Stochastic Processes in T-cell Signaling

by

Ming Yang

B.Eng. (First Class Honors), National University of Singapore (2006)
M.S.CEP., Massachusetts Institute of Technology (2008)

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

Doctor of Philosophy in Chemical Engineering

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 2012

Author .....
Department of Chemical Engineering
June 4, 2012

Certified by...
Arup K. Chakraborty
Robert T. Haslam Professor of Chemical Engineering
Professor of Chemistry and Biological Engineering
Thesis Supervisor

Accepted by .....
Patrick S. Doyle
Professor of Chemical Engineering
Chairman of the Committee for Graduate Students

Stochastic Processes in T-cell Signaling

by

Ming Yang

Submitted to the Department of Chemical Engineering
on June 4, 2012, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy in Chemical Engineering

Abstract

T cells are orchestrators for adaptive immunity. Antigen recognition by T cells is mediated by the interactions between T-cell receptors (TCRs) and peptide-MHC (pMHC) molecules. How T cells can translate stimulatory external cues (e.g., TCR-pMHC interactions where the peptides are derived from foreign proteins) to functional responses (e.g., proliferation), while not responding to self-pMHC has been a puzzle for decades. The ability to discriminate foreign antigens from self antigens demands extraordinary intricacy for the design and operation of signaling pathways. This problem is a special challenge since fluctuations at the microscopic level are ubiquitous in biochemical networks, due to stochastic nature of reactions and uncertainties in protein expressions. The prevalence of noise imposes further challenges for T cells to deliver biological functions reliably. The overarching theme of this thesis is to understand the role of stochasticity in T-cell signaling.

Four problems have been selected for presentation in this thesis:

1. Fluctuation-driven transitions can drive cellular systems out of stable states and lead to spurious responses. We proposed a theoretical and computational framework to identify key reactions or species that are responsible for regulating such stochastic transitions. The identification of these critical components for network stability not only pinpoints key deleterious protein mutations, but also helps intelligently select drug targets. The semi-analytical method we derived using large-deviation theory and calculus of variations agrees well with computational costly brute-force simulations. Additionally, our framework unveils qualitative characteristics of key reactions regulating stochastic transitions. We believe that we have developed the first method to carry out fully stochastic sensitivity analyses using analytical calculations.

2. In collaboration with Dr. Jeroen Roose's lab at UCSF, we investigated the roles of RasGRP and SOS in the activation of ERK and P38 MAP kinase pathways. We extended established computational models developed in our lab and predicted that SOS' allosteric pocket is important for the magnitude and bimodal pattern of ERK

activation, which was confirmed experimentally. The synergy between computational modeling and experimental studies enabled us to propose mechanistic models that incorporate features such as co-operativity and non-linearity thresholding to study P38 activation. These models were consistent with the experimental findings that SOS is preferentially more important than RasGRP for Rac-P38 activation and SOS' allosteric pocket has little effect on Rac-P38 pathways, and also generated numerous experimentally testable hypotheses.

3. While rare events, such as escapes from stable basins, take a long time (waiting time) to occur, they take little time to complete once they have started. We showed that for Markov processes characterized by detailed balance, successful transitions, on average, complete exactly as quickly as transitions in the opposite (non-rare) direction. We first provide a general proof by invoking time reversibility, and then elaborate the proof by considering two specific dynamics, namely, continuous-time Markov Chains with detailed balance and one-dimensional Langevin Dynamics. We employ ideas from measure theory and stochastic calculus. We conclude that rare events, once they happen, happen quickly, and speculate about extensions to non-equilibrium systems, such as viral escape.

4. While microscopic fluctuations complicate reliable functioning of biochemical networks, stochastic noise also offers enormous information about the underlying network that generates such noise. We present an effort to exploit the non-random structure of random noise for network topology identification. In particular, we applied linear noise approximations to two three-node network motifs, namely, incoherent feed-forward loop (IFF) and negative feedback loop (NFB), and obtained correlation functions governing the fluctuations of species copy numbers at steady state. We identified two signatures that can be used to discriminate IFF from NFB. This endeavor represents a first step toward understanding how, and to what extent, time-series data with fine time and length resolutions can be used to infer network structures.

Thesis Supervisor: Arup K. Chakraborty
Title: Robert T. Haslam Professor of Chemical Engineering
Professor of Chemistry and Biological Engineering

To those who have loved me and whom I have loved

Acknowledgments

Reflecting upon the years I spent in Cambridge, Massachusetts, I am full of gratitude and thanksgiving. First, I would like to extend my deepest appreciation to my thesis advisor, Prof. Arup K. Chakraborty. In the past six years, Arup has been not only my academic teacher, whose devotion to and passion for science have greatly shaped my intellectual development, but also my life mentor, whose value and philosophy have made a life-long impact on me. I thank Arup for his far-reaching vision of applying physical sciences to biomedical disciplines and his generosity of encouraging me to pursue fundamental sciences and mathematics beyond traditional chemical engineering, which enables me to integrate knowledge and techniques from diverse disciplines to study immunological systems. While I feel regretful that academia is not my immediate career goal, I do hope that at some point down my life trajectory, I can contribute all I have inherited and learned here to science, as Arup has always hoped and encouraged.

I am blessed to have Prof. Paul I. Barton and Prof. Darrell J. Irvine on my thesis committee. Their diverse expertise and perspectives have greatly enriched my educational and research experience; their willingness to accommodate my tight schedule has enabled my defense to proceed in a timely manner.

During my PhD years at MIT, I had the fortune to work closely with three renowned experimental biologists, namely, Prof. Ed Palmer at University Hospital Basel and Prof. Arthur Weiss and Prof. Jeroen Roose, both at UCSF. I thank these collaborators for tolerating my naivety and teaching me biology tirelessly, and more importantly, for being a model of pursuing science with single-heartedness and caring for junior researchers with kindness.

This thesis would not be possible without the support, both emotionally and professionally, from the former and present members of the Chakraborty group at MIT. It was my privilege to work with each one of them as a team, and specifically, I would like to give my wholehearted thanks to Dr. Steve M. Abel, Dr. Andrew Ferguson, and Dr. Christopher C. Govern. Steve, words are insufficient to express

how grateful I am to have you as "the big brother" in the group. Thank you for accompanying me in walking through the darkest valley in my PhD journey. Andy, thank you for the life lessons you shared with me at Meadhall. I wish you great success as a faculty member at UIUC and sincerely wish our life paths can cross again in the future. Finally, Chris, my long-time officemate and collaborator, knowing and working with you has been one of my highlights at MIT. I could never come to this stage without you motivating, guiding, and comforting me. My thoughts and prayers will be always with you, no matter where you are.

Many of my intellectual stimuli have come from brilliant minds outside the Department. My special thanks go to Prof. Leonid Kogan from the MIT Sloan School of Management, who taught me stochastic calculus and change of measure, both being major breakthroughs in my intellectual understanding and playing critical roles in my PhD research; also Prof. Anna Mikusheva from the Department of Economics at MIT, who educated me on time series analysis and offered generous help during my job hunting. I would like to thank Mr. Chong Wang, my roommate and PhD candidate in the MIT condensed matter physics group, Dr. Hung-An Chang at the MIT Computer Science and Artificial Intelligence Laboratory (CSAIL), Mr. Ying-zong Huang at the MIT Research Laboratory of Electronics (RLE), and Ms. Mengdi Wang at the Department of Electrical Engineering and Computer Science at MIT. Chong, your ingenious insight into physics and extraordinary command of mathematics have transformed my PhD research experience from experimentally-driven computational modeling to theory-based framework development. I hope you will continue your journey with physics, with unquenchable enthusiasms and unleashed talents. Hung-An, thank you for teaching me machine learning and programming. I will never forget those prayers we shared in the fall of 2011; I wish you all the best in Seattle. Ying-zong, knowing you is the honor of my life. Despite celebrating the same birthday, I always have so much to learn from you, not only a sophisticated and well-rounded mind, but a humble and tranquil heart. I am confident that you will apply yourself and contribute maximally to society whatever career you decide to choose. Mengdi, thank you for standing behind me during the moment of sorrows

and tears. I am so glad that our encounter does not end here at MIT. I also want to extend my special thanks to four friends outside the MIT community, namely, Dr. Lixin (Jane) Dai in the Stanford Physics Department, Dr. Xiaobai Chen in the Princeton Computer Science Department, Dr. Yi Zhang in the CMU Machine Learning Department, and Dr. Jing Zhao in the USC Computer Science Department. For each one of you, your speaking mind, and more importantly, your shining soul have shaped my life.

My years at Boston have witnessed not only my maturity in intellectual capacity, but much more, my growth in life. I am deeply indebted to the following families and individuals in my church, for their love and care outpoured on me: Henry and Carol Hwang, Peng and Ester Yi, Daniel and Tracy Chen, Isaac and Tammy Chu, Vic and Ivy Law, Gustavo and Sabrina Gonzalez, Peter and Rebecca Tsui, Saul and Tina Becker, Thomas and Cynthia Cheng, Gideon and Christina Lin, Bruce and Peggy Light, Andy and Crystal Chang, David and Sarah Hwang, Jim and Judy Danker, Paul and Conwen Lee, Helen Wang, Frank Meng, Fan Zhao, David Liu, Randall Briggs, Ryan Pester, Brian Ferreira, Kevin Bagnall, Lucas Petrash, Andrew Koo, Titus Jahng, Melvin Zhang, and many others. In particular, I would like to mention Brother Dave Bekker, who passed away last September and has been the individual with the biggest impact on me in the past seven years. These dear ones have rendered not only pleasure, but meaning, to my life.

The final note is to the people I owe the most – my family. It has been twelve years since I went overseas for education. Many times, I fail the duty of the only son, a grandson, and a family member. It is those sacrifices beyond my imaginations borne by my family that have made me this far. To my little cousin "*xiaomei*", Ling Yang, I can see a lot of my self in you, despite that you are much more gifted and diligent. I hope you can enjoy a more balanced and joyful life. I look forward to seeing you in this city. To you, "*xiaomei*", I dedicate this thesis.

Cambridge, Massachusetts, United States

June 2012

Contents

Chapter 1

Introduction

1.1 Background and motivations

T cells are orchestrators of adaptive immunity. They can recognize and eliminate invading pathogens, and also develop memory against the same types of pathogens [1]. T cells scan pathogens on the surface of antigen-presenting cells (APCs). Antigen-presenting cells can internally process self and foreign proteins, cut them into short peptides, and present those peptides on major histocompatibility complex (MHC) molecules. T cell recognize antigens through the interaction between T-cell receptor (TCR), expressed on the T-cell surface, and peptide-MHC (pMHC) presented on the APC surface.

How T cells translate stimulatory external signals (e.g., TCR-pMHC interactions) to functional responses (e.g., proliferation, cytokine secretion, or apoptosis) has been the focus for immunologists for decades. While too modest pathogen-specific immune responses will fail to clear viruses effectively, erroneous functioning of adaptive immunity will attack self issues and hence lead to autoimmunity. The ability to differentiate quantitatively similar signals and result in qualitatively distinct response demands extraordinary intricacy for signaling pathways responsible for making functional decisions.

Such a stringent requirement is further complicated by the prevalence of stochasticity in signaling networks. Noise or randomness at the microscopic level arises from

at least two reasons: (1) biological reactions, by nature, are stochastic; (2) cell-to-cell variability in protein concentrations is ubiquitous. The small copy numbers of reacting species in any real biochemical network exacerbate stochastic fluctuations. Therefore, it is crucially important to understand in the face of microscopic fluctuations, how T cells regulate stochastic noise and deliver biological functions reliably.

The overarching theme of this thesis is to understand the role of stochasticity in T-cell signaling networks.

1.2 Thesis outline

This thesis consists of four projects which encompass, although not exhaustively, the extent of my work as a PhD student in the Department of Chemical Engineering at MIT.

In Chapter 2, in collaboration with Dr. Christopher Govern, a former PhD student of the Chakraborty group, I studied stochastic transitions due to microscopic fluctuations in biochemical networks that are able to drive cellular systems out of stable states and lead to spurious responses [2, 3]. We proposed a theoretical and computational framework to identify key reactions and species that are responsible for regulating such stochastic transitions. We believe such an understanding is important, because it enables us to identify the weak spot (i.e., reactions or species) in a given network. The mutation of those weak spots can easily lead to false triggering, spurious activation, or cancer [4]. On the other hand, the successful identification of key components that are effective in suppressing noise enables us to intelligently select potential candidates for drug targets. The semi-analytical method we derived using large deviation theory and calculus of variations agrees well with computational costly brute-force simulations, when both are applied to an important T-cell signaling module. Additionally, our framework reveals qualitative features of key reactions for network stability.

In Chapter 3, in collaboration with Dr. Jesse Jun and Dr. Jeroen Roose, both from USCF, we investigated the roles of RasGRP and SOS in the activation of ERK

and P38 MAP kinase pathways. Previously, by the synergy between computational modeling and experimental studies, a significant amount of knowledge was gained regarding how the interplay of RasGRP and SOS regulates the activation of Ras [5, 6]. However, little was known regarding how RasGRP and SOS regulate MAP kinase signaling cascades, which were reported to be crucial in diverse cellular processes, such as cell proliferation, differentiation or apoptosis [7]. In this work, I extended established computational models [6, 8, 9] to study the effect of SOS on ERK activation and hypothesized that SOS' allosteric pocket is important for the magnitude and pattern of ERK activation, which was later confirmed by experiments. I implemented the Hartigan statistical test [10] to confirm bimodality reported in flow cytometry data. Based on the experimental observation that SOS is preferentially more important than RasGRP for Rac-P38 activation and the mutation of SOS' allosteric pocket has little effect on that, I proposed several mechanistic models, incorporating features such as co-operativity and non-linearity thresholding. These models were able to explain observed experimental data, and more importantly, generate experimentally testable hypotheses.

In Chapter 4, in collaboration with Dr. Christopher Govern, instead of focusing on the average waiting time for a rare event to occur (mean first-passage time, or MFPT), as we did in Chapter 2, we examined the averaging duration for a rare event to complete once it has started (conditional mean first-passage time, or CMFPT). We showed that for Markov processes with detailed balance (such systems at steady state are known as equilibrium systems), the successful transition, though takes a long time to wait for its arrival, completes exactly as quickly as a transition in the opposite (non-rare) direction. We first provided a general proof for equilibrium Markov processes by invoking the property of time reversibility. We substantiated this proof to two specific dynamics, namely, continuous-time Markov Chain with detailed balance and one-dimensional Langevin Dynamics. In the proof of latter case, we employed machineries from measure theory and stochastic calculus and verified our results by numerical calculations. We concluded by extending physical insights obtained from proving cases in equilibrium systems to non-equilibrium systems (e.g., most real bio-

chemical networks). We claimed that in general, rare events, once happen, happen quickly. Suppressing undesirable rare transitions by adding competing forces (e.g., mounting immune pressure or administering drug treatment in order to quench viral escape), while prolongs the waiting time for rare events to occur, actually fastens the completion of such rare transitions once they have started.

In Chapter 5, I presented an effort of exploiting the non-random underlying structure of random noise for network topology identification. While the prevalence of stochasticity at microscopic level of biological systems imposes challenges for reliably translating external signal cues to biological functions, as seen in Chapter 2, noise also offers enormous information about the biochemical networks that generate such fluctuations. In particular, I focused on two important three-node network motifs, namely, incoherent feedforward loop (IFF) and negative feedback loop (NFB), which can both achieve adaptation, and posed the question whether one can discriminate these two by examining time series data of species copy numbers. Specifically, I applied linear noise approximation (also known as system-size expansion [11]) to both topologies and obtained correlation functions governing the fluctuations of species copy numbers at steady state. I identified two signatures that can be used to differentiate IFF from NFB. This endeavor can be the first step to the answer to a broader question: how or to what extent people can utilize time-series data with fine time and length resolutions to infer underlying networks.

1.3 References to published work and work outside the thesis scope

The work presented in Chapter 2 has been published in Physical Review Letters [12]. The work presented in Chapter 3 is being submitted to Proceedings of the National Academy of Sciences during the writing of this thesis. The materials presented in Chapters 4 and 5 are at the final stage of preparation for submission.

Besides the work presented in this thesis, I also participated in the following

projects during my PhD career:

1. In collaboration with Dr. Ed Palmer at University Hospital Basel, I studied the role of co-receptor in T-cell receptor triggering and its implications to thymic maturation. Specifically, α-CPM, a connecting peptide motif on the α-chain of TCR, is reported to be crucial for the interaction between TCR and co-receptor CD8 [13]. It has been experimentally observed that, surprisingly, the mutation on α-CPM impairs T-cell maturation more severely than CD8 deletion [14, 13]. We built computational models to provide a mechanistic explanation for this. We performed finite-state Markov Chain analysis, coupled with spatially-resolved stochastic simulations, to highlight the importance of rebinding between MHC and co-receptor in enhancing LCK recruitment.

2. In an ongoing collaboration with Dr. Arthur Weiss at UCSF, we built a mechanistic model to understand the regulation of kinase LCK by phosphatase CD45 and kinase CSK. It has been experimentally observed that the activity of LCK is tuned by the fine balance between CD45 and CSK, and while high level of CD45 is required for basal TCR signaling, low level of CD45 is preferred for inducible TCR triggering [15]. Our current computational model, by examining the steady-state distribution of LCK among its inactive, primed, and active states, is able to provide mechanistic understanding of such experimental observation. Through continued synergy between computational modeling and experimental studies, we hope to gain further insights on CSK regulation and understand how sensitive TCR triggering is to changes in CSK and CD45 concentrations.

3. I contributed minimally to the work led by a former post-doctoral researcher of the Chakraborty group, Dr. Jayajit Das, in collaboration with Dr. Jeroen Roose and Dr. Arthur Weiss, in which SOS-mediated positive feedback was shown to be the origin of digital signaling in T-cell activation [6]. In this work, I performed stochastic simulations of receptor-induced Ras activation. The computational model developed in this work later on served as a basic building

block for Ras-SOS signaling module, as seen in Chapters 2 and 3.

4. In collaboration with Dr. Jeroen Roose, we adapted the computational model used in [6] to identify Rasgrp1 as a Ras activator contributing to oncogenesis of T-cell lymphoma. Combining *in silico* and *in vitro* approaches, we demonstrated that Rasgrp1 efficiently synergizes with receptor input to simulate the Akt and NKκB pathways and T-ALL outgrowth *in vivo*. This work is currently under the review of Science Signaling.

Chapter 2

Identifying Dynamical Bottlenecks of Stochastic Transitions in Biochemical Networks

In biochemical networks, identifying key proteins and protein-protein reactions that regulate fluctuation-driven transitions leading to pathological cellular function is an important challenge. Using large deviation theory, we develop a semi-analytical method to determine how changes in protein expression and rate parameters of protein-protein reactions influence the rate of such transitions. Our formulas agree well with computationally costly direct simulations, and are consistent with experiments. Our approach reveals qualitative features of key reactions that regulate stochastic transitions.

2.1 Introduction

Biochemical reactions underlie the function of living cells. For example, they translate stimulatory external cues to functional responses. At the microscopic level, fluctuations in reaction rates and protein concentrations are ubiquitous [16]. Such noise can drive cells out of stable states (e.g., the "resting" state of unstimulated cells), leading to spurious responses [2, 3]. Discovering the key mechanisms that are re-

sponsible for regulating such stochastic fluctuations in cellular reaction networks is important for uncovering design principles of biological signaling networks [17]. In particular, understanding how changes in protein concentrations or mutations that change rate parameters promote fluctuation-driven transitions to pathological cellular states (e.g., cancer [4]) can help identify key deleterious protein mutations and inform efforts to manipulate specific proteins (drug targets) that would reverse such aberrant regulation.

Toward obtaining such mechanistic pictures, we calculate how small changes in rate parameters and protein concentrations affect the average time it takes for a biologically meaningful stochastic transition to occur – either a (potentially undesirable) fluctuation out of a stable state to an unstable state with biological significance, or a stochastic switch between different stable states. Shorter transition times correspond to more unstable networks.

Determining these sensitivities can be computationally complex for two main reasons. First, direct simulations to determine transition times can be costly, because transitions away from a stable state can be rare and therefore hard to sample [18, 19, 20]. Second, realistic biochemical networks often involve many reaction rates and concentrations, each of which must be perturbed to determine if it significantly affects the transition.

Here we develop a semi-analytical technique that aims to overcome these challenges by exploiting ideas from large deviation theory (LDT) [21, 22, 23, 24, 25, 26, 27, 28, 29, 30]. The semi-analytical expressions reveal qualitative features that characterize the key components that affect network stability. For a nontrivial biological model, we demonstrate that our technique gives results quantitatively consistent with trajectory-based simulation results. Computationally, the advantage of our approach is that it requires only a single deterministic simulation to determine the effect of perturbing all rate constants and species concentrations, as long as the transitions are rare and the small perturbations do not lead to phase transitions.

2.2 Analytical expression of action

Consider a system of N_S different chemical species, whose copy numbers evolve stochastically according to a reaction network with N_R reactions, characterized by a stoichiometric matrix E (of dimension $N_S \times N_R$) and rate constants $\mathbf{k}$ (of dimension N_R) according to pre-defined rate laws (e.g., mass action). The system may have one or more stable steady states for species concentrations, corresponding to different stable cellular states. We assume that the volume V is sufficiently large to make large excursions away from any particular fixed point rare relative to relaxation to the fixed point, and derivatives with respect to copy numbers (or concentrations) are well defined. In this limit, transitions away from a fixed point, specified by a vector of concentrations ($\mathbf{c}_A^{FP}$) to another fixed point $\mathbf{c}_B^{FP}$ or to any other biologically relevant (set of) states, can be described by a single rate K. The percentage change in K when a rate constant k_i is perturbed by a small percentage can be quantified by $\dfrac{\partial \ln K}{\partial \ln k_i}$. Similarly, the percentage change in K when the concentration of species i at some arbitrary time t, $c_i(t)$, is perturbed by adding molecules of species i to the system is $\dfrac{\partial \ln K}{\partial c_i}$. (The absolute change in c_i is used because there is no unique time-independent concentration scale.)

First consider transitions between stable steady states $\mathbf{c}_A^{FP}$ and $\mathbf{c}_B^{FP}$. We describe a possible transition path by the actual reaction propensities at each time $\hat{\boldsymbol{\nu}}(t)$ (that is, the number of reactions of each type that occur per volume per time). Given a starting point $\mathbf{c}(0)$ (e.g., $\mathbf{c}_A^{FP}$), $\mathbf{c}(t)$ is uniquely defined by $\hat{\boldsymbol{\nu}}(t)$ through Eq (2.1):

$$\mathbf{c}(t) = \mathbf{c}(0) + E \int_0^t \hat{\boldsymbol{\nu}}(t')dt' \tag{2.1}$$

Suppose a system is evolving according to a particular path $\hat{\boldsymbol{\nu}}(t)$ and that the system is therefore at a particular point $\mathbf{c}(t)$ at time t. Over the next differential time interval $[t, t+\Delta t]$, the actual number of reactions of type i that occur is $n_i = \hat{\nu}_i(t)V\Delta t$. However, the expected number of reactions of type i is $\lambda_i = \nu_i V \Delta t$, where ν_i, the deterministic propensities, are given by the pre-defined rate law as a function of the state $\mathbf{c}(t)$ and the rate constant k_i. The probability of observing n_i reactions of type

i over the differential time Δt follows a Poisson distribution characterized by λ_i.

Hence, the probability of observing $\mathbf{n} = \hat{\boldsymbol{\nu}}(t)V\Delta t$ reactions is:

$$\mathbb{P}(\mathbf{n}) = \prod_{i=1}^{N_R} \frac{\lambda_i^{n_i}}{n_i!} e^{-\lambda_i} = \prod_{i=1}^{N_R} \frac{(\nu_i V\Delta t)^{\hat{\nu}_i V\Delta t}}{(\hat{\nu}_i V\Delta t)!} e^{-\nu_i V\Delta t} \tag{2.2}$$

The first equality holds assuming the time interval is sufficiently small so that the species concentrations do not change significantly over the interval. By Stirling's approximation and in the continuum limit, the probability density of the path $\hat{\boldsymbol{\nu}}(t)$, not necessarily a transition path connecting the endpoints, with t from 0 to τ, is proportional to $\exp(-VS(\hat{\boldsymbol{\nu}}, \tau))$, where

$$S(\hat{\boldsymbol{\nu}}, \tau) = \sum_i^{N_R} \int_0^\tau dt(\hat{\nu}_i \ln \frac{\hat{\nu}_i}{\nu_i} - \hat{\nu}_i + \nu_i) \tag{2.3}$$

$S(\hat{\boldsymbol{\nu}}, \tau)$ serves as the action (or the rate function) of the path $\hat{\boldsymbol{\nu}}$ over $[0, \tau]$. Eq (2.3) parses the action of a path in a complex reaction network into contributions from individual reactions (it has been derived in a different way by Liu [31]). Hence, it can potentially identify reactions that can most effectively alter rates of rare transitions.

The minimal action S^* for the transition, and the corresponding most probable (least action) transition path $\hat{\boldsymbol{\nu}}^*$, can be found by minimizing Eq (2.3) over all the paths that originate from $\mathbf{c}_A^{FP}$ and reach $\mathbf{c}_B^{FP}$ at time τ, and then minimizing over τ: $S^* = \min_\tau \min_{\hat{\boldsymbol{\nu}}} S(\hat{\boldsymbol{\nu}}, \tau)$ [21, 27].

The transition rate K of a rare event is, from large deviation theory (e.g., WKB approximation [21, 22]), $K = A\exp(-VS^*)$, where A depends sub-exponentially on the volume and both A and S^* depend on the parameters describing the system (e.g., $\mathbf{k}$ and initial concentrations). In the large-V limit, simple calculation obtains the sensitivity of K to different parameters to be: $\frac{\partial \ln K}{\partial \ln k_i} \approx -V\frac{\partial S^*}{\partial \ln k_i}$, and $\frac{\partial \ln K}{\partial c_i} \approx -V\frac{\partial S^*}{\partial c_i}$.

2.3 Perturbation of rate parameters

The perturbation of rate constant k_i by Δk_i changes the minimal action S^*. As shown in the supplement, the action of the new optimal path can be parsed in terms of that of the original optimal path under the new k_i, plus corrections for the fact that the two optimal paths are different (not only because their endpoints, which are steady states of the system, change with k_i). However, the corrections are of lower order, $O(\Delta k_i^2)$(proof in the Supplement, Fig. 2-2). Hence, to compute $\frac{\partial S^*}{\partial \ln k_i}$, and therefore $\frac{\partial \ln K}{\partial \ln k_i}$, we only have to evaluate the change of action (Eq. (3)) along the unperturbed optimal path $\hat{\boldsymbol{\nu}}^*$ when k_i changes:

$$\frac{\partial S^*(\mathbf{k})}{\partial \ln k_i} = \int_0^{\tau^*} \frac{\partial \mathcal{L}}{\partial \nu_i} \frac{\partial \nu_i}{\partial \ln k_i}\bigg|_{\hat{\boldsymbol{\nu}}^*} dt = k_i \int_0^{\tau^*} (1 - \frac{\hat{\nu}_i^*}{\nu_i^*}) \frac{\partial \nu_i^*}{\partial k_i} dt \tag{2.4}$$

where $\mathcal{L} = \hat{\nu}_i \ln \frac{\hat{\nu}_i}{\nu_i} - \hat{\nu}_i + \nu_i$. If ν_i is linear with k_i (e.g., mass action), the sensitivities can be further simplified:

$$\frac{\partial \ln K}{\partial \ln k_i} \approx -V \frac{\partial S^*(\mathbf{k})}{\partial \ln k_i} = V \int_0^{\tau^*} (\hat{\nu}_i^* - \nu_i^*) dt \tag{2.5}$$

For reactions that must occur more frequently than they would have deterministically in order to make the transition occur, RHS of Eq (2.5) is positive, suggesting that the transition rate K increases with the rate parameters of such reactions. For rate constants that participate in multiple reactions (e.g., dephosphorylations by the same phosphatase), the right hand side of Eq (2.5) will contain a summation over all such reactions.

Eq (2.5) provides a way to calculate the effect of perturbing rate constants on the transition time given a single input, the unperturbed optimal path $\hat{\boldsymbol{\nu}}^*$. This input path can be determined numerically using the efficient geometric minimal action method (gMAM; [27]).

2.4 Perturbation of initial concentrations

Adding molecules of a species not governed by conservation laws does not alter the rate of rare transitions (e.g., steady states do not change). For species governed by conservation laws (as specified by E in Eq (2.1)), perturbing concentrations by adding molecules to the system can change the steady states $\mathbf{c}_A^{FP}$ and $\mathbf{c}_B^{FP}$, the optimal path $\hat{\boldsymbol{\nu}}^*$, and the deterministic propensities $\boldsymbol{\nu}^*$ along the optimal path. We show that ignoring terms of $O(\Delta c_i^2)$ and higher, the change in minimal action is due to the change in deterministic propensities evaluated along the unperturbed optimal path $\hat{\boldsymbol{\nu}}^*$ (proof in Supplement and Fig. 2-3). Arguments analogous to the perturbation in k_i lead to

$$\frac{\partial \ln K}{\partial c_i} \approx -V \sum_j^{N_R} \int_0^{\tau^*} (1 - \frac{\hat{\nu}_j^*}{\nu_j^*}) \frac{\partial \nu_j^*}{\partial c_i} dt \tag{2.6}$$

where $\frac{\partial \nu_i^*}{\partial c_i}$ can be evaluated at each time point given the rate law. For mass action kinetics, the RHS of Eq (2.6) can be further simplified to be $V \sum_j^{N_R} \int_0^{\tau^*} E_{ij} \frac{\hat{\nu}_j^* - \nu_j^*}{c_i} dt$.

To study the escape from a stable state to a predefined condition other than a steady state (e.g., the number of molecules of a particular species acquires a specific value), Eqs. (2.5) and (2.6) still apply. The only modification is that the optimal path $\hat{\boldsymbol{\nu}}^*$ as the input into Eqs. (2.5) and (2.6) needs a further optimization over all the possible end states consistent with this condition, as described algorithmically in [27]. Note that Eq (2.6) does not directly apply, as derived, to the case where the perturbed species participates in the definition of the specified end condition (see the Supplement).

2.5 Results and Discussion

To test the accuracy of the methods developed above, we apply them to a biochemical reaction network that is present in several cell-signaling networks. Upon external stimulation, in many cell types, Ras proteins are converted from their inactive,

GDP-bound state to an active GTP-bound state. Active Ras can stimulate several downstream pathways, and aberrant regulation of Ras activation underlies many cancers. We study a particular version of a simple model for Ras activation described previously [6, 8, 32]. This model has 26 reactions (and associated rate constants), 14 species governed by 5 conservation equations, and in our study has about 400 molecules involved. The copy number of individual species is as small as 10 in our model studies. The dynamics do not observe detailed balance.

The main feature of the model is that Ras can be activated via two pathways mediated by proteins called Rasgrp and SOS, and deactivated by RasGAP. The activation by SOS is governed by a positive feedback loop: SOS's catalytic activity increases significantly by the binding of the active form of Ras, RasGTP to its allosteric site (denoted by GTP-SOS; in this paper, we use x-SOS (SOS-x) to denote species x bound to the SOS allosteric (catalytic) pocket, where x = GTP or GDP, with "Ras" omitted). This enables the system to exhibit bistability when SOS concentration is at an intermediate level [6]. Meanwhile, at a low SOS level, only one stable state exists, characterized by low level of RasGTP. Thus, the model is rich enough to investigate how cells control stochastic transitions between multiple stable states (at intermediate SOS level) and suppress fluctuations from a single stable state (at low SOS level), preventing spurious activation to an undesired state with a high level of active Ras. We used Eqs. (2.5) and (2.6) to predict the sensitivities of the transition time for both these situations. To obtain the unperturbed optimal path $\hat{\boldsymbol{\nu}}^*$, we implemented the gMAM [27].

To test the accuracy of Eqs. (2.5) and (2.6), we explicitly calculated the transition time under the unperturbed parameters and under perturbed values of each parameter (1% increase in rate constants; 1 molecule increase in each conserved species). We chose forward flux sampling (FFS; [33]) as the trajectory-based simulation method, and RasGTP level as its thresholding parameter.

As shown in Fig. 2-1A and B, Eqs. (2.5) and (2.6) predict a correctly ordered list of key reactions and species that influence the transition time. Additionally, the predictions are in remarkable quantitative agreement with the direct simulation results.

Our technique correctly identifies the key reactions for regulating Ras activation to be RasGAP catalytic activity, GTP-SOS catalytic activity, and Rasgrp catalytic activity. The most important species are RasGAP, SOS, and Rasgrp, which are, as expected, associated with these key reactions. These predictions are consistent with recent experimental results in T-cell cancers[32], where the mutation of Ras that abrogates RasGAP activity is observed to significantly change active Ras level; i.e., spurious Ras activation is very sensitive to RasGAP activity. When the SOS level decreases, the relative importance of Rasgrp increases (compare Fig. 2-1A and B to Fig. 2-4 A and B). This is consistent with the experimental finding that Rasgrp over-expression is a major cause for human cancer when little stimulus is present [34].

The quantitative discrepancies between predictions from Eqs. (2.5) and (2.6) and direct simulations (Fig. 2-1A and B) are due to the finite size of the simulation system, the finite change in rate parameters and initial concentrations, and statistical uncertainties in FFS results. Note that for RasGAP, the species in Fig. 2-1B to which the transition time is the most sensitive, an increase of one molecule represents a relatively large (10%) change in concentration, which explains the deviation between the prediction and the direct simulation. Additional calculations (data not shown) suggest that for this model, our method can predict the qualitative order of reactions or species (but not the quantitative values of their sensitivities), even when larger perturbations are applied, as long as the perturbations do not lead to phase transitions.

Both direct simulation and our method require one simulation for the unperturbed transition: for direct simulation, to determine the unperturbed transition rate K, and for our method, to determine the unperturbed optimal path $\hat{\boldsymbol{\nu}}^*$. To compute the sensitivity, direct simulation requires additional simulations to explore all possible perturbations of the parameters. For the simple model we studied, this corresponds to 31 additional simulations. Our method requires no additional simulations.

Furthermore, the semi-analytical forms of Eqs. (2.3) and (2.5) enable us to draw qualitative conclusions about key reactions responsible for network stability. It might be expected that the network will be the most sensitive to reactions that require the

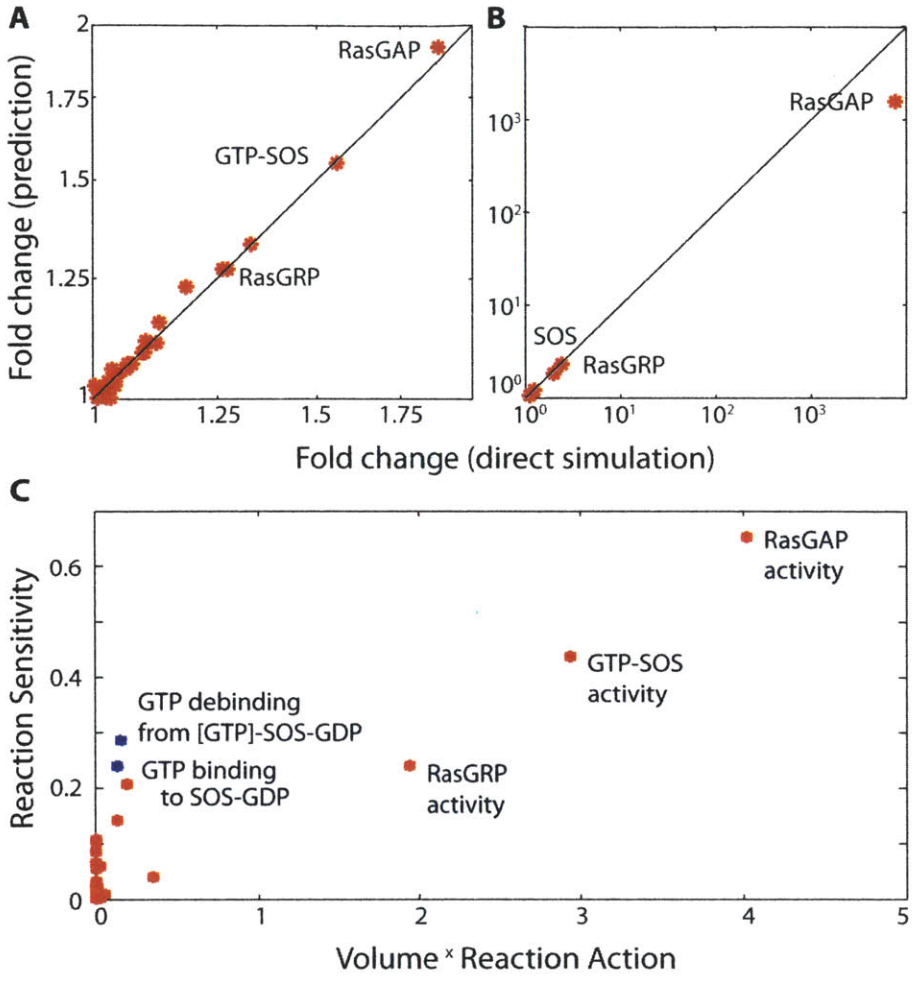


Figure 2-1: (A, B) Quantitative consistency between predictions and direct simulation results at SOS = 50 (intermediate). Fold change in transition rate (i.e., defined as a larger rate divided by a smaller one) (A) when each rate constant is increased by 1% and (B) when one molecule of each type is added to the system. Each red dot represents the perturbation of (A) a different rate constant and (B) a different species concentration. The direct simulation results are the averages of ten independent FFS runs, and for each FFS run 10^3 points were stored on each surface. Only the results for the key reactions, namely RasGAP catalytic activity, GTP-SOS catalytic activity, and RasGRP catalytic activity are labelled in (A), and only the results for the key species, namely RasGAP, SOS, and RasGRP are labelled in (B). Similar calculations are done at SOS = 15 (low) and reported in Fig. 2-4. (C) Reaction action is a dominant but not sole factor for sensitivity. Each dot represents the sensitivity and action of one reaction at SOS = 50. The rankings of reactions by action and by sensitivity are similar. However, note that the binding (unbinding) reaction of GTP to (from) SOS-GDP, indicated by the blue dot, has small action but large sensitivity. Only the reactions with relatively large sensitivities are labelled.

rarest fluctuations to drive the transition. Note that a reaction's contribution to the rarity of the transition is quantified by reaction action $S_i^* = \int_0^{\tau^*} (\hat{\nu}_i^* \ln \frac{\hat{\nu}_i^*}{\nu_i^*} - \hat{\nu}_i^* + \nu_i^*) dt$ (as seen from Eq (2.3)). By ranking reactions by their sensitivities and by their actions, we see that rarity has a dominant effect on sensitivity, but it is not the only effect (Fig. 2-1C).

Eqs. (2.3) and (2.5) show that (at least) three factors determine the sensitivity to a particular reaction: (a) the rarity of its required fluctuation, as measured by S_i^*, (b) the frequency of the reaction as measured by its deterministic propensity ν_i^*, and (c) the uniformity of the fluctuations distributed over the time course of the transition. Holding the other two factors constant, the reaction sensitivity increases with each of these factors.

The importance of (b) can be seen more vividly by noticing that the second-order approximation of S_i^* yields $S_i^* \approx \int_0^{\tau^*} (\hat{\nu}_i^* - \nu_i^*)^2 / 2\nu_i^* dt$; hence given the same action S_i^*, the reaction with larger ν_i^* has a larger sensitivity. This explains why in Fig. 2-1C the binding (unbinding) reaction of RasGTP to (from) the allosteric pocket of SOS with RasGDP bound to its catalytic pocket (i.e., SOS-GDP) has small action but relatively large sensitivity, compared to RasGRP activity. Our calculation shows that the deterministic propensities of these two reactions are about 30 times bigger than that of RasGRP activity. For (c), note that S_i^* is convex with respect to $\hat{\nu}_i^*/\nu_i^*$ (a similar functional form also appears in the Boltzmann H-theorem), while the reaction sensitivity is linear; we can show that by Jensen's inequality, given two reactions with the same action and the same frequency, the reaction with sustained modest fluctuations, rather than a short large burst, has a larger sensitivity (proof in Supplement). The action also vanishes faster than the sensitivity because of the above difference in scaling. This may be why we see a cluster of reactions with moderate sensitivities, but small actions, in Fig. 2-1C.

We have applied a LDT method to a finite system to obtain accurate results, which offers substantial computational advantages compared to direct simulation. It also provides qualitative insights into the features of biochemical reactions important for regulating fluctuation-driven transitions. These features should enable studies of real

biological systems. For example, we could seek to determine how spurious transitions caused by cancerous mutations may be quenched by drugs that target certain species involved in key reactions.

We thank Chong Wang for helpful discussions. This work was supported by NCI (PSOC) and P01 AI091580-01.

2.6 Supplementary Materials

2.6.1 Perturbation of rate parameters

Let $\mathbf{c}_A^{FP}$ and $\mathbf{c}_B^{FP}$ be the steady states under rate parameter $\mathbf{k}$. Let the optimal path connecting $\mathbf{c}_A^{FP}$ to $\mathbf{c}_B^{FP}$ be $\hat{\boldsymbol{\nu}}^*_{A\to B}$; let its duration be $\tau^*_{A\to B}$; and let the corresponding minimal action be $S^*(\mathbf{k}) = S(\hat{\boldsymbol{\nu}}^*_{A\to B}, \tau^*_{A\to B}; \mathbf{k})$, as defined in Eq (2.3) in the main text (Fig. 2-2A). Note that here we explicitly include $\mathbf{k}$ as an argument of S; this reflects the dependence of S on $\mathbf{k}$ through the deterministic propensity $\boldsymbol{\nu}^*$ (which depends on the rate constants), independent of the path $\hat{\boldsymbol{\nu}}^*$ and the time τ^*.

Now perturb $\mathbf{k}$ by a small amount $\Delta\mathbf{k}$ such that the perturbed rate parameter is $\mathbf{k}' = \mathbf{k} + \Delta\mathbf{k}$. Let $\mathbf{c}_{A'}^{FP}$ and $\mathbf{c}_{B'}^{FP}$ be the new steady states under $\mathbf{k}'$ and $\hat{\boldsymbol{\nu}}'^*_{A'\to B'}$ be the new optimal path from $\mathbf{c}_{A'}^{FP}$ to $\mathbf{c}_{B'}^{FP}$ in time $\tau'^*_{A'\to B'}$, so that the new minimal action is $S^*(\mathbf{k}') = S(\hat{\boldsymbol{\nu}}'^*_{A'\to B'}, \tau'^*_{A'\to B'}; \mathbf{k}')$ (Fig. 2-2A). In the following derivation, we assume the perturbation in $\mathbf{k}$ is done by perturbing a single rate constant, k_i; the derivation can be easily generalized to the case of perturbing multiple rate constants because small perturbations are additive. (Of course, this is no longer true when the deviations in k_i are larger.) Note the use of $*$ to denote an optimal path, time, or action, and the use of $'$ to denote quantities calculated under the perturbed rate constant $\mathbf{k}'$. (The notation $'*$ indicates a quantity optimized under $\mathbf{k}'$.)

As shown in Fig. 2-2A (right), the proof involves the construction of two new paths: we connect $\mathbf{c}_{A'}^{FP}$ to $\mathbf{c}_A^{FP}$ via the optimal path under $\mathbf{k}'$, $\hat{\boldsymbol{\nu}}'^*_{A'\to A}$, which takes time $\tau'^*_{A'\to A}$; we also connect $\mathbf{c}_B^{FP}$ to $\mathbf{c}_{B'}^{FP}$ via the optimal path under $\mathbf{k}'$, $\hat{\boldsymbol{\nu}}'^*_{B\to B'}$, which takes time $\tau'^*_{B\to B'}$. We refer to these paths as extension paths.

To compute the sensitivity of the minimal action S^* to k_i, we express the derivative as:

$$\frac{\partial S^*(\mathbf{k})}{\partial k_i} = \lim_{\Delta k_i \to 0} \frac{S^*(\mathbf{k}') - S^*(\mathbf{k})}{\Delta k_i} = \lim_{\Delta k_i \to 0} \frac{S(\hat{\boldsymbol{\nu}}'^*_{A' \to B'}, \tau'^*_{A' \to B'}; \mathbf{k}') - S(\hat{\boldsymbol{\nu}}^*_{A \to B}, \tau^*_{A \to B}; \mathbf{k})}{\Delta k_i} \tag{2.7}$$

We illustrate how the numerator can be approximated and parsed in Figs. 2-2B-F. We approximate $S(\hat{\boldsymbol{\nu}}'^*_{A' \to B'}, \tau'^*_{A' \to B'}; \mathbf{k}')$ by $S(\hat{\boldsymbol{\nu}}^*_{A \to B}, \tau^*_{A \to B}; \mathbf{k}')$ as shown in Figs. 2-2C-E, by demonstrating that the difference between these two terms is $O(\Delta k_i^2)$ (Fig. 2-2F):

$$S(\hat{\boldsymbol{\nu}}'^*_{A' \to B'}, \tau'^*_{A' \to B'}; \mathbf{k}') - S(\hat{\boldsymbol{\nu}}^*_{A \to B}, \tau^*_{A \to B}; \mathbf{k}') = O(\Delta k_i^2) \tag{2.8}$$

Eq (2.8) is proved by the following steps:

1. We construct a new path by concatenating three existing paths, the optimal path under $\mathbf{k}$ and the two extension paths: $\hat{\boldsymbol{\nu}}'_{A' \to B'} = \{\hat{\boldsymbol{\nu}}'^*_{A' \to A}, \hat{\boldsymbol{\nu}}^*_{A \to B}, \hat{\boldsymbol{\nu}}'^*_{B \to B'}\}$, which takes $\tau'_{A' \to B'} = \tau'^*_{A' \to A} + \tau^*_{A \to B} + \tau'^*_{B \to B'}$ to transition from $\mathbf{c}^{FP}_{A'}$ to $\mathbf{c}^{FP}_{B'}$. Note that we now have two paths connecting the same endpoints $\mathbf{c}^{FP}_{A'}$ and $\mathbf{c}^{FP}_{B'}$: the path $\hat{\boldsymbol{\nu}}'^*_{A' \to B'}$ is the optimal path under $\mathbf{k}'$, by definition; the concatenated path $\hat{\boldsymbol{\nu}}'_{A' \to B'}$ is a valid but not optimal path under $\mathbf{k}'$.

   The distance between these two paths vanishes when $\Delta k_i \to 0$, provided that there is no phase transition (i.e., $\mathbf{c}^{FP}_{A'}$ converges to $\mathbf{c}^{FP}_{A}$, $\mathbf{c}^{FP}_{B'}$ to $\mathbf{c}^{FP}_{B}$, and $\hat{\boldsymbol{\nu}}^*_{A \to B}$ to $\hat{\boldsymbol{\nu}}^*_{A' \to B'}$).

   Since the action under $\mathbf{k}'$ achieves its minimum at the optimal path $\hat{\boldsymbol{\nu}}'^*_{A' \to B'}$, by calculus of variations, $\delta S / \delta \hat{\boldsymbol{\nu}}' = 0$ at $\hat{\boldsymbol{\nu}}'^*_{A' \to B'}$ for any t. Further, because we assume there is no phase transition, $\partial \hat{\boldsymbol{\nu}}' / \partial k_i$ does not diverge at $\hat{\boldsymbol{\nu}}'^*_{A' \to B'}$. Hence by the chain rule, we can formally establish the following equation:

$$\lim_{\Delta k_i \to 0} \frac{S(\hat{\boldsymbol{\nu}}'^*_{A' \to B'}, \tau'^*_{A' \to B'}; \mathbf{k}') - S(\hat{\boldsymbol{\nu}}'_{A' \to B'}, \tau'_{A' \to B'}; \mathbf{k}')}{\Delta k_i}$$
$$= \int_0^{\tau^\ddagger} \left(\frac{\delta S}{\delta \hat{\boldsymbol{\nu}}'} \frac{\partial \hat{\boldsymbol{\nu}}'}{\partial k_i} \right) \Bigg|_{\hat{\boldsymbol{\nu}}'^*_{A' \to B'}} dt = 0$$

   where $\tau^\ddagger = \max(\tau'^*_{A' \to B'}, \tau'_{A' \to B'})$. (Note that $\tau^\ddagger$ is introduced because paths

$\hat{\boldsymbol{\nu}}'^{*}_{A'\to B'}$ and $\hat{\boldsymbol{\nu}}'_{A'\to B'}$ take different times to transition from $\mathbf{c}^{FP}_{A'}$ to $\mathbf{c}^{FP}_{B'}$ (i.e., $\tau'^{*}_{A'\to B'} \neq \tau'_{A'\to B'}$). $\tau^{\ddagger}$ can be defined as such for the following reason: without loss of generality, if $\Delta t = \tau'^{*}_{A'\to B'} - \tau'_{A'\to B'} > 0$, we can extensionally define $\hat{\boldsymbol{\nu}}'_{A'\to B'}(t)$ on the time interval $[-\Delta t, 0]$ to be the deterministic propensities at $\mathbf{c}^{FP}_{A'}$, so that the system remains at $\mathbf{c}^{FP}_{A'}$ for $t \in [-\Delta t, 0]$. This extension in time does not alter the action of $\hat{\boldsymbol{\nu}}'_{A'\to B'}$ since in the extended time actual and deterministic propensities being equal leads to zero action. We then shift the time of $\hat{\boldsymbol{\nu}}'_{A'\to B'}(t)$ by Δt, resulting in a path $\hat{\boldsymbol{\nu}}^{\ddagger}_{A'\to B'}(t)$ on $t \in [0, \tau'^{*}_{A'\to B'}]$; note that $S(\hat{\boldsymbol{\nu}}'_{A'\to B'}, \tau'_{A'\to B'}; \mathbf{k}') = S(\hat{\boldsymbol{\nu}}^{\ddagger}_{A'\to B'}, \tau'^{*}_{A'\to B'}; \mathbf{k}')$. Now when we compare $S(\hat{\boldsymbol{\nu}}'^{*}_{A'\to B'}, \tau'^{*}_{A'\to B'}; \mathbf{k}')$ to $S(\hat{\boldsymbol{\nu}}^{\ddagger}_{A'\to B'}, \tau'^{*}_{A'\to B'}; \mathbf{k}')$, the following two conditions are still satisfied, namely at $\hat{\boldsymbol{\nu}}'^{*}_{A'\to B'}$, $\delta S/\delta \hat{\boldsymbol{\nu}}' = 0$ and $\partial \hat{\boldsymbol{\nu}}'/\partial k_i$ does not diverge, and therefore the above functional derivative can be established.) This gives rise to (Fig. 2-2C):

$$S(\hat{\boldsymbol{\nu}}'^{*}_{A'\to B'}, \tau'^{*}_{A'\to B'}; \mathbf{k}') = S(\hat{\boldsymbol{\nu}}'_{A'\to B'}, \tau'_{A'\to B'}; \mathbf{k}') + O(\Delta k_i^2)$$

2. Since actions are additive, the action of the concatenated path is just the sum of the actions of its component paths (Fig. 2-2F):

$$S(\hat{\boldsymbol{\nu}}'_{A'\to B'}, \tau'_{A'\to B'}; \mathbf{k}') = S(\hat{\boldsymbol{\nu}}'^{*}_{A'\to A}, \tau'^{*}_{A'\to A}; \mathbf{k}')$$
$$+S(\hat{\boldsymbol{\nu}}^{*}_{A\to B}, \tau^{*}_{A\to B}; \mathbf{k}') + S(\hat{\boldsymbol{\nu}}'^{*}_{B\to B'}, \tau'^{*}_{B\to B'}; \mathbf{k}')$$

3. Note that the action under $\mathbf{k}'$ of one extension path, $\hat{\boldsymbol{\nu}}'^{*}_{A'\to A}$, is $O(\Delta k_i^2)$ (Fig. 2-2E), for the following reason: under $\mathbf{k}'$, the minimal action from steady state $\mathbf{c}^{FP}_{A'}$ to a neighboring state $\mathbf{c}^{FP}_{A}$, as shown in [22], can be expressed as $S^*(\mathbf{r}_A)$, where $\mathbf{r}_A = \mathbf{c}_A - \mathbf{c}_{A'}$. By Taylor expansion, $S^*(\mathbf{r}_A) = S^*(\mathbf{0}) + \mathbf{r}_A \partial S^*(\mathbf{0})/\partial \mathbf{r} + O(\mathbf{r}_A^2)$, where $S^*(\mathbf{0}) = 0$ by definition. Since the minimal action $S^*(\mathbf{r}_A)$ achieves its minimum at $\mathbf{r}_A = \mathbf{0}$, corresponding to the steady state $\mathbf{c}^{FP}_{A'}$, $\partial S^*(\mathbf{0})/\partial \mathbf{r} = \mathbf{0}$. Therefore, $S(\hat{\boldsymbol{\nu}}'^{*}_{A'\to A}, \tau'^{*}_{A'\to A}, \mathbf{k}') = S^*(\mathbf{r}_A) = O(\Delta k_i^2)$, given that $\mathbf{r}_A = O(\Delta k_i)$ (we have assumed there is no phase transition).

Last, the action of the other extension path, $\hat{\boldsymbol{\nu}}'^*_{B\to B'}$, is zero, since the optimal path from $\mathbf{c}_B^{FP}$ to the steady state $\mathbf{c}_{B'}^{FP}$ under $\mathbf{k}'$ is a deterministic path, which has zero action [22, 27](Fig. 2-2E).

This concludes the proof for Eq (2.8) (i.e., Fig. 2-2F). By substituting Eq (2.8) into Eq (2.7), we see that to compute the sensitivity of minimal action to k_i, we only have to compute $S(\hat{\boldsymbol{\nu}}^*_{A\to B}, \tau^*_{A\to B}; \mathbf{k}') - S(\hat{\boldsymbol{\nu}}^*_{A\to B}, \tau^*_{A\to B}; \mathbf{k})$; that is, to evaluate the change in action due to the change in k_i along the original optimal path $\hat{\boldsymbol{\nu}}^*$ (Fig. 2-2G). This proves Eq (2.4).

To study the escape from a single stable state to a set of states corresponding to a predefined condition, the same proof applies, with the only exception that we do not need to connect the optimal endpoint under $\mathbf{k}$, $\mathbf{c}_B^{opt}$, to the optimal endpoint under $\mathbf{k}'$, $\mathbf{c}_{B'}^{opt}$. This is because $\mathbf{c}_B^{opt}$ lies on the hyper-surface corresponding to the allowed end states, and therefore $\hat{\boldsymbol{\nu}}'_{A'\to B} = \{\hat{\boldsymbol{\nu}}'^*_{A'\to A}, \hat{\boldsymbol{\nu}}^*_{A\to B}\}$ is a legitimate path connecting $\mathbf{c}_{A'}^{FP}$ to the target hyper-surface. Thus, Eq (2.8) still holds.

2.6.2 Perturbation of initial concentrations

Consider perturbing the concentration of species i by Δc_i. Assume the concentration of species i is governed by conservation equations. Before perturbation, let $\mathbf{c}_A^{FP}$ and $\mathbf{c}_B^{FP}$ denote the steady states, $\hat{\boldsymbol{\nu}}^*_{A\to B}$ denote the optimal path, and $S^*(\mathbf{c}_A^{FP}) = S(\hat{\boldsymbol{\nu}}^*_{A\to B}, \tau^*_{A\to B}; \mathbf{c}_A^{FP})$ denote the minimal action. We suppress the dependence of S on $\mathbf{k}$, since the latter is held constant; we explicitly include the starting point as one argument, since the deterministic propensity $\boldsymbol{\nu}^*$ depends on the starting point, as seen in Eq (2.1) of the main text. Because there are conservation equations, any path in species concentration space $\mathbf{c}(t)$ evolves on a subspace $\mathbf{C}$ with dimension $(N_S - N_C)$, where N_C is the number of species conservation equations.

After perturbation, let the steady states be $\mathbf{c}_{A'}^{FP}$ and $\mathbf{c}_{B'}^{FP}$; and let the minimal action be $S^*(\mathbf{c}_{A'}^{FP}) = S(\hat{\boldsymbol{\nu}}'^*_{A'\to B'}, \tau'^*_{A'\to B'}; \mathbf{c}_{A'}^{FP})$. After perturbation, paths $\mathbf{c}'(t)$ evolve on a subspace $\mathbf{C}'$. Similar to the case of perturbation of $\mathbf{k}$, we assume in this sec-

tion the perturbation in $\mathbf{c}$ is done by perturbing a single species, c_i, and it can be easily generalized to the case of perturbing multiple species concentrations by small amounts.

As illustrated in Fig. 2-3A, first we shift $\mathbf{c}_A^{FP}$ by Δc_i so that its image $\mathbf{c}_{A''}$ lands in the subspace $\mathbf{C}'$. Starting from $\mathbf{c}_{A''}$, generate a dynamical path according to $\hat{\boldsymbol{\nu}}^*_{A\to B}$. Note that although $\hat{\boldsymbol{\nu}}^*_{A\to B}$ was optimized on the subspace $\mathbf{C}$, it still describes a valid sequence of reactions in the subspace $\mathbf{C}'$. Furthermore, the resulting path in the species concentration space stays in the subspace $\mathbf{C}'$, since it starts in this subspace and evolves according to reactions that obey the conservation laws. Denote the endpoint of this path in the subspace $\mathbf{C}'$ as $\mathbf{c}_{B''}$. Note the use of $'$ to denote quantities obtained after the perturbation in concentration, and the use of $''$ to denote quantities of the shifted path.

As we did in the case of perturbing k_i, we construct two extension paths: connect $\mathbf{c}_{A'}^{FP}$ to $\mathbf{c}_{A''}$ using the optimal path $\hat{\boldsymbol{\nu}}'^*_{A'\to A''}$ in time $\tau'^*_{A'\to A''}$, and $\mathbf{c}_{B''}$ to $\mathbf{c}_{B'}^{FP}$ using the optimal path $\hat{\boldsymbol{\nu}}'^*_{B''\to B'}$ in time $\tau'^*_{B''\to B'}$ (Fig. 2-3B).

Now to compute the sensitivity of the minimal action S^* to c_i, we have:

$$\frac{\partial S^*}{\partial c_i} = \lim_{\Delta c_i\to 0}\frac{S^*(\mathbf{c}_{A'}^{FP}) - S^*(\mathbf{c}_A^{FP})}{\Delta c_i} = \lim_{\Delta c_i\to 0}\frac{S(\hat{\boldsymbol{\nu}}'^*_{A'\to B'}, \tau'^*_{A'\to B'}; \mathbf{c}_{A'}^{FP}) - S(\hat{\boldsymbol{\nu}}^*_{A\to B}, \tau^*_{A\to B}; \mathbf{c}_A^{FP})}{\Delta c_i} \tag{2.9}$$

To compute the numerator (illustrated in Fig. 2-3C), we approximate its first term by $S(\hat{\boldsymbol{\nu}}^*_{A\to B}, \tau^*_{A\to B}; \mathbf{c}_{A''})$ through the steps illustrated in Figs. 2-3D-F, applying the same argument for the case of perturbing k_i, showing that (i.e., Fig. 2-3G):

$$S(\hat{\boldsymbol{\nu}}'^*_{A'\to B'}, \tau'^*_{A'\to B'}; \mathbf{c}_{A'}^{FP}) - S(\hat{\boldsymbol{\nu}}^*_{A\to B}, \tau^*_{A\to B}; \mathbf{c}_{A''}) = O(\Delta c_i^2) \tag{2.10}$$

To compute Eq (2.9), we only have to compute $S(\hat{\boldsymbol{\nu}}^*_{A\to B}, \tau^*_{A\to B}; \mathbf{c}_{A''}) - S(\hat{\boldsymbol{\nu}}^*_{A\to B}, \tau^*_{A\to B}; \mathbf{c}_A^{FP})$. Note that these two paths in the species concentration space are parallel and pointwise different by Δc_i. To evaluate the change in action, we can simply calculate the action change due to the shift of the original optimal path. This proves Eq (2.6).

Eq (2.6) can be directly applied to study the escape from a single stable state

to a predefined set of states, if the perturbed species does not participate in the definition of the set of endpoints. If the perturbed species specifies the set of allowed end states, the original endpoint after the parallel shift (i.e., $\mathbf{c}_{B''}$) is no longer a legitimate endpoint of the perturbed system, as it does not lie on the hyper-surface corresponding to the predefined condition.

2.6.3 Quantitative consistency between predictions and direct simulation results

At SOS = 15, we consider stochastic escape from the stable state to the hyperplane where 60% of Ras are activated. As seen in Fig. 2-4, the predictions based on Eqs. (2.5) and (2.6) of the main text are quantitatively consistent with the direct simulation results.

2.6.4 Fluctuations being uniform increases reaction sensitivity

With reaction action and deterministic propensity being equal, the reaction with evenly distributed fluctuations has a larger sensitivity. Note that the reaction action is $S_i^* = \int_0^{\tau^*} (\frac{\hat{\nu}_i^*}{\nu_i^*} \ln \frac{\hat{\nu}_i^*}{\nu_i^*} - \frac{\hat{\nu}_i^*}{\nu_i^*} + 1)\nu_i^* dt$. We define $x_i(t)$ as $x_i(t) = \frac{\hat{\nu}_i^*(t)}{\nu_i^*(t)} - 1$, which measures the normalized deviation of the actual propensity from the deterministic one. Furthermore, we define a function y as $y(x_i(t)) = (x_i + 1)\ln(1 + x_i) - x_i$. With these two new notations, the reaction action can be rewritten as $S_i^* = \int_0^{\tau^*} y(x_i)\nu_i^* dt$. If time t is viewed as a random variable on $[0, \tau^*]$ with probability density $\nu_i^*/\mathbb{Z}_i$ (where $\mathbb{Z}_i$ is the normalization factor $\mathbb{Z}_i = \int_0^{\tau^*} \nu_i^* dt$), then S_i^* is exactly $\mathbb{E}_t[y(x_i(t))]\mathbb{Z}_i$, i.e., the expected value of $y(x_i(t))$ with proper normalization; note here the expectation is taken with respect to time t. Alternatively, we can treat x_i as a random variable, and express S_i^* as an expectation of $y(x_i)$ with respect to x_i:

$$S_i^* = \mathbb{E}_t[y(x_i(t))]\mathbb{Z}_i = \mathbb{E}_{x_i}[y(x_i)]\mathbb{Z}_i$$

where the probability density of x_i is formally $\frac{dt}{dx_i}\nu_i^*/\mathbb{Z}_i$ (if there are multiple t that give the same x_i, sum the above quantity over all such t).

Noticing that the function $y(x_i)$ is a convex function of x_i, we apply Jensen's inequality to obtain the following:

$$S_i^* = \mathbb{E}_{x_i}[y(x_i)]\mathbb{Z}_i \geq y(\mathbb{E}_{x_i}[x_i])\mathbb{Z}_i \tag{2.11}$$

where in the second relation the equal sign can be achieved when x_i is a constant (i.e., zero variance). The argument of the function y in the last expression, $\mathbb{E}_{x_i}[x_i] = \int_0^{\tau^*} x_i\nu_i^* dt/\mathbb{Z}_i = \int_0^{\tau^*}(\hat{\nu}_i^* - \nu_i^*)dt/\mathbb{Z}_i$. Compared to Eq (2.5), $\mathbb{E}_{x_i}[x_i]$ is exactly the sensitivity of reaction i divided by a factor of $\mathbb{Z}_i V$.

Now consider two reactions i and j with equal actions (i.e., $S_i^* = S_j^*$) and equal deterministic propensities (i.e., $\nu_i^* = \nu_j^*$ and hence $\mathbb{Z}_i = \mathbb{Z}_j$). Without loss of generality, assume that reaction i has more uniformly distributed fluctuations (i.e., $\text{var}(x_i) < \text{var}(x_j)$). As seen from Eq (2.11), it suggests that $y(\mathbb{E}_{x_i}[x_i]) > y(\mathbb{E}_{x_j}[x_j])$. Since $y(x)$ monotonically increases on $x \in [0, +\infty]$ and monotonically decreases on $x \in (-1, 0]$, if sensitivities of reactions i and j (i.e., $\mathbb{E}_{x_i}[x_i]$ and $\mathbb{E}_{x_j}[x_j]$) are both positive or both negative, then $\mathbb{E}_{x_i}[x_i]$ is larger in absolute value than $\mathbb{E}_{x_j}[x_j]$. So, the reaction with more uniformly distributed fluctuations ($\text{var}(x_i) < \text{var}(x_j)$, and so $y(\mathbb{E}_{x_i}[x_i]) > y(\mathbb{E}_{x_j}[x_j])$) has a larger absolute sensitivity ($\mathbb{E}_{x_i}[x_i]VZ_i$).

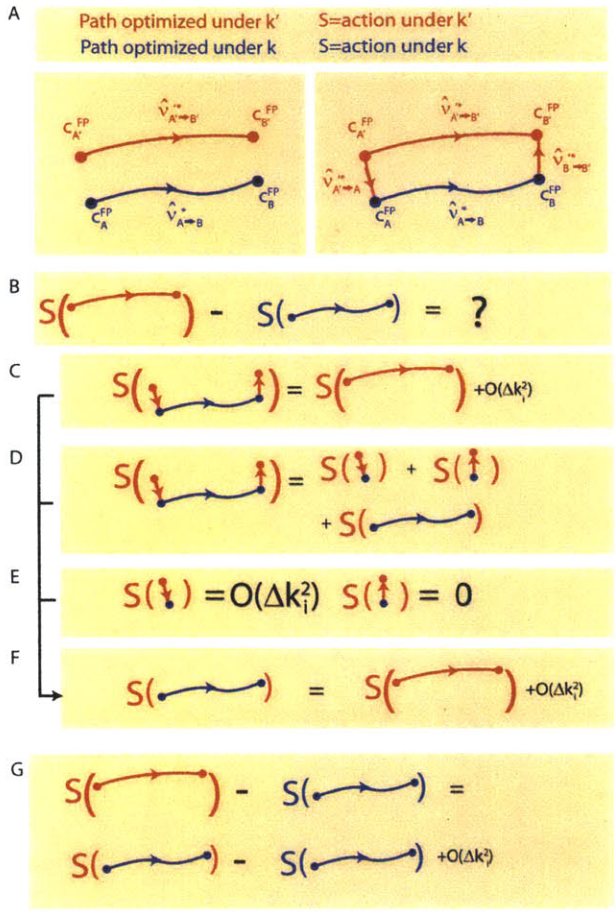


Figure 2-2: A sketch showing the calculation of the sensitivity to rate constants. (A) The optimal paths connecting the steady states before (blue) and after (red) the perturbation in $\mathbf{k}$ (left); the extension paths (right). (B) The change in the minimal action due to the perturbation in $\mathbf{k}$ (i.e., the numerator of Eq (2.7)). (C) The action of the concatenated path is the same as the minimal action after perturbation to the order of $O(\Delta k_i^2)$. (D) The action of the concatenated path is the sum of the actions of its component paths. (E) The actions of the extension paths are $O(\Delta k_i^2)$ and zero. (F) Therefore, the action, evaluated under $\mathbf{k}'$, of the unperturbed optimal path is the same as the minimal action after the perturbation to the order of $O(\Delta k_i^2)$. (G) The change in the minimal action due to the perturbation in $\mathbf{k}$ is the change in action due to the change in $\mathbf{k}$ along the original optimal path, to the order of $O(\Delta k_i^2)$.

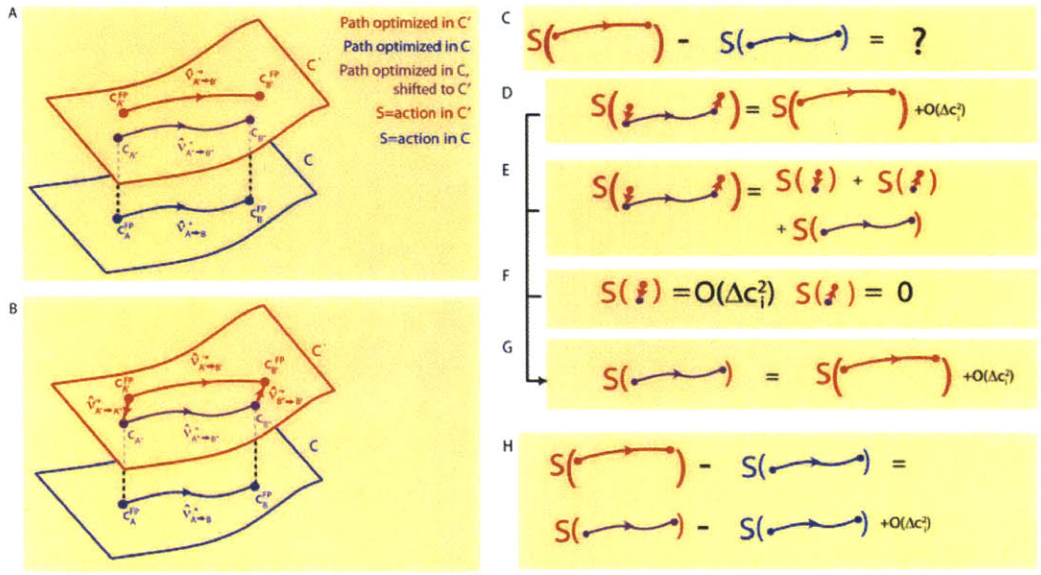


Figure 2-3: A sketch showing the calculation of the sensitivity to species concentrations. (A) The optimal paths connecting the steady states before (blue) and after (red) the perturbation in $\mathbf{c}$; the optimal path before perturbation (blue) shifted pointwise by $\Delta\mathbf{c}$ (purple). (B) The sketch in (A) with extension paths added. (C)The change in the minimal action due to the perturbation in $\mathbf{c}$. (D) The action of the concatenated path is the same as the minimal action after perturbation to the order of $O(\Delta c_i^2)$. (E) The action of the concatenated path is the sum of the actions of its component paths. (F) The actions of the extension paths are $O(\Delta c_i^2)$ and zero. (G) Therefore, the action of the shifted path is the same as the minimal action after the perturbation to the order of $O(\Delta c_i^2)$. (H) The change in the minimal action due to the perturbation in $\mathbf{c}$ is the change in action due to the shift of the original optimal path, to the order of $O(\Delta c_i^2)$.

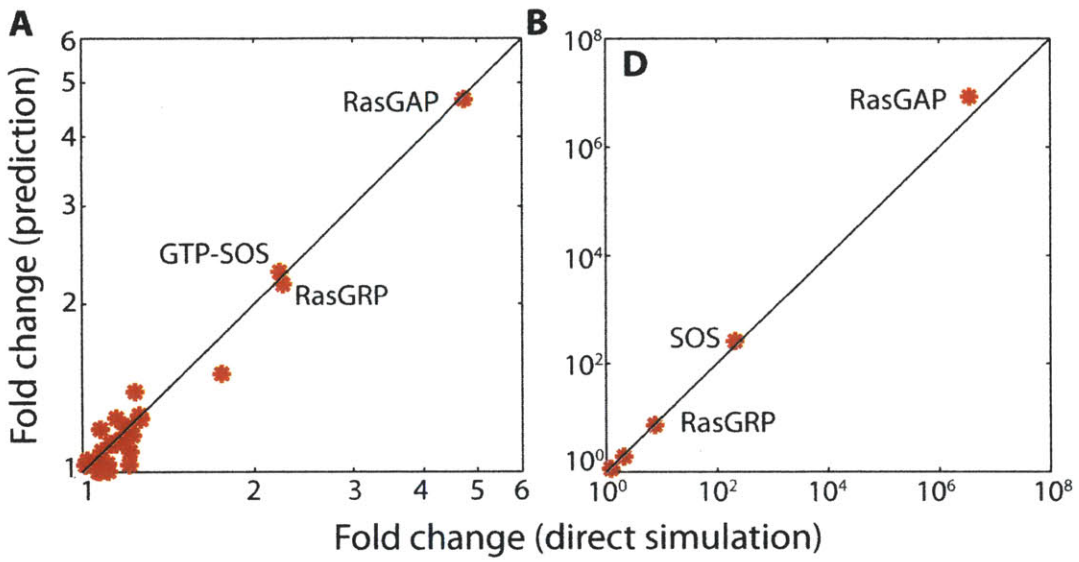


Figure 2-4: Quantitative consistency between predictions and direct simulation results at SOS = 15 (low). (A) Fold change in transition rate (i.e., defined as a larger rate divided by a smaller one) due to 1% increase of rate constants. The transition is considered as stochastic escape from the stable state to the hyperplane where 60% of Ras are activated. Each red dot represents the perturbation of a different rate constant. The direct simulation results are the averages of ten independent FFS runs, and for each FFS run 10^3 points were stored on each surface. Only the results for the key reactions, namely RasGAP catalytic activity, GTP-SOS catalytic activity, and RasGRP catalytic activity are labelled. (B) Fold change in transition rate when one molecule of each type is added to the system. Only the results for the key species, namely RasGAP, SOS, and RasGRP are labelled.

Chapter 3

T-cell Receptor Induced Activation of ERK but not of P38 MAP Kinase Pathways Requires Allosteric Activation of SOS

Activation of the small GTPase Ras is a pivotal regulatory point of antigen receptor-induced MAPK activation in lymphocytes. For Ras regulation, lymphocytes utilize at least two families of Ras Guanine nucleotide Exchange Factors (RasGEFs), namely, SOS and RasGRP. Previously, we have found that RasGRP alone induces analog Ras-ERK activation, while SOS and RasGRP synergize to activate Ras-ERK in a digital pattern. Here we show that binding of Ras-GTP to SOS1 allosteric pocket is critical for this synergy by computational modeling and biological experiments. BCR-induced ERK activation in allosteric pocket mutant SOS1 (W729E)-expressing DT40 cells resembles that of parental SOS-deficient DT40 B cells, displaying unimodal analog ERK activation pattern regardless of dosage and timing of stimulation. Inhibition of RasGRP1 expression and/or activation have great impact on antigen receptor-stimulated ERK activation, but P38 activation is only modestly affected. Optimal activation of P38 requires the presence of SOS1 but not its allosteric pocket.

Redundancy of having two RasGEFs in lymphocytes is important for cells (1) to display greater flexibility and distinct response thresholds to graded stimuli, and (2) to ensure correct regulation of other MAPKs such as P38.

3.1 Introduction

MAP kinase (MAPK) signaling cascades are conserved pathways that can be activated by a wide variety of stimuli and play a role in diverse cellular processes, like cell proliferation, differentiation, or apoptosis [7]. Early on it was recognized that activation patterns of MAPK can vary and can have different biological effects [35]. Specifically, stimulation of Rat adrenal pheochromocytoma (PC-12) cells with neuronal growth factor (NGF) results in sustained activation of the MAPK ERK and differentiation, whereas stimulation with epidermal growth factor (EGF) elicits transient ERK activation and cell proliferation [35]. Rewiring of the feedback loops in the NGF- and EGF-ERK networks subsequently alters the cell fate, further demonstrating a causative link between mode of MAPK activation and cell biological outcome [36]. Similarly, cooperating feedback loops underlie the conversion of graded input signal into binary MAPK activation and an all-or-none biological response, maturation, of the stimulated Xenopus oocytes [37].

In the thymus a stringent selection process ensures the development of functional T cells, while potentially auto-reactive T cells are deleted [38]. Differential activation of the ERK-, P38-, and JNK-MAPK pathways appears to correlate with these opposing cell fates: life of death [38, 39]. The RAF-MEK-ERK signaling cascade downstream of the activated form of the small GTPase Ras has been extensively studied, including studies in T lymphocytes and thymocytes [40, 41]. Ras cycles between GTP-associated active state and GDP-bound inactive state. Active Ras-GTP recruits and activates Raf kinase, which initiates the sequential phosphorylation cascade of RAF-MEK-ERK. Lymphocytes utilize at least two families of Ras guanine nucleotide exchange factors (RasGEFs) for immunoreceptor-induced GTP-loading of Ras, RasGRP and SOS [42]. RasGRP family GEFs are predominantly expressed

in lymphoid cells and mediate calcium- and/or diacylglycerol-responsive Ras activation [42]. SOS RasGEFs are recruited to the plasma membrane where Ras resides via the small adapter molecule Grb2, which binds to phosphorylated tyrosines in adapter molecules or in receptor tyrosine kinases [42]. In T cells, RasGRP1 plays a non-redundant role in activation of Ras/ERK pathway and positive selection of developing T cells and TCR-induced ERK phosphorylation is impaired in Rasgrp1$^{-/-}$ mice [43]. Similarly, deletion of RasGRP1 and RasGRP3 in the chicken DT40 model B cell line greatly impairs B cell receptor (BCR)-induced ERK activation [44, 5]. SOS 1 and SOS2 consist of the second GEF family in lymphocytes. In knockout mice, SOS1 is essential for normal placental development and survival of embryo, while SOS2 is dispensable. Conditional inactivation of SOS1 in thymocytes impairs receptor-induced ERK activation only moderately. By combining cellular and biochemical approaches with computational models, we established that RasGRP and SOS synergize to activate Ras/ERK pathway in lymphocytes. Significantly, we demonstrated that RasGRP activates ERK in a graded, or analog, manner, whereas SOS (with RasGRP) does so in a digital, or bimodal way [6]. The synergy arises from RasGTP produced by RasGRP which subsequently can bind to, and allosterically activate, SOS [5, 6]. The allosteric binding pocket for Ras had been previously identified in the crystal structure of the catalytic domain of SOS1 with Ras [45] and loading of this pocket by RasGTP induces a conformational switch [46]. The resultant RasGTP-SOS-Ras positive feedback loop operates in both EGF-stimulated Hela cells 13 that do not express RasGRP1 [5] and in a very efficient manner in TCR- or BCR-stimulated lymphocyte cell lines that are positive for RasGRP1 expression [5, 6]. We hypothesized that coinciding expression of RasGRP and SOS in lymphocytes offers the capacity to these cells to signal to ERK in two modes, analog and digital, the latter as a result from passing a threshold of signal input so that SOS is allosterically activated [47]. It is not know how disruption of the feedback to SOS allosteric pocket might affect ERK activation.

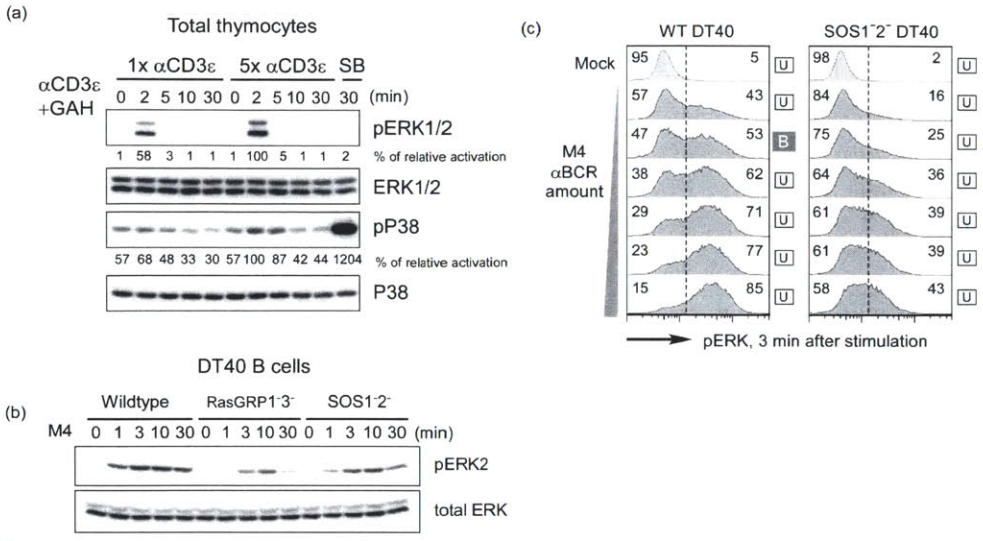
Activation of the P38 and JNK MAPK pathways is also critical for developing T cells, but the molecular mechanism of activation in lymphoid cells is largely unknown

[48, 49]. Analyses of mouse models have provided fascinating evidence suggesting that different Ras-MAPK pathways might drive positive versus negative selection [50]. In Grb2 heterozygous mice P38 and JNK activation and negative selection is impaired, while ERK phosphorylation and positive selection is intact [51]. Conversely, ERK-1 and-2 doubly deficient thymocytes demonstrate impaired positive selection [52], but negative selection is intact [53]. Most notably, pharmacological inhibition of P38 MAPK bocks negative selection of thymocytes in fetal thymic organ cultures. The mechanistic details of how upstream signaling molecules might orchestrate coordinate activation of these MAPK pathways are unknown. Here we utilized biochemical approaches with model cell lines and primary lymphocytes in combination with computational approaches to examine the potential role of allosteric activation of SOS regulating the quantity and quality of Ras and ERK-, P38-, and JNK-MAPK signal output.

3.2 Results

3.2.1 Both RasGRP and SOS contribute to optimal ERK activation

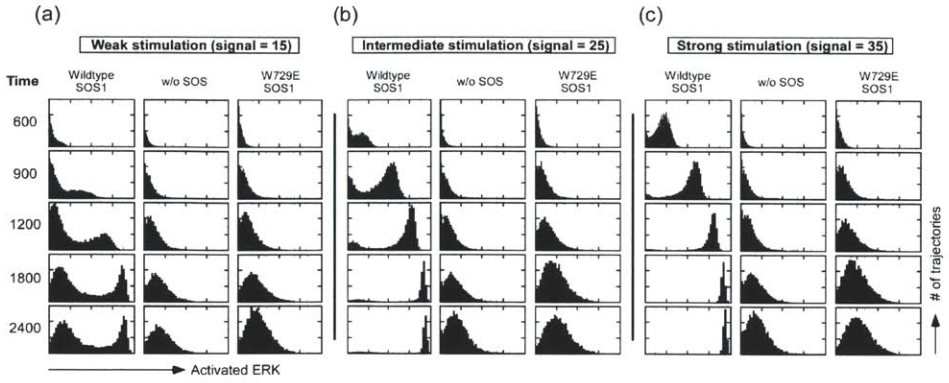
It has been suggested that developing T cells (thymocytes) require lower levels of TCR signal input to activate ERK kinases and higher levels of TCR stimulation to also activate P38 and JNK [38, 49, 54]. To test this we isolated thymocytes from mice and stimulated these with a low or high dose of anti-CD3ϵ crosslinking antibody, which simulates a TCR stimulus. Whereas ERK phosphorylation was readily induced by low levels of TCR stimulation, phosphorylation of P38 critically depended on a high dose of stimulating anti-CD3ϵ antibody. Phosphorylation of JNK was efficiently induced by sorbitol exposure but was not triggered by high doses of anti-CD3ϵ antibody (Fig. 3-1a). Similarly, wild-type chicken DT40 B cells activate ERK in a very sensitive manner, following anti-BCR stimulating antibody. Both the presence of RasGRP or SOS RasGEFs is required for triggering high levels of ERK activation at the



Jun et. *al*, Figure 1 **MAPK signaling depends on stimuli dosage and SOS**

Figure 3-1: MAPK ERK activation pattern depends on stimuli dosage and SOS. (a) TCR stimulation of total mouse thymocytes results in activation of MAPK ERK and P38. Note that ERK activation is much more robust. SB indicates 30 min stimulation with 0.4 M sorbitol as a positive control for phospho-P38. (b) Deficiency in either RasGRP or SOS leads to impaired ERK activation in DT40 B cells. Wild-type, RasGRP1-3- and SOS1-2- DT40 B cells were activated with maximal aBCR stimulation for indicated time. (c) SOS1/2-deficient DT40 B cells do not demonstrate a bimodal P-ERK pattern (indicated as B) after 3 min BCR stimulation, regardless of strength of BCR stimulus (range is 1:32,000 to 1:1,000 dilution). Bimodality is statistically tested by Hartigans analysis [10]. Such analyses indicated that only in wt DT40 cells bimodal pattern of ERK phosphorylation can be achieved.

population level (Fig. 3-1b). We previously reported that the bimodality in ERK activation disappears in the absence of SOS1/SOS2, following a time course with two different doses of BCR-stimulating antibody [6]. Expanding on these findings, we established that not a single dose or timepoint of BCR stimulation resulted in bimodal ERK activation when SOS RasGEFs are missing (Fig. 3-1c and data not shown). We hypothesized that the lack of bimodal ERK activation in SOS deficient cells is caused by the absence of the allosteric feedback loop to SOS and tested this hypothesis next.



Jun et. *al*, Figure 2 **Predicted patterns of pERK - allosteric activation**

Figure 3-2: In silico prediction of ERK activation pattern in cells with wild-type SOS1, without SOS or with allosteric mutant W729E SOS1 expression. ERK activation is projected at different time points with three different stimuli levels, namely, (a) weak stimulation, (b) intermediate stimulation, and (c) strong stimulation.

3.2.2 In Silico modeling: Influece of SOS1 allosteric pocket on ERK activation

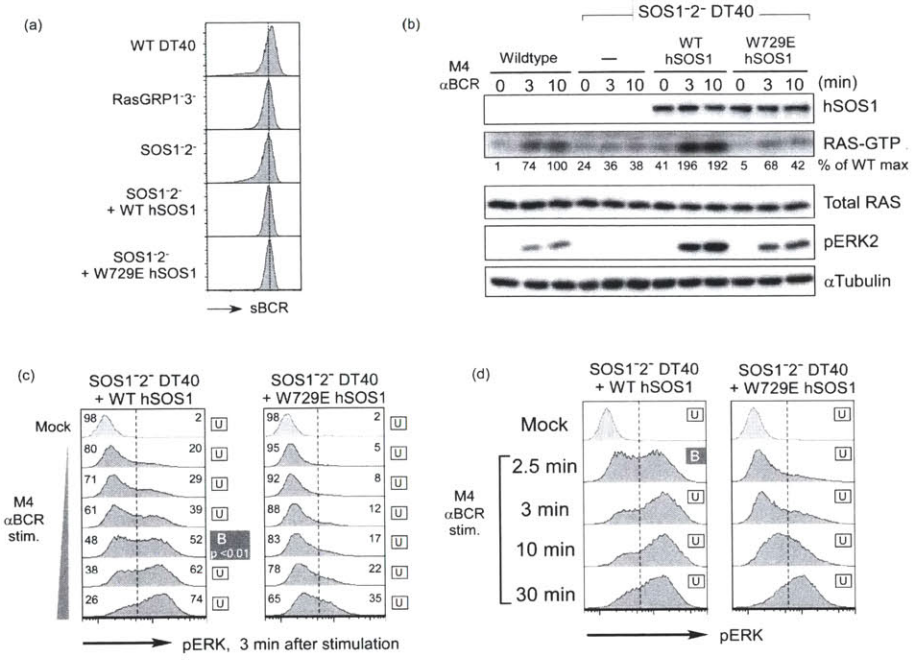
Here we utilized established computational models (Fig. 3-9) that stochastically simulate the Ras-ERK activation [6, 8, 9] to explore the effect of SOS, in particular, SOS' allosteric pocket, in ERK activation.

We modeled the ERK activation in cells with wild-type SOS1, without SOS, or with allosteric-mutant SOS (i.e., W729E) by abrogating the binding of RasGTP to SOS' allosteric pocket. We investigated ERK activation profiles at different time points with three different stimuli levels, namely, (a) weak stimulation, (b) intermediate stimulation, and (c) strong stimulation (Fig. 3-2). While the bimodal pattern of ERK activation is observed in cells with wild-type SOS1, bimodality is not seen at any dosage level and any time points without SOS or without intact SOS' allosteric pocket.

3.2.3 An intact allosteric pocket in SOS1 is required for bimodal ERK activation

Based on structural analysis of SOS protein, tryptophan residue 729 in SOS1 (W729) was identified as a key residue coordinating Ras binding within allosteric pocket. Single amino acid alteration of tryptophan into glutamic acid (W729E) cripples allosteric activation of nucleotide exchange activity of the isolated catalytic domain of SOS1 [6]. We hypothesize that BCR-induced digital ERK activation requires the synergy between RasGRP and SOS, and that SOS allosteric pocket enables positive feedback enhancement of RasGEF activity primed by RasGRP. To directly test this directly, we stably reconstituted SOS1-2- DT40 cells with either WT or W729E (allosteric pocket mutant) human SOS1 (hSOS1). Selected stable clones expressed comparable levels of surface B cell receptor, similar to that on WT DT40 and other DT40 mutants (Fig. 3-3a) and WT- or W729E-hSOS1 was expressed at comparable levels (Fig. 3-3b). We first analyzed Ras- and ERK-activation at the population levels and established that WT hSOS1 reconstitution not only rescued the BCR-induced Ras and ERK activation defect in SOS1-2- DT40 cells, but also augmented the magnitude of activation (Fig. 3-3b). By contrast, W729E hSOS1-reconstituted cells demonstrated significantly lower levels of activated Ras and ERK compared to the WT hSOS1 stable clone (Fig. 3-3b).

Moreover, BCR stimulation of stably reconstituted SOS1-2- -WT hSOS1 cells resulted in a rescue of the bimodal ERK activation pattern, revealed by dose-responses and time courses of BCR-induced phospho-ERK flow cytometry and Hartigans analyses of the resulting phospho-ERK histograms (Fig. 3-3c and d) [10]. By contrast, reminiscent of parental SOS1-2- DT40, bimodal ERK activation was never observed in SOS1-2- -W729E hSOS1 DT40 cells (Fig. 3-3c and d). Expression of W729E hSOS1 did not lead to a general impairment of ERK activation as stimulation of cells with the synthetic analog of DAG (PMA) resulted in very similar ERK activation in WT DT40, SOS1-2- -WT hSOS1, and SOS1-2- -W729E hSOS1 cells (Supplemental figure). Likewise, the observed results did not reflect uniquely selected stable clones,



Jun et. *al*, Figure 3 SOS1 allosteric pocket is required for magnitude and pattern of Ras-ERK signaling

Figure 3-3: Efficient and bimodal Ras-ERK signaling requires an intact allosteric pocket in SOS1. (a)Surface BCR levels of various DT40 cell lines. SOS1/2-deficient DT40 cells stably reconstituted with wildtype or W729E hSOS1 express comparable surface BCR level to wildtype or other mutant DT40 cells. (b) BCR induced Ras and ERK activation depends on SOS and on SOS allosteric pocket. Note the decrease in W729E hSOS1, compared to wt hSOS1. (c/d) Reconstitution of SOS1SOS2 deficient DT40 B cells with the W729E allosteric pocket mutant does not lead to rescue of the bimodal pattern, regardless of dose or time course of stimulation. Bimodality statistically supported by Hartigans analysis is marked as B.

additional WT and W729E hSOS1 stable clones were tested in parallel and exhibited similar dynamics and patterns of BCR-induced P-ERK (Data not shown).

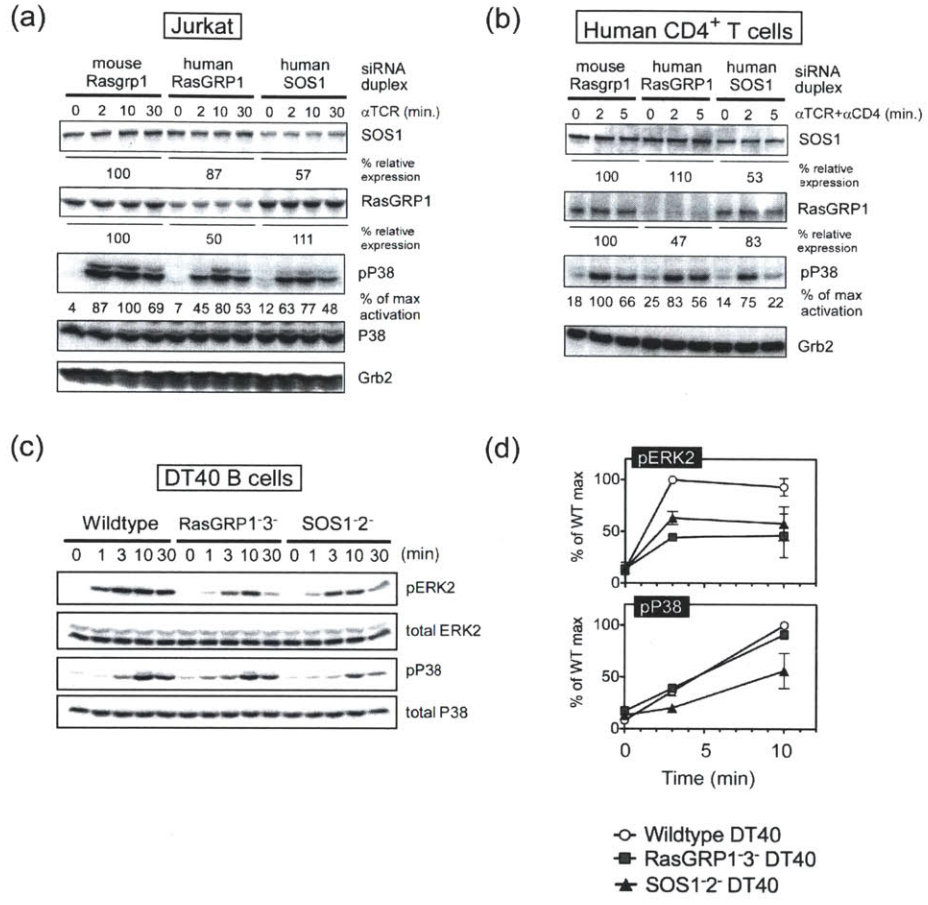
3.2.4 Optimal P38 activation preferentially requires SOS1

Pharmacological inhibition of P38 MAPK bocks negative selection of thymocytes in fetal thymic organ cultures. Negative selection and P38 activation has been reported to be unaffected in Rasgrp1 deficient mice [55]. However, these assays were performed with thymocytes that were stimulated with PMA, which robustly activates the RasGRP pathway but does not engage the SOS pathway. Thus it is difficult to

make a definitive conclusion based on these assays. In Grb2 heterozygous thymocytes, TCR-induced P38 activation is impaired [51] suggestive of a possible role for SOS that is recruited by Grb2. However, the effects of targeted deletion of Sos1 in thymocytes on P38 activation have not been explored [56]. It thus remains to be determined if and how these RasGEFs might affect the P38 MAPK pathway and how this compares to their roles on the ERK pathway. To address this question we determined TCR-induced P38 activation in Jurkat T cells or human CD4+ T cells in which RasGRP1 or SOS1 expression were targeted by siRNA duplexes. In contrast to the reported more dominant role for RasGRP1 in terms of ERK activation in these cells compared to SOS1 [5], we found the opposite here. TCR-induced activation of P38 was consistently more impaired by the reduction in SOS1 expression than by the reduction in RasGRP1 expression in both Jurkat T cells and human CD4+ T cells (Fig. 3-4a and b). Complete absence of RasGRP-1 and -3 or of SOS-1 and -2 in our RasGRP1-3- and SOS1-2- DT40 cell lines illustrated this point more clearly, the SOS RasGEFs made a larger contribution towards BCR-induced P38 activation than the RasGRP RasGEFs, whereas the opposite holds true for activation of ERK MAPK (Fig. 3-4c). Given the differences between RasGRP and SOS in defective ERK versus P38 activation, one hypothesis would be that activation of ERK and P38 occurs mechanistically different downstream of these RasGEFs. We next set out to test this hypothesis.

3.2.5 The PLCγ-Rasgrp1 axis contributes only minimally to TCR-induced P38 activation

We next utilized a pharmacological approach to inhibit phospholipase C (PLC) with U73122, the enzyme that produces the second messengers IP3 and DAG that are essential for RasGRP1 membrane recruitment and activation in thymocytes and T cell lines [42]. As expected based on the established role of the DAG-RasGRP1-Ras pathway towards ERK activation, exposure of total mouse thymocytes to U73122 reduced the level of ERK activation following TCR stimulation, whereas the U73343 inert ana-



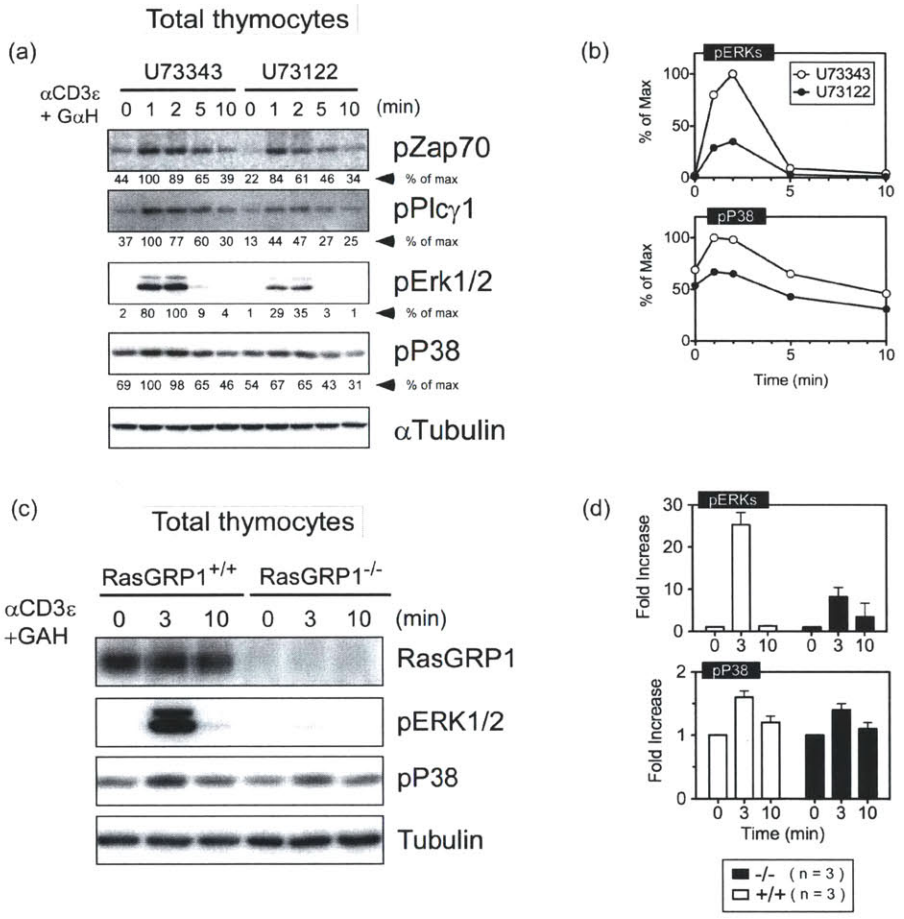
Jun et. *al*, Figure 4 Optimal P38 activation requires both SOS1 and RasGRP1

Figure 3-4: Optimal P38 activation preferentially requires SOS1. (a) siRNA-driven knockdown of SOS1 in Jurkat T cells or (b) in human peripheral blood CD4 T cells results in a reduction of TCR-triggered P38 activation that is more pronounced than when RasGRP1 expression levels are reduced by siRNA-driven knockdown. Note these P38 results are diametrically opposite from what we have reported for ERK activation where RasGRP1 is more dominant. (c/d) Impaired BCR-induced P38 activation in the absence of SOS. Normal BCR-induced P38 activation in the absence of RasGRP.

log as a control had no measurable effects (Fig. 3-5a and b). Significantly, U73122 had a much more modest effect of the TCR-induced activation of P38 (Fig. 3-5a and b). We next took a genetic approach by comparing thymocytes from Rasgrp1$^{-/-}$ mice with those from β2M/MHC class II doubly deficient mice. In both cases thymocyte development is arrested and thymi typically consist of a rather uniform cell population of CD4+CD8+ thymocytes 23. In Rasgrp1$^{-/-}$ mice the defect is intrinsic to the thymocytes, whereas in the β2M/MHC class II doubly deficient mice the defect lies in the epithelial cells and not in the thymocytes. As previously reported, ERK activation is severely impaired in TCR-stimulated CD4+CD8+ thymocytes that are deficient for Rasgrp1 [43] (Fig. 3-5c and d). By contrast the modest TCR-induced P38 activation that can be observed in wild type CD4+CD8+ thymocytes is only minutely/not significantly affected by the genetic deletion of Rasgrp1 (Fig. 3-5c and d). Thus the mechanism of P38 activation appears to be very different from ERK activation and occurs mostly independent of Rasgrp1.

3.2.6 SOS1 allosteric pocket is dispensable for BCR-stimulated P38 activation

We have previously demonstrated that RasGTP produced by RasGRP1 can allosterically activate SOS1 to produce more RasGTP, which subsequently leads to a digital ERK signal downstream [5, 6]. Having observed that RasGRP1 plays only a minimal role in antigen receptor-induced P38 activation, we focused on SOS next and tested if the feedback loop to the allosteric pocket in SOS1 might be involved and how uncoupling this loop via the W729E point-mutation in SOS1 affects the activation of P38. Under conditions of strong BCR stimulation, the earlier described defect in P38 activation in SOS1$^{-/-}$SOS2$^{-/-}$ DT40 B cells was restored by reconstitution with WT hSOS1. Unexpectedly, a very similar level of restoration was obtained in SOS1$^{-/-}$SOS2$^{-/-}$ -W729E-hSOS1 cells (Fig. 3-6a). W729E-hSOS1 reconstituted cells demonstrated P38 activation that was very similar to that in WT hSOS1 reconstituted cells regardless of the dose or length of the BCR stimulation,



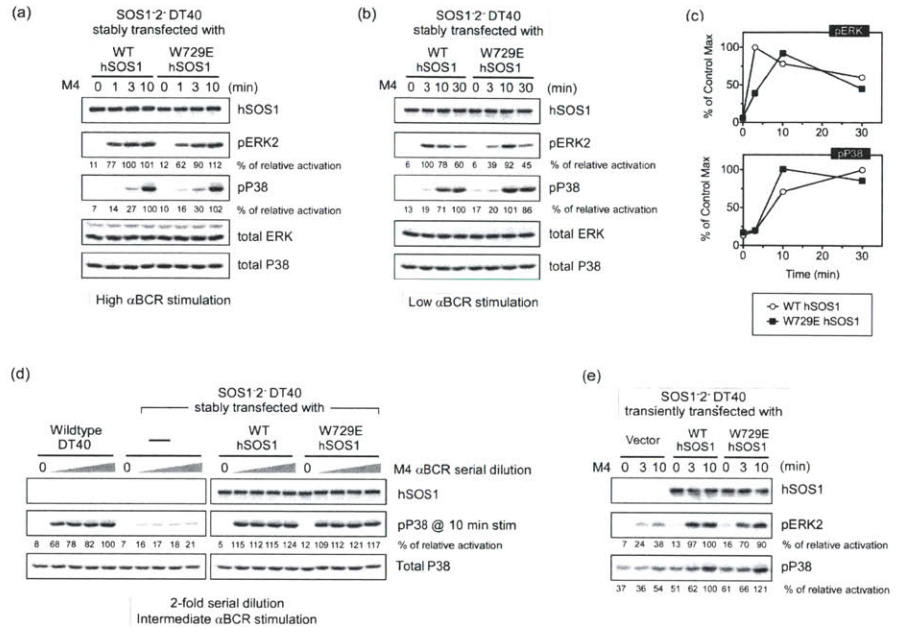
Jun et. *al*, Figure 5 PLCγ-RasGRP1 axis contributes only minimally to P38 activations in thymocytes

Figure 3-5: PLCg-RasGRP1 axis contributes only minimally to P38 activation in thymocytes. (a/b) Pharmacological inhibition of (DAG-producing) PLCγ by U73122 results in a modest reduction of P38 activation compared to the reduction in ERK activation in TCR-stimulated total thymocytes. U73343, inert analog is used as a treatment control. (c/d) Minute reduction in TCR-induced P38 activation in Rasgrp1-deficient thymocytes. RasGRP1$^{+/+}$ thymocytes were obtained from MHCI/II-double deficient mice.

demonstrating that the allosteric pocket in SOS1 is dispensable for optimal BCR-triggered P38 activation (Fig. 3-6b, c, and d). These results were further validated with transient reconstitution of $SOS1^{-/-}SOS2^{-/-}$ cells with vector, WT hSOS1, or W729E-hSOS1. Again, W729E-hSOS1 rescued the P38 activation defect with the same efficiency as WT hSOS1 (Fig. 3-6e). Importantly, ERK activation was always lower in W729E-hSOS1 compared to WT hSOS1 expressing cells, reiterating the importance of allosteric activation of SOS1 for optimal ERK activation (Figs. 6a-6d). The transient transfection approach also offered the opportunity to investigate if Ras exchange activity of SOS1 is required for P38 activation. Despite numerous attempts, we were unable to generate $SOS1^{-/-}SOS2^{-/-}$ cells stably expression F929A-hSOS1, the RasGEF mutant form of hSOS1.

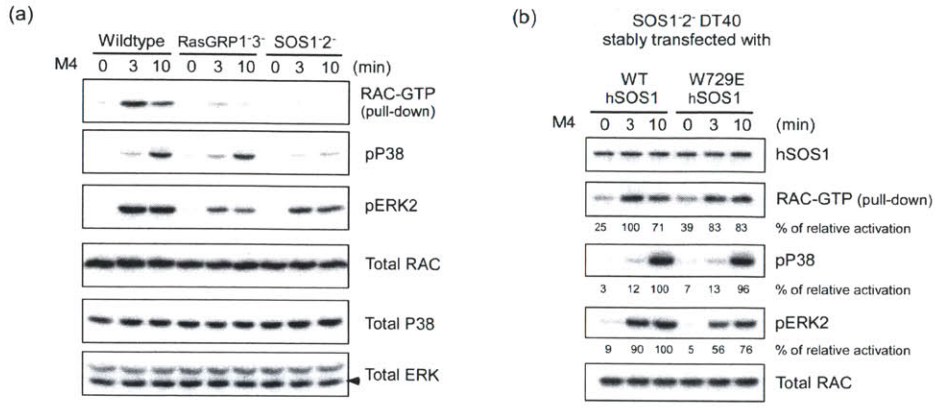
3.2.7 SOS1, Rac, and P38 activation

As an initial start to map a pathway from SOS1 towards P38 in lymphoid cells we first examined activation of the small GTPase Rac. Rac is typically positioned upstream of the MKK3/6-P38 MAPK pathway, although most research has been done in non-lymphoid cell types [48, 54]. In agreement with the impairment in BCR-induced P38 activation, GTP-loading of Rac induced by BCR stimulation was reduced in $SOS1^{-/-}SOS2^{-/-}$ DT40 B cells (Fig. 3-7a). We also detected a Rac activation defect in $RasGRP1^{-/-}RasGRP3^{-/-}$ cells, but this was more modest than in the $SOS1^{-/-}SOS2^{-/-}$ DT40 B cells, paralleling the requirements for P38 activation and contrasting the need for RasGRP or SOS in terms of ERK activation (Fig. 3-7a). Moreover, stable reconstitution of $SOS1^{-/-}SOS2^{-/-}$ cells with W729E-hSOS1 resulted in a significant rescue of BCR-induced Rac activation, approximating the levels obtained in WT hSOS1 reconstituted cells (Fig. 3-7b).



Jun et. *al*, Figure 6 SOS1 allosteric pocket is dispensable for optimal P38 activation

Figure 3-6: SOS1s allosteric pocket is dispensable for optimal P38 activation. (a/b/d) W729E hSOS1 and wt hSOS1 similarly rescue the SOS1SOS2 deficient DT40 in terms of P38 activation regardless of stimuli dosage. Note that ERK activation defect is only obvious at early time point in response to maximal αBCR stimulation (a). At lowest αBCR stimulation, ERK activation defect is more evident. (d) P38 activation is measured in response to two fold serial dilution of BCR antibody at a range of intermediate dosage. (e) SOS1-2- DT40 cells were transiently transfected with empty vector, WT hSOS1 or W729E hSOS1 and stimulated as indicated.



Jun et. *al*, Figure 7 **SOS is required for BCR-induced RAC activation**

Figure 3-7: BCR-induced RAC activation requires SOS but not allosteric activation of SOS1. (a) RAC activation in response to high αBCR dosage is measured by RAC pull-down assay. (b) RAC activation in stably hSOS1-reconstituted DT40s is measured by RAC pull-down assay. Low αBCR dosage is used for stimulation.

3.2.8 Computational prediction: roles of SOS in Rac-P38 activation

Lastly, we built a coarse-grained computational model to speculate the roles of SOS in Rac-P38 activation. First, notice the differences in activation profiles of Rac and P38 (Fig. 3-7a): RasGRP1$^{-/-}$RasGRP3$^{-/-}$ cells had a significant defect in Rac activation but a fairly robust P38 activation; in contrast, SOS1$^{-/-}$SOS2$^{-/-}$ cells had even lower Rac activation and that amount of Rac activation failed to activate P38 effectively. We hypothesized two possible mechanisms to explain this: (1) Rac activated by RasGRP-mediated pathways could be localized in a different compartment from that activated by SOS-mediated pathways, and SOS-activated Rac is more efficient in activating P38; (2) the amount of active Rac might need to overcome a threshold to effectively activate P38, and RasGRP1$^{-/-}$RasGRP3$^{-/-}$ can generate enough active Rac to cross that threshold, but SOS1$^{-/-}$SOS2$^{-/-}$ cells cannot. While the first hypothesis can be tested by imaging experiments that track the location of active Rac, the second can be realized by introducing non-linearity in Rac-P38 activation. We adopted the second

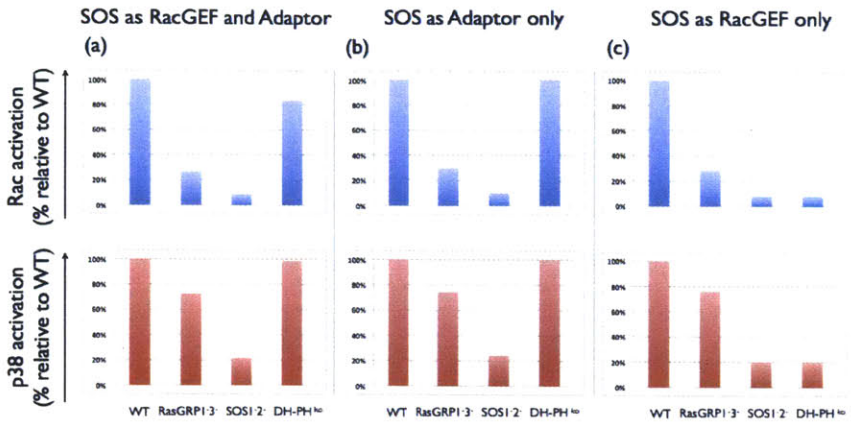


Figure 3-8: Computational predictions for SOS activating Rac-P38 pathway. When SOS functions as (a) both a RacGEF and an adaptor, (b) an adaptor only, and (c) a RacGEF only, the activation profile of Rac-P38 for cells with DH-PH domain mutant SOS is (a) between that for wild-type SOS and that for $SOS1^{-/-}SOS2^{-/-}$, (b) the same as wild-type SOS, and (c) the same as $SOS1^{-/-}SOS2^{-/-}$.

hypothesis in our computational modeling, and a Hill-coefficient of two was sufficient to produce the thresholding effect observed in Fig. 3-7a.

Second, we investigated whether SOS functioned as a RacGEF, an adaptor (i.e., a structural component connecting BCR-generated signals to RacGEF activation), or both. We speculated that SOS might possess RacGEF activity through its DH-PH domain [57] and/or help the recruitment and stabilization of RacGEF (such as Vav) at the LAT signalisome. Furthermore, we assumed that RasGRP facilitates the stable formation of LAT signalisome by co-operativity [47] and hence regulates Rac activation via RacGEF (such as Vav) that is recruited to LAT. By computational modeling, we predicted the Rac-P38 activation profiles for SOS with DH-PH mutations (i.e., abrogating SOS RacGEF activity) under three scenarios (Fig. 3-8): (a) SOS is both a RacGEF and an adaptor, (b) SOS is only an adaptor, (c) SOS is only

a RacGEF. For (a), the DH-PH mutation leaves SOS adaptor functionality intact and hence the Rac-P38 activation profile for DH-PH mutants falls between that for wild-type and that for $SOS1^{-/-}SOS2^{-/-}$; for (b), the DH-PH mutation does not alter SOS functions and the DH-PH mutant has the same Rac-P38 activation profile as wild-type; for (c), the DH-PH domain mutation completely destroys SOS functions and the DH-PH mutant is as deleterious as $SOS1^{-/-}SOS2^{-/-}$. Therefore, by experimentally testing the effect of DH-PH mutation on the Rac-P38 activation, we may unveil the role of SOS in regulating the Rac-P38 pathway.

3.3 Discussion

In response to gradual increase of stimuli, Ras/MAPK signaling displays a great flexibility in ERK signaling pattern and dynamics. MAPK cascade is suggested to be intrinsically capable of showing switch-like highly cooperative activation [58, 59]. Additionally, MAPK cascade-extrinsic mechanisms also contribute to flexible shaping of MAPK signaling outcome. These mechanisms include Ras nano-clusters [60], dual negative feedback control by SHP1 [61], sub-cellular location of cascade activity [62], or scaffold-mediated signal quality change [63]. Previously, we also proposed SOS allosteric pocket-mediated interplay between RasGRP and SOS generates highly sensitive digital ERK activation in lymphocyte antigen receptor signaling [5, 6].

In the present study, we directly tested our proposed model by both computer-based simulations and biological experiments involving reconstitution of full-length human SOS1 (hSOS1) into $SOS1^{-}2^{-}$ DT40 B cells. As experiment readouts, we measured ERK phosphorylation in both population (Western blot) and in individual cell level (phospho-flow cytometry). Our results clearly demonstrated that disruption of allosteric activation of SOS is sufficient to alter the quantity and quality of MAPK ERK signaling in DT40 B cells. In spite of comparable level of sBCR and transfected protein expression, W729E mutant SOS1 cells signal to ERK in an analogous manner, much reminiscent of $SOS1^{-}2^{-}$ cells (Fig. 3-3). DT40 B cells expressing SOS1 bearing mutated allosteric pocket did not show robust bimodal ERK activation pat-

tern regardless of dosage nor the duration of stimulation (Figs. 3-3C and D). Our independent computer simulation also supports experimental findings (Fig. 3-2).

We and others previously demonstrated RasGTP binding to SOS allosteric pocket potentiates Ras-ERK pathway signal output [6, 64]. But, our present study shows for the first time the role of SOS allosteric activation in regulating highly sensitive (digital) MAPK ERK activation in lymphocytes. Both DT40 B cells and mouse primary thymocytes show greater dependence on RasGRP1 for robust activation of ERK [43, 44, 5]. In our tested model, RasGRP activation temporally precedes the activation of SOS, catalyzing initial RasGTP accumulation priming SOS via allosteric activation. The W729E mutation within SOS1 REM domain disrupts RasGTP binding to allosteric pocket and uncouples positive feedback loop. However, we cannot completely rule out factors, other than defective positive feedback loop and suboptimal catalytic capability of SOS1, can also contribute to attenuated ERK signal output in W729E mutant SOS1-expressing DT40 B cells. For example, plasma membrane Ras nano-clusters have been identified and are suggested to convey analog stimuli input into a given unit of Ras activity, effectively functioning as analog-digital converter [60, 65]. Impaired binding of Ras at the allosteric pocket of SOS1 may destabilize SOS1 at the plasma membrane, perturbing SOS1 interaction with Ras nano-clusters.

Previous studies in non-lymphoid cells anecdotally suggested correlation between SOS and Rac GEF activity [66, 67, 68]. However, it was not known whether SOS connects antigen receptor to activation of RAC and its downstream MAPK P38. Only hints came from studies of mouse thymocytes with reduced expression or complete loss of Grb2, that show impaired P38 and JNK activation [51, 69]. In present study, BCR-induced Rac and P38 activation in SOS1^{-}2^{-} DT40 cells is severe reduced (Fig. 3-7A). This was surprising because P38 activation in RasGRP1^{-}3^{-} DT40 cells is relatively normal in spite of more severe defect in activating ERK (Fig. 3-7A). Similarly, thymocytes deficient of RasGRP1 exhibited comparable CD3-stimulated P38 activation to RasGRP1+ thymocytes although ERK activation was greatly diminished without RasGRP1 expression (Fig. 3-5 C&D). Inhibition of RacGTP accumulation correlates with impaired BCR-induced P38 and JNK activity in DT40 cells [70]. Therefore,

wildtype-comparable P38 activation in RasGRP1^{-}3^{-} DT40 cells, in spite of reduced Rac-GTP accumulation, was unexpected (Fig. 3-7). Additionally, allosteric SOS1 activation appears to be dispensable for optimal BCR-induced P38 activation although Rac activation is not fully rescued without intact distal Ras binding pocket (Figs. 3-6 & 3-7B). Based on our results, sub-optimal Rac-GTP level is enough to support full activation of P38 in response to BCR stimulation as long as full-length SOS1 protein with intact catalytic Ras binding site is present.

Due to low sensitivity and robustness of P38 activation and difficulty with RNAi in primary cells, we could not fully examine whether SOS-RAC-P38 connection stays true only in DT40 B cells. Our finding has to be further validated in SOS-deficient primary lymphocytes to rule out whether DT40 B cell lines show particular dependence on SOS for BCR-RAC-P38 pathway. Nevertheless, it is interesting to contemplate how SOS couples BCR to RAC-P38 pathway. As a mean of unbiased evaluation, we used computer-based simulations to test three possible scenarios: (1) SOS itself as a RacGEF, (2) SOS as a structural component, and (3) SOS as both a RacGEF and a structural component for other RacGEFs.

Structurally SOS has dual GEF potential both as Ras GEF and Rac GEF in which SOSs atypical Dbl homology domain might perform the exchange reaction of GDP for GTP [57]. However, purified SOS1 alone has never been shown to catalyze GTP loading of RAC (personal communication with John Kuriyan). Instead, distinct SOS1 complex in association with EPS8 and E3b1/Abi-1 co-factors was shown to have in vitro RacGEF activity [67]. To serve as a RacGEF, SOS has to dissociate from Grb2 and form alternate complex with EPS8 and E3b1[71]. However, to this date it remains unknown whether such mechanism is also operating in antigen receptor-induced RAC and MAPK activations in lymphocytes.

Alternatively, SOS effect on Rac could be indirect, through the Rac GEF activity of another molecule(s) in a SOS-dependent manner. Complex of Grb2 and SOS1 at 2:1 ratio was recently shown to mediate clustering of multiple LATs and associated proteins in T cells [72]. Attenuation of LAT clustering impairs TCR signaling more especially at low grade TCR stimulation. Due to negative cooperativity, dual Grb2-

SOS1 complex only forms when Grb2 exists in molar excess compared to SOS1, therefore molar quantity of Grb2 may serve as a threshold/bottleneck determining degree of LAT aggregation. It remains to be verified whether Grb2-SOS1 facilitates similar clustering of LAB in B cells, but failure to promote LAB clustering may lead to severe impairment in RAC-P38 activation pathway. In this respect, impaired JNK and P38 activation in Grb2-heterozygote null thymocytes is note-worthy [51].

In the second scenario we modeled, SOS provides a structural role in coupling upstream antigen receptor signal to downstream activation of RacGEF(s). Candidate RasGEFs include Tiam1, Vav and Dock2 [73, 74]. First, Tiam1 is so far described as a RacGEF required for chemokine-induced Rac activation in mouse T cell, and it remains to be revealed whether Tiam1 is also connected to antigen receptor [75]. However, Tiam1 is unique in that its RacGEF activity is suggested to be directly stimulated by binding to RasGTP [76]. Allosteric pocket mutant hSOS1 expression incompletely rescues BCR-stimulated RAC-GTP accumulation in SOS1-2- DT40 B cells (Fig. 3-7B). We can speculate RacGEF connected to SOS might take inputs from upstream Ras activity as well. In this regard, Ras activity-sensing RacGEF like Tiam1 is a good candidate.

Second, Vav family proteins are the best studied Rac-GEF downstream of lymphocyte antigen receptor [77]. Unlike Tiam1, Vav is not directly responsive to RasGTP levels. However, Vavs function relies on membrane-recruitment where it can activate Rac, and its GEF activity is regulated indirectly by Ras via PI3Kinase [73, 78]. Additionally, several studies indicate crosstalk between Vav and RasGRP1 [79, 80, 81]. Third, CDM (Caenorhabditis elegans Ced-5, mammalian DOCK180 and Drosophila melanogaster myoblast city) family proteins have non classical RhoGTPase GEF activity [82]. Dock2 is a member of CDM family highly expressed in lymphocytes [74]. Dock2 is required for TCR-induced Rac activation [83]. Interestingly Dock2 is rapidly tyrosine phosphorylated following TCR stimulation [84]. These facts taken together suggests Dock2 is spatially in close proximity to LAT-SLP76 signalosome, making it a plausible RacGEF connected to SOS through TCR/BCR.

3.4 Materials & Methods

3.4.1 Cell lines, mice, stimulations and inhibitor treatment.

Cultures of human Jurkat leukemic T cells, chicken DT40 B cell lines and DT40-derived lines were maintained as described before [44, 5, 6]. All experimental mice were used at the age of 6-7 weeks. RasGRP1-deficient mouse was obtained from James Stone. Age and sex-matched MHCI/II-double deficient (Abb/b2m) mice were purchased from Taconic (Hudson, NY). For cell stimulation, harvested cells were rested for 30 minutes in PBS or plain RPMI at 37°C. For PLCγ inhibition, cells were preloaded for 20 minutes with U73122 inhibitor (Calbiochem) or its inactive analog U73343 at 5 μM (Calbiochem). Stimulations of DT40 cells were carried out in PBS or RPMI at 37°C with the indicated doses of M4 aBCR antibody hybridoma ascites fluid preparation.

3.4.2 Hartigan's analysis

We previously employed Hartigans analyses of P-ERK histograms to label these as unimodal or bimodal [6]. Here we used the average of cell counts in the neighboring five gates to smooth the cell count data in each gate in the pERK flow cytometry experiment. We re-populated the pERK levels uniformly within each gate to generate a continuous distribution of pERK expression pattern and applied Hartigans test to the generated distribution.

3.4.3 Plasmids, Stable & Transient Transfection

Plasmids used in this study were described before [5, 6, 64]. For transient co-transfection of human SOS1 with hCD16/7 fusion construct, $300\mu l$ of SOS^{1-2-} DT40 B cell suspension at 66 $\times$ 106 cells/ml in plain RPMI was electro-transfected with $10\mu g$ of each plasmid by using a Biorad electroporator (Biorad) set at square wave, 300V, 10ms single pulse. Transfected cells were cultured for 16-20 hr prior to anti-hCD16-biotin MACS purifications following manufacturers guideline (Miltenyi Biotec).

For stable reconstitution of human SOS1, SOS^{1-2-} DT40 B cells were transfected with $10\mu g$ hSOS1 plasmid by electroporation set at exponential decay, 250V, $950\mu F$. Transfected cells were selected by $10\mu g$/ml final blasticidin S (Fischer Biotech) from 24 hr post-transfection. Single cell-derived clones were isolated from a set of serially diluted culture on 96-well-plates after 1-2 weeks of seeding. Isolated clones were screened for hSOS1 and surface BCR (sBCR) expression by Western blot and FACS respectively.

3.4.4 Western blot analysis of cell lysates

Levels of various proteins were measured by Western blot as previously described [5, 6]. Following antibodies were purchased from Cell Signaling: Phospho-p44/42 MAPK ERK1/2 (pT204/pY204), total ERK1/2, Phospho-P38 (pT180/pY182, clone 3D7), total P38, Phospho-PLCγ1 (pY783) and Phospho-ZAP70 (pY319). Other antibodies are human SOS1 (BD Transduction Lab.), mRasGRP1 (Clone m199), Grb2 (SC-255, Santa Cruz Biotech), αTubulin (Sigma), Rac1 (Clone 23A8, Millipore) and Pan-Ras (Clone Ab-3, Calbiochem) for detection of chicken Ras. Proteins were visualized using ECL Western blot substrate (Pierce) and LAS-4000 image system (Fujifilm Life Science). Densitometric analysis was performed using Multi Guage V3.0 (Fujifilm Life Science). Shown results are representative of minimum two or more independent experiments.

3.4.5 Intracellular FACS staining for pERK

BCR-induced ERK phosphorylation was measured by intracellular FACS staining as previously described [6]. In brief, DT40 cells were resuspended/rested in PBS at 2.0 $\times$ 106 cells/75 μl (per well of 96 well plate). Cells were stimulated for desired time by adding 75 μl of 2$\times$ stimulation mix and subsequently fixed for 20 minutes at room temperature by mixing with 150 μl fixation buffer (Cytofix/Cytoperm, BD Biosciences). Cells were washed and permeabilized for at least 30 minutes with 90% methanol at -20^{o}C. Fixed/permeabilized cells were stained with anti-Phospho-p44/42

MAP Kinase (Thr202/Tyr204) antibody at RT in the presence of 2% (v/v) normal goat serum (Jackson Immunoresearch Laboratories). Subsequently, cells were washed twice and stained with PE- or APC-conjugated AffiniPure F(ab)2 fragment Donkey Anti-Rabbit IgG (Jackson Immunoresearch Laboratories). Stained cells were washed 3 times and directly analyzed by FACS.

3.4.6 Ras and Rac activation assays

BCR induced Ras and Rac activation was analyzed by a Ras/RacGTP pull-down assays according to the manufacturers instructions.

3.5 Supplementary Materials

3.5.1 Computational Models

For Fig. 3-2, we used an established computation model of Ras-activation via receptor input, as described in Fig. 1A of Riese et al. [8] with the extension of Ras-Raf-MEK-ERK activation cascade [9], as depicted in Fig. 3-9.

We applied the standard Gillespie algorithm [85] to simulate the described signaling network stochastically. In all stochastic simulations, we used a spatially homogeneous simulation box of size V = area ($4mm^2$) $\times$ height ($0.02mm$). This choice of the system size ensures that the system is well-mixed. The initial concentrations and the rate constants are those in Tables S1-2 and S3-4 of Das et al. [6], Tables 1-2 of Riese et al. [8], and supplementary materials of Locasale et al. [9], except that we used 36 molecules/(μm^2) for PIP_2 to match their published results [8]. More details of the simulation technique and choice of parameters can be found in the supplemental material of Das et al. [6] and Riese et al..[8].

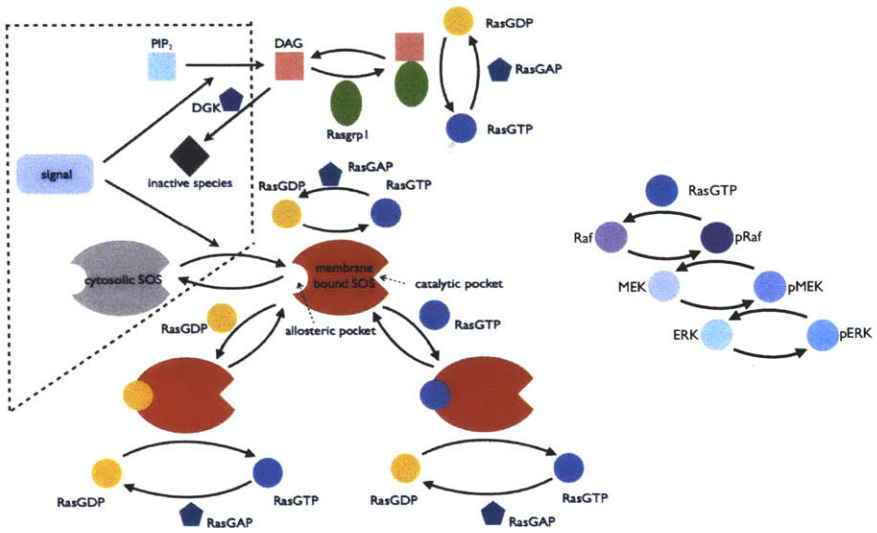


Figure 3-9: Coarse-grained model of ERK activation [6, 8, 9]. Ras can be activated via both DAG-Rasgrp1-mediated and SOS-mediated pathways. The catalytic rate of SOS depends on the state of its allosteric pocket: empty, bound to RasGDP, or bound to RasGTP, with increasing level of catalytic activities. In particular, the catalytic rate of SOS with RasGTP bound to its allosteric pocket is much larger than that with RasGDP bound to its allosteric pocket. This constitutes a SOS-mediated positive feedback for Ras activation. Ras can be deactivated by RasGAP, while DAG can be metabolized by DGK. Activated Ras can trigger the activation cascade of Ras-Raf-MEK-ERK.

Chapter 4

Rare Events Happen Suddenly

4.1 Introduction

Many important biological transitions occur rarely, either because the intrinsic dynamics of the process are slow relative to the time scales of observation, or because key intermediate steps, however fast they occur, occur only with low probabilities. The latter case, generally referred to as "rare events", includes transitions out of a stable basin (e.g., viral escape) and transitions between stable basins of reaction networks (e.g., potentially, in cancer). Abstractly, such transitions can be visualized by imagining a random walker moving from a stable basin to an unstable peak (en route to another stable basin). Such transitions are characterized by many unsuccessful attempts, followed, eventually, by a successful attempt. As a signature of their rarity, the expected time for the occurrence of such a successful attempt is exponential in the height of the barrier, and is inversely proportional to the probability of the transition.

A remarkable property of such transitions is that, although the system waits a long time for an arrival of a successful attempt, the successful attempt itself completes rapidly. This final excursion takes little time. In particular, we show that for Markov processes with detailed balance (such processes at steady state are known as at equilibrium), the successful transition occurs as quickly as the reverse. The successful excursion is not merely fast relative to the total waiting time for the rare transition, which is to be expected, but rather it is just as fast as a successful excursion in the

opposite (non-rare) direction.

One reason this is surprising is that the waiting time for the rare transition is quite long, much longer than the transition downhill, which occurs successfully rapidly, with few failed attempts. Perhaps it is easy to confuse the duration of the successful attempt with the total waiting time. However, the results are surprising for two more nuanced reasons. First, the transitions uphill are associated with slower intrinsic dynamics than the transition downhill, which is why they are improbable (i.e., rare). It seems natural to think that they will take longer to occur, even when successful. Second, a trajectory going uphill seems more likely to take a step backward en route to the final destination, since such back-tracking steps are more probable than the uphill ones. If these steps take the trajectory all the way back to the origin, the trajectory is unsuccessful and not relevant; however, successful trajectories can step back as long as they do not go so far as the origin, and these backward steps would consume time.

This interesting property of rare transition is well-known in a particular context: in the limit that stochastic fluctuations relative to deterministic drift approach zero, it has been shown by large-deviation theory that the rare transition undergoes the deterministic path from the opposite direction in reversed time [26]; as a consequence, the last excursion times are identical in both directions. We are not aware of a physical understanding of the generality of this phenomenon. Here we present different approaches to proving this property in Markov processes with detailed balance, to gain physical insight and to understand to what extent this property can hold in non-equilibrium systems, such as viral escape.

4.2 General proof for equilibrium Markov processes

First, we provide a proof for Markov processes at equilibrium in general by invoking a key property – time reversibility. Consider the trajectory taken from a Markov process at steady state (i.e., the current state of the system does not depend on the initial preparation of the system), as shown in Fig. 4-1. The object of computation is

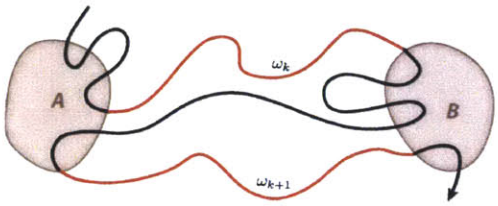


Figure 4-1: Schematic illustration of a long trajectory oscillating between two states A and B. The segments ω (indexed by k) from A to B without touching either state in between are as shown in red.

the conditional mean first-passage time (CMFPT) from state A to state B, denoted by $\mathbb{E}[\tau_{A\to B}|\mathcal{A}]$, that is, the expected value of the time from A to B (denoted by $\tau_{A\to B}$) conditioned on that once the system leaves state A, it does not return to state A before it reaches state B (the condition denoted by $\mathcal{A}$). To compute $\mathbb{E}[\tau_{A\to B}|\mathcal{A}]$, one natural way is to construct a "forward" trajectory ensemble by extracting all segments of the trajectory that start in A and reach B without touching either state in between (colored in red); if we label these segments by index k, and assume the segment ω_k leaves A at time t_k and arrives at B at $t_k + \tau_k$, then $\mathbb{E}[\tau_{A\to B}|\mathcal{A}]$ is simply $\sum_{k=1}^{k=N} \tau_k/N$, where N is the total number of such segments. We claim that the "forward" trajectory ensemble $\{\omega_k\}$ is sufficient to compute the CMFPT from B to A, $\mathbb{E}[\tau_{B\to A}|\mathcal{B}]$. For each "forward" trajectory ω_k, we can define its time-reversed trajectory $\bar{\omega}_k$ by $\bar{\omega}_k(t) = \omega_k(\tau_k - t)$. Note that such a defined $\bar{\omega}_k(t)$ is a valid "backward" trajectory, i.e., it starts from B and ends at A without touching either state in between. Because the system is at steady state and the dynamic obeys detail balance by assumption, the system is time-reversible, that is, the probability of observing ω_k is the same as that of observing $\bar{\omega}_k$. Therefore, the "forward" trajectory ensemble $\{\omega_k\}$, constructed as such, constitutes the "backward" trajectory ensemble. Furthermore, the "backward"

CMFPT $\mathbb{E}[\tau_{B\to A}|\mathcal{B}]$ is $\sum_{k=1}^{k=N}\tau_k/N$ as well.

Next, we substantiate this general proof into two specific dynamics, namely, continuous-time Markov Chain and one-dimensional Langevin dynamics, both of which are restricted to preserve detailed balance.

4.3 Proof for continuous-time Markov Chain with detailed balance

Consider a particular path ω connecting state A to state B, where the transition from state i to $i+1$ takes time τ_i and $s_1 = A$ and $s_M = B$. Given that the waiting time τ_i at state i follows an exponential distribution characterized by the propensity of leaving state i, r_i, we can express the probability density of this path ω conditioned on that it starts from A as follows:

$$\mathbb{P}(\omega|s_1 = A) = p_{s_1\to s_2}p_{s_2\to s_3}\exp(-r_2\tau_2)\times\ldots\times p_{s_{M-1}\to s_M}\exp(-r_{M-1}\tau_{M-1}) \quad (4.1)$$

where $p_{s_i\to s_{i+1}}$ is the transition probability from s_i to s_{i+1}. The clock starts when the trajectory exits state A and stops as soon as it enters state B; hence neither the time spent at state A or that at state B in included.

Next, consider a path $\bar{\omega}$, the time-reversed path of ω, which passes through the same set of states as path ω, but in the reverse order, and with the same waiting times at each each state as path ω. The probability density of the path $\bar{\omega}$ conditioned on that it starts from state B is

$$\mathbb{P}(\bar{\omega}|s_1 = B) = p_{s_M\to s_{M-1}}p_{s_{M-1}\to s_{M-2}}\exp(-r_{M-1}\tau_{M-1})\times\ldots\times p_{s_2\to s_1}\exp(-r_2\tau_2) \quad (4.2)$$

Since we assume the dynamic obeys detailed balance, it is possible to assign an energy E_i to each of the states such that:

$$\frac{p_{i\to j}}{p_{j\to i}} = \exp[-\beta(E_j - E_i)] \quad (4.3)$$

where β is an energy scale (i.e., the temperature in a physical system).

Combining Eqs (4.1), (4.2), and (4.3), we can obtain the ratio of the conditional probability densities of the forward and backward paths:

$$\frac{\mathbb{P}(\omega|s_1 = A)}{\mathbb{P}(\bar{\omega}|s_1 = B)} = \exp\left[-\beta(E_B - E_A)\right] \tag{4.4}$$

The probabilities of the transition times cancel because the transition propensity r_i is a function of state i only, and the paths ω and $\bar{\omega}$ spend the same time at state i. The energies of intermediate states also cancel as in a telescoping series. The resulting ratio of the conditional probability densities of the forward and backward paths is only a function of the energy difference between state A and state B.

This property in Eq (4.4) can be used to prove that the CMFPT from A to B, $\mathbb{E}[\tau_{A\to B}|\mathcal{A}]$, is identical to that from B to A, $\mathbb{E}[\tau_{B\to A}|\mathcal{B}]$. Applying Baye's rule, we have the following:

$$\mathbb{E}[\tau_{A\to B}|\mathcal{A}] = \frac{\int_{\{\omega_{A\to B}\}} d\omega \tau(\omega)\mathbb{P}(\omega)}{\int_{\{\omega_{A\to B}\}} d\omega \mathbb{P}(\omega)} \tag{4.5}$$

where $\tau(\omega)$ is the time taken by the path ω (i.e., $\tau(\omega) = \sum_{i=2}^{M-1} \tau_i$) and the integrals include al paths that connect A to B. Since $\tau(\omega) = \tau(\bar{\omega})$ and at steady state there is one-to-one mapping between ω and $\bar{\omega}$, we have from Eqs (4.4) and (4.5) that

$$\mathbb{E}[\tau_{A\to B}|\mathcal{A}] = \frac{\int_{\{\bar{\omega}_{B\to A}\}} d\bar{\omega} \tau(\bar{\omega})\mathbb{P}(\bar{\omega})}{\int_{\{\bar{\omega}_{B\to A}\}} d\bar{\omega} \mathbb{P}(\bar{\omega})} = \mathbb{E}[\tau_{B\to A}|\mathcal{B}] \tag{4.6}$$

Note that the energy term in Eq (4.4), which is independent of paths, would appear in both the numerator and denominator of Eq (4.6), and therefore cancels.

4.4 Proof for Langevin Dynamics with detailed balance

Now consider a system governed by one-dimensional Langevin dynamics expressed in the form of stochastic differential equations (SDEs):

$$dX = r(X)dt + s(X)dW \tag{4.7}$$

where X is the state variable describing the position of the system, $r(X)$ is the state-dependent drift term, $s(X)$ is the state-dependent diffusivity, and W is the standard Brownian Motion. Note the Langevin Dynamics described by Eq (4.7) is equivalent to the following form:

$$\frac{dX}{dt} = r(X) + s(X)\xi(t) \tag{4.8}$$

where $\xi(t)$ is a delta-correlated stationary Gaussian process with zero-mean, satisfying $\langle \xi(t) \rangle = 0$ and $\langle \xi(t)\xi(t') \rangle = \delta(t - t')$.

Without loss of generality, we denote state A as $X = a$ and state B as $X = b$, where $a < b$. We define the "forward" CMFPT from A to B as $\mathbb{E}(\tau(\omega_{A\to B})|\mathcal{A})$, and the "backward" CMFPT from B to A as $\mathbb{E}(\tau(\omega_{B\to A})|\mathcal{B})$, where $\tau(\omega_{i\to j})$ is the first-passage time of trajectory ω from state i to state j.

The goal here is to show that regardless of the functional forms of $r(X)$ and $s(X)$,

$$\mathbb{E}(\tau(\omega_{A\to B})|\mathcal{A}) = \mathbb{E}(\tau(\omega_{B\to A})|\mathcal{B}) \tag{4.9}$$

In this section, we specialize to the case where the diffusivity is constant:

$$dX = r(X)dt + sdW \tag{4.10}$$

The generalization to the state-dependent diffusivity can be found in Chapter 4.7.1. Furthermore, the above results can be extended to multi-dimensional Langevin dynamics with detailed balance, which is presented in Chapter 4.7.2.

The general machinery used here is change of measure, more specifically, the conditional Girsanov's theorem, adapted from measure theory. An equivalent formulation can be developed by path integrals. The essence of change of measure is that by reassigning different weights (i.e., probabilities) to different trajectories, Brownian motion with drift (such as that in Eq (4.10)) will behave just like Brownian motion without drift.

Denote the probability measure under which Brownian motion in Eq (4.10) is defined measure P. Note that the backward CMFPT, $\mathbb{E}(\tau(\omega_{B\to A})|\mathcal{B})$ governed by Eq (4.10) is the same as the forward CMFPT governed by the following dynamics:

$$dX = -r(c - X)dt + sdW \tag{4.11}$$

where $c = a + b$ to ensure symmetry. Now to prove Eq (4.10), we have to prove

$$\mathbb{E}^{(10)}(\tau(\omega_{A\to B})|\mathcal{A}) = \mathbb{E}^{(11)}(\tau(\omega_{A\to B})|\mathcal{A}) \tag{4.12}$$

where the superscript is to denote under which equation the system dynamic is defined.

Tilt the probability measure of Eq (4.10) such that $dX = sdW^Q$ under measure Q. By the Girsanove theorem, such a probability measure Q can be obtained by defining the Radon-Nikodym derivative expressed as follows:

$$\xi_\tau = \left(\frac{dP}{dQ}\right)_{\mathbb{F}_\tau} = \exp\{\int_0^\tau \frac{r(X)}{s}dW^Q - \frac{1}{2}\int_0^\tau \frac{r^2(X)}{s^2}dt\} = \exp\{\int_0^\tau \frac{r(X)}{s^2}dX - \frac{1}{2}\int_0^\tau \frac{r^2(X)}{s^2}dt\} \tag{4.13}$$

where $\mathcal{F}_\tau$ is the associated Brownian filtration indexed by the stopping time τ (i.e., the first-passage time of the trajectory ω).

We construct a potential function $U(x)$ such that $r(x) = -\dfrac{dU(x)}{dx}$ (an important note: for a one-dimensional system, such a construction of potential functions always exists; this is no longer true for high-dimensional Langevin Dynamics, where additional constraints on the functional form of the drift vector are needed for such a

potential landscape construction), then by Ito's calculus,

$$\int_0^\tau \frac{r(X)}{s^2} dX(t) = \frac{U(a) - U(b)}{2s^2} - \int_0^\tau \frac{1}{2} \frac{\partial r(X)}{\partial X} dt \tag{4.14}$$

In Eq (4.14), the emergence of the second term on the RHS is due to the fact that the integral on the LHS is a stochastic Ito's integral, instead of a Riemann integral. With this, we can simplify Eq (4.13) and obtain the expression for ξ_τ:

$$\xi_\tau = \left(\frac{dP}{dQ}\right)_{\mathbb{F}_\tau} = \exp\{\frac{U(a) - U(b)}{s^2}\} \exp\{-\frac{1}{2} \int_0^\tau (\frac{r^2(X)}{s^2} + \frac{\partial r(X)}{\partial X}) dt\} \tag{4.15}$$

Hence the LHS of Eq (4.12) can be evaluated under measure Q:

$$\mathbb{E}^{(10)}(\tau(\omega_{A\to B})|\mathcal{A}) = \frac{\mathbb{E}^Q(\xi_\tau \tau|\mathcal{A})}{\mathbb{E}^Q(\xi_\tau|\mathcal{A})} = \frac{\mathbb{E}^Q\Big(\tau \exp\{-\frac{1}{2}\int_0^\tau (\frac{r^2(X)}{s^2} + \frac{\partial r(X)}{\partial X}) dt\}|\mathcal{A}\Big)}{\mathbb{E}^Q\Big(\exp\{-\frac{1}{2}\int_0^\tau (\frac{r^2(X)}{s^2} + \frac{\partial r(X)}{\partial X}) dt\}|\mathcal{A}\Big)} \tag{4.16}$$

Note that the denominator term in 4.16 exists to ensure the proper normalization under the conditional probability. The potential energy term cancels in the same way as in Eq (4.6).

Likewise, one can tilt the probability measure of Eq (4.11) such that under another measure Q', $dX = sdW^{Q'}$. In this case, the Radon-Nikodym derivative of measure P relative to Q' is

$$\begin{aligned}
\xi'_\tau &= \left(\frac{dP}{dQ'}\right)_{\mathbb{F}_\tau} = \exp\{\int_0^\tau \frac{-r(c-X)}{s} dW^{Q'} - \frac{1}{2}\int_0^\tau \frac{r^2(c-X)}{s^2} dt\} \\
&= \exp\{\int_0^\tau \frac{-r(c-X)}{s^2} dX - \frac{1}{2}\int_0^\tau \frac{r^2(c-X)}{s^2} dt\} \\
&= \exp\{\frac{U(b) - U(a)}{s^2}\} \exp\{-\frac{1}{2}\int_0^\tau (\frac{r^2(c-X)}{s^2} + \frac{\partial r(c-X)}{\partial X}) dt\}
\end{aligned}$$

As a result, the RHS of Eq (4.12) equals

$$\mathbb{E}^{(11)}(\tau(\omega_{A\to B})|\mathcal{A}) = \frac{\mathbb{E}^{Q'}(\xi'_\tau \tau|\mathcal{A})}{\mathbb{E}^{Q'}(\xi'_\tau|\mathcal{A})} = \frac{\mathbb{E}^{Q'}\left(\tau \exp\{-\frac{1}{2}\int_0^\tau (\frac{r^2(c-X)}{s^2} + \frac{\partial r(c-X)}{\partial X})dt\}|\mathcal{A}\right)}{\mathbb{E}^{Q'}\left(\exp\{-\frac{1}{2}\int_0^\tau (\frac{r^2(c-X)}{s^2} + \frac{\partial r(c-X)}{\partial X})dt\}|\mathcal{A}\right)} \tag{4.17}$$

The remaining proof is done by matching trajectories. Let $\phi(t)$ and $\phi^*(t)$ denote a dynamical path performed the Brownian walker defined by Eq (4.10) and by Eq (4.11), respectively. Note that there is a one-to-one correspondence between $\phi(t)$ with duration τ and $\phi^*(t)$ with the same duration τ, and the mapping is simply $\phi^*(t) = c - \phi(\tau - t)$. Furthermore, under the changed measure Q and Q', respectively, both dynamics originally defined by Eq (4.10) and by Eq (4.11) become unbiased Brownian motion, which means $\phi(t)$ and $\phi^*(t) = c - \phi(\tau - t)$ occur with equal probabilities under Q and Q', respectively. With that, the numerator and denominator of Eq 4.16 are equal to those of Eq 4.17, respectively. That concludes the proof.

On a separate note, it is important to include the second term on the right-hand side of Eq 4.14 [86] . The emerge of this fluctuation correction term is due to Ito's interpretation. Without this term, the above argument would have been valid when the sign of the drift is reversed (i.e., from $r(X)$ to $-r(X)$), and one would have wrongly concluded that the mean first-passage time from an arbitrary starting point X_0 ($a \leq X_0 < b$) to $x = b$ conditioned on that it does not reach first $X = a$ is the same as that with an opposite drift. However, in the presence of this term, such an argument would work only when $\partial r(X)/\partial X = 0$ (i.e., the drift is constant).

4.5 Numerical verification

Our results for the case of one-dimensional Langevin dynamics can be verified numerically. To compute the expected time that a particle governed by Eq 4.7, starting from X_0 reaches $X = b$ before it reaches $X = a$ (denoted by $\tau(X_0)$), we first calculate the splitting probability that this particle reaches $X = a$ before $X = b$ (denoted by

$\mathbb{P}(X_0)$). $\mathbb{P}(X_0)$ is governed by the following ODE:

$$r\frac{d\mathbb{P}}{dX} + \frac{s^2}{2}\frac{d^2\mathbb{P}}{dX^2} = 0 \tag{4.18}$$

subject to the boundary conditions that $\mathbb{P} = 0$ evaluated at $X_0 = a$ and $\mathbb{P} = 1$ evaluated at $X_0 = b$ [87]. With that, we can obtain $\tau(X_0)$ by solving the following ODE:

$$r\frac{d(\tau\mathbb{P})}{dX} + \frac{s^2}{2}\frac{d^2(\tau\mathbb{P})}{dX^2} + \mathbb{P} = 0 \tag{4.19}$$

subject to the boundary conditions that $\tau\mathbb{P} = 0$ evaluated at $X_0 = a$ or $X_0 = b$ [87]. To compute the CMFPT from a to b (or from b to a), we let X_0 be infinitesimally close to a (or b).

We computed the forward and backward CMFPT for two dynamics, namely $dX = mdt + dW$ and $dX = mXdt + dW$ with $a = 0$, $b = 1$, and varying values of m. From Fig. 4-2A&B, it is clear that the forward and backward CMFPT are identical. Furthermore, we let $X_0 = 0.5$ and computed the CMPFT starting from X_0 under positive and negative drifts, respectively. From Fig. 4-2C&D, reversing the sign of drift does not change the CMFPT when the drift and diffusivity are constant, but it does when the drift is state-dependent. These are consistent with our theoretical derivations.

4.6 Discussion and Extension to non-equilibrium systems

The proceeding results demonstrate that the conditional time to go up an energy hill is equal to the time to go down in equilibrium systems. This precise identity will not extend in general to Markov systems without detailed balance (i.e., whose steady states are nonequilibrium). A toy example is to consider a discrete-time Markov Chain with transition probabilities $p_{A\to B} = p_{B\to 3} = p_{3\to 1} = 1$. Due to the presence of cycles, CMFPT from state A to state B is 1, while that from B to A is 2. However,

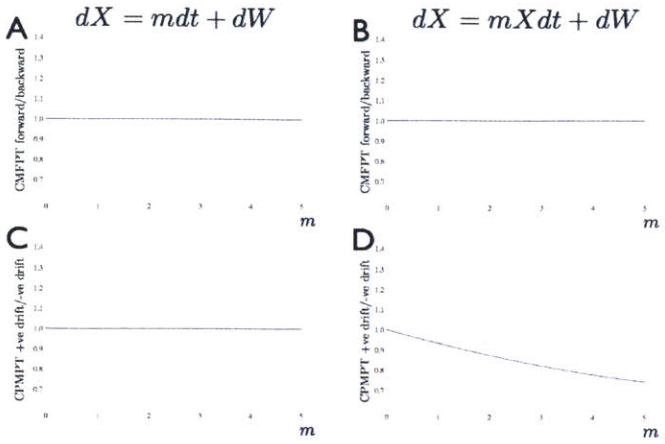


Figure 4-2: Numerical verifications. (A, B) The ratio between forward and backward CMFPTs for a particle on the interval [0,1] following (A) $dX = mdt + dW$ or (B) $dX = mXdt + dW$ with varying m. (C, D) The ratio between CMPFTs for a particle following (C) $dX = mdt + dW$ or (D) $dX = mXdt + dW$, starting from $X_0 = 0.5$ reaches $X = 1$ before $X = 0$ with positive and negative drifts.

a more general interpretation of the result, that is, rare events occur suddenly, does extend in general. Consider a particular sequence of states s_i connecting an energy well to a peak. One's instinct might be that the expected time to transition following this sequence of states is $\sum_i 1/\lambda_{s_i \to s_{i+1}}$, where $\lambda_{s_i \to s_{i+1}}$ is the transition rate from state i to its next state in sequence. However, a surprising, but well-known fact from Poisson processes states that the expected time of traversing these sequence of states is really $\sum_i (1/\sum_j \lambda_{s_i \to s_j})$, where the second summation is over all the states j accessible from the current state i. Thus, the less probable the required transition from s_i to s_{i+1} (i.e., the smaller $\lambda_{s_i \to s_{i+1}}$ is relative to $\sum_j \lambda_{s_i \to s_j}$), the shorter the conditional expected time of the transition is to what might have been assumed. Alternatively, making a path less probable by adding fast competing transitions to states along the transition path actually decreases the expected conditional time of traversing the path. This understanding may have significant implications to non-equilibrium systems: for example, to quench viral escape, one may consider mounting

immune pressure or administering drugs to suppress the occurrence of rare events; however, with these measures implemented, if the virus ever successfully escapes, the last excursion is even faster than before.

4.7 Supplemental Materials

4.7.1 Generalized 1D Langevin Dynamics

Now we prove the most general case for 1D Langevin Dynamics, where both the drift and the diffusivity are state-dependent, as described by Eq (4.7).

Define $Y(X) = \int_d^X \frac{1}{s(u)} du$, where d is a constant and assume such a Y is well-defined. Also assume that $s(u) > 0 \; \forall u \in [d, X]$ such that there is a one-to-one correspondence between Y and X. Apply Ito's lemma, one has

$$dY(X) = \frac{\partial Y}{\partial X} dX + \frac{1}{2} \frac{\partial^2 Y}{\partial X^2} (dX)^2 + \frac{\partial Y}{\partial t} dt \tag{4.20}$$

The last term vanishes since Y is not an explicit function of t. By the construction of $Y(X)$,

$$dY(X) = \frac{1}{s(X)} dX + \frac{1}{2} \left(- \frac{s'(X)}{s^2(X)} \right) s^2(X) dt = \left(\frac{r(X)}{s(X)} - \frac{1}{2} s'(X) \right) dt + dW \tag{4.21}$$

To prove Eq (4.9), we only have to show that the CMFPT from $Y(X = a)$ to $Y(X = b)$ is the same as the CMFPT from $Y(X = b)$ to $Y(X = a)$. And this has been proven in Chapter 4.4 since the one-dimensional Langevin Dynamics describing Y has a constant diffusivity.

4.7.2 Multi-dimensional Langevin Dynamics with detailed balance

In this last section, we consider the N-dimensional Langevin Dynamics, where the drift vector can be expressed as the gradient of a well-defined potential energy landscape and the diffusivity tensor is diagonal, isotropic and state-independent (note

that these are required for the Langevin Dynamics to fulfill detailed balance):

$$d\mathbf{X} = \mathbf{r}(\mathbf{X})dt + \mathbf{s}d\mathbf{W} \tag{4.22}$$

where $\mathbf{X}$ is an N-dimensional state variable, $\mathbf{W}$ is a standard N-dimensional Brownian motion, $\mathbf{s}$ is an isotropic and state-independent $N \times N$ diagonal diffusivity tensor, and $\mathbf{r}(\mathbf{X})$ is an N-dimensional drift vector that satisfies

$$\mathbf{r}(\mathbf{X}) = -\nabla U(\mathbf{X}) \tag{4.23}$$

where $U(\mathbf{X})$ is a well-defined N-dimensional potential energy landscape.

We are interested in proving that the CMFPT from $\mathbf{X} = \mathbf{a}$ to $\mathbf{X} = \mathbf{b}$ is the same as the CMFPT from $\mathbf{X} = \mathbf{b}$ to $\mathbf{X} = \mathbf{a}$. One can apply the multidimensional Girsanov theorem, which states that:
Let $\Theta(t) = \{\theta_1(t), \theta_2(t), ..., \theta_N(t)\}$ be an N-dimensional process that is $\mathbb{F}_t$-measurable and satisfies the following condition under measure P: $\mathbb{E}\exp(\frac{1}{2}\int_0^t \sum_{i=1}^N \theta_i^2(u)du) < \infty$. Define a random process $Z(t)$:

$$Z(t) = \exp\left\{\sum_{i=1}^{n}\left(-\int_0^t \theta_i(u)dW_i(u) - \frac{1}{2}\int_0^t \theta_i^2(u)du\right)\right\}$$

where $\mathbf{W}(t) = \{W_1(t), W_2(t), ..., W_N(t)\}$ is a standard N-dimensional Brownian motion under measure P. Then under measure Q, the following process is a standard N-dimesional Brownian motion:

$$\widetilde{\mathbf{W}}(t) = \mathbf{W}(t) + \int_0^t \Theta(u)du$$

and the probability measure Q is defined as

$$\left(\frac{dQ}{dP}\right)_{\mathbb{F}_t} = Z(t)$$

The proof in this section is analogous to that in Chapter 4.4. One can define the

Radon-Nikodym derivative of measure P relative to measure Q as

$$\xi_T = \left(\frac{dP}{dQ}\right)_{\mathbb{F}_T} = \exp\left\{\sum_{i=1}^{n}\left(\int_0^T \theta_i(u)dW_i^Q(u) - \frac{1}{2}\int_0^T \theta_i^2(u)du\right)\right\}$$

where $\theta_i = r_i/s$ and s is the diagonal entry of the diffusivity tensor. One can recognize that

$$\begin{aligned}
\xi_\tau &= \exp\left\{\sum_{i=1}^{n}\left(\int_0^\tau \theta_i(u)dW_i^Q(u) - \frac{1}{2}\int_0^\tau \theta_i^2(u)du\right)\right\} \\
&= \exp\left\{\sum_{i=1}^{n}\left(\int_0^\tau \frac{r_i(u)}{s^2}dX_i(u) - \frac{1}{2}\int_0^\tau \theta_i^2(u)du\right)\right\} \\
&= \exp\left(\frac{U(\mathbf{X}=\mathbf{A}) - U(\mathbf{X}=\mathbf{B})}{s^2}\right)\exp\left\{\sum_{i=1}^{n}\left(-\frac{1}{2}\int_0^\tau \theta_i^2(u)du\right) + \frac{1}{2}\int_0^\tau dX_t^T \cdot \nabla_{X_t}^2 U \cdot dX_t\right\}
\end{aligned}$$

And the rest of the proof is done by matching trajectories just as in Chapter 4.4.

Chapter 5

Applying Linear-Noise Approximation to Network Topology Identification

5.1 Introduction

Our understanding of biological systems is limited by the time scale and length scale of experimental techniques available. The advancement of real-time molecular-level experimental measurement and the emergence of ample experimental data with fine time solution provide promises to probe and identify molecular networks, letting them be gene regulatory networks, protein-protein interaction networks, or signaling transduction networks. At microscopic levels, fluctuations are ubiquitous and significant due to the low copy numbers of molecules of interest and the stochastic nature of their interactions [11, 88]. The presence of such noise, on one hand, brings statistical uncertainty to experimental data observed, and on the other hand, offers enormous information through the nonrandom underlying structure of the noise, which would be otherwise ignored at macroscopic levels.

A related but different question is how to meaningfully represent complex biological networks in lower dimensions by grouping relevant nodes according to their

connectivity or better according to their similarities in functionalities. The concept of network motifs, connectivity-patterns or sub-graphs that occur much more often than they do in random networks, as advocated by Uri Alon et al. [89], emerges as an attempt to address this question. These basic building blocks, each carrying certain topology, functions, and design principles, are found in biological networks in organisms with different levels of complexities and as well as in social networks. Besides network motifs, various clustering techniques primarily based on machine learning have been recently proposed to intellectually coarse-grain biological network representations.

Figure 5-1: Two three-node network motifs: (A) incoherent feedforward loop (abbreviated as IFF) and (B) negative feedback loop (abbreviated as NFB). (A) Node A activates Node C directly and deactivates Node C mediated via Node B. (B) Node A activates Node C and Node C forms a negative feedback loop mediated via Node B. In both (A) and (B), Node B is deactivated by self-decay.

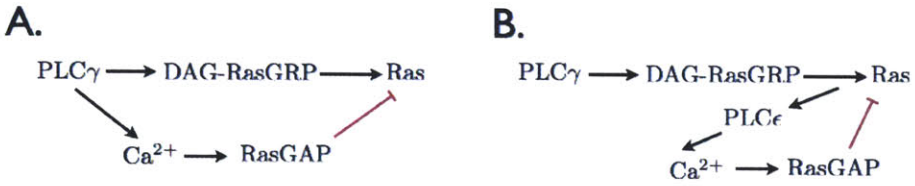


Figure 5-2: Biological examples of IFF and NFB. (A) PLCγ regulates Ras activation by IFF: PLCγ activates Ras via DAG-RasGRP-mediated pathway and deactivates Ras via Ca^{2+}-RasGAP-mediated pathway. (B) PLCγ regulates Ras activation by NFB: PLCγ activates Ras, which in turn deactivate itself via PLCϵ-Ca^{2+}-RasGAP-mediated negative feedback loop.

In this work, we focus on two commonly-seen three-node motifs, namely, incoherent feedforward loop (abbreviated as IFF) and negative feedback loop (abbreviated

as NFB), as depicted in Fig. 5-1, which are the only two three-node topologies that perform adaptation (i.e., the ability to reset to a new steady state after responding to a stimulus). We ask the question whether one can differentiate these two from molecular-level time series data of copy numbers. Such an understanding is important to dissect complex reaction networks and pinpoint dominant pathways. For instance, as depicted in Fig. 5-2, the activation of PLCγ can activate Ras via DAG-RasGRP pathway and at the same time deactivate Ras via Ca^{2+}-RasGAP pathway, which constitutes an IFF (Fig. 5-2A); an alternative is that Ras activated via DAG-RasGRP pathway activates PLCϵ and subsequently triggers Ca^{2+}-RasGAP pathway to deactivate Ras, in which case PLCγ regulates Ras activation through an NFB (Fig. 5-2B). The ability to discriminate IFF and NFB will shed light on how exactly Ras is regulated by Ca^{2+}-RasGAP pathway [90].

Here, we apply the linear noise approximation (abbreviated as LNA, also named van Kampen's system-size expansion) to explore features in correlation functions of molecular copy numbers and identify signatures, under certain parameter regions, that can differentiate IFF from NFB. We hope this endeavor can be the first step to utilize time-series data with fine time and length resolutions to unveil signaling network topologies.

5.2 LNA applied to IFF

5.2.1 Problem formulation and preliminary analysis

Here we assume the simplest reaction kinetic model for the IFF topology, as depicted in Fig. 5-1A. Consider the toy model defined as follows:

$$C_{inactive} \xrightarrow{k_1[A]} C_{active}$$

$$B_{inactive} \xrightarrow{k_2[A]} B_{active}$$

$$C_{active} \xrightarrow{k_3[B]_{active}} C_{inactive}$$

$$B_{active} \xrightarrow{k_4} B_{inactive}$$

Fix the total concentrations of species B and C, respectively (i.e., $[B]_T = [B]_{active} + [B]_{inactive}$ and $[C]_T = [C]_{active} + [C]_{inactive}$). Assume the amount of species A does not fluctuate. For the simplicity of notation, we use $[B]$ and $[C]$ to represent $[B]_{active}$ and $[C]_{active}$, respectively, and keep track of $[B]$ and $[C]$ throughout the analysis. The deterministic equations describing the time-evolution of the species concentrations are

$$\frac{d[B]}{dt} = k_2[A]([B]_T - [B]) - k_4[B] \tag{5.1}$$

$$\frac{d[C]}{dt} = k_1[A]([C]_T - [C]) - k_3[B][C] \tag{5.2}$$

Define the stoichiometric matrix υ, where its entry $\upsilon_{i,j}$ is the stoichiometric coefficient of species j (for species B, $j = 1$; for species C, $j = 2$) in the reaction i. The stoichiometric matrix υ for this model is

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \\ 0 & -1 \\ -1 & 0 \end{bmatrix} \tag{5.3}$$

It is easy to see that there is only one steady-state solution $[B]^s, [C]^s$ to the system defined by Eqs (5.1) and (5.2). Define the Jacobian matrix A, where its entry $A_{i,j} = \sum_{k=1}^{R} \upsilon_{ki} \left(\frac{\partial f_k}{\partial \phi_j}\right)$, where ϕ_j represents the concentration of species j, and f_k the propensity of reaction k. The Jacobian matrix A for the toy model is

$$\begin{bmatrix} -k_2[A] - k_4 & 0 \\ -k_3[C] & -k_1[A] - k_3[B] \end{bmatrix} \tag{5.4}$$

The two eigenvalues for this Jacobian matrix A are $-k_2[A] - k_4$ and $-k_1[A] - k_3[B]$, both of which are negative at steady state, suggesting that the steady state of the system is stable.

5.2.2 System-size expansion applied to IFF

Now we apply the system-size expansion to the system at steady state. The goal is to obtain correlation functions governing molecular fluctuations of species B and C.

Assume the system has a constant volume Ω. Expressing Eqs (5.1) and (5.2) in terms of numbers of molecules, one has

$$\frac{dn_B}{dt} = \Omega^{-1}k_2 n_A[(n_B)_T - n_B] - k_4 n_B \tag{5.5}$$

$$\frac{dn_C}{dt} = \Omega^{-1}k_1 n_A[(n_C)_T - n_C] - \Omega^{-1}k_3 n_B n_C \tag{5.6}$$

where $n_j = \Omega c_j$ (c_j is the concentration of species j). Hence, the Master Equation describing the evolution of probability distribution function $P(n_B, n_C, t)$ is

$$\begin{aligned}\frac{dP(n_B,n_C,t)}{dt} &= \Omega^{-1}k_1(\mathbf{E}_C^{-1} - 1)n_A[(n_C)_T - n_C]P + \Omega^{-1}k_3(\mathbf{E}_C^{1} - 1)n_B n_C P \\ &\quad +\Omega^{-1}k_2(\mathbf{E}_B^{-1} - 1)n_A[(n_B)_T - n_B]P + k_4(\mathbf{E}_B^{1} - 1)n_B P\end{aligned} \tag{5.7}$$

where $\mathbf{E}$ is the step operator, which gives $\mathbf{E}_i^1 f(n_i) = f(n_i + 1)$ and $\mathbf{E}_i^{-1} f(n_i) = f(n_i - 1)$.

One would expect the number of molecules n_j has the mean scaled with Ω and the standard deviation scaled with $\Omega^{1/2}$. Hence, we set $n_B = \Omega\phi(t) + \Omega^{1/2}\zeta(t)$ and $n_C = \Omega\psi(t) + \Omega^{1/2}\eta(t)$. The dynamics of $(\phi(t), \psi(t))$ is expected to be governed by the macroscopic equations (5.1) and (5.2).

We write the probability distribution function $P(n_B, n_C, t)$ as a function of (ζ, η, t): $P(n_B, n_C, t) = \Pi(\zeta, \eta, t)$. Since $\mathbf{E}_B^1 f(n_B) = f(n_B + 1)$ and therefore $\mathbf{E}_B^1 g(\zeta) = g(\zeta + \Omega^{-1/2})$, one can approximate the step operator $\mathbf{E}_B^1$ by the Taylor Expansion

$$\mathbf{E}_B^1 = 1 + \Omega^{-1/2}\frac{\partial}{\partial\zeta} + 1/2\Omega^{-1}\frac{\partial^2}{\partial\zeta^2} + ... \tag{5.8}$$

Similarly, other step operators can be approximated as partial differential operators

$$\mathbf{E}_B^{-1} = 1 - \Omega^{-1/2}\frac{\partial}{\partial\zeta} + 1/2\Omega^{-1}\frac{\partial^2}{\partial\zeta^2} + ... \tag{5.9}$$

$$\mathbf{E}_C^1 = 1 + \Omega^{-1/2}\frac{\partial}{\partial \eta} + 1/2\Omega^{-1}\frac{\partial^2}{\partial \eta^2} + ... \tag{5.10}$$

$$\mathbf{E}_C^{-1} = 1 - \Omega^{-1/2}\frac{\partial}{\partial \eta} + 1/2\Omega^{-1}\frac{\partial^2}{\partial \eta^2} + ... \tag{5.11}$$

The time derivative in the Master Equation (5.7) is taken with constant n_B and n_C. This means that $d\zeta/dt = -\Omega^{1/2}d\phi/dt$ and $d\eta/dt = -\Omega^{1/2}d\psi/dt$. By the definition of total derivative, the time derivate can be written as

$$\frac{dP(n_B, n_C, t)}{dt} = \frac{\partial \Pi}{\partial t} - \Omega^{1/2}\frac{d\phi}{dt}\frac{\partial \Pi}{\partial \zeta} - \Omega^{1/2}\frac{d\psi}{dt}\frac{\partial \Pi}{\partial \eta} \tag{5.12}$$

With Eqs (5.8) to (5.12), assuming the number of molecule A does not fluctuate, the Master Equation (5.7) can be expanded as follows

$$\begin{aligned}
&\frac{\partial \Pi}{\partial t} - \Omega^{1/2}\frac{d\phi}{dt}\frac{\partial \Pi}{\partial \zeta} - \Omega^{1/2}\frac{d\psi}{dt}\frac{\partial \Pi}{\partial \eta} \\
= \;& k_1[A](-\Omega^{-1/2}\frac{\partial}{\partial \eta} + 1/2\Omega^{-1}\frac{\partial^2}{\partial \eta^2})[\Omega\psi_T - \Omega\psi(t) - \Omega^{1/2}\eta]\Pi \\
+ \;& \Omega^{-1}k_3(\Omega^{-1/2}\frac{\partial}{\partial \eta} + 1/2\Omega^{-1}\frac{\partial^2}{\partial \eta^2})(\Omega\phi(t) + \Omega^{1/2}\zeta)(\Omega\psi(t) + \Omega^{1/2}\eta)\Pi \\
+ \;& k_2[A](-\Omega^{-1/2}\frac{\partial}{\partial \zeta} + 1/2\Omega^{-1}\frac{\partial^2}{\partial \zeta^2})(\Omega\phi_T - \Omega\phi(t) - \Omega^{1/2}\zeta)\Pi \\
+ \;& k_4(\Omega^{-1/2}\frac{\partial}{\partial \zeta} + 1/2\Omega^{-1}\frac{\partial^2}{\partial \zeta^2})(\Omega\phi(t) + \Omega^{1/2}\zeta)\Pi
\end{aligned} \tag{5.13}$$

Now collecting the $O(\Omega^{1/2})$ terms, we obtain the following two equations describing the dynamics of $\phi(t)$ and $\psi(t)$.

$$\frac{d\phi}{dt} = k_2[A](\phi_T - \phi(t)) - k_4\phi(t) \tag{5.14}$$

$$\frac{d\psi}{dt} = k_1[A](\psi_T - \psi(t)) - k_3\phi(t)\psi(t) \tag{5.15}$$

With no surprise, Eqs (5.14) and (5.15) are the same as Eqs (5.1) and (5.2), the deterministic description of the reaction network. Given the initial condition, they can be solved analytically. Let ϕ^s and ψ^s denote the steady-state solution.

Collecting the $O(\Omega^0)$ terms, we obtain the following equation describing the time-

evolution of probability $\Pi(\zeta, \eta, t)$:

$$\begin{aligned}\frac{\partial \Pi}{\partial t} &= (k_2[A] + k_4)\frac{\partial}{\partial \zeta}(\zeta\Pi) + (k_1[A] + k_3\phi(t))\frac{\partial}{\partial \eta}(\eta\Pi) + k_3\psi(t)\frac{\partial}{\partial \eta}(\zeta\Pi) \\ &+ \tfrac{1}{2}[k_1[A](\psi_T - \psi(t)) + k_3\phi(t)\psi(t)]\frac{\partial^2}{\partial \eta^2}\Pi + \tfrac{1}{2}[k_2[A](\phi_T - \phi(t)) + k_4\phi(t)]\frac{\partial^2}{\partial \zeta^2}\Pi\end{aligned} \tag{5.16}$$

Recognizing that Eq (5.14) is a multivariate linear Fokker-Planck equation, which has been solved in [11], we can readily write down the following equations describing the decay of the first moments of the noise

$$\partial_t\langle\zeta\rangle = -(k_2[A] + k_4)\langle\zeta\rangle \tag{5.17}$$

$$\partial_t\langle\eta\rangle = -k_3\psi(t)\langle\zeta\rangle - (k_1[A] + k_3\phi(t))\langle\eta\rangle \tag{5.18}$$

Since we are interested to study the fluctuations in the linear regime of the steady state, one can approximate Eqs (5.17) and (5.18) by having $\phi(t) = \phi^s$ and $\psi(t) = \psi^s$. Notice that Eqs (5.17) and (5.18) are the same as the linearization of the deterministic equations (5.1) and (5.2) around their steady state, and we can write them in a more compact Matrix form

$$\partial_t\langle c\rangle = A\langle c\rangle \tag{5.19}$$

where $\langle c\rangle$ is the vector of the first moments of the fluctuations, and A is the Jacobian matrix as defined in (5.4). This shows that the decays of the first moments of the fluctuations are governed by the Jacobian matrix A derived from the deterministic analysis, and the time scales associated with the decays are simply the eigenvalues of A.

In the case of IFF, since the evolution of species B is not influenced by species C, the A_{12} entry in the Jacobian matrix is zero, and the eigenvalues of A are simply $A_{11} = -(k_2[A] + k_4)$ and $A_{22} = -(k_1[A] + k_3\phi^s)$. The solution to Eqs (5.17) and (5.18) is

$$\langle\zeta\rangle = \langle\zeta\rangle_0 e^{A_{11}t} \tag{5.20}$$

$$\langle\eta\rangle = \langle\eta\rangle_0 e^{A_{22}t} + A_{21}\langle\zeta\rangle_0\frac{e^{A_{11}t} - e^{A_{22}t}}{A_{11} - A_{22}} \tag{5.21}$$

From Eq (5.13), we can write down equations for the second moments of the fluctuations. Define the diffusivity matrix B as $B_{jk} = \sum_{i=1}^{R} f_i \upsilon_{ij} \upsilon_{ik}$, where f_i is the propensity of reaction i, and υ_{ij} is the stoichiometric matrix defined by (5.3). In the case of this toy model, the diffusivity matrix B is simply

$$\begin{bmatrix} B_{11} = f_2 + f_4 = k_2[A](\phi_T - \phi(t)) + k_4\phi(t) & B_{12} = 0 \\ B_{21} = 0 & B_{22} = f_1 + f_3 = k_1[A](\psi_T - \psi(t)) + k_3\phi(t)\psi(t) \end{bmatrix} \tag{5.22}$$

For the second moments,

$$\partial_t \langle \zeta^2 \rangle = 2A_{11}\langle \zeta^2 \rangle + B_{11} \tag{5.23}$$

$$\partial_t \langle \eta^2 \rangle = 2A_{21}\langle \eta\zeta \rangle + 2A_{22}\langle \eta^2 \rangle + B_{22} \tag{5.24}$$

$$\partial_t \langle \eta\zeta \rangle = A_{21}\langle \zeta^2 \rangle + A_{22}\langle \eta\zeta \rangle + A_{11}\langle \eta\zeta \rangle \tag{5.25}$$

In general, if let Ξ be the variance-covariance matrix for the numbers of molecules, then $\Xi(t)$ is governed by the following dynamics

$$\partial_t \Xi = A\Xi + \Xi A^T + B \tag{5.26}$$

To quantify the variance-covariance of the species at the stationary state, we can simply set $\partial_t \Xi = 0$ and solve the matrix equation

$$A\Xi + \Xi A^T + B = 0 \tag{5.27}$$

Notice that the diffusivity matrix B is symmetric by its construction, and it is easy to convince oneself that Ξ that satisfies Eq (5.27) is symmetric as well, fulfilling the definition of a variance-covariance matrix.

For this IFF toy model, the "sizes" of the fluctuations at the stationary state, measured by the variance-covariance matrix, are as follows

$$\langle \zeta^2 \rangle^s = -\frac{B_{11}}{2A_{11}} > 0 \tag{5.28}$$

$$\langle \zeta\eta \rangle^s = \frac{A_{21}B_{11}}{2(A_{11}+A_{22})A_{11}} < 0 \tag{5.29}$$

$$\langle \eta^2 \rangle^s = -\frac{B_{22}+2A_{21}\langle \eta\zeta \rangle^s}{2A_{22}} > 0 \tag{5.30}$$

The sign of the covariance $\langle \zeta\eta \rangle^s$ is worth special attention. The denominator is the product of two negative terms; in the numerator, B_{11} is strictly positive by definition, and A_{21}, representing the effect of species B on C, is negative. As the result, the covariance of B and C at the stationary state is negative, regardless of the rate parameter region. Physically, this is because B negatively regulates C while C has no effect on B. As one will see later, this can be a signature differentiating an incoherent feedforward loop from a negative feedback loop, where B and C regulate each other.

Last but not the least, one can calculate the correlation function matrix. In general, the time scales governing the decay of correlation functions are the same as those in Eqs (5.20) and (5.21), namely, the eigenvalues of the Jacobian matrix A. For example, to compute $\langle \zeta_0\eta_t \rangle^s$,

$$\begin{aligned}
& \langle \zeta_0\eta_t \rangle^s = \mathbb{E}[\zeta_0\eta_t] = \mathbb{E}[\mathbb{E}[\zeta_0\eta_t \,|\zeta_0\eta_0]] \\
= & \int\int \mathbb{E}[\zeta_0\eta_t \,|\zeta_0=\zeta,\eta_0=\eta] P(\zeta_0=\zeta,\eta_0=\eta)\, d\zeta\, d\eta \\
= & \int\int \zeta\, \mathbb{E}[\eta_t \,|\zeta_0=\zeta,\eta_0=\eta] P(\zeta_0=\zeta,\eta_0=\eta)\, d\zeta\, d\eta \\
= & \int\int \{\zeta\eta e^{A_{22}t} + A_{21}\zeta^2 \frac{e^{A_{11}t}-e^{A_{22}t}}{A_{11}-A_{22}}\} P(\zeta_0=\zeta,\eta_0=\eta)\, d\zeta\, d\eta \\
= & \langle \zeta\eta \rangle^s e^{A_{22}t} + A_{21}\langle \zeta^2 \rangle^s \frac{e^{A_{11}t}-e^{A_{22}t}}{A_{11}-A_{22}}
\end{aligned} \tag{5.31}$$

The first equality holds because the noise at the stationary state, by definition, has zero mean; the second equality is by the law of total expectation; the third is by the definition of the conditional expectation; the fifth is by the substitution of Eq (5.21); and the final equality is by the definition of variance and covariance.

The rest three correlation functions can be obtained in the similar way.

$$\langle \zeta_0\zeta_t \rangle^s = \langle \zeta^2 \rangle^s e^{A_{11}t} \tag{5.32}$$

$$\langle \eta_0 \zeta_t \rangle^s = \langle \zeta \eta \rangle^s e^{A_{11} t} \tag{5.33}$$

$$\langle \eta_0 \eta_t \rangle^s = \langle \eta^2 \rangle^s e^{A_{22} t} + A_{21} \langle \zeta \eta \rangle^s \frac{e^{A_{11} t} - e^{A_{22} t}}{A_{11} - A_{22}} \tag{5.34}$$

Notice that there is only one time scale involved in the decay of $\langle \zeta_0 \zeta_t \rangle^s$ and $\langle \eta_0 \zeta_t \rangle^s$, because there is only time scale governing the decaying of species B. This will not be the case if species B is coupled with species C, which is another signature different from a negative feedback loop.

5.3 LNA applied to NFB

A similar analysis is applied to a toy model of Negative FeedBack Loop (NFB), defined as follows:

$$C_{inactive} \xrightarrow{k_1[A]} C_{active}$$

$$B_{inactive} \xrightarrow{k_2[C]_{active}} B_{active}$$

$$C_{active} \xrightarrow{k_3[B]_{active}} C_{inactive}$$

$$B_{active} \xrightarrow{k_4} B_{inactive}$$

The only difference between this NFB model and the IFF model previously studied is Reaction 2. As a result, the stoichiometric matrix of this NFB model is the same as that defined in (5.3), and hence the diffusivity matrix B for this NFB model is the same as that defined in (5.22). The Jacobian matrix A for this NFB model is defined as follows

$$\begin{bmatrix} -k_2[C] - k_4 & k_2([B]_T - [B]) \\ -k_3[C] & -k_1[A] - k_3[B] \end{bmatrix} \tag{5.35}$$

Notably, the entry A_{12} is positive, due to the fact that species C up-regulates species B. It is easy to verify that the system has only one physically feasible steady state; the two eigenvalues λ_1, λ_2 are both negative, since the determinant of the Jacobian is positive while the trace is negative.

Solve Eq (5.27) for the variance-covariance matrix Ξ, notably for the covariance

term

$$\Xi_{12} = \Xi_{21} = \frac{B_{11}A_{21}A_{22} + B_{22}A_{12}A_{11}}{2(A_{11} + A_{22})(A_{11}A_{22} - A_{12}A_{21})} = \frac{B_{11}A_{21}A_{22} + B_{22}A_{12}A_{11}}{2\,\mathrm{trace}(A)\det(A)} \tag{5.36}$$

In the case of IFF, $A_{12} = 0$ reduces Eq (5.36) to Eq (5.29). Here, however, while the denominator is negative, the sign of the numerator is inconclusive: the first term in the numerator is positive, and the second is negative. The physical intuition is that while the negative regulation of species C from B gives a negative covariance between B and C, the positive regulation of species B from C gives a positive covariance; the final sign of the covariance depends on the relative strengths of these two opposing effects.

One can solve the correlation function matrix for this NFB system, and all of the four functions will have two time scales, characterized by λ_1 and λ_2. This is different from the IFF case, where two of the four correlation functions are governed by a single timescale Eqs (5.32) and (5.33). One can imagine that when correlation functions become experimentally accessible, this can be another discriminator between IFF and NFB.

5.4 Conclusion

In this work, by applying linear noise approximations to explore correlation functions of species copy numbers in two important three-node motifs, which share similar biological functions, we have successfully identified two signatures that can differentiate these two network topologies. This endeavor can be viewed as a first step to reverse engineer network topology by exploiting time series of protein concentrations.

Bibliography

[1] R. A. G. Thomas J. Kindt, Barbara A. Osborne, *Kuby Immunology* (W. H. Freeman and Company, 6th edition, 2006).

[2] N. Q. Balaban *et al.*, Science **305**, 1622 (2004).

[3] M. Thattai and A. van Oudenaarden, Genetics **167**, 523 (2004).

[4] P. B. Warren, Phys. Rev. E **80**, 030903 (2009).

[5] J. P. Roose *et al.*, Molecular and Cellular Biology **27**, 2732 (2007).

[6] J. Das *et al.*, Cell **136**, 337 (2009).

[7] L. Chang and M. Karin, Nature **410**, 37 (2001).

[8] M. J. Riese *et al.*, Journal of Biological Chemistry **286**, 5254 (2011).

[9] J. W. Locasale, A. S. Shaw, and A. K. Chakraborty, Proceedings of the National Academy of Sciences **104**, 13307 (2007).

[10] J. A. Hartigan and P. M. Hartigan, Annals of Statistics **13**, 70 (1985).

[11] N. V. Kampen, *Stochastic Processes in Physics and Chemistry* (North Holland, Elsevier, 2007).

[12] C. C. Govern, M. Yang, and A. K. Chakraborty, Phys. Rev. Lett. **108**, 058102 (2012).

[13] M. Mallaun *et al.*, The Journal of Immunology **180**, 8211 (2008).

[14] A. W. Goldrath, K. A. Hogquist, and M. J. Bevan, Immunity **6**, 633 (1997).

[15] J. Zikherman and A. Weiss, Immunity **32**, 342 (2010).

[16] A. Raj and A. van Oudenaarden, Cell **135**, 216 (2008).

[17] I. Lestas, G. Vinnicombe, and J. Paulsson, Nature **467**, 174 (2010).

[18] A. Dickson and A. R. Dinner, Annual Review of Physical Chemistry **61**, 441 (2010).

[19] P. G. Bolhuis, D. Chandler, C. Dellago, and P. L. Geissler, Annual Review of Physical Chemistry **53**, 291 (2002).

[20] W. E and E. Vanden-Eijnden, Annual Review of Physical Chemistry **61**, 391 (2010).

[21] H. Touchette, Physics Reports **478**, 1 (2009).

[22] M. I. Dykman, E. Mori, J. Ross, and P. M. Hunt, Journal of Chemical Physics **100**, 5735 (1994).

[23] R. S. Maier and D. L. Stein, Phys. Rev. E **48**, 931 (1993).

[24] E. Aurell and K. Sneppen, Phys. Rev. Lett. **88**, 048101 (2002).

[25] M. Heymann and E. Vanden-Eijnden, Phys. Rev. Lett. **100**, 140601 (2008).

[26] E. Vanden-Eijnden and M. Heymann, Journal of Chemical Physics **128**, 061103 (2008).

[27] M. Heymann, Ph.D. thesis, NYU, New York University, 2007.

[28] M. I. Dykman, I. B. Schwartz, and A. S. Landsman, Phys. Rev. Lett. **101**, 078101 (2008).

[29] M. Khasin and M. I. Dykman, Phys. Rev. Lett. **103**, 068101 (2009).

[30] M. Assaf, E. Roberts, and Z. Luthey-Schulten, Phys. Rev. Lett. **106**, 248102 (2011).

[31] D. Liu, Journal of Computational Physics **227**, 8672 (2008).

[32] C. Hartzell *et al.* (unpublished).

[33] R. J. Allen, P. B. Warren, and P. R. ten Wolde, Phys. Rev. Lett. **94**, 018104 (2005).

[34] J. O. Lauchle *et al.*, Nature **461**, 411 (2009).

[35] C. J. Marshall, Cell **80**, 179 (1995).

[36] S. D. Santos, P. J. Verveer, and P. I. Bastiaens, Nat Cell Biol **9**, 324 (2007).

[37] J. Ferrell, J. E. and E. M. Machleder, Science **280**, 895 (1998).

[38] T. K. Starr, S. C. Jameson, and K. A. Hogquist, Annu Rev Immunol **21**, 139 (2003).

[39] E. Palmer, Nat Rev Immunol **3**, 383 (2003).

[40] E. Genot and D. A. Cantrell, Curr Opin Immunol **12**, 289 (2000).

[41] D. D'Ambrosio *et al.*, Eur J Immunol **24**, 616 (1994).

[42] J. C. Stone, Genes Cancer **2**, 320 (2011).

[43] N. A. Dower *et al.*, Nat Immunol **1**, 317 (2000).

[44] M. Oh-hora *et al.*, The Journal of Experimental Medicine **198**, 1841 (2003).

[45] S. M. Margarit *et al.*, Cell **112**, 685 (2003).

[46] T. S. Freedman *et al.*, Proc Natl Acad Sci U S A **103**, 16692 (2006).

[47] A. Prasad *et al.*, Proceedings of the National Academy of Sciences **106**, 528 (2009).

[48] M. Rincon and G. Pedraza-Alva, Immunol Rev **192**, 131 (2003).

[49] G. Werlen, B. Hausmann, D. Naeher, and E. Palmer, Science **299**, 1859 (2003).

[50] T. J. Yun and M. J. Bevan, Nat Immunol **2**, 13 (2001).

[51] Q. Gong *et al.*, Nat Immunol **2**, 29 (2001).

[52] A. M. Fischer *et al.*, Immunity **23**, 431 (2005).

[53] M. A. McGargill *et al.*, J Immunol **183**, 4838 (2009).

[54] C. Dong, R. J. Davis, and R. A. Flavell, Annu Rev Immunol **20**, 55 (2002).

[55] J. J. Priatel *et al.*, Immunity **17**, 617 (2002).

[56] R. L. Kortum *et al.*, Proc Natl Acad Sci U S A **108**, 12407 (2011).

[57] A. Nimnual and D. Bar-Sagi, Sci. STKE **2002**, pe36 (2002).

[58] N. I. Markevich, J. B. Hoek, and B. N. Kholodenko, The Journal of Cell Biology **164**, 353 (2004).

[59] E. C. O'Shaughnessy, S. Palani, J. J. Collins, and C. A. Sarkar, Cell **144**, 119 (2011).

[60] T. Tian *et al.*, Nat Cell Biol **9**, 905 (2007).

[61] G. Altan-Bonnet and R. N. Germain, PLoS Biol **3**, e356 (2005).

[62] A. Harding *et al.*, Current Biology **15**, 869 (2005).

[63] C. J. Bashor, N. C. Helman, S. Yan, and W. A. Lim, Science **319**, 1539 (2008).

[64] S. Boykevisch *et al.*, Curr Biol **16**, 2173 (2006).

[65] A. Harding and J. F. Hancock, Cell Cycle **7**, 127 (2008).

[66] A. S. Nimnual, B. A. Yatsula, and D. Bar-Sagi, Science **279**, 560 (1998).

[67] G. Scita *et al.*, Nature **401**, 290 (1999).

[68] A. Mettouchi *et al.*, Molecular Cell **8**, 115 (2001).

[69] I. K. Jang *et al.*, Proceedings of the National Academy of Sciences **107**, 10620 (2010).

[70] A. Hashimoto *et al.*, The Journal of Experimental Medicine **188**, 1287 (1998).

[71] M. Innocenti *et al.*, The Journal of Cell Biology **160**, 17 (2003).

[72] J. C. D. Houtman *et al.*, Nat Struct Mol Biol **13**, 798 (2006).

[73] H. C. Welch, W. Coadwell, L. R. Stephens, and P. T. Hawkins, FEBS Letters **546**, 93 (2003), ¡ce:title¿Signal Transduction Special Issue¡/ce:title¿.

[74] V. L. J. Tybulewicz and R. B. Henderson, Nat Rev Immunol **9**, 630 (2009).

[75] A. Gerard *et al.*, Blood **113**, 6138 (2009).

[76] J. M. Lambert *et al.*, Nat Cell Biol **4**, 621 (2002).

[77] M. Turner and D. D. Billadeau, Nat Rev Immunol **2**, 476 (2002).

[78] J. Han *et al.*, Science **279**, 558 (1998).

[79] M. J. Caloca *et al.*, EMBO J **22**, 3326 (2003).

[80] L. F. Reynolds *et al.*, Journal of Biological Chemistry **279**, 18239 (2004).

[81] J. L. Zugaza, M. J. Caloca, and X. R. Bustelo, Oncogene **23**, 5823 (2004).

[82] V. M. M. Braga, Nat Cell Biol **4**, E188 (2002).

[83] T. Sanui *et al.*, Immunity **19**, 119 (2003).

[84] C. Brockmeyer *et al.*, Journal of Biological Chemistry **286**, 7535 (2011).

[85] D. T. Gillespie, The Journal of Physical Chemistry **81**, 2340 (1977).

[86] P. Faccioli, M. Sega, F. Pederiva, and H. Orland, Phys. Rev. Lett. **97**, 108101 (2006).

[87] *A Guide to to First-Passage Processes* (Cambridge University Press, August 6, 2001).

[88] C. Gardiner, *Handbook of Stochastic Methods: for Physics, Chemistry and the Natural Sciences* (Springer; 3rd edition, Springer Series in Synergetics, 2004).

[89] U. Alon, *An Introduction to Systems Biology: Design Principles of Biological Circuits* (Chapman and Hall/CRC, Mathematical and Computational Biology, 2006).

[90] M. A. Daniels *et al.*, Nature **444**, 724 (2006).