UNIVERSITY OF CALIFORNIA, SAN DIEGO

Mathematical modeling of signaling and synthetic networks in single cells

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

in

Bioengineering

by

Jangir Selimkhanov

Committee in charge:

Jeff Hasty, Chair
Alex Hoffmann
Scott Rifkin
Gabriel Silva
Lev Tsimring
Kun Zhang

2014

The dissertation of Jangir Selimkhanov is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

_____

_____

_____

_____

_____
Chair

University of California, San Diego

2014

DEDICATION

To my family and Angie for their unwaivering support.

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

AU ..... arbitrary units

CFP ..... cyan fluorescent protein

DNA ..... deoxyribonucleic acid

E. coli ..... *Escherichia coli*

Fig. ..... Figure

GFP ..... green fluorescent protein

IPTG ..... isopropyl β-D-1-thiogalactopyranoside

mRNA ..... messenger ribonucleic acid

ODE ..... ordinary differential equation

YFP ..... yellow fluorescent protein

LIST OF FIGURES

ACKNOWLEDGEMENTS

I would like to thank all of my applied mathematics professors for helping me discover the power of math and its countless applications. I would especially like to acknowledge my undergraduate advisor, Mary Silber, without whose guidance I would probably never have considered coming to USCD to pursue a graduate degree. I would like to thank Jeff Hasty for letting me join his lab where mathematical modeling is held in high regard and where I had the opportunity and support to work on variety of projects and further expand on my mathematical skills in the process. I would not have been able to succeed in my research without the help of Lev Tsimring, who has been most influential and helpful with every project that I have worked on at UCSD. Together, Jeff and Lev, have taught me a great deal about research and how best to approach and solve difficult problems. Finally, I would like to thank everyone in the Biodynamics lab, BioCircuits Institute, and SDCSB for their help and support throughout the past five years.

Chapter 2 contains material originally published as Prindle, A.*, Selimkhanov, J.*, Li, H., Razinkov, I., Tsimring, L. S., and Hasty, J. 2014: Rapid and tunable post-translational coupling of genetic circuits. *Nature*. (*equal contribution). Copyright permission to republish here was granted by Nature Publishing Group.

Chapter 3 contains material originally published as Prindle, A., Selimkhanov, J., Danino, T., Tsimring L., and Hasty, J. 2012: Genetic circuits in *Salmonella typhimurium*. *ACS Synthetic Biology*, 10: 458-464. Copyright permission to republish here was granted by ACS.

Chapter 4 contains material originally published as Longo, D. M.*, Selimkhanov, J.*, Kearns, J. D., Hasty, J., Hoffmann, A., and Tsimring, L. S. 2013: Dual delayed feedback provides sensitivity and robustness to the NF-κB signaling module. *PLoS Computational Biology*, 9(6), e1003112. (*equal contribution). Permission to republish here is not required by PLOS.

Chapter 6 contains material being prepared for publication as Selimkhanov, J.*,

Taylor, B.*, Yao J., Pilko A., Albeck J., Hoffmann A., Tsimring L.S., Wollman R.. Accurate information transmission in dynamic biochemical signaling networks.

VITA

2008	Bachelor of Science in Biomedical Engineering and Engineering Sciences and Applied Mathematics Northwestern University
2012	Master of Science in Bioengineering University of California, San Diego
2014	Doctor of Philosophy in Bioengineering University of California, San Diego

FIELDS OF STUDY

Major Field: Bioengineering (Mathematical Modeling)

Studies in Biological Dynamics
Professor Jeff Hasty and Dr. Lev S. Tsimring

PUBLICATIONS

Peer Reviewed Journal Articles

Mondragon-Palomino O., Danino T., Selimkhanov J., Tsimring L., and Hasty J.. Entrainment of a population of synthetic genetic oscillators. *Science*, 333(6047): 1315-9. (2011)

Selimkhanov J., Hasty J., and Tsimring L.S.. Recent advances in single-cell studies of gene regulation. *Current Opinion Biotechnology*, 23(1): 34-40 (2012)

Prindle A., Selimkhanov J., Danino T., Samayoa P., Goldberg A., Bhatia S.N., and Hasty J.. Genetic circuits in Salmonella typhimurium. *ACS Synthetic Biology*, 1(10): 458-464 (2012)

Longo, D. M.*, Selimkhanov, J.*, Kearns, J. D., Hasty, J., Hoffmann, A., and Tsimring, L. S.. Dual delayed feedback provides sensitivity and robustness to the NF-κB signaling module. *PLoS Computational Biology*, 9(6), e1003112. (2013)

Prindle A.*, Selimkhanov J.*, Li H., Razinkov I., Tsimring L. S., Hasty J.. Rapid and tunable post-translational coupling of genetic circuits. *Nature*, (2014)

ABSTRACT OF THE DISSERTATION

Mathematical modeling of signaling and synthetic networks in single cells

by

Jangir Selimkhanov

Doctor of Philosophy in Bioengineering

University of California, San Diego, 2014

Professor Jeff Hasty, Chair

Recent advances in quantification methods of regulatory and signaling gene networks has lead to an increasing amount of data that has opened the door for improved understanding of cell behavior. The key to that understanding is through the use of mathematical models that can explain existing data as well as help generate new hypotheses through prediction. Refinement of these models with new experimental data creates a feedback loop, where modeling drives experiments while newly generated data constrain the model. Mathematical principles underlying various models can then give us insight into basic biological principles that describe network dynamics. In this thesis, several different applications of mathematical modeling are used to help further our understanding of signaling and synthetic gene networks. First, mathematical modeling is used to explain the underlying mechanisms in coupling of two synthetic gene oscillators to each other as well as to the host environment, which leads to the observed non-trivial biological behavior. Second, focusing on a specific signaling protein network, charac-

terized by transcription factor nuclear factor kappa B (NF-κB), mathematical modeling is used to understand how the underlying the cell-to-cell variability leads to variability in the response of the system to gradually increasing levels of the network-activating tumor necrosis factor alpha (TNFα). Finally, information-theoretic approach is applied to three different signaling networks to help gain insight into the role that various sources of noise and various forms of network responses play in signal transduction.

Chapter 1

Introduction

In studies of gene networks, mathematical modeling has been an instrumental tool in helping us understand various biological systems and design synthetic ones. Use of ordinary differential equations, for example, has been key to describing cell population dynamics ([77;112;70]), while probability theory has helped us understand the nature of noise in single cell behavior[38;114]. The key to effective modeling of biological systems is to define the role that the model will have in helping us understand something about the system of interest. In many cases, it turns out that simpler models, with fewer number of parameters, may lead to more insightful conclusions that are less dependent on understanding of every single detail of the network. These models provide us with general principles underlying the specific biological system in question.

One important aspect of biological systems that mathematical models help us with is the dynamics of gene regulation inside the cell. The cell is never in a steady state as it constantly responds to internal and external stimuli. Mathematical modeling gives us the theory, which we can use to capture dynamic behavior and understand the effect gene network architecture has on signal propagation and processing. Models can also help us to design synthetic gene networks with specific functions, acting as blue prints that guide engineering of synthetic circuits. They can also help troubleshoot those same circuits, when they do not work in a specific system, by providing an easy test system

that does not require a huge time investment. In the the next few chapters, I hope to illustrate how many of these benefits of mathematical modeling have come though in helping us understand variety of biological systems.

Chapter 2

Rapid and tunable post-translational coupling of genetic circuits

2.1 Introduction

One promise of synthetic biology is the creation of genetic circuitry that enables the execution of logical programming in living cells. Such "wet programming" is positioned to transform a wide and diverse swath of biotechnology ranging from therapeutics and diagnostics to water treatment strategies. While progress in the development of a library of genetic modules continues apace[96;125;136;152], a major challenge for their integration into larger circuits is the generation of sufficiently fast and precise communication between modules[80;31]. An attractive approach is to integrate engineered circuits with host processes that facilitate robust cellular signaling[100]. In this context, recent studies have demonstrated that bacterial protein degradation can trigger a precise response to stress by overloading a limited supply of intracellular proteases[43;92;21]. Here, we use protease competition to engineer rapid and tunable coupling of genetic circuits across multiple spatial and temporal scales. We characterize coupling delay times that are more than an order of magnitude faster than standard transcription-factor based coupling methods (less than one minute compared with $\sim$20-40 minutes) and demonstrate

tunability through manipulation of the linker between the protein and its degradation tag. We use this mechanism as a platform to couple genetic clocks at the intracellular and colony level, then synchronize the multi-colony dynamics to reduce variability in both clocks. We show how the coupled clock network can be used to encode independent environmental inputs into a single time series output, thus enabling the possibility of frequency multiplexing in a genetic circuit context. Our results establish a general framework for the rapid and tunable coupling of genetic circuits through the use of native queueing processes such as protein degradation.

2.2 Fast degradation coupling

In order to engineer rapid coupling between synthetic genetic modules, we developed a post-translational coupling platform that operates via shared degradation by the ClpXP protease (Fig. 2.1a). In this scheme, all LAA-tagged components[68] are dynamically linked via competition for a limited number of proteases[21;49], such that tagged modules remain tightly aligned (1$\pm$1 min, GFP-CFP curve pairs in Fig. a) despite significant induction delay (31$\pm$5 min, inducer-GFP offset in Fig. a). This coupling method produces delays that are more than an order of magnitude faster than standard transcription-factor based coupling methods ($\sim$20-40 min)[118;62]. To illustrate directly the response time that can be achieved by coordinating module output via modulating ClpXP activity, we show that low levels (90 μM) of externally provided H_2O_2 "inducer" rapidly (< 2 min, our experimental timestep) and reversibly modulates the concentration of constitutively expressed GFP in a ClpXP-dependent manner (Fig. b). Here, H_2O_2 reduces the native substrate load on ClpXP by obstructing RssB, the adapter protein that targets the alternative sigma factor σ^S for degradation by ClpXP[94;43;92]. Since σ^S is continuously produced and degraded by ClpXP, inactivating its rate-limiting adapter protein results in an instantaneous increase in the effective ClpXP degradation rate for LAA-tagged proteins[109].

We systematically explored the coupling mechanism by driving a constitutive

module with a quorum-sensing (c). As the pacemaker, the quorum clock generates density-dependent synchronous oscillations at the colony level via acyl-homoserine lactone (AHL), a small molecule capable of synchronizing cellular behavior across distances up to 100 μm[27]. Using microfluidic devices[40] we observed the colony-level expression of the constitutive module, finding oscillating expression synchronized to the quorum clock (Fig. 2.1c, top right). We then constructed a library of degradation tags by adding a series of variable-length spacer regions between the downstream protein and its degradation tag. Spacer regions contained between one and five copies of the amino acid sequence "TS" and their effects on offset time compared to that of a previously published alternate degradation tag (Fig. 2.5b-f). While all spacer sequences produced synchronous activation dynamics, the degradation dynamics of the downstream module were offset depending on the length of the linker sequence, where longer linkers produced greater GFP-CFP offset time (Fig. 2.1C, bottom). Thus, our ClpXP coupling platform rapidly links genetic modules via shared degradation, where the strength and timing of coupling can be tuned by changing the degradation kinetics of individual modules.

2.3 Degradation coupled genetic oscillators

In order to engineer coupling between genetic modules capable of generating their own dynamics, we designed a circuit containing the quorum clock and a variant of a previously described intracellular clock (Fig. 2.2a)[127]. This $p_{lac/ara-1}$ intracellular clock variant retains the fast dynamics and simple genetic architecture of the published $P_{LlacO-1}$ negative feedback oscillator, yet its period is tunable by both isopropyl β-D-1-thiogalactopyranoside (IPTG) and arabinose in the presence of chromosomal araC. We first used small microfluidic devices (100 cells) and observed fast and asynchronous intracellular clock oscillations without quorum clock contribution, since the quorum clock requires a critical colony size to function. In larger devices (5,000 cells), we observed a transition from asynchronous oscillations to identical intracellular/quorum clock oscil-

lations as the population grew larger (Fig. 2.2b). In the case of the larger population, the substrate load on ClpXP during the quorum clock pulse is sufficient to shift the intracellular clock out of its oscillatory regime, enabling complete linkage between the two clocks despite their vastly different spatial and temporal scales. Thus, despite lacking a mode of cell-cell communication itself, the intracellular clock is effectively synchronized at the colony level via ClpXP-mediated coupling with the quorum clock.

We found that changing the intracellular clock period of individual cells indirectly tuned the quorum clock period, where IPTG values associated with longer intracellular clock periods inversely produced shorter quorum clock periods (Fig. 2.2c). We developed a computational model of the oscillator network involving a form of load-mediated pulse frequency modulation to explain this effect (Fig. 2.2d-f). Between coupled pulses, the intracellular clock accelerates the quorum pulse onset via load-mediated decreases in the degradation rate of LuxI, where larger intracellular clock load produces higher levels of the AHL-synthase (Fig. 2.2e, left and Fig. 2.6a-e). During the coupled pulse, contributions of the intracellular clock leave the duration of the pulse itself unchanged (Fig. 2.2e, left: model and right: experimental). Linking the intracellular and quorum clocks via degradation also yielded an expansion in the oscillatory regime for the coupled system with respect to flow rate compared to the quorum clock alone (Fig. 2.2f). In this way, the intracellular clock continually excites the quorum clock to fire, enabling more robust function at higher external flow rates (Fig. 2.7a-c).

2.4 Multispectral genetic frequency encoding

With a platform for rapidly coupling genetic clocks at multiple scales, we sought to engineer a system capable of frequency encoding information from both clocks into the multispectral time series of a single reporter (Fig. 2.3a). Here, the measured output of the intracellular clock reporter contains contributions from its own fast intracellular clock dynamics between slow quorum clock bursts. Since the range of natural periods for the faster $p_{lac/ara\text{-}1}$ intracellular clock is fully separated from the slower quorum

clock[127;27;107], both IPTG/arabinose and flow rate inputs can be encoded into frequency-modulated oscillations in the time domain where they can be independently extracted by Fourier transform. Thus, the measurement of a single clock history reveals the activities both underlying clock networks.

We began by characterizing the frequency response curves for both the intracellular and quorum clocks in isolation, finding ranges of 7-25 min and 55-95 min, respectively, when sweeping IPTG/arabinose and flow rate inputs (Fig. 2.3b, top: intracellular clock in araC+ strain and bottom: quorum clock, original study data[27]). We then measured trajectories taken from the coupled clock system and extracted the frequency components of both clocks by Fourier transform (Fig. 2.3c). In sweeping IPTG/arabinose inducers, we found the frequency response of the intracellular clock contribution to the multispectral reporter to be unchanged by the inclusion of the quorum clock, where the intracellular frequency response to IPTG/arabinose was equivalent to the isolated clock (Fig. 2.3d, top: coupled and Fig. 2.3b, top: isolated). We then swept flow rates at 3 fixed inducer levels, finding distinct response curves for the quorum clock contribution to the multispectral reporter shifted in accordance with our model for ClpXP-mediated frequency modulation by the intracellular clock (Fig. 2.3d, bottom). Thus, to decode a given pair of IPTG/arabinose and flow rate inputs, we first recover the intracellular clock frequency as a measure of IPTG/arabinose and then use the corresponding quorum clock response curve to measure flow rate.

2.5 Array level oscillator coupling

To extend rapid coupling to greater spatial scales, we added a genetic H_2O_2 signaling[106] cassette to the network and observed synchronization at the multi-colony level (Fig. 2.4a). In conducting these experiments, we also observed H_2O_2 -mediated interaction between the native stress response network and our synthetic circuit at ClpXP (Fig. 2.4b). In the original design, H_2O_2 synchronized quorum clock oscillations by transcriptional upregulation of the lux promoter via the aerobic response control sys-

tem ArcAB[106]. In addition to transcriptional increase (Fig. 2.4c, top), we found an increase in the apparent degradation rate with H_2O_2 (Fig. 2.4c, bottom and Fig. 2.8a-b), consistent with increased ClpXP activity in response to externally provided H_2O_2 . The coupled increases in transcriptional output and effective ClpXP degradation rate in response to H_2O_2 also tightens the period distribution at the multi-colony level by mitigating the effects of period variation in an individual colony (Fig. 2.4c, top and Fig. 2.8c-d).

Engineering synthetic circuits composed of interacting modules is an ongoing effort[96;125;136;152] that has generally relied on transcription and translation, with less attention paid to post-translational coupling mechanisms[51]. Protease competition offers the advantages of rapid response, modularity with distinct recognition sequences, and simultaneous control over multiple circuits with protease adapters[89;50]. More generally, in natural biological networks, competition for cellular resources (e.g., metabolites, enzymes, transcription factors, binding sites) produces nonlinear coupling effects that serve to reduce noise, increase sensitivity to input concentrations, and discriminate between multiple inputs[49;18;97;17;129]. We envision that coordinating engineered circuits via built-in cellular processes—what we term "host-linked" coupling—has the potential to produce more sophisticated circuits by facilitating robust signaling between synthetic modules.

2.6 Methods and additional experimental results

2.6.1 Data analysis

Single cell and individual trap fluorescent trajectories were obtained from time-lapse images using our previously developed algorithms[?] [106] and builtin MATLAB® functions. We identified peaks and troughs from these trajectories and used these values to calculate periods and amplitudes. To calculate the coupling delay in Figure 1A and offset time in 1C we measured the difference between the 10% amplitude points

of trajectory pairs. The induction time was measured from induction start time to 10% amplitude of the induced module. To extract both frequencies from time series data, we performed Fourier transforms using the Lomb-Scargle algorithm. We used two sequential transforms to isolate each component separately. First, we used a band-pass filter (5 - 25 min) to extract the fast intracellular clock component. Then, we filtered out these fast frequencies using a second band-pass filter (75 - 150 min) to extract the slower quorum clock component. Finally, we overlay the 2 power spectra, preserving the relative amplitude of the peaks.

2.6.2 Degradation tag experiments

In addition to exploring the effect of variable-length linker (TS repeats) on the phase-shift in module degradation (Fig. 2.5c-f), we tested a well characterized AAV degradation tag[3]. In Andersen *et al*, GFP-AAV was shown to have 50% higher half-life than GFP-LAA. In this study, downstream module (CFP-AAV) showed a delay in degradation relative to the driver module (GFP-LAA) that was similar to that of the 2 TS-linker sequence (Fig. 2.5b bottom). Further characterization is required to determine the differences in the mechanism of action between variable-length TS linker sequence before the SspB binding region and the AAV degradation tag. While CFP to GFP bleed-over is more significant than GFP to CFP bleed-over, the CFP to GFP bleed-over is not relevant to our experiment in Figure 1a, where the induced protein (GFP) drives the protein level of the coupled protein (CFP). Thus, we performed an experiment to test the potential for bleeding from sfGFP into CFP fluorescence channel by activation sfGFP with 10nM AHL in a strain that lacked CFP fluorophore. We saw no change in CFP fluorescence while sfFGP increased as expected (Fig. 2.5b top).

2.6.3 NFB helps H_2O_2 synchronize oscillations between colonies

We defined the inter-pulse (wait) time as the time between the 10% downlsope point of one peak and 10% upsope point of the following peak (Fig. 2.5a).The mean

QS inter-pulse time decreased with addition of IPTG (0.5mM) to the coupled system, while the time of each pulse stayed constant. In addition, we find that QS trajectories from the coupled oscillator system showed significantly lower variability without IPTG as compared to 0.5mM IPTG (Fig. 2.7a-b). These results suggests that stronger NFB (0mM IPTG) associated with higher NFB protein production[127] leads to shorter and more robust inter-pulse behavior in the coupled system. In large biopixel devices, less robust colony-level oscillations prevent H_2O_2 from effectively coupling neighboring pixels, resulting in unsynchronized QS oscillations (No NFB in Fig. 2.7c). NFB reduces inter-pulse duration noise, which allows H_2O_2 to synchronize QS oscillations in neighboring colonies in biopixel devices (0.1mM IPTG in Fig. 2.7c). Increasing NFB strength, further

2.6.4 H_2O_2 increases protein degradation rate

Our analysis of H_2O_2 synchronized quorum clock trajectories showed decrease in the period and increase in the amplitude of oscillations (Fig. 2.4b Top). H_2O_2 synchronization leads to clear reduction of the degradation time in these trajectories (Fig. 2.8a). One of the significant contributors to the decrease in the period is the increase in the activity of ClpXP targeted proteins, which we quantified as the rate of CFP fluorescence decrease from the peak time to the 10% downslope time. Figure 2.8b shows a significant increase in the ClpXP degradation rate (3X) due to H_2O_2 coupling.

2.7 Model formulation

2.7.1 QS oscillator

To describe dynamic behavior of uncoupled QS oscillator, we expanded on the delay-differential equation model presented in[27]. In addition to the equations for LuxI (I), AiiA (A), internal AHL (H_i), external AHL (H_e), we included AHL substrate (S), consisting of acyl-ACPs and S-adenosylmethionine (SAM)[?] , to account for the slow-

ing down of H_i production while the number of LuxI molecules is still on the rise. Transcription, translation, and maturation rate of proteins are combined into a single time-delay parameter τ_H. Transcriptional activation by the LuxR and AHL complex (2 of each LuxR and AHL molecules) give delayed production term $P(\tau_H)$, which depends on the past concentration of internal AHL, $H_i(t-\tau_H)$. We assumed a constant level of LuxR since it is not tagged for fast degradation and has a large amount of genetic copies on the plasmid (it is on colE1 twice and p15A once). We used hill coefficient of 4 in accordance with[?] to account for high AHL cooperativity possibly due to AHL-LuxR polymerazation. Diffusion of AHL through cell membrane is described by terms proportional to D, while dilution of external AHL is described by the term proportional to μ. Cell density parameter d was incorporated into the system to account for the difference in the total cell volume and media volume. Enzymatic degradation terms proportional to γ_I and γ_A describe enzymatic degradation of LuxI and AiiA respectively through Michaelis-Mentent kinetics. Different values of k_I and k_A represent different preferential binding dynamics of LuxI and AiiA to ClpXP.

$$\frac{\partial A}{\partial t} = C_A P(\tau_H) - \frac{\gamma_A(A/k_A)}{1+A/k_A+I/k_I} \tag{2.1}$$

$$\frac{\partial I}{\partial t} = C_I P(\tau_H) - \frac{\gamma_I(I/k_I)}{1+A/k_A+I/k_I} \tag{2.2}$$

$$\frac{\partial H_i}{\partial t} = \frac{bI(S/k_S)}{1+S/k_S} - \frac{\gamma_H A(H_i/k_H)}{1+H_i/k_H} + D(H_e-H_i) \tag{2.3}$$

$$\frac{\partial H_e}{\partial t} = -\frac{d}{1-d}D(H_e-H_i) - \mu H_e \tag{2.4}$$

$$\frac{\partial S}{\partial t} = S_0 - S - \frac{bI(S/k_S)}{1+(S/k_S)} \tag{2.5}$$

$$P(\tau_H) = \alpha_0 + \frac{\alpha_H(H(t-\tau_H)/h0)^4}{1+(H(t-\tau_H)/h0)^4}$$

Experimentally relevant scaled parameters used with this model are described in "Model parameter values".

2.7.2 NFB oscillator

To describe dynamic behavior of NFB oscillator, we used a single delay-differential equation for LacI (L) based on[86]. Transcription, translation, and maturation of proteins are lumped together into time-delay parameter τ_L. Transcriptional inactivation of LacI gives the delayed production term $Q(\tau_L)$, which depends on the past concentration of LacI, $L(t-\tau_L)$. Enzymatic degradation of LacI is described by the term proportional to γ_L through Michaelis-Mentent kinetics. Parameter C in production expression Q represents the effect of IPTG on the strength of LacI repression.

$$\begin{aligned} \frac{\partial L}{\partial t} &= Q(\tau_L) - \frac{\gamma_L(L/k_L)}{1+L/k_L} \\ Q(\tau_L) &= \tfrac{\alpha_L}{1+(L(t-\tau_L)/C)^2} \end{aligned} \tag{2.6}$$

The dynamics of the above model accounted for most of the experimental results. To resolve the amplitude increase in the NFB oscillator when coupled to the QS oscillator during the QS pulse we had to include reporter dynamics with equations for YFP precursor (Y_p) and mature YFP (Y_m). These additional equations are not required to explain the QS dynamics in the coupled system. Experimentally relevant scaled parameters used with this model are shown in "Model parameter values".

$$\frac{\partial L}{\partial t} = Q(\tau_L) - \frac{\gamma_L(L/k_L)}{1+L/k_L+Y_p/k_L+Y_p/k_L} \tag{2.7}$$

$$\frac{\partial Y_p}{\partial t} = Q(\tau_L) - \frac{\gamma_L(Y_p/k_L)}{1+L/k_L+Y_p/k_L+Y_m/k_L} - Y_p \tag{2.8}$$

$$\frac{\partial Y_m}{\partial t} = Y_p - \frac{\gamma_L(Y_m/k_L)}{1+L/k_L+Y_p/k_L+Y_m/k_L} \tag{2.9}$$

$$Q(\tau_L) = \tfrac{\alpha_L}{1+(L(t-\tau_L)/C)^2}$$

2.7.3 Coupled NFB and QS oscillators

Coupling of the two oscillators was accomplished by increasing the effective "queueing" effect through CplXP degradation[21]. In the uncoupled case, the degradation

of the two oscillator components would be independent, $\frac{ClpXP}{1+QS} + \frac{ClpXP}{1+NFB}$, while in the coupled scenario, $\frac{ClpXP}{1+QS+NFB}$, the degraded components end up in the same degradation term. To couple NFB and QS oscillators through ClpXP degradation, we added LuxI and AiiA from QS system to the degradation expressions in NFB system and LacI (L) from NFB system to the degradation expression in QS system.

$$\frac{\partial A}{\partial t} = C_A P(\tau_H) - \frac{\gamma_A (A/k_A)}{1 + A/k_A + I/k_I + L} \tag{2.10}$$

$$\frac{\partial I}{\partial t} = C_I P(\tau_H) - \frac{\gamma_I (I/k_I)}{1 + A/k_A + I/k_I + L} \tag{2.11}$$

$$\frac{\partial H_i}{\partial t} = \frac{bI(S/k_S)}{1 + S/k_S} - \frac{\gamma_H A(H_i/k_H)}{1 + H_i/k_H} + D(H_e - H_i) \tag{2.12}$$

$$\frac{\partial H_e}{\partial t} = -\frac{d}{1-d} D(H_e - H_i) - \mu H_e \tag{2.13}$$

$$\frac{\partial S}{\partial t} = S_0 - S - \frac{bI(S/k_S)}{1 + (S/k_S)} \tag{2.14}$$

$$\frac{\partial L}{\partial t} = Q(\tau_L) - \frac{\gamma_L (L/k_L)}{1 + L/k_L + A + I} \tag{2.15}$$

$$P(\tau_H) = \alpha_0 + \frac{\alpha_H (H(t-\tau_H)/h0)^4}{1+(H(t-\tau_H)/h0)^4}$$

$$Q(\tau_L) = \frac{\alpha_L}{1+(L(t-\tau_L)/C)^2}$$

Experimentally relevant scaled parameters used with this model are described in "Model Parameter Values". We varied the flow μ, IPTG concentration C, and arabinose concentration α_L to recapture many of the experimental findings.

Leader cell wait time shortening

To understand the multicellular dynamics of QS pulse activation we constructed a model with two identical cells that share external AHL (H_e). We first considered a QS only system consisting of two cells with slightly different constitutive production of AiiA and LuxI. In this system, the slower cell couples to the faster one, suggesting that cells whose QS pulse fires first cause QS pulse activation in the nearby cells through AHL cell-to-cell communication (Fig. 2.6a). Next we added NFB to cell 1 in a two-cell

system, resulting in period shortening of that cell. As the result, when the two cells were linked through external AHL, the slower cell 2 (without NFB), coupled to the faster cell 1 (Fig. 2.6b). Consequently, even though NFB might be out of phase in different cells, the onset of QS pulse in the faster cells can initiate the propagation of the QS pulse through the rest of the cells in the nearby region. This effect further reduces cell-cell QS variability, which we see from period variability reduction in a 20-cell model (Fig. 2.6d). We added noise to constitutive production of AiiA and LuxI proteins ($\alpha_0 = 0.6 \pm 0.1$) of each of the 20 cells and showed period variability reduction in synched vs unsynched cells (Fig. 2.6e).

2.7.4 QS and H_2O_2 coupled through queueing

To describe dynamic behavior of QS oscillator in response to H_2O_2 produced during LuxI fluorescent reporter expression, we added a differential equation describing production and degradation of H_2O_2 (V_i and V_e) to the QS oscillator delay-differential equation model. We assumed that the production of H_2O_2 is dependent on the concentration of LuxI, which is under the same promoter as the CFP fluorescent protein. Degradation of H_2O_2 by catalase is proportional to its concentration. H_2O_2 affects the QS oscillator in two characteristic ways. First, ArcA, which is under normal conditions partially represses Lux promoter, is inactivated under oxidizing conditions triggered by H_2O_2 , relieving Lux repression and increasing LuxI and AiiA production. We model this phenomenon by adding a multiplier to the production term that is dependent on H_2O_2 concentration. Second, H_2O_2 has been shown to reduce ClpXP load, leading to increased rate of AiiA and LuxI degradation. Again, we model this behavior by adding a multiplier in front of the degradation term, dependent on H_2O_2 concentration. Finally, H_2O_2 can freely diffuse across cell membrane, which we describe a diffusion term characterized by diffusion parameter D_V. Extracellular H_2O_2 (V_e) can further leave the system with the rate proportional to its concentration.

$$\frac{\partial A}{\partial t} = C_A P(\alpha_H, \tau) - (1 + V_i)\frac{\gamma_A (A/k_A)}{1 + A/k_A + I/k_I} \quad (2.16)$$

$$\frac{\partial I}{\partial t} = C_I P(\alpha_H, \tau) - (1 + V_i)\frac{\gamma_I (I/k_I)}{1 + A/k_A + I/k_I} \quad (2.17)$$

$$\frac{\partial H_i}{\partial t} = \frac{bI(S/k_S)}{1 + S/k_S} - \frac{\gamma_H A(H_i/k_H)}{1 + H_i/k_H} + D(H_e - H_i) \quad (2.18)$$

$$\frac{\partial H_e}{\partial t} = -\frac{d}{1-d} D(H_e - H_i) - \mu H_e \quad (2.19)$$

$$\frac{\partial S}{\partial t} = S_0 - S - \frac{bI(S/k_S)}{1 + (S/k_S)} \quad (2.20)$$

$$\frac{\partial V}{\partial t} = \frac{\delta(I/C_I)}{1 + I/C_I} - V_i + D_V(V_e - V_i) \quad (2.21)$$

$$\frac{\partial V_e}{\partial t} = \frac{d}{1-d} D_V(V_e - V_i) - \mu_V * V_e \quad (2.22)$$

$$P(\tau_H) = (1 + f_p V)(\alpha_0 + \tfrac{\alpha_H (H(t-\tau_H)/h0)^4}{1+(H(t-\tau_H)/h0)^4})$$

H_2O_2 increases QS period robustness

As we have mentioned before, reduction in inter-pulse duration leads to reduction in period variability arising from noise. Incorporating H_2O_2 effects on QS oscillator into our model (see above) results in several major changes in QS trajectory. First, as expected the amplitude of QS and the downslope time of QS decrease with addition of H_2O_2 (Fig. c). The result of these two effects also results in shortening of inter-pulse duration, which leads to more robust QS oscillations (Fig. 2.8d). We simulated the model to obtain at least 50 period measurement for period CV calculation. The noise was introduced into the model through addition of a noisy production term ($\alpha_v = \pm 0.1$) to the delayed production term $P(\tau_H) = \alpha_v + (1 + f_p V)(\alpha_0 + \frac{\alpha_H (H(t-\tau_H)/h0)^4}{1+(H(t-\tau_H)/h0)^4})$.

Interestingly, our model shows that individual effects of H_2O_2 activation of lux promoter and increase in ClpXP activity result in the increase the CV of the QS period (Fig. 2.8d). With respect to increased ClpXP activity, higher CV is mainly due to the resulting longer inter-pulse duration (Fig. 2.8c green). Increased lux promoter activity, however, leads to more variable degradation due to higher pulse amplitude variability.

The two countering H_2O_2 effects seem to cancel each other's variability generating more robust QS oscillations.

Fitting model parameters to experimental results

To fit the NFB period data from experiments we used the following parameter scaling functions for the LacI production term ($Q(\tau_L) = \frac{\alpha_L}{1+(L(t-\tau_L)/C)^2}$) to fit IPTG and arabinose (ARA) concentrations:

$$alpha_L \propto A_A + D_A \frac{(\frac{ARA}{C_A})^{H_A}}{(1+\frac{ARA}{C_A})^{H_A}}$$

$$A_A = 0.2758, D_A = 1.6291, C_A = 0.5638, H_A = 0.9029$$

$$C \propto A_C + D_C \frac{(\frac{IPTG}{C_C})^{H_C}}{(1+\frac{IPTG}{C_C})^{H_C}}$$

$$A_C = 0.0968, D_C = 60.8510, C_C = 8.2451, H_C = 0.4334$$

Similarly we fit the model flow term μ to the experimental flow values using the following function

$$\mu = A_\mu \mu^2 + B_\mu \mu + C$$

$$A_\mu = 1.2e-7, B_\mu = 0.0022, C_\mu = -0.11$$

Model parameter values

C_A = 1 (AiiA copy number); C_I = 4 (LuxI copy number); γ_A = 8 (ClpXP degradation of AiiA); γ_I = 8 (ClpXP degradation of LuxI); K_A = 1 (AiiA binding affinity to ClpXP); K_I = 0.2 (LuxI binding affinity to ClpXP); α_0 = 0.6 (Lux promoter basal production); α_H = 3 (Lux promoter AHL induced production); $h0$ = 0.1 (AHL promoter binding affinity); τ_H = 1 (delay in LuxI and AiiA production); b = 1 (AHL synthesis rate by LuxI); k_S = 25 (AHL substrate binding affinity to LuxI) ; S_0 = 50 (basal AHL substrate production); γ_H = 1 (AHL degradation rate by AiiA); k_H = 0.1 (AHL binding affinity to AiiA); D = 0.8 (AHL diffusion across the membrane); d = 0.1 (cell density); μ

= 0.5 (flow rate); α_L = 1 (LacI/YFP production rate); C = 0.0025 (LacI promoter binding affinity); τ_L = 0.7 (delay in LacI/YFP production); k_L = 0.001 (LacI/YFP binding affinity to ClpXP); γ_L = 0.05 (ClpXP degradation of LacI/YFP); δ = 1 (H_2O_2 production due to QS fluorophores); C_I = 2 (Michaelis constant); f_p = 1.3 (strength of H_2O_2 activation of LuxI promoter); D_V = 8 (H_2O_2 diffusion across membrane); μ_V = 0 (extracellular H_2O_2 dilution)

2.8 Acknowledgments

Chapter 2 contains material originally published as Prindle, A.*, Selimkhanov, J.*, Li, H., Razinkov, I., Tsimring, L. S., and Hasty, J. 2014: Rapid and tunable post-translational coupling of genetic circuits. *Nature*. (*equal contribution). Copyright permission to republish here was granted by Nature Publishing Group.

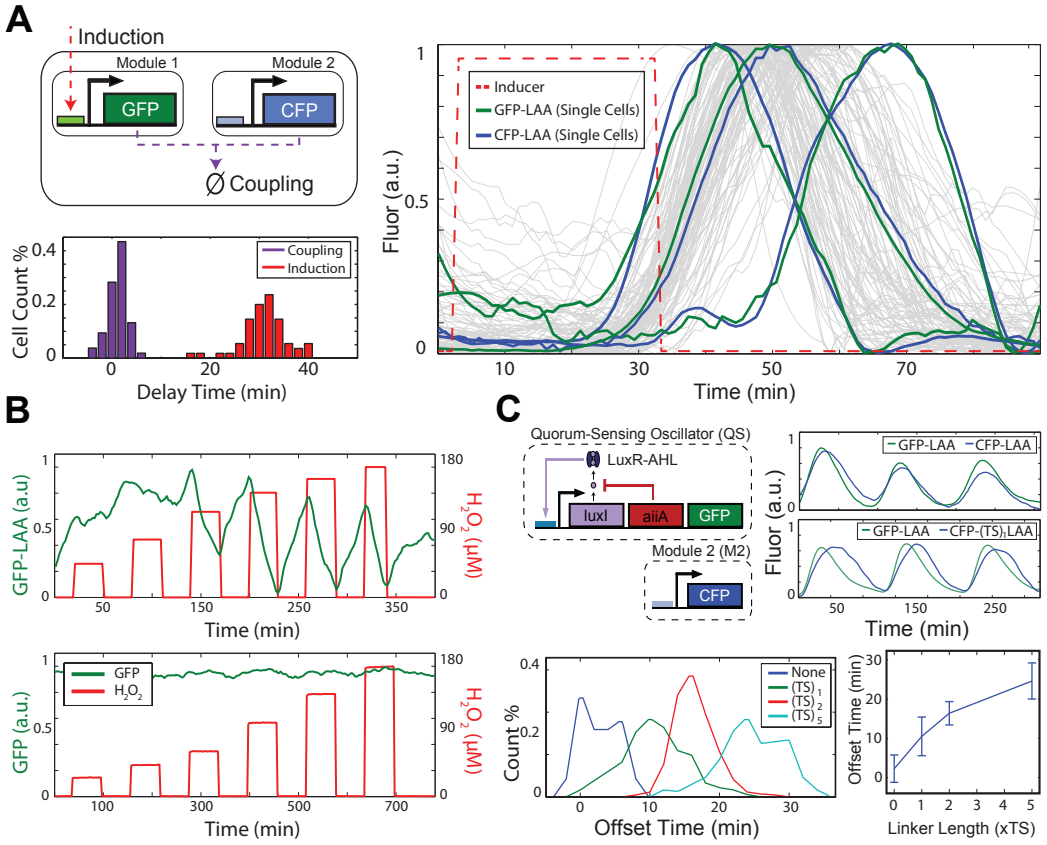

Figure 2.1: A rapid post-translational coupling platform based on shared degradation. (A) We measured the delays associated with module-module coordination by ClpXP (1 ± 1 min) and input-output response via transcription/translation (31 ± 5 min) in a single experiment by inducing the *lux* promoter and tracking the response of sfGFP-LAA (*lux* promoter) and CFP-LAA ($p_{lac/ara-1}$promoter) in single cells (55 cell trajectories). (B) Rapid (< 2 min, our experimental timestep) induction of protein degradation by externally provided H_2O_2 produces reversible changes in ClpXP load in response to obstruction of RssB[94;43;92]. (C) To use post-translational coupling to drive downstream modules, we linked a quorum clock to a constitutively expressed fluorescent protein via the addition of identical LAA tags. With identical degradation tags, the constitutive module couples tightly to the quorum pacemaker. The addition of a variable-length linker (TS repeats) before the degradation tag phase-shifts the degradation dynamics, where longer linkers produced longer delays. The error bars indicate s.d. of offset time, centered at the mean (50-200 cells for each TS-linker length).

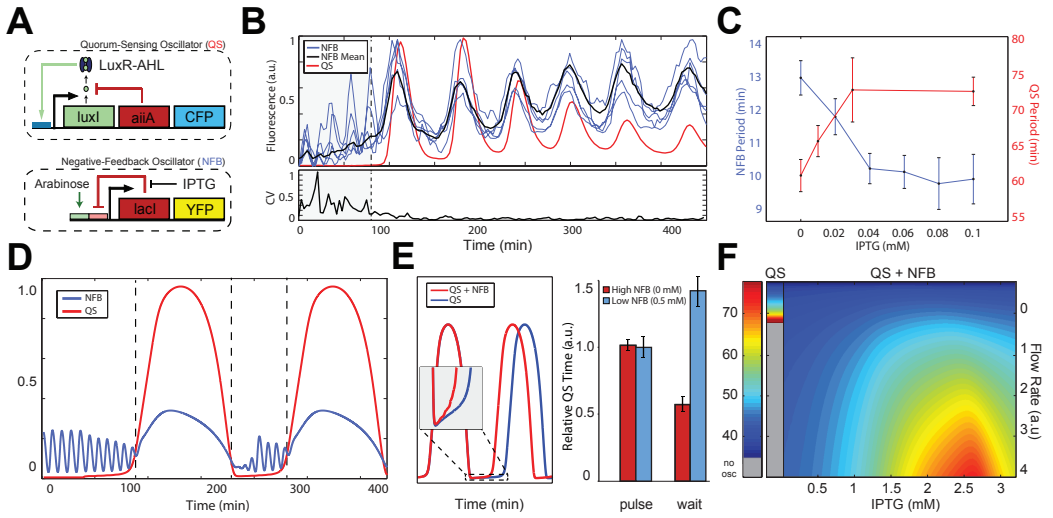

Figure 2.2: Post-translationally linked genetic clocks at multiple scales. (A) The network is composed of coupled intracellular[127] and quorum clocks[27]. The intracellular clock oscillates as a result of delayed negative feedback on its own promoter and its period is tunable by IPTG/arabinose. Quorum clock oscillations are tunable by media flow rate and are synchronized via AHL at the colony level. (B) The coupled intracellular-quorum clock system oscillates asynchronously in small populations and transitions to synchronized oscillations in larger populations once the quorum clock fires. Despite lacking a mode of cell-cell communication itself, the coefficient of variation of the intracellular clock drops markedly via host-linked coupling with the quorum clock (bottom, data from 28 single cell traces). (C) IPTG reduces the intracellular clock period in small cell populations without the quorum clock (blue) and increases the coupled period in larger populations with the quorum clock (red). Each data point taken from 10-30 oscillatory peaks. The error bars indicate s.e.m. of the period, centered at the mean. (D) In our computational model, load-mediated coupling allows the intracellular clock to modulate the quorum clock period via degradation coupling at ClpXP, where the intracellular clock continues oscillating between coupled pulses and accelerates the pulse onset. (E) This adaptive form of pulse frequency modulation ensures that the pulse dynamics remain unchanged while the inter-pulse duration is adjusted (left: model and right: experimental, 6-9 oscillatory peaks) The error bars indicate s.e.m. of relative quorum clock period. (F) This mechanism also makes the coupled system more robust by enabling oscillation at higher media flow rates.

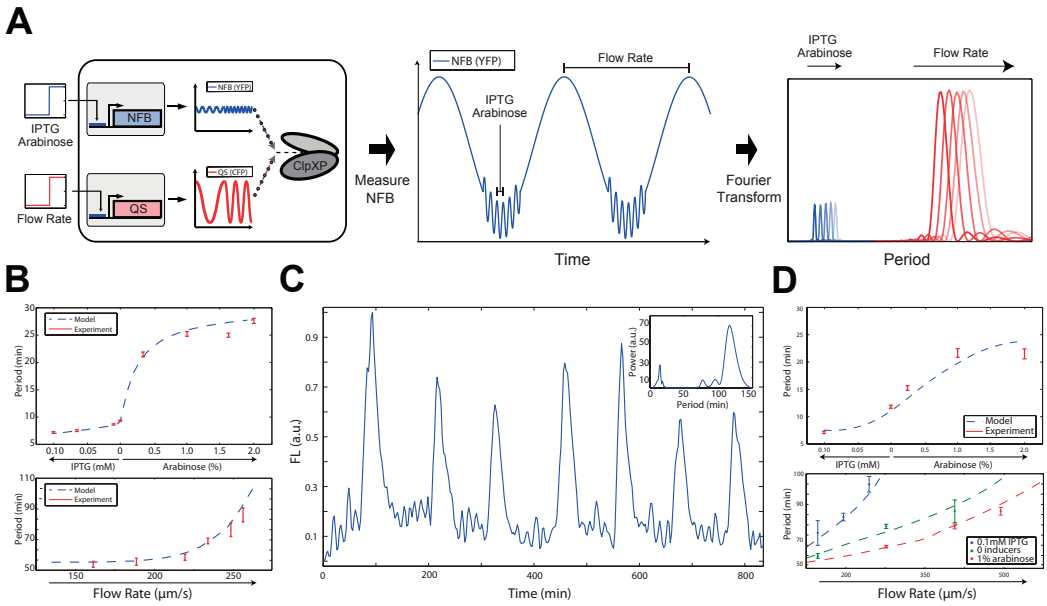

Figure 2.3: Genetic multispectral encoding. (A) Separate IPTG/arabinose and flow rate inputs are encoded into frequency-modulation oscillations that can be measured from the time series of the reporter for the intracellular clock. This engineered system is capable of encoding information from two underlying networks into a single multispectral time series. (B) Frequency response curves generated from experimental data and computational models for the intracellular clock (top, data from 30 single cell traces each) and quorum clock (bottom, model applied to data from the original study[27]) in isolation. The error bars indicate s.e.m. of the period, centered at the mean. (C) In the coupled system, frequency-modulated oscillations from both clocks can be observed in the output of the intracellular clock and extracted by inverse Fourier transform (inset). (D) Independent recovery of both IPTG/arabinose and flow rate inputs, where the frequency response of the intracellular clock to IPTG/arabinose is equivalent to the isolated clock (top) and the frequency response of the quorum clock is shifted by the intracellular clock (bottom). Periods calculated from 5-10 single cell traces for each condition. The error bars indicate s.e.m. of the period, centered at the mean.

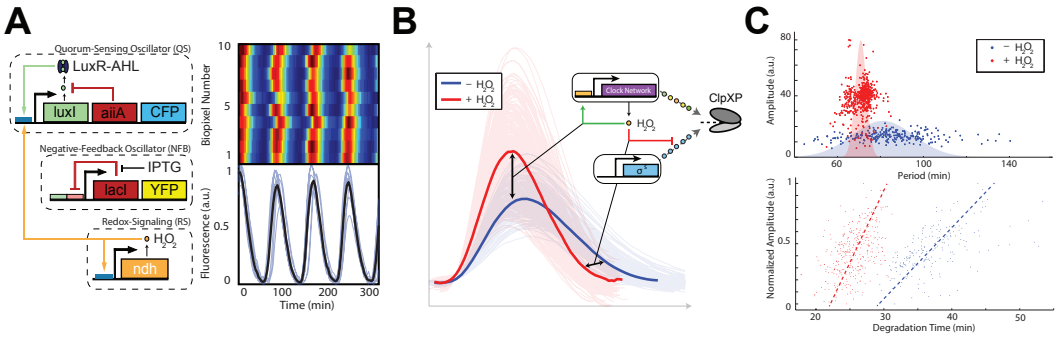

Figure 2.4: Post-translational coupling at the multi-colony level. (A) At the multi-colony level, interaction of H_2O_2 generated by redox signaling with the cellular stress response network synchronizes quorum clock oscillations between colonies. Traces taken from 10 separate colonies across the array. (B) Host-linked oscillations change distinct aspects of the waveform in response to H_2O_2 produced by the enzymatic activity of NDH. With H_2O_2 , oscillations have larger amplitudes and steeper downslopes, revealing increases in both transcription and degradation produced by the interaction of the synthetic clock network with the native stress response. (C) H_2O_2 increases the oscillatory amplitude while decreasing the required degradation time, revealing an increase in ClpXP activity. This increase in ClpXP capacity in response to H_2O_2 serves to mitigate the effects of transcriptional noise by minimizing the effects of amplitude variation on the period, resulting in a tightening of the period distribution with H_2O_2 (model: Fig. 2.8c-d).

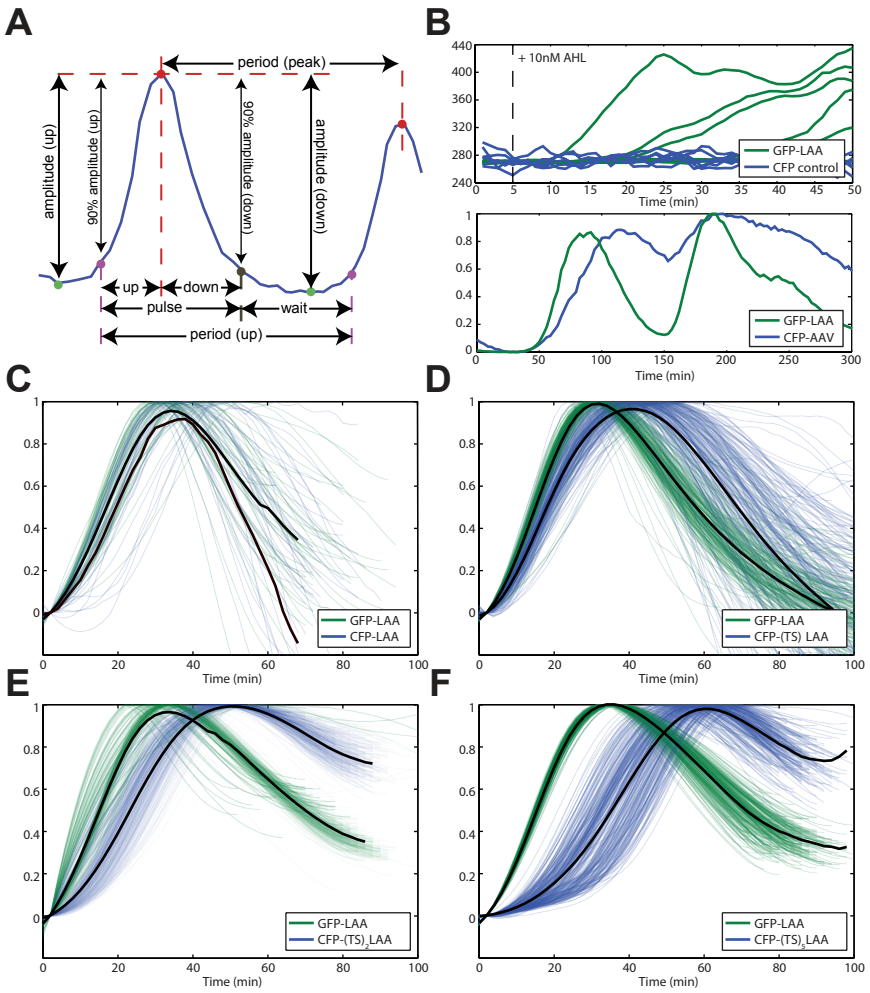

Figure 2.5: Increasing length of the TS linker sequence results in increasing downstream module degradation delay. (A) Detailed breakdown of single fluorescent trajectory analysis. Peaks are identified in red, troughs in green, upslope 10% points in purple and downslope 10% points in dark beige. The two period measurements are peak to peak and the time between two successive 10% upslope points. (B) Top: sfGFP does not show bleed-over into CFP fluorescence channel. Induction of sfGFP with 10nM AHL (dashed line) showed increase in fluorescence of sfGFP, which was not detected in CFP channel. Bottom: the use of the published AAV degradation tag[3] shows delay in the downstream module degradation of 15min. (C) Without the TS linker sequence, there is very little delay in downstream module degradation. (D) Single TS linker sequence results in 10 min delay. (E) Double TS linker sequence results in 16 min delay, similar to that of AAV degradation sequence. (F) 5-TS linker sequence results in 25 min delay (data shown in panels c-f was used to generate Fig. 2.1C).

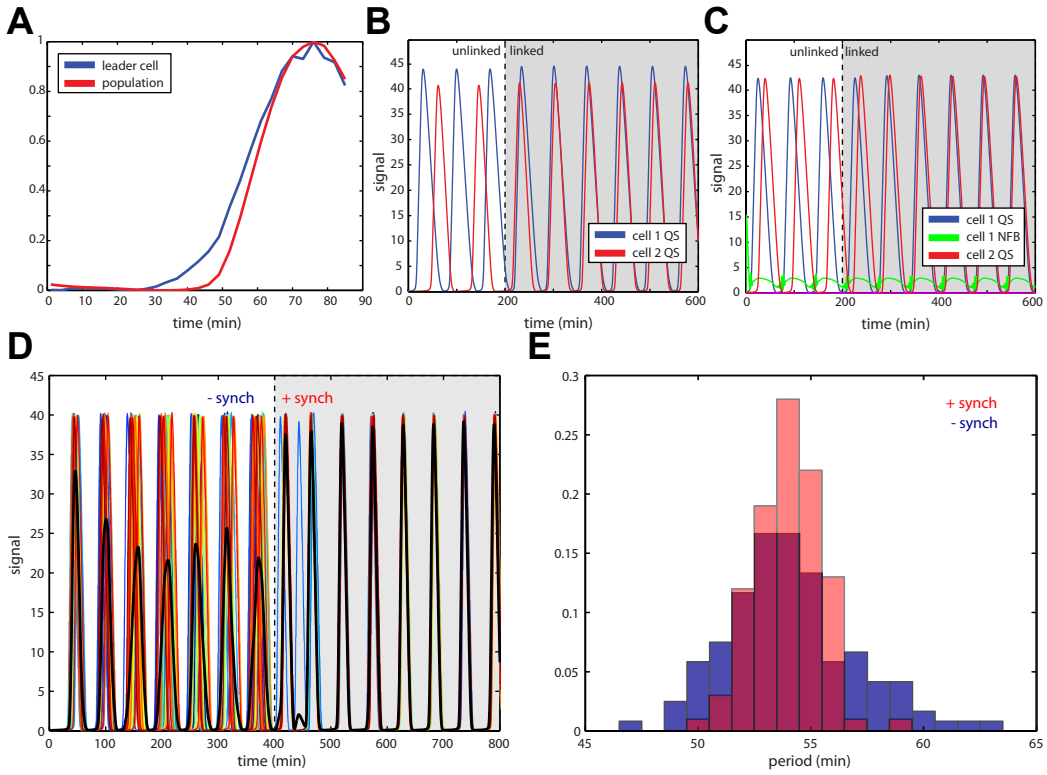

Figure 2.6: Cell-cell communication by AHL reduces variability in the quorum clock. (A) Individual "leader" cells show early activation of quorum clock proteins relative to the mean population response. (B) In a 2-cell simulation, cells 1 and 2 start out unlinked with slightly different constitutive production of AiiA and LuxI (α_0). At $t = 100$min the two cells are linked through external AHL in the media, showing the cell with slower dynamics (2) linking up to cell 1 with shorter periods. (C) Cells 1 and 2 start out unlinked with cell 1 including intracellular clock dynamics (green) that result in higher frequency oscillations in cell 1. When the cells are linked ($t = 100$), the slower cell 2, without the intracellular clock, links on to the faster cell through external AHL communication between the cells. (D) Trajectories of 20 cells with noisy constitutive production at lux promoter synchronize when their external AHL pool is mixed at $t = 400$min. Mean trajectory is shown in black. (E) Period variability after cell synching (red) is lower than in individual cells (blue).

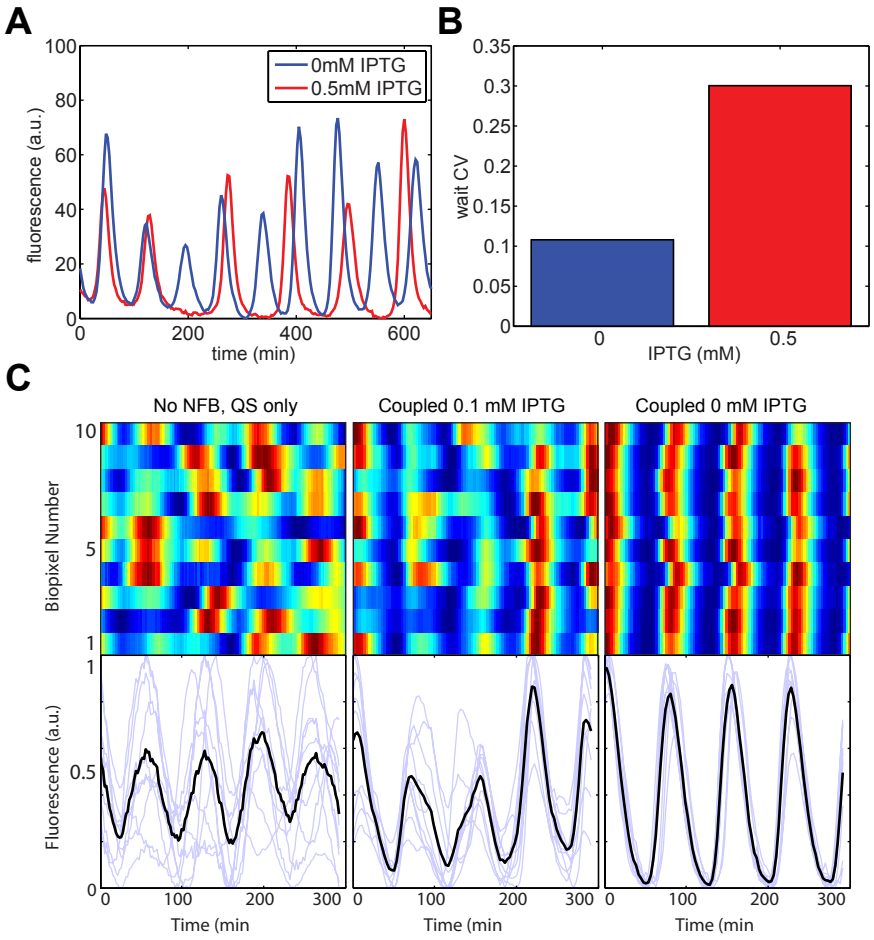

Figure 2.7: The intracellular clock increases robustness in the coupled oscillator system by reducing the period of the quorum clock. (A) Removal of IPTG, which increases intracellular clock strength, leads to more regular oscillations (experimental). (B) The decrease in variability of the inter-pulse time of the coupled oscillator without IPTG suggests that the intracellular clock plays a significant role in the inter-pulse dynamics (experimental). (C) At very high flow rate, the quorum clock oscillates irregularly. Tuning up the intracellular clock reduces the quorum clock period, restoring regular oscillations and allowing for global level synchronization between colonies due to H_2O_2 biopixel couping. Genetic addition of the intracellular clock (0.1mM IPTG) helps synchronize the quorum clock at high flows ($430 \mu m/s$). Increasing the strength of the intracellular clock with removal of IPTG further enhances H_2O_2 inter-colony synchronization (experimental).

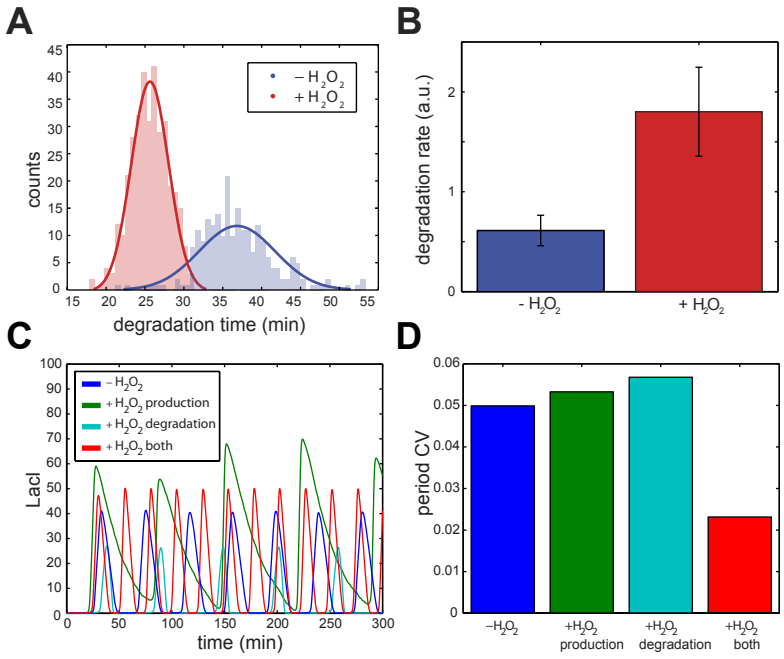

Figure 2.8: H_2O_2 increases degradation rate by ClpXP that, in combination with transcriptional increase at the *lux* promoter, decreases variability in the oscillator period. (A) There is a significant decrease in the degradation time due to H_2O_2 (experimental). (B) This is due to effective increase in ClpXP degradation rate (experimental). (C) H_2O_2 activation of lux promoter alone would only increase the amplitude of quorum clock oscillations. Similarly, H_2O_2 -dependent increase in ClpXP activity results only in steeper degradation and longer inter-pulse duration. Combination of the two effects leads to increase in amplitude and decrease in inter-pulse duration, which matches experiments (model). (D) Individually the two H_2O_2 effects do little to lower the quorum clock period CV, which is reduced when both are present (model).

Chapter 3

Genetic circuits in *Salmonella typhimurium*

3.1 Introduction

Synthetic biology has rapidly progressed over the last decade and is now positioned to impact important problems in health and energy. In the clinical arena, the field has thus far focused primarily on the use of bacteria and bacteriophages to overexpress therapeutic gene products. The next generation of multi-gene circuits will control the triggering, amplitude, and duration of therapeutic activity *in vivo*. This will require a host organism that is easy to genetically modify, leverages existing successful circuit designs, and has the potential for use in humans. Here, we show that gene circuits that were originally constructed and tested in *E. coli* translate to *Salmonella typhimurium*, a therapeutically relevant microbe with attenuated strains that have exhibited safety in several human clinical trials. These strains are essentially non-virulent, easy to genetically program, and specifically grow in tumor environments. Developing gene circuits on this platform could enhance our ability to bring sophisticated genetic programming to cancer therapy, setting the stage for a new generation of synthetic biology in clinically relevant microbes.

An explosion of DNA sequencing [63], synthesis [87], and manipulation [46] technologies has driven the development of synthetic genetic programs of increasing complexity in living cells [148;100;69]. Underlying this work is the hope that engineered biological systems will be used to solve important problems in energy and health over the coming years. Initially inspired by electronic circuits, researchers began by designing small transcriptional switches [45] and oscillators [37]. These early successes fostered a growing population of physicists, computer scientists, and engineers that aimed to apply an engineering-based methodology to the design of biological systems. In the past decade, substantial success has been achieved using this genetic circuits approach termed *synthetic biology* [148;100;55].

Multi-gene logic gates capable of integrating environmental signals have been constructed in bacteria [133], yeast [115], and mammalian cells [52]. Electronics-inspired networks have included counters [44], pulse generators [10], filters [126;62], and communication modules [154;10]. Sophisticated circuits can now be controlled by light, yielding genetic programs readily tunable both *in vitro* [132] and *in vivo* in live animals [153]. Dynamic genetic clocks have been constructed that function at the single-cell [128], colony [28], and multi-colony [106] level in growing bacterial populations, and even in mammalian cells [136]. In a recent study, redox signaling mediated by H_2O_2 vapor permitted the synchronization of millions of oscillating bacteria across an LCD-like sensor array [106].

Early efforts toward clinical applications have utilized bacteria [4;151;35;113;36] and bacteriophages [82;83] (viruses that infect bacteria) to perform therapeutic functions in vivo. Commensal bacteria have been engineered to fight diabetes [35], HIV [113], and cholera [36] by producing and delivering therapeutic agents directly in the human microbiome. Because certain bacteria grow preferentially in hypoxic environments, a number of studies have engineered cancer-fighting bacteria to selectively attack tumors [4;151]. Toward still another application, a pair of studies has engineered phages to produce foreign enzymes, making them far more potent than their unmodified counterparts at dispersing bacterial biofilms [82;83].

In most of these cases, the genetic programs involved were responsible for over-

expressing target genes, similar to traditional genetic engineering where genes are added, removed, or modified one at a time in a stepwise fashion. To truly achieve its clinical potential, synthetic biology must continue to do what has made it successful: engineer progressively more complex, multi-input networks in which the triggering, amplitude, and duration of therapeutic activity is controllable. This will require using hosts that are easy to genetically modify and compatible with the clinical requirements regarding safety, immunogenicity, and drug resistance. While bacteriophage and adenovirus have their advantages, viruses have smaller genomes and therefore have a narrower range of genetic modifications, frequently induce host resistance, and are highly cell-type specific[93;108].

As one potential bridge between organisms such as *E. coli* and clinically relevant microbes, *Salmonella typhimurium* is a bacterial anti-cancer platform that is closely related to *E. coli*, has been extensively studied in vivo for therapeutic applications[41;105;58;78;42;81] and has been shown to be safe in human clinical trials[57;139;102]. The development of attenuated strains has utilized auxotrophy and phoPQ deletions to suppress virulence cell invasion and virulence[58]. Lipid A mutations have been generated to reduce immunogenicity, stimulating a much weaker immune response than wild-type strains[81]. Despite this reduced potency, systemically injected *S. typhimurium* cells retain their ability to target and selectively replicate within tumors, displaying a thousand-fold growth preference relative to other organs[41;105;42;81;58]. Their motility allows them to follow chemical gradients and penetrate deep into the tumor vasculature[65;66], much further than passively diffusing small molecules[41]. And many of these strains also display innate oncolytic activity, regressing tumors simply by growing in them[105;99;81;64].

Perhaps the most important property of *S. typhimurium* for synthetic biology is the ease of genetic modification. It is a model organism whose genome is sequenced[88], has knockout collections, and the genetic tools are almost identical to *E. coli*. *S. typhimurium* is capable of stably expressing recombinant DNA from plasmid-based circuits *in vivo*. This approach has already been used to produce a number of therapeutic compounds directly within tumors, but most often via “always on“ expression of well-

established genes[58;53;103]. This work has laid the foundation for more sophisticated functionality, such as programmed delivery profiles that take advantage of plasmid instability[26]. Such a focus will merge the dynamic sensing, production, and delivery capabilities of genetic circuits with the native tumor seeking and penetration of *S. typhimurium*.

3.2 Experimental results

In order to test the degree to which existing synthetic circuits function in *S. typhimurium*, we transformed the attenuated strain ELH430 (SL1344 ΔphoPQ, gift of Elizabeth Hohmann, MGH)[61] with several genetic oscillator constructs. First, we tested a single-plasmid variant of a published single-cell gene oscillator[128]. Using our microfluidic platform[40;23], we observed robust oscillations for all *S. typhimurium* cells over many generations (Fig. 3.1A,B). While the qualitative period-inducer relationship was similar to *E. coli*, the curve was shifted toward faster periods as compared to *E. coli* strain JS006 (MG1655 ΔaraC,lacI) (Fig. 3.1C). In contrast, we initially expected *S. typhimurium* to oscillate slower since longer division times generally result in period lengthening[128]. When we measured the dependence of oscillatory period on temperature in *S. typhimurium*, we found the trend qualitatively similar to *E. coli*, where lower temperatures (and therefore longer doubling times) resulted in longer oscillatory periods (Fig. 3.1D). We therefore hypothesized that the faster oscillations in *S. typhimurium* are not due to growth rate differences, but rather a strain-dependent factor such as mean promoter level, transcription rate, or enzymatic degradation rate.

To explore this quantitatively, we used automated single-cell tracking using a previously developed algorithm[95] to compare a large number of single-cell time courses from *S. typhimurium* and *E. coli* . Oscillators are an ideal circuit to quantify strain-specific parameters such as transcription and degradation rates since they allow for hundreds of measurements in a single experiment. For each oscillatory period, the trough-to-peak and peak-to-trough slopes were measured. Since the ClpXP degradation machinery is likely saturated[22], the peak-to-trough slope yields an estimate for the

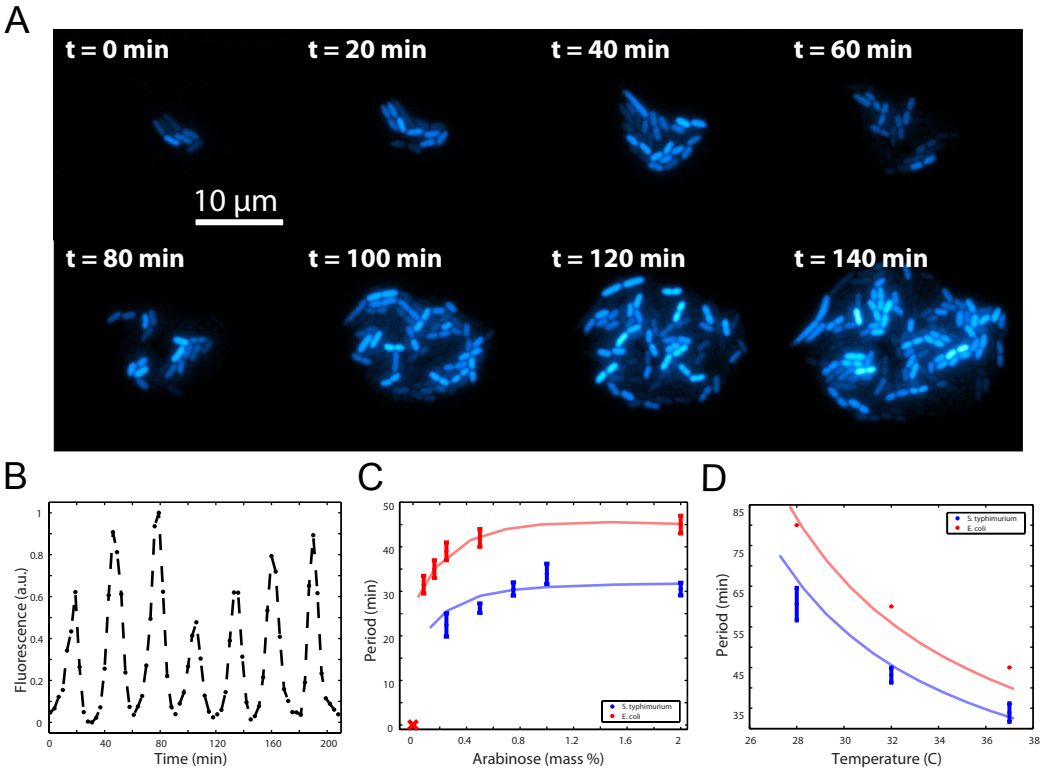

Figure 3.1: A fast, robust, and tunable genetic oscillator in *S. typhimurium*. (A) Timelapse fluorescence microscopy depicting asynchronous oscillations in a growing colony of *S. typhimurium*. (B) A single-cell trajectory extracted from image data. (C) Period vs. inducer concentration for *S. typhimurium* compared to original data taken in *E. coli*. The trends are qualitatively similar yet *S. typhimurium* is shifted toward shorter periods. Points are experimental measurements fit to a line generated by computational modeling. (D) Period vs. temperature for *S. typhimurium* compared to original data taken in *E. coli* with similar trends.

zeroth-order enzymatic degradation rate in degrade-and-fire oscillators[85]. Interestingly, we found that the apparent enzymatic degradation rate in *S. typhimurium* was roughly 1.5-fold that of *E. coli* (Fig. 3.4A). In our computational model of the oscillator, this increase reproduced the experimentally observed period-inducer relationship (Fig. B).

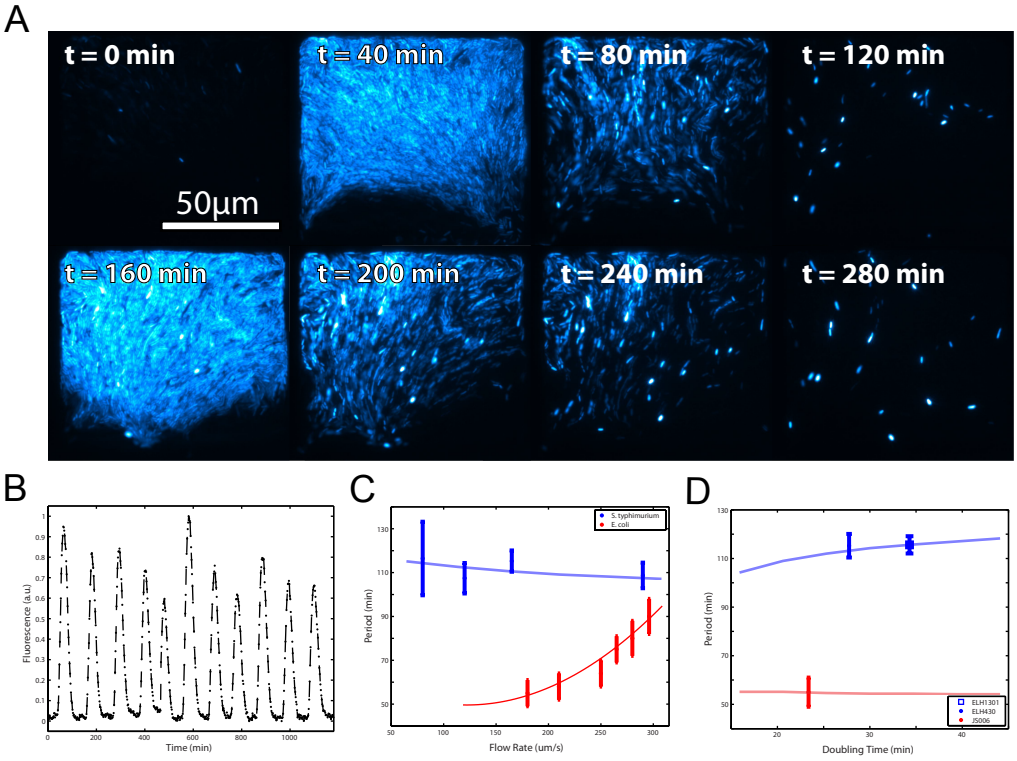

Figure 3.2: A synchronized quorum of genetic clocks in *S. typhimurium*. (A) Timelapse fluorescence microscopy depicting coherent oscillations at the colony-level for a growing colony of *S. typhimurium*. (B) A colony trajectory extracted from image data that illustrates the regularity of oscillations over time. (C) Period vs. flow rate for *S. typhimurium* compared to original data taken in *E. coli*. *S. typhimurium* displays much higher periods that appear to be independent of flow rate.

Next, we transformed *S. typhimurium* with a quorum-sensing oscillator that had been previously characterized in *E. coli*[28], and observed coherent, colony-level oscillations for more than 48 hours (Fig. 3.2A,B). Here, we found that the period-flow rate dependence was markedly different in *S. typhimurium* than in the original study, where oscillatory period was much longer and changed very little across a wide range of flow rates (Fig. 3.4C). Interestingly, while increased degradation rate resulted in faster os-

cillations for single-cells (Fig. 3.4B), our computational model correctly predicts the opposite trend for the quorum-sensing oscillator when degradation is increased (Fig. 3.4C).

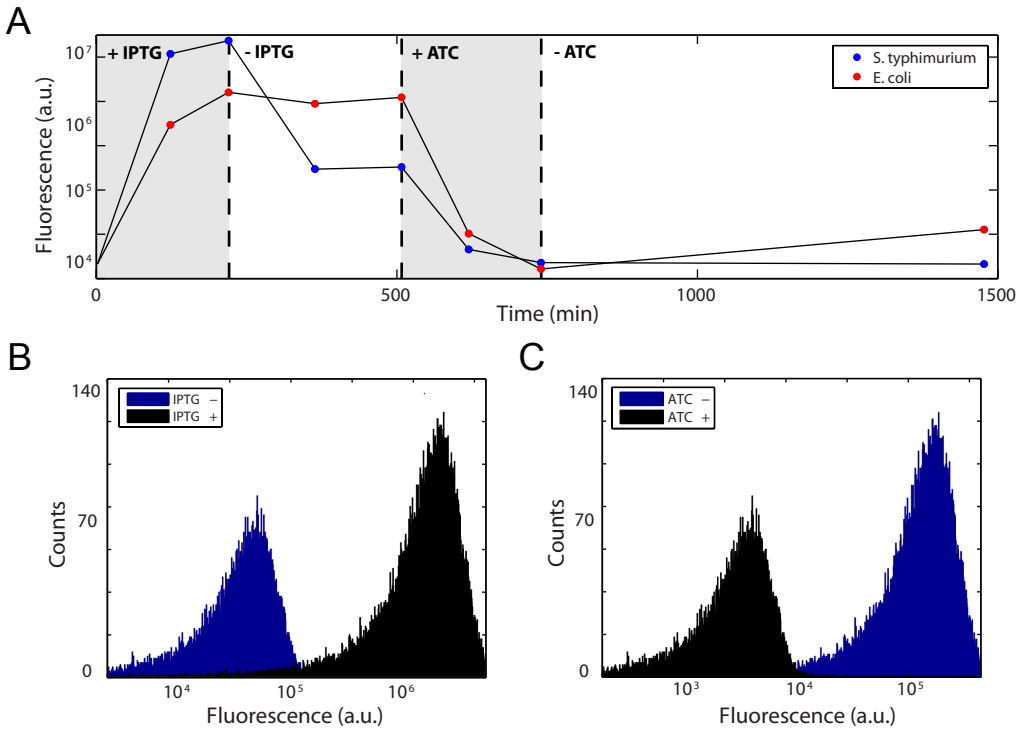

Figure 3.3: A genetic toggle switch in *S. typhimurium*. (A) A time course of fluorescence output that illustrates switching by both IPTG and ATC quantified by flow cytometry in periodically diluted batch culture experiments. (B) Raw flow cytometer data illustrating switching by 2mM IPTG and (C) 500 ng/μl ATC.

Finally, we tested the original genetic toggle switch, plasmid pIKE107[45]. In this circuit, a transient pulse of IPTG inducer turns the switch ON and reporter expression is maintained at a high level. A second pulse of ATC inducer turns the switch OFF, dropping reporter expression indefinitely. In periodically diluted batch culture experiments similar to the original study, we used flow cytometry to observe robust switching and bistability when inducing with either 2 mM IPTG or 500 ng/ul dox in cultures growing at 37 C (Fig. 3.3A-C). Interestingly, the fluorescence level at which *S. typhimurium* settled after we removed IPTG was lower than the same circuit in *E. coli* (Fig. 3.3A). We

suspected that the differences in apparent degradation and expression rates (Fig. 3.4A) might explain this change, since the steady-state repressor balance would be adjusted.

To test this hypothesis, we used the original computational model of the toggle switch[45] and quantified the steady-state expression level over time for strain parameters measured in *E. coli* and *S. typhimurium*. We found that the *S. typhimurium* parameters reproduced the experimentally observed curves, where expression rises to a higher level when switched ON then decays to a lower steady-state when IPTG is removed (Fig. 3.4D). While these parameters are particularly important for dynamic circuits, they can also impact the performance of stable switches since repressors are continuously being produced and degraded.

3.3 Conclusions and outlook

A central issue in the design of genetic circuits is the degree to which native and engineered networks should be integrated. Synthetic biology began by fully isolating itself from the strain background, using it solely to supply energy, enzymatic machinery, and a cellular volume in which to function. In contrast, industrial applications in medicine and energy have commonly utilized a variety of microbes for their native networks[120;148;2]. As our biological knowledge of native networks and our ability to engineer new circuits has improved, it has become increasingly possible to blend these two strategies[100].

S. typhimurium is an ideal strain for clinical synthetic biology since it is closely related to *E. coli*, well studied *in vivo*, has safety precedence for clinical trials in humans, and displays a thousand-fold growth preference for tumor environments[41;105;58;78]. Moving to other microbes for clinical and industrial purposes will require the determination of the critical strain parameters that define the space of bacteria capable of hosting genetic circuits. Next steps will involve measurement of these parameters and testing circuits in strains of interest that are further removed in the phylogenetic tree[150]. One such roadmap would begin with more distantly related gamma proteobacteria like *Pseu-*

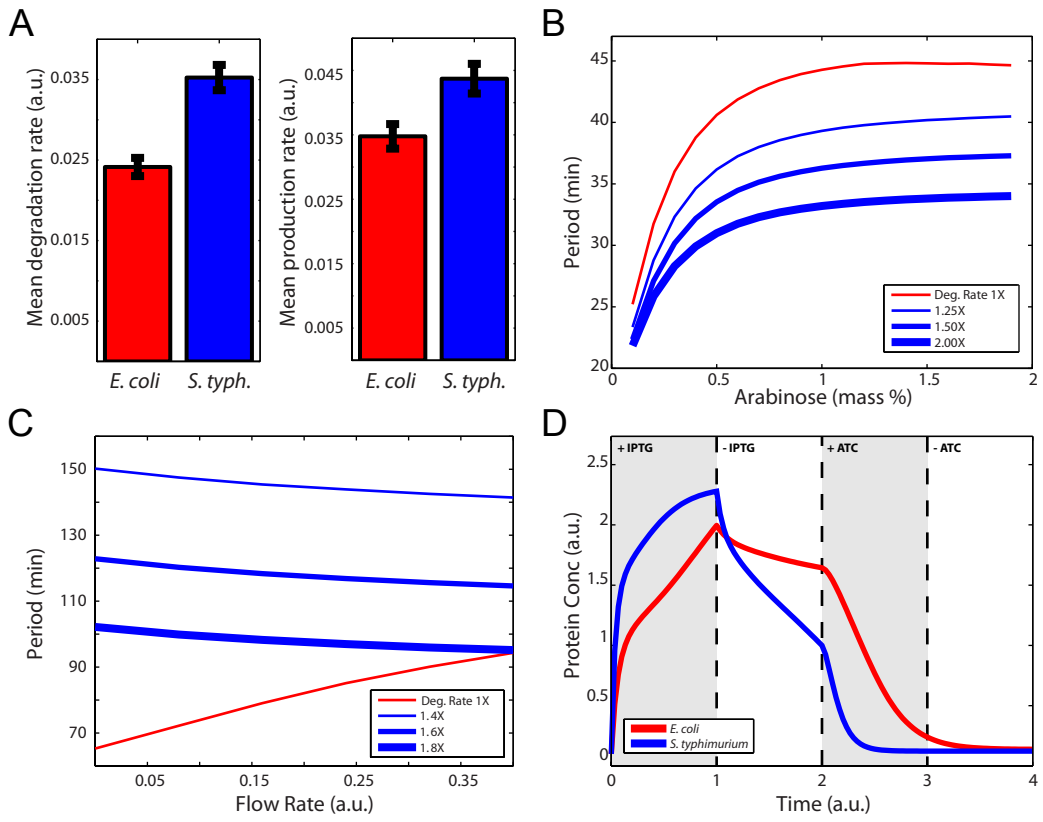

Figure 3.4: Computational modeling of *S. typhimurium* genetic circuits. (A) Comparison of enzymatic degradation rate between *S. typhimurium* and *E. coli* generated from automated single-cell tracking. Degradation rate is approximately 1.5X higher in *S. typhimurium*. (B) A higher degradation rate results in the shorter periods observed experimentally for the single-cell oscillator. (C) In contrast, increased degradation rate results in longer periods for the quorum-sensing oscillator that are comparatively unchanged with flow rate. (D) Increased degradation and expression rates produce the experimentally observed behavior for the *S. typhimurium* toggle switch.

domonas aeruginosa before moving outside the phylum to alpha proteobacteria such as *Calubacter crescentus*. Additionally, individual components and modules can also receive a "portability" score that estimate the degree to which they translate to other hosts. For example, while lacI- and tetR-based circuits are nearly universal, more generally the function of other components are likely to be more sensitive to strain-specific parameters. This work will enable synthetic biology to move beyond *E. coli* into a diverse range of microbes for clinical and industrial applications.

3.4 Microscopy and microfluidics

We used a microscopy system similar to our recent studies[28]. Fluorescent images were taken at 4X every 30 seconds using the EMCCD camera (20ms exposure, 97% attentuation) or 2 minutes (2s exposure, 90% attenuation) using a standard CCD camera to prevent photobleaching or phototoxicity.

In each device, *E. coli* cells are loaded from the cell port while keeping the media port at sufficiently higher pressure than the waste port below to prevent contamination. Cells were loaded into the cell traps by manually applying pressure pulses to the lines to induce a momentary flow change. The flow was then reversed and allowed for cells to receive fresh media with 0.075% Tween which prevented cells from adhering to the main channels and waste ports.

To measure fluid flow rate before each experiment, we measured the streak length of fluorescent beads (1.0 μm) upon 100 ms exposure to fluorescent light. We averaged at least 1,000 data points for each.

We used several microfluidic devices over the course of the study. For single-cell oscillators (Fig. 3.1), we used a previously described device consisting of a trapping region and a dynamic switch[95]. Traps have dimensions 40 μm wide x 50 μm long x 0.95 μm high, with the long sides open to media flow. Since *E. coli* and *S. typhimurium* cells have a 1 μm diameter, the trap maintains growing cells in a monolayer. For colony oscillators (Fig. 3.2), we used a previously described device consisting of arrays of

square trapping regions[28;106]. Trap dimensions were always 100 μm x 85 μm x 1.65 μm high and spacing between traps was 25 μm. This size allows cells to grow in a colony arrangement rather than a monolayer, while still allowing quantitative measurement of colony fluorescence.

3.5 Degradation and production rate quantification

Single cell fluorescence trajectories were obtained from time-lapse movies using custom software previously developed in MATLAB[95]. Each cell fluorescence trajectory represents the median GFP fluorescence signal inside that cell over time. Using built-in MATLAB functions we identified the peaks and troughs for each trajectory. The degradation rate was calculated by taking the amplitude change from peak to the successive trough and dividing by the time change between the peak and the trough. These peak-to-trough sections of the trajectory represent the time when the production of GFP is repressed and the observed dynamics are solely driven by degradation of GFP. Similarly we calculated the net production rate, by calculating the amplitude change from trough to successive peak and dividing by the time change between the trough and the peak. The measurement gives the net production rate, which includes the degradation of the protein.

We find the mean degradation rate for *E. coli* to be 0.024 (0.001 s.e.) and mean net production rate to be 0.035 (0.002 s.e.). For *S. typhimurium* the mean degradation rate was estimated to be 0.035 (0.002 s.e.) and mean net production rate to be 0.044 (0.002 s.e.).

3.6 Modeling

To generate the plot in Figure 3.4D, we used previously described genetic toggle switch model[45]. We included three additional parameters to model the effects of IPTG (C_{IPTG}), ATC (C_{ATC}), and dilution (D) on the synthesis and degradation of proteins:

$$\frac{\partial u}{\partial t} = \frac{C_{IPTG_{(0,1)}}\alpha_u}{1+v^n} - (\gamma_u + D)u$$
$$\frac{\partial v}{\partial t} = \frac{C_{ATC_{(0,1)}}\alpha_v}{1+u^n} - (\gamma_v + D)v$$

In this model, we set n=2 to allow for cooperativity of repression of both promoters. C_{IPTG_0} and C_{ATC_0} were set to 1 for the case of no inducers present. Next, we used metropolis algorithm to find the rest of the parameters to fit the qualitative nature of the curves from Figure 3.1A. The parameters found to generate the *E. coli*curve were: $C_{IPTG_1} = 1.25$,$C_{ATC_1} = 1.68$,$\alpha_u = 4.28$,$\alpha_v = 5.80$,$\gamma_u = 1.76$,$\gamma_v = 2.37$,$D = 0.11$. The parameters found to generate the *S. typhimurium*curve were: $C_{IPTG_1} = 1.25$, $C_{ATC_1} = 1.68$,$\alpha_u = 11.00$,$\alpha_v = 8.36$,$\gamma_u = 4.86$,$\gamma_v = 3.21$,$D = 0.08$. It is interesting to note that the optimized parameters show higher production and degradation as well as lower dilution for *S. typhimurium*curve relative to *E. coli*curve, which correlates well with our experimental measurements.

The dynamics of single cell oscillator were modeled using previously described model for activator (a_2) and repressor (r_4) proteins[128]. The production and degradation of these proteins is described by the following set of reactions:

$$
\begin{aligned}
P_{0,0}^{a/r} &\xrightarrow{b_{a/r}} P_{0,0}^{a/r} + m_{a/r} \\
P_{1,0}^{a/r} &\xrightarrow{\alpha b_{a/r}} P_{1,0}^{a/r} + m_{a/r} \\
m_a &\xrightarrow{t_a} m_a + a_{uf} \\
m_r &\xrightarrow{t_r} m_r + r_{uf} \\
a_{uf} &\xrightarrow{k_{fa}} a \\
r_{uf} &\xrightarrow{k_{fr}} r \\
a + a &\underset{k_{-da}}{\overset{k_{da}}{\rightleftharpoons}} a_2 \\
r + r &\underset{k_{-dr}}{\overset{k_{dr}}{\rightleftharpoons}} r_2 \\
r_2 + r_2 &\underset{k_{-t}}{\overset{k_t}{\rightleftharpoons}} r_4 \\
a_{uf} &\xrightarrow{\lambda f(X)} \emptyset \\
r_{uf} &\xrightarrow{f(X)} \emptyset \\
a &\xrightarrow{\lambda f(X)} \emptyset \\
r &\xrightarrow{f(X)} \emptyset \\
a_2 &\xrightarrow{\lambda f(X)} \emptyset \\
r_2 &\xrightarrow{f(X)} \emptyset \\
r_4 &\xrightarrow{f(X)} \emptyset
\end{aligned}
$$

We updated the degradation function $F(X)$ to include dilution as follows:

$$f(X) = \frac{\gamma}{c_e + X} + DX$$

Here, X is the total number of $ssrA$ tags in the system (one for each monomeric version, two for dimers, and four for tetramers, including proteins bound to operator sites). We varied the parameter γ from 1x to 2x to evaluate the effect of degradation difference between *E. coli*and *S. typhimurium*on the period of oscillation calculated from

single cell model simulations. Dilution rate was calculated from experimentally measured cell half life as $\frac{ln(2)}{T_{\frac{1}{2}}}$.

To model the dynamics of the quorum-sensing oscillator, we used our previously described model for intracellular concentrations of LuxI (I), AiiA (A), internal AHL (H_i), and external AHL (H_e)[28],

$$\frac{\partial A}{\partial t} = C_A[1-(d/d_0)^4]\,G(\alpha,\tau) - \frac{\gamma_A A}{1+f(A+I)} - DA \tag{3.1}$$

$$\frac{\partial I}{\partial t} = C_I[1-(d/d_0)^4]\,G(\alpha,\tau) - \frac{\gamma_I I}{1+f(A+I)} - DI \tag{3.2}$$

$$\frac{\partial H_i}{\partial t} = \frac{bI}{1+kI} - \frac{\gamma_H A H_i}{1+gA} + D(H_e - H_i) - DH_i \tag{3.3}$$

$$\frac{\partial H_e}{\partial t} = -\frac{d}{1-d}D(H_e - H_i) - \mu H_e + D_1\frac{\partial^2 H_e}{\partial x^2} \tag{3.4}$$

To model the difference in periods of oscillation between *E. coli*and *S. typhimurium* we varied the degradation parameters γ_A and γ_I. We looked at the changes in the period over different values of the flow rate parameter μ, while varying the degradation prapameters from 1x to 2x of the original model value. To account for the difference in doubling time between the two strains, we introduce exponential decay terms into the model to account for dilution in addition to the enzymatic degradation terms. We add terms $-DI$, $-DH_i$, and $-DH$ to the first three equations respectively, with $D = \frac{ln(2)}{T_{\frac{1}{2}}}$. We then looked at how the change in doubling time affected the period of both strains shown in Figure 3.2D.

3.7 Acknowledgments

Chapter 3 contains material originally published as Prindle, A., Selimkhanov, J., Danino, T., Tsimring L., and Hasty, J. 2012: Genetic circuits in *Salmonella typhimurium*. *ACS Synthetic Biology*, 10: 458-464. Copyright permission to republish here was granted by ACS.

Chapter 4

Dual delayed feedback provides sensitivity and robustness to the NF-κB signaling module

4.1 Introduction

Many cellular stress-responsive signaling systems exhibit highly dynamic behavior with oscillatory features mediated by delayed negative feedback loops. What remains unclear is whether oscillatory behavior is the basis for a signaling code based on frequency modulation (FM) or whether the negative feedback control modules have evolved to fulfill other functional requirements. Here, we use experimentally calibrated computational models to interrogate the negative feedback loops that regulate the dynamic activity of the transcription factor NF-κB. Linear stability analysis of the model shows that oscillatory frequency is a hard-wired feature of the primary negative feedback loop and not a function of the stimulus, thus arguing against an FM signaling code. Instead, our modeling studies suggest that the two feedback loops may be tuned to provide for rapid activation and inactivation capabilities for transient input signals of a wide range of durations; by minimizing late phase oscillations response durations

may be fine-tuned in a graded rather than quantized manner. Further, in the presence of molecular noise the dual delayed negative feedback system minimizes stochastic excursions of the output to produce a robust NF-κB response.

Many important signal transduction pathways contain a negative feedback motif consisting of an activator that activates its own repressor. Activated repression is capable of generating oscillatory behavior[141] and has been observed to do so in biological systems such as the Hes1 regulatory protein which controls neuronal differentiation[124], the p53-Mdm2 system that mediates the DNA damage response[75], and the NF-κB (Q04207) signaling network that governs the immune response and inflammation[59;101].

The role of activated repression is well understood in the context of transient signaling as functioning to limit the duration of the induced activity. Indeed, misregulation of the negative feedback mechanisms that control NF-κB and p53 has been shown to generate prolonged inflammatory or genotoxic stress responses, respectively, that lead to cell death or chronic disease[143;134]. Further, negative feedback can sensitize and speed-up responses to weak or transient input signals[122] when compared to constitutive attenuation mechanisms.

In contrast, the physiological role of oscillatory signaling behavior remains poorly understood. Recent work has shown that, in the calcium stress pathway in yeast, the frequency of nuclear localization of a stress-response transcription factor can be modulated by the magnitude of the extracellular calcium concentration, and this frequency modulation results in a coordinated expression of target genes[19]. In the NF-κB and p53 signaling systems, the function of oscillations is still unknown. Oscillations in p53 activity were proposed to represent a counting mechanism that quantizes the response, ensuring a robust but appropriate amount of activity for a specific degree of DNA damage[84]. An alternate view was proposed in which oscillations of the p53-controlling ATM kinase activity allow for periodic sampling of the damaged DNA to track its repair and, if necessary, drive further p53 signaling to sustain the repair programs[11]. Oscillations in NF-κB activity were proposed to determine which genes would be transcriptionally induced, thereby representing a temporal code that conveys information about the stim-

ulus to gene promoters[101]. However, it is not clear whether or not the frequency encodes information in this systems as no differences in NF-κB target gene expression were observed between oscillating and non-oscillating genetic variants[9].

Recent work has demonstrated that oscillations in NF-κB activity can be generated by pulsatile stimulation with TNFα (P06804)[6]. However, an analysis of the repeated activation of NF-κB that is driven by an oscillating signal provides little information about the role of oscillations that naturally arise with persistent stimulation. Thus, the role(s) of oscillations in NF-κB activity remains unclear and several questions are still unanswered: Do these oscillations convey information encoded in the frequency to downstream processes? Do they function to generate a periodically recurring phase of sensitivity to stimuli or regulatory crosstalk representing a potential "counting" mechanism? Do they "quantize" the output signal, thus specifying robust units of activity? Or, are the oscillations caused by persistent signaling simply a non-functional by-product of the requirement for the negative feedback architecture to enable sensitive, fast responses to transient stimuli?

Mathematical models comprised of a small number of equations have led to a greater understanding of biological processes in terms of molecular interactions, diffusion, dose responses, gradient sensing, the role stochasticity in gene expression and in fate decisions[76;74;5;7]. Although several models of networks with autoregulation have been developed[77;112;70], most of these networks do not incorporate delays. In signaling, however, such elegant models often do not faithfully reproduce the dynamic behavior of the signaling system because actual biological networks involve many molecular interactions that tend to slow overall signal processing. Larger models comprised of many molecular species and parameters have proven useful in exploring dynamic signaling behavior via computational simulations in conjunction with experimental studies, but they are analytically intractable and therefore do not provide the degree of conceptual insights that small models do.

Here we pursue an alternative approach to modeling NF-κB signaling. We construct a new model that replaces cascading reactions with a single but delayed compound

reaction that enables both recapitulation of experimentally observed dynamics and the use of powerful analytical tools. With these tools, we explore the physiological function of the dynamic behavior of NF-κB produced by the activated repression mechanism mediated by its inducible inhibitors, IκBα (Q9Z1E3) and IκBϵ (O54910). The mathematical analysis results in predictions that are addressed experimentally and thus lead to fundamental insights about the function and origins of this signaling system.

4.2 Results

4.2.1 NF-κB model formulation

The basic structure of the NF-κB signaling module is shown in Figure 4.1*A*[59]. In resting cells, NF-κB is sequestered in the cytoplasm by IκB proteins. Cellular stimulation leads to activation of the IκB kinase (IKK) which phosphorylates IκB proteins thus targeting them for degradation. Upon degradation of IκB proteins, NF-κB moves into the nucleus and activates hundreds of target genes including the predominant IκB isoform, IκBα. Synthesized IκBα enters the nucleus, binds to NF-κB, and the IκBα-NF-κB complex is exported back to the cytoplasm. Thus, the core feature of the NF-κB signaling module is a negative feedback loop mediated by IκBα. This can be abstracted to a simple motif in which x (NF-κB) activates y (IκBα), y represses x, and repression of x by y is relieved by K (active IKK) (Figure 4.1*B*).

Using this motif as a guide, we formulated our model of the IκBα-mediated NF-κB response as a set of 9 reactions and 6 variables. Specifically, the model assumes that the total number of the NF-κB molecules (X) is conserved, however they can exist either in free/nuclear form (x) or sequestered outside of nucleus within the IκBα-NF-κB complex ($[xy]$). The model contains non-delayed reactions for the binding of free NF-κB to the unbound IκBα promoter (d_{0y}) to form the bound IκBα promoter (d_{1y}), binding of IκBα protein (y) to free NF-κB to form the IκBα-NF-κB complex, constitutive degradation of IκBα, and induced degradation of free and bound IκBα proteins by

the active IκB kinase IKK (K) producing free NF-κB. In contrast, a compound delayed reaction describes the synthesis of IκBα protein. This reaction involves a time delay τ_y, which represents the time needed for transcription, translation, nuclear import and export, and protein-protein interactions.

Using experimentally validated assumptions, we reduced the set of mass-action kinetics equations for the 9 reactions to a single delay-differential equation:

$$\dot{Y}(y) = a_y d_{0y}(y_\tau) + b_y d_{1y}(y_\tau) - g_y y - K r_y Y(y) \tag{4.1}$$

where $Y(y) = y[1 + c_y X/(1 + c_y y)]$ is the total IκBα concentration (the sum of free IκBα (y) and IκBα bound to NF-κB), $d_{0y}(y) = F d_y/[F + X/(1 + c_y y)]$, $d_{1y}(y) = X d_y/[F(1 + c_y y) + X]$, are the probabilities for the IκBα promoter to be free or bound to NF-κB, respectively, $F = f_1/f_0$, $c_y = k_y/(k_{-y} + K r_y)$, and the subscript τ denotes the variable taken at time $t - \tau_y$ (see Methods for details of the derivation).

Mirroring the biological system, the non-dimensional time-dependent parameter $K(t)$, which characterizes the active IKK concentration, is used as the proxy input signal. The first term in the r.h.s. of Eq. 4.1 represents constitutive synthesis from the unbound IκBα promoter, the second term represents induced synthesis from the NF-κB-bound IκBα promoter, the third term represents constitutive degradation of IκBα protein, and the fourth term represents IKK-induced degradation of IκBα. Values of $K \sim 1$ correspond to the rate of IKK-induced degradation of NF-κB-IκBα complex which is of the same magnitude as unbound IκBα. Nuclear NF-κB level x at any time can be determined directly from IκBα levels via $x = X/(1 + c_y y)$. The time delay τ_y is incorporated in the synthesis terms: we assume that the rate of production of new proteins at time t depends on the state of the system at time $t - \tau_y$. Incorporating this time delay allows us to explore the behavior of the negative feedback loop without simulating the full set of reactions associated with it. We obtained values for the time delay and for the other model parameters by calibrating the behavior of the model with experimental results. As a starting point, we used parameter values from biochemical

measurements[67]. However, some modifications were necessary because these values represent the rates of single reaction steps and the model contains compound reactions.

To validate the model, we compared it to experiments. In response to a persistent input signal (starting at time $t = 0$), our simulations of the IκBα-mediated negative feedback system show pronounced oscillations in nuclear NF-κB levels with an oscillation period of about 90 minutes (Figure 4.1*C*). Oscillations with a similar period were observed experimentally when mutant cells containing only the IκBα feedback loop were persistently stimulated with the inflammatory cytokine TNF (Figure 4.1*D*).

To address the dynamics of the wild-type NF-κB system that feature both IκBα and IκBϵ feedback loops, we expanded the model to include an additional 9 reactions and 4 variables involving IκBϵ (Tables 3, 4). Following the same reduction procedure (see Methods for derivation), we derived a deterministic model consisting of two coupled delay-differential equations for the concentrations of the two IκB isoforms, IκBα (y) and IκBϵ (z),

$$\dot{Y} = a_y d_{0y}(y_\tau, z_\tau) + b_y d_{1y}(y_\tau, z_\tau) - g_y y - K r_y Y \tag{4.2}$$

$$\dot{Z} = m\left[a_z d_{0z}(y_\tau, z_\tau) + b_z d_{1z}(y_\tau, z_\tau)\right] - g_z z - K r_y Z \tag{4.3}$$

where $Y = y[1 + c_y X/(1 + c_y y + c_z z)]$, $Z = z[1 + c_z X/(1 + c_y y + c_z z)]$, $d_{0y,0z} = F d_{y,z}/(F + x)$, $d_{1y,1z} = x d_{y,z}/[F + x]$, $x = X/(1 + c_y y + c_z z)$, $c_{y,z} = k_{y,z}/(k_{-y,-z} + K r_{y,z})$, and $y_\tau = y(t - \tau_y)$, $z_\tau = z(t - \tau_z)$. Parameter $0 < m < 1$ here is the scaling factor which characterizes the relative strength of the secondary feedback loop.

In Eqs. 4.2 and 4.3, Y represents total IκBα (the sum of free IκBα (y) and IκBα bound to NF-κB ($[xy]$), and Z represents total IκBϵ (the sum of free IκBϵ (z) and IκBϵ bound to NF-κB ($[xz]$)). The terms in the r.h.s. of Eqs. 4.2 and 4.3 again represent constitutive synthesis from the identical unbound IκBα and IκBϵ promoters, induced synthesis from the NF-κB-bound promoters, constitutive degradation of IκBα and IκBϵ proteins, and IKK-induced degradation of IκBα and IκBϵ. Nuclear NF-κB levels are determined directly by IκBα and IκBϵ levels. Parameter values for the IκBα-mediated reactions were determined in the previous section. For the IκBϵ feedback reactions, we

use the same parameter values except for the constitutive synthesis and the constitutive degradation rates, which were chosen based on experimental measurements[67]

4.2.2 Is the oscillation period a function of the stimulus?

The advantage of our modeling approach is that it allows for analytical studies of the network dynamics. Here, we perform a linear stability analysis of the delay-differential equation (4.1) to identify the characteristic period and decay rate of NF-κB oscillations produced when input signal is present ($K > 0$). For sufficiently large K, induced synthesis and degradation are much stronger than basal ones, so the latter can be neglected ($a_y = g_y = 0$).

Expressing y via Y and substituting it into d_{0y}, d_{1y} yields a closed equation for Y in the form

$$\dot{Y} = G(Y_\tau) - Kr_y Y \tag{4.4}$$

where $Y_\tau = Y(t - \tau_y)$ and the function $G(\cdot)$ has the form

$$G(Y) = d_y \frac{b_y c_y (X - Y) - b_y + b_y \sqrt{(c_y(X - Y) - 1)^2 + 4c_y X}}{2c_y F + c_y(X - Y) - 1 + \sqrt{(c_y(X - Y) - 1)^2 + 4c_y X}} \tag{4.5}$$

The fixed point Y_s (stationary solution) of this equation is given by the algebraic equation

$$G(Y_s) = Kr_y Y_s \tag{4.6}$$

The stability of this solution is determined by the eigenvalue of the linearized equation (4.4) linearized near the fixed point Y_s (see Methods for details). The corresponding eigenvalue can be found in terms of the Lambert function $W(z)$ defined via $We^W = z$,

$$\lambda = \tau_y^{-1} W(B\tau_y e^{Kr_y\tau_y}) - Kr_y \tag{4.7}$$

The imaginary part of λ gives the oscillation frequency $\omega = 2\pi f$, and the (negative) real part of λ gives the decay rate δ of oscillations. Plotting the period ($T = 1/f$) (Figure 4.2*A*) and decay (δ) (Figure 4.2*B*) of the oscillations as a function of the delay

reveals a strong dependence. In contrast, the signaling perturbation K (the active IKK kinase) that acts as the input for the model determines the amplitude of the response but only negligibly affects the period or the oscillation decay (Figure 4.2*B*). The mathematical reason for this asymmetry is that the imaginary part of the Lambert function $W(x)$ for negative values of its argument changes very weakly for arguments below -2 ($ImW(-2)] = 1.67...$, $ImW(-20) = 2.27...$) and asymptotically approaches π for very large negative values of the argument. This is why the period of dampened oscillations ($2\pi\tau_y/Im[W(B\tau_y e^{K\tau_y\tau_y})]$) depends strongly on delay τ_y and only very weakly on K. Meanwhile, the real part of the eigenvalue λ (the decay rate) is linearly proportional to K because of the second term in Eq.(4.7) and also strongly depends on τ_y because of the first term. Thus, we find that the period is highly dependent on the delay but is rather insensitive to changes in the input level. This is confirmed by direct simulations of the full nonlinear equation (4.1), where time series of x are plotted for several different values of τ_y and K (Figure 4.9). Since variations of stimulus do not lead to significant frequency modulation of NF-κB activity, oscillations of NF-κB are unlikely to encode information about the stimulus.

4.2.3 Damping of oscillations in a dual delayed feedback loop system

The main qualitative difference between the one-loop system considered in the previous section, and the wild-type NF-κB module is the presence of another IκB isoform, IκBϵ, which also provides negative feedback regulation on NF-κB activity (Figure 4.3*A, B*), however with slower kinetics[67]. Experimental and computational work has shown that IκBϵ-mediated feedback can cause damping of IκBα-mediated oscillations[67] and (Figure 4.3*C*). More recent computational work has predicted that IκBϵ-mediated feedback desynchronizes oscillations but does not dampen oscillations in single cells[6]. Thus, the mechanism by which IκBϵ-mediated feedback produces damped oscillations at the population level is not well established. Furthermore, it is unknown whether the

damping function of the IκBϵ-mediated feedback loop has evolved to achieve a specific regulatory function or may simply be a secondary consequence of another function. We hypothesize that the primary role of the second feedback loop is to mitigate oscillatory behavior produced by the first feedback loop.

To address our hypothesis that IκBϵ-mediated feedback specifically evolved to dampen IκBα-mediated oscillations, we performed a parameter optimization procedure on the wild-type model (Eqs. 4.2 and 4.3) to determine the IκBϵ synthesis parameters that result in maximum damping. To characterize the degree of damping, we chose the maximum peak-trough difference after 6 hrs as a metric for the persistence of oscillations. According to the definition of this performance metric, "optimal damping" occurs when this metric is minimized. In our optimization procedure, we varied two important parameters, the time delay of the second feedback loop τ_z and the scaling factor m which simultaneously varies the rates of constitutive and induced synthesis of IκBϵ. Choosing $m = 0$ is equivalent to the complete removal of the IκBϵ-mediated negative feedback loop while $m = 1$ represents the case in which the inducible synthesis rates for IκBϵ are the same as for IκBα. The two-dimensional optimization search is shown in a color map (Figure 4.3*D*) indicating that the performance metric is minimized at $m = 0.3, \tau_z = 72$. Time course simulations with the optimized parameter set show a high degree of damping (Figure 4.3*G*) similar to what is observed experimentally (Figure 4.3*C*).

To determine whether these optimized parameter values correspond to observations, we measured relevant parameter values experimentally. The synthesis delays for IκBα and IκBϵ were determined by measuring IκBα and IκBϵ mRNA levels in a time course of TNF-treated murine embryonic fibroblasts (MEFs) in multiple independent experiments (Figure 4.3*E*, S2*A,B*). The measured delay for IκBα was $25.8 \pm 5.4\mathrm{min}$, and $59.4 \pm 12.8\mathrm{min}$ for IκBϵ, which agrees well with the model prediction for optimal damping.

Since it is difficult to measure the promoter strength experimentally, we employed an implicit way of comparing experiment with the model. To relate the parameter value m to experimental measurements, we set $m = 0.3$ in the model and calculated

the ratio of peak values for IκBα and IκBϵ proteins R_m, which we found to be equal 3.9. Then we measured the ratios of basal (unstimulated) to peak protein levels for IκBα and IκBϵ in experiment via quantitative Western blots of whole cell lysates generated during a TNF time course. These were compared to recombinant protein standards to derive absolute molecule number per cell. Peak IκBα protein levels were measured to be 379,800 molecules per cell, and IκBϵ 71,300 molecules per cell, with both values being subject to an estimated 25% error (Figure 4.3*F*, S2*C,D*). These protein levels correspond to the experimental peak values ratio $R_e = 5.3$ which is close to the model prediction $R_m = 3.9$.

4.2.4 Duration encoding in a dual delayed negative feedback loop system

We next addressed why the NF-κB signaling module may have evolved to produce oscillatory behavior if the oscillation frequency is not a function of the stimulus and does not constitute a signaling code. We first simulated persistent stimulation of a variant NF-κB system without feedback (we assume that IκBα is constitutively produced, so $d_{0y} = 1$, $d_{1y} = 0$ in Eq. 4.1) and found that this system produces long term, non-oscillatory NF-κB activity (Figure 4.4*A Top, blue line*). As TNF is secreted in bursts and therefore perceived by surrounding cells as transient or pulse stimulation, we then performed stimulations of pulses 15, 30, and 45 min in duration. In the negative feedback-deficient NF-κB system, the pulses resulted in transient responses that were attenuated very slowly. Faster attenuation can be achieved by increasing the constitutive synthesis rate, a_y. Increasing a_y by two orders of magnitude results in pulse NF-κB responses to transient stimuli, but the responsiveness (in amplitude) is much reduced (Figure 4.12).

We then performed similar simulations in a single negative feedback loop NF-κB system and found that this network topology allows for a rapid shutdown of NF-κB activity for transient inputs (Figure 4.4*A Middle*). This suggests that the NF-κB net-

work may have evolved from a pathway without feedback to a pathway with a single negative feedback loop to allow for a more sensitive transient response. Although the negative feedback indeed allows for greater sensitivity, a secondary consequence is that pronounced oscillations arise when the input signal persists for a long time period (Figure 4.4*A Middle, blue line*). The addition of a second negative feedback loop with a different time delay can help to dampen these oscillations, while preserving the responsiveness of the signaling module to transient stimuli (Figure 4.4*A Bottom*).

By plotting the duration of the response (above a given threshold) we investigated what may be called "temporal dose response curves" of the single and dual feedback systems (Figure 4.4*B*). The dual feedback system has a response duration close to 60 min for short pulses (< 100 min), and a duration proportional to the input duration for longer pulses. The single feedback system has the same behavior as the dual feedback system for short inputs, but for longer inputs the single feedback system produces a quantized response with the same output duration for several different input durations. Our analysis indicates that a dual feedback system is able to produce temporally graded responses, whereas a single feedback system that oscillates does not. Given that the duration of the second phase of the NF-κB response to TNF is a critical determinant of gene expression programs[59], we suggest that the NF-κB system has evolved a dual feedback system that allows for NF-κB activity whose duration is more closely related to the duration of the cytokine stimulus.

This fine temporal control, achieved via dual negative feedback, may be critical for complex cytokine-mediated cell-to-cell interactions involved in the adaptive immune response present in vertebrates, but may not be necessary for innate patogen-induced immune responses. We hypothesized that, on an evolutionary timescale, the appearance of dual negative feedback loops that regulate NF-κB activity may coincide with the transition from an innate to an adaptive immune system. To address this hypothesis, we used BLASTP with an E-value cutoff of 1e-25 to search for homologs of the mouse IκBα and IκBϵ protein sequences in other organisms (see Methods). We found homologs for both IκBα and IκBϵ, not only in other mammals (such as chimp, dog, platypus), but

also in other vertebrate classes including fish, amphibians, and birds (Figure 4.5). Thus, dual negative feedback regulation of NF-κB activity appears to be present in all organisms with adaptive immunity. In contrast, we did not find any invertebrate organisms with homologs for both IκBα and IκBϵ (Figure 4.5). Therefore, invertebrates, which lack adaptive immunity, also appear to lack the potential for dual negative feedback regulation of NF-κB mediated by IκBα and IκBϵ suggesting that the temporal control achieved with this regulatory architecture is not necessary for innate immune responses.

4.2.5 Robustness to fluctuations in a dual delayed negative feedback loop system

Thus far, we have examined the response of the network to transient stimulation in the absence of fluctuations. However, it is well known that noise in gene expression can cause significant variability in cellular responses[135;131;39;119;77;144]. Sometimes this variability can be beneficial[56], but in most cases, noise has a detrimental effect on the robustness of cellular functions. Mechanisms have presumably evolved to mitigate the unwanted effects of noise, especially in signaling pathways. In this section we examine the variability in the response of the NF-κB module that arises due to intrinsic and extrinsic noise, and we demonstrate that the dual-feedback loop architecture allows for a more robust response than the single feedback loop system. Further, we investigate how the relative contribution of intrinsic and extrinsic fluctuations depends on the size of the system.

The concentration of signaling molecules such as NF-κB can vary significantly between cells[34]. This variability in protein levels represents a source of extrinsic noise. We examined the variability in the response of the network to fluctuations in the total level of NF-κB and fluctuations in the IKK input level by simulating the network behavior with total NF-κB levels and active IKK levels distributed within a certain rage around their nominal values. The coefficient of variation (CV) in peak nuclear NF-κB levels and the CV in late-phase nuclear NF-κB levels is defined as $CV =$

$2(x_{max} - x_{min})/(x_{max} + x_{min})$ where x_{max} (x_{min}) are the maximum (minimum) values of NF-κB at the peak or during the late phase. NF-κB late-phase response is defined as the nuclear NF-κB level following the trough after the first peak response. In Text S1 we compare the extrinsic CV in the peak and the late phase for various values of IKK and NF-κB (see Figure 4.13).

Intrinsic noise arises from the stochastic nature of biochemical processes such as transcription and translation[39]. To examine the response of the NF-κB signaling module in the presence of intrinsic genetic noise, we used the Gillespie algorithm[47] modified according to[15] to perform stochastic simulations of both regular and delayed biochemical reactions included in our delayed feedback model. These latter reactions are initiated at times dictated by their respective rates, but the numbers of molecules are only updated after the time delay since the reaction initiation.

We ran stochastic simulations of both a single and dual feedback system and estimated the ensemble average $\langle X \rangle$ of the number of NF-κB molecules X and the magnitude of fluctuations as characterized by the standard deviation $\Delta X = [\langle X - \langle X \rangle \rangle^2]^{1/2}$ and the coefficient of variation $CV = \Delta X / \langle X \rangle$. To determine how the variability in the response varies with the magnitude of the input and the size of the system, we determined the CV in peak nuclear NF-κB levels and the CV in late-phase nuclear NF-κB levels for several values of IKK (Figure 4.6*A,C*) and for systems with up to 100,000 NF-κB molecules (Figure 4.6*B,D*). In Figure 4.6, we also plot CV values for extrinsic variations ($\pm 20\%$) in total NF-κB at several values of IKK (Figure 4.6*A,C*) and CV values for extrinsic variations in IKK ($\pm 20\%$) for several different system sizes (Figure 4.6*B,D*). We find that, even with this relatively low level ($\pm 20\%$) of extrinsic variability in IKK and NF-κB protein levels[34], variability in the response of the network is dominated by extrinsic noise for large systems ($> 10,000$ NF-κB molecules).

The CV in late-phase nuclear NF-κB levels is similar for extrinsic and intrinsic noise when the size of the system is reduced to 1000 NF-κB molecules. Next, we investigated the behavior of the NF-κB signaling module in this regime where intrinsic noise levels become significant by analyzing stochastic simulations produced with a system

with total NF-κB levels set to 1000 molecules. We ran stochastic simulations of all three systems studied deterministically above: no-feedback, single negative feedback, and dual negative feedback (Figure 4.7). Note that ensemble-averaged time series agree with the deterministic simulations very well (Figure 4.14). In the case of no feedback (Figure 4.7*A*) there is a strong robust response to the incoming persistent signal as characterized by the low values of the coefficient of variation. However, as we have seen above in Figure 4.4*A*, the major flaw of this system is its slow response to the pulse-like signals. Next, we simulated the 9 biochemical reactions included in the IκBα-mediated single negative feedback loop (Figure 4.7*B*). In single runs the first peak in nuclear NF-κB levels appears to be very robust, as illustrated by Figure 4.7*B Top*. The CV is lowest (< 0.2) during the first peak in nuclear NF-κB indicating that this portion of the response is very robust. Subsequent peaks in this undamped system lead to higher CV (> 0.5) in the later portion of the response.

Next, we performed stochastic simulations of the 18 biochemical reactions included in the dual delayed feedback model (with both IκBα- and IκBϵ-mediated feedback) (Figure 4.7*C*). In the dual feedback model, as in the single IκBα-mediated feedback model, there is a very robust first peak. However, unlike the single IκBα-mediated feedback model, in the dual feedback system the noise levels remain at a low level (< 0.5) following the first peak in nuclear NF-κB (Figure 4.7*C Bottom*). Thus, the dual feedback architecture allows for lower noise levels also in the later portion of the response.

What is the underlying reason for the robustness of the initial response from this circuit? The main source of intrinsic noise lies in the transcription and translation of IκB isoforms, since they are transcribed from single genes. In contrast, fluctuations in protein degradation and transport processes are relatively small, because the copy numbers of the corresponding molecules are large. In the NF-κB network, the peak in nuclear NF-κB levels that occurs following stimulation is produced via the degradation of IκB proteins that bind and sequester NF-κB in the cytoplasm. Thus, we argue that robustness of the initial response of the NF-κB circuit is explained by the fact that it

uses the sequestering mechanism and does not rely on the protein production.

To test this hypothesis, we simulated the behavior of an alternative network that relies on transcription of auto-repressor, rather than the degradation of inhibitor proteins, for signaling (Figure 4.7*D*). This system can be modeled with two variables: x, the number of repressor molecules, and d, the binary state of the promoter ($d = 0$ corresponds to the unbound promoter and $d = 1$ corresponds to bound promoter), and with four reactions (binding and unbinding of the repressor to the promoter, degradation of the repressor, and delayed synthesis of the repressor with rate $K(t)(1-d)$ where $K(t)$ is the external signal (Tables S2, S3). The input signal activates the production of the auto-repressor which after a certain time delay binds to the promoter and terminates further synthesis. Deterministically, this circuit also provides a desired response to a persistent stimulation with a large well-defined first peak. However, stochastic simulations reveal significant differences in the noise performance of this design as compared with the NF-κB circuit (note that the agreement between deterministic and stochastic simulations is less accurate in this case because of the strong promoter fluctuations (Figure 4.14*D*). Activation of the auto-repressor network is much less robust than the activation of the NF-κB network (cf. Figure 4.7*D* and Figures 4.7*B,C*). In fact, in the auto-repressor network, the coefficient of variation is highest (> 1) during the initial peak (Figure 4.7*D Bottom*). These results confirm our conjecture that the sequestering mechanism incorporated in the design of the NF-κB network gives rise to a much more robust activation of NF-κB than alternative networks that rely on transcription for activation and signaling. This finding is in accord with recent work[33] where the sequestering of Cdc20 protein was also implicated in the noise resistance of the spindle assembly checkpoint.

As we mentioned previously, recent computational work has suggested that persistent oscillations are present in wild-type cells with both IκBα- and IκBϵ-mediated feedback but stochastic variability leads to desynchronization among individual cells and therefore produces damped oscillations at the population level[6;104]. Our computational results demonstrate that, although stochastic oscillations are still present in individual cells with both IκBα- and IκBϵ-mediated feedback (Figure 4.7*C*), the oscillatory

propensity can be greatly reduced by the second feedback loop in the wild-type NF-κB signaling module. Further, stochastic simulations of the dual-feedback network reveal highly synchronized damped oscillations (Figure 4.15*C*) with cellular variations due to intrinsic noise becoming significant only when the system size is drastically reduced (Figure 4.7*C*).

To show that our results are not limited to the conceptual NF-κB model introduced above, we simulated the more detailed stochastic NF-κB model formulated in [104], which explicitly incorporates IKKK/IKK signaling cascade and NF-κB shuttling between the nucleus and the cytoplasm (see Methods and Figure 4.16*A*). One of the key assumptions made in the model [104] is that the strong stochasticity of the NF-κB response is caused by the slow and stochastic binding/dissociation of NF-κB to the corresponding promoters of IκBα, IκBϵ, and A20 target genes. The slow rates chosen by the authors for these reactions lead to the high variability of oscillatory dynamics among cells (Figure 4.16*B*). However, there is experimental evidence that the binding time of NF-κB may be significantly shorter, at least in certain types of cells. According to Fluorescence Recovery After Photobleaching (FRAP) measurements in HeLa cells [14], the typical time scale of NF-κB binding to the target promoters is on the order of a second rather than minutes, suggesting more rapid equilibration between the NF-κB-bound promoters and the pool of unbound nuclear NF-κB molecules. We found that increasing the binding and dissociation rates by $10^2 ... 10^3$ times profoundly changes the dynamics of the signaling system. NF-κB trajectories become more regular, suggesting that the behavior of individual cells translates more directly into the behavior of the population (Figure 4.16*C*). After adjusting the binding/dissociation rates along with a few other parameters, the updated model recapitulated the population response to chronic TNFα stimulation under various genetic conditions (WT, IκB$\epsilon^{-/-}$, and A20$^{-/-}$) (Figure 4.17) in agreement with earlier experimental results [59;67;149].

To quantify the magnitude of the late oscillatory NF-κB response to a chronic TNFα stimulation, we chose as a metric the average maximum peak-trough difference 5 hrs after initial stimulation. This quantity can be computed in two different ways.

The mean single-cell variability can be characterized by the magnitude M_s found by computing the maximum peak-trough differences for individual trajectories, and then averaging them over all trajectories:

$$M_s = \langle \max_{t>5h}\{N(t)\} - \min_{t>5h}\{N(t)\} \rangle \tag{4.8}$$

The population-level variability can be characterized by the magnitude M_p which is found by first computing an average trajectory and then computing its maximum peak-trough difference:

$$M_p = \max_{t>5h}\{\langle N(t) \rangle\} - \min_{t>5h}\{\langle N(t) \rangle\} \tag{4.9}$$

If the stochasticity is small, these two measures are similar, however for strong stochasticity they may differ significantly. Using these metrics, we first confirmed that for the parameter values adopted by [104], the model shows significant single-cell oscillations both in the IκB$\epsilon^{-/-}$ and in the WT, independently of the time delay in the IκBϵ loop (M_s, Figure 4.8*A*), but the population-averaged response shows significant oscillation dampening for the time delay around 45 min (M_p, Figure 4.8*C*). However, for our re-parameterized model with fast binding/dissociation, the stochasticity of individual trajectories is small, and both metrics show similar trend: the amplitude of oscillations in the WT is strongly suppressed at the optimal time delay of 45 min both for the population average (Figure 4.8*B*) and the individual cells (Figure 4.8*D*), which falls within the margin of error of our experimental results (Figure 4.3*D*).

4.3 Discussion

In this work we have developed a minimal model of the NF-κB signaling pathway that uses a small number of reactions (some of them compound) thus making it amenable to mathematical analysis. Previously, another simplified model of NF-κB signaling was developed in which a massive overshoot in IκBα resulted in an effective slowing of signaling dynamics [73], and produced spiky oscillations that are not seen in physiological conditions. Our model, which utilizes an explicit time delay, recapitulates

experimentally observed signaling behavior. It demonstrates that models with explicit time delays can be useful for investigating the mechanistic basis of the dynamic behavior of signaling pathways.

Using this model, we explored the potential role of NF-κB oscillations which are observed in a variant of the NF-κB signaling module with the secondary negative feedback loop involving IκBϵ, disabled. In particular, we addressed the question of whether the frequency of these oscillations contains information, as in neurons which sometimes encode information in the frequency of action potentials[12] and in the activation of the transcription factor NF-AT which is responsive to the number of Ca^{2+} pulses[32]. By analyzing the oscillatory response of a system regulated solely by the IκBα-mediated negative feedback loop, we found that both the frequency and the decay rate of the oscillations produced by this system are highly dependent on the internal parameters of the circuit, but are not sensitive to changes in the input signal levels. This result suggests that the oscillatory frequency does not encode information about the stimulus. Hence, stimulus-specific gene expression is unlikely determined by stimulus-specific frequencies of NF-κB oscillations. If there is a temporal code for stimulus-specific gene expression it is unlikely to involve frequency modulation, but may involve amplitude modulation over time.

When a second feedback regulator, IκBϵ, is added to the model, the oscillations caused by a persistent stimulation are significantly dampened, in agreement with our earlier findings[67]. By performing an optimization procedure, we determined that the specific experimentally observed parameter values for the synthesis delay and peak protein abundance of both IκB isoforms correspond to maximal efficiency of damping. These findings suggest that the second feedback (IκBϵ) has evolved to produce damping of the oscillatory behavior of the first feedback (IκBα). Furthermore, we demonstrated that this finding is not limited to our simple model, but can be expanded to more complex models. For example, in a recent model by[104] with fast binding/unbinding rates of NF-κB the secondary IκBϵ feedback leads to a reduction in NF-κB oscillations in individual cells. However, cell-cell variability and extrinsic noise can further reduce NF-κB

oscillations on a population level.

From the evolutionary perspective, we have a peculiar situation in which a signaling module apparently first developed a negative feedback loop that made it prone to oscillations, and then added a secondary loop which mitigated these oscillations. This brings the question, if oscillatory responses are not beneficial to the cell, why has the primary negative feedback appeared in the system in the first place? By comparing transient response of several variants of signaling modules (0-, 1- and 2-feedback loop designs) in the presence of stochastic fluctuations we showed that the primary negative feedback loop involving the release of sequestered NF-κB proteins created a strong, rapid, and robust response to short pulses of active IKK signal. However, for longer signals a single-feedback-loop system exhibits a suboptimal "temporal dose response behavior" that leads to a quantized response to signals of different durations. In contrast, the dual feedback network generates response durations that are proportional to the stimulus input durations. Fine-tuning of the response duration may be reflective of a signaling code in which duration of NF-κB activity may be a key determinant of stimulus-specific gene expression program.

Cytokines such as TNFα facilitate adaptive responses at the effector stages[90]. The evolution of cytokines is associated with the evolution of an adaptive immune system to allow for coordination of various cell types[98]. Unlike pathogen exposure, cytokines are produced during varying amounts of time thereby generating time-varying signals. Our analysis showed that the dual negative feedback module is more capable at distinguishing differences in the duration of incoming signals. This function is important for the transduction of cytokine signals, but not pathogen signals. Our BLASTP analysis indeed demonstrates that the evolution of the dual negative feedback system may correlate with the evolution of adaptive immunity.

4.4 Methods

4.4.1 Derivation of the deterministic model

Using mass action kinetics, the full set of reactions for the dual feedback loop NF-κB system (Tables 2, 4) can be expressed by the following ODEs:

$$\dot{x} = -f_0 d_{0y} x + f_1 d_{1y} - k_y yx + k_{-y}[xy] + Kr_y[xy] - f_0 d_{0z} x + f_1 d_{1z} - k_z zx + k_{-z}[xz] + Kr_z[xz] \tag{4.10}$$

$$\dot{[xy]} = k_y yx - k_{-y}[xy] - Kr_y[xy] \tag{4.11}$$

$$\dot{y} = a_y d_{0y}(t - \tau_y) + b_y d_{1y}(t - \tau_y) - k_y yx + k_{-y}[xy] - g_y y - Kr_y y \tag{4.12}$$

$$\dot{d_{0y}} = -f_0 d_{0y} x + f_1 d_{1y} \tag{4.13}$$

$$\dot{d_{1y}} = f_0 d_{0y} x - f_1 d_{1y} \tag{4.14}$$

$$\dot{[xz]} = k_z zx - k_{-z}[xz] - Kr_z[xz] \tag{4.15}$$

$$\dot{z} = a_z d_{0z}(t - \tau_z) + b_z d_{1z}(t - \tau_z) - k_z zx + k_{-z}[xz] - g_z z - Kr_z z \tag{4.16}$$

$$\dot{d_{0z}} = -f_0 d_{0z} x + f_1 d_{1z} \tag{4.17}$$

$$\dot{d_{1z}} = f_0 d_{0z} x - f_1 d_{1z} \tag{4.18}$$

The total number of κB binding sites on each promoter is conserved:

$$d_{0y} + d_{1y} = d_y \tag{4.19}$$

$$d_{0z} + d_{1z} = d_z \tag{4.20}$$

We assume that the total amount of NF-κB in the cell X is conserved

$$X = x + d_{1y} + d_{1z} + [xy] + [xz] = const. \tag{4.21}$$

Since the number of binding sites available for NF-κB protein is small, we can neglect the amount of NF-κB bound to the IκBα and IκBϵ promoters, so

$$X = x + [xy] + [xz] \tag{4.22}$$

Solving Eq. 4.22 for x yields:

$$x = X - [xy] + [xz] \tag{4.23}$$

DNA binding reactions are usually fast, so we can assume that they are at quasi-equilibrium at all times,

$$f_1 d_{1y} = f_0 d_{0y} x \tag{4.24}$$

$$f_1 d_{1z} = f_0 d_{0z} x \tag{4.25}$$

Using Eqs. 4.19 and 4.20, substituting into Eqs. 4.24 and 4.25, and solving for d_{0y}, d_{1y}, d_{0z}, d_{1z} yields:

$$d_{0y} = \frac{F d_y}{F + x} \tag{4.26}$$

$$d_{1y} = \frac{x d_y}{F + x} \tag{4.27}$$

$$d_{0z} = \frac{F d_z}{F + x} \tag{4.28}$$

$$d_{1z} = \frac{x d_z}{F + x} \tag{4.29}$$

where $F = f_1/f_0$.

We also assume quasi-equilibrium for IκBNF-κB binding reactions,

$$k_y y x = k_{-y}[xy] + K r_y [xy] \tag{4.30}$$

$$k_z z x = k_{-z}[xz] + K r_z [xz] \tag{4.31}$$

Substituting $[xy]$ and $[xz]$ from Eqs. 4.30 and 4.31 into Eq. 4.23 yields:

$$x = X - [k_y y x/(k_{-y} + K r_y)] - [k_z z x/(k_{-z} + K r_z)] \tag{4.32}$$

Now we can solve Eq. 4.32 for x

$$x = \frac{X}{1 + k_y y/(k_{-y} + K r_y) + k_z z/(k_{-z} + K r_z)} \tag{4.33}$$

and substitute it in Eqs. 4.12 and 4.16. These equations contain both fast and slow terms. However, it is easy to see that rate equations for variables $Y = y + [xy]$ and $Z = z + [xz]$ contain only slow terms:

$$\dot{Y} = a_y d_{0y}(t - \tau_y) + b_y d_{1y}(t - \tau_y) - g_y y - K r_y Y \tag{4.34}$$

$$\dot{Z} = a_z d_{0z}(t - \tau_z) + b_z d_{1z}(t - \tau_z) - g_z z - K r_z Z \tag{4.35}$$

Y and Z can in turn be expressed via y and z by:

$$Y = y[1 + c_y X/(1 + c_y y + c_z z)] \tag{4.36}$$

$$Z = z[1 + c_z X/(1 + c_y y + c_z z)] \tag{4.37}$$

where $c_y = k_y/(k_{-y} + K r_y)$ and $c_z = k_z/(k_{-z} + K r_z)$. Equations 4.34–4.35 combined with definitions Eqs. 4.26–4.29, 4.33, 4.36, and 4.37 represent a closed system of two delay-differential equations 2, 3 for the dual-feedback NF-κB module. Setting $Z = z = 0$ in these equations leaves us with a single delay-differential equation for the single feedback loop system Eq. 1.

4.4.2 Details of the linear stability analysis

The fixed point Y_s of Eq. (4.4) is given by the algebraic equation (4.6). Unfortunately, Eq. (4.6) does not permit finding Y_s in explicit form. However, this calculation can be significantly simplified if the total number of NF-κB proteins is large, so $c_y X \gg 1$, then y can be neglected as compared with total Y. Then $x = X - Y$, and $d_{1y} = d_y(X - Y)/(F + X - Y)$, and expression (4.5) for $G(Y)$ simplifies:

$$G(Y) = \frac{b_y d_y (X - Y_\tau)}{F + X - Y_\tau} \tag{4.38}$$

Now the stationary level of Y can be obtained explicitly

$$Y_s = \frac{b_y d_y + K r_y (F + X) - \sqrt{(b_y d_y + K r_y (F + X))^2 - 4 K r_y b_y d_y X}}{2 K r_y} \tag{4.39}$$

The stability of this stationary solution is determined by the linearized equation (4.4) for a small perturbation ξ near Y_s,

$$\dot{\xi} = B\xi_\tau - Kr_y\xi \tag{4.40}$$

where $\xi = Y - Y_s$, subscript τ again indicates the delayed value of ξ taken at time $t-\tau$, and $B = dG(Y_s)/dY$. Using formula (4.38) we obtain

$$B = \frac{b_y d_y F}{(F + X - Y_s)^2} \tag{4.41}$$

where Y_s is given by Eq. (4.39). The eigenvalue λ of the linearized equation (4.40) is found by substituting $\xi = \xi_0 \exp(\lambda t)$, yielding the transcendental equation

$$\lambda = Be^{-\lambda\tau_y} - Kr_y \tag{4.42}$$

whose solution is given by Eq. (4.7).

4.4.3 Stochastic model formulation

For the analysis of a full NF-κB system, we adopted the basic structure of the NF-κB model formulated in[104] which in turn was based on the population-level model first proposed in[59]. The structure of the model is shown in Figure 4.16*A*. In resting cells, NF-κB is sequestered in the cytoplasm by IκB proteins. In response to TNFα stimulation, IKKK protein becomes active, and activates IKK kinase. IKK phosphorylates IκB proteins targeting them for degradation. Upon degradation of IκB proteins, NF-κB moves into the nucleus and activates hundreds of target genes. In the model, we focus on the dynamics of three genes associated with the negative feedback of the system. Following NF-κB activation, synthesized A20 proteins attenuate TNFα signal by repressing IKKK and IKK transitions into their active states. NF-κB also binds IκBα and IκBϵ protein promoters, which following translation in the cytoplasm, translocate back into the nucleus and bind free NF-κB sequestering it out of the nucleus. In addition, IκB proteins are directly responsible for NF-κB dissociation from the DNA.

The biological processes in the model were interpreted through stochastic and deterministic representations similar to[104]. Nuclear transport, complex formation, synthesis, transcription, and translation were described through a set of ordinary differential equations. Regulation of gene activity through NF-κB binding and dissociation from DNA was modeled using stochastic representation. The time-evolution of the system was accomplished through a hybrid simulation algorithm that uses Gillespie algorithm[47] to evaluate the state of stochastic processes and an ODE solver to compute the state of deterministic processes.

4.4.4 Details of the BLASTP search for IκBα and IκBϵ homologs

We performed two BLASTP searches (using default parameters) to search for IκBα and IκBϵ homologs. The mouse IκBα protein sequence (gi28386026) was used as the query for the first search. The mouse IκBϵ protein sequence (gi2739158) was used as the query for the second search. We used an E-value of 1e-25 as a cutoff for both searches.

Note that we selected only unique homologs for both IκBα and IκBϵ in all vertebrates. We did not find unique IκBϵ homologue for several vertebrates. We expect that this is due to the fact that complete genomes are not currently available for these organisms.

4.4.5 Cell culture experiments

Immortalized murine embryonic fibroblasts[59] were chronically stimulated with 10ng/mL TNF (Roche) and IκBα and IκBϵ mRNA and protein levels were monitored by RNase Protection Assay (RPA) and Western Blot, respectively, as previously described[67]. RPA results for each time course were quantitated using ImageQuant software (GE Healthcare) and used to determine the time of half-maximal inducibility between basal and peak mRNA levels for IκBα and IκBϵ (Figure 4.10 *A,B*). Western Blot results were also quantitated with ImageQuant software and used to determine the time point

of peak expression. The basal abundances of IκBα and IκBϵ protein were determined via comparison to a standard curve of recombinant IκB protein (R Tsu, JD Kearns, C Lynch, D Vu, K Ngo, S Basak, G Ghosh, A Hoffmann *in preparation*). The peak abundances of IκBα and IκBϵ were determined via multiplication of the basal value by the fold inducibility at the peak time point (Figure 4.10 *C,D*). Experimental levels of nuclear NF-κB in cells with only the IκBα-mediated negative feedback loop intact and in wild-type cells containing both IκBα- and IκBϵ-mediated negative feedback were determined by EMSAs in[59].

4.4.6 Extrinsic noise in dual negative feedback loop system

We find that the CV in peak nuclear NF-κB increases linearly with extrinsic variation in total NF-κB and with extrinsic variation in IKK with identical CV values for both the single and dual feedback models (Figure 4.13*A,B*). In contrast, the CV in late-phase (asymptotic) NF-κB levels are significantly lower in the dual feedback system than in the single feedback system. We varied the magnitude of extrinsic noise by changing the spread of parameters (total NF-κB and IKK) from 0% to 50%. The CV in late-phase NF-κB for the dual feedback system increases linearly from 0 to approximately 1 as the range of total NF-κB (Figure 4.13*C*) and IKK (Figure 4.13*D*) is increased to $\pm 50\%$, while the CV in late-phase NF-κB for the single feedback system increases from approximately 1.6 to 1.9 (Figure 4.13*C,D*). Thus, in the presence of extrinsic variations in IKK and total NF-κB, the dual feedback system allows for a late-phase response which is more robust than the response produced by the single feedback system.

4.4.7 Details of the full stochastic model

For the analysis of a full NF-κB system, we adopted the basic structure of the NF-κB model formulated in Paszek et al.[104]. The structure of the model is shown in Figure 4.16*A*. The biological processes in the model were interpreted through stochastic and deterministic representations. Nuclear transport, complex formation, synthe-

sis, transcription, and translation were described through a set of ordinary differential equations (ODEs). Regulation of gene activity through NF-κB binding and dissociation from DNA was modeled using stochastic representation. The time-evolution of the system was accomplished through a hybrid simulation algorithm that uses Gillespie algorithm[48] to evaluate the state of stochastic processes and an ODE solver to compute the state of deterministic processes. Full details of the model can be found in Longo et al.[79].

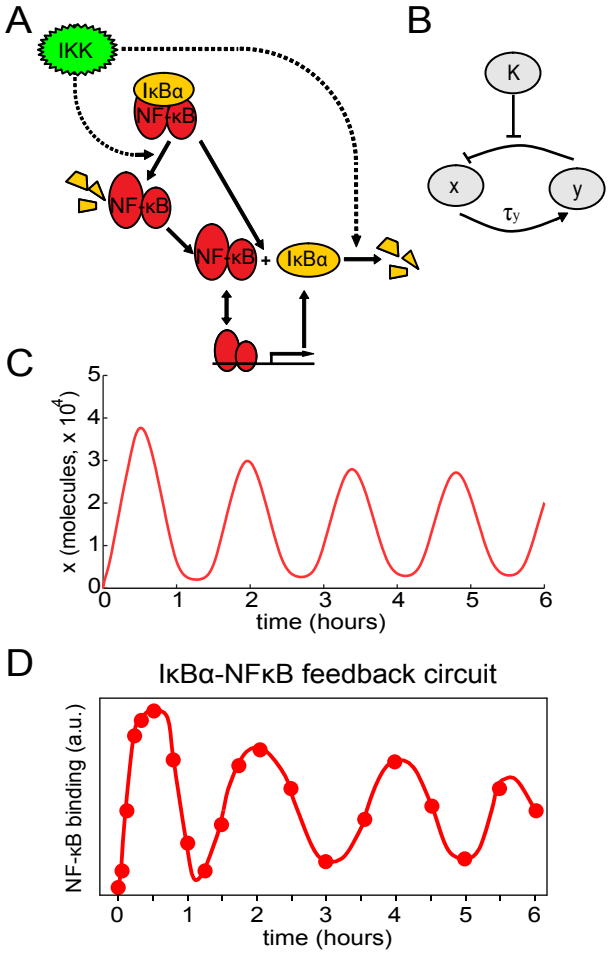

Figure 4.1: Oscillatory behavior from a system with a single negative feedback loop. (A) Diagram of the IκBα-NF-κB signaling module. (B) Diagram of a system with a single delayed negative feedback loop. (C) Nuclear NF-κB levels (x) in response to persistent stimulation as a function of time produced using our delayed feedback model. (D) Experimental levels of nuclear NF-κB (determined by EMSAs) in cells with only the IκBα-mediated negative feedback loop intact (data from[59]).

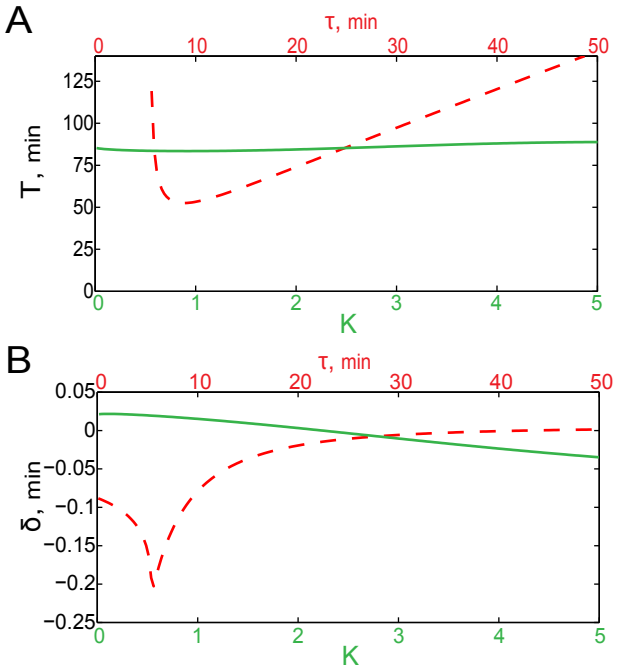

Figure 4.2: Period and decay rate of oscillations produced by the IκBα-mediated negative feedback system. (A) The oscillation period T as a function of K with $\tau_y = 25$ min (green line) and as a function of τ_y with $K = 2$ (red dashed line). (B) The oscillation decay rate δ as a function of K with $\tau_y = 25$ min (green line) and as a function of τ_y with $K = 2$ (red dashed line).

4.5 Acknowledgments

Chapter 4 contains material originally published as Longo, D. M.*, Selimkhanov, J.*, Kearns, J. D., Hasty, J., Hoffmann, A., and Tsimring, L. S. 2013: Dual delayed feedback provides sensitivity and robustness to the NF-κB signaling module. *PLoS Computational Biology*, 9(6), e1003112. (*equal contribution). Permission to republish here is not required by PLOS.

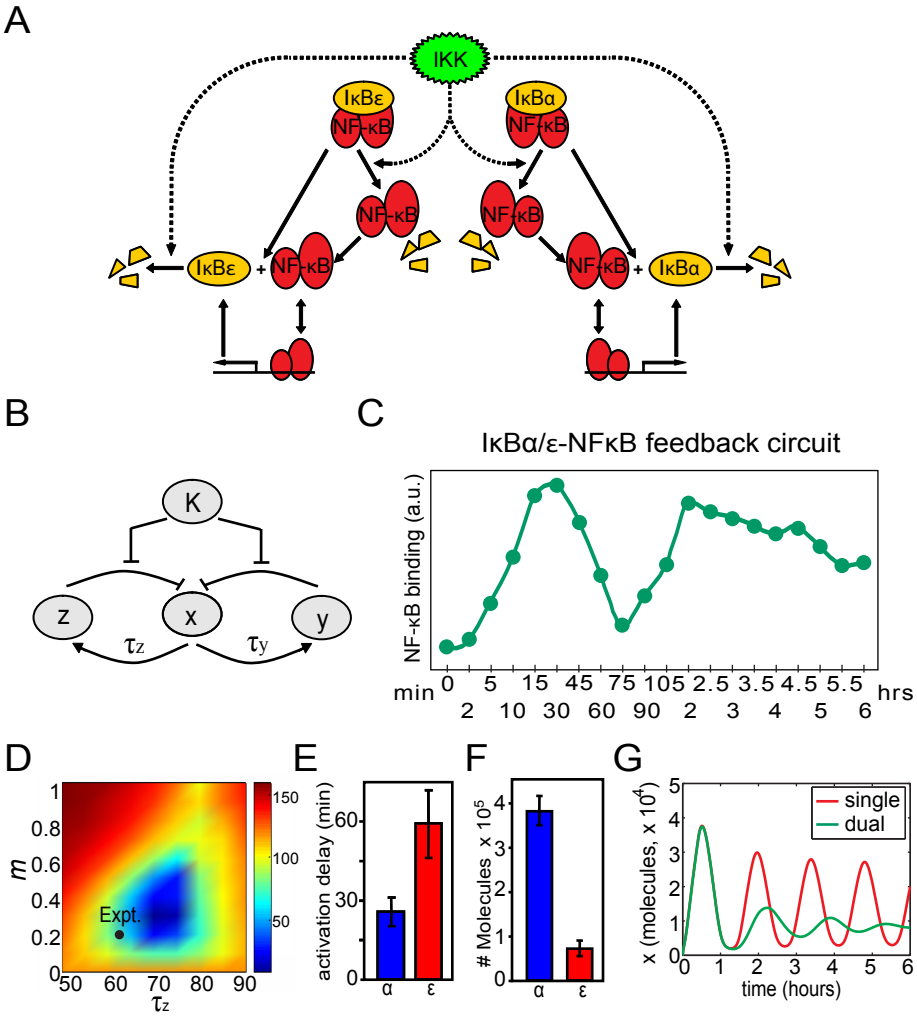

Figure 4.3: Damped oscillations with a dual negative feedback system. (A) Diagram of the dual feedback system with both IκBα- and IκBϵ-mediated negative feedback loops. (B) Diagram of a system with dual delayed negative feedback loops. (C) Experimental levels of nuclear NF-κB (determined by EMSAs) in wild-type cells containing both IκBα-and IκBϵ-mediated negative feedback (data from[59]). (D) Optimization of the parameters of the second feedback loop m and τ_z towards maximizing the oscillations damping. The optimization method minimizes peak-minus-trough differences six hours after the onset of stimulation, the global minimum occurs at $\tau_z = 72$ min, $m = .3$. The black dot indicates the experimentally measured parameter values ($\tau_z = 59$ min, $m = .2$). Note that m was not measured directly. The value of m corresponding to the experimentally measured value of R_e was determined with the model (Figure 4.11). (E) Experimental measurements of IκBα and IκBϵ synthesis delays. (F) Experimental values for peak IκBα and IκBϵ protein levels. (G) Simulated time course of nuclear NF-κB levels (x) for the single feedback system and for the optimized dual feedback system in response to persistent stimulation with $K = 2$.

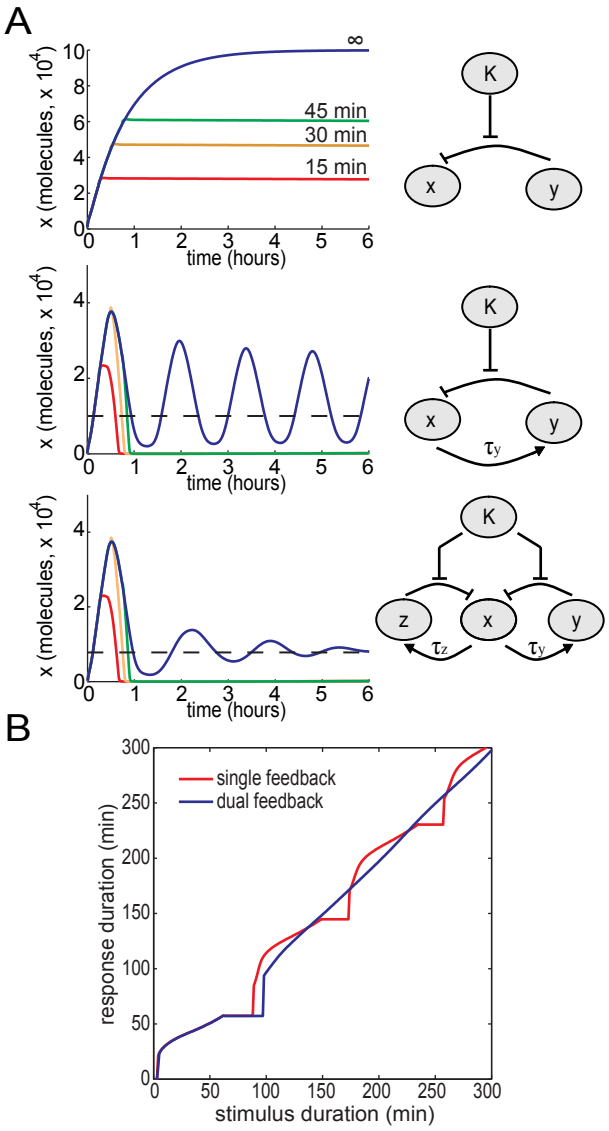

Figure 4.4: Response of the NF-κB signaling module to transient inputs with magnitude $\mathbf{K = 2}$. (A) Time series of x for a system with all feedback removed (top), a system with IκBα-mediated negative feedback (middle), and a system with both IκBα- and IκBϵ-mediated negative feedback (bottom) in response to 15 min (red), 30 min (orange), 45 min (green), and persistent (blue) stimulation. (B) The response duration as a function of the stimulus duration for the single feedback and dual feedback systems. The response duration is the amount of time x exceeds a threshold level of 50 (as indicated by the dashed black lines in the graphs shown in (A).

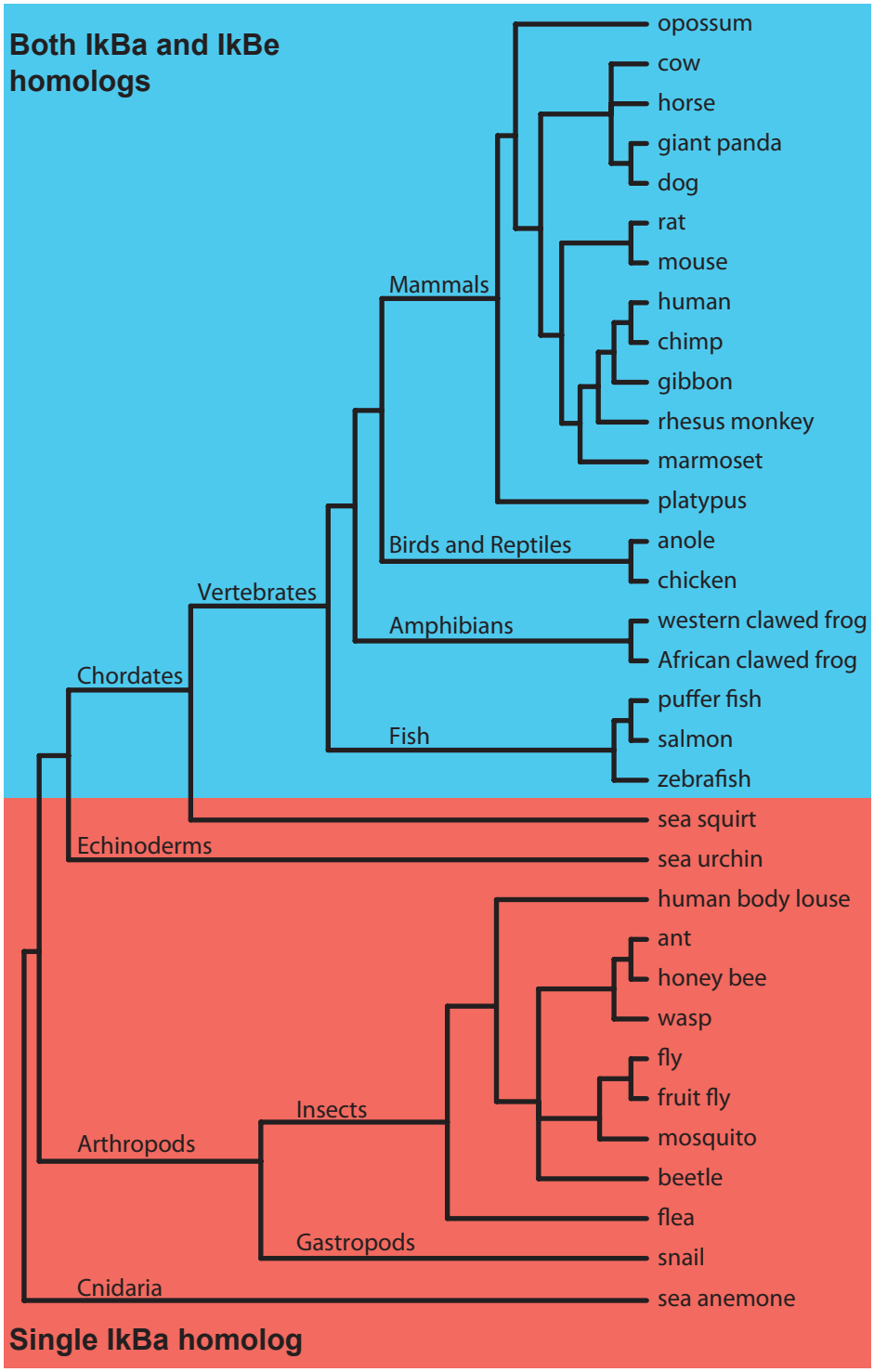

Figure 4.5: Schematic of a phylogenetic tree showing organisms in which IκB homologs were found using BLASTP. Organisms with homologs for both IκBα and IκBϵ are in blue shaded region and organisms with a single homolog are in red shaded region. The branches in the schematic phylogenetic tree are not drawn to scale. (For simplicity, not all organisms with single or dual homologs are shown here.

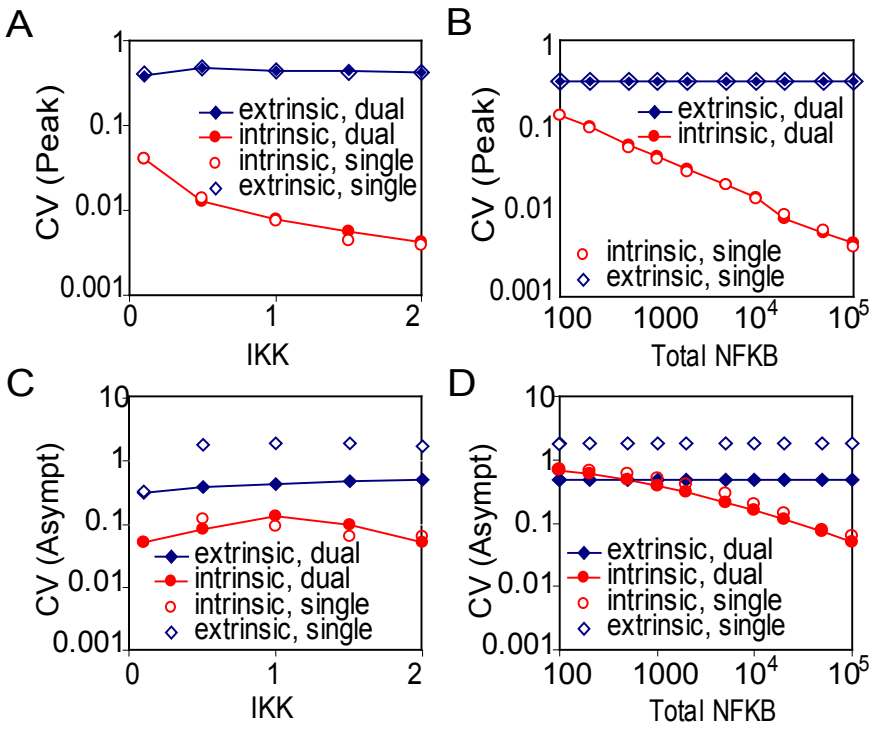

Figure 4.6: The coefficient of variation (CV) in nuclear NF-κB levels due to extrinsic and intrinsic fluctuations. The CV was calculated for peak (A,B) and late-phase (C,D) nuclear NF-κB levels for both single and dual feedback systems. The CV due to intrinsic fluctuations was determined from at least 50 runs of the stochastic simulations at each value of IKK (A,C) and total NF-κB (B,D). The CV due to extrinsic fluctuations in total NF-κB and IKK levels was determined by varying the total NF-κB level by $\pm 20\%$ for each value of IKK (A,C) and by varying IKK by $\pm 20\%$ for value of total NF-κB (B,D).

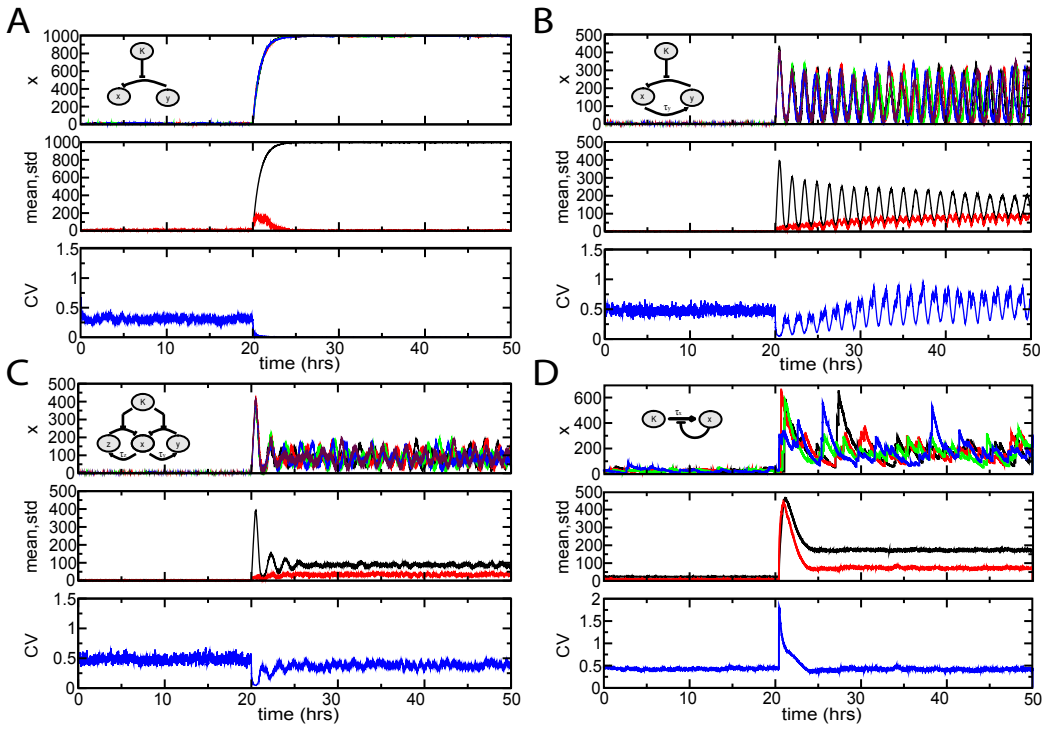

Figure 4.7: Stochastic model simulation results for various network architectures (with 1000 total NF-κB molecules). The architectures analyzed are the NF-κB network with no feedback loops (A), only IκBα-mediated negative feedback (B), the NF-κB network with both IκBα- and IκBϵ-mediated negative feedback (C), and an alternative auto-repressive network (D). The top panel in each group shows four typical runs of stochastic simulations for each network, the middle panel shows the mean and standard deviation for 200 runs of each network, and the bottom panel shows the corresponding coefficient of variation. The input signal, K(t), is switched from $K = 0$ to $K = K_{max}$ at $t = 20$ hrs. In A-C, the magnitude of external signal $K_{max} = 2$, in D, $K_{max} = 50$.

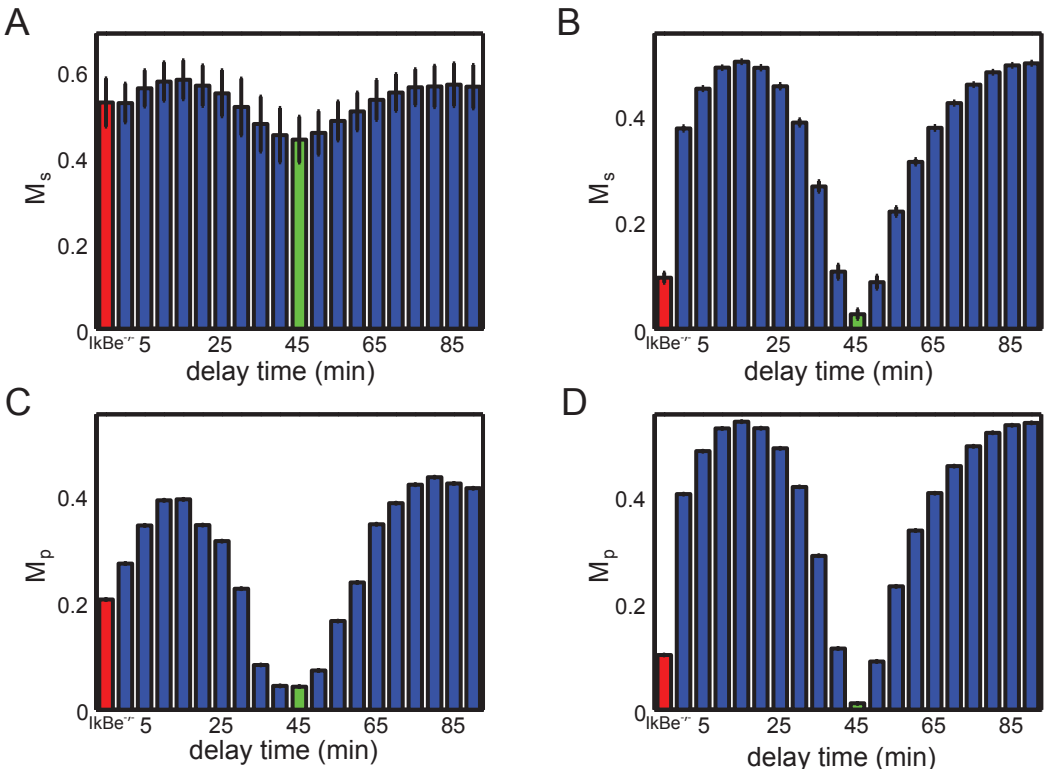

Figure 4.8: Effect of delay time on damped NF-κB oscillations in the detailed NF-κB model. Magnitudes of single-cell oscillations 5 hours after stimulation (A,B) and of population averaged oscillations (C,D) are shown for the IκB$\epsilon^{-/-}$ knockout (red bar) and in the WT for different time delays in the IκBϵ feedback loop (blue bars) with original (A,C) and adjusted (B,D) parameter values. The optimal time delay of 45 min is shown by the green bar. Each bar represents the average of nuclear NF-κB variation for 500 single cell trajectories. Error bars represent $\pm$ one standard deviation.

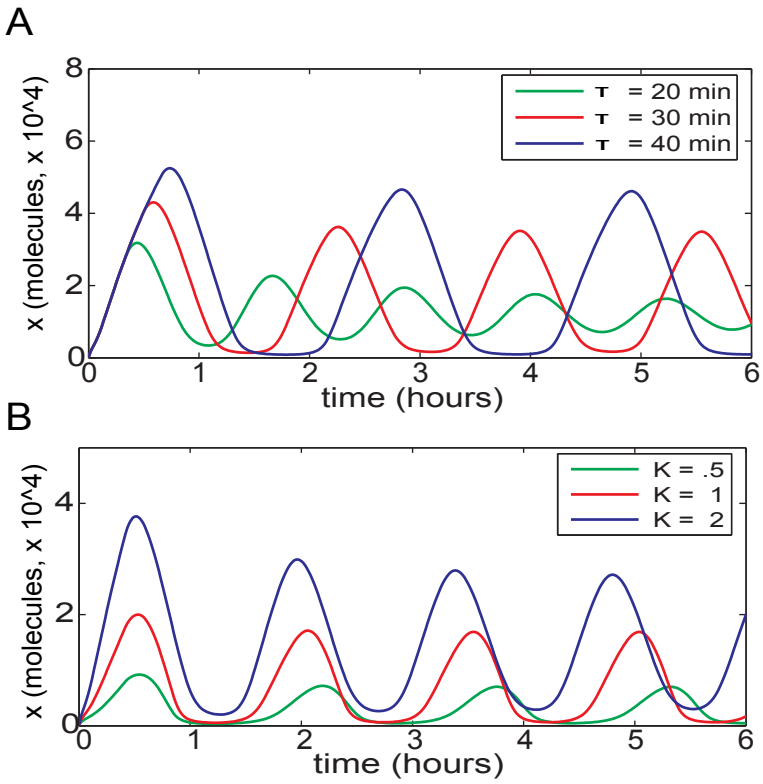

Figure 4.9: Oscillations produced by the IκBα-mediated negative feedback system. (A) $K = 2$ for $\tau_y = 20$ min, $\tau_y = 30$ min, and $\tau_y = 40$ min and with (B) $\tau_y = 25$ min for $K = .5$, $K = 1$, and $K = 2$.

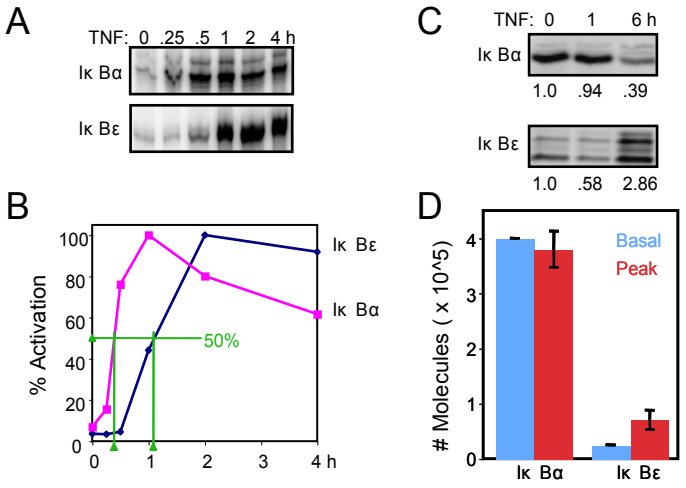

Figure 4.10: Representative experimental data for IκB$\alpha$and IκBϵsynthesis delays and feedback strengths. (A) mRNA synthesis for IκBα and IκBϵ were measured by RNase Protection Assay in wild-type immortalized MEF cells in response to 10ng/mL TNF chronic stimulation. (B) The RPA results were quantitated to determine the intensity of each band in the gel (ImageQuant, GE Healthcare). The highest intensity band in each set was set to 100% Activation and the other bands were normalized accordingly. The delay time to reach half maximal synthesis was calculated as the time at which the activation curve crossed the 50% level. A set of N=10 replicate experiments were performed to calculate the global average. (C) The protein abundances for IκBα and IκBϵ at their respective activation peaks in wild-type immortalized MEF cells chronically stimulated with 10ng/mL TNF were measured by Western Blot analysis (IκBα at 1 h and IκBϵ at 6 h) . Fold induction vs. basal state are shown below each gel and were calculated by quantitation of the band intensities and normalization to the 0 h band. (D) Bar plot of the average protein abundances from multiple Western Blot experiments for peak levels of IκBα (N=7) and IκBα (N=5). The basal state abundances were measured by comparison to a standard curve of recombinant IκBα or IκBϵ protein (JD Kearns, S Basak, C Lynch, A Hoffmann *in preparation*). Peak abundances were calculated by multiplying the quantitated fold induction (as in C) by the basal abundance. Error bars on the peak bars represent one standard deviation.

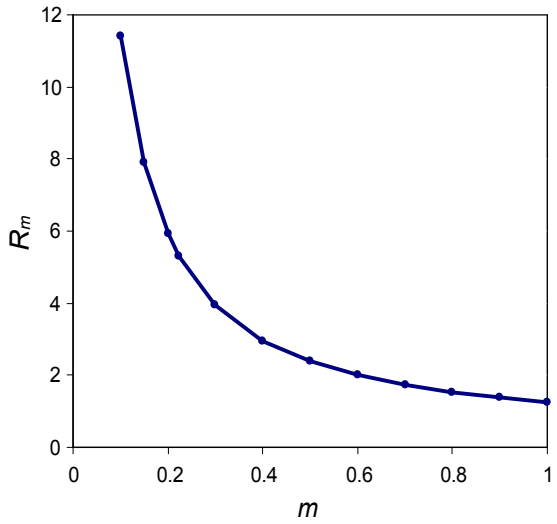

Figure 4.11: The ratio of peak IκBα protein levels to peak IκBϵ protein levels (R_m) versus m. R_m was determined for several values of m in the model simulations to determine the value of m corresponding to $R_m = 5.3$ (The experimentally measured value for the ratio is $R_e = 5.3$). This value of m ($m = 0.2$) was used to plot the point in Figure 3D which indicates the experimental values of m and τ_z.

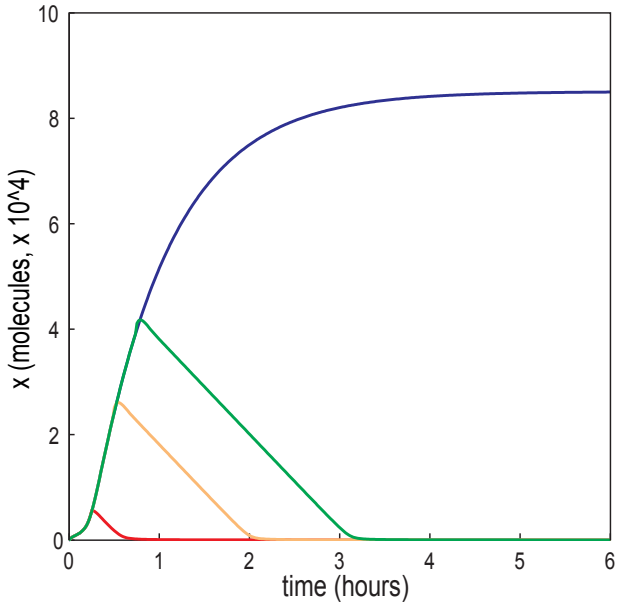

Figure 4.12: Response of a no-feedback system with constitutive IκBα synthesis increased from a_y = 0.00185 nM/min to a_y = 0.3 nM/min. The time series of x is shown for 15 min (red), 30 min (orange), 45 min (green), and persistent (blue) stimulation.

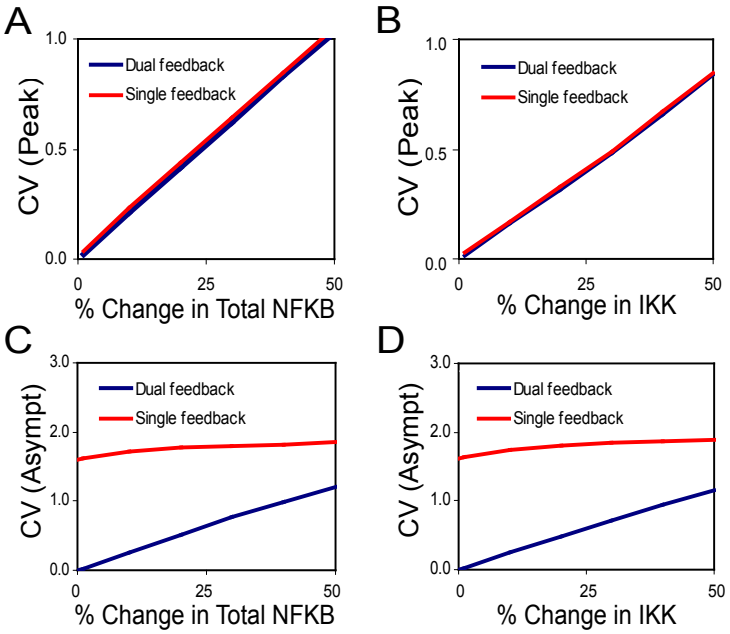

Figure 4.13: Coefficient of variation (CV) of nuclear NF-κB levels. Comparison of the CV for peak and late-phase nuclear NF-κB levels due to extrinsic variability in total NF-κB (A,C) and IKK (B,D) for the single feedback and dual feedback network.

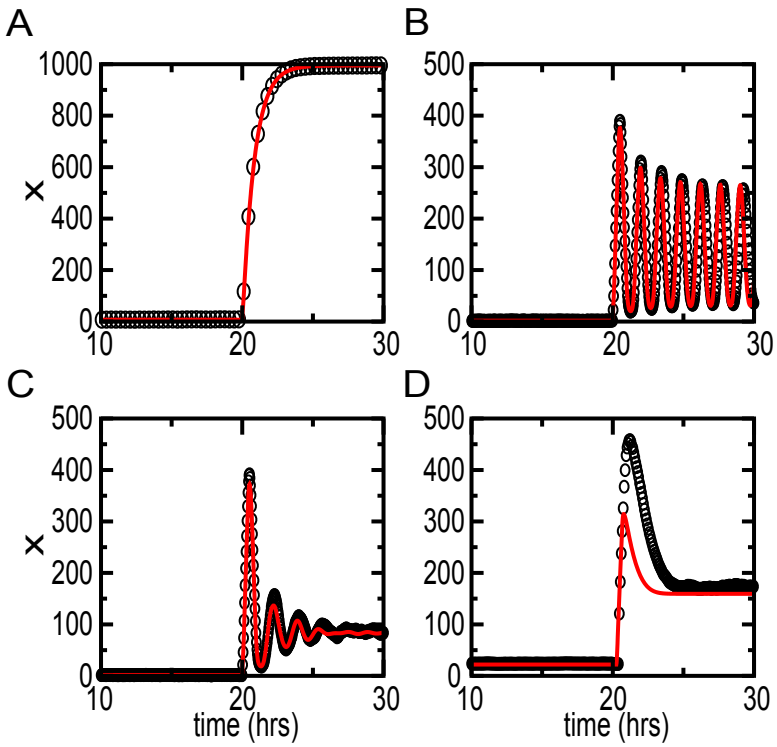

Figure 4.14: Comparison of ensemble-averaged runs of stochastic simulations with deterministic simulations for four different circuits. (A) no-feedback model, (B) single negative feedback model, (C) dual feedback system, (D) and auto-repressor system. Lines - deterministic simulations, symbols - stochastic simulations averaged over 200 runs.

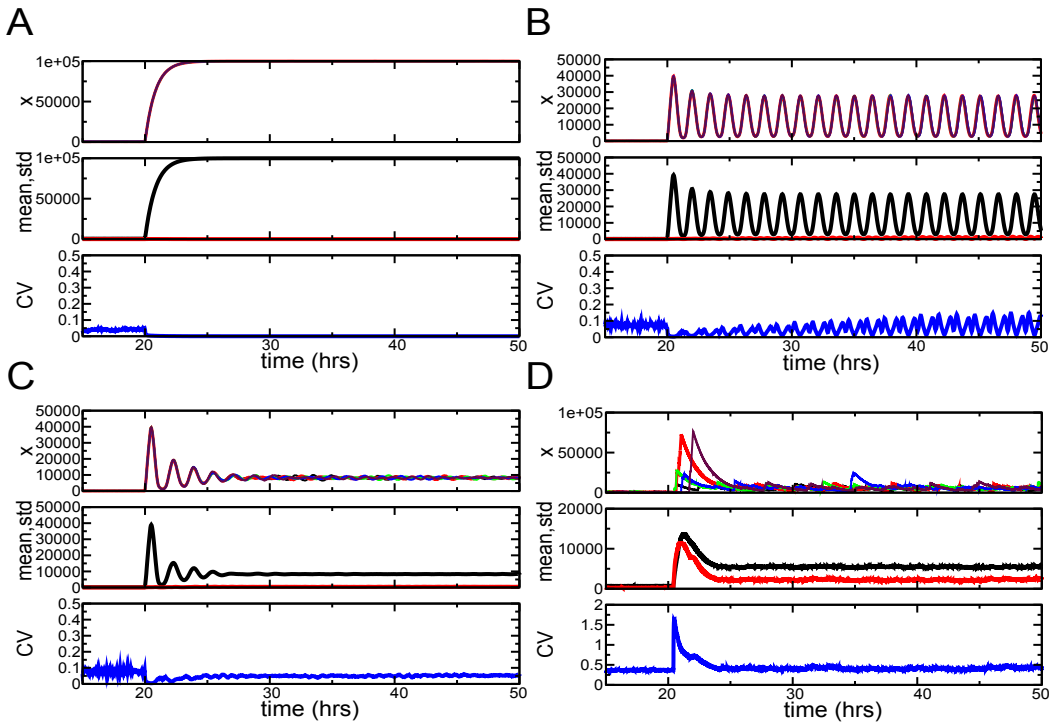

Figure 4.15: Stochastic simulation results with 100,000 total NF-κB molecules. Four NF-κB networks were considered: no feedback loops (A), only IκBα-mediated negative feedback (B), the NF-κB network with both IκBα- and IκBϵ-mediated negative feedback (C), and an alternative auto-repressive network (D). The top panel in each group shows four typical runs of stochastic simulations for each network, the middle panel shows the mean and standard deviation for 200 runs of each network, and the bottom panel shows the corresponding coefficient of variation. The input signal, K(t), is switched from $K = 0$ to $K = K_{max}$ at $t = 20$ hrs. In A-C, the magnitude of external signal $K_{max} = 2$, in D, $K_{max} = 50$.

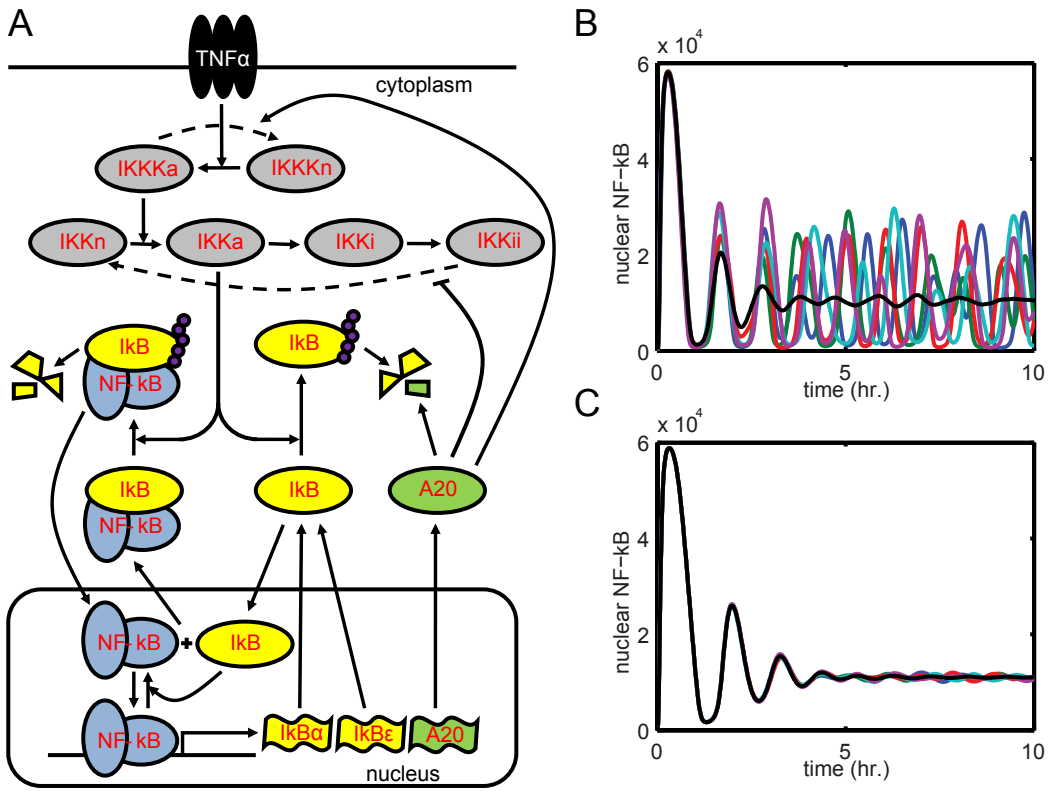

Figure 4.16: Oscillatory behavior from NF-κB signaling system. (A) Diagram of the NF-κB signaling network model adopted from Paszek *et. al* (2010). In resting cells, NF-κB is sequestered in the cytoplasm by IκB proteins. In response to TNFα stimulation, IKKK protein becomes active, activating IKK kinase. In turn, IKK phosphorilates IκB proteins targeting them for degradation. Upon degradation of IκBα, NF-κB moves into the nucleus and activates hundreds of target genes. In the model, we describe the dynamics of three genes associated with the negative feedback of the system. Following NF-κB activation, synthesized A20 proteins attenuate TNFα signal by repressing IKKK and IKK transitions into their active states. NF-κB also binds IκBα and IκBϵ protein promoters, which following translation in the cytoplasm, translocate back into the nucleus and bind up free NF-κB sequestering it out of the nucleus. In addition, IκB proteins are directly responsible for NF-κB dissociation from the DNA. (B) Nuclear NF-κB levels in response to persistent stimulation as a function of time computed using Paszek *et. al* (2010) wildtype model. (C) Nuclear NF-κB levels in response to persistent stimulation as a function of time computed using our re-parameterized model. Black curve represents an average of 500 cell trajectories.

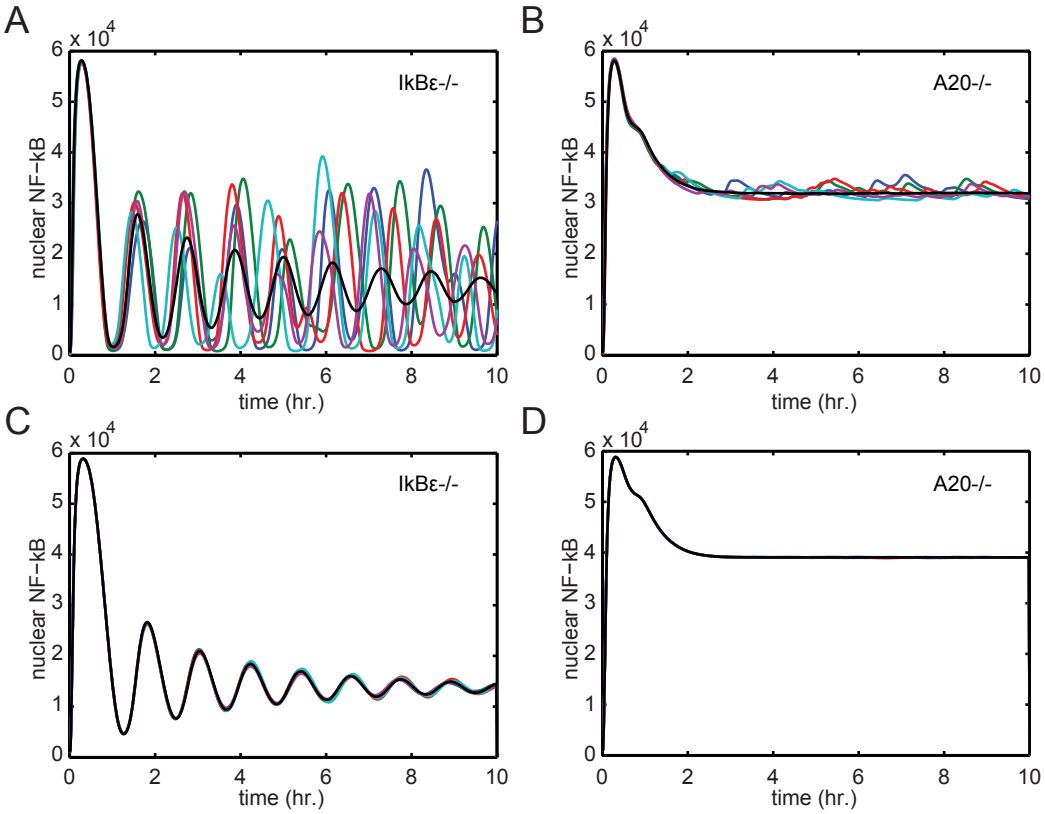

Figure 4.17: Nuclear NF-κB response in A20 and IκBϵ knockout models to chronic TNFα stimulation. The results for the updated model (C,D) shows close similarity, in NF-κB population dynamics, to the results obtained using Paszek *et. al* (2010) wildtype model (A,B) for both IκBϵ knockout (A,C) and A20 knockout (B,D) models. Black trajectories represent the average of 500 cell trajectories.

Chapter 5

NF-κB response to in-vivo mode of TNFα induction

5.1 Introduction

Nuclear factor kappa B (NF-κB) is a well-studied global regulator of gene expression that coordinates the cellular response to a variety of external stimuli such as tumor necrosis factor alpha (TNFα), which is critical in inflammation and immunity. NF-κB is normally sequestered in the cytoplasm by IκB proteins, but it translocates into the nucleus upon TNFα stimulation and acts to regulate a variety of downstream genes before it is shuttled out of the nucleus through IκB-binding back into the cytoplasm. Oscillation dynamics of NF-κB shuttling have been implicated in the functional dynamics of subsequent gene expression but it remains to be determined to what extent dynamic stimulation of the system affects nuclear-cytoplasmic NF-κB shuttling. Until recently, due to the limitations of experimental technology, most NF-κB activity has been studied in response to "step" function of TNFα induction. In-vivo, cells are normally exposed to gradual increase in external TNFα concentration[25]. With recently developed mammalian microfluidic culture platform[71], we can now consider how gradual "ramp" TNFα-induction affects NF-κB signaling. Comparing the difference in NF-κB dynam-

ics in response to ”step” and ”ramp” TNFα induction allows us to gain insight into the properties of the signaling network. Using mathematical modeling of the network, we are further able to understand the possible noise aspects of the network.

5.2 Experimental results

In order to study the effects of different TNFα induction profiles on the dynamics of NF-κB we used a cell line expressing a fluorescent p65-dsRED fusion protein (fused to relA component of NF-κB) and a nuclear H2B-GFP marker. The latter was required to effectively track single cell NF-κB profiles for up to 10hrs. Using recently developed microfluidic cell culture device, which allows for delivery of controlled TNFα dynamic profile, we were able to induce the cells with a persistent step and gradually increasing linear ramp stimulation profiles[71]. In addition, two different levels of TNFα (1ng/ml and 10ng/ml) were applied to the cells. Overall, four induction conditions were considered: 10ng/ml TNFα step, 10ng/ml TNFα ramp (5.8hr), 1ng/ml TNFα step, and 1ng/ml TNFα ramp (3.1hr). The difference in the time of the ramps was due to the difficulty of setting up ramp time without exposing cell to TNFα prior to the experiment. Over 80 single cell NF-κB response trajectories were obtained for each of the conditions. NF-κB response was measured by the mean nuclear fluorescence (p65-dsRED).

Virtually all cells responded to all four TNFα induction conditions ($> 90\%$) with some ($> 50\%$) showing oscillatory behavior (> 3 oscillatory peaks). Of the responders, the general form of the response consisted of TNFα-condition-dependent initial peak amplitude and relatively TNFα-independent secondary peak response (Fig. 5.1 and 5.5A). Cells showed much stronger and more variable NF-κB initial peak amplitude response to step TNFα induction relative to ramp profile (Fig. 5.1AB and 5.2AC) Initial response peak for ramp experiments was slower and more variable than for step stimulations (Fig 5.2CD). The timing of the initial peak was slower for the lower concentration of TNFα induction, while the amplitude of the response was higher for higher TNFα concentration. Overall, we find that the variability of the initial amplitude response did

not match the variability of the timing of initial response.

5.3 Modeling results

To help understand the variability associated with initial peak response, we adopted a simplified deterministic version of earlier described model of TNFα-induced NF-κB activation (Chapter 4). Based on our previous discussion, the intrinsic stochasticity of protein production was found to be insignificant. We removed IKKK complex from the model and instead reformulated IKK complex dynamics with a modified 4-state description. The active form of IKK, which is responsible for IκB degradation, was defined by two different states, with one being more active (IKKa+) than then other (IKKa). IKKa is part of a loop, where the activation of IKKn is proportional to TNFα concentration. We included a variable threshold parameter (r) for TNFα activation, which leads to variable temporal response observed in experimental data. IKKa is then converted to an inactive form IKKi, which cannot target IκB for degradation, before returning to neutral state, IKKn. IKKa can also reversibly enter a more active state, IKKa+, which represses its own formation. This dynamic is responsible for TNFα independent IκB degradation response. The dynamics downstream of IKK activation describe IKK dependent degradation of IκB proteins leading to NF-κB translocation into the nucleus, where NF-κB binding to DNA results in production of IκB proteins, which in turn helps remove NF-κB from the nucleus (Fir 5.1A). The model consists of 8 ordinary differential equations, and one delayed differential equation describing the production of IκB. There are 25 parameters, of which, 17 associated with NF-κB dynamics were adopted from the earlier described model. The rest of the parameters, upstream of NF-κB loop, were fitted to the mean amplitude and timing of initial response and the amplitude of the secondary response using well known metropolis algorithm.

We then simulated extrinsic noise, by varying the values of total IKK, total NF-κB, delay, and TNFα threshold (r) parameters according to uniform distributions (500 simulated trajectories). We fitted the range of variability for each of the parameters,

obtaining NF-κB initial and secondary response distributions that are qualitatively very similar to our experimental findings (Fig. 5.4 and 5.5B). While the fit is not perfect, due to simplicity of the model, the model is able to explain the increased variability in the timing of the initial NF-κB response to ramp TNFα induction and the variable amplitude of the NF-κB initial response to step TNFα induction profile. Our model shows more variable initial peak response to ramp TNFα activation, which is most likely due to our ability to detect very low amplitude response in our model relative to experimental data, where measurement noise plays a significant factor.

5.4 Discussion

Using novel microfluidic experimental setup, we were able to characterize single-cell NF-κB initial response using dynamic TNFα stimulation profile. Our findings, corroborated by our model, show that the effect of variability in cell-to-cell parameters on the variability of response timing is abated when cells are exposed to TNFα in a stepwise manner. In ramp experiments that mimic an in-vivo profile of induction, we observe enhanced variability as each cell is gradually exposed to its own minimum level of TNFα required to elicit NF-κB translocation response. Furthermore, the cell-to-cell variability contributes to the variability in amplitude response regardless of the TNFα concentration. These results indicate that there are significant quantitative differences in cell responses to step versus linear ramp activation be TNFα. Since ramp profile TNFα activation is a more accurate representation of the signal cells experience in vivo, the focus on the timing of initial NF-κB response requires further attention. The proposed model also suggests a new IKK activation mechanism consisting of two different active states for IκB kinase, which requires further experiment validation. Another key question raised by our findings is what is the key biological role for cell-to-cell variability in the timing of the response?

5.5 Model equations

$$
\begin{aligned}
\dot{NF\kappa B} &= c5a \cdot I\kappa B\_NF\kappa B + kd1a \cdot I\kappa B\_NF\kappa B - ka1a \cdot NF\kappa B \cdot I\kappa B + \\
&\quad + kc2a(IKKa + \epsilon \cdot IKKa+) \cdot I\kappa B\_NF\kappa B - \\
&\quad - ki1 \cdot NF\kappa B + ke1 \cdot NF\kappa B \qquad (5.1) \\
\dot{nNF\kappa B} &= -ka1a \cdot kv \cdot nI\kappa B \ \cdot nNF\kappa B + kd1a \cdot nI\kappa B\_NF\kappa B - \\
&\quad - ke1 \cdot nNF\kappa B + + c5a \cdot nI\kappa B\_NF\kappa B + ki1 \cdot NF\kappa B \qquad (5.2) \\
\dot{I\kappa B} &= c1a \cdot \frac{q1 \cdot nNF\kappa B(\tau)}{q1 \cdot nNF\kappa B(\tau) + q2a \cdot nI\kappa B(\tau)} - c4a \cdot I\kappa B \\
&\quad - kc1a(IKKa + \epsilon \cdot IKKa+) \cdot I\kappa B - ka1a \cdot I\kappa B \cdot NF\kappa B - \\
&\quad - ki3a \cdot I\kappa B + ke3a \cdot nI\kappa B + kd1a \cdot I\kappa B\_NF\kappa B \qquad (5.3) \\
\dot{nI\kappa B} &= -ka1a \cdot kv \cdot nI\kappa B \cdot nNF\kappa B + ki3a \cdot I\kappa B - ke3a \cdot nI\kappa B + \\
&\quad + kd1a \cdot nI\kappa B\_NF\kappa B \qquad (5.4) \\
\dot{I\kappa B\_NF\kappa B} &= ka1a \cdot I\kappa B \cdot NF\kappa B - c5a \cdot I\kappa B\_NF\kappa B - \\
&\quad - kc2a \cdot (IKKa + \epsilon \cdot IKKa+) \cdot I\kappa B\_NF\kappa B - \\
&\quad - kd1a \cdot I\kappa B\_NF\kappa B + ke2a \cdot nI\kappa B\_NF\kappa B \qquad (5.5) \\
\dot{nI\kappa B\_NF\kappa B} &= ka1a \cdot kv \cdot nI\kappa B \cdot NF\kappa B - ke2a \cdot nI\kappa B\_NF\kappa B - \\
&\quad - kd1a \cdot nI\kappa B\_NF\kappa B - c5a \cdot nI\kappa B\_NF\kappa B \qquad (5.6) \\
\dot{IKKn} &= kp \cdot (KNN - IKKn - IKKa - IKKa^{+}) - \\
&\quad (TNF\alpha > r) \cdot TNF\alpha \cdot ka \cdot IKKn \qquad (5.7) \\
\dot{IKKa} &= (TNF\alpha > r) \cdot TNF\alpha \cdot ka \cdot IKKn - \\
&\quad ki \cdot IKKa - \frac{ks1 \cdot IKKa}{(1 + IKKa^{+}/S)^2} + ks2 \cdot IKKa^{+} \qquad (5.8) \\
\dot{IKKa^{+}} &= \frac{ks1 \cdot IKKa}{(1 + IKKa^{+}/S)^2} - ks2 \cdot IKKa^{+} \qquad (5.9)
\end{aligned}
$$

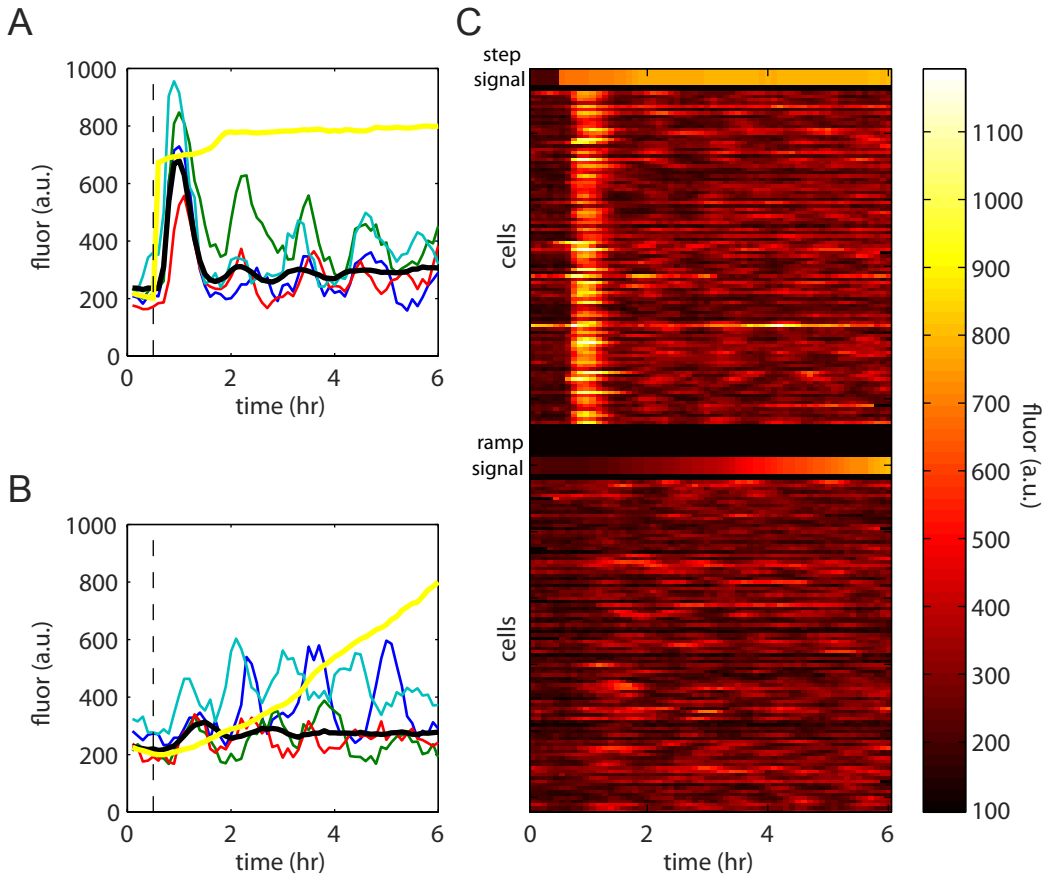

Figure 5.1: NF-κB nuclear localization in response to step and ramp TNFα activation. (A) Single cell response (color) to 10 ng/ml TNFαstep induction (yellow). Mean NF-κB response is in black. (B) Single cell response (color) to 10 ng/ml TNFαramp induction (yellow). Mean NF-κB response is in black. (C) NF-κB response of 50 cells to step and ramp TNFα activation shows fast and strong initial peak response to step relative to ramp induction.

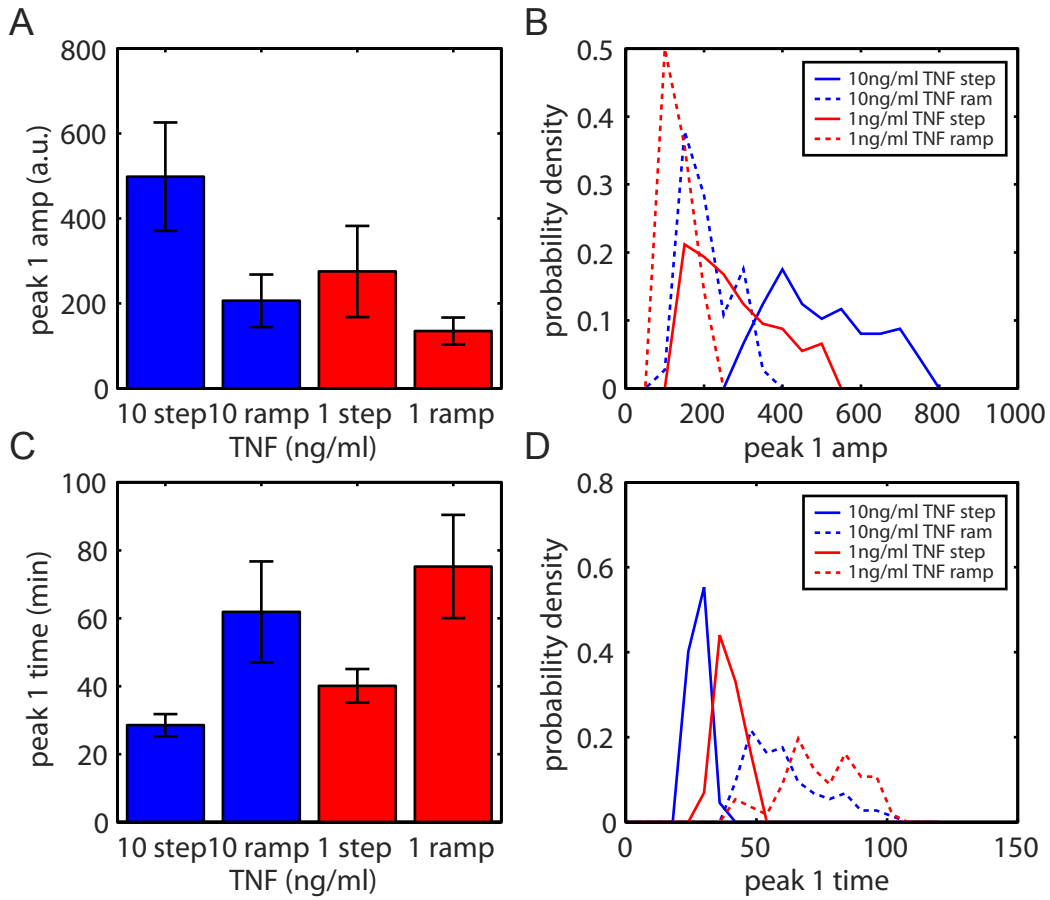

Figure 5.2: Amplitude and timing statistics of initial peak NF-κB response. (A,B) Amplitude of initial NF-κB response is greater for step TNFα activation with increased variability among cells relative to ramp activation. Higher concentration of TNFα results in higher amplitude response. The small difference between ramp responses between different TNFα concentrations is possibly due to much longer ramp time for the 10ng/ml TNFα concentration (5.8hrs vs 3.1hrs). (C,D) Timing of initial peak NF-κB response is significantly slower in response to ramp TNFα activation. Higher concentration of TNFα leads to faster and less variable timing of NF-κB response.

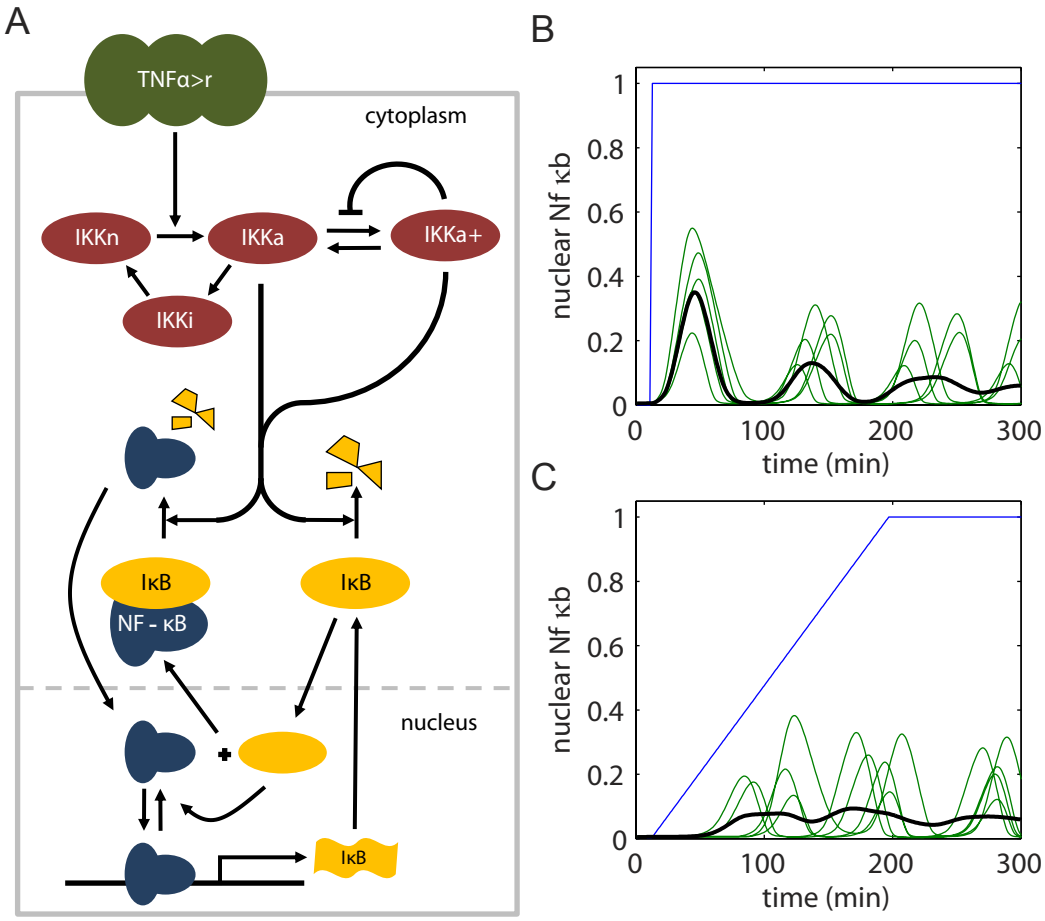

Figure 5.3: NF-κB model with simulated single cell nuclear NF-κB trajectories. (A) Simple NF-κB model is characterized to well known negative feedback loop associated with IκB binding and sequestering NF-κB molecules in the nucleus. Activation of NF-κB was modeled through TNFα threshold activation of IKK complex, which exhibits two different active states that target IκB for degradation. The IKKa+ state has negative feedback and exhibits stronger IκB degradation that IKKa state. The different states allow for NF-κBactivation that is independent of initial TNFα concentration. The model consists of 9 ODEs and 30 parameters with explicit delay in IκB production. Simulated single cell nuclear NF-κB trajectories (green) in response to step (B) and ramp (C) TNFα input (blue) are similar to the experimental trajectories (Fig 5.1A) (mean of 500 simulated trajectories is in black).

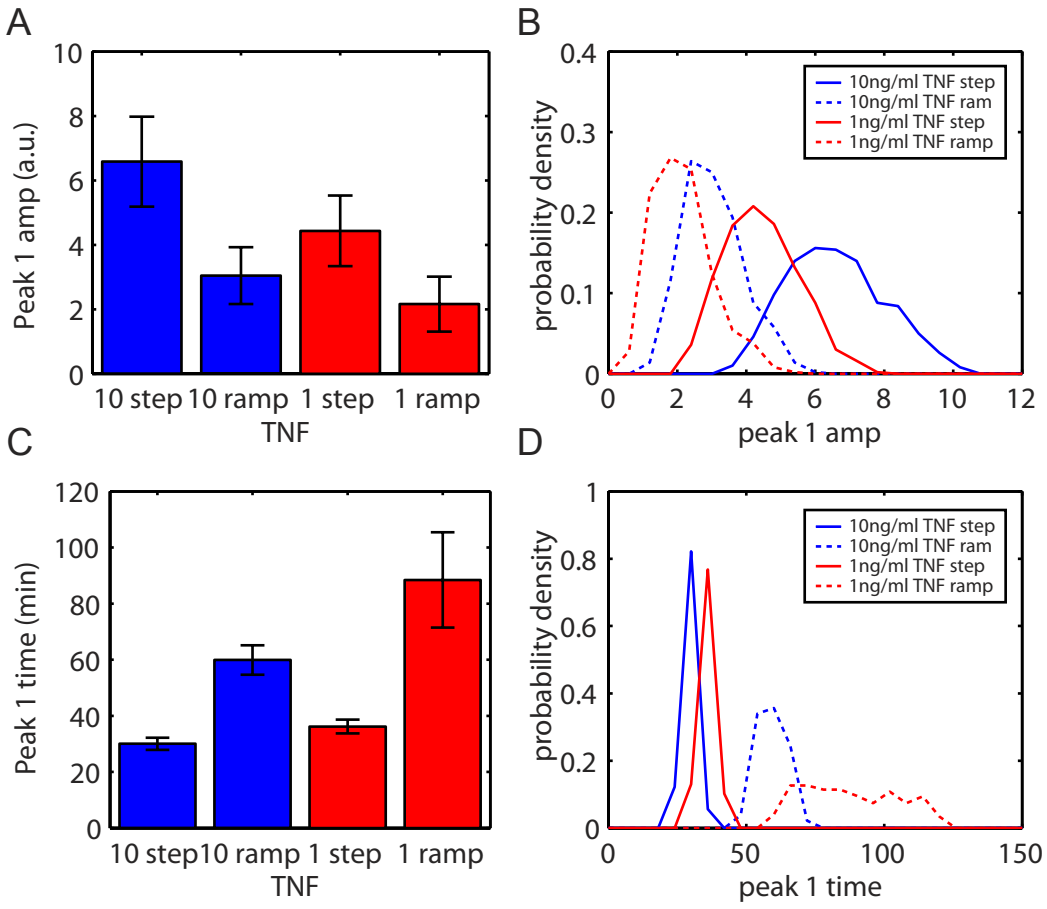

Figure 5.4: Extrinsic noise in the model recapitulates experimental amplitude and timing statistics with simulated initial peak NF-κB response. (A,B) Qualitatively the model is able to recapture both the amplitude (A,B) and timing (C,D) of initial NF-κB response. Unlike the experiment, the deterministic trajectories allow us to capture much lower amplitude response for ramp TNFα induction, which results in broader amplitude distribution relative to what we see in the experiment. The TNFα threshold parameter in the model is able to explain the broader initial peak time distribution of NF-κB response to ramp TNFα activation.

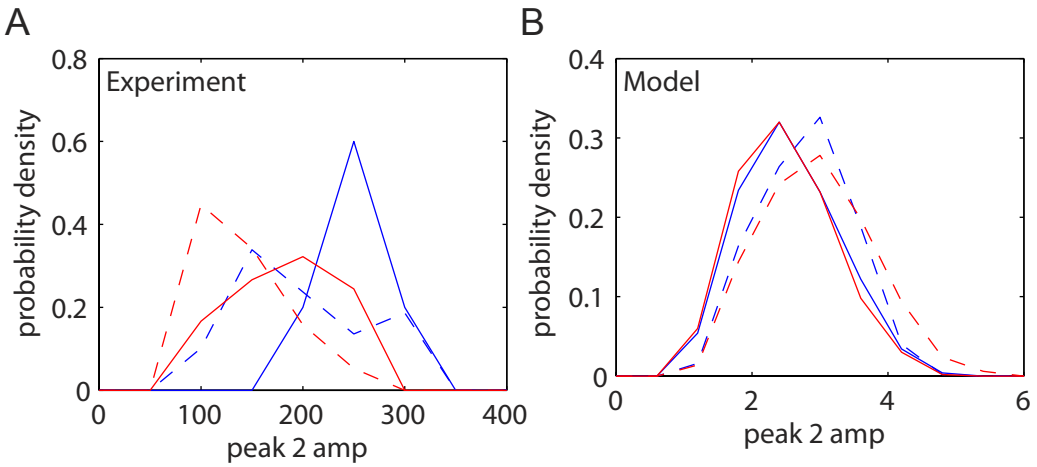

Figure 5.5: Secondary peak amplitude response distribution. (A) The amplitude of the secondary peak response shows much less separation for different levels of TNFα and induction profiles. The slight differences could be accounted for by measurement differences between associated with experimental setup. (B) The model supports the claim that secondary response is independent of TNFα level and induction profile.

Chapter 6

Accurate Information Transmission in Dynamic Biochemical Signaling Networks

6.1 Introduction

Stochasticity inherent to biochemical reactions (intrinsic noise) and variability in cellular states (extrinsic noise) can degrade information transmitted through biochemical signaling networks. Using a novel algorithm we analyzed the ability of temporal signal modulation, i.e. dynamics, to reduce noise-induced information loss. In three signaling pathways, Erk, Ca^{2+}, and NF-κB, considering response dynamics reveals significantly higher information transmission capacities than when these responses are reduced to static signals. Theoretical analysis using information-theory formalism identified a key role for dynamics in overcoming extrinsic noise in signal transduction. Numerical simulation and experimental measurement of information transmission in the Erk network under partial inhibition confirm our theoretical predictions and show that signaling dynamics mitigate, and can potentially eliminate, information loss due to extrinsic noise. By curbing the information-degrading effects of cellular variability dynamic responses

substantially increase the accuracy of information transmission in biochemical signaling networks.

The role of biological signaling networks is to reliably transmit specific information about the extracellular environment to multiple intracellular downstream effectors, allowing the cell to adjust its physiological state to changing conditions. The stochasticity of molecular interactions (a.k.a noise in biological systems) can interfere with signal transduction and degrade the transmitted information. How well signaling networks can perform their core functions in the presence of noise is a fundamental question in signal transduction. Information-theoretic approaches allow the estimation of the information transmission capacity of noisy biochemical networks[16;116;147;140;30;29;137;146;91;117;54;20;13]. Previous application of such method to the analysis of signaling networks suggested that due to noise, cells lose most of the information about the concentration of ligands to the extent that they can only reliably identify the presence of an activating ligand[20;142]. The suggestion that noise degrades most of the information about the activating ligand was surprising given the demonstrated ability of key signaling hubs to differentially respond to multiple classes of ligands[60;121;110;54], which indicates higher than binary information transmission capacity. Furthermore, the recent application of an optogentic tool to measure single-cell dose responses has shown high precision in cellular Erk response that persists over multiple hours[138]. Therefore it is unclear to what extent does noise actually limits the information transmission capacity of a biochemical signaling networks.

6.2 Results

One possible resolution of these conflicting observations is based on the fact that the information-theoretic analyses of signaling networks so far have solely been based on a scalar static measurements done at a single time point, whereas in many cases the information on activating ligands is encoded using a multivariate vector that contains the single cell response at multiple time points[60;121;54;111]. Therefore, it is plausible that the vector dynamic response has higher information transmission capacity than the pre-

vious estimates based on singe time point measurements. To test this hypothesis, we performed dynamical single cell measurements of three key signaling pathways (Figure 6.1): Erk activity (Figure 6.1C) as reported by the EKARev FRET biosensor[1;72] in MCF10A cells in response to Epidermal Growth Factor (EGF), Ca2+ levels (Figure 6.1D) in MCF10A cells in response to ATP, and NF-κBactivity (Figure 6.1E) as reported by the dynamic nucleo-cytoplasmic shuttling of EYFP-p65 in Raw264.7 cells in response to Lipopolysaccharides. Fully automated computational image analysis allowed us to measure the response of 910,121 cells. The large sample size was instrumental for the analysis of high-dimensional multivariate dynamic responses. In all three pathways there was substantial variability within the dynamic (Figure 6.1CDE) and static (Figure 6.1F) single cell responses across multiple levels of activating ligands.

To analyze the implication of noise on information loss we utilized an information-theoretic approach to calculate the information transmission capacity of a dynamic signaling network. The information transmission capacity (also referred to as channel capacity[24]) can be measured through the calculation of the maximal mutual information between the measured response and the activating ligand level. To calculate the mutual information of the dynamic response, we developed a new estimation algorithm that allowed us to calculate the mutual information between a multivariate response (a vector) and the activating ligand (a scalar). The algorithm uses continuous multi-dimensional response data and is based on the estimation of the conditional probability density for each cells response using k-nearest-neighbors approach. Using our new algorithm we estimated the information transmission capacity of the dynamic response and several types of static responses. For single time point static responses we saw that the transmission capacity (< 1 bit) is indeed similar to what was found previously[20;142] (Figure 6.2A) for all single time point estimates. However, in all three pathways, the transmission capacity increases for a multivariate dynamic response (Figure 6.2B). Overall, across all three signaling pathways, the dynamic response had significantly higher information transmission capacity than several possible static responses (Figure 6.2C, p-value < 0.05 for all comparisons).

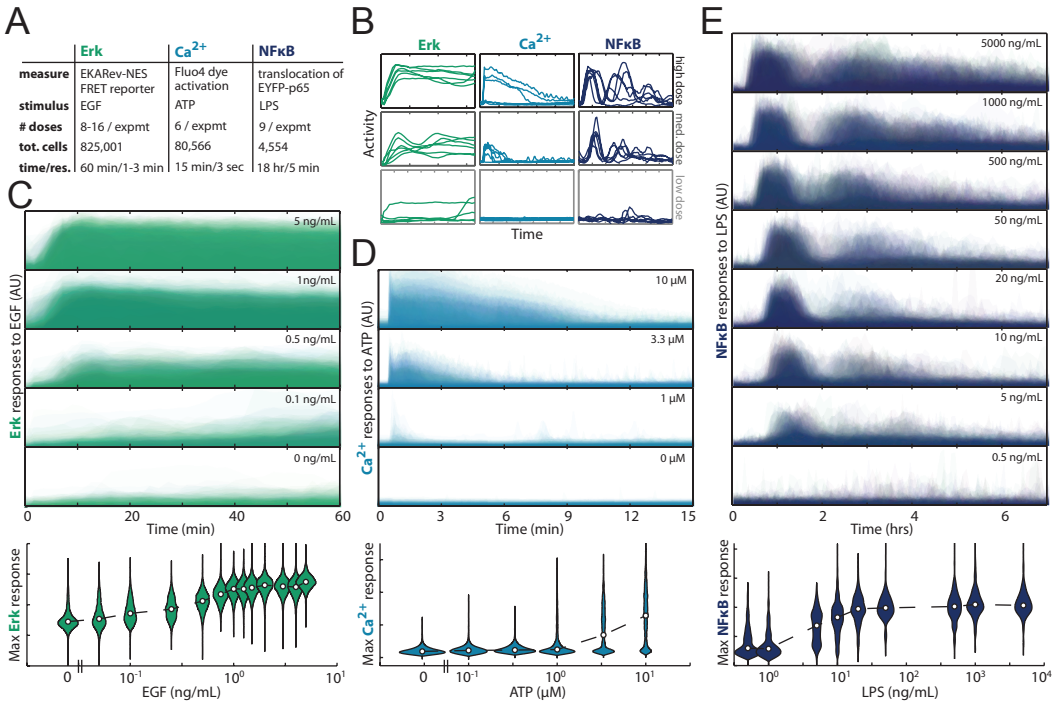

Figure 6.1: Single cell measurement of the dynamic response of Erk, Ca2+ and NFkB. A. Overview of single cell data analyzed in this work. B. Examples of single cell response dynamic trajectories. CDE. Temporal histograms of several representative dosages for Erk (C) Ca2+ (D) and NFkB (E) where the intensity of color reflects the probability density of a cellular response at that magnitude at that time point. The scale of the y-axis in B and CDE is the same for each pathway. The Arbitrary Units represent the FRET/CFP ratio reported by the EKARev Erk biosensor (C), the nuclear intensity of Ca2+ indicator dye Flou-4 (D), and the ratio of nuclear to cytoplasmic localization of an EYFP-p65 reporter (E). F. dose response curves of maximal response in the three signaling pathways.

To elucidate the origins of the enhanced information transmission capacity of dynamic signaling response we developed a mathematical theory using information theoretic formalism (Selimkhanov et al., in preparation). The key advancement of our theory is that it explicitly accounts for the differences between the information-degrading effects of intrinsic and extrinsic noise sources in the context of multivariate responses. Intrinsic noise adds to the uncertainty in all dimensions (i.e. time points) independently from each other. In contrast, extrinsic noise components are constrained by the signaling network that generates the dynamics. Therefore the fluctuating components generated by extrinsic noise are deterministically dependent on one another. As a result, intrinsic

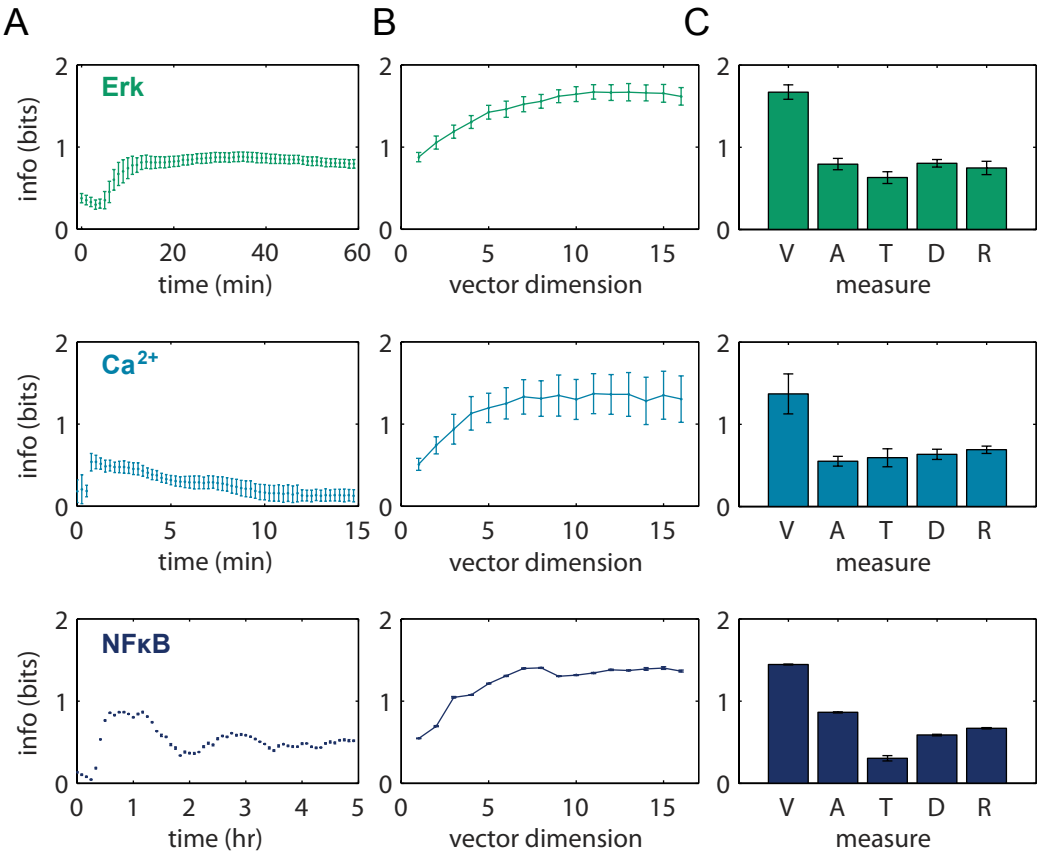

Figure 6.2: Information transmission capacity of static and dynamic Erk, Ca2+ and NFkB responses. A. Information transmission capacity calculated from static univariate response distribution based on all single time point measurements. B Information transmission capacity calculated from multivariate dynamic responses as a function of the dimension of the multivariate vector. The multivariate vector was subsampled using a uniform grid centered on the middle time point. C Comparison of the multivariate vector (V) measurement to several univariate and bivariate responses: max response amplitude (A), max response time (T), max rate of response (D), ratio of max response amplitude to initial response amplitude (R). Errorbars are s.e.m from 6 (4) biological replicas for Erk (Ca2+) and jackknife standard deviation for NFkB (5 iterations). The multivariate vector was significantly higher than all univariate measures (p-values < 0.05, all comparisons).

and extrinsic noise sources have different effects on the information transmission capacity of multivariate responses. In the case of intrinsic noise, additional independent measurements only contribute logarithmically to the overall information due to simple ensemble averaging (in accord with earlier findings by Cheong et al.[20]). In the case of extrinsic noise, dynamical measurements can provide complete information about the a

priori uncertain internal state of the cell and therefore lead to a substantial gain in the information about the activating ligand. Consequently, while an infinite number of measurements are required to completely eliminate information loss due to intrinsic noise, complete elimination of the information loss due to extrinsic noise only requires that the effective dimension of the dynamic response (number of linearly independent response measurements) is higher than the number of independent uncertain parameters which determine the dynamics of the signaling network. Figure 6.3A shows how the increase of response dimension allows cells to overcome the degrading effects of intrinsic and extrinsic noise sources.

To test our analytical theory and demonstrate the ability of signaling dynamics to overcome extrinsic noise we used computer simulations of Erk responses based on a published model[130]. The mathematical model recapitulated the experimental data on Erk dynamics (Fig. 6.9) and allowed us to create Erk response distribution resulting solely from extrinsic noise. Our theoretical analysis predicts that the multivariate dynamic response can completely overcome the information loss that results from introduction of extrinsic noise. To test this prediction, we created sets of simulated Erk trajectories given an increasing number of input levels when the model parameters for Erk and Mek varied according to uniform distribution ($+/-20\%$ mean value). The simulation results were used to calculate the mutual information between Erk response and activating ligand. Our analysis (Figure 6.3B) supports the analytical prediction and shows that indeed the dynamic multivariate response can transmit the complete information on the activating ligand. In contrast, the information transmission capacity of a univariate response based maximal Erk levels saturated and therefore was unable to faithfully transmit the complete information about activating ligand. An intuitive demonstration for the limitation of univariate response and the ability of multivariate response to overcome extrinsic noise is shown in Figure 6.3CD. Figure 6.3C shows the simulated time series of Erk_pp for two input levels of EGF with a single parameter Mek randomly chosen from a uniform distribution ($+/-20\%$ mean value). Superficially, the trajectories appear overlapping, but in fact, they are completely separable via the procedure known

as embedding[123]. Plotting the frequency of Erk-pp at $t = 9$ and $t = 24$ minutes on a two-dimensional plane (Figure 6.3D) shows that since there was only a single parameter that was varied, the responses to a single input level lie on a one dimensional manifold within the two-dimensional embedding space. The two one-dimensional manifolds for different inputs are completely separated from each other (inset), but overlap considerably in any one-dimensional projection. This simple example demonstrates how the extrinsic variability of a single parameter can be completely eliminated with only two time points.

The ratio of desired to spurious responses (also known as the signal to noise ratio, or SNR) is an important property of communication channels and signaling network. The mutual information between the input strength and the response is generally determined by the systems SNR. But on a more quantitative level, the mutual information depends on the properties of noise (intrinsic vs. extrinsic) and the type of response. Our analytical theory describes mutual information for three different types of responses as a function of SNR: (1) scalar responses that do not distinguish between intrinsic and extrinsic noise, (2) redundant vector responses (repeated measurements) that can only reduce intrinsic noise, and (3) dynamic responses that combine all the benefits of redundant measurements with very efficient mitigation of extrinsic variability. To test these results experimentally we varied the SNR in the Erk network using partial inhibition of the Erk kinase MEK using six different dosages of the inhibitor U0126. At each MEK inhibitor level we measured Erk response to eight EGF levels. Total 48 conditions were measured in four biological replicas. At each MEK inhibitor level we calculated the mutual information and the SNR using the single cell responses. Overall Figure 6.4 is based on the single cell response of 535,107 cells overall. The theoretical estimate of the mutual information for the static response, which only needs the overall SNR, is in very good agreement with our experimental measurements. The theoretical prediction of mutual information for the redundant case additionally requires the knowledge of the ratio of intrinsic to extrinsic noise (IER). IER was estimated in two separate ways: (1) by quantifying the fluctuations in the later (quasi-stationary) portion of the response

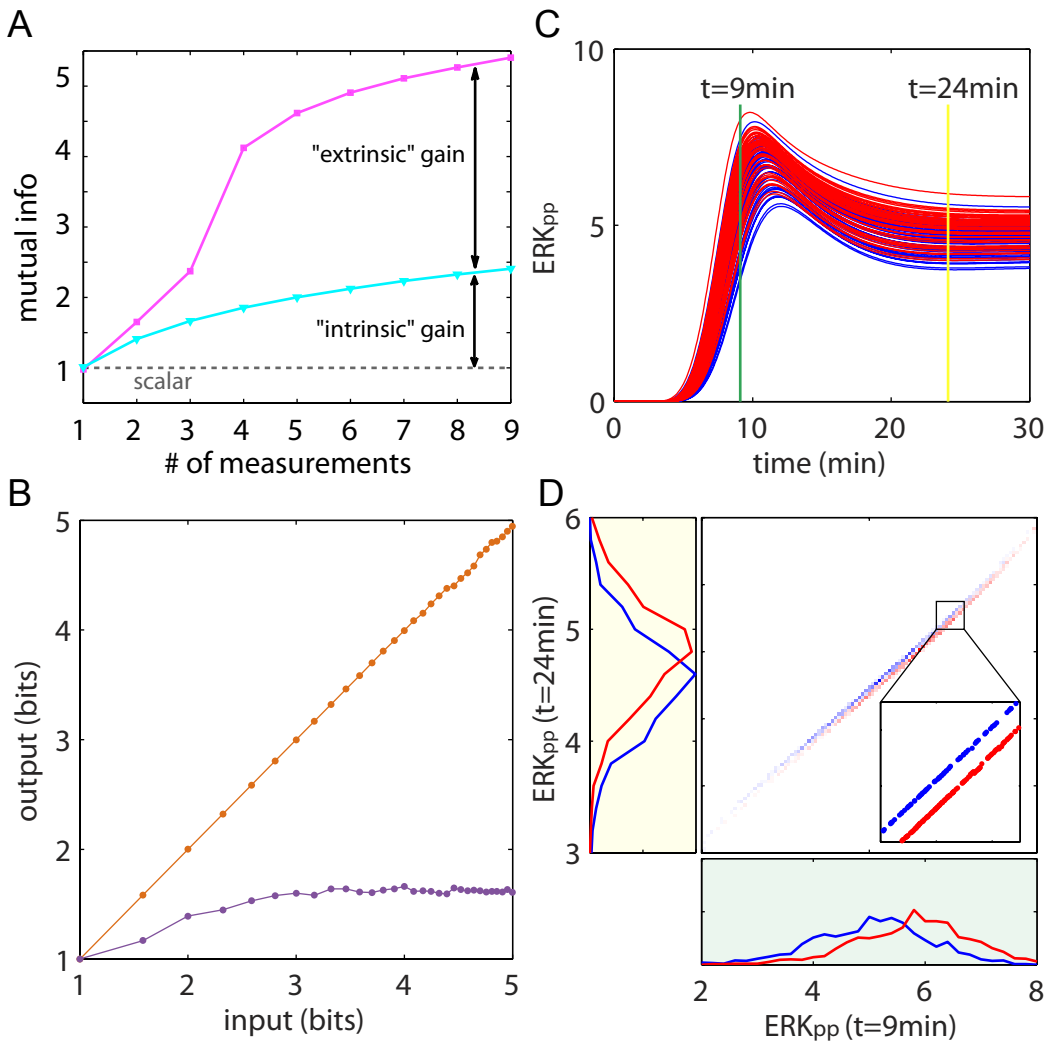

Figure 6.3: Theoretical decomposition of information loss caused by intrinsic and extrinsic noises. A graphical representation of the analytical expression for the gain from overcoming intrinsic (cyan) and extrinsic (magenta) noise sources obtained form random systems with three parameters. B information transmission capacity of dynamic (blue) and static (maxmial response, green) calculated based on computational model of Erk where only extrinsic noise contributed to cell response variability. C Example of Erk trajectory variability for two different inputs levels (red and blue). Variability was generated using a normal distribution of a single parameter, Mek concentration. D Two dimensional histogram (center) and marginal distributions (left and bottom) for the two input levels (shown in red and blue) at two time points ($T = 9$ & 24min) from the trajectories shown in C. Since only a single parameter was varried, the responses vary on a one dimensional manifold. As a result, while the univariate marginal distributions show substantial response overlap, the two dimensional distribution shows completely seperated response levels (inset).

time series of our measured Erk data (Fig. 6.13A). (2) Using the data from Toettcher et al.[138] that performed repeated measure of a single cell response (Fig. 6.13A). The predicted mutual information based on redundant responses using either of the IER estimates are significantly below the experimentally measured values. On the other hand, the measured mutual information are in good agreement with the values predicted from the dynamic model that was based on an Erk model and required additional estimation of the effective number of unobserved system parameters (Fig. 6.10). Overall this analysis demonstrates that the information gain from multivariate measurements is indeed the direct result of the dynamic nature of Erk response.

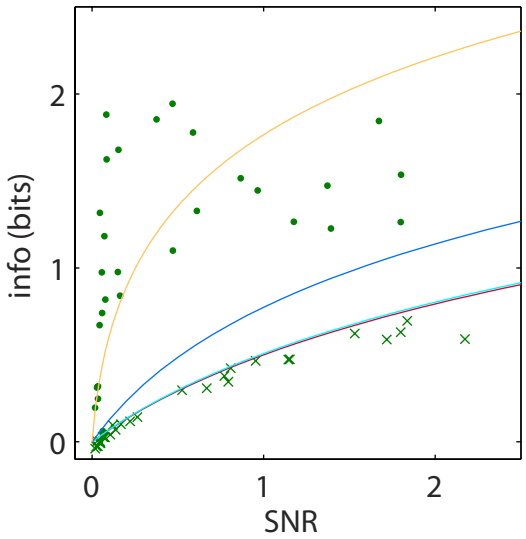

Figure 6.4: Measured information gain is a result of Erk dynamics ability to mitigate extrinsic noise. Experimental measurement of the mutual information between Erk response and EGF measured as a function of the response magnitude to noise ratio. Each pair of markers (dot and cross) is based on calculations of SNR and mutual information from the (dynamic and static) responses of 18,500 cells from an 8-well dose response experiment. Overall 29 experiments that were done under six levels of the MEK inhibitor U0126 are shown that included total of 535,107 single cells responses. Theoretical predictions of the mutual information as a function of SNR for three types of responses: static scalar (red line), redundant measurements where the multivariate response has no dynamics (dark and light blue lines) calculated based on two independent estimates of IER, and dynamic response (orange) that can mitigate both intrinsic and extrinsic noise.

6.3 Discussion

The robustness of biological systems is epitomized by their ability to function in the presence of large uncertainties[145;8]. A major source of uncertainty is the variability in cellular states, e.g. protein concentration within individual cells. We showed that signaling dynamics allow biochemical networks to mitigate this major source of extrinsic noise and thereby maximize the information transmission capacity of signaling networks. Our estimates for the information transmission capacities should be considered as lower bounds since part of what we may refer to as intrinsic noise may in fact be a result of experimental imperfections. As a result, the intrinsic to extrinsic ratio may be lower than our estimates which would only further support the overall claim that dynamic response increases the information transmission capacity. Furthermore, the information about the input ligand may be encoded in the dynamics of multiple signaling molecules. While the theory and observations presented in this work focused on the information transmission capacity of the dynamics of a single signaling molecule, the extension of our results to the case of multiple signaling molecules responding to the same ligand is straightforward. It is important to note that not all of the information contained in the multivariate dynamical responses may actually be used by cells. Yet, since reliable information transmission is a fundamental function of cellular signaling networks, it is plausible that evolutionary pressures shaped the cellular machinery to maximize the reliable decoding of multivariate dynamic signals.

6.4 Information transfer estimation from experimental data

6.4.1 Algorithm derivation

To derive the algorithm for estimation of information transfer, we first consider the type of experimental data that we have acquired, which will guide our general ap-

proach. In our signal transduction networks, the input signal S is defined by m discrete levels of extracellular ligand concentration ($S = [s_1, s_2, ..., s_m]$). For each input signal s_i we have n_i output protein trajectories ($R_i = [r_{i1}, r_{i2}, ..., r_{in_i}]$), with each trajectory occupying a single point in continuous Euclidean space of dimension d, where d is the number of time points in each output trajectory. Combined, we have $N = \sum_i^m n_i$ trajectories in our response R array. The general breakdown of the data is as follows:

$$S = \begin{bmatrix} s_1 \\ s_2 \\ \vdots \\ s_i \\ \vdots \\ s_m \end{bmatrix}, R = \begin{bmatrix} R_1 \\ R_2 \\ \vdots \\ R_i \\ \vdots \\ R_m \end{bmatrix}, s_i \rightarrow R_i = \begin{bmatrix} r_{i1} \\ r_{i2} \\ \vdots \\ r_{ij} \\ \vdots \\ r_{in_i} \end{bmatrix}, r_{ij} = [r_{ij,1}, r_{ij,2}, \ldots, r_{ij,d}]$$

To estimate the information transfer (I) between an input (S) and an output (R) using well known formula

$$I(R; S) = H(R) - H(R|S). \tag{6.1}$$

we need to calculate Shannon entropies $H(R)$ and $H(R|S)$. The general scheme of our approach is shown in Fig 6.5. First, given that our data is continuous, we need to define how we will estimate these entropies. For a continuous probability density $f(x)$ of some observable X, the Shannon entropy is defined as differential entropy

$$H_{\text{diff}}(X) = -\int_{-\infty}^{\infty} f(x) log_2(f(x)) dx. \tag{6.2}$$

Following change of variable of integration, Equation 6.2 becomes

$$H_{\text{diff}}(X) = -\int_0^1 log_2(f(x)) dy. \tag{6.3}$$

where $y = \int_{-\infty}^{x} f(t))dt$ is the cumulative probability density. We can estimate y by the cumulative probability distribution of N_x observations using

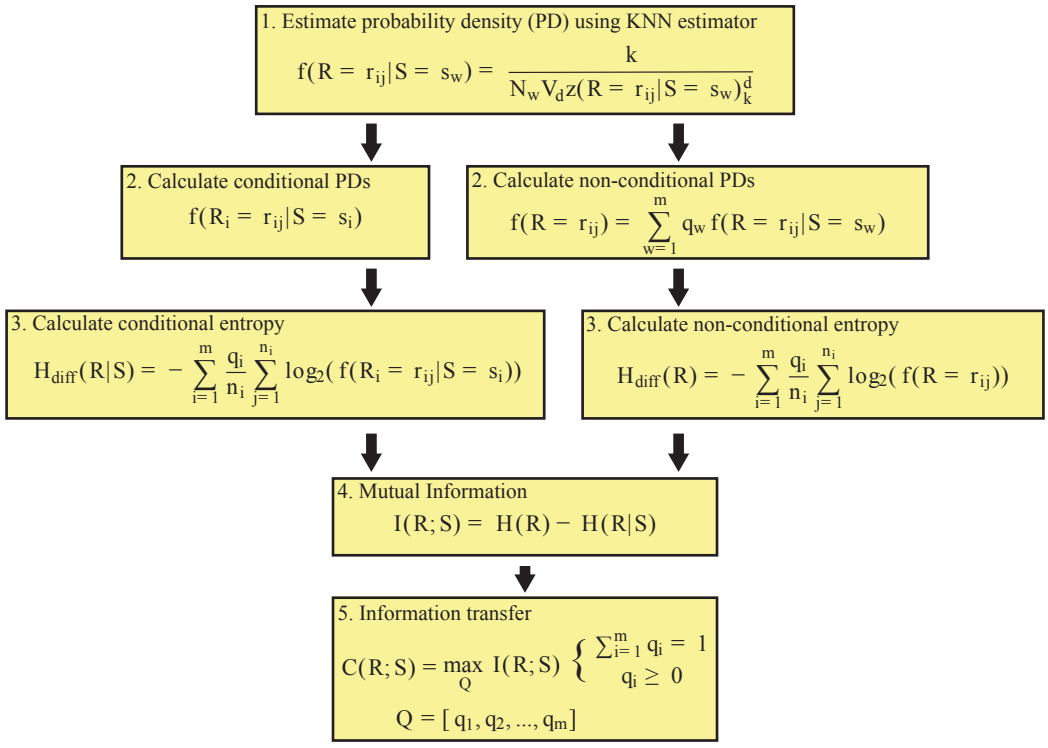

Figure 6.5: General scheme for estimation of information transmission based on experimentally obtained conditional responses.

$$H_{\text{diff}}(X) = -\sum_{j=1}^{N_x} \delta_j log_2(f(x_j)), \tag{6.4}$$

where δ_j is the probability of observing x_j, $P(X = x_j)$.

Using Equation 6.4 as our basis, we will now illustrate how to obtain $H_{\text{diff}}(R|S)$ and $H_{\text{diff}}(R)$, given that our experimental data only contains conditional probabilities of a responses.

For the conditional case, $H_{\text{diff}}(R|S)$, since all n_i responses in R_i are equally likely, $\delta_j = \frac{1}{n_i}$, we can estimate probability density of a single response r_{ij} directly from all the other responses to $S = s_i$,

$$H_{\text{diff}}(R_i|S = s_i) = -\sum_{j=1}^{n_i} \frac{1}{n_i} log_2(f(R_i = r_{ij}|S = s_i)). \tag{6.5}$$

Here $f(R_i = r_{ij}|S = s_i)$ represents the probability density of response r_{ij} in R_i given all the other responses r_i (in R_i) to the signal $S = s_i$. We will explain how to estimate $f(R_i = r_{ij}|S = s_i)$ later in the derivation. With the probability of a given signal, $q_i = P(S = s_i)$, we can then sum the conditional entropies of each signal to get overall conditional entropy,

$$H_{\text{diff}}(R|S) = \sum_{i=1}^{m} q_i H_{\text{diff}}(R_i|S = s_i) = -\sum_{i=1}^{m} q_i \sum_{j=1}^{n_i} \frac{1}{n_i} log_2(f(R_i = r_{ij}|S = s_i)). \tag{6.6}$$

The case of estimating $H_{\text{diff}}(R)$ requires special attention, since we do not have access to non-conditional probabilities of responses. The difficulty arises from the fact that our estimate of non-conditional density of a single response, $f(R = r)$ is dependent on the probability of the input signals that generated all other responses. One approach is to estimate probability density that a given response r occurred in response to a given input signal (s_w), for each of the signals in S, by effectively placing that response into R_w and estimating the probability density for r is if it were also a response to s_w. Using total probability, for every response r in R, we can estimate the probability density within each set of responses R_w and sum over m such densities multiplied by the probability q_w of the signal that generated responses in R_w, as follows

$$f(R = r) = \sum_{w=1}^{m} q_w f(R = r|S = s_w). \tag{6.7}$$

Plugging 6.7 into 6.4, we get

$$H_{\text{diff}}(R) = -\sum_{i=1}^{m} \sum_{j=1}^{n_i} \delta_{ij} log_2(f(R = r_{ij})). \tag{6.8}$$

The key difference between $f(R_i = r_{ij}|S = s_i)$ and $f(R = r_{ij})$, is that the former calculates the conditional probability density of a response among all other responses to the same signal, while the latter estimates non-conditional probability density of a response combining probability theory and conditional probability density of the response belonging to each of the subsets of responses R_i. To get δ_k we must account

for the different probabilities associated with observing responses from different input signals and the number of responses n_i obtained for each input signal:

$$H_{\text{diff}}(R) = -\sum_{i=1}^{m} \frac{q_i}{n_i} \sum_{j=1}^{n_i} log_2(f(R = r_{ij})). \tag{6.9}$$

Now that we have formulas for $H_{\text{diff}}(R)$ and $H_{\text{diff}}(R|S)$, we need to estimate the probability densities $f(R = r|S = s_i)$ in 6.7 and $f(R_i = r_{ij}|S = s_i)$ in 6.5. This can be accomplished with the k-nearest neighbor (KNN) estimator,

$$f(x_j|X) = \frac{k}{N_x V_d z(x_j|X)_k^d} \tag{6.10}$$

where

$$V_d = \frac{\pi^{\frac{d}{2}}}{\Gamma(\frac{d}{2} + 1)} \tag{6.11}$$

is the the volume of a unit sphere of dimension d (also dimension of x_j), N_x is the number of x_j in X, and $z(x_j)_k$ is the Euclidean distance to the kth nearest neighbor in X from x_j (loftsgaarden quesenberry 1965).

Applying this estimator to Equations 6.5 and 6.7, we get

$$H_{\text{diff}}(R|S) = -\sum_{i=1}^{m} \frac{q_i}{n_i} \sum_{j=1}^{n_i} log_2(\frac{k}{n_i V_d z(r_{ij}|R_i)_k^d}). \tag{6.12}$$

$$H_{\text{diff}}(R) = -\sum_{i=1}^{m} \frac{q_i}{n_i} \sum_{j=1}^{n_i} log_2(\sum_{w=1}^{m} q_w \frac{k}{n_w V_d z(r_{ij}|R_w)_k^d}) \tag{6.13}$$

where $z(r_{ij}|R_i)_k$ is the distance from response r_j in R_i to the kth nearest neighbor in R_i, while $z(r_{ij}|R_w)_k$ is the distance from response r_{ij} in R_i to the kth nearest neighbor in R_w (Figure 6.7).

Without the knowledge of q_i, we are unable to estimate the information transfer $I(R;S)$ using Equations 6.13 and 6.12. However, the maximum information transfer (C) can be calculated with

$$C(R;S) = \max_{Q}\{I(R;S)\}, \tag{6.14}$$

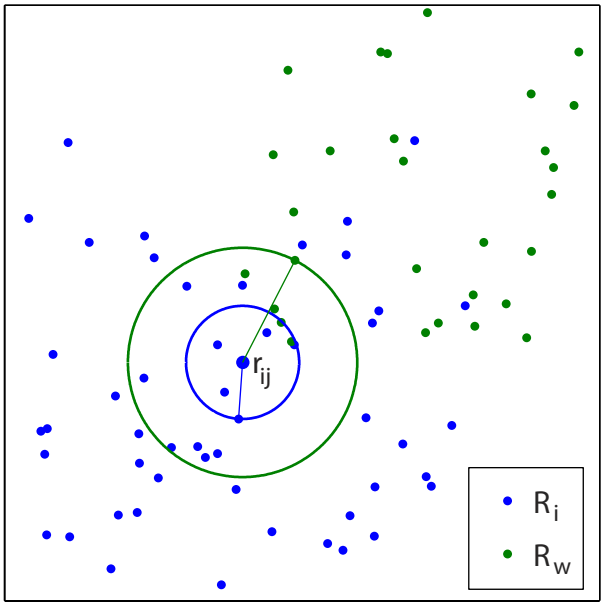

Figure 6.6: Representation of k-nearest neighbor calculation for $k = 5$. The blue circle radius is the distance to the fifth closest neighbor within the same input response represented by blue points. The green circle radius is the distance to the fifth closest neighbor to a different input response (green points).

where $Q = [q_1, q_2, ..., q_m]$, such that $\sum_{i=1}^{m} q_i = 1$ and $q_i \geq 0$ (Elements of Information Theory, 2nd ed.). This corresponds to the maximum possible informatin transfer between input S and output R.

6.4.2 Information transfer estimate validation

To test the accuracy of our algorithm, we first calculated the difference between our entropy estimate of a multivariate Gaussian distribution (5000 samples) and the exact analytical value of the entropy (Figure 6.7 A). For $k = 10$, the KNN density estimate differs by less than 5% from true value of probability density of a multivariate Gaussian distribution of up to dimension $d = 20$. To calculate the maximum information transfer C, we used fmincon function in Optimization Toolbox™ (MATLAB®) to optimize over possible q_i. Using our algorithm ($k = 10$), we calculated C of two

unit-variance Gaussian responses ($5,000$ samples each), for increasing separation between their means (Figure 6.7 B,C red curve). For comparison, we also used exact values for entropies to calculate C, showing little difference between our method and the exact information transfer C (Figure 6.7 C blue curve). In addition, we looked at how well our method does in the case of different number of responses to individual signals as well as different variance in the distribution of responses to different signals (Fig. 6.7D,E). In all cases, our algorithm does very well relative to true probability densities of responses. Furthermore, we compared our "bin-less" algorithm in the case of one dimensional Gaussians against "binned" algorithm described by Cheong et al., for increasing separation between two Gaussians, again finding very comparable results (Fig 6.8A).

To correct for sample size bias, we adapted jackknife sampling procedure similar to Cheong et al.[20]. Specifically, we computed information transfer (C) for sampled fractions of the data ($60\% - 100\%$). Then, plotting (C) relative to the inverse of sample size, we fitted a straight line to obtain the y-intercept corresponding to the information transfer for infinite sample size (Fig. 6.8B).

6.5 Additional calculations for model fitting

6.5.1 Erk model simulations

For numerical simulations, we adopted ODE model of Erk signaling network from Sturm et al.[130]. The model incorporates dynamics from RasGTP through Raf and Mek down to Erk phosphorilation. We used the input concentration of RasGTP as a proxy for extracellular EGF, varying its value over several orders of magnitude. The early dynamics of ERKpp generated by the model closely resemble the dynamics of FRET signal recorded in experiments (Figure 6.9). Using this model, we were then able to test some of the results predicted by the earlier discussed theory.

Simulations with extrinsic noise

To illustrate the effect of using multi-dimensional measurements to eliminate the contribution of extrinsic noise to the information transfer of the system, we calculated MI using model simulation trajectories of ERKpp as the response and the input RasGTP. The range of input RasGTP (2500 to 22500) was chosen to minimize saturation at the both ends of response. Except for ERK and MEK, model parameter values were kept consistent with Strum et al.for all simulations. To generate extrinsic noise we randomly sampled ERK and MEK values from uniform distributions ($\pm 20\%$) centered around the nominal values presented in Strum et al.. The model was allowed to reach a steady state with the chosen ERK and MEK parameters at the lowest value of RasGTP (2500), before applying inducing amount of RasGTP. The model was simulated for 30min. As the number of input levels of RasGTP within the input range increased, we found that the multi-dimensional measurement of ERKpp ($0 : 3 : 30min$) resulted in MI equivalent to the number of input levels (blue curve in Fig 6.3B). In stark contrast, the single measurement MI shows a saturation around 2 bits (green curve in Fig 6.3B), which is consistent with theoretical prediction that at least 3 measurements are required to completely remove extrinsic noise.

To further demonstrate the underlying principles that allow for elimination of extrinsic noise, we plot 50 ERKpp trajectories generated from two nearby input values of RasGTP (Fig 6.3C). The overlap between these trajectories might suggest that it would be practically impossible to distinguish between the two input signal values. Furthermore, considering two different time points (green lines), this is exactly the case given the overlap between ERKpp response values at those time points (left and bottom 1-D histograms in Fig 6.3D). If we consider those two points together, however, we can clearly see a separation between the two input levels as the 2-D histogram in Fig 6.3D indicates.

Fitting vector MI vs SNR data

To fit dynamic Erk response experimental inhibition data in Figure 3B (green filled circles), we applied theoretical description of mutual information and SNR (see Section IIC: Extrinsic and intrinsic noise) to our Erk model. We calculated sensitivities of Erk response at 10 equally spaced time points on the trajectory with respect to the signal (Ras_GTP) and 7 model protein parameters (Mek, Erk, Raf, Phase1, Phase2, Phase3, Phase4) near the middle induction level and nominal parameter values. Constructing Gram matrices with these sensitivities, we were then able to calculate MI as a function of SNR. We assumed equal coefficient of variation (CV) for all parameters. Intrinsic noise was calculated based on experimental IER ratio. Varying the number of parameters contributing to the extrinsic noise in the model, we were able to fit the mean MI vs SNR curve to the experimental data (Supp Fig. 6.10. For a given number of parameters, we generated the mean MI vs SNR curve for all combinations of parameters. The best fit was obtained with two parameters accounting for the extrinsic noise in the model. This could be thought of as the number of effective system parameters contributing to extrinsic noise that our dynamic measurements can overcome.

6.5.2 Sampling dimension for vector response

To select the time points that should be part of the dynamic response vector for increasing dimension of the vector in Figure 4B, we used a simple strategy. For a given time frame of the trajectory (60min for ERK, 15min for calcium, 5hrs for NF-κB) and chosen dimension N, we selected $N + 2$ equally spaced response values throughout the given time frame, removing the first and the last values. For example, for vector of dimension 1, we chose response value at the center of the given time frame, while for dimension 2, we chose values located at the one third and two thirds points of the time frame (Fig. 6.11).

6.5.3 Experimental noise analysis

Signal-to-Noise Ratio (SNR)

To calculate Erk signal-to-noise ratio (SNR), we defined the signal magnitude σ_r^2 as the variance of average responses over all m input levels of EGF:

$$\sigma_r^2 = \frac{1}{m}\sum_{i=1}^{m}((\frac{1}{m}\sum_{w=1}^{m}\frac{1}{n_w}\sum_{j=1}^{n_w} r_{wj}) - \frac{1}{n_i}\sum_{j=1}^{n_i} r_{ij}))^2 \tag{6.15}$$

Noise magnitude was defined as the average of the variances of n_i responses to a single input level of EGF:

$$\sigma_n^2 = \frac{1}{m}\sum_{i=1}^{m}(\frac{1}{n_i}\sum_{j=1}^{n_i}(\frac{1}{n_i}\sum_{w=1}^{n_i} r_{iw} - r_{ij})^2 \tag{6.16}$$

SNR is then σ_r^2/σ_n^2.

Autocorrelation of Erk response

We performed autocorrelation analysis on Erk trajectories to gain insight into our time sampling of the data. According to the analysis, decay of autocorrelation function shows that on average self-correlation is lost after 11min (Fig. 6.12). This suggests that Erk dynamics can on average be characterized by 6 points over 60min time frame.

Intrinsic-to-Extrinsic Ratio (IER)

To calculate Erk intrinsic-to-extrinsic noise ratio (IER) from our data, we defined intrinsic noise as combination of stochasticity inherent to biochemical reactions in signal transduction and measurement noise, while extrinsic noise was defined by the variability in individual cell states. To estimate the upper bound on the experimental IER, we used the fact that our sampling of Erk response was faster than Erk dynamics (Autocorrelation of Erk response), to calculate intrinsic noise. Using second portion of Erk trajectory, where Erk levels did not change significantly between successive time

point measurements, we estimated the intrinsic noise (σ_ξ^2) as the variance of the differences in Erk expression between successive time points (Supp Fig. 6.13 red). To get the extrinsic noise (σ_e^2), we estimated the total noise according to equation 6.16 (Supp Fig. 6.13 cyan) for the second portion of Erk trajectory, and simply subtracted the intrinsic noise from the total noise. The mean IER (σ_ξ^2/σ_e^2) for all experimental conditions (69) was estimated to be 0.024.

To estimate IER from Toettcher et al.[138] data, we used a slightly different approach. Using hill function fit, presented in Toettcher et al., for each cell, we estimated intrinsic noise as variance of the differences between experimental values and the model fit (Fig. 6.13 red). Similarly, to get the total noise, we fit the same function to all of the points and calculated the variance of the differences between experimental values and the model fit (Supp Fig. 6.13 cyan). To get extrinsic noise, we simply subtracted the intrinsic noise from the total noise. The mean ration of intrinsic to extrinsic noise was estimated to be 1.14.

6.6 Acknowledgments

Chapter 6 contains material being prepared for publication as Selimkhanov, J.*, Taylor, B.*, Yao J., Pilko A., Albeck J., Hoffmann A., Tsimring L.S., Wollman R.. Accurate information transmission in dynamic biochemical signaling networks.

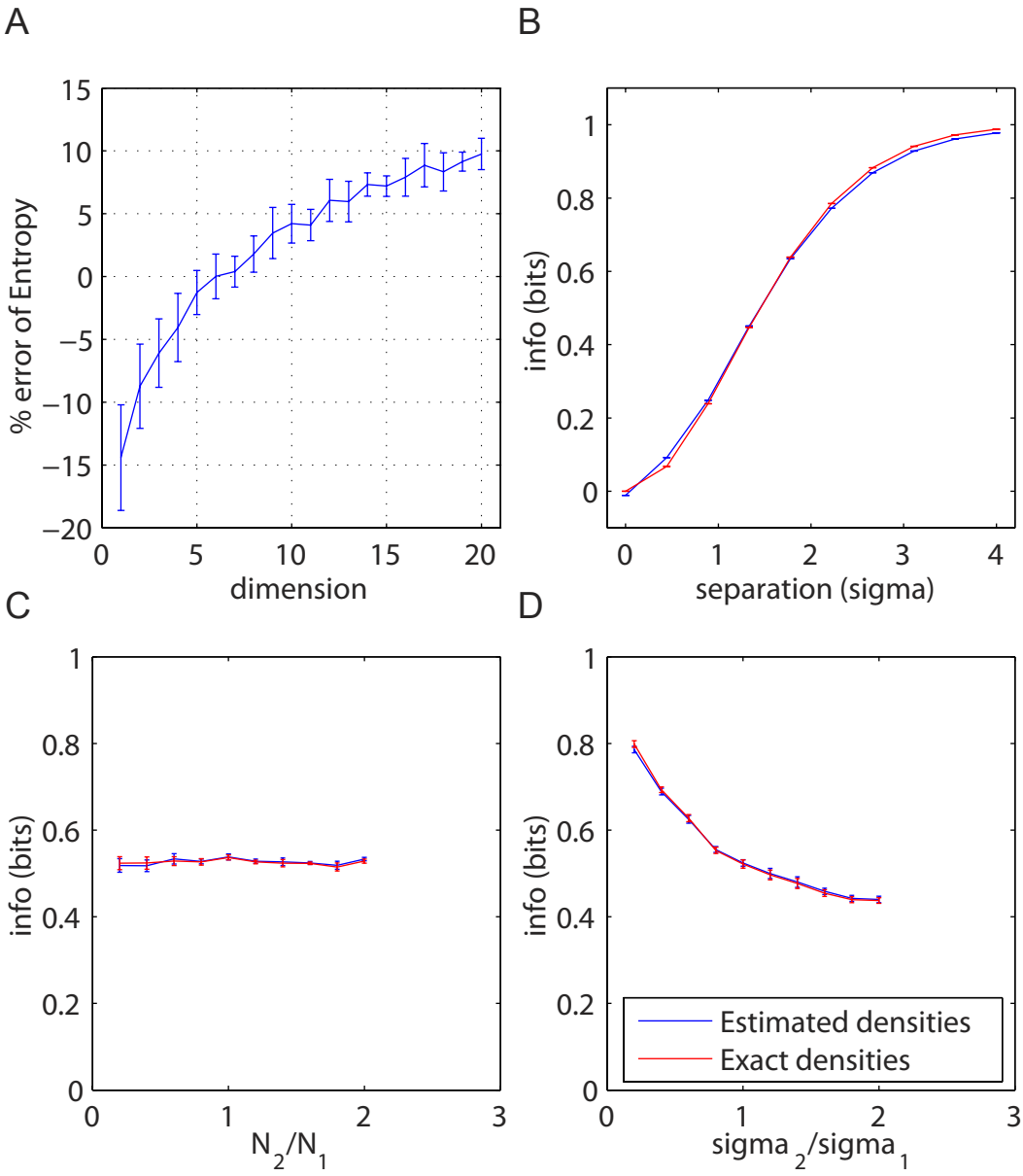

Figure 6.7: Information transfer algorithm validation. (A) Error of our entropy calculation for multivariate Gaussian distribution ($5000 samples$) using k-nearest neighbor density estimator ($k = 10$) for increasing dimension.(B) Sample 2 Gaussian (2D). (C) Information transfer C between two unit-variance Gaussian ($5,000$ samples each), calculated using our entropy estimated algorithm $k = 10$ (green) and exact densities (blue). The insert shows how our density estimate compares to true density for each of the 5000 points within one Gaussians. Vary distance between two normal distributions ($N = 1000, \sigma = 1$). Vary size of one of the two normal distributions ($\mu_1 - \mu_2 = 1.5, \sigma = 1$). Vary sigma of one of the two normal distributions ($\mu_1 - \mu_2 = 1.5, N = 1000$)

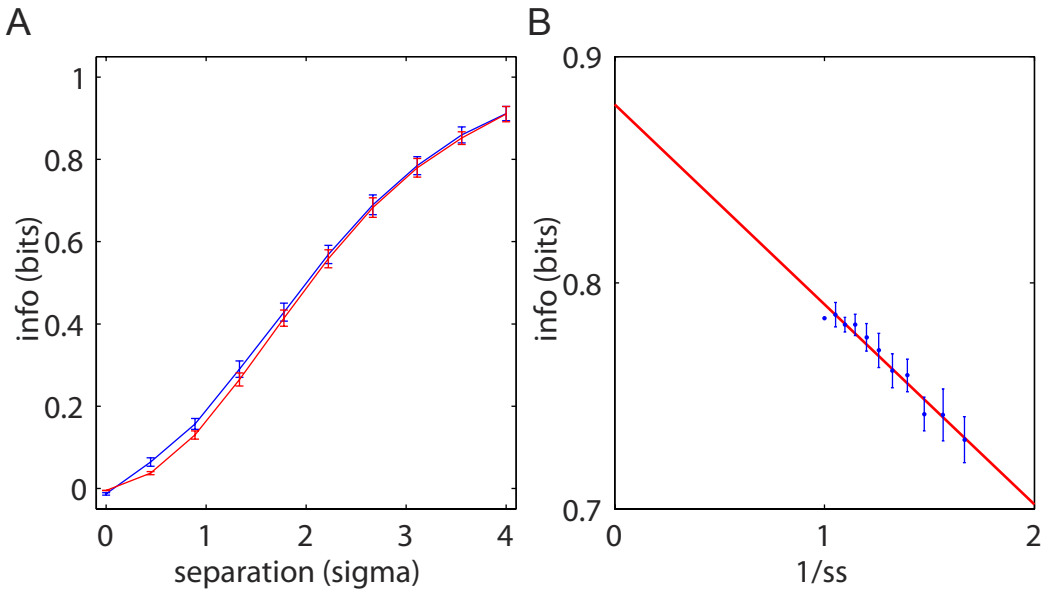

Figure 6.8: (A) Comparison of our binless channel capacity calculation versus binned method (Cheong et al) for 2 Gaussians (Var=1) for increasing separation between their means showed similar results. (B) Applying jackknife sampling to to the data, we can extrapolate the information transfer (C) at infinite sample size.

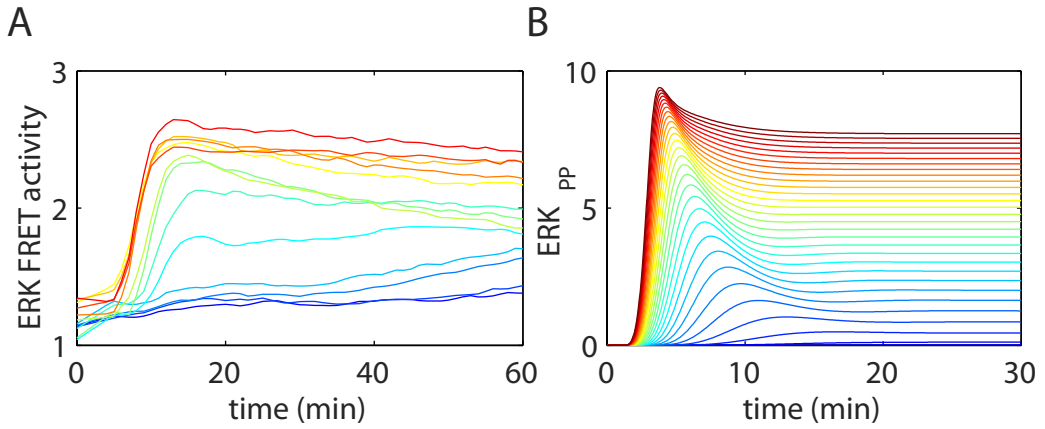

Figure 6.9: Model simulation comparison to experimental Erk FRET trajectories. (A) Mean response of Erk FRET sensor to persistent EGF input. (B) ERKpp response trajectories from simulations of the Erk model (Strum et al) for increasing amounts of RasGTP.

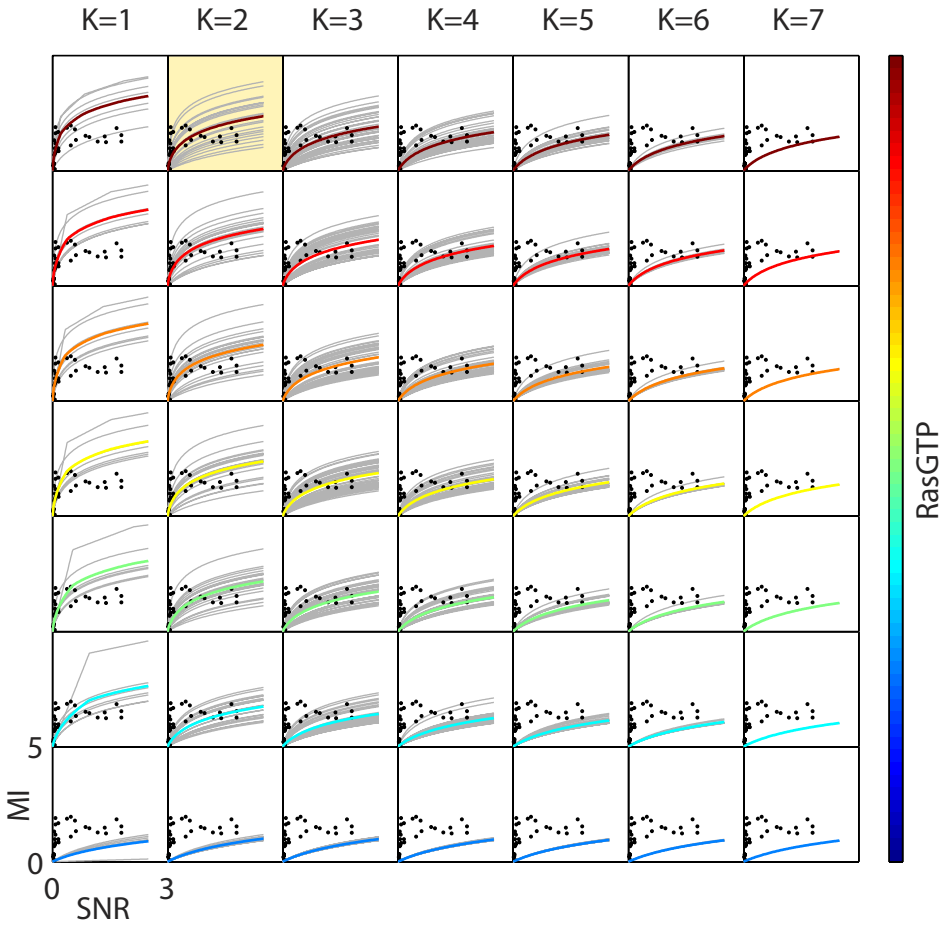

Figure 6.10: Fitting Erk model to dynamic MI vs SNR data using analytic theory approach. We used K number of parameters and RasGTP input level as the two fitting parameters for the model. Grey MI vs SNR curves were obtained for are all combinations of K parameters (out of 7). Colored curves (input RasGTP level) correspond to the mean of all of those combinations of K parameters. The black dots represent experimentally obtained values for vector MI shown in Figure 4, that the model (colored curves) were fit to. Yellow highlighted box corresponds to the best fit for K and RasGTP level.

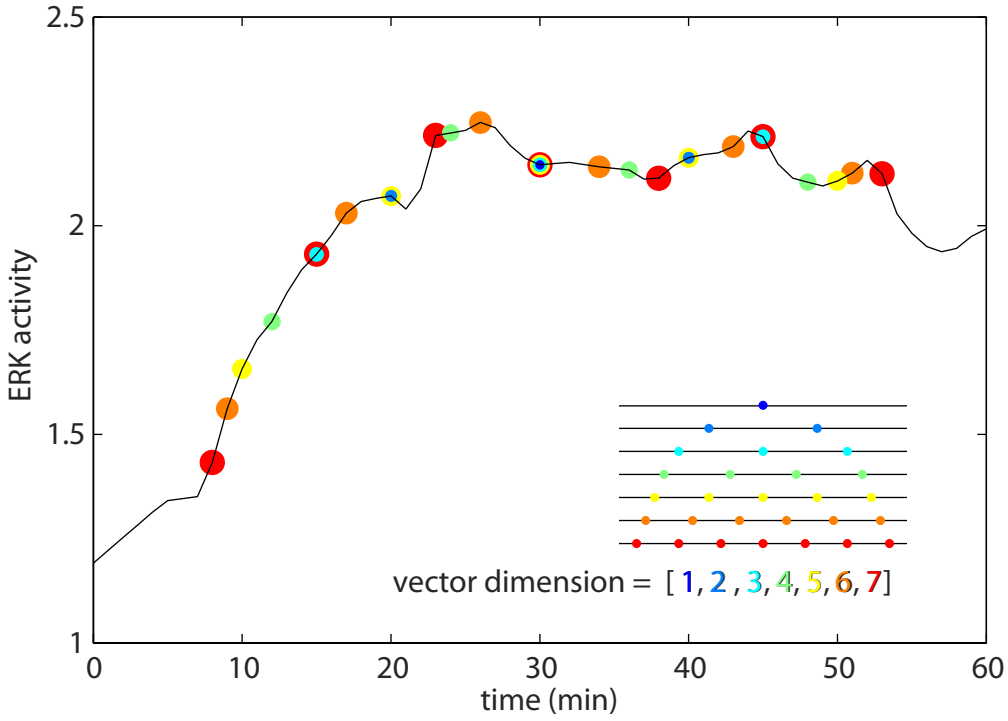

Figure 6.11: Dimension sampling approach for vector dimension in Fig 6.2B.

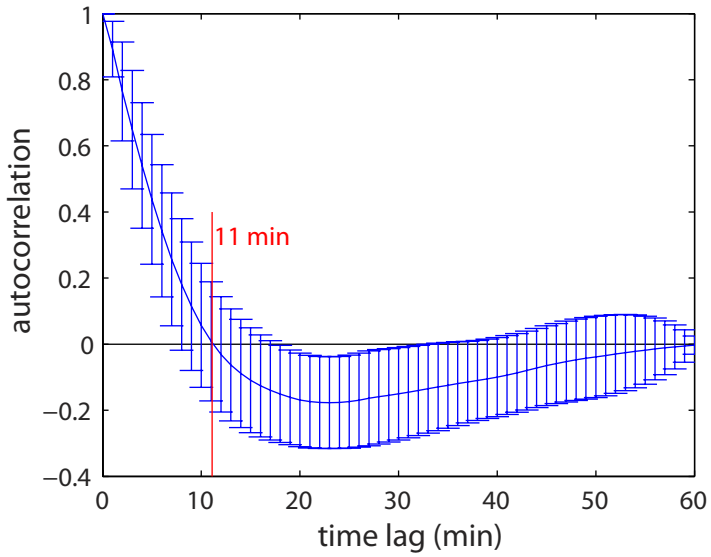

Figure 6.12: Mean autocorrelation of Erk response trajectories (bars represent s.d.). Decay of autocorrelation function shows that on average self-correlation is lost after 11min.

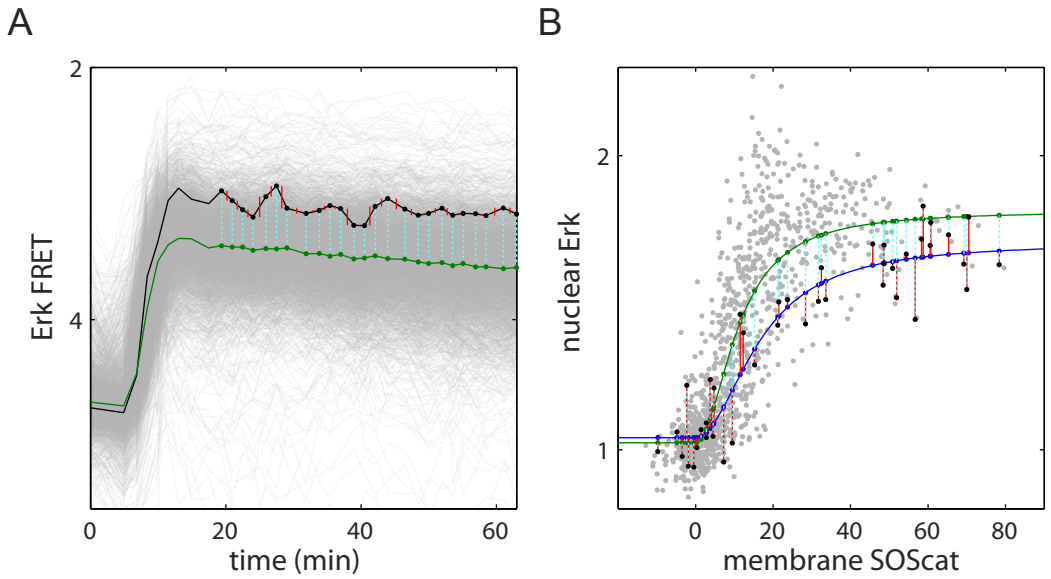

Figure 6.13: Estimate of extrinsic and intrinsic noise in Erk data. (A) Using our data, intrinsic noise was estimated by the mean of squared errors between successive Erk trajectory points (red). Total noise was estimated by the mean of squared errors (cyan) between single Erk trajectory and average of all trajectories (green). Extrinsic ratio was obtained from the difference between total noise and intrinsic noise. The mean ratio of intrinsic to extrinsic noise was estimated to be 0.024. (B) Using Toettcher et al.data, we fit a hill function to the data and calculated the mean squared error between the fit for each cell (intrinsic noise) and between the fit for all points and each cell (total noise). The IER was estimated to be 1.14

References

[1] Albeck, J. G., Mills, G. B., and Brugge, J. S., 2013: Frequency-modulated pulses of erk activity transmit quantitative proliferation signals. *Molecular cell*, **49**(2), 249–261.

[2] Alper, H., and Stephanopoulos, G., 2009: Engineering for biofuels: exploiting innate microbial capacity or importing biosynthetic potential? *Nature Reviews Microbiology*, **7**(10), 715–723.

[3] Andersen, J. B., Sternberg, C., Poulsen, L. K., Bjørn, S. P., Givskov, M., and Molin, S., 1998: New unstable variants of green fluorescent protein for studies of transient gene expression in bacteria. *Applied and environmental microbiology*, **64**(6), 2240–2246.

[4] Anderson, C., Clarke, E., Arkin, A., and Voigt, C., 2006: Environmentally controlled invasion of cancer cells by engineered bacteria. *Journal of Molecular Biology*, **355**(4), 619–627. ISSN 00222836. doi:10.1016/j.jmb.2005.10.076.

[5] Arkin, A., Ross, J., and McAdams, H. H., 1998: Stochastic kinetic analysis of developmental pathway bifurcation in phage λ-infected escherichia coli cells. *Genetics*, **149**(4), 1633–1648.

[6] Ashall, L., Horton, C. A., Nelson, D. E., Paszek, P., Harper, C. V., Sillitoe, K., Ryan, S., Spiller, D. G., Unitt, J. F., Broomhead, D. S., Kell, D. B., Rand, D. A., Se, V., and White, M. R. H., 2009: Pulsatile stimulation determines timing and specificity of NF-κB-dependent transcription. *Science*, **324**(5924), 242–246. doi: 10.1126/science.1164860.

[7] Aurell, E., Brown, S., Johanson, J., and Sneppen, K., 2002: Stability puzzles in phage λ. *Phys. Rev. E*, **65**(5), 051914. doi:10.1103/PhysRevE.65.051914.

[8] Barkal, N., and Leibler, S., 1997: Robustness in simple biochemical networks. *Nature*, **387**(6636), 913–917.

[9] Barken, D., Wang, C. J., Kearns, J., Cheong, R., Hoffmann, A., and Levchenko, A., 2005: Comment on "oscillations in nf-κb signaling control the dynamics of gene expression". *Science*, **308**(5718), 52. doi:10.1126/science.1107904.

[10] Basu, S., Mehreja, R., Thiberge, S., Chen, M., and Weiss, R., 2004: Spatiotemporal control of gene expression with pulse-generating networks. *Proceedings of the National Academy of Sciences of the United States of America*, **101**(17), 6355.

[11] Batchelor, E., Mock, C., Bhan, I., Loewer, A., and Lahav, G., 2008: Recurrent initiation: a mechanism for triggering p53 pulses in response to dna damage. *Molecular cell*, **30**(3), 277.

[12] Bean, B. P., 2007: The action potential in mammalian central neurons. *Nature Reviews Neuroscience*, **8**(6), 451–465.

[13] Behar, M., and Hoffmann, A., 2013: Tunable signal processing through a kinase control cycle: the ikk signaling node. *Biophysical journal*, **105**(1), 231–241.

[14] Bosisio, D., Marazzi, I., Agresti, A., Shimizu, N., Bianchi, M. E., and Natoli, G., 2006: A hyper-dynamic equilibrium between promoter-bound and nucleoplasmic dimers controls nf-[kappa]b-dependent gene activity. *EMBO J*, **25**(4), 798–810. doi:10.1038/sj.emboj.7600977.

[15] Bratsun, D., Volfson, D., Tsimring, L. S., and Hasty, J., 2005: Delay-induced stochastic oscillations in gene regulation. *Proc Natl Acad Sci U S A*, **102**(41), 14593–14598. doi:10.1073/pnas.0503858102.

[16] Brennan, M. D., Cheong, R., and Levchenko, A., 2012: How information theory handles cell signaling and uncertainty. *Science (New York, NY)*, **338**(6105).

[17] Buchler, N. E., and Louis, M., 2008: Molecular titration and ultrasensitivity in regulatory networks. *Journal of molecular biology*, **384**(5), 1106–1119.

[18] Burger, A., Walczak, A. M., and Wolynes, P. G., 2010: Abduction and asylum in the lives of transcription factors. *Proceedings of the National Academy of Sciences*, **107**(9), 4016–4021.

[19] Cai, L., Dalal, C. K., and Elowitz, M. B., 2008: Frequency-modulated nuclear localization bursts coordinate gene regulation. *Nature*, **455**(7212), 485–490.

[20] Cheong, R., Rhee, A., Wang, C. J., Nemenman, I., and Levchenko, A., 2011: Information transduction capacity of noisy biochemical signaling networks. *science*, **334**(6054), 354–358.

[21] Cookson, N. A., Mather, W. H., Danino, T., Mondragón-Palomino, O., Williams, R. J., Tsimring, L. S., and Hasty, J., 2011: Queueing up for enzymatic processing: correlated signaling through coupled degradation. *Molecular systems biology*, **7**(1).

[22] Cookson, N. A., Mather, W. H., Danino, T., Mondragon-Palomino, O., Williams, R. J., Tsimring, L. S., and Hasty, J., 2011: Queueing up for enzymatic processing: Correlated signaling through coupled degradation. *Molecular Systems Biology*.

[23] Cookson, S., Ostroff, N., Pang, W. L., Volfson, D., and Hasty, J., 2005: Monitoring dynamics of single-cell gene expression over multiple cell cycles. *Mol. Syst. Biol.*, **1**, msb4100032–E1–6.

[24] Cover, T. M., and Thomas, J. A., 2006: Elements of information theory 2nd edition.

[25] Cui, W., Morrison, D. C., and Silverstein, R., 2000: Differential tumor necrosis factor alpha expression and release from peritoneal mouse macrophages in vitro in response to proliferating gram-positive versus gram-negative bacteria. *Infection and immunity*, **68**(8), 4422–4429.

[26] Danino, T., Lo, J., Prindle, A., Hasty, J., and Bhatia, S., 2012: In vivo gene expression dynamics of tumor-targeted bacteria. *ACS Synthetic Biology*.

[27] Danino, T., Mondragón-Palomino, O., Tsimring, L., and Hasty, J., 2010: A synchronized quorum of genetic clocks. *Nature*, **463**(7279), 326–330.

[28] Danino, T., Mondragon-Palomino, O., Tsimring, L., and Hasty, J., 2010: A synchronized quorum of genetic clocks. *Nature*, **463**, 326–330.

[29] de Ronde, W., Tostevin, F., and Ten Wolde, P., 2012: Feed-forward loops and diamond motifs lead to tunable transmission of information in the frequency domain. *Physical Review E*, **86**(2), 021913.

[30] de Ronde, W. H., Tostevin, F., and Ten Wolde, P. R., 2010: Effect of feedback on the fidelity of information transmission of time-varying signals. *Physical Review E*, **82**(3), 031914.

[31] Del Vecchio, D., Ninfa, A. J., and Sontag, E. D., 2008: Modular cell biology: retroactivity and insulation. *Molecular systems biology*, **4**(1).

[32] Dolmetsch, R. E., Xu, K., and Lewis, R. S., 1998: Calcium oscillations increase the efficiency and specificity of gene expression. *Nature*, **392**(6679), 933–935.

[33] Doncic, A., Ben-Jacob, E., and Barkai, N., 2006: Noise resistance in the spindle assembly checkpoint. *Molecular systems biology*, **2**(1), 2006.0027.

[34] Dower, S., and Qwarnstrom, E., 2003: Signalling networks, inflammation and innate immunity. *Biochemical Society Transactions*, **31**, 1462–1471.

[35] Duan, F., Curtis, K., and March, J., 2008: Secretion of insulinotropic proteins by commensal bacteria: Rewiring the gut to treat diabetes. *Applied and environmental microbiology*, **74**(23), 7437–7438.

[36] Duan, F., and March, J., 2010: Engineered bacterial communication prevents vibrio cholerae virulence in an infant mouse model. *Science's STKE*, **107**(25), 11260.

[37] Elowitz, M. B., and Leibler, S., 2000: A synthetic oscillatory network of transcriptional regulators. *Nature*, **403**(6767), 335–8.

[38] Elowitz, M. B., Levine, A. J., Siggia, E. D., and Swain, P. S., 2002: Stochastic gene expression in a single cell. *Science*, **297**(5584), 1183–6.

[39] Elowitz, M. B., Levine, A. J., Siggia, E. D., and Swain, P. S., 2002: Stochastic gene expression in a single cell. *Science Signalling*, **297**(5584), 1183–1186.

[40] Ferry, M., Razinkov, I., and Hasty, J., 2011: Microfluidics for synthetic biology from design to execution. *Methods Enzymol*, **497**, 295.

[41] Forbes, N., 2010: Engineering the perfect (bacterial) cancer therapy. *Nature Reviews Cancer*, **10**(11), 785–794.

[42] Forbes, N., Munn, L., Fukumura, D., and Jain, R., 2003: Sparse initial entrapment of systemically injected salmonella typhimurium leads to heterogeneous accumulation within tumors. *Cancer research*, **63**(17), 5188.

[43] Fredriksson, Å., Ballesteros, M., Peterson, C. N., Persson, Ö., Silhavy, T. J., and Nyström, T., 2007: Decline in ribosomal fidelity contributes to the accumulation and stabilization of the master stress response regulator σs upon carbon starvation. *Genes & development*, **21**(7), 862–874.

[44] Friedland, A., Lu, T., Wang, X., Shi, D., Church, G., and Collins, J., 2009: Synthetic gene networks that count. *Science*, **324**(5931), 1199.

[45] Gardner, T. S., Cantor, C. R., and Collins, J. J., 2000: Construction of a genetic toggle switch in *Escherichia coli*. *Nature*, **403**(6767), 339–42.

[46] Gibson, D., Glass, J., Lartigue, C., Noskov, V., Chuang, R., Algire, M., Benders, G., Montague, M., Ma, L., and Moodie, M., 2010: Creation of a bacterial cell controlled by a chemically synthesized genome. *Science*, **329**(5987), 52.

[47] Gillespie, D. T., 1977: Exact stochastic simulation of coupled chemical reactions. *Journal of Physical Chemistry*, **81**(25), 2340–2361.

[48] Gillespie, D. T., 1977: Exact stochastic simulation of coupled chemical-reactions. *J. Phys. Chem.*, **81**(25), 2340–61.

[49] Goldbeter, A., and Koshland, D. E., 1981: An amplified sensitivity arising from covalent modification in biological systems. *Proceedings of the National Academy of Sciences*, **78**(11), 6840–6844.

[50] Griffith, K. L., and Grossman, A. D., 2008: Inducible protein degradation in bacillus subtilis using heterologous peptide tags and adaptor proteins to target substrates to the protease clpxp. *Molecular microbiology*, **70**(4), 1012–1025.

[51] Grünberg, R., and Serrano, L., 2010: Strategies for protein synthetic biology. *Nucleic acids research*, **38**(8), 2663–2675.

[52] Guido, N. J., Wang, X., Adalsteinsson, D., McMillen, D., Hasty, J., Cantor, C. R., Elston, T. C., and Collins, J. J., 2006: A bottom-up approach to gene regulation. *Nature*, **439**(7078), 856–60.

[53] Guo, H., Zhang, J., Inal, C., Nguyen, T., Fruehauf, J. H., Keates, A. C., and Li, C. J., 2011: Targeting tumor gene by shrna-expressing salmonella-mediated rnai. *Gene Ther*, **18**(1), 95–105. doi:10.1038/gt.2010.112.

[54] Hao, N., Budnik, B. A., Gunawardena, J., and O'Shea, E. K., 2013: Tunable signal processing through modular control of transcription factor translocation. *Science*, **339**(6118), 460–464.

[55] Hasty, J., McMillen, D., and Collins, J. J., 2002: Engineered gene circuits. *Nature*, **420**(6912), 224–30.

[56] Hasty, J., Pradines, J., Dolnik, M., and Collins, J. J., 2000: Noise-based switches and amplifiers for gene expression. *Proceedings of the National Academy of Sciences*, **97**(5), 2075–2080.

[57] Heimann, D., and Rosenberg, S., 2003: Continuous intravenous administration of live genetically modified salmonella typhimurium in patients with metastatic melanoma. *Journal of immunotherapy (Hagerstown, Md.: 1997)*, **26**(2), 179.

[58] Hoffman, R., 2011: Tumor-seeking salmonella amino acid auxotrophs. *Current Opinion in Biotechnology*.

[59] Hoffmann, A., Levchenko, A., Scott, M. L., and Baltimore, D., 2002: The IκB-NF-κB signaling module: Temporal control and selective gene activation. *Science*, **298**(5596), 1241–1245. doi:10.1126/science.1071914.

[60] Hoffmann, A., Levchenko, A., Scott, M. L., and Baltimore, D., 2002: The iκb-nf-κb signaling module: temporal control and selective gene activation. *Science*, **298**(5596), 1241–1245.

[61] Hohmann, E., Oletta, C., and Miller, S., 1996: Evaluation of a phop/phoq-deleted, aroa-deleted live oral salmonella typhi vaccine strain in human volunteers. *Vaccine*, **14**(1), 19–24.

[62] Hooshangi, S., Thiberge, S., and Weiss, R., 2005: Ultrasensitivity and noise propagation in a synthetic transcriptional cascade. *Proceedings of the National Academy of Sciences of the United States of America*, **102**(10), 3581–3586.

[63] Jasny, B., and Zahn, L., 2011: A celebration of the genome, part i. *Science*, **331**(6017), 546.

[64] Jia, L. J., Wei, D. P., Sun, Q. M., Huang, Y., Wu, Q., and Hua, Z. C., 2007: Oral delivery of tumor-targeting salmonella for cancer therapy in murine tumor models. *Cancer Sci*, **98**(7), 1107–1112. doi:10.1111/j.1349-7006.2007.00503.x.

[65] Kasinskas, R. W., and Forbes, N. S., 2006: Salmonella typhimurium specifically chemotax and proliferate in heterogeneous tumor tissue in vitro. *Biotechnol Bioeng*, **94**(4), 710–721. doi:10.1002/bit.20883.

[66] Kasinskas, R. W., and Forbes, N. S., 2007: Salmonella typhimurium lacking ribose chemoreceptors localize in tumor quiescence and induce apoptosis. *Cancer Res*, **67**(7), 3201–3209. doi:10.1158/0008-5472.CAN-06-2618.

[67] Kearns, J. D., Basak, S., Werner, S. L., Huang, C. S., and Hoffmann, A., 2006: IκBϵ provides negative feedback to control NF-κB oscillations, signaling dynamics, and inflammatory gene expression. *The Journal of Cell Biology*, **173**(5), 659–664. doi:10.1083/jcb.200510155.

[68] Keiler, K., Waller, P., and Sauer, R., 1996: Role of a peptide tagging system in degradation of proteins synthesized from damaged messenger rna. *Science*, **271**(5251), 990.

[69] Khalil, A., and Collins, J., 2010: Synthetic biology: applications come of age. *Nature Reviews Genetics*, **11**(5), 367–379.

[70] Kim, J.-R., Yoon, Y., and Cho, K.-H., 2008: Coupled feedback loops form dynamic motifs of cellular networks. *Biophysical Journal*, **94**(2), 359 – 365. ISSN 0006-3495. doi:DOI:10.1529/biophysj.107.105106.

[71] Kolnik, M., Tsimring, L. S., and Hasty, J., 2012: Vacuum-assisted cell loading enables shear-free mammalian microfluidic culture. *Lab on a chip*, **12**(22), 4732–4737.

[72] Komatsu, N., Aoki, K., Yamada, M., Yukinaga, H., Fujita, Y., Kamioka, Y., and Matsuda, M., 2011: Development of an optimized backbone of fret biosensors for kinases and gtpases. *Molecular biology of the cell*, **22**(23), 4647–4656.

[73] Krishna, S., Jensen, M. H., and Sneppen, K., 2006: Minimal model of spiky oscillations in nf-κb signaling. *Proceedings of the National Academy of Sciences*, **103**(29), 10840–10845.

[74] Kruse, K., Pantazis, P., Bollenbach, T., Jlicher, F., and Gonzlez-Gaitn, M., 2004: Dpp gradient formation by dynamin-dependent endocytosis: receptor trafficking and the diffusion model. *Development*, **131**(19), 4843–4856. doi:10.1242/dev.01335.

[75] Lahav, G., Rosenfeld, N., Sigal, A., Geva-Zatorsky, N., Levine, A. J., Elowitz, M. B., and Alon, U., 2004: Dynamics of the p53-mdm2 feedback loop in individual cells. *Nature genetics*, **36**(2), 147–150.

[76] Lander, A. D., Nie, Q., and Wan, F. Y., 2002: Do morphogen gradients arise by diffusion? *Developmental Cell*, **2**(6), 785 – 796. ISSN 1534-5807. doi:DOI:10.1016/S1534-5807(02)00179-X.

[77] Lipshtat, A., Perets, H. B., Balaban, N. Q., and Biham, O., 2005: Modeling of negative autoregulated genetic networks in single cells. *Gene*, **347**(2), 265–271.

[78] Lm, Z., Luo, X., Feng, M., Li, Z., Ittensohn, M., Trailsmith, M., Bermudes, D., Lin, S., and King, I., 2001: Tumor amplified protein expression therapy: Salmonella as a tumor-selective protein delivery vector. *Oncology Research Featuring Preclinical and Clinical Cancer Therapeutics*, **12**(3), 127–135.

[79] Longo, D. M., Selimkhanov, J., Kearns, J. D., Hasty, J., Hoffmann, A., and Tsimring, L. S., 2013: Dual delayed feedback provides sensitivity and robustness to the nf-κb signaling module. *PLoS computational biology*, **9**(6), e1003112.

[80] Lou, C., Stanton, B., Chen, Y.-J., Munsky, B., and Voigt, C. A., 2012: Ribozyme-based insulator parts buffer synthetic circuits from genetic context. *Nature biotechnology*, **30**(11), 1137–1142.

[81] Low, K., Ittensohn, M., Le, T., Platt, J., Sodi, S., Amoss, M., Ash, O., Carmichael, E., Chakraborty, A., Fischer, J., Lin, S., Luo, X., Miller, S., Zheng, L., King, I., Pawelek, J., and Bermudes, D., 1999: Lipid a mutant salmonella with suppressed virulence and tnfα induction retain tumor-targeting in vivo. *Nature biotechnology*, **17**(1), 37–41.

[82] Lu, T., and Collins, J., 2007: Dispersing biofilms with engineered enzymatic bacteriophage. *Proceedings of the National Academy of Sciences*, **104**(27), 11197.

[83] Lu, T., and Collins, J., 2009: Engineered bacteriophage targeting gene networks as adjuvants for antibiotic therapy. *Proceedings of the National Academy of Sciences*, **106**(12), 4629.

[84] Ma, L., Wagner, J., Rice, J. J., Hu, W., Levine, A. J., and Stolovitzky, G. A., 2005: A plausible model for the digital response of p53 to dna damage. *Proceedings of the National Academy of Sciences of the United States of America*, **102**(40), 14266–14271.

[85] Mather, W., Bennett, M., Hasty, J., and Tsimring, L., 2009: Delay-Induced Degrade-and-Fire Oscillations in Small Genetic Circuits. *Biophys. J Phys Rev Lett*, **102**, 068105.

[86] Mather, W., Bennett, M. R., Hasty, J., and Tsimring, L. S., 2009: Delay-induced degrade-and-fire oscillations in small genetic circuits. *Physical review letters*, **102**(6), 068105.

[87] Matzas, M., Stähler, P., Kefer, N., Siebelt, N., Boisguérin, V., Leonard, J., Keller, A., Stähler, C., Häberle, P., Gharizadeh, B., Barzadeh, F., and Church, G., 2010: High-fidelity gene synthesis by retrieval of sequence-verified dna identified using high-throughput pyrosequencing. *Nature biotechnology*, **28**(12), 1291–1294.

[88] McClelland, M., Sanderson, K., Spieth, J., Clifton, S., Latreille, P., Courtney, L., Porwollik, S., Ali, J., Dante, M., Du, F., Hou, S., Layman, D., Leonard, S., Nguyen, C., Scott, K., Holmes, A., Grewal, N., Mulvaney, E., Ryan, E., Sun, H., Florea, L., Miller, W., Stoneking, T., Nhan, M., Waterston, R., and Wilson, R., 2001: Complete genome sequence of salmonella enterica serovar typhimurium lt2. *Nature*, **413**(6858), 852–856. ISSN 0028-0836. doi:10.1038/35101614.

[89] McGinness, K. E., Baker, T. A., and Sauer, R. T., 2006: Engineering controllable protein degradation. *Molecular cell*, **22**(5), 701–707.

[90] Medzhitov, R., and Janeway Jr, C. A., 1997: Innate immunity: impact on the adaptive immune response. *Current opinion in immunology*, **9**(1), 4–9.

[91] Mehta, P., Goyal, S., Long, T., Bassler, B. L., and Wingreen, N. S., 2009: Information processing and signal integration in bacterial quorum sensing. *Molecular systems biology*, **5**(1).

[92] Merrikh, H., Ferrazzoli, A. E., Bougdour, A., Olivier-Mason, A., and Lovett, S. T., 2009: A dna damage response in escherichia coli involving the alternative sigma factor, rpos. *Proceedings of the National Academy of Sciences*, **106**(2), 611–616.

[93] Merril, C., Scholl, D., and Adhya, S., 2003: The prospect for bacteriophage therapy in western medicine. *Nature Reviews Drug Discovery*, **2**(6), 489–497.

[94] Mika, F., and Hengge, R., 2005: A two-component phosphotransfer network involving arcb, arca, and rssb coordinates synthesis and proteolysis of σs (rpos) in e. coli. *Genes & development*, **19**(22), 2770–2781.

[95] Mondragon-Palomino, O., Danino, T., Selimkhanov, J., Tsimring, L., and Hasty, J., 2011: Entrainment of a population of synthetic genetic oscillators. *Science*, **333**.

[96] Moon, T. S., Lou, C., Tamsir, A., Stanton, B. C., and Voigt, C. A., 2012: Genetic programs constructed from layered logic gates in single cells. *Nature*, **491**(7423), 249–253.

[97] Mukherji, S., Ebert, M. S., Zheng, G. X., Tsang, J. S., Sharp, P. A., and van Oudenaarden, A., 2011: Micrornas can generate thresholds in target gene expression. *Nature genetics*, **43**(9), 854–859.

[98] Murphy, K. M., Travers, P., and Walport, M., 2012: *Janeway's immunobiology*, volume 7. Garland Science New York, NY, USA.

[99] Nagakura, C., Hayashi, K., Zhao, M., Yamauchi, K., Yamamoto, N., Tsuchiya, H., Tomita, K., Bouvet, M., and Hoffman, R. M., 2009: Efficacy of a genetically-modified salmonella typhimurium in an orthotopic human pancreatic cancer in nude mice. *Anticancer Res*, **29**(6), 1873–1878.

[100] Nandagopal, N., and Elowitz, M. B., 2011: Synthetic biology: Integrated gene circuits. *Science*, **333**(6047), 1244–1248.

[101] Nelson, D. E., Ihekwaba, A. E. C., Elliott, M., Johnson, J. R., Gibney, C. A., Foreman, B. E., Nelson, G., See, V., Horton, C. A., Spiller, D. G., Edwards, S. W., McDowell, H. P., Unitt, J. F., Sullivan, E., Grimley, R., Benson, N., Broomhead, D., Kell, D. B., and White, M. R. H., 2004: Oscillations in NF-κB signaling control the dynamics of gene expression. *Science*, **306**(5696), 704–708. doi: 10.1126/science.1099962.

[102] Nemunaitis, J., Cunningham, C., Senzer, N., Kuhn, J., Cramm, J., Litz, C., Cavagnolo, R., Cahill, A., Clairmont, C., and Sznol, M., 2003: Pilot trial of genetically modified, attenuated salmonella expressing the e. coli cytosine deaminase gene in refractory cancer patients. *Cancer gene therapy*, **10**(10), 737–744.

[103] Nguyen, V. H., Kim, H. S., Ha, J. M., Hong, Y., Choy, H. E., and Min, J. J., 2010: Genetically engineered salmonella typhimurium as an imageable therapeutic probe for cancer. *Cancer Res*, **70**(1), 18–23. doi:10.1158/0008-5472. CAN-09-3453.

[104] Paszek, P., Ryan, S., Ashall, L., Sillitoe, K., Harper, C., Spiller, D., Rand, D., and White, M., 2010: Population robustness arising from cellular heterogeneity. *Proc Natl Acad Sci U S A*, **107**(25), 11644–11649. doi:10.1073/pnas.0913798107.

[105] Pawelek, J., Low, K., and Bermudes, D., 1997: Tumor-targeted salmonella as a novel anticancer vector. *Cancer research*, **57**(20), 4537.

[106] Prindle, A., Samayoa, P., Razinkov, I., Danino, T., Tsimring, L. S., and Hasty, J., 2011: A sensing array of radically coupled genetic/biopixels/'. *Nature*, **481**(7379), 39–44.

[107] Prindle, A., Selimkhanov, J., Danino, T., Samayoa, P., Goldberg, A., Bhatia, S. N., and Hasty, J., 2012: Genetic circuits in salmonella typhimurium. *ACS synthetic biology*, **1**(10), 458–464.

[108] Projan, S., 2004: Phage-inspired antibiotics? *Nature biotechnology*, **22**(2), 167–168.

[109] Pruteanu, M., and Hengge-Aronis, R., 2002: The cellular level of the recognition factor rssb is rate-limiting for σs proteolysis: implications for rssb regulation and signal transduction in σs turnover in escherichia coli. *Molecular microbiology*, **45**(6), 1701–1713.

[110] Purvis, J. E., Karhohs, K. W., Mock, C., Batchelor, E., Loewer, A., and Lahav, G., 2012: p53 dynamics control cell fate. *Science*, **336**(6087), 1440–1444.

[111] Purvis, J. E., and Lahav, G., 2013: Encoding and decoding cellular information through signaling dynamics. *Cell*, **152**(5), 945–956.

[112] Ramsey, S. A., Smith, J. J., Orrell, D., Marelli, M., Petersen, T. W., de Atauri, P., Bolouri, H., and Aitchison, J. D., 2006: Dual feedback loops in the gal regulon suppress cellular heterogeneity in yeast. *Nature genetics*, **38**(9), 1082–1087.

[113] Rao, S., Hu, S., McHugh, L., Lueders, K., Henry, K., Zhao, Q., Fekete, R., Kar, S., Adhya, S., and Hamer, D., 2005: Toward a live microbial microbicide for hiv: commensal bacteria secreting an hiv fusion inhibitor peptide. *Proceedings of the National Academy of Sciences of the United States of America*, **102**(34), 11993.

[114] Raser, J. M., and O'Shea, E. K., 2004: Control of stochasticity in eukaryotic gene expression. *Science*, **304**(5678), 1811–4.

[115] Regot, S., Macia, J., Conde, N., Furukawa, K., Kjellén, J., Peeters, T., Hohmann, S., De Nadal, E., Posas, F., and Solé, R., 2010: Distributed biological computation with multicellular engineered networks. *Nature*, **469**(7329), 207–211.

[116] Rhee, A., Cheong, R., and Levchenko, A., 2012: The application of information theory to biochemical signaling systems. *Physical biology*, **9**(4), 045011.

[117] Richard, C. Y., Pesce, C. G., Colman-Lerner, A., Lok, L., Pincus, D., Serra, E., Holl, M., Benjamin, K., Gordon, A., and Brent, R., 2008: Negative feedback that improves information transmission in yeast signalling. *Nature*, **456**(7223), 755–761.

[118] Rosenfeld, N., and Alon, U., 2003: Response delays and the structure of transcription networks. *Journal of molecular biology*, **329**(4), 645–654.

[119] Rosenfeld, N., Young, J. W., Alon, U., Swain, P. S., and Elowitz, M. B., 2005: Gene regulation at the single-cell level. *Science Signalling*, **307**(5717), 1962–1965.

[120] Ruder, W. C., Lu, T., and Collins, J. J., 2011: Synthetic biology moving into the clinic. *Science*, **333**(6047), 1248–1252.

[121] Santos, S. D., Verveer, P. J., and Bastiaens, P. I., 2007: Growth factor-induced mapk network topology shapes erk response determining pc-12 cell fate. *Nature cell biology*, **9**(3), 324–330.

[122] Savageau, M. A., 1974: Optimal design of feedback control by inhibition. *Journal of Molecular Evolution*, **4**(2), 139–156.

[123] Schreiber, T., 1999: Interdisciplinary application of nonlinear time series methods. *Physics reports*, **308**(1), 1–64.

[124] Shimojo, H., Ohtsuka, T., and Kageyama, R., 2008: Oscillations in notch signaling regulate maintenance of neural progenitors. *Neuron*, **58**(1), 52–64.

[125] Siuti, P., Yazbek, J., and Lu, T. K., 2013: Synthetic circuits integrating logic and memory in living cells. *Nature biotechnology*.

[126] Sohka, T., Heins, R., Phelan, R., Greisler, J., Townsend, C., and Ostermeier, M., 2009: An externally tunable bacterial band-pass filter. *Proceedings of the National Academy of Sciences*, **106**(25), 10135.

[127] Stricker, J., Cookson, S., Bennett, M., Mather, W., Tsimring, L., and Hasty, J., 2008: A fast, robust and tunable synthetic gene oscillator. *Nature*, **456**(7221), 516–519.

[128] Stricker, J., Cookson, S., Bennett, M. R., Mather, W. H., Tsimring, L. S., and Hasty, J., 2008: A fast, robust and tunable synthetic gene oscillator. *Nature*, **456**(7221), 516–9.

[129] Strogatz, S., 2001: Nonlinear dynamics and chaos: with applications to physics, biology, chemistry and engineering.

[130] Sturm, O. E., Orton, R., Grindlay, J., Birtwistle, M., Vyshemirsky, V., Gilbert, D., Calder, M., Pitt, A., Kholodenko, B., and Kolch, W., 2010: The mammalian mapk/erk pathway exhibits properties of a negative feedback amplifier. *Science signaling*, **3**(153), ra90.

[131] Swain, P. S., Elowitz, M. B., and Siggia, E. D., 2002: Intrinsic and extrinsic contributions to stochasticity in gene expression. *Proceedings of the National Academy of Sciences*, **99**(20), 12795–12800.

[132] Tabor, J., Salis, H., Simpson, Z., Chevalier, A., Levskaya, A., Marcotte, E., Voigt, C., and Ellington, A., 2009: A synthetic genetic edge detection program. *Cell*, **137**(7), 1272–1281. ISSN 0092-8674.

[133] Tamsir, A., Tabor, J., and Voigt, C., 2010: Robust multicellular computing using genetically encoded nor gates and chemical/wires/'. *Nature*, **469**(7329), 212–215.

[134] Tergaonkar, V., 2006: NFκB pathway: A good signaling paradigm and therapeutic target. *The International Journal of Biochemistry & Cell Biology*, **38**(10), 1647 – 1653. ISSN 1357-2725. doi:DOI:10.1016/j.biocel.2006.03.023.

[135] Thattai, M., and Van Oudenaarden, A., 2001: Intrinsic noise in gene regulatory networks. *Proceedings of the National Academy of Sciences*, **98**(15), 8614–8619.

[136] Tigges, M., Marquez-Lago, T., Stelling, J., and Fussenegger, M., 2009: A tunable synthetic mammalian oscillator. *Nature*, **457**(7227), 309–312.

[137] Tkačik, G., Callan, C. G., and Bialek, W., 2008: Information flow and optimization in transcriptional regulation. *Proceedings of the National Academy of Sciences*, **105**(34), 12265–12270.

[138] Toettcher, J. E., Weiner, O. D., and Lim, W. A., 2013: Using optogenetics to interrogate the dynamic control of signal transmission by the ras/erk module. *Cell*, **155**(6), 1422–1434.

[139] Toso, J., Gill, V., Hwu, P., Marincola, F., Restifo, N., Schwartzentruber, D., Sherry, R., Topalian, S., Yang, J., Stock, F., Freezer, L., Morton, K., Seipp, C., Haworth, L., Mavroukakis, S., White, D., MacDonald, S., Mao, J., Sznol, M., and Rosenberg, S., 2002: Phase i study of the intravenous administration of attenuated salmonella typhimurium to patients with metastatic melanoma. *Journal of clinical oncology*, **20**(1), 142–152.

[140] Tostevin, F., and Ten Wolde, P. R., 2010: Mutual information in time-varying biochemical systems. *Physical Review E*, **81**(6), 061917.

[141] Tyson, J. J., Chen, K. C., and Novak, B., 2003: Sniffers, buzzers, toggles and blinkers: dynamics of regulatory and signaling pathways in the cell. *Curr. Opin. Cell Biol.*, **15**(2), 221–31.

[142] Uda, S., Saito, T. H., Kudo, T., Kokaji, T., Tsuchiya, T., Kubota, H., Komori, Y., Ozaki, Y.-i., and Kuroda, S., 2013: Robustness and compensation of information transmission of signaling pathways. *Science*, **341**(6145), 558–561.

[143] Vogelstein, B., Lane, D., and Levine, A. J., 2000: Surfing the p53 network. *Nature*, **408**(6810), 307–310. doi:10.1038/35042675.

[144] Volfson, D., Marciniak, J., Blake, W. J., Ostroff, N., Tsimring, L. S., and Hasty, J., 2005: Origins of extrinsic variability in eukaryotic gene expression. *Nature*, **439**(7078), 861–864.

[145] Von Dassow, G., Meir, E., Munro, E. M., and Odell, G. M., 2000: The segment polarity network is a robust developmental module. *Nature*, **406**(6792), 188–192.

[146] Walczak, A. M., Tkačik, G., and Bialek, W., 2010: Optimizing information flow in small genetic networks. ii. feed-forward interactions. *Physical Review E*, **81**(4), 041905.

[147] Waltermann, C., and Klipp, E., 2011: Information theory based approaches to cellular signaling. *Biochimica et Biophysica Acta (BBA)-General Subjects*, **1810**(10), 924–932.

[148] Weber, W., and Fussenegger, M., 2011: Emerging biomedical applications of synthetic biology. *Nature Reviews Genetics*.

[149] Werner, S. L., Kearns, J. D., Zadorozhnaya, V., Lynch, C., ODea, E., Boldin, M. P., Ma, A., Baltimore, D., and Hoffmann, A., 2008: Encoding NF-κB temporal control in response to TNF: distinct roles for the negative regulators IκBα and a20. *Genes & Development*, **22**(15), 2093–2101. doi:10.1101/gad.1680708.

[150] Wu, D., Hugenholtz, P., Mavromatis, K., Pukall, R., Dalin, E., Ivanova, N., Kunin, V., Goodwin, L., Wu, M., Tindall, B., Hooper, S., Pati, A., Lykidis, A., Spring, S., Anderson, I., D'haeseleer, P., Zemla, A., Singer, M., Lapidus, A., Nolan, M., Copeland, A., Han, C., Chen, F., Cheng, J., Lucas, S., Kerfeld, C., Lang, E., Gronow, S., Chain, P., Bruce, D., Rubin, E., Kyrpides, N., Klenk, H., and Eisen, J., 2009: A phylogeny-driven genomic encyclopaedia of bacteria and archaea. *Nature*, **462**(7276), 1056–1060.

[151] Xiang, S., Fruehauf, J., and Li, C. J., 2006: Short hairpin rna-expressing bacteria elicit rna interference in mammals. *Nat Biotechnol*, **24**(6), 697–702. doi:10.1038/nbt1211.

[152] Xie, Z., Wroblewska, L., Prochazka, L., Weiss, R., and Benenson, Y., 2011: Multi-input rnai-based logic circuit for identification of specific cancer cells. *Science Signaling*, **333**(6047), 1307.

[153] Ye, H., Daoud-El Baba, M., Peng, R. W., and Fussenegger, M., 2011: A synthetic optogenetic transcription device enhances blood-glucose homeostasis in mice. *Science*, **332**(6037), 1565–1568. doi:10.1126/science.1203535.

[154] You, L., Cox, I., R. S., Weiss, R., and Arnold, F. H., 2004: Programmed population control by cell-cell communication and regulated killing. *Nature*, **428**(6985), 868–71.