matrix $\mathbf{K}$. For a fixed-precision numeric matrix, this would take at most $O(d_x^3)$; however, since $\mathbf{K}$ has symbolic

entries rather than numerical ones, the sizes of the entries grow with the number of row operations. In fact, as Equation 2.11 shows, the number of valid King-Altman patterns is precisely the number of terms in the polynomial expansion of the minors. Thus even a few species, if highly connected, can generate thousands of terms and easily overwhelm conventional memory architectures.

To evaluate the performance of *py*-substitution versus *KAPattern*, we generated random models with six species and anywhere from 10 to 20 first-order reactions between them. Three distinct realizations were generated for each model. Models for which *KAPattern* failed – typically because the stoichiometric matrix described a disjoint network – were discarded. The command-line version of Matlab 2010b was used to derive the steady state concentration vector for each model using *py*-substitution and *KAPattern*, and for *py*-substitution the command-line version of Maple 14 was used as well. Internal memory was cleared prior to each derivation to prevent caching. The architecture used was a commodity netbook PC running Windows XP SP3 with an Intel 1.7 GHz Atom processor and 1 GB RAM. The derivation was repeated in triplicate for each realization to reduce variance introduced by the CPU scheduler. Execution times include initialization of the symbolic variables and coefficient matrix, kernel calculation, and derivation of $\overline{\mathbf{y}}$ in the case of *py*-substitution, and all steps prior to file writing in the case of *KAPattern*.

Results from the simulation are given in Figure 2.4. The data show that using Matlab, *KAPattern* provides consistently better performance and better scaling with respect to the number of reactions. This is likely because *KAPattern* uses Wang algebra to avoid explicit representation of the fully expanded minors in memory [81]. In contrast, Gaussian elimination of the coefficient matrix uses MuPAD, the Matlab symbolic engine, which is memory intensive and sensitive to expression swell. Models of even modest degree exhaust physical memory and cause "thrashing", resulting in poor runtime performance for models larger than 15 reactions. However, using Maple, direct solution of the steady state equation is typically an order of magnitude faster than *KAPattern* and exhibits identical scaling. This is likely because Maple's symbolic solver considers equations in increasing order of

their memory footprint. This data therefore argues that the King-Altman method is not more efficient than direct solution of the steady state equation.

Py-substitution is more general than King-Altman

To solve the steady state equation, the King-Altman method requires that the stoichiometric matrix $\mathbf{S}$ satisfies Equation 2.3. As we saw above, for $\mathbf{S}$ to satisfy Equation 2.3 we must be able to partition $\mathcal{X}$ into two disjoint sets, a set $\mathcal{E}$ of "enzymes" and a complementary set $\mathcal{E}^c$ of "substrates". The partition must be such that every reaction $r \in \mathcal{R}$ consumes a single species in $\mathcal{E}$ and produces a single, different species in $\mathcal{E}$. All other species produced or consumed by r must be in $\mathcal{E}^c$. The concentrations of these substrates are assumed to be time-invariant. As such, rows in $\mathbf{S}$ that correspond to substrates can be removed, and the substrate concentrations can be incorporated into the kinetics of the reactions. By inspection, the only such partition for the fumarase model is Equation 2.35, analyzed above.

By comparison, *py*-substitution does not require that the stoichiometric matrix satisfies Equation 2.3. The substrates $x_6, \ldots, x_9$ can therefore remain variable with respect to time and incorporated into the steady state solution, of which there are many. Without recourse to pseudospecies, the six bimolecular reaction velocities require that x_6, x_7, x_9 and x_1, x_3, x_5 be partitioned separately into sets $\mathcal{X}_p$ and $\mathcal{X}_{lin}$, or vice-versa. One such partition is

$$\begin{aligned}
\mathcal{K}_p &= \{k_1, \ldots, k_5, k_{12}\} \\
\mathcal{K}_{lin} &= \{k_6, \ldots, k_{11}\} \\
\mathcal{X}_p &= \{x_1, \ldots, x_5\} \\
\mathcal{X}_{lin} &= \{x_6, \ldots, x_9\}.
\end{aligned}$$

The resulting coefficient matrix has a five-dimensional null space, consistent with Equation 2.17 since $d_y = 10$ and rank $\mathbf{S} = 5$. A basis for this null space is given by the columns in $\mathbf{N}$, where

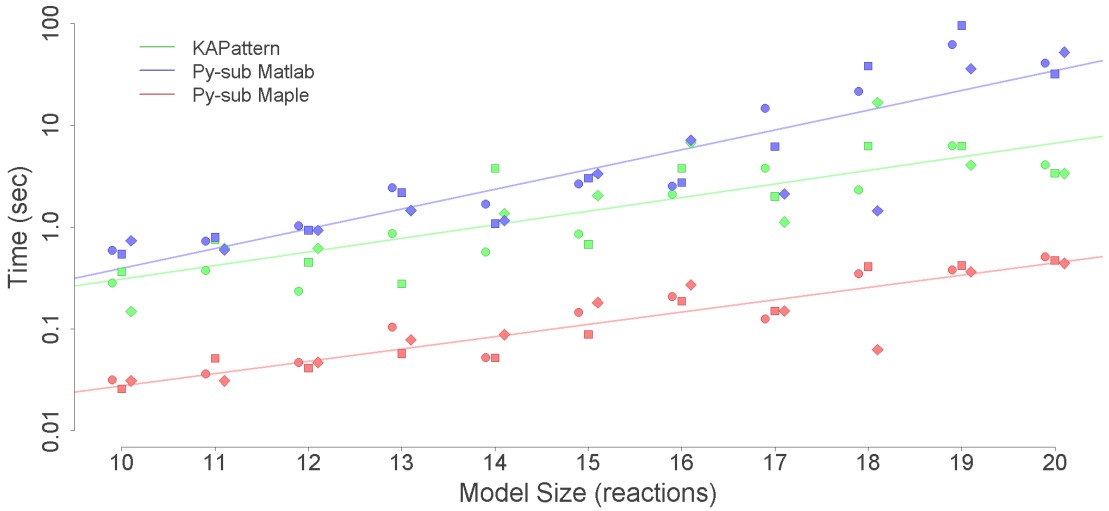


Figure 2.4: **Computational performance of *KAPattern* versus *py-substitution*, implemented in either Matlab or Maple.** Given a first-order model with six species and the number of reactions indicated by the x-axis, the time required to derive an expression for the steady state of the model is indicated by the y-axis. Three random realizations were used for every model size. Every calculation was performed in triplicate, but the error in calculation time was negligible.

$$\mathbf{N} = \left[\begin{array}{rrrrr} 0 & 0 & 0 & 0 & (p_6p_7)/p_8 \\ -p_{10}/p_9 & (p_4p_{11}+p_1p_7)/p_9 & 0 & 0 & 0 \\ -1 & (p_4p_{11})/p_{10} & (p_2p_9)/p_{10} & 0 & 0 \\ p_{10}/p_{11} & -p_4 & (p_3p_7)/p_{11} & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & (p_5p_{10})/p_8 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{array}\right].$$

The set $\mathcal{Y}_q$ must therefore contain five elements. We may select these elements with some flexibility by our choice of basis vector coefficients. The simplest choice, $\mathbf{q} = [q_1, \ldots, q_5]^{\mathrm{T}}$, yields the steady state mapping ψ_{ss}:

$$\begin{array}{llll} k_1 \mapsto k_1 & k_6 \mapsto k_{12}x_1x_9/x_2 & x_1 \mapsto x_1 & x_6 \mapsto x_6 \\ k_2 \mapsto k_2 & k_7 \mapsto (x_6(k_1x_1 + k_4x_5) - (k_{10}x_4))/x_3 & x_2 \mapsto x_2 & x_7 \mapsto x_7 \\ k_3 \mapsto k_3 & k_8 \mapsto (k_2x_3x_7 + k_4x_5x_6)/x_4 - k_{10} & x_3 \mapsto x_3 & x_8 \mapsto x_8 \\ k_4 \mapsto k_4 & k_9 \mapsto (k_{10}x_4 + k_3x_1x_7)/x_5 - k_4x_6 & x_4 \mapsto x_4 & x_9 \mapsto x_9 \\ k_5 \mapsto k_5 & k_{10} \mapsto k_{10} & x_5 \mapsto x_5 & \\ k_{12} \mapsto k_{12} & k_{11} \mapsto k_5x_4x_8/x_2 & & \end{array}$$

Other maps are available, however. By Equation 2.21, the submatrix formed by taking any 5 linearly independent rows of $\mathbf{N}$ produces a different vector of coefficients, and thus a different partition of $\mathcal{Y}$. For our particular choice of ψ_p above, 72 partitions are possible, calculated by testing which combinations of 5 rows in $\mathbf{N}$ are linearly independent. As an illustration, consider the case where the rate constants k_8 and k_{11} are easier to measure than substrates x_7 and x_8. Because of this, we would prefer x_7 and x_8 to be dependent variables. Equivalently, we

want $\psi_{py}(x_7)$, $\psi_{py}(x_8) \in \mathcal{Y}_q^c$, and $\psi_{py}(k_8)$, $\psi_{py}(k_{11}) \in \mathcal{Y}_q$. Since $\psi_p(k_8) = y_3$ and $\psi_p(k_{11}) = y_6$, any 5×5 submatrix of $\mathbf{N}$ containing rows 3 and 6 whose determinant is not zero will accomplish this. Below is the vector $\mathbf{q}'$ calculated from the matrix formed by rows 3, 5, 6, 7, and 10.

$$\mathbf{q}' = \begin{bmatrix} q_2 \\ q_4 \\ (q_1 p_{10} + q_2 p_{10} - q_4 p_4 p_{11})/(p_2 p_9) \\ (q_3 p_8)/(p_5 p_{10}) \\ q_5 \end{bmatrix},$$

This results in the desired steady state mapping, ψ'_{ss}:

$$\begin{aligned}
k_1 &\mapsto k_1 & x_1 &\mapsto x_1 \\
k_2 &\mapsto k_2 & x_2 &\mapsto x_2 \\
k_3 &\mapsto k_3 & x_3 &\mapsto x_3 \\
k_4 &\mapsto k_4 & x_4 &\mapsto x_4 \\
k_5 &\mapsto k_5 & x_5 &\mapsto x_5 \\
k_6 &\mapsto (k_{12} x_1 x_9)/x_2 & x_6 &\mapsto x_6 \\
k_7 &\mapsto (x_6(k_1 x_1 + k_4 x_5))/x_3 - (k_{10} x_4)/x_3 & x_7 &\mapsto (k_{10} x_4 + k_8 x_4 - \\
k_8 &\mapsto k_8 & & \quad k_4 x_5 x_6)/(k_2 x_3) \\
k_9 &\mapsto (k_{10} x_4)/x_5 - k_4 x_6 + & x_8 &\mapsto (k_{11} x_2)/(k_5 x_4) \\
& \quad (k_3 x_1 (k_{10} x_4 + k_8 x_4 - k_4 x_5 x_6))/(k_2 x_3 x_5) & x_9 &\mapsto x_9 \\
k_{10} &\mapsto k_{10} \\
k_{11} &\mapsto k_{11} \\
k_{12} &\mapsto k_{12}
\end{aligned}$$

This offers another illustration of how the choice of substitution strategy and null space basis vectors allow one to choose independent parameters flexibly among sets $\mathcal{K}$ and $\mathcal{X}$ when solving for steady state. See "fum2.m.trace.pdf" in Protocol S1 for details of this derivation.

2.3.3 Steady state establishes a threshold for drug-induced cell death

Finally, we sought to use *py*-substitution to characterize the relationship between steady state and the response to the cancer drug, dulanermin. Dulanermin is a recombinant human form of the endogenous ligand TRAIL, whose mechanism for triggering cell death is modeled in version 1.0 of the *extrinsic apoptosis reaction model*, or EARM [71]. This model considers the biochemical events following engagement of the death receptors 4 and 5 (DR4/5), including receptor-induced cleavage of initiator caspases, positive-feedback by effector caspases, and feed-forward amplification by the mitochondrial pathway following outer membrane permeabilization, or MOMP (Figure 2.5). The EARM model was trained on data derived from HeLa cells co-treated with cyclohexamide, an inhibitor of protein synthesis that results in hypersensitivity to TRAIL [97]. Accordingly, any amount of ligand in the EARM model results in cell death. The abundance of ligand still affects the time of death, defined for example by the time t_{PARP} at which half of the caspase 3 target protein PARP has been cleaved (Figure 2.6A, left) [98]. Note in this section we refer to the abundance of a species rather than its concentration, as these are the units chosen by the original authors.

In the absence of cyclohexamide, however, HeLa cells do not all die following exposure to TRAIL. Rather, a fraction of cells persist, and this resistance is a function of the proteomic state prior to stimulation [66]. To capture this phenomenon, we extended the EARM model so that proteins continued to be synthesized and degraded following exposure to TRAIL. Specifically, we introduced 43 new synthesis and degradation fluxes as well as 2 protein inactivation reactions (see "xearm.mpl" in Protocol S1). These reactions were chosen so that every species is subject to at least one efflux. We refer to our extended model as xEARM. Because xEARM satisfies our definition of a mass action model, we use *py*-substitution to identify an analytical expression for its steady state. To derive this expression, a mapping function ψ_p was chosen so that every non-zero parameter in EARM was mapped to an independent parameter in $\mathcal{P}$. As a result, we were able to preserve the snap-action dynamics of MOMP that is central to the original model (Fig-

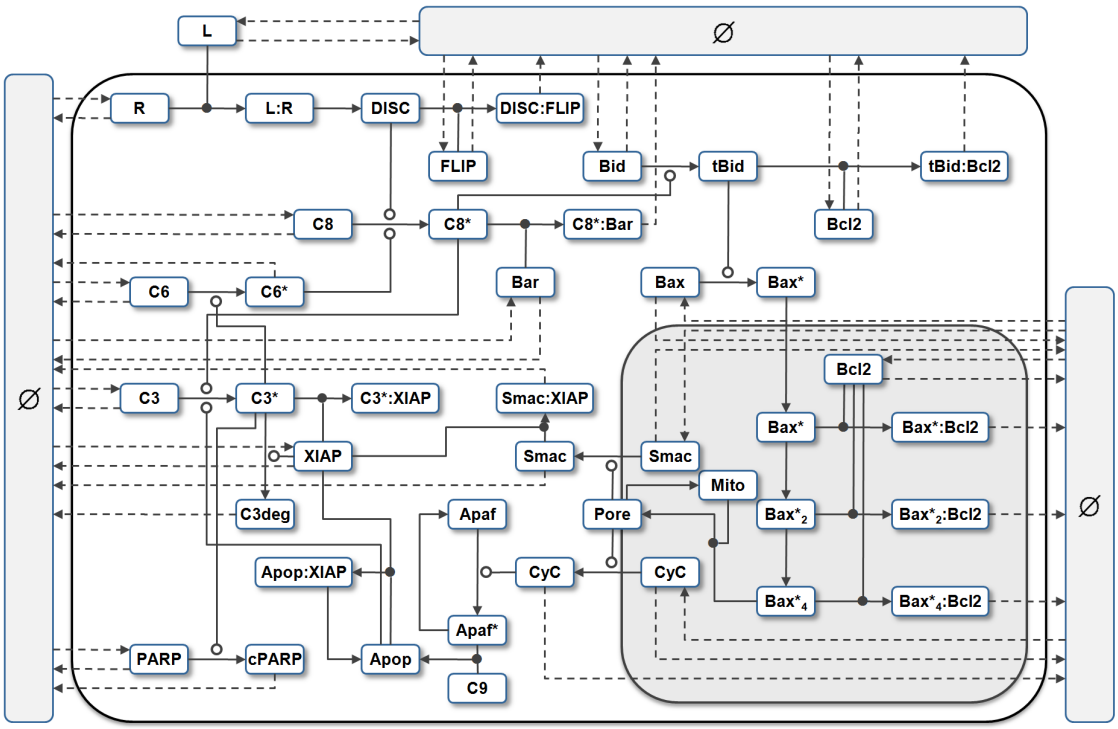


Figure 2.5: **Reaction diagram for the xEARM model.** Reactions new to this version include all fluxes to or from a source node, indicated by dashed lines to or from a Ø. In addition, the activation of *Apaf* was made reversible, as were the formation of mitochondrial pores. The complete model contains 58 species and 115 reactions. See [71] for a description of the original EARM model, and "xearm.mpl" in Protocol S1 for a complete description of xEARM.

ure 2.6A, right). Honoring the published parameters required that we introduce two pseudospecies, one for each the di- and tetrameric forms of Bax (variables x_{32} and x_{35}, respectively),

$$\hat{x}_{59} = x_{32}^2 \tag{2.39}$$

$$\hat{x}_{60} = x_{35}^2. \tag{2.40}$$

The coefficient matrix $\mathbf{C}$ and null space basis matrix $\mathbf{N}$ were calculated as before, with the latter calculation requiring less than a minute on our benchmark PC. The null space of $\mathbf{C}$ has 17-dimensions, resulting in a matrix of basis vectors of the form

$$\mathbf{N} = \begin{bmatrix} \mathbf{n}_1 & \mathbf{n}_2 & \mathbf{n}_3 & \mathbf{n}_4 & \dots & \mathbf{n}_{17} \end{bmatrix}.$$

Basis vectors $\mathbf{n}_4$ to $\mathbf{n}_{17}$ preserve the steady state ratios of paired synthesis and degradation reactions. Vector $\mathbf{n}_{17}$, for example, ensures that a change δ in k_{114} results in a change δx_{42} in k_{113}, where x_{42} is the abundance of Cytochrome C in the mitochondria and k_{113} and k_{114} are its rates of synthesis and degradation, respectively. The vector $\mathbf{n}_3$ scales the steady state abundances of mitochondrial Bax and Bcl2 complexes with respect to changes in the rate of Bcl2 synthesis. Vectors $\mathbf{n}_1$ and $\mathbf{n}_2$ are algebraically intractable and thus defy simple biochemical interpretation. Two of these vectors, $\mathbf{n}_1$ and $\mathbf{n}_3$, are constrained by the pseudospecies $\hat{x}_{59}$ and $\hat{x}_{60}$. To resolve these constraints, note that Equations 2.39 and 2.40 require that

$$\psi_p(\hat{x}_{59}) = \psi_p(x_{32}^2) \tag{2.41}$$

$$\psi_p(\hat{x}_{60}) = \psi_p(x_{35}^2). \tag{2.42}$$

By our mapping function ψ_{py} (see "xearm.mpl.trace.pdf" in Protocol S1, pp. 120–121), Equations 2.41 and 2.42 become

$$y_{42} = y_{21}^2 \tag{2.43}$$

$$y_{43} = y_{23}^2, \tag{2.44}$$

where $\psi_y(y_{21}), \psi_y(y_{23}), \psi_y(y_{42}),$ and $\psi_y(y_{43}) \in \text{span}_{\mathbb{Q}(\mathcal{P})}(\{q_1, q_2, q_3\})$. Solving Equation 2.44 for q_3 gives

$$q_3 = \frac{b_1}{b_2} q_2^2, \tag{2.45}$$

where $b_1, b_2 \in \mathbb{Q}[\mathcal{P}]$. Substituting Equation 2.45 into Equation 2.43 and solving for q_1 gives

$$q_1 = \frac{-a_2 \pm \sqrt{a_2^2 - 4a_1a_3}}{2a_3}, \tag{2.46}$$

where $a_1, a_2,$ and $a_3 \in \mathbb{Q}(\mathcal{P})[q_2]$ (see "xearm.mpl.trace.pdf" in Protocol S1, pp. 121–126).

Obviously, Equation 2.46 identifies an explicit bistability in the xEARM model. Basis vector coefficient q_1 — and by Equation 2.45, q_3 — can take either of two values for any numerical realization of the model. By examination of ψ_{py}, we find that these two coefficients affect all modified and compound species, as well as synthesis rates for proteins within and upstream of the mitochondria. Using the parameter values supplied in [71], however, we find that one of the solutions to Equation 2.46 is negative. The corresponding steady state is therefore infeasible and the solution was discarded.

In addition to parameters in [71], a full numerical realization of the xEARM model requires values for parameters $p_{71}, \ldots, p_{86}$ and $q_2, q_4, \ldots, q_{17}$. All but three of these elements represent first-order degradation rate constants, to which we assigned values equivalent to a half-life of one hour. This value was based on global quantifications of protein turnover in mammalian cells, which revealed that signaling proteins tend to be short-lived [62]. Two of the elements, p_{77} and p_{78}, represent first-order inactivation fluxes, which we assumed to be ten times faster than protein degradation. The final element q_2 is the steady state abundance of

the mitochondrial Bax2:Bcl2 complex, which we set to 20 molecules. Six of the elements were then modified from their initial values to better match the dynamics of caspase activation and PARP cleavage, as reported in [71]. The complete table of parameter values required to initialize and numerically integrate the xEARM model is given in Table S2.

For comparison, Table S3 lists the steady state abundances of species in the original and extended EARM models, sorted in order of decreasing difference. As expected, every species in EARM with a non-zero abundance has precisely the same abundance in xEARM, since these are independent parameters in the steady state solution. Among species with zero abundance in EARM, the mitochondrial Bax:Bcl2 complex exhibits the greatest disparity, with the steady state abundance in xEARM being in the low thousands of molecules. Ubiquitinated, cleaved caspase 3 and cleaved PARP are also in the low hundreds of molecules, but this represents only a small fraction of their total cellular abundance. A full 25 species with zero abundance in the EARM model have an abundance of less than 1 molecule in xEARM. This indicates that, even though the steady state reaction velocities are markedly different between EARM and xEARM, by using *py*-substitution we were able to engineer a steady state where the species abundances are appreciably similar between the two models.

Next we asked whether the xEARM model remained viable in the presence of low doses of TRAIL, but still exhibited MOMP when stimulated with high doses of TRAIL. To do so we created a numerical realization of the model using the parameters from Table S2, then perturbed the model from its steady state using a step increase in the abundance of TRAIL (variable x_1). The magnitude of the step ranged from 1 to 100-fold and was followed by numerical integration of the mass balance equations out to 48 hours. As shown in Figure 2.6A, MOMP is only observed in xEARM when TRAIL is increased by $10^{1.25}$-fold or more. We label this minimum dose of TRAIL required for MOMP L_{thresh}. Increments less than L_{thresh} result in a small and transient change in cleaved PARP abundance, followed by a return to the pre-stimulated steady state. By comparison, any magnitude dose of TRAIL causes MOMP in the original EARM model.

This ability of xEARM to distinguish between low and high doses of TRAIL, in conjunction with an analytical expression for its steady state, allowed us to systematically perturb the steady state and ask how these perturbations affect the sensitivity to TRAIL. To illustrate this capability we varied the steady state abundance of each major xEARM species over a 100-fold range, centered about each species' wildtype value as reported in Table S2. For each variation, we performed a binary search to identify L_{thresh}. The results from this procedure are plotted in Figure 2.6B. As expected, increases in XIAP, Bcl2, FLIP, and Bar result in reduced sensitivity to TRAIL stimulation, while increases in Procaspase 8, TRAIL receptor DR4/5, Bax, and Bid result in increased sensitivity [21]. What is interesting, however, is the following. First, TRAIL sensitivity is most affected by changes in the abundance of Procaspase 8 and Bar, an inhibitor of active caspase 8 [99]. The ability to activate caspase 8, then, appears to be a critical determinant of TRAIL sensitivity, as previously suggested [100, 101]. Second, the abundances of Procaspase 3, 6, and 9 have little effect on the sensitivity to TRAIL. This observation is in good agreement with the model-based prediction that induction of MOMP does not require positive-feedback via this caspase loop [71].

A common metric for describing how model parameters affect the sensitivity to TRAIL is to calculate the change in time at which death occurs in response to a small change in each parameter [66, 102, 98]. It is conceivable, however, that changes in the time of death do not accurately reflect changes in the threshold of TRAIL at which death occurs. Therefore, to test this assumption we calculated parameter sensitivity coefficients for the ligand threshold, $\partial L_{\text{thresh}}/\partial p$, and the time at which death occurs, $\partial t_{\text{PARP}}/\partial p$, using the xEARM and EARM models, respectively. The numerators $\partial L_{\text{thresh}}$ and ∂t_{PARP} were calculated by backward finite difference approximation and all sensitivities were normalized to the maximum observed sensitivity for each metric (Figure 2.6C). The data show good agreement for positive regulators of TRAIL sensitivity, but some disparity in the negative regulators. Specifically, while t_{PARP} is particularly sensitive to changes in XIAP and Bcl2, L_{thresh} is most sensitive to changes in Bar. This result argues that some caution should be taken when equating changes in the time of death with changes

in TRAIL sensitivity.

2.4 Discussion

We have described a simple but flexible method for deriving analytical expressions for the steady states of mass action models. Central to our method is the observation that mass action models are systems of polynomial equations that are generally no greater than degree 2. This permits a partitioning of rate constants and species concentrations into disjoint sets of quantities, $\mathcal{P}$ and $\mathcal{Y}$, where the reaction velocity vector is linear with respect to the variables in $\mathcal{Y}$. If the cardinality of $\mathcal{Y}$ is greater than the rank of the stoichiometric matrix, then the steady state equation can be solved analytically using simple linear methods.

There is considerable benefit to deriving an analytical expression for the steady state of a model. An analytical expression can be used to identify network ultrasensitivity [103], robustness [104], multistationarity [105], and invariants [106]. For enzyme catalytic models that have no true steady state but nevertheless satisfy the assumptions for quasi-steady state, an analytical expression can relate the rate of product formation to the initial concentrations of the substrates and enzyme [107]. Critically, these properties do not depend on the numerical values of the parameters, which may be difficult to measure [108]. In our companion manuscript, we show that analytical steady state expressions can be used to identify changes in the kinetic rate constants that do not alter the species concentrations. These *isostatic perturbations* can be used to characterize the dynamic plasticity of a system, and also how changes in the rates of protein turnover can affect the response to perturbation, independently of changes to steady state concentrations.

Even if numerical interrogation is ultimately intended and all parameters must be assigned values, deriving an analytical expression for the steady state still confers a number of benefits. First, including steady state constraints can facilitate the construction of a model [109]. As illustrated by our treatment of the Open Michaelis-Menten model, *py*-substitution affords considerable flexibility in selecting which quantities are independent — thus requiring numerical values prior to

simulation — and which quantities can be derived from the independent quantities. This partly transforms the problem of parameterizing a model from one of numerically fitting the rate constants to available data [110], to one of identifying the steady state expression that maximizes incorporation of known quantities into the independent set of parameters. Second, incorporating steady state concentration measurements can reduce the total number of parameters required. In the traditional approach to parameterization, every rate constant is assigned a value prior to simulation, as well as the abundance of any species not subject to synthesis and degradation. Using *py*-substitution, only independent quantities must be assigned a value. This number is equal to the total number of species and reactions, minus the rank of the stoichiometric matrix. As the stoichiometric matrix approaches full rank, this number converges to the number of species. Since most systems have more reactions than species, *py*-substitution often requires fewer parameters than the traditional approach. This can be observed in the xEARM model, where 119 parameters are required for simulation after deriving a steady state expression using *py*-substitution (100 rate constants, 18 species, and the mitochondrial volume), versus 133 parameters required for traditional parameterization (115 rate constants, 17 species, and the mitochondrial volume).

Further, in the case of the xEARM model, we have demonstrated that an analytical expression of the steady state allows systematic characterization of its effect on the response to perturbation. This was made possible in two ways. First, it allowed the model to operate at a non-trivial steady state. In the original EARM model, infinite sensitivity to TRAIL is caused by unbalanced reactions. Once the receptor is engaged, caspase cleavage and pore formation proceed deterministically to completion. As a result, for cells to be "alive" prior to stimulation, the model must assume a trivial steady state in which the abundance of TRAIL and all reaction velocities are zero. Using *py*-substitution, we were able to engineer a non-trivial steady state that is viable at low doses of TRAIL. Second, we were able to apply systematic changes to the steady state concentrations. By virtue of the mapping function ψ_{ss}, these resulted in compensating changes to the kinetic rate constants such that steady state was preserved. For each modification, we

were then able to calculate the number of TRAIL molecules required to induce cell death, as well as the sensitivity of this threshold to changes in the steady state concentrations of different species.

Previous studies with models operating at trivial steady states employed sensitivity metrics that were with respect to the time at which death occurs, and not whether it occurs [66, 98]. These studies suggested that the dynamics of TRAIL-induced cell death depend critically on Bcl-2 [98]. Also, whether cell death proceeds to completion depends on XIAP [98], and whether the mitochondrial feed-forward loop is required depends on the ratio of XIAP to Procaspase 3 [111]. In contrast, our analysis indicates that whether cell death occurs is primarily determined by the ratio of Procaspase 8 to its negative regulator, Bar. Our sensitivity analysis with respect to the threshold at which death occurs is therefore related to but distinct from analyses that consider only the timing of death, and may relate better to clinical applications since we don't assume co-treatment with cyclohexamide.

For all these reasons, an analytical expression for the steady state of a model can be of general benefit to cell systems modeling. Indeed, other methods have previously addressed the challenge of deriving analytical steady state expressions, most notably the King-Altman method. Prior to the advent of modern computers, the authors realized that for a particular class of mass action models, the laborious calculation of steady state enzyme ratios could be achieved by a conceptually simpler graphical method. As we have shown, however, this simpler approach is no longer more efficient. More significantly, the King-Altman method requires that all reactions be first- or pseudo-first order in the time-varying species. Without this stipulation, Equation 2.3 no longer holds and the reaction network can no longer be described by a graph. This requirement is often stated as a pair of assumptions: 1) that no enzyme is itself a substrate and 2) that all substrates remain constant over the time scale of steady state formation [78]. The second of these can be considered common to any method that treats time-varying species as constants when solving the steady state equation. The first of these, however, is violated by any cascade of post-translational modifications, for example the well-known MAP

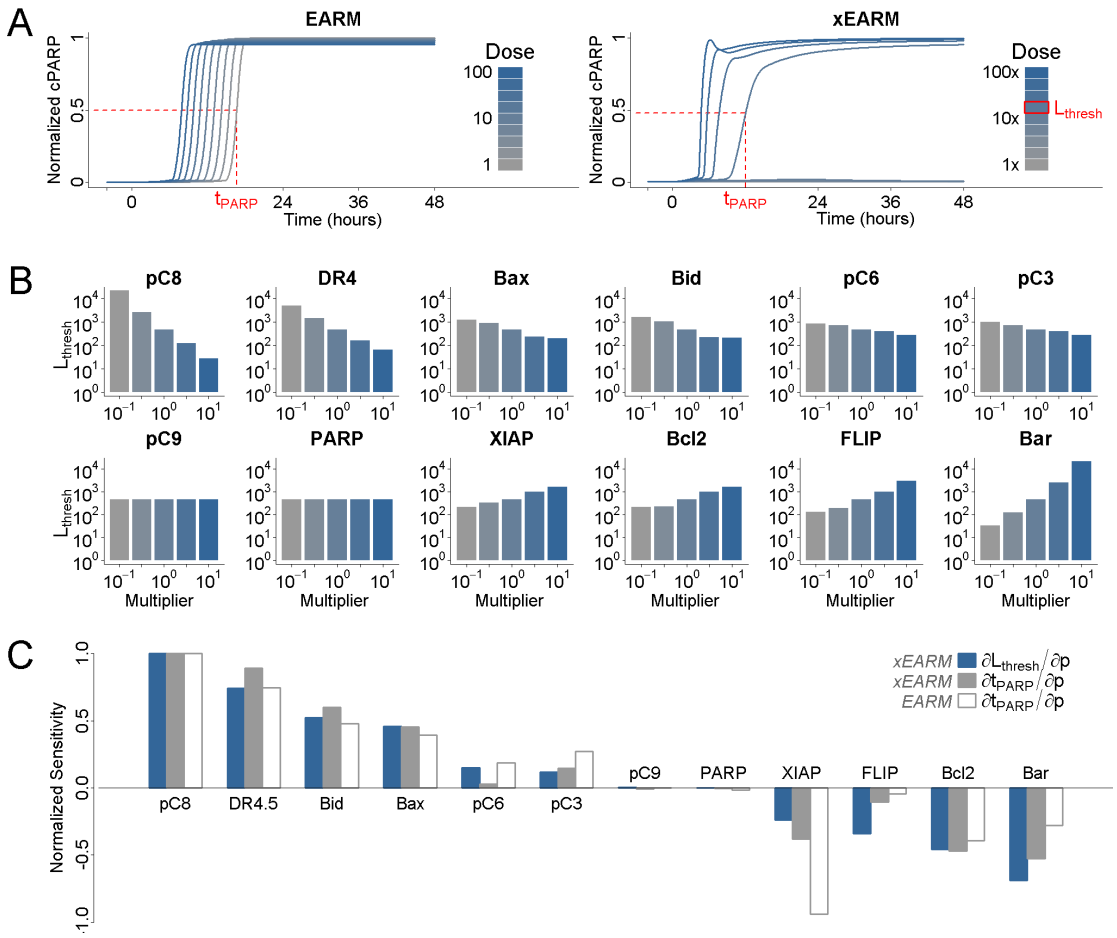


Figure 2.6: **Steady state determinants of sensitivity of TRAIL-induced cell death.** **A.** The dynamics of PARP cleavage are shown for EARM (left) and xEARM (right), in response to increasing doses of the TRAIL ligand (gray to blue). The abundance of cleaved PARP for each model has been normalized to the maximum observed abundance. For each model, for a particular dose of TRAIL, the time t_{PARP} at which PARP is 50% cleaved is indicated by the dashed red lines. For xEARM, the minimum abundance of TRAIL required to observe MOMP, L_{thresh}, is indicated on the color scale at right. **B.** Changes in L_{thresh} in response to changes in the steady state abundance of 12 primary xEARM species. Species have been sorted from left to right in order of those for which an increase in abundance results in the greatest increase in TRAIL sensitivity, to those for which an increase in abundance result in the greatest decrease in sensitivity. **C.** Normalized sensitivity coefficients for L_{thresh}, calculated using the xEARM model (blue), and t_{PARP}, calculated using both EARM (white) and xEARM (gray), for each of the 12 primary species in panel (B).

kinase cascade [112].

Although recent methods relax these assumptions [84, 85], in the contemporary systems biology literature, analytical derivation of the steady state rarely, if ever, precedes numerical interrogation of a model. Since this derivation is of considerable value, we sought to develop a method that was simple, scalable, and general to mass action models. First, we described our method using only concepts from linear algebra, and we have provided complete code for all seven examples described in this manuscript, with implementations in either Matlab or Maple. Second, we show that *py*-substitution scales well. The xEARM model has 58 species and 115 reactions, and we were able to derive a steady state expression in less than a minute on a conventional desktop computer. Finally, we demonstrated that *py*-substitution can be generally applied to chemical reaction networks whose reaction velocities are modeled by mass action kinetics. This is a considerably broader class of models than can be addressed using the King-Altman and other methods, which require that the reaction network exhibit specific structural properties.

This does, however, open up an interesting avenue for further research: precisely what properties must a mass action model exhibit for its steady state to be derived using *py*-substitution? How many different steady state expressions are possible, and which of these is the "best"? As we have shown with the fumarase model, even after the rate constants and species concentrations were partitioned into sets $\mathcal{P}$ and $\mathcal{Y}$, 72 different steady state expressions were possible. These different expressions arose from flexibility in selecting the pivot columns in the coefficient matrix, since the pivot vs. free columns partition the linear variables into dependent vs. independent variables. Equivalently, these different expressions arise from flexibility in ordering the linear variables, since different orderings permute the columns of the coefficient matrix and result in a different reduced row echelon form. Since the number of possible steady state expressions is large but finite, a combinatorial optimization strategy ought to be able to identify the best steady state expression, where the difference between any two expressions could take into account measurement uncertainty in the independent quantities, as well as computational complexity in deriving the final steady state expression.

Finally, we consider that the steady state may not be the only state of interest, but perhaps specified dynamic states as well. Essentially, this replaces the zero vector in Equation 2.1 with a vector of non-zero values. From linear algebra, we know that the solution to this dynamic equation can be expressed as the sum of a particular solution to the dynamic equation and an arbitrary point in the null space of the coefficient matrix. The solution is thus straightforward, raising the possibility of incorporating specific dynamic states into the parameterization of a model as well.

Chapter 2, in part, is a reprint of material as it appears in PLoS Computational Biology. Loriaux, Paul M; Tesler, Glenn; Hoffmann, Alexander, Public Library of Science 2013. The dissertation author was the primary investigator and author of this paper.

Chapter 3

A Protein Turnover Signaling Motif Controls the Stimulus-Sensitivity of Stress Response Pathways

Stimulus-induced perturbations from the steady state are a hallmark of signal transduction. In some signaling modules, the steady state is characterized by rapid synthesis and degradation of signaling proteins. Conspicuous among these are the p53 tumor suppressor, its negative regulator Mdm2, and the negative feedback regulator of NF-κB, IκBα. We investigated the physiological importance of this turnover, or *flux*, using a computational method that allows flux to be systematically altered independently of the steady state protein abundances. Applying our method to a prototypical signaling module, we show that flux can precisely control the dynamic response to perturbation. Next, we applied our method to experimentally validated models of p53 and NF-κB signaling. We find that high p53 flux is required for oscillations in response to a saturating dose of ionizing radiation (IR). In contrast, high flux of Mdm2 is not required for oscillations but preserves p53 sensitivity to sub-saturating doses of IR. In the NF-κB system, degradation of NF-κB-bound IκB by the IκB kinase (IKK) is required for activation in response

to TNF, while high IKK-independent degradation prevents spurious activation in response to metabolic stress or low doses of TNF. Our work identifies flux pairs with opposing functional effects as a signaling motif that controls the stimulus-sensitivity of the p53 and NF-κB stress-response pathways, and may constitute a general design principle in signaling pathways.

3.1 Introduction

Eukaryotic cells must constantly recycle their proteomes. Of the approximately 10^9 proteins in a typical mouse L929 fibrosarcoma cell, 10^6 are degraded every minute [113]. Assuming first-order degradation kinetics, this rate of constitutive protein turnover, or flux, imposes an average half-life of 24 hours. Not all proteins are equally stable, however. Genome-wide quantifications of protein turnover in HeLa cells [114, 115] and 3T3 murine fibroblasts [62] show that protein half-lives can span several orders of magnitude. Thus while some proteins exist for months and even years [116], others are degraded within minutes. Gene ontology terms describing signaling functions are highly enriched among short-lived proteins [115, 117, 118], suggesting that rapid turnover is required for proper signal transduction. Indeed, defects in protein turnover are implicated in the pathogenesis of cancer and other types of human disease [119, 120].

Conspicuous among short-lived signaling proteins are those that regulate the p53 and NF-κB stress response pathways. The p53 protein itself, for example, has a half-life of less than 30 minutes [121, 122]. Mdm2, the E3 ubiquitin ligase responsible for regulating p53, has a half-life of 45 minutes [62]. And the half-life of unbound IκBα, the negative feedback regulator of NF-κB, is less than 15 minutes [123, 23] (see Figure 3.1), requiring that 6,500 new copies of IκBα be synthesized every minute [23]. Given the energetic costs of protein synthesis, we hypothesized that rapid turnover of these proteins is critical to the stimulus-response behavior of their associated pathways.

To test our hypothesis we developed a method to systematically alter the rates of protein turnover in mass action models without affecting their steady state

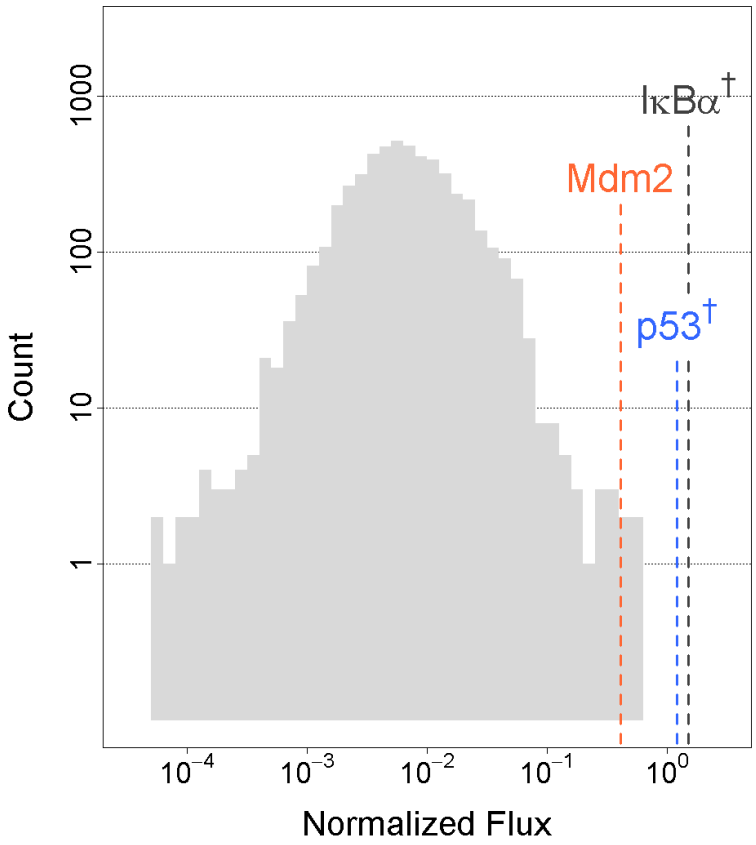


Figure 3.1: **A genome-wide distribution of protein flux.** A histogram of protein flux was generated from data in [62] (N=5030). Assuming first-order degradation kinetics, the published half-life for each protein was used in conjunction with its steady state abundance to calculate its rate of synthesis. This rate was then divided by the steady state abundance to derive each protein's normalized flux, that is, the fraction of its steady state population that is synthesized every hour. Normalized flux values for Mdm2, p53, and unbound IκBα are indicated by the dashed lines. Daggers denote proteins whose half-lives are extrinsic to the dataset.

abundances. Our method requires an analytical expression for the steady state of a model, which we derive using the *py*-substitution method described in the previous chapter. From this expression, changes in parameter values that do not affect the steady state are found in the null space of the matrix whose elements are the partial derivatives of the species abundances with respect to the parameters. We call this vector space the *isostatic subspace*. After deriving a basis for this subspace, linear combinations of basis vectors identify *isostatic perturbations* that modify specific reactions independently of all the others, for example those that control protein turnover. By systematic application of these isostatic perturbations to a model operating at steady state, the effects of flux on stimulus-responsiveness can be studied in isolation of changes to steady state abundances (see Section 3.4).

We first apply our method to a prototypical negative feedback module in which an activator controls the expression of its own negative regulator. We show that reducing the flux of either the activator or its inhibitor slows the response to stimulation. However, reducing the flux of the activator lowers the magnitude of the response, whereas reducing the flux of the inhibitor increases it. This complementarity allows the activator and inhibitor fluxes to exert precise control over the module's response to stimulation.

Given this level of control, we hypothesized that rapid turnover of p53 and Mdm2 must be required for p53 signaling. A hallmark of p53 is that it responds to DNA damage in a series of digital pulses [124, 125, 36, 35, 126]. These pulses are important for determining cell fate [127, 128, 9]. To test whether high p53 and Mdm2 flux are required for p53 pulses, we applied our method to a model in which exposure to ionizing radiation (IR) results in oscillations of active p53 [35]. By varying each flux over three orders of magnitude, we show that high p53 flux is indeed required for oscillations. In contrast, high Mdm2 flux is not required, but rather controls the refractory time in response to transient stimulation. If the flux of Mdm2 is low, a second stimulus after 22 hours does not result in appreciable activation of p53.

In contrast to p53, the flux of NF-κB turnover is very low, while the flux of its inhibitor, IκB, is very high. Prior to stimulation, most NF-κB is sequestered

in the cytoplasm by IκB. Upon stimulation by an inflammatory signal like tumor necrosis factor alpha (TNF), IκB is phosphorylated and degraded, resulting in rapid but transient translocation of NF-κB to the nucleus and activation of its target genes [129, 130, 131]. Two separate pathways are responsible for the turnover of IκB [123]. In one, IκB bound to NF-κB is phosphorylated by the IκB kinase (IKK) and targeted for degradation by the ubiquitin-proteasome system. In the other pathway, unbound IκB is targeted for degradation and requires neither IKK nor ubiquitination [40, 132]. We call these the "productive" and "futile" fluxes, respectively. Applying our method to a model of NF-κB activation, we show that the futile flux acts as a negative regulator of NF-κB activation while the productive flux acts as a positive regulator. We find that turnover of bound IκB is required for NF-κB activation in response to TNF, while high turnover of unbound IκB prevents spurious activation of NF-κB in response to low doses of TNF or ribotoxic stress caused by ultraviolet light (UV). As with p53 then, juxtaposition of a positive and negative regulatory flux govern the sensitivity of NF-κB to different stimuli, and may constitute a common signaling motif for controlling stimulus-specificity in diverse signaling pathways.

3.2 Results

3.2.1 Activator and inhibitor fluxes can precisely control the dynamics of signaling

To examine the effects of flux on stimulus-responsiveness, we return to the prototypical negative feedback model introduced in Chapter 1 (Figure 3.2.1A). Recall that this model has four kinetic parameters, the rates of synthesis and degradation of the activator X and its inhibitor, Y (k_1, k_2, k_3, and k_4). The dynamics of the activator in response to stimulation can be described by its maximum amplitude, A and the time T at which A is observed (Figure 3.2.1B).

To understand how these kinetic parameters shape the response of the activator, we applied systematic changes to the fluxes of X and Y prior to stimulation

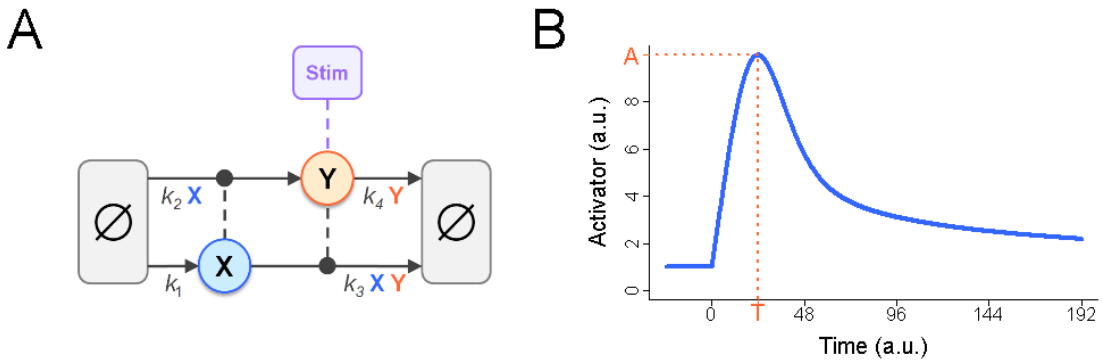


Figure 3.2: **A prototypical negative feedback module. A.** In this simple model of negative feedback control, an activator X is constitutively produced but catalytically degraded by an inhibitor, Y. Y is constitutively degraded but its synthesis requires X. Each of these four reactions is modeled using mass action kinetics. To stimulate the model and activate X, the steady state abundance of Y is instantaneously depleted. **B.** In response to stimulation, the abundance of X increases until activator-induced synthesis of Y forces a return to steady state. This response can be characterized by A, the maximum abundance of X following stimulation, and T, the time at which A is observed. Parameters were chosen for this model such that the steady state abundances of X and Y equal one arbitrary unit and the stimulus-induced amplitude of X is 10 a.u. at time $T = 24$.

and plotted the resulting values of A and T. Multiplying the flux of X over the interval $[10^{-1}, 10^{+1}]$ showed, as expected, that the value of A increases while the value of T decreases (Figure 3.2.1A). In other words, a high activator flux results in a strong, fast response to stimulation. If we repeat the process with the inhibitor, we find that both A and T decrease as the flux increases; a high inhibitor flux results in a fast but weak response (Figure 3.2.1B). This result illustrates that fluxes of different regulators can have different but complementary effects on stimulus-induced signaling dynamics.

Complementarity suggests that changes in flux can be identified such that A is altered independently of T, or T independently of A. Indeed, if both activator and inhibitor fluxes are increased in equal measure, A is held fixed while the value of T decreases (Figure 3.2.1C). Increasing both fluxes thus simultaneously reduces the timescale of the response without affecting its magnitude. An equivalent relationship can be found such that T remains fixed while A is affected (Figure 3.2.1D). Because an increase in either flux will reduce T, to alter A without affecting T requires an increase in one flux but a decrease in the other. Also, T is more sensitive to changes in the inhibitor flux versus the activator flux; small changes in the former must be paired with larger changes in the latter. This capability to achieve any value of A or T indicates that flux can precisely control the response to stimulation, without requiring any changes to steady state protein abundance.

3.2.2 High p53 and Mdm2 flux is required for p53 responsiveness to ionizing radiation

Given that flux precisely controls the dynamic response to stimulation in a prototypical signaling module, we hypothesized that for p53, oscillations in response to DNA damage require the high rates of turnover reported for p53 and Mdm2. To test this, we applied our method to a published model of p53 activation in response to ionizing gamma radiation (IR), a common DNA damaging agent (Figure 3.2.2A) [35]. Because the model uses arbitrary units, we rescaled it so that the steady state abundances of p53 and Mdm2, as well as their rates of synthesis

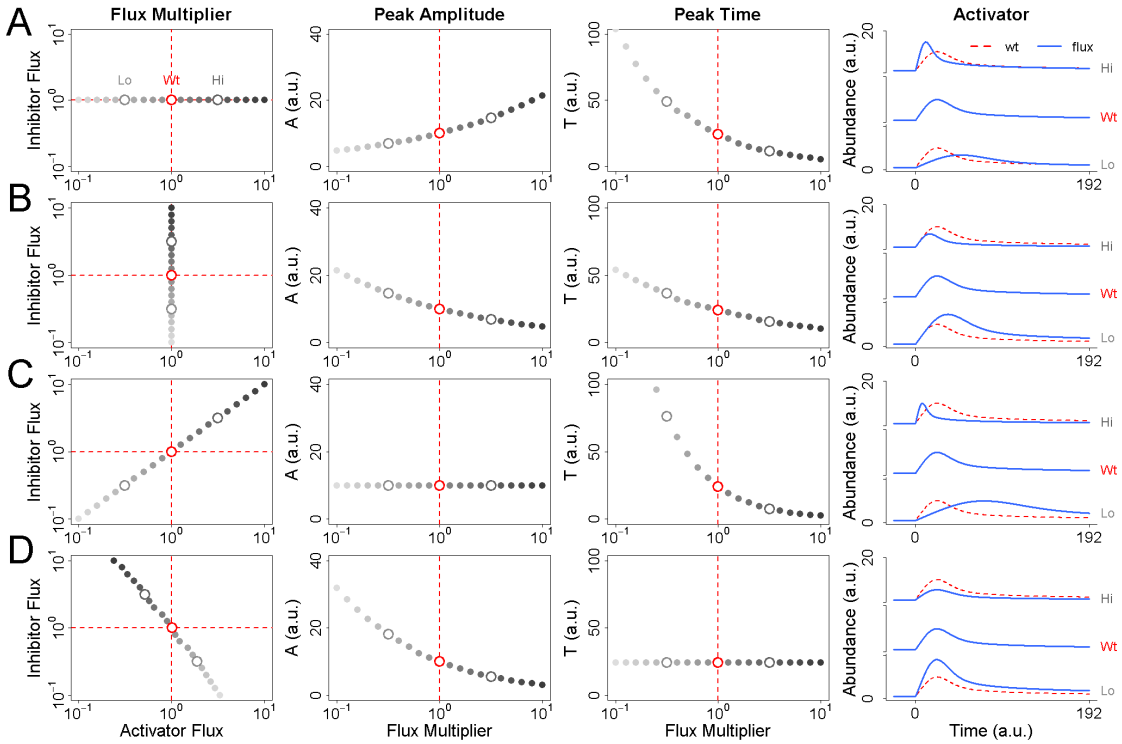


Figure 3.3: **Effects of flux on the dynamic response to stimulation.** **A.** The magnitude of the activator flux is varied between 10^{-1} (light gray) and 10^{+1} (dark gray) times its nominal steady state value prior to stimulation. The peak amplitude A of X in response to stimulation is observed to increase with the flux of X while the time T at which the peak occurs is observed to decrease. Representative profiles of the activator at low, wildtype, and high values of the flux are shown at right. The dashed red line indicates the nominal wildtype response. **B.** The magnitude of the inhibitor flux is varied between 10^{-1} and 10^{+1} times its nominal steady state value prior to stimulation. Both A and T are observed to decrease. **C.** The fluxes of both X and Y are varied simultaneously between 10^{-1} and 10^{+1} times their nominal wildtype values. As a result, A is held constant while T is reduced. **D.** The magnitude of the inhibitor flux is varied between 10^{-1} and 10^{+1} times its nominal steady state value prior to stimulation. For each value of this flux, the value of activator flux is calculated such that T is held constant. As in panel (B), above, A is observed to decrease as the magnitude of the flux of Y increases.

.

and degradation, matched published values (see Table 3.2.2). We note that these values are also in good agreement with the consensus parameters reported in [36].

Next we implemented a multiplier of Mdm2-independent p53 flux and let it take values on the interval $[10^{-2}, 10^{+1}]$. For each value we simulated the response to IR using a step function in the production of the upstream Signal molecule, k_{11}, as previously described [35]. To characterize the p53 response we let A_∞ be the amplitude of stable oscillations in phosphorylated p53 (Figure 3.2.2B), and use this as a metric for p53 sensitivity. Where $A_\infty > 0$, we say the module is sensitive to IR stimulation. We find that A_∞ is greater than zero only when the flux of p53 is near its observed value or higher (Figure 3.2.2A). If the flux of p53 is reduced by 2-fold or more, p53 no longer stably oscillates in response to stimulation, but exhibits damped oscillations instead.

Interestingly, repeating this analysis with a multiplier for the Mdm2 flux over the same interval reveals that Mdm2 flux has little bearing on p53 oscillations (Figure 3.2.2B). For any value of the multiplier chosen, $A_\infty > 0$. As with p53, this multiplier alters the Signal-independent flux of Mdm2 but does not affect Signal-induced Mdm2 degradation. If oscillations are already compromised by a reduced p53 flux, no concomitant reduction in Mdm2 flux can rescue the oscillations (Figure 3.2.2C). We therefore conclude that the flux of p53, but not Mdm2, is required for IR-sensitivity in the p53 signaling module. What then is the physiological relevance of high Mdm2 flux? In the model, signal-mediated Mdm2 auto-ubiquitination [134] is a major contributor to Mdm2 degradation after stimulation. If Signal production is transient, Mdm2 protein levels must be restored solely via Signal-independent degradation. We therefore hypothesized that if the flux of Mdm2 is low, Mdm2 protein levels would remain elevated after stimulation and compromise sensitivity to subsequent stimuli.

To test this hypothesis we again let the Mdm2 flux multiplier take values over the interval $[10^{-2}, 10^{+1}]$. For each value we stimulated the model with a 2-hour pulse of Signal production, followed by 22 hours of rest, followed by a second 2-hour pulse (Figure 3.2.2B). We defined A_1 to be the amplitude of the first peak of phosphorylated p53 and A_2 to be the amplitude of the second peak.

Table 3.1: Parameters used in model of p53 oscillations.

Param	Meaning	Value	Units[1]	Pub[2]	Ref[3]
x_1	Initial p53 inactive concentration	1.5e4	Cs	2.0e4	[133]
x_3	Initial Mdm2 concentration	1.0e4	Cs	1.0e4	[62]
k_1	Inactive p53 degradation rate	1.5	h^{-1}	1.18	[121]
k_2	Mdm2 degradation rate	0.75	h^{-1}	0.92	[62]
k_3	Inhibitor degradation rate	0.525	h^{-1}		
k_4	Inhibitor-independent Signal degradation rate	5.63	h^{-1}		
k_5	Mdm2-dependent p53inactive degradation rate	7.5e-5	$Cs^{-1}h^{-1}$		
k_6	Mdm2-dependent p53active degradation rate	2.1e-5	$Cs^{-1}h^{-1}$		
k_7	Signal-dependent Mdm2 inactivation rate	7.5e-6	$Cs^{-1}h^{-1}$		
k_8	Saturating Inhibitor-dependent Signal degradation rate	37.5	h^{-1}		
k_9	p53 inactive production rate	3.38e4	$Cs^{-1}h^{-1}$	2.4e4	[121][4]
k_{10}	p53-independent Mdm2 production rate	7.5e3	$Cs^{-1}h^{-1}$	4.2e3	[62]
k_{12}	p53-dependent Mdm2 production rate	0.675	h^{-1}		
k_{13}	Inhibitor production rate	0.15	h^{-1}		
k_{14}	Saturating production rate of p53active	7.5	h^{-1}		
c_1	Hill coefficient of Signal degradation by Inhibitor	4.0	unitless		
c_2	Hill coefficient of active p53 production by Signal	4.0	unitless		
c_3	Inhibitor concentration for half-maximal Signal degradation	1.0e4	Cs		
c_4	Signal concentration for half-maximal p53 production	5.0e4	Cs		
c_5	Time delay in Mdm2 production	0.933	h		
c_6	Time delay in Inhibitor production	1.6	h		

[1] 1 Cs is equivalent to 1 copy of the indicated molecule. 1 h is 1 hour.

[2] Published value. Half-lives were converted to rate constants by inverting and multiplying by ln(2).

[3] Citation for published value.

[4] Value calculated by multiplying p53 steady state abundance ($2e4$) by degradation rate constant (1.18).

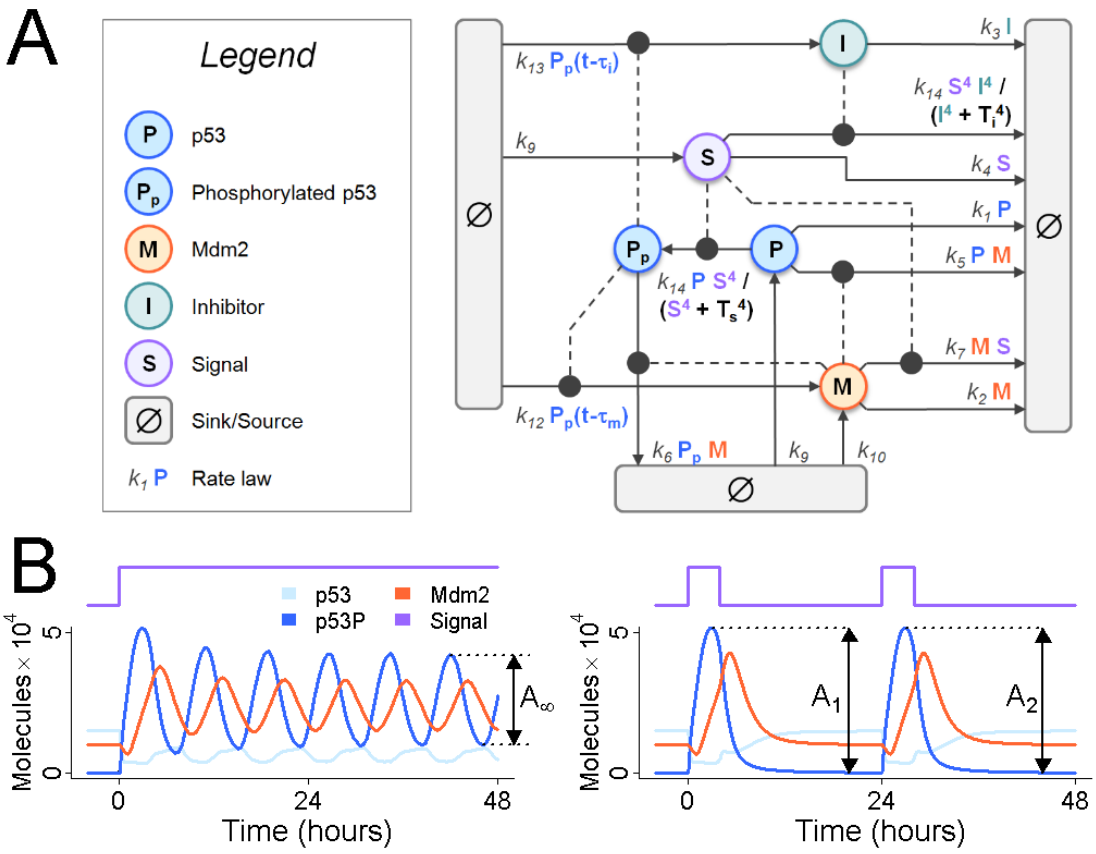
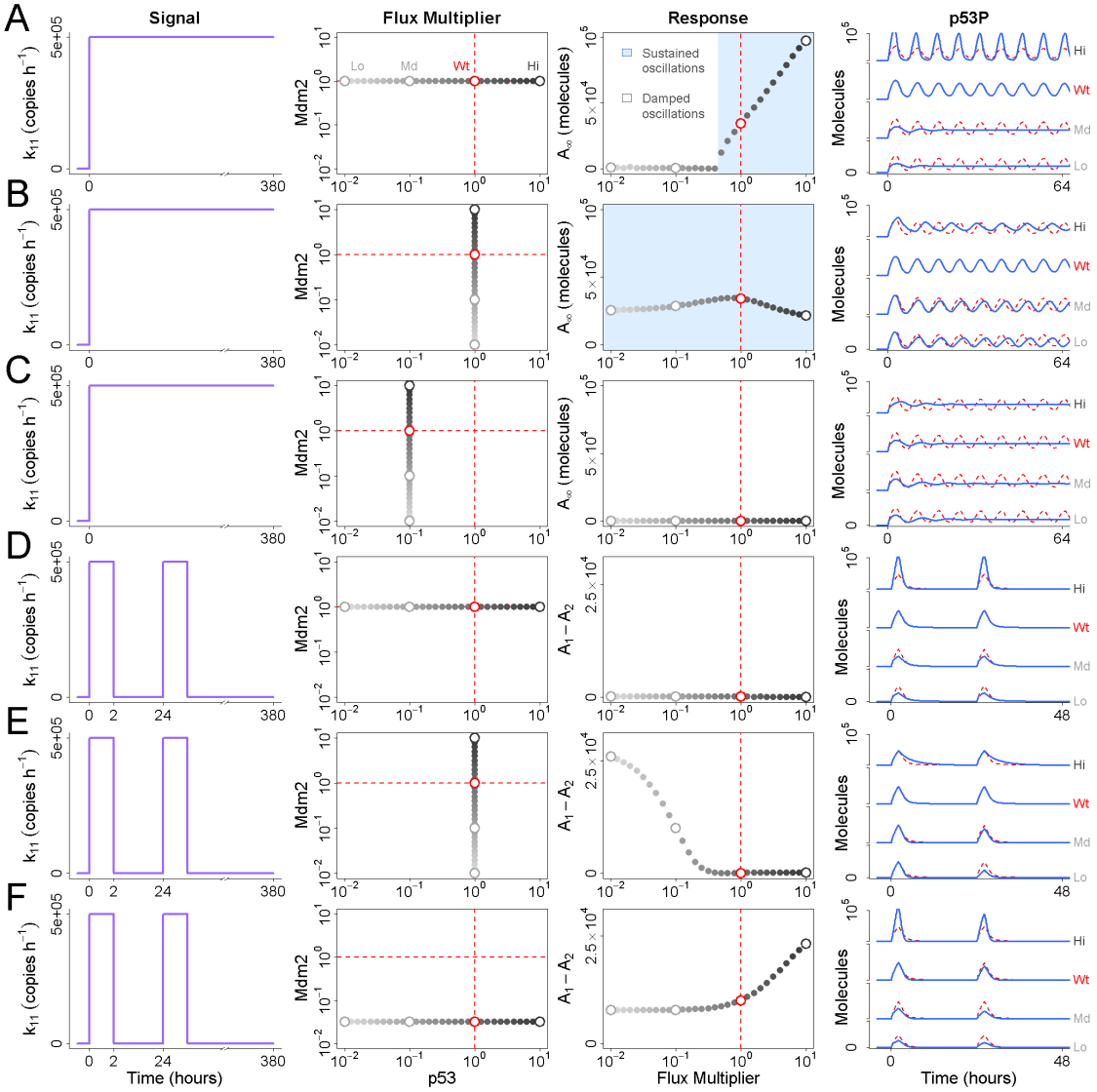


Figure 3.4: **A model of p53 oscillations in response to ionizing radiation.** **A.** The model shown here is structurally identical to [35], but parameter values have been scaled to match published rates of synthesis and degradation for p53 and Mdm2 as well as their steady state abundances (see Section 3.4, Methods). **B.** Ionizing radiation is modeled as an increase in synthesis of the Signal molecule (left; model parameter k_{11}) [35]. In response to a step increase in Signal production, phosphorylated p53 is observed to oscillate. We define A_∞ to be the amplitude of the stable oscillations. In response to a 2-hour pulse in Signal production (right), p53 exhibits a transient peak in phosphorylation, as does Mdm2. We define A_1 to be the amplitude of phosphorylated p53, and A_2 to be its amplitude in response to a second, identical pulse, 22 hours after the first pulse.

Figure 3.5: **Effects of flux on the p53 response to ionizing radiation.** **A.** The Mdm2-independent flux of p53 is varied between 10^{-2} and 10^{+1} times its wildtype value prior to a step increase in Signal production (model parameter k_{11}; light gray to dark gray) [35]. The magnitude of A_∞ is plotted as a function of this p53 flux multiplier. Values that give rise to stable oscillations are shaded in blue. At right, representative profiles of phosphorylated p53 are shown for high, wildtype, moderate, and low values of the multiplier. Note that the wildtype flux is indicated by the dashed line in red. **B.** As panel (A)), above, but now the p53-independent flux of Mdm2 is varied between 10^{-2} and 10^{+1} times its wildtype value (light gray to dark gray). Stable oscillations are observed for all values of the Mdm2 flux multiplier. **C.** As panel (B), above, but for all simulations the flux of p53 is at one-tenth its nominal wildtype value. Instead of sustained oscillations, damped oscillations are observed for all values of the Mdm2 flux multiplier. **D.** The flux of p53 is varied between 10^{-2} and 10^{+1} times its wildtype value prior to a 2-hour pulse in Signal production, followed by 22 hours of rest, followed by a second 2-hour pulse. No difference is observed in the amplitude of phosphorylated p53 in response to the first and second pulse. **E.** As panel (D), above, but now the flux of Mdm2 is varied instead of p53. At lower values of the Mdm2 flux multiplier, a significant difference is observed between the amplitude of phosphorylated p53 in response to the first and second pulse. **F.** As panel (D), above, but while the p53 flux is allowed to vary, the flux of Mdm2 is held constant at $10^{-1.5}$ times its wildtype value. This concomitant reduction of the p53 flux is not able to rescue the Mdm2-compromised response to the second pulse.


A
B
C
D
E
F
Signal
Flux Multiplier
Response
p53P
Lo
Md
Wt
Hi
Sustained oscillations
Damped oscillations
Time (hours)
p53
Flux Multiplier
Time (hours)

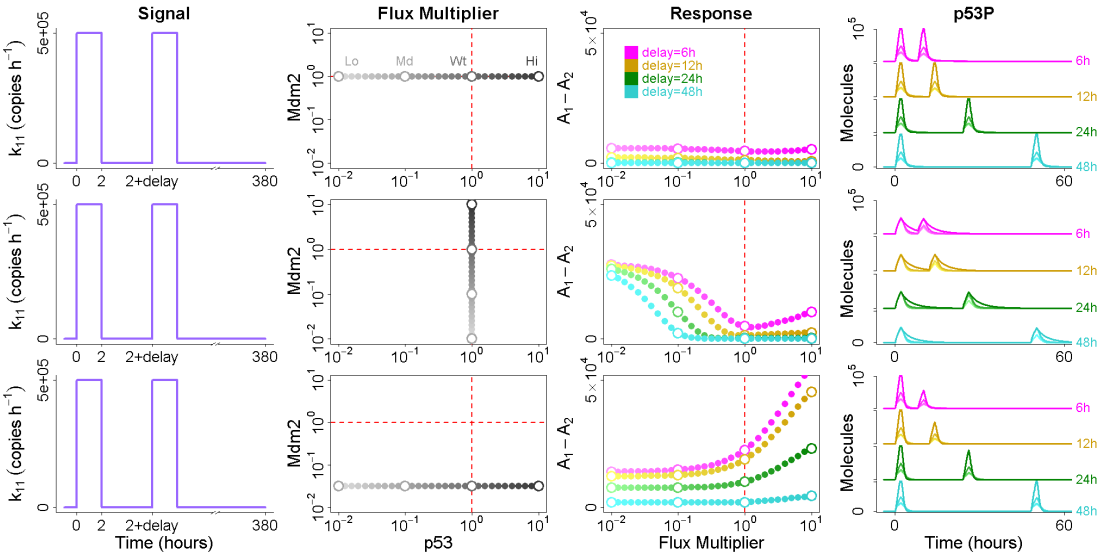


Figure 3.6: **Choice of interval time does not affect the role of Mdm2 flux in p53 refractory time.** This plot is identical to Figure 3.2.2D-E, except that the interval between pulses is taken to be 6 (magenta), 12 (yellow), 24 (green), or 48 hours (cyan). Representative traces at right are grouped according to interval time.

Sensitivity to the second pulse is defined as the difference between A_1 and A_2, with $A_1 - A_2 = 0$ indicating full sensitivity. As seen in Figures 3.2.2D and E, the flux of p53 has no bearing on the sensitivity to the second pulse while the flux of Mdm2 strongly affects it. At one one-hundredth the observed Mdm2 flux — corresponding to protein half-life of 3-days — over 20,000 fewer molecules of p53 are phosphorylated, representing more than a two-fold reduction in sensitivity (Figure 3.2.2E). This result is robust with respect to the interval of time chosen between pulses (Figure 3.2.2). If the sensitivity to the second pulse is already compromised by a reduced Mdm2 flux, a concomitant reduction in p53 flux fails to rescue it, while an increase in p53 flux still further reduces it (Figure 3.2.2F). We therefore conclude that the flux of Mdm2, and not p53, controls the system's refractory time, and a high Mdm2 flux is required to re-establish sensitivity after transient stimulation.

3.2.3 High IκB flux buffers NF-κB from activation in response to UV and low doses of TNF

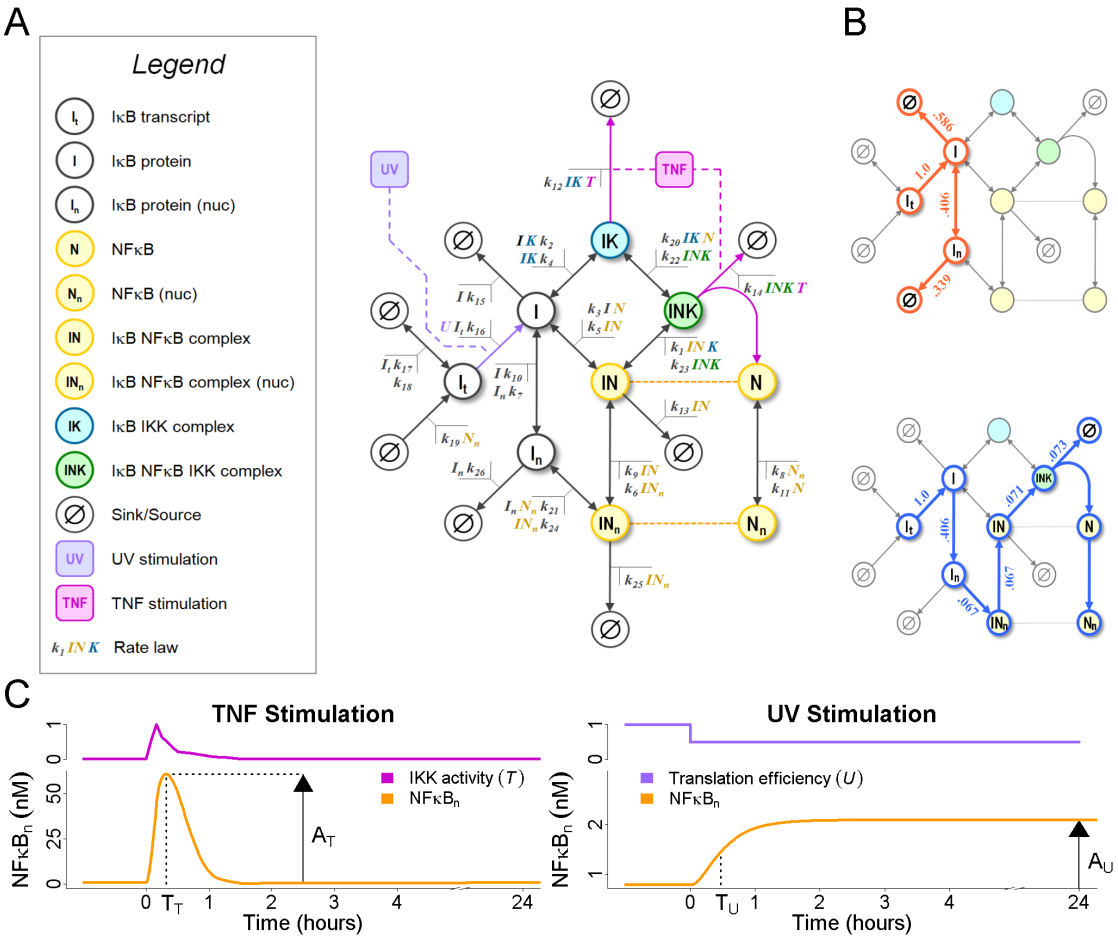
A second major stress-response pathway is that of NF-κB. NF-κB is potently induced by the inflammatory cytokine TNF, but shows a remarkable resistance to internal metabolic perturbations or ribotoxic stresses induced by ultraviolet light (UV) [23], or to triggers of the unfolded protein response (UPR) [135]. Like p53, the dynamics of NF-κB activation play a major role in determining target gene expression programs [10, 12]. Although NF-κB is considered stable, the flux of IκBα — the major feedback regulator of NF-κB — is conspicuously high. We hypothesized that turnover of IκB controls the stimulus-responsiveness of the NF-κB signaling module.

Beginning with a published model of NF-κB activation [23], we removed the beta and epsilon isoforms of IκB, leaving only the predominant isoform, IκBα (hereafter, simply "IκB"; Figure 3.2.3A). Steady state analysis of this model supported the observation that almost all IκB is degraded by either of two pathways: a "futile" flux, in which IκB is synthesized and degraded as an unbound monomer; and a "productive" flux, in which free IκB enters the nucleus and binds to NF-κB,

shuttles to the cytoplasm, then binds to and is targeted for degradation by IKK (Figure 3.2.3B). These two pathways account for 92.5% and 7.3% of the total IκB flux, respectively. The inflammatory stimulus TNF was modeled as before, using a numerically-defined IKK activity profile derived from *in vitro* kinase assays [12] (Figure 3.2.3A, variable T). Stimulating with TNF results in strong but transient activation of NF-κB. A second stimulus, ribotoxic stress induced by UV irradiation, was modeled as 50% reduction in translation and results in only modest activity [23]. As above, we let A_T be the amplitude of activated NF-κB in response to TNF and T_T the time at which A_T is observed. Analogously, we let A_U be the amplitude of NF-κB in response to UV, and T_U the time at which NF-κB activation equals one-half A_U (see Figure 3.2.3C). We then implemented multipliers for the futile and productive flux and let each multiplier take values on the interval $[10^{-2}, 10^{+1}]$. For each value we simulated the NF-κB response to TNF and UV and plotted the effects on A and T.

The results show that reducing the productive flux yields a slower, weaker response to TNF (Figure 3.2.3A). By analogy to Figure 3.2.1, this indicates that the productive flux of IκB is a positive regulator of NF-κB activation. In contrast, the futile flux acts as a negative regulator of NF-κB activity, though its effects on A_T and T_T are more modest (Figure 3.2.3B). Thus, similar to p53, the activation of NF-κB is controlled by a positive and negative regulatory flux. In response to UV, a reduction in either flux delays NF-κB activation, but reducing the futile flux results in a significant increase in A_U while reducing the productive flux has almost no effect (Figure 3.2.3C and D). Conversely, while an increase in the futile flux has no effect on A_U, an increase in the productive flux results in a significant increase. If we now define NF-κB to be sensitive to TNF or UV when A_T or A_U are ten-fold higher than its active but pre-stimulated steady state abundance, then TNF sensitivity requires a productive flux multiplier $> 10^{-1.6}$, while UV insensitivity requires a productive flux multiplier $< 10^{0.7}$ and a futile flux multiplier $> 10^{-0.8}$. This suggests that the flux pathways of IκB may be optimized to preserve NF-κB sensitivity to external inflammatory stimuli while minimizing sensitivity to internal metabolic stresses.

Figure 3.7: **A model of IκB-regulated NF-κB activation. A.** An IκB-centric diagram of NF-κB regulation. IκB is transcribed in an NF-κB-dependent and -independent manner. Translated IκB may bind to IKK (cyan), NF-κB (yellow), or both (green), or it may shuttle to the nucleus and bind to NF-κB there. Degradation of IκB is possible from any state, though only when bound to IKK can degradation be enhanced by IKK activity. Activation of NF-κB is achieved by the time-dependent numerical inputs T (magenta) and U (violet). T represents the activity of IKK kinase while U is the efficiency of mRNA translation. Both are defined over the interval $[0, 1]$, with $T(t < 0) = 0.01$ and $U(t < 0) = 1.0$ being their wildtype, unstimulated values. **B.** The futile (red) and productive (blue) IκB degradation fluxes. The fraction of total IκB flux through each reaction is listed next to the corresponding reaction arrow. **C.** Two stimuli used in our analysis of NF-κB activation and the effects of IκB flux. Stimulation by TNF is modeled using the time-dependent IKK activation profile described in [12] and results in strong but transient activation of NF-κB. Stimulation by UV is modeled as a 50% reduction of translational efficiency, as described in [23], and results in modest but sustained activation. As with p53, we define A_T and A_U to be he maximum activity of NF-κB in response to TNF and UV, respectfully, and T_T to be the time at which A_T is observed. Because A_T is observed infinitely often, we define T_U to be the time at which NF-κB activation reaches one-half A_U.


A
Legend
IκB transcript
IκB protein
IκB protein (nuc)
NFκB
NFκB (nuc)
IκB NFκB complex
IκB NFκB complex (nuc)
IκB IKK complex
IκB NFκB IKK complex
Sink/Source
UV stimulation
TNF stimulation
Rate law
UV
TNF
IK
INK
IN
N
B
C
TNF Stimulation
IKK activity (T)
NFκB_n
NFκB_n (nM)
Time (hours)
UV Stimulation
Translation efficiency (U)
NFκB_n
NFκB_n (nM)
Time (hours)

In contrast to p53, the negative regulatory flux of IκB dominates the positive flux. We hypothesized that this imbalance must affect the sensitivity of NF-κB to weak stimuli. To test this hypothesis we generated dose-response curves for TNF and UV using the following multipliers for the futile flux: 10^{-2}, 10^{-1}, 10^{0}, and 10^{+1}. The results confirm that reducing the futile flux of IκB results in hypersensitivity at low doses of TNF (Figure 3.3A). At one one-hundredth the wildtype flux, a ten-fold weaker TNF stimulus yields an equivalent NF-κB response to the full TNF stimulus at the wildtype flux. Similarly, a high futile flux prevents strong activation of NF-κB in response to UV (Figure 3.2.3B). At 10^{-1} and 10^{-2} times the futile flux, UV stimulation results in a 20-fold increase in NF-κB activity, compared to just a 2-fold increase at the wildtype flux. We therefore conclude that turnover of unbound IκB controls the EC50 of the NF-κB signaling module, and that rapid turnover renders NF-κB resistant to metabolic and spurious inflammatory stimuli.

3.3 Discussion

Previous studies have shown that the fluxes of p53 [121, 122], its inhibitor Mdm2 [136, 137], and the unbound negative regulator of NF-κB, IκB [123], are remarkably high. To investigate whether rapid turnover of these proteins is required for the stimulus-response behavior of the p53 and NF-κB stress response pathways, we developed a computational method to alter protein turnover, or flux, independently of steady state protein abundance.

For p53, we show that high flux is required for sensitivity to sustained stimulation after ionizing radiation (Figure 3.2.2A). Interestingly, inactivating mutations in p53 have long been known to enhance its stability [138], either by interfering with Mdm2-catalyzed p53 ubiquitination [139, 140], or by affecting p53's ability to bind DNA and induce the expression of new Mdm2 [141, 142, 143, 144]. Inactivation of p53 also compromises the cell's sensitivity to IR [15, 14, 145, 146]. Our results offer an intriguing explanation for this phenomenon, that p53 instability is required for oscillations in response to IR. Indeed, IR sensitivity was shown to correlate with p53 mRNA abundance [147, 148, 149], a likely determinant of p53

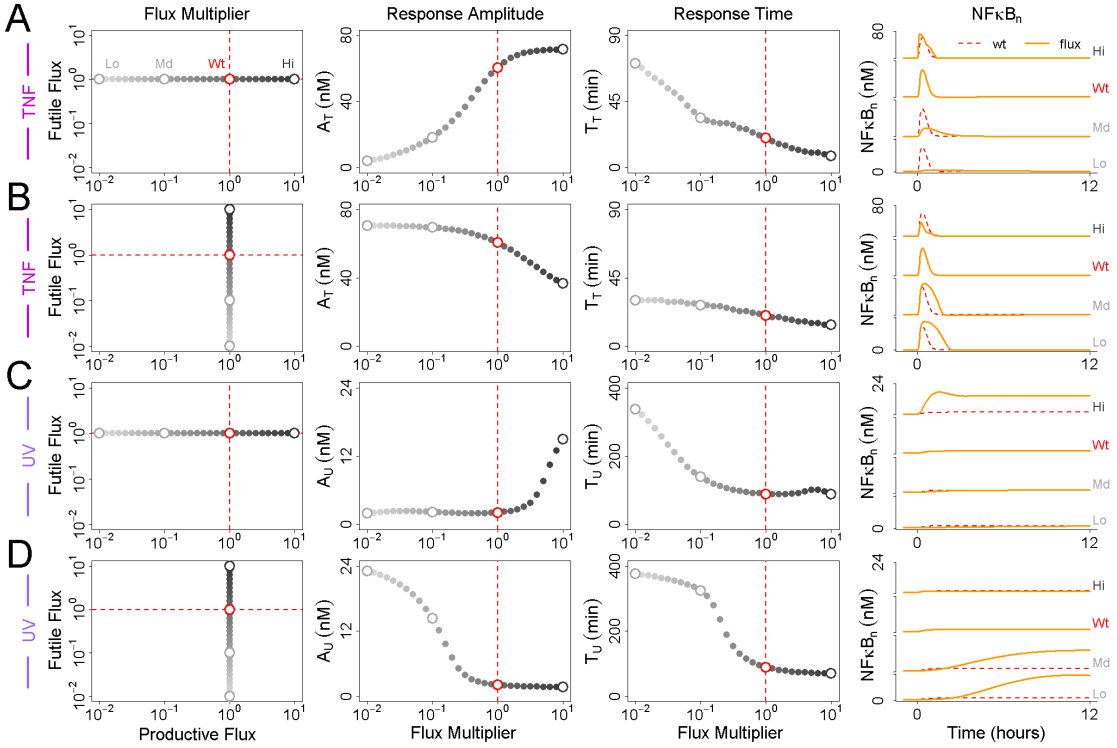


Figure 3.8: **Effects of IκB flux on the NF-κB response to stimulation.** **A.** The productive flux of IκB was varied between 10^{-2} and 10^{+1} times its wildtype value prior to stimulation by TNF (light gray to dark gray), and the resulting NF-κB response values A_T and T_T plotted in columns 2 and 3. Representative nuclear NF-κB profiles for low, moderate, wildtype, and high values of the flux multiplier are shown at right. Again, the wildtype productive flux is indicated by the dashed line in red. **B.** The futile flux of IκB was varied between 10^{-2} and 10^{+1} times its wildtype value prior to stimulation by TNF and the resulting NF-κB response values A_T and T_T plotted in columns 2 and 3. **C** and **D.** As (A) and (B), above, but the response to UV stimulation is plotted instead of TNF.

protein flux. In further support of this hypothesis, mouse embryonic fibroblasts lacking the insulin-like growth factor 1 receptor (IGF-1R) exhibit reduced p53 synthesis and degradation, but normal protein abundance. These cells were also shown to be insensitive to DNA damage, caused by the chemotherapeutic agent etoposide [137].

Like p53, increased stability of Mdm2 has been observed in human leukemic cell lines [150], and Mdm2 is a strong determinant of IR sensitivity [16, 17]. Again our results suggest these observations may be related. Activation of p53 in response to IR is mediated by the ATM kinase ("Signal" in Figure 3.2.2) [151, 152]. Batchelor *et al.* show that saturating doses of IR result in feedback-driven pulses of ATM, and therefore p53 [35]. In Figure 3.2.2B we show that these are independent of Mdm2 flux. However, sub-saturating doses of IR (10 Gy versus 0.5 Gy) [153, 133] cause only transient activation of ATM [154], after which constitutive Mdm2 synthesis is required to restore p53 sensitivity (Figure 3.2.2E). This suggests that high Mdm2 flux is required for sensitivity to prolonged exposure to sub-saturating doses of IR. Indeed, this inverse relationship between flux and refractory time has been observed before. In Ba/F3 pro-B cells, high turnover of the Epo receptor maintains a linear, non-refractory response over a broad range of ligand concentrations [22].

For NF-κB, our method revealed that an isostatic reduction in the half-life of IκB sensitizes NF-κB to TNF (Figure 3.3A), as well as to ribotoxic stress agents like UV (Figure 3.3B). This observation agrees with previous theoretical studies using a dual kinase motif, where differential stability in the effector isoforms can modulate the dynamic range of the response [155]. For NF-κB, the flux of free IκB acts as a kinetic buffer against weak or spurious stimuli, similar to serial post-translational modifications on the T cell receptor [156], or complementary kinase-phosphatase activities in bacterial two-component systems [157]. In contrast, increasing the half-life of IκBα alone — without a coordinated increase in its rate of synthesis — increases the abundance of free IκBα and actually dampens the activity of NF-κB in response to TNF [40]. This difference highlights the distinction between isostatic perturbations and traditional, unbalanced

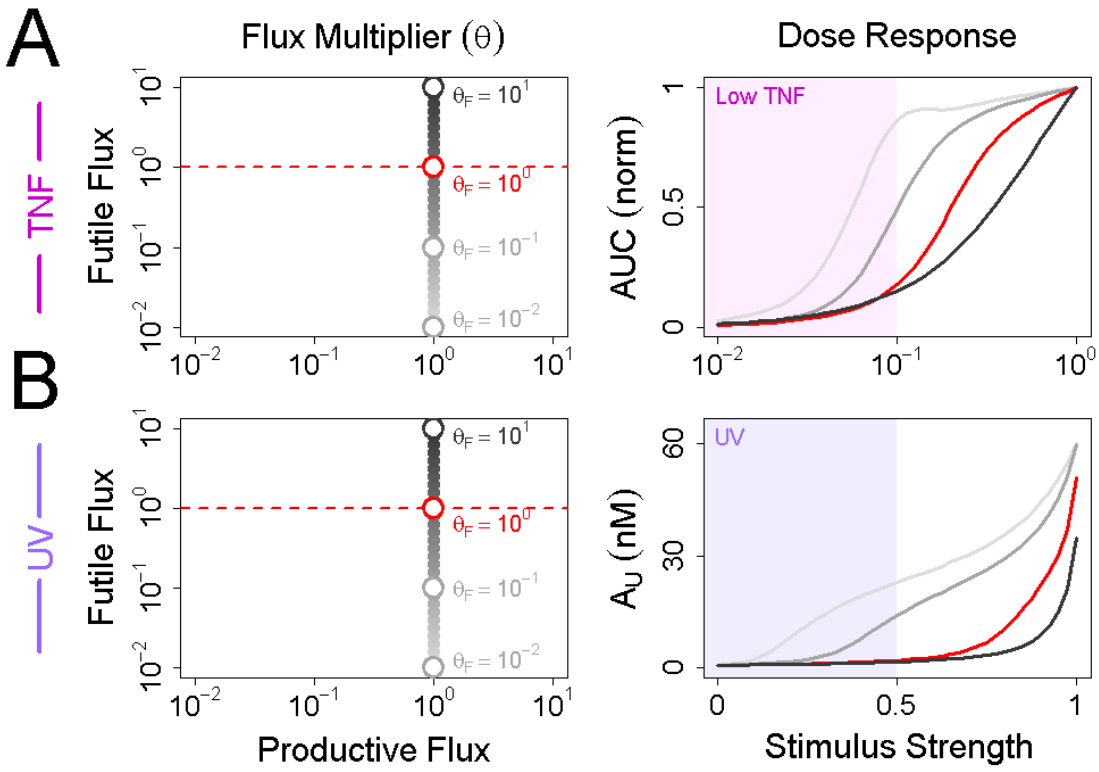


Figure 3.9: **IκB flux controls the sensitivity of NF-κB to stimulation by TNF and UV. A.** The futile flux of IκB was varied between 10^{-2} and 10^{+1} times its wildtype value prior to stimulation with variable doses of TNF (see Methods). For low, medium, high, and wildtype values of the futile flux, the area under the NF-κB activation curve is plotted as a function of TNF dose. The region of the plot corresponding to low doses of TNF, where the activation of IKK does not exceed 10%, is shaded in pink. (B) As above, but variable doses of UV are used instead of TNF. Because the response to UV is sustained and not transient, we have plotted the value of A_U as a function of UV dose instead of the area under the NF-κB activation curve.

.

perturbations that also affect the steady state abundances. It also calls attention to a potential hazard when trying to correlate stimulus-responsiveness with protein abundance measurements: observed associations between responses and protein abundances do not rule out implied changes in kinetic parameters as the causal link. Indeed static, and not kinetic measurements, are the current basis for molecular diagnosis of clinical specimens. Thus while nuclear expression of p53 [158, 159, 160, 161, 162, 163, 164, 165] and NF-κB [166, 167, 168] have been shown to correlate with resistance to treatment in human cancer, the correlation is not infallible [15, 169, 170, 171, 172, 173]. If stimulus-responsiveness can be controlled by protein turnover independently of changes to steady state abundance, then correlations between abundance and a therapeutic response may be masked by isostatic heterogeneity between cells.

For p53 and NF-κB, we show that stimulus sensitivity can be controlled by a paired positive and negative regulatory flux. We propose that this pairing may constitute a common regulatory motif in cell signaling. In contrast to other regulatory motifs [174, 175], the "flux motif" described here does not have a unique structure. The positive p53 flux, for example, is formed by the synthesis and degradation of p53 itself, while the positive flux in the NF-κB system includes the nuclear import of free NF-κB and export of NF-κB bound to IκB. For p53, the negative flux is formed by synthesis and degradation of Mdm2, while for NF-κB it is formed by the synthesis, shuttling, and degradation of cytoplasmic and nuclear IκB. Thus the reaction structure for each flux is quite different, but they nevertheless form a regulatory motif that is common to both pathways (Figure 3.3). And since the mathematical models used here are only abstractions of the underlying network, the true structure of the p53 and NF-κB flux motifs are in reality even more complex.

Finally, in this study we have examined the effects of flux on stimulus-responsiveness, but in a typical signaling module, many other isostatic perturbations exist. For example, the isostatic subspace of our NF-κB model has 18 dimensions, of which only a few were required by the analysis presented here. By simultaneously considering all isostatic perturbations, some measure of the

dynamic plasticity of a system can be estimated, perhaps as a function of its steady state. Such an investigation can inform diagnosis of biological samples, and whether information from a single, static observation is sufficient to predict the response to a particular chemical treatment, or whether live-cell measurements are required as well. As we have shown that protein turnover can be a powerful determinant of stimulus-sensitivity, we anticipate that kinetic measurements will be useful predictors of sensitivity to chemical therapeutics.

3.4 Methods

3.4.1 Modeling isostatic perturbations in protein turnover

To begin, we assume that the system of interest has been modeled using mass action kinetics and that the steady state abundance of every biochemical species is a known function of input parameters. In other words,

$$\overline{\mathbf{x}} = f(\mathbf{p})$$

such that

$$\frac{d\overline{\mathbf{x}}}{dt} = \mathbf{0}. \tag{3.1}$$

Equation 3.1 is the well-known steady state equation; $\mathbf{p}$ is a vector of independent parameters, and $\mathbf{x}$ is the vector of species abundances. As stated in Chapter 1, we use an overbar to denote a vector $\mathbf{x}$ that satisfies Equation 3.1. For excellent reviews on mass action models and their limitations, see [81-83]. For a method on finding analytical solutions to the steady state equation, see our accompanying manuscript. Next, we wish to find a change $\Delta\mathbf{p}$ in the input parameters such that the resulting change $\Delta\mathbf{x}$ in the species abundances is zero, where $\Delta\mathbf{x}$ is defined as

$$\Delta\mathbf{x} = \overline{\mathbf{x}}(\mathbf{p} + \Delta\mathbf{p}) - \overline{\mathbf{x}}(\mathbf{p}).$$

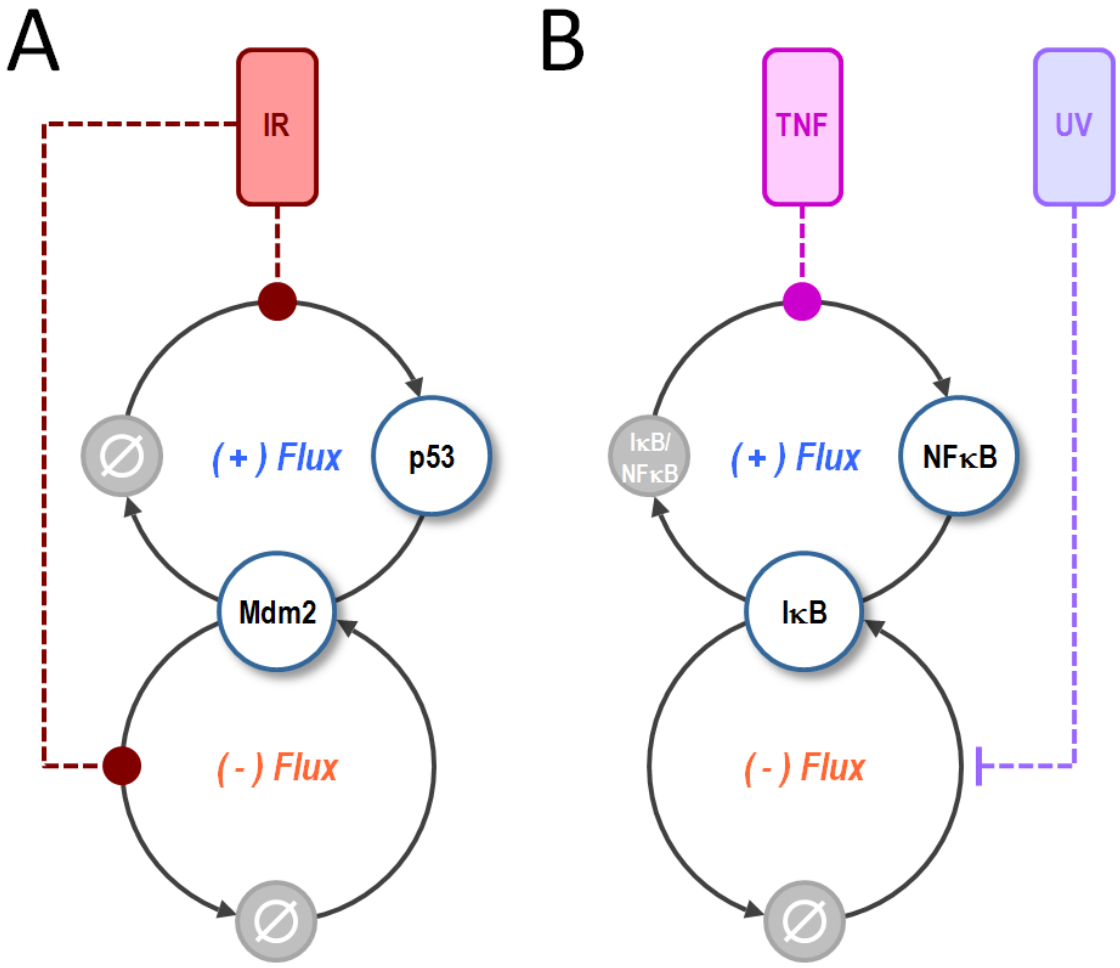


Figure 3.10: **A paired positive and negative flux motif controls stimulus-sensitivity in the p53 and NF-κB stress response pathways. A.** For p53, the positive (+) flux is formed by the synthesis and degradation of p53 itself. The negative (-) flux is formed by synthesis and degradation of Mdm2. Together these fluxes control the sensitivity of p53 to IR-stimulation, which acts by inducing the synthesis of p53 and the degradation of Mdm2. **B.** For NF-κB the (+) flux is formed by association and dissociation of NF-κB from its negative regulator, IκB. The (-) flux is formed by synthesis and degradation of IκB. These fluxes control the sensitivity of NF-κB to TNF-stimulation, which induces the dissociation of NF-κB from IκB, and UV-stimulation, which inhibits the synthesis of IκB.

.

Thus for $\Delta\mathbf{x} = \mathbf{0}$, we require that

$$\overline{\mathbf{x}}(\mathbf{p}) = \overline{\mathbf{x}}(\mathbf{p} + \Delta\mathbf{p}).$$

The right-hand side of this equation can be approximated by a truncated Taylor series, as follows:

$$\overline{\mathbf{x}}(\mathbf{p} + \Delta\mathbf{p}) \approx \overline{\mathbf{x}}(\mathbf{p}) + \mathbf{J}_x\Delta\mathbf{p},$$

where $\mathbf{J}_x$ is the Jacobian matrix whose elements are the partial derivatives of each species with respect to each parameter. Thus, for $\Delta\mathbf{x} = \mathbf{0}$ we require that

$$\mathbf{J}_x\Delta\mathbf{p} = \mathbf{0}.$$

In other words, $\Delta\mathbf{p}$ must lie in the null space of $\mathbf{J}_x$. We call this the *isostatic subspace* of the model — parameter perturbations in this subspace will not affect any of the steady state species abundances. If $\Delta\mathbf{p}$ lies within the isostatic subspace, it is an *isostatic perturbation vector*. Let $\mathbf{N}_x$ be a matrix whose columns form a basis for the isostatic subspace. Then a general expression for an isostatic perturbation vector is simply

$$\mathbf{p} = \mathbf{N}_x\mathbf{q}, \tag{3.2}$$

where $\mathbf{q}$ is a vector of unknown basis vector coefficients. Finally, Equation 3.2 can be solved for a specific linear combination of basis vectors that achieves the desired perturbation. In our case we identified those combinations that result in changes to protein turnover.

3.4.2 A prototypical negative feedback model

Our prototypical negative feedback model consists of two species, an activator X and an inhibitor Y, and four reactions, illustrated in Figure 3.2.1A. Let $\bar{x}_1$ denote the steady state abundance of the activator and $\bar{x}_2$ denote the abundance

of the inhibitor. An analytical expression for the steady state of this model was identified by solving Equation 3.1 for the rates of synthesis, giving

$$k_1 = k_3 \bar{x}_1 \bar{x}_2 \tag{3.3}$$

$$k_2 = k_4 \frac{\bar{x}_2}{\bar{x}_1}. \tag{3.4}$$

To parameterize the model we first let $\bar{x}_1 = \bar{x}_2 = 1$ a.u.. Degradation rate constants were then calculated such that $A = 10$ a.u. at time $T = 24$, where again A is the maximum amplitude of the response. Activation was achieved by letting $x_2(t = 0) \to 0$. To modify the flux, we defined flux multipliers θ_1 and θ_2 such that $k_3' = \theta_1 k_3$ and $k_4' = \theta_2 k_4$. Note that by virtue of Equations 3.3 and 3.4, values for θ_1 and θ_2 other than unity result in commensurate changes in k_1 and k_2 such that steady state is preserved. Figures 3.2.1A and B were achieved by letting θ_1 and θ_2 vary over the interval $[10^{-1}, 10^{+1}]$, then calculating the altered vector of rate constants $\mathbf{k}'$ and simulating the model's response to stimulation. Figure 3.2.1C required letting θ_1 vary over this same interval while having $\theta_2 = \theta_1$. Finally, Figure 3.2.1D was achieved by letting θ_2 vary over the same interval, and for each value of θ_2, numerically calculating the value of θ_1 that gave $T = 24$.

3.4.3 A model of p53 oscillations

All species, reactions, and rate equations required by our model of p53 oscillations are as previously described [35]. Our only modification was to scale the parameter values so that the rates of p53 and Mdm2 synthesis and degradation, as well as their steady state abundances, matched published observations (see Table 3.2.2). Specifically we let 1 Cs $= 5 \times 10^4$ molecules and 1 hour (Batchelor *et al.*) be 1.33 hours. To derive a steady state solution for this model, we solved Equation 3.1 for the steady state abundance of Mdm2 and the rate of Mdm2-independent p53 degradation, giving

$$k_1 = \frac{k_9}{P} - \frac{k_5 k_{10}}{k_2}, \text{ and}$$
$$M = \frac{k_{10}}{k_2}.$$

To simulate the response to ionizing radiation we used the (scaled) stimulus given in [35]. Namely, at time $t = 0$ we let the rate of Signal production, k_{11}, go to 5×10^5 molecules hour^{-1}. This stimulus was either maintained indefinitely (Figures 3.2.2A-C) or for just 2 hours, followed by 22 hours of rest, followed by a second 2 hour stimulation (Figures 3.2.2D-F). Changes in p53 or Mdm2 flux were achieved as above, by defining modifiers θ_P and θ_M such that

$$k_9' = k_5 P M(\theta_P - 1), \tag{3.5}$$
$$k_2' = \theta_M k_2, \text{ and} \tag{3.6}$$
$$k_{10}' = \theta_M k_{10}. \tag{3.7}$$

Prior to stimulation, we let one modifier take values on the interval $[10^{-2}, 10^{-1}]$ while holding the other modifier constant. Equations 3.6 and 3.7 ensure that the p53-independent flux of Mdm2 is modified without affecting its steady state abundance. Equation 3.5, which is slightly more complicated, results in changes to the rate of Mdm2-indpendent p53 degradation, k_1, by modifying the independent parameter k_9, which controls the rate of p53 synthesis. This yields the desired

$$k_1' = \theta_P k_1.$$

Numerical integration was carried out to time $t = 384$ hours. After each integration, we defined A_∞ to be the minimum vertical distance between any adjacent peak and trough in phosphorylated p53, and A_1 and A_2 to be the amplitudes of the first and second peak, respectively. For more information on the time delay parameters τ_i and τ_m, and their role in generating oscillations, see [176, 177].

3.4.4 A model of NF-κBactivation

Our model of NF-κB activation is similar to the one described in [23], except the beta and epsilon isoforms of IκB have been removed. Our model has 10 species and 26 reactions, the majority of which are illustrated in Figure 3.2.3A. Rate equations and parameter values are identical to those in [23]. An analytical expression for the steady state of this model was found by solving Equation 3.1 for the following dependent variables: I, IK, INK, I_n, and IN, and the rate constants k_{11}, k_{16}, and k_{19}.

Activation of NF-κB is achieved by either of two, time-dependent numerical input variables, T and U. T modifies the activity of IKK while U modifies the efficiency of IκB translation. Both have a finite range of $[0, 1]$ and have unstimulated, wildtype values of $T(t < 0) = 0.01$ and $U(t < 0) = 1.0$, respectively. The inflammatory stimulus TNF is modeled using a unique function $T(t)$ derived from *in vitro* kinase assays [12]. Since these assays only measured IKK activity out to 4 hours, we extended each stimulus by assuming the value $T(t = 4$ hours$)$ is maintained out to $t = 24$ hours. Justification for this can be found in the 24-hour kinase assays in [86], which shows no IKK activity between 8 and 24 hours after TNF stimulation. UV stimulation is modeled using a step decrease in the value of U from 1.0 to 0.5 for the entire 24 hours. This mimics the 50% reduction in translational efficiency observed in [23].

Steady state analysis of this model revealed that over 99% of all IκB was degraded via either of two pathways, futile (92%) and productive (7%). See Figure3.2.3B for the composition of these pathways. To modify the flux through either pathway without altering any of the steady state abundances, the algebraic method described above proved absolutely necessary. Specifically, we solved Equation 3.2 for the unique set of basis vector coefficients such that the following conditions held: (1) only reaction rate constants involved in the targeted pathway were modified; (2) if a reaction on the pathway was reversible, its ratio of forward to reverse rate constants was preserved; and (3) the magnitude of an alteration was relative to the bottleneck reaction. For the futile flux this was k_{26}, the degradation of unbound IκB. For the productive flux it was k_6, the export of NF-κB-bound IκB.

As in the p53 models above, we then defined multipliers θ_F and θ_P such that

$$k'_{26} = \theta_F k_{26}$$
$$k'_6 = \theta_P k_6.$$

Finally, to generate Figure 3.2.3 we let the appropriate multiplier take values on the interval $[10^{-2}, 10^{+1}]$ prior to stimulation with TNF or UV.

Dose response curves in Figure 3.3 were generated by letting θ_F take values in $\{10^{-2}, 10^{-1}, 10^{0}, 10^{+1}\}$ and simulating the response to varying doses of TNF and UV. To vary the TNF dose, we scaled the displacement of the numerical IKK activation curve above its basal value of 1% using log-spaced multipliers on the interval $[10^{-2}, 10^{0}]$. We call this multiplier the "stimulus strength". A stimulus strength of $10^{-1.0}$, for example, yields the same basal IKK activity as the full TNF dose used in Figure 3.2.3, but a peak activity whose magnitude is just one-tenth that of the full dose. To measure the TNF response, we calculated an area under the curve (AUC) by subtracting NF-κB basal activity from the TNF-induced NF-κB activation curve, then integrated this curve from the point of stimulus to the time at which it becomes less than one-tenth the basal activity. All AUCs were normalized to the full TNF dose. To vary the UV dose we varied the magnitude of the displacement of U from unity. A stimulus strength of 0.1, for example, results in a step decrease in U from 1.0 to 0.9. Because the response to UV is sustained instead of transient, we plotted A_U as a function of stimulus strength instead of the area under the curve.

Chapter 3, in full, is a reprint of material as it appears in PLoS Computational Biology. Loriaux, Paul M; Hoffmann, Alexander, Public Library of Science 2013. The dissertation author was the primary investigator and author of this paper.

Chapter 4

Kinetic network features are better predictors of TRAIL-induced cell death than static features

TNF-related apoptosis inducing ligand, or TRAIL, is a promising anti-cancer agent for its ability to selectively kill transformed cells. Intercellular heterogeneity, however, even in a clonal population of cells, causes variability in the response to TRAIL and poses a significant challenge to the discovery of prognostic biomarkers. To identify functionally relevant predictors of TRAIL sensitivity in a heterogeneous population of cells, we extensively sampled a validated mass action model of TRAIL-induced cell death using empirically-derived distributions of biochemical parameters. By quadratic programming feature selection, we find that just four features can emulate the full ODE model to within 78% accuracy. Remarkably, these four features are kinetic, not static. That is, they describe the rates of synthesis and degradation of Bar and XIAP, but not the abundances of the proteins themselves. This result is robust to large changes in parameter values, and argues that significant predictive information is lost when clinical specimens are fixed after biopsy instead of being used for live cell assays.

4.1 Introduction

The development of prognostic biomarkers to predict the cellular response to perturbation will revolutionize the diagnosis and treatment of human cancer [178, 179]. Overexpression of the estrogen or progesterone receptor, for example, or expression of human epidermal growth factor receptor 2, identifies breast cancers that are effectively treated by the adjuvants Trastuzumiab or Tamoxifen [180, 181]. In spite of these and other examples, however [182], identifying clinically useful biomarkers remains a considerable challenge [183, 184]. First, many candidate biomarkers are identified using high-throughput modalities that are prone to overfitting [185]. These modalities cannot easily distinguish between causative abnormalities and consequences of disease progression or the host response [186]. Second, at the time of detection, most cancers exhibit significant genetic and biochemical heterogeneity. Chemical treatment of a heterogeneous cell population results in a fractional response [187], and may select for aggressive subpopulations that are resistant to further treatment [188, 189]. For these reasons, identification of cellular features that are both functionally relevant to the intended therapy and predictive across a heterogeneous population of cells will significantly improve biomarker validation.

To address this challenge we developed a computational framework that combines bottom-up network analysis with top-down machine learning to identify a parsimonious set of biochemical features that accurately predict the response to chemical perturbation in a heterogeneous population of cells. We illustrate our framework using the anti-cancer therapeutic, rhTRAIL/APO2L (hereafter, "TRAIL"). TRAIL is a TNF family member that preferentially induces apoptosis in transformed cells [190]. Several TRAIL analogues are currently in clinical trials [67], but the response to TRAIL is heterogeneous and cell-type specific [191, 68]. A recent study has further shown that sister cells exhibit a strong correlation in their sensitivity to TRAIL, and that this correlation decays over time [66]. Biochemical variability caused by stochastic gene expression is therefore sufficient to explain the variability in response to TRAIL, and argues that accurate prediction of the response is feasible if the right biochemical features could be measured *a priori*.

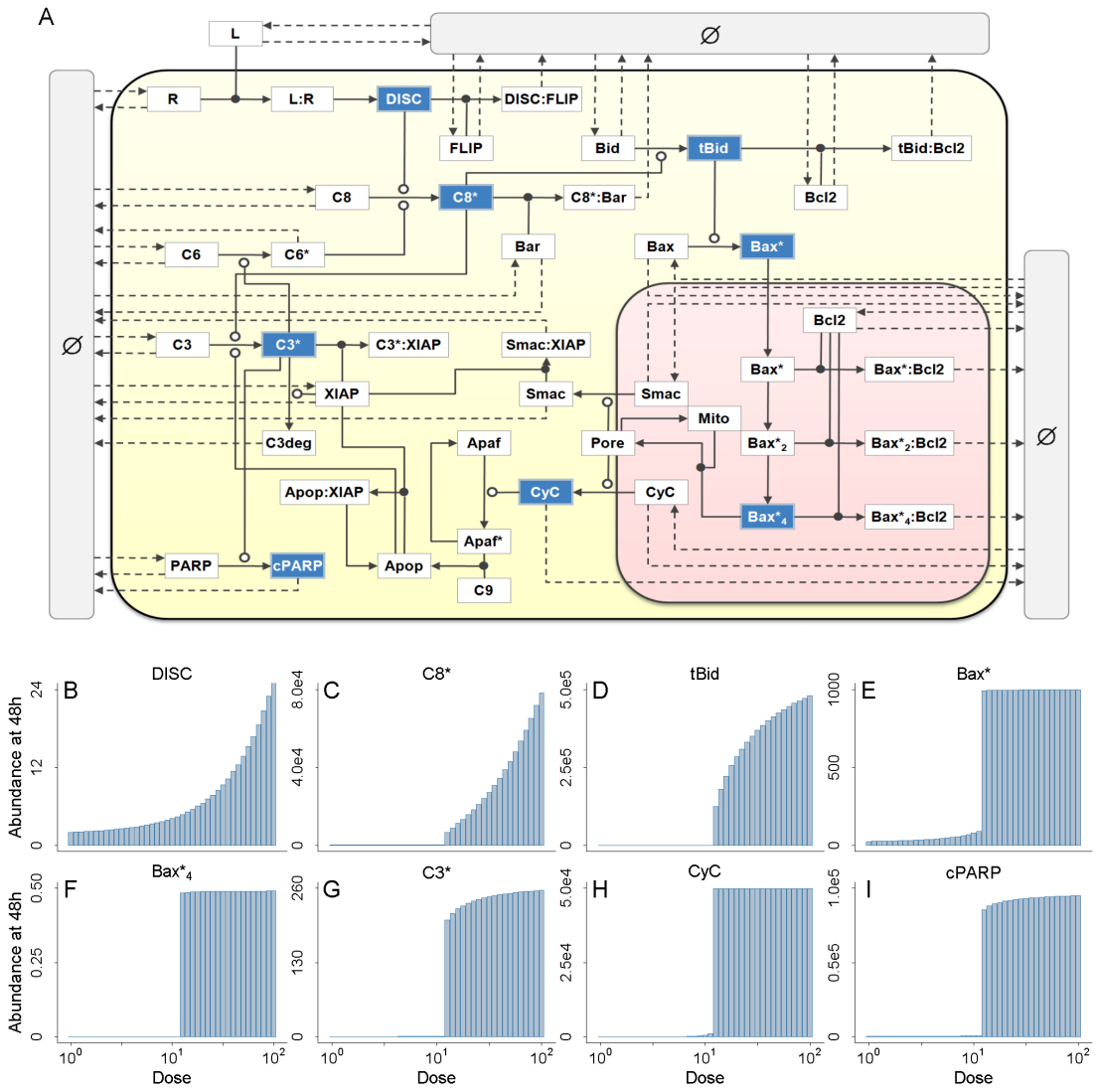
To identify these features, we used a validated model of TRAIL-induced cell death based on ordinary differential equations (ODEs) [71], extended to include protein synthesis and degradation and assigned an analytical, non-trivial steady state [192]. By letting protein abundance and turnover parameters be described by random variables conforming to empirically-derived distributions, we were able to rapidly evaluate the response to TRAIL in a heterogeneous population of cells. Using quadratic programming feature selection, we first confirm that steady state can indeed serve as an accurate predictor of the response to TRAIL, independent of stimulus-induced dynamics. Second, we found that just four steady state features can achieve 78% of the accuracy of the full ODE model. This represents a 30-fold reduction in parametric complexity, and only a modest reduction in accuracy. It also supports our conjecture that predicting the outcome of a complex biochemical network in response to perturbation can be achieved with only a few high-quality measurements. Third, these features are overwhelmingly kinetic rather than static. That is, we show that the rates of synthesis and degradation of a few key regulatory molecules are considerably more informative than the absolute abundances of those or all other molecules combined. If correct, this result has strong implications for clinical diagnosis of drug sensitivity; it argues that significant predictive information is lost when clinical specimens are fixed after biopsy instead of being used for live cell assays.

4.2 Results

4.2.1 A heterogeneous model of TRAIL-induced cell death

To begin, we use ordinary differential equations to model the time-dependent behavior of 58 molecular species in response to stimulation by TRAIL. Our model is based on [71] and is described further in [192]. For a review of TRAIL-induced signaling, see [193, 194]. Briefly, we model TRAIL-mediated assembly of the death-inducing signaling complex (DISC) and subsequent activation of caspase 8. Active caspase 8 cleaves and activates caspase 3, but also Bid, which in turn activates cytoplasmic Bax. Activated Bax then enters the mitochondria to form tetrameric

Figure 4.1: **A bistable model of TRAIL-induced cell death. A.** A model of TRAIL-induced cell death that includes the original 58 species and 70 reactions introduced in [71], as well as 15 zero-order synthesis reactions, 28 first-order degradation reactions, and 2 first-order back-reactions resulting in deactivation of Mito and Apaf. The 43 synthesis and degradation reactions are represented by dashed lines emanating from or terminating in a gray "null" compartment. This compartment is intended to symbolize the boundaries of the system, not the extracellular environment. Other symbols are as described in [71]. The cytoplasm and mitochondria are shaded in yellow and rose, respectively. Species selected for dose-response analysis are shaded in blue. Dose-response curves for the eight species in (A) shaded in blue are **B.** DISC, **C.** active caspase 8, **D.** truncated Bid, **E.** active Bax monomers in the cytoplasm **F.** tetrameric Bax in the mitochondria, **G.** active caspase 3, **H.** Cytochrome C in the cytoplasm, and **I.** cleaved PARP. Each panel depicts the absolute abundance of the species at 48 hours after stimulation with TRAIL. The dose of TRAIL ranges from 1 to 100-fold its ambient abundance.


A
L
R
L:R
DISC
DISC:FLIP
FLIP
Bid
tBid
tBid:Bcl2
C8
C8*
C8*:Bar
Bcl2
C6
C6*
Bar
Bax
Bax*
C3
C3*
C3*:XIAP
Smac:XIAP
XIAP
Smac
Mito
C3deg
Apaf
Pore
Bax*2
Bax*2:Bcl2
Apop:XIAP
CyC
Bax*4
Bax*4:Bcl2
Apaf*
PARP
cPARP
Apop
C9
B
DISC
C
C8*
D
tBid
E
Bax*
F
Bax*4
G
C3*
H
CyC
I
cPARP
Abundance at 48h
Dose

pores through which Cytochrome C is released into the cytoplasm. This event is commonly called mitochondrial outer membrane permeabilization, or MOMP. Once in the cytoplasm, Cytochrome C catalyzes the formation of the Apoptosome, which results in degradation of XIAP and feed-forward activation of caspase 3. Activation of caspase 3 is required for chromatin condensation and DNA fragmentation (not modeled), and ultimately results in cell death [195, 196]. Because MOMP is irreversible [197], we used the abundances of tetrameric Bax (Bax4) and caspase 3-cleaved poly (ADP-ribose) polymerase (cPARP) as indicators of cell fate [198, 42] (Figure 4.2.1A). In order to treat the steady state as an independent variable, we used py-substitution to derive an analytical expression for the steady state [199, 192]. The resulting expression has 119 independent parameters, 18 of which are species abundances, 100 of which are kinetic rate constants, and 1 of which describes the mitochondrial volume. The remaining 40 abundances and 15 rate constants are rational polynomials in the independent parameters.

To verify that our model was capable of distinguishing between high and low doses of TRAIL, we generated dose-response curves over a 100-fold increase in ligand abundance. At each dose we simulated the model to 48 hours and recorded the abundance of eight informative species (Figure 4.2.1B-I). The results clearly show two distinct dose-response regimes. At low doses there is a linear response in the DISC, but no caspase activation, Cytochrome C release, nor cleavage of PARP. At high doses there is complete Cytochrome C release, activation of caspases 3 and 8, and accumulation of cPARP. The model can therefore distinguish between perturbations that do and do not result in cell death. Next, we wanted to model a heterogeneous population of cells. Recent single-cell experiments suggest that protein abundances are gamma distributed [200]. This distribution arises naturally from Poisson production of messenger RNA, followed by exponentially-distributed bursts of protein translation [202]. A remarkable conclusion from this work is that for highly expressed proteins, the variance is proportional to the square of the mean. Using the abundances given in [71], we therefore calculated the variance, shape, and scale parameters of 14 independent species in our model (Figure 4.2.1A-B). By virtue of being constrained to the steady state, 40 depen-

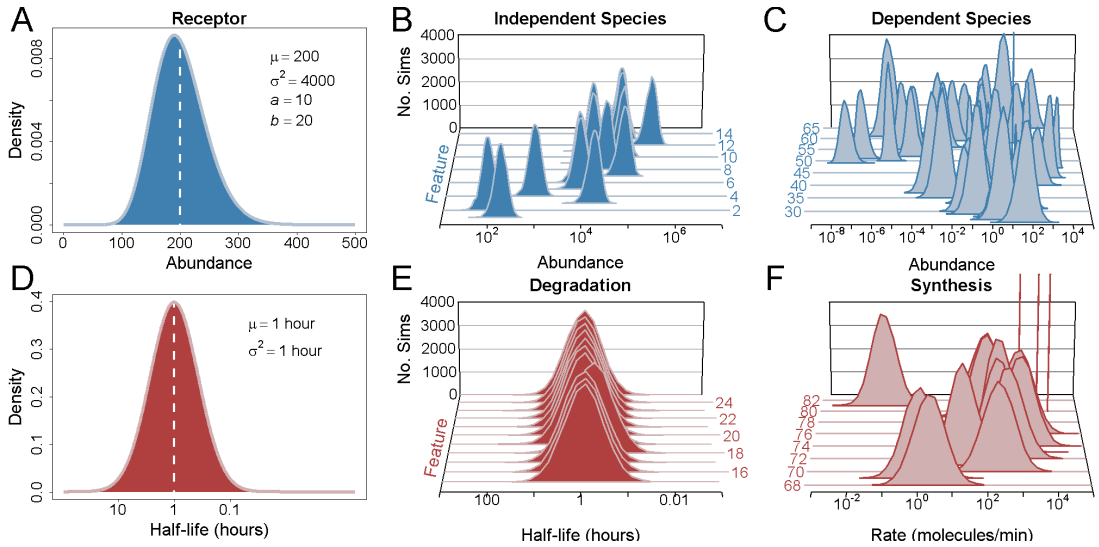


Figure 4.2: **Introducing heterogeneity into the bistable model of TRAIL-induced cell death. A.** The *a priori* probability density function used for the abundance of the TRAIL receptor. Like other independent species abundances, we let this function be defined by a gamma distribution. The mean μ of the distribution is taken to be the value given in [71]; its variance σ^2 is taken to be one-tenth the square of the mean, i.e. the extrinsic noise limit described in [200]. From these, a shape a and scale parameter b was calculated according to the definition of the gamma distribution. **B.** *A priori* probability density functions for all 14 independent species abundances. Species are plotted front to back according to their index. **C.** *A posteriori* probability density functions for all 40 dependent species abundances. Each dependent species abundance is a rational polynomial in the independent parameters. **D.** *A priori* probability density function for the half-life of the TRAIL receptor. This and other protein half-lives were assigned a standard log-normal distribution with a nominal half-life of one hour, consistent with the observation that signaling proteins experience rapid turnover [201]. **E.** *A priori* probability density functions for all 11 primary protein half-lives, plotted according to their feature index. **F.** *A posteriori* probability density functions for the kinetic rate constants describing the efflux of 15 modified proteins and protein complexes.

dent species also assumed a probability distribution, albeit with higher variance (Figure 4.2.1C). In addition, protein half-lives in murine 3T3 cells are known to be log-normally distributed [62]. We therefore let 11 degradation rate constants follow a log-normal distribution with a coefficient of variation (CV) equal to 0.368, equivalent to a variance of 1 hour in the log-normal distribution of protein half-lives (Figure 4.2.1D-E). This short half-life was motivated by the observation that signaling proteins tend to be short-lived [201]. Again by virtue of steady state, 15 synthesis rate constants assumed a probability distribution as well (Figure 4.2.1F).

Next we sampled the model 20,000 times and simulated its response to an ambiguous dose of TRAIL (1,000 ligands per cell). Examining the abundance of four species – active caspase 3, cPARP, tetrameric Bax (Bax4), and cytoplasmic Cytochrome C – at 48 hours after stimulation revealed two distinct subpopulations (Figure 4.2.1A-D). Cells that experienced MOMP achieved the hyperactive steady state indicative of cell death. Cells that did not experience MOMP returned to their prestimulated steady states. Due to the symmetry and distance of the two subpopulations of Bax4 at 48 hours, we chose this as our primary response variable. We note that the abundance of Bax4 in the responding population is between 1 and 10 tetramers, in good agreement with the observation that only a few mitochondrial pores are required for MOMP [203]. After fitting with a two-component Gaussian mixture model, those that returned to their prestimulated steady states were scored as "alive" and assigned a response value of zero. Those that did not were scored "dead" and assigned a response value of one.

4.2.2 Steady state is an accurate predictor of the response to TRAIL

After binarizing the response, we returned to the prior empirical parameter distributions and calculated their correlation with the response. Note that random-valued parameters in the forward numerical integration problem are now "features" in the backward model selection problem, and hereafter we refer to them as such. We also distinguish between static and kinetic features. Static features are the steady state abundances of the molecular species. Kinetic features are the

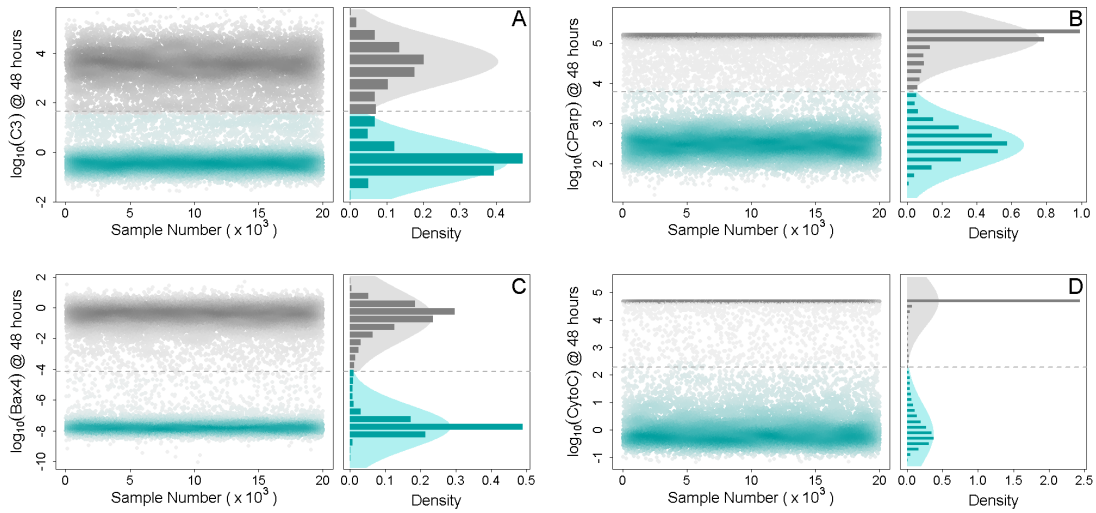


Figure 4.3: **Response of the heterogeneous model to an ambiguous dose of TRAIL.** Shown here are the simulated abundances of four species at 48 hours following the addition of 1,000 molecules of TRAIL, for each of the 20,000 samples. The four species are **A.** active caspase 3, **B.** cleaved Parp, **C.** tetrameric Bax, and **D.** cytoplasmic Cytochrome C. For each panel, the left plot show the absolute abundance of each species as a function of sequential sample number. Points are shaded by density and assigned a color post-analytically based on whether that abundance corresponds to a positive response (gray; cell dies), or a negative response (cyan; cell lives). The right plot in each panel shows a 1-D histogram of the response as well as the pair of Gaussians fitted to its density estimate (see Methods). The saddle point between the two Gaussians is indicated by the dashed line and distinguishes responsive from unresponsive samples.

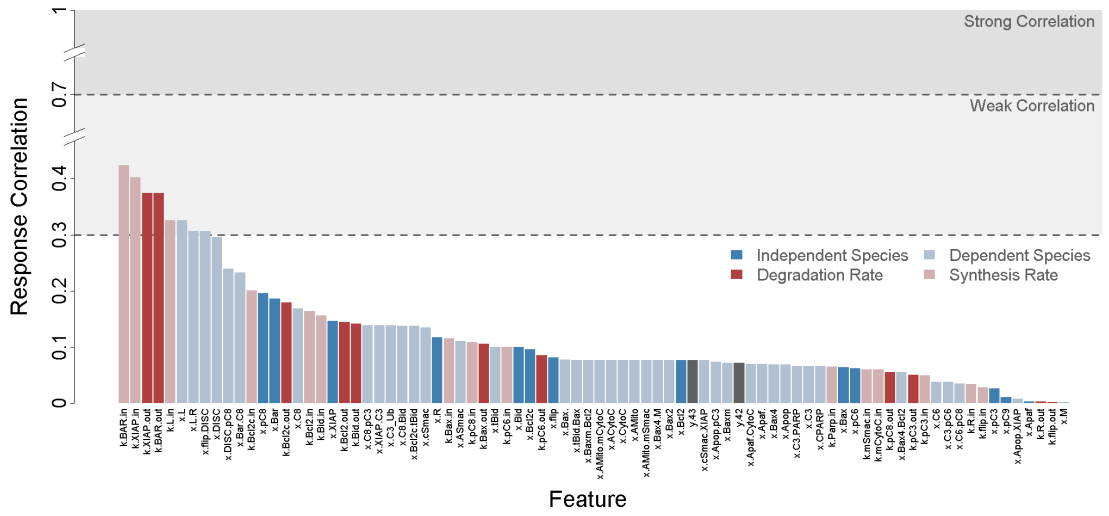


Figure 4.4: **Correlation between features and the binary response.** Pearson correlation coefficients were calculated between each feature and the binary response variable derived from the abundance of tetrameric Bax at 48 hours (see Methods). The absolute values of the coefficients are shown here, sorted in decreasing order and color-coded to indicate whether the feature is an independent species (dark blue) or degradation rate constant (dark red), or a dependent species (light blue) or rate of synthesis (light red). Light and dark gray rectangular regions in the plot illustrate regions of weak and strong correlation, respectively.

rate constants that relate species abundances to the reaction velocities in which they participate. As shown in Figure 4.2.2, no single feature correlates well with the response to TRAIL. This is in agreement with the conclusion drawn by [66], where no single protein exhibits strong correlation with the time elapsed between administration of TRAIL and MOMP, unless artificially overexpressed.

We therefore asked whether we could accurately predict the response to TRAIL using a small subset of features. A confounding issue here is that the feature set is highly cross-correlated. In other words, any two features may be redundant; knowing both would be no more informative than knowing either feature in isolation. To account for redundancy we calculated the cross-correlation matrix (Figure 4.2.2). As expected, every feature perfectly cross-correlates with itself, as illustrated along the diagonal. Because the rates of protein synthesis were constrained to steady state, paired synthesis and degradation rates are highly correlated as well (clusters K1 and K2). Cluster K2 contains synthesis rates that

are well-correlated with the 40 dependent species abundances. Among the static features, we observe four clusters. Cluster S4 contains the independent species abundances, all of which exhibit low correlation with every other feature. Clusters S1, S2, and S3 are the receptor-proximal, post-mitochondrial, and mitochondrial clusters, respectively. The high degree of cross-correlation observed in cluster S3 argues that if the state of the mitochondria is a good predictor of TRAIL sensitivity, then only one mitochondrial feature will need to be measured. Similar arguments can be made for clusters S1 and S2. Figure 4.2.2 also illustrates that dependent features can cross-correlate if their steady state expressions are dominated by the same independent features. Thus, species that are well separated in the biochemical network, e.g. activated caspase 6 and cytoplasmic Bcl-2, are nevertheless strongly correlated due to the constraints imposed by steady state.

Using the correlation between each feature with the response, and the cross-correlation between every pair of features, we sought to identify a subset of maximally predictive, minimally redundant features for predicting the response to TRAIL. To do this we used quadratic programming feature selection [204], or QPFS. QPFS expresses this objective as a quadratic program and finds a vector of weights on the features that minimizes redundancy while maximizing relevance (see Methods). If we examine the QPFS-weighted correlations between each feature and the response, we see a dramatic reduction in complexity (Figure 4.2.2). In particular, above the weighted correlation thresholds of 0.3 and 0.12, we observe only four and twelve features, respectively. These are the rates of synthesis and degradation of BAR and XIAP, and above 0.12 the rates of synthesis and degradation of cytoplasmic Bcl-2, mitochondrial Bcl-2, Bid, and the steady state abundances of procaspase 8 and the TRAIL receptor.

To confirm the ability of these features to predict the response of the full ODE model, we used logistic regression to model the log-odds ratio of the probability of responding to TRAIL as a linear combination of the four or twelve features. To do this we divided the 20,000 simulated responses into equal sized training and test datasets. Regression coefficients were derived by maximum likelihood estimation on the training data using the statistical software package, R.

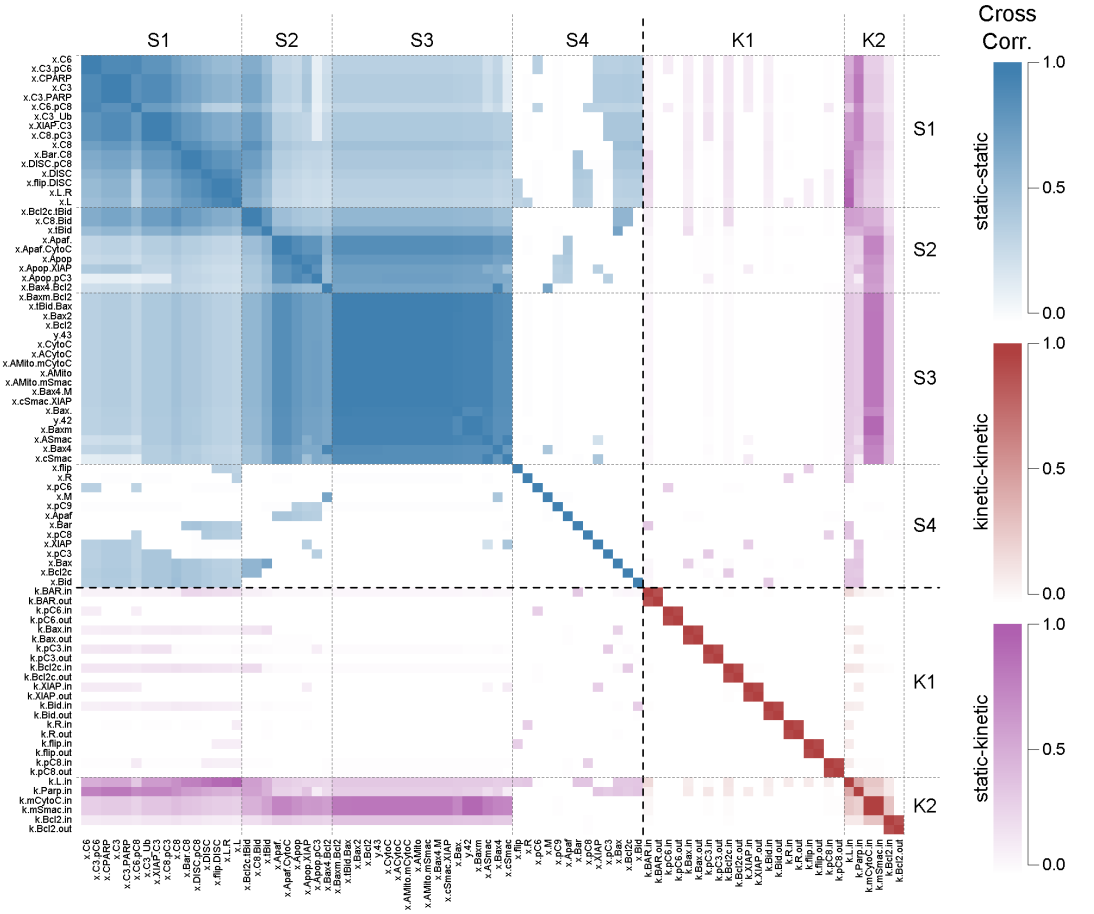


Figure 4.5: **Cross-correlation between features.** Pearson correlation coefficients were calculated between all pairs of features and plotted as a matrix. Hierarchical clustering was used to order the rows and columns, with clustering being performed separately for static versus kinetic features. Six clusters were identified after sorting, S1-S4, K1, and K2. We use hue to distinguish between cross-correlations within static features (blue), within kinetic features (red), and between static and kinetic features (violet). Color saturation is used to indicate the strength of the correlation, from weak (low saturation) to strong (high saturation).

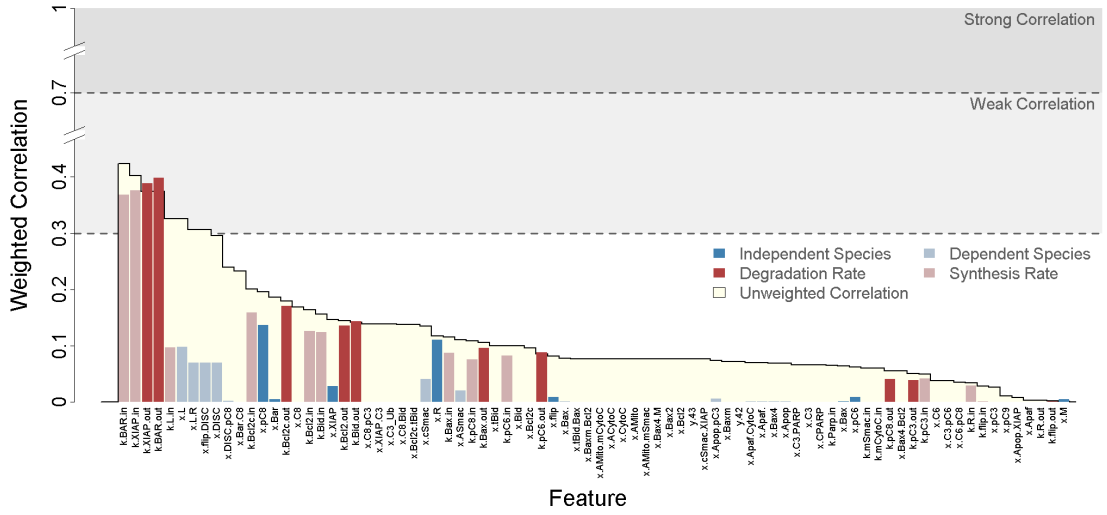


Figure 4.6: **QPFS weighted correlation between features and the binary response to TRAIL.** Features are plotted in the same order as in Figure 4.2.2, but bar height now represents the QPFS-weighted correlation between each feature and the response to TRAIL rather than the raw, unweighted correlation. For comparison, the unweighted correlations are indicated here by the yellow silhouetted region. As before, features are color-coded to indicate whether they are independent species (dark blue), dependent species (light blue), degradation rate constants (dark red), or synthesis rates (light red).

We then compared predictions from the logistic regression models with the test data and found that they achieved 78% and 84% percent accuracy, respectively (Figure 4.2.3). This indicates that, when predicting the response to TRAIL, a substantial reduction in model complexity can be achieved with only a modest loss in accuracy.

4.2.3 Kinetic features outperform static features as predictors of the response to TRAIL

Our results strongly suggest that kinetic features are more powerful predictors of TRAIL-responsiveness than static features. Indeed, all seven of the best features are kinetic. Since only 33

An implicit parameter in the calculation of model accuracy is the threshold at which an unknown system is classified as responsive. By convention this threshold is zero, that being when the predicted probability of responding equals the probability of not responding. By altering this threshold, the fraction of true to false positives can be adjusted, yielding the well-known receiver operating characteristic (ROC) curve. To test whether kinetic regression models simply outperform static models at a particular classification threshold, we generated ROC curves for all 80 logistic regression models. Again, the results confirm that kinetic regression models are far more discriminatory than static regression models. As with the previous result, a kinetic regression model containing only the rates of synthesis and degradation of XIAP outperforms the static regression model incorporating all 54 steady state species abundances (Figure 4.3B).

Finally, it was recently shown that determinants of TRAIL-sensitivity are context dependent (Gaudet et al, 2012). For example, in the context of Bcl-2 overexpression, the steady state abundance of Bax is the primary regulator of TRAIL-responsiveness. With respect to our own observations, we wondered whether the dominance of kinetic features was an artifact of the short protein half-lives used in our model. To test this, we repeated our analysis using the slower half-lives reported in (Gaudet et al, 2012). We found that in this context, kinetic features still outperform static features when predicting the response to TRAIL. However, the

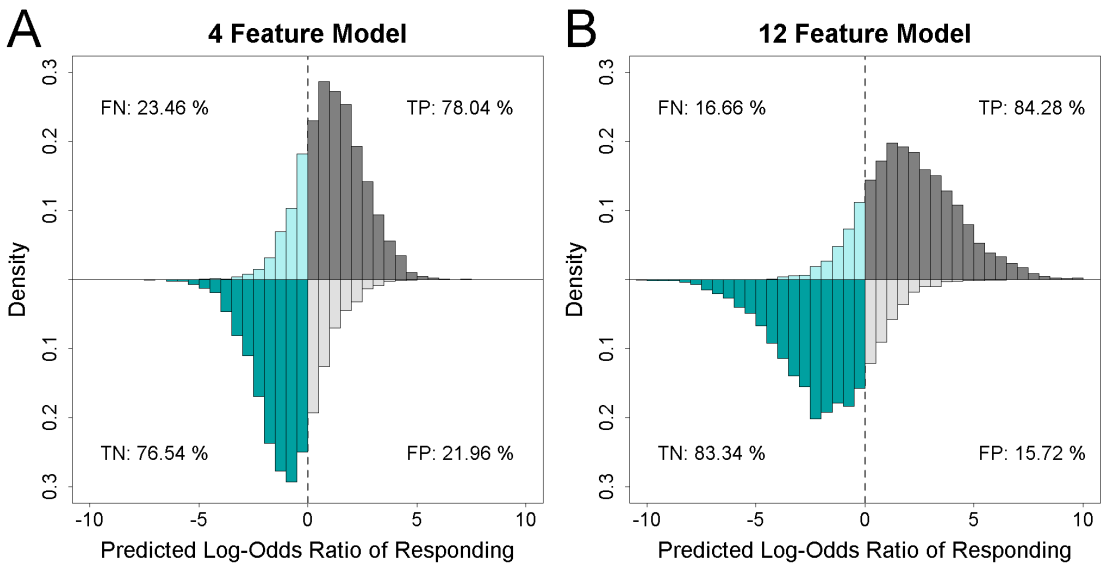


Figure 4.7: **Prediction accuracy of a 4- and 12-feature logistic regression model. A.** The four top-ranked features as identified by QPFS were used to build a logistic regression model. In decreasing order, those features are the rates of synthesis of BAR and XIAP, and the rates of degradation of XIAP and BAR. The model was trained on one half of the dataset of 20,000 simulations and tested on the other. The abscissa gives the predicted log-odds ratio of the probability of responding to TRAIL versus not responding. A value of zero corresponds to equal probability. Values greater than zero are predicted to respond to TRAIL (cells die; gray), while values less than zero are predicted not to respond (cells live; cyan). Samples in the first and second quadrant are those that actually responded to TRAIL according to the ODE model (positives). Samples in the third and fourth quadrant are those did not respond to TRAIL (negatives). **B.** As above, but the twelve top-ranked features as identified by QPFS were used instead. These include: the rates of synthesis and degradation of cytoplasmic Bcl-2, mitochondrial Bcl-2, Bid, and the steady state abundances of procaspase 8 and the TRAIL receptor.

features identified by QPFS now describe the turnover of Bax and Bcl-2 instead of Bar and XIAP. Thus while kinetic features are generally stronger predictors of TRAIL-responsiveness, feature identify is partly determined by the rate of global protein turnover.

4.3 Discussion

We have described a framework for identifying a parsimonious set of biochemical features that accurately predicts the response to chemical perturbation in a heterogeneous population of cells. This framework is made possible by the following, recent advances: 1) an experimentally validated, reaction network model of the biological process in question [71], 2) computational methods for deriving an analytical expression for the steady state of the network [192] and incorporating kinetic variability [201], 3) single cell experiments that describe the distributions of proteins in a heterogeneous population [200], and 4) a method to identify subsets of features that are simultaneously predictive and non-redundant [204].

The potential application of model reduction techniques to biomarker discovery cannot be overstated. To date, over 300 curated, quantitative models of cell decision and signaling processes have been developed [205]. Reducing these models to "representative kernels" can help quantify the functional importance of candidate markers and filter out false-positives [206]. Typical model reduction techniques either prune out inconsequential reactions [207, 208] or coarse-grain the dynamic variables using low-dimensional surrogates [209]. While both these techniques eliminate insensitive variables, the resulting model is still a system of differential equations that requires numerical integration. An alternative approach is to "emulate" the network model using a statistical meta-model [210, 211]. Since numerical integration is not required for these emulators, changes in input parameters can be rapidly mapped to changes in the response. Our method extends this approach by considering only the biologically relevant response (cell death) and incorporating physiological distributions on the input parameters. As a result, we can build an informative classifier for a heterogeneous population of cells using

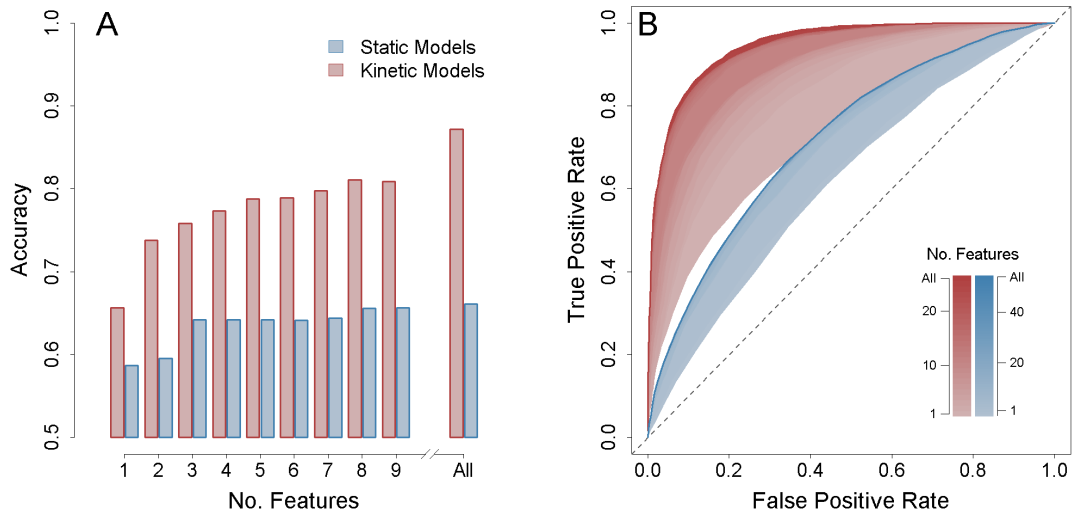


Figure 4.8: **Kinetic features outperform static features. A.** Static-only (blue) or kinetic-only (red) features were iteratively added to a logistic regression model in order of their QPFS ranking and the models trained and tested as before. The accuracy of each model as a function of its size is shown here, where "All" indicates either all 56 static features or all 26 kinetic features. An accuracy of 0.5 is equivalent to random guessing. **B.** For each of the 56 static and 26 kinetic regression models, a receiver operating characteristic curve was generated by varying the classification threshold for responders and non-responders over the entirety of its range.

only $\approx 10^3$ simulations.

After applying our framework to a model of TRAIL-induced cell death, our first conclusion is that the resting state of a signaling system can indeed serve as a good predictor of its response to perturbation. That is, a logistic regression model that considers only four steady state features will, four times out of five, accurately predict the response to TRAIL compared to the full ODE model. *A priori*, there is no reason to think that this should be possible. The dynamics of TRAIL-induced signaling effectors are highly non-linear. That the response to TRAIL can be modeled as a linear combination of just four steady state features is therefore surprising. Of course, it will be interesting to see whether this conclusion generalizes to other biological systems. One hallmark of TRAIL-induced cell death is that it does not appear to be sensitive to stochastic fluctuations [66]. In systems that are sensitive to stochastic fluctuation, we would expect this correlation

between steady state and stimulus-responsiveness to be compromised, but to what degree requires further study.

For TRAIL, we find that the rates of protein synthesis and degradation, or flux, of XIAP and Bar are the strongest predictors of the response. These are followed by the fluxes of Bid and Bcl-2 and the steady state abundances of the receptor and procaspase 8. The locality of these features in the reaction network lends credibility to this finding. Bar and XIAP are at the points of bifurcation and convergence of the mitochondrial feed forward loop, respectively. The original ODE model was trained on data derived from HeLa cells, for which MOMP is normally required for cell death [212]. The turnover of proteins that control this bifurcation are therefore good predictors of the response. It is interesting that predictive power migrates to Bax and Bcl-2 when global protein stability is increased by 60-fold. This suggests a greater role for the mitochondria in the context of slow protein turnover, and may constitute an additional determinant of Type I versus Type II behavior in these cells, in addition to the steady state ratio of caspase 3 abundance to XIAP [111].

In contrast, some features are conspicuously low in the QPFS rankings. For example, it has been observed that cells can be sensitized to TRAIL by up-regulation of the receptor [213] or down-regulating the anti-apoptotic FLIP [214]. However, neither of these features figure prominently in the QPFS rankings. Two explanations are possible. The first is that these molecules are indeed critical, but this effect is mediated through other features that themselves are better predictors of the response. For example, FLIP inhibits procaspase 8 processing. Downregulation of FLIP may therefore lead to decreased abundance of procaspase 8, which QPFS identifies as being a strong predictor of TRAIL sensitivity. The second possibility is that these features may acquire predictive power depending on the cell type and condition, as observed above and in [111, 98].

Finally, we conclude that kinetic features are stronger predictors of TRAIL-induced cell death than static features. This conclusion does not appear to be sensitive to the mean of the kinetic feature distribution. Whether the global protein-half life is modeled at 1 or 66 hours, QPFS identifies the fluxes of Bar and XIAP, or

Bax and Bcl-2 respectively, as the most powerful predictors of TRAIL responsiveness. In contrast, when the CV of the kinetic feature distribution is reduced from 0.386 to 0.25 (as reported in [98], we see a slight reduction in the dominance of kinetic features (Supplementary Figure 4.2.1). This observation is not surprising, however. In the limit where the variance of a feature is so small as to be effectively constant, knowing that feature will have no impact on classification. Since genome-wide measurements of protein half-life suggest a CV between 0.48 and 2.28 [114, 115, 62], we conclude that, in general, kinetic features dominate static features when predicting the response to TRAIL. Indeed, static features that were treated as independent random variables in the present study have been shown to be cross-correlated [98]. Incorporating these empirical cross-correlations into the QPFS rankings would further dilute their predictive power and contribute to the overall dominance of kinetic features.

In support of this conclusion, work from our laboratory shows that steady state flux can exert significant control over the dynamic response to perturbation [201]. In other words, isostatic signaling networks have a high degree of dynamic plasticity, and only when certain protein flux parameters are constrained can a particular stimulus-response behavior be observed. What is surprising, however, is that both synthesis and degradation rates of the same species are relevant and non-redundant predictors. In the case of TRAIL, the six best kinetic predictors are the synthesis and degradation of XIAP, Bar, and cytoplasmic Bcl-2. Why this might be warrants further study, but also suggests an improvement to the framework presented here. If both synthesis and degradation are good predictors, than the ratio of these two features may be a better predictor still. This suggests that simple functions of primary features should themselves be considered for ranking by QPFS, and indeed may reduce the size of the resulting regression model even further.

We envision that this work can help steer the development of next generation diagnostics. If the results with TRAIL hold for other systems, and indeed kinetic features are generally more powerful predictors than static features, then this necessitates development of sensitive assays for measuring kinetic parameters

in primary cells, perhaps via pulse-labeling with non-radioactive isotopes followed by quantitative mass spectrometry or RNA sequencing [215, 216, 62]. It also argues that valuable information is lost as soon as cells are fixed, and that diagnoses on primary human tissue samples should be performed using live cells, prior to fixation. Alternatively, it may be possible to infer the values of kinetic features from other, more easily addressable static features. For example, the rate of degradation of a particular protein may be inferable from the abundance of the E3 ubiquitin ligase that targets it for degradation. The rate of protein synthesis is largely determined by mRNA abundance [217], which itself is inferable from the state of the corresponding promoter [218]. Thus while our framework uses information about the biochemical mechanism of drug action and physiological distributions of biochemical parameters to identify a parsimonious set of predictive features, further work may be required to translate these findings into clinically feasible diagnostics.

4.4 Methods

4.4.1 Model Construction

To construct a bi-responsive model of TRAIL-induced apoptosis, the Albeck model was extended to include 15 synthesis reactions, 28 degradation reactions, and deactivation reactions for the species Mito and Apaf. An analytical expression for the steady state was derived using Maple version 14 and a py-substitution strategy that preserved all internal reaction kinetics and steady state species abundances as reported in the Albeck model [192]. The 45 new reactions required 31 additional parameters, 28 of them being degradation rate constants. For these we imposed a nominal half-life of one hour, justified by the observation that signaling proteins tend to be short lived [201]. Several species half-lives were then manually adjusted to better fit published dynamic profiles for active caspase 8, caspase 3, and cleaved Parp. They are: the Bar-caspase 8 complex, caspase 6, cleaved Parp, the TRAIL ligand, cytoplasmic Cytochrome C, and Bar. Inactivation of Mito and Apaf were assumed to be 10 times faster than protein degradation.

Heterogeneity in the steady state was achieved by letting 14 independent

steady state abundances be gamma distributed and 11 independent degradation rate constants be log normally distributed. Variance in the species abundances was set equal to one-tenth the square of the mean, i.e. the "extrinsic noise limit" observed in [200]. Mean abundances were taken from [71]. For each gamma distribution, the scale parameter was calculated from the square of the mean over the variance, and the shape parameter from the variance over the mean. Degradation rate constants were assumed to have a CV of 0.368, equivalent to a variance of 1 in the log-normal distribution of protein half-lives. 40 species and 15 kinetic parameters likewise assumed a probability distribution by virtue of being constrained to steady state (Figure 4.2.1A). Only the internal reaction kinetics, as well as degradation rates of complexes and modified species, remained constant.

4.4.2 Feature Selection and Regression

To sample the heterogeneous population described by our TRAIL model, values were chosen for each of the 25 independent, random parameters according to their prescribed probability density functions. These values were then used to calculate the 57 dependent parameters whose values were constrained by steady state. To simulate each sample's response to stimulation, we instantaneously added 1000 molecules of TRAIL and numerically integrated the system to 48 hours post-stimulation using SciLab version 5.3, using the "stiff" backward differentiation formula method.

To score the outcome of the simulation, the amount of tetrameric Bax or cleaved PARP at 48 hours was recorded and fit to a mixture of two univariate Gaussians using expectation maximization. The abundance of tetrameric Bax at 48 hours for each steady state sample was given a Z-score for each Gaussian and assigned to the population for which the Z-score was smaller. Samples that were assigned to the Gaussian with the higher mean were assigned a response variable of 1, indicating a positive response to TRAIL. Those that were assigned to the Gaussian with the lower mean were assigned a response variable of 0, indicating a negative response to TRAIL.

Once binarized, the Pearson correlation statistic was calculated between

each of the 82 features and the response. Again note that for classification, we refer to the 82 random valued parameters as features. Also, the same Pearson correlation statistic was calculated between every pair of features. The absolute values of these correlation statistics yielded a feature relevance vector F and a redundancy, or cross-correlation matrix Q, respectively. From these, a weight vector x was calculated that minimizes the quadratic program

$$x = \min\{\frac{1}{2}(1-\alpha)x^T Q x - \alpha F^T x\}$$

where the weighting factor α was calculated empirically by dividing the mean of F by the mean of Q. Further, because the matrix Q was singular, we approximated it using its eigenvalue decomposition as suggested in [204] using a cumulative eigenvalue threshold of 0.999.

Once the optimal vector of weights was identified, we used this as an ordering by which to incorporate features into a logistic regression model. Specifically, we modeled the log-odds ratio of the probability of responding to TRAIL versus not responding as a linear combination of the steady state features. Maximum likelihood estimates of the regression coefficients was achieved using the glm function in R, with the link function "logit". Model accuracy was calculated as the fraction of true positive and true negatives over all predictions. ROC curves were obtained by iterating the log-odds ratio threshold over its full range of possible values, and at each value calculating the true and false positive rates, i.e. the ratio of true to total positives, and the ratio of false to total negatives, respectively.

Chapter 4, in part, is currently being prepared for submission for publication. Loriaux, Paul M; Elkan, Charles; Hoffmann, Alexander. The dissertation author was the primary investigator and author of this material.

Chapter 5

Conclusion

In this chapter I reflect on the origins and evolution of the projects discussed in Chapters 2, 3, and 4, and discuss how the results described therein could be expanded upon.

5.1 Chapter 2

The *py*-substitution method was borne out of attempts to parameterize a comprehensive model of TNF-induced cell death via activation of NF-κB, MAPK, and caspase 8. Rather than resort to nonlinear optimization heuristics [32, 56, 219], I hoped to relate the parameters in this model to "observables" that could be measured in the laboratory. This approach is illustrated in Figure 5.1, where the parameterization of a simple two-reaction system are shown to require four observable quantities.

Figure 5.1 first appears in a presentation given during lab meeting on September 18, 2007, but itself is a refined version of a figure presented on June 5$^{\text{th}}$ of that same year. Even then the trade-offs of *py*-substitution were well-recognized. A slide from June 5$^{\text{th}}$ lists "takes a while to parameterize" and "adds inertia to model structure" as detriments of the method, but "rate constants removed from model formulation" and "steady state becomes a model input" as potential benefits. These trade-offs still hold true six years later. As shown in Chapter 2, *py*-substitution can indeed eliminate inscrutable rate constants from the set of in-

Consider the bimolecular reaction in which species A is converted to B by species C, and the opposing, unimolecular deconversion of B back into C:

$$v_1: \; A + C \rightarrow B + C$$
$$v_2: \; B \rightarrow A$$

The dynamic mass balance equations are

$$\frac{dA}{dt} = v_2 - v_1$$
$$\frac{dB}{dt} = v_1 - v_2$$

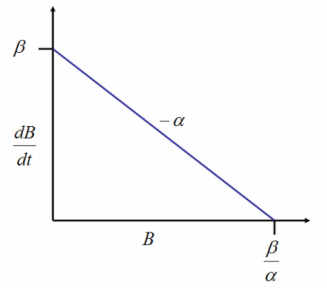
From which we may observe that the sum $A + B$ is constant. We will denote this quantity by the observable $o_1 = A + B$. Assuming simple mass action kinetics, we may write the dynamic mass balance equation for B as

$$\frac{dB}{dt} = k_1 AC - k_2 B = k_1 C(o_1 - B) - k_2 B$$
$$= k_1 o_1 C - B(k_1 C + k_2)$$

For convenience define $\alpha = k_1 C + k_2$ and $\beta = k_1 o_1 C$, then

$$\frac{dB}{dt} = \beta - \alpha B.$$

Thus the rate of change of B decreases linearly with α from $(0, \beta)$ to $(\alpha, \beta/\alpha)$:

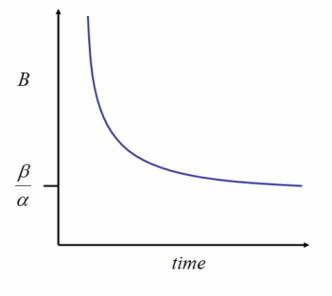


An analytical solution for B can be derived by multiplying through by $e^{\alpha t}$:

$$B(t) = \frac{\beta}{\alpha} + ce^{\alpha t}.$$

Again we see that the steady state concentration of B is given by

$$\lim_{t \to \infty} \left(\frac{\beta}{\alpha} + ce^{\alpha t} \right) = \frac{\beta}{\alpha}.$$



Let observables o_2 and o_3 be the steady state concentrations of B and C, respectively. Then

$$o_2 = \frac{\beta}{\alpha}$$

and from the original definitions of α and β:

$$k_1 o_1 o_3 = o_2(k_1 o_3 + k_2), \text{ giving}$$

$$k_2 = \frac{k_1 o_3 (o_1 - o_2)}{o_2}.$$

If we are only interested in steady state concentrations, we may choose k_1 arbitrarily and be done. To constrain the kinetics of this system, we must specify the half-life of the exponential term. Let this half-life be given by the observable o_4, then

$$\alpha = -\frac{\ln(1/2)}{o_4} = k_1 o_3 + k_2, \text{ or}$$
$$k_1 = -\frac{\ln(1/2) - k_2 o_4}{o_3 o_4}.$$

In summary, from four observable quantities:

o_1 = total concentration of $A + B$
o_2 = s.s. concentration of B
o_3 = s.s. concentration of C
o_4 = half life of exponential decay from B' to o_2

allow us to fully parameterize the system as follows:

$$A_{ss} = o_1 - o_2$$
$$B_{ss} = o_2$$
$$C_{ss} = o_3$$
$$k_1 = -\frac{o_2 \ln(1/2)}{o_1 o_3 o_4}$$
$$k_2 = -\frac{\ln(1/2)(o_1 - o_2)}{o_1 o_4}.$$

Figure 5.1: This figure illustrates the rational parameterization of a simple, reversible chemical conversion.

dependent parameters. And in Section 2.3.3, we show that using *py*-substitution, the steady state can be systematically altered to characterize its effect on the cellular response to stress. These, in addition to a reduction in the total number of parameters required, are cited in [192] as being three benefits of having an analytical steady state expression.

Even with modern symbolic solvers, however, deriving such an expression can be cumbersome. For this reason, even small changes in model structure must be made judiciously. So while *py*-substitution has now been applied to over 20 different models, there are still avenues for improvement. First, in Section 2.2 we show that a particular steady state solution is the linear combination of basis vectors spanning the null space of the coefficient matrix. Identifying these basis vectors requires a row reduction of the coefficient matrix. Because symbolic expressions do not simplify like numeric ones do (the symbolic sum of two variables "a" and "b" is a new variable "a+b" requiring twice the memory; numerically, the sum of two floating point numbers, 2.2 and 7.1, is just another floating point number, 9.3), each row operation during the reduction effectively doubles the number of terms in the reduced row. The result is "term explosion", which causes some steady state expressions to exceed physical memory (see Section 2.3.2). Because row reduction is deterministic, however, the complexity of the steady state solution could be calculated *a priori*, using information about the structure of the reaction network and the substitution strategy.

At a minimum, this calculation could be used identify impractical network architectures and substitution strategies. Or it could help identify a memory architecture that supports a full symbolic solution. An alternative, however, may be to use a hybrid approach to reduce the overall complexity of the solution. Recall from above that systems of numerical equations are far less burdensome than symbolic ones. This suggests that, after the initial mapping ψ_{py}, a subset of independent parameters in $\mathcal{P}$ could be mapped to real numbers. These numerical quantities can be combined arithmetically during row reduction, yielding a significant reduction in overall solution complexity. These real-valued parameters must be constant for all subsequent analyses, however, so this mapping must be made judiciously, on

parameters whose values are known with some confidence.

Finally, as discussed in Section 2.3.2, there are multiple but finite solutions to the steady state equation for a given model. This has ramifications where perturbations to the steady state are concerned. Consider a simple model with a single protein that is synthesized and degraded. The steady state abundance of this protein can be perturbed by adjusting its rate of synthesis, the degradation rate constant, or both. As Chapter 3 convincingly shows, the choice of perturbation will affect the model's response to stimulation. So which is correct? The answer is that, without full knowledge of what causes the change in abundance, all solutions are possible. In the work just described, dynamic effects caused by changes in steady state were conditioned by our particular choice of perturbation. Thus Figure 1.4D illustrates the effect on dynamics caused by changes in the steady state abundance of the inhibitor, Y, due to changes in its rate of synthesis. The same changes in the steady state abundance of Y due to changes to its rate of degradation may yield a different result. This suggests that a comprehensive exploration of the effects of steady state on stimulus-responsiveness should consider *all* steady state perturbations, not just specific ones. This calls for exhaustive identification of all steady state solutions, rather than the manual partition of model elements into sets $\mathcal{P}$ and $\mathcal{Y}$, as described in Chapter 2.

5.2 Chapter 3

Our manuscript, "A Protein Turnover Signaling Motif Controls the Stimulus-Sensitivy of Stress Response Pathways", followed directly from the work of a previous student in the lab, Dr. Ellen O'Dea Mercado [123, 23]. This work identified the instability of unbound IκBα as a critical mediator of canonical NF-κB sensitivity in response to stress. This prompted the question: what is the functional distinction between rapid protein turnover in a negative regulator like IκBα, versus a positive regulator like p53? Our manuscript explores this question in some detail, (see, e.g., Section 3.2.1), but falls short of any universal claims because "responsiveness" in the p53 and NF-κB pathways could not be rigorously defined.

Nevertheless, our extension of *py*-substitution to identify isostatic perturbations allowed us to demonstrate that the flux of IκBα, p53, and Mdm2 exert significant control over the sensitivities of these pathways, and may explain why the physiological rates of turnover of these molecules are so high.

For this reason, early versions of the *py*-substitution manuscript — dating back to January 2010 — include flux analyses of p53 and IκB, as these were intended to be the "methodology" and "results" sections of a single manuscript. Two observations prompted its dissolution into separate projects. First, contemporary publications from the laboratories of Drs. Scott Diamond [109], Jeremy Gunawardena [78], and Dan Beard[82], indicated that there was sufficient interest in the community for methods that identify or manipulate the steady states of mass action models. Second, our own description of *py*-substitution required so much space that it seemed to warrant a manuscript in its own right. By February 2012, a mature draft of the protein turnover manuscript was written, but it was not submitted until August 17$^{\text{th}}$, the delay being caused primarily by the *py*-substitution manuscript which we intended to co-submit as its "companion". Retrospectively, attempting to coordinate the co-submission of these two manuscripts may have been one of the greatest misjudgments of my graduate career. The motivation at the time, however, was that this protein turnover manuscript would help carry *py*-substitution into PLoS Computation Biology, which at the time, frowned on submissions whose focus was methodological (this policy has since been reversed).

There are several ways to extend the results presented in Chapter 3. The first is obviously to pursue some of the computational predictions. For example, we mentioned that mouse embryonic fibroblasts lacking IGF-1R exhibit reduced p53 synthesis and degradation, but normal protein abundance, and are insensitive to DNA damage. Our theoretical analysis suggests that pulses of p53 in response to IR require rapid turnover of p53. A simple experiment might therefore be to express GFP-tagged p53 in *igf1r*$^{-/-}$ cells, then see whether p53 oscillates in response to IR. If not, as we would predict, this may explain why these cells are insensitive to DNA damage. For IκBα, rapid turnover of the unbound monomer was shown to be mediated by a single tyrosine residue in its C-terminal PEST domain (Y289)

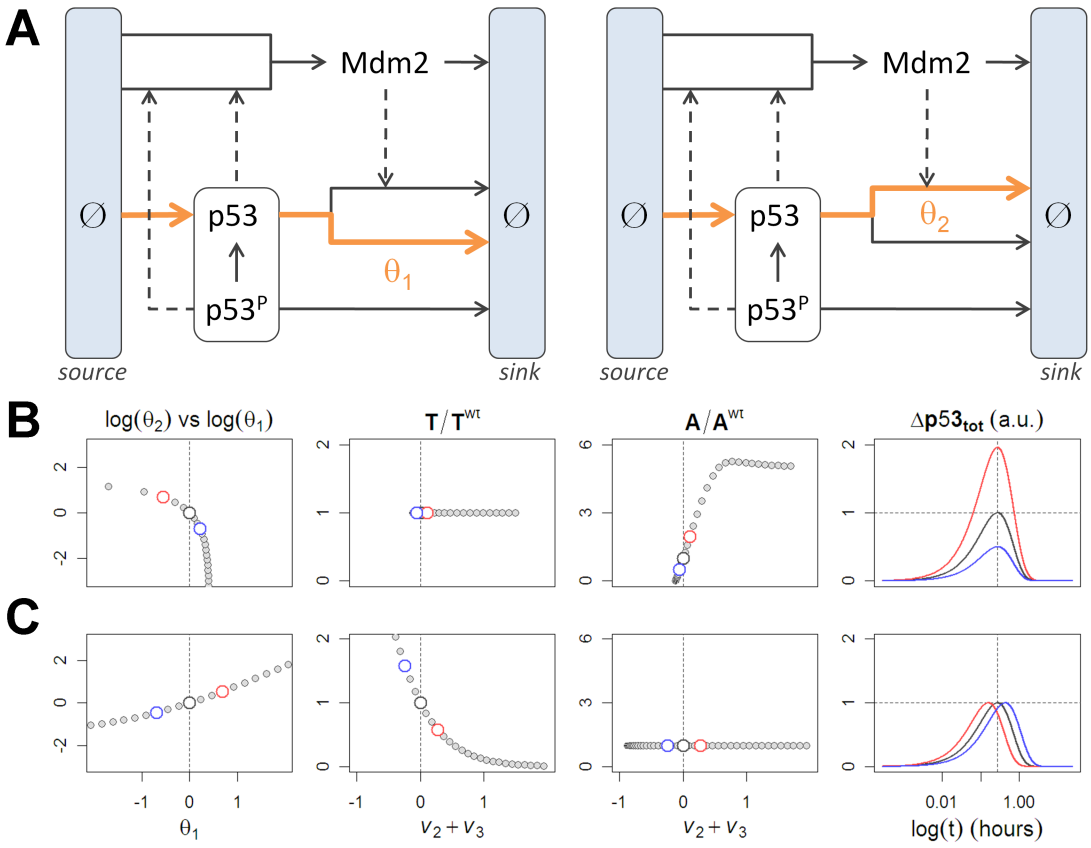


Figure 5.2: **Isostatic control of p53 activation by two p53 degradation pathways. A.** Two pathways for p53 degradation, Mdm2-independent and -dependent. The rates of p53 degradation through these two pathways are controlled by the multipliers θ_1 and θ_2, respectively. **B.** Decreasing θ_1 while increasing θ_2 results in higher amplitudes of p53 activity but with the same timing. **C.** Increasing both θ_1 and θ_2 results in earlier activation of p53, but no change in amplitude.

[40, 132]. Stable expression of a Y289F mutant IκBα in an *nfkbia*$^{-/-}$ background under a weak promoter would result in reduced IκBα flux but normal steady state abundance. Our computational analysis predicts that these cells would be hypersensitive to stimulation by UV or low doses of TNF. For molecules for which natural "flux-mutations" do not exist, we suggest in Chapter 3 that one could be engineered using a combination of tet-responsive promoters and clpXP protease recognition sequences.

A second extension to our work is theoretical. First, although Chapter 3 focuses on the control of p53 and NF-κB activation through a paired positive and negative regulatory flux motif, these are not the only two isostatic perturbations that exist in these models. Indeed, as illustrated in Figure 5.2, the dynamic response of p53 to stimulation can also be controlled by the Mdm2-dependent and -independent degradation pathways — independently, that is, of the flux of Mdm2 itself. This suggests that any number of dynamic response characteristics can be controlled by an equal number of isostatic perturbations. As mentioned in Section 3.3, the isostatic subspace for a model contains as many perturbations as there are reaction loops. Synthesis and degradation cycles are just another example of this. Reversible reactions are another. If boundaries can be assigned to the flux through these loops, then we can simulate the "dynamic plasticity" of a model around its physiological steady state. This would constitute an extension of previous work on extreme pathways [220] and steady state sampling [221] into the non-equilibrium behavior of models with nonlinear dynamics.

Finally, an admitted shortcoming of our conclusion in Chapter 3 is that we were not able to rigorously define "sensitivity" and "specificity". In response to reviewers' comments on the corresponding manuscript, an attempt to do so was made, using the prototypical negative feedback model where $\bar{x}_1 = \bar{x}_2 = 1$, and $k_1 = k_2 = k_3 = k_4 = 1$. We then examined the effects of all isostatic and anisostatic perturbations on the response to all four possible stimuli (Figure 5.2). Initial results are shown in Figure 5.3 using the area under the activation curve relative to wildtype as a dynamic response variable. Prior to each stimulus, the steady state was perturbed by reducing (red traces) or increasing (green traces) the

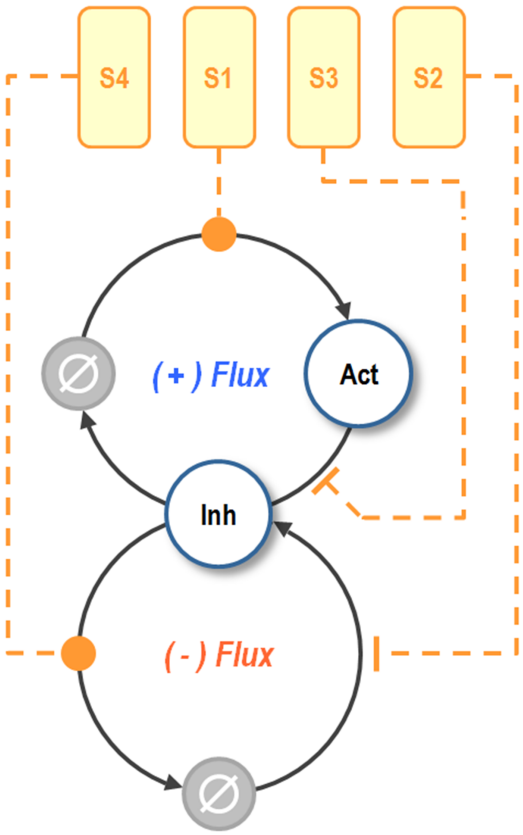


Figure 5.3: **A prototypical negative feedback model with four stimuli.** The prototypical negative feedback model shown in Figure 3.1 was expanded to include all four possible stimuli. These are: increased synthesis of the Activator (S1), decreased synthesis of the Inhibitor (S2), decreased degradation of the Activator (S3), and increased degradation of the Inhibitor (S4).

Activator flux (row 1), Inhibitor flux (row 2), Activator synthesis (row 3), Inhibitor synthesis (row 4), Activator degradation (row 5), or Inhibitor degradation (row 6). The stimulus parameter γ gives the $\log_2$ ratio of stimulus duration to stimulus intensity. The final column, "Entropy", provides a measure of how well the four stimuli can be distinguished for a particular perturbation and value of γ.

Though preliminary, some conclusions can be drawn from these early analyses. First, as γ increases (stimuli go from short and intense to long and weak), the ability to distinguish between stimuli vanishes. This suggests that stimulus specificity is a function of the non-equilibrium behavior of the network, a topic that is currently being explored in greater detail by Marcelo Behar in our laboratory (unpublished results). The second observation is that stimuli S2-S4, as well as perturbations 3-5 (Activator and Inhibitor synthesis, and Activator degradation), yield remarkably similar results. Thus, some stimuli and steady state perturbations are degenerate with respect to the stimulus-induced dynamics and have no bearing on stimulus-specificity. Finally, however, it isn't immediately obvious why evolution would select for high flux and not high flux plus increased protein abundance. Even though we show that high flux is important for stimulus sensitivity, we have yet to identify the functional constraint on protein abundance.

5.3 Chapter 4

The origins of our work with TRAIL are slightly obscure. In May, 2010, I contacted Dr. Glenn Tesler to request his help in solving the steady state equation for the original EARM model by John Albeck [71]. This led to the use of Maple over Matlab for our symbolic operations, and much of the material now discussed in Section 2.3.2. Three months later, in a presentation during lab meeting, I postulated that simulated data could be used in place of real data for the purposes of predicting the response to TRAIL. The advantage of simulated data is that it is abundant and can expose mechanistic insight. The disadvantage is that it this data is blind to mechanisms that aren't included in the model. Steady state sampling of our extended TRAIL model [192] began in February 2011, which, under

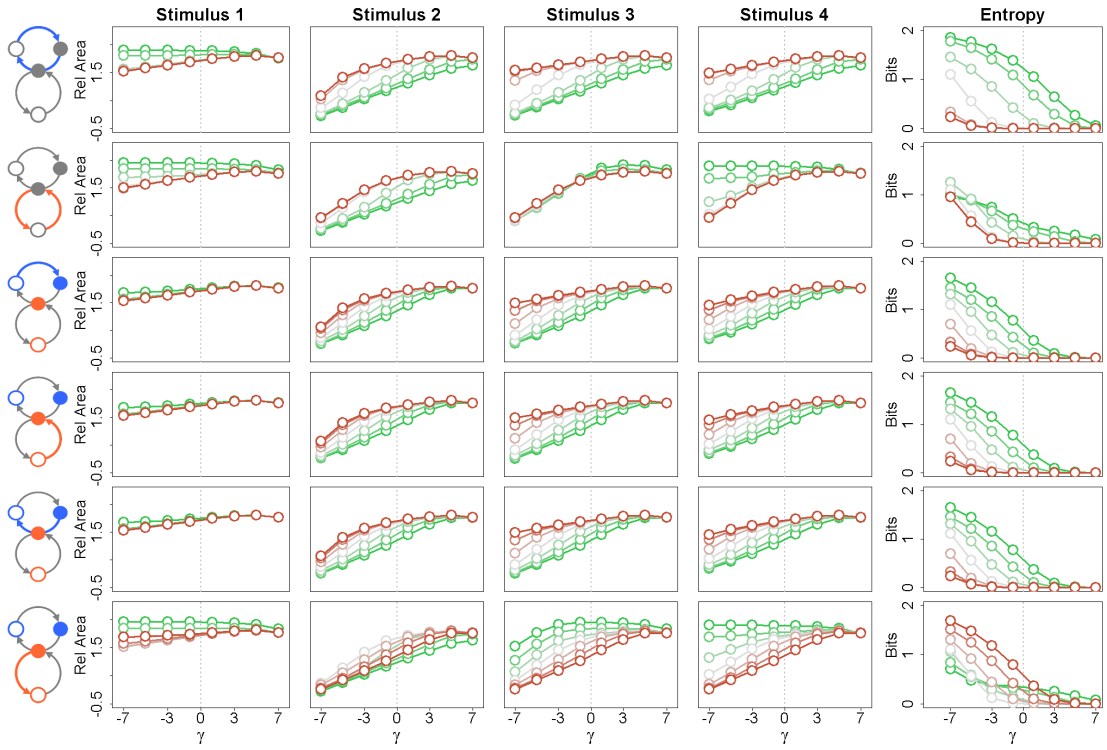


Figure 5.4: **Exhaustive analysis of steady state control of stimulus-responsiveness in a prototypical negative feedback model.** The prototypical negative feedback model shown in Figure 5.2 was stimulated by each of four possible stimuli, S1-S4. Prior to stimulation, the steady state was perturbed by reducing (red traces) or increasing (green traces) the Activator flux (row 1), Inhibitor flux (row 2), Activator synthesis (row 3), Inhibitor synthesis (row 4), Activator degradation (row 5), or Inhibitor degradation (row 6). For each stimulus, the area under the activation curve is shown ($\int_{t=0}^{\infty} \mathrm{Act}(t) - \mathrm{Act}(t=0) \quad dt$), as a function of the stimulus shape parameter γ, where the value 2^{γ} gives the ratio of stimulus duration to stimulus intensity. The Entropy column provides a measure of how well the four stimuli can be distinguished for a particular perturbation and value of γ.

the guidance of Dr. Charles Elkan, developed over the course of the following month into a masters research project in the Department of Computer Science and Engineering. In August of that year, the core of the current manuscript was presented during lab meeting, and a mature draft was completed by October, 2011. Submission of the manuscript was delayed first for possible coordination with a similar manuscript from the laboratory of Dr. Peter Sorger [98], and then in reaction to that manuscript after it was accepted to PLoS Computation Biology. A revised manuscript was submitted to Nature Molecular Systems Biology in October 2012, but was not sent out for review. Subsequent work has focused on defining the context in which our conclusions from Chapter 4 are valid.

To that end, our first thought was to vary the global mean and variance of the protein abundance and flux parameters. To do so we introduced "gross physiological parameters" to control the global protein abundance (GPA), extrinsic noise limit (XNL), global protein half-life (GPH), and half-life standard deviation (HSD). These acted as modifiers on the wildtype parameterization of our extended TRAIL model, where for a given steady state abundance $\bar{x}$ and degradation rate constant k, their modified values are given by

$$\bar{x}' \sim \Gamma(\text{GPA} \cdot \bar{x},\ 0.1 \cdot \text{XNL}) \tag{5.1}$$

$$k' \sim -\frac{1}{2}\ln N(\text{GPH},\ 0.1 \cdot \text{HSD}). \tag{5.2}$$

The parameters GPA, XNL, and HSD all took values on the interval $[2^{-3}, 2^3]$, while GPH took values on the interval $[2^0, 2^6]$, because our wildtype half-life of 1 hour was already quite short. For each value of these parameters, we sampled our model 2,000 times and generated the QPFS-weighted rankings as described in Chapter 4. Figure 5.3 shows the sum of kinetic (red) versus static (blue) weights for a given value of each of the four gross physiological parameters. The results show that kinetic features do indeed generally outperform static features, as suggested by our wildtype parameterization. This dominance is particularly pronounced for high values of global protein abundance and low values of the extrinsic noise limit, as well as high values of the half-life standard deviation. These last two results

make sense, in that, as the width of a feature distribution vanishes to zero, so does it's ability to predict the response. One particularly interesting result, however, is that the global protein half-life has no bearing on kinetic versus static dominance. The reason for this observation has not been investigated.

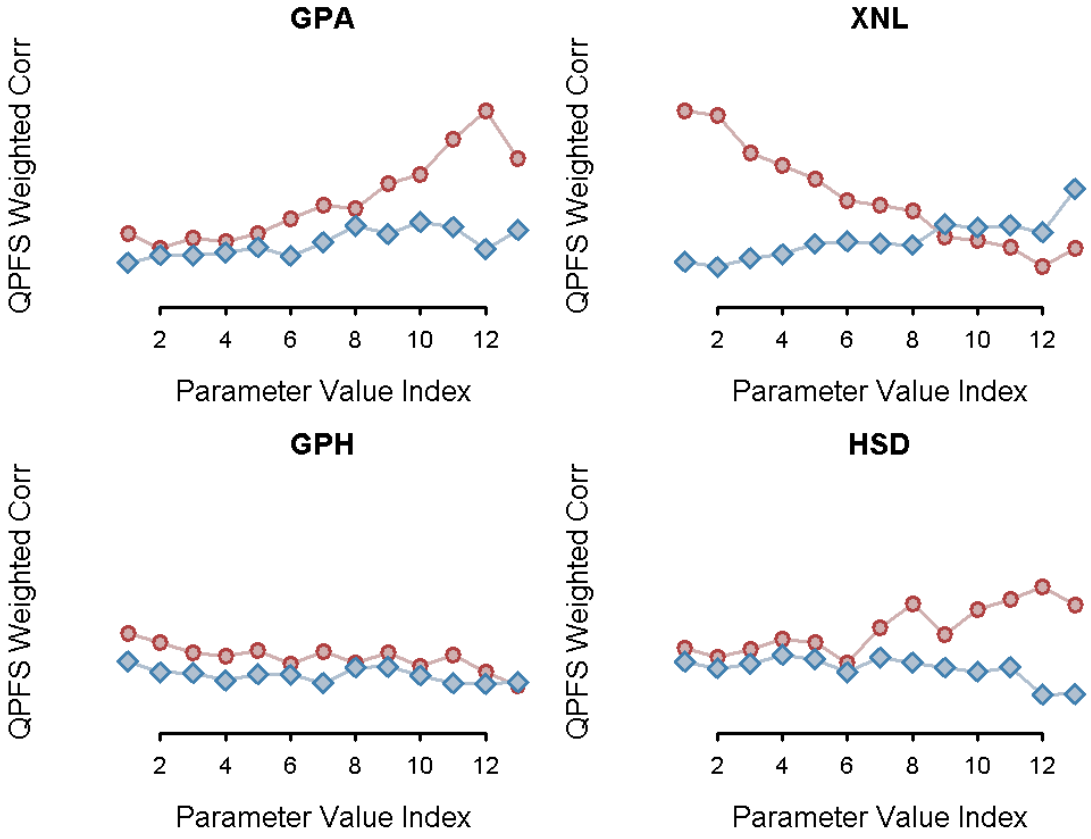


Figure 5.5: **Sensitivity of kinetic superiority in the QPFS weighted rankings to gross changes in model parameters.** The global protein abundance (GPA), extrinsic noise limit (XNL), global protein half-life (GPH), and half-life standard deviation (HSD) were varied over a 2^6-fold range and the QPFS weighted rankings of all kinetic versus all static features generated as described in Chapter 4. Equations 5.1 and 5.2 describe how these parameters affect the other model parameters.

Bibliography

[1] Feng Chen, Allyson Evans, John Pham, and Brian Plosky. Cellular stress responses: a balancing act. *Molecular Cell*, 40(2):175, October 2010.

[2] Simone Fulda, Adrienne M Gorman, Osamu Hori, and Afshin Samali. Cellular stress responses: cell survival and cell death. *International journal of cell biology*, 2010:214074, January 2010.

[3] Nicholas D Lakin and Stephen P Jackson. Regulation of p53 in response to DNA damage. *Oncogene*, 18(53):7644–55, December 1999.

[4] Heike L Pahl. Activators and target genes of Rel/NF-kappaB transcription factors. *Oncogene*, 18(49):6853–66, November 1999.

[5] Bharat B Aggarwal. Signalling pathways of the TNF superfamily: a double-edged sword. *Nature Reviews Immunology*, 3(9):745–56, September 2003.

[6] Marcelo Behar and Alexander Hoffmann. Understanding the temporal codes of intra-cellular signals. *Current Opinion in Genetics & Development*, 20(6):684–93, December 2010.

[7] Annette Schneider, Ursula Klingmüller, and Marcel Schilling. Short-term information processing, long-term responses: Insights by mathematical modeling of signal transduction. Early activation dynamics of key signaling mediators can be predictive for cell fate decisions. *BioEssays*, 34(7):542–50, July 2012.

[8] Jeremy E Purvis and Galit Lahav. Encoding and decoding cellular information through signaling dynamics. *Cell*, 152(5):945–56, February 2013.

[9] Jeremy E Purvis, Kyle W Karhohs, Caroline Mock, Eric Batchelor, Alexander Loewer, and Galit Lahav. P53 Dynamics Control Cell Fate. *Science*, 336(6087):1440–4, June 2012.

[10] Alexander Hoffmann, Andre Levchenko, Martin L Scott, and David Baltimore. The IkappaB-NF-kappaB signaling module: temporal control and selective gene activation. *Science*, 298(5596):1241–5, November 2002.

[11] Markus W Covert, Thomas H Leung, Jahlionais E Gaston, and David Baltimore. Achieving stability of lipopolysaccharide-induced NF-kappaB activation. *Science*, 309(5742):1854–7, September 2005.

[12] Shannon L Werner, Derren Barken, and Alexander Hoffmann. Stimulus specificity of gene expression programs determined by temporal control of IKK activity. *Science*, 309(5742):1857–61, September 2005.

[13] Atsushi Matsuzawa, Ping-Hui Tseng, Sivakumar Vallabhapurapu, Jun-Li Luo, Weizhou Zhang, Haopeng Wang, Dario a a Vignali, Ewen Gallagher, and Michael Karin. Essential cytoplasmic translocation of a cytokine receptor-assembled signaling complex. *Science*, 321(5889):663–8, August 2008.

[14] Jonathon M. Lee and Alan Bernstein. P53 Mutations Increase Resistance To Ionizing Radiation. *Proceedings of the National Academy of Sciences of the United States of America*, 90(12):5742–6, June 1993.

[15] Andrei V. Gudkov and Elena A. Komarova. The role of p53 in determining sensitivity to radiotherapy. *Nature Reviews Cancer*, 3(2):117–29, February 2003.

[16] Susan M Mendrysa, Matthew K McElwee, Jennifer Michalowski, Kathleen A O'Leary, Karen M Young, and Mary Ellen Perry. Mdm2 Is critical for inhibition of p53 during lymphopoiesis and the response to ionizing irradiation. *Molecular and Cellular Biology*, 23(2):462–72, January 2003.

[17] Mary Ellen Perry. Mdm2 in the response to radiation. *Cancer Research*, 2(1):9–19, January 2004.

[18] Shannon L Werner, Jeffrey D Kearns, Victoria Zadorozhnaya, Candace Lynch, Ellen O'Dea, Mark P Boldin, Averil Ma, David Baltimore, and Alexander Hoffmann. Encoding NF-kappaB temporal control in response to TNF: distinct roles for the negative regulators IkappaBalpha and A20. *Genes & Development*, 22(15):2093–101, August 2008.

[19] S Hinz, A Trauzold, L Boenicke, C Sandberg, S Beckmann, E Bayer, H Walczak, H Kalthoff, and H Ungefroren. Bcl-XL protects pancreatic adenocarcinoma cells against CD95- and TRAIL-receptor-mediated apoptosis. *Oncogene*, 19(48):5477–86, November 2000.

[20] Simone Fulda, Eric Meyer, and Klaus-Michael Debatin. Inhibition of TRAIL-induced apoptosis by Bcl-2 overexpression. *Oncogene*, 21(15):2283–94, April 2002.

[21] Lidong Zhang and Bingliang Fang. Mechanisms of resistance to TRAIL-induced apoptosis in cancer. *Cancer Gene Therapy*, 12(3):228–37, March 2005.

[22] Verena Becker, Marcel Schilling, Julie Bachmann, Ute Baumann, Andreas Raue, Thomas Maiwald, Jens Timmer, and Ursula Klingmüller. Covering a broad dynamic range: information processing at the erythropoietin receptor. *Science*, 328(5984):1404–8, June 2010.

[23] Ellen L O'Dea, Jeffrey D Kearns, and Alexander Hoffmann. UV as an amplifier rather than inducer of NF-kappaB activity. *Molecular Cell*, 30(5):632–41, June 2008.

[24] Jonathan G. Izant and Harold Weintraub. Inhibition of thymidine kinase gene expression by anti-sense RNA: a molecular approach to genetic analysis. *Cell*, 36(4):1007–15, April 1984.

[25] Andrew Fire, SiQun Xu, Mary K. Montgomery, Steven A. Kostas, Samuel E. Driver, and Craig C. Mello. Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans. *Nature*, 391(6669):806–11, February 1998.

[26] Derek W Bartlett and Mark E Davis. Insights into the kinetics of siRNA-mediated gene silencing from live-cell and live-animal bioluminescent imaging. *Nucleic Acids Research*, 34(1):322–33, January 2006.

[27] Angela Reynolds, Emily M Anderson, Annaleen Vermeulen, Yuriy Fedorov, Kathryn Robinson, Devin Leake, Jon Karpilow, William S Marshall, and Anastasia Khvorova. Induction of the interferon response by siRNA is cell type- and duplex length-dependent. *RNA (New York, N.Y.)*, 12(6):988–93, June 2006.

[28] Manfred Gossen and Hermann Bujard. Tight control of gene expression in mammalian cells by tetracycline-responsive promoters. *Proceedings of the National Academy of Sciences of the United States of America*, 89(12):5547–51, June 1992.

[29] Manfred Gossen, Sabine Freundlieb, Gabriele Bender, Gerhard Müller, Wolfgang Hillen, and Hermann Bujard. Transcriptional activation by tetracyclines in mammalian cells. *Science*, 268(5218):1766–9, June 1995.

[30] Chris Grilly, Jesse Stricker, Wyming Lee Pang, Matthew R Bennett, and Jeff Hasty. A synthetic gene network for tuning protein degradation in Saccharomyces cerevisiae. *Molecular Systems Biology*, 3(127):127, January 2007.

[31] Yannick Rondelez. Competition for Catalytic Resources Alters Biological Network Dynamics. *Physical Review Letters*, 108(1):018102, January 2012.

[32] Kwang-Hyun Cho and Olaf Wolkenhauer. Analysis and modelling of signal transduction pathways in systems biology. *Biochemical Society Transactions*, 31(Pt 6):1503–9, December 2003.

[33] Byron Goldstein, James R Faeder, and William S Hlavacek. Mathematical and computational models of immune-receptor signalling. *Nature Reviews Immunology*, 4(6):445–56, June 2004.

[34] Edda Klipp and Wolfram Liebermeister. Mathematical modeling of intracellular signaling pathways. *BMC Neuroscience*, 7 Suppl 1:S10, January 2006.

[35] Eric Batchelor, Caroline S Mock, Irun Bhan, Alexander Loewer, and Galit Lahav. Recurrent initiation: a mechanism for triggering p53 pulses in response to DNA damage. *Molecular Cell*, 30(3):277–89, May 2008.

[36] Naama Geva-Zatorsky, Nitzan Rosenfeld, Shalev Itzkovitz, Ron Milo, Alex Sigal, Erez Dekel, Talia Yarnitzky, Yuvalal Liron, Paz Polak, Galit Lahav, and Uri Alon. Oscillations and variability in the p53 system. *Molecular Systems Biology*, 2:2006.0033, January 2006.

[37] Jeffrey D Kearns, Soumen Basak, Shannon L Werner, Christine S Huang, and Alexander Hoffmann. IkappaBepsilon provides negative feedback to control NF-kappaB oscillations, signaling dynamics, and inflammatory gene expression. *The Journal of Cell Biology*, 173(5):659–64, June 2006.

[38] Raymond Cheong, Alexander Hoffmann, and Andre Levchenko. Understanding NF-kappaB signaling via mathematical modeling. *Molecular Systems Biology*, 4(192):192, January 2008.

[39] Jeffrey D Kearns and Alexander Hoffmann. Integrating computational and biochemical studies to explore mechanisms in NF-{kappa}B signaling. *The Journal of Biological Chemistry*, 284(9):5439–43, February 2009.

[40] Erika Mathes, Ellen L O'Dea, Alexander Hoffmann, and Gourisankar Ghosh. NF-kappaB dictates the degradation pathway of IkappaBalpha. *The EMBO Journal*, 27(9):1357–67, May 2008.

[41] Stefan Legewie, Nils Blüthgen, and Hanspeter Herzel. Mathematical modeling identifies inhibitors of apoptosis as mediators of positive feedback and bistability. *PLoS Computational Biology*, 2(9):e120, September 2006.

[42] John G Albeck, John M Burke, Bree B Aldridge, Mingsheng Zhang, Douglas a Lauffenburger, and Peter K Sorger. Quantitative analysis of pathways controlling extrinsic apoptosis in single cells. *Molecular Cell*, 30(1):11–25, April 2008.

[43] Jonathan R. Karr, Jayodita C. Sanghvi, Derek N. Macklin, Miriam V. Gutschow, Jared M. Jacobs, Benjamin Bolival, Nacyra Assad-Garcia, John I. Glass, and Markus W Covert. A whole-cell computational model predicts phenotype from genotype. *Cell*, 150(2):389–401, July 2012.

[44] Stuart A. Kauffman. Metabolic stability and epigenesis in randomly constructed genetic nets. *Journal of Theoretical Biology*, 22(3):437–67, March 1969.

[45] Leon Glass and Stuart A. Kauffman. The logical analysis of continuous, non-linear biochemical control networks. *Journal of Theoretical Biology*, 39(1):103–29, April 1973.

[46] Steven Watterson, Stephen Marshall, and Peter Ghazal. Logic models of pathway biology. *Drug Discovery Today*, 13(9-10):447–56, May 2008.

[47] Melody K Morris, Julio Saez-Rodriguez, Peter K Sorger, and Douglas a Lauffenburger. Logic-based models for the analysis of cell signaling networks. *Biochemistry*, 49(15):3216–24, April 2010.

[48] L.A. Zadeh. Fuzzy logic = computing with words. *IEEE Transactions on Fuzzy Systems*, 4(2):103–111, May 1996.

[49] Melody K. Morris, Julio Saez-Rodriguez, David C. Clarke, Peter K Sorger, and Douglas a. Lauffenburger. Training signaling pathway maps to biochemical data with constrained fuzzy logic: quantitative analysis of liver cell responses to inflammatory stimuli. *PLoS Computational Biology*, 7(3):e1001099, March 2011.

[50] Bree B Aldridge, Julio Saez-Rodriguez, Jeremy L Muhlich, Peter K Sorger, and Douglas a Lauffenburger. Fuzzy logic analysis of kinase pathway crosstalk in TNF/EGF/insulin-induced signaling. *PLoS Computational Biology*, 5(4):e1000340, April 2009.

[51] Bernhard ØPalsson. *Systems Biology: Properties of Reconstructed Networks*. Cambridge University Press, Cambridge, 2006.

[52] Nathan D Price, Jennifer L Reed, and Bernhard ØPalsson. Genome-scale models of microbial cells: evaluating the consequences of constraints. *Nature Reviews Microbiology*, 2(11):886–97, November 2004.

[53] Douglas McCloskey, Bernhard ØPalsson, and Adam M Feist. Basic and applied uses of genome-scale metabolic network reconstructions of Escherichia coli. *Molecular Systems Biology*, 9(661):661, April 2013.

[54] Bree B Aldridge, John M Burke, Douglas a Lauffenburger, and Peter K Sorger. Physicochemical modelling of cell signalling pathways. *Nature Cell Biology*, 8(11):1195–203, November 2006.

[55] Ramon Grima and Santiago Schnell. Modelling reaction kinetics inside cells. *Essays in Biochemistry*, 45:41–56, January 2008.

[56] Sree N Sreenath, Kwang-Hyun Cho, and Peter Wellstead. Modelling the dynamics of signalling pathways. *Essays in Biochemistry*, 45:1–28, January 2008.

[57] Michael Hucka, Andrew Finney, Herbert M. Sauro, Hamid Bolouri, John C. Doyle, Hiroaki. Kitano, Adam P. Arkin, Ben J. Bornstein, Dennis Bray, Athel. Cornish-Bowden, Autumn A. Cuellar, Sergey Dronov, Ernst Dieter Gilles, Martin Ginkel, Victoria Gor, Igor I. Goryanin, Warren J. Hedley, T. Charles Hodgman, Jan-Hendrik Hofmeyr, Peter J. Hunter, Nick S. Juty, Jay L. Kasberger, Andreas Kremling, Ursula Kummer, Nicolas Le Novere, Leslie M. Loew, Daniel Lucio, Pedro Mendes, Eric Minch, Eric D. Mjolsness, Yoichi Nakayama, Melanie R. Nelson, Poul F. Nielsen, Takeshi Sakurada, James C. Schaff, Bruce E. Shapiro, Thomas S. Shimizu, Hugh D. Spence, Jörg Stelling, Kouichi Takahashi, Masaru Tomita, John M. Wagner, and Jian Wang. The systems biology markup language (SBML): a medium for representation and exchange of biochemical network models. *Bioinformatics (Oxford, England)*, 19(4):524–531, March 2003.

[58] Benjamin J Bornstein, Sarah M Keating, Akiya Jouraku, and Michael Hucka. LibSBML: an API library for SBML. *Bioinformatics (Oxford, England)*, 24(6):880–1, March 2008.

[59] Hiromu Takizawa, Kazushige Nakamura, Akito Tabira, Yoichi Chikahara, Tatsuhiro Matsui, Noriko Hiroi, and Akira Funahashi. LibSBMLSim: a reference implementation of fully functional SBML simulator. *Bioinformatics (Oxford, England)*, 29(11):1474–1476, June 2013.

[60] Nicolas Le Novère, Benjamin Bornstein, Alexander Broicher, Mélanie Courtot, Marco Donizelli, Harish Dharuri, Lu Li, Herbert M. Sauro, Maria Schilstra, Bruce Shapiro, Jacky L Snoep, and Michael Hucka. BioModels Database: a free, centralized database of curated, published, quantitative kinetic models of biochemical and cellular systems. *Nucleic Acids Research*, 34(Database issue):D689–91, January 2006.

[61] Chen Li, Marco Donizelli, Nicolas Rodriguez, Harish Dharuri, Lukas Endler, Vijayalakshmi Chelliah, Lu Li, Enuo He, Arnaud Henry, Melanie I Stefan, Jacky L Snoep, Michael Hucka, Nicolas Le Novère, and Camille Laibe.

BioModels Database: An enhanced, curated and annotated resource for published quantitative kinetic models. *BMC Systems Biology*, 4:92, January 2010.

[62] Björn Schwanhäusser, Dorothea Busse, Na Li, Gunnar Dittmar, Johannes Schuchhardt, Jana Wolf, Wei Chen, and Matthias Selbach. Global quantification of mammalian gene expression control. *Nature*, 473(7347):337–42, May 2011.

[63] Takashi Nakakuki, Marc R Birtwistle, Yuko Saeki, Noriko Yumoto, Kaori Ide, Takeshi Nagashima, Lutz Brusch, Babatunde a Ogunnaike, Mariko Okada-Hatakeyama, and Boris N Kholodenko. Ligand-specific c-Fos expression emerges from the spatiotemporal control of ErbB network dynamics. *Cell*, 141(5):884–96, May 2010.

[64] Eric Batchelor, Alexander Loewer, Caroline Mock, and Galit Lahav. Stimulus-dependent dynamics of p53 in single cells. *Molecular Systems Biology*, 7(488):488, May 2011.

[65] Kathryn Miller-Jensen, Kevin a Janes, Yun-Ling Wong, Linda G Griffith, and Douglas a Lauffenburger. Adenoviral vector saturates Akt pro-survival signaling and blocks insulin-mediated rescue of tumor necrosis-factor-induced apoptosis. *Journal of cell science*, 119(Pt 18):3788–98, September 2006.

[66] Sabrina L Spencer, Suzanne Gaudet, John G Albeck, John M Burke, and Peter K Sorger. Non-genetic origins of cell-to-cell variability in TRAIL-induced apoptosis. *Nature*, 459(7245):428–32, May 2009.

[67] Avi Ashkenazi and Roy S Herbst. To kill a tumor cell: the potential of proapoptotic receptor agonists. *The Journal of Clinical Investigation*, 118(6):1979–90, June 2008.

[68] Devalingam Mahalingam, Corina N Oldenhuis, Eva Szegezdi, Francis J Giles, Elisabeth G E de Vries, Steven de Jong, and Steffan T Nawrocki. Targeting TRAIL towards the clinic. *Current Drug Targets*, 12(14):2079–90, December 2011.

[69] Mario Niepel, Sabrina L Spencer, and Peter K Sorger. Non-genetic cell-to-cell variability and the consequences for pharmacology. *Current Opinion in Chemical Biology*, 13(5-6):556–61, December 2009.

[70] John J Tyson, Katherine C Chen, and Bela Novak. Sniffers, buzzers, toggles and blinkers: dynamics of regulatory and signaling pathways in the cell. *Current Opinion in Cell Biology*, 15(2):221–31, April 2003.

[71] John G Albeck, John M Burke, Sabrina L Spencer, Douglas A Lauffenburger, and Peter K Sorger. Modeling a snap-action, variable-delay switch controlling extrinsic cell death. *PLoS Biology*, 6(12):2831–52, December 2008.

[72] John G Monroe. Ligand-independent tonic signaling in B-cell receptor function. *Current Opinion in Immunology*, 16(3):288–95, June 2004.

[73] Daniel J. Gough, Nicole L. Messina, Christopher J P Clarke, Ricky W. Johnstone, and David E. Levy. Constitutive type I interferon modulates homeostatic balance through tonic signaling. *Immunity*, 36(2):166–74, February 2012.

[74] Javier Macia, Sergi Regot, Tom Peeters, Núria Conde, Ricard Solé, and Francesc Posas. Dynamic signaling in the Hog1 MAPK pathway relies on high basal signal transduction. *Science Signaling*, 2(63):ra13, January 2009.

[75] Mykyta Artomov, Mehran Kardar, and Arup K Chakraborty. Only signaling modules that discriminate sharply between stimulatory and nonstimulatory inputs require basal signaling for fast cellular responses. *The Journal of chemical physics*, 133(10):105101, September 2010.

[76] Dinesh Kumar Singh, Chin-Jen Ku, Chonlarat Wichaidit, Robert J Steininger, Lani F Wu, and Steven J Altschuler. Patterns of basal signaling heterogeneity can distinguish cellular populations with different drug sensitivities. *Molecular Systems Biology*, 6(369):369, May 2010.

[77] Edward L. King and Carl Altman. A schematic method of deriving the rate laws for enzyme-catalyzed reactions. *The Journal of Physical Chemistry*, 60(10):1375–1378, 1956.

[78] Matthew Thomson and Jeremy Gunawardena. The rational parameterization theorem for multisite post-translational modification systems. *Journal of Theoretical Biology*, 261(4):626–36, December 2009.

[79] Mikhail V. Volkenstein and Boris N. Goldstein. A new method for solving the problems of the stationary kinetics of enzymological reactions. *Biochimica et Biophysica Acta (BBA) - General Subjects*, 115(2):471–477, February 1966.

[80] Sungman Cha. A simple method for derivation of rate equations for enzyme-catalyzed reactions under the rapid equilibrium assumption or combined assumptions of equilibrium and steady state. *The Journal of Biological Chemistry*, 243(4):820–5, February 1968.

[81] Chan F Lam and David G Priest. Enzyme kinetics. Systematic generation of valid King-Altman patterns. *Biophysical Journal*, 12(3):248–56, March 1972.

[82] Feng Qi, Ranjan K Dash, Yu Han, and Daniel a. Beard. Generating rate equations for complex enzyme systems by a computer-assisted systematic method. *BMC Bioinformatics*, 10:238, January 2009.

[83] Mercedes Pérez Millán, Alicia Dickenstein, Anne Shiu, and Carsten Conradi. Chemical reaction systems with toric steady states. *Bulletin of Mathematical Biology*, 74(5):1027–65, May 2012.

[84] Elisenda Feliu, Michael Knudsen, Lars N Andersen, and Carsten Wiuf. An algebraic approach to signaling cascades with N layers. *Bulletin of Mathematical Biology*, 74(1):45–72, January 2012.

[85] Elisenda Feliu and Carsten Wiuf. Variable elimination in post-translational modification reaction networks with mass-action kinetics. *Journal of Mathematical Biology*, February 2012.

[86] Matthew A Oberhardt, Arvind K Chavali, and Jason A Papin. Flux balance analysis: interrogating genome-scale metabolic networks. *Methods in molecular biology (Clifton, N.J.)*, 500:61–80, January 2009.

[87] Erwin P Gianchandani, Arvind K Chavali, and Jason a Papin. The application of flux balance analysis in systems biology. *Wiley Interdisciplinary Reviews. Systems Biology and Medicine*, 2(3):372–82, 2010.

[88] Jeffrey D Orth, Ines Thiele, and Bernhard ØPalsson. What is flux balance analysis? *Nature Biotechnology*, 28(3):245–8, March 2010.

[89] Iman Famili, Radhakrishnan Mahadevan, and Bernhard ØPalsson. k-Cone analysis: determining all candidate values for kinetic parameters on a network scale. *Biophysical Journal*, 88(3):1616–25, March 2005.

[90] Neema Jamshidi and Bernhard ØPalsson. Formulating genome-scale kinetic models in the post-genome era. *Molecular Systems Biology*, 4(171):171, January 2008.

[91] Stephen J Fromm and Herbert J Fromm. A two-step computer-assisted method for deriving steady-state rate equations. *Biochemical and Biophysical Research Communications*, 265(2):448–52, November 1999.

[92] Douglas Poland. King-Altman-Hill diagram method for open systems. *The Journal of Physical Chemistry*, 93(9):3605–3612, May 1989.

[93] David Cox, John Little, and Donal OShea. *Ideals, Varieties, and Algorithms*. Undergraduate Texts in Mathematics. Springer New York, New York, NY, third edition, 2007.

[94] Arjun Kumar Manrai and Jeremy Gunawardena. The geometry of multisite phosphorylation. *Biophysical Journal*, 95(12):5533–43, December 2008.

[95] Ivan Martínez-Forero, Antonio Peláez-López, and Pablo Villoslada. Steady state detection of chemical reaction networks using a simplified analytical method. *PLoS One*, 5(6):e10823, January 2010.

[96] J. Norman Hansen, Eugene C. Dinovo, and Paul D Boyer. Initial and equilibrium 18O, 14C, 3H, and 2H exchange rates as probes of the fumarase reaction mechanism. *The Journal of Biological Chemistry*, 244(22):6270–9, November 1969.

[97] Frank Mühlenbeck, Elvira Haas, Ralph Schwenzer, Gisela Schubert, Matthias Grell, Craig Smith, Peter Scheurich, and Harald Wajant. TRAIL/Apo2L activates c-Jun NH2-terminal kinase (JNK) via caspase-dependent and caspase-independent pathways. *The Journal of Biological Chemistry*, 273(49):33091–8, December 1998.

[98] Suzanne Gaudet, Sabrina L. Spencer, William W. Chen, and Peter K Sorger. Exploring the contextual sensitivity of factors that determine cell-to-cell variability in receptor-mediated apoptosis. *PLoS Computational Biology*, 8(4):e1002482, January 2012.

[99] Hong Zhang, Qunli Xu, Stanislaw Krajewski, Maryla Krajewska, Zhihua Xie, Sally Fuess, Shinichi Kitada, Krzysztof Pawlowski, Adam Godzik, and John C. Reed. BAR: An apoptosis regulator at the intersection of caspases and Bcl-2 family proteins. *Proceedings of the National Academy of Sciences of the United States of America*, 97(6):2597–602, March 2000.

[100] Angelika Eggert, Michael A Grotzer, Tycho J Zuzak, Barbara R Wiewrodt, Ruth Ho, Naohiko Ikegaki, and Garrett M Brodeur. Resistance to tumor necrosis factor-related apoptosis-inducing ligand (TRAIL)-induced apoptosis in neuroblastoma cells correlates with a loss of caspase-8 expression. *Cancer Research*, 61(4):1314–9, February 2001.

[101] T M Ganten, T L Haas, J Sykora, H Stahl, M R Sprick, S C Fas, A Krueger, M a Weigand, A Grosse-Wilde, W Stremmel, P H Krammer, and H Walczak. Enhanced caspase-8 recruitment to and activation at the DISC is critical for sensitisation of human hepatocellular carcinoma cells to TRAIL-induced apoptosis by chemotherapeutic drugs. *Cell death and differentiation*, 11 Suppl 1(April):S86–96, July 2004.

[102] Monica Schliemann, Eric Bullinger, Steffen Borchers, Frank Allgower, Rolf Findeisen, and Peter Scheurich. Heterogeneity Reduces Sensitivity of Cell Death for TNF-Stimuli. *BMC Systems Biology*, 5(1):204, December 2011.

[103] Michael Knudsen, Elisenda Feliu, and Carsten Wiuf. Exact analysis of intrinsic qualitative features of phosphorelays using mathematical models. *Journal of Theoretical Biology*, 300:7–18, May 2012.

[104] Guy Shinar and Martin Feinberg. Structural sources of robustness in biochemical reaction networks. *Science*, 327(5971):1389–91, March 2010.

[105] Matthew Thomson and Jeremy Gunawardena. Unlimited multistability in multisite phosphorylation systems. *Nature*, 460(7252):274–7, July 2009.

[106] Jeremy Gunawardena. Distributivity and processivity in multisite phosphorylation can be distinguished through steady-state invariants. *Biophysical Journal*, 93(11):3828–34, December 2007.

[107] L A Segel. On the validity of the steady state assumption of enzyme kinetics. *Bulletin of Mathematical Biology*, 50(6):579–93, January 1988.

[108] J E Bailey. Complex biology with no parameters. *Nature Biotechnology*, 19(6):503–4, June 2001.

[109] Jeremy E. Purvis, Ravi Radhakrishnan, and Scott L Diamond. Steady-state kinetic modeling constrains cellular resting states and dynamic behavior. *PLoS Computational Biology*, 5(3):e1000298, March 2009.

[110] William W Chen, Birgit Schoeberl, Paul J Jasper, Mario Niepel, Ulrik B Nielsen, Douglas a Lauffenburger, and Peter K Sorger. Input-output behavior of ErbB signaling pathways as revealed by a mass action model trained against dynamic data. *Molecular Systems Biology*, 5(239):239, January 2009.

[111] Bree B Aldridge, Suzanne Gaudet, Douglas a Lauffenburger, and Peter K Sorger. Lyapunov exponents and phase diagrams reveal multi-factorial control over TRAIL-induced apoptosis. *Molecular Systems Biology*, 7(553):553, January 2011.

[112] C Y Huang and James E Ferrell. Ultrasensitivity in the mitogen-activated protein kinase cascade. *Proceedings of the National Academy of Sciences of the United States of America*, 93(19):10078–83, September 1996.

[113] Michael F Princiotta, Diana Finzi, Shu-Bing Qian, James Gibbs, Sebastian Schuchmann, Frank Buttgereit, Jack R Bennink, and Jonathan W Yewdell. Quantitating protein synthesis, degradation, and endogenous antigen processing. *Immunity*, 18(3):343–54, March 2003.

[114] François-Michel Boisvert, Yasmeen Ahmad, Marek Gierlinski, Fabien Charrière, Douglas Lamont, Michelle Scott, Geoff Barton, and Angus I Lamond. A quantitative spatial proteomics analysis of proteome turnover in human cells. *Molecular & cellular proteomics : MCP*, 44(0), December 2011.

[115] Sidney B Cambridge, Florian Gnad, Chuong Nguyen, Justo Lorenzo Bermejo, Marcus Krüger, and Matthias Mann. Systems-wide proteomic analysis in mammalian cells reveals conserved, functional protein turnover. *Journal of proteome research*, 10(12):5275–84, December 2011.

[116] Jeffrey N Savas, Brandon H Toyama, Tao Xu, John R Yates, and Martin W Hetzer. Extremely long-lived nuclear pore proteins in the rat brain. *Science*, 335(6071):942, March 2012.

[117] Hsueh-Chi Sherry Yen, Qikai Xu, Danny M Chou, Zhenming Zhao, and Stephen J Elledge. Global protein stability profiling in mammalian cells. *Science*, 322(5903):918–23, November 2008.

[118] John C Price, Shenheng Guan, Alma Burlingame, Stanley B Prusiner, and Sina Ghaemmaghami. Analysis of proteome dynamics in the mouse brain. *Proceedings of the National Academy of Sciences of the United States of America*, 107(32):14508–13, August 2010.

[119] Aaron Ciechanover and Alan L Schwartz. Ubiquitin-mediated degradation of cellular proteins in health and disease. *Hepatology (Baltimore, Md.)*, 35(1):3–6, January 2002.

[120] Keiichi I Nakayama and Keiko Nakayama. Ubiquitin ligases: cell-cycle control and cancer. *Nature Reviews Cancer*, 6(5):369–81, May 2006.

[121] Warren Maltzman and Linda Czyzyk. UV irradiation stimulates levels of p53 cellular tumor antigen in nontransformed mouse cells. *Molecular and Cellular Biology*, 4(9):1689–94, September 1984.

[122] E Reihsaus, M Kohler, S Kraiss, M Oren, and M Montenarh. Regulation of the level of the oncoprotein p53 in non-transformed and transformed cells. *Oncogene*, 5(1):137–45, January 1990.

[123] Ellen L O'Dea, Derren Barken, Raechel Q Peralta, Kim T Tran, Shannon L Werner, Jeffrey D Kearns, Andre Levchenko, and Alexander Hoffmann. A homeostatic model of IkappaB metabolism to control constitutive NF-kappaB activity. *Molecular Systems Biology*, 3(111):111, January 2007.

[124] R Lev Bar-Or, R Maya, L a Segel, Uri Alon, a J Levine, and M Oren. Generation of oscillations by the p53-Mdm2 feedback loop: a theoretical and experimental study. *Proceedings of the National Academy of Sciences of the United States of America*, 97(21):11250–5, October 2000.

[125] Galit Lahav, Nitzan Rosenfeld, Alex Sigal, Naama Geva-Zatorsky, Arnold J Levine, Michael B Elowitz, and Uri Alon. Dynamics of the p53-Mdm2 feedback loop in individual cells. *Nature Genetics*, 36(2):147–50, February 2004.

[126] Eric Batchelor, Alexander Loewer, and Galit Lahav. The ups and downs of p53: understanding protein dynamics in single cells. *Nature Reviews Cancer*, 9(5):371–7, May 2009.

[127] Kazunari Iwamoto, Hiroyuki Hamada, and Masahiro Okamoto. Mechanism of cell cycle disruption by multiple p53 pulses. *Genome informatics. International Conference on Genome Informatics*, 25(1):12–24, January 2011.

[128] Xiao-Peng Zhang, Feng Liu, and Wei Wang. Two-phase dynamics of p53 in the DNA damage response. *Proceedings of the National Academy of Sciences of the United States of America*, 108(22):8990–5, May 2011.

[129] Alexander Hoffmann and David Baltimore. Circuitry of nuclear factor kappaB signaling. *Immunological reviews*, 210:171–86, April 2006.

[130] Matthew S Hayden and Sankar Ghosh. Shared principles in NF-kappaB signaling. *Cell*, 132(3):344–62, February 2008.

[131] Ellen L O'Dea and Alexander Hoffmann. NF-κB signaling. *Wiley Interdisciplinary Reviews. Systems Biology and Medicine*, 1(1):107–15, 2009.

[132] Erika Mathes, Lily Wang, Elizabeth Komives, and Gourisankar Ghosh. Flexible regions within I{kappa}B{alpha} create the ubiquitin-independent degradation signal. *The Journal of biological chemistry*, 285(43):32927–36, October 2010.

[133] Lan Ma, John Wagner, John Jeremy Rice, Wenwei Hu, Arnold J Levine, and Gustavo a Stolovitzky. A plausible model for the digital response of p53 to DNA damage. *Proceedings of the National Academy of Sciences of the United States of America*, 102(40):14266–71, October 2005.

[134] Jayne M Stommel and Geoffrey M Wahl. Accelerated MDM2 autodegradation induced by DNA-damage kinases is required for p53 activation. *The EMBO Journal*, 23(7):1547–56, April 2004.

[135] Arvin B Tam, Ellen L Mercado, Alexander Hoffmann, and Maho Niwa. ER Stress Activates NF-κB by Integrating Functions of Basal IKK Activity, IRE1 and PERK. *PLoS One*, 7(10):e45078, January 2012.

[136] Jijie Gu, Hidehiko Kawai, Linghu Nie, Hiroyuki Kitao, Dmitri Wiederschain, Aart G Jochemsen, John Parant, Guillermina Lozano, and Zhi-Min Yuan. Mutual dependence of MDM2 and MDMX in their functional inactivation of p53. *The Journal of Biological Chemistry*, 277(22):19251–4, May 2002.

[137] Lei Xiong, Fei Kou, Ying Yang, and Jiarui Wu. A novel role for IGF-1R in p53-mediated apoptosis through translational modulation of the p53-Mdm2

feedback loop. *The Journal of Cell Biology*, 178(6):995–1007, September 2007.

[138] Nanda R Rodrigues, Andrew Rowan, M E Smith, I B Kerr, Walter F Bodmer, Julian V Gannon, and David P Lane. p53 mutations in colorectal cancer. *Proceedings of the National Academy of Sciences of the United States of America*, 87(19):7555–9, October 1990.

[139] C A Midgley and D P Lane. p53 protein stability in tumour cells is not determined by mutation but is dependent on Mdm2 binding. *Oncogene*, 15(10):1179–89, September 1997.

[140] Seiichi Nakamura, Jack A Roth, and Tapas Mukhopadhyay. Multiple lysine mutations in the C-terminal domain of p53 interfere with MDM2-dependent protein degradation and ubiquitination. *Molecular and Cellular Biology*, 20(24):9391–8, December 2000.

[141] S E Kern, K W Kinzler, S J Baker, J M Nigro, V Rotter, A J Levine, P Friedman, C Prives, and B Vogelstein. Mutant p53 proteins bind DNA abnormally in vitro. *Oncogene*, 6(1):131–6, January 1991.

[142] S E Kern, J A Pietenpol, S Thiagalingam, A Seymour, K W Kinzler, and B Vogelstein. Oncogenic forms of p53 inhibit p53-regulated gene expression. *Science*, 256(5058):827–30, May 1992.

[143] N P Pavletich, K a Chambers, and C O Pabo. The DNA-binding domain of p53 contains the four conserved regions and the major mutation hot spots. *Genes & Development*, 7(12B):2556–64, December 1993.

[144] Y Cho, S Gorina, P D Jeffrey, and N P Pavletich. Crystal structure of a p53 tumor suppressor-DNA complex: understanding tumorigenic mutations. *Science*, 265(5170):346–55, July 1994.

[145] S W Lowe, S Bodis, A McClatchey, L Remington, H E Ruley, D E Fisher, D E Housman, and T Jacks. p53 status and the efficacy of cancer therapy in vivo. *Science*, 266(5186):807–10, November 1994.

[146] R R Wallace-Brodeur and S W Lowe. Clinical implications of p53 mutations. *Cellular and Molecular Life Sciences*, 55(1):64–75, January 1999.

[147] Arie Rogel, Malka Popliker, Cynthia G Webb, and M Oren. p53 cellular tumor antigen: analysis of mRNA levels in normal adult tissues, embryos, and tumors. *Molecular and Cellular Biology*, 5(10):2851–5, October 1985.

[148] Elena A. Komarova, M V Chernov, R Franks, K Wang, G Armin, C R Zelnick, D M Chin, S S Bacus, G R Stark, and Andrei V Gudkov. Transgenic

mice with p53-responsive lacZ: p53 activity varies dramatically during normal development and determines radiation and drug sensitivity in vivo. *The EMBO Journal*, 16(6):1391–400, March 1997.

[149] Elena A. Komarova, K Christov, a I Faerman, and Andrei V Gudkov. Different impact of p53 and p21 on the radiation response of mouse tissues. *Oncogene*, 19(33):3791–8, August 2000.

[150] Yi Pan and Dale S Haines. The pathway regulating MDM2 protein degradation can be altered in human leukemic cells. *Cancer Research*, 59(9):2064–7, May 1999.

[151] S Banin, L Moyal, S Shieh, Y Taya, C W Anderson, L Chessa, N I Smorodinsky, C Prives, Y Reiss, Y Shiloh, and Y Ziv. Enhanced phosphorylation of p53 by ATM in response to DNA damage. *Science*, 281(5383):1674–7, September 1998.

[152] C E Canman, D S Lim, K A Cimprich, Y Taya, K Tamai, K Sakaguchi, E Appella, M B Kastan, and J D Siliciano. Activation of the ATM kinase by ionizing radiation and phosphorylation of p53. *Science*, 281(5383):1677–9, September 1998.

[153] Christopher J Bakkenist and Michael B Kastan. DNA damage activates ATM through intermolecular autophosphorylation and dimer dissociation. *Nature*, 421(6922):499–506, January 2003.

[154] Giacomo Buscemi, Paola Perego, Nives Carenini, Makoto Nakanishi, Luciana Chessa, Junjie Chen, Kumkum Khanna, and Domenico Delia. Activation of ATM and Chk2 kinases in relation to the amount of DNA strand breaks. *Oncogene*, 23(46):7691–700, October 2004.

[155] Tae Lee, Guang Yao, Joseph Nevins, and Lingchong You. Sensing and integration of Erk and PI3K signals by Myc. *PLoS Computational Biology*, 4(2):e1000013, February 2008.

[156] T W McKeithan. Kinetic proofreading in T-cell receptor signal transduction. *Proceedings of the National Academy of Sciences of the United States of America*, 92(11):5042–6, May 1995.

[157] Eli S Groban, Elizabeth J Clarke, Howard M Salis, Susan M Miller, and Christopher a Voigt. Kinetic buffering of cross talk between bacterial two-component sensors. *Journal of molecular biology*, 390(3):380–93, July 2009.

[158] G. Cattoretti, F. Rilke, S. Andreola, L. D'Amato, and D. Delia. P53 expression in breast cancer. *International journal of cancer. Journal international du cancer*, 41(2):178–83, February 1988.

[159] J L Ostrowski, A Sawan, L Henry, C Wright, J a Henry, C Hennessy, T J Lennard, B Angus, and C H Horne. P53 Expression in Human Breast Cancer Related To Survival and Prognostic Factors: an Immunohistochemical Study. *The Journal of pathology*, 164(1):75–81, May 1991.

[160] Jorma Isola, Tapio Visakorpi, Kaija Holli, and O P Kallioniemi. Association of overexpression of tumor suppressor protein p53 with rapid cell proliferation and poor prognosis in node-negative breast cancer patients. *Journal of the National Cancer Institute*, 84(14):1109–14, July 1992.

[161] Hilary M. Martin, M. Isabel Filipe, Richard W. Morris, David P. Lane, and Frederico Silvestre. p53 expression and prognosis in gastric carcinoma. *International journal of cancer. Journal international du cancer*, 50(6):859–62, April 1992.

[162] Dennis C Quinlan, Ann G Davidson, Carol L Summers, Herbert E Warden, and Himanshu M Doshi. Accumulation of p53 protein correlates with a poor prognosis in human lung cancer. *Cancer Research*, 52(17):4828–31, September 1992.

[163] a Yamaguchi, Y Kurosaka, S Fushida, M Kanno, Y Yonemura, K Miwa, and I Miyazaki. Expression of p53 protein in colorectal cancer and its relationship to short-term prognosis. *Cancer*, 70(12):2778–84, December 1992.

[164] Valerie Rusch, David Klimstra, Ennapadam Venkatraman, Julie Oliver, Nael Martini, Richard Gralla, Mark Kris, and Ethan Dmitrovsky. Aberrant p53 expression predicts clinical resistance to cisplatin-based chemotherapy in locally advanced non-small cell lung cancer. *Cancer Research*, 55(21):5038–42, November 1995.

[165] Makoto Sohda, Hitoshi Ishikawa, Norihiro Masuda, Hiroyuki Kato, Tatsuya Miyazaki, Masanobu Nakajima, Minoru Fukuchi, Ryokuhei Manda, Yasuyuki Fukai, Hideyuki Sakurai, and Hiroyuki Kuwano. Pretreatment evaluation of combined HIF-1alpha, p53 and p21 expression is a useful and sensitive indicator of response to radiation and chemotherapy in esophageal cancer. *International journal of cancer. Journal international du cancer*, 110(6):838–44, July 2004.

[166] L Lessard, A-M Mes-Masson, L Lamarre, L Wall, J-B Lattouf, and F Saad. NF-kappa B nuclear localization and its prognostic significance in prostate cancer. *BJU international*, 91(4):417–20, March 2003.

[167] Laurent Lessard, Pierre I Karakiewicz, Pascale Bellon-Gagnon, Mona Alam-Fahmy, Hazem a Ismail, Anne-Marie Mes-Masson, and Fred Saad. Nuclear localization of nuclear factor-kappaB p65 in primary prostate tumors is highly

predictive of pelvic lymph node metastases. *Cancer Research*, 12(19):5741–5, October 2006.

[168] D Wynford-Thomas. P53 in tumour pathology: can we trust immunocytochemistry? *The Journal of pathology*, 166(4):329–30, April 1992.

[169] Julie G Izzo, Usha Malhotra, Tsung-Teh Wu, Joe Ensor, Rajyalakshmi Luthra, Jeffrey H Lee, Stephen G Swisher, Zhongxing Liao, K S Clifford Chao, Walter N Hittelman, Bharat B Aggarwal, and Jaffer a Ajani. Association of activated transcription factor nuclear factor kappab with chemoradiation resistance and poor outcome in esophageal carcinoma. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology*, 24(5):748–54, February 2006.

[170] D S Hawkins, G William Demers, and D A Galloway. Inactivation of p53 enhances sensitivity to multiple chemotherapeutic agents. *Cancer Research*, 56(4):892–8, March 1996.

[171] Donald E Henson. Loss of p53-immunostaining intensity in breast cancer. *Journal of the National Cancer Institute*, 88(15):1015–6, August 1996.

[172] Timothy W Jacobs, J E Prioleau, Isaac E Stillman, and Stuart J Schnitt. Loss of tumor marker-immunostaining intensity on stored paraffin slides of breast cancer. *Journal of the National Cancer Institute*, 88(15):1054–9, August 1996.

[173] H Eid, M Van der Looij, E Institoris, L Géczi, I Bodrogi, E Oláh, and M Bak. Is p53 expression, detected by immunohistochemistry, an important parameter of response to treatment in testis cancer? *Anticancer research*, 17(4A):2663–9, 1997.

[174] Ron Milo, S Shen-Orr, S Itzkovitz, N Kashtan, D Chklovskii, and Uri Alon. Network motifs: simple building blocks of complex networks. *Science*, 298(5594):824–7, October 2002.

[175] Uri Alon. Network motifs: theory and experimental approaches. *Nature Reviews Genetics*, 8(6):450–61, June 2007.

[176] N MacDonald. Time lag in a model of a biochemical reaction sequence with end product inhibition. *Journal of Theoretical Biology*, 67(3):549–56, August 1977.

[177] William Mather, Matthew R Bennett, Jeff Hasty, and Lev S Tsimring. Delay-induced degrade-and-fire oscillations in small genetic circuits. *Physical review letters*, 102(6):068105, February 2009.

[178] Charles Auffray, Zhu Chen, and Leroy Hood. Systems medicine: the future of medical genomics and healthcare. *Genome Medicine*, 1(1):2, January 2009.

[179] Leroy Hood and Stephen H Friend. Predictive, personalized, preventive, participatory (P4) cancer medicine. *Nature Reviews Clinical Oncology*, 8(3):184–7, March 2011.

[180] Edward H Romond, Edith A Perez, John Bryant, Vera J Suman, Charles E Geyer, Nancy E Davidson, Elizabeth Tan-Chiu, Silvana Martino, Soonmyung Paik, Peter A Kaufman, Sandra M Swain, Thomas M Pisansky, Louis Fehrenbacher, Leila A Kutteh, Victor G Vogel, Daniel W Visscher, Greg Yothers, Robert B Jenkins, Ann M Brown, Shaker R Dakhil, Eleftherios P Mamounas, Wilma L Lingle, Pamela M Klein, James N Ingle, and Norman Wolmark. Trastuzumab plus adjuvant chemotherapy for operable HER2-positive breast cancer. *The New England journal of medicine*, 353(16):1673–84, October 2005.

[181] Martine J Piccart-Gebhart, Marion Procter, Brian Leyland-Jones, Aron Goldhirsch, Michael Untch, Ian Smith, Luca Gianni, Jose Baselga, Richard Bell, Christian Jackisch, David Cameron, Mitch Dowsett, Carlos H Barrios, Günther Steger, Chiun-Shen Huang, Michael Andersson, Moshe Inbar, Mikhail Lichinitser, István Láng, Ulrike Nitz, Hiroji Iwata, Christoph Thomssen, Caroline Lohrisch, Thomas M Suter, Josef Rüschoff, Tamás Suto, Victoria Greatorex, Carol Ward, Carolyn Straehle, Eleanor McFadden, M Stella Dolci, and Richard D Gelber. Trastuzumab after adjuvant chemotherapy in HER2-positive breast cancer. *The New England journal of medicine*, 353(16):1659–72, October 2005.

[182] Joseph a Ludwig and John N Weinstein. Biomarkers in cancer staging, prognosis and treatment selection. *Nature Reviews Cancer*, 5(11):845–56, November 2005.

[183] Eleftherios P Diamandis. Cancer biomarkers: can we turn recent failures into success? *Journal of the National Cancer Institute*, 102(19):1462–7, October 2010.

[184] George Poste. Bring on the biomarkers. *Nature*, 469(7329):156–7, January 2011.

[185] David F Ransohoff. Rules of evidence for cancer molecular-marker discovery and validation. *Nature Reviews Cancer*, 4(4):309–14, April 2004.

[186] Magdalena Chechlinska, Magdalena Kowalewska, and Radoslawa Nowak. Systemic inflammation as a confounding factor in cancer biomarker discovery and validation. *Nature Reviews Cancer*, 10(1):2–3, January 2010.

[187] H E Skipper. Kinetics of mammary tumor cell growth and implications for therapy. *Cancer*, 28(6):1479–99, December 1971.

[188] D. Shibata. Heterogeneity and Tumor History. *Science*, 336(6079):304–305, April 2012.

[189] Timothy A Yap, Marco Gerlinger, P Andrew Futreal, Lajos Pusztai, and Charles Swanton. Intratumor heterogeneity: seeing the wood for the trees. *Science translational medicine*, 4(127):127ps10, March 2012.

[190] S R Wiley, K Schooley, P J Smolak, W S Din, C P Huang, J K Nicholl, G R Sutherland, T D Smith, C Rauch, and C A Smith. Identification and characterization of a new member of the TNF family that induces apoptosis. *Immunity*, 3(6):673–82, December 1995.

[191] Klaus W Wagner, Elizabeth A Punnoose, Thomas Januario, David A Lawrence, Robert M Pitti, Kate Lancaster, Dori Lee, Melissa von Goetz, Sharon Fong Yee, Klara Totpal, Ling Huw, Viswanatham Katta, Guy Cavet, Sarah G Hymowitz, Lukas Amler, and Avi Ashkenazi. Death-receptor O-glycosylation controls tumor-cell sensitivity to the proapoptotic ligand Apo2L/TRAIL. *Nature Medicine*, 13(9):1070–7, September 2007.

[192] Paul Michael Loriaux, Glenn Tesler, and Alexander Hoffmann. Characterizing the Relationship between Steady State and Response Using Analytical Expressions for the Steady States of Mass Action Models. *PLoS Computational Biology*, 9(2):e1002901, February 2013.

[193] Alexandru Almasan and Avi Ashkenazi. Apo2L/TRAIL: apoptosis signaling, biology, and potential for cancer therapy. *Cytokine & growth factor reviews*, 14(3-4):337–48, June 2003.

[194] Shulin Wang and Wafik S El-Deiry. TRAIL and apoptosis induction by TNF-family death receptors. *Oncogene*, 22(53):8628–33, November 2003.

[195] M Woo, R Hakem, M S Soengas, G S Duncan, A Shahinian, D Kägi, A Hakem, M McCurrach, W Khoo, S a Kaufman, G Senaldi, T Howard, S W Lowe, and T W Mak. Essential contribution of caspase 3/CPP32 to apoptosis and its associated nuclear changes. *Genes & Development*, 12(6):806–19, March 1998.

[196] a G Porter and R U Jänicke. Emerging roles of caspase-3 in apoptosis. *Cell Death and Differentiation*, 6(2):99–104, February 1999.

[197] Clare Sheridan and Seamus J Martin. Commitment in apoptosis: slightly dead but mostly alive. *Trends in Cell Biology*, 18(8):353–7, August 2008.

[198] Scott H Kaufmann, Serge Desnoyers, Yvonne Ottaviano, Nancy E Davidson, and Guy G Poirier. Specific proteolytic cleavage of poly(ADP-ribose) polymerase: an early marker of chemotherapy-induced apoptosis. *Cancer Research*, 53(17):3976–85, September 1993.

[199] Paul M Loriaux and Alexander Hoffmann. *A framework for modeling the relationship between cellular steady-state and stimulus-responsiveness.*, volume 110. Elsevier Inc., January 2012.

[200] Yuichi Taniguchi, Paul J Choi, Gene-Wei Li, Huiyi Chen, Mohan Babu, Jeremy Hearn, Andrew Emili, and X. Sunney Xie. Quantifying E. coli proteome and transcriptome with single-molecule sensitivity in single cells. *Science*, 329(5991):533–8, July 2010.

[201] Paul Michael Loriaux and Alexander Hoffmann. A protein turnover signaling motif controls the stimulus-sensitivity of stress response pathways. *PLoS Computational Biology*, 9(2):e1002932, February 2013.

[202] Nir Friedman, Long Cai, and X Sunney Xie. Linking stochastic dynamics to population distribution: an analytical framework of gene expression. *Physical review letters*, 97(16):168302, October 2006.

[203] H Düssmann, M Rehm, C G Concannon, S Anguissola, M Würstle, S Kacmar, P Völler, H J Huber, and J H M Prehn. Single-cell quantification of Bax activation and mathematical modelling suggest pore formation on minimal mitochondrial Bax accumulation. *Cell Death and Differentiation*, 17(2):278–90, February 2010.

[204] I. Rodriguez-Lujan, R. Huerta, C. Elkan, and C. Santa Cruz. Quadratic Programming Feature Selection. *Journal of Machine Learning Research*, 11:1491–1516, 2010.

[205] Nicolas Le Novère, Benjamin Bornstein, Alexander Broicher, Mélanie Courtot, Marco Donizelli, Harish Dharuri, Lu Li, Herbert M. Sauro, Maria Schilstra, Bruce Shapiro, Jacky L Snoep, and Michael Hucka. BioModels Database: a free, centralized database of curated, published, quantitative kinetic models of biochemical and cellular systems. *Nucleic Acids Research*, 34(Database issue):D689–91, January 2006.

[206] Jeong-Rae Kim, Junil Kim, Yung-Keun Kwon, Hwang-Yeol Lee, Pat Heslop-Harrison, and Kwang-Hyun Cho. Reduction of complex signaling networks to a representative kernel. *Science Signaling*, 4(175):ra35, January 2011.

[207] Zuyi (Jacky) Huang, Yunfei Chu, and Juergen Hahn. Model simplification procedure for signal transduction pathway models: An application to IL-6 signaling. *Chemical Engineering Science*, 65(6):1964–1975, March 2010.

[208] Tom Quaiser, Anna Dittrich, Fred Schaper, and Martin Mönnigmann. A simple work flow for biologically inspired model reduction–application to early JAK-STAT signaling. *BMC Systems Biology*, 5(1):30, January 2011.

[209] A Dokoumetzidis and L Aarons. Proper lumping in systems biology models. *IET systems biology*, 3(1):40–51, January 2009.

[210] Stefano Conti and Anthony OHagan. Bayesian emulation of complex multi-output and dynamic computer models. *Journal of Statistical Planning and Inference*, 140(3):640–651, March 2010.

[211] Kristin Tø ndel, Ulf G Indahl, Arne B Gjuvsland, Jon Olav Vik, Peter Hunter, Stig W Omholt, and Harald Martens. Hierarchical cluster-based partial least squares regression (HC-PLSR) is an efficient tool for metamodelling of nonlinear dynamic models. *BMC Systems Biology*, 5(1):90, January 2011.

[212] M Mandal, S B Maggirwar, N Sharma, S H Kaufmann, S C Sun, and R Kumar. Bcl-2 prevents CD95 (Fas/APO-1)-induced degradation of lamin B and poly(ADP-ribose) polymerase and restores the NF-kappaB signaling pathway. *The Journal of Biological Chemistry*, 271(48):30354–9, November 1996.

[213] Nathan G Dolloff, Patrick A Mayes, Lori S Hart, David T Dicker, Robin Humphreys, and Wafik S El-Deiry. Off-target lapatinib activity sensitizes colon cancer cells through TRAIL death receptor up-regulation. *Science translational medicine*, 3(86):86ra50, June 2011.

[214] Sarah Shirley and Olivier Micheau. Targeting c-FLIP in cancer. *Cancer Letters*, pages 1–10, November 2010.

[215] Mary K Doherty, Dean E Hammond, Michael J Clague, Simon J Gaskell, and Robert J Beynon. Turnover of the human proteome: determination of protein intracellular stability by dynamic SILAC. *Journal of proteome research*, 8(1):104–12, January 2009.

[216] Caroline C Friedel, Lars Dölken, Zsolt Ruzsics, Ulrich H Koszinowski, and Ralf Zimmer. Conserved principles of mammalian transcriptional regulation revealed by RNA half-life. *Nucleic Acids Research*, 37(17):e115, September 2009.

[217] Nicholas T Ingolia, Sina Ghaemmaghami, John R S Newman, and Jonathan S Weissman. Genome-wide analysis in vivo of translation with nucleotide resolution using ribosome profiling. *Science*, 324(5924):218–23, April 2009.

[218] Sylvia C Tippmann, Robert Ivanek, Dimos Gaidatzis, Anne Schöler, Leslie Hoerner, Erik van Nimwegen, Peter F Stadler, Michael B Stadler, and Dirk Schübeler. Chromatin measurements reveal contributions of synthesis and decay to steady-state mRNA levels. *Molecular Systems Biology*, 8(593):593, January 2012.

[219] Kyoung Ae Kim, Sabrina L Spencer, John G Albeck, John M Burke, Peter K Sorger, Suzanne Gaudet, and Do Hyun Kim. Systematic calibration of a cell signaling network model. *BMC bioinformatics*, 11:202, January 2010.

[220] C H Schilling, D Letscher, and B O Palsson. Theory for the systemic definition of metabolic pathways and their use in interpreting metabolic function from a pathway-oriented perspective. *Journal of Theoretical Biology*, 203(3):229–48, April 2000.

[221] Jan Schellenberger and Bernhard ØPalsson. Use of randomized sampling for analysis of metabolic networks. *The Journal of biological chemistry*, 284(9):5457–61, February 2009.