Spring 5-17-2010

Interplay of Extrinsic and Intrinsic Cues in Cell-Fate Decisions

Santhosh Palani
University of Pennsylvania, santhosh.palani@gmail.com

Interplay of Extrinsic and Intrinsic Cues in Cell-Fate Decisions

Abstract
A cell's decision making process is coordinated by dynamic interplay between its extracellular environment and its intracellular milieu. For example, during stem cell differentiation, fate decisions are believed to be ultimately controlled by differential expression of lineage-specific transcription factors, but cytokine receptor signals also play a crucial instructive role in addition to providing permissive proliferation and survival cues.

Here, we present a minimal computational framework that integrates the intrinsic and extrinsic regulatory elements implicated in the commitment of hematopoietic progenitor cells to mature red blood cells (Chapter 2). Our model highlights the importance of bidirectional interactions between cytokine receptors and transcription factors in conferring properties such as ultrasensitivity and bistability to differentiating cells. These system-level properties can induce a switch-like characteristic during differentiation and provide robustness to the mature state. We then experimentally test predictions from this lineage commitment model in a model system for studying erythropoiesis (Chapter 3). Our experiments show that hemoglobin synthesis is highly switch-like in response to cytokine and cells undergoing lineage commitment possess memory of earlier cytokine signals. We show that erythrocyte-specific receptor and transcription factor are indeed synchronously co-upregulated and the heterogeneity in their expression is positively correlated during differentiation, confirming the presence of autofeedback and receptor-mediated positive feedback loops.

To evaluate the possibility of employing this minimal topology as a synthetic "memory module" for cell engineering applications, we constructed this topology synthetically in Saccharomyces cerevisiae by integrating Arabidopsis thaliana signaling components with an endogenous yeast pathway (Chapter 4). Our experiments show that any graded and unimodal signaling pathway can be rationally rewired to achieve our desired topology and the resulting network immediately attains high ultrasensitivity and bimodality without tweaking. We further show that this topology can be tuned to regulate system dynamics such as activation/deactivation kinetics, signal amplitude, switching threshold and sensitivity.

We conclude with a computational study to explore the generality of this interplay between extrinsic and intrinsic cues in hematopoiesis. We extend our minimal model analysis in Chapter 2 to examine the more complex fate decisions in bipotent and multipotent progenitors, particularly how these cells can make robust decisions in the presence of multiple extrinsic cues and intrinsic noise (Chapter 5). Our model provides support to both the instructive and stochastic theories of commitment: cell fates are ultimately driven by lineage-specific transcription factors, but cytokine signaling can strongly bias lineage commitment by regulating these inherently noisy cell-fate decisions with complex, pertinent behaviors such as ligand-mediated ultrasensitivity and robust multistability. The simulations further suggest that the kinetics of differentiation to a mature cell state can depend on the starting progenitor state as well as on the route of commitment that is chosen. Lastly, our model shows good agreement with lineage-specific receptor expression kinetics from microarray experiments and provides a computational framework that can integrate both classical and alternative commitment paths in hematopoiesis that have been observed experimentally.

Degree Type
Dissertation

Degree Name
Doctor of Philosophy (PhD)

Graduate Group
Bioengineering

First Advisor
Dr. Casim A. Sarkar

Keywords
Stem Cell Engineering, Synthetic Biology, Systems Biology, Cell Differentiation, Nonlinear Dynamics

Subject Categories
Biomedical Engineering and Bioengineering

INTERPLAY OF EXTRINSIC AND INTRINSIC CUES IN CELL-FATE DECISIONS

Santhosh Palani

A DISSERTATION

in

Bioengineering

Presented to the Faculties of the University of Pennsylvania in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy 2010

_____

Supervisor of Dissertation

Casim A. Sarkar, Ph.D., Assistant Professor of Bioengineering

_____

Graduate Group Chairperson

Susan S. Margulies, Ph.D., Professor of Bioengineering

Dissertation Committee

Daniel A. Hammer, Ph.D., Professor of Bioengineering (Committee Chair)

Mark Goulian, Ph.D., Associate Professor of Biology

Matthew J. Lazzara, Ph.D., Assistant Professor of Chemical and Biomolecular Engineering

DEDICATION

To Sid,

This work wouldn’t have been possible without his encouragement and support.

ACKNOWLEDGEMENTS

I am profoundly grateful to my advisor Casim Sarkar, whose scientific rigor, academic scholarship and timely insights are matched only by his patience, teaching and generosity. He has been the most significant influence in my scientific development and I feel truly fortunate to have him as a mentor.

I would like to thank my committee members – Dan Hammer, Mark Goulian and Matt Lazzara for reading the thesis and providing me with helpful comments and suggestions.

I feel privileged to have worked with highly gifted individuals - Pam, Ellen, Najaf, Daphne, Daina and Christine. Thanks for all your help and support.

Finally, I thank my family – Appa, Amma, Sekar, Baskar and Geetha for their unconditional love.

ABSTRACT

INTERPLAY OF EXTRINSIC AND INTRINSIC CUES IN CELL-FATE DECISIONS

Santhosh Palani

Supervisor: Casim A. Sarkar, Ph.D.

A cell's decision making process is coordinated by dynamic interplay between its extracellular environment and its intracellular milieu. For example, during stem cell differentiation, fate decisions are believed to be ultimately controlled by differential expression of lineage-specific transcription factors, but cytokine receptor signals also play a crucial instructive role in addition to providing permissive proliferation and survival cues.

Here, we present a minimal computational framework that integrates the intrinsic and extrinsic regulatory elements implicated in the commitment of hematopoietic progenitor cells to mature red blood cells (Chapter 2). Our model highlights the importance of bidirectional interactions between cytokine receptors and transcription factors in conferring properties such as ultrasensitivity and bistability to differentiating cells. These system-level properties can induce a switch-like characteristic during differentiation and provide robustness to the mature state. We then experimentally test predictions from this lineage commitment model in a model system for studying erythropoiesis (Chapter 3). Our experiments show that hemoglobin synthesis is highly switch-like in response to cytokine and cells undergoing lineage commitment possess memory of earlier cytokine signals. We show that erythrocyte-specific receptor and transcription factor are indeed synchronously co-upregulated and the heterogeneity in their expression is positively

correlated during differentiation, confirming the presence of autofeedback and receptor-mediated positive feedback loops.

To evaluate the possibility of employing this minimal topology as a synthetic "memory module" for cell engineering applications, we constructed this topology synthetically in *Saccharomyces cerevisiae* by integrating *Arabidopsis thaliana* signaling components with an endogenous yeast pathway (Chapter 4). Our experiments show that any graded and unimodal signaling pathway can be rationally rewired to achieve our desired topology and the resulting network immediately attains high ultrasensitivity and bimodality without tweaking. We further show that this topology can be tuned to regulate system dynamics such as activation/deactivation kinetics, signal amplitude, switching threshold and sensitivity.

We conclude with a computational study to explore the generality of this interplay between extrinsic and intrinsic cues in hematopoiesis. We extend our minimal model analysis in Chapter 2 to examine the more complex fate decisions in bipotent and multipotent progenitors, particularly how these cells can make robust decisions in the presence of multiple extrinsic cues and intrinsic noise (Chapter 5). Our model provides support to both the instructive and stochastic theories of commitment: cell fates are ultimately driven by lineage-specific transcription factors, but cytokine signaling can strongly bias lineage commitment by regulating these inherently noisy cell-fate decisions with complex, pertinent behaviors such as ligand-mediated ultrasensitivity and robust multistability. The simulations further suggest that the kinetics of differentiation to a mature cell state can depend on the starting progenitor state as well as on the route of commitment that is chosen. Lastly, our model shows good agreement with lineage-specific receptor expression kinetics from microarray experiments and provides a computational framework that can integrate both classical and alternative commitment paths in hematopoiesis that have been observed experimentally.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1

Introduction

‘It’s high time molecular biology became quantitative, it cries out to a physicist… for modeling. Modeling isn’t a crutch, it’s the opposite; it’s a way of suggesting experiments to do, to fill gaps in our understanding.’

- John Maddox, Editor of *Nature*, 1966-73 and 1980-95.

1.1. From genes to behaviors

Biological species show great diversity and exhibit a wide array of complex behaviors. The displayed behaviors in each species, the phenotype, have been traditionally attributed to its genetic framework, the genotype. Elucidation of this correlation between genotype and phenotype is among the most actively studied problems in the biological sciences. The crux of this challenge lies in the fact that complexity in biological systems is evolved rather than designed and therefore our knowledge of the underlying framework is unlikely to be complete. With the advent of molecular biology, the genotype-phenotype problem was reduced to the understanding of the molecular details of the genetic elements and their role in exhibiting robust cellular states. The sequencing of the human genome has led us to uncover most of the molecular elements and classical genetic and biochemical approaches have successfully revealed most of the cellular states. However,

we still lack the mechanistic understanding of how molecular interactions (e.g., protein-protein or protein-DNA) regulate the processes (e.g., differentiation, proliferation, or apoptosis) involved in the achievement of cellular phenotypes. It is now well known that genetic elements are connected through large biochemical networks and their interactions tend to be highly non-linear. Technological advancement in quantifying gene expression at the single cell level has led us to identify various biochemical responses that dictate cellular phenotypes. By studying the network-response dynamics, we can potentially gain mechanistic insights into cellular behaviors.

1.2. Analytical modeling

Compared to the physical sciences, the role of mathematics in the biological sciences is far less appreciated. As explained above, non-linearity lies at the heart of biological problems and studying only the properties of individual molecules or interactions will not help us in predicting or understanding behaviors. Mathematical modeling and analysis has traditionally proven potent in understanding non-linear systems and in combination with classical experiments can provide an invaluable tool in studying biological processes[1].

In the biological community, there is a great deal of skepticism about the utility of mathematical models in solving complex biological problems. The most popular criticism among biologist is that any attempt to quantitatively model biological systems is flawed, as we have not yet identified all of the molecular players or characterized all of the existing biochemical interactions. This concern is completely valid as we saw from the

recent discovery of the previously unknown world of short RNAs[2]. However, this limitation only affects those mathematical models that focus on accurately matching experimental data by exhaustively incorporating from incomplete lists of molecules and interactions. Even though these models can be closely fitted to data, they often prove less useful in predicting newer experiments or in uncovering underlying principles if fundamental mechanisms are missing from the model description.

An alternate modeling approach, which does not suffer from the limitations of exhaustive modeling is analytical modeling or in a more ideal case, minimal modeling that targets only the most important players in a given biological process and shows how topological connections between these molecules may influence the overall property of the system studied (Figure 1.1). These models, when constructed with reasonable approximations, tend to be predictive and can identify key processes that regulate the overlying behavior.

1.3. Cell differentiation

Cell differentiation is a process in which a progenitor cell commits and morphs into a more lineage-restricted progenitor or a mature cell. This process has been intensely studied for decades by cell and developmental biologists, therefore making it an attractive dynamic system to model. As shown in Figure 1.2, studies have characterized several network topologies and responses that lead to various phenotypic behaviors[3]. With respect to cell differentiation, the most well recognized response is the “all-or-none” switch-like response and the most studied players are the lineage-specific cytokines and

transcription factors (this is particularly true in the hematopoietic system, discussed further in Chapter 2). Lineage-specific cytokines are cell-extrinsic molecules that bind to their cognate receptors on the cell surface and transmit signals through a cascade to the nucleus that promotes survival, proliferation and differentiation of the cell to a particular lineage. Lineage-specific transcription factors are cell-intrinsic, DNA-binding elements that bind to the promoter elements of the lineage-specific gene and regulate their expression. A primary focus of this dissertation is to examine how known topological connections between cell-extrinsic cues (e.g., cytokines) and cell-intrinsic factors (e.g., transcription factors) can give rise to all-or-none switch-like responses during differentiation.

1.4. All-or-none response

Cell differentiation was originally thought to be a process in which an undifferentiated cell gradually transitions into a mature state by traversing through a series of stable intermediate states. This theory originated when cell differentiation was studied with classical population averaging experiments like western blotting or RT-PCR. As seen in Figure 1.3A, when undifferentiated cells are treated with various levels of differentiation-inducing stimulus and the population average of a lineage-specific gene is measured, the experiment shows a graded response, i.e., the expression of the lineage-specific gene increased linearly with stimulus concentration. This effect can be explained by the proposed theory that each cell rests on a stable intermediate state based on the concentration of the stimuli (Figure 1.3B). However, with the advent of single cell

measurements, when the same experiment was repeated, it was noticed that the cells do not navigate through stable intermediate states, whereas the percentage of cells resting in the initial state and final differentiated state changes with the concentration of the stimuli (Figure 1.3C). According to this observation, each cell either remains in the undifferentiated state or differentiates completely to the mature state, hence the name all-or-none or bistable response[4-8].

1.5. Significance of a bistable response

There can be several physiologically significant advantages in possessing a bistable response instead of a graded response during cell differentiation. Figure 1.4 shows the steady-state expression of a lineage-specific gene to varying concentration of stimulus for a graded and a bistable response. By comparing the plots, we can see that the bistable response is less noisy and requires a threshold concentration of stimulus to exhibit a strong expression. Also, the off-state and on-state, classified based on the expression of the gene, are discrete in a bistable response. Moreover, the high-expression on-state achieved in the bistable response is much more robust to fluctuations in stimulus concentration than graded response due to the presence of hysteresis or memory (discussed further in Chapter 2).

1.6. Positive feedback loops in cell differentiation

It is not that apparent how the graded response curve arising from cytokine-receptor binding can give rise to an all-or-none response during cell differentiation. The answer may lie in the non-linear wiring of the signaling topology that transmits information into and out of the cell. Networks associated with cell differentiation that are capable of exhibiting bistable response most often require positive feedback loops for achieving non-linearity. Figure 1.5 shows examples of minimal topologies that exhibit different responses. Let us consider S as the stimulus and A as a transcription factor that activates a series of lineage-specific genes B, C, and D. If S activates A linearly and A correspondingly activates B, C and D, the system shows a graded response (Figure 1.5A). When an auto-activation loop is added to the first topology, the system shows an ultrasensitive response (Figure 1.5B). Ultrasensitivity is a system level property in which a less-than 81-fold increase in stimulus drives the system form 10% to 90% response[9]. Since the transition curvature is steeper in an ultrasensitive network compared to a graded network, they tend to be less noisy; however, they do not possess hysteresis or memory of the stimulus. When the auto-activation loop requires cooperative binding of the transcription factor A, the system shows bistability (Figure 1.5C). Transcription autofeedback loops with cooperativity are indeed the most recognized topology for showing all-or-none response during cell-differentiation[10].

1.7. Significance of the research

Throughout the normal lifespan of an organism, certain cells and tissues require adult stem cells for turnover and repair. Adult stem cells provide an excellent alternative to embryonic stem cells in basic research and clinical treatment. Also, using adult stem cells for research and clinical purposes avoids any ethical or legal issues that are associated with embryonic stem cells. However, the number of cell types that can be generated by multipotent adult stem cells compares poorly with pluripotent embryonic stem cells. Studies of adult stem cells can provide valuable information about complex signaling events occurring during tissue maintenance and repair. One of the most important motivations in these studies is understanding how undifferentiated cells differentiate and stay differentiated. Dysregulated cell division and differentiation can lead to various physiological disorders including various forms of cancer and tissue abnormalities. Understanding the molecular details of cell cycle progression and stem cell differentiation can suggest new therapies for such diseases.

One of the most important potential applications of adult stem cells is cell-based therapies. The need for transplantable organs and tissues far surpasses the available supply. Stem cells directed to differentiate into specific cell types can offer an ever-lasting source of replacement cells and tissues to treat diseases including anemia, leukopenia, thrombocytopenia, Type I diabetes, Parkinson's, Alzheimer's, spinal cord injury, heart disease, stroke, arthritis, burns and immunodeficiency diseases. For example, in anemia and thrombocytopenia, patients have lower-than-normal blood counts

of erythrocytes and platelets, respectively. Directed differentiation of stem cells to produce erythrocytes and megakaryocytes can provide potential therapies through autologous or allogeneic transplants.

It is well appreciated that understanding and manipulating cellular function requires much more than identifying cellular components and establishing component-component interactions. There is a constant need to understand how these components work together to produce complex, often emergent, behaviors. Systems biology has been successfully used to study biological pathways and its interactions as non-linear networks. Mathematical modeling of these networks can help us gain a quantitative and predictive understanding of a biological system and can further assist in establishing the core topologies that can be used to re-wire or re-engineer native pathways to attain desired phenotypes. Systems biology can thus provide a design framework within which synthetic biology can operate. Synthetic biology is still an emerging field, which has great potential to contribute to a wide-range of applications from therapeutics to alternate fuels. Synthetic biologists are currently at the stage of constructing and testing out fundamental motifs, which can provide a wide range of complex behaviors including ultrasensitivity, bistability, hysteresis, oscillations, noise-reduction and logic gates. In the near future, these regulatory modules can be interfaced with various sensory inputs (environmental, physical or chemical) to confer desired, complex biological responses in processes ranging from metabolism to protein synthesis to cell differentiation.

1.8. Thesis statement

We will use experimental and computational techniques to understand how integrating cell-extrinsic and cell-intrinsic elements can regulate lineage commitment during cell differentiation and can also be used to create tunable synthetic molecular switches for cell engineering applications.

1.9. Chapter layout

In Chapter 2, we present a deterministic model of a cytokine signaling pathway that plays a role in committing progenitor cells to red blood cells, the differentiation process known as erythropoiesis. This model connects erythropoietin (Epo), an erythrocyte-specific cytokine and GATA1, an erythrocyte-specific transcription factor. We show that, due to the presence of multiple topological connections between the cytokine receptor and transcription factor, the system can exhibit a bistable response to the cytokine even without cooperativity. Furthermore, we identify a minimal topology within this erythrocyte model that still gives rise to an all-or-none response. In Chapter 3, we experimentally demonstrate a positive correlation between EpoR and GATA1 expression during human progenitor cell differentiation and we explore the dynamics and phenotypic consequences of this correlation. In Chapter 4, to assess the possibility of using the minimal model discovered in Chapter 1 as a tunable bistable switch, we construct it synthetically in *Saccharomyces cerevisiae* using signaling elements from *Arabidopsis*

thaliana. We demonstrate that key features of our synthetic switch can be rationally tuned using predictions from our minimal model. In Chapter 5, we extend the minimal model in Chapter 1 to a multilineage commitment model and provide a stochastic framework for understanding both canonical and alternative pathways of lineage commitment in hematopoiesis.

Figure 1.1

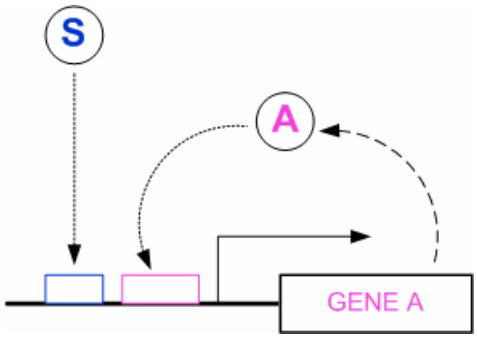


Figure 1-1 Topology of an autofeedback loop

Stimulus S binds to the response element present upstream of gene A and initiates its transcription. Once expressed, A can regulate its own expression through an autofeedback loop.

Figure 1.2

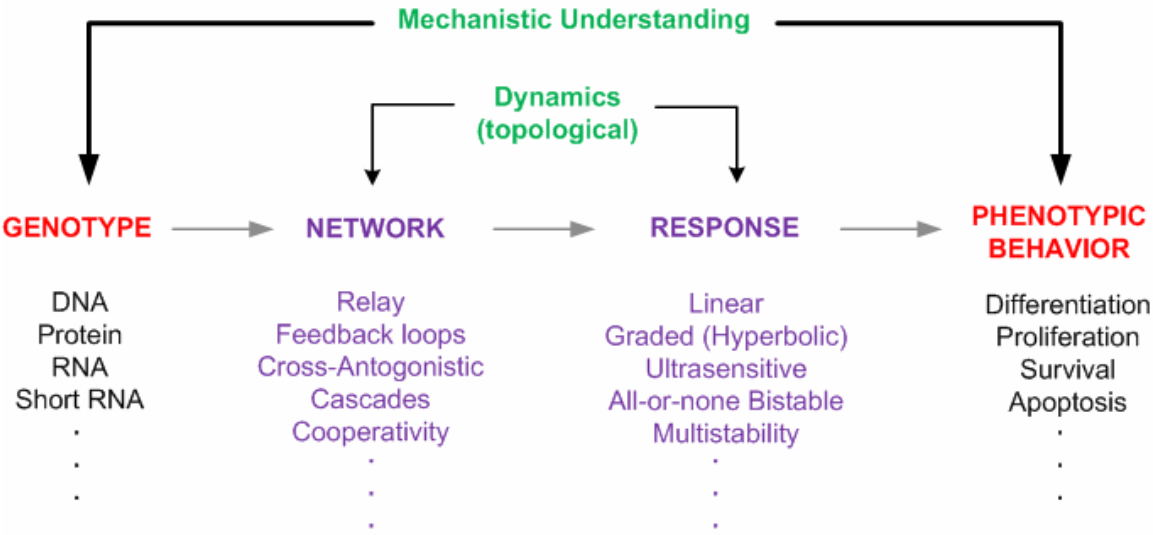


Figure 1-2 List of well studied network topologies and responses

Mechanistic understanding of genotype-phenotype association can be achieved by studying the underlying network-response dynamics.

Figure 1.3

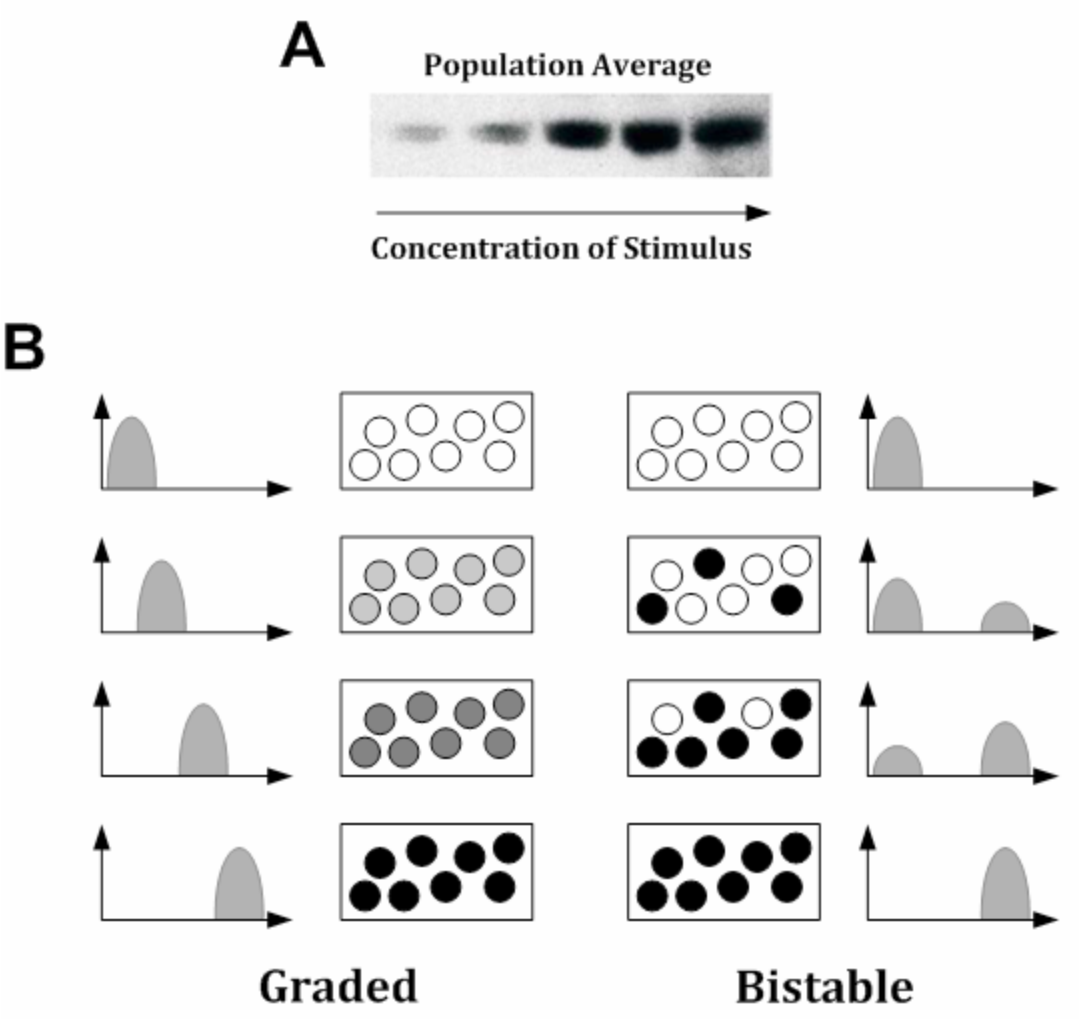


Figure 1-3 Graded vs bistable response in cell differentiation

(A) Western blot showing linear increase in the population-average expression of a gene of interest with change in concentration of stimulus. (B) Two different models explaining the observed result in A; graded response (left) and bistable response (right)

Figure 1.4

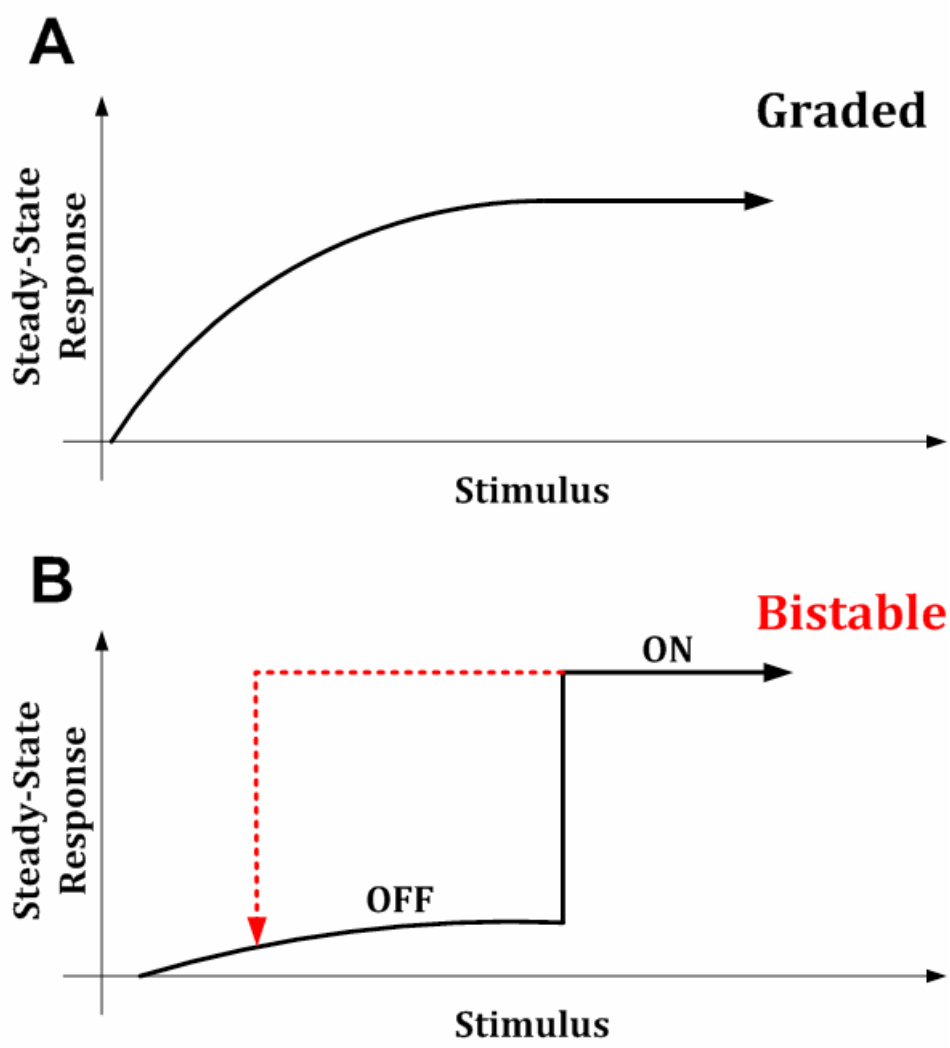


Figure 1-4 Steady-state plots for graded and bistable responses

For each stimulus concentration, the system is allowed to reach steady state and the corresponding response or expression value is plotted. (A) Graded response shows linear change in the levels of response with increase in stimulus until the system reaches saturation. (B) Bistable response stays low for stimulus concentrations less the threshold concentration. Above threshold level, the system switches to the high-expression or on-state. Once the on-state is reached, the system can sustain the response, even when the stimulus is lowered from the initial threshold concentration due to hysteresis (dotted red line).

Figure 1.5

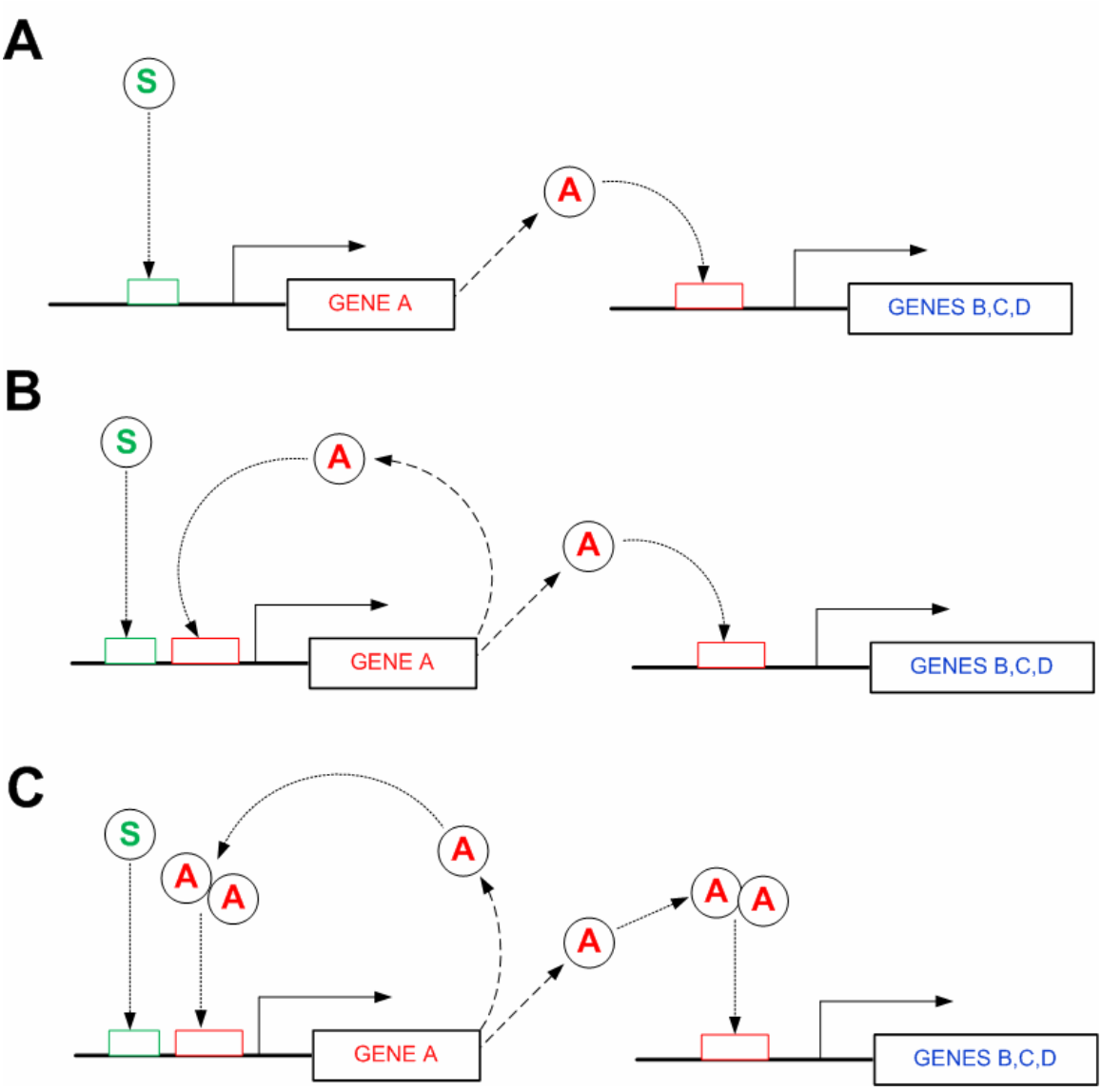


Figure 1-5 Topologies that give rise to graded, ultrasensitive and bistable responses

(A) Stimulus S activates a transcription factor A, which in turn activates several downstream genes (graded response). (B) Topology A with an additional transcriptional autofeedback loop (ultrasensitive response). (C) Topology B with cooperative binding of transcription factor (bistable response).

1.10. References

1. Callard, R.E. & Yates, A.J. Immunology and mathematics: crossing the divide. *Immunology* **115**, 21-33 (2005).
2. Ghildiyal, M. & Zamore, P. Small silencing RNAs: an expanding universe. *Nat Rev Genet* **10**, 94-108 (2009).
3. Santos, S., Verveer, P. & Bastiaens, P. Growth factor-induced MAPK network topology shapes Erk response determining PC-12 cell fate. *Nat Cell Biol* **9**, 324-330 (2007).
4. Bagowski, C.P. & Ferrell, J.E., Jr. Bistability in the JNK cascade. *Current biology : CB* **11**, 1176-1182 (2001).
5. Becskei, A., Seraphin, B. & Serrano, L. Positive feedback in eukaryotic gene networks: cell differentiation by graded to binary response conversion. *The EMBO journal* **20**, 2528-2535 (2001).
6. Ozbudak, E.M., Thattai, M., Lim, H.N., Shraiman, B.I. & Van Oudenaarden, A. Multistability in the lactose utilization network of Escherichia coli. *Nature* **427**, 737-740 (2004).
7. Gardner, T.S., Cantor, C.R. & Collins, J.J. Construction of a genetic toggle switch in Escherichia coli. *Nature* **403**, 339-342 (2000).
8. Xiong, W. & Ferrell, J.E., Jr. A positive-feedback-based bistable 'memory module' that governs a cell fate decision. *Nature* **426**, 460-465 (2003).
9. Huang, C.Y. & Ferrell, J.E., Jr. Ultrasensitivity in the mitogen-activated protein kinase cascade. *Proceedings of the National Academy of Sciences of the United States of America* **93**, 10078-10083 (1996).
10. Ferrell, J.E. & Xiong, W. Bistability in cell signaling: How to make continuous processes discontinuous, and reversible processes irreversible. *Chaos* **11**, 227-236 (2001).

Chapter 2

Receptor Feedback Loop Can Generate Bistability in GATA-1 Expression

(Adapted from Palani S. and Sarkar C.A. (2008) Positive Receptor Feedback during Lineage Commitment Can Generate Ultrasensitivity to Ligand and Confer Robustness to a Bistable Switch, Biophysical Journal, Volume 95, Issue 4, 1575-1589)

2.1. Introduction

The process of cellular differentiation entails a complex series of events through which an uncommitted progenitor can morph into a stable specialized cell. While many of the critical individual molecular components involved in specific differentiation processes have been identified, the complex interactions and topology of signaling and transcriptional networks can lead to non-intuitive behavior. Mathematical modeling and analysis can provide insights into the system-level properties that arise from such an array of interactions.

In cellular processes in which a binary decision must be made, bistability can be an important system-level property that arises from the corresponding signaling

networks. Changes in the system input can toggle a bistable system between two steady states; additionally, the system can display memory by sustaining a high (or low) steady-state response after significant reduction (or increase) in the magnitude of the stimulus[1]. Biological examples of bistability include cell-cycle regulation in *Xenopus* oocytes[2] and *Saccharomyces cerevisiae*[3], self-sustaining biochemical memory[4], synthetic genetic switches[5-9], and differentiation of common myeloid precursors into macrophages and neutrophils[10]. Bistability is often accompanied by ultrasensitivity to a stimulus, another common property of nonlinear systems[11-13]. Since, there is growing evidence that cell differentiation is an all-or-none 'switch-like' event, rather than a continuous transition of an unspecialized cell into a mature one[2, 14], mathematical modeling of the commitment process is attractive because the switch-like response and cellular memory implicit in the biological process arise naturally in the formulation of such nonlinear models.

Hematopoiesis, the formation of blood cells, takes place in two distinct stages: primitive differentiation and terminal differentiation. During primitive differentiation, a hematopoietic stem cell differentiates into a multipotent or bipotent progenitor cell, which, upon terminal differentiation, gives rise to a mature cell. Lineage-specific cytokines (extrinsic) and transcription factors (intrinsic) are believed to be the important molecular components that affect cell survival, proliferation, and commitment during terminal differentiation[15, 16].

Erythropoietin (Epo) is a hematopoietic cytokine responsible for the proliferation, survival, and differentiation of erythroid cells[17]. The Epo receptor (EpoR) has a single transmembrane domain, an extracellular domain for Epo binding, and an intracellular domain for signaling[18]. In the absence of ligand, Epo receptors exist predominantly as

inactive homodimers on the cell surface[19]. Binding of Epo to the receptor homodimer changes the orientation of the receptor subunits, which leads to activation of several signaling cascades including the PI3K/AKT, STAT5-Bcl_{XL}, and Ras/MAPK pathways[20]. Erythroid progenitors lacking functional EpoR do not mature into erythrocytes and show phenotypic abnormalities[21, 22].

The zinc-finger GATA-1 is a transcription factor that plays a critical role in erythroid differentiation[23, 24]. It binds as a monomer to the consensus sequence (A/T)GATA(A/G), which is present in the promoter and enhancer regions of virtually all erythroid-specific genes[25-28]. GATA-1 undergoes several post-translational modifications (acetylation, phosphorylation, sumoylation, and ubiquitination) that may be critical for its optimal transcriptional activity[29]. Analysis of the promoter regions of the EpoR gene shows no TATA or CAAT box, but does reveal the presence of a GATA-1 binding motif, thus providing a meaningful link between a lineage-specific transcription factor and a lineage-restricted receptor[30-32]. Active GATA-1 also binds to the regulatory region of its own gene, thereby enhancing its total expression through a positive feedback loop[33-36]. Disruption of the GATA-1 gene in murine embryonic stem cells by homologous recombination blocks erythroid development, emphasizing the absolute need for GATA-1 in red blood cell maturation[37].

Common myeloid progenitors give rise to erythroid burst-forming units (BFU-E), the earliest known erythroid precursor cells. BFU-E mature into erythroid colony-forming units (CFU-E); this is accompanied by an increase in EpoR expression and the cells become increasingly dependent on Epo[38, 39]. EpoR and GATA-1 levels both rise in parallel and reach their maximum during CFU-E maturation into proerythroblasts and

their subsequent differentiation into early basophilic erythroblasts[40, 41]. Both GATA-1 and EpoR levels fall during further maturation from the basophilic stage to the polychromatic stage as cells synthesize large amounts of globins[38, 41]. Further differentiation from polychromatic erythrocytes to reticulocytes is independent of EpoR and GATA-1, as their levels fall sharply and the cells also show a decrease in globin expression. Hence, it is during the temporal window from an early CFU-E to a basophilic erythroblast that EpoR and GATA-1 may act in concert to drive commitment of the erythroid precursor to terminal differentiation and induce the synthesis of globins.

Recent evidence suggests several modes of crosstalk between EpoR signaling and GATA-1 transcriptional activity, and analysis of these interactions may offer insights into the commitment program during erythroid differentiation. In brief, EpoR signaling via AKT can lead to GATA-1 activation; in return, active GATA-1 can upregulate synthesis of both itself and EpoR. Epo activates AKT by phosphorylating this kinase at Ser-473 in a PI3K-dependent manner[42]. The importance of AKT signaling in erythropoiesis was demonstrated in JAK2-deficient fetal liver progenitor cells: erythroid differentiation can be supported in these cells by overexpressing active AKT and it can also be inhibited by downmodulating AKT using RNA interference[43]. Active AKT appears to have a significant role in enhancing GATA-1 transcriptional activity by mediating some of its post-translational modifications, including phosphorylation and acetylation. AKT phosphorylates GATA-1 at Ser-310 and enhances its transcriptional activity in primary fetal liver cells[42]. However, mice with a S310A mutation in GATA-1 showed no hematopoietic abnormalities during normal or stress erythropoiesis, indicating that

phosphorylation of GATA-1 is dispensable for red blood cell differentiation and may only be required for maximal activity[44].

p300 and CBP acetyltransferases (AT) acetylate GATA-1 at lysine residues present in the C-terminal tail of its zinc fingers[45-47]. *In vivo* ChIP assays show that lysine to alanine mutations at the acetylation residues dramatically impair GATA-1 association with chromatin[48], suggesting that acetylation is critical for GATA-1-mediated gene expression. p300 and CBP also have histone acetyltransferase (HAT) activity and may play a role in enhanceosome stability by acetylating GATA-1 and histones[49, 50]. AKT phosphorylates p300 at Ser-1834 and this has been shown to be essential for AT, HAT, and transcriptional activity of p300[51-53]. Interestingly, Ser-1834 lies in the E1A binding domain that is necessary for binding of p300/CBP to GATA-1. It has also been suggested that phosphorylation may aid in GATA-1 binding to CBP, since the Ser-310 residue of GATA-1 is within the C-terminal acetylation motif of GATA-1. Taken together, these results suggest an additional role for Epo (other than providing survival and proliferation cues) in erythroid precursor commitment and differentiation by activating GATA-1 through the PI3K/AKT pathway and influencing the intrinsic signals that lead to commitment and differentiation.

Based on this experimental evidence, we present a deterministic model of the upregulation and activation of the erythrocyte-specific transcription factor GATA-1, a ‘master regulator’ of erythrocyte commitment. Lineage specification models previously reported suggest that erythrocyte differentiation from erythroid/myeloid bipotent precursor can arise due to the differential expression of antagonistic transcription factors (upregulation of GATA-1 and downregulation of PU.1) driven primarily by cell-intrinsic

events[54, 55]. These models provide insight into the dynamics of a binary cell-fate decision from the viewpoint of 'multilineage priming,' auto-stimulation, and reciprocal repression.

The present work focuses on erythrocyte commitment rather than differentiation and examines how both intracellular and extracellular factors may influence the cell-fate decision. As depicted in Figure 2.1, the topology of our model captures the essential elements of outside-in signaling (Epo-mediated activation of GATA-1), intracellular signal amplification (GATA-1-mediated upregulation of GATA-1 synthesis), and inside-out signaling (GATA-1-mediated upregulation of EpoR). Using this model, we show that upregulation of EpoR in erythroid precursor cells upon Epo addition can generate ultrasensitivity to ligand as well as robust bistability in GATA-1 expression during commitment and this may provide 'switch-like' differentiation characteristics.

Further analysis of a generalized minimal model confirms that the topological connectivity of the two feedback loops alone is both necessary and sufficient for generating the overall system dynamics. Although there are several ways of achieving bistability[1, 56], feedback loops are the most commonly identified mechanism; however, feedback loops that give rise to robust bistability in purely deterministic models have, to date, been shown to be highly cooperative in at least one reaction[57-60]. Here, we present a novel way of achieving robust bistability in cell signaling networks without molecular cooperativity through two linked positive feedback loops. This topology may have general implications for cellular decision-making.

2.2. Model Development

2.2.1. Model construction and description

The core reaction of the proposed erythrocyte commitment model is the activation of GATA-1 by AKT through EpoR signaling (Figure 2.1; light gray background). Our model concentrates exclusively on the two positive feedback loops that serve to increase the concentrations of the reactant species (AKTpp and GATA-1) in this core reaction, which leads to greater accumulation of GATA-1*, the activated form of a 'master regulator' of erythrocyte-specific genes. The model specifically incorporates the following components/motifs in the feedback loops which may have an effect on the overall system behavior. *1) EpoR homodimerization:* Unlike many other cytokine-receptor systems, EpoR homodimerizes (but does not signal) before Epo addition, which should therefore confer ultrasensitivity to the number of receptor dimers available to bind Epo. This effect was modeled as a two-step process of EpoR binding to JAK2 and EpoRJ dimerizing to form EpoRJD. Alternatively, EpoR homodimerization could be treated as a single-step process without considering the effect of JAK2. *2) PI3K/AKT pathway:* Signaling in the MAPK cascade has been shown to convert graded signals into ultrasensitive responses[61]; therefore, the similar cascade structure in the PI3K/AKT pathway might ultrasensitize the signals from the cell surface to GATA-1. *3) Double phosphorylation of AKT:* Recent reports have shown that bistability in signaling circuits can arise from multisite phosphorylation[56]; hence, we explicitly modeled AKT activation as two phosphorylation steps. *4) Transcription and translation:* Delay in feedback loops have been shown to generate interesting behaviors in signaling networks[62], so these two processes were modeled as separate steps. Additionally, explicit inclusion of mRNA

species in the model facilitates comparisons with experimental microarray data (see Figure 2.2). The construction of the full model therefore included the following reaction steps in Figure 2.1.

We have used a deterministic, ODE-based approach to model this signal transduction/transcriptional network. While this modeling framework represents an ideal approximation of the true intracellular milieu[63], it can still provide useful information regarding the system dynamics, particularly for nonlinear systems of the type studied here[64]. In step 1, JAK2 binds to the intracellular domain of EpoR to form a receptor-JAK complex (EpoRJ). EpoRJ dimerizes to form EpoRJD in step 3. EpoRJ and EpoRJD undergo constitutive receptor endocytosis (steps 2 and 4). In step 5, Epo binds to the extracellular domain of EpoRJD forming the activated complex (EpoRJD*) and the endocytosis of the complex is shown in step 6. PI3K is activated (PI3K*) by the complex and is deactivated by a phosphatase (steps 7 and 8). PI3K* converts PIP_2 to PIP_3 in step 9. PIP_3 binds to the PH domain of AKT and phosphorylates AKT on Ser-473 and Thr-308 (steps 11-15). This doubly phosphorylated form of AKT (AKTpp) catalyzes the activation of GATA-1 (step 16). Activated GATA-1 (GATA-1*) is deactivated and degraded in steps 17 and 26, respectively. Monomeric GATA-1* enhances transcriptional synthesis of nuclear EpoR mRNA (EpoRmRNAn) and GATA-1 mRNA (GATA1mRNAn) in steps 18 and 19. The nuclear mRNAs (EpoRmRNAn and GATA1mRNAn) are translocated to the cytoplasm (EpoRmRNAc and GATA1mRNAc, respectively), where they are either translated to their corresponding protein forms or degraded (steps 20 – 25).

EpoR and GATA-1 are present at basal levels in progenitor cells before the addition of Epo. The basal expression of Epo receptor may be independent of GATA-1 as there is also a Sp1 binding site on the 1.7 kb 5'-flanking region of the EpoR gene. Based on current evidence, it appears that, as EpoR is transported to the cell membrane, it is rapidly bound by JAK2 and homodimerizes. Accordingly, we have assumed 90% of EpoR to be initially present in the dimeric state, 9% to be monomers bound by JAK2, and 1% to be free receptors. Activation and deactivation reactions of PI3K, PIP_2, GATA-1, and AKT are assumed to have Michaelis-Menten kinetics. AKT phosphorylation is modeled as a two-step process[65]. Double phosphorylation of AKT by 3'-phosphoinositide-dependent protein kinase 1 (PDK1) is necessary for its complete activation[66, 67]. Dephosphorylation of PI3K*, PIP_3, and AKTpp are implicitly modeled without considering the rate of change of the phosphatases involved. The role of AKTpp in GATA-1 activation is modeled as a single enzymatic step, encompassing both direct (e.g., phosphorylation) and indirect (e.g., acetylation) mechanisms. It is important to note that Epo may activate GATA-1 by AKT-independent mechanisms, but this does not change the qualitative nature of the model (see minimal model below). The mRNA transcription rate is assumed to saturate hyperbolically with active transcription factor concentration, a rapid-equilibrium approximation[68]. The rate of translation is approximated to be proportional to the concentration of the cytoplasmic mRNA[69]. All degradation reactions are modeled with first-order kinetics.

2.2.2. Positive feedback loops

There are two feedback loops considered in this model. Since GATA-1* positively autoregulates its own transcriptional rate, reactions 19, 21, and 24 drive the first positive

feedback loop. This loop increases the concentration of inactive GATA-1 in the cell. The strength of this feedback is governed by the parameter $F_1V_{19,}$ the maximal transcriptional rate of the GATA-1 gene, shown in reaction 19. GATA-1* is also shown to regulate the synthesis rate of EpoRJD through reactions 18, 20, 22, 1, and 3, which start the second positive feedback loop, whose strength is denoted by $F_2V_{18,}$ the maximal rate of production of EpoR mRNA. In this loop, GATA-1* upregulates the expression of EpoRJD which in turn increases the number of complexes formed on the cell surface and leads to the increase in the concentration of activated AKT kinase (AKTpp). In this model, F_1 and F_2 determine the relative strengths of the feedback loops as V_{18} and V_{19} are kept equal and constant. The core reaction in the model is the activation of GATA-1 by AKTpp (reaction 16) and the two feedback loops work synchronously to drive this reaction and produce GATA-1*, which in turn drives both of the feedback loops and also regulates the transcription of other erythrocyte specific genes.

2.2.3. Nondimensionalization and computation

The full model, which consists of 18 ordinary differential equations derived from 27 reactions with 44 parameters, is given in Table 2.1. To simplify parameter estimation and mathematical analyses, the model was completely nondimensionalized; the nondimensional forms of the differential equations and the parameters are given in the Tables 2.2 and 2.3, respectively). In the nondimensional model, each reactant concentration is normalized by the total concentration of its respective basal inactive form. The species used in the mathematical analyses are Epo receptor homodimer (EpoRJD), complex (EpoRJD*), activated AKT (AKTpp), and activated GATA-1 (GATA-1*). The respective nondimensional forms of these reactants are:

$$[RJD]=\frac{[EpoRJD]}{[EpoR_T]_0},[C]=\frac{[EpoRJD^*]}{[EpoR_T]_0},[App]=\frac{[AKTpp]}{[AKT_T]},[GA]=\frac{[GATA\text{-}1^*]}{[GATA\text{-}1_T]_0}.$$

A complete list of the nondimensional reactants and their initial conditions is provided in Table 2.4. The nondimensional equations were solved using the Systems Biology Toolbox (SBtoolbox) for MATLAB (The MathWorks, Natick, MA)[70]. Parameter sensitivity and parameter estimation were also performed with SBtoolbox. MATLAB was used for analyzing bistability and ultrasensitivity through steady-state response plots.

2.2.4. Parameter estimation and sensitivity analysis

Of the 44 parameters present in the model, 29 parameters were incorporated directly from the literature, 8 parameters were refined from values provided in the literature, and the remaining 7 parameters (V_{16}, K_{16}, V_{17}, K_{17}, F_1, F_2, k_{26}) were estimated to fit time course measurements of GATA-1 DNA binding activity during erythroid precursor commitment and differentiation as reported by Dalyot et al. Of these 7 parameters, the steady-state values of the reactants in the model are highly sensitive only to F_1, F_2, and k_{26}. To initially compare the model to these experimental data, a negative feedback loop was added to account for the degradation of GATA-1* after progenitor commitment. This was necessary since the experimental data covers a much broader temporal window of the differentiation process, from GATA-1* production in progenitors to complete GATA-1 degradation in mature erythrocytes. We have assumed that the change in GATA-1 DNA binding activity is due to corresponding changes in the levels of GATA-1*. The fitted parameters were then used in mathematical analyses performed without the negative feedback loop, as our model is only intended to analyze the commitment decision of the progenitor cells much earlier in the differentiation process and not account for larger-

scale phenotypic changes that are observed in mature erythrocytes after commitment. Tables 2.5 and 2.6 give the values of the estimated parameters and the initial conditions of the reactants in the model.

Sensitivity analysis was performed for the Epo receptor dimer, complex, activated AKT, and activated GATA-1 by perturbing all 44 parameters and obtaining the normalized steady-state sensitivities (ranging from 0-1). EpoRJD steady-state levels are most sensitive to the transcriptional rate of *EPOR* (F_2V_{18}) and to the translation and degradation rate constants of cytosolic EpoR mRNA (k_{22} and k_{23}, respectively) (Figure 2.3A). Epo receptor dimer is also highly sensitive to complex formation (k_5) and to a lesser extent to complex dissociation (k_{-5}) and endocytosis (k_6). Epo-EpoR complex (Figure 2.3B) has a sensitivity profile similar to that of the receptor dimer, except that it is more sensitive to the internalization rate constant of the complex (k_6) and less sensitive to binding kinetics (k_5 and k_{-5}). Activated AKT levels are sensitive to the activation and deactivation of PIP_3 and AKT [V_9-V_{15}, K_9-K_{15} and k_{11}] and to the transcriptional rate of *EPOR* (F_2V_{18}) (Figure 2.3C). GATA-1* is most sensitive to the transcriptional rate of *GATA1* (F_1V_{19}) and to the translation and degradation rate constants of cytosolic GATA-1 mRNA (k_{24} and k_{25}, respectively). GATA-1* is also highly sensitive to its degradation rate constant (k_{26}) (Figure 2.3D). It is important to note that the sensitivity analysis was performed while keeping the system in the biologically relevant on-state ($F_1 = 0.04$, $F_2 = 0.123$), which is F_1-limited (cf. Figure 2.9B). Therefore, perturbation of F_2 should not affect the steady-state levels of GATA-1*. From the sensitivity analysis we can see that the steady-state values of the receptor dimer, complex, and activated AKT are influenced by feedback 2 (EpoR synthesis), whereas GATA-1* steady-state levels are influenced

directly by feedback 1 (GATA-1 synthesis) and indirectly by feedback 2 (GATA-1 activation via EpoR synthesis).

Estimation of the unknown rate constants was performed by fitting the GATA-1* time course with GATA-1 DNA binding ability during erythroid differentiation in primary human liquid cultures. To fit the model with the experimental data, a simple negative feedback was added to account for the decrease in the levels of GATA-1 after progenitor commitment. The rate equations and the rate constants of the negative feedback loop are given below:

Rate equations and initial condition:

$$\frac{d[GATA\text{-}1^*]}{dt} = -V_{28}[GATA\text{-}1^* NegFeed][GATA\text{-}1^*]/(K_{28} + [GATA\text{-}1^*])$$

$$\frac{d[GATA\text{-}1^* NegFeed]}{dt} = V_{29}[GATA\text{-}1^*]/(K_{29} + [GATA\text{-}1^*])$$

$$[GATA\text{-}1^* NegFeed]_0 = 0$$

Rate constants:

$V_{28} = 6.90 \times 10^{-5}\ s^{-1}$; $V_{29} = 3.92 \times 10^{-5}$ nM/s; $K_{28} = 293.03$ nM; $K_{29} = 516.49$ nM

The EpoR and GATA-1 mRNA profiles from this EpoR/GATA-1 model with negative feedback were then compared with entirely independent microarray data from differentiating $CD34^+$ human hematopoietic progenitor cells (NCBI Gene Expression Omnibus database (#GDS2431))[71]. As seen in Figure 2.2, the model trends compare favorably with the experimental data. EpoR and GATA-1 mRNA levels are low at the start of terminal differentiation. Upon Epo induction (day 0), both EpoR and GATA-1 mRNA expression levels increase and reach a peak at days 8-9 (late Basophilic

Erythroblast/early Polychromatic Erythroblast stage) and their levels fall with further maturation.

2.2.5. Identification of a generalized minimal model

To ascertain what topological features in the EpoR/GATA-1 model are responsible for its robust bistability, the model was systematically reduced to a minimal form by stepwise elimination of various linear topological motifs, including EpoR homodimerization, PI3K/AKT cascading, multi-site phosphorylation, and individual transcription and translation steps (data not shown). Conversely, both feedback loops were critical for robust bistability (see Results).

The minimal model (Figure 2.4) consists of the following reactions. The cell-surface receptor and the inactive lineage specific transcription factor (InactiveTF) are expressed at basal (ligand-independent) levels in the naïve cell. After addition of ligand, a fraction of the cell-surface receptors become complexes (step 2) and transmit a downstream signal to enzymatically activate the transcription factor (step 4). Constitutive receptor endocytosis, complex internalization, and InactiveTF degradation reactions are shown in steps 1, 3, and 8, respectively. The active transcription factor (ActiveTF) can then upregulate the expression of both receptor and inactive transcription factor (steps 6 and 7, respectively). ActiveTF can be deactivated or degraded (steps 5 and 9, respectively). The activation of transcription factor by complex and its deactivation are modeled as single enzymatic steps and are assumed to have Michaelis-Menten kinetics. Complex internalization and all degradation reactions are modeled to have first-order kinetics. The transcription and translation reactions are modeled as a single step, where the rate of protein formation is assumed to saturate hyperbolically with the concentration

of active transcription factor. The state of the system is represented by the concentration of ActiveTF; high levels denote the on-state (committed state) and low levels denote the off-state (naïve state).

2.2.6. Nondimensionalization and computation of the minimal model

The dimensional and nondimensional forms of the minimal model, each consisting of four differential equations, are provided in Tables 2.7 and 2.8, respectively. The species present in the minimal model are ligand (L, time-invariant), receptor (R), complex (C), inactive transcription factor (ITF), and active transcription factor (ATF). The nondimensionalization was performed in a manner analogous to the EpoR/GATA-1 model (Tables 2.9 and 2.10):

$$[L]=\frac{[Ligand]}{K_d},\ [R]=\frac{[Receptor]}{[Receptor]_0},\ [C]=\frac{[Complex]}{[Receptor]_0},\ [ITF]=\frac{[InactiveTF]}{[InactiveTF]_0},\ [ATF]=\frac{[ActiveTF]}{[InactiveTF]_0}.$$

The system of differential equations was solved analytically using Maple (Maplesoft, Waterloo, Canada) and the full solution for all reactants is given in Table 2.11. Bistability and bifurcation analyses were performed using MATLAB. The values of the kinetic parameters in this model are given in Table 2.12.

Sensitivity analysis was performed for the reactants in the minimal model by perturbing all 16 parameters and obtaining the normalized steady-state sensitivities. The most sensitive parameters for each of the reactants are shown in Figure 2.5. Receptor (R) steady-state levels were primarily sensitive to its transcriptional rate (F_2V_7) and complex formation (k_2). Complex (C) steady-state levels were most sensitive to receptor transcriptional rate (F_2V_7) and complex endocytosis (k_3). Inactive Transcription factor

(ITF) levels were mainly sensitive to receptor transcription rate (F_2V_7), ITF transcription rate (F_1V_6), complex endocytosis (k_3), and activation of ITF by complex (V_4, K_4). Active transcription factor (ATF) steady-state levels were largely sensitive to its degradation (k_9) and ITF transcription rate (F_1V_6). From this analysis, we observe that receptor and complex steady-state levels are sensitive only to feedback 2; ITF levels are sensitive to both feedback 1 and feedback 2; and, ATF levels are sensitive solely to feedback 1.

2.3. Results

2.3.1. Bistability and ultrasensitivity in the EpoR/GATA-1 network

Stimulus/response plots have been used to predict bistability, hysteresis, and ultrasensitivity in molecular networks. The system is induced over a wide range of input stimuli and the corresponding responses are obtained after the system reaches steady state. The state of our EpoR/GATA-1 network is represented by the concentration of GATA-1*; high levels (obtained from both accumulation and activation of GATA-1) denote the on-state (*committed state*) and low levels denote the off-state (*uncommitted state*). In this simulation, Epo was considered to be the stimulus and the responses of important downstream effectors activated by the ensuing signals were analyzed. In Figure 2.6, the steady-state values of the nondimensionalized reactants [RJD, C, App, GA] are plotted against the concentration of Epo normalized to its dissociation constant (K_d = 58 pM[72]). When [Epo] = 0, the system is in the off-state, with RJD at its basal steady-state value of 0.45 and C, App, and GA all at zero, as there are no complexes. As the Epo concentration increases from 0, the steady-state value of the Epo receptor dimer

decreases (Figure 2.6A) as a result of complex formation (Figure 2.6B) and there is a subsequent marginal increase in App and GA (Figure 2.6C and 2.6D). As the concentration of Epo is further increased to $0.96K_d$, the number of complexes formed increases, but this is still not sufficient to maintain the positive feedback loops and the system remains in the off-state. Only when the input stimulus exceeds $0.96K_d$ does the system switch to the on-state, as the complexes can then generate enough AKTpp for GATA-1* levels to exceed the threshold concentration needed to sustain the feedback loops. Therefore, the system exhibits ultrasensitivity for a small perturbation in the concentration of Epo about $0.96K_d$. The on-state is accompanied by a large burst of GATA-1*, an event known to precede the accumulation of various erythroid specific genes. The system continues to remain in the on-state with further increase in Epo levels.

To explore whether this network can exhibit memory, the system was taken to the on-state by increasing the concentration of Epo to its K_d value. The stimulus was then reduced to $0.96K_d$ and the system was allowed to reach steady-state. It can be seen from the plots that the system remains in the on-state as the active positive feedback loops can sustain the system in the committed state. Thus, the downstream effectors in the system exhibit hysteresis with respect to cytokine stimulus. As the Epo concentration is further reduced from $0.96K_d$ to $0.008K_d$, the steady-state value of RJD increases since less complexes are formed and, in turn, there is a reduction in the levels of App; nonetheless, the number of complexes is still sufficient to sustain the feedback loops and to maintain high levels of GATA-1*. As Epo levels are reduced below $0.008K_d$, the system switches back to the off-state due to a lack of sufficient Epo-mediated signaling. When the system is in the committed state, removal of the stimulus below the threshold level does not

immediately bring the system back to the off-state, instead exhibiting bistability over a large range of stimulus concentration. This bistable expression of GATA-1* can reduce the sensitivity of the system to noise by necessitating a high Epo concentration to initially achieve the on-state and, thereafter, by providing marked robustness to the active state. Though the on-state is still maintained when the stimulus level is reduced approximately 120-fold below the threshold concentration, further decreases in Epo concentration drive the system back to the off-state, suggesting that it is not completely irreversible. This is in accord with the phenotypic change observed after commitment during which the cell becomes increasingly independent of EpoR signaling and GATA-1 levels start to fall. The high expression of GATA-1 achieved by Epo induction at commitment can initiate chromatin rearrangements and expression of lineage specific genes, thereby 'locking' the cell in the mature state. The steady states plotted in Figure 2.6 are only the stable values; the unstable steady states are omitted, as they are not experimentally accessible.

2.3.2. Pretreatment can change the threshold concentration of the stimulus

The steady-state response plot of GA (Figure 2.6D) shows that the Epo concentration has to be greater than $0.96K_d$ for the system to be in the on-state. Is there a way to attain the on-state for values of Epo less than the threshold concentration? Given the memory implicit in this network, we hypothesized that transient pretreatment of cells with high concentrations of Epo should influence their commitment decision since the switch to the on-state is determined by the number of complexes needed to sustain the positive feedback loops. If the cell were pretreated with a high concentration of Epo for a fixed amount of time, it should still be possible to achieve the on-state even if the Epo concentration were then reduced to a level lower than the threshold concentration

($0.96K_d$), since there would be an appropriate accumulation of multiple activated species during pretreatment. To test this using our model, the concentration of Epo during pretreatment was fixed at its K_d value and was then reduced to the value given on the x-axis in Figure 2.7. The minimum pretreatment time required for the system to attain the on-state for a range of constant Epo concentrations lower than the threshold concentration is plotted in Figure 2.7. The corresponding plots of RJD, C, App, and GA requirements to achieve the on-state for lower Epo concentrations are similar and are given in Figure 2.8. For Epo concentrations greater than $0.96K_d$, the threshold concentration, the cell does not require pretreatment for commitment. As the Epo concentration is reduced from $0.96K_d$ to $0.008K_d$ (a range that corresponds precisely to the bistable window in Figure 2.6), the pretreatment time required to accumulate sufficient GATA-1* to attain the on-state increases exponentially. Reducing the Epo concentration below $0.008K_d$, does not bring the system to the on-state for any pretreatment time, as the system is in the monostable off-state below this Epo concentration (see Figure 2.6).

2.3.3. Double positive feedback loops lead to robust bistability

The EpoR/GATA-1 network consists of two positive feedback loops that coordinate to create a burst of GATA-1*, an event critical for erythrocyte commitment. The first feedback loop is the transcription of *GATA1* by GATA-1*, which increases the concentration of inactive GATA-1 (substrate), and the second feedback loop is the transcription of *EPOR* by GATA-1*, which leads to an increase in the levels of AKTpp (enzyme) in the presence of Epo. Parameters F_1 and F_2 govern the maximum transcriptional rates of *GATA1* and *EPOR*, respectively, and hence represent the strength of the GATA-1*/GATA-1 and GATA-1*/EpoR/AKTpp feedback loops, respectively.

The parameter-fitted values of F_1 and F_2 are 0.04 and 0.123, which correspond to a steady-state value of GA = 295 as seen in Figure 2.6D. In addition to other epigenetic factors, one possible explanation for the difference in the transcriptional rates of *GATA1* and *EPOR* could be the distinct mechanisms by which GATA-1 binds to its consensus sequence present in the promoter regions of these genes. It should be noted that the two positive feedback loops are interdependent (linked via the GATA-1 activation reaction; reaction with light gray background in Figure 2.1) and are necessary for the commitment decision to accumulate GATA-1*. When $F_1 = 0$, the cell cannot make more inactive GATA-1, and can only activate the existing low levels of GATA-1, so the system stays in the off-state for any value of $F_2 > 0$; similarly, when $F_2 = 0$, the cell cannot make enough surface complexes to activate GATA-1 via AKTpp, so the system remains in the off-state for any physiologically reasonable value of $F_1 > 0$. For very large values of F_1, however, high levels of GATA-1* can be achieved, albeit in a manner that does not impart memory to the system (the stimulus/response plot in this case is hyperbolic and monostable everywhere).

Figure 2.9A shows a 3D plot of the steady-state value of GA as a function of F_1 and F_2 when the Epo concentration is equal to its K_d. Changing the values of F_1 and F_2 can switch the system from the off-state to the on-state as well as change the set point of the reactants – specifically GA – in the on-state. As seen from the plot; for $F_1 = 0$, $F_2 > 0$ or $F_1 > 0$, $F_2 = 0$, the system is in the off-state. For the estimated value of $F_1 = 0.04$, as we increase F_2 from 0 the system remains in the off-state until F_2 reaches 0.118. Any increase of F_2 over 0.118 causes the system to switch to the on-state with a GA set point value of 295. Further increasing F_2 does not change the value of GA and the system

remains in the on-state. For the estimated value of F_2 = 0.123, as we increase F_1 from 0 the system remains in the off-state until F_1 = 0.01, at which point the system switches to the on-state with a low GA set point value. As we further increase F_1 to 0.04, the system stays in the on-state and increases the GA steady-state value to the estimated value of 295. When F_1 is increased beyond 0.04, the steady-state value of GA increases and saturates at an F_1 value of 1.2.

A top view of the 3D plot is given in Figure 2.9B to address the effect of changes in F_1 and F_2 on the set point of GA in the on-state. In this phase diagram, the on-state is divided into two regions: F_1-limiting, where an increase in F_1 (but not F_2) will increase the set point of GA in the on-state, and F_2-limiting, where an increase in F_2 (but not F_1) will increase the set point of GA in the on-state. It can also be seen that the critical value of F_2 above which the system attains the on-state slightly decreases as we increase F_1 and the critical line eventually asymptotes at F_2 = 0.05 for very high F_1. The EpoR/GATA-1 system is likely to always be F_1-limited because of the extremely high GATA-1 levels required to be F_2-limited.

Figure 2.10 shows the steady-state response plots of GA for several values of F_1 and F_2 (analogous to Figure 2.6D). In Figure 2.10A, F_1 is kept constant at 1.8 and F_2 levels are varied between 0.15 and 2.00, to span both the F_1 and F_2-limiting regions (see Figure 2.9B). For F_2 = 0.15, the threshold concentration of Epo needed to switch the system to the on-state is $0.85K_d$. As F_2 is increased, the threshold Epo concentration decreases and the system reaches the on-state for lower stimulus values. It can also be seen that the set point value of GA in the on-state is dependent on F_2 only in the F_2-limiting region. The bistable window shown here is comprised of a constant and a

variable on-state GA set point region. The constant on-state GA region is at its maximum width in the F_1-limiting region (F_2 = 2.0) and decreases for F_2 values closer to the F_2-limiting region (F_2 = 0.40, 0.25). For F_2 values in the F_2-limiting region (F_2 = 0.20, 0.15), the on-state GA set point varies throughout the bistable window and decreases as the Epo concentration is decreased. In Figure 2.10B, F_2 is kept constant at 0.20 and F_1 is varied from 0.1 and 2.0. Increasing F_1 value does not change the threshold concentration of Epo needed to achieve the on-state; however it increases the set point of GA in the on-state. Similar to Figure 2.10A, the constant on-state GA region in the bistable window decreases when moving from an F_1-limiting region to an F_2-limiting region.

2.3.4. Bistable expression of GATA-1*

The steady-state response plot in Figure 2.6D shows the wide range of Epo concentrations in which GATA-1* exhibits bistable expression for the fitted F_1 and F_2 values. To understand the influence of the two positive feedback loops in defining the bistable window, we plotted the monostable (either ON or OFF) and bistable (ON and OFF) GATA-1* regions as a function of Epo concentration and feedback strength. Here, the F_1 and F_2 values are chosen to cover both the F_1- and F_2-limiting regions as shown in Figure 2.9B. In Figure 2.11A, a log-log plot of F_1 vs. $[Epo]/K_d$, with F_2 constant (0.20), shows the regions of monostable and bistable expression of GATA-1*. At low F_1 values, the system only achieves bistability for a narrow range of Epo concentrations. As we increase F_1, the bistable window increases and remains constant for larger F_1 values. The increase in the bistable window is only due to the decrease in the bistable-on to monostable-off transition concentration, as the bistable-off to monostable-on threshold concentration remains constant for all values of F_1. This reveals that the Epo

concentration at which the system initially switches to the on-state is independent of F_1. However, F_1 governs the extent of memory in the system by changing the Epo concentration at which the system switches from the on-state to the off-state. Figure 2.11B shows a semi-log plot of F_2 vs. $[Epo]/K_d$, with F_1 constant (1.8). For values of F_2 less than 0.08, the system remains in the off-state for all Epo concentrations. The system attains the on-state for higher F_2 values and also exhibits bistability for a wide range of Epo concentrations. In contrast to Figure 2.11A, the bistable window in Figure 2.11B shifts as we increase F_2 due to a substantial decrease in the bistable-on to monostable-off transition concentration as well as a smaller decrease in the bistable-off to monostable-on threshold concentration. This indicates that F_2 plays a role in determining the Epo concentration at which the system reaches on-state as well as in governing the magnitude of memory in the system.

Since basal levels of inactive GATA-1 are low, the system needs both F_1 (for accumulation of GATA-1) and F_2 (for activation of GATA-1) to attain the on-state (accumulation of activated GATA-1). For systems having a high basal expression level of inactive transcription factor or lineage specific receptor (though neither is the case for the erythrocyte differentiation problem), it becomes relevant to examine how bistability may be achieved. Can such systems potentially attain the on-state* (activation, no accumulation) even if $F_1 = 0$ or $F_2 = 0$? We tested this using our EpoR/GATA-1 model, with F_1 fixed at 0. Figure 2.11C shows the bistable expression of the active transcription factor in the absence of F_1. For a given F_2, the Epo concentration at which the system attains the on-state does not change when compared with Figure 2.11B, but the respective concentration at which the system switches back to the monostable off-state is increased

dramatically, thus narrowing the bistable window, or memory, in the system. A system which has feedback 1 (upregulation of GATA-1 by GATA-1*) but no feedback 2 (upregulation of EpoR by GATA-1*) does not exhibit bistability for any value of F_1, confirming the observation that autoregulating positive feedback loops without cooperativity do not show bistability in deterministic models; bistability in a system lacking feedback 2 can be recovered by incorporating the need for transcription factor dimerization for activation (data not shown). In summary, this shows that the bistability and ultrasensitivity achieved in the Epo-GATA-1 model were primarily due to the presence of feedback 2, and that feedback 1 only plays a role in increasing the extent of memory in the system. The values of F_1 and F_2 may also change during the differentiation process, thus dynamically modulating the robustness of the system, though this time-dependence was not considered here.

2.3.5. Construction of a generalized minimal model

The EpoR/GATA-1 model exhibits ultrasensitivity and bistability for a wide range of Epo, F_1, and F_2 values. The structural aspects of the EpoR/GATA-1 model are the two linked positive feedback loops, receptor homodimerization, PI3K/AKT signaling pathway, double phosphorylation of AKT, and the transcription and translation steps. We systematically developed and tested various sub-models of the parent model to identify the dispensable steps and to obtain a generalized minimal model that still retains the ultrasensitivity and robust bistability of the parent model (data not shown). This analysis revealed that the two positive feedback loops were both necessary and sufficient for recapitulating the overall system behavior of the full EpoR/GATA-1 model. This reduced lineage-specific receptor/transcription factor model (Figure 2.4) includes only four time-

dependent species: receptor (R), complex (C), and inactive (ITF) and active (ATF) transcription factor. This model was solved analytically and the exact solution for each of the four reactants was determined. The steady-state response plots for these reactants for selected values of F_1 and F_2 are given in Figure 2.12. The bistability plots of ATF (Figure 2.13) in the minimal model closely mimic those in the EpoR/GATA-1 model (Figure 2.11).

2.3.6. Bifurcation analysis of the minimal model

Unlike the EpoR/GATA-1 model, the minimal model can be solved analytically, which can prove useful in understanding the contributions of each of the two positive feedback loops to the overall behavior of the system. The solution curves of ATF are plotted against L (normalized to its K_d) for various values of F_2 holding F_1 constant at 20 as shown in Figure 2.14A. The solid lines and the dotted lines denote the stable and unstable roots, respectively. For low values of F_2 (0.01), the system has only two real roots (one stable and one unstable) and is purely monostable for all ligand concentrations. As F_2 value is increased to 0.1, the expression of the ActiveTF becomes narrowly bistable, with the endpoints of this bistable window defined by two saddle-node bifurcations that appear to depend on F_2 but not on F_1 (see below). The degree of bistability increases dramatically as F_2 is increased to 1 and then to 10. It can be seen that, for these F_2 values, the solution curves intersect to form a transcritical bifurcation. The transcritical set point, which constrains the maximum theoretical value of ATF, seems to be independent of F_2. The transcritical bifurcation also divides the bistable window into F_1-limiting (right; constant on-state value) and F_2-limiting regions (left; variable on-state value). For low F_2 values, the bistable region is completely F_2-limited and as we increase F_2, the bistable

region becomes increasingly F_1-limited. Finally, it can also be observed that the threshold ligand concentration to achieve the on-state decreases as we increase the F_2 value.

Figure 2.14B shows the bifurcation diagrams for ATF plotted against L by varying F_1 and keeping F_2 constant at 10. Unlike the previous case (Figure 2.14A), the system can achieve bistability over a narrow range of L by forming two saddle-node bifurcations even when F_1 is zero. As we increase F_1, the size of the bistable window and the on-state set point value both increase. At a critical value of F_1 (here 4.3), the maximum bistable window is achieved, coincident with the appearance of an apparent subcritical pitchfork bifurcation at the lowest value of L at which the system is still bistable. At this F_1 value, the set point of the ATF in the bistable region is completely F_1-limited. As F_1 is increased beyond 4.3, the solution curves form a transcritical bifurcation similar to that seen in Figure 2.14A. As we further raise the value of F_1 to 20 and then 200, the value of L at which the transcritical bifurcation occurs shifts from low to high, making the bistable region increasingly F_2-limited. This is in contrast to Figure 2.14A, where the transcritical bifurcation point moves from right to left and the bistable region becomes increasingly F_1-limited as we increase F_2. Importantly, increasing F_1 augments the maximum on-state set point value of ATF but has no effect on the threshold ligand concentration necessary for achieving the on-state.

By comparing the bifurcation plots of ATF in Figure 2.14 with various plots of activated GATA-1 (Figure 2.6D, Figure 2.10, and Figure 2.11), the following conclusions can be deduced for the EpoR/GATA-1 system: the width of the bistable region and the range of GATA-1* values in the on-state can both depend on F_1 (under F_1-limited conditions) and/or F_2 (under F_2-limited conditions); the maximum GATA-1* value in the

on-state is determined by F_1; the threshold Epo concentration at which the system switches to the on-state is dictated by F_2; the maximum bistable window achievable is set by F_2; and, the system requires an F_2 value above a critical threshold in order to exhibit bistability.

2.4. Discussion

EpoR and GATA-1 are both essential for erythrocyte precursor commitment and differentiation, and we present here a deterministic model that bidirectionally links the lineage-specific receptor and transcription factor. Based on recent biochemical data, we chose the PI3K/AKT cascade as the signaling pathway that connects EpoR and GATA-1. The model accounts for basal expression of EpoR and GATA-1, Epo binding to EpoR to activate the PI3K/AKT pathway, activation of GATA-1 by phosphorylated AKT, positive autoregulation of GATA-1 expression by GATA-1*, and upregulation of EpoR expression by GATA-1*.

In order to gain mechanistic insights into system behavior, we chose to focus on this small set of critical molecular effectors implicated in erythropoiesis. However, it should be noted that our explicitly modeled topology represents only a fraction of the full regulatory network and, therefore, inferring cell fate from the level of a single metric (e.g., GATA-1*) represents an approximation of a high-dimensional attractor[55, 73]. Signaling pathways that were excluded from the present model include JAK2/STAT5/Bcl_{XL}, which provides anti-apoptotic signals during erythrocyte differentiation[74], and Ras/MAPK, which is involved in cell survival[75], cell-cycle

regulation[76], and the degradation of DNA-bound GATA-1[77]. Also, the JAK2/STAT5 pathway activated by Epo can initiate a negative feedback loop on the PI3K/AKT pathway by activating SOCS proteins which can suppress Epo receptor signaling[78, 79]. Our model, despite neglecting these additional complexities, can nevertheless effectively capture the system dynamics observed experimentally (see Figure 2.2).

Through steady-state response plots (Figure 2.6) and bistability plots (Figure 2.11), it was revealed that the EpoR/GATA-1 network can exhibit ultrasensitivity and bistability. Since these properties may play important roles in erythrocyte commitment, it was informative to probe the role of positive feedback in such a topology. As shown in Figure 2.14A, positive receptor feedback can ultrasensitize the system to ligand and can generate a considerable memory effect once the on-state is achieved. Other transcription factors (e.g., GATA-3[80]) are believed to be intracellularly amplified through a classical autoregulatory positive feedback loop: synthesis of the new transcription factor is followed by dimerization (or higher order oligomerization) and the complex is then transcriptionally active. If this is sufficient for programming cell fate, why, then, might a transcription factor such as GATA-1 have evolved to upregulate a lineage-specific receptor as well?

The answer may lie in the different modes of activation. While the dimerization step closes the positive feedback loop for some transcription factors, experimental evidence suggests that GATA-1 binds DNA as a monomer[81, 82] and shows no detectable DNA-binding ability before the addition of Epo. Thus, EpoR signaling may be necessary to close the GATA-1 autoregulatory loop by activating the transcription factor via AKT. By upregulating EpoR to increase its own activation, GATA-1 can effectively mimic the

molecular cooperativity of other transcription factors in generating robust network bistability without employing any cooperative reactions. (The importance of cooperativity in achieving bistability is restricted to the class of deterministic models discussed in the present work; it is indeed possible to achieve steep sigmoidal responses through nonidealities such as molecular crowding[83], stochastic focusing[84], and dimensionally-restricted reactions.)

Two unique elements of the EpoR/GATA-1 model should be highlighted. First, by decoupling the synthesis and activation steps in the positive GATA-1 autoregulatory loop, a cell may be able to independently tune the switching threshold, the on-state expression level, and the extent of memory in the network by separately modulating F_1 and F_2 (e.g., epigenetically). Second, there is an external checkpoint (Epo) that modulates this autoregulatory loop. This is attractive because it provides a novel and meaningful link between canonically extrinsic (cytokine) and intrinsic (transcription factor) signals in regulating not only cell survival but also maturation.

The strengths of the positive feedback loops are governed by the rates of transcription of *GATA1* (F_1) and *EPOR* (F_2). The estimated values of F_1 and F_2 are 0.04 and 0.123, respectively. The difference in the rates of transcription of *EPOR* (chromosome 19p13.2) and *GATA1* (chromosome Xp11.23) may be due to the distinct binding mechanisms of GATA-1 to these promoters, dissimilarities in the ease of accession of the GATA-1 binding sites, and the recruitment of other cofactors that may regulate *EPOR* and *GATA1* expression differently. GATA-1 also interacts with other factors, notably the ubiquitous transcription factor Sp1, erythroid restricted factor EKLF,

and friend of GATA-1 (FOG1) that may alter its transcriptional activity among the various GATA-1 regulated genes[29].

Though treated as constants in our model, F_1 and F_2 may also change temporally during commitment and differentiation due to additional biophysical (e.g., chromatin remodeling) and biochemical (e.g., co-factor upregulation/downregulation) processes. Accordingly, the values of F_1 and F_2 may also vary substantially between primary cells and immortalized lines, and may even differ among cell lines, depending upon how far a cell line is from commitment towards the erythrocyte lineage, relative expression of GATA-1 cofactors, basal levels of EpoR and GATA-1 expression, and expression of antagonistic transcription factors driving other lineages. Cell-specific feedback strengths that differ significantly from those used in our models may serve to attenuate or amplify the actual effects on the network.

Finally, system-level properties such as bistability and ultrasensitivity that may be generally applicable to lineage commitment can be experimentally corroborated. Pretreatment of progenitor cells with ligand, as outlined in the Results (Figure 2.7), can be performed to show expected hystereses in transcription factor activation and lineage commitment. Additionally, the models elucidate how the steady-state response profiles of activated transcription factor can be influenced by F_1 and F_2, and these can be experimentally validated by using pharmacological inhibitors or RNA interference to exogenously manipulate the values of F_1 and F_2. While the results presented here are motivated by the EpoR/GATA-1 network and its critical role in erythropoiesis, it will be interesting to see whether similar topologies are uncovered in other cell systems which enable their hosts to make robust decisions in response to external stimuli.

Figure 2.1

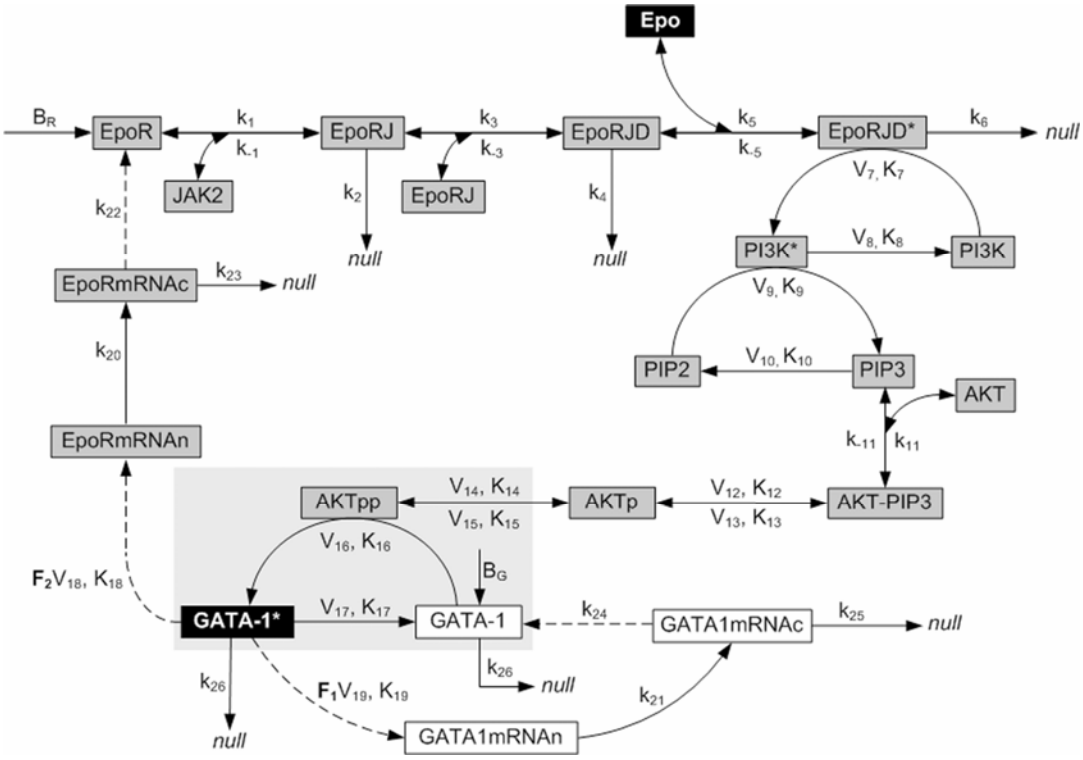
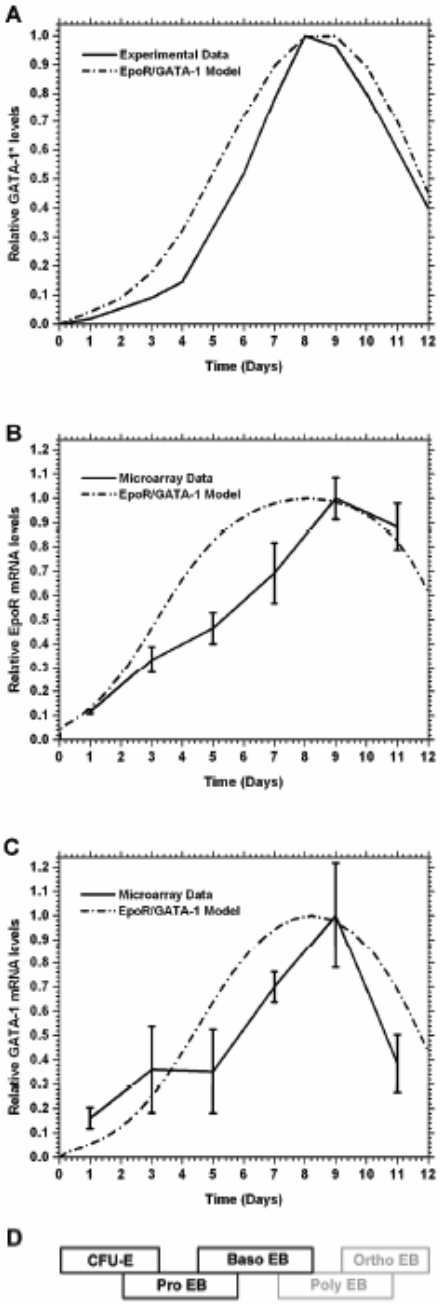


Figure 2-1 EpoR-GATA1 model

Kinetic model of the EpoR/GATA-1 network implicated in erythrocyte progenitor commitment. EpoR and GATA-1 are present at low basal levels before the addition of Epo. Epo binds to EpoR homodimers (EpoRJD), thereby activating AKT through the PI3K/AKT pathway. Doubly phosphorylated AKT (AKTpp) activates GATA-1 directly and indirectly through covalent modifications (modeled here as a single step). GATA-1*, the activated form of the transcription factor, upregulates GATA-1 and EpoR gene expression by binding to GATA motifs present in the response elements of their corresponding genes. The core reaction of the model is the activation of GATA-1 by AKTpp and is highlighted with a light gray background. The reactants in the white boxes comprise one feedback process (with transcriptional strength F_1), a synthesis loop that generates more inactive GATA-1 (substrate), and the species in the dark gray boxes represent a second feedback process (with transcriptional strength F_2), an Epo-regulated loop that makes more AKTpp (enzyme). Both feedback loops provide inputs to the core reaction to form GATA-1* (product). All reactants except Epo and JAK2 are time-variant. Reactions (1-6, 11, 20-21, 23, and 25-26), (7-10, 12-17, 22, and 24), and (18 and 19) are modeled with mass- action, Michaelis-Menten, and rapid-equilibrium kinetics, respectively. The species names ending with mRNAc and mRNAn denote cytoplasmic and nuclear mRNA, respectively. Double-headed ⟷ and single-headed ⟶ solid arrows indicate reversible and irreversible reactions respectively. Dashed ⇢ arrows specify irreversible reactions (transcription, translation) in which reactants are not consumed. *, p, and pp denote the activated, singly phosphorylated, and doubly phosphorylated forms of species. All reactions going to *null* denote first-order degradation processes. The values of the rate constants shown in the figure are given in Table 2.5.

Figure 2.2



CFU-E: Colony Forming Units – Erythroid; Pro EB: Pro Erythroblasts; Baso EB: Basophilic Erythroblast; Poly EB: Polychromatic Erythroblast; Ortho EB: Orthochromatic Erythroblast.

Figure 2-2 Comparison of the EpoR/GATA-1 model with experimental data

(A) GATA-1* time course plot generated from the EpoR/GATA-1 model is fitted to the experimental data to obtain the unknown kinetic parameters. (B & C) Relative EpoR and GATA-1 mRNA plots obtained after parameter estimation are compared with microarray data during $CD34^+$ progenitor differentiation. (D) Characterized stages of red blood cell development during terminal differentiation are drawn to scale using the x-axis in plots A, B, and C. The EpoR/GATA-1 model is relevant only for erythrocyte commitment (CFU-E to Baso EB) and does not cover the latter stages of differentiation (from Poly EB (grayed out))

Figure 2.3

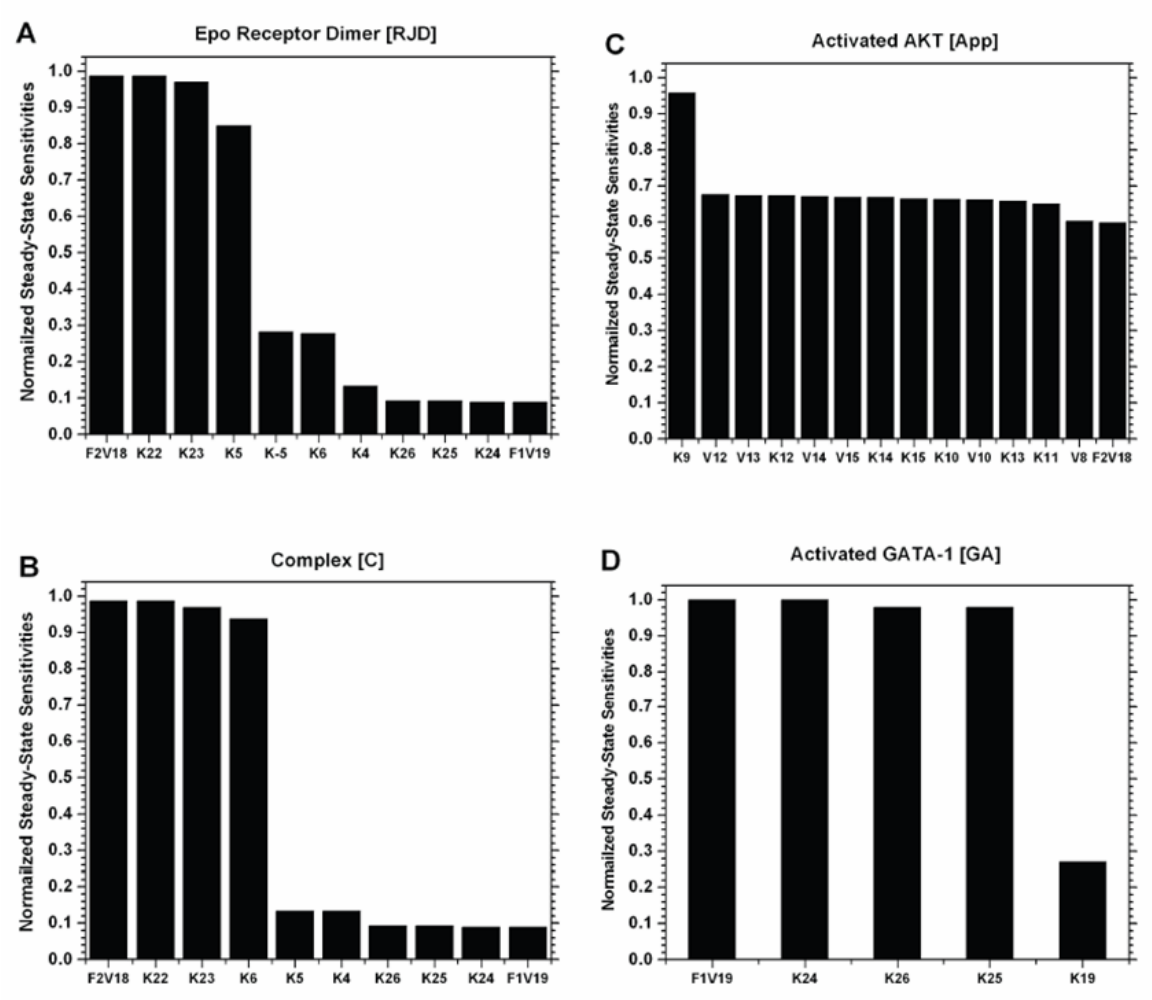


Figure 2-3 Parameter sensitivity analysis

(A) Epo receptor dimer [RJD]. (B) Epo-Epo receptor complex [C]. (C) Activated AKT [App]. (D) Activated GATA-1 [GA].

Figure 2.4

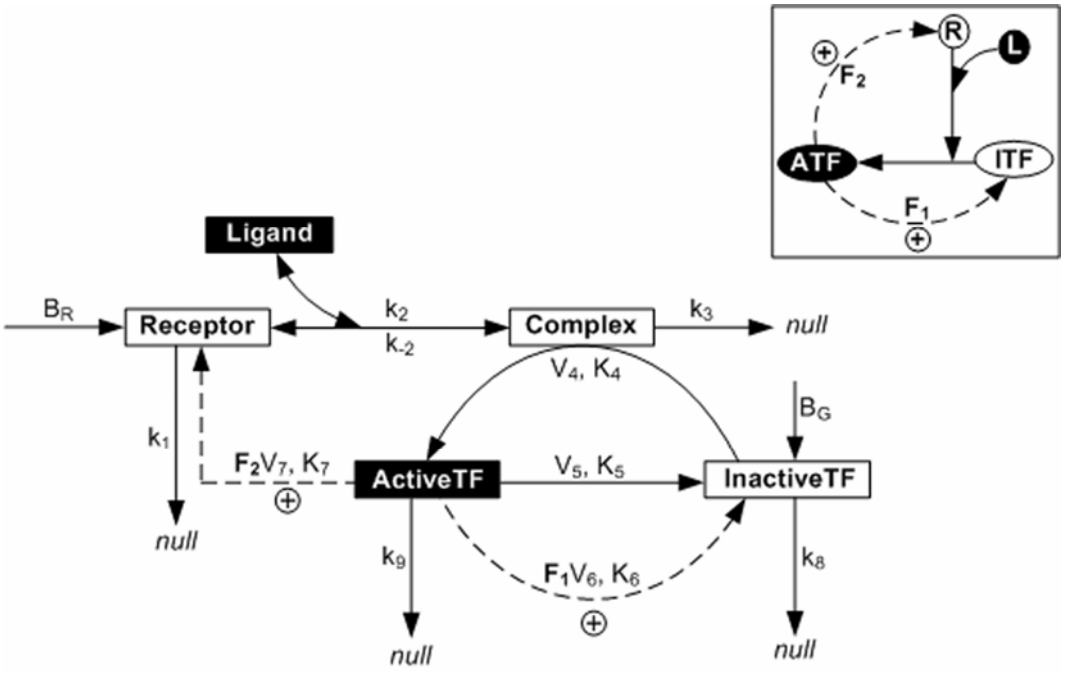


Figure 2-4 Generalized minimal model for lineage commitment

Receptor and Inactive Transcription Factor (InactiveTF) are present at basal levels before the addition of ligand. Ligand binds to Receptor to form Complex and activates InactiveTF to form Active Transcription Factor (ActiveTF). ActiveTF upregulates the levels of InactiveTF and Receptor through two positive feedback loops. Reactions (1-3, 8, and 9), (4 and 5), and (6 and 7) are modeled with mass-action, Michaelis-Menten, and rapid-equilibrium kinetics, respectively. Double-headed ⟷ and single- headed ⟶ solid arrows indicate reversible and irreversible reactions, respectively. Dashed ⇢ arrows specify irreversible transcriptional activation and translation reactions (modeled as a single step). All reactants except ligand are time variant. All reactions going to *null* denote first-order degradation processes. The values of the rate constants shown in the figure are given in Table 2.12. *(inset)* A further simplified schematic of the minimal model highlighting the two feedback loops. L, R, ITF, and ATF denote the nondimensional forms of Ligand, Receptor, InactiveTF, and ActiveTF. Basally expressed R converts basally expressed ITF to ATF only in the presence of L. ATF upregulates itself by inducing the expression of both ITF (with transcriptional strength F_1) and R (with transcriptional strength F_2). It should be noted that positive feedback to ITF is intrinsically regulated, whereas positive feedback via R (to activate ITF) is dependent on the external stimulus L.

Figure 2.5

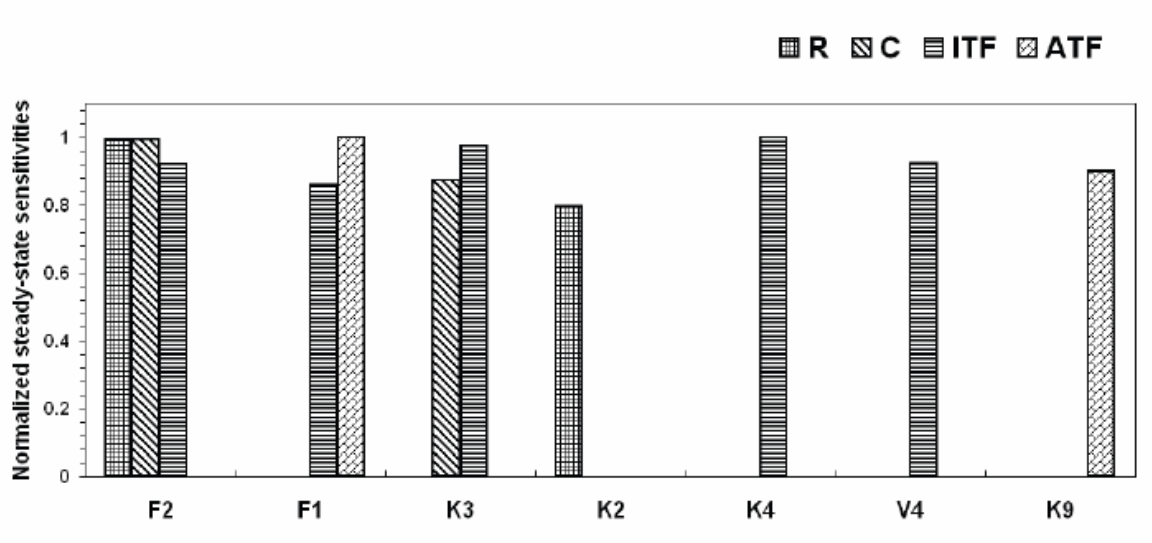


Figure 2-5 Normalized steady-state sensitivities of the reactants in the minimal model

Figure 2.6

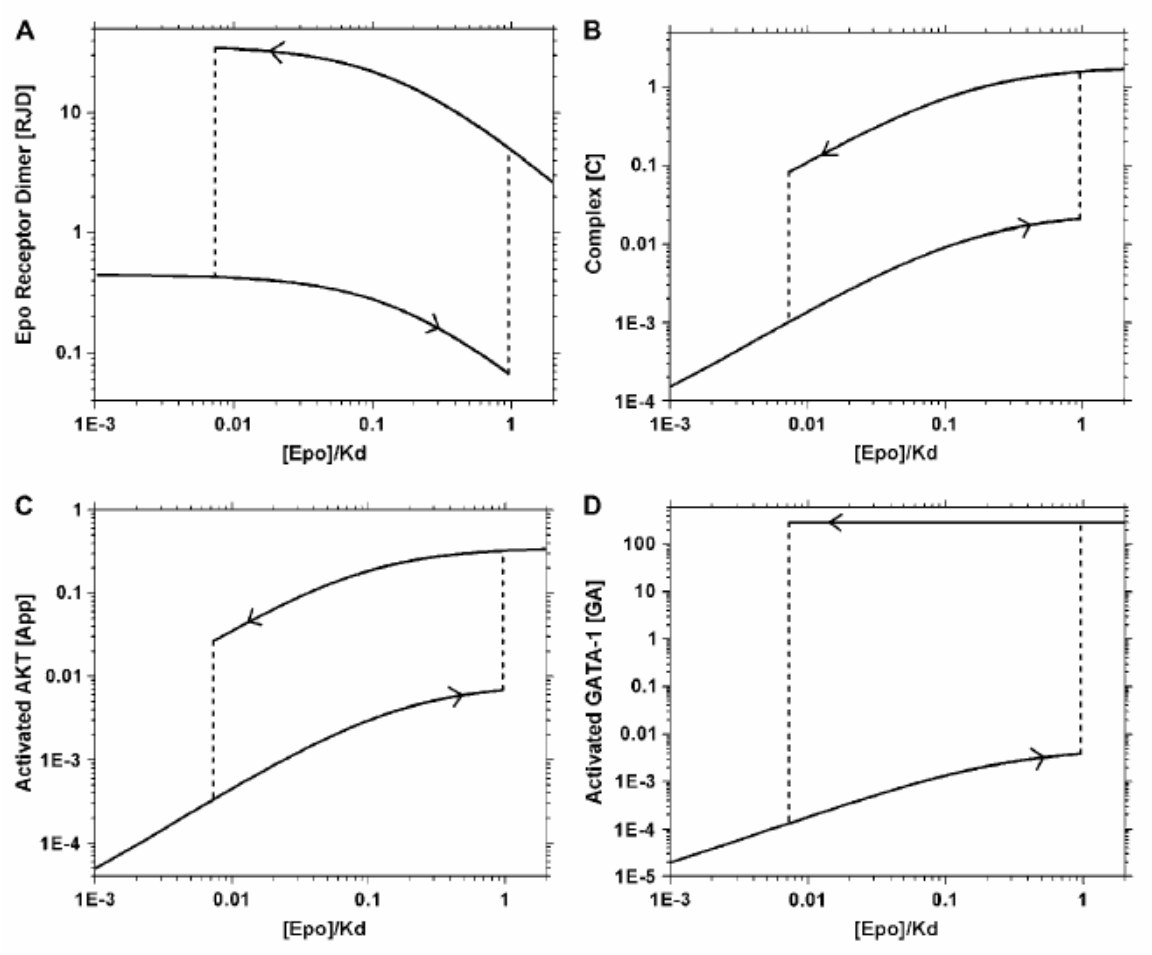


Figure 2-6 Nondimensionalized steady-state response plots

(A) Epo receptor dimer [RJD]. (B) Epo-Epo receptor complex [C]. (C) Activated AKT [App]. (D) Activated GATA-1 [GA]. The stimulus, Epo, is normalized to its K_d value and each downstream effector is normalized to the total concentration of its respective basal inactive form. The plots show that, for the fitted values of F_1 (0.04) and F_2 (0.123), the system is ultrasensitive to Epo and exhibits bistability for a wide range of Epo concentrations ($0.008K_d$ to $0.96K_d$).

Figure 2.7

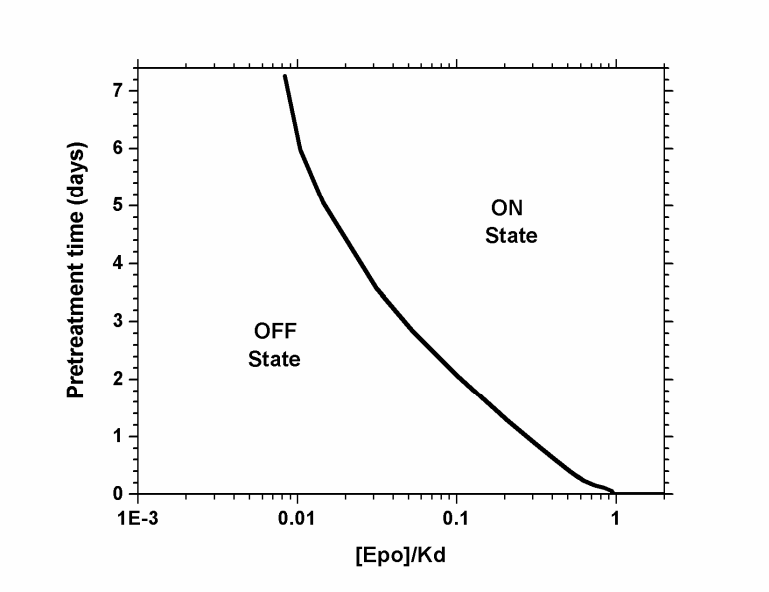


Figure 2-7 Pretreatment time plot

Minimum pretreatment time required for the system to attain the on-state for a range of Epo values lower than the threshold concentration ($0.96K_d$). Epo is normalized to its K_d. The pretreatment concentration of Epo is kept at its K_d value and is thereafter reduced to the value given on the x-axis. The plot suggests that for Epo concentrations greater than $0.96K_d$, the cell should not require pretreatment and, for values less than the threshold concentration, the pretreatment time increases dramatically with decreases in Epo concentration. For Epo values less than $0.008K_d$, the system will always remain in the monostable off-state for any pretreatment time (cf. Figure 2.6).

Figure 2.8

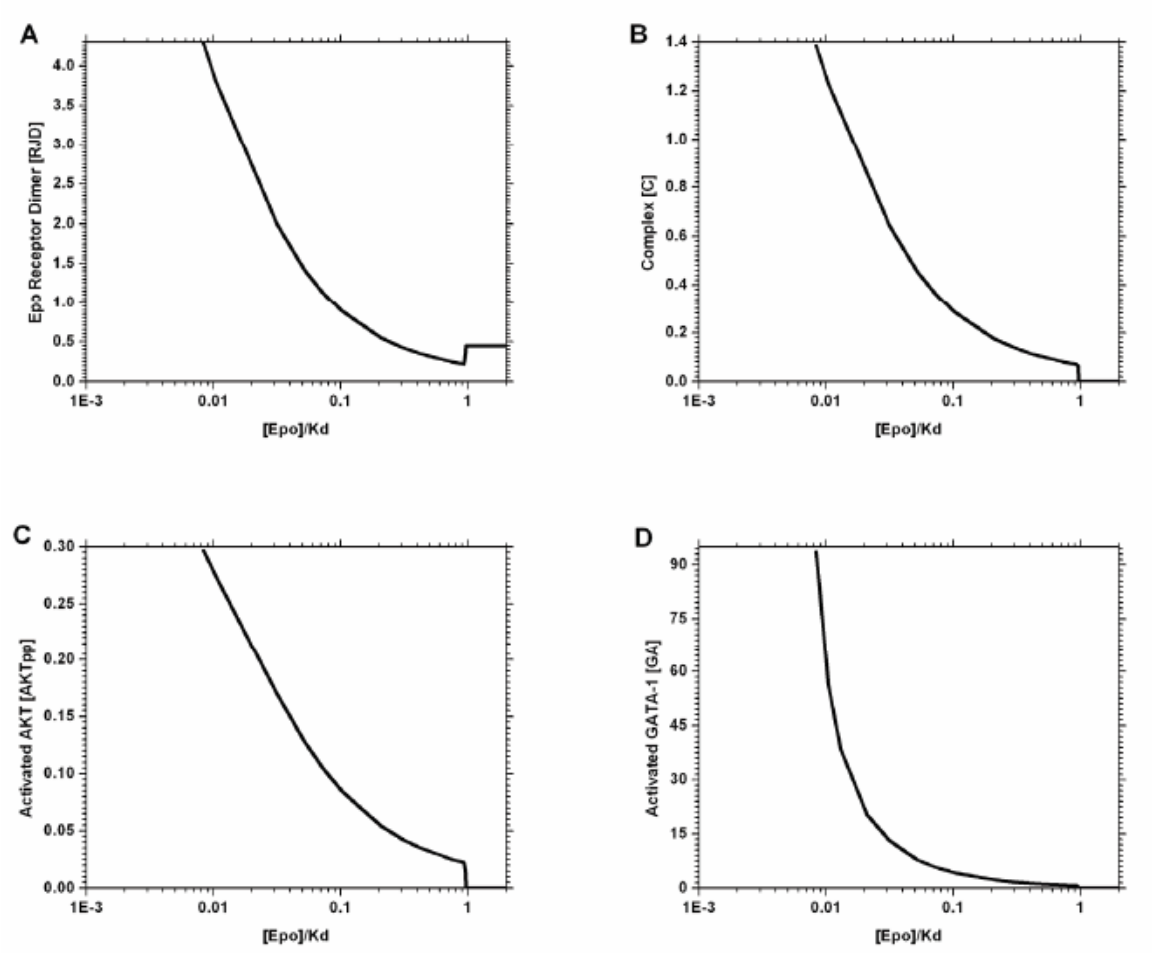


Figure 2-8 Minimum reactant concentration plot

Nondimensional minimum initial effector concentrations required for the system to achieve the on-state for a range of Epo concentrations lower than the threshold concentration ($0.96K_d$). Epo is normalized to its K_d value. (A) Epo receptor dimer [RJD]. (B) Epo-Epo receptor complex [C]. (C) Activated AKT [App]. (D) Activated GATA-1 [GA].

Figure 2.9

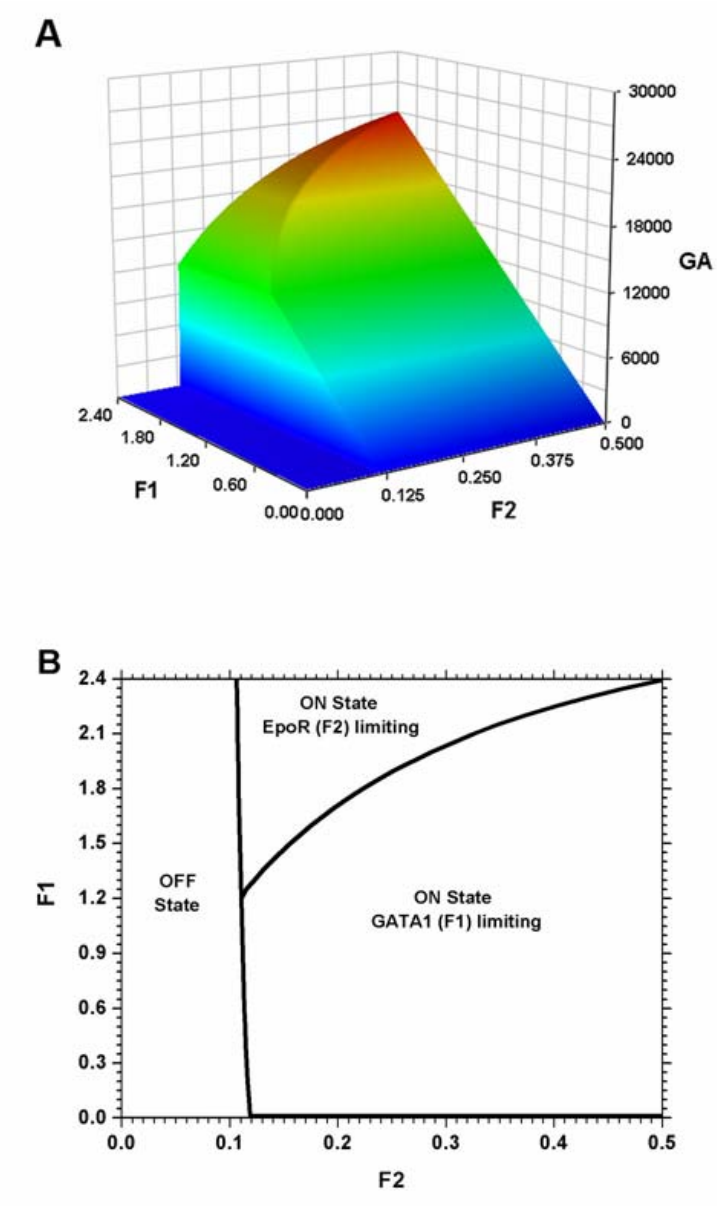


Figure 2-9 Effect of the two positive feedback loops on the on-state GA value

(A) Steady-state GA values as a function of F_1 and F_2; Epo is kept at its K_d value. For the estimated values of F_1 (0.04) and F_2 (0.123), the system is strongly F_1-limited. (B) Corresponding phase diagram of the 3D plot showing the off-state region, the F_1-limited on-state, and the F_2-limited on-state. Increasing the values of F_1 and F_2 increases the on-state set point of GA in the F_1-limited and F_2-limited regions, respectively. The EpoR/GATA-1 system is likely to behave as an F_1-limited system due the high GA values required to be F_2 limiting.

Figure 2.10

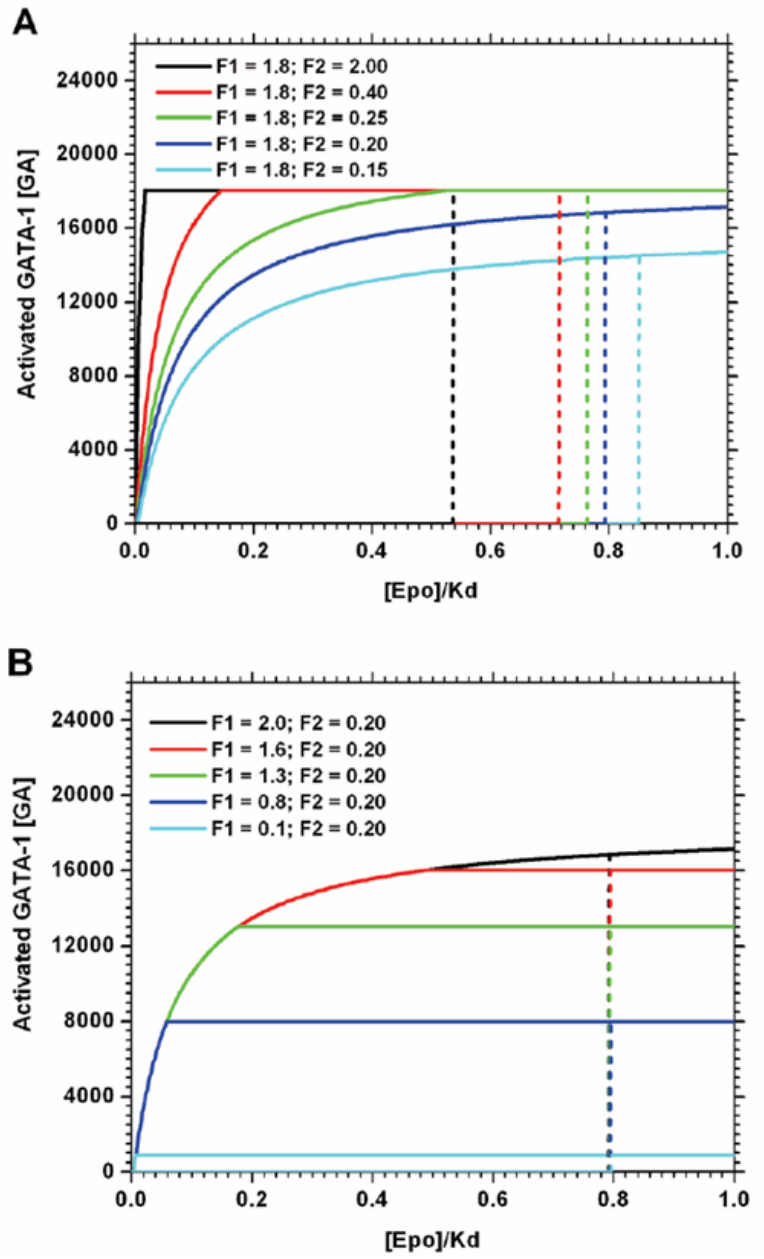


Figure 2-10 Steady-state response plots of GA for various values of F_1 and F_2

Steady-state response plots of GA for various values of F_1 and F_2, spanning both the F_1- and F_2-limiting regions. (A) Changing F_2 keeping F_1 constant at 1.8. (B) Changing F_1 keeping F_2 constant at 0.20.

Figure 2.11

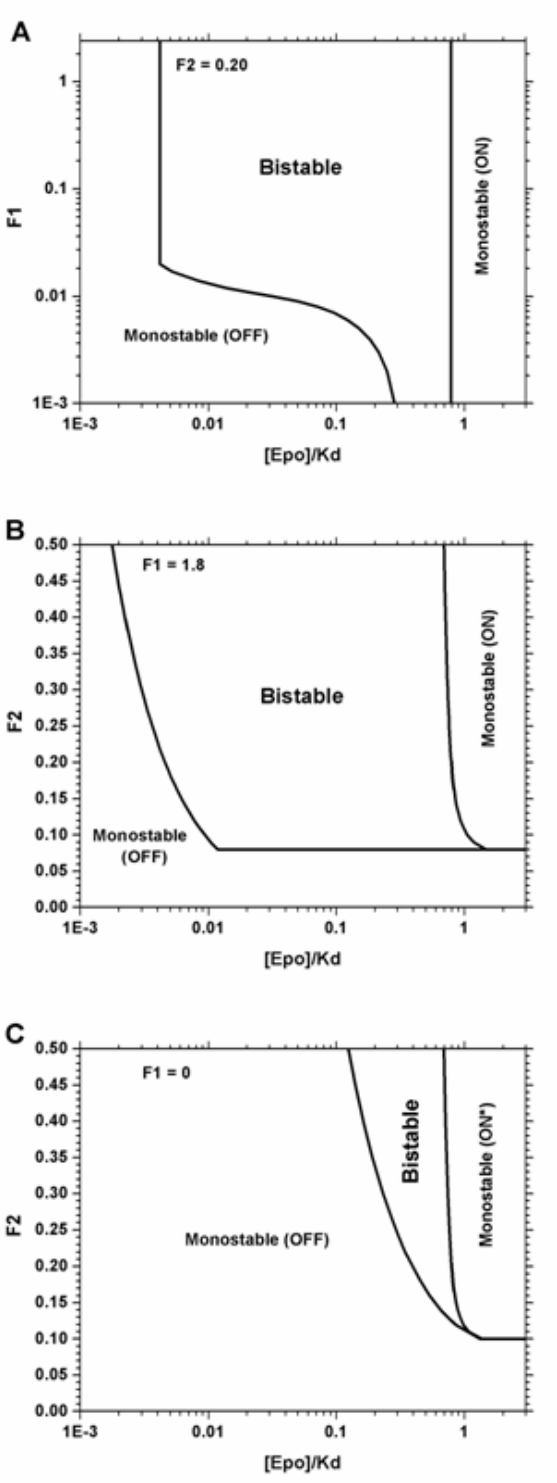


Figure 2-11 Bistable expression of GATA-1*

(A) A log-log plot showing the change in the bistable expression of GATA-1* for varying F_1 (F_2 = 0.20). The threshold Epo concentration needed to achieve the on-state appears to be independent of F_1. The maximum bistable window achievable is dependent on F_2; however, the width of the bistable window is F_1-dependent for lower F_1 values. (B) A semi-log plot showing the bistable expression of GATA-1* for changing F_2 values (F_1 = 1.8). There appears to be a threshold F_2 value below which the system is purely monostable. Increasing F_2 increases the width of the bistable window and, to a lesser extent, decreases the threshold Epo concentration required to reach the on-state. (C) Narrow bistable expression of GATA-1* for various F_2 values when F_1 = 0. The ON* state denotes an on-state due to activation but no accumulation. This state would not commit a cell to differentiate due to low basal levels of GATA-1. When F_2 = 0, the system is monostable for any F_1 value (e.g., the x-axis in Figure 2.11B).

Figure 2.12

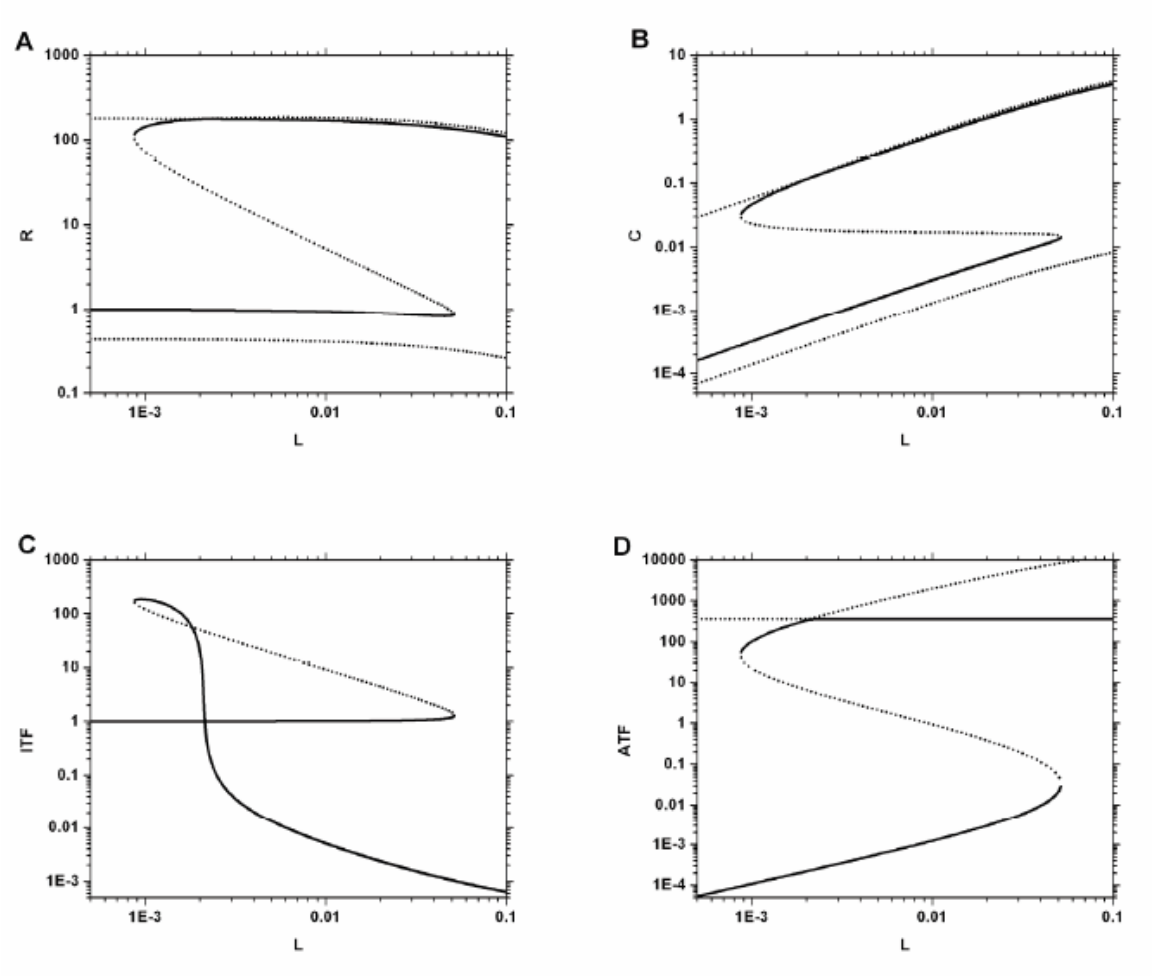


Figure 2-12 Nondimensionalized steady-state response plots for the minimal model

Nondimensionalized steady-state response plots showing both the stable (solid) and unstable (dotted) roots of the reactants in the minimal model: (A) Receptor [R]. (B) Complex [C]. (C) Inactive Transcription Factor [ITF]. (D) Active Transcription Factor [ATF]. The nondimensionalized stimulus, L, is the ligand concentration divided by its K_d value and each downstream effector is normalized to the total concentration of its respective basal inactive form.

Figure 2.13

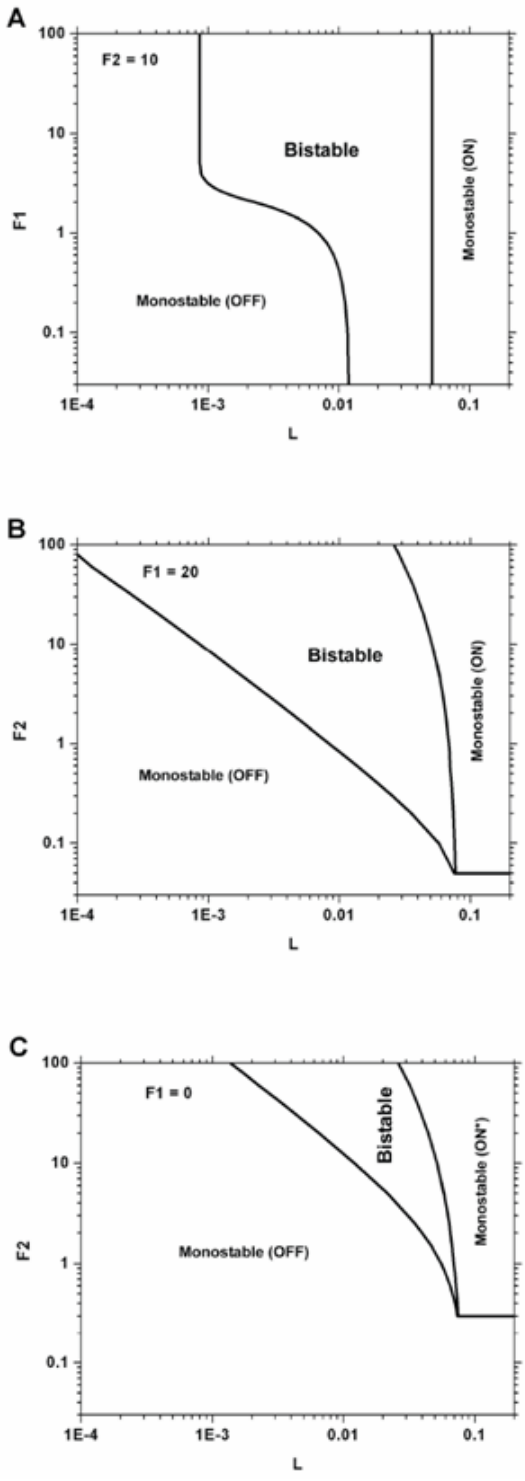


Figure 2-13 Bistability plots of ActiveTF

(A) A log-log plot showing the change in the bistable expression of ActiveTF for varying F_1 (F_2 = 10). (B) A log-log plot showing the bistable expression of ActiveTF for varying F_2 (F_1 = 20). (C) Bistable expression of ActiveTF for various F_2 values when F_1 = 0. The ON* state denotes an on-state due to activation but no accumulation. This state would not commit a cell to differentiate due to low basal levels of InactiveTF. When F_2 = 0, the system is monostable for any F_1 value (not shown).

Figure 2.14

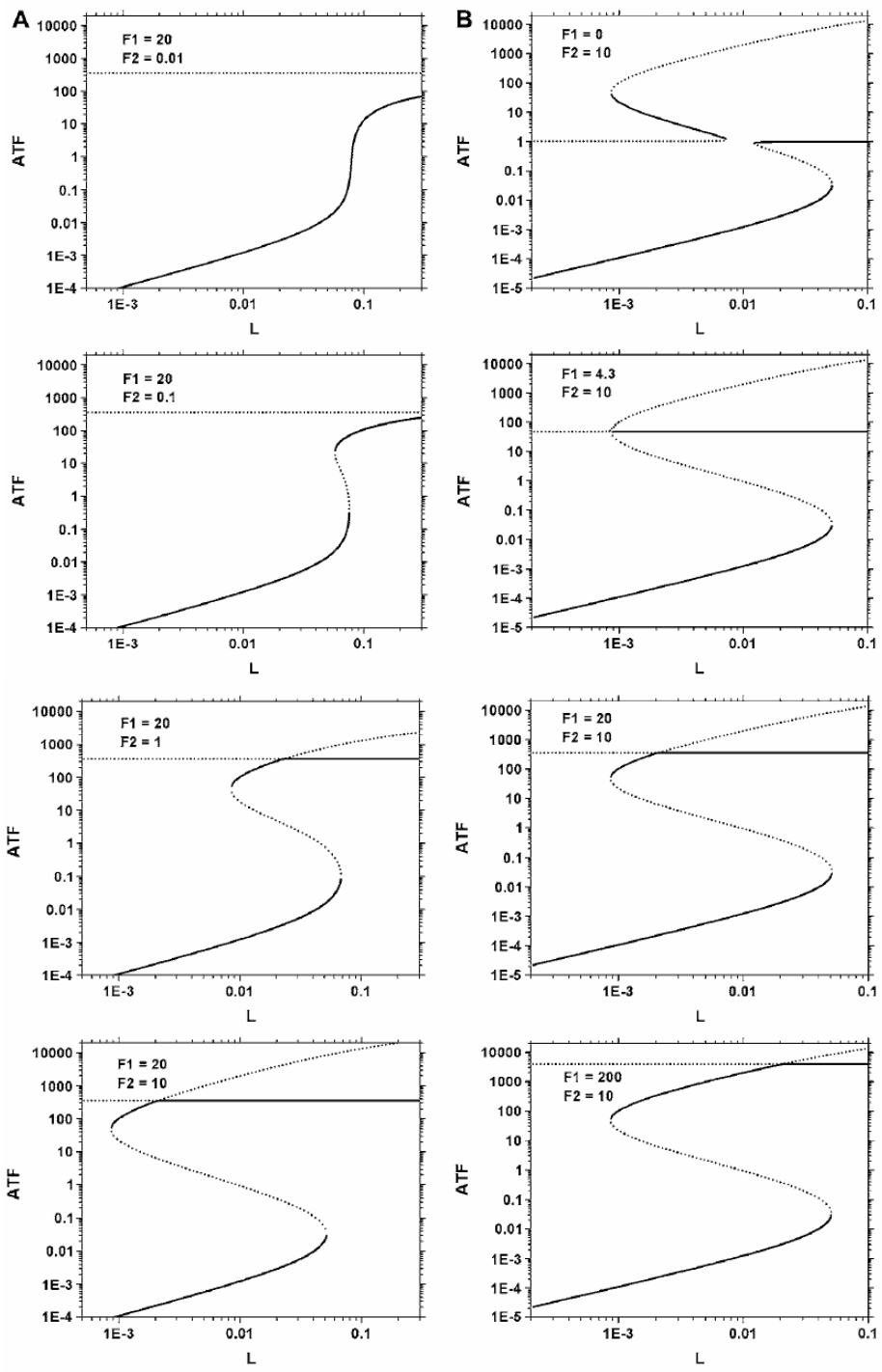


Figure 2-14 Bifurcation analysis for the minimal model

(A) Solution curves obtained from the analytical solution of nondimensionalized Active Transcription Factor (ATF) are plotted against ligand concentration normalized to its K_d (L) by varying F_2 (0.01, 0.1, 1, and 10); F_1 is held constant at 20. The solid lines and the dotted lines denote the stable and unstable roots respectively. For low F_2 values, the system is purely monostable. As we increase F_2, two saddle-node bifurcations emerge and they determine the width of the bistable window and the threshold ligand concentration necessary to reach the on-state. As we further increase F_2, the solution curves intersect to form a transcritical bifurcation. The saddle-node set points and the transcritical set point (maximum achievable on-state value) are F_2- and F_1-dependent, respectively. (B) F_1 is varied (0, 4.3, 20, and 200); F_2 is held constant at 10. The system can achieve bistable expression of ATF even when $F_1 = 0$. As we increase F_1, the bistable window increases and reaches a maximum width, which is determined only by the F_2 value. Increasing F_1 further increases the on-state set point value of ATF but has no effect on the threshold ligand concentration required to reach the on-state.

Table 2.1

Table 2-1 Rate equations for EpoR/GATA-1 erythrocyte commitment model

$$\frac{d[EpoR]}{dt} = R_{22} - R_1 + B_R$$
$$\frac{d[EpoRJ]}{dt} = R_1 - R_2 - 2R_3$$
$$\frac{d[EpoRJD]}{dt} = R_3 - R_4 - R_5$$
$$\frac{d[EpoRJD^*]}{dt} = R_5 - R_6$$
$$\frac{d[PI3K]}{dt} = R_8 - R_7$$
$$\frac{d[PI3K^*]}{dt} = R_7 - R_8$$
$$\frac{d[PIP2]}{dt} = R_{10} - R_9$$
$$\frac{d[PIP3]}{dt} = R_9 - R_{10} - R_{11}$$
$$\frac{d[AKT]}{dt} = -R_{11}$$
$$\frac{d[AKT\text{-}PIP3]}{dt} = R_{11} - R_{12} + R_{13}$$
$$\frac{d[AKTp]}{dt} = R_{12} - R_{13} - R_{14} + R_{15}$$
$$\frac{d[AKTpp]}{dt} = R_{14} - R_{15}$$
$$\frac{d[GATA\text{-}1]}{dt} = R_{17} - R_{16} + R_{24} + B_G - R_{26}$$
$$\frac{d[GATA\text{-}1^*]}{dt} = R_{16} - R_{17} - R_{27}$$
$$\frac{d[EpoRmRNAn]}{dt} = R_{18} - R_{20}$$
$$\frac{d[EpoRmRNAc]}{dt} = R_{20} - R_{23}$$
$$\frac{d[GATA1mRNAn]}{dt} = R_{19} - R_{21}$$
$$\frac{d[GATA1mRNAc]}{dt} = R_{21} - R_{25}$$

$$1: k_1[EpoR][JAK2] - k_{-1}[EpoRJ]$$
$$2: k_2[EpoRJ]$$
$$3: (k_3/2)[EpoRJ]^2 - k_{-3}[EpoRJD]$$
$$4: k_4[EpoRJD]$$
$$5: k_5[Epo][EpoRJD] - k_{-5}[EpoRJD^*]$$
$$6: k_6[EpoRJD^*]$$
$$7: V_7[EpoRJD^*][PI3K]/(K_7 + [PI3K])$$
$$8: V_8[PI3K^*]/(K_8 + [PI3K^*])$$
$$9: V_9[PI3K^*][PIP2]/(K_9 + [PIP2])$$
$$10: V_{10}[PIP3]/(K_{10} + [PIP3])$$
$$11: k_{11}[PIP3][AKT] - k_{-11}[AKT\text{-}PIP3]$$
$$12: V_{12}[AKT\text{-}PIP3]/(K_{12} + [AKT\text{-}PIP3])$$
$$13: V_{13}[AKTp]/(K_{13} + [AKTp])$$
$$14: V_{14}[AKTp]/(K_{14} + [AKTp])$$
$$15: V_{15}[AKTpp]/(K_{15} + [AKTpp])$$
$$16: V_{16}[AKTpp][GATA\text{-}1]/(K_{16} + [GATA\text{-}1])$$
$$17: V_{17}[GATA\text{-}1^*]/(K_{17} + [GATA\text{-}1^*])$$
$$18: F_2V_{18}[GATA\text{-}1^*]/(K_{18} + [GATA\text{-}1^*])$$
$$19: F_1V_{19}[GATA\text{-}1^*]/(K_{19} + [GATA\text{-}1^*])$$
$$20: k_{20}[EpoRmRNAn]$$
$$21: k_{21}[GATA1mRNAn]$$
$$22: k_{22}[EpoRmRNAc]$$
$$23: k_{23}[EpoRmRNAc]$$
$$24: k_{24}[GATA1mRNAc]$$
$$25: k_{25}[GATA1mRNAc]$$
$$26: k_{26}[GATA\text{-}1]$$
$$27: k_{26}[GATA\text{-}1^*]$$

Table 2.2

Table 2-2 Nondimensional rate equations for EpoR/GATA-1 model

$$\frac{d[R]}{dT} = M37 + M1 - M2 + M3$$

$$\frac{d[RJ]}{dT} = M2 - M3 - M4 + M5 - M6$$

$$\frac{d[RJD]}{dT} = M7 - M8 - M9 - M10 + M11$$

$$\frac{d[C]}{dT} = M10 - M11 - M12$$

$$\frac{d[PA]}{dT} = M13 - M14$$

$$\frac{d[PI]}{dT} = M15 - M16$$

$$\frac{d[PIA]}{dT} = M16 - M15 - M17 + M18$$

$$\frac{d[A]}{dT} = -M19 + M20$$

$$\frac{d[Ap]}{dT} = M21 - M22 - M23 + M24$$

$$\frac{d[App]}{dT} = M23 - M24$$

$$\frac{d[G]}{dT} = M33 + M25 - M26 + M27 - M28$$

$$\frac{d[GA]}{dT} = M26 - M25 - M29$$

$$\frac{d[mGn]}{dT} = M30 - M31$$

$$\frac{d[mGc]}{dT} = M31 - M32$$

$$\frac{d[mRn]}{dT} = M34 - M35$$

$$\frac{d[mRc]}{dT} = M35 - M36$$

$$[P] = 1 - [PA]$$

$$[API] = 1 - ([A] + [Ap] + [App])$$

$M1 = BR$

$M2 = \eta_1 * R$

$M3 = RJ$

$M4 = \eta_3 * (RJ\hat{}2)$

$M5 = \eta_{-3} * RJD$

$M6 = \eta_2 * RJ$

$M7 = \eta_3 * (RJ\hat{}2)/2$

$M8 = \eta_{-3} * RJD/2$

$M9 = \eta_4 * RJD$

$M10 = \eta_5 * RJD$

$M11 = \eta_{-5} * C$

$M12 = \eta_6 * C$

$M13 = R_1 * \gamma_7 * C * P/(\eta_7 + P)$

$M14 = \gamma_8 * PA/(\eta_8 + PA)$

$M15 = T_1 * (\gamma_{10} * PIA/(\eta_{10} + PIA))$

$M16 = T_1 * (\gamma_9 * R_2 * PA * PI/(\eta_9 + PI))$

$M17 = T_1 * \eta_{11} * A * PIA$

$M18 = T_1 * R_3 * AP1$

$M19 = (T_1/R_3) * (\eta_{11} * A * PIA)$

$M20 = T_1 * AP1$

$M21 = T_1 * (\gamma_{12} * AP1/(\eta_{12} + AP1))$

$M22 = T_1 * (\gamma_{13} * AP2/(\eta_{13} + AP2))$

$M23 = T_1 * (\gamma_{14} * AP2/(\eta_{14} + AP2))$

$M24 = T_1 * (\gamma_{15} * AP3/(\eta_{15} + AP3))$

$M25 = T_2 * (\gamma_{17} * GA)/(\eta_{17} + GA)$

$M26 = T_2 * (\gamma_{16} * R_4 * AP3 * G/(\eta_{16} + G))$

$M27 = T_2 * BG$

$M28 = T_2 * G$

$M29 = T_2 * GA$

$M30 = T_2 * F_1 * \gamma_{19} * GA/(\eta_{19} + GA)$

$M31 = T_2 * \eta_{21} * mGn$

$M32 = T_2 * \eta_{25} * mGc$

$M33 = T_2 * \eta_{24} * mGc$

$M34 = T_2 * F_2 * \gamma_{18} * GA/(\eta_{18} + GA)$

$M35 = T_2 * \eta_{20} * mRn$

$M36 = T_2 * \eta_{23} * mRc$

$M37 = T_2 * \eta_{22} * mRc$

Table 2.3

Table 2-3 Nondimensional parameters for EpoR/GATA-1 model

$$T = t * k_{-1};\ BR = \frac{B_R}{k_{-1}[EpoR_T]_0};\ BG = \frac{B_G}{k_{26}[GATA\text{-}1_T]_0};\ \eta_1 = \frac{k_1[JAK2]}{k_{-1}};$$

$$\eta_2 = \frac{k_2}{k_{-1}};\ \eta_3 = \frac{k_3[EpoR_T]_0}{k_{-1}};\ \eta_{-3} = \frac{2k_{-3}}{k_{-1}};\ \eta_4 = \frac{k_4}{k_{-1}};\ \eta_5 = \frac{k_5[Epo]}{k_{-1}};$$

$$\eta_{-5} = \frac{k_{-5}}{k_{-1}};\ \eta_6 = \frac{k_6}{k_{-1}};\ R_1 = \frac{[EpoR_T]_0}{[PI3K_T]};\ \gamma_7 = \frac{V_7}{k_{-1}};\ \eta_7 = \frac{K_7}{[PI3K_T]};$$

$$\gamma_8 = \frac{V_8}{k_{-1}[PI3K_T]};\ \eta_8 = \frac{K_8}{[PI3K_T]};\ R_2 = \frac{[PI3K_T]}{[PIP2_T]};\ T_1 = \frac{k_{-11}}{k_{-1}};\ \gamma_9 = \frac{V_9}{k_{-11}};$$

$$\eta_9 = \frac{K_9}{[PIP2_T]};\ \gamma_{10} = \frac{V_{10}}{k_{-11}[PIP2_T]};\ \eta_{10} = \frac{K_{10}}{[PIP2_T]};\ \eta_{11} = \frac{k_{11}[AKT_T]}{k_{-11}};$$

$$R_3 = \frac{[AKT_T]}{[PIP2_T]};\ \gamma_{12} = \frac{V_{12}}{k_{-11}[AKT_T]};\ \eta_{12} = \frac{K_{12}}{[AKT_T]};\ \gamma_{13} = \frac{V_{13}}{k_{-11}[AKT_T]};$$

$$\eta_{13} = \frac{K_{13}}{[AKT_T]};\ \gamma_{14} = \frac{V_{14}}{k_{-11}[AKT_T]};\ \eta_{14} = \frac{K_{14}}{[AKT_T]};\ \gamma_{15} = \frac{V_{15}}{k_{-11}[AKT_T]};$$

$$\eta_{15} = \frac{K_{15}}{[AKT_T]};\ T_2 = \frac{k_{26}}{k_{-1}};\ R_4 = \frac{[AKT_T]}{[GATA\text{-}1_T]_0};\ \gamma_{16} = \frac{V_{16}}{k_{26}};\ \eta_{16} = \frac{K_{16}}{[GATA\text{-}1_T]_0};$$

$$\gamma_{17} = \frac{V_{17}}{k_{26}[GATA\text{-}1_T]_0};\ \eta_{17} = \frac{K_{17}}{[GATA\text{-}1_T]_0};\ \gamma_{18} = \frac{V_{18}}{k_{26}[EpoR_T]_0};$$

$$\eta_{18} = \frac{K_{18}}{[GATA\text{-}1_T]_0};\ \gamma_{19} = \frac{V_{19}}{k_{26}[GATA\text{-}1_T]_0};\ \eta_{19} = \frac{K_{19}}{[GATA\text{-}1_T]_0};\ \eta_{20} = \frac{k_{20}}{k_{26}};$$

$$\eta_{21} = \frac{k_{21}}{k_{26}};\ \eta_{22} = \frac{k_{22}}{k_{26}};\ \eta_{23} = \frac{k_{23}}{k_{26}};\ \eta_{24} = \frac{k_{24}}{k_{26}};\ \eta_{25} = \frac{k_{25}}{k_{26}}$$

Table 2.4

Table 2-4 Nondimensional ratios and initial conditions of the reactants

$$[R]=\frac{[EpoR]}{[EpoR_T]_0};\ [RJ]=\frac{[EpoRJ]}{[EpoR_T]_0};\ [RJD]=\frac{[EpoRJD]}{[EpoR_T]_0};[C]=\frac{[EpoRJD^*]}{[EpoR_T]_0};$$

$$[P]=\frac{[PI3K]}{[PI3K_T]};\ [PA]=\frac{[PI3K^*]}{[PI3K_T]};\ [PI]=\frac{[PIP2]}{[PIP2_T]};\ [PIA]=\frac{[PIP3]}{[PIP2_T]};$$

$$[A]=\frac{[AKT]}{[AKT_T]};\ [API]=\frac{[AKT\text{-}PIP3]}{[AKT_T]};\ [Ap]=\frac{[AKTp]}{[AKT_T]};\ [App]=\frac{[AKTpp]}{[AKT_T]};$$

$$[G]=\frac{[GATA\text{-}1]}{[GATA\text{-}1_T]_0};\ [GA]=\frac{[GATA\text{-}1^*]}{[GATA\text{-}1_T]_0};\ [mGn]=\frac{[GATA1mRNAn]}{[GATA\text{-}1_T]_0};$$

$$[mGc]=\frac{[GATA1mRNAc]}{[GATA\text{-}1_T]_0};\ [mRn]=\frac{[EpoRmRNAn]}{[EpoR_T]_0};\ [mRc]=\frac{[EpoRmRNAc]}{[EpoR_T]_0}$$

$$[R]_0=0.01;\ [RJ]_0=0.09;\ [RJD]_0=0.45;\ [C]_0=0;[P]_0=1;\ [PA]_0=0;\ [PI]_0=1;$$

$$[PIA]_0=0;\ [A]_0=1;\ [API]_0=0;\ [Ap]_0=0;\ [App]_0=0;[G]_0=1;\ [GA]_0=0;$$

$$[mGn]_0=0;\ [mGc]_0=0;[mRn]_0=0;[mRc]_0=0$$

Table 2.5

Table 2-5 Values of the kinetic parameters in the EpoR/GATA-1 model

Michaelis-Menten constants ($K_7 – K_{10}$, $K_{12} – K_{19}$) are given in nM; V_7, V_9, and V_{16} are given in s^{-1}; V_8, V_{10}, $V_{12} – V_{15}$, $V_{17} – V_{19}$ are given in nM/s. First- and second-order rate constants are given in s^{-1} and $nM^{-1}s^{-1}$ respectively. Numbers in brackets denote the values given in the literature

Parameter Value	Reference	Parameter Value	Reference
$k_1 = 0.0379$	(1), (2)	$k_{-1} = 0.05$	(1)
$k_3 = 1.3339$	(1), (2)	$k_{-3} = 0.1$	(1)
$k_2 = 5 \times 10^{-5}$ [8.3×10^{-5}]	Estimation, (3)	$k_4 = 5 \times 10^{-5}$ [8.3×10^{-5}]	Estimation, (3)
$k_5 = 0.0083$	(4)	$k_{-5} = 4.83 \times 10^{-4}$	(4)
$V_7 = 0.3$	(5)	$K_7 = 78$	(5)
$V_8 = 46.2$	(5)	$K_8 = 117$	(5)
$V_9 = 16.9$	(6)	$K_9 = 39.1$	(6)
$V_{10} = 17000$	(6)	$K_{10} = 9.02$	(6)
$k_{11} = 507$	(6)	$k_{-11} = 234$	(6)
$V_{12} = 2 \times 10^4$	(7)	$K_{12} = 80000$	(7)
$V_{13} = 1.2198$	(6)	$K_{13} = 4.35$	(6)
$V_{14} = 2 \times 10^4$	(7)	$K_{14} = 80000$	(7)
$V_{15} = 2.7097$ [2.4054]	Estimation, (6)	$K_{15} = 299.62$ [12]	Estimation, (6)
$V_{16} = 0.171$	Estimation	$K_{16} = 0.189$	Estimation
$V_{17} = 0.0368$	Estimation	$K_{17} = 0.0852$	Estimation
$V_{18} = 0.01$	(1)	$K_{18} = 212.12$ [400]	Estimation, (1)
$V_{19} = 0.01$	(1)	$K_{19} = 1092$ [400]	Estimation, (1)
$k_{20} = 0.002$ [0.001]	Estimation, (1)	$k_{21} = 0.0001$ [0.001]	Estimation, (1)
$k_{22} = 0.015$ [0.01]	Estimation, (1)	$k_{23} = 0.0006$	(1)
$k_{24} = 0.0194$ [0.01]	Estimation, (1)	$k_{25} = 0.00046$	(1)
$k_6 = 0.001$	(4)	$k_{26} = 5 \times 10^{-5}$	Estimation
$F_1 = 0.04$	Estimation	$F_2 = 0.123$	Estimation

1. Yamada, S., S. Shiono, A. Joo and A. Yoshimura. 2003. Control mechanism of JAK/STAT signal transduction pathway. *FEBS Lett.* 534:190-196.
2. Lu, X., A.W. Gross and H.F. Lodish. 2006. Active conformation of the erythropoietin receptor: Random and cysteine-scanning mutagenesis of the extracellular juxtamembrane and transmembrane domains. *J. Biol. Chem.* 281:7002-7011.
3. Sarkar, C.A. and D.A. Lauffenburger. 2003. Cell-level pharmacokinetic model of granulocyte colony-stimulating factor: Implications for ligand lifetime and potency in vivo. *Mol. Pharmacol.* 63:147-158.
4. Gross, A. W. and H.F. Lodish. 2006. Cellular trafficking and degradation of erythropoietin and novel erythropoiesis stimulating protein (NESP). *J. Biol. Chem.* 281:2024-2032.
5. Koh, G., H.F. Teong, M.V. Clement, D. Hsu and P.S. Thiagarajan. 2006. A decompositional approach to parameter estimation in pathway modeling: A case study of the akt and MAPK pathways and their crosstalk. *Bioinformatics.* 22:e271-80.
6. Hatakeyama, M., S. Kimura, T. Naka, T. Kawasaki, N. Yumoto, M. Ichikawa, J.H. Kim, K. Saito, M. Saeki, M. Shirouzu, S. Yokoyama and A. Konagaya. 2003. A computational model on the modulation of mitogen-activated protein kinase (MAPK) and akt pathways in heregulin-induced ErbB signalling. *Biochem. J.* 373:451-463.
7. Biondi, R.M., P.C. Cheung, A. Casamayor, M. Deak, R.A. Currie and D.R. Alessi. 2000. Identification of a pocket in the PDK1 kinase domain that interacts with PIF and the C-terminal residues of PKA. *EMBO J.* 19:979-988.

Table 2.6

Table 2-6 Initial concentrations of the reactants before Epo addition

Reactant	Initial Concentration (nM)
EpoR (Epo receptor)	0.0833
JAK2	12
EpoRJ (JAK2 bound Epo receptor)	0.7500
EpoRJD (EpoRJ homodimer)	3.7504
$EpoR_{T0}$ (Total receptor)	8.3341
$PI3K_T$	10
$PIP2_T$	800
AKT_T	10
$GATA\text{-}1_{T0}$	10

All other reactants have values of zero initially.

Table 2.7

Table 2-7 Rate equations for the minimal model

$$\frac{d[Receptor]}{dt} = R_7 - R_1 - R_2 + B_R$$
$$\frac{d[Complex]}{dt} = R_2 - R_3$$
$$\frac{d[InactiveTF]}{dt} = B_G - R_8 - R_4 + R_5 + R_6$$
$$\frac{d[ActiveTF]}{dt} = R_4 - R_5 - R_9$$

$$R_1 : k_1[Receptor]$$
$$R_2 : k_2[Receptor][Ligand] - k_{-2}[Complex]$$
$$R_3 : k_3[Complex]$$
$$R_4 : V_4[Complex][InactiveTF]/(K_4 + [InactiveTF])$$
$$R_5 : V_5[ActiveTF]/(K_5 + [ActiveTF])$$
$$R_6 : F_1V_6[ActiveTF]/(K_6 + [ActiveTF])$$
$$R_7 : F_2V_7[ActiveTF]/(K_7 + [ActiveTF])$$
$$R_8 : k_8[InactiveTF]$$
$$R_9 : k_9[ActiveTF]$$

Table 2.8

Table 2-8 Nondimensional rate equations for the minimal model

$$\frac{d[R]}{dT} = M_9 - M_1 + M_2 + M_3 - M_4$$
$$\frac{d[C]}{dT} = M_1 - M_2 - M_5$$
$$\frac{d[ITF]}{dT} = -M_6 + M_7 + M_8 + M_{10} - M_{11}$$
$$\frac{d[ATF]}{dT} = M_6 - M_7 - M_{12}$$

$$M_1 = \eta_2 * R$$
$$M_2 = \eta_{-2} * C$$
$$M_3 = BR$$
$$M_4 = R$$
$$M_5 = \eta_3 * C$$
$$M_6 = T_1 * \gamma_4 * R_1 * C * ITF / (\eta_4 + ITF)$$
$$M_7 = T_1 * \gamma_5 * ATF / (\eta_5 + ATF)$$
$$M_8 = T_1 * F_1 * \gamma_6 * ATF / (\eta_6 + ATF)$$
$$M_9 = F_2 * \gamma_7 * ATF / (\eta_7 + ATF)$$
$$M_{10} = T_1 * BG$$
$$M_{11} = T_1 * ITF$$
$$M_{12} = T_1 * \eta_9 * ATF$$

Table 2.9

Table 2-9 Nondimensional parameters for the minimal model

$$T = t^{*} k_1;\ BR = \frac{B_R}{k_1[Receptor]_0};\ BG = \frac{B_G}{k_8[InactiveTF]_0};\ \eta_2 = \frac{k_2{}^{*}[Ligand]}{k_1};$$

$$\eta_{-2} = \frac{k_{-2}}{k_1};\ \eta_3 = \frac{k_3}{k_1};\ R_1 = \frac{[Receptor]_0}{[InactiveTF]_0};\ \gamma_4 = \frac{V_4}{k_8};\ \eta_4 = \frac{K_4}{[InactiveTF]_0};$$

$$\gamma_5 = \frac{V_5}{k_8{}^{*}[InactiveTF]_0};\ \eta_5 = \frac{K_5}{[InactiveTF]_0};\ T_1 = \frac{k_8}{k_1};$$

$$\gamma_6 = \frac{V_6}{k_8{}^{*}[InactiveTF]_0};\ \eta_6 = \frac{K_6}{[InactiveTF]_0};$$

$$\gamma_7 = \frac{V_7}{k_1{}^{*}[Receptor]_0};\ \eta_7 = \frac{K_7}{[InactiveTF]_0};\ \eta_9 = \frac{k_9}{k_8}$$

Table 2.10

Table 2-10 Nondimensional ratios and initial conditions for the minimal model

$$[R]=\frac{[Receptor]}{[Receptor]_0};\ [C]=\frac{[Complex]}{[Receptor]_0};\ [ITF]=\frac{[InactiveTF]}{[InactiveTF]_0};\ [ATF]=\frac{[ActiveTF]}{[InactiveTF]_0}$$

$$[R]_0=1;\ [C]_0=0;\ [ITF]_0=1;\ [ATF]_0=0$$

Table 2.11

Table 2-11 Exact solution of the minimal model

$$[ATF]^5 + S_4[ATF]^4 + S_3[ATF]^3 + S_2[ATF]^2 + S_1[ATF] + S_0 = 0$$

$$S_0 = \frac{N_7 * N_{12}}{N_{14}^2};\ S_1 = \frac{N_7 * N_{13} + N_8 * N_{12} - \eta_4 * \eta_6 * N_{10}}{N_{14}^2};$$

$$S_2 = \frac{N_8 * N_{13} + N_7 * N_{14} + N_9 * N_{12} - \eta_4 * \eta_6 * N_{11} - \eta_4 * N_{10}}{N_{14}^2};$$

$$S_3 = \frac{N_9 * N_{13} + N_8 * N_{14} + N_{14} * N_{12} + \eta_4 * \eta_6 * N_{14} - \eta_4 * N_{11}}{N_{14}^2};$$

$$S_4 = \frac{N_9 * N_{14} + N_{14} * N_{13} + \eta_4 * N_{14}}{N_{14}^2}$$

$$N_1 = \left(\frac{\eta_{-2} + \eta_3}{\eta_2}\right) + \eta_3$$

$$N_2 = \frac{F_2 * \gamma_7 + BR}{N_1}$$

$$N_3 = \frac{BR * \eta_7}{N_1}$$

$$N_4 = \gamma_4 * R_1 * N_2$$

$$N_5 = \gamma_4 * R_1 * N_3$$

$$N_6 = \gamma_5 + \eta_9 * \eta_5$$

$$N_7 = N_5 * \eta_5$$

$$N_8 = \eta_5 * N_4 + N_5 - \eta_7 * N_6$$

$$N_9 = N_4 - N_6 - \eta_9 * \eta_7$$

$$N_{10} = \eta_7 * N_6$$

$$N_{11} = N_6 + \eta_7 * \eta_9$$

$$N_{12} = BG * \eta_6$$

$$N_{13} = BG + F_1 * \gamma_6 - \eta_9 * \eta_6$$

$$N_{14} = -\eta_9$$

$$[C] = \frac{N_2 * [ATF] + N_3}{\eta_7 + [ATF]}$$

$$[R] = \left(\frac{\eta_{-2} + \eta_3}{\eta_2}\right) * [C]$$

$$[ITF] = \frac{N_{12} + N_{13} * [ATF] + N_{14} * [ATF]^2}{\eta_6 + [ATF]}$$

Table 2.12

Table 2-12 Values of the kinetic parameters mentioned in the minimal model

$B_R = 5 \times 10^{-4}$ nM/s; $B_G = 5 \times 10^{-4}$ nM/s; $k_1 = 5 \times 10^{-5}$ s^{-1}; $k_2 = 0.0083$ $nM^{-1}s^{-1}$;
$k_{-2} = 4.83 \times 10^{-4}$ s^{-1}; $k_3 = 0.001$ s^{-1}; $V_4 = 0.1715$ s^{-1}; $K_4 = 0.189$ nM; $V_5 = 0.0289$ nM/s;
$K_5 = 0.0571$ nM; $V_6 = 0.01$ nM/s; $K_6 = 400$ nM; $V_7 = 0.01$ nM/s; $K_7 = 400$ nM;
$k_8 = 0.00005$ s^{-1}; $k_9 = 0.00005$ s^{-1}; $F_1 = 20$; $F_2 = 10$
$[Receptor]_0 = 10$ nM; $[Complex]_0 = 0$ nM; $[InactiveTF]_0 = 10$ nM; $[ActiveTF]_0 = 0$ nM

2.5. References

1. Ferrell, J.E. & Xiong, W. Bistability in cell signaling: How to make continuous processes discontinuous, and reversible processes irreversible. *Chaos* **11**, 227-236 (2001).
2. Ferrell, J.E., Jr. & Machleder, E.M. The biochemical basis of an all-or-none cell fate switch in Xenopus oocytes. *Science* **280**, 895-898 (1998).
3. Cross, F.R., Archambault, V., Miller, M. & Klovstad, M. Testing a mathematical model of the yeast cell cycle. *Molecular biology of the cell; Molecular biology of the cell* **13**, 52-70 (2002).
4. Xiong, W. & Ferrell, J.E., Jr. A positive-feedback-based bistable 'memory module' that governs a cell fate decision. *Nature* **426**, 460-465 (2003).
5. Gardner, T.S., Cantor, C.R. & Collins, J.J. Construction of a genetic toggle switch in Escherichia coli. *Nature* **403**, 339-342 (2000).
6. Hasty, J., Pradines, J., Dolnik, M. & Collins, J.J. Noise-based switches and amplifiers for gene expression. *Proceedings of the National Academy of Sciences of the United States of America* **97**, 2075-2080 (2000).
7. Samoilov, M., Plyasunov, S. & Arkin, A.P. Stochastic amplification and signaling in enzymatic futile cycles through noise-induced bistability with oscillations. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 2310-2315 (2005).
8. Rosenfeld, N., Young, J.W., Alon, U., Swain, P.S. & Elowitz, M.B. Gene regulation at the single-cell level. *Science (New York, N.Y.)* **307**, 1962-1965 (2005).
9. Rosenfeld, N. & Alon, U. Response delays and the structure of transcription networks. *Journal of Molecular Biology* **329**, 645-654 (2003).
10. Laslo, P. et al. Multilineage transcriptional priming and determination of alternate hematopoietic cell fates. *Cell; Cell* **126**, 755-766 (2006).
11. Legewie, S., Bluthgen, N., Schafer, R. & Herzel, H. Ultrasensitization: switch-like regulation of cellular signaling by transcriptional induction. *PLoS computational biology* **1**, e54 (2005).
12. Goldbeter, A. & Koshland, D.E., Jr. Sensitivity amplification in biochemical systems. *Quarterly reviews of biophysics; Quarterly reviews of biophysics* **15**, 555-591 (1982).
13. Koshland, D.E., Jr., Goldbeter, A. & Stock, J.B. Amplification and adaptation in regulatory and sensory systems. *Science; Science* **217**, 220-225 (1982).
14. Chang, H.H., Oh, P.Y., Ingber, D.E. & Huang, S. Multistable and multistep dynamics in neutrophil differentiation. *BMC cell biology* **7**, 11 (2006).
15. Metcalf, D. Lineage commitment and maturation in hematopoietic cells: the case for extrinsic regulation. *Blood; Blood* **92**, 345-347; discussion 352 (1998).
16. Enver, T., Heyworth, C.M. & Dexter, T.M. Do stem cells play dice? *Blood; Blood* **92**, 348-351; discussion 352 (1998).
17. Krantz, S.B. Erythropoietin. *Blood; Blood* **77**, 419-434 (1991).

18. D'Andrea, A.D., Fasman, G.D. & Lodish, H.F. Erythropoietin receptor and interleukin-2 receptor beta chain: a new receptor family. *Cell; Cell* **58**, 1023-1024 (1989).
19. Lu, X., Gross, A.W. & Lodish, H.F. Active conformation of the erythropoietin receptor: random and cysteine-scanning mutagenesis of the extracellular juxtamembrane and transmembrane domains. *Journal of Biological Chemistry* **281**, 7002-7011 (2006).
20. Constantinescu, S.N., Ghaffari, S. & Lodish, H.F. The Erythropoietin Receptor: Structure, Activation and Intracellular Signal Transduction. *Trends in endocrinology and metabolism: TEM* **10**, 18-23 (1999).
21. Ghaffari, S. et al. Erythropoiesis in the absence of janus-kinase 2: BCR-ABL induces red cell formation in JAK2(-/-) hematopoietic progenitors. *Blood* **98**, 2948-2957 (2001).
22. Wu, H., Liu, X., Jaenisch, R. & Lodish, H.F. Generation of committed erythroid BFU-E and CFU-E progenitors does not require erythropoietin or the erythropoietin receptor. *Cell* **83**, 59-67 (1995).
23. Evans, T. & Felsenfeld, G. The erythroid-specific transcription factor Eryf1: a new finger protein. *Cell* **58**, 877-885 (1989).
24. Martin, D.I. & Orkin, S.H. Transcriptional activation and DNA binding by the erythroid factor GF-1/NF-E1/Eryf 1. *Genes & development* **4**, 1886-1898 (1990).
25. Welch, J.J. et al. Global regulation of erythroid gene expression by transcription factor GATA-1. *Blood* **104**, 3136-3147 (2004).
26. Orkin, S.H. Globin gene regulation and switching: circa 1990. *Cell* **63**, 665-672 (1990).
27. Mignotte, V., Eleouet, J.F., Raich, N. & Romeo, P.H. Cis- and trans-acting elements involved in the regulation of the erythroid promoter of the human porphobilinogen deaminase gene. *Proceedings of the National Academy of Sciences of the United States of America* **86**, 6548-6552 (1989).
28. Cantor, A.B. & Orkin, S.H. Transcriptional regulation of erythropoiesis: an affair involving multiple partners. *Oncogene* **21**, 3368-3376 (2002).
29. Tang, X.B., Liu, D.P. & Liang, C.C. Regulation of the transcription factor GATA-1 at the gene and protein level. *Cellular and molecular life sciences : CMLS* **58**, 2008-2017 (2001).
30. Chiba, T., Ikawa, Y. & Todokoro, K. GATA-1 transactivates erythropoietin receptor gene, and erythropoietin receptor-mediated signals enhance GATA-1 gene expression. *Nucleic acids research* **19**, 3843-3848 (1991).
31. Kuramochi, S., Ikawa, Y. & Todokoro, K. Characterization of murine erythropoietin receptor genes. *Journal of Molecular Biology; Journal of Molecular Biology* **216**, 567-575 (1990).
32. Zon, L.I., Youssoufian, H., Mather, C., Lodish, H.F. & Orkin, S.H. Activation of the erythropoietin receptor promoter by transcription factor GATA-1. *Proceedings of the National Academy of Sciences of the United States of America* **88**, 10638-10641 (1991).
33. Hannon, R., Evans, T., Felsenfeld, G. & Gould, H. Structure and promoter activity of the gene for the erythroid transcription factor GATA-1. *Proceedings of*

the National Academy of Sciences of the United States of America **88**, 3004-3008 (1991).
34. Tsai, S.F., Strauss, E. & Orkin, S.H. Functional analysis and in vivo footprinting implicate the erythroid transcription factor GATA-1 as a positive regulator of its own promoter. *Genes & development* **5**, 919-931 (1991).
35. Zon, L.I. & Orkin, S.H. Sequence of the human GATA-1 promoter. *Nucleic acids research* **20**, 1812 (1992).
36. Iwasaki, H. et al. GATA-1 converts lymphoid and myelomonocytic progenitors into the megakaryocyte/erythrocyte lineages. *Immunity* **19**, 451-462 (2003).
37. Pevny, L. et al. Erythroid differentiation in chimaeric mice blocked by a targeted mutation in the gene for transcription factor GATA-1. *Nature* **349**, 257-260 (1991).
38. Wickrema, A., Krantz, S.B., Winkelmann, J.C. & Bondurant, M.C. Differentiation and erythropoietin receptor gene expression in human erythroid progenitor cells. *Blood* **80**, 1940-1949 (1992).
39. Fibach, E., Manor, D., Oppenheim, A. & Rachmilewitz, E.A. Proliferation and maturation of human erythroid progenitors in liquid culture. *Blood* **73**, 100-103 (1989).
40. Broudy, V.C., Lin, N., Brice, M., Nakamoto, B. & Papayannopoulou, T. Erythropoietin receptor characteristics on primary human erythroid cells. *Blood* **77**, 2583-2590 (1991).
41. Dalyot, N. et al. Erythropoietin triggers a burst of GATA-1 in normal human erythroid cells differentiating in tissue culture. *Nucleic acids research* **21**, 4031-4037 (1993).
42. Zhao, W., Kitidis, C., Fleming, M.D., Lodish, H.F. & Ghaffari, S. Erythropoietin stimulates phosphorylation and activation of GATA-1 via the PI3-kinase/AKT signaling pathway. *Blood* **107**, 907-915 (2006).
43. Ghaffari, S. et al. AKT induces erythroid-cell maturation of JAK2-deficient fetal liver progenitor cells and is required for Epo regulation of erythroid-cell differentiation. *Blood* **107**, 1888-1891 (2006).
44. Rooke, H.M. & Orkin, S.H. Phosphorylation of Gata1 at serine residues 72, 142, and 310 is not essential for hematopoiesis in vivo. *Blood* **107**, 3527-3530 (2006).
45. Blobel, G.A., Nakajima, T., Eckner, R., Montminy, M. & Orkin, S.H. CREB-binding protein cooperates with transcription factor GATA-1 and is required for erythroid differentiation. *Proceedings of the National Academy of Sciences of the United States of America* **95**, 2061-2066 (1998).
46. Boyes, J., Byfield, P., Nakatani, Y. & Ogryzko, V. Regulation of activity of the transcription factor GATA-1 by acetylation. *Nature* **396**, 594-598 (1998).
47. Hung, H.L., Lau, J., Kim, A.Y., Weiss, M.J. & Blobel, G.A. CREB-Binding protein acetylates hematopoietic transcription factor GATA-1 at functionally important sites. *Molecular and cellular biology* **19**, 3496-3505 (1999).
48. Lamonica, J.M., Vakoc, C.R. & Blobel, G.A. Acetylation of GATA-1 is required for chromatin occupancy. *Blood* **108**, 3736-3738 (2006).
49. Blobel, G.A. CREB-binding protein and p300: molecular integrators of hematopoietic transcription. *Blood* **95**, 745-755 (2000).

50. Letting, D.L., Rakowski, C., Weiss, M.J. & Blobel, G.A. Formation of a tissue-specific histone acetylation pattern by the hematopoietic transcription factor GATA-1. *Molecular and cellular biology* **23**, 1334-1340 (2003).
51. Liu, Y., Denlinger, C.E., Rundall, B.K., Smith, P.W. & Jones, D.R. Suberoylanilide hydroxamic acid induces Akt-mediated phosphorylation of p300, which promotes acetylation and transcriptional activation of RelA/p65. *The Journal of biological chemistry* **281**, 31359-31368 (2006).
52. Vojtek, A.B. et al. Akt regulates basic helix-loop-helix transcription factor-coactivator complex formation and activity during neuronal differentiation. *Molecular and cellular biology* **23**, 4417-4427 (2003).
53. Huang, W.C. & Chen, C.C. Akt phosphorylation of p300 at Ser-1834 is essential for its histone acetyltransferase and transcriptional activity. *Molecular and cellular biology* **25**, 6592-6602 (2005).
54. Roeder, I. & Glauche, I. Towards an understanding of lineage specification in hematopoietic stem cells: a mathematical model for the interaction of transcription factors GATA-1 and PU.1. *Journal of theoretical biology* **241**, 852-865 (2006).
55. Huang, S., Guo, Y.P., May, G. & Enver, T. Bifurcation dynamics in lineage-commitment in bipotent progenitor cells. *Developmental biology* **305**, 695-713 (2007).
56. Ortega, F., Garces, J.L., Mas, F., Kholodenko, B.N. & Cascante, M. Bistability from double phosphorylation in signal transduction. Kinetic and structural requirements. *The FEBS journal* **273**, 3915-3926 (2006).
57. Angeli, D., Ferrell, J.E., Jr. & Sontag, E.D. Detection of multistability, bifurcations, and hysteresis in a large class of biological positive-feedback systems. *Proceedings of the National Academy of Sciences of the United States of America; Proceedings of the National Academy of Sciences of the United States of America* **101**, 1822-1827 (2004).
58. Muzzey, D. & van Oudenaarden, A. When it comes to decisions, myeloid progenitors crave positive feedback. *Cell; Cell* **126**, 650-652 (2006).
59. Lai, K., Robertson, M.J. & Schaffer, D.V. The sonic hedgehog signaling system as a bistable genetic switch. *Biophysical journal* **86**, 2748-2757 (2004).
60. Ferrell, J.E., Jr. Self-perpetuating states in signal transduction: positive feedback, double-negative feedback and bistability. *Current opinion in cell biology* **14**, 140-148 (2002).
61. Huang, C.Y. & Ferrell, J.E., Jr. Ultrasensitivity in the mitogen-activated protein kinase cascade. *Proceedings of the National Academy of Sciences of the United States of America* **93**, 10078-10083 (1996).
62. Bhalla, U.S. & Iyengar, R. Robustness of the bistable behavior of a biological signaling feedback loop. *Chaos* **11**, 221-226 (2001).
63. Savageau, M.A. Michaelis-Menten mechanism reconsidered: implications of fractal kinetics. *Journal of theoretical biology* **176**, 115-124 (1995).
64. Callard, R.E. & Yates, A.J. Immunology and mathematics: crossing the divide. *Immunology* **115**, 21-33 (2005).

65. Hatakeyama, M. et al. A computational model on the modulation of mitogen-activated protein kinase (MAPK) and Akt pathways in heregulin-induced ErbB signalling. *The Biochemical journal* **373**, 451-463 (2003).
66. Alessi, D.R. et al. Characterization of a 3-phosphoinositide-dependent protein kinase which phosphorylates and activates protein kinase Balpha. *Current biology : CB; Current biology : CB* **7**, 261-269 (1997).
67. Andjelkovic, M., Maira, S.M., Cron, P., Parker, P.J. & Hemmings, B.A. Domain swapping used to investigate the mechanism of protein kinase B regulation by 3-phosphoinositide-dependent protein kinase 1 and Ser473 kinase. *Molecular and cellular biology; Molecular and cellular biology* **19**, 5061-5072 (1999).
68. Smolen, P., Baxter, D.A. & Byrne, J.H. Frequency selectivity, multistability, and oscillations emerge from models of genetic regulatory systems. *The American Journal of Physiology* **274**, C531-542 (1998).
69. Yamada, S., Shiono, S., Joo, A. & Yoshimura, A. Control mechanism of JAK/STAT signal transduction pathway. *FEBS letters* **534**, 190-196 (2003).
70. Schmidt, H. & Jirstrand, M. Systems Biology Toolbox for MATLAB: a computational platform for research in systems biology. *Bioinformatics (Oxford, England)* **22**, 514-515 (2006).
71. Keller, M.A. et al. Transcriptional regulatory network analysis of developing human erythroid progenitors reveals patterns of coregulation and potential transcriptional regulators. *Physiological genomics* **28**, 114-128 (2006).
72. Gross, A.W. & Lodish, H.F. Cellular trafficking and degradation of erythropoietin and novel erythropoiesis stimulating protein (NESP). *Journal of Biological Chemistry* **281**, 2024-2032 (2006).
73. Huang, S., Eichler, G., Bar-Yam, Y. & Ingber, D.E. Cell fates as high-dimensional attractor states of a complex gene regulatory network. *Physical Review Letters* **94**, 128701 (2005).
74. Battle, T.E. & Frank, D.A. The role of STATs in apoptosis. *Current Molecular Medicine* **2**, 381-392 (2002).
75. Yu, Y.L. et al. MAPK-mediated phosphorylation of GATA-1 promotes Bcl-XL expression and cell survival. *The Journal of biological chemistry* **280**, 29533-29542 (2005).
76. Richmond, T.D., Chohan, M. & Barber, D.L. Turning cells red: signal transduction mediated by erythropoietin. *Trends in cell biology* **15**, 146-155 (2005).
77. Hernandez-Hernandez, A. et al. Acetylation and MAPK phosphorylation cooperate to regulate the degradation of active GATA-1. *The EMBO journal* **25**, 3264-3274 (2006).
78. Starr, R. et al. A family of cytokine-inducible inhibitors of signalling. *Nature* **387**, 917-921 (1997).
79. Endo, T.A. et al. A new protein containing an SH2 domain that inhibits JAK kinases. *Nature* **387**, 921-924 (1997).
80. Hofer, T., Nathansen, H., Lohning, M., Radbruch, A. & Heinrich, R. GATA-3 transcriptional imprinting in Th2 lymphocytes: a mathematical model. *Proceedings of the National Academy of Sciences of the United States of America* **99**, 9364-9368 (2002).

81. Nalefski, E.A., Nebelitsky, E., Lloyd, J.A. & Gullans, S.R. Single-molecule detection of transcription factor binding to DNA in real time: specificity, equilibrium, and kinetic parameters. *Biochemistry* **45**, 13794-13806 (2006).
82. Omichinski, J.G. et al. NMR structure of a specific DNA complex of Zn-containing DNA binding domain of GATA-1. *Science* **261**, 438-446 (1993).
83. Savageau, M.A. in Journal of Molecular Recognition : JMR, Vol. 6 149-157ENGLAND; 1993).
84. Paulsson, J., Berg, O.G. & Ehrenberg, M. Stochastic focusing: fluctuation-enhanced sensitivity of intracellular regulation. *Proceedings of the National Academy of Sciences of the United States of America* **97**, 7148-7153 (2000).

Chapter 3

Heterogeneity in the expression of EpoR and GATA1 is positively correlated during erythropoiesis

3.1. Introduction

Cell differentiation is the process during which a progenitor cell commits towards a particular lineage and remodels its transcriptome towards that lineage-restricted state. The process of lineage-commitment is believed to precede the drastic phenotype change during differentiation. During commitment, the progenitor cell makes a fate-decision between proliferation and one of the several available differentiation states.

Lineage-specific cytokines and transcription factors are shown to be absolutely necessary to robustly reach mature cell states[1-7]. Cytokines are cell-extrinsic factors that bind to their cognate receptors on the progenitor cell surface and promote signals that help in the survival, proliferation and differentiation towards a specific lineage. Lineage-specific transcription factors are cell-intrinsic factors that transcriptionally regulate most of the lineage-specific genes. It has been heavily debated whether the cell-fate commitment in hematopoiesis is an instructive or a stochastic process[1, 8, 9]. According to

the instructive theory, external cues such as cytokines regulate the process of cell commitment and bias the progenitor cell to a specific lineage, whereas the stochastic theory suggests that each progenitor cell has been pre-committed by the differential expression of lineage-specific transcription factors and cytokines merely provide survival and proliferation signals during differentiation. Recent groundbreaking studies using single-cell imaging techniques have unambiguously shown that cytokines can instruct cells during differentiation[10]. However, the extent of this instruction and the topological links between lineage-specific cytokine receptor signaling and lineage-specific transcription factor upregulation are still unclear.

Progenitor cells tend to express low levels of all of the receptors and transcription factors that are specific to the various lineages to which they can potentially commit, a process known as transcriptional priming[11]. During differentiation, a specific transcription factor gets upregulated and, through cross-antagonism, downregulates other lineage-specific transcription factors[2]. Due to this, the heterogeneity in the expression of the transcription factors is shown to be directly correlated with their lineage choices[3]. However, correlations between lineage-specific receptor expression and the corresponding transcription factor or lineage specification have not been studied. Here, we study these previously unexplored correlations using erythropoiesis as a model system.

Erythropoietin (Epo) is a cytokine that binds to the Epo receptor (EpoR) present on erythroid progenitor cells and provides signals of survival, proliferation and commitment towards erythropoiesis[12-14]. Of the several erythroid-specific transcription factors, GATA1 is believed to be the master regulator of erythropoiesis as it binds to

response elements upstream of most of the enzymes necessary for hemoglobin synthesis and upregulates their expression[15-17]. Previous studies have shown that GATA1 can autoregulate its own transcription as well as regulate the transcription of EpoR[18-24]. Also, recent biochemical evidence shows that EpoR signaling can directly result in post-transcriptional modification of GATA1 (acetylation and/or phosphorylation) and enhance its DNA binding and transactivation capabilities[25-32].

Based upon these experimental observations, we previously developed a minimal model to analyze the topological connections between GATA1 and EpoR[33]. The minimal topology for accumulation of active GATA1 contains two positive feedback loops: an autofeedback loop that produces inactive GATA1 and an Epo-dependent, EpoR-mediated feedback loop that activates GATA1 (Figure 3.1). Dynamical systems modeling of this topology suggests that the presence of these two feedback loops can not only offer a direct positive correlation in the upregulation of GATA1 and EpoR, but can also provide cellular memory with respect to external cues, thereby establishing robustness in lineage commitment during differentiation.

In this study, using an erythrocyte progenitor cell line, we experimentally demonstrate that hemoglobin synthesis is highly switch-like in response to Epo and cells undergoing lineage commitment possess memory of earlier Epo signals. We show that total GATA1 and cell-surface EpoR are indeed co-upregulated and follow a synchronous expression pattern during differentiation, confirming the presence of autofeedback and receptor-mediated positive feedback loops, respectively. Our results also show that the heterogeneity in EpoR expression during early-stage lineage commitment is positively correlated with GATA1 expression, reinforcing the topological connection between

lineage-specific receptor and transcription factor even in the presence of significant noise. Moreover, due to this positive association and its consequent effect on cell phenotype, we show that differential expression of EpoR in progenitor cells can be a marker for erythrocyte commitment, with cells expressing high levels of EpoR having significantly accelerated kinetics of differentiation. Our study highlights the interplay between cell-extrinsic and cell-intrinsic factors in determining lineage commitment. These findings may be of general applicability to other lineage-specific receptor and transcription-factor pairs during differentiation, particularly in hematopoiesis.

3.2. Materials and Methods

3.2.1. Cell culture

UT-7/GM cells were kindly provided by Dr. Norio Komatsu (Juntendo University School of Medicine) and Dr. Kenneth Kaushansky (University of California, San Diego). Recombinant human GM-CSF was purchased from Peprotech. UT-7/GM cells were maintained in Iscove's modified Dulbecco's medium (IMDM, Invitrogen) supplemented with 10% fetal bovine serum (FBS, HyClone) and 1 ng/ml GM-CSF. Cell viability was quantified through Trypan blue (Mediatech) dye exclusion, observed under light microscope (10x magnification).

3.2.2. Erythroid differentiation

Recombinant human erythropoietin (Epo) was purchased from AppliChem. To induce erythroid differentiation, UT-7/GM cells were growth factor starved for 16 hours and then cultured in IMDM containing 10% FBS and 1U/ml Epo.

3.2.3. Dianisidine staining

Hemoglobin positive cells were identified by incubating cells in serum-free IMDM medium containing 0.2% 3,3'-dimethoxybenzidine, fast blue B (dianisidine, Sigma), 0.3% acetic acid and 0.3% hydrogen peroxide for 20 minutes at room temperature. Cells were observed under light microscope (40x magnification) and the cells that stained dark blue were marked positive for hemoglobin.

3.2.4. Analysis of surface EpoR expression by immunofluorescence

Approximately $1x10^5$ cells were extracted from culture and centrifuged to remove the medium. Cells were then washed twice with 1xPBS containing 0.5% bovine serum albumin (PBS-B). Cells were resuspended in 30 μl of PBS-B and incubated with 1 μg of human IgG antibody (R&D Systems) for 15 minutes at room temperature. Monoclonal anti-human EpoR antibody conjugated with phycoerythrin (150 ng) was added to the sample and incubated at 4°C for 45 minutes. Cells were then washed again with PBS-B and resuspended in 200 μl of PBS-B. The cells were then analyzed for EpoR expression using a Guava Flow Cytometer.

3.2.5. Analysis of GATA1 expression by western blotting

Approximately $1x10^7$ cells were extracted and centrifuged to remove culture media. Cells were then washed with 1xPBS and incubated in 500 μl lysis buffer (Invitrogen) supplemented with PMSF, protease inhibitors (Sigma) and phosphatase inhibitors (Sigma) for 30 minutes at 4°C. The samples were centrifuged at 13000 g for 15 minutes at 4°C. Total protein concentrations in the samples were estimated by BCA protein assay (Pierce) according to the manufacturer's instructions. After the addition of 4x reducing agent (Invitrogen) and 10x sample buffer (Invitrogen), 20 μg of protein sample was boiled at 95°C for 5 minutes, separated by SDS-polyacrylamide gel electrophoresis (160 volts, 95 minutes, 4°C), transferred onto nitrocellulose membrane (30 volts, 60 minutes, 4°C) and blocked with Odyssey blocking buffer (LI-COR, 60 minutes, room temperature). The membrane was incubated with rat monoclonal GATA-1 (N6) antibody from Santa Cruz (sc-265, dilution 1:1000) and then washed five times with 1xPBS containing 0.2% Tween-20 (Bio-Rad). Lastly, the membrane was treated with rabbit anti-rat IRdye-800 antibody from Rockland Immunochemicals (612-431-026, dilution 1:15000) for 45 minutes at room temperature. The protein signals were detected using a LI-COR infrared imager and GATA1 bands were quantified with the analysis software provided by the manufacturer.

3.2.6. Cell sorting based on surface EpoR expression

Approximately $3x10^7$ cells were extracted from culture, centrifuged to remove medium and washed with PBS-B. Cells were incubated with monoclonal anti-human EpoR antibody conjugated with phycoerythrin and prepared for sorting as described in the

section 3.2.4. The EpoR-antibody bound cells were resuspended in PBS-B and sorted for EpoR expression in a Becton Dickinson Digital Vantage cell sorter. EpoR (low, med, high) cells were classified as being the bottom, middle and top 5% of the unsorted cells, respectively. The sorted cells were either re-cultured for differentiation experiments or lysed to extract proteins for GATA1 western blot experiments.

3.3. Results and Discussion

3.3.1. Epo-induced lineage commitment and differentiation

UT-7/GM cells, which are derived from a parental human megakaryoblastic leukemia cell line (UT-7), have been previously shown to be a successful model system to study erythrocyte differentiation[34, 35]. UT-7/GM is a growth factor dependent cell line and requires granulocyte-macrophage colony stimulating factor (GM-CSF) for its survival and proliferation[36]. The cells remain undifferentiated in the presence of GM-CSF and erythrocyte differentiation can be induced by replacing GM-CSF with Epo. The differentiation process closely mirrors that of other human erythrocyte progenitor model systems. One reason that the dynamics of receptor and transcription factor upregulation may have been overlooked to date is that a large number of studies have been performed in mouse models that possess fast differentiation kinetics (~ 2 days). By contrast, human UT-7/GM cells require two weeks in culture with 1U/ml Epo to reach their differentiation maximum (as assessed by hemoglobin staining). Figure 3.2A shows UT-7/GM cells grown in GM-CSF (left panel) and Epo (right panel) for 14 days and stained with dianisidine, a dye that indicates the presence of hemoglobin (dark blue spots). As seen in

this figure, Epo-induced cells show hemoglobin synthesis compared to the uninduced cells.

We performed an Epo-dose response experiment to study the relationship between Epo induction and hemoglobin production (Figure 3.2B). UT-7/GM cells cultured in GM-CSF were growth-factor starved for 16 hours, and then cultured in media containing different concentrations of Epo. The cells were stained with dianisidine, after 14 days of culture with Epo, to determine the percentage of hemoglobin positive cells. The plot shows an ultrasensitive, or switch-like, response, where a 10-fold increase in Epo level (from 0.001U to 0.01U) switches most of the undifferentiated cells to the differentiated state. This response is much steeper than a graded, hyperbolic response, which requires an 81-fold increase in stimulus to attain a change in response from 10% to 90%. Increasing the stimulus over 0.01U did not show significant improvement in differentiation efficiency. This switch-like response is also consistent with the presence of positive feedback loops in the differentiation circuitry.

We quantified the percentage of differentiated cells at different time points during the two-week period with Epo-induction (1U, 0.01U, 0.001U) (Figure 3.2C). The plot suggests that for Epo concentrations of 1U and 0.01U, the cells remained mostly undifferentiated for the first 4 days and then showed a linear increase in differentiation efficiency from day 4 to day 10. This is particular interesting when we compare these results to the cell viability data for the same time points during differentiation (Figure 3.2D). The profile shows an initial decrease in cell viability, possibly due to the selection of cells expressing a minimum threshold level of EpoR since Epo signaling is absolutely necessary for cell survival. The cells appear to recover by day 4, suggesting that lineage-

selection and differentiation are independent events, with selection precedes differentiation. Compared to cultures with 1U Epo, cultures with 0.01U Epo show decreased viability during initial selection as well as a lower percentage of hemoglobin-positive cells. Cells cultured at 0.001U show a marked decrease in viability during the initial selection and fail to recover. Due to the poor viability, these cultures showed a significantly lower percentage of differentiated cells during the two-week period.

There seems to be a threshold concentration of Epo (0.01U higher bound) necessary to achieve differentiation of large percentage of cells (approximately 75%). Since lineage commitment should occur during the initial stages of differentiation, we devised an experiment to pretreat the cells with a high concentration of Epo (0.01U for 3 or 6 days) and then reduce the Epo concentration to a sub-threshold level for the remainder of the differentiation period. Pretreated cells achieved a significantly higher percentage of differentiation when compared to control cells grown at 0.001U Epo (Figure 3.2E). Cells pretreated for 6 days differentiated better than the 3 day pretreated cells, but their efficiency was significantly lower than the culture grown at 0.01U. The corresponding cell viability assay showed that pretreated cells achieved a greater percentage viability compared to the control cells grown at 0.001U. This experiment suggests that the initial pretreatment commits cells towards differentiation and that lowering of Epo concentration did not affect a large percentage of the committed cells as they retained the memory of high Epo concentration and continued to differentiate.

3.3.2. Synchronous upregulation of GATA1 and EpoR during differentiation

The memory seen during lineage commitment as well as the existence of a switch-like response (from section 3.3.1) can be due to the presence of the two positive feedback

loops between GATA1 and EpoR, induced by Epo. It is well recognized that lineage-specific transcription factor expression gets upregulated during differentiation due to the need to upregulate all lineage-specific genes and to downregulate other lineage-specific transcription factors. We confirmed the presence of the autofeedback look by showing the upregulation of total GATA1 during Epo-induced differentiation of UT-7/GM cells through quantitative western blots (Figure 3.3A). We also quantified cell-surface EpoR levels from viable cells during differentiation by flow cytometry (Figure 3.3B). EpoR levels were expected to decrease with Epo-induction since Epo/EpoR complex endocytosis is much faster than EpoR constitutive endocytosis. Indeed, we noticed an initial decrease in EpoR expression in the first few hours after Epo induction and the receptor levels quickly reached steady state (data not shown); however, after 24 hours, the EpoR levels slowly began to climb and, by day 8, reached a maximum of approximately 12-fold that of the initial steady-state. This shows that EpoR is significantly upregulated, confirming the presence of an Epo-induced positive feedback loop through the receptor during differentiation. This result is interesting, because as seen from the viability plot (Figure 3.2D), the initial selection happens in the first few days of differentiation, when the EpoR levels are still low. After selection, the cells should express enough EpoR to provide survival signals and do not have the need to increase EpoR expression; hence, the upregulation of EpoR should be an event necessary for commitment rather than selection.

3.3.3. Heterogeneity in EpoR and GATA1 expression is positively correlated in differentiating cells

After establishing the presence of the two positive feedback loops during Epo-induced differentiation, we wanted to analyze whether EpoR and GATA1 expression are correlated during differentiation. UT-7/GM cells were induced with Epo and the heterogeneity in EpoR levels in viable cells was measured by flow cytometry (Figure 3.4A). Histograms show significant broadening on day 2 when compared with the day-0, day-4 or day-7 populations. This deviation from unimodal distribution in EpoR expression seems to be driven by Epo and possibly through the two positive feedback loops that regulate GATA1 and EpoR.

In section 3.3.2, the average expression levels of EpoR and GATA1 were shown to be positively correlated by synchronous upregulation during differentiation. We wanted to further analyze whether this correlation existed even within subpopulations of the same culture on day 2, when a broadening of the EpoR expression histogram was observed (Figure 3.4A). Cells cultured with Epo for two days were sorted into three different populations based on the surface EpoR levels – $EpoR^{low}$, bottom 5%; $EpoR^{med}$, middle 5%; and $EpoR^{high}$, top 5% of the unsorted cells (Figure 3.4B). The sorted populations were lysed and the total GATA1 level in each of the populations was determined through quantitative western blotting. As seen in Figure 3.4C, the sorted fractions also showed a positive correlation between EpoR and GATA1. This suggests that EpoR and GATA1 may be part of the same biochemical network, as the change in GATA1 levels due to the heterogeneity in its expression reflects on EpoR expression. This shows that cell-intrinsic noise can be transferred to cell-extrinsic components and,

conversely and perhaps more importantly, extrinsic factors can potentially regulate or bias intrinsic factors during lineage commitment.

3.3.4. Differential expression of EpoR as a marker for erythrocyte progenitor commitment

Erythroid progenitors that have committed to differentiate with Epo induction start to express hemoglobin-synthesizing genes. Synthesis of hemoglobin seems to be preceded by the upregulation of GATA1 and EpoR. Since EpoR and GATA1 expression are correlated as shown in Figures 3.3 and 3.4, selecting progenitor cells based on differential EpoR levels should alter the kinetics of commitment. We sorted day 2 Epo-induced cells into $EpoR^{low}$, $EpoR^{med}$ and $EpoR^{high}$ populations as in section 3.3.3. The sorted populations were immediately transferred to Epo-containing media to continue the differentiation program. The surface EpoR level in all three populations was quantified on day 4 and day 7 and compared to the unsorted population (Figure 3.5A). $EpoR^{low}$ cells could not be propagated, likely due to insufficient EpoR signaling for survival. Comparison of EpoR levels revealed no significant change in the mean between the unsorted population and the $EpoR^{med}$ population, whereas the receptor expression level in the $EpoR^{high}$ population was significantly higher than in the other two samples. From the dianisidine assay, $EpoR^{high}$ and $EpoR^{med}$ populations showed approximately 4.0-fold and 2.5-fold increases in positive cells, respectively, compared to the unsorted population on day 4 (Figure 3.5B). On day 7, $EpoR^{high}$ and $EpoR^{med}$ populations showed approximately 2.0-fold and 1.6-fold increase in positive cells, respectively, compared to the unsorted population (Figure 3.5B). This result suggests that progenitor cells expressing high levels of EpoR are intrinsically primed towards erythrocyte commitment and possess enhanced

kinetics of differentiation. Hence, due to the positive correlation between GATA1 and EpoR, differential expression of EpoR in progenitor cells can serve as a marker for sorting and recovering erythrocyte-committed cells.

Figure 3.1

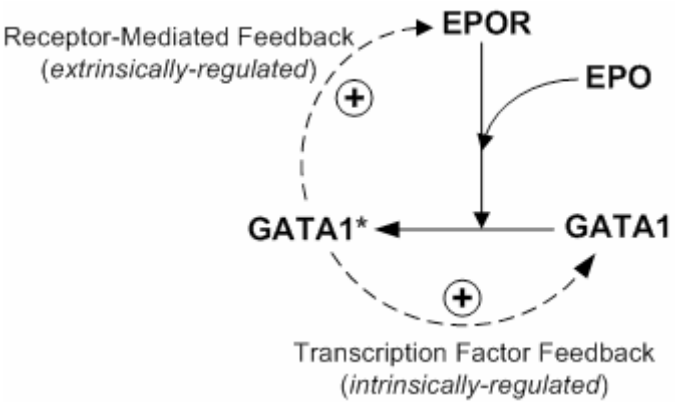


Figure 3-1 Positive feedback loops connecting EpoR and GATA1

A model showing the topological connections between EpoR and GATA1: Epo binds to EpoR to activate signaling pathways that can post-transcriptionally activate GATA1. GATA1 can upregulate its own inactive form through an autofeedback loops as well as upregulate the expression of EpoR, thereby enhancing its own activation. The autofeedback loop is intrinsically regulated, whereas Epo extrinsically regulates the receptor-mediated feedback loop.

Figure 3.2

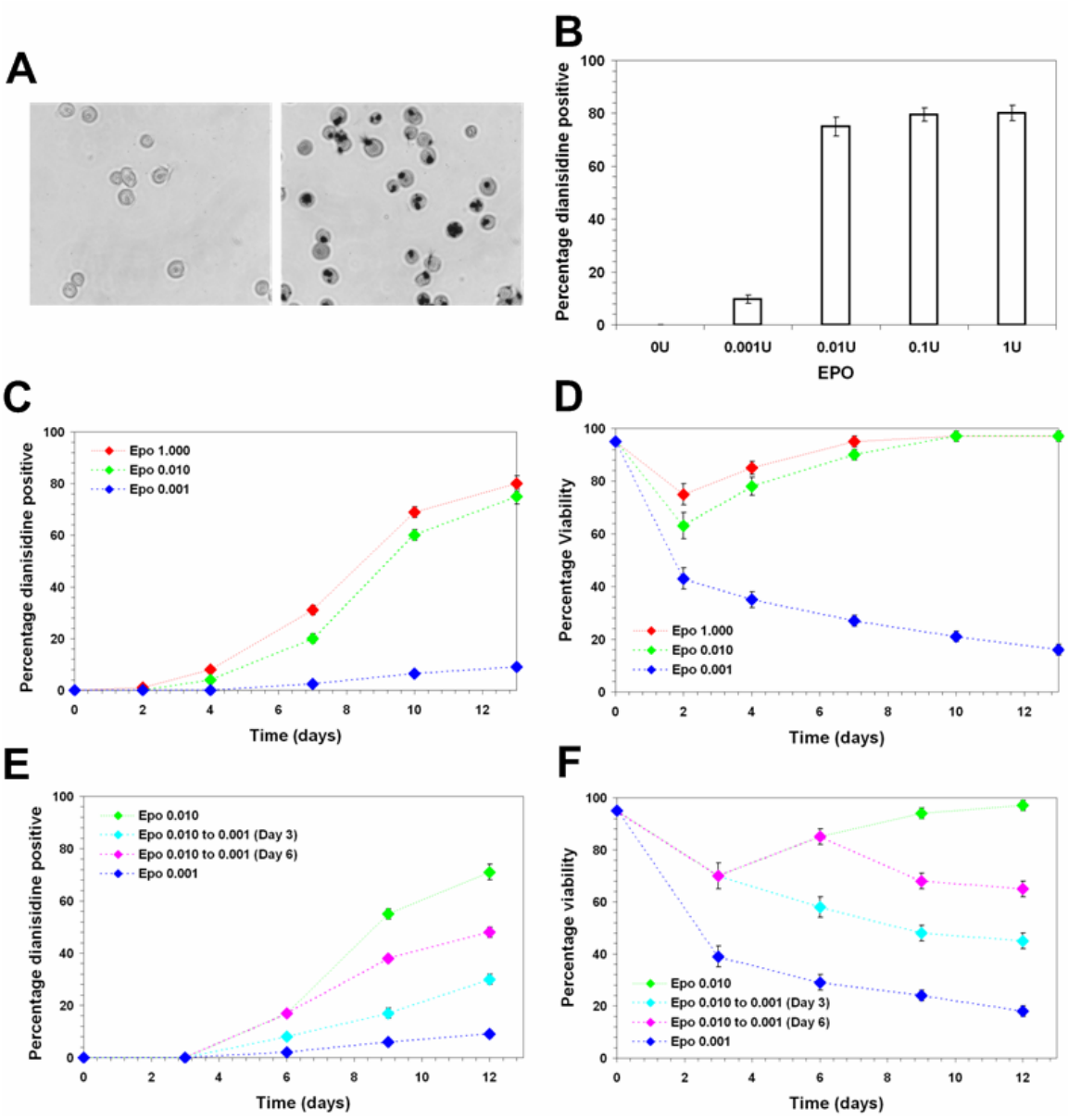


Figure 3-2 Epo-induced differentiation of UT-7/GM cells

(A) Dianisidine staining: Control cells growing in GM-CSF showing no synthesis of hemoglobin (left panel), Cells induced with Epo (1U/ml) for 14 days showing the presence of hemoglobin through the dark spots (right panel). (B) Epo dose response curve: Cells were growth factor starved and treated with different concentrations of Epo and stained with dianisidine after 14 days. Percentage of dianisidine positive cells (presence of hemoglobin) is plotted against Epo concentration. The plot shows an ultrasensitive, or switch-like, response in differentiation to Epo concentration. (C) Kinetic of differentiation: Cells were growth factor starved and treated with 1U/ml, 0.01U/ml and 0.001U/ml of Epo and the percentage of dianisidine positive cells were quantified at different time points during the differentiation period. High percentage of cells with Epo 1U/ml and 0.01 U/ml differentiated by day 13, whereas cells grown with Epo 0.001 U/ml remained largely undifferentiated. (D) Cell viability during differentiation: During the differentiation kinetics experiment in C, differentiating cells were also checked for percentage viability through Trypan blue dye exclusion assay. Cell viability decreases sharply in the first two days (~75% for Epo 1U/ml and ~65% for Epo 0.01U/ml) and then recovers to initial levels (~95%). Cells grown with Epo 0.001 U/ml showed sustained decrease in viability and did not recover. (E) Pretreatment experiment: Cells were grown with Epo 0.01 U/ml for 3 or 6 days and then switched to 0.001 U/ml Epo. Cells grown all through with 0.01 or 0.001 U/ml of Epo are shown as controls. (F) Cell viability during the pretreatment assay: For the data point in E, cells were also analyzed to quantify the percentage viability.

Figure 3.3

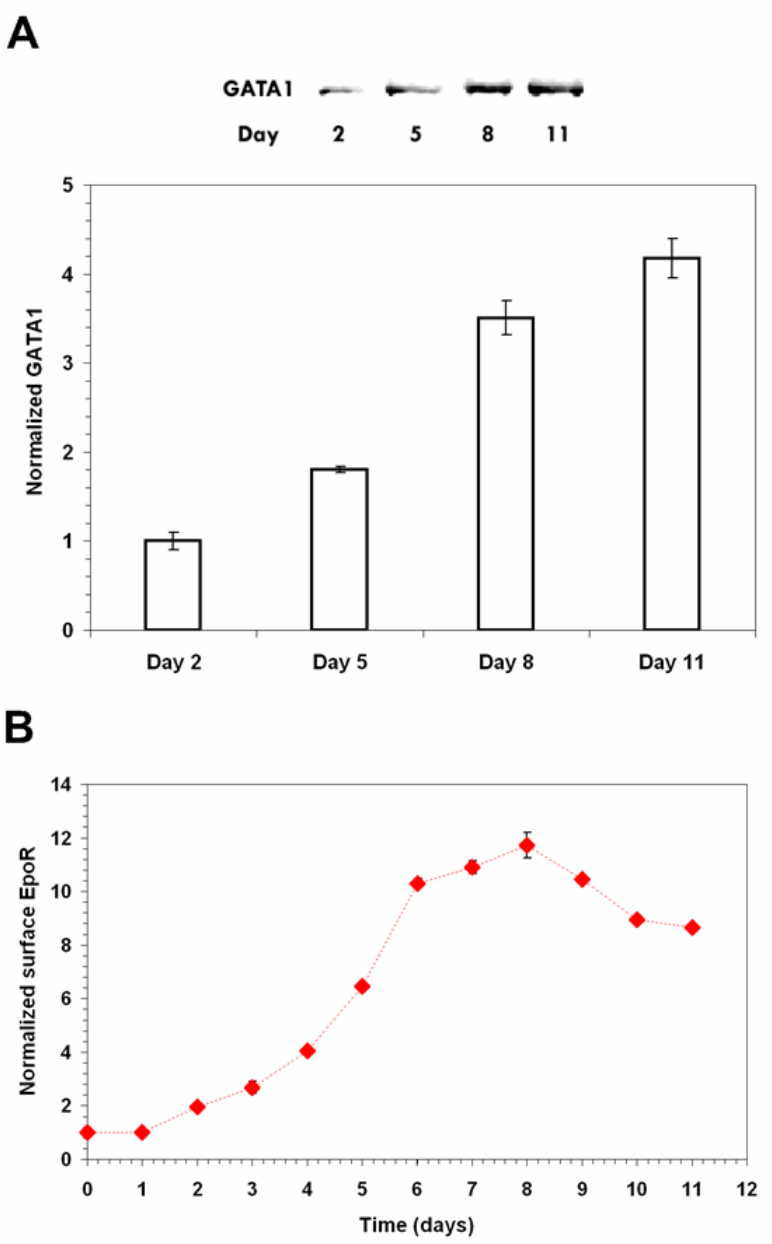


Figure 3-3 Synchronous upregulation of GATA1 and EpoR

(A) Upregulation of total GATA1: Cell lysates collected at different time points during differentiation were blotted for total GATA1. The quantified protein bands show an upregulation in total GATA1 levels. (B) Upregulation of surface EpoR: Cells from differentiating cultures were treated with monoclonal antibody to EpoR conjugated with phycoerythrin and surface EpoR levels were quantified by flow cytometry. Surface EpoR levels show upregulation during Epo-induced differentiation.

Figure 3.4

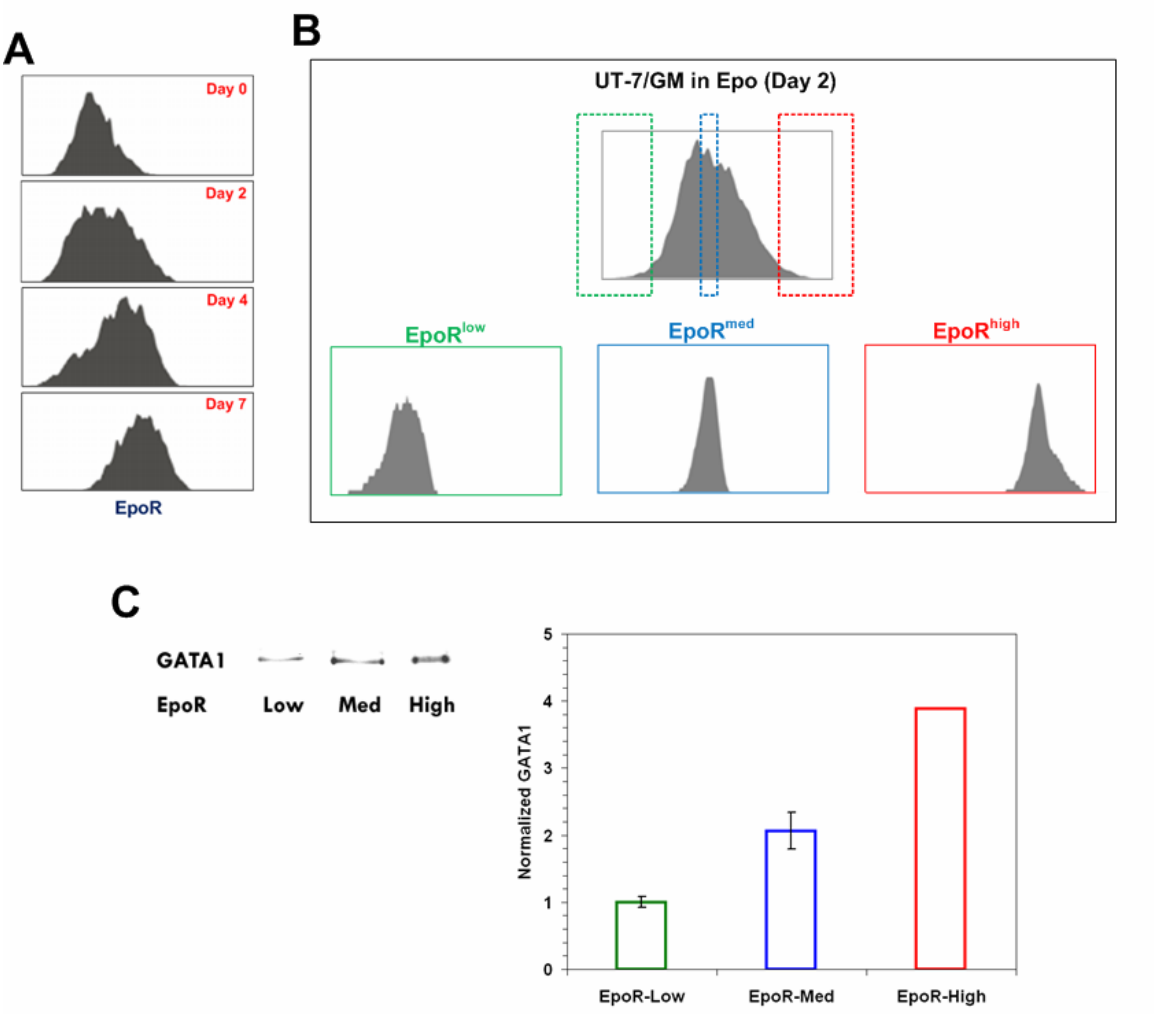


Figure 3-4 Heterogeneity in EpoR and GATA1 expression is positively correlated

(A) Histograms showing the heterogeneity in surface EpoR expression at different time points (days 0, 2, 4 and 7) during differentiation. (B) Cell sorting based on surface EpoR expression: Cells grown in Epo for two days were sorted into three distinct populations: bottom 5% ($EpoR^{low}$), middle 5% ($EpoR^{med}$) and top 5% ($EpoR^{high}$). (B) Total GATA1 levels in each of the three sorted populations were measured by quantitative western blots. The noise in EpoR expression is correlated with GATA1 levels.

Figure 3.5

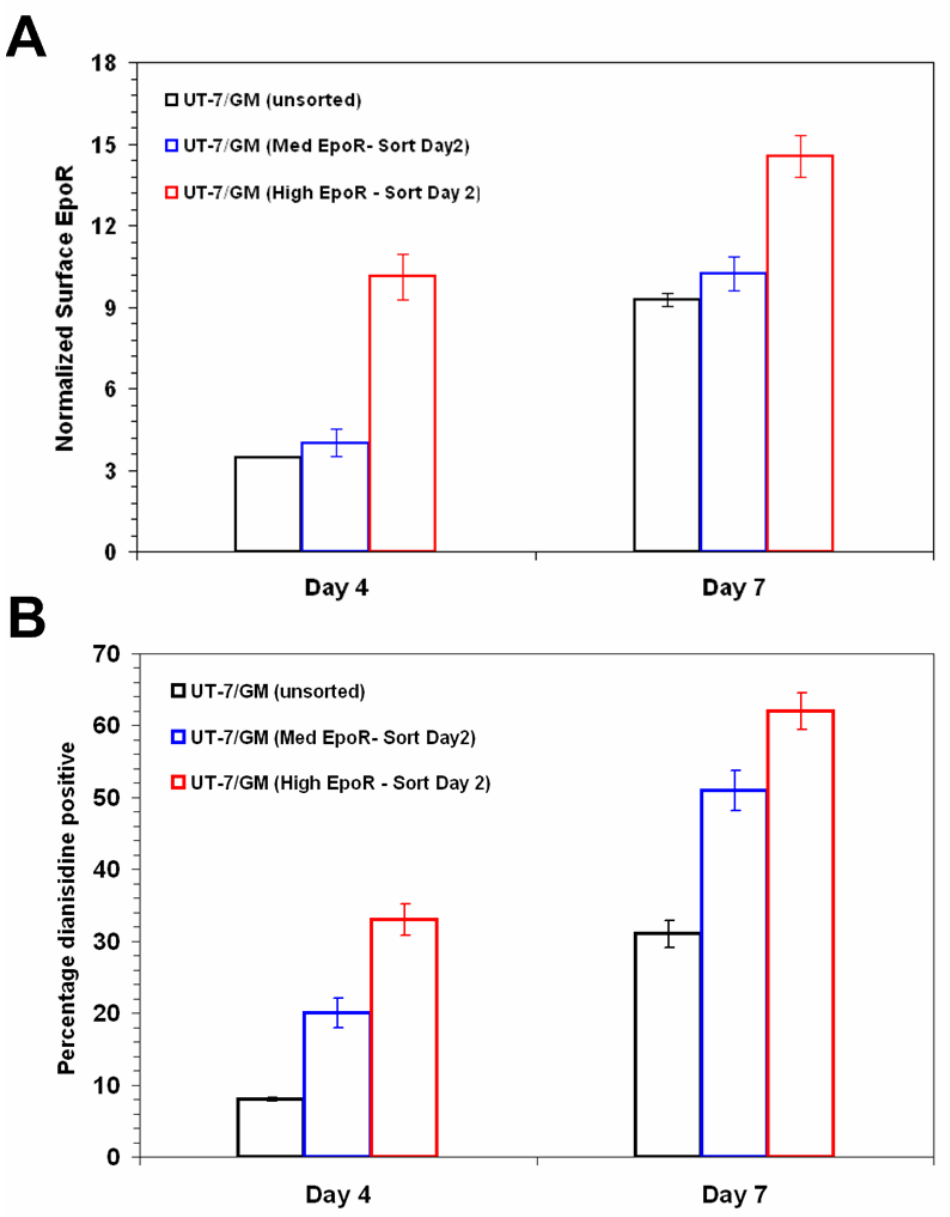


Figure 3-5 Differential expression of EpoR as a marker for progenitor commitment

(A) Surface EpoR expression during differentiation: Sorted populations ($EpoR^{low}$, $EpoR^{med}$ and $EpoR^{high}$) from Figure 3.4 were immediately re-cultured with Epo (1U/ml). $EpoR^{low}$ cells could not be propagated, likely due to the lack of survival signals from Epo. Surface EpoR levels in $EpoR^{med}$ and $EpoR^{high}$ cells were compared to the unsorted cells on day 4 and day 7. (B) Hemoglobin positive cells: Percentage of differentiation on day 4 and day 7 from the sorted cells ($EpoR^{med}$ and $EpoR^{high}$) were quantified by dianisidine staining and compared to the unsorted cells.

3.4. References

1. Metcalf, D. Hematopoietic cytokines. *Blood* **111**, 485-491 (2008).
2. Cantor, A.B. & Orkin, S.H. Hematopoietic development: a balancing act. *Current opinion in genetics & development; Current opinion in genetics & development* **11**, 513-519 (2001).
3. Eckfeldt, C.E., Mendenhall, E.M. & Verfaillie, C.M. The molecular repertoire of the 'almighty' stem cell. *Nature reviews.Molecular cell biology* **6**, 726-737 (2005).
4. Akashi, K. et al. Lymphoid development from stem cells and the common lymphocyte progenitors. *Cold Spring Harbor symposia on quantitative biology* **64**, 1-12 (1999).
5. Akashi, K., Traver, D., Miyamoto, T. & Weissman, I.L. A clonogenic common myeloid progenitor that gives rise to all myeloid lineages. *Nature* **404**, 193-197 (2000).
6. Forsberg, E.C., Bhattacharya, D. & Weissman, I.L. Hematopoietic stem cells: expression profiling and beyond. *Stem cell reviews* **2**, 23-30 (2006).
7. Kondo, M. et al. Biology of hematopoietic stem cells and progenitors: implications for clinical application. *Annual Review of Immunology* **21**, 759-806 (2003).
8. Enver, T., Heyworth, C.M. & Dexter, T.M. Do stem cells play dice? *Blood; Blood* **92**, 348-351; discussion 352 (1998).
9. Metcalf, D. Lineage commitment and maturation in hematopoietic cells: the case for extrinsic regulation. *Blood; Blood* **92**, 345-347; discussion 352 (1998).
10. Rieger, M., Hoppe, P., Smejkal, B., Eitelhuber, A. & Schroeder, T. Hematopoietic cytokines can instruct lineage choice. *Science* **325**, 217-218 (2009).
11. Laslo, P. et al. Multilineage transcriptional priming and determination of alternate hematopoietic cell fates. *Cell; Cell* **126**, 755-766 (2006).
12. Krantz, S.B. Erythropoietin. *Blood; Blood* **77**, 419-434 (1991).
13. Ghaffari, S. et al. Erythropoiesis in the absence of janus-kinase 2: BCR-ABL induces red cell formation in JAK2(-/-) hematopoietic progenitors. *Blood* **98**, 2948-2957 (2001).
14. Wu, H., Liu, X., Jaenisch, R. & Lodish, H.F. Generation of committed erythroid BFU-E and CFU-E progenitors does not require erythropoietin or the erythropoietin receptor. *Cell* **83**, 59-67 (1995).
15. Welch, J.J. et al. Global regulation of erythroid gene expression by transcription factor GATA-1. *Blood* **104**, 3136-3147 (2004).
16. Cantor, A.B. & Orkin, S.H. Transcriptional regulation of erythropoiesis: an affair involving multiple partners. *Oncogene* **21**, 3368-3376 (2002).
17. Pevny, L. et al. Erythroid differentiation in chimaeric mice blocked by a targeted mutation in the gene for transcription factor GATA-1. *Nature* **349**, 257-260 (1991).
18. Chiba, T., Ikawa, Y. & Todokoro, K. GATA-1 transactivates erythropoietin receptor gene, and erythropoietin receptor-mediated signals enhance GATA-1 gene expression. *Nucleic acids research* **19**, 3843-3848 (1991).

19. Kuramochi, S., Ikawa, Y. & Todokoro, K. Characterization of murine erythropoietin receptor genes. *Journal of Molecular Biology; Journal of Molecular Biology* **216**, 567-575 (1990).
20. Zon, L.I. & Orkin, S.H. Sequence of the human GATA-1 promoter. *Nucleic acids research* **20**, 1812 (1992).
21. Zon, L.I., Youssoufian, H., Mather, C., Lodish, H.F. & Orkin, S.H. Activation of the erythropoietin receptor promoter by transcription factor GATA-1. *Proceedings of the National Academy of Sciences of the United States of America* **88**, 10638-10641 (1991).
22. Hannon, R., Evans, T., Felsenfeld, G. & Gould, H. Structure and promoter activity of the gene for the erythroid transcription factor GATA-1. *Proceedings of the National Academy of Sciences of the United States of America* **88**, 3004-3008 (1991).
23. Tsai, S.F., Strauss, E. & Orkin, S.H. Functional analysis and in vivo footprinting implicate the erythroid transcription factor GATA-1 as a positive regulator of its own promoter. *Genes & development* **5**, 919-931 (1991).
24. Iwasaki, H. et al. GATA-1 converts lymphoid and myelomonocytic progenitors into the megakaryocyte/erythrocyte lineages. *Immunity* **19**, 451-462 (2003).
25. Ghaffari, S. et al. AKT induces erythroid-cell maturation of JAK2-deficient fetal liver progenitor cells and is required for Epo regulation of erythroid-cell differentiation. *Blood* **107**, 1888-1891 (2006).
26. Zhao, W., Kitidis, C., Fleming, M.D., Lodish, H.F. & Ghaffari, S. Erythropoietin stimulates phosphorylation and activation of GATA-1 via the PI3-kinase/AKT signaling pathway. *Blood* **107**, 907-915 (2006).
27. Rooke, H.M. & Orkin, S.H. Phosphorylation of Gata1 at serine residues 72, 142, and 310 is not essential for hematopoiesis in vivo. *Blood* **107**, 3527-3530 (2006).
28. Blobel, G.A., Nakajima, T., Eckner, R., Montminy, M. & Orkin, S.H. CREB-binding protein cooperates with transcription factor GATA-1 and is required for erythroid differentiation. *Proceedings of the National Academy of Sciences of the United States of America* **95**, 2061-2066 (1998).
29. Blobel, G.A. CREB-binding protein and p300: molecular integrators of hematopoietic transcription. *Blood* **95**, 745-755 (2000).
30. Liu, Y., Denlinger, C.E., Rundall, B.K., Smith, P.W. & Jones, D.R. Suberoylanilide hydroxamic acid induces Akt-mediated phosphorylation of p300, which promotes acetylation and transcriptional activation of RelA/p65. *The Journal of biological chemistry* **281**, 31359-31368 (2006).
31. Vojtek, A.B. et al. Akt regulates basic helix-loop-helix transcription factor-coactivator complex formation and activity during neuronal differentiation. *Molecular and cellular biology* **23**, 4417-4427 (2003).
32. Huang, W.C. & Chen, C.C. Akt phosphorylation of p300 at Ser-1834 is essential for its histone acetyltransferase and transcriptional activity. *Molecular and cellular biology* **25**, 6592-6602 (2005).
33. Palani, S. & Sarkar, C.A. Positive receptor feedback during lineage commitment can generate ultrasensitivity to ligand and confer robustness to a bistable switch. *Biophysical journal* **95**, 1575-1589 (2008).

34. Miura, Y., Komatsu, N. & Suda, T. Growth and differentiation of two human megakaryoblastic cell lines; CMK and UT-7. *Progress in clinical and biological research; Progress in clinical and biological research* **356**, 259-270 (1990).
35. Komatsu, N. et al. Establishment and characterization of a human leukemic cell line with megakaryocytic features: dependency on granulocyte-macrophage colony-stimulating factor, interleukin 3, or erythropoietin for growth and survival. *Cancer research; Cancer research* **51**, 341-348 (1991).
36. Komatsu, N. et al. In vitro development of erythroid and megakaryocytic cells from a UT-7 subline, UT-7/GM. *Blood* **89**, 4021-4033 (1997).

Chapter 4

Converting a Linear Signaling Pathway into an Externally-Regulated, Tunable, and Reversible Switch

4.1. Introduction

In response to extracellular cues, natural biological systems can evoke dynamic internal responses that can be critical for achieving robust phenotypic changes. Natural systems can exhibit a wide array of responses with modularity and specificity by integrating signaling elements at the cell surface, cytoplasm and nucleus. Cell-surface receptors enable cue-specific recognition and signaling, cytoplasmic messengers provide signal processing modules and transcriptional elements in the nucleus regulate complex changes in gene expression. The signal-response circuitry in natural biological systems is extremely unwieldy due to evolved complexity. Synthetic biology provides a means to dissect these complex signaling motifs in order to better understand the design principles underlying the dynamics of a network of interest and, ultimately, the phenotypic response[1-3].

Synthetic biology as a field possesses tremendous potential[4-8] in improving drug delivery, producing alternative fuels, detecting cancer, engineering tissues, and enhancing gene therapy; however, it has faced major challenges due to the inefficiency of individual modules, incompatibility between interacting parts, unpredictability in large circuitry and inconsistency due to cellular variability[9]. The solution to some of these drawbacks can arise from designing generalizable methodologies to invoke specific responses from any given circuitry with minimal modifications.

Changes in gene expression are fundamental to any phenotypic response and two molecular components that invariably regulate this process are cell surface receptors and nuclear transcription factors. These two signaling elements are ubiquitously present in most natural systems and control survival, apoptosis, proliferation, lineage commitment and differentiation[10-12]. Here, we present an externally-regulated autofeedback topology, inspired from lineage commitment networks in stem cells and progenitors[13, 14], which is capable of converting any linear receptor-to-transcription factor signaling pathway into a reversible switch.

Most of the synthetic systems developed to date utilize autofeedback loops activated by membrane diffusible small molecules to achieve a bimodal response[15-20]. These systems are typically irreversible and require extensive tuning to achieve the desired response. In contrast, natural systems can regulate their autofeedback loops through external ligand and achieve robustness in the exhibited response by incorporating signaling modules between the receptors and transcription factors. Previously, we developed a mathematical model of a minimal topology that can generate robust bistable responses for any linearly connected receptor/transcription factor pair[13, 14]. In this

topology, a ligand binds to its cognate cell-surface receptor and transmits a signal that activates a downstream transcription factor. The activated transcription factor activates its own transcription as well as that of the cognate cell-surface receptor. Since the transcription factor requires a post-transcriptional modification for activation, the displayed response is reversible and externally regulatable. Furthermore, our synthetic approach only requires rewiring of existing components (as opposed to introducing new components), and therefore requires minimal tweaking to achieve a robust response.

As proof of principle, we tested our approach in *Saccharomyces cerevisiae* using a pathway that involves synthetic receptor (CRE1 from Arabidopsis thaliana) signaling to an endogenous yeast transcription factor (SKN7)[21]. This basic pathway, which is graded and unimodal, was rationally rewired to achieve our desired topology and the resulting network immediately showed high ultrasensitivity and bimodality without tweaking. We further show that this topology can be tuned to regulate system dynamics such as activation/deactivation kinetics, signal amplitude, switching threshold and sensitivity.

4.2. Materials and Methods

4.2.1. Plasmids

Plasmids and their properties are provided in Table 4.1. *E. coli* strain XL1-Blue was used for all plasmid construction. Yeast integration plasmids pRS403 and pRS405 were purchased from Stratagene. Yeast centromere plasmids DQ232595, DQ232596 and DQ232600 were kindly provided by Ron Weiss (Massachusetts Institute of Technology). The SKN7 gene was PCR amplified from the *S. cerevisiae* genome. Plasmids were

constructed as follows: for pSP001, DQ232600 was cut with Pvu1 to remove CEN6/ARSH4 component and ligated into pRS405 digested with Pvu1; for pSP002 and pSP003, DQ232600 was digested with BamH1 to excise AtCRE1a-P_{SSRE} region and cloned into DQ232595 and DQ232596. CEN6/ARSH4 and HIS3 component were removed from the modified DQ232595 and DQ232596 plasmids (Pvu1 site) and ligated into pRS405; for pSP004, the SKN7 gene was cloned to replace yEGFP3 in DQ232596 using the BamH1/Asc1 restriction sites. The CEN6/ARSH4 component in the modified DQ232596 vector was removed (Pvu1 site) and ligated into pRS403 plasmid. All plasmid constructs were verified by sequencing.

4.2.2. Yeast strains, genomic screens and culture

Yeast strains and their properties are listed in Table 4.2. Yeast transformations were performed using the high-efficiency LiAc/SS carrier DNA/PEG method with the standard synthetic dropout (SD) media. The no-feedback strain cRcTF was constructed by integrating a single copy of pSP001 into the LEU2 locus of the TM182 strain. Receptor feedback strains sRcTF and tRcTF were built by integrating a single copy of pSP002 and pSP003, respectively, into the LEU2 locus of TM182 strain. The transcription factor feedback strain cRtTF was constructed by integrating a single copy of pSP004 into the HIS3 locus of the cRcTF strain. Double feedback strains sRtTF and tRtTF were built by integrating a single copy of pSP004 into the HIS3 locus of the sRcTF and tRcTF strains, respectively. Yeast genomic DNA isolation was performed using the phenol:chloroform:isoamyl alcohol (PCI) method. The isolated genomic DNA was probed for stable single integration of the transformed gene at the desired locus through PCR.

The wild-type TM182 strain was used as a negative control in all experiments. Yeast strains were grown at 30ºC in their respective SD Ura$^-$/His$^-$/Leu$^-$ medium with 2% galactose. IP was purchased from Sigma-Aldrich.

4.2.3. Analysis of GFP expression

For all IP-induction experiments, GFP levels were quantified by fluorescence-activated cell sorting (FACS) using a Guava flow cytometer containing a 488 nm argon excitation laser and a 525/30 nm emission filter. For each sample, 10,000 events were acquired and the fluorescence data was analyzed using FlowJo 7.5.

4.3. Results and discussion

4.3.1. Network design

A cell-surface receptor from *A. thaliana* has previously been coupled to signaling components from *S. cerevisiae* to successfully create a synthetic linear signaling pathway[21]. In this work, we have rationally constructed several topological variations of this pathway to achieve switch-like bimodality in system response, as well as tunability in set-point, EC50 and ultrasensitivity. Previous work with these signaling elements used yeast centromere plasmids for pathway construction[21]. Due to the difference in copy number of transformed genes in the cells, we observed a high degree of heterogeneity in the expression of the signaling elements (data not shown). To minimize this variability, we chose to use yeast integrating plasmids and performed genomic screens to obtain clones with stable, single integrations of the transformed genes.

The basic pathway without feedback loops (cRcTF) was created by integrating a single copy of AtCRE1, an *A. thaliana* cytokinin receptor gene[22], driven by the constitutive CYC1 promoter and a single copy of yEGFP3, yeast green fluorescent gene driven by the SSRE promoter. The synthetic SSRE promoter contains the binding site for the yeast endogenous transcription factor SKN7. The plant cytokinin isopentenyl adenine[23] (IP) was used as the ligand to activate the AtCRE1 receptor. The activated receptor signals via a yeast endogenous pathway to phosphorylate SKN7[24, 25]. Phospho-SKN7 can bind to the SSRE promoter and activate GFP expression (Figure 4.1).

Networks with only receptor feedback (sRcTF and tRcTF) were created by replacing the CYC1 promoter of AtCRE1 with SSRE or TR-SSRE (contains two binding sites for phosphorylated SKN7), thereby creating an IP-regulated positive feedback loop through the receptor. The network with only transcription-factor feedback (cRtTF) was created by integrating a single copy of the SKN7 gene, driven by the TR-SSRE promoter, to the basic pathway. Networks with both receptor and transcription-factor feedback (sRtTF and tRtTF) were constructed by integrating a single copy of the SKN7 gene (driven by TR-SSRE promoter) and by replacing the CYC1 promoter of AtCRE1 with either SSRE or TR-SSRE.

4.3.2. Set-point and synthesis kinetics

Expression of GFP is dependent on the concentration of phosphorylated SKN7 (phospho-SKN7). Synthesis of phospho-SKN7 requires the expression of SKN7 and its subsequent posttranslational modification via IP:AtCRE1 complex signaling. Hence, the expression of GFP can be limited either by the insufficiency of SKN7 (transcription factor limited) or IP:AtCRE1 complex (receptor limited).

All six strains were grown overnight in their respective SD media to reach steady-state. The cultures were then diluted to an OD_{660} reading of 0.1 and induced with 1μM IP. Aliquots from the growing cultures were extracted at different time points to quantify GFP levels by flow cytometry.

Change in receptor feedback strengths (CYC1, SSRE, TR-SSRE) did not change the achieved steady-state set-point of GFP when the transcription factor SKN7 was driven only by the constitutive promoter. This is probably due to the fact that receptor expression from all the promoters was sufficient to activate the endogenous levels of SKN7; hence, these systems are transcription-factor limited (Figure 4.2A).

When we added an exogenous copy of the SKN7 gene, driven by a TR-SSRE promoter, to the network, the system became receptor-limited. Strain cRtTF was most limited by the synthesis of AtCRE1, which was significantly overcome by replacing the constitutive CYC1 promoter with SSRE promoter (Figure 4.2B). A further increase in GFP expression was obtained by switching the SSRE promoter to the stronger TR-SSRE promoter (Figure 4.2B). In contrast to the strains with constitutive (and limiting) expression of SKN7 (cRcTF, sRcTF and tRcTF), strains cRtTF and sRtTF are limited by the synthesis of AtCRE1.

Also, the kinetics of production of GFP seems to vary based on the network topology due to the variation in the rate of synthesis of Phospho-SKN7. In Figure 4.2C, the kinetic profiles of all the strains were normalized to their steady-state value. The strain with the basic pathway is the fastest to reach steady-state (60% level is reached around ~4 hours), followed by the receptor-feedback strains (60% level in ~5 hours for

sRcTF and ~ 6 hours for tRcTF). The transcription-factor feedback strain and sRtTF double feedback strain reached 60% level ~ 7.5 hours; whereas the tRtTF strain was the slowest and attained the 60% mark ~ 9 hours post IP-induction.

This shows that by operating the system with different strengths of feedback at the transcription factor and receptor level, we can achieve tunability in the steady state set-point as well on the time to reach steady-state.

4.3.3. Ultrasensitivity and EC50

All six strains were grown overnight in their respective SD medium and then diluted back to an OD_{660} of 0.1 and induced with different concentrations of IP spanning from 0.01 to 10 μM. The steady-state levels of GFP were plotted against IP concentration to analyze the dynamics of the response. We noticed that the basic pathway and the networks with only receptor-mediated feedback showed a graded response with respect to IP concentration, whereas the strains with transcription-factor feedback and with both feedback loops showed a highly switch-like response to change in IP (Figure 4.3). No feedback and receptor feedback strains reached set-point saturation around 1 μM IP, whereas the other three strains reached saturation at 10-fold lower IP (Figure 4.3).

The steady-state dose-response curves show that the ultrasensitivity was lowest in the no-feedback (Hill coefficient ~2) and receptor-feedback networks (Hill coefficient ~2), significantly higher in the transcription-factor-feedback network (Hill coefficient ~11) and, higher still in the double-feedback networks (Hill coefficient ~24). Also, the networks with transcription-factor feedback and double feedback showed 2-fold lower EC50 when compared with no feedback and receptor-feedback networks (Figure 4.3).

This shows that the basic pathway in this system is ultrasensitive and its sensitivity can be increased by the incorporation of transcription factor feedback, and further improved by the addition of receptor-feedback loops. Also, with feedback loops, the system can be tuned to switch on at lower IP concentrations than those that activate the basic pathway.

4.3.4. Bimodality and degradation kinetics

Degradation kinetics for the no feedback control strain cRcTF and double feedback strain tRtTF were compared in Figure 4.4A. The two strains were grown in 1 μM IP for 24 hours to achieve saturated levels of GFP. The cells were then spun down, washed three times with regular media and cultured in their respective SD media without IP. Aliquots from the cultures were extracted and quantified for GFP levels. The figure shows that both systems are reversible and therefore ultimately shut off; however, there is a delay in the degradation kinetics of tRtTF when compared to cRcTF (Figure 4.4A). This could possibly be due to the higher accumulation of phospho-SKN7 in tRtTF than in cRcTF at steady state. Another possible explanation is that, due to the higher cell-surface receptor expression on the cells with receptor feedback, more residual IP may be bound on tRtTF cells than on cRcTF cells.

When IP was only reduced to a sub-threshold concentration (rather than completely removed), the degradation kinetics of the GFP signal became markedly different between the two strains. When IP was lowered to 0.025 μM, the tRtTF strain retained ~50% of its maximal response after 10 hours whereas the cRcTF strain only sustained ~10% of its response (Figure 4.4B). When IP was lowered to 0.05 μM, the

tRtTF strain sustained ~70% of its response but the cRcTF strain only retained ~10% of its maximal response (Figure 4.4C). This shows that the tRtTF topology confers memory to the cell with respect to the receptor-specific ligand (here, IP). The initial induction with high IP triggers this memory, which helps to sustain the response even after lowering the IP to sub-threshold levels. We analyzed the GFP histograms from time points in Figure 4.4C. As seen from the plot, the entire histogram for the cRcTF strain shifted to lower mean fluorescence at each time point when a step change in IP (from 1 μM to 0.05 μM) was introduced into the system (Figure 4.4D). This shows a typical monostable response, as each cell with the basic pathway lost its fluorescence gradually over time once the IP concentration was lowered. By contrast, the tRtTF strain showed bimodality in GFP expression as the IP concentration was lowered to 0.05 μM; a high percentage of cells remained in the high-GFP state and a low percentage of cells completely lost their fluorescence, with virtually no cells exhibiting intermediate fluorescence intensities (Figure 4.4D). The aggregate decrease in fluorescence of a population of tRtTF cells does not result from a decrease in the mean fluorescence intensity of the overall population but, rather, from a shift in the relative populations of the high-GFP and low-GFP states.

In conclusion, we have presented here a method by which any linear signaling pathway with incorporation of two positive feedback loops can be rewired to exhibit a bimodal response. This system is completely reversible and by changing the strengths of the feedback loops, we show tunability in kinetics, set-point, EC50 and ultrasensitivity.

Figure 4.1

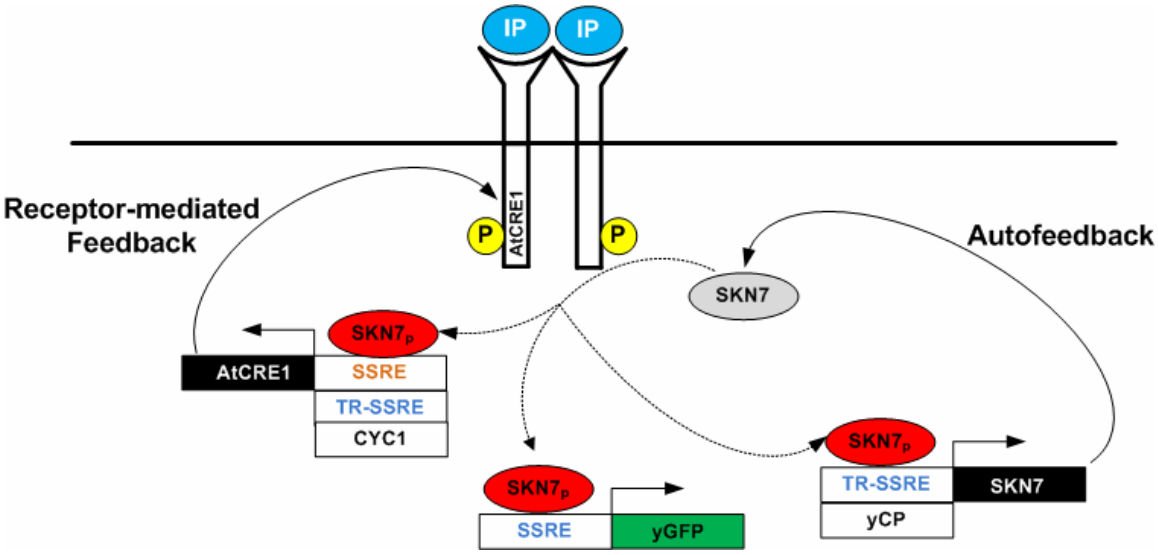
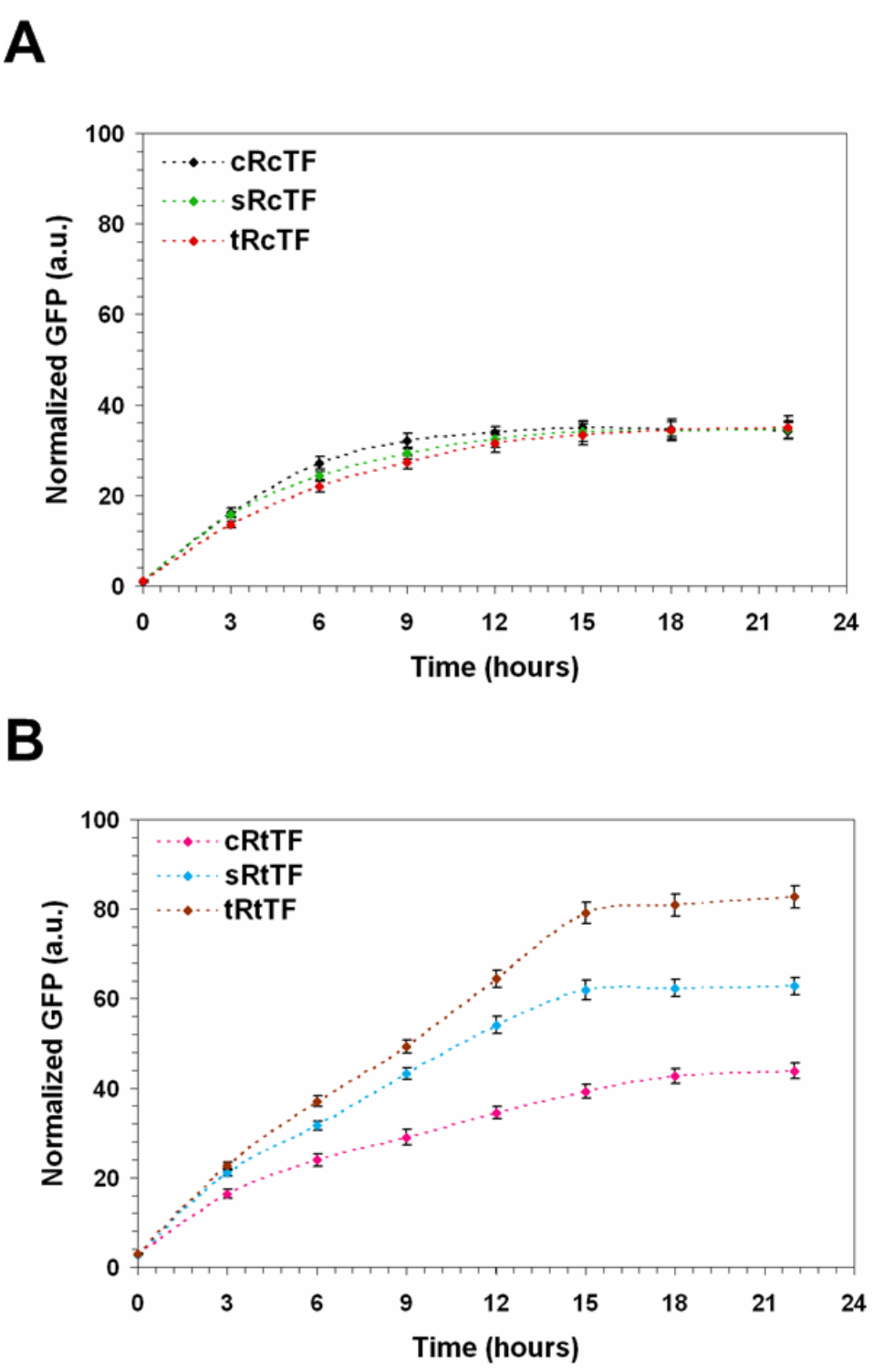


Figure 4-1 Network design

Plant cytokinin IP binds and activates the plant receptor AtCRE1. Signaling from this receptor results in phosphorylation of the yeast endogenous transcription factor SKN7. Phospho-SKN7 binds to synthetic promoters SSRE and TR-SSRE and expresses the downstream genes. AtCRE1 is either expressed constitutively or expressed via SSRE or TR-SSRE promoters to form a receptor feedback loop. SKN7 is expressed constitutively and in some networks, an additional copy is regulated through the TR-SSRE promoter to form an autofeedback loop. GFP is only expressed through the SSRE promoter.

Figure 4.2


A
100
80
60
40
20
0
Normalized GFP (a.u.)
cRcTF
sRcTF
tRcTF
0
3
6
9
12
15
18
21
24
Time (hours)
B
100
80
60
40
20
0
Normalized GFP (a.u.)
cRtTF
sRtTF
tRtTF
0
3
6
9
12
15
18
21
24
Time (hours)

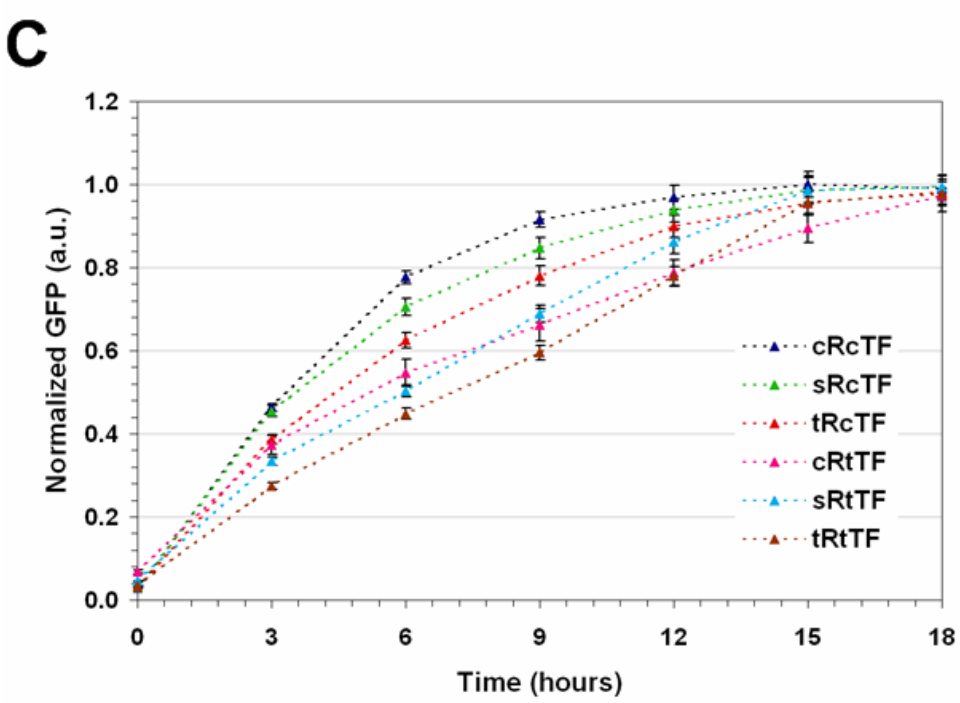


Figure 4-2 Variation in set-point and synthesis rate of GFP

All strains were grown in 1 µM IP until GFP levels reached saturation. Aliquots from the cultures at different time points were taken to determine the synthesis kinetics of GFP (A) Expression of GFP limited by transcription-factor expression: Kinetic profiles of strains with variable strength of receptor feedback loop and constitutive expression of transcription factor. AtCRE1 receptor expression from cRcTF, sRcTF and tRcTF strains were sufficient to activate the endogenous levels of SKN7. (B) Expression of GFP limited by receptor expression: Kinetic profiles of strains with variable strengths of receptor feedback loop and autofeedback driven expression of transcription factor. When an additional copy of SKN7, driven by the TR-SSRE promoter, was added to the system, the strains cRtTF and sRtTF became receptor limited. (C) Kinetic profiles of the strains are normalized to their steady-state set-point. All six strains achieve steady-state at different times; cRcTF seems to be the fastest and tRtTF the slowest.

Figure 4.3

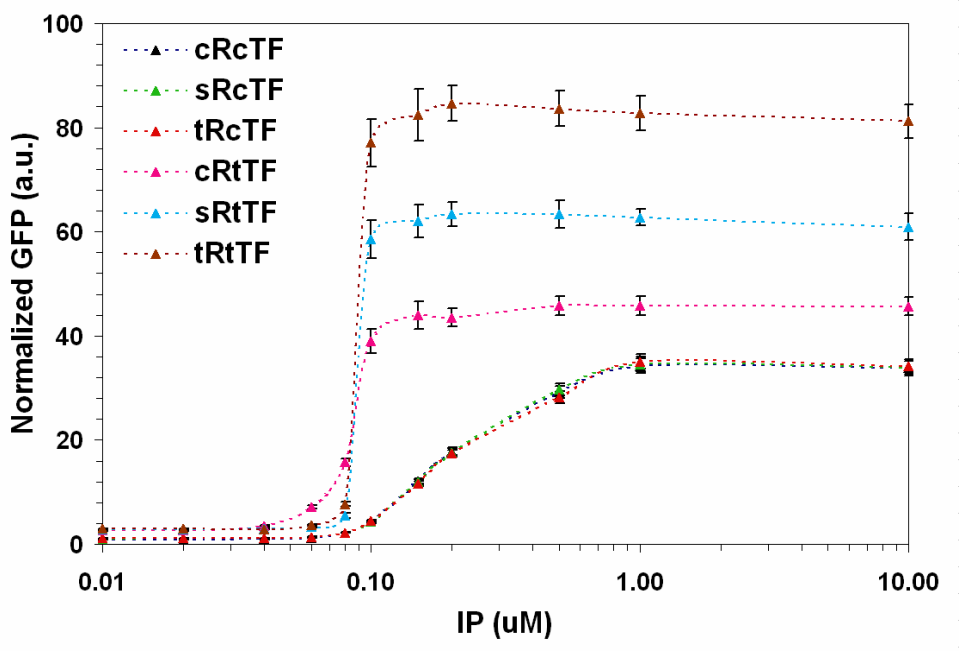
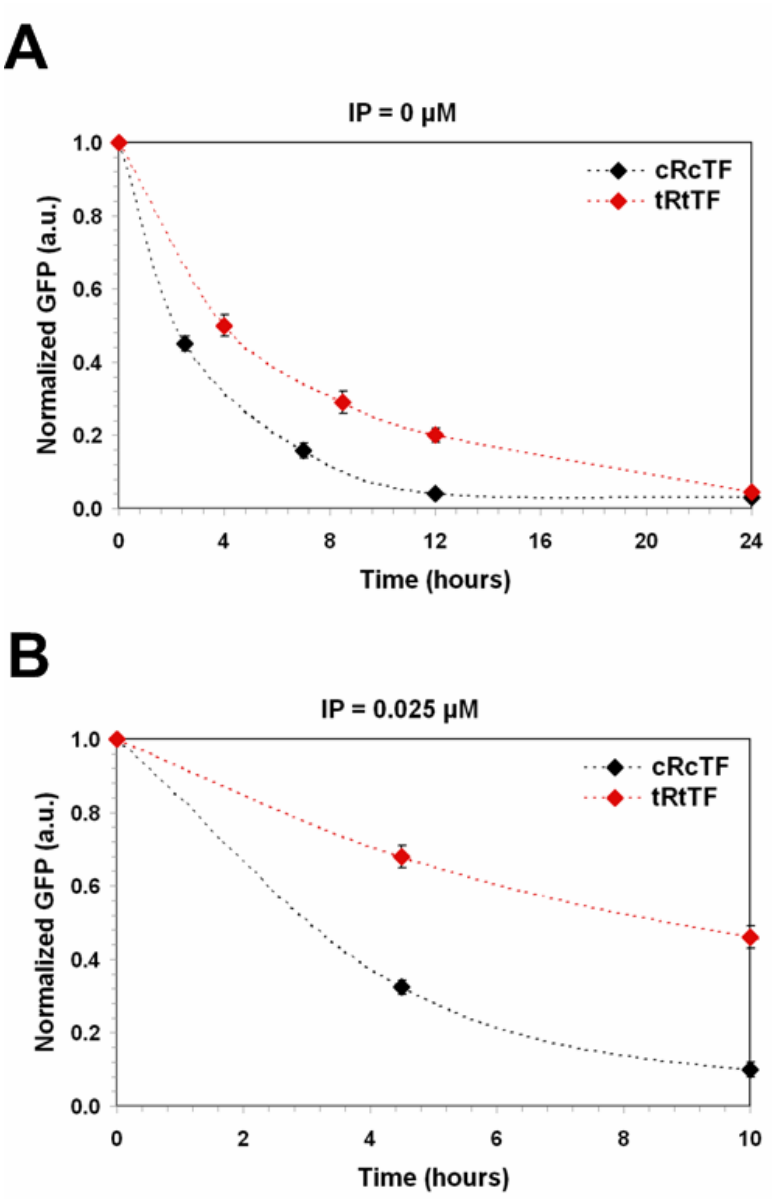


Figure 4-3 Tunability of set-point, EC50 and ultrasensitivity

All six strains were grown at different IP concentrations spanning from 0.01 µM to 10 µM. cRcTF, sRcTF and tRcTF showed graded responses to changes in IP, whereas cRtTF, sRtTF and tRtTF exhibited switch-like responses. Also, cRcTF, sRcTF and tRcTF, when compared to cRtTF, sRtTF and tRtTF, needed two-fold higher IP to achieve the half-maximal response and 10-fold higher IP to achieve the full response.

Figure 4.4


A
IP = 0 µM
Normalized GFP (a.u.)
1.0
0.8
0.6
0.4
0.2
0.0
0
4
8
12
16
20
24
Time (hours)
cRcTF
tRtTF
B
IP = 0.025 µM
Normalized GFP (a.u.)
1.0
0.8
0.6
0.4
0.2
0.0
0
2
4
6
8
10
Time (hours)
cRcTF
tRtTF

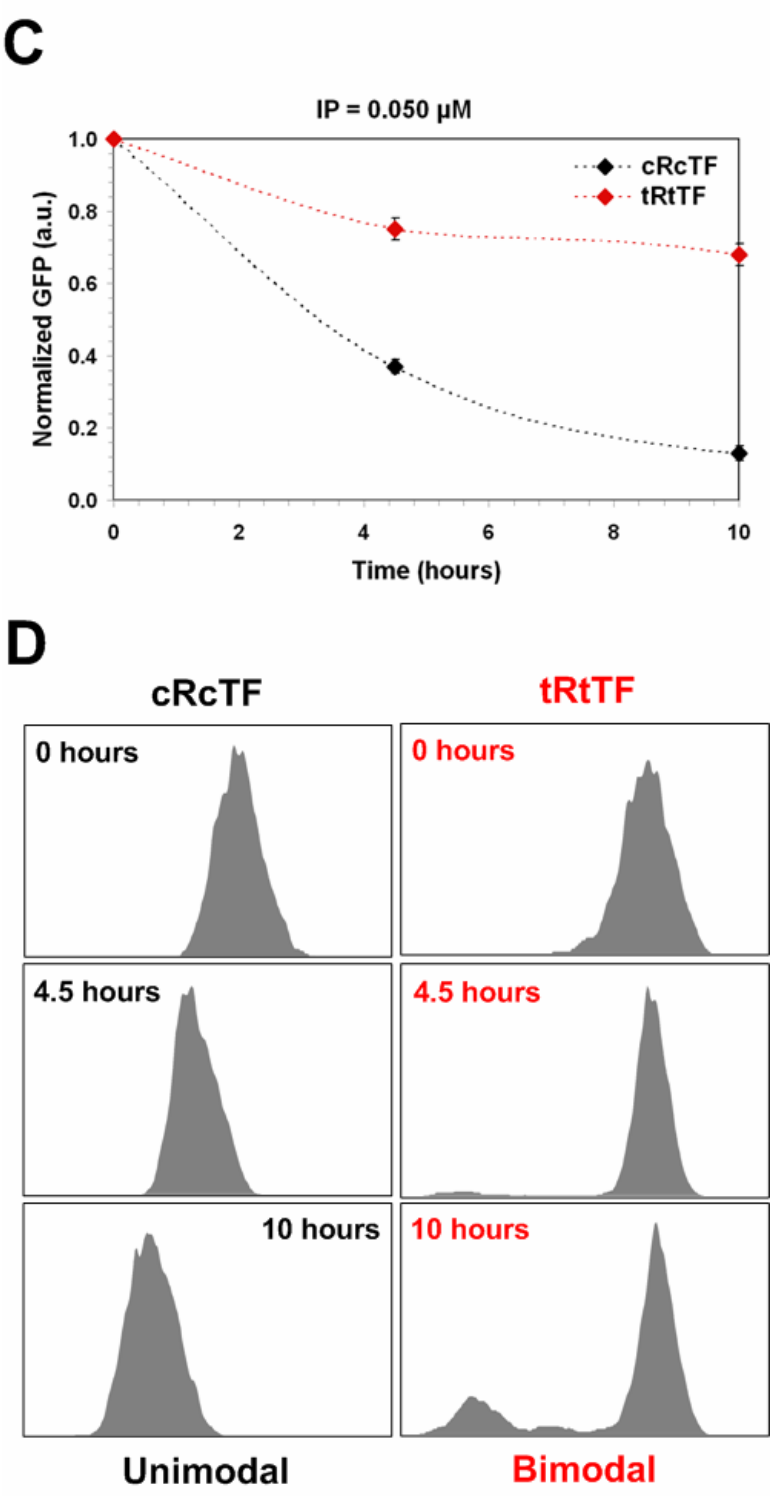


Figure 4-4 Degradation kinetics and bimodality

Strains cRcTF and tRtTF were grown in their respective SD media with 1 μM IP for 24 hours. After reaching saturation in GFP set-point, the cultures were washed and moved into media with different sub-threshold IP concentrations: (A) IP = 0, both the strains show reversibility by switching off GFP expression; however, the degradation kinetics of tRtTF are slower than those of cRcTF. (B) IP = 0.025 μM, after 10 hours, tRtTF sustained ~50% of maximal response, whereas cRcTF managed only ~10%. (C) IP = 0.05 μM, after 10 hours, tRtTF retained ~70% of its steady-state response, whereas cRcTF managed only ~10%. (D) Histograms of GFP expression from (C) at time points 0, 4.5 and 10 hours. Strain cRcTF shows a unimodal response, whereas tRtTF shows a bimodal response.

Table 4.1

Table 4-1 Plasmids used in this study

Plasmid	Properties	Reference
pRS403	HIS3, integration vector	Stratagene
pRS405	LEU2, integration vector	Stratagene
DQ232595	CEN6/ARSH4, HIS3, P_{SSRE}-yEGFP3-T_{ADH1}	(1)
DQ232596	CEN6/ARSH4, HIS3, $P_{TR\text{-}SSRE}$-yEGFP3-T_{ADH1}	(1)
DQ232600	CEN6/ARSH4, LEU2, P_{CYC1}-AtCRE1a, P_{SSRE}-yEGFP3-T_{ADH1}	(1)
pSP001	LEU2, P_{CYC1}-AtCRE1a, P_{SSRE}-yEGFP3-T_{ADH1}	This study
pSP002	LEU2, P_{SSRE}-AtCRE1a, P_{SSRE}-yEGFP3-T_{ADH1}	This study
pSP003	LEU2, $P_{TR\text{-}SSRE}$-AtCRE1a, P_{SSRE}-yEGFP3-T_{ADH1}	This study
pSP004	HIS3, $P_{TR\text{-}SSRE}$-SKN7-T_{ADH1}	This study

1. Chen, M.T. & Weiss, R. Artificial cell-cell communication in yeast Saccharomyces cerevisiae using signaling elements from Arabidopsis thaliana. *Nature biotechnology* **23**, 1551-1555 (2005).

Table 4.2

Table 4-2 Yeast strains used in this study

Yeast Strain	Properties	Reference
TM182	MAT leu2 his3; sln1::hisG; P_{GAL1}-PTP2 (ura3)	(1)
cRcTF	TM182 + pSP001	This study
sRcTF	TM182 + pSP002	This study
tRcTF	TM182 + pSP003	This study
cRtTF	NoFb1 + pSP004	This study
sRtTF	ReFb1 + pSP004	This study
tRtTF	ReFb1 + pSP004	This study

1. Maeda, T., Wurgler-Murphy, S.M. & Saito, H. A two-component system that regulates an osmosensing MAP kinase cascade in yeast. *Nature* **369**, 242-245 (1994).

4.4. References

1. Kiel, C., Yus, E. & Serrano, L. Engineering signal transduction pathways. *Cell* **140**, 33-47 (2010).
2. Agapakis, C. & Silver, P. Synthetic biology: exploring and exploiting genetic modularity through the design of novel biological networks. *Mol Biosyst* **5**, 704-713 (2009).
3. Haynes, K. & Silver, P. Eukaryotic systems broaden the scope of synthetic biology. *J Cell Biol* **187**, 589-596 (2009).
4. Burrill, D. & Silver, P. Making cellular memories. *Cell* **140**, 13-18 (2010).
5. Gore, J. & van Oudenaarden, A. Synthetic biology: The yin and yang of nature. *Nature* **457**, 271-272 (2009).
6. Mukherji, S. & van Oudenaarden, A. Synthetic biology: understanding biological design from synthetic circuits. *Nat Rev Genet* **10**, 859-871 (2009).
7. Alper, H. & Stephanopoulos, G. Engineering for biofuels: exploiting innate microbial capacity or importing biosynthetic potential? *Nat Rev Microbiol* **7**, 715-723 (2009).
8. Ro, D. et al. Production of the antimalarial drug precursor artemisinic acid in engineered yeast. *Nature* **440**, 940-943 (2006).
9. Kwok, R. Five hard truths for synthetic biology. *Nature* **463**, 288-290 (2010).
10. Pearce, L., Komander, D. & Alessi, D. The nuts and bolts of AGC protein kinases. *Nat Rev Mol Cell Biol* **11**, 9-22 (2010).
11. Levy, D. & Darnell, J.J. Stats: transcriptional control and biological impact. *Nat Rev Mol Cell Biol* **3**, 651-662 (2002).
12. Kolch, W. Coordinating ERK/MAPK signalling through scaffolds and inhibitors. *Nat Rev Mol Cell Biol* **6**, 827-837 (2005).
13. Palani, S. & Sarkar, C.A. Positive receptor feedback during lineage commitment can generate ultrasensitivity to ligand and confer robustness to a bistable switch. *Biophysical journal* **95**, 1575-1589 (2008).
14. Palani, S. & Sarkar, C. Integrating extrinsic and intrinsic cues into a minimal model of lineage commitment for hematopoietic progenitors. *PLoS Comput Biol* **5**, e1000518 (2009).
15. Gardner, T.S., Cantor, C.R. & Collins, J.J. Construction of a genetic toggle switch in Escherichia coli. *Nature* **403**, 339-342 (2000).
16. Hasty, J., McMillen, D. & Collins, J.J. Engineered gene circuits. *Nature* **420**, 224-230 (2002).
17. Guido, N.J. et al. A bottom-up approach to gene regulation. *Nature* **439**, 856-860 (2006).
18. Deans, T.L., Cantor, C.R. & Collins, J.J. A tunable genetic switch based on RNAi and repressor proteins for regulating gene expression in mammalian cells. *Cell* **130**, 363-372 (2007).
19. Ozbudak, E.M., Thattai, M., Lim, H.N., Shraiman, B.I. & Van Oudenaarden, A. Multistability in the lactose utilization network of Escherichia coli. *Nature* **427**, 737-740 (2004).

20. Becskei, A., Seraphin, B. & Serrano, L. Positive feedback in eukaryotic gene networks: cell differentiation by graded to binary response conversion. *The EMBO journal* **20**, 2528-2535 (2001).
21. Chen, M.T. & Weiss, R. Artificial cell-cell communication in yeast Saccharomyces cerevisiae using signaling elements from Arabidopsis thaliana. *Nature biotechnology* **23**, 1551-1555 (2005).
22. Inoue, T. et al. Identification of CRE1 as a cytokinin receptor from Arabidopsis. *Nature* **409**, 1060-1063 (2001).
23. Mok, D.W. & Mok, M.C. Cytokinin Metabolism and Action. *Annual Review of Plant Physiology and Plant Molecular Biology* **52**, 89-118 (2001).
24. Li, S. et al. The yeast histidine protein kinase, Sln1p, mediates phosphotransfer to two response regulators, Ssk1p and Skn7p. *The EMBO journal* **17**, 6952-6962 (1998).
25. Maeda, T., Wurgler-Murphy, S.M. & Saito, H. A two-component system that regulates an osmosensing MAP kinase cascade in yeast. *Nature* **369**, 242-245 (1994).

Chapter 5

Integrating Extrinsic and Intrinsic Signals into a Multilineage Cell-Fate Model

(Adapted from Palani S. and Sarkar C.A. Integrating Extrinsic and Intrinsic Cues into a Mininal Model of Lineage Commitment for Hematopoietic Progenitors (2009) PLoS Computational Biology, September 2009, Volume 5, Issue 9, e1000518)

5.1 Introduction

Multipotent stem cells have the ability to both self-renew and differentiate, thus sustaining the stem cell pool and giving rise to mature, specialized cells, respectively. The hematopoietic stem cell (HSC), located in the adult bone marrow, is well characterized and has served as a popular model system for understanding self-renewal, lineage commitment, and differentiation[1]. HSCs are responsible for producing the entire repertoire of blood cells through the process of hematopoiesis. During hematopoiesis, HSCs lose the capacity to self-renew and differentiate into common myeloid (CMP) and common lymphoid (CLP) progenitors[2, 3]. Multipotent progenitors undergo further lineage-restricted differentiation to give rise to mature cells via bipotent progenitors. In

addition to this classical commitment paradigm in hematopoiesis, alternative pathways are emerging. For example, it has also been observed that HSCs and multipotent progenitors can bypass canonical intermediate states during commitment[4, 5]. The exact molecular events that direct lineage commitment at the stem cell stage or at the multipotent progenitor level remain elusive, but it is well appreciated that lineage-specific transcription factors and cytokine receptors play critical roles.

Lineage-specific transcription factors have been identified as master regulators of commitment and differentiation. They drive the expression of pertinent lineage-specific genes, thereby initiating the phenotypic change in the progenitor cell down a specific differentiation path[6, 7]. Developmental potency of a multipotent progenitor is reflected by the co-expression of multiple lineage-specific transcription factors at low levels, a phenomenon known as transcriptional priming[8]. This promiscuous gene expression pattern in the progenitor cell necessitates that, during cell differentiation, a specific transcription factor is upregulated, chiefly by positive autoregulation[9, 10], and other lineage transcription factors are downregulated, primarily through cross-antagonism[11-13].

In addition to lineage-specific transcription factors, cell differentiation is also believed to be tightly regulated by cytokines. Cytokines signal via their cognate receptors whose cytoplasmic domains activate various pathways involved in survival, proliferation, and differentiation[14-16]. It has been extensively debated whether cell fate during differentiation is a stochastic or an instructive process. The stochastic theory claims that the differential expression of lineage-specific transcription factors due to intrinsic noise in progenitor cells dictates the commitment decision[17-19], whereas the instructive theory argues that the absolute dependence on lineage-specific cytokine receptor signals during

differentiation shows that cell-fate decisions are regulated by extrinsic growth factor cues[20, 21]. An underlying question evoked by both of these theories is whether cytokines provide instructive cues or select lineage-committed progenitors by providing permissive survival and proliferation signals. The instructive model does not account for the occurrence of certain mature cell types even when their lineage-specific receptors are knocked out. The predetermined distribution of the heterogeneous progenitor population into mature cells, as suggested by the stochastic model fails to explain how specific cell types can be enriched during stress or how homeostasis is restored after infections or therapy. A recent groundbreaking study utilizing bioimaging techniques at the single-cell level suggests that there is validity to both of these theories[20]. These authors showed that lineage-specific cytokines can strongly instruct lineage choice, although differentiation was still possible in the absence of lineage-specific cytokines.

A more comprehensive understanding of lineage commitment may emerge by analyzing the biochemical associations that coordinate cell-extrinsic and cell-intrinsic events. The promiscuous gene expression pattern during differentiation is observed not only in lineage-specific transcription factors, but also in lineage-specific receptors. A critical commitment signal during differentiation is the upregulation of the transcription factor, which aids in expressing the lineage-specific genes; however, the need to upregulate the lineage-specific receptor, an event also integral to commitment, is still unclear. This is particularly confounding since the low number of lineage-specific receptors present in a progenitor cell is sufficient for providing permissive survival cues. During lineage commitment, the expression of the cytokine receptor mirrors the expression of the transcription factor, often due to the presence of transcription factor

binding domains in the promoter region of the receptor gene[22-25]. The advantage of regulating the lineage-specific receptor expression through the lineage-specific transcription factor is not apparent. Recent biochemical evidence also suggests that cytokines can provide signals to functionally activate lineage-specific transcription factors through post-translational modifications[26] and can also regulate the expression of transcription factors during cell differentiation[27].

Cell differentiation is believed to be an all-or-none "switch-like" event rather than a gradual transition of a precursor cell to a stable, mature cell. Mathematical modeling and analysis have been successfully used to provide insights into the biological networks that give rise to such switch-like behaviors[28]. Typically, the networks involved in lineage specification seem to engender cellular memory through nonintuitive behaviors, such as bistable response profiles. The components that generate bistability, the toggling of the system between two stable steady states, include nonlinear feedback loops[29, 30], external noise[31], and multi-site covalent modifications[32]. Previous lineage commitment models have suggested that transcriptionally primed multipotent progenitors are capable of exhibiting bistability purely via cell intrinsic events of autoregulation and cross-antagonism[33, 34], but these models have assumed the existence of cooperative positive feedback loops to achieve bistability and do not consider the role of extracellular cues.

While cooperativity is a widely recognized biological mechanism that may play an important role in lineage commitment, alternative mechanisms can generate similar switch-like behavior in networks where cooperativity has not been observed. For example, we have previously shown that cytokine-regulated, positive feedback of receptor can generate robust bistability to stimulus without cooperativity in a

deterministic model for unilineage commitment[35] (Chapter 2). Furthermore, even in networks with cooperativity, receptor-mediated feedback may provide additional robustness to the system behavior and, perhaps more importantly, offer an external mode of regulation of cell-fates.

Here, we present a minimal model that integrates the bidirectional regulation between lineage-specific cytokines and transcription factors with previously explored autofeedback loops and cross-antagonism to understand the interplay between cell-extrinsic and cell-instrinic factors in fate decisions of hematopoietic progenitors. The model is formulated under the premise that cell-fate decisions are stochastic in nature, but external cues such as lineage-specific cytokines can heavily bias this stochasticity, effectively instructing lineage choice (as has now been experimentally observed). Our model shows that the strength of cross-antagonism can be a critical determinant in achieving multistability. The analyzed network exhibits a "bilayer" of memory with respect to external stimuli to provide robustness to both the bipotent and committed states. The model suggests that noise in the network can enable stochastic switching between the stable states; however, the distribution of the uncommitted population among the various states during differentiation can still be strongly biased by external cues. Furthermore, this modeling framework captures both classical and alternative modes of lineage commitment seen in hematopoiesis. Although discrete cell fates are likely to represent high-dimensional attractors[36], our minimal model may provide an initial step towards understanding how extrinsic factors integrate with intrinsic factors and may elucidate new mechanisms that underlie cell-fate decisions.

5.2. Methods

5.2.1. Deterministic model

The minimal model shown in Figure 5.1 represents a regulatory network for lineage commitment of a multipotent progenitor to lineages A and B. The multipotent progenitor expresses basal levels of both lineage-specific transcription factors TF_A and TF_B (present in their inactive forms ITF_A and ITF_B) and lineage-specific receptors R_A and R_B before the addition of ligand. Addition of L_A to the system leads to receptor-ligand complex C_A formation. Complex C_A activates signaling pathways that lead to the activation of ITF_A to form ATF_A. Even though a mechanistic understanding of how this occurs via cytokine-mediated signaling has not fully emerged, we have modeled it to be rapidly regulated at the protein level (e.g., by post-translational modification). There may be other mechanisms involved (e.g., transcriptional and translational regulation) that are not considered here. The activated form of the transcription factor, ATF_A, upregulates the transcription of its own gene through positive autoregulatory feedback loop, enhancing production of ITF_A. ATF_A also upregulates the expression of the lineage-specific receptor R_A forming a ligand-regulated positive feedback loop. The model also accounts for basal synthesis of R_A and ITF_A, degradation of R_A, C_A, ITF_A and ATF_A and inactivation of ATF_A (not explicitly shown in Figure 5.1). For simplification, we consider the network topology in the commitment of the two lineages to be symmetric: the reactions involved in the activation of ITF_B to ATF_B by ligand L_B and the formation of the two positive feedback loops are analogous to those described in lineage A. To account for the cross-antagonism between the transcription factors TF_A and TF_B, ATF_A and ATF_B are modeled to downregulate the induced expression of [ITF_B, R_B] and [ITF_A, R_A] by competitively

inhibiting the binding of ATF_B and ATF_A to the regulatory domains present upstream of their lineage-specific receptor and TF genes. This multilineage commitment network led to a deterministic model with eight ordinary differential equations (ODEs) as shown in Table 5.1. The initial conditions and the values of the rate constants are provided in Table 5.2. A single-compartmental homogenous system is assumed and the pathways involved in TF activation and in the synthesis of TF and receptor are lumped as single-step reactions.

5.2.2. Stochastic version of the deterministic model

The Gillespie stochastic algorithm was employed to simulate a stochastic version of the ODE model[37]. The stochastic reactions and their probability functions are given in Table 5.3. Conversion of the deterministic model to its stochastic form was performed by using composite Michaelis-Menten type rate expressions in the propensity function instead of decomposing the minimal model into a series of elementary reactions; this was done to directly compare the dynamics of both the approaches[38, 39].

The probability functions and reactions for the stochastic model are provided in Table 5.3. The model was run either from the uncommitted state or from the bipotent state to reach the committed state. The initial conditions for both starting states are also provided in Table 5.2.

To achieve values close to the true stable steady states, the simulations for both the deterministic and the stochastic models were run for 100,000 mins or ~70 days. As an example, for one representative set of parameter values, the deterministic model reached 99% of its 100,000-minute value at approximately 2500 mins. Thus, even for conditions

that took significantly longer to reach steady state, our simulation time was more than sufficient. Also, since the time point that we picked is arbitrary, we performed 100 repetitions for each simulation condition to account for random fluctuations. Around this time point, the noise distribution had low variance and was relatively constant.

Small perturbations in the initial conditions for active species have no effect on the steady-state values or the distribution among the final states. Large perturbations also have no effect on the steady-state values, but can significantly alter the final-state distribution by priming the system with active receptors and/or transcription factors.

Parameters used to generate Figure 5.5

All simulations were started from the uncommitted state with $L_A|L_B$ values of 0|350, 100|250, 175|175, 250|100 and 350|0.

Initial conditions of the reactants:

$[R_A]_0 = 10$; $[R_B]_0 = 10$; $[C_A]_0 = 0$; $[C_B]_0 = 0$; $[ITF_A]_0 = 10$; $[ITF_B]_0 = 10$; $[ATF_A]_0 = 0$; $[ATF_B]_0 = 0$

For no inhibition condition: $K_{IA} = K_{IB} = \infty$

For moderate inhibition condition (in molecules): $K_{IA} = K_{IB} = 400$

For strong inhibition condition (in molecules): $K_{IA} = K_{IB} = 50$

All other parameter values are given in Table 5.2. For each condition, the simulations were run for 100,000 minutes and repeated 10,000 times.

Parameters used to generate Figure 5.6

All simulations were run using the moderate inhibition condition (in molecules): $K_{IA} = K_{IB} = 400$

Initial conditions of the reactants starting from the uncommitted state:

$[R_A]_0 = 10$; $[R_B]_0 = 10$; $[C_A]_0 = 0$; $[C_B]_0 = 0$; $[ITF_A]_0 = 10$; $[ITF_B]_0 = 10$; $[ATF_A]_0 = 0$; $[ATF_B]_0 = 0$

Initial conditions of the reactants starting from the bipotent state:

$[R_A]_0 = 30$; $[R_B]_0 = 30$; $[C_A]_0 = 125$; $[C_B]_0 = 125$; $[ITF_A]_0 = 6$; $[ITF_B]_0 = 6$; $[ATF_A]_0 = 273$; $[ATF_B]_0 = 273$

$L_A|L_B$ values for different trajectories:

Uncommitted to committed A – 250|5

Uncommitted to committed B – 5|250

Uncommitted to bipotent – 250|250

Bipotent to committed A – 250|5

Bipotent to committed B – 5|250

All other parameter values are given in Table 5.2. For each condition, the simulations were run for 100 hours and repeated 100 times.

Parameters used to generate Figures 5.7B and 5.7C

All simulations were started from the uncommitted state with $L_A|L_B$ values of 175|175.

Initial conditions of the reactants starting from the uncommitted state:

$[R_A]_0 = 10$; $[R_B]_0 = 10$; $[C_A]_0 = 0$; $[C_B]_0 = 0$; $[ITF_A]_0 = 10$; $[ITF_B]_0 = 10$; $[ATF_A]_0 = 0$; $[ATF_B]_0 = 0$

For moderate inhibition condition (in molecules): $K_{IA} = K_{IB} = 400$

For strong inhibition condition (in molecules): $K_{IA} = K_{IB} = 50$

All other parameter values are given in Table 5.2. For each condition, the simulations were run for 100,000 minutes and repeated 200 times.

5.2.3. Computational methods

The ODE-based deterministic model was solved using the numerical stiff solver ode15s in MATLAB (The Mathworks, Natick, MA). Time course, steady-state response and multistability plots were also created using MATLAB. The Gillespie algorithm for the stochastic model was programmed in C++. Histograms, phase plots and time trajectories of the stochastic simulations were created using the open-source statistical package R.

5.2.4. Microarray analysis

Normalized microarray data were generously provided by Tariq Enver (University of Oxford)[40]. The detailed experimental procedures for the microarray experiments and analyses are provided elsewhere[40]. EPOR, GCSFR and TPOR mRNA levels extracted from the data were further normalized to their basal levels present in the uninduced FDCP-mix. The inherent heterogeneity in the differentiating populations at each time point was overcome by weighting the contribution of each cell population to the average expression of the gene of interest.

Cell samples collected from FDCP-mix cells that were induced to differentiate along neutrophil, erythrocyte, and megakaryocyte lineages are inherently heterogeneous.

To overcome the problem of analyzing aggregate microarray data from these samples, we performed a simple deconvolution by generating a weighting function to account for the non-uniform contribution from each sub-population to the average signal. This is possible since the population fractions at each time point for each sample were previously determined.

Weighting function:

$$wBi \times fB + wNi \times fN + wEbi \times fEb + wEri \times fEr + wMi \times fM$$

B – blast (CMP)

N – neutrophil

Eb – erythroblast

Er – erythrocyte

M – megakaryocyte

wBi – contribution of blasts to the expression of gene i

fB – fraction of blasts in the population

wNi – contribution of neutrophils to the expression of gene i

fN – fraction of neutrophils in the population

wEbi – contribution of erythroblasts to the expression of gene i

fEb – fraction of erythroblasts in the population

wEri – contribution of erythrocytes to the expression of gene i

fEr – fraction of erythrocytes in the population

wMi – contribution of megakaryocytes to the expression of gene i

fM – fraction of megakaryocytes in the population

The weighted parameters (wXi) were obtained from a global fit to the microarray data for all time points across all experiments using a least squares approach and are listed in the Table 5.4. The kinetic expression profile of each gene for each lineage was then obtained by only considering blasts and the cells that are pertinent to the lineage of interest.

For erythrocytes: blasts, erythroblasts and erythrocytes

For megakaryocytes: blasts and megakaryocytes

For neutrophils: blasts and neutrophils

After removing the fractional contribution of cells from other lineages, the remaining populations were normalized to obtain a total fraction of 1.

5.3. Results

5.3.1. Model formulation

Different cell states in our model are identified by the relative expression levels of lineage-specific receptors and transcription factors. An uncommitted (or 'off-state') cell, such as a common myeloid progenitor (CMP), is one that expresses lineage-specific receptors and transcription factors for multiple lineages at low levels. It is primed to differentiate into several lineages, but not yet committed to any specific lineage. A bipotent (or 'intermediate-state') cell, such as a megakaryocyte-erythrocyte progenitor (MEP), is one that is restricted to exactly two lineages, but not yet committed to either of

them. Lineage-specific receptors and transcription factors for the two lineages are expressed at intermediate levels. A committed (or 'on-state') cell, such as a proerythroblast, is one that expresses the receptor and transcription factor of a single lineage at a high level and will eventually terminally differentiate into the corresponding mature cell.

The topology of our minimal model for multilineage commitment was informed by various experimental studies on lineage-specific receptors and transcription factors. The cytokines Epo, Tpo, GCSF, and MCSF have been shown to offer instructive cues to uncommitted and bipotent cells to differentiate into committed cells, which then terminally differentiate into erythrocytes, megakaryocytes, neutrophils, and macrophages, respectively. Lineage-specific transcription factors GATA-1, PU.1, T-bet, and GATA-3 orchestrate the differentiation program of erythrocytes, neutrophils, Th1, and Th2 cells, respectively, by regulating the expression of their lineage-specific genes. Transcription factors GATA-1 and PU.1 have been shown to autoregulate their gene expression by binding to the promoter region of their own genes. Erythrocytic transcription factor GATA-1 has been shown to transactivate the Epo receptor (EPOR) gene and the neutrophilic transcription factor PU.1 has been observed to regulate the expression of the GCSF receptor (GCSFR). A transcription factor can prevent another transcription factor from binding to DNA either by competitively binding to response elements (as in the case of GATA-1 and GATA-2) or by binding to the DNA-binding domain of the transcription factor itself (for example, GATA-1 and PU.1).

The topology shown in Figure 5.1 represents a generalized minimal network of these observed connections between the cytokine and lineage-specific transcription factor

during lineage commitment. The model assumes that the fate decision of an uncommitted cell to either lineage A or lineage B is determined solely by the concentrations of the active forms of the respective lineage-specific transcription factors, ATF_A and ATF_B. The components that drive the formation of each ATF are the inactive transcription factor (ITF), which serves as the substrate, and the ligand (L)-receptor (R) complex (C), which serves as the enzyme. The strong upregulation of ATF during lineage commitment is achieved through two positive feedback loops that upregulate ITF and R, respectively. Transcription factor feedback is a cell-intrinsic autofeedback loop and receptor feedback is an externally (ligand) regulated positive feedback loop. F_{1A} and F_{2A} (expressed in molecules/min) denote the strengths of the transcription factor and receptor feedback loop for lineage A, respectively; F_{1B} and F_{2B} represent the corresponding feedback strengths for lineage B. During commitment, a lineage-specific transcription factor gets upregulated and other lineage transcription factors get downregulated due to cross-antagonism. The mechanism of cross-antagonism between the transcription factors is modeled to be competitive inhibition in binding to response elements present upstream of the transcription factor and receptor genes, thereby affecting the strengths of the two positive feedback loops. While cell fates are likely to represent high dimensional attractors and this higher level of complexity is not considered here, our minimal model framework may be useful in elucidating the interplay among extrinsic and intrinsic factors in lineage commitment and differentiation. The deterministic (ordinary differential equations) and the stochastic (probability functions) versions of the model along with the kinetic parameters and initial conditions are provided in Tables 5.1-5.3.

5.3.2. Double positive feedback loops, coupled with moderate transcriptional cross-antagonism, can lead to multistability

To explore the role of the two positive feedback loops in lineage commitment, we first considered the case with no competitive inhibition between the transcription factors. The inhibitor dissociation constants K_{IA} (inhibitory effect of A on B) and K_{IB} (inhibitory effect of B on A) are kept infinite. Figure 5.2A shows the steady-state values of ATF_A as the strength of two autofeedback loops, F_{1A} and F_{1B}, are changed. The strengths of the receptor-mediated feedback loops and the ligand levels are kept constant ($F_{2A} = F_{2B} = 3$ molecules/min, $L_A = L_B = 100$ molecules). We can see that the system rests in the uncommitted state when $F_{1A} = 0$ for the chosen F_2 and L values. As we increase F_{1A}, the system switches to the on-state (committed state) for lineage A. Since, F_1 constitutes the strength of the autofeedback loop in A, increasing F_{1A} over the threshold value will increase the set point of ATF_A in the on-state, provided F_{2A} is not limiting. To consider the effect of receptor-mediated feedback on the steady-state values of ATF_A, the strength of the autofeedback loops and ligand are kept constant ($F_{1A} = F_{1B} = 3$ molecules/min, $L_A = L_B = 100$ molecules). Similar to F_{1A}, there seems to be a critical value for F_{2A} at which the system switches to the on-state (Figure 5.2B). As F_2 controls the activation loop, increasing F_{2A} beyond the critical level will not change the on-state set point value of ATF_A, provided F_{2A} is not limiting. As expected, F_{1B} and F_{2B} have no effect on ATF_A since we have assumed no crosstalk between the two pathways.

The above analysis was repeated with moderate inhibition ($K_{IA} = K_{IB} = 400$ molecules). Similar to the no inhibition case, there appear to be critical values for F_{1A} and F_{2A} at which the system switches to the on-state (Figures 5.2C and 5.2D). However,

increasing F_{1B} and F_{2B} increases the switching values of F_{1A} and F_{2A}, due to the negative feedback from ATF_B on ATF_A. It is interesting to note that for high values of F_{1B} and F_{2B}, the system reaches a stable, intermediate state at which the concentration of ATF_A is higher than that in the uncommitted state, but less than that in the committed state (by symmetry, the same effect is observed for ATF_B; see Figure 5.3). As in the committed state, the set point in this intermediate state increases with F_1, but not with F_2. To better visualize the intermediate state, cross-sections of F_{1B} and F_{2B} from Figures 5.2C and 5.2D for various values of F_{1A} and F_{2A} are given in Figures 5.2E and 5.2F, respectively. For strong inhibition ($K_{IA} = K_{IB} = 50$ molecules), the system achieves commitment to lineage A for F_{1A} and F_{2A} values above the threshold levels. When F_{1B} and F_{2B} are increased over the critical value, the system requires concomitantly larger increases in F_{1A} and F_{2A} values to switch from the uncommitted state compared to the moderate inhibition condition (Figures 5.2G and 5.2H). Also, strong mutual inhibition between the transcription factors destroys the stable intermediate state, so the cells can rest only in the uncommitted or committed state. Since the model is symmetric with respect to lineages A and B, the steady-state responses of ATF_B with respect to changes in F_1 and F_2 are analogous to the results shown for ATF_A (see Figure 5.3). It should be noted that the system is capable of achieving multistability for a given F_1 and F_2 (results not shown); however, only the stable solution attained without the memory of strong feedback (i.e., the simulations were always started from the off-state) is plotted in Figure 5.2.

5.3.3. "Bilayer" memory in a tristable system

External regulation provides a practical way to control the dynamics of the network without the need to alter the internal control elements of the system. We analyzed how

cell commitment might be influenced in the presence of conflicting ligands with the strength of the positive feedback loops held constant ($F_1 = F_2 = 3$ molecules/min) for the moderate inhibition case. As seen from the phase plots in Figure 5.4, increasing L_A for low L_B cases commits the uncommitted cell to lineage A (red region in Figure 5.4A), increasing L_B for low L_A commits the cell to lineage B (red region in Figure 5.4B), and for high values of both L_A and L_B the system rests at a third, bipotent state that is primed but not committed to either of the lineages (overlapping yellow regions in Figures 5.4A and 3B). For low L_A and L_B (both less than ~40 molecules), the system remains in the uncommitted state (overlapping blue regions in Figures 5.4A and 5.4B).

To explore the robustness of the bipotent and committed states, we tested the system for memory to external stimulus. From the phase plots, we chose $L_B = 300$ to analyze the robustness of the bipotent state. The steady-state response plots of ATF_A and ATF_B for $L_B = 300$ are given in Figures 5.4C and 5.4D. In Figure 5.4C, increasing L_A switches the system from the committed B state to the bipotent state (solid red line). After reaching the bipotent state, the ligand concentration can be decreased far below the initial switching concentration while still maintaining the system in the bipotent state (dotted red line). However, complete removal of L_A switches the system back to the committed B state. For the ligand concentrations spanned by the dotted red line, the system is bistable. Considering the steady-state response of ATF_B in the same simulation, we see that for low L_A values, the system is already committed to lineage B (Figure 5.4D). However, increasing L_A can decommit the cell to a bipotent state (solid blue line). Decreasing L_A after reaching the bipotent state maintains the cell in that state for values of L_A much lower than the decommitment concentration (dotted blue line). So, the system is also

bistable for ATF_B expression (inversely correlated to ATF_A expression) and can exist either in the committed state for lineage B or in the bipotent state based on the memory of L_A.

To analyze the switching of the system across three states, we chose $L_B = 100$ based again upon the phase plots in Figures 5.4A, B. In Figure 5.4E, a modest increase in L_A switches the system to the bipotent state and a further increase in L_A, switches the system to the committed A state (solid red line). If the ligand concentration is lowered after the system reaches either the bipotent state or the committed state, the system remains in the current state (dotted and dot-dash red lines). This hysteresis is greater for the committed state than for the bipotent state, suggesting that the committed state is more robust to changes in the ligand concentration. For $L_B = 100$ and for $10 < L_A < 75$, the system exhibits tristability (i.e., it can exist in committed state A, committed state B, or the bipotent state). The steady-state response plot of ATF_B for $L_B = 100$ (Figure 5.4F) shows that a committed B cell decommits to the bipotent state and then further to lineage A with an increase in L_A (solid blue line). As in Figure 5.4E, the bipotent and lineage A states are robust with respect to decreases in L_A (dotted and dot-dash blue lines) and the system exhibits tristability for the same concentration range of L_A as in Figure 5.4E. It should also be noted that the ligand-dependent multistability seen for a lineage-specific transcription factor is the same for the corresponding lineage-specific receptor, thus simultaneously generating memory in cell-extrinsic and cell-intrinsic signals.

5.3.4. Extrinsic cues can regulate stochastic switching

We developed a stochastic version of the ordinary differential equation (ODE)-based deterministic model to analyze how noise in the network might affect the fate decision of

an uncommitted cell (i.e., one that initially contains no ATF_A or ATF_B) and how external signals might regulate these stochastic transitions. The stochastic model was initialized with several $L_A|L_B$ combinations (0|350; 100|250; 175|175; 250|100; 350|0) for the no inhibition, moderate inhibition, and strong inhibition conditions. In each of 10,000 simulations, the system was allowed to reach steady-state and steady-state ATF_A and ATF_B levels for the first three ligand combinations listed above are shown as 3D histograms in Figure 5.5 (since the model is symmetric, the 250|100 and 350|0 plots are virtual mirror images of the 100|250 and 0|350 plots, respectively, in Figure 5.5). Unlike the deterministic model, which only provided a population average of the four attainable steady states (uncommitted, bipotent, lineage A, lineage B) for any $L_A|L_B$, the stochastic simulations elucidated the relative populations of these multiple steady states for a given $L_A|L_B$. For the no inhibition condition, an uncommitted cell can reach any of four distinct stable states given the appropriate extracellular cues: uncommitted, A, B, and a committed AB state with high ATF_A and ATF_B values (though this last state is simply a consequence of having no inhibition and likely has little relevance in biological mechanisms specific to cell commitment decisions). When ATF_A and ATF_B can moderately inhibit each other, the uncommitted, A, B, and bipotent states can all be populated, even for a single $L_A|L_B$ combination (e.g., middle plot in Figure 5.5). However, when the transcription factors exhibit strong cross-antagonism, this bipotent state is no longer realizable and cells only commit fully to one lineage or the other or stayed uncommitted. The stochastic simulations with various combinations of conflicting ligand concentrations and for different levels of competitive inhibition show that all of the populations obtained from the deterministic model are stable and distinct even with

the introduction of noise. For conditions in which only one ligand was present (e.g., 0|350), the cells committed only to the induced lineage for all levels of inhibition. A small fraction of the initial population remained uncommitted for all conditions for the chosen steady-state time point. When external cues of equal strength were provided (175|175), cells in the absence of inhibition primarily reached the committed AB state; with strong inhibition, they attained nearly equal levels of the committed A and B states; and with moderate inhibition, the cells were roughly evenly distributed across the bipotent, A, and B states. When high but unequal ligand levels were used (e.g., 100|250), cells in the no inhibition model commit almost exclusively to the AB state since the effects of L_A and L_B are entirely uncoupled. However for the strong and moderate inhibition conditions, the initial population committed predominantly to the lineage corresponding to the higher ligand value. This shows that, while the noise in the system is capable of distributing the initial population to all available steady states for any ligand concentration above a minimum threshold, a dominant external signal can still strongly bias the system to its specific lineage.

5.3.5. Time trajectories during lineage commitment

From 100 individual stochastic trajectories, we calculated the average time for an uncommitted cell to reach lineage A, lineage B, or the bipotent state and, in separate simulations, the average time for a bipotent progenitor to reach lineage A or B. A phase plot of the total transcription factors (tTF = ITF + ATF) shows that it takes ~36 hours for the uncommitted cell to reach lineage A, lineage B, or the bipotent state; however, when ligand concentrations that destabilize the bipotent progenitor are applied, it only takes ~24 hours for the bipotent progenitor to reach either of the committed states (Figure

5.6A). This effect is even more pronounced when we look at the phase plots for ATF (Figure 5.6B); the time to reach the high level of active transcription factor(s) from the uncommitted cell is still ~36 hours, however it takes much less time (~14 hours) for the bipotent progenitor to reach lineage A or B. The kinetics of reaching new steady-state levels for total receptor (tR = R + C) and complex (Figures 5.6D and 5.6E) are faster than those for tTF and ATF, respectively, but the trend of reaching commitment faster from the bipotent state compared to the uncommitted state is similar to the transcription factor plots. Figures 5.6C and 5.6F respectively show the mean C_A and ATF_A values with respect to time (in hours) for transitions from the uncommitted state to lineage A (blue line), uncommitted state to bipotent state (orange line), and bipotent state to lineage A (green line). The error bars represent the standard deviation of the trajectories from the mean values. The red lines show the decreases in C_B and ATF_B as the bipotent cell follows the trajectory to commit to lineage A. A primed bipotent cell reaches either committed state faster than an uncommitted cell does, primarily due to the fact that accumulation of new transcription factor molecules (protein synthesis) is a much slower process than deactivation of existing active transcription factor molecules. Furthermore, cytokine signaling has been shown to accelerate differentiation, so the dynamics of activated receptors and transcription factors are likely to influence the kinetics of differentiation.

5.3.6. Comparison to experiments

Figure 5.7A shows a widely accepted branching diagram for differentiation from the common myeloid progenitor (CMP). CMPs undergo lineage-restricted differentiation to form either granulocyte-macrophage progenitors (GMPs) or megakaryocyte-erythrocyte

progenitors (MEPs). GMPs give rise to neutrophils or macrophages, whereas MEPs differentiate into megakaryocytes or erythrocytes. It has been recently demonstrated that alternative routes of differentiation are possible in hematopoiesis: HSCs and multipotent progenitors can bypass canonical intermediate states in reaching mature states, suggesting that these lineage-restricting steps may be more complex than a series of simple binary decisions. We have shown alternative trajectories analogous to these in Figure 5.7A (gray arrows). In Figure 5.7B, the light green and light red lines represent 200 individual stochastic trajectories from the strong inhibition model that committed to lineage A and lineage B, respectively. The dark green and dark red lines show the average of these trajectories. As the strong inhibition model cannot generate a bipotent state, all of the trajectories are directed towards single-lineage populations (A or B). In Figure 5.7C, the light blue, gray, and light red lines denote 200 individual stochastic trajectories from the moderate inhibition model that committed to lineage A, the bipotent state, and lineage B, respectively. The dark blue line denotes the average of all trajectories committing to either lineage A or the bipotent state. The dark red line denotes the average of all trajectories committing to either lineage B or the bipotent state.

To qualitatively compare the receptor dynamics predicted by our model to those seen in experiments, we compared our simulations to lineage-specific receptor expression from microarray data[40] (graciously provided by Tariq Enver, University of Oxford); the data were collected at multiple time points during differentiation of multipotent myeloid progenitors (FDCP-mix, which are CMP-like progenitors) across three lineages (neutrophil, erythrocyte, and megakaryocyte)[40]. The relative mRNA expression levels of the lineage-specific receptors – erythropoietin receptor (EPOR), granulocyte colony-

stimulating factor receptor (GCSFR), and thrombopoietin receptor (TPOR) – were extracted from the processed microarray data. Since the phenotypic heterogeneity was also quantified at each time point in these microarray experiments, we were able to perform a simple deconvolution to estimate the contribution of each distinct cell type to the overall signal (see Table 5.4). Therefore, the receptor expression trajectory for a given lineage in Figures 5.7D-F represents the average of only those multipotent, bipotent, and committed cells that lie along that specific lineage path (as is also the case for the average computational trajectories shown in bold lines in Figures 5.7B-C) and excludes those cells that belong to other commitment paths (for example, the TPOR trajectory includes blasts and megakaryocytes, but excludes erythroblasts, erythrocytes, and neutrophils which were also present in the *in vitro* cultures used for microarray analysis). The level of receptor was normalized to the basal levels in the CMP state. The error bars show the standard error of the mean from three independent experiments.

We constructed phase plots of EPOR and GCSFR showing the receptor trajectories (t = 0 to 7 days) as CMPs differentiate into either erythrocytes or neutrophils (Figure 5.7D). Induction of CMPs with EPO or GCSF drives cell commitment to the erythrocytic (red line) or the neutrophilic (green line) lineage, respectively. During neutrophil commitment, GCSFR expression is significantly upregulated, but EPOR expression stays at or below basal levels; conversely, during erythrocyte commitment, EPOR expression is increased and GCSFR expression is unchanged or slightly reduced. Figure 5.7E shows the experimental phase plot of TPOR and GCSFR expression when CMPs are induced to differentiate into megakaryocyte or neutrophil lineages by stimulating with TPO (blue line) and GCSF (green line), respectively. As in Figure 5.7D,

receptor expression corresponding to the induced lineage is upregulated and the receptor expression corresponding to the other lineage is unchanged or even slightly downregulated. Figure 5.7F shows the phase plot from the differentiation experiments to erythrocytic and megakaryocytic lineages. Induction of CMPs with Epo or Tpo drives CMPs to either the erythrocytic (red line) or the megakaryocytic (blue line) lineage. Interestingly, during erythrocyte and megakaryocyte commitment, EPOR and TPOR are co-upregulated; however, the observed increase was higher for the receptor corresponding to the specific lineage that was predominantly generated. Statistical analysis was performed to deduce positive receptor correlation for the receptor pairs in Figures 5.7D, 5.7E, and 5.7F by comparing the overall slope of each trajectory (inverted to lie along the x-axis, if appropriate) at both the 3-day and 7-day time points to a value of zero (no correlation) by a one-sample, one-tailed t-test. The correlation in receptor expression for EPOR-GCSFR and TPOR-GCSFR was either negative or not statistically significant. However, the EPOR-TPOR receptor pair showed a positive correlation with statistical significance. The symbols in Figure 5.7F denote the 3-day (†, *) and 7-day (‡, #) time points during erythrocyte and megakaryocyte differentiation from the CMP (p-values: † (0.027), * (0.009), ‡ (0.060), # (0.008)).

Comparing experimental results to the model simulations, we note that the trajectories in the erythrocyte-neutrophil (Figure 5.7D) and neutrophil-megakaryocyte (Figure 5.7E) plots compare well with the strong inhibition model (Figure 5.7B) and the trajectories from the erythrocyte-megakaryocyte plot (Figure 5.7F) show agreement with the moderate inhibition model (Figure 5.7C). This inference is validated by the widely accepted observation that the transcription factors for the erythrocytic and

megakaryocytic lineages are strongly cross-antagonistic to the transcription factor for the neutrophil lineage. Other than evolutionary constraints, the model suggests that the strength of the transcriptional cross-antagonism can dictate whether two distinct lineage-specific receptors (and the corresponding lineage-specific transcription factors) can be co-upregulated, which in turn can influence the nature of the instructive, possibly conflicting, cues that the cell receives. This paradigm may highlight different modes of receptor regulation, and corresponding transcriptional activity, in various stages and branches of hematopoiesis (e.g., Figure 5.7A).

5.4. Discussion

Mathematical models of lineage commitment during hematopoiesis have generally analyzed cell-fate decisions from an intrinsic standpoint. Here, we show how extrinsic regulation can play a role in instructing lineage choice and, furthermore, how a cell might process and respond to conflicting extracellular cues. It has been extensively debated whether cytokines play an instructive or permissive role during lineage commitment. In this work, we have assumed that cell-fate decisions can be stochastic but that external cues can strongly bias this stochasticity and instruct cells to specific lineages. A recent publication definitively demonstrated an instructive role for cytokines in hematopoiesis[20]. This strongly underscores our need to understand how extracellular cues, either in isolation or in combination, influence hematopoiesis. Our model also suggests a possible alternative mode of commitment, whereby an uncommitted multipotent progenitor may

commit directly to a mature lineage without transitioning through a bipotent state. This potential plasticity has been seen experimentally in HSCs[4] and multipotent progenitors[2].

The initial cell state that is modeled here is a common multipotent progenitor that expresses multiple lineage-specific receptors and transcription factors at low levels and is capable of differentiating along several lineages. In particular, two lineages that may exhibit different levels of transcriptional cross-antagonism are analyzed. The lineage commitment decision is modeled to be driven by the accumulation of the functionally active form of the lineage-specific transcription factor. This event is driven through two positive feedback loops, a synthesis loop that produces the transcription factor and a regulatory loop that aids in the activation of the transcription factor. This two-step positive feedback mechanism provides a means to externally regulate the classical autofeedback loop and can be of general significance in cell-fate decision models. In our lineage commitment model, the regulatory loop targets the cell-surface receptor, but analogous topologies may be seen in systems where the regulation is achieved extracellularly (upregulating the ligand) or intracellularly (upregulating a rate-limiting enzyme in the signaling pathway). Also, it should be noted that even though we have considered the external stimuli to be cytokines, they may also be cell-cell interactions, cell-matrix interactions, mechanical cues, or other diffusible factors.

Through steady-state response plots, we have shown that the system exhibits ultrasensitivity to ligand and can achieve multistability in active transcription factor levels (Figure 5.4). Here, ultrasensitivity to ligand confers switch-like behavior in cell-fate specification. Multistability provides memory to both the intermediate (bipotent) and committed cell states, enabling the system to robustly sustain its current state even when

external stimuli to sub-threshold levels after switching. Although the system modeled here represents a reversible switch, irreversibility during differentiation can be achieved by epigenetic means such as chromatin remodeling.

Hematopoiesis is traditionally viewed as a process in which a HSC undergoes several lineage-restricted binary decisions to reach the mature state. However, hematopoiesis may be more fluid than the classical paradigm suggests, as recent experiments have shown that hematopoietic stem cells can completely bypass canonical progenitor states during lineage commitment. While we acknowledge that the classical view of lineage specification invokes binary decisions, it is possible that binary cell fate-choices are a special case of a more general decision-making strategy. For example, in hematopoiesis, the common lymphoid progenitor appears to be oligopotent in producing T cells, B cells, and NK cells without any observed bipotent intermediate. Our network topology is capable of generating both binary and ternary cell-fate decisions, depending on the strength of transcriptional cross-antagonism. For example, in Figure 5.5, strong inhibition enables only a binary cell-fate choice; however, simply relaxing the strength of the inhibition to moderate levels (without changing the topology) enables three possible fates from the uncommitted state.

In support of the stochastic theory of commitment, our model suggests that, irrespective of the strength of external factors, intrinsic noise in transcriptional networks can switch a significant percentage of cells to a committed state or the bipotent state; however, in support of the instructive theory, extrinsic cues can still strongly bias the majority of the uncommitted cell population to the final state induced by the higher ligand signal (e.g., see Figure 5.5). Our model suggests a new paradigm that integrates

classical and alternative modes of lineage commitment and also accommodates both stochastic and instructive roles in hematopoiesis (Figure 5.8). It is generally appreciated that upstream commitment events are more stochastic in nature while downstream events are more instructive. Stochastic events in HSCs and multipotent progenitors can potentially lead to the generation of all mature cell types, explaining 'normal' hematopoiesis even when a lineage-specific receptor is knocked out (although other non-canonical extrinsic cues may also play compensatory roles). In parallel, instructive cytokine signaling in multipotent progenitors and bipotent progenitors, which can strongly bias and accelerate lineage commitment, may drive stress responses and restore homeostasis. Furthermore, emerging alternative commitment paths suggest that decision-making in hematopoietic progenitors need not be purely binary. HSCs have been shown to bypass multipotent progenitors and directly produce bipotent MEPs and common lymphoid progenitors appear to directly generate T cells, B cells, and NK cells. The model presented in this work suggests a framework in which both binary and ternary decisions may be possible in multipotent CMPs. Such bypass mechanisms in commitment may also provide important redundancies that ensure mature cell production if a specific intermediate state becomes dysregulated.

Many of the predictions from our minimal, multipotent commitment model can be experimentally verified. Multipotent and bipotent progenitors can be identified and sorted with multi-color flow cytometry, using specific cell-surface markers for the lineages of interest. Cytokine-induced time course experiments conducted on these bipotent cells can corroborate whether they reach mature states faster than the corresponding multipotent progenitors. Experiments with conflicting extracellular ligand cues can be useful in

determining the strength of the instructive cues, the degree of transcriptional cross-antagonism between lineages, and the existence of a bipotent progenitor. For example, to analyze the differentiation paths of erythrocytes and neutrophils from a common progenitor, FDCP-mix cells can be induced with both Epo and GCSF and the trajectories of the expression of the lineage-specific transcription factors (GATA1, PU.1) and receptors (EPOR, GCSFR) can be determined by sensitive flow cytometry measurements. Furthermore, groundbreaking new bioimaging techniques which enable observation of single cells over an extended period should mitigate technical difficulties that have hampered such analyses and should help to further elucidate the roles of extrinsic and intrinsic regulation on cell commitment decisions.

Figure 5.1

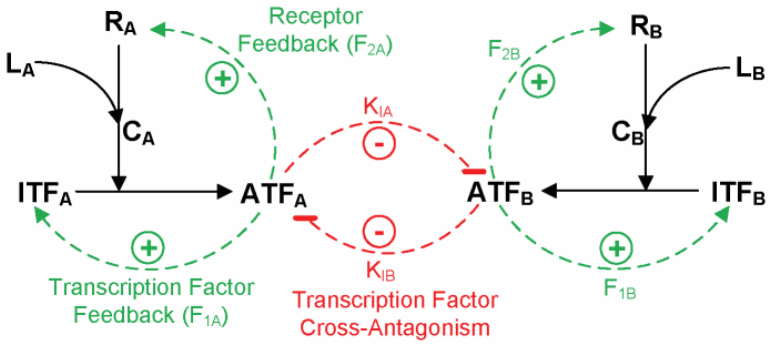


Figure 5-1 A minimal model of multilineage commitment

A multipotent progenitor expresses lineage-specific receptors (R_A and R_B) and inactive transcription factors (ITF_A and ITF_B) at low levels with the potential to differentiate into lineage A or B. Addition of ligand (L_A, L_B) leads to complex formation (C_A, C_B), which activates the corresponding lineage-specific transcription factor. Active TF (ATF_A, ATF_B) binds to the response elements present upstream of the transcription factor and receptor genes and induces two positive feedback loops (dashed green arrows). To account for cross-antagonism between the lineages, the active transcription factors are modeled to competitively inhibit the activation of the positive feedback loops in the other lineage (dashed red lines). F_{1A} and F_{2A} denote the respective strengths of the transcription factor and receptor feedback loops for lineage A; similarly, F_{1B} and F_{2B} represent the corresponding feedback strengths for lineage B. Inhibitor dissociation constants K_{IA} and K_{IB} denote the inhibitory effect of A on B and B on A, respectively.

Figure 5.2

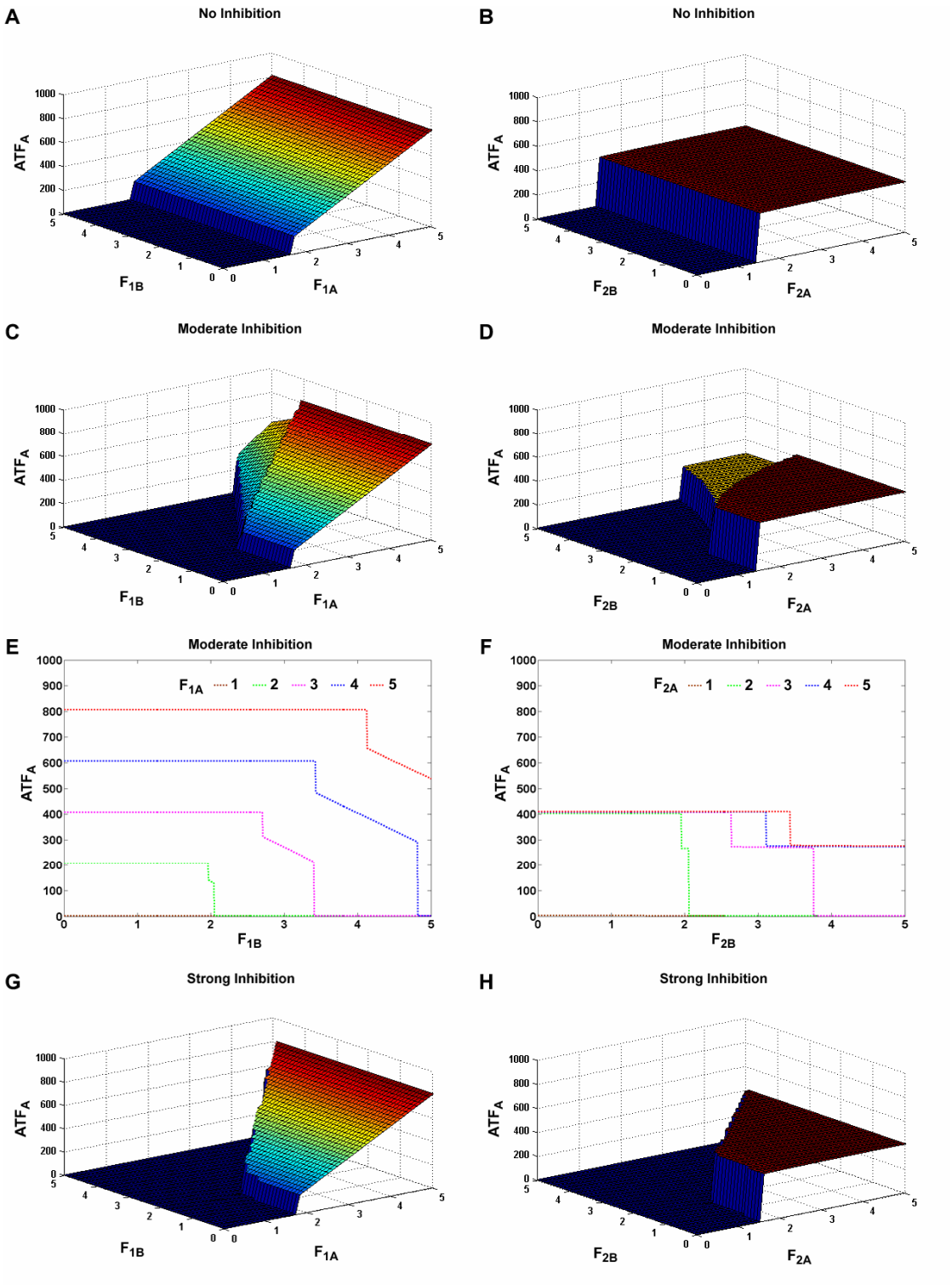


Figure 5-2 Effect of the positive feedback loops on the on-state ATF_A levels

A. Strengths of the autofeedback loops (F_{1A} and F_{1B}) are varied for both lineages and the steady-state values of ATF_A are plotted for the no inhibition condition ($K_{IA} = K_{IB} = \infty$), keeping the strength of the receptor feedbacks (F_2) constant. B. Strengths of the receptor feedback loops (F_{2A} and F_{2B}) are varied and the values of ATF_A are plotted for the no inhibition condition, keeping the strength of the auto feedbacks (F_1) constant. C. Same as part A except with moderate inhibition ($K_{IA} = K_{IB} = 400$ molecules). D. Same as part B except with moderate inhibition. E. Cross-sectional plot from C for various values of F_{1A}. F. Cross-sectional plot from D for various values of F_{2A}. G. Same as part A except with strong inhibition ($K_{IA} = K_{IB} = 50$ molecules). H. Same as part B except with strong inhibition. No inhibition and strong inhibition give rise to only on or off populations, whereas moderate inhibition can generate a third intermediate population.

Figure 5.3

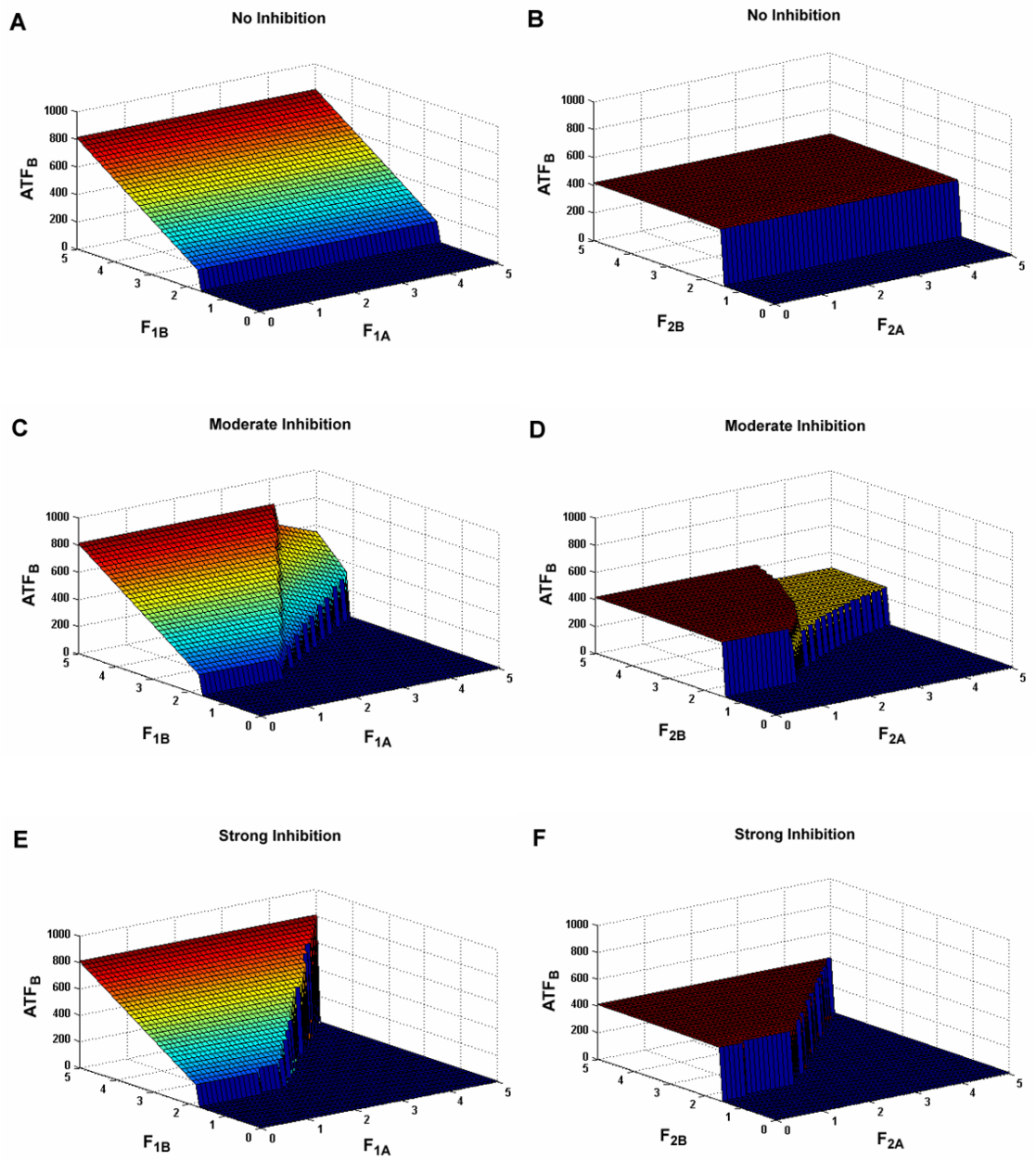


Figure 5-3 Effect of the positive feedback loops on the on-state ATF_B levels

Strengths of the autofeedback loops (F_{1A} and F_{1B}) are varied for both lineages and the steady-state values of ATF_B are plotted for no (A), moderate (C) and strong (E) inhibition, keeping the strength of receptor feedback (F_{2A} and F_{2B}) constant. Strengths of the receptor feedback loops (F_{2A} and F_{2B}) are varied and the values of ATF_B are plotted for no (B), moderate (D) and strong (F) inhibition, keeping the strength of autofeedback (F_{1A} and F_{1B}) constant.

Figure 5.4

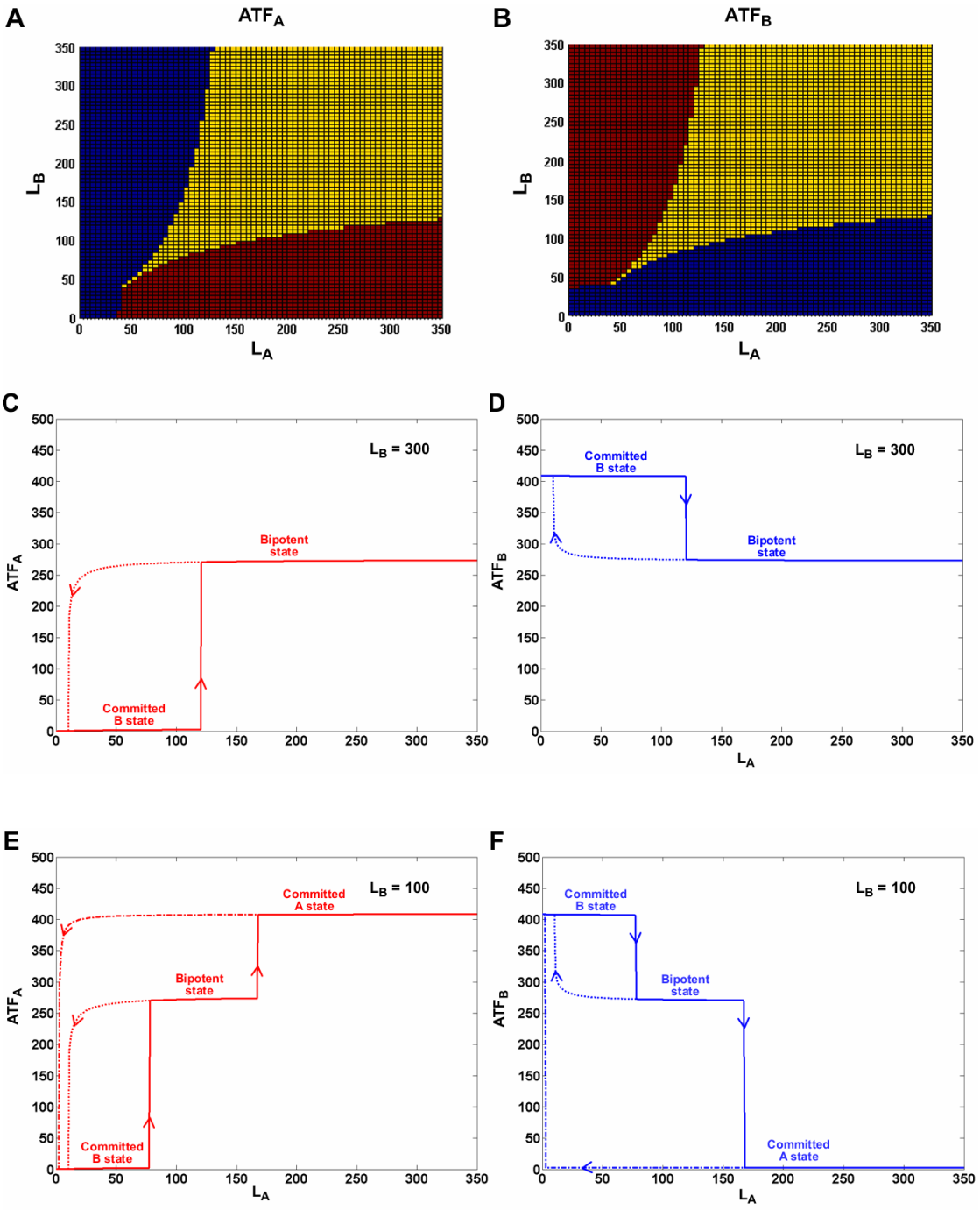


Figure 5-4 Effect of ligand on the on-state ATF levels

A. Phase plot showing the steady-state ATF_A levels (blue – low, yellow – medium, red – high) when L_A and L_B values are varied. B. Phase plot showing the steady-state ATF_B levels when L_A and L_B values are varied. Low L_A and low L_B do not commit the uncommitted cell to either lineage (overlapping blue region in A and B). Low L_A and high L_B values commit the cell to lineage B (blue region in A and red region in B). High L_A and low L_B values commit the cell to lineage A (red region in A and blue region in B). High L_A and high L_B commit the cell to the bipotent state (overlapping yellow region in A and B). Steady-state response plots: C. Increasing L_A from 0, with L_B constant at 300, abruptly switches the cell from the committed B state to the bipotent state (increase in ATF_A to intermediate level) after reaching a threshold concentration (solid red line). After achieving the bipotent state, decreasing L_A to sub-threshold values does not immediately switch the cell state, suggesting significant memory in the system (dotted red line). D. Increasing L_A from 0, with L_B constant at 300, decommits the cell to the bipotent state (decrease in ATF_B to intermediate level) after reaching the threshold concentration (solid blue line). After achieving the bipotent state, decreasing L_A to sub-threshold values does not immediately switch the cell state, again suggesting significant memory (dotted blue line). E. Increasing L_A from 0, with L_B constant at 100, abruptly switches the committed B cell to the bipotent state (increase in ATF_A to intermediate level) and then again to the committed A state (increase in ATF_A to high level) after reaching the corresponding threshold concentrations (solid red line). After achieving the bipotent state or the committed state, decreasing L_A to sub-threshold values does not immediately switch the cell response, suggesting significant memory in both states (dotted and dot-dash red line). F. Increasing L_A from 0, with L_B constant at 100, decommits the cell to the bipotent state (decrease in ATF_B to intermediate level) and then again to the committed lineage A state (decrease in ATF_B to low level) after reaching the corresponding threshold concentrations (solid blue line). After achieving the bipotent state or the committed lineage A state, decreasing L_A to sub-threshold values does not immediately switch the cell response, suggesting significant memory in both states (dotted and dot-dash blue line). Plots C and D show bistable expression of ATF_A and ATF_B; plots E and F exhibit both bistable and tristable expression of the transcription factors.

Figure 5.5

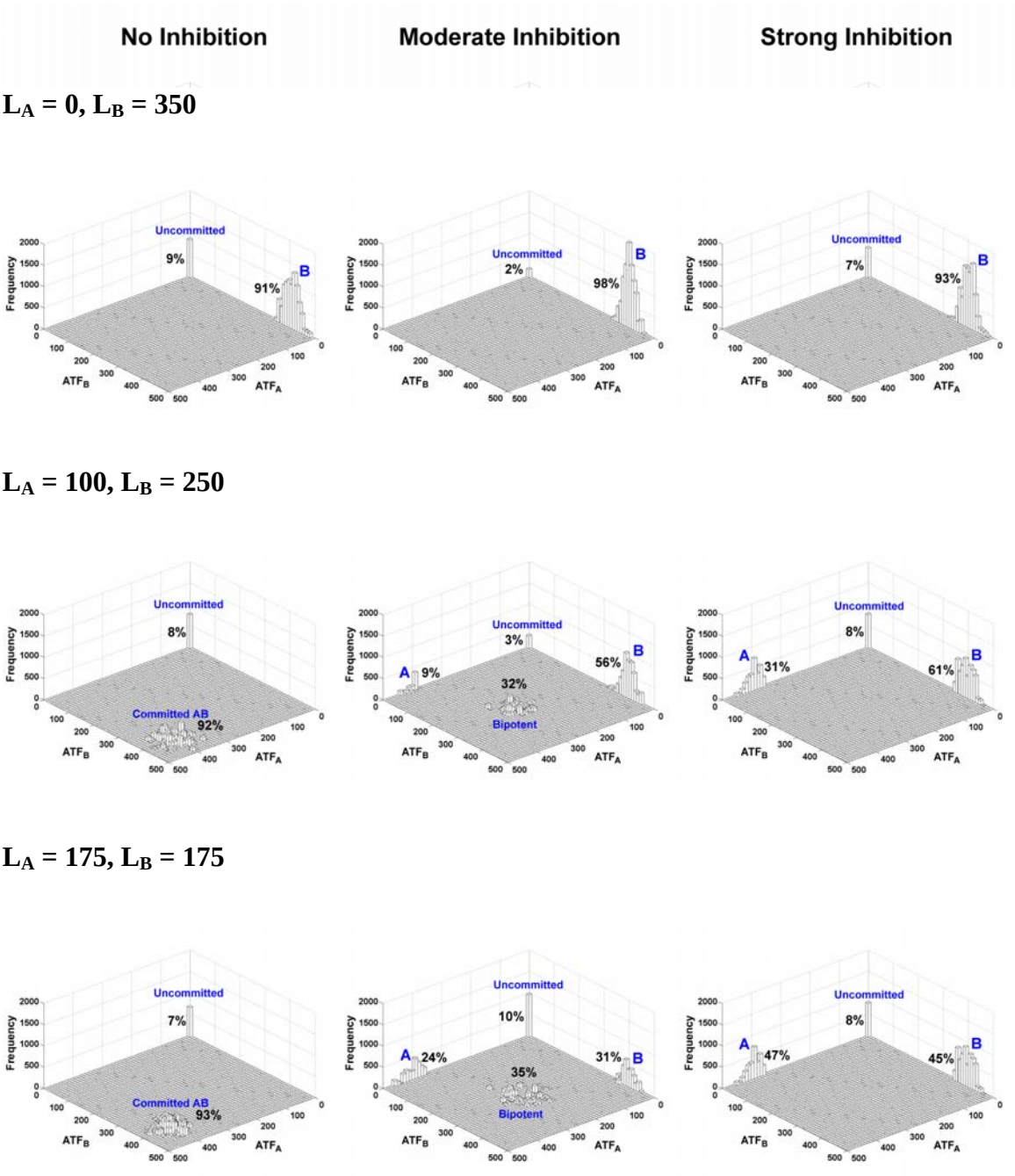


Figure 5-5 External regulation of stochastic transitions

Three different $L_A|L_B$ combinations (0|350, 100|250, and 175|175) were run using the stochastic version of the model with no, moderate, or strong inhibition conditions and the system was allowed to reach steady state. ATF_A and ATF_B values from 10,000 runs for each condition are plotted here as three-dimensional histograms. With strong inhibition, the system cannot achieve the intermediate, bipotent state that is seen with moderate inhibition. When induced with only one ligand (e.g., 0|350), the initial population, for all inhibition conditions, commits predominantly to the lineage corresponding to that ligand. When the uncommitted state is stimulated with equal values of ligand (175|175), the no inhibition condition primarily results in a state that corresponds to high activation of both transcription factors (unlikely to be a biologically relevant state for cell-commitment decisions); the strong and the moderate inhibition conditions result in significant population of all of the available states except the uncommitted state. When one ligand value is higher (e.g., 100|250), in the presence of inhibition, the majority of the cells committed to the lineage corresponding to the higher ligand concentration. The number next to each individual population denotes the percentage of the total population when treated with the given combination of L_A and L_B.

Figure 5.6

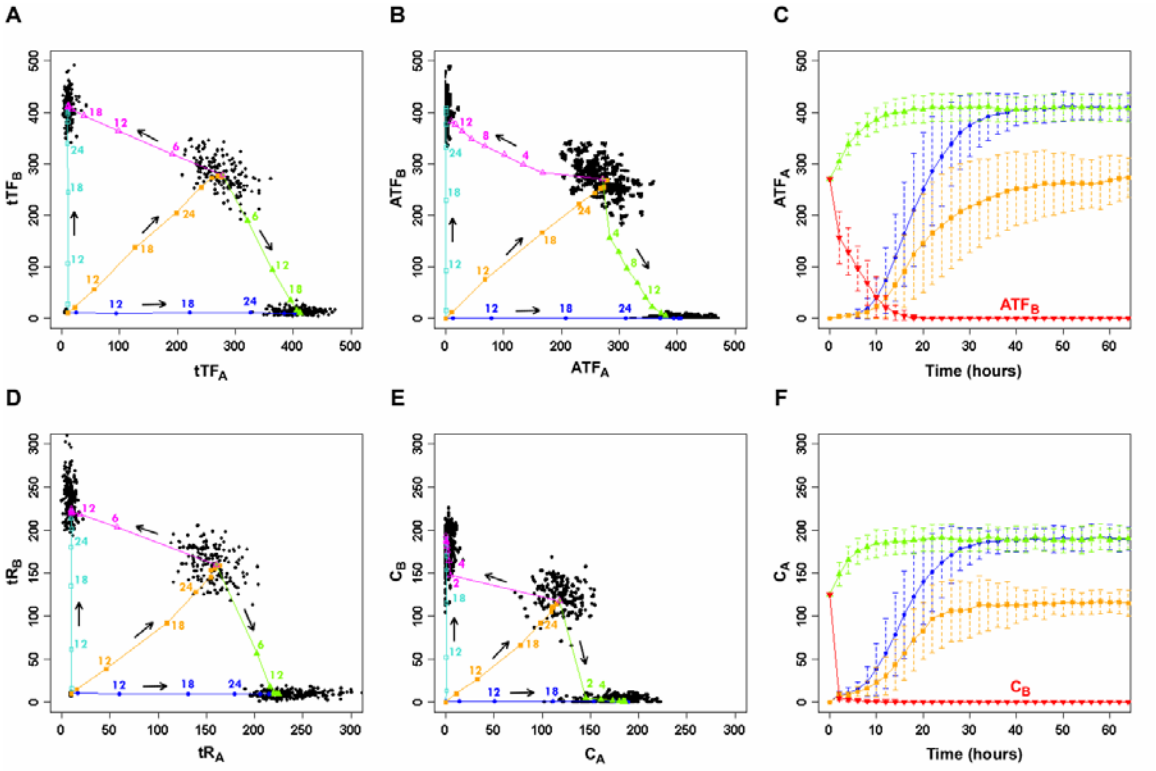


Figure 5-6 Time trajectories during lineage commitment

A. Phase plot of total transcription factor (ITF+ATF) for the four steady-state populations (uncommitted, A, B, and bipotent). B. Phase plot of active transcription factor (ATF). C. Time trajectories for ATF_A in panel B for the transition from the uncommitted cell to committed A state (blue line) and bipotent state (orange line) and from the bipotent state to committed A state (green line) is shown as a time course plot. The error bars represent the standard deviation of the mean. The red line shows the level of ATF_B as the bipotent cell transitions to the committed A state. D. Phase plot of total receptor (R+C). E. Phase plot of active complex (C). F. Time trajectories for C_A in panel E for the transition from the uncommitted cell to committed A state (blue line) and bipotent state (orange line) and from the bipotent state to committed A state (green line) is shown as a time course plot. The error bars represent the standard deviation of the mean. The red line shows the level of C_B as the bipotent cell transitions to the committed A state. In the phase plots, the arrows indicate the direction of commitment (averaged over 200 stochastic runs each): from the uncommitted state, the three possible commitment trajectories lead to pure lineage A, pure lineage B, and the bipotent state. In separate simulations starting with the bipotent state and with initial ligand concentrations sufficient to destabilize this state, the two possible commitment trajectories lead to pure lineage A and pure lineage B. Each trajectory has several nodes and the number at each node denotes the average time (in hours) it takes to reach the node from the initial state. Each black dot in A, B, D and E represents the endpoint (100,000 min) of an individual stochastic trajectory.

Figure 5.7

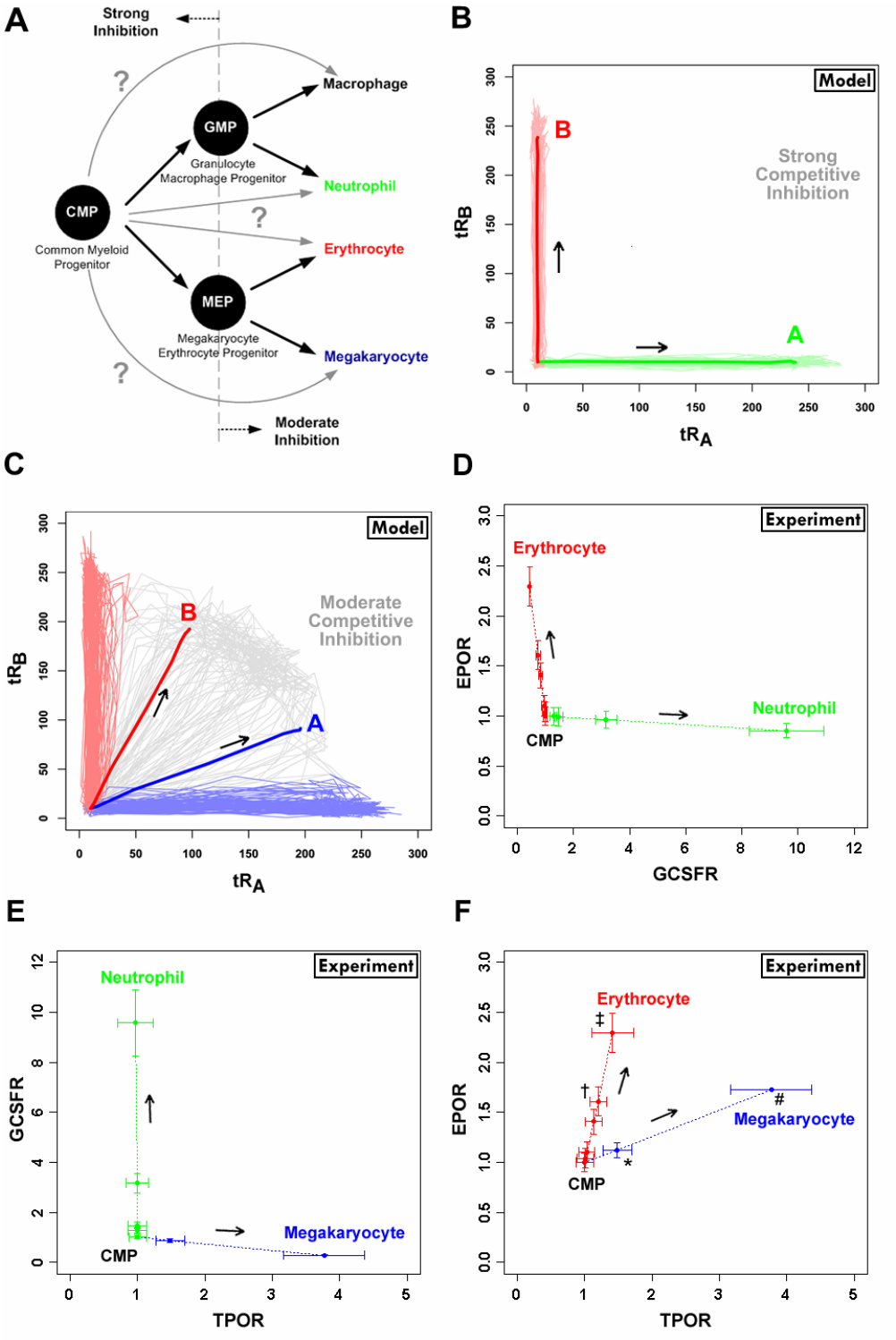


Figure 5-7 Comparison of multilineage commitment model to experimental data

A. The classical model of hematopoiesis is given here as a branching diagram showing the differentiation paths from the common myeloid progenitor (CMP) to four distinct myeloid lineages (megakaryocyte, erythrocyte, neutrophil, and macrophage) via bipotent progenitors (GMP – granulocyte/macrophage progenitor and MEP – megakaryocyte/erythrocyte progenitor). Other possible alternate routes of commitment, bypassing the bipotent state, are shown as gray lines. B. Stochastic simulations of total receptor levels under strong competitive inhibition. Light green and red lines indicate the individual trajectories from the uncommitted cell to lineages A and B, respectively. The dark red and green lines denote the averaged trajectories of all stochastic runs. C. Stochastic simulation for total receptor levels under moderate competitive inhibition condition. Light blue, light red, and gray lines indicate the individual trajectories from the uncommitted cell to A, B, and the bipotent state, respectively. The dark blue line denotes the average value of all stochastic runs that commit to either lineage A or the bipotent state; the dark red line denotes the average value of all stochastic runs that commit to either lineage B or the bipotent state. D. Trajectories from microarray data showing upregulation of EPOR and GCSFR during erythrocyte (red) and neutrophil (green) commitment from the CMP, respectively. E. Trajectories from microarray data showing upregulation of TPOR and GCSFR during megakaryocyte (blue) and neutrophil (green) commitment from the CMP, respectively. F. Trajectories from microarray data showing upregulation of EPOR and TPOR during erythrocyte (red) and megakaryocyte (blue) commitment from the CMP. The data in D-F represent the average of the multipotent, bipotent, and mature cells for any single lineage (see Table S4), thus enabling a direct comparison to the model simulations. The error bars in D-F show the standard error of the mean. The symbols in F denote the 3-day (†, *) and 7-day (‡, #) time points during erythrocyte and megakaryocyte differentiation from the CMP, respectively. Statistical analysis was performed to deduce positive correlation in receptor pair upregulation by comparing the overall slope of each trajectory (inverted to lie along the x-axis, if appropriate) at both the 3-day and 7-day time points to a value of zero (no correlation) by a one-sample, one-tailed t-test (p-values: † (0.027), * (0.009), ‡ (0.060), # (0.008)).

Figure 5.8

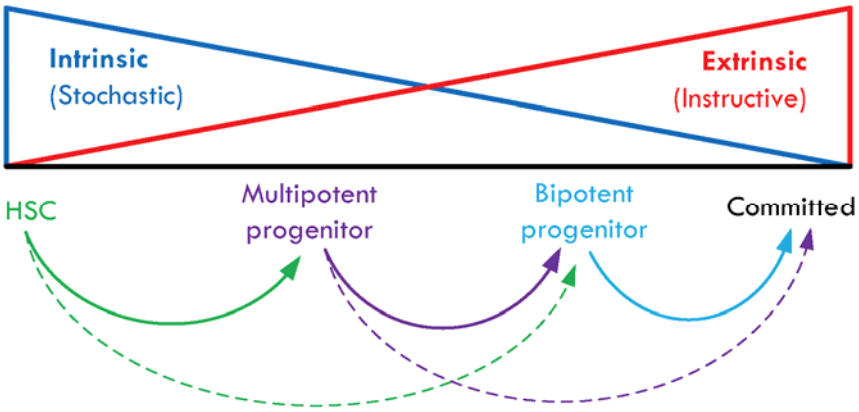


Figure 5-8 Proposed paradigm for hematopoiesis

Extrinsic (instructive) and intrinsic (stochastic) cues can both play roles in commitment of progenitor cells. In addition to classical pathways of commitment (solid arrows), bypass mechanisms have been reported for HSCs (dashed green arrow) and our model suggests that this may be possible for multipotent progenitors as well (dashed purple arrow).

Table 5.1

Table 5-1 Ordinary differential equations for the deterministic model

$$\frac{d(R_A)}{dt} = B_A^R - kdegR_A * R_A - kon_A * L_A * R_A + koff_A * C_A + \frac{F_{2A} * ATF_A}{K_{DA} * \left[1 + \frac{ATF_B}{K_{IB}}\right] + ATF_A}$$

$$\frac{d(C_A)}{dt} = kon_A * L_A * R_A - koff_A * C_A - ke_A * C_A$$

$$\frac{d(ITF_A)}{dt} = B_A^{ITF} - kdegTF_A * ITF_A - \frac{k_{1A} * C_A * ITF_A}{K_{M1A} + ITF_A} + \frac{k_{2A} * P_A * ATF_A}{K_{M2A} + ATF_A} + \frac{F_{1A} * ATF_A}{K_{DA} * \left[1 + \frac{ATF_B}{K_{IB}}\right] + ATF_A}$$

$$\frac{d(ATF_A)}{dt} = \frac{k_{1A} * C_A * ITF_A}{K_{M1A} + ITF_A} - \frac{k_{2A} * P_A * ATF_A}{K_{M2A} + ATF_A} - kdegTF_A * ATF_A$$

$$\frac{d(R_B)}{dt} = B_B^R - kdegR_B * R_B - kon_B * L_B * R_A + koff_B * C_B + \frac{F_{2B} * ATF_B}{K_{DB} * \left[1 + \frac{ATF_A}{K_{IA}}\right] + ATF_B}$$

$$\frac{d(C_B)}{dt} = kon_B * L_B * R_B - koff_B * C_B - ke_B * C_B$$

$$\frac{d(ITF_B)}{dt} = B_B^{ITF} - kdegTF_B * ITF_B - \frac{k_{1B} * C_B * ITF_B}{K_{M1B} + ITF_B} + \frac{k_{2B} * P_B * ATF_B}{K_{M2B} + ATF_B} + \frac{F_{1B} * ATF_B}{K_{DB} * \left[1 + \frac{ATF_A}{K_{IA}}\right] + ATF_B}$$

$$\frac{d(ATF_B)}{dt} = \frac{k_{1B} * C_B * ITF_B}{K_{M1B} + ITF_B} - \frac{k_{2B} * P_B * ATF_B}{K_{M2B} + ATF_B} - kdegTF_B * ATF_B$$

L_A : Ligand for lineage A (constant)
R_A : Lineage-specific receptor for the A lineage
C_A : Ligand-receptor complex for the A lineage
ITF_A : Lineage-specific inactive transcription factor for the A lineage
ATF_A : Active transcription factor for the A lineage
P_A : Phosphatase deactivating ATF_A (constant)

L_B : Ligand for lineage B (constant)
R_B : Lineage-specific receptor for the B lineage
C_B : Ligand-receptor complex for the B lineage
ITF_B : Lineage-specific inactive transcription factor for the B lineage
ATF_B : Active transcription factor for the B lineage
P_B : Phosphatase deactivating ATF_B (constant)

Table 5.2

Table 5-2 Rate constants and initial conditions for the deterministic and stochastic models

Rate constant	Value	Description
B_A^R	0.05 molecules/min	Basal synthesis rate of R_A
$kdegR_A$	0.005 min^{-1}	Degradation rate of R_A
kon_A	0.001 $molecules^{-1}min^{-1}$	Association rate constant for L_A and R_A binding
$koff_A$	0.05 min^{-1}	Dissociation rate constant for L_A and R_A binding
ke_A	0.01 min^{-1}	Endocytic rate constant for C_A
B_A^{ITF}	0.05 molecules/min	Basal synthesis rate of ITF_A
$kdegTF_A$	0.005 min^{-1}	Degradation rate of ITF_A and ATF_A
k_{1A}	60 min^{-1}	Rate of activation of ITF_A
K_{M1A}	100 molecules	Michaelis constant for activation of ITF_A
K_{2A}	30 min^{-1}	Rate of deactivation of ATF_A
K_{M2A}	30 molecules	Michaelis constant for deactivation of ATF_A
F_{1A}	3 molecules/min	Strength of transcription factor feedback for lineage A
F_{2A}	3 molecules/min	Strength of receptor feedback for lineage A
K_{DA}	200 molecules	Equilibrium dissociation constant for ATF_A binding to DNA
K_{IB}	400 molecules	Inhibitor dissociation constant (effect of B on A)
B_B^R	0.05 molecules/min	Basal synthesis rate of R_B
$kdegR_B$	0.005 min^{-1}	Degradation rate of R_B
kon_B	0.001 $molecules^{-1}min^{-1}$	Association rate constant for L_B and R_B binding
$koff_B$	0.05 min^{-1}	Dissociation rate constant for L_B and R_B binding
ke_B	0.01 min^{-1}	Endocytic rate constant for C_B
B_B^{ITF}	0.05 molecules/min	Basal synthesis rate of ITF_B
$kdegTF_B$	0.005 min^{-1}	Degradation rate of ITF_B and ATF_B
k_{1B}	60 min^{-1}	Rate of activation of ITF_B
K_{M1B}	100 molecules	Michaelis constant for activation of ITF_B
K_{2B}	30 min^{-1}	Rate of deactivation of ATF_B
K_{M2B}	30 molecules	Michaelis constant for deactivation of ATF_B
F_{1B}	3 molecules/min	Strength of transcription factor feedback for lineage B
F_{2B}	3 molecules/min	Strength of receptor feedback for lineage B
K_{DB}	200 molecules	Equilibrium dissociation constant for ATF_B binding to DNA
K_{IA}	400 molecules	Inhibitor dissociation constant (effect of A on B)

Phosphatases P_A and P_B are held constant at 15 molecules for all simulations.

Initial conditions when starting from the uncommitted or off-state (in molecules):
$[R_A]_0 = 10$; $[R_B]_0 = 10$; $[C_A]_0 = 0$; $[C_B]_0 = 0$; $[ITF_A]_0 = 10$; $[ITF_B]_0 = 10$; $[ATF_A]_0 = 0$; $[ATF_B]_0 = 0$

Initial conditions when starting from the bipotent or intermediate state (in molecules):
$[R_A]_0 = 30$; $[R_B]_0 = 30$; $[C_A]_0 = 125$; $[C_B]_0 = 125$; $[ITF_A]_0 = 6$; $[ITF_B]_0 = 6$; $[ATF_A]_0 = 273$; $[ATF_B]_0 = 273$

Table 5.3

Table 5-3 Probability functions and reactions for the stochastic model

$\xrightarrow{P_1} R_A$	$P_1 = B_A^R$
$R_A \xrightarrow{P_2}$	$P_2 = kdegR_A * R_A$
$R_A \xrightarrow{P_3} C_A$	$P_3 = kon_A * L_A * R_A$
$C_A \xrightarrow{P_4} R_A$	$P_4 = koff_A * C_A$
$C_A \xrightarrow{P_5}$	$P_5 = ke_A * C_A$
$\xrightarrow{P_6} ITF_A$	$P_6 = B_A^{ITF}$
$ITF_A \xrightarrow{P_7}$	$P_7 = kdegTF_A * ITF_A$
$ITF_A \xrightarrow{P_8} ATF_A$	$P_8 = \frac{k_{1A} * C_A * ITF_A}{K_{M1A} + ITF_A}$
$ATF_A \xrightarrow{P_9} ITF_A$	$P_9 = \frac{k_{2A} * P_A * ATF_A}{K_{M2A} + ATF_A}$
$ATF_A \xrightarrow{P_{10}}$	$P_{10} = kdegTF_A * ATF_A$
$\xrightarrow{P_{11}} R_A$	$P_{11} = \frac{F_{2A} * ATF_A}{K_{DA} * \left[1 + \frac{ATF_B}{K_{IB}}\right] + ATF_A}$
$\xrightarrow{P_{12}} ITF_A$	$P_{12} = \frac{F_{1A} * ATF_A}{K_{DA} * \left[1 + \frac{ATF_B}{K_{IB}}\right] + ATF_A}$
$\xrightarrow{P_{13}} R_B$	$P_{13} = B_B^R$
$R_B \xrightarrow{P_{14}}$	$P_{14} = kdegR_B * R_B$
$R_B \xrightarrow{P_{15}} C_B$	$P_{15} = kon_B * L_B * R_B$
$C_B \xrightarrow{P_{16}} R_B$	$P_{16} = koff_B * C_B$
$C_B \xrightarrow{P_{17}}$	$P_{17} = ke_B * C_B$
$\xrightarrow{P_{18}} ITF_B$	$P_{18} = B_B^{ITF}$
$ITF_B \xrightarrow{P_{19}}$	$P_{19} = kdegTF_B * ITF_B$
$ITF_B \xrightarrow{P_{20}} ATF_B$	$P_{20} = \frac{k_{1B} * C_B * ITF_B}{K_{M1B} + ITF_B}$
$ATF_B \xrightarrow{P_{21}} ITF_B$	$P_{21} = \frac{k_{2B} * P_B * ATF_B}{K_{M2B} + ATF_B}$
$ATF_B \xrightarrow{P_{22}}$	$P_{22} = kdegTF_B * ATF_B$
$\xrightarrow{P_{23}} R_B$	$P_{23} = \frac{F_{2B} * ATF_B}{K_{DB} * \left[1 + \frac{ATF_A}{K_{IA}}\right] + ATF_B}$
$\xrightarrow{P_{24}} ITF_B$	$P_{24} = \frac{F_{1B} * ATF_B}{K_{DB} * \left[1 + \frac{ATF_A}{K_{IA}}\right] + ATF_B}$

Table 5.4

Table 5-4 Parameter fitting of microarray data

GENE	Blast	Neutrophil	Erythroblast	Erythrocyte	Megakaryocyte
PU.1	1.000	4.009	0.379	0.250	0.000
GATA1	1.000	0.693	4.233	1.801	1.414
EPOR	1.000	0.837	3.021	2.028	2.014
TPOR	1.000	0.967	1.679	1.317	4.875
GCSFR	1.000	10.425	0.000	0.644	0.000

GENE	Blast	Neutrophil	Erythroblast	Erythrocyte	Megakaryocyte
PU.1	1.104	4.071	0.323	0.383	0.000
GATA1	1.172	0.754	4.464	2.111	1.545
EPOR	1.092	0.906	3.311	2.187	1.979
TPOR	1.133	1.240	1.477	1.822	5.662
GCSFR	1.071	11.875	0.503	0.377	0.000

GENE	Blast	Neutrophil	Erythroblast	Erythrocyte	Megakaryocyte
PU.1	0.896	3.947	0.434	0.118	0.000
GATA1	0.828	0.631	4.001	1.491	1.283
EPOR	0.908	0.769	2.731	1.869	2.048
TPOR	0.867	0.694	1.881	0.812	4.088
GCSFR	0.929	8.964	0.000	0.626	0.000

5.5. References

1. Eckfeldt, C.E., Mendenhall, E.M. & Verfaillie, C.M. The molecular repertoire of the 'almighty' stem cell. *Nature reviews.Molecular cell biology* **6**, 726-737 (2005).
2. Kondo, M., Weissman, I.L. & Akashi, K. Identification of clonogenic common lymphoid progenitors in mouse bone marrow. *Cell* **91**, 661-672 (1997).
3. Akashi, K., Traver, D., Miyamoto, T. & Weissman, I.L. A clonogenic common myeloid progenitor that gives rise to all myeloid lineages. *Nature* **404**, 193-197 (2000).
4. Adolfsson, J. et al. Identification of Flt3+ lympho-myeloid stem cells lacking erythro-megakaryocytic potential a revised road map for adult blood lineage commitment. *Cell* **121**, 295-306 (2005).
5. Kondo, M. et al. Biology of hematopoietic stem cells and progenitors: implications for clinical application. *Annual Review of Immunology* **21**, 759-806 (2003).
6. Rosenbauer, F. & Tenen, D.G. Transcription factors in myeloid development: balancing differentiation with transformation. *Nature reviews.Immunology* **7**, 105-117 (2007).
7. Iwasaki, H. et al. GATA-1 converts lymphoid and myelomonocytic progenitors into the megakaryocyte/erythrocyte lineages. *Immunity* **19**, 451-462 (2003).
8. Laslo, P. et al. Multilineage transcriptional priming and determination of alternate hematopoietic cell fates. *Cell; Cell* **126**, 755-766 (2006).
9. Chen, H. et al. PU.1 (Spi-1) autoregulates its expression in myeloid cells. *Oncogene* **11**, 1549-1560 (1995).
10. Tsai, S.F., Strauss, E. & Orkin, S.H. Functional analysis and in vivo footprinting implicate the erythroid transcription factor GATA-1 as a positive regulator of its own promoter. *Genes & development* **5**, 919-931 (1991).
11. Cantor, A.B. & Orkin, S.H. Hematopoietic development: a balancing act. *Current opinion in genetics & development; Current opinion in genetics & development* **11**, 513-519 (2001).
12. Grass, J.A. et al. GATA-1-dependent transcriptional repression of GATA-2 via disruption of positive autoregulation and domain-wide chromatin remodeling. *Proceedings of the National Academy of Sciences of the United States of America* **100**, 8811-8816 (2003).
13. Liew, C.W. et al. Molecular analysis of the interaction between the hematopoietic master transcription factors GATA-1 and PU.1. *The Journal of biological chemistry* **281**, 28296-28306 (2006).
14. Robb, L. Cytokine receptors and hematopoietic differentiation. *Oncogene* **26**, 6715-6723 (2007).
15. Metcalf, D. Hematopoietic cytokines. *Blood* **111**, 485-491 (2008).
16. Murphy, K. Fate vs choice: the immune system reloaded. *Immunol Res* **32**, 193-200 (2005).
17. Abkowitz, J., Catlin, S. & Guttorp, P. Evidence that hematopoiesis may be a stochastic process in vivo. *Nat Med* **2**, 190-197 (1996).
18. Enver, T., Heyworth, C.M. & Dexter, T.M. Do stem cells play dice? *Blood; Blood* **92**, 348-351; discussion 352 (1998).

19. Losick, R. & Desplan, C. Stochasticity and cell fate. *Science* **320**, 65-68 (2008).
20. Rieger, M., Hoppe, P., Smejkal, B., Eitelhuber, A. & Schroeder, T. Hematopoietic cytokines can instruct lineage choice. *Science* **325**, 217-218 (2009).
21. Metcalf, D. Lineage commitment and maturation in hematopoietic cells: the case for extrinsic regulation. *Blood; Blood* **92**, 345-347; discussion 352 (1998).
22. Chiba, T., Ikawa, Y. & Todokoro, K. GATA-1 transactivates erythropoietin receptor gene, and erythropoietin receptor-mediated signals enhance GATA-1 gene expression. *Nucleic acids research* **19**, 3843-3848 (1991).
23. Zhang, D.E., Hetherington, C.J., Chen, H.M. & Tenen, D.G. The macrophage transcription factor PU.1 directs tissue-specific expression of the macrophage colony-stimulating factor receptor. *Molecular and cellular biology* **14**, 373-381 (1994).
24. Hohaus, S. et al. PU.1 (Spi-1) and C/EBP alpha regulate expression of the granulocyte-macrophage colony-stimulating factor receptor alpha gene. *Molecular and cellular biology* **15**, 5830-5845 (1995).
25. Smith, L.T., Hohaus, S., Gonzalez, D.A., Dziennis, S.E. & Tenen, D.G. PU.1 (Spi-1) and C/EBP alpha regulate the granulocyte colony-stimulating factor receptor promoter in myeloid cells. *Blood* **88**, 1234-1247 (1996).
26. Zhao, W., Kitidis, C., Fleming, M.D., Lodish, H.F. & Ghaffari, S. Erythropoietin stimulates phosphorylation and activation of GATA-1 via the PI3-kinase/AKT signaling pathway. *Blood* **107**, 907-915 (2006).
27. Dahl, R. et al. Regulation of macrophage and neutrophil cell fates by the PU.1:C/EBPalpha ratio and granulocyte colony-stimulating factor. *Nature immunology* **4**, 1029-1036 (2003).
28. Callard, R.E. & Yates, A.J. Immunology and mathematics: crossing the divide. *Immunology* **115**, 21-33 (2005).
29. Ferrell, J.E. & Xiong, W. Bistability in cell signaling: How to make continuous processes discontinuous, and reversible processes irreversible. *Chaos* **11**, 227-236 (2001).
30. Ferrell, J.E., Jr. Self-perpetuating states in signal transduction: positive feedback, double-negative feedback and bistability. *Current opinion in cell biology* **14**, 140-148 (2002).
31. Samoilov, M., Plyasunov, S. & Arkin, A.P. Stochastic amplification and signaling in enzymatic futile cycles through noise-induced bistability with oscillations. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 2310-2315 (2005).
32. Markevich, N.I., Hoek, J.B. & Kholodenko, B.N. Signaling switches and bistability arising from multisite phosphorylation in protein kinase cascades. *The Journal of cell biology* **164**, 353-359 (2004).
33. Roeder, I. & Glauche, I. Towards an understanding of lineage specification in hematopoietic stem cells: a mathematical model for the interaction of transcription factors GATA-1 and PU.1. *Journal of theoretical biology* **241**, 852-865 (2006).
34. Huang, S., Guo, Y.P., May, G. & Enver, T. Bifurcation dynamics in lineage-commitment in bipotent progenitor cells. *Developmental biology* **305**, 695-713 (2007).

35. Palani, S. & Sarkar, C.A. Positive receptor feedback during lineage commitment can generate ultrasensitivity to ligand and confer robustness to a bistable switch. *Biophysical journal* **95**, 1575-1589 (2008).
36. Chang, H., Hemberg, M., Barahona, M., Ingber, D. & Huang, S. Transcriptome-wide noise controls lineage choice in mammalian progenitor cells. *Nature* **453**, 544-547 (2008).
37. Gillespie, d. Exact stochastic simulation of coupled chemical-reactions. *Journal of physical chemistry* **81**, 2340-2361 (1977).
38. Gonze, D., Halloy, J. & Goldbeter, A. Deterministic versus stochastic models for circadian rhythms. *Journal of biological physics* **28**, 637-653 (2002).
39. Song, H., Smolen, P., Av-Ron, E., Baxter, D.A. & Byrne, J.H. Dynamics of a minimal model of interlocked positive and negative feedback loops of transcriptional regulation by cAMP-response element binding proteins. *Biophysical journal* **92**, 3407-3424 (2007).
40. Bruno, L. et al. Molecular signatures of self-renewal, differentiation, and lineage choice in multipotential hemopoietic progenitor cells in vitro. *Mol Cell Biol* **24**, 741-756 (2004).