

# SPATIOTEMPORAL COORDINATION OF SIGNALING AT SINGLE MOLECULE RESOLUTION

Michael Rhys Griarte Pablo

A dissertation submitted to the faculty at the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Chemistry

Chapel Hill  
2020

Approved by:

Timothy C. Elston

Klaus M. Hahn

Dorothy A. Erie

Nancy L. Thompson

Richard Superfine II

Gary J. Pielak

© 2020  
Michael Rhys Griarte Pablo  
ALL RIGHTS RESERVED

## ABSTRACT

Michael Rhys Griarte Pablo: *Spatiotemporal coordination of signaling at single molecule resolution*  
(Under the direction of Timothy Elston and Klaus Hahn)

Advances in live-cell single-molecule imaging and modeling over the past decade have invited the closer study of biological structure and dynamics at the nanoscale. The higher resolution of these single-molecule experiments results in finely-grained datasets that can feed detailed quantitative models. Likewise, single-molecule models can account for microscopic details such as noise and heterogeneity inherent to diffusional and chemical processes, which are often neglected in models based on bulk concentrations. Examining microscale biological structures at single molecule resolution in living cells has led to new findings, such as the dynamic regulation of nanoscale structure. I cover three topics from the perspective of single molecules. Chapters 1-3 are on modeling the spatiotemporal coordination of both spontaneous and pheromone-guided yeast polarity establishment. Chapter 4 is on computational modeling and analysis for a technique called Binder/Tag, which we applied to study the conformational dynamics of the protein Src kinase in living cells. Chapter 5 is on modeling clustering-mediated activation of immunoreceptors, using the phagocytic receptor Fc $\gamma$ RIIA as a prototypical example.

This is dedicated to family, friends, and teachers  
who believed in me when I did not believe in myself.



## ACKNOWLEDGEMENTS

I am forever grateful for the support my family has given me throughout my PhD. Your calls, texts, and visits were sources of kindness, love, and energy that lifted me up. No matter what happened during my time here at UNC, I could rely on you. I've held onto all the mail you sent me over the past five years as a physical reminder of your support. Thank you so much.

I've had the fortune to be surrounded by many wonderful people who have been sources of friendship and inspiration. Amy Pomeroy and Kimiko Suzuki, it was a pleasure to spend the past five years together as labmates in spirit, even if not in location. J. Cody Herron, thank you for working with me on our shared projects; I hand off the torch to you! Sam Ramirez, Bei Liu, Orrin Stone, Khem Ghusinga, Kaiyun Guan, Nick Elston and Nick Henderson – thank you for collaborating with me and improving my work. Thank you as well to many of the other Elston and Hahn lab members who have offered both advice and jokes over the years, including Vinal Lakhani, Callie Miller, Jeff Snell, Shuang Li, Takashi Watanabe, Ellen O'Shaughnessy, Dan Marston, Nick Pinkin, and Neha Pankow, Ana Nogueira, Shiqiong Hu, Fred Pimenta, and Joe Szulczewski. Beyond the lab, I thank Sam Stadtmiller, Emily Lentz, and Ben Morgan, for our many fun get-togethers spanning food, art, games, and TV. I am also incredibly lucky to have met my partner Tom Swiderski while at UNC. In the past year, you have been supportive, inspiring, and a joy to be around.

Last but not least, I have been blessed with brilliant mentors. Tim and Klaus, thank you for taking a chance on me, despite my lack of experience in either of your fields of expertise. Your guidance and support over the years have helped me to grow both as a person and a scientist. Thank you also to Danny Lew at Duke for his support and collaboration, and likewise to Beverly Errede and Henrik Dohlman for their insightful comments in lab meetings over the years. I also need to thank the Molecular and Cellular

Biophysics Training program, and especially Lisa Phillipie. Her support and the program's support made it possible for me to take my first steps towards the work presented here.

## TABLE OF CONTENTS

LIST OF TABLES .....	xi
LIST OF FIGURES .....	xii
LIST OF ABBREVIATIONS .....	xv
Chapter 1: Noise drives organization in a Turing-type model of yeast polarity establishment .....	1
Overview .....	1
1.1. Introduction .....	2
1.2. Establishing a connection between macroscopic and microscopic rate constants .....	6
1.3. Previously reported kinetic parameters are strongly diffusion-limited .....	8
1.4. Microscopic fluctuations promote polarity establishment .....	10
1.4.1. Enhanced polarization within 2D simulations .....	10
1.4.2. Enhanced polarization within quasi-3D simulations .....	14
1.5. Discussion .....	19
1.6. Methods .....	22
1.6.1. The molecular circuit for polarity establishment .....	22
1.6.2. Particle-based simulations in 2D .....	25
1.6.3. Particle-based simulations in quasi-3D .....	26
1.6.4. Reaction-diffusion partial differential equation simulations .....	28
1.6.5. Empirically mapping macroscopic to microscopic rate constants .....	28
1.7. Supplemental Methods and Derivations .....	33
1.7.1. Quantifying polarization .....	33

1.7.2.	Determining bifurcation points .....	35
1.7.3.	Estimating the 2D diffusion limit.....	37
1.7.4.	Deriving and validating the 2D $\lambda - \varrho$ theory.....	39
1.7.5.	Deriving integrals for quasi-3D injection and ejection .....	44
1.7.6.	Loss of polarity with increasing membrane diffusivity .....	46
1.7.7.	An alternative to the quasi-3D system .....	46
REFERENCES .....		48
Chapter 2: Ratiometric GPCR signaling enables directional sensing in yeast.....		52
Overview.....		52
2.1.	Introduction.....	53
2.2.	Nonuniform GPCR distributions can mislead G-protein activation .....	54
2.3.	Ratiometric GPCR signaling robustly directs G-protein activation.....	56
2.4.	Ratiometric GPCR signaling can sharpen the G-protein gradient .....	59
2.5.	Discussion .....	60
2.6.	Methods.....	62
2.6.1.	Particle-based simulations of ratiometric and nonratiometric gradient sensing.....	62
2.6.2.	Establishing receptor density and activity gradients.....	63
2.6.3.	Calibrating G-protein inactivation rates.....	63
2.6.4.	Particle-based simulations of receptor gradient degradation .....	64
REFERENCES .....		66
Chapter 3: Yeast polarization toward mating partners.....		69
Overview.....		69
3.1.	Introduction.....	70
3.2.	Stable polarization is insensitive to shallow gradients.....	72

3.3.	Indecisive polarization is sensitive to shallow gradients .....	75
3.4.	Can polarization during the indecisive phase guide a stable polarity site? .....	78
3.5.	Pheromone gradients under mating conditions may be highly non-linear .....	78
3.6.	Discussion .....	80
3.7.	Methods .....	81
3.7.1.	Particle-based simulations of GPCR-coupled polarity establishment .....	81
3.7.2.	Quantifying patch movement with frame-to-frame correlation .....	83
3.7.3.	Analyzing polarization relative to the gradient .....	83
3.7.4.	Simulating pheromone gradients experienced by a mating pair .....	83
	REFERENCES .....	85
Chapter 4: Probing conformational changes of single molecules in living cells with Binder/Tag .....		87
	Overview .....	87
4.1.	Introduction .....	89
4.2.	Single particle tracking of total and active Src with Binder/Tag .....	90
4.3.	Accumulation of active Src in adhesions in a size-dependent process .....	94
4.4.	Nanoscale clusters of Src are dynamic hotspots of activity .....	97
4.5.	Inferring <i>in vivo</i> biochemical kinetics .....	102
4.6.	Discussion .....	107
4.7.	Methods .....	110
4.7.1.	Single molecule microscopy .....	110
4.7.2.	Diffusional analysis .....	111
4.7.3.	Correlating adhesion imaging with single particle tracking .....	112
4.7.4.	Cluster analysis .....	112
4.7.5.	Co-diffusion events and kinetic modeling .....	114
4.7.6.	Statistical tests .....	115

REFERENCES .....	116
Chapter 5: Modeling clustering-induced activation of receptor immunotyrosine activation motifs ....	120
Overview.....	120
5.1. Introduction.....	121
5.2. Clustering-mediated receptor activation with tunable response thresholds .....	123
5.3. Investigating the role of specific parameters in controlling the clustering threshold .....	126
5.4. Full and committed vs. partial and indecisive cluster activation .....	129
5.5. Discussion and Future Directions .....	131
5.5.1. Simulation speed and model simplicity .....	132
5.5.2. Receptor co-clustering .....	132
5.5.3. Receptor pre-clustering.....	133
5.5.4. Downstream feedback regulation of clustering and activation .....	133
5.5.5. Noise in receptor abundance and activation.....	134
5.6. Methods.....	136
5.6.1. Particle-based simulation .....	136
5.6.2. Parameter optimization via evolutionary algorithm.....	137
REFERENCES .....	138

## LIST OF TABLES

Table 1.1. A list of common computational approaches for simulating spatial and temporal stochasticity in biochemical reaction networks.....	5
Table 1.2. Microscopic parameters and effective macroscopic rate constants for reversible/irreversible bimolecular reactions of the form $A + B \leftrightarrow C$ .....	10
Table 1.3. Parameters used to perform simulations described in the main text.....	24
Table 1.4. Microscopic parameters and effective macroscopic parameters for reversible/irreversible bimolecular reactions of the form $A + B \leftrightarrow C$ , individual fits to each initial condition.....	32
Table 3.1 Model parameters. ....	82

## LIST OF FIGURES

Figure 1.1. Computational modeling schematics.....	5
Figure 1.2. Illustration of the different reaction regimes (reaction-limited, diffusion-influenced, and diffusion-limited) and the range of validity for the 2D $\lambda$ - $q$ theory. ....	7
Figure 1.3. Empirical estimates for macroscopic rate constants in the yeast polarity model for the two different parameter sets. ....	9
Figure 1.4. Simulations of polarity establishment within the Turing unstable regime. ....	11
Figure 1.5. Variability in 2D polarization from microscopic fluctuations. ....	13
Figure 1.6. Stochasticity facilitates polarization. ....	14
Figure 1.7. Quasi-3D particle-based simulations of the polarity establishment model. ....	15
Figure 1.8. Quantitative comparisons of polarization in quasi-3D particle-based simulations and corresponding RDEs. ....	17
Figure 1.9. The effect of Cdc42 concentration on polarization for quasi-3D particle-based simulations. .	18
Figure 1.10. Reservoir approach schematics and validation. ....	27
Figure 1.11. Comparisons of deterministic rate equations in 2D and the polarity establishment network in a Turing stable regime.....	29
Figure 1.12. Fitting simulations using 2D parameters to the deterministic rate equations. ....	30
Figure 1.13. Fitting simulations using quasi-3D parameters to the deterministic rate equations. ....	31
Figure 1.14. Different choices of $r$ in $H(r)$ do not change the qualitative features of our results. ....	34
Figure 1.15. Representative $H(r)$ curves for the particle-based and RDE simulations showing similarity in polarization. ....	34
Figure 1.16. Polarity establishment is still slower for the RDE model if equations are seeded with later distributions from the particle-based simulations. ....	35
Figure 1.17. Bifurcation point identification with linear stability analysis.....	36
Figure 1.18. Quantifying $H(r)$ at $t=1800s$ for pre-polarized q3D-RDEs is a reasonable marker for maintenance vs. loss of polarity.....	37
Figure 1.19. Accepted parameters for the yeast polarization model appear supra-diffusive in a 2D context. ....	39
Figure 1.20. Comparisons between the deterministic rate equations in 2D and multiple realizations of the particle-based simulation for the simple reversible bimolecular reaction $A+B \leftrightarrow C$ . ....	43



Figure 1.21. Comparisons between the deterministic rate equations in 2D and multiple realizations of the particle-based simulation for the polarity establishment network under reaction-limited, non-polarizing conditions. ....	44
Figure 1.22. Loss of polarization with increasing membrane diffusivity. ....	46
Figure 1.23. Considering the volume-adjusted, two-compartment RDE system makes no qualitative difference for our observations. ....	47
Figure 2.1. GPCR density variation along the cell membrane can mislead G protein activation. ....	56
Figure 2.2. Ratiometric sensing allows cells to orient up-gradient despite uneven receptor density. ....	58
Figure 2.3. Ratiometric sensing amplifies the gradient signal and improves accuracy even when receptors are distributed uniformly. ....	60
Figure 2.4. Calibration of G-protein inactivation rates for model comparison, and effect of diffusion-limited versus reaction-limited regimes. ....	64
Figure 2.5. Robustness of simulation results to varying receptor abundance and diffusion. ....	65
Figure 2.6. Effects of receptor diffusion in particle simulations. ....	65
Figure 3.1. GPCR signaling couples pheromone sensing to the Cdc42-mediated polarity machinery. ....	73
Figure 3.2. Stable polarization is insensitive to shallow gradients. ....	74
Figure 3.3. Stable polarization is insensitive to shallow gradients. ....	76
Figure 3.4. The indecisive polarity phenotype is sensitive to shallow gradients. ....	77
Figure 3.5. Pheromone gradients experienced by a simulated emitter and receiver cell. ....	79
Figure 4.1. Applying Binder/Tag to probe Src conformations with single particle tracking (SPT). ....	92
Figure 4.2. Working expression levels of tagSrc and Binder are non-perturbing. ....	93
Figure 4.3. Two independent diffusional analyses reveal slow-diffusing open-form Src. ....	94
Figure 4.4. Interweaving adhesion imaging and dual-color single particle tracking of tagSrc and Binder reveals accumulation of slow-moving active Src. ....	96
Figure 4.5. Observing nanoscale clusters of Src. ....	98
Figure 4.6. Src clusters are hotspots of activity. ....	100
Figure 4.7. Src clusters are dynamic, heterogenous structures. ....	101
Figure 4.8. Inferring <i>in vivo</i> Src regulatory kinetics through co-diffusion analysis. ....	103
Figure 4.9. Controls for analyzing <i>in vivo</i> kinetics. ....	105

Figure 5.1. Simulating clustering-mediated activation of receptors. ....	125
Figure 5.2. Parameter distributions obtained at different levels of clustering-triggered activation. ....	126
Figure 5.3. Effects of perturbing specific parameters one-by-one. ....	127
Figure 5.4. Full and committed versus partial and indecisive modes of cluster activation.....	130
Figure 5.5. Even under strongly-activating conditions, full activation dominates at the low activation threshold (“dimer”), but disappears at higher activation thresholds (“decamer”).....	130

## LIST OF ABBREVIATIONS

2D	Two dimensional
3D	Three dimensional
ANOVA	Analysis of Variance
BCR	B cell receptor
CI	Confidence Interval
EGFR	Epidermal growth factor receptor
FcR	Fc receptor, or fragment, crystallizable receptor
FRET	Förster resonance energy transfer
GAP	GTPase-activating protein
GDI	Guanine nucleotide dissociation inhibitor
GDP	Guanosine diphosphate
GEF	Guanine nucleotide exchange factor
GPCR	G-protein coupled receptor
GTP	Guanosine triphosphate
HMM	Hidden Markov Model
Ig	Immunoglobulin
ITAM	Immunotyrosine activation motif
KDE	Kernel density estimate
MAT	Mating type
MEF	Mouse embryonic fibroblast
MSD	Mean squared displacement
MSS	Moment scaling spectrum
PALM	Photo-activated localization microscopy
PB	Particle-based

PDE	Partial differential equation
PM	Plasma membrane
q3D	Quasi-three dimensional
RDE	Reaction-diffusion equation
SFK	Src family kinase
SH2	Src homology 2
SH3	Src homology 3
SPT	Single particle tracking
TCR	T cell receptor
TIRF	Total internal reflection fluorescence
Tyr	Tyrosine

## CHAPTER 1:NOISE DRIVES ORGANIZATION IN A TURING-TYPE MODEL OF YEAST POLARITY ESTABLISHMENT<sup>1</sup>

### Overview

Cell polarity is the asymmetric localization of signaling molecules. Polarity provides directionality in fundamental processes such as migration, growth, and division. The changes in cell shape inherent to such processes are driven in part by the local accumulation of active Rho-family GTPases, which regulate cytoskeletal remodeling (Etienne-Manneville, 2004). In the budding yeast *Saccharomyces cerevisiae*, polarization of the Rho GTPase Cdc42 is required during both the budding and mating reproductive programs (Slaughter et al., 2009). The protein Cdc42 is a master polarity regulator conserved among Ascomycetes, including budding yeast and fission yeast, as well as in more complex organisms such as humans (Chiou et al., 2017). In *S. cerevisiae*, Cdc42 is the only Rho GTPase important for polarization, but in other systems, other Rho GTPases can play complimentary and antagonistic roles. In addition, *S. cerevisiae* has been historically favored as a model organism because of its amenability to genetic manipulation (Duina et al., 2014). Thus, because of its experimental tractability and biochemical simplicity, *S. cerevisiae* is a powerful model for studying polarity establishment.

Cell polarization is often guided by spatial cues. *S. cerevisiae* (hereafter, yeast) polarize up gradients of pheromone to find mates, the social amoebae *Dictyostelium discoideum* polarize up gradients of cyclic adenosine monophosphate to aggregate into fruiting bodies, and the axons of hippocampal neurons grow either up or down gradients of Netrin-1 depending upon the overall concentration (Bhattacharjee and Folch, 2017; Rappel and Edelstein-Keshet, 2017; von Philipsborn and Bastmeyer, 2007). However, cell

---

<sup>1</sup> This chapter previously appeared as an article in the journal PLoS Computational Biology. The original citation is as follows: Pablo M, Ramirez SA, Elston TC. “Particle-based Simulations of Polarity Establishment Reveal Stochastic Promotion of Turing Pattern Formation,” *PLoS Comp. Biol.* 14(3): e1006016.

polarization can occur independent of gradients. Yeast can polarize in the absence of pheromone gradients and bud guidance machinery. *Dictyostelium* pseudopods can form at rates independent of gradients (Andrew and Insall, 2007). Hippocampal neurons can randomly grow dendrites and axons when grown on a plate in the absence of gradients (Tahirovic and Bradke, 2009). A possible explanation for cue-less polarization is the amplification of noise by positive feedback loops, which allow symmetry-breaking molecular fluctuations to initiate polarity establishment.

Investigating the spatiotemporal dynamics of the Rho GTPase circuitry is fundamental to understanding polarity establishment. Historically, mathematical models and numerical simulations of the biochemistry and biophysics have provided insights into the intrinsic properties of signaling circuitry, complementary to experimental investigations. Famously, Turing-type models provide a mathematical framework for understanding how spontaneous patterns can emerge within biochemical reaction-diffusion systems (Turing, 1952). These models are classically seeded with an initial noisy distribution and allowed to evolve deterministically, which neglects the constitutive fluctuations over time intrinsic to individual molecules. Conceptually, this noise could disrupt nascent polarity sites. Here we demonstrate instead that the noise inherent to diffusional and chemical processes can enhance cell-scale polarity establishment.

## **1.1. Introduction**

Yeast polarization is typically guided by spatial cues such as bud scars and pheromone gradients. Polarization is still possible when these cues are removed, albeit in random directions (Slaughter et al., 2009). Weak asymmetries created by molecular noise may provide the impetus for spontaneous polarization, as proposed by mathematical models of spontaneous patterning in biochemical systems (Gierer and Meinhardt, 1972; Turing, 1952). These models use diffusion-driven instabilities of the chemical reaction dynamics to generate symmetry breaking without relying on mechanisms such as diffusional barriers, directed transport, and molecular cross-linking. Instead, these systems require: (1) positive feedback to amplify local fluctuations; (2) chemical species that diffuse at different rates; and (3) a mechanism for limiting the growth of the polarity site. Models in which patterning can be induced by an

arbitrarily weak perturbation (e.g. molecular-level fluctuations) are called Turing-type. Goryachev and Pokhilko were the earliest to use a Turing-type model to study yeast polarization (Goryachev and Pokhilko, 2008). Other, non-Turing type models of polarity require perturbations of finite strength to induce pattern formation, but still require the same fundamental ingredients: positive feedback, differences in diffusion rates, and a way to limit the growth of the polarity site (Mori et al., 2008).

A common approach to modeling the spatiotemporal dynamics of polarizing biochemical systems is to use deterministic reaction-diffusion equations (RDEs) in the form of non-linear partial differential equations (PDEs). Deterministic RDEs ignore stochastic effects intrinsic to chemical reactions and thermal diffusion. However, stochastic effects are known to enhance pattern formation in some biochemical systems (Dziekan et al., 2012; Lemarchand and Nowakowski, 2011). There are many approaches to study stochastic effects in biological signaling models. In roughly increasing order of granularity, there are stochastic differential equations, such as chemical Langevin equations (Gillespie, 2000; McKane et al., 2014); spatially discretized, temporally-continuous approaches, such as the spatial Gillespie algorithm (Fange et al., 2010; Gillespie, 1977; Hattne et al., 2005); direct particle-based simulations, as implemented in Smoldyn and MCell (Andrews and Bray, 2004; Kerr et al., 2008); and exact Brownian dynamics, such as Green's function reaction dynamics (van Zon and ten Wolde, 2005a, 2005b). Rather than attempting to cover the full spectrum of approaches and computational tools here, we refer the reader to excellent reviews that describe the theoretical underpinnings and software implementations of these methods (Agbanusi and Isaacson, 2014; Angermann et al., 2012; Schöneberg et al., 2014; Sept and Carlsson, 2014). The advantages and limitations for some of the more common methods are described in Table 1.1. Hybrid approaches that mix particle simulations with a deterministic partial differential equation solver are most similar to the approach taken here (Schaff et al., 2016).

A variety of stochastic methods have been used to study the effects of noise in non-Turing models of yeast polarization (Altschuler et al., 2008; Schaff et al., 2016; Walther et al., 2012). Some models, such as the neutral drift polarity model, used particle-based approaches (Jilkine et al., 2011); others, like models based on wave-pinning, used Gillespie or stochastic PDE-based approaches (Ozbudak et al., 2005;

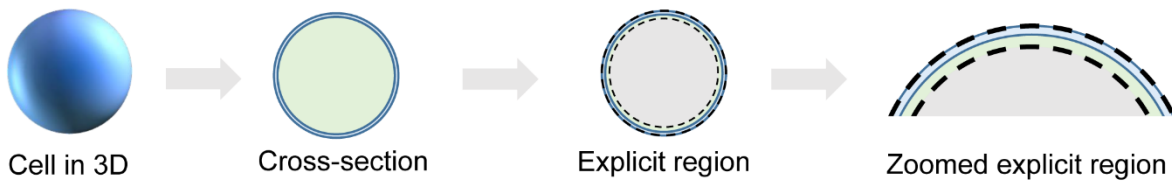
Wedlich-Soldner et al., 2003). Other investigations of stochasticity in polarization with more detailed signaling models, including the Turing-type Goryachev–Pokhilko model, leveraged Gillespie and stochastic PDE approaches (Freisinger et al., 2013; Wu et al., 2015). Here, we present stochastic particle-based simulations of the Goryachev–Pokhilko model, and compare them to deterministic RDE simulations of the same system to evaluate stochastic effects on polarization. In this model, bimolecular reactions can occur between two reactants on the membrane, and between a reactant on the membrane and a reactant in the cytoplasm. Molecules can exchange between the membrane and cytoplasm. These simulations explicitly track molecules at and near the cell membrane, where polarization occurs, and implicitly handle molecules away from the membrane. We present two different scenarios. In the first, the cell membrane and the nearby cytoplasm are assumed to be purely two-dimensional (2D), ignoring the remaining bulk cytoplasm. In the second, the bulk cytoplasm is approximated by attaching a molecular reservoir that tracks molecular abundances without explicitly simulating individual molecules. Mass is stochastically exchanged between the 2D particle-based domain and the reservoir with rates determined by diffusion, thus creating a quasi-3D system (Figure 1.1).



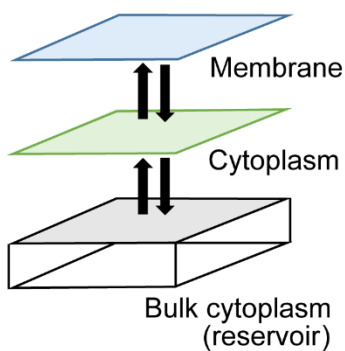
Approach	Spatially discretized?	Temporally discretized?	Comment
Stochastic partial differential equations	Yes	Yes	More efficient than particle-based simulations. Breaks down in low concentration limits.
Spatial Gillespie	Yes	No	Can be more efficient than particle-based simulations. Can suffer from artifacts due to spatial discretization.
Hybrid particle-based-PDE methods (this work)	No	Yes	Approximation of full particle-based methods by explicitly modeling only a portion of the domain of interest, and implicitly modeling the remainder using either deterministic or stochastic methods.
Fully particle-based	No	Yes	More accurate than PDE and Gillespie approaches. Computationally demanding.
Exact Brownian dynamics	No	No	More accurate than particle-based simulation, but can come at a higher computational demand. Also called an “event-driven” approach, but this can cause confusion with Gillespie-type approaches.

**Table 1.1. A list of common computational approaches for simulating spatial and temporal stochasticity in biochemical reaction networks.**

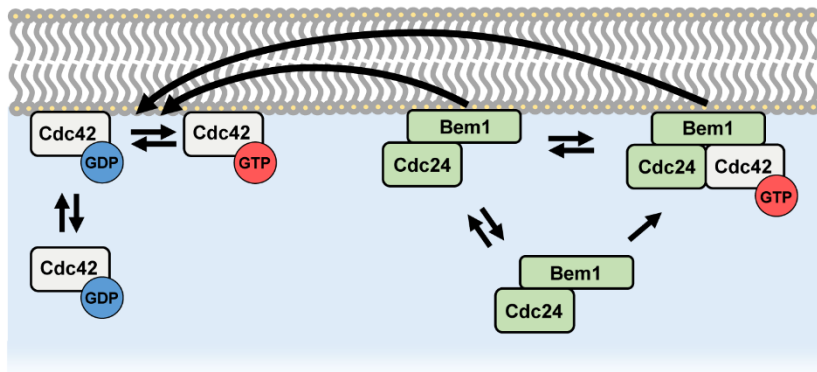
**A**



**B**



**C**



**Figure 1.1. Computational modeling schematics.**

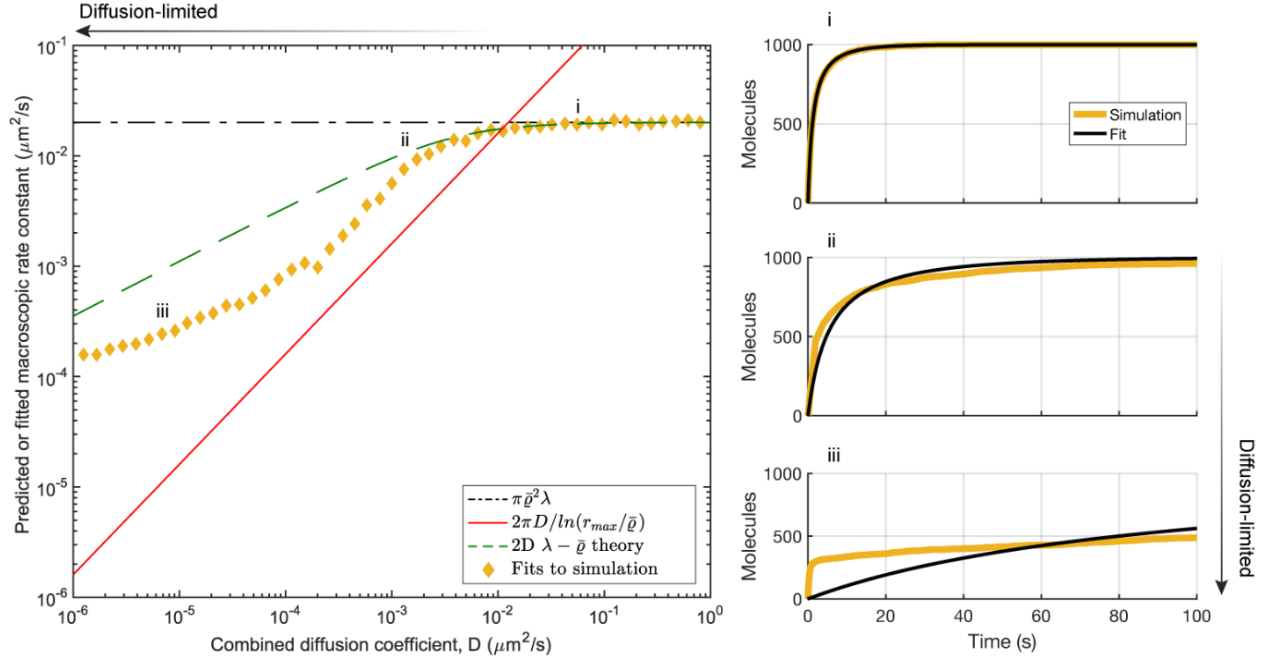
(A) Molecules at the cell membrane and a thin slice of adjacent cytoplasm are simulated explicitly. Both compartments are modeled as 2D surface. (B) In the quasi-3D simulations, a well-stirred compartment representing bulk cytoplasm is added to approximate 3D effects. (C) Reaction scheme for Turing-type Cdc42-dependent yeast polarity establishment, based on the Goryachev and Polkhilko model.

## 1.2. Establishing a connection between macroscopic and microscopic rate constants

Investigating the role of molecular fluctuations in polarity establishment requires a way to compare stochastic particle-based simulation results to the behavior of the system in the deterministic macroscopic limit, where the spatiotemporal dynamics of biochemical concentrations are governed by RDEs.

Therefore, it is critical to relate microscopic parameters to macroscopic rate constants. For first-order reactions, this is trivial, and follows the relation  $P = 1 - \exp(-k\Delta t)$ , where  $P$  is the probability of the reaction for an individual molecule,  $k$  is the macroscopic rate constant, and  $\Delta t$  is the simulation timestep. The situation is more complicated for second-order reactions, because all of the reactions in this system involve the cell membrane.

In chemical kinetic theory, there are two limiting regimes for second-order reactions: the diffusion limit and the reaction limit. The diffusion limit is the maximum rate at which a second-order reaction can proceed, and is where two particles react upon their first encounter. In solution or in the cytoplasm (i.e. in 3D), it is possible to define a macroscopic rate constant in the diffusion limit by considering the diffusional flux through an absorbing sphere of radius  $\bar{q}$  located at the origin, when the concentration  $C$  of the reactant is held fixed at infinity (Erban and Chapman, 2009; Lipková et al., 2011). The flux is given by  $J = 4\pi D\bar{q}C$ , where  $D$  is the sum of the diffusion coefficients of the reactants, and the diffusion-limited rate constant is  $k = 4\pi D\bar{q}$ . However, on membranes (i.e. in 2D), diffusion-limited rate constants are not well-defined (Fange et al., 2010; Yogurtcu and Johnson, 2015). As an estimate of the appropriate time scale, the flux through an absorbing circle of radius  $\bar{q}$  when the computational domain remains finite (see 1.7.3 Estimating the 2D diffusion limit). The flux is given by  $J = 2\pi DC / \ln(r_{max}/\bar{q})$ , where  $r_{max}$  characterizes the size of the computational domain. In contrast to the 3D case, in the limit  $r_{max} \rightarrow \infty$ , the 2D flux goes to zero. We used the flux on a finite domain to estimate a time scale for second-order diffusion-limited reactions,  $k_{DL} = 2\pi D / \ln(r_{max}/\bar{q})$ , which has the units of a 2D second-order rate constant. This expression is represented by the red curve in Figure 1.2.



**Figure 1.2. Illustration of the different reaction regimes (reaction-limited, diffusion-influenced, and diffusion-limited) and the range of validity for the 2D  $\lambda - \bar{q}$  theory.**

The left panel shows estimated rate constants (yellow diamonds) for the 2D second order reaction  $A+B \rightarrow C$  obtained by fitting chemical rate equations (black curves, right panels) to results from particle-based simulations (yellow curves, right panels). The reaction limit,  $k_{RL} = \pi\bar{q}^2\lambda$ , is indicated by the black dot-dashed line and the estimate for the diffusion limit  $k_{DL} = 2\pi D / \ln(r_{max}/\bar{q})$  is represented by the red curve. The results from the 2D  $\lambda - \bar{q}$  theory are shown as the green-dashed line. Parameters chosen are  $D_A = D_B = \frac{1}{2}D$  (x-axis of left panel),  $\lambda = 2.5554 \text{ s}^{-1}$ ,  $r_{max} = 2.5 \text{ } \mu\text{m}$ ,  $\bar{q} = 0.05 \text{ } \mu\text{m}$ . Simulations were conducted on a  $L = 5 \text{ } \mu\text{m}$  domain.

In the reaction limit, multiple encounters on average are required before the reaction occurs. We computed a rate constant in the reaction limit by assuming the reactants are uniformly distributed. In 2D, this produces an overall reaction rate of  $(\pi\bar{q}^2/A)\lambda N_A N_B$ , where  $\pi\bar{q}^2$  is the capture area,  $A$  is the area of the system,  $\lambda$  is the microscopic reaction rate, and  $N_A$  and  $N_B$  are the particle numbers for the two reactants. This leads to a rate constant of  $k_{RL} = \pi\bar{q}^2\lambda$ , plotted as the black dashed line in Figure 1.2.

The macroscopic rate constants for bimolecular reactions can be calculated from the underlying microscopic parameters using  $\lambda - \bar{q}$  theory. The theory assumes the two reactants have a summed diffusivity  $D$ , and that reactions proceed with rate  $\lambda$  if the two reactants are within  $\bar{q}$  of one another. In 3D, the  $\lambda - \bar{q}$  theory can be used in both the reaction and diffusion limit; i.e. regardless of the reactants' diffusion coefficients (Erban and Chapman, 2009; Lipková et al., 2011). In 2D, the  $\lambda - \bar{q}$  theory can map

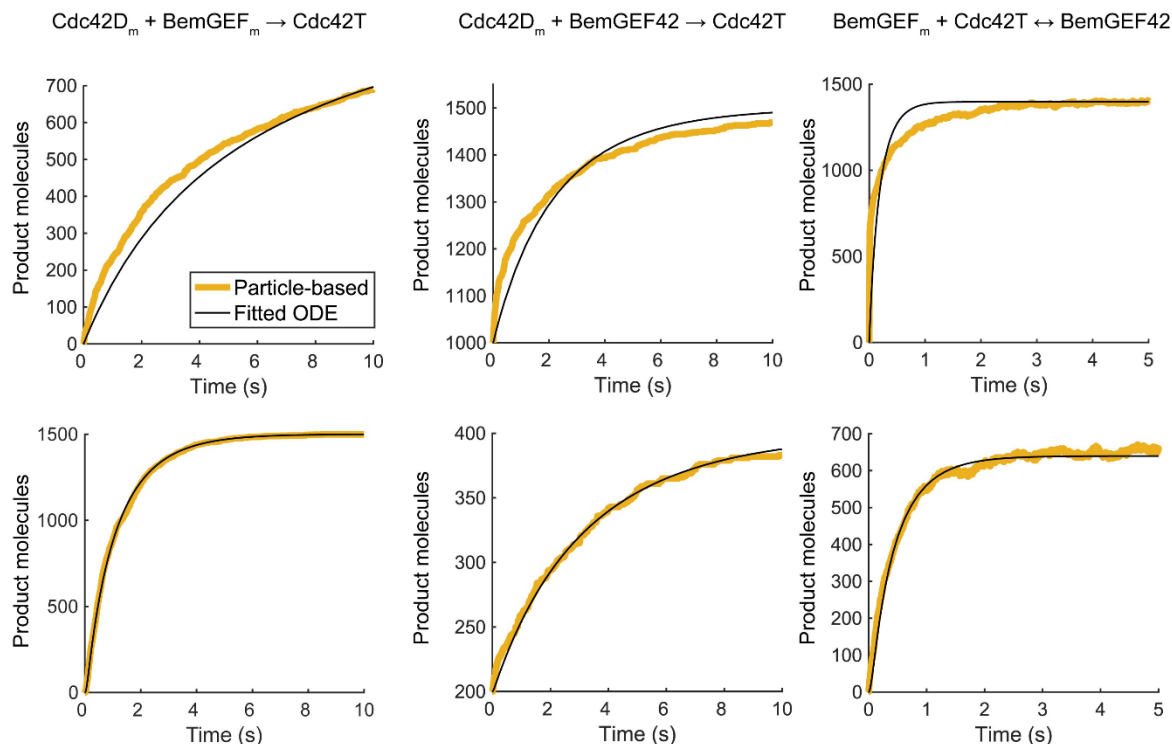
macro- and microscopic parameters only in the reaction limit. Despite that, we used the  $\lambda - \bar{q}$  formalism to compute 2D rate constants (see 1.7.4 Deriving and validating the 2D  $\lambda - \bar{q}$  theory). This was a reasonable prediction of reaction kinetics (green dashed line vs. yellow diamonds, left panel, Figure 1.2) if the system was not too far from the reaction limit (black dot-dashed line).

### 1.3. Previously reported kinetic parameters are strongly diffusion-limited

We attempted to estimate microscopic rate constants  $\lambda$  from macroscopic rate constants  $k$  used in published models of yeast polarity establishment. In doing so, we found that several published second-order rate constants were orders of magnitude larger than our estimate for the diffusion limit,  $k_{DL}$  (Figure 1.19). The most likely explanation is that  $k_{DL}$  does not account for the membrane-cytoplasm exchange of polarity molecules. In the biochemical network, reactive chemical species can exchange between the membrane, where diffusion is relatively slow, and the cytoplasm, where diffusion is relatively fast. The different diffusion coefficients associated with these cellular compartments complicates the mapping between microscopic and macroscopic parameters. Furthermore, the reactivity of these species also changes depending upon whether they are in the membrane or cytoplasm.

To work around this issue, we estimated effective rate constants by fitting chemical rate equations to results from particle-based simulations. The simulations contained reversible second-order association reactions that accounted for mass exchange between the cytoplasm and membrane in a purely 2D system. For each bimolecular reaction, we started with previously published rate constants from RDE models for polarity establishment (Howell et al., 2012; Woods et al., 2016; Wu et al., 2015), used the  $\lambda - \bar{q}$  formalism to estimate  $\lambda$ 's, performed corresponding particle simulations, then fit chemical rate equations to compute the macroscopic rate constants. For purely 2D simulations, significant changes to the published parameter values were made to facilitate polarization for benchmarking purposes. For whole cell, quasi-3D simulations, parameters were held close to published values with exceptions for the bimolecular reactions obtained from the fitting procedure. The fits for the purely 2D and quasi-3D cases are shown (Figure 1.3, Figure 1.12, Figure 1.13), and the resulting rate constants are provided in Table 1.2

and Table 1.4. Fitting the simulation results to appropriate chemical rate equations produced good estimates for the quasi-3D case and reasonable ones for the purely 2D case. Additional analyses of the polarity network, discussed below, further supported the validity of the mapping.



**Figure 1.3. Empirical estimates for macroscopic rate constants in the yeast polarity model for the two different parameter sets.**

Results from particle-based simulations that include membrane exchange are shown as yellow curves. Fits to the simulation results using appropriate rate equations are shown as black curves. Top row, parameters used for purely 2D simulation. Bottom row, parameters used for quasi-3D simulation.

Usage	Reaction	Membrane on/off rates (1/s)				Fitted bimolecular association rates	
		$k_{A,on}$	$k_{A,off}$	$k_{B,on}$	$k_{B,off}$	$k_f$	$k_r$
2D	Cdc42Dm + BemGEFm → Cdc42T	36	13	10	40	$0.040 \mu\text{m}^2\text{s}^{-1}$	-
	Cdc42Dm + BemGEF42 → Cdc42T	36	13	-	-	$0.184 \mu\text{m}^2\text{s}^{-1}$	-
	BemGEFm + Cdc42T ↔ BemGEF42	10	40	-	-	$0.054 \mu\text{m}^2\text{s}^{-1}$	$31.4 \text{ s}^{-1}$
q3D	Cdc42Dm + BemGEFm → Cdc42T	36	0.65	10	10	$0.16 \mu\text{M}^{-1}\text{s}^{-1}$	-
	Cdc42Dm + BemGEF42 → Cdc42T	36	0.65	-	-	$0.16 \mu\text{M}^{-1}\text{s}^{-1}$	-
	BemGEFm + Cdc42T ↔ BemGEF42	10	10	-	-	$0.79 \mu\text{M}^{-1}\text{s}^{-1}$	$0.37 \text{ s}^{-1}$

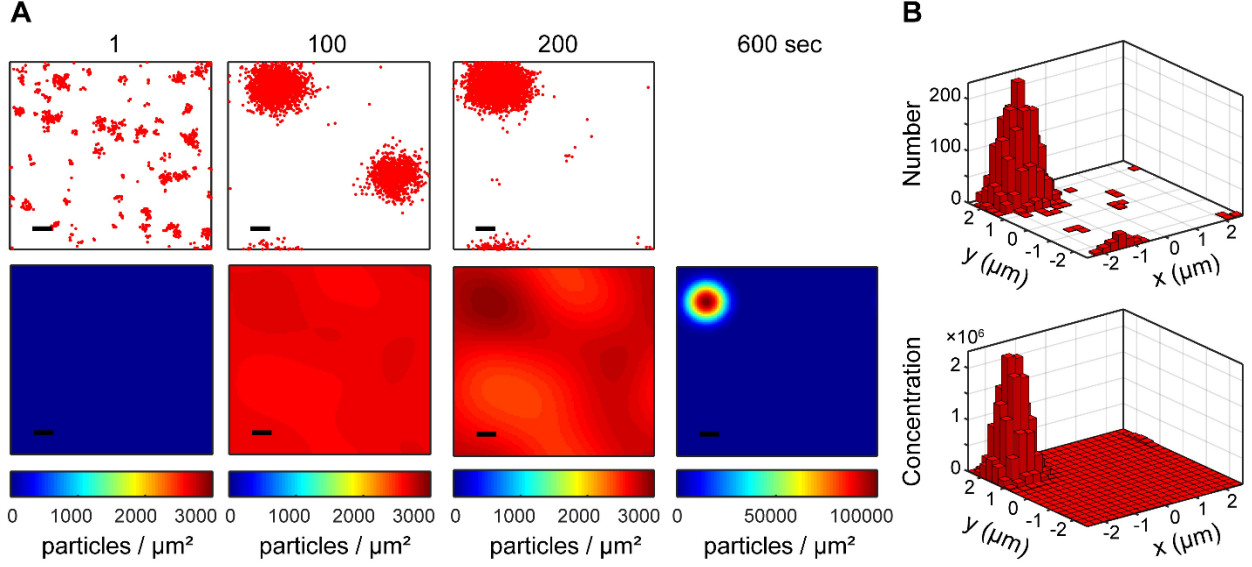
**Table 1.2. Microscopic parameters and effective macroscopic rate constants for reversible/irreversible bimolecular reactions of the form  $A + B \leftrightarrow C$ .**

The rate constants  $k_{A,on}$  and  $k_{B,on}$  are the rates at which reactants A and B associate with the membrane, and the rate constants  $k_{A,off}$  and  $k_{B,off}$  are the corresponding dissociation rates. The reported fitted rates are the averaged results from fitting to simulations with five separate initial conditions. See Table 1.4 for more details.

## 1.4. Microscopic fluctuations promote polarity establishment

### 1.4.1. Enhanced polarization within 2D simulations

Stochastic particle-based simulations were first performed in a purely 2D system (see 1.6.2 Particle-based simulations in 2D). The particle-based simulations were initialized with all molecules inactive, in the cytoplasm, and unbound. As expected, stochastic fluctuations permitted escape from this spatially homogeneous initial state, ultimately leading to polarization (Figure 1.4). To fairly compare particle simulation results with solutions of the RDEs, molecular distributions from particle-based simulations at  $t = 1$  second were used as initial conditions for the RDEs. The two simulation methods generated similar polarized distributions.



**Figure 1.4. Simulations of polarity establishment within the Turing unstable regime.**

Snapshots of total Cdc42-GTP (both Cdc42-GTP and Bem1-GEF-Cdc42-GTP). Top: Particle-based simulations. Red dots represent individual molecules. Bottom: Reaction-diffusion partial differential equation simulations. (A) Individual molecules in particle-based simulations, and individual pixels in 100x100 grid RDE simulations. Scale bar, 0.5  $\mu\text{m}$ . (B) To compare the polarity patches, 2D histograms of the final polarized states were computed, where both distributions were binned on coarsened 20x20 grids.

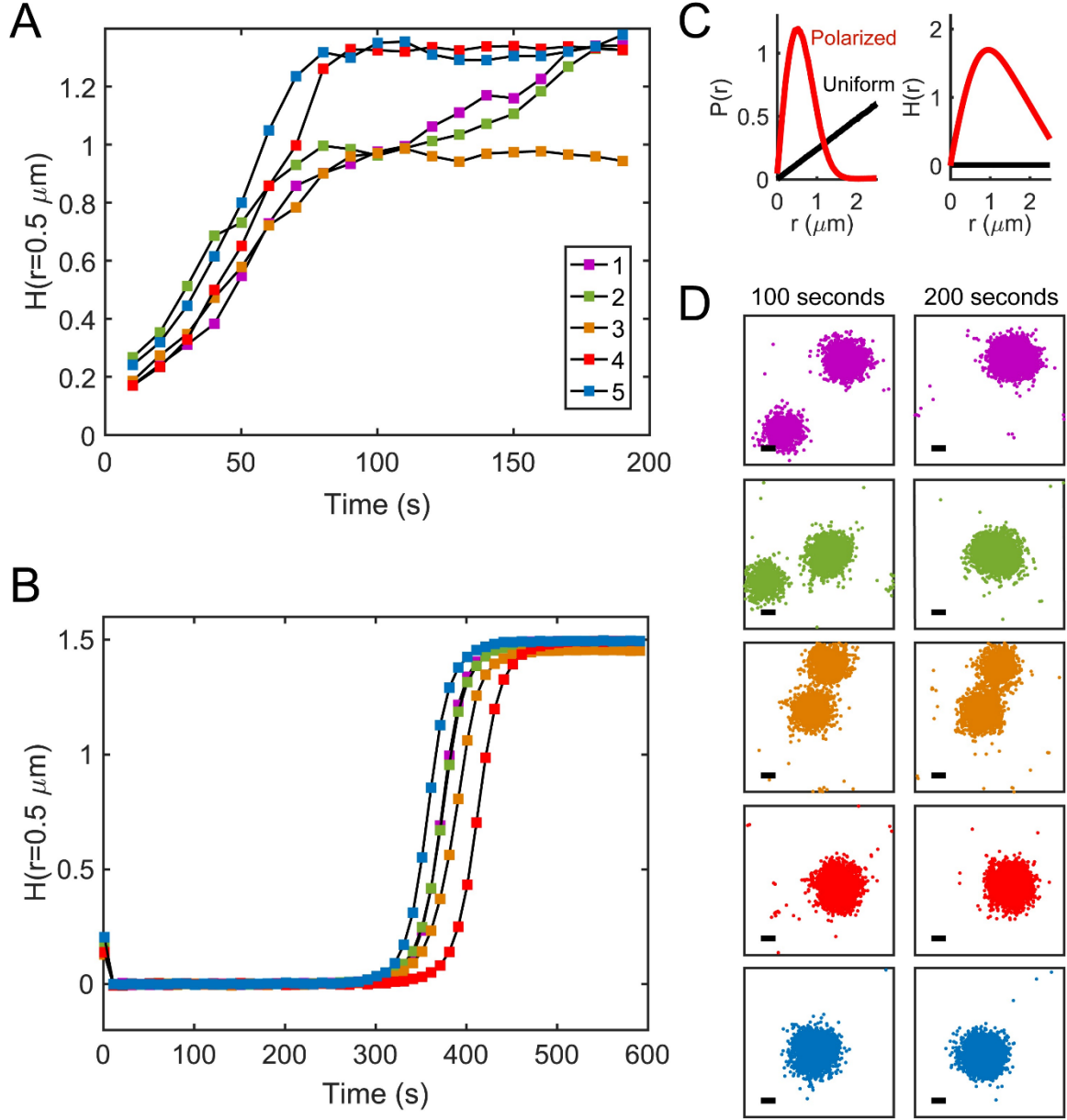
Quantitative comparison of polarization was based on the function  $H(r)$ , which measures the deviation of a particle distribution from a uniform distribution based on the pairwise distance distribution (1.7.1 Quantifying polarization).  $H(r)$  and the related metric, Ripley's  $K$ -function, have been used frequently to study clustering in biology (Owen et al., 2010; Wehrens et al., 2014). Positive values of  $H(r)$  correspond to increased particle density at distances  $r$ , and a maximum in  $H(r)$  denotes a characteristic size. Rather than choosing the  $r$  that maximizes  $H(r)$  under different conditions,  $r = 0.5 \mu\text{m}$  was used for all analyses. This value allowed comparisons across all data sets, including those where the simulation domain size was varied. Qualitative features of the results do not depend on the choice of  $r$ , nor on the particle-based time point used to initialize the RDEs (Figure 1.14 and Figure 1.16).

At steady-state, polarized distributions from the particle-based and RDE simulations had similar  $H(r)$  curves (Figure 1.15), suggesting the two systems were equivalently parameterized.  $H(r)$  over time was also calculated to quantitatively compare the polarization dynamics for particle-based and RDE simulations. Examining multiple realizations of single simulation conditions (Figure 1.5) revealed that

metastable multi-polar states could emerge from initially unpolarized distributions, consistent with prior theoretical and experimental work (Goryachev and Pokhilko, 2008; Howell et al., 2009, 2012). In one realization, resolution into a single polarity site did not occur by 200 seconds (Figure 1.5, Simulation 3). For other realizations of the same parameter set, the simulation yielded a unique polarity site in half the time. The particle-based simulations polarized more rapidly than the RDEs, which were completely unpolarized at  $t = 200$ s. This indicates that molecular fluctuations increased the rate at which polarity establishment occurred. Additionally, the RDEs did not exhibit transient plateaus in  $H(r)$ , indicating metastable multi-patch states did not emerge.

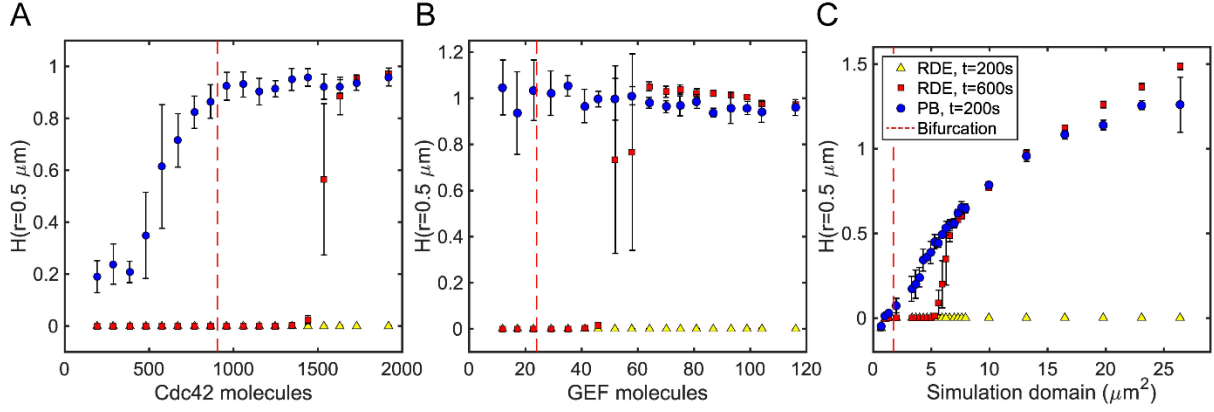
Sufficiently strong fluctuations can cause polarization outside of the Turing unstable regime (Dziekan et al., 2012; Goryachev and Leda, 2017; Lemarchand and Nowakowski, 2011; Trong et al., 2014). These investigations relied on simplified models or phenomenological methods for introducing noise into the system. To test if intrinsic fluctuations are sufficient to produce “noise-induced” polarity outside of the Turing unstable regime, we examined 2D polarity establishment as a function of Cdc42 concentration, Bem1-Cdc24 (interchangeably, BemGEF or GEF in this chapter) concentration, and total particle number at fixed concentration, generating bifurcation diagrams for these parameters. Linear stability analysis of the RDEs was used to determine the bifurcation point at which the spatially homogenous solution goes through a Turing instability as molecular abundances and system size were varied (see 1.7.2 Determining bifurcation points and Figure 1.17). This analysis established threshold values at which the RDEs no longer polarize, i.e. the homogeneous stable regime. The bifurcation plots are shown in Figure 1.6.





**Figure 1.5. Variability in 2D polarization from microscopic fluctuations.**

(A) Measurements of  $H(r = 0.5 \mu\text{m})$  at 10 second intervals across  $n = 5$  particle-based simulation realizations. (B) Measurements of  $H(r = 0.5 \mu\text{m})$  across the corresponding RDE simulations. (C) The pairwise distance distribution  $P(r)$  and our polarity metric  $H(r)$  for polarized (red) and uniform (black) particle distributions. (D) Snapshots of total Cdc42-GTP for each particle-based realization at  $t = 100$  and  $t = 200$  seconds. In some cases, particle coordinates were re-centered after simulation to keep polarity patches from visually wrapping around to the other side of the periodic domain. Scale bars  $0.5 \mu\text{m}$ .



**Figure 1.6. Stochasticity facilitates polarization.**

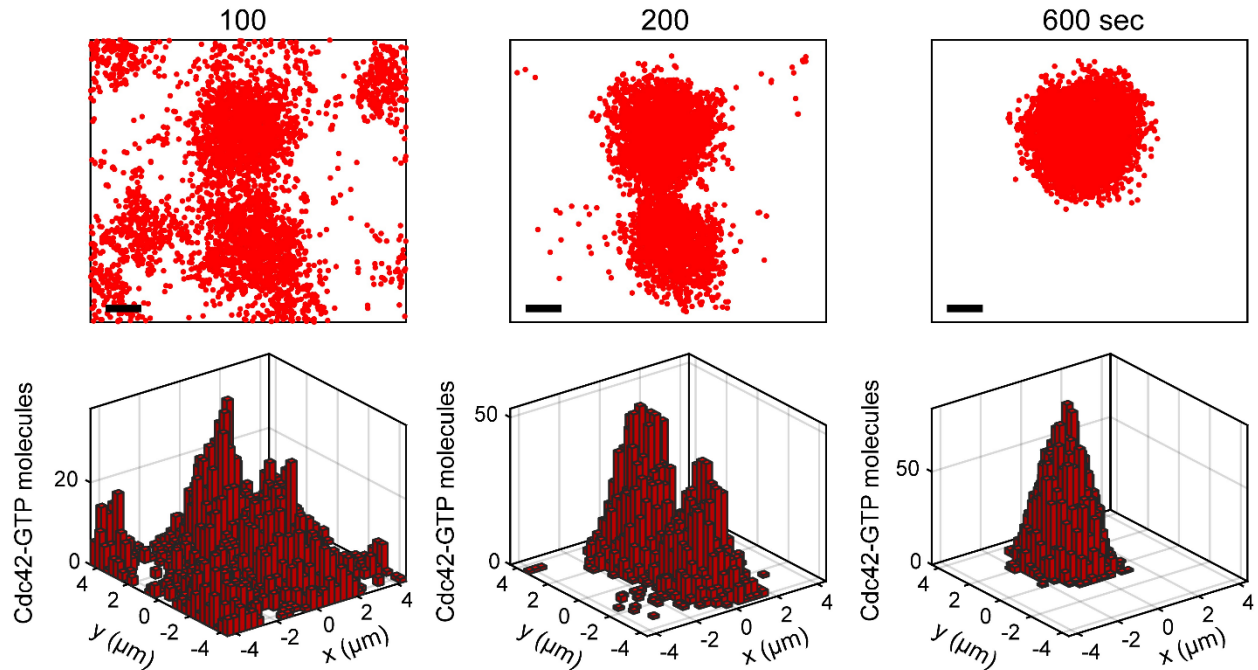
Bifurcation plots showing polarization, measured by  $H(r = 0.5 \mu\text{m})$ , versus parameters influencing the total particle numbers in the simulation. (A) Varying Cdc42 concentration on a fixed domain. (B) Varying GEF concentration on a fixed domain. (C) Varying the simulation area and particle numbers at constant concentrations. Bifurcations were found via linear stability analysis of the deterministic RDEs.

Across all parameters tested, none of the RDE simulations polarized to a measurable degree after 200 seconds. In contrast, most particle-based simulations exhibited polarity by then. Within the Turing unstable regime, the RDE simulations show similar levels of polarization around 600 seconds compared to the particle-based simulations. However, near the bifurcation point within the Turing unstable regime, the RDEs did not polarize even after 600s, consistent with the slowed patterning expected from bifurcation theory. In this parameter regime, the particle-based simulations still clearly exhibited polarity within 200 seconds. Furthermore, for the Cdc42 and GEF bifurcation diagrams, the particle-based simulations showed polarization below the critical point, in the Turing stable regime, showing that molecular fluctuations can increase the range over which polarity establishment occurs. Together, our observations reveal that stochastic effects facilitate polarization in this 2D instance of the Turing-type model by decreasing time to polarize and expanding the parameter space in which polarity can occur.

#### 1.4.2. Enhanced polarization within quasi-3D simulations

Quasi-3D (q3D) simulations of a whole yeast cell were performed by combining the 2D particle-based approach with stochastic exchange to and from a molecular reservoir representing the bulk cytoplasm (for detail, see 1.6.3 Particle-based simulations in quasi-3D). Empirical estimation of rate

constants was again performed by fitting rate equations to the particle-based simulations. Simulations used 0.050 to 0.3  $\mu\text{M}$  Cdc42, and 0.06  $\mu\text{M}$  BemGEF ( $N_{\text{Cdc42}} = 1,970$  to 11,820 and  $N_{\text{BemGEF}} = 2364$  assuming a volume corresponding to a spherical cell with a 5  $\mu\text{m}$  diameter). Quantitative Western blotting experiments support 5,000–10,000 Cdc42 copies per cell, consistent with the simulated concentration range (Watson et al., 2014), while previous models assumed Cdc42 concentrations ranging from 19.3 nM (Klunder et al., 2013) to 5  $\mu\text{M}$  (Wu et al., 2015). Models specify BemGEF concentrations ranging from 0.017  $\mu\text{M}$  (Goryachev and Pokhilko, 2008; Wu et al., 2015) to 0.06  $\mu\text{M}$  (Woods et al., 2016). Since prior experimental work showed that multi-polar states can resolve within 2 minutes (Howell et al., 2009, 2012), particle-based simulations were initially limited to 200 seconds. This simulation time was insufficient for complete polarization, as multiple or misshapen patches were often observed (Figure 1.7). This is possibly because the initial conditions of the model conservatively place all molecules in the cytoplasm, in inactive states.

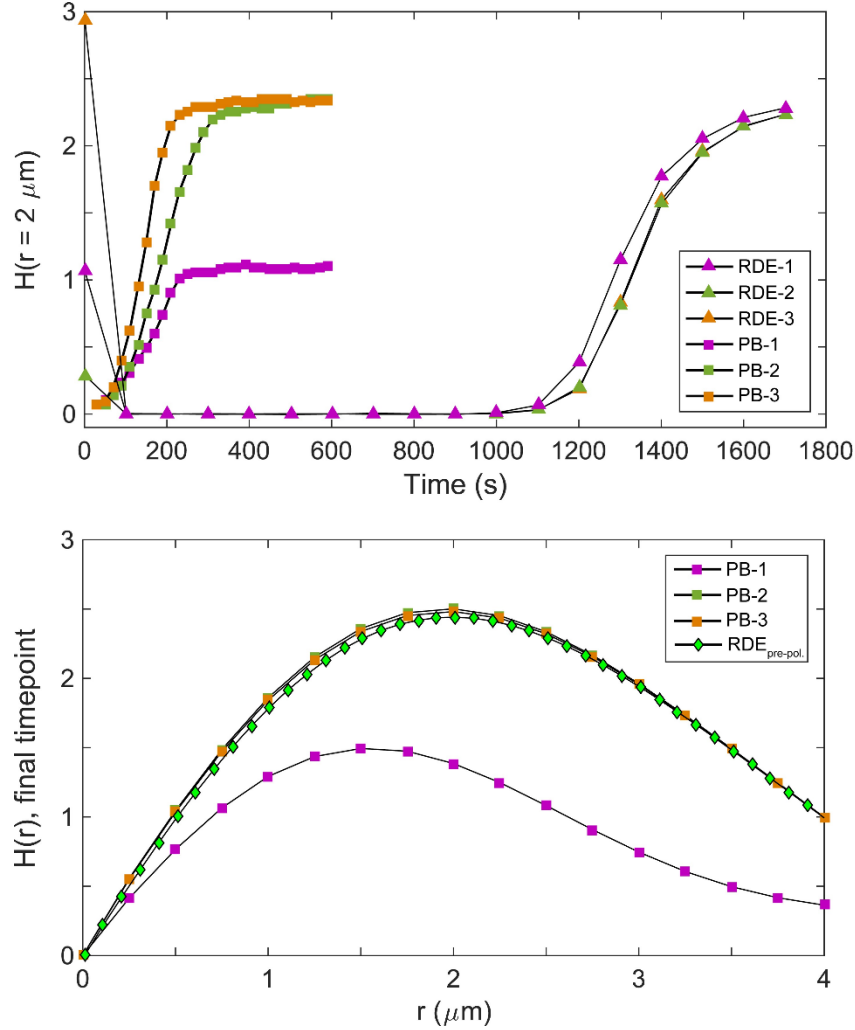


**Figure 1.7. Quasi-3D particle-based simulations of the polarity establishment model.**

Shown are snapshots of total Cdc42-GTP. Scale bar, 1.0  $\mu\text{m}$ . Corresponding 2D histograms of the local number of molecules are shown.

Extending the q3D particle-based simulations was computationally expensive, so to identify specific simulation conditions of interest, the bifurcation point as a function of Cdc42 concentration was determined. The reservoir equations made linear stability analysis problematic, so instead pre-polarized distributions were generated and allowed to either decay towards homogeneity or persist (Figure 1.18). Under these conditions,  $[Cdc42] \geq 0.055 \mu M$  was sufficient for polarization, but  $[Cdc42] = 0.050 \mu M$  could not sustain polarity, so particle-based simulations with  $[Cdc42] = 0.050, 0.055, 0.060, 0.150$ , and  $0.155 \mu M$  were extended for another 400 seconds. This simulation time was sufficient to tighten misshapen polarity sites and resolve metastable multi-polar states in some cases (Figure 1.7 and Simulations 2 and 3 in Figure 1.8). This suggests biologically relevant competition time scales can be obtained purely through stochastic molecular fluctuations. The time scale for competition observed here is consistent with prior theoretical work on this signaling model, where about 5 minutes was needed to resolve two-patch competition in the context of an RDE with Gaussian noise added (Wu et al., 2015).

To compare with the deterministic case, q3D RDE simulations were run for 1800s total, initialized with molecular distributions from  $t = 1$  s of the q3D particle-based simulations. Polarization dynamics were quantified using  $H(r = 2 \mu m)$ , which matched the size of a fully-formed polarity site. Similar to the purely 2D case, fully polarized particle-based simulations were quantitatively consistent with fully polarized RDE simulations, and the RDE simulations took much longer to polarize than the corresponding particle-based simulations (Figure 1.8). No multi-patch states emerged in the RDEs, but it is likely that multi-patch states would also compete slowly. This highlights the importance of molecular fluctuations in using a Turing-type model to capture appropriate polarization timescales.

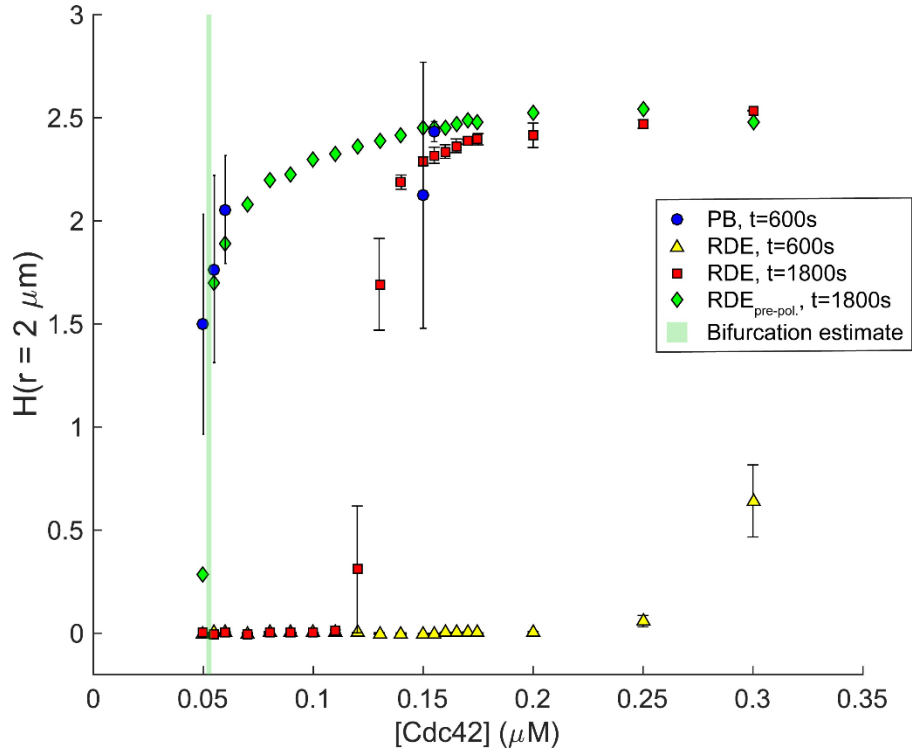


**Figure 1.8. Quantitative comparisons of polarization in quasi-3D particle-based simulations and corresponding RDEs.**

Top: time courses of  $H(r = 2 \mu\text{m})$ . Results across multiple realizations of  $[\text{Cdc42}] = 0.150 \mu\text{M}$  are shown. Bottom: Plots of  $H(r)$  at final time points. By 1800s, the q3D RDEs did not fully polarize, so the  $H(r)$  starting from a pre-polarized distribution is shown instead.

Finally, to examine the robustness of this behavior over realistic concentration regimes, polarization in the q3D particle-based and q3D RDE systems was compared as a function of Cdc42 concentration. The results were consistent with the purely 2D model. Particle-based simulations at  $t = 600\text{s}$  exhibited clear polarization, even at  $[\text{Cdc42}] = 0.050 \mu\text{M}$ , outside the deterministically non-polarizing region (Figure 1.9). At the highest concentration, q3D RDE simulations exhibited partial polarization at  $t = 600\text{s}$ , but by  $t = 1800\text{s}$ , most of the RDEs beyond the bifurcation exhibited measurable polarization.

The macroscopic system studied here represents a 3-compartment model (membrane, near-membrane, and bulk cytoplasm). A prior study of polarity competition in the presence of noise reported a similar competition time scale, but it utilized a volume-adjusted, two-compartment model of the RDEs (Wu et al., 2015). To facilitate comparison, we performed particle-based simulations to examine the volume-adjusted, two-compartment system's bifurcation diagram with respect to Cdc42 concentration. There is qualitatively no change in the results, and linear stability analysis of the volume-adjusted, two-compartment system is similar to the numerically determined bifurcation point for the q3D RDEs (Figure 1.23).



**Figure 1.9. The effect of Cdc42 concentration on polarization for quasi-3D particle-based simulations.**

A bifurcation diagram comparing polarity, measured via  $H(r = 2 \mu\text{m})$ , in the particle-based and reaction-diffusion simulations as a function of Cdc42 concentration. Simulations with pre-polarized RDEs were used to identify an estimated range for the bifurcation point. All other points are given by the mean  $\pm$  1 s.d. ( $n = 5$  realizations, except for  $t = 600\text{s}$  particle-based simulations at  $[\text{Cdc42}] = 0.150 \mu\text{M}$ ,  $n = 3$ , and  $0.155 \mu\text{M}$ ,  $n = 4$ ).

## 1.5. Discussion

Strong positive feedback to amplify heterogeneities in molecular distributions is an important component of many models of cellular polarity establishment. Given the stochastic nature of biochemical reactions involved in the polarity circuit, local heterogeneities are expected to emerge everywhere along the cell. Work in both non-Turing type (Altschuler et al., 2008; Jilkine et al., 2011; Schaff et al., 2016), and Turing-type systems (Dziekan et al., 2012; Goryachev and Leda, 2017; Lemarchand and Nowakowski, 2011; Trong et al., 2014) has shown that stochasticity can aid pattern formation. Here, we provide the first simulations of particle-based Turing-type yeast polarity establishment. Both our 2D and quasi-3D particle-based simulations capture microscopic stochastic effects, which indeed facilitate polarization. As anticipated, differences between the particle-based and reaction-diffusion approaches were most obvious around the bifurcation point (Figure 1.6 and Figure 1.9). Stochastic fluctuations allowed for polarization outside of the Turing unstable regime and more rapid polarity establishment across all parameters tested. Turing-type patterning mechanisms have been described as slow relative to other hypothesized patterning mechanisms, such as wave-pinning, making it a less likely biological mechanism in some contexts. Our simulations highlight that molecular fluctuations can alleviate such issues. Given our simulations do not include other sources of fluctuations, such as endocytic and exocytic events (Dyer et al., 2013; McClure et al., 2015), our results represent the minimal level of variability expected to be observed in polarity establishment. This minimal variability is sufficient to generate significant variations in competition times across multiple realizations of a single parameter set (Figure 1.5 and Figure 1.8), even at molecular abundances representative of whole yeast cells. Therefore, particle-based simulations are an important computational tool for understanding the dynamics and control of biological pattern formation.

Polarity establishment is often modeled using reaction-diffusion equations that ignore the discrete nature of biomolecules, and treat concentrations of molecular species as continuous variables. The chemical rate constants that appear in these equations represent macroscopic quantities that depend on microscopic properties, such as diffusion coefficients and molecular size. In three-dimensional domains

where particles diffuse with a single diffusion coefficient, theories for computing macroscopic rate constants from the underlying microscopic dynamics are well established (Andrews and Bray, 2004; Erban and Chapman, 2009). However, for two-dimensional systems, second-order rate constants in the diffusion limit are not well-defined (Yogurtcu and Johnson, 2015). Additionally, in the polarity system, molecular species transition between the cytoplasm, where diffusion is relatively fast, to the plasma membrane, where diffusion is relatively slow. Developing theories for computing appropriate rate constants under these conditions is an active area of research, and a theoretical framework is not provided here. An empirical approach was used instead, estimating effective second-order rate constants by fitting rate equations to the results of particle-based simulations of isolated reactions (Figure 1.3, Figure 1.12, and Figure 1.13). This approach allowed fair comparisons between our particle-based and RDE simulation simulations, as evidenced by quantitative similarities in polarization (Figure 1.8 and Figure 1.15) and equivalent kinetics under non-polarizing conditions (Figure 1.11 and Figure 1.21). Still, this empirical approach to estimating rate constants cannot capture the correct kinetics under all conditions: in general, a single rate constant is inappropriate for describing 2D diffusion-limited reactions (Yogurtcu and Johnson, 2015). While this discrepancy presents challenges for comparing particle-based simulations to RDEs, it also highlights an advantage of particle-based simulations: the real behavior of a system might not be well-described with macroscopic approximations. Note that many polarity models based on RDEs employ effective kinetics, such as Michaelis-Menten or Hill kinetics. To perform particle-based simulations of these models requires “unpacking” these effective kinetic schemes into their elementary chemical steps. Doing so not only allows an investigation into the effects of molecular-level noise, but also provides a rigorous test for the validity of the approximate reaction schemes, whose derivations typically rely on a separation of time scales.

The slow diffusivity of membrane-bound species is important. Other reports in the literature have used a faster membrane diffusion coefficient,  $D_m = 0.03 \mu\text{m}^2\text{s}^{-1}$  (Klunder et al., 2013). Simulations with this diffusion coefficient lose polarization in the pure 2D system if all other parameters are fixed (Figure



1.22). This occurs even though  $D_m = 0.03 \mu\text{m}^2\text{s}^{-1}$  maintains more than two orders of magnitude difference from the cytoplasmic diffusivity  $D_c$ .

It is important to acknowledge some limitations of the present approach. First, while treating the membrane and adjacent cytoplasm as a single 2D plane seems reasonable, it ignores effects from 3D curvature, which can play a role in the polarization process (Giese et al., 2015; Ramirez et al., 2015). Additionally, the implied geometry of the system, a rectangular prism, means that cytoplasmic protein abundances are overestimated near the cell membrane. To illustrate, a typical yeast cell has a diameter of  $5 \mu\text{m}$ . This corresponds to a surface area of  $78.5 \mu\text{m}^2$ , treating the cell as a sphere. Mapping this surface area to a square produces a square side length of  $8.86 \mu\text{m}$ . The volume for a spherical  $d = 5 \mu\text{m}$  yeast cell is  $26.2 \mu\text{m}^3$ . To achieve an equal-volume rectangular prism, with a top face surface area of  $78.5 \mu\text{m}^2$ , the depth of the prism must be  $0.833 \mu\text{m}$ . This is much smaller than the cell radius of  $2.5 \mu\text{m}$ . Aside from the geometry of the system, this approach also neglects gradients that might develop between the cell membrane and interior of the cell, either via chemical means or sufficiently slow cytoplasmic diffusion. The polarity network studied here does not involve reactions between two cytoplasmic species, so the reservoir component of the simulation is chemically inert. If reactions did occur within the cytoplasm, our particle-based approach could be extended to include chemical rate equations for the concentrations of the reservoir species, and our method for injection and ejection of particles would still be sufficient as long as cytoplasmic gradients were not of interest. However, if gradients of cytoplasmic components were required, then the reservoir would need to be modeled with PDEs, and the methods defining particle injection and ejection would need to be suitably adapted, along the lines of work done in (Franz et al., 2013). Full treatment of the reservoir with a PDE approach would make the approach presented here more similar to hybrid methods such as (Schaff et al., 2016). Having acknowledged these limitations, we have also performed these particle-based simulations on a spherical cell in true 3D and observed qualitatively similar features, as shown in Chapter 3.

In summary, we have shown that molecular stochasticity can facilitate cellular polarity establishment by promoting the speed of polarization and expanding the effectively Turing unstable regime. This

phenomenon was examined in the context of a Turing-type model of yeast signaling involving Cdc42 and Bem1-Cdc24 in a positive feedback loop. In particular, polarization within the quasi-3D system appears to occur roughly on biologically relevant timescales, which does not seem possible with deterministic RDEs. We also have highlighted general considerations for comparing the spatiotemporal dynamics of membrane-bound proteins at molecular, particle-based scales and at coarser, concentration-based scales. Symmetry breaking in many contexts involves guiding cues not considered here, such as a pheromone gradients or bud scars in yeast (Slaughter et al., 2009). However, these cues can be surprisingly weak: a computational study of yeast pheromone receptors in a pheromone gradient predicted differences in receptor occupancy as small as  $45 \pm 50$  molecules between the front (towards with the gradient) and the back (Lakhani and Elston, 2017). I visit how this may affect polarity establishment in Chapter 3.

## 1.6. Methods

### 1.6.1. The molecular circuit for polarity establishment

The molecular signaling network used in this study, illustrated in Figure 1.1C, is based on earlier work (Goryachev and Pokhilko, 2008; Wu et al., 2015). The network contains a positive feedback loop because Cdc42-GTP can bind a Bem1-Cdc42 complex to increase the GEF's catalytic activity. Cdc24 is a GEF, while Bem1 is a scaffold protein. It is assumed that Cdc24 and Bem1 function as essentially a single unit (Woods et al., 2016). Table 1.3 provides parameters used for simulations. The corresponding reaction-diffusion equations (RDEs) that govern the system are as follows:

$$\begin{aligned} \frac{\partial Cdc42T}{\partial t} = & (k_{2a}BemGEF_m + k_3BemGEF42) \cdot Cdc42D_m - (k_{2b} + k_{4a}BemGEF_m + k_7BemGEF_c) \\ & \cdot Cdc42T + k_{4b}BemGEF42 + D_m \Delta Cdc42T \end{aligned}$$

$$\begin{aligned} \frac{\partial Cdc42D_m}{\partial t} = & k_{2b}Cdc42T - (k_{2a}BemGEF_m + k_3BemGEF42 + k_{5b}) \cdot Cdc42D_m + k_{5a}Cdc42D_c \\ & + D_m \Delta Cdc42D_m \end{aligned}$$

$$\frac{\partial BemGEF42}{\partial t} = (k_{4a}BemGEF_m + k_7BemGEF_c) \cdot Cdc42T - k_{4b}BemGEF42 + D_m \Delta BemGEF42$$

$$\frac{\partial BemGEF_m}{\partial t} = k_{1a}BemGEF_c + k_{4b}BemGEF_{42} - (k_{1b} + k_{4a}Cdc42T)BemGEF_m + D_m\Delta BemGEF_m$$

$$\begin{aligned} \frac{\partial BemGEF_c}{\partial t} = & k_{1b}BemGEF_m - (k_{1a} + k_7Cdc42T) \cdot BemGEF_c + k_{inj}BemGEF_{c,res} - k_{ejc}BemGEF_c \\ & + D_c\Delta BemGEF_c \end{aligned}$$

$$\frac{dBemGEF_{c,res}}{dt} = \int_{\Omega \in \text{Reservoir}} [-k_{inj}BemGEF_{c,res} + k_{ejc}BemGEF_c] d\Omega$$

$$\frac{\partial Cdc42D_c}{\partial t} = k_{5b}Cdc42D_m - k_{5a}Cdc42D_c + k_{inj}Cdc42D_{c,res} - k_{ejc}Cdc42D_c + D_c\Delta Cdc42D_c$$

$$\frac{dCdc42D_{c,res}}{dt} = \int_{\Omega \in \text{Reservoir}} [-k_{inj}Cdc42D_{c,res} + k_{ejc}Cdc42D_c] d\Omega$$

$$k_{inj} = \int_0^{z_{max}-z_{impl}} P_{inj}(z) dz$$

$$k_{ejc} = \int_{z_{max}-z_{impl}}^{z_{max}} P_{ejc}(z) dz$$

where  $\Delta$  here denotes the two-dimensional Laplacian, and the terms containing the rates  $k_{inj}$  and  $k_{ejc}$  are for the cytoplasmic reservoir. In the purely 2D form of the RDEs, these reservoir terms are absent. The quasi-3D form directly follows schematics shown in Figure 1.1B and Figure 1.10, using a 2D membrane compartment, a 2D cytoplasmic compartment, and a reservoir to and from which mass is deterministically exchanged. The reservoir is assumed perfectly mixed, but the explicitly modeled cytoplasmic compartment is not. With this formulation, spatial gradients are possible in the  $xy$  plane (i.e. along the cell membrane), but they are ignored along  $z$  (i.e. moving into the cell). Previous work has considered cytoplasmic diffusion coefficients from  $1 \mu\text{m}^2/\text{s}$  up to infinity (i.e., perfectly well-mixed). A finite diffusion coefficient was chosen here. An RDE system that treats the membrane and cytoplasm as a two-compartment system was also considered, as in (Wu et al., 2015). Results from this system were qualitatively similar to our 3-compartment model (Figure 1.23).

Description	Parameter	For purely 2D simulations	For quasi-3D simulations	Reference
$\text{BemGEF}_c \rightarrow \text{BemGEF}_m$	$k_{1a}$	$10 \text{ s}^{-1}$	$10 \text{ s}^{-1}$	(Kuo et al., 2014)
$\text{BemGEF}_m \rightarrow \text{BemGEF}_c$	$k_{1b}$	$40 \text{ s}^{-1}$	$10 \text{ s}^{-1}$	(Kuo et al., 2014)
$\text{Cdc42D}_m + \text{BemGEF}_m \rightarrow \text{Cdc42T}$	$k_{2a}$	Target $0.032 \mu\text{m}^2\text{s}^{-1}$ Fitted $0.040 \mu\text{m}^2\text{s}^{-1}$	Target $0.16 \mu\text{M}^{-1}\text{s}^{-1}$ Fitted $0.16 \mu\text{M}^{-1}\text{s}^{-1}$	(Kuo et al., 2014)
	$\lambda_{2a}$	$5.3 \text{ s}^{-1}$	$5.3 \text{ s}^{-1}$	---
$\text{Cdc42T} \rightarrow \text{Cdc42D}_m$	$k_{2b}$	$0.35 \text{ s}^{-1}$	$0.32 \text{ s}^{-1}$	(Howell et al., 2012)
$\text{Cdc42D}_m + \text{BemGEF42} \rightarrow \text{Cdc42T}$	$k_3$	Target $0.280 \mu\text{m}^2\text{s}^{-1}$ Fitted $0.184 \mu\text{m}^2\text{s}^{-1}$	Target $0.35 \mu\text{M}^{-1}\text{s}^{-1}$ Fitted $0.16 \mu\text{M}^{-1}\text{s}^{-1}$	(Kuo et al., 2014)
	$\lambda_3$	$180 \text{ s}^{-1}$	$15.7 \text{ s}^{-1}$	---
$\text{BemGEF}_m + \text{Cdc42T} \rightarrow \text{BemGEF42}$	$k_{4a}$	Target $0.050 \mu\text{m}^2\text{s}^{-1}$ Fitted $0.054 \mu\text{m}^2\text{s}^{-1}$	Target $10 \mu\text{M}^{-1}\text{s}^{-1}$ Fitted $0.79 \mu\text{M}^{-1}\text{s}^{-1}$	(Kuo et al., 2014)
	$\lambda_{4a}$	$9.6 \text{ s}^{-1}$	$8250 \text{ s}^{-1}$	---
$\text{BemGEF42} \rightarrow \text{BemGEF}_m + \text{Cdc42T}$	$k_{4b}$	Target $40 \text{ s}^{-1}$ Fitted $31.4 \text{ s}^{-1}$	Target $10 \text{ s}^{-1}$ Fitted $0.37 \text{ s}^{-1}$	(Kuo et al., 2014)
$\text{Cdc42D}_c \rightarrow \text{Cdc42D}_m$	$k_{5a}$	$36 \text{ s}^{-1}$	$36 \text{ s}^{-1}$	(Kuo et al., 2014)
$\text{Cdc42D}_m \rightarrow \text{Cdc42D}_c$	$k_{5b}$	$13 \text{ s}^{-1}$	$0.65 \text{ s}^{-1}$	(Kuo et al., 2014)
$\text{BemGEF}_c + \text{Cdc42T} \rightarrow \text{BemGEF42}$	$k_7$	$2.0014 \mu\text{m}^2\text{s}^{-1}$	$10 \mu\text{M}^{-1}\text{s}^{-1}$	(Kuo et al., 2014)
	$\lambda_7$	$256 \text{ s}^{-1}$	$256 \text{ s}^{-1}$	---
Diffusion coefficient in cytoplasm	$D_{\text{cyto}}$	$15 \mu\text{m}^2\text{s}^{-1}$	$15 \mu\text{m}^2\text{s}^{-1}$	---
Diffusion coefficient on membrane	$D_{\text{memb}}$	$0.0025 \mu\text{m}^2\text{s}^{-1}$	$0.0025 \mu\text{m}^2\text{s}^{-1}$	(Kuo et al., 2014)
Membrane to cytoplasm volume ratio	$\eta$	1	0.01006	(Kuo et al., 2014)
Membrane surface area	A	$0.21\text{-}10.5\pi$	$25\pi$	(Kuo et al., 2014)
Molecular interaction radii	$\bar{\sigma}, \bar{\varrho}$	$0.050 \mu\text{m}$	$0.050 \mu\text{m}$	---
Total Cdc42		$14.5\text{-}145.5$ particles/ $\mu\text{m}^2$	$0.05\text{-}0.30 \mu\text{M}$	---
Total BemGEF		$0.87\text{-}8.8$ particles/ $\mu\text{m}^2$	$0.06 \mu\text{M}$	(Woods et al., 2016)

**Table 1.3. Parameters used to perform simulations described in the main text.**

References are for the quasi-3D parameters. “Target” and “fitted” values for  $k_{2a}$ ,  $k_3$ ,  $k_{4a}$ , and  $k_{4b}$  exist because of the empirical fitting described in the main text. The target was used as the input to 2D  $\lambda - \bar{\varrho}$  theory as the starting point, producing the corresponding microscopic rates  $\lambda_{2a}$ ,  $\lambda_3$ ,  $\lambda_{4a}$ , and  $\lambda_7$ . Fitted macroscopic rates were obtained after fitting as described earlier. Particle simulations used  $\Delta t = 0.1 \text{ ms}$ , and RDE simulations used  $\Delta t = 1 \text{ ms}$ . Quasi-3D simulations assumed a cell volume corresponding to a  $5 \mu\text{m}$  diameter sphere.

### 1.6.2. Particle-based simulations in 2D

We first considered a purely 2D computational domain representing molecules in the cell membrane and a thin volume of cytoplasm adjacent to the membrane. The membrane and cytoplasmic molecules were differentiated by their diffusivity and reactivity. The spatial coordinates of molecules  $x$  and  $y$  were treated as continuous variables, while time was discretized in intervals of  $\Delta t$ . Thermal diffusion was handled using the Euler-Maruyama method (Higham, 2001).

$$x(t + \Delta t) = x(t) + \xi_i \sqrt{2D\Delta t}$$

$$y(t + \Delta t) = y(t) + \xi_j \sqrt{2D\Delta t}$$

where  $\xi_i$  and  $\xi_j$  are normally distributed random numbers with mean zero and variance 1. Molecules that undergo association or dissociation events are not updated by the Euler-Maruyama method.

First-order or unimolecular reactions were assigned probabilities of occurring in  $\Delta t$  given by  $P_i = 1 - \exp(-k_i \Delta t)$ , where  $k_i$  was the rate constant for the  $i$ -th reaction. If the first-order reaction involved the dissociation of two molecules, then the two products were placed a distance of  $\bar{\sigma}$  apart, with one of the molecules located at the position of the complex, and the orientation angle chosen at random from a uniform distribution. For second-order or bimolecular reactions, it was assumed that two molecules react with probability  $P_\lambda = \lambda \Delta t$  when they are within a distance  $\bar{q}$ . Thus, if the two reactants are within a reactive range  $\bar{q}$ , they react with an average rate  $\lambda$ . This approach is based on the Doi method (Doi, 1976). It is distinct from the classic diffusion-limited Smoluchowski approach, where molecules react upon finding one another for the first time and molecular radii are adjusted to reach the desired kinetics (Smoluchowski, 1917). When two molecules bind, their positions are updated by moving one of the particles to the exactly same position as its binding partner. Periodic boundary conditions are assumed in both spatial directions, so that intermolecular distances for various calculations were solved as the minimum Euclidean distance along the 2D surface of a torus.

### 1.6.3. Particle-based simulations in quasi-3D

We next expanded our approach to approximate a whole cell by introducing a molecular reservoir the account for contributions from the bulk cytoplasm, yielding a quasi-3D approach (Figure 1.10). The cytoplasmic reservoir was treated implicitly, tracking only the number of molecules in the reservoir, instead of the dynamics of individual particles. To simulate stochastic exchange between the explicitly-modeled and implicitly-modeled regions of the cytoplasm, an approach similar to (Lakhani and Elston, 2017) was used, using diffusional probability distributions to determine the number of molecules injected into ( $n_{inj}$ ) and ejected from ( $n_{ejc}$ ) the explicitly-modeled cytoplasm at each time step. The resulting injection and ejection probabilities  $P_{inj}$  and  $P_{ejc}$  correspond to the probability that a single molecule at a depth  $z$  diffuses the distance required to enter ( $z_{impl}-z$ ) or exit ( $z-z_{impl}$ ) the explicit simulation region (see 1.7.5. Deriving integrals for quasi-3D injection and ejection).

$$P_{inj}(z) = \frac{1}{2} \left[ \operatorname{erf} \left( \frac{z_{max} - z}{\sqrt{4D\Delta t}} \right) - \operatorname{erf} \left( \frac{z_{impl} - z}{\sqrt{4D\Delta t}} \right) \right]$$

$$P_{ejc}(z) = \frac{1}{2} \left[ \operatorname{erf} \left( \frac{z}{\sqrt{4D\Delta t}} \right) - \operatorname{erf} \left( \frac{z - z_{impl}}{\sqrt{4D\Delta t}} \right) \right]$$

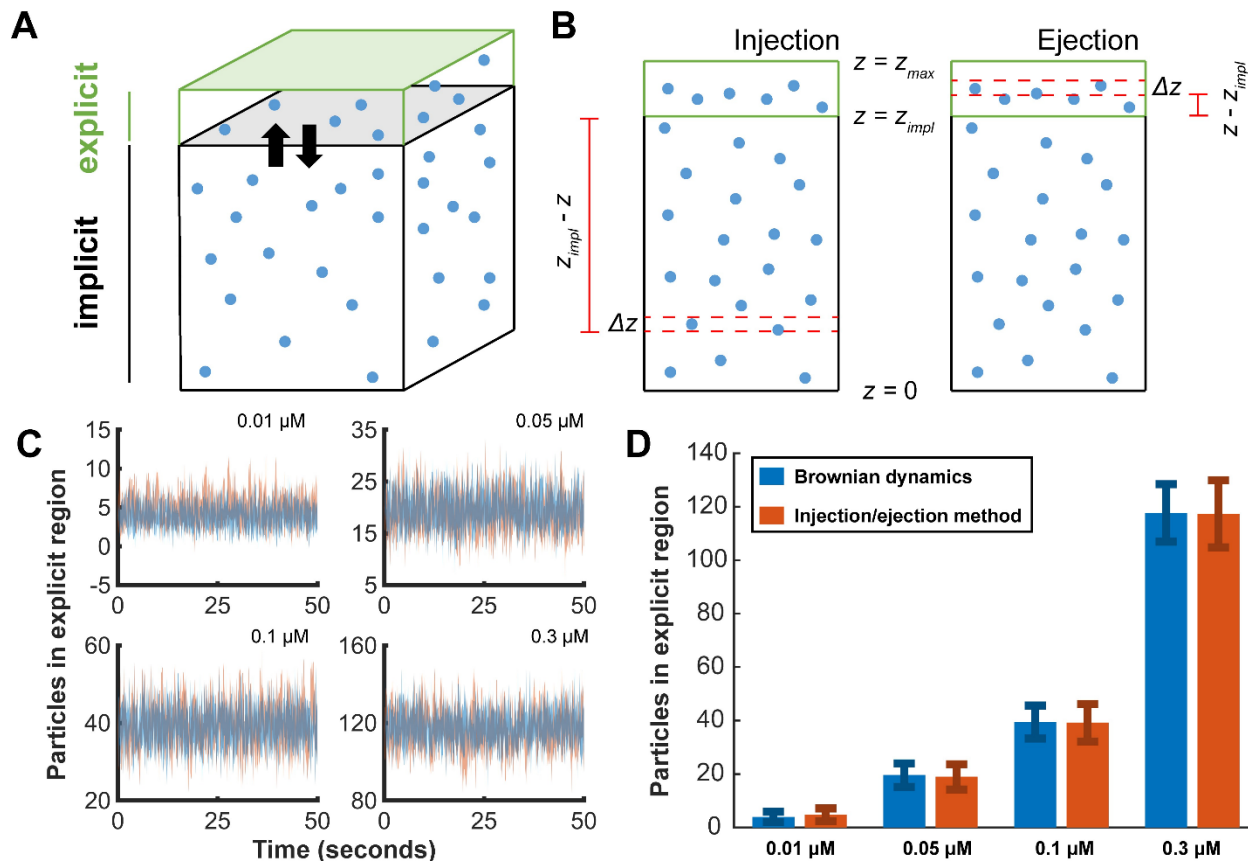
where  $z_{max}$  is the total height of the implicit and explicit domains, and  $z_{impl}$  is the height of the implicit domain.  $P_{inj}(z)$  and  $P_{ejc}(z)$  are approximations, since the probability densities in the derivation correspond to a freely diffusing particle on an infinite domain. Next, to calculate the mean number of particles that are injected and ejected, the injection and ejection probability densities were integrated over the appropriate domain, and multiplied by the 3D concentration ( $c$ ) and the surface area ( $A_r$ ).

$$\langle n_{inj} \rangle(t) = c_{impl}(t) \cdot A_r \cdot \int_0^{z_{max}-z_{impl}} P_{inj}(z) dz$$

$$\langle n_{ejc} \rangle(t) = c_{expl,cyto}(t) \cdot A_r \cdot \int_{z_{max}-z_{impl}}^{z_{max}} P_{ejc}(z) dz$$

Finally, to approximate the stochastic fluctuations introduced by particles diffusing in and out of the explicit simulation domain, Poisson distributions with means  $\langle n_{inj} \rangle$  and  $\langle n_{ejc} \rangle$  were sampled at each time step. Coupling this reservoir to the cytoplasmic layer of the 2D particle-based method yielded the quasi-

3D particle-based approach. Comparisons between this approximate method and Brownian dynamics simulations of diffusing particles showed that the molecular reservoir approach was consistent with both the mean and standard deviation for particle number over time (Figure 1.10).



**Figure 1.10. Reservoir approach schematics and validation.**

(A) Molecules can diffuse in and out of the reservoir. Although distinct molecules are shown for illustration, the reservoir is perfectly mixed. (B) Particles at a depth  $z$  must diffuse a distance of either  $z_{impl} - z$  to enter, or  $z - z_{impl}$  to exit, the explicit simulation domain. The integrals are solved numerically over discrete slices with thickness  $\Delta z$ . (C) Time courses of the number of molecules in the explicit domain, comparing our approach and a non-reactive Brownian dynamics simulation. The shaded regions represent the mean  $\pm 1$  S.D. over 5 realizations. (D) Time-averaged comparisons, mean  $\pm 1$  S.D. of fluctuations, over 500s, 1 realization.

Computing  $\langle n_{inj} \rangle(t)$  and  $\langle n_{ejc} \rangle(t)$  in practice involved numerically integrating over 100,000 discrete slices using the trapezoidal rule over the appropriate domains. Instead of re-calculating the integrals each simulation step, the integrals were pre-calculated and stored in a table to save computation time. This yields the mean behavior, which is sufficient for the q3D RDEs. However, to stochastically perform

injection and ejection in the particle-based simulations, draw uniform random numbers  $x \in [0,1]$  and use the inverse Poisson cumulative distribution function to obtain a Poisson-distributed particle number. To ensure physical validity, the tails of these probability distributions are cut off at the number of available particles in each compartment each time step.

#### 1.6.4. Reaction-diffusion partial differential equation simulations

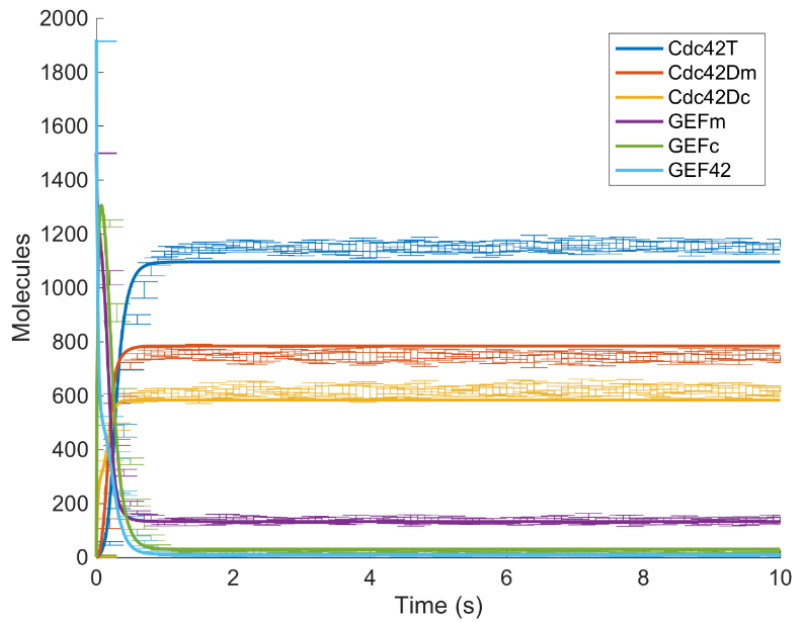
Initial conditions were generated by binning molecular distributions at  $t = 1$  sec from each realization of the particle-based simulation at each simulation condition. These pixellations were obtained on  $100 \times 100$  grids using MATLAB's `histcounts2` function, to be consistent with the grid size used for RDE simulation. Simulations were conducted using an operator splitting method, where reaction terms were solved using an adaptive Euler step, and diffusion terms were solved using a Fourier transform-based approach. As in the particle-based simulations, periodic boundary conditions were taken.

#### 1.6.5. Empirically mapping macroscopic to microscopic rate constants

A starting macroscopic rate constant was used to estimate an initial microscopic rate parameter  $\lambda$  via  $2D \lambda - \bar{q}$  theory (Table 1.3, “Target” values for bimolecular reactions). Then, particle-based simulations of the individual (ir)reversible bimolecular reactions were performed, allowing the reactants to undergo membrane-cytoplasm exchange as appropriate for each target reaction. Whether the parameter set was intended for 2D or quasi-3D simulations, these calibration simulations were performed on a 2D domain. For bimolecular reactions of the form  $A_m + B_m \leftrightarrow C$ , with membrane-cytoplasm exchange reactions  $A_c \leftrightarrow A_m$  and  $B_c \leftrightarrow B_m$ , time courses for particle numbers of  $A_m$ ,  $A_c$ ,  $B_m$ ,  $B_c$ , and  $C$  were extracted from each simulation. The membrane-cytoplasm exchange rates (specified during the particle-based simulation) were fixed during the fitting procedure, so that only  $k_f$  and  $k_r$  were fit. Fitting was done using MATLAB's built-in function `fminsearch`, where the sum of squared errors along normalized time courses (each scaled so that each species' maximum value in the time course was 1) was minimized. The rate constant  $k_7$  did not need to be fit, as it involves a cytoplasmic reactant and is not diffusion-limited. A height  $h = 0.0083$

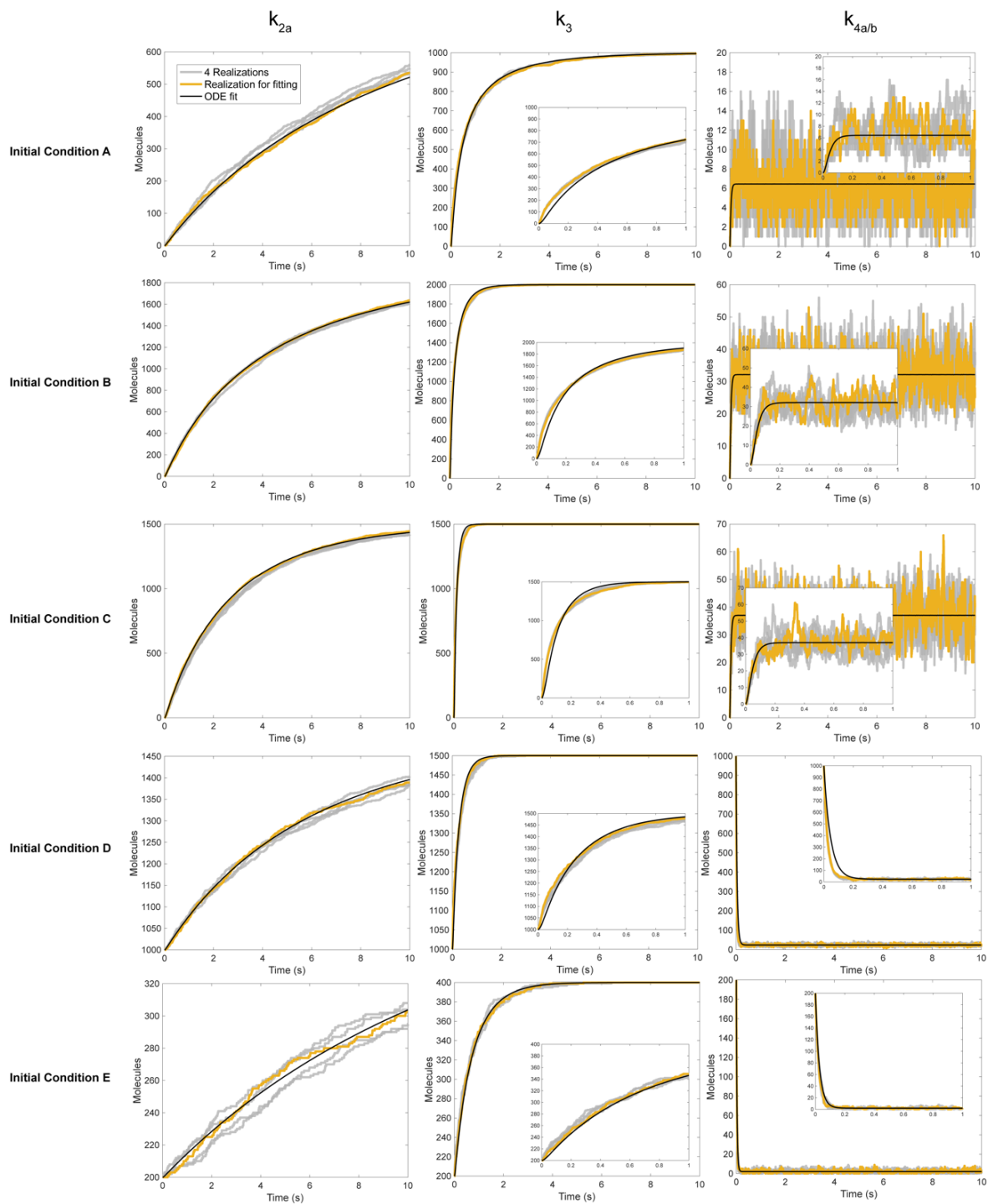


$\mu\text{m}$  for the cytoplasmic volume adjacent to the cell membrane was assumed, consistent with parameterizations of the membrane-to-cytoplasm volume ratio  $\eta$  used previously for this system (Wu et al., 2015). Scaling by  $h$  converts between 3D bimolecular rate constants  $k_{3D}$  ( $\mu\text{m}^3\text{s}^{-1}$ ) and 2D rate constants  $k_{2D}$  ( $\mu\text{m}^2\text{s}^{-1}$ ). The empirical mapping is a fair comparison between the particle-based and RDE systems based on the quantitative similarities between polarity sites in the two methods (Figure 1.8 and Figure 1.15), as well as the consistency of species time courses in a Turing-stable regime conditions (Figure 1.11 and Figure 1.21).



**Figure 1.11. Comparisons of deterministic rate equations in 2D and the polarity establishment network in a Turing stable regime.**

Linear stability analysis was used to identify a high  $[\text{Cdc42}]$  regime, consistent with parameters used in Figure 1.6. Here,  $N_{\text{Cdc42}} = 1921$  and  $N_{\text{GEF}} = 1500$ . All other parameters are the same as in the main text. Error bars are the mean  $\pm$  1 s.d. from  $n=5$  realizations of the particle-based simulation, solid lines are the solutions of the ordinary differential equations.



**Figure 1.12. Fitting simulations using 2D parameters to the deterministic rate equations.** Only the amount of product C formed is shown. Each ODE was fit to a single realization (gold). Additional realizations of each condition are provided to show the amount of intrinsic variability (gray).

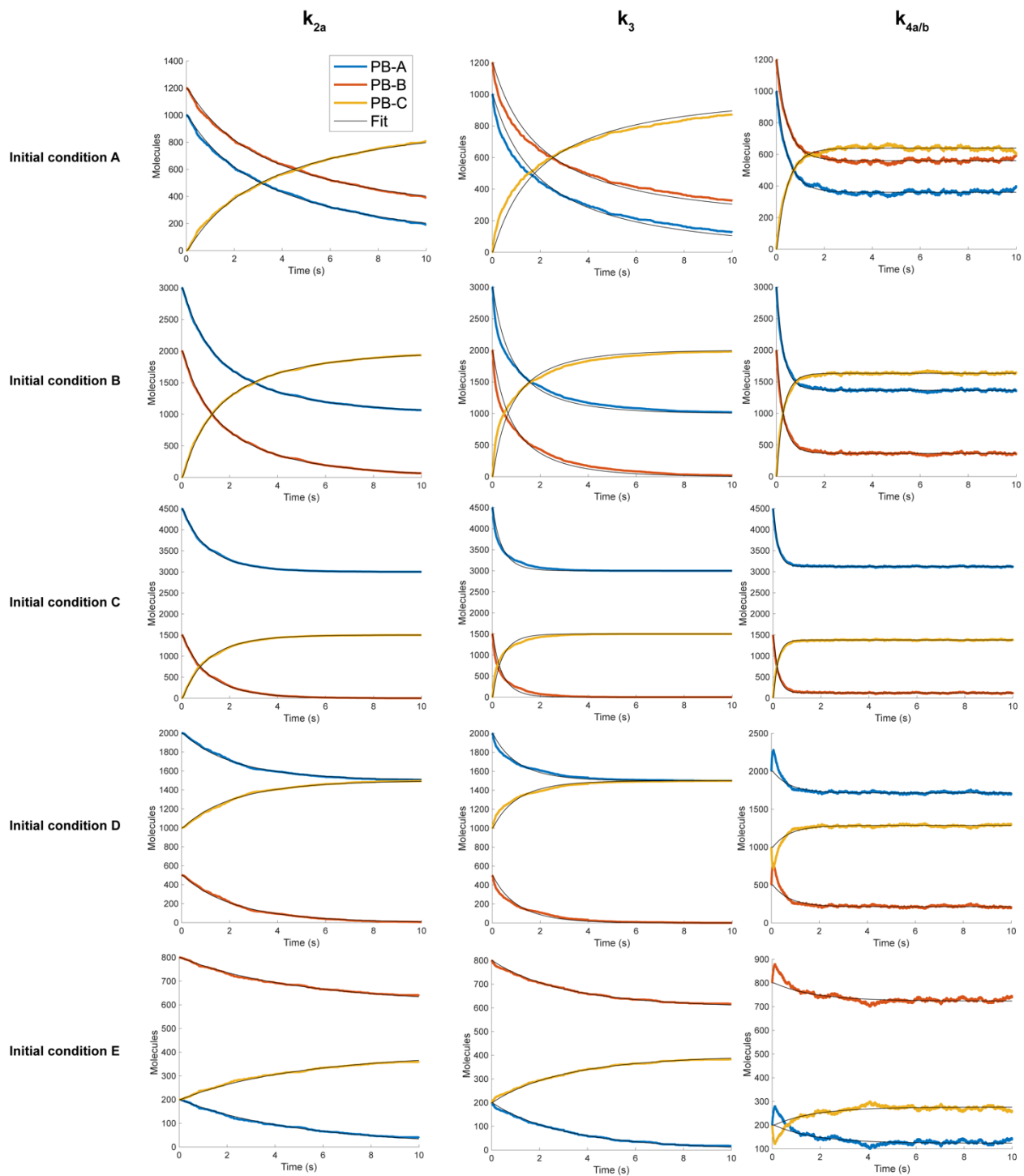


Figure 1.13. Fitting simulations using quasi-3D parameters to the deterministic rate equations.

	Reaction	IC	Membrane on/off rates				Microscopic parameters		Fitted	
			$k_{Aon}$ (1/s)	$k_{Aoff}$ (1/s)	$k_{Bon}$ (1/s)	$k_{Boff}$ (1/s)	$\lambda$ (1/s)	$k_r$ (1/s)	$k_f$ (2D: $\mu m^2/s$ ; 3D: $\mu M^{-1}/s$ )	$k_r$ (1/s)
Parameters for quasi-3D simulation	$Cdc42D_m + BemGEF_m \rightarrow Cdc42T$	A	36	0.65	10	10	5.30	-	0.1664	-
		B						-	0.1584	-
		C						-	0.1599	-
		D						-	0.1634	-
		E						-	0.1679	-
	$Cdc42D_m + BemGEF42 \rightarrow Cdc42T$	A	36	0.65	-	-	15.7	-	0.14444	-
		B						-	0.1484	-
		C						-	0.2054	-
		D						-	0.1669	-
		E						-	0.1374	-
	$BemGEF_m + Cdc42T \leftrightarrow BemGEF42$	A	10	10	-	-	8245	10	0.9943	0.492
		B							1.0043	0.480
		C							1.2541	0.546
		D							0.3997	0.179
		E							0.3098	0.157
Parameters for purely 2D simulation	$Cdc42D_m + BemGEF_m \rightarrow Cdc42T$	A	36	13	10	40	5.30	-	0.0367	-
		B						-	0.0392	-
		C						-	0.0411	-
		D						-	0.0398	-
		E						-	0.0432	-
	$Cdc42D_m + BemGEF42 \rightarrow Cdc42T$	A	36	13	-	-	178	-	0.162	-
		B						-	0.181	-
		C						-	0.220	-
		D						-	0.192	-
		E						-	0.166	-
	$BemGEF_m + Cdc42T \leftrightarrow BemGEF42$	A	10	40	-	-	9.60	40	0.0570	32.8
		B							0.0586	33.3
		C							0.0595	32.8
		D							0.0379	21.9
		E							0.0593	36.1

**Table 1.4. Microscopic parameters and effective macroscopic parameters for reversible/irreversible bimolecular reactions of the form  $A + B \leftrightarrow C$ , individual fits to each initial condition.**

$D_m = 0.0025 \mu m^2/s$  and  $D_c = 15 \mu m^2/s$  were used in the simulations. The domain area was  $64 \mu m^2$ . Fits to rate equations were performed to five separate particle-based simulations using different initial conditions (ICs). The fitted rates were then averaged. Initial conditions were as follows. **Set A:**  $A_0 = 1000$ ,  $B_0 = 1200$ ,  $C_0 = 0$ . **Set B:**  $A_0 = 3000$ ,  $B_0 = 2000$ ,  $C_0 = 0$ . **Set C:**  $A_0 = 4500$ ,  $B_0 = 1500$ ,  $C_0 = 0$ . **Set D:**  $A_0 = 2000$ ,  $B_0 = 500$ ,  $C_0 = 1000$ . **Set E:**  $A_0 = 200$ ,  $B_0 = 800$ ,  $C_0 = 200$ .

## 1.7. Supplemental Methods and Derivations

### 1.7.1. Quantifying polarization

Polarization was measured using the  $H$ -function  $H(r)$ , a rescaled version of Ripley's  $K$ -function. Ripley's  $K$ -function is a commonly used metric in experimental biology, and has been applied to study both experimental and simulated protein clustering (Kiskowski et al., 2009; Owen et al., 2010; Wehrens et al., 2014). The  $H(r)$  is related to the cumulative distribution of pairwise inter-particle distances  $P(r)$ :

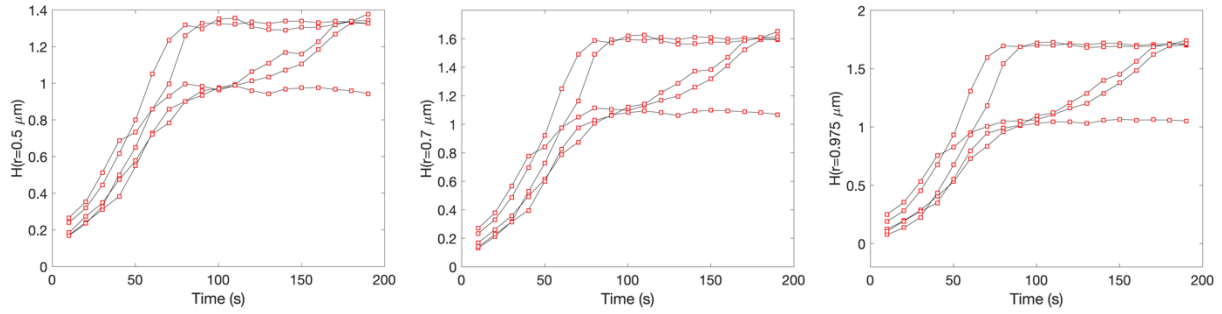
$$H(r) = \sqrt{\frac{L^2}{\pi} \int_0^r P(r') dr'} - r$$
$$P(r) = \frac{1}{N(N-1)} \sum_{i=1}^N m_i(r)$$

where  $N$  is the total number of particles subjected to cluster analysis,  $m_i(r)\Delta r$  is the number of particles at a distance  $d$  from particle  $i$  such that  $r-\Delta r/2 \leq d \leq r + \Delta r/2$ , and  $L$  is the length along either the  $x$ - or  $y$ -axis of the square simulation domain. Since the simulation domain boundaries are periodic,  $d$  is the minimum Euclidean distance along the two-dimensional surface of a torus.  $P(r)$  can also be related to the well-known pair correlation function  $g(r)$  by normalizing  $m_i(r)\Delta r$  by the expected density of particles within the area defined by  $r \pm \Delta r/2$ . See Figure 1.5C for an example.

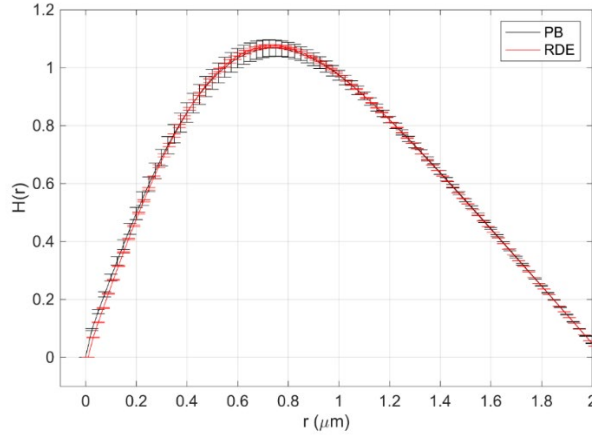
$H(r) = 0$  if particles are distributed according to a uniform distribution.  $H(r) > 0$  indicate the presence of more particles at inter-particle distances  $r$  as compared to a uniform distribution, while negative values of  $H(r)$  indicate fewer particles. Positive values of  $H(r)$  therefore indicate spatial clustering, or polarization, and the value of  $r$  for which  $H(r)$  is maximized reflects the characteristic cluster size. When "total Cdc42-GTP" levels were quantified or visualized, this meant both Cdc42-GTP and Bem1-GEF-Cdc42-GTP complexes, to reflect the total population of Cdc42-GTP molecules.  $H(r = 0.5 \mu\text{m})$  was typically used for comparisons. Qualitative features of our results do not depend on our choice of  $r$ , nor on the particle-based time point used to initialize the RDEs.

$H(r)$  was quantified in the RDE simulations by producing discrete particle distributions from continuous concentration data. A cumulative distribution function corresponding to the concentrations in

each spatially discretized bin was calculated, then pairs of random numbers  $p_1, p_2 \in \text{Unif}(0,1)$  were drawn to randomly select values from the cumulative distribution function, thus choosing two of the bins. Because the PDE assumes that molecules are uniformly distributed within each bin, four more uniform random numbers were drawn to pick locations within the two bins. This generated two sets of coordinates  $(x_1, y_1)$  and  $(x_2, y_2)$  such that  $(x_1, y_1)$  fits within the first bin, and  $(x_2, y_2)$  fits within the second. The pairwise distance between the points was recorded, and the process repeated many times to ensure sufficient sampling, typically  $n = 500,000$ . The resultant pairwise distance distribution is analogous to  $P(r)$ , permitting similar steps as above to compute a reaction-diffusion analogue of  $H(r)$ , facilitating direct quantitative comparison of polarization between the stochastic and deterministic simulations.

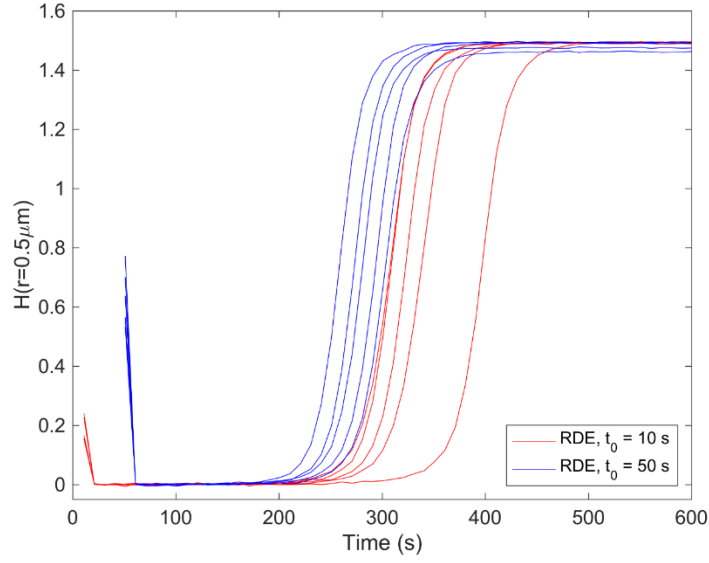


**Figure 1.14. Different choices of  $r$  in  $H(r)$  do not change the qualitative features of our results.** Figure 1.5 has been replicated here alongside versions with  $H(r = 0.7 \mu\text{m})$  and  $H(r = 0.975 \mu\text{m})$ .



**Figure 1.15. Representative  $H(r)$  curves for the particle-based and RDE simulations showing similarity in polarization.**

Quantification was performed at steady-state polarity ( $t=200\text{s}$  for particle-based simulation,  $t=600\text{s}$  for RDE simulation.) Uses the parameters in Figure 1.6A-B with  $N_{\text{Cdc42}} = 1921$  and  $N_{\text{GEF}} = 116$ .



**Figure 1.16. Polarity establishment is still slower for the RDE model if equations are seeded with later distributions from the particle-based simulations.**

The corresponding particle-based simulations are all polarized by 200s under these conditions, showing that the RDEs are consistently slow in comparison to the particle-based simulations.

#### 1.7.2. Determining bifurcation points

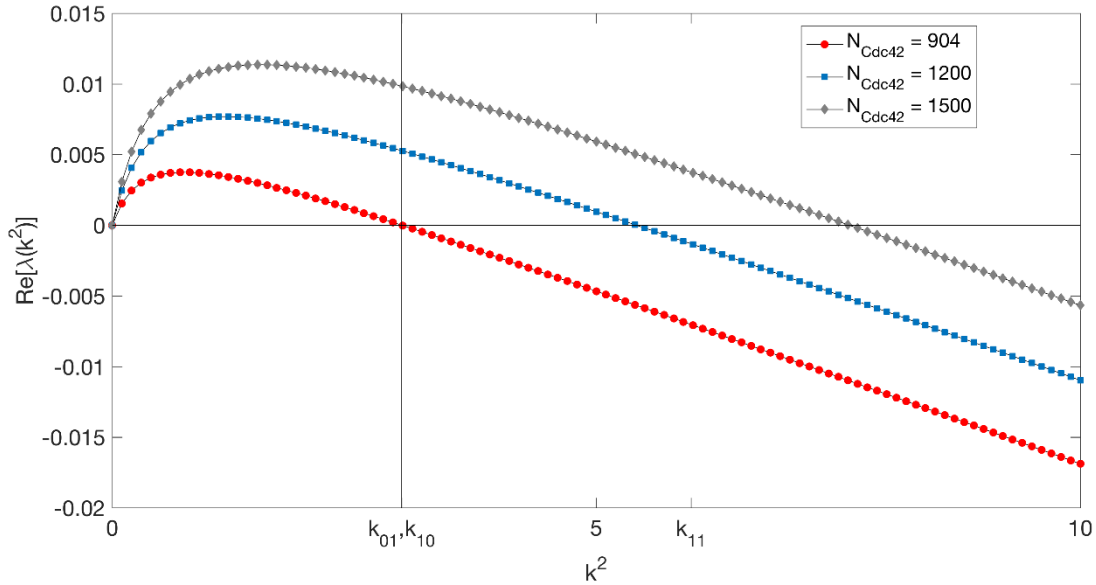
For the 2D reaction-diffusion system, linear stability analysis was used to determine conditions where the homogeneous steady state was Turing unstable (Murray, 2003). In this analysis, the full reaction diffusion equations are linearized around the homogenous steady-state, and the effect of an arbitrary small spatial perturbation is evaluated. The small perturbation is represented as a linear combination of a particular set of spatial functions, which are eigenfunctions (modes) of the Laplacian operator and are subject to appropriate boundary conditions. Because of the linearity, the initial growth of each mode is proportional to where the eigenvalue  $\lambda_n$  depends on the corresponding wave number  $k_n^2$ . A small spatial perturbation will grow if  $\text{Re}[\lambda]$  is greater than zero for at least one mode. For a square domain of side  $L$  with periodic boundary conditions, the eigenfunctions are (Haberman, 1998):

$$W_{nm}(x, y) = \left( a_n \cos \frac{2n\pi x}{L} + b_n \sin \frac{2n\pi x}{L} \right) \left( a_m \cos \frac{2m\pi y}{L} + b_m \sin \frac{2m\pi y}{L} \right)$$

where

$$k_{nm}^2 = \left( \frac{2n\pi}{L} \right)^2 + \left( \frac{2m\pi}{L} \right)^2$$

and  $n$  and  $m$  are integers. The coefficients  $a_n$ ,  $b_n$ ,  $a_m$ , and  $b_m$  are constants determined by the initial perturbation and are not relevant in this analysis. The dispersion relation  $\text{Re}[\lambda(k_{nm}^2)]$  is calculated for the set of PDEs that define the signaling network. The system is unstable if  $\text{Re}[\lambda(k_{nm}^2)] > 0$  for any mode  $n, m$ . If the system is in a Turing-unstable state, then decreasing the concentration of a particular species will induce a bifurcation where the system becomes stable to spatial perturbations. For the cases examined here (decreasing domain size or species concentration) the bifurcation occurs when  $\text{Re}[\lambda(k_{nm}^2)]$  becomes zero for the wave numbers  $k_{01}^2$  and  $k_{10}^2$  because these are the smallest relevant wave numbers (the mode  $n = 0, m = 0$  corresponds to a uniform function and is not relevant in this analysis).



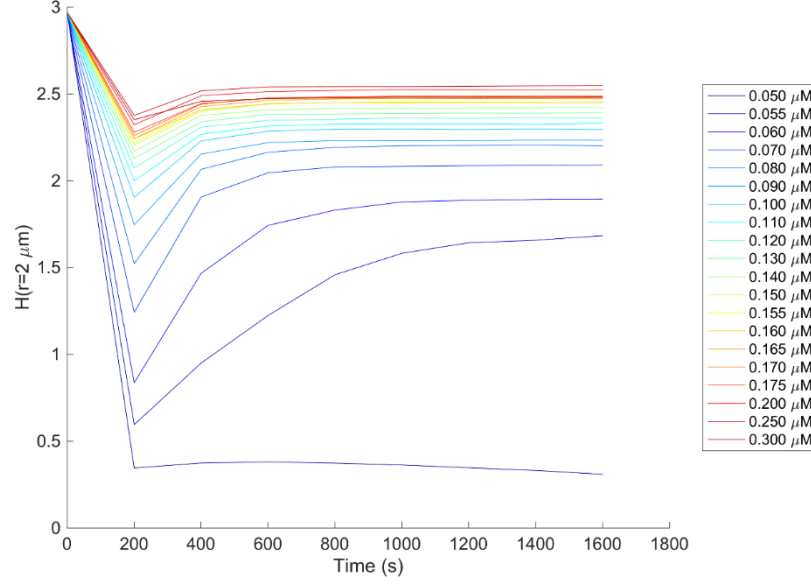
**Figure 1.17. Bifurcation point identification with linear stability analysis.**

The dispersion relation  $\text{Re}[\lambda(k^2)]$  is plotted against wave numbers  $k^2$ . The smallest number of Cdc42 molecules such that both  $\text{Re}[\lambda(k_{01}^2)] > 0$  and  $\text{Re}[\lambda(k_{10}^2)] > 0$  is the bifurcation point. This result corresponds to Figure 1.6A.

For the quasi-3D reaction-diffusion system, where linear stability analysis was more difficult, a simpler, numerical approach was used to determine the Cdc42 concentration threshold below which the system cannot polarize. The initial conditions were set up as follows: all the Cdc42 and BemGEF was set inactive in the “explicit” cytoplasmic compartment, except for a  $L/5 \times L/5$  square centered on the origin, where  $L$  is the domain length, in which 90% of the Cdc42 was converted into the active membrane-bound



state. Simulations to determine the bifurcation point by pre-polarization were carried out for 600s, which was sufficient to distinguish loss or maintenance of polarity on the basis of  $H(r = 2 \mu\text{m})$  (Figure 1.18).



**Figure 1.18. Quantifying  $H(r)$  at  $t=1800\text{s}$  for pre-polarized q3D-RDEs is a reasonable marker for maintenance vs. loss of polarity.**

### 1.7.3. Estimating the 2D diffusion limit

Consider a circular capture zone, centered at the origin of a circular domain. Let  $\bar{q}$  denote the radius of the capture zone, and  $r_{max}$  denote the radius of the domain. Then, introduce a species with diffusivity  $D$  with Brownian diffusion. This species, whose non-dimensionalized concentration is denoted by  $c(r)$ , is absorbed at  $\bar{q}$  and has a fixed concentration at  $r_{max}$ . The system is described at steady state by:

$$D \left( \frac{d^2 c}{dr^2} + \frac{1}{r} \frac{dc}{dr} \right) = 0, \quad \text{for } \bar{q} \leq r \leq r_{max}$$

with boundary conditions:

$$c(r) = 0, \quad \text{for } 0 \leq r \leq \bar{q}$$

$$c(r) = C, \quad \text{for } r = r_{max}$$

The general solution to the differential equation is:

$$c(r) = a_1 + a_2 \ln(r), \quad \text{for } \bar{q} \leq r \leq r_{max}$$

Using the boundary conditions:

$$a_1 = -\frac{C \ln(\bar{\varrho})}{\ln\left(\frac{r_{max}}{\bar{\varrho}}\right)} \quad \text{and} \quad a_2 = \frac{C}{\ln\left(\frac{r_{max}}{\bar{\varrho}}\right)}$$

$$c(r) = \frac{C(\ln(r) - \ln(\bar{\varrho}))}{\ln(r_{max}) - \ln(\bar{\varrho})}, \quad \text{for } \bar{\varrho} \leq r \leq r_{max}$$

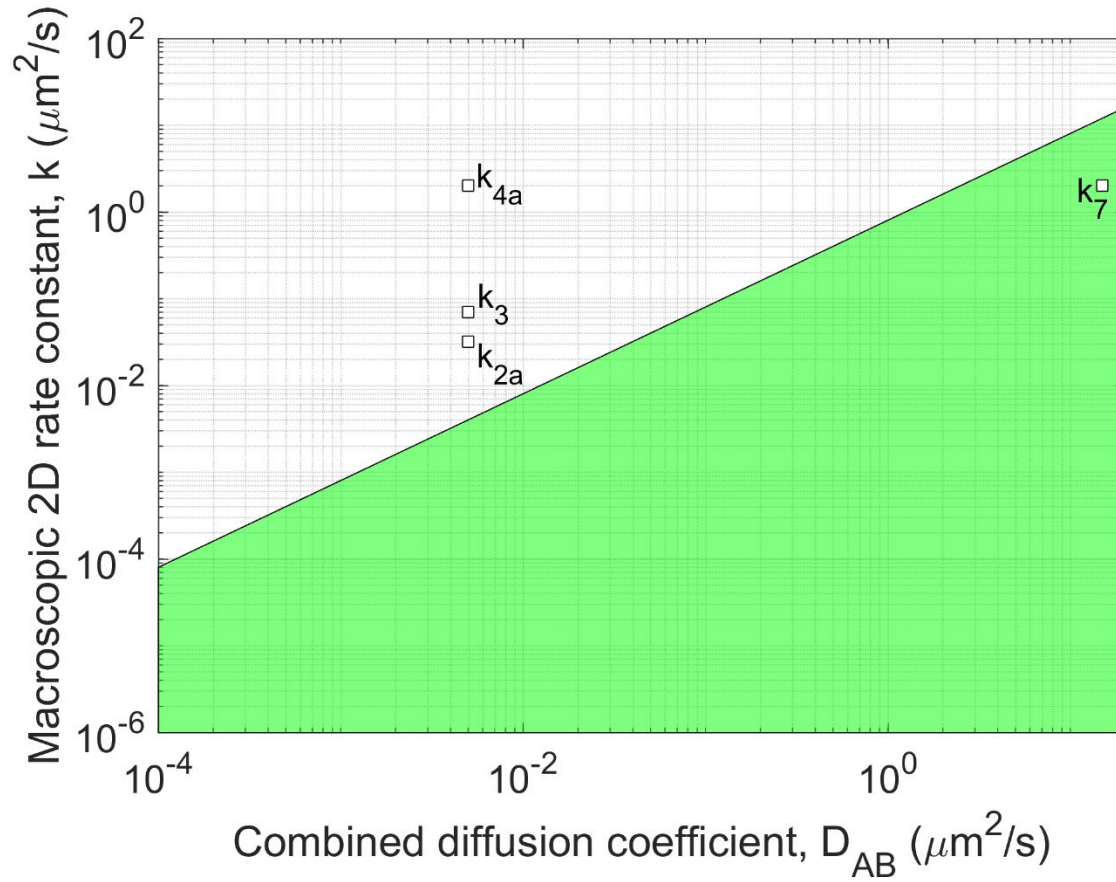
The capture rate is the total flux into  $r_B$ , which can be used as an estimate for the time scale of the diffusion-limited 2D reaction:

$$\text{Total Flux} = J = 2\pi D \bar{\varrho} \left. \frac{dc}{dr} \right|_{r=\bar{\varrho}} = \frac{2\pi D C}{\ln(r_{max}/\bar{\varrho})}$$

Thus, the estimated time scale for the diffusion-limited second-order rate constant is:

$$k_{DL} = \frac{2\pi D}{\ln(r_{max}/\bar{\varrho})}$$

This estimate can be used to assess whether the diffusion-limited conditions are relevant to published parameters for the yeast polarization network. Several of the rate constants appear to exceed the  $k_{DL}$  (Figure 1.19). However, this calculation assumes molecules can only find one another by diffusion along the plasma membrane, and does not take into account the effect of particle exchange between the membrane and cytoplasm, which can lead to significantly larger effective rate constants.



**Figure 1.19.** Accepted parameters for the yeast polarization model appear supra-diffusive in a 2D context.

The blank line is an estimate of the macroscopic rate constant using  $2\pi D / \ln(r_{\max}/\bar{\rho})$ . This neglects membrane-cytoplasm exchange. It is important to remember that, strictly, there is no well-defined diffusion-limited rate constant. Parameters:  $r_{\max} = 2.5 \mu\text{m}$ ,  $\bar{\rho} = 0.05 \mu\text{m}$ . Simulations were conducted on a  $L = 5 \mu\text{m}$  domain.

#### 1.7.4. Deriving and validating the 2D $\lambda - \bar{\rho}$ theory

The ratio between the unbinding and binding radii  $\alpha \equiv \bar{\sigma}/\bar{\rho}$  is crucial in defining the appropriate relation that describes the microscopic rate constant  $\lambda$  for the reversible bimolecular reaction  $A + B \rightleftharpoons C$ . In a manner similar to (Lipková et al., 2011), we begin by considering a coordinate system with a particle of B placed at the origin. In this frame of reference, the independent diffusion of A and B can be equivalently described as the random walk of a molecule A with diffusion coefficient  $D_A + D_B$ . When a molecule A diffuses within a distance  $\bar{\rho}$  of B, it is removed with a rate  $\lambda$ . The reverse dissociation reaction can be described with the introduction of new molecules of A at a distance  $\bar{\sigma}$  from the origin. In our 2D

system, we must work with 2D bimolecular rate constants – to convert a 3D rate constant to a 2D rate constant, we simply divide by the assumed depth of our explicit region (0.00833  $\mu\text{m}$ ). Let  $c(r)$  be the equilibrium concentration of molecules of A at distance  $r$  from the origin. We non-dimensionalize the problem by defining the reduced quantities:

$$\begin{aligned}\beta &\equiv \bar{\varrho} \sqrt{\frac{\lambda}{D_A + D_B}} & \hat{r} &\equiv \frac{r}{\bar{\varrho}} \\ \kappa &\equiv \frac{k_1}{(D_A + D_B)} & \hat{c} &\equiv \frac{c}{c_\infty}\end{aligned}$$

we have scaled lengths by  $\bar{\varrho}$ , time by  $\bar{\varrho}^2 (D_A + D_B)^{-1}$ , and concentrations by the bulk concentration:

$$\lim_{r \rightarrow \infty} c(r) = c_\infty$$

We consider the case where  $\alpha > 1$ , and ignore  $\alpha \leq 1$ , as the second case turns out to be unusable. If  $\alpha > 1$ , the dissociation radius  $\bar{\sigma}$  is greater than the binding radius  $\bar{\varrho}$ . The equations describing the microscopic reaction-diffusion system in polar coordinates are:

$$(D_A + D_B) \left( \frac{d^2 c}{dr^2} + \frac{1}{r} \frac{dc}{dr} \right) - \lambda c = 0, \quad \text{for } r \leq \bar{\varrho}$$

$$(D_A + D_B) \left( \frac{d^2 c}{dr^2} + \frac{1}{r} \frac{dc}{dr} \right) + Q(r - \bar{\sigma}) = 0, \quad \text{for } r \geq \bar{\varrho}$$

where  $Q(r - \bar{\sigma})$  is a Dirac-like distribution describing the production of molecules at  $r = \bar{\sigma}$ . In terms of the non-dimensional variables, we have:

$$\frac{d^2 \hat{c}}{d\hat{r}^2} + \frac{1}{\hat{r}} \frac{d\hat{c}}{d\hat{r}} - \beta^2 \hat{c} = 0, \quad \text{for } \hat{r} \leq 1 \tag{Equation 1.1}$$

$$\frac{d^2 \hat{c}}{d\hat{r}^2} + \frac{1}{\hat{r}} \frac{d\hat{c}}{d\hat{r}} + \omega \delta(\hat{r} - \alpha) = 0, \quad \text{for } \hat{r} \geq 1 \tag{Equation 1.2}$$

where  $\delta(\hat{r} - \alpha)$  is a Dirac-delta function, and  $\omega$  is the rate of creation of molecules at  $\hat{r} = \alpha$ . The average number of molecules produced by the forward and backward reactions must balance at equilibrium, and therefore we have:

$$2\pi\alpha\omega = \kappa = 2\pi \left. \frac{d\hat{c}}{d\hat{r}} \right|_{\hat{r}=1} \quad \text{Equation 1.3}$$

where  $\kappa$  represents the association rate constant. This  $\kappa$  is directly equal to the flux through the circle of radius 1, because it corresponds to a macroscopic bimolecular rate constant, in contrast to the synthesis term  $\omega$ . Substituting  $\omega = \kappa/(2\pi\alpha)$  into **Equation 1.2**, we obtain:

$$\frac{d^2\hat{c}}{d\hat{r}^2} + \frac{1}{\hat{r}} \frac{d\hat{c}}{d\hat{r}} + \frac{\kappa\delta(\hat{r} - \alpha)}{2\pi\alpha} = 0, \quad \text{for } \hat{r} \geq 1 \quad \text{Equation 1.4}$$

We can write down the general solutions of Equation 1.1 and Equation 1.4 in the form:

$$\hat{c}(\hat{r}) = a_1 I_0(\beta\hat{r}) + a_2 K_0(-\beta\hat{r}), \quad \text{for } \hat{r} \leq 1$$

$$\hat{c}(\hat{r}) = a_3 + a_4 \ln(\hat{r}) + \text{sgn}(\alpha - \hat{r}) \frac{\kappa}{4\pi} \ln\left(\frac{\hat{r}}{\alpha}\right), \quad \text{for } \hat{r} \geq 1$$

where  $a_1$ ,  $a_2$ ,  $a_3$ , and  $a_4$  are real constants to be determined,  $I_0(x)$  and  $K_0(x)$  are zeroth-order modified Bessel functions of the 1<sup>st</sup> and 2<sup>nd</sup> kind, respectively, and  $\text{sgn}(x)$  is the signum function. We specify the constants by using particular boundary conditions, as follows.

Going out to infinity, the concentration approaches the bulk concentration

$$\lim_{\hat{r} \rightarrow \infty} \hat{c}(\hat{r}) = 1$$

$\hat{c}(0)$  is a finite value

$$\hat{c}(0) = \hat{c}_0$$

The dimensionless concentration equations must agree at the boundary  $\hat{r} = 1$

$$\hat{c}_{\hat{r} \leq 1}(1) = \hat{c}_{\hat{r} \geq 1}(1)$$

To determine the integration constants, begin by using the first boundary condition.

$$\lim_{\hat{r} \rightarrow \infty} a_3 + a_4 \ln(\hat{r}) + \text{sgn}(\alpha - \hat{r}) \frac{\kappa}{4\pi} \ln\left(\frac{\hat{r}}{\alpha}\right) = 1$$

$$a_3 - 1 + \lim_{\hat{r} \rightarrow \infty} a_4 \ln(\hat{r}) - \frac{\kappa}{4\pi} [\ln \hat{r} - \ln \alpha] = 0$$

$$a_3 - 1 + \frac{\kappa}{4\pi} \ln \alpha + \lim_{\hat{r} \rightarrow \infty} a_4 \ln(\hat{r}) - \frac{\kappa}{4\pi} [\ln \hat{r}] = 0$$

$$a_3 - 1 + \frac{\kappa}{4\pi} \ln \alpha + \lim_{\hat{r} \rightarrow \infty} \ln(\hat{r}) \left[ a_4 - \frac{\kappa}{4\pi} \right] = 0$$

This requires that  $a_4 = \frac{\kappa}{4\pi}$ , which then makes it apparent that  $a_3 = 1 - \frac{\kappa}{4\pi} \ln(\alpha)$ . Next, note that at  $\hat{c}(0) = \hat{c}_0$ ,  $K_0(\beta\hat{r})$  is infinite, requiring that  $a_2$  must be zero. Finally, requiring that the two solutions are equal at  $\hat{r} = 1$  determines  $a_1$ :

$$a_1 I_0(\beta) = 1 + \frac{\kappa}{4\pi} \ln\left(\frac{\hat{r}}{\alpha}\right) [1 + \text{sgn}(\alpha - \hat{r})]$$

Since  $\hat{r} = 1$  and  $\hat{\alpha} > 1$ ,

$$a_1 I_0(\beta) = 1 + \frac{\kappa}{2\pi} \ln\left(\frac{1}{\alpha}\right)$$

$$a_1 = \frac{1 - \frac{\kappa}{2\pi} \ln(\alpha)}{I_0(\beta)}$$

Substitution of all constants into the equations for  $\hat{c}(\hat{r})$  yields:

$$\hat{c}(\hat{r}) = \frac{I_0(\beta\hat{r})}{I_0(\beta)} \left[ 1 - \frac{\kappa \ln(\alpha)}{2\pi} \right], \quad \text{for } \hat{r} \leq 1$$

$$\hat{c}(\hat{r}) = 1 + \frac{\kappa}{4\pi} \ln\left(\frac{\hat{r}}{\alpha}\right) [1 + \text{sgn}(\alpha - \hat{r})], \quad \text{for } \hat{r} \geq 1$$

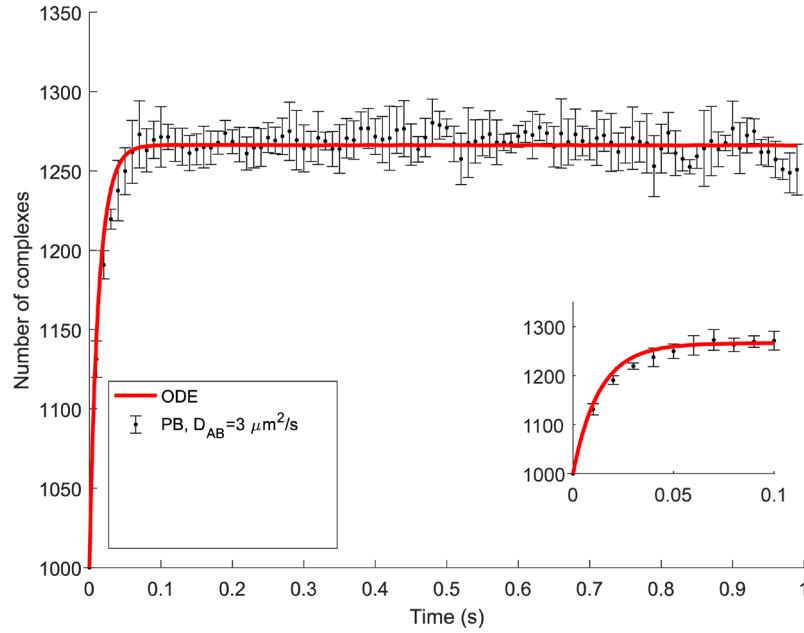
To get an expression for  $\kappa$  in terms of  $\beta$ , which contains our microscopic rate  $\lambda$ , we evaluate the derivative  $\hat{c}'(\hat{r})$  at  $\hat{r} = 1$  and set it equal to  $\kappa/2\pi$ .

$$\kappa = \frac{2\pi\beta \frac{I_1(\beta)}{I_0(\beta)}}{\left(1 + \ln(\alpha) \beta \frac{I_1(\beta)}{I_0(\beta)}\right)}, \quad \alpha > 1$$

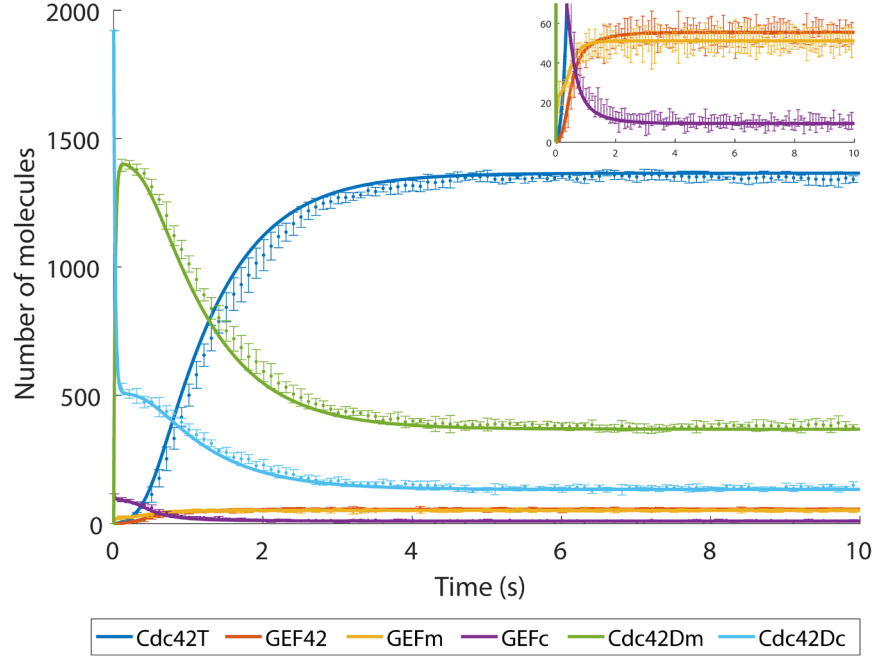
Note that, if we take the limit as  $\alpha \rightarrow 1$  and hold  $\beta$  constant,  $\kappa$  increases. This is consistent with the idea that, if the probability of reacting within  $\bar{\varrho}$  is kept the same, as we shrink the release radius, molecules of A will react with B more rapidly. Returning to experimentally measurable quantities  $k_I$ ,  $D_A$ ,  $D_B$ , and the model parameters  $\bar{\varrho}$ ,  $\bar{\sigma}$ , and  $\lambda$ :

$$k_1 = (D_{AB}) \frac{2\pi\bar{q}\sqrt{\frac{\lambda}{D_{AB}}} \left( \frac{I_1\left(\bar{q}\sqrt{\frac{\lambda}{D_{AB}}}\right)}{I_0\left(\bar{q}\sqrt{\frac{\lambda}{D_{AB}}}\right)} \right)}{\left( 1 + \ln\left(\frac{\bar{\sigma}}{\bar{q}}\right) \bar{q}\sqrt{\frac{\lambda}{D_{AB}}} \frac{I_1\left(\bar{q}\sqrt{\frac{\lambda}{D_{AB}}}\right)}{I_0\left(\bar{q}\sqrt{\frac{\lambda}{D_{AB}}}\right)} \right)}, \quad \alpha > 1$$

where  $D_{AB} \equiv D_A + D_B$ , and  $I_0(x)$  and  $I_1(x)$  are the zeroth- and first-order modified Bessel functions of the first kind. Although there is no analytic inverse of  $I_v(x)$ , we can numerically solve for the microscopic rate constant  $\lambda$ . To validate this calculation, we can compare particle-based simulations using the 2D  $\lambda - \bar{q}$  theory to deterministic rate equations based on macroscopic theory. We tested the simple reversible bimolecular association reaction  $A + B \rightleftharpoons C$  (Figure 1.20), as well as the more complicated Cdc42 signaling circuitry in a non-polarizing context (Figure 1.21).



**Figure 1.20. Comparisons between the deterministic rate equations in 2D and multiple realizations of the particle-based simulation for the simple reversible bimolecular reaction  $A+B \leftrightarrow C$ .** The error bars for the particle-based simulations represent the mean $\pm$ 1s.d.,  $n=5$ .



**Figure 1.21. Comparisons between the deterministic rate equations in 2D and multiple realizations of the particle-based simulation for the polarity establishment network under reaction-limited, non-polarizing conditions.**

The error bars for the particle-based simulations represent the mean  $\pm$  1 s.d.,  $n=5$ . The solid lines are solutions of the partial differential equations, using  $D_m = D_c = 1.5 \mu\text{m}^2/\text{s}$ .

#### 1.7.5. Deriving integrals for quasi-3D injection and ejection

We begin by considering the diffusional probability density defined by the diffusion equation in one dimension. We assume that diffusing particles are not confined – while this is not truly the case, we demonstrate that this assumption turns out to be reasonable.

$$\partial_t p(z, t | x_0, t_0) = D \partial_z^2 p(z, t | z_0, t_0)$$

$$p(z, t_0 | z_0, t_0) = \delta(z - z_0)$$

$$p(z \rightarrow \pm\infty, t | z_0, t_0) = 0$$

The corresponding Green's function solution is:

$$p(z, t | z_0, t_0) = \frac{1}{\sqrt{4\pi D(t - t_0)}} \exp \left[ -\frac{(z - z_0)^2}{4D(t - t_0)} \right] \quad \text{Equation 1.5}$$

If we consider diffusion distributions  $p_i$  emerging from Dirac-Delta distributions at  $z_i$  within the cytoplasm, then each probability distribution  $p_i$  satisfies **Equation 1.5** at each time step. We further



assume that the explicit and implicit cytoplasmic domains are perfectly well-mixed along the  $z$ -axis. We rewrite this solution after substituting  $\Delta t = t - t_0$  and also redefine  $z = z - z_0$  for convenience.

$$p(z) = \frac{1}{\sqrt{4\pi D\Delta t}} \exp\left[-\frac{z^2}{4D\Delta t}\right]$$

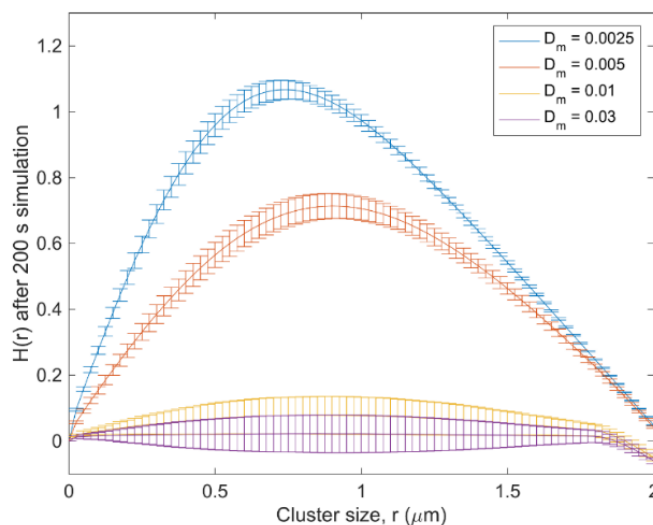
To compute the probability of diffusing from one compartment to the other, we integrate  $p$  ranging over the distance that must be traveled. We either assume that diffusion must halt at the maximal distance (subsequently denoted the ‘exact case’), or include diffusional probabilities that bypass the maximal distance towards infinity (the ‘approximate case’). Throughout the manuscript, we use the exact case, but the approximate case can be intuitively examined for correctness, so we present it here.

$$\begin{aligned} P_{inj}(z_i) &= \int_{z_{impl}-z}^{z_{max}-z} \frac{1}{\sqrt{4\pi D\Delta t}} \exp\left[-\frac{z^2}{4D\Delta t}\right] dz \approx \int_{z_{impl}-z}^{\infty} \frac{1}{\sqrt{4\pi D\Delta t}} \exp\left[-\frac{z^2}{4D\Delta t}\right] dz \\ P_{inj}(z) &= \frac{1}{2} \left[ \operatorname{erf}\left(\frac{z_{max}-z}{\sqrt{4D\Delta t}}\right) - \operatorname{erf}\left(\frac{z_{impl}-z}{\sqrt{4D\Delta t}}\right) \right] \approx \frac{1}{2} \left[ 1 - \operatorname{erf}\left(\frac{z_{impl}-z}{\sqrt{4D\Delta t}}\right) \right] \\ P_{ejc}(z) &= \int_{z-z_{impl}}^z \frac{1}{\sqrt{4\pi D\Delta t}} \exp\left[-\frac{z^2}{4D\Delta t}\right] dz \approx \int_{z-z_{impl}}^{-\infty} \frac{1}{\sqrt{4\pi D\Delta t}} \exp\left[-\frac{z^2}{4D\Delta t}\right] dz \\ P_{ejc}(z) &= \frac{1}{2} \left[ \operatorname{erf}\left(\frac{z}{\sqrt{4D\Delta t}}\right) - \operatorname{erf}\left(\frac{-(z_{impl}-z)}{\sqrt{4D\Delta t}}\right) \right] \approx \frac{1}{2} \left[ 1 - \operatorname{erf}\left(\frac{-(z_{impl}-z)}{\sqrt{4D\Delta t}}\right) \right] \end{aligned}$$

The term  $z$  is the depth of the particle in coordinates where  $z = 0$  is the bottom of the implicit reservoir,  $z_{impl}$  is the depth marking the interface between the implicit and explicit domains, and  $z_{max}$  is the total height of the implicit and explicit domains, as illustrated in Figure 1.10. In the approximate forms, we note that if we consider a slice at the explicit-implicit boundary  $z = z_{impl}$ , we have  $P_{inj} = P_{ejc} = 0.5$ , which reflects that a Brownian particle is equally likely to diffuse up or diffuse down. Cutting the integrals at the cytoplasmic boundaries alters this probability slightly, but not by a noticeable amount. We choose to use the exact integrals since there is very minimal computational advantage to using the approximate form.

### 1.7.6. Loss of polarity with increasing membrane diffusivity

Other reports in the literature have assumed faster diffusion coefficients on the membrane (Kl nder et al., 2013). Doing the same in our model without changing any other parameters causes a loss in polarity (Figure 1.22).

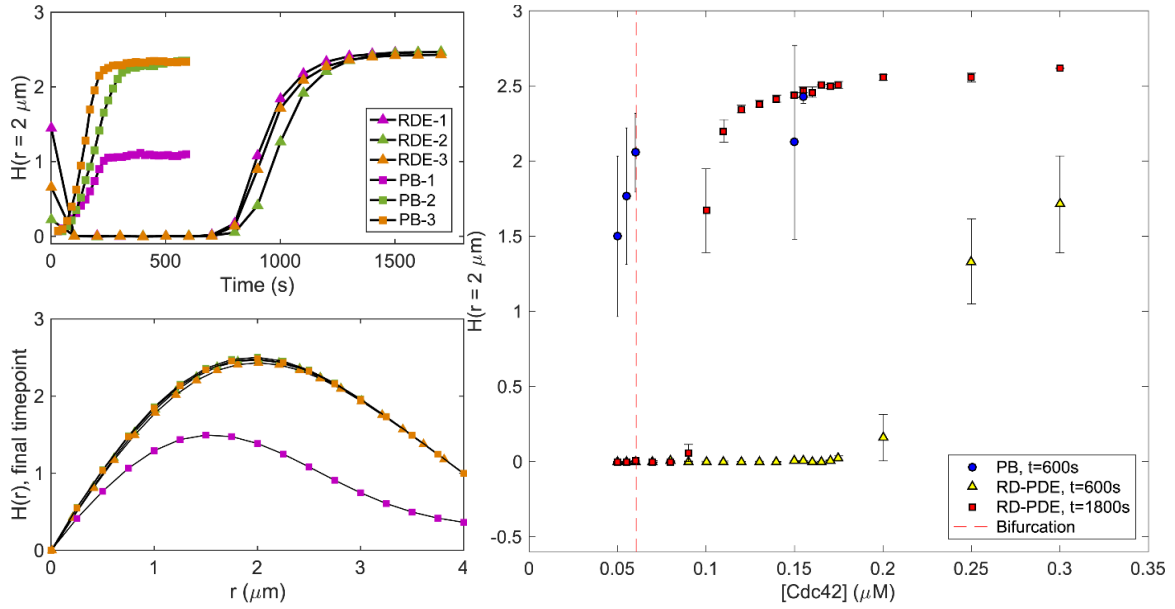


**Figure 1.22. Loss of polarization with increasing membrane diffusivity.**

The parameters in Figure 1.6A-B with  $N_{Cd42} = 1921$  and  $N_{GEF} = 116$ , were used to create the results for  $D_m = 0.0025 \mu\text{m}^2/\text{s}$ . We then increased the diffusivity on the membrane while keeping all other parameters fixed. Each curve represents the mean  $\pm$  1 s.d. from  $n=5$  realizations.

### 1.7.7. An alternative to the quasi-3D system

In addition to the quasi-3D, reservoir-based approximation of the cell cytoplasm, we also considered a volume-adjusted, two-compartment RDE system, as had been previously reported in the literature (Wu et al., 2015) to complement our quasi-3D simulation results. We observe no qualitative change our results: the particle-based simulations still exhibit accelerated polarization relative to the RDE simulations, and also still polarize past the deterministically-predicted bifurcation point (Figure 1.23).



**Figure 1.23. Considering the volume-adjusted, two-compartment RDE system makes no qualitative difference for our observations.**

Left: polarization dynamics and quantitative polarity site similarity, as in Figure 1.8. Here, 1800s was sufficient for full polarization, so the RDE  $H(r)$  curves were overlaid with the PB simulations. Right: Bifurcation diagram. The bifurcation point was calculated using linear stability analysis.

## REFERENCES

- Agbanusi, I.C., and Isaacson, S.A. (2014). A comparison of bimolecular reaction models for stochastic reaction-diffusion systems. *Bull Math Biol* 76, 922–946.
- Altschuler, S.J., Angenent, S.B., Wang, Y., and Wu, L.F. (2008). On the spontaneous emergence of cell polarity. *Nature* 454, 886–889.
- Andrew, N., and Insall, R.H. (2007). Chemotaxis in shallow gradients is mediated independently of PtdIns 3-kinase by biased choices between random protrusions. *Nat. Cell Biol.* 9, 193–200.
- Andrews, S.S., and Bray, D. (2004). Stochastic simulation of chemical reactions with spatial resolution and single molecule detail. *Phys. Biol.* 1, 137–151.
- Angermann, B.R., Klauschen, F., Garcia, A.D., Prustel, T., Zhang, F., Germain, R.N., and Meier-Schellersheim, M. (2012). Computational modeling of cellular signaling processes embedded into dynamic spatial contexts. *Nat. Methods* 9, 283–289.
- Bhattacharjee, N., and Folch, A. (2017). Large-scale microfluidic gradient arrays reveal axon guidance behaviors in hippocampal neurons. *Microsyst. Nanoeng.* 3, 17003.
- Chiou, J.-G., Balasubramanian, M.K., and Lew, D.J. (2017). Cell polarity in yeast. *Annu. Rev. Cell Dev. Biol.* 33, 77–101.
- Doi, M. (1976). Stochastic theory of diffusion-controlled reaction. *J. Phys. A, Math. Gen.* 9, 1479–1495.
- Duina, A.A., Miller, M.E., and Keeney, J.B. (2014). Budding yeast for budding geneticists: a primer on the *Saccharomyces cerevisiae* model system. *Genetics* 197, 33–48.
- Dyer, J.M., Savage, N.S., Jin, M., Zyla, T.R., Elston, T.C., and Lew, D.J. (2013). Tracking shallow chemical gradients by actin-driven wandering of the polarization site. *Curr. Biol.* 23, 32–41.
- Dziekan, P., Lemarchand, A., and Nowakowski, B. (2012). Particle dynamics simulations of Turing patterns. *J. Chem. Phys.* 137, 074107.
- Erban, R., and Chapman, S.J. (2009). Stochastic modelling of reaction-diffusion processes: algorithms for bimolecular reactions. *Phys. Biol.* 6, 046001.
- Etienne-Manneville, S. (2004). Cdc42--the centre of polarity. *J. Cell Sci.* 117, 1291–1300.
- Fange, D., Berg, O.G., Sjöberg, P., and Elf, J. (2010). Stochastic reaction-diffusion kinetics in the microscopic limit. *Proc. Natl. Acad. Sci. USA* 107, 19820–19825.
- Franz, B., Flegg, M.B., Chapman, S.J., and Erban, R. (2013). Multiscale Reaction-Diffusion Algorithms: PDE-Assisted Brownian Dynamics. *SIAM J Appl Math* 73, 1224–1247.
- Freisinger, T., Klünder, B., Johnson, J., Müller, N., Pichler, G., Beck, G., Costanzo, M., Boone, C., Cerione, R.A., Frey, E., et al. (2013). Establishment of a robust single axis of cell polarity by coupling multiple positive feedback loops. *Nat. Commun.* 4, 1807.
- Gierer, A., and Meinhardt, H. (1972). A theory of biological pattern formation. *Kybernetik* 12, 30–39.

- Giese, W., Eigel, M., Westerheide, S., Engwer, C., and Klipp, E. (2015). Influence of cell shape, inhomogeneities and diffusion barriers in cell polarization models. *Phys. Biol.* 12, 066014.
- Gillespie, D.T. (1977). Exact stochastic simulation of coupled chemical reactions. *J. Phys. Chem.* 81, 2340–2361.
- Gillespie, D.T. (2000). The chemical Langevin equation. *J. Chem. Phys.* 113, 297.
- Goryachev, A.B., and Leda, M. (2017). Many roads to symmetry breaking: molecular mechanisms and theoretical models of yeast cell polarity. *Mol. Biol. Cell* 28, 370–380.
- Goryachev, A.B., and Pokhilko, A.V. (2008). Dynamics of Cdc42 network embodies a Turing-type mechanism of yeast cell polarity. *FEBS Lett.* 582, 1437–1443.
- Haberman, R. (1998). *Elementary applied partial differential equations.*
- Hattne, J., Fange, D., and Elf, J. (2005). Stochastic reaction-diffusion simulation with MesoRD. *Bioinformatics* 21, 2923–2924.
- Higham, D.J. (2001). An algorithmic introduction to numerical simulation of stochastic differential equations. *SIAM Rev.* 43, 525–546.
- Howell, A.S., Savage, N.S., Johnson, S.A., Bose, I., Wagner, A.W., Zyla, T.R., Nijhout, H.F., Reed, M.C., Goryachev, A.B., and Lew, D.J. (2009). Singularity in polarization: rewiring yeast cells to make two buds. *Cell* 139, 731–743.
- Howell, A.S., Jin, M., Wu, C.-F., Zyla, T.R., Elston, T.C., and Lew, D.J. (2012). Negative feedback enhances robustness in the yeast polarity establishment circuit. *Cell* 149, 322–333.
- Jilkine, A., Angenent, S.B., Wu, L.F., and Altschuler, S.J. (2011). A density-dependent switch drives stochastic clustering and polarization of signaling molecules. *PLoS Comput. Biol.* 7, e1002271.
- Kerr, R.A., Bartol, T.M., Kaminsky, B., Dittrich, M., Chang, J.-C.J., Baden, S.B., Sejnowski, T.J., and Stiles, J.R. (2008). Fast monte carlo simulation methods for biological reaction-diffusion systems in solution and on surfaces. *SIAM J Sci Comput* 30, 3126.
- Kiskowski, M.A., Hancock, J.F., and Kenworthy, A.K. (2009). On the use of Ripley’s K-function and its derivatives to analyze domain size. *Biophys. J.* 97, 1095–1103.
- Klunder, B., Freisinger, T., Wedlich-Söldner, R., and Frey, E. (2013). GDI-mediated cell polarization in yeast provides precise spatial and temporal control of Cdc42 signaling. *PLoS Comput. Biol.* 9, e1003396.
- Kuo, C.-C., Savage, N.S., Chen, H., Wu, C.-F., Zyla, T.R., and Lew, D.J. (2014). Inhibitory GEF phosphorylation provides negative feedback in the yeast polarity circuit. *Curr. Biol.* 24, 753–759.
- Lakhani, V., and Elston, T.C. (2017). Testing the limits of gradient sensing. *PLoS Comput. Biol.* 13, e1005386.
- Lemarchand, A., and Nowakowski, B. (2011). Do the internal fluctuations blur or enhance axial segmentation? *EPL* 94, 48004.

- Lipková, J., Zygalakis, K.C., Chapman, S.J., and Erban, R. (2011). Analysis of brownian dynamics simulations of reversible bimolecular reactions. *SIAM J Appl Math* 71, 714–730.
- McClure, A.W., Minakova, M., Dyer, J.M., Zyla, T.R., Elston, T.C., and Lew, D.J. (2015). Role of polarized G protein signaling in tracking pheromone gradients. *Dev. Cell* 35, 471–482.
- McKane, A.J., Biancalani, T., and Rogers, T. (2014). Stochastic pattern formation and spontaneous polarisation: the linear noise approximation and beyond. *Bull Math Biol* 76, 895–921.
- Mori, Y., Jilkine, A., and Edelstein-Keshet, L. (2008). Wave-pinning and cell polarity from a bistable reaction-diffusion system. *Biophys. J.* 94, 3684–3697.
- Murray, J.D. (2003). *Mathematical Biology II* (New York: Springer).
- Owen, D.M., Rentero, C., Rossy, J., Magenau, A., Williamson, D., Rodriguez, M., and Gaus, K. (2010). PALM imaging and cluster analysis of protein heterogeneity at the cell surface. *J. Biophotonics* 3, 446–454.
- Ozbudak, E.M., Becskei, A., and van Oudenaarden, A. (2005). A system of counteracting feedback loops regulates Cdc42p activity during spontaneous cell polarization. *Dev. Cell* 9, 565–571.
- Ramirez, S.A., Raghavachari, S., and Lew, D.J. (2015). Dendritic spine geometry can localize GTPase signaling in neurons. *Mol. Biol. Cell* 26, 4171–4181.
- Rappel, W.-J., and Edelstein-Keshet, L. (2017). Mechanisms of cell polarization. *Current Opinion in Systems Biology* 3, 43–53.
- Schaff, J.C., Gao, F., Li, Y., Novak, I.L., and Slepchenko, B.M. (2016). Numerical Approach to Spatial Deterministic-Stochastic Models Arising in Cell Biology. *PLoS Comput. Biol.* 12, e1005236.
- Schöneberg, J., Ullrich, A., and Noé, F. (2014). Simulation tools for particle-based reaction-diffusion dynamics in continuous space. *BMC Biophys.* 7, 11.
- Sept, D., and Carlsson, A.E. (2014). Modeling large-scale dynamic processes in the cell: polarization, waves, and division. *Q Rev Biophys* 47, 221–248.
- Slaughter, B.D., Smith, S.E., and Li, R. (2009). Symmetry breaking in the life cycle of the budding yeast. *Cold Spring Harb. Perspect. Biol.* 1, a003384.
- Smoluchowski, M. (1917). Versuch einer mathematischen Theorie der Koagulationskinetik kolloider Lösungen. *Z. Phys. Chem.* 92, 129–168.
- Tahirovic, S., and Bradke, F. (2009). Neuronal polarity. *Cold Spring Harb. Perspect. Biol.* 1, a001644.
- Trong, P.K., Nicola, E.M., Goehring, N.W., Kumar, K.V., and Grill, S.W. (2014). Parameter-space topology of models for cell polarity. *New J Phys* 16, 065009.
- Turing, A.M. (1952). The chemical basis of morphogenesis. *Philosophical Transaction of the Royal Society of London. Series B.* 237, 37–72.
- van Zon, J.S., and ten Wolde, P.R. (2005a). Simulating biochemical networks at the particle level and in time and space: Green’s function reaction dynamics. *Phys. Rev. Lett.* 94, 128103.

- van Zon, J.S., and ten Wolde, P.R. (2005b). Green's-function reaction dynamics: a particle-based approach for simulating biochemical networks in time and space. *J. Chem. Phys.* 123, 234910.
- von Philipsborn, A., and Bastmeyer, M. (2007). Mechanisms of gradient detection: a comparison of axon pathfinding with eukaryotic cell migration. *Int Rev Cytol* 263, 1–62.
- Walther, G.R., Marée, A.F.M., Edelstein-Keshet, L., and Grieneisen, V.A. (2012). Deterministic versus stochastic cell polarisation through wave-pinning. *Bull Math Biol* 74, 2570–2599.
- Watson, L.J., Rossi, G., and Brennwald, P. (2014). Quantitative analysis of membrane trafficking in regulation of Cdc42 polarity. *Traffic* 15, 1330–1343.
- Wedlich-Soldner, R., Altschuler, S., Wu, L., and Li, R. (2003). Spontaneous cell polarization through actomyosin-based delivery of the Cdc42 GTPase. *Science* 299, 1231–1235.
- Wehrens, M., ten Wolde, P.R., and Mugler, A. (2014). Positive feedback can lead to dynamic nanometer-scale clustering on cell membranes. *J. Chem. Phys.* 141, 205102.
- Woods, B., Lai, H., Wu, C.-F., Zyla, T.R., Savage, N.S., and Lew, D.J. (2016). Parallel Actin-Independent Recycling Pathways Polarize Cdc42 in Budding Yeast. *Curr. Biol.* 26, 2114–2126.
- Wu, C.-F., Chiou, J.-G., Minakova, M., Woods, B., Tsygankov, D., Zyla, T.R., Savage, N.S., Elston, T.C., and Lew, D.J. (2015). Role of competition between polarity sites in establishing a unique front. *Elife* 4.
- Yogurtcu, O.N., and Johnson, M.E. (2015). Theory of bi-molecular association dynamics in 2D for accurate model and experimental parameterization of binding rates. *J. Chem. Phys.* 143, 084117.

## CHAPTER 2: RATIOMETRIC GPCR SIGNALING ENABLES DIRECTIONAL SENSING IN YEAST<sup>2</sup>

### Overview

As described in Chapter 1, accurately sensing extracellular chemical gradients is essential in yeast polarization during mating. Gradient sensing is also important for the aggregation of social amoebae, the directed growth of neurons, the migration of sperm towards eggs, and the clearance of pathogens by neutrophils (Alvarez et al., 2014; Bhattacharjee and Folch, 2017; Rappel and Edelstein-Keshet, 2017; Swaney et al., 2010; von Philipsborn and Bastmeyer, 2007). In most cases, cells sense external signals via G-protein-coupled receptors (GPCRs), leading to cytoskeletal reorganization that produces directional growth or movement (Insall, 2013).

The yeast *Saccharomyces cerevisiae* uses GPCRs to sense extracellular gradients of pheromone secreted by potential mates. Pheromone-bound GPCRs activate intracellular G proteins, initiating a cascade to orient the polarity machinery. In collaboration with researchers at Duke University, we showed that yeast successfully bias their polarity sites up-gradient despite pheromone receptors that are asymmetrically distributed along the cell surface (Henderson et al., 2019). Receptor asymmetry can result in a ligand-bound gradient of receptors that do not accurately reflect the external pheromone gradient, implying yeast must have an error correction mechanism.

Here, I present simulations predicting that yeast can robustly decode gradients by responding to the ratio of occupied to free receptors, rather than the abundance of occupied receptors. The biochemical mechanism involves stimulation of signaling by bound receptors together with inhibition of signaling by

---

<sup>2</sup> This chapter has appeared as part of an article in the journal PLoS Biology. The original citation is as follows: Henderson NT, Pablo M, Ghose D, Clark-Cotton MR, Zyla TR, Nolen J, Elston TC, Lew DJ. “Ratiometric GPCR signaling enables directional sensing in yeast.” *PLoS Biol.* 17(10): e3000484.



free receptors. These predictions were tested and verified by our collaborators, whose results are provided in the journal publication and not here. This mechanism can also sharpen the gradient of activated G protein. Chapter 1 investigated yeast polarity establishment by modeling a Cdc42-centric positive feedback loop. This chapter begins to consider polarization in the context of mating: how are extracellular gradients of pheromone accurately converted into intracellular gradients? Then in Chapter 3, we will model how such intracellular gradients influence Cdc42-mediated polarization.

## 2.1. Introduction

The sequence of molecular events that convert extracellular chemical gradients to directional outputs is perhaps best understood in the genetically tractable budding yeast *S. cerevisiae*. Yeast cells of mating type **a** (MAT $\mathbf{a}$ ) can mate with cells of mating type  $\alpha$  (MAT $\alpha$ ). The cells secrete peptide pheromones that bind GPCRs on cells of the opposite type ( $\alpha$ -factor is sensed by Ste2 in **a** cells, and **a**-factor is sensed by Ste3 in  $\alpha$  cells) (Wang and Dohlman, 2004). Once bound by pheromone, these GPCRs activate heterotrimeric G proteins to generate GTP-G $\alpha$  and G $\beta\gamma$ . Pheromone binding is linked to polarity establishment through G $\beta\gamma$ -mediated recruitment of Far1 and the formation of a G $\beta\gamma$ -Far1-Cdc24 complex, which is capable of activating the master polarity regulator Cdc42. Our model in this chapter stops at G protein activation, and Chapter 3 describes the subsequent steps in more detail.

Like other eukaryotic cells, yeast are thought to compare ligand concentrations across the cell in order to determine the orientation of the gradient (Arkowitz, 2009). If the cellular distribution of pheromone-activated receptors reflects the pheromone gradient, then G $\beta\gamma$ -Far1-Cdc24 complexes will be enriched up-gradient, spatially biasing activation of Cdc42 to kick off positive feedback at the right location for mating. However, the small size of yeast ( $\sim 5\ \mu\text{m}$  diameter) makes such global gradient sensing difficult (Berg and Purcell, 1977). Indeed, simulations constrained by experimental data on binding and diffusion parameters have indicated that the process is inaccurate (Lakhani and Elston, 2017). When yeast are exposed to artificial, calibrated pheromone gradients, the initial polarization often occurs in the wrong direction (Moore et al., 2008; Segall, 1993). These cells are nevertheless able to re-orient the polarity site

(Dyer et al., 2013; Hegemann and Peter, 2017; Kelley et al., 2015), through actin and vesicle-dependent mechanisms (McClure et al., 2015; Savage et al., 2012). The moving patch has been theorized to operate like a sensor for changes in the local pheromone concentration (Ayscough and Drubin, 1998; Suchkov et al., 2010). It is unclear to what extent yeast cells rely on global sensing to orient the formation of a polarity site versus local sensing to re-orient the polarity site. A recent study found that when cells were placed in an artificial pheromone gradient in a microfluidics device, initial site formation was essentially random with respect to the gradient, and orientation occurred almost entirely by exploratory polarization (Hegemann et al., 2015).

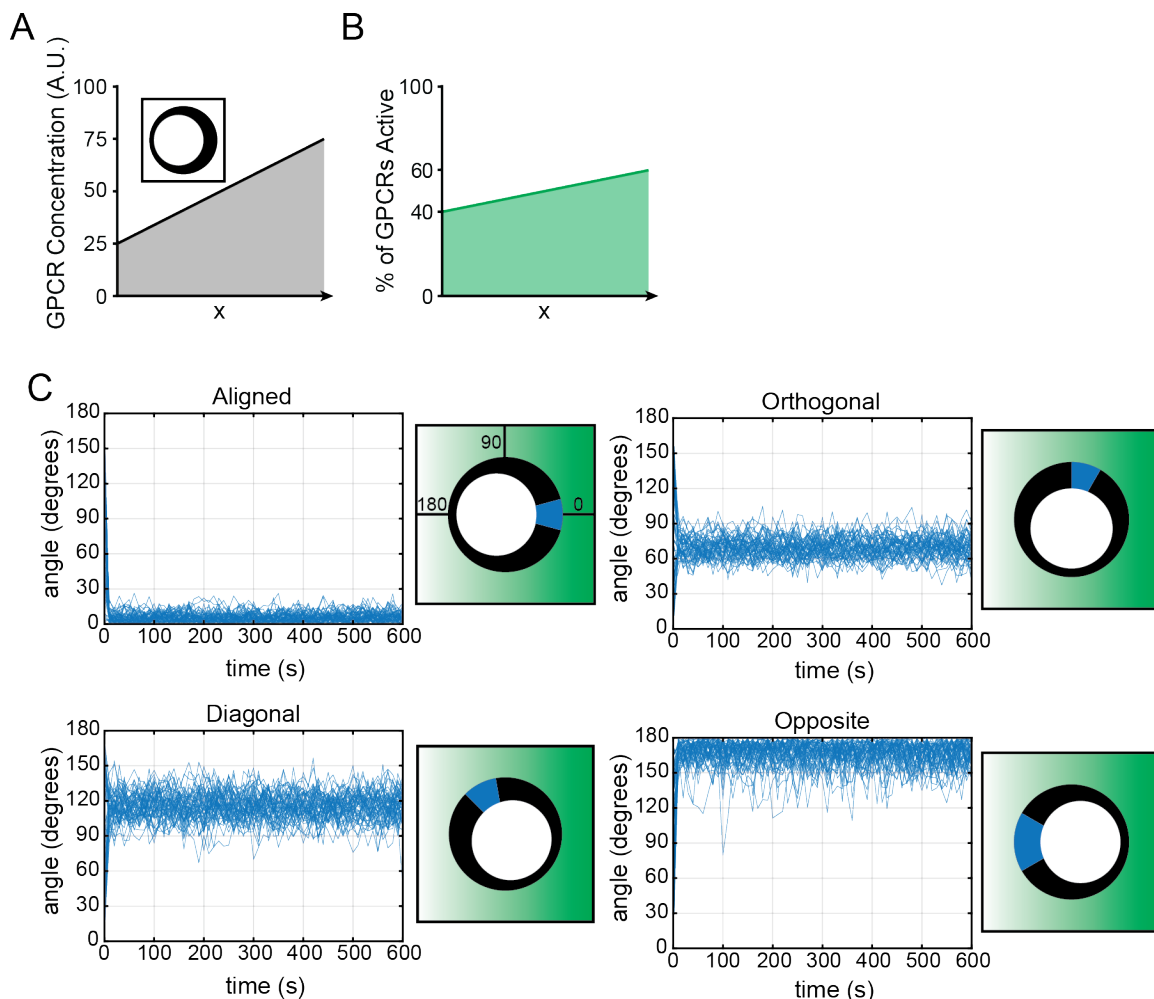
To study spatial sensing in a more natural context, our collaborators imaged mating events in mixed populations of MAT $\alpha$  and MAT $\alpha$  cells and found evidence for both global spatial sensing and subsequent local error correction. Importantly, they observed nonrandom initial clustering of polarity proteins biased towards eventual mating partners. The initial polarization was surprisingly accurate given that it occurred despite a highly nonuniform, and thus potentially misleading, distribution of Ste2 receptors. In principle, this would disrupt global sensing. In the following, I present computational simulations of Ste2 signaling that show how cells could correct for asymmetric receptor distributions.

## **2.2. Nonuniform GPCR distributions can mislead G-protein activation**

The Ste2 distribution along the cell surface was highly heterogeneous in growing cells that were not exposed to  $\alpha$ -factor, with on average 3-fold variation in local concentration. This nonuniform receptor distribution poses a potential problem for accurate gradient sensing: one might expect cells to be preferentially sensitive to pheromone on the side where receptors are enriched, which may not correspond to the side facing a mating partner. To illustrate the problem, we conducted particle-based simulations of a model spherical cell with receptors distributed unevenly over the surface with a 3-fold difference from one side of the cell to the other (Figure 2.1A). We assume that heterotrimeric G proteins diffuse at the membrane and become activated when they encounter a ligand-bound receptor. G-protein inactivation occurs at a rate that is the same everywhere (see 2.6. Methods). A stable pheromone gradient was

simulated by assuming that the probability that a receptor is active is 1.5-fold higher on the up-gradient side than the down-gradient side of the cell (Figure 2.1B). Receptor diffusion does not significantly blur this gradient (Figure 2.6). While pheromone gradients may not be stable under mating conditions, we use them in our model to allow for straightforward interpretation.

At any given time, our simulations provide the locations of all active G proteins on the cell surface. From that, we calculated a resultant vector for active G protein and plotted the angle between this vector and the imposed pheromone gradient (Figure 2.1C). When the receptor density gradient and the pheromone gradient were aligned, the simulated cells identified the correct direction. However, when the density gradient was not aligned with the pheromone gradient, the simulated cells were easily misled, with active G proteins accumulating in regions with high receptor density (Figure 2.1C). Thus, without some compensatory mechanism, we would expect yeast cells to have difficulty decoding the pheromone gradient in the presence of an uncorrelated receptor density gradient.



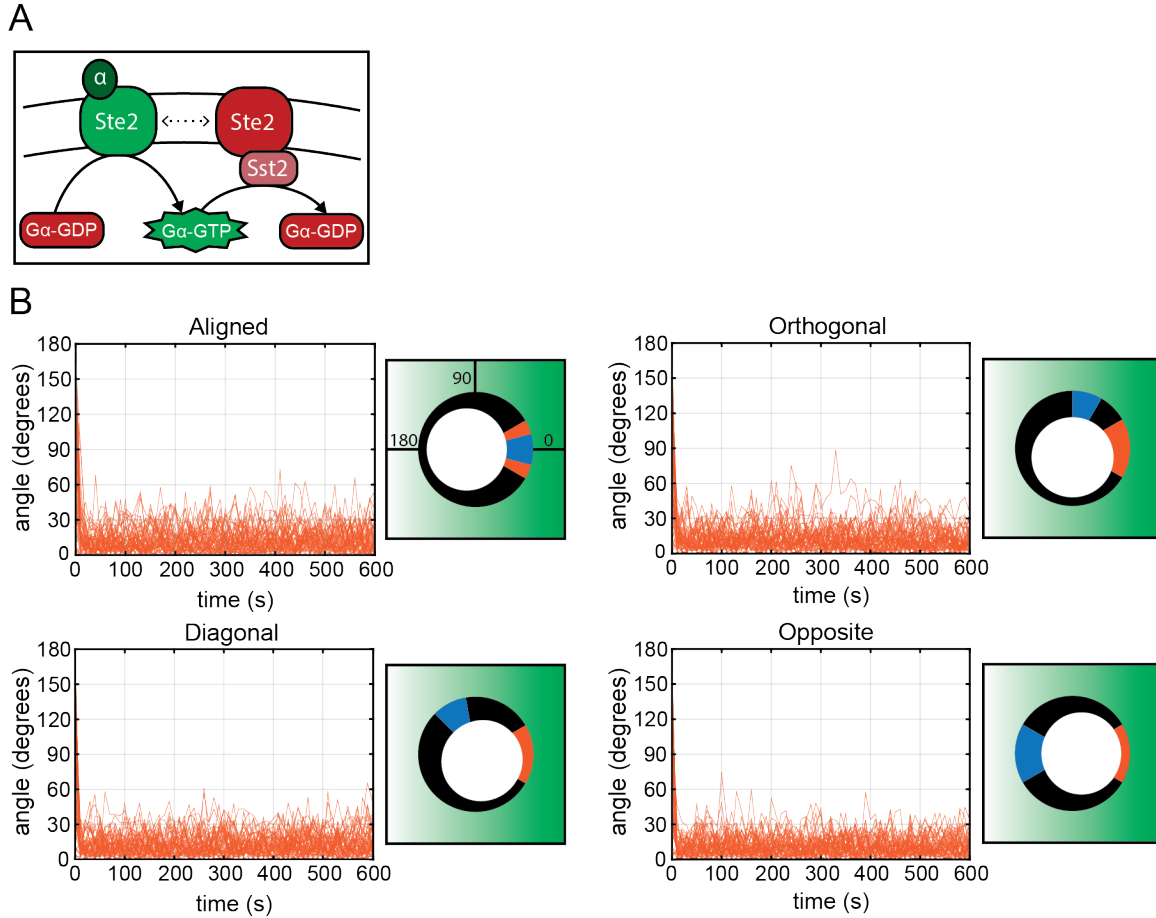
**Figure 2.1. GPCR density variation along the cell membrane can mislead G protein activation.** (A) Simulated receptors were distributed unevenly: receptor density is indicated by the thickness of the black line (inset) and reflects a 3-fold gradient, similar to the experimental Ste2 distribution. (B) A 1.5-fold pheromone gradient was simulated along the x axis by varying the % of active receptors from 40% to 60% across the cell diameter. (C) Simulations were conducted with receptor activity and density gradients oriented as in the illustrations. The locations of all the active G proteins were used to calculate a G-protein vector, whose angle to the direction of the pheromone gradient is plotted (y axis) against time (x axis) (left). 0° indicates perfect orientation: active G-protein vector in the same direction as the applied receptor activity gradient. The approximate range of G-protein vectors (blue wedge) is shown on the cartoon on the right, along with the pheromone gradient (green shading) and receptor density (as in A).

### 2.3. Ratiometric GPCR signaling robustly directs G-protein activation

A potential compensatory mechanism that could correct for the presence of more active receptors at sites of high receptor density exploits the fact that receptor-dense regions would contain more unbound receptor as well as ligand-bound receptor. If unbound receptor counteracts G-protein activation, that

could cancel out the higher rate of G-protein inactivation by ligand-bound receptors. In yeast cells, the regulator of G-protein signaling (RGS) protein Supersensitive 2 (Sst2) that inactivates the G protein is recruited to the cell membrane via binding to unoccupied Ste2 (Ballon et al., 2006). A recent study insightfully suggested that this would cause cells to measure the ratio of ligand-bound to unbound receptors, i.e. ratiometric sensing (Bush et al., 2016). Pheromone-bound Ste2 loads GTP on  $G\alpha$ , whereas unbound Ste2-Sst2 promotes GTP hydrolysis by  $G\alpha$ , so the level of activated  $G\alpha$  depends on the ratio between the pheromone-bound and unbound Ste2 rather than the absolute level of bound Ste2 (Figure 2.2A). Here, we explore the possibility that such ratiometric sensing would also lead to measurement of the spatial distribution of the ratio of active/total receptors so that differences in the local receptor density would not distort a cell's ability to determine the orientation of a pheromone gradient.

We repeated the simulations in which G proteins decode a pheromone gradient that is distorted by the presence of uneven receptor density. The central difference was that instead of a single G-protein deactivation rate regardless of spatial position, G-protein deactivation occurred when an active G protein encountered an unbound receptor. For fair comparison, the G-protein inactivation rate constants for the nonratiometric and ratiometric simulations were empirically calibrated to produce similar levels of active G protein at the midpoint of the gradient (see 2.6. Methods). For the ratiometric model, the simulated cell correctly identified the direction of the pheromone gradient no matter what the receptor density distribution (Figure 2.2B). These simulations assumed that bimolecular reactions were diffusion-limited, but similar results were obtained assuming reaction-limited kinetics (Figure 2.4B). Thus, ratiometric sensing provides a robust mechanism for prevent cells from being misled by uneven receptor density.



**Figure 2.2. Ratiometric sensing allows cells to orient up-gradient despite uneven receptor density.** (A) Proposed ratiometric pheromone sensing mechanism. The G protein is activated by pheromone-bound receptor (Ste2 +  $\alpha$ -factor) and inactivated by the RGS protein Sst2. Sst2 associates with pheromone-free Ste2. Thus, G-protein activity reflects the ratio of bound to unbound receptors. (B). Particle-based simulations were repeated as in Figure 2.1, except that instead of spontaneous inactivation, G proteins were inactivated upon encountering inactive receptors. These “ratiometric” simulations (orange) were plotted as in Figure 2.1C. For comparison, both the nonratiometric (blue) and ratiometric (orange) results are depicted in the cartoons.

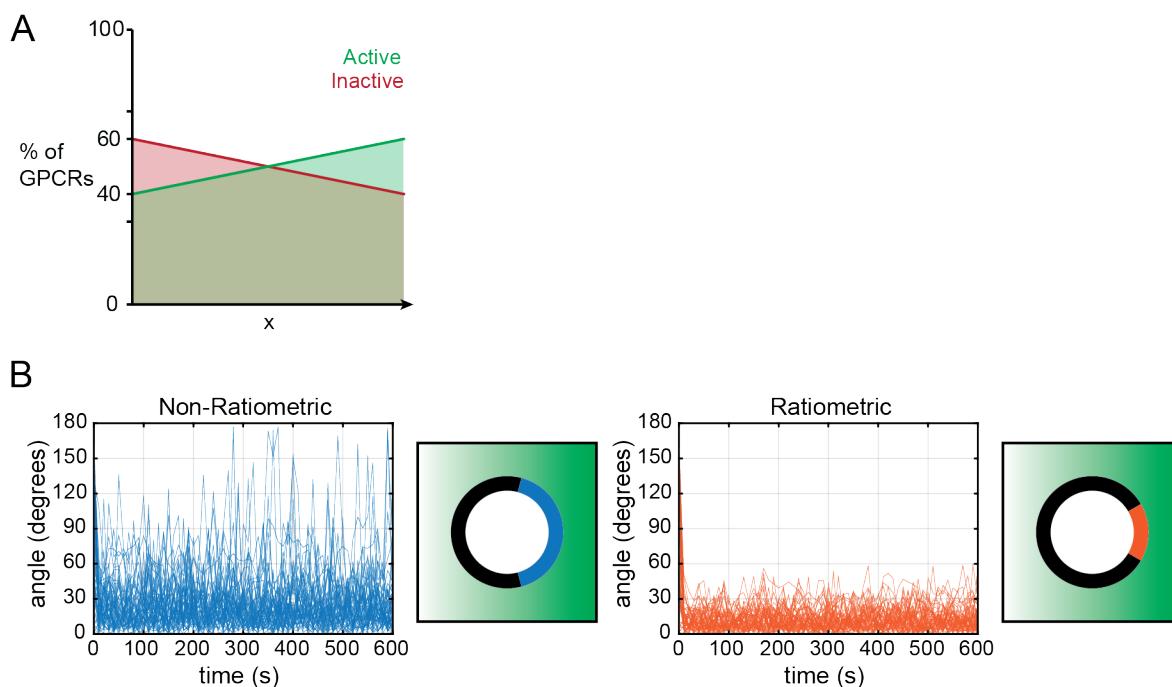
Experimentally breaking ratiometric signaling by replacing the negative regulator Sst2 with a human paralog hsRGS4 that does not associate with Ste2 resulted in cells whose polarity sites correlated with Ste2 receptor density rather than mating partners. Furthermore, in agreement with the model, the partner-sensing accuracy of non-ratiometric cells could be rescued replacing wildtype Ste2 with mutants that had impaired endocytosis and thus a more uniform membrane distribution. Therefore, we infer that yeast cells

use Sst2-dependent local ratiometric sensing of receptor occupancy to extract accurate information from the pheromone gradient despite having nonuniform receptor density.

#### **2.4. Ratiometric GPCR signaling can sharpen the G-protein gradient**

In addition to protecting cells from being misled by uneven receptor density, ratiometric sensing could, in principle, confer a benefit even in cells that had uniformly distributed receptors. This is because a gradient of pheromone would generate both a gradient in the concentration of ligand-bound receptors and an opposing gradient in the concentration of unoccupied receptors (Figure 2.3A). Consider a gradient of active receptor rising from left to right across the cell, with 50% active receptor in between. If we compare ratiometric and nonratiometric sensing models matched so that the G-protein deactivation rate in both models is equal when 50% of the receptor is bound to ligand, then the active G-protein concentration on the right side will be higher for the ratiometric model because the inactivation rate (mediated by inactive receptor, which is <50% on this side) is lower. Similarly, the active G-protein concentration on the left side will be lower for the ratiometric model because the inactivation rate is higher. Thus, the difference in active G-protein concentration between the two ends of the cell will always be larger in the ratiometric model.

To examine the benefit of ratiometric sensing, we conducted particle-based simulations with uniform receptor density. As predicted, the ratiometric model reduced the noise in the direction of the gradient as compared to the nonratiometric model (Figure 2.3B). This benefit was preserved with receptor diffusion (Figure 2.5A), and ratiometric sensing outperformed nonratiometric sensing even with as few as 1,000 total receptors (Figure 2.5B). Thus, ratiometric sensing can, in principle, provide a significant benefit even to cells with uniform receptors.



**Figure 2.3. Ratiometric sensing amplifies the gradient signal and improves accuracy even when receptors are distributed uniformly.**

(A) When receptor distribution is uniform, a gradient of active receptors automatically implies an opposing gradient of inactive receptors. (B) Simulations with uniform receptor density. The ratiometric (orange) and nonratiometric (blue) models were simulated as in Figure 2.2B.

## 2.5. Discussion

The rapid diffusion of peptide pheromones and the small size of the yeast cell led to the expectation that there would be only a small difference in pheromone concentration between the up- and down-gradient sides of the cell. This poses a fundamental difficulty in extracting accurate directional information in the face of molecular noise (Berg and Purcell, 1977). Indeed, cells responding to a 0.5 nM/ $\mu\text{m}$  pheromone gradient were reported to orient initial polarity clusters almost at random (Hegemann et al., 2015). Moreover, the polarity circuit in yeast contains strong positive feedback (Chiou et al., 2017; Johnson et al., 2011), which allows cells to polarize in random directions when treated with uniform pheromone concentrations (Dyer et al., 2013; Strickfaden and Pryciak, 2008). This would be expected to enable noise-driven polarization in random directions in cells exposed to shallow gradients (Chou et al., 2008). Making matters even worse, we documented significant receptor asymmetry, with (on average) 3-



fold more concentrated receptors on one side of the cell than the other. This creates a receptor density gradient that is significantly steeper than the assumed pheromone gradient. Because the receptor density gradient is randomly oriented with respect to the mating partner, this poses a serious hurdle in accurate gradient detection. Despite the difficulties enumerated above, the location of initial polarity factor clustering in mating mixtures is highly nonrandom and surprisingly accurate, with more than 50% of cells clustering within 30° of the correct direction and less than 5% of cells clustering in the opposite segment (a random process would have 17% of cells polarizing in each segment).

One way to avoid being misled by an asymmetric receptor distribution would be to compare the local ratio of occupied and unoccupied receptors, rather than simply the density of occupied receptors, across the cell surface. An elegant mechanism to extract such information was proposed by (Bush et al., 2016). Because the RGS protein Sst2 binds to unoccupied receptors (Ballon et al., 2006), those receptors promote GTP hydrolysis by  $G\alpha$ . Conversely, occupied receptors catalyze GTP-loading by  $G\alpha$ . Thus, the net level of GTP- $G\alpha$  reflects the fraction (and not the number) of occupied receptors on the cell (Bush et al., 2016). For this mechanism to promote local ratiometric sensing, it requires additionally that a pheromone-bound receptor diffuse slowly relative to its lifetime at the surface (approximately 10 min) (Jenness and Spatrick, 1986) so that information about where receptors were when they bound to pheromone is not lost. We found that receptors do indeed diffuse very slowly ( $D < 0.0005 \mu\text{m}^2/\text{s}$ ) at the yeast plasma membrane. Moreover, ratiometric gradient sensing requires that the ratio of active to inactive receptors is measured locally rather than globally. Simulations with realistic numbers of receptors and G proteins demonstrated that this mechanism has the potential to extract unbiased information about the pheromone gradient even in the face of uneven receptor density.

When RGS function was delocalized by replacing Sst2 (which binds unoccupied receptors) with an equivalently active amount of hsRGS4 (which binds the plasma membrane), the accuracy of initial polarity clustering was severely compromised. Instead of polarizing towards potential partners, these cells assembled polarity clusters at regions where receptors were concentrated. Thus, abrogating the Sst2-based ratiometric sensing mechanism allowed cells to be misled by the asymmetric receptor distribution.

Accurate orientation could be restored to these cells by making receptor distribution more uniform. In sum, our findings suggest that local ratiometric sensing compensates for uneven receptor distribution and allows more accurate polarization towards mating partners.

An additional benefit of ratiometric sensing in terms of gradient detection is that this mechanism exploits both the gradient in ligand-bound receptors and the complementary gradient in unoccupied receptors to sharpen the downstream G-protein gradient. This feature would be beneficial even in cells with uniform receptor density. Whereas for yeast, the main function of ratiometric sensing appears to be to correct for the uneven receptor distribution, we speculate that in other systems in which receptors are distributed more uniformly, ratiometric sensing would still be beneficial as a gradient amplification mechanism. Interactions between mammalian RGS proteins and specific GPCRs, analogous to the Sst2–Ste2 interaction in yeast, have been identified in many contexts (Georgoussi et al., 2006; Ghil et al., 2014; Neitzel and Hepler, 2006). Indeed, one such interaction involved CXC chemokine receptor 2 (CXCR2), which mediates chemotactic responses in leukocytes (Snow et al., 1998). It will be interesting to determine whether other GPCRs exploit ratiometric sensing to sharpen gradient detection.

## **2.6. Methods**

### **2.6.1. Particle-based simulations of ratiometric and nonratiometric gradient sensing**

Simulations of the ratiometric and nonratiometric models were performed using the Smoldyn software (v2.56) on Mac (3.4 GHz Intel processor) and Linux systems (2.50 GHz and 2.30 GHz Intel processors, Longleaf cluster at UNC Chapel Hill, Chapel Hill, NC, USA) (Andrews and Bray, 2004; Andrews et al., 2010). The main components of the code are publicly available at <https://github.com/mikepab/ratiometric-gpcr-particle-sims>. Unless otherwise noted, the simulations were performed using the following conditions: 1) 10,000 receptor molecules and 2,500 G proteins diffusing as point particles on a sphere with diameter 5  $\mu\text{m}$ ; 2) the G-protein diffusion coefficient was  $D = 0.002 \mu\text{m}^2/\text{s}$ , and receptors were not allowed to diffuse (but see section 2.6.4 Particle-based simulations of receptor gradient degradation); and 3) for second-order reactions, the lambda-rho algorithm, with a fixed

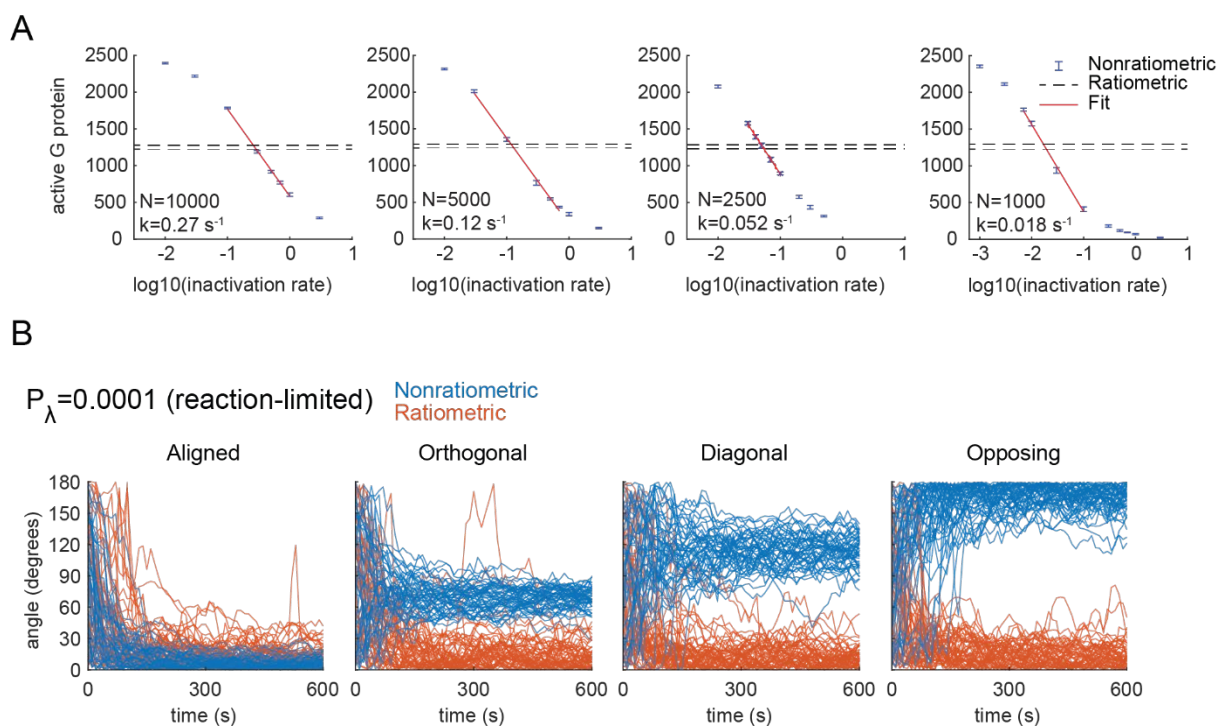
reactive radius ( $\rho = 4$  nm) and fixed reaction probability ( $P_\lambda = 1$  per simulation step), was used to compute rate constants. A reaction probability of 1 results in diffusion-limited reactions. We also studied reaction-limited versions of our models and found similar results (Figure 2.4B). The simulation time step was set to 100 ms so that the root-mean-squared displacements were below the reactive radius. The ratiometric models had a bimolecular G-protein inactivation reaction dependent upon inactive (pheromone-free) receptor, while the nonratiometric models had a unimolecular G-protein inactivation reaction that occurred with a single rate constant throughout the cell.

### 2.6.2. Establishing receptor density and activity gradients

Receptor density and activity gradients were established prior to performing simulations using inverse transform sampling. A desired gradient (receptor or activity) was used to produce a probability distribution as a function of the spatial coordinates. A random number  $P_i \sim \text{Unif}(0,1)$  was drawn for each receptor with proposed coordinates  $(x_i, y_i, z_i)$ , and if  $P_i < P(x_i, y_i, z_i)$ , a receptor was placed at the specified location (for density gradients) or was activated (for activity gradients).

### 2.6.3. Calibrating G-protein inactivation rates

We determined inactivation rates for the nonratiometric model that produced active G protein equivalent to the inactivation rates specified for the ratiometric model (Figure 2.4A). Simulation-based calibration was used to determine these first-order rates rather than analytic equations for relating microscopic reaction probabilities and macroscopic rates because such equations to relate the two quantities can break down on membranes in the diffusion limit (Mahmutovic et al., 2012; Yogurtcu and Johnson, 2015). For consistency, the same calibration process was done for the reaction-limited versions of our simulations.



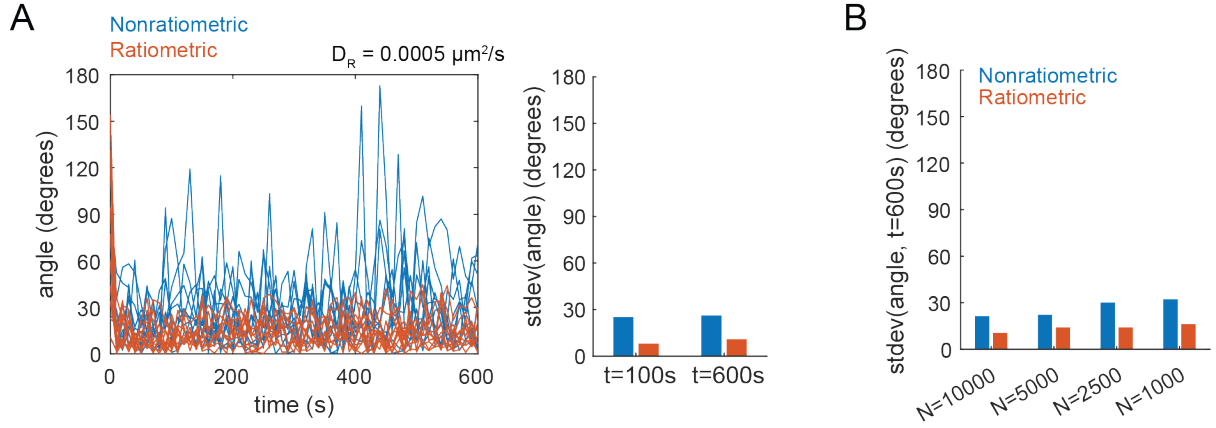
**Figure 2.4. Calibration of G-protein inactivation rates for model comparison, and effect of diffusion-limited versus reaction-limited regimes.**

(A) G-protein inactivation rate constant calibration, relating the nonratiometric and ratiometric models. The results shown are for the mean of 10 simulations for each condition, and the error bars represent  $\pm 1$  SD. Changing the number of receptor molecules (N) requires recalibration of the inactivation rate in the nonratiometric model. (B) Effect of decreasing the reaction rates to a reaction-limited regime ( $P_\lambda = 0.0001$  per time step). The corresponding nonratiometric G-protein inactivation rate was  $k = 0.0031 \text{ s}^{-1}$ . The results shown are for 50 realizations of each model. Although it now takes longer for simulations to reach steady state, once at steady state, the G-protein distributions are similar to those in the diffusion-limited scenario.

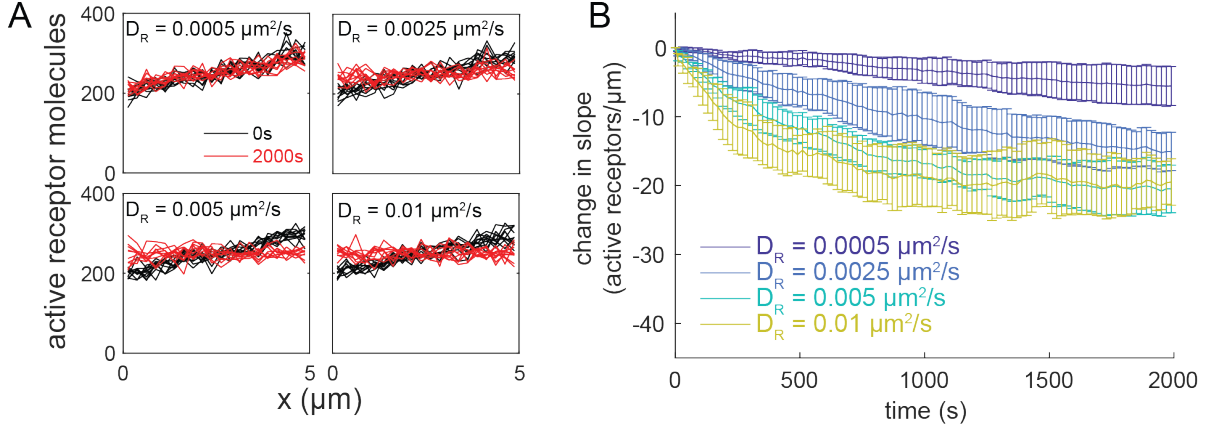
#### 2.6.4. Particle-based simulations of receptor gradient degradation

Neither the ratiometric nor nonratiometric simulations exhibited noticeable loss in gradient sensing capability when the receptor diffusion was increased from  $D = 0$  to  $D = 0.0005 \mu\text{m}^2/\text{s}$  (Figure 2.5A), leading us to question whether the receptor gradient was actually degraded by diffusion over the 10-min timescale of interest. To test this, we removed the G proteins from the simulations to reduce computational costs and varied the receptor diffusivity in extended simulations (2,000 seconds, or  $>30$  min) with a 40% to 60% receptor activity gradient and no density gradient (Figure 2.6A). The active

receptor gradient was measured by linear regression of the number of molecules detected in 250 nm bins along the direction of the initial gradient (Figure 2.6B).



**Figure 2.5. Robustness of simulation results to varying receptor abundance and diffusion.** (A) Accuracy of G-protein activity gradients for the nonratiometric (blue) and ratiometric (orange) models with uniform receptor density, as in Figure 2.3B but allowing receptor diffusion at  $D = 0.0005 \mu\text{m}^2/\text{s}$ . Left: illustrative simulation with measurements every 10 seconds. Right: Variability in orientation angle from 10 simulations of each model, at  $t = 100 \text{ s}$  and  $600 \text{ s}$  snapshots (SD). (B) Effect of decreasing receptor abundance. Variability in orientation angle from 50 simulations of each condition.



**Figure 2.6. Effects of receptor diffusion in particle simulations.** (A) Snapshots of the active receptor gradient at  $t = 0$  (black) and  $2,000 \text{ s}$  (red) for different values of the diffusion coefficient. Each curve represents a histogram with 250 nm bins derived from a single simulation. (B) Decay of the active receptor gradient as measured by the slopes of linear regressions fitted to the data in (A). The results show the mean of 10 realizations  $\pm 1 \text{ SD}$  for the four diffusion coefficients tested.

## REFERENCES

- Alvarez, L., Friedrich, B.M., Gompper, G., and Kaupp, U.B. (2014). The computational sperm cell. *Trends Cell Biol.* 24, 198–207.
- Andrews, S.S., and Bray, D. (2004). Stochastic simulation of chemical reactions with spatial resolution and single molecule detail. *Phys. Biol.* 1, 137–151.
- Andrews, S.S., Addy, N.J., Brent, R., and Arkin, A.P. (2010). Detailed simulations of cell biology with Smoldyn 2.1. *PLoS Comput. Biol.* 6, e1000705.
- Arkowitz, R.A. (2009). Chemical gradients and chemotropism in yeast. *Cold Spring Harb. Perspect. Biol.* 1, a001958.
- Ayscough, K.R., and Drubin, D.G. (1998). A role for the yeast actin cytoskeleton in pheromone receptor clustering and signalling. *Curr. Biol.* 8, 927–930.
- Ballon, D.R., Flanary, P.L., Gladue, D.P., Konopka, J.B., Dohlman, H.G., and Thorner, J. (2006). DEP-domain-mediated regulation of GPCR signaling responses. *Cell* 126, 1079–1093.
- Berg, H.C., and Purcell, E.M. (1977). Physics of chemoreception. *Biophys. J.* 20, 193–219.
- Bhattacharjee, N., and Folch, A. (2017). Large-scale microfluidic gradient arrays reveal axon guidance behaviors in hippocampal neurons. *Microsyst. Nanoeng.* 3, 17003.
- Bush, A., Vasen, G., Constantinou, A., Dunayevich, P., Patop, I.L., Blaustein, M., and Colman-Lerner, A. (2016). Yeast GPCR signaling reflects the fraction of occupied receptors, not the number. *Mol. Syst. Biol.* 12, 898.
- Chiou, J.-G., Balasubramanian, M.K., and Lew, D.J. (2017). Cell polarity in yeast. *Annu. Rev. Cell Dev. Biol.* 33, 77–101.
- Chou, C.-S., Nie, Q., and Yi, T.-M. (2008). Modeling robustness tradeoffs in yeast cell polarization induced by spatial gradients. *PLoS One* 3, e3103.
- Dyer, J.M., Savage, N.S., Jin, M., Zyla, T.R., Elston, T.C., and Lew, D.J. (2013). Tracking shallow chemical gradients by actin-driven wandering of the polarization site. *Curr. Biol.* 23, 32–41.
- Georgoussi, Z., Leontiadis, L., Mazarakou, G., Merkouris, M., Hyde, K., and Hamm, H. (2006). Selective interactions between G protein subunits and RGS4 with the C-terminal domains of the mu- and delta-opioid receptors regulate opioid receptor signaling. *Cell Signal.* 18, 771–782.
- Ghil, S., McCoy, K.L., and Hepler, J.R. (2014). Regulator of G protein signaling 2 (RGS2) and RGS4 form distinct G protein-dependent complexes with protease activated-receptor 1 (PAR1) in live cells. *PLoS One* 9, e95355.
- Hegemann, B., and Peter, M. (2017). Local sampling paints a global picture: Local concentration measurements sense direction in complex chemical gradients. *Bioessays* 39.

- Hegemann, B., Unger, M., Lee, S.S., Stoffel-Studer, I., van den Heuvel, J., Pelet, S., Koepl, H., and Peter, M. (2015). A cellular system for spatial signal decoding in chemical gradients. *Dev. Cell* 35, 458–470.
- Henderson, N.T., Pablo, M., Ghose, D., Clark-Cotton, M.R., Zyla, T.R., Nolen, J., Elston, T.C., and Lew, D.J. (2019). Ratiometric GPCR signaling enables directional sensing in yeast. *PLoS Biol.* 17, e3000484.
- Insall, R. (2013). The interaction between pseudopods and extracellular signalling during chemotaxis and directed migration. *Curr. Opin. Cell Biol.* 25, 526–531.
- Jenness, D.D., and Spatrick, P. (1986). Down regulation of the alpha-factor pheromone receptor in *S. cerevisiae*. *Cell* 46, 345–353.
- Johnson, J.M., Jin, M., and Lew, D.J. (2011). Symmetry breaking and the establishment of cell polarity in budding yeast. *Curr. Opin. Genet. Dev.* 21, 740–746.
- Kelley, J.B., Dixit, G., Sheetz, J.B., Venkatapurapu, S.P., Elston, T.C., and Dohlman, H.G. (2015). RGS proteins and septins cooperate to promote chemotropism by regulating polar cap mobility. *Curr. Biol.* 25, 275–285.
- Lakhani, V., and Elston, T.C. (2017). Testing the limits of gradient sensing. *PLoS Comput. Biol.* 13, e1005386.
- Mahmutovic, A., Fange, D., Berg, O.G., and Elf, J. (2012). Lost in presumption: stochastic reactions in spatial models. *Nat. Methods* 9, 1163–1166.
- McClure, A.W., Minakova, M., Dyer, J.M., Zyla, T.R., Elston, T.C., and Lew, D.J. (2015). Role of polarized G protein signaling in tracking pheromone gradients. *Dev. Cell* 35, 471–482.
- Moore, T.I., Chou, C.-S., Nie, Q., Jeon, N.L., and Yi, T.-M. (2008). Robust spatial sensing of mating pheromone gradients by yeast cells. *PLoS One* 3, e3865.
- Neitzel, K.L., and Hepler, J.R. (2006). Cellular mechanisms that determine selective RGS protein regulation of G protein-coupled receptor signaling. *Semin. Cell Dev. Biol.* 17, 383–389.
- Rappel, W.-J., and Edelstein-Keshet, L. (2017). Mechanisms of cell polarization. *Current Opinion in Systems Biology* 3, 43–53.
- Savage, N.S., Layton, A.T., and Lew, D.J. (2012). Mechanistic mathematical model of polarity in yeast. *Mol. Biol. Cell* 23, 1998–2013.
- Segall, J.E. (1993). Polarization of yeast cells in spatial gradients of alpha mating factor. *Proc. Natl. Acad. Sci. USA* 90, 8332–8336.
- Snow, B.E., Hall, R.A., Krumins, A.M., Brothers, G.M., Bouchard, D., Brothers, C.A., Chung, S., Mangion, J., Gilman, A.G., Lefkowitz, R.J., et al. (1998). GTPase activating specificity of RGS12 and binding specificity of an alternatively spliced PDZ (PSD-95/Dlg/ZO-1) domain. *J. Biol. Chem.* 273, 17749–17755.

- Strickfaden, S.C., and Pryciak, P.M. (2008). Distinct roles for two Galpha-Gbeta interfaces in cell polarity control by a yeast heterotrimeric G protein. *Mol. Biol. Cell* 19, 181–197.
- Suchkov, D.V., DeFlorio, R., Draper, E., Ismael, A., Sukumar, M., Arkowitz, R., and Stone, D.E. (2010). Polarization of the yeast pheromone receptor requires its internalization but not actin-dependent secretion. *Mol. Biol. Cell* 21, 1737–1752.
- Swaney, K.F., Huang, C.-H., and Devreotes, P.N. (2010). Eukaryotic chemotaxis: a network of signaling pathways controls motility, directional sensing, and polarity. *Annu. Rev. Biophys.* 39, 265–289.
- von Philipsborn, A., and Bastmeyer, M. (2007). Mechanisms of gradient detection: a comparison of axon pathfinding with eukaryotic cell migration. *Int Rev Cytol* 263, 1–62.
- Wang, Y., and Dohlman, H.G. (2004). Pheromone signaling mechanisms in yeast: a prototypical sex machine. *Science* 306, 1508–1509.
- Yogurtcu, O.N., and Johnson, M.E. (2015). Theory of bi-molecular association dynamics in 2D for accurate model and experimental parameterization of binding rates. *J. Chem. Phys.* 143, 084117.



## CHAPTER 3: YEAST POLARIZATION TOWARD MATING PARTNERS<sup>3</sup>

### Overview

As described in Chapters 1 and 2, yeast cells must accurately sense the orientation of extracellular pheromone gradients to polarize and grow towards a mate. Gradient sensing is important in many other cells, including neurons, social amoebae, sperm, and neutrophils (Alvarez et al., 2014; Bhattacharjee and Folch, 2017; Rappel and Edelstein-Keshet, 2017; Swaney et al., 2010; von Philipsborn and Bastmeyer, 2007). The events converting extracellular gradients into a directed response are perhaps best understood in yeast, which sense pheromone gradients in order to polarize and chemotropically grow towards mating partners.

Sensing shallow gradients can be challenging for small cells such as yeast, because they may experience little difference in ligand concentration on the up-gradient and down-gradient sides of the cell (Berg and Purcell, 1977). In artificial, calibrated pheromone gradients as weak as 0.1 nM/ $\mu\text{m}$ , yeast polarization is essentially random (Hegemann et al., 2015; Lakhani and Elston, 2017; Moore et al., 2008; Segall, 1993). Polarity establishment in these shallow gradients is often corrected by actin-dependent re-localization of the polarity site, or “wandering”, orienting the cell up-gradient (Dyer et al., 2013; Hegemann and Peter, 2017; Kelley et al., 2015). The situation is different within yeast mating mixtures, where pheromone gradients are generated by yeast rather than artificially controlled. In these more natural conditions, though the initial polarity site is still sometimes incorrectly oriented, it tends to be biased in the correct direction. Importantly, a different error correction process takes place instead of wandering

---

<sup>3</sup>Parts of Section 3.2 of this chapter were done in collaboration with an undergraduate student, Kaiyun Guan. Parts of Section 3.3 of this chapter are part of a submission to the journal PLoS Computational Biology. (*In preparation*. Ramirez SA, Pablo M, Ghose D, Burk S, Lew DJ, Elston TC. *A novel stochastic simulation approach enables exploration of mechanisms to regulate polarization dynamics*). Section 3.5 was motivated by conversations with Prof. Daniel J Lew.

polarity. “Indecisive” polarity occurs, in which polarity factors cluster into weak, erratic clusters that rapidly fluctuate over the surface of the cell (Henderson et al., 2019).

After a variable period of indecisive behavior, a strong, stable polarity site emerges, oriented towards a mate. The duration of the indecisive behavior is shorter if the initial polarity site was accurately placed, suggesting that the indecisive behavior is indeed related to error correction. However, it is unclear whether the presence of weakly clustered, erratically moving polarity factors are functionally important in establishing an accurate stable polarity site, or if they are simply a side effect of other signaling processes. Furthermore, it is unclear why a different error correction process might occur in artificial calibrated gradients versus mating mixtures.

Here I present preliminary investigations on how weakly clustered polarity factors could contribute to more accurate polarity establishment. We have hypothesized that indecisive polarity functions as a way to improve the initial guess that the cell makes during the formation of a stable polarity site. Specifically, that the indecisive polarity phenotype allows the cell to accumulate active Cdc42 at the cell surface, biased upgradient, without yet committing itself to the polarity program. Our approach is to take the Cdc42-centric polarity model in Chapter 1, extend it to model indecisive polarity in 3D, and couple it to the pheromone-sensing model in Chapter 2. We find that simulations of stable polarity site were insensitive to shallow pheromone gradients, whereas simulations of indecisive polarity allowed biased Cdc42 activation along shallower gradients. Future work includes simulating the transition from indecisive to stable polarity and assessing changes in gradient perception. I also present here simulations of the pheromone concentration profile experienced by a cell near a mating partner. Our results suggest that the gradients experienced by cells near mates are highly nonlinear and dynamic, providing a clue to why the error correction process in a mating mixture would be different from a static linear gradient.

### **3.1. Introduction**

As described in Chapter 2, yeast use GPCRs (Ste2 and Ste3) to decode gradients of pheromone ( $\alpha$ -factor and **a**-factor) generated by cells of the opposite mating type (MAT**a** and MAT $\alpha$ ) (Wang and

Dohlman, 2004). Once bound to pheromone, the GPCRs activate heterotrimeric G proteins, generating GTP-G $\alpha$  and G $\beta\gamma$ . Pheromone-free Ste2 can inactivate GTP-G $\alpha$  via the negative regulator RGS2, driving re-association of the heterotrimer. Free G $\beta\gamma$  recruits two key scaffold proteins, Ste5 and Far1, from the cell interior to the membrane (Butty et al., 1998; Nern and Arkowitz, 1999; Pryciak and Huntress, 1998). Ste5 recruitment leads to activation of a mitogen-activated protein kinase cascade, which induces transcription of mating-related genes, arrests the cell cycle in G1 in preparation for mating, and promotes cytoskeletal polarization (Pryciak and Huntress, 1998). However, our primary focus is on Far1, which provides spatial information to Cdc42, the master regulator of cell polarity in yeast. Recruitment of Far1 orients the cytoskeleton towards the mating partner by providing spatial information to the conserved Rho-family GTPase Cdc42, which is the master regulator of cell polarity in yeast (Bi and Park, 2012; Butty et al., 1998; Nern and Arkowitz, 1998, 1999). Specifically, Far1 can directly bind the guanine nucleotide exchange factor Cdc24 (Butty et al., 1998; Nern and Arkowitz, 1998, 1999; Wiget et al., 2004). The G $\beta\gamma$ -Far1-Cdc24 complex is thought to enhance Cdc42 activation at sites with elevated levels of G $\beta\gamma$ . Though mutations that disrupt Far1-Cdc24 binding do not abolish polarity establishment, they do disrupt proper orientation of polarity up the pheromone gradient (Butty et al., 1998; Nern and Arkowitz, 1999). Thus, Far1 provides a direct spatial connection between pheromone receptor-pheromone binding and downstream Cdc42 activation, allowing yeast to sense pheromone gradients to find partners. Weak gradients, however, may be ignored in favor of noise-driven spontaneous polarity establishment, which was described in Chapter 1.

Studies with artificial, calibrated gradients have suggested that the shallowest gradients yeast can respond to are in the range of 0.1 nM/ $\mu$ m to 0.5 nM/ $\mu$ m (Hegemann et al., 2015; Moore et al., 2008; Segall, 1993). Yeast cells in such shallow gradients polarize essentially randomly, then exhibit subsequent error correction (Dyer et al., 2013; Hegemann and Peter, 2017; Kelley et al., 2015). Constrained by estimates of the binding affinity between pheromone ( $\alpha$ -factor) and its receptor (Ste2), simulations of these gradients have predicted that the difference in average occupancy of receptors at the front versus back of the cell is as little as 50 molecules. This is only 50 receptors out of thousands of

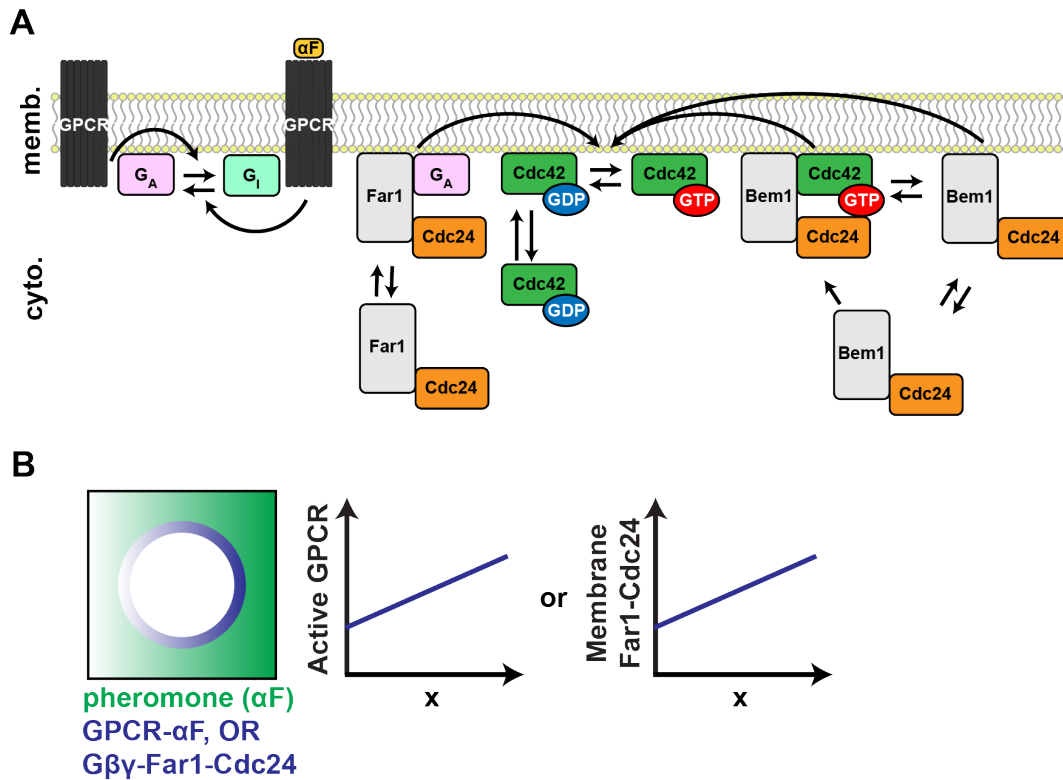
receptors on the cell surface, suggesting the process is highly noisy, and the fluctuations caused by reaction and diffusion processes imply that cells must somehow integrate this information over time (Lakhani and Elston, 2017). Indeed, polarity establishment in these shallow artificial gradients is often corrected over time through actin-dependent wandering of the polarity site, orienting the cell up-gradient (Dyer et al., 2013; Hegemann and Peter, 2017; Kelley et al., 2015).

Studies using mating mixtures of MAT $\alpha$  and MAT $\alpha$  cells are more physiologically relevant, but the strength and shape of the pheromone gradient is uncertain. In mating mixtures, the initial polarity site is often incorrectly oriented, but it is biased in the correct direction: approximately 40% of cells polarize within 30° of the mating partner, compared to an expected 17% under a random process (Henderson et al., 2019). Furthermore, yeast in these mating mixtures do not exhibit patch wandering as an error correction mechanism. Instead, an indecisive polarity phenotype is observed, where erratic fluctuations of weakly clustered polarity factors rapidly form and dissipate across the surface of the cell. This indecisive phase leads to the emergence of a single strong, stable polarity site oriented towards a mate, and is shorter if the initial polarity site was accurately placed. However, are the weak polarity clusters observed during the indecisive phase functionally important? It is possible that the indecisive phenotype is merely a side-effect of other signaling processes. The transition to stable polarization correlates with the accumulation of MAPK activity, so perhaps accurate stable polarization merely relies on detecting sufficient pheromone over a prolonged period. Alternatively, indecisive polarization may allow the simple accumulation of polarity factors upgradient prior to committing to stable polarization, allowing cells to make a better-informed guess.

### **3.2. Stable polarization is insensitive to shallow gradients**

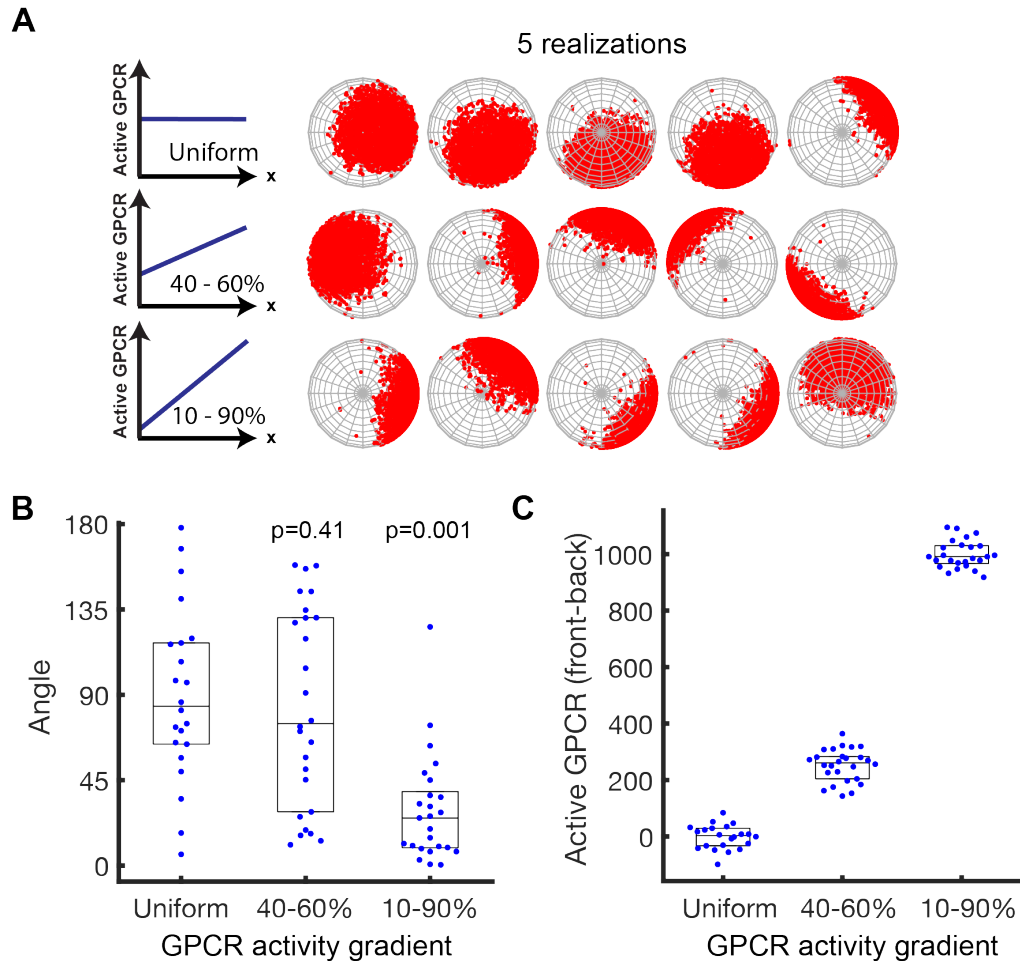
In weak pheromone gradients, initial polarization essentially follows a random process. This is likely because the polarity circuitry amplifies molecular noise. To demonstrate the problem, we conducted particle-based simulations of polarity establishment in a stable gradient of pheromone-bound receptors on a spherical yeast cell (see Methods). To link the polarity machinery to the GPCR signaling model

described in Chapter 2, we assume that active G protein recruits from a pool of intracellular Far1-Cdc24. At the membrane, both Far1-Cdc24 and Bem1-Cdc24 can activate Cdc42. Signaling in the polarity model works as in Chapter 1: both Cdc42 and the activator Bem1-Cdc24 can exchange between the membrane and cytoplasm. Bem1-Cdc24 can bind to active Cdc42, creating a positive feedback loop. Altogether, the Far1-Cdc24 complex allows activation in the direction of the pheromone gradient, while the Bem1-Cdc24 positive feedback loop allows a polarity site to form. We omit the dynamics of G protein subunit dissociation and association, the dynamics of binding between Far1 and Cdc24, and the dynamics of binding between Bem1 and Cdc24. The model is summarized in Figure 3.1.



**Figure 3.1. GPCR signaling couples pheromone sensing to the Cdc42-mediated polarity machinery.** (A) Model signaling diagram. Pheromone-bound GPCRs convert inactive G proteins to active G proteins while pheromone-free GPCRs catalyze the opposite reaction. Activated G protein recruits Far1-Cdc24 to the membrane, where it can catalyze the activation of Cdc42-GDP. Cdc42-GDP can spontaneously cycle between the cytoplasm and membrane, but can only become activated at the membrane. Bem1-Cdc24 can also spontaneously translocate to the plasma membrane. Once on the membrane, it can activate Cdc42-GDP. Both cytosolic and membrane-associated Bem1-Cdc24 can bind to active Cdc42-GTP, creating a positive feedback loop. (B) Cartoon of an extracellular pheromone gradient. Instead of directly simulating a pheromone gradient, we simulate either a pre-formed gradient of GPCR activity (shown here) or a pre-formed gradient of G protein-Far1-Cdc24 activity (not shown here).

Simulating this model suggested that, even with optimistic estimates of the parameters of GPCR, G-protein, and Far1-Cdc24 rate constants, a stable polarity site can only be guided by steep gradients of extracellular pheromone (see Table 3.1 for parameters). Simulated yeast only formed correctly-oriented polarity sites when approximately 1,000 pheromone-bound receptors were upgradient (Figure 3.2). This is consistent with experimental reports that yeast polarize in random directions when in shallow, calibrated gradients (Hegemann et al., 2015; Moore et al., 2008; Segall, 1993) and reflects the dominance of noise-driven spontaneous polarization over gradient-guided polarization. However, the parameters result in very large polarity sites, so re-parameterization may be needed to find more reasonably-sized patches.



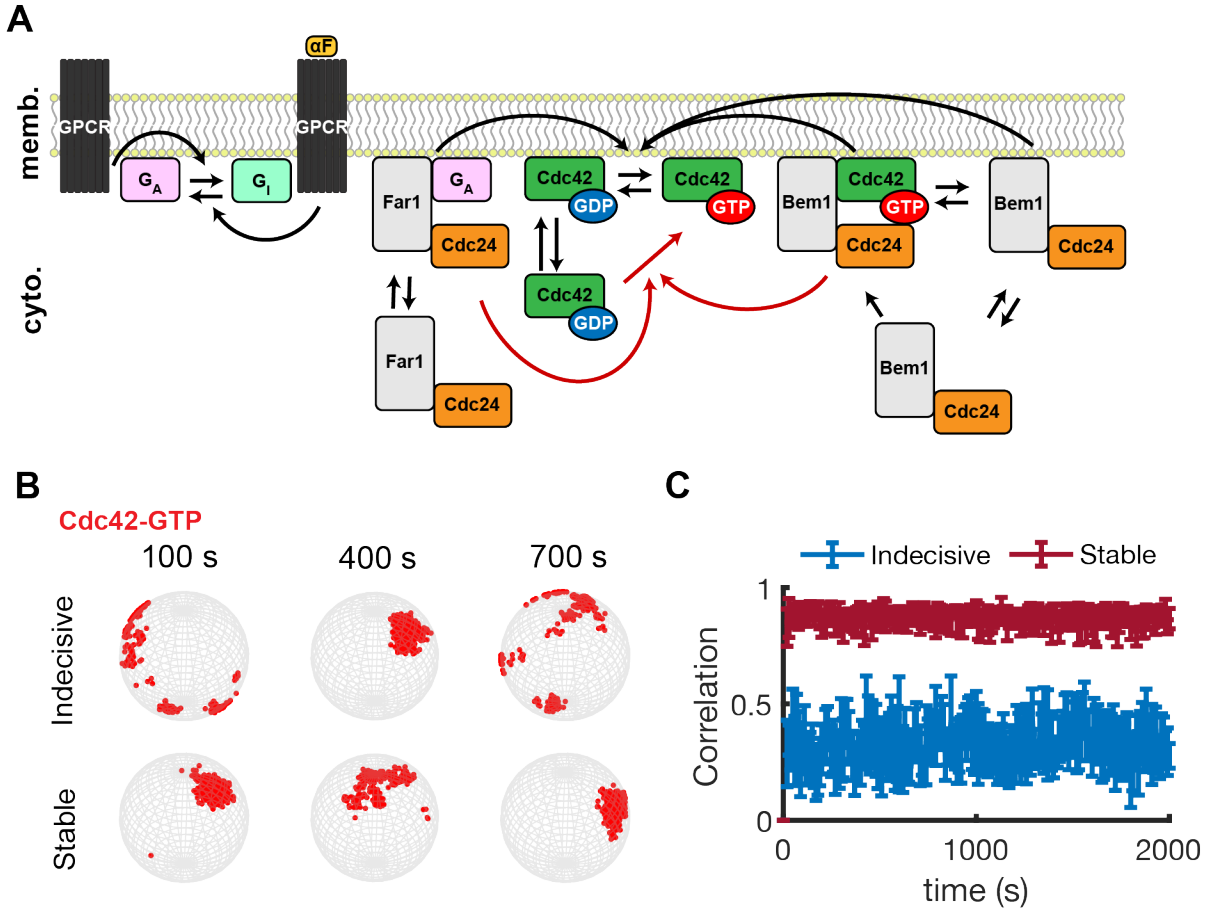
**Figure 3.2.** Stable polarization is insensitive to shallow gradients.

(A) Representative simulation results using different active GPCR gradients (i.e. pheromone gradients) (B) The angle of the patch with respect to the pheromone gradient. *P*-values obtained from unpaired *t*-tests against the uniform condition. (C) The number of active GPCRs at the front vs. back of the cell. Uniform,  $n=22$ ; 40-60%,  $n=26$ ; 10-90%,  $n=25$  realizations. Boxplots show 25<sup>th</sup>, 50<sup>th</sup>, and 75<sup>th</sup> percentiles.

In spontaneous polarization, stochastic association of Cdc42 and Bem1-Cdc24 to the plasma membrane results in tiny variations in the activity of Cdc42 across the plasma membrane. Positive feedback amplifies these variations, which grow and compete for polarity factors until a single, dominant patch remains. Far1-Cdc24 does not participate in this feedback loop. Weak gradients of  $G\beta\gamma$ -Far1-Cdc24 (i.e. weak gradients of pheromone) are thus likely limited to influencing the initial stages of polarity establishment. This may be problematic, as illustrated by our simulations. A possible solution would be to use indecisive polarization as an initial form of polarity establishment: in this scenario, the pheromone-mediated gradient of Cdc42 activation may be able to compete with Bem1-mediated positive feedback.

### **3.3. Indecisive polarization is sensitive to shallow gradients**

To model indecisive polarity, we allowed the direct activation of cytosolic Cdc42 by both the Bem1-Cdc24-Cdc42-GTP complex and the  $G\beta\gamma$ -Far1-Cdc24 complex. This decision was motivated by extensive parameter exploration performed by a much more efficient stochastic simulation algorithm, described in a manuscript in preparation<sup>3</sup>. Biochemically, it is unclear how this additional reaction would function. While MAPK activity appears to set a threshold for indecisive and committed polarization (Henderson et al., 2019), we do not know how it would introduce direct activation of cytosolic Cdc42. Perhaps the guanine nucleotide dissociation inhibitor (GDI), which regulates Cdc42 exchange between the membrane and cytoplasm and is not explicitly included in our models, is involved. Nonetheless, simulations with the added reaction are able to recapitulate the indecisive polarity phenotype (Figure 3.3, see Table 3.1 for parameters). These results are purely driven by changes in the polarity network, and are independent of the pheromone-sensing side of the network. Indeed, the simulations in Figure 3.3 were performed in the absence of any GPCR, G protein, or Far1-Cdc24.

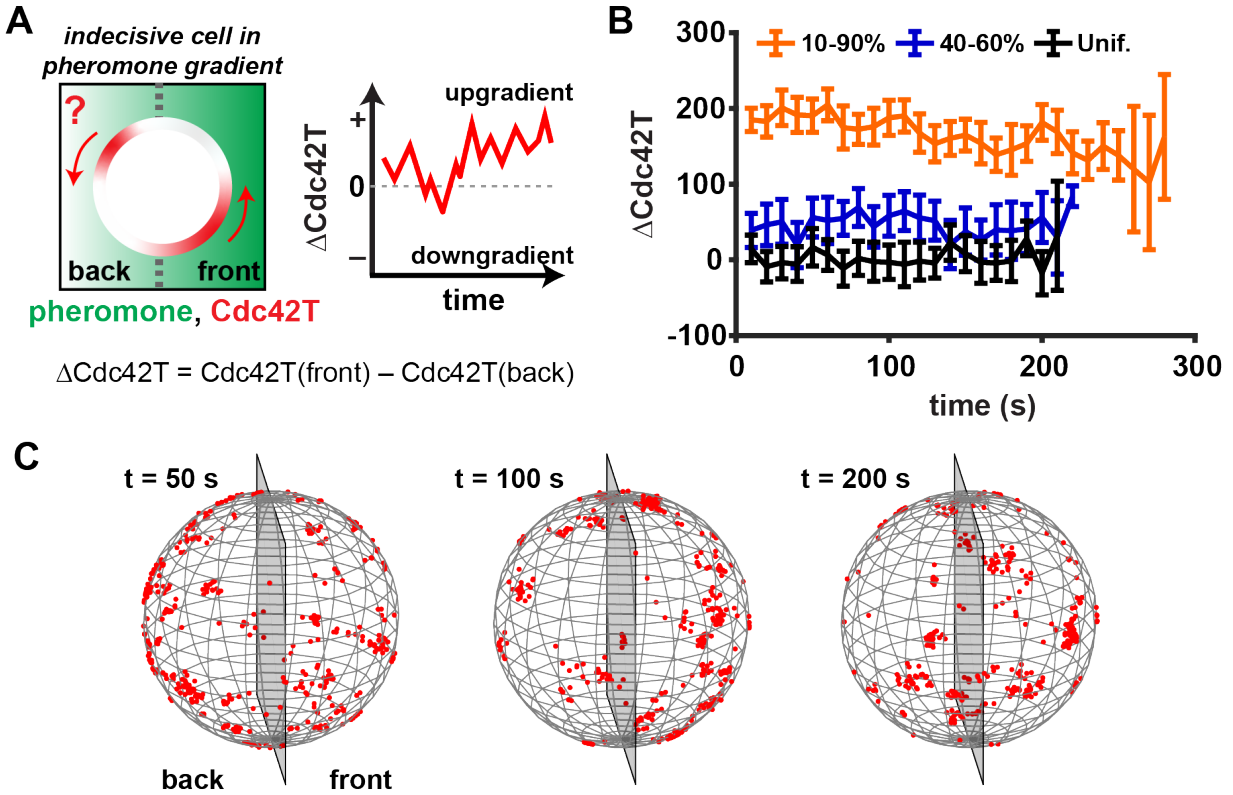


**Figure 3.3. Stable polarization is insensitive to shallow gradients.**

(A) Signaling network diagram as in Figure 3.1, with added reaction arrows shown in red. (B) Snapshots of dynamic and stable polarity establishment with the new model. (C) Quantification of the movement of the patch by frame-to-frame correlation of local active Cdc42 (mean  $\pm$  1 s.d.,  $n=5$  per condition). High correlation indicates the patch is not moving.

Next, we asked whether adding the pheromone-sensing side of the network would bias the active Cdc42 distribution upgradient. The erratic, poorly-localized nature of Cdc42-GTP under these conditions means we cannot calculate an angle between the polarity site and the gradient, but we can readily check the difference in Cdc42-GTP abundance between the front and back of the cell. Quantifying the indecisive simulations in this way shows that the weak polarity clusters can be biased upgradient versus downgradient (Figure 3.4). Problematically, we observed that introducing the pheromone-sensing component of the signaling pathway lead to even weaker, more erratic polarity sites than in the pheromone-free indecisive model.





**Figure 3.4. The indecisive polarity phenotype is sensitive to shallow gradients.**

(A) Modeling strategy. A cell exhibiting indecisive Cdc42-GTP polarization (red) is simulated in an approximated pheromone gradient (green). Pheromone is approximated by setting up a fixed gradient of pheromone-bound (active) and pheromone-free (inactive) GPCR. The number of Cdc42-GTP molecules at the front vs. back of the cell ( $\Delta\text{Cdc42T}$ ) is measured over time to assess gradient sensing. (B) Quantification of  $\Delta\text{Cdc42T}$  at the front vs. back of the simulated cell for varying GPCR gradient slopes. Error bars represent 95% confidence intervals. (C) Representative snapshots of a cell simulated in the 40-60% GPCR activity gradient, with a plane separating the front and back.

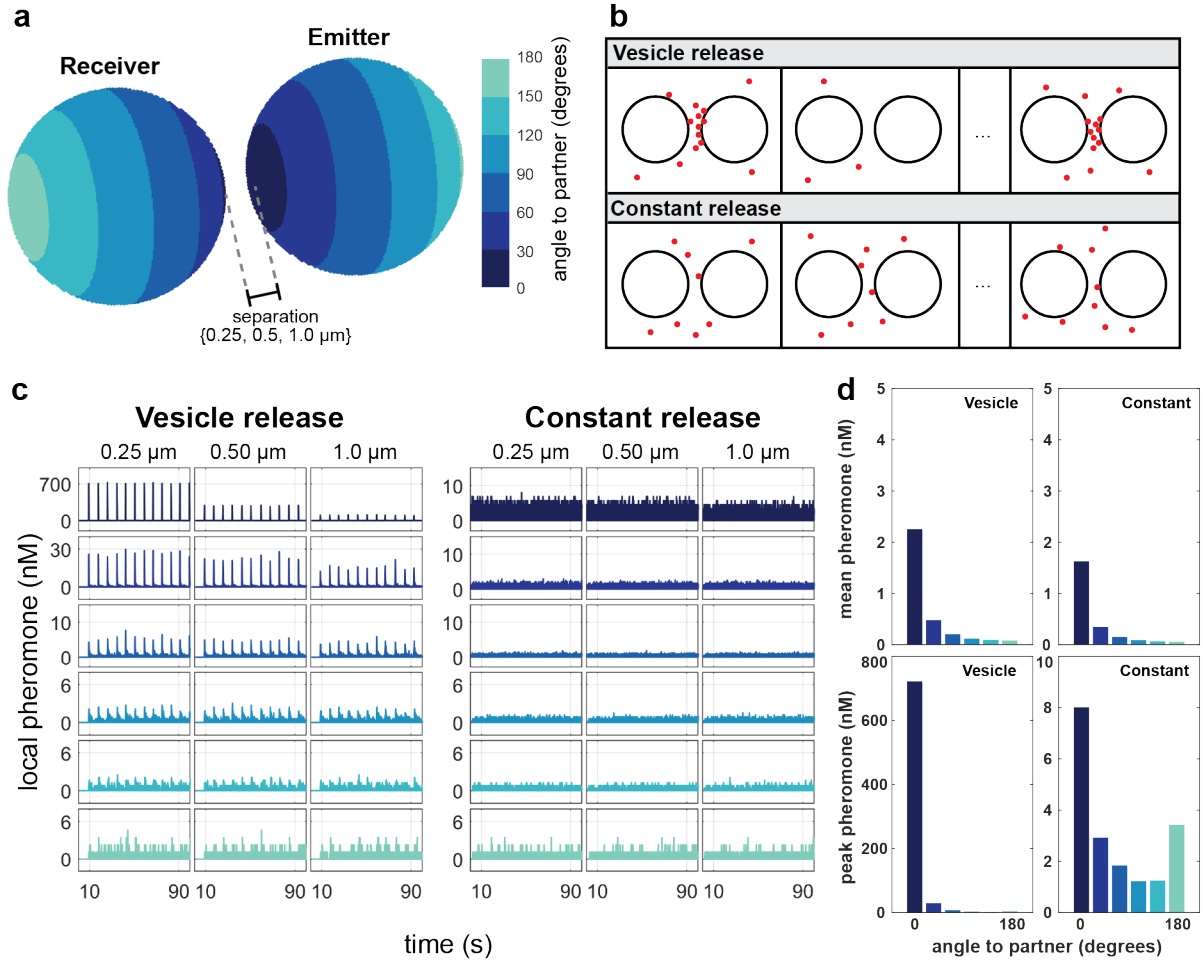
The pheromone-free behavior observed at 400 seconds in Figure 3.3C is clearly different from the pheromone-biased behavior in Figure 3.4C. Somehow adding the pheromone signaling part of the pathway further destabilized the indecisive polarity phenotype, resulting in nearly global activation of Cdc42. Therefore, optimization of model parameters is necessary to recapitulate indecisive polarization when combined with the pheromone pathway. This might be solved by strengthening the positive feedback from Bem1-Cdc24, or by weakening the activation from the Far1-Cdc24.

### **3.4. Can polarization during the indecisive phase guide a stable polarity site?**

After parameter optimization, the next step for this work is to determine whether an initial phase of indecisive polarity, followed by a transition into stable polarity, could improve the accuracy of the final Cdc42-GTP patch with respect to the gradient. This might be done by initiating simulations in the indecisive regime, then slowly tuning reaction parameters (i.e. strengthening the positive feedback) to shift the model into the stable polarity regime.

### **3.5. Pheromone gradients under mating conditions may be highly non-linear**

The work in the preceding sections assumed that gradients of active GPCR molecules or of G $\beta\gamma$ -Far1-Cdc24 complexes were linear, to approximate the nature of artificial calibrated pheromone gradients. However, it is unclear what the shape of the pheromone gradient actually is for two proximal mating partners. In particular, yeast pheromone secretion may occur in localized bursts (Martin, 2019; Merlini et al., 2016). The pheromone ( $\alpha$ -factor) transporter protein Ste6 is found both in vesicles and concentrated towards the polarity site (Kuchler et al., 1993; Michaelis, 1993), suggesting that it is trafficked to the membrane along vesicles during polarization. It is possible that pheromone is also inside of Ste6-containing vesicles, which could cause temporal spikes in local pheromone abundance. We developed simple particle-based models to investigate the perceived pheromone concentration profile, by only modeling the 3D Brownian diffusion of pheromone between and around two cells: one secreting pheromone (the emitter), and one not (the receiver). We measured local pheromone concentrations around the emitter and the receiver (Figure 3.5). For reference, we also compared the vesicle-release simulation to simulations where pheromone was released deterministically at a corresponding averaged rate.



**Figure 3.5. Pheromone gradients experienced by a simulated emitter and receiver cell.**

**(A,B)** Model schematic. **(A)** The receiver and emitter are separated by 0.25, 0.5, or 1.0  $\mu\text{m}$ . Local concentrations are measured in angular bins (colors). **(B)** Comparison of pheromone production modes. Vesicle release is simulated as bursts of 660 pheromone molecules at a Poisson rate  $k=0.83\text{ s}^{-1}$ , from the emitter surface at the point nearest to the receiver. Constant release is simulated by producing 1 pheromone molecule every 1.8 milliseconds precisely. Over time the two modes have similar average levels of pheromone production. **(C)** Time series of local pheromone concentrations at the receiver. Each column represents a different separation distance, and each row represents a different angular bin (top: nearest, bottom: farthest). Local volumes were calculated based on a 0.25  $\mu\text{m}$  shell off of the surface of the receiver. **(D)** Mean and peak pheromone concentrations observed over time.

The profiles are very clearly nonlinear in the vesicle-release simulations, but the nonlinearity is transient, decaying as the pheromone rapidly diffuses away ( $D = 150\text{ }\mu\text{m}^2/\text{s}$ ). This is very different from a temporally-stable static pheromone gradients often used microfluidic devices. However, this simple model omits pheromone receptors that could bind molecules at the surface, and thus ignores how the kinetics of receptor binding would shape the gradient. The affinity between the receptor Ste2 and  $\alpha$ -factor

is 6 nM, with a relatively slow dissociation rate ( $k_{off} \sim 10^{-2} - 10^{-3} \text{ s}^{-1}$ , corresponding to roughly a half-life of 70 - 700 seconds for a pheromone-GPCR complex), so we anticipate the spikes in the local pheromone concentration (Figure 3.5C, left) to be much broader due to receptor binding/unbinding. If completely averaged over time, the gradient is much less steep, though still nonlinear (Figure 3.5D).

In future work, accounting for receptor-pheromone binding interactions would give us a clearer picture of the pheromone gradient experienced by a pair of mating cells. It would be computationally challenging to explicitly simulate both the external pheromone gradient and receptor binding. Therefore, it would be best to implicitly model the pheromone gradient. The pheromone gradient could be implicitly modeled by calculating a probability distribution  $P(x,y,z,t)$  that  $n$  pheromone molecules are in a particular region of space, and use that to compute a spatially-dependent receptor “activation” rate  $k_{on}(x,y,z)$ . Calibration simulations would be necessary to ensure this rate is consistent with the 6 nM affinity.

### 3.6. Discussion

Yeast polarization towards pheromone in mating mixtures appears qualitatively different from yeast polarization within artificial gradients, suggesting that yeast have underappreciated mechanisms for detecting and decoding natural gradients. This chapter presents preliminary results towards understanding how indecisive polarity establishment, a behavior seen in mating mixtures and not artificial gradients, may contribute towards accurate placement of a polarity site. Models of stable polarization could bias Cdc42 activation accurately when in very strong gradients, while models of indecisive polarization were able to bias Cdc42 activation even in weak gradients. However, improved parameterization of the indecisive polarity model is necessary, as Cdc42 activation became too disperse when the pheromone sensing components of the pathway were added. More broadly, we would want to investigate the sensitivity of our results to variations in the other kinetic parameters.

Simulating pheromone emission towards a nearby mate predict that the pheromone gradient itself is highly nonlinear and dynamic. This argues that artificial gradients are qualitatively different from gradients in mating mixtures, and suggests that strong, transient pheromone signals might be part of the

cause for indecisive polarization. Consideration of pheromone-receptor interactions, which might hold onto pheromone that otherwise rapidly diffuses away, are necessary to more accurately judge the gradient. The pheromone simulations could be modified to including multiple emitting source cells as well as a floor and/or a ceiling in the simulation (mimicking a surface, or glass slides) to further interrogate possible gradient profiles. Finally, considering the effect of dynamic pheromone-receptor interactions would also be important in the polarity models: by using a fixed GPCR activity gradient as an approximation, we have assumed that only the equilibrium behavior was relevant.

### 3.7. Methods

#### 3.7.1. Particle-based simulations of GPCR-coupled polarity establishment

The simulations of polarity establishment in Chapter 1 relied on custom code and approximated a fully 3D cell with a quasi-3D approach. The simulations of GPCR signaling in Chapter 2, however, used the better-optimized Smoldyn simulation software (v2.56-v2.61), in part because support for the lambda-rho algorithm, described in Chapters 1 and 2, was added. Each molecule in the simulation was represented as a point particle subject to Brownian motion. Cytosolic molecules diffused within the sphere, and had rate-dependent membrane association probabilities upon collision with the sphere surface. Membrane-associated molecules diffused along the surface of the sphere. All bimolecular reactions had reaction probabilities defined by  $P = \lambda \Delta t$ , where  $\lambda$  is the rate of reaction after reactant collision, and  $\Delta t$  is the timestep. Parameters for indecisive polarity were found by exploratory analysis using a more efficient stochastic simulation method, detailed in a manuscript in preparation<sup>3</sup>. Pheromone gradients were approximated by a fixed GPCR activation gradient, similar to Chapter 2. Briefly, we assumed that half the receptors are active (pheromone-bound) at the midpoint (i.e. the implicit pheromone concentration is at the  $K_d$  of the receptor). We neglect receptor diffusion as described previously (see Section 2.6.4). G proteins diffusing on the membrane were activated upon encountering an active receptor, and are inactivated upon encountering an inactive receptor. Linear gradients were generated along the x-axis by sampling from a probability distribution:

$$P(x) = \left( \frac{P_{max} - P_{min}}{2R} \right) x + P_{min}$$

where  $P_{max}$  and  $P_{min}$  are the maximum and minimum probability of having an active molecule (or placing a G $\beta\gamma$ -Far1-Cdc24 molecule), and  $R$  is the cell radius.

Parameter	Description	Indecisive Pol. (- pher.)	Stable Pol. (+ pher.)	Indecisive Pol. (+ pher.)
R	Cell radius	2.25675	2.5	2.25675
$\rho$	Reaction radius	0.05	0.05	0.05
$k_{1a}$	BemGEF <sub>c</sub> $\rightarrow$ BemGEF <sub>m</sub>	5	10	5
$k_{1b}$	BemGEF <sub>m</sub> $\rightarrow$ BemGEF <sub>c</sub>	10	10	10
$\lambda_{2a}$	Cdc42D <sub>m</sub> + BemGEF <sub>m</sub> $\rightarrow$ Cdc42T	0.0637	5.3	0.0637
$k_{2b}$	Cdc42T $\rightarrow$ Cdc42D <sub>m</sub>	3	0.32	3
$\lambda_3$	Cdc42D <sub>m</sub> + BemGEF42 $\rightarrow$ Cdc42T	0	15.7	0
$\lambda_{4a}$	BemGEF <sub>m</sub> + Cdc42T $\rightarrow$ BemGEF42	0.00163	8250	0.00163
$k_{4b}$	BemGEF42 $\rightarrow$ BemGEF <sub>m</sub> + Cdc42T	20	10	20
$k_{5a}$	Cdc42D <sub>c</sub> $\rightarrow$ Cdc42D <sub>m</sub>	0.4	36	0.4
$k_{5b}$	Cdc42D <sub>m</sub> $\rightarrow$ Cdc42D <sub>c</sub>	0.65	0.65	0.65
$\lambda_7$	BemGEF <sub>c</sub> + Cdc42T $\rightarrow$ BemGEF42	1436.7	256	1436.7
$\lambda_8$	Cdc42D <sub>c</sub> + BemGEF42 $\rightarrow$ Cdc42T	2873.4	n/a	2873.4
$\lambda_9$	Cdc42D <sub>c</sub> + BemGEF $\rightarrow$ Cdc42T	0	n/a	0
$\lambda_{10}$	Cdc42D <sub>c</sub> + Far1GEFGa $\rightarrow$ Far1GEFGa42	n/a	n/a	0.1
$\lambda_{11}$	Far1GEF + Ga $\rightarrow$ Far1GEFGa	n/a	0.1	100
$\lambda_{12}$	Far1GEFGa $\rightarrow$ Far1GEF + Ga	n/a	10	10
z	Adsorption scaling factor	0.7523	0.01	0.7523
D <sub>m1</sub>	Memb. diffusion coeff. 1	0.01	0.0025	0.01
D <sub>m2</sub>	Memb. diffusion coeff. 2	n/a	0.002	0.002
D <sub>m3</sub>	Memb. diffusion coeff. 3	n/a	0	0
D <sub>c</sub>	Cytosolic diffusion coeff.	15	15	15
N <sub>Cdc42</sub>	Num. Cdc42 molecules	5000	5000	5000
N <sub>BemGEF</sub>	Num. BemGEF molecules	500	2364	500
N <sub>Far1GEF</sub>	Num. Far1GEF molecules	-	1000	500
N <sub>Gprotein</sub>	Num. G protein molecules	-	2500	2500
N <sub>GPCR</sub>	Num. GPCR molecules	-	5000	5000
$\Delta t$	Timestep	0.0001	0.0001	0.0001

**Table 3.1 Model parameters.**

D<sub>m1</sub> applies to Cdc42D<sub>m</sub>, Cdc42T, BemGEF42, BemGEF, Far1GEF<sub>m</sub>, and Far1GEFGa<sub>m</sub>. D<sub>m2</sub> applies to G proteins. D<sub>m3</sub> applies to GPCRs. The membrane association rate constants (time<sup>-1</sup>) were converted to adsorption coefficients (length time<sup>-1</sup>)  $\kappa = zk$  in Smoldyn, with  $z = 0.7523 \mu\text{m}$ .

### 3.7.2. Quantifying patch movement with frame-to-frame correlation

To compare the dynamics of stable and indecisive polarity, we computed frame-to-frame Pearson's correlation coefficients,  $r$ , of the local active Cdc42 concentration on the surface of the cell. High correlations reflected stable behavior, while weak correlations reflected dynamic or erratic behavior. Local concentrations were determined by binning particles on the surface with a spherical triangulation (1280 faces generated by a 3-time subdivision of an icosahedron).

### 3.7.3. Analyzing polarization relative to the gradient

The orientation of stable polarity sites relative to the gradient was quantified using the following:

$$\text{angle} = \arccos\left(\frac{\overrightarrow{\text{COM}}_{xy} \cdot \langle 1,0 \rangle}{\|\overrightarrow{\text{COM}}_{xy}\|}\right)$$

where  $\text{COM}_{xy}$  is the  $xy$  coordinates of the active Cdc42 center of mass. This two-dimensional calculation for the angle is used because surface areas of the spherical sectors for the 2D angles are all equal, while the spherical sectors for the 3D angles are not. This avoids artificial over/under-representation at the poles of the sphere. The orientation of weak, erratic polarity clusters undergoing indecisive polarization was quantified differently because the active Cdc42 center of mass was a poor measure of localization (since, i.e. multiple weak clusters could form). Instead, the difference in active Cdc42 abundance at the front vs. back of the cell was calculated.

### 3.7.4. Simulating pheromone gradients experienced by a mating pair

We set up simulations with two spheres, one denoted the emitter, and one denoted the receiver. The emitter released pheromone from a point on its surface nearest to the receiver. Vesicle-based secretion of pheromone was modeled as Poisson-distributed events ( $k = 0.83 \text{ s}^{-1}$ ) generating 660 pheromone molecules from a point on the emitter surface closest to the receiver. For comparison, steady secretion of pheromone was modeled as the release of 1 molecule every 0.018 seconds from the same location. The two cells were simulated within a spherical domain of 50 micron diameter with absorbing boundary conditions to remove pheromone that diffused far away.

Local pheromone concentrations were computed in volumes defined by 3D angular bins extending  $r_{extend} = 0.25 \mu\text{m}$  beyond the surface of the  $r = 2.5 \mu\text{m}$  receiver sphere. The volume of the region varies as the angle sweeps across the surface. The volumes  $V_i$  defined by six angular bins  $\{0-30^\circ, 30-60^\circ, 60-90^\circ, 90-120^\circ, 120-150^\circ, 150-180^\circ\}$  can be calculated according to:

$$V_i = \frac{2\pi}{3} [(r + r_{extend})^3 - r^3] \left[ 1 - \cos\left(\frac{\pi i}{6}\right) \right] - \sum_{j=1}^{i-1} V_j$$

Then, the molecules counted per bin are simply converted to nanomolar:

$$[\text{pheromone}], \text{nM} = \frac{N_{\text{pheromone}}}{V_i \times 10^{-15} \frac{\text{L}}{\mu\text{m}^3}} \frac{10^9 \text{ nM/M}}{6.02 \times 10^{23} \text{ molecules/mol}}$$



## REFERENCES

- Alvarez, L., Friedrich, B.M., Gompper, G., and Kaupp, U.B. (2014). The computational sperm cell. *Trends Cell Biol.* 24, 198–207.
- Berg, H.C., and Purcell, E.M. (1977). Physics of chemoreception. *Biophys. J.* 20, 193–219.
- Bhattacharjee, N., and Folch, A. (2017). Large-scale microfluidic gradient arrays reveal axon guidance behaviors in hippocampal neurons. *Microsyst. Nanoeng.* 3, 17003.
- Bi, E., and Park, H.-O. (2012). Cell polarization and cytokinesis in budding yeast. *Genetics* 191, 347–387.
- Butty, A.C., Pryciak, P.M., Huang, L.S., Herskowitz, I., and Peter, M. (1998). The role of Far1p in linking the heterotrimeric G protein to polarity establishment proteins during yeast mating. *Science* 282, 1511–1516.
- Dyer, J.M., Savage, N.S., Jin, M., Zyla, T.R., Elston, T.C., and Lew, D.J. (2013). Tracking shallow chemical gradients by actin-driven wandering of the polarization site. *Curr. Biol.* 23, 32–41.
- Hegemann, B., and Peter, M. (2017). Local sampling paints a global picture: Local concentration measurements sense direction in complex chemical gradients. *Bioessays* 39.
- Hegemann, B., Unger, M., Lee, S.S., Stoffel-Studer, I., van den Heuvel, J., Pelet, S., Koeppl, H., and Peter, M. (2015). A cellular system for spatial signal decoding in chemical gradients. *Dev. Cell* 35, 458–470.
- Henderson, N.T., Pablo, M., Ghose, D., Clark-Cotton, M.R., Zyla, T.R., Nolen, J., Elston, T.C., and Lew, D.J. (2019). Ratiometric GPCR signaling enables directional sensing in yeast. *PLoS Biol.* 17, e3000484.
- Kelley, J.B., Dixit, G., Sheetz, J.B., Venkatapurapu, S.P., Elston, T.C., and Dohlman, H.G. (2015). RGS proteins and septins cooperate to promote chemotropism by regulating polar cap mobility. *Curr. Biol.* 25, 275–285.
- Kuchler, K., Dohlman, H.G., and Thorner, J. (1993). The a-factor transporter (STE6 gene product) and cell polarity in the yeast *Saccharomyces cerevisiae*. *J. Cell Biol.* 120, 1203–1215.
- Lakhani, V., and Elston, T.C. (2017). Testing the limits of gradient sensing. *PLoS Comput. Biol.* 13, e1005386.
- Martin, S.G. (2019). Molecular mechanisms of chemotropism and cell fusion in unicellular fungi. *J. Cell Sci.* 132.
- Merlini, L., Khalili, B., Bendežú, F.O., Hurwitz, D., Vincenzetti, V., Vavylonis, D., and Martin, S.G. (2016). Local Pheromone Release from Dynamic Polarity Sites Underlies Cell-Cell Pairing during Yeast Mating. *Curr. Biol.* 26, 1117–1125.
- Michaelis, S. (1993). STE6, the yeast a-factor transporter. *Semin. Cell Biol.* 4, 17–27.
- Moore, T.I., Chou, C.-S., Nie, Q., Jeon, N.L., and Yi, T.-M. (2008). Robust spatial sensing of mating pheromone gradients by yeast cells. *PLoS One* 3, e3865.

- Nern, A., and Arkowitz, R.A. (1998). A GTP-exchange factor required for cell orientation. *Nature* 391, 195–198.
- Nern, A., and Arkowitz, R.A. (1999). A Cdc24p-Far1p-Gbetagamma protein complex required for yeast orientation during mating. *J. Cell Biol.* 144, 1187–1202.
- Pryciak, P.M., and Huntress, F.A. (1998). Membrane recruitment of the kinase cascade scaffold protein Ste5 by the Gbetagamma complex underlies activation of the yeast pheromone response pathway. *Genes Dev.* 12, 2684–2697.
- Rappel, W.-J., and Edelstein-Keshet, L. (2017). Mechanisms of cell polarization. *Current Opinion in Systems Biology* 3, 43–53.
- Segall, J.E. (1993). Polarization of yeast cells in spatial gradients of alpha mating factor. *Proc. Natl. Acad. Sci. USA* 90, 8332–8336.
- Swaney, K.F., Huang, C.-H., and Devreotes, P.N. (2010). Eukaryotic chemotaxis: a network of signaling pathways controls motility, directional sensing, and polarity. *Annu. Rev. Biophys.* 39, 265–289.
- von Philipsborn, A., and Bastmeyer, M. (2007). Mechanisms of gradient detection: a comparison of axon pathfinding with eukaryotic cell migration. *Int Rev Cytol* 263, 1–62.
- Wang, Y., and Dohlman, H.G. (2004). Pheromone signaling mechanisms in yeast: a prototypical sex machine. *Science* 306, 1508–1509.
- Wiget, P., Shimada, Y., Butty, A.-C., Bi, E., and Peter, M. (2004). Site-specific regulation of the GEF Cdc24p by the scaffold protein Far1p during yeast mating. *EMBO J.* 23, 1063–1074.

## CHAPTER 4: PROBING CONFORMATIONAL CHANGES OF SINGLE MOLECULES IN LIVING CELLS WITH BINDER/TAG<sup>4</sup>

### Overview

Protein conformational changes can elicit dramatic functional changes. Common outcomes of conformational changes include shifts in protein stability, the exposure of binding sites, the formation of catalytic sites, and the re-localization of the protein itself. These outcomes ultimately cause changes in cellular behavior. Connecting a protein's conformational changes to specific cellular outcomes requires understanding both the protein's structures as well as its subcellular context. Methods such as X-ray crystallography and cryo-electron microscopy have provided valuable atomic resolution structures of proteins, but are difficult or impossible to apply to living cells. Because X-ray crystallography requires protein crystals, it is generally restricted to *in vitro* static snapshots, with some notable exceptions where intracellular crystals form (Schönherr et al., 2018). Cryo-electron microscopy has been used to look within cells for well over a decade (Medalia et al., 2002), but requires freezing. Because these techniques are unsuited to living cells, they may miss dynamic changes in cellular behavior. Nuclear magnetic resonance (NMR) can measure both structure and dynamics in solution and even in living cells (Freedberg and Selenko, 2014; Smith et al., 2015). However, NMR is not well-suited to reporting subcellular localization.

Förster resonance energy transfer (FRET)-based biosensors are a widely used approach to fluorescently report the conformational changes and subcellular locations of proteins in living cells. In FRET, light is used to excite a donor fluorophore, and conformational states are inferred through the

---

<sup>4</sup>This chapter is part of a manuscript in preparation for submission to *Cell* in which I am one of three co-first authors (together with Bei Liu and Orrin Stone). My contributions to the work were purely computational, but the results draw heavily on experimental work by my co-authors. Therefore, I include experimental data, discussion, and methods necessary to understanding the computational contributions I made.

distance- and dipole-dependent transfer of the donor emission to an acceptor fluorophore (Roy et al., 2008). The indirect excitation of the acceptor results in poor signal-to-noise relative to a directly excited fluorophore, thus ensemble measurements of many proteins are required to produce meaningful readouts. As a result, it is difficult to investigate differences in protein conformation at the level of individual molecules. Single molecule FRET (smFRET) allows researchers to examine conformational changes of individual molecules, but this technique is largely performed with purified protein samples. Very few attempts have been made to apply smFRET in living cells due to technical difficulties, such as the need for site specific labeling *in vivo* and low signal-to-noise (Murakoshi et al., 2004; Sakon and Weninger, 2010). Another fluorescence-based approach is to use fluorophore-labeled affinity reagents. Affinity reagents are engineered based on existing small protein domains that bind selectively to a particular conformation of a protein. In contrast to smFRET, affinity reagents can have good signal-to-noise through direct excitation of the fluorophore (Hodgson et al., 2008). However, affinity reagents can be difficult to design and require significant optimization for each target protein conformation.

Researchers in Dr. Klaus Hahn's lab developed the Binder/Tag technique, which generalizes the affinity reagent approach (Stone, 2018). We applied Binder/Tag to study the conformational dynamics of Src, an important protein in motility, proliferation, and cancer (Playford and Schaller, 2004; Roskoski, 2015; Wheeler et al., 2009). In particular, we studied Src at single-molecule resolution, revealing the nanoscale organization and dynamic coordination of Src conformations in living cells. These single particle tracking experiments generated hundreds of thousands of trajectories, requiring sophisticated computational approaches to interpret. I developed and applied computational approaches to analyze data from the Binder/Tag technique for visualizing and tracking the conformational states of individual Src molecules in living cells. We found that active Src was concentrated in large adhesions, in part driven by enhanced recruitment and activation of Src. We also found tiny (~200 nm diameter) clusters of Src molecules with centralized binding sites, enriched in the active conformation. The clusters were transient structures, and captured Src both from the cytoplasm and from nearby membrane. Our observations imply selective binding and/or generation of active Src at clusters. Fitting stochastic, single-molecule models of

Src conformational dynamics to the data allowed us to extract kinetic rates of activation, inactivation, and plasma membrane dissociation in a live-cell context. We show that open-conformation Src has a significantly slower rate of dissociation from the plasma membrane compared to closed-conformation Src. These findings shed new light on Src activation and architecture at the nanoscale, and demonstrate new approaches to investigate single molecule kinetics and spatial organization in living cells.

#### **4.1. Introduction**

The Binder/Tag technique was developed to probe the dynamics and subcellular localization of protein conformational changes within living cells. Specifically, we applied Binder/Tag to autoinhibited proteins, which have an autoinhibitory domain that suppresses function by keeping the protein in an inactive conformation, regulating ligand binding, subcellular localization, and/or enzymatic activity (Pufall and Graves, 2002). Binder/Tag relies on two key molecules: the Tag, a seven amino acid peptide, and the Binder, an 18 kDa protein. Binder (also known as SspB) and Tag (also known as SsrA) are two bacterially-derived molecules that bind with specificity and high affinity (16 nM) (Wah et al., 2002). The Tag is engineered into a protein of interest such that when the protein is in the closed conformation, the Tag is masked. In the protein's open conformation, the Tag is exposed and the Binder can associate. Thus, the Binder acts as an affinity reagent for the protein's open conformation. Fluorescently labeling both the tagged protein of interest and the Binder allows us to follow both the localization and conformation of our protein of interest.

We applied the Binder/Tag approach to Src kinase. Src owes its name to the word sarcoma, and has been extensively studied due to its association with cancer. The viral protein v-Src is encoded by the oncogene of Rous sarcoma virus, which was discovered over a century ago in chickens (Rous, 1911). The normal cellular homologue Src was characterized in the 1970s (Stehelin et al., 1976). Src is an autoinhibited protein that has a closed inactive state and an open active state, which are primarily regulated by phosphorylation of a C-terminal tyrosine (Tyr527). Phosphorylation of the C-terminal tyrosine results in an intramolecular pTyr-SH2 interaction, locking the protein in the closed inactive

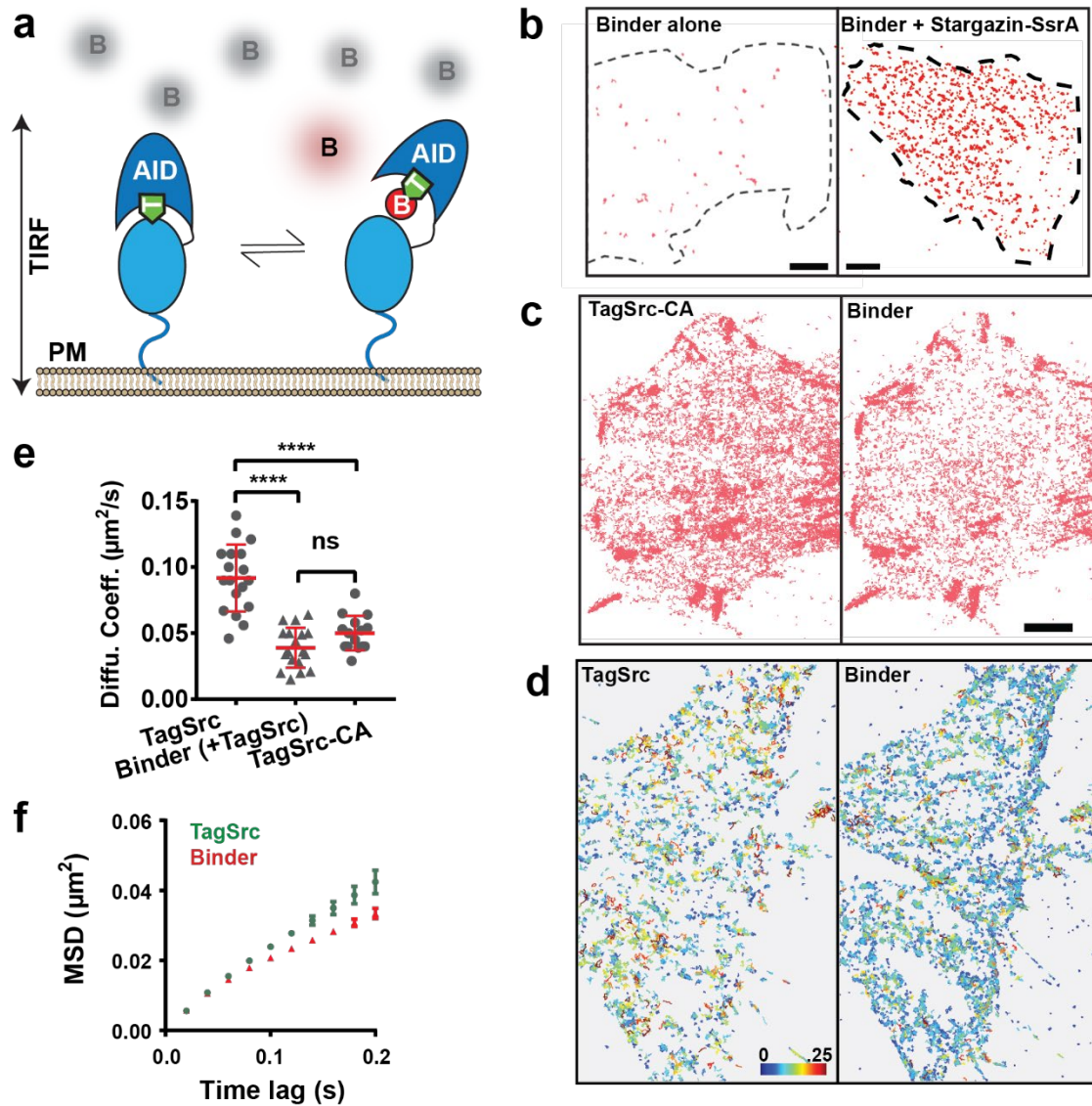
conformation. Dephosphorylation and/or competitive binding for the SH2 domain results in open conformation Src, which can become kinase-active through subsequent phosphorylation in its kinase domain (Tyr416) (Kmieciak and Shalloway, 1987; Roskoski, 2005). v-Src is essentially a constitutively active form of Src, as it lacks seven amino acid residues at the C-terminus including the critical autoinhibitory tyrosine (Cooper et al., 1986). Extensive research has revealed that mutations in human Src are not a primary driver of tumorigenesis, but Src is nonetheless an important regulator in many related pathways including cell motility, proliferation, and survival (Playford and Schaller, 2004; Roskoski, 2015; Wheeler et al., 2009). Furthermore, Src is the eponymous member of the Src family kinase (SFK) proteins, which are important in immunological and neurological contexts (Parsons and Parsons, 2004).

The activation of Src and Src family kinases results in dynamic micro- and nanoscale subcellular localization. Work with constitutively active Src mutants showed that active Src diffused more slowly in the plasma membrane, and exhibited enrichment within adhesion structures (Fincham et al., 1996; Shvartsman et al., 2007). However, these mutants may be non-representative of native Src behavior. Early iterations of Src affinity reagents supported these ideas, though they had problems with specificity versus Src family kinases and other related proteins (Ting et al., 2001; Wang et al., 2005). Advances in protein engineering shed light on the microscale subcellular localization of active Src, showing activity-dependent localization within adhesions (Chu et al., 2014; Karginov et al., 2014; Koudelková et al., 2019). The development of superresolution microscopy led to nanoscale, sub-diffraction imaging of Src and Src mutants within fixed cells, revealing the existence of tiny nanoscale Src clusters (Githaka et al., 2016; Owen et al., 2010; Smith et al., 2016). Using Binder/Tag, we can investigate the dynamics and localization of Src activation at single molecule resolution in living cells, without relying on Src mutants or on non-specific affinity reagents.

## **4.2. Single particle tracking of total and active Src with Binder/Tag**

We performed dual-color single particle tracking (SPT) of tagSrc and Binder to report the dynamics of both total Src and open Src (Figure 4.1A). Several controls were performed to validate the technique.

First, since Binder is a cytosolic protein with little non-specific binding to the plasma membrane (PM), it should rapidly diffuse in the cytosol in the absence of a membrane-associated Tag. Accordingly, SPT of fluorescently labeled Binder expressed alone yields very few tracks in comparison to co-expression with Tag fused to the transmembrane receptor Stargazin (Figure 4.1B, top). Second, we considered the constitutively active mutant Src-Y527F. The Y527F mutation results in a C-terminal tail that cannot bind the SH2 domain, leaving Src in the open conformation and priming it for activation by Y416 phosphorylation. Dual-color SPT of tagSrc-Y527F and Binder produced similar SPT maps, indicating the technique is working as expected (Figure 4.1B, middle). This experiment also showed that constitutively active Src accumulates on adhesion-like structures (confirmed by visualizing adhesions, see Section 4.3), which is consistent with biosensor-based studies of Src conformation (Koudelková et al., 2019) and with chemogenetic studies of Src kinase activation (Chu et al., 2014; Karginov et al., 2014). Although the Binder and tagSrc-Y527F SPT maps are similar, the majority of individual tracked molecules did not co-diffuse. This is because only a small fraction of the tagged protein of interest and the Binder was fluorescently labeled, in order to resolve individual fluorophores for SPT. Next, to further demonstrate that Binder is reporting on the open conformation of Src, we used diffusional analyses to show that Binder co-expressed with tagSrc-Y527F diffuses similar to tagSrc-Y527F itself, both of which are slower than tagSrc-WT. This is consistent with studies using Src-Y527F (Machiyama et al., 2015; Shvartsman et al., 2007). SPT maps of tagSrc and Binder reconstructed from a 60-second movie are shown in Figure 4.1D, with diffusional analyses shown in Figure 4.1E, F.

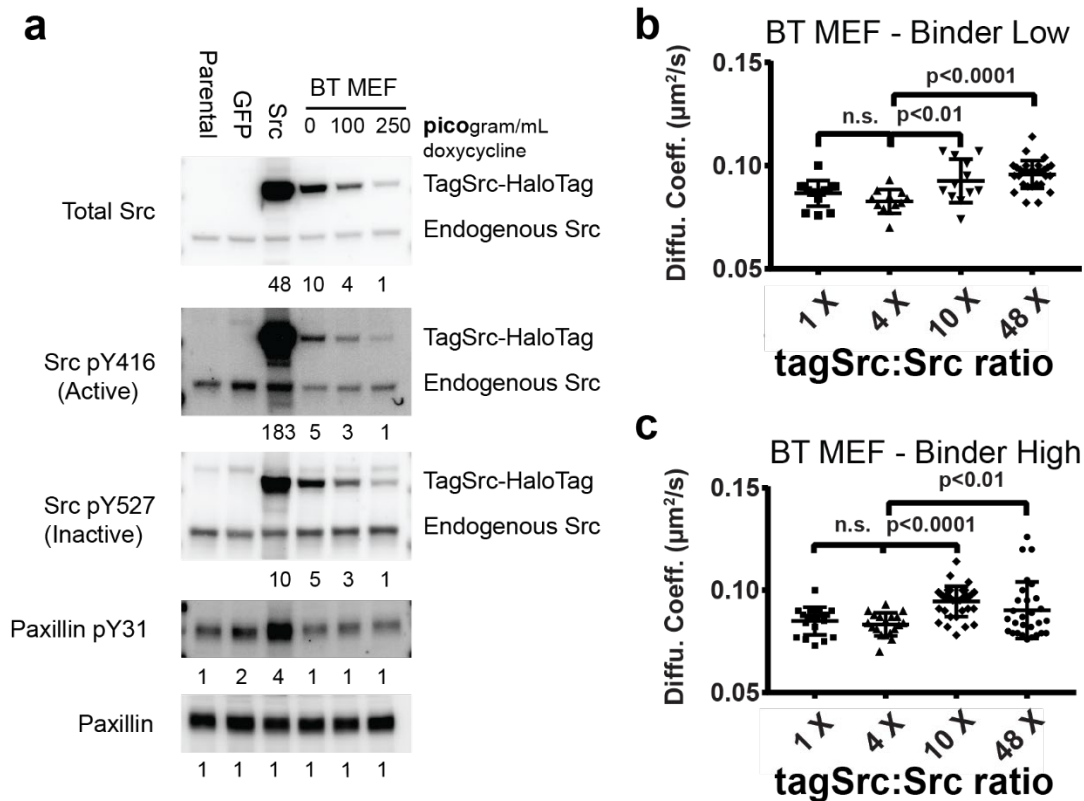


**Figure 4.1. Applying Binder/Tag to probe Src conformations with single particle tracking (SPT).** (A) Reporting conformational states with Binder/Tag under total internal reflection fluorescence (TIRF) microscopy. T: Tag, B: Binder, AID: Autoinhibitory domain, PM: plasma membrane. Cytosolic Binder away from the TIRF plane is not observed, and cytosolic Binder within the TIRF plane diffuses rapidly relative to membrane-associated Binder. (B,C) Control SPT experiments in COS-7 cells with transient transfection, scale bars 5  $\mu\text{m}$ . (B) Binder alone compared to Binder co-transfected with tagStargazin. (C) Constitutively active tagSrc (tagSrc-Y527F) compared to Binder, scale bar 5  $\mu\text{m}$ . (D) SPT experiments in mouse embryonic fibroblast (MEF) cells, stably expressing tagSrc and Binder. Trajectories colored by diffusion coefficient ( $\mu\text{m}^2/\text{s}$ ). (E) Diffusion coefficients extracted from mean squared displacement (MSD) analysis. (F) MSD curves for tagSrc and Binder, 95% CI shown.

Because the overexpression of native proteins and introduction of mutant proteins into living cells can significantly perturb normal cell signaling, we also assessed tagSrc abundance, diffusion, and



signaling over a range of tagSrc and Binder expression levels (Figure 4.2). Using mouse embryonic fibroblasts (MEFs) stably expressing tagSrc and Binder under the control of a doxycycline-responsive promoter (named BT MEF cells), we found that tagSrc expression at about 4-fold of endogenous Src produced sufficient signal for both the tagSrc and Binder channels without perturbing Src diffusion or activity. Our chosen levels of Binder expression did not affect Src activation or diffusion.

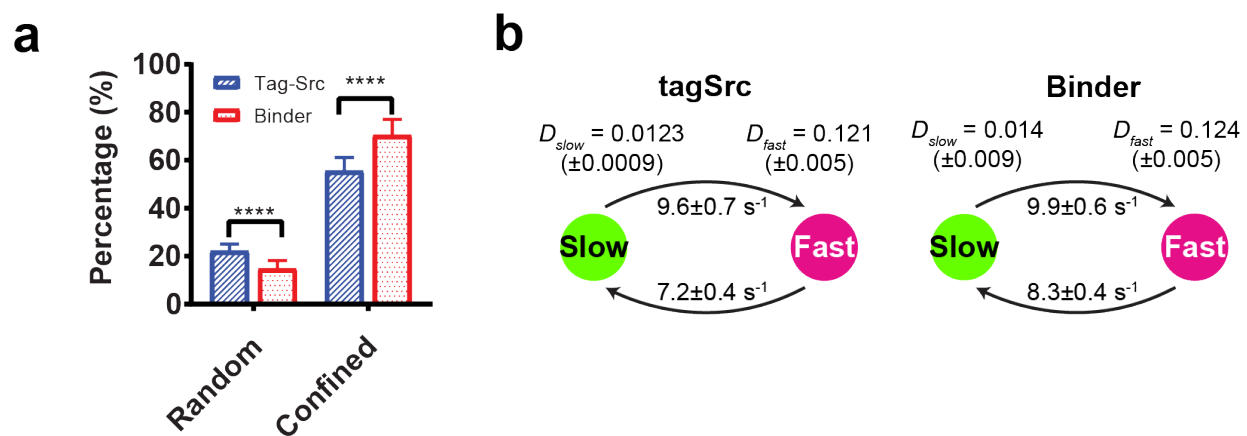


**Figure 4.2. Working expression levels of tagSrc and Binder are non-perturbing.**

(A) Cell lysate immunoblots measuring tagSrc expression and phosphorylation relative to endogenous Src, Paxillin phosphorylation as a readout of Src activity, and total Paxillin abundance. Lanes: Parental, parent MEF cell line; GFP, transient transfection of GFP; Src, transient transfection of Src; BT MEF dosed with doxycycline to tune expression of tagSrc and Binder. (B,C) Measuring differences in tagSrc diffusion as a function of both tagSrc and Binder expression.

To characterize the slowed diffusion of open conformation Src, we used two complementary diffusional analysis approaches. Moment scaling spectrum analysis (Ewers et al., 2005; Ferrari et al., 2001) revealed that Src can exist in both free and confined diffusional states. Binder shows a statistically significant higher percentage of confined diffusion tracks than tagSrc (Figure 4.3A), indicating more

interactions with slow-diffusing binding partners and/or enclosure within cytoskeletal corrals (Kusumi et al., 2005). Note, however, that the Binder also exhibits some free (random) diffusion, indicating not all open-conformation Src is bound to slow-diffusing partners. An independent hidden Markov model approach (HMM) was also used to identify slow and fast diffusional states (Persson et al., 2013). In support of the MSS analysis, the HMM indicated that the Binder more readily entered the slow-diffusing state than tagSrc, though tagSrc and Binder had otherwise similar diffusive states and similar slow-to-fast transitions (Figure 4.3B). Altogether, these results support a model where open-conformation Src transiently interacts with slow-diffusing binding partners.



**Figure 4.3. Two independent diffusional analyses reveal slow-diffusing open-form Src.** (A) Moment scaling spectrum analysis of tagSrc and Binder. (B) Two-state Hidden Markov model analysis of tagSrc and Binder. Mean  $\pm$  1s.d. shown.

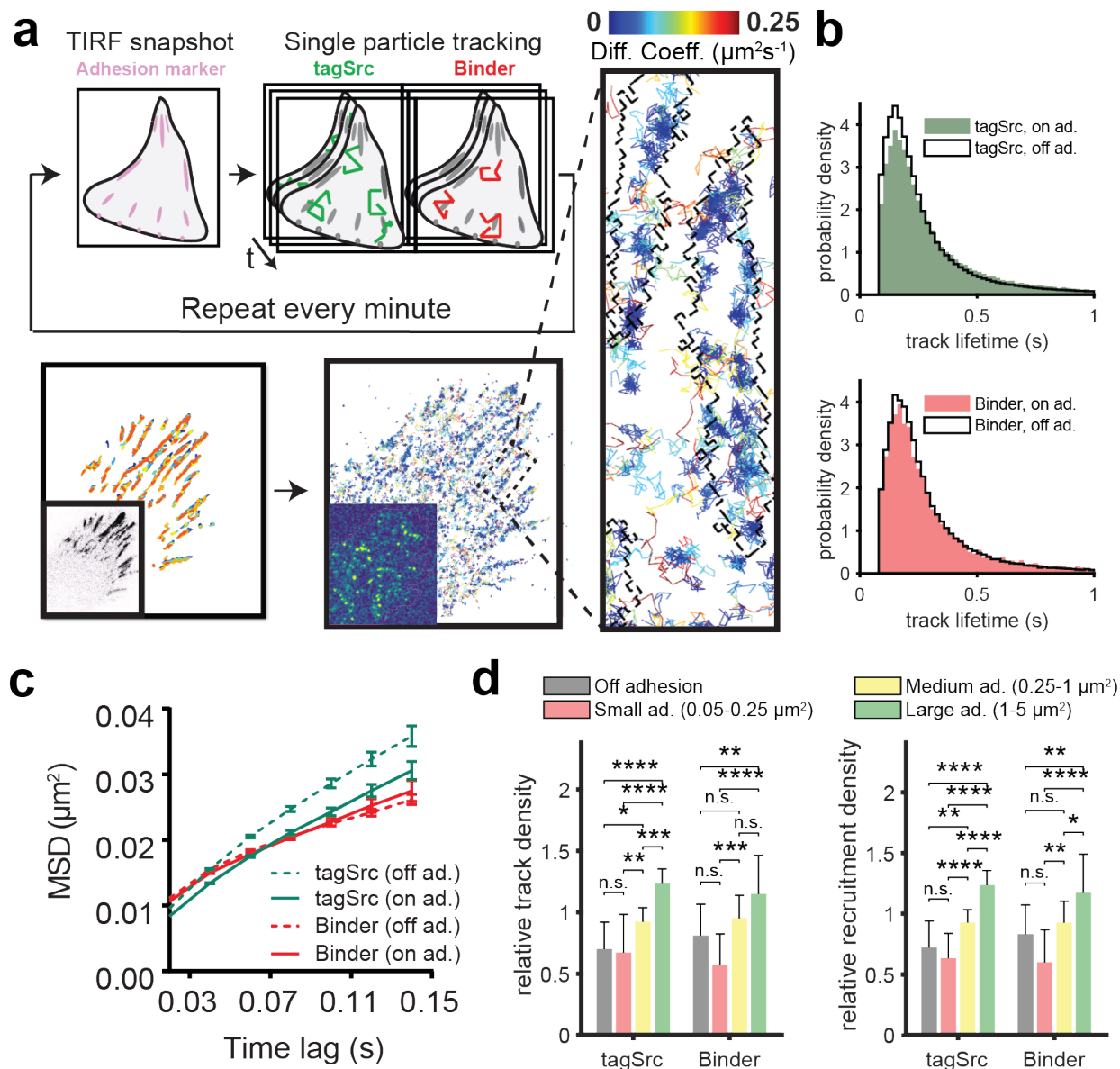
#### 4.3. Accumulation of active Src in adhesions in a size-dependent process

Our results showed that active Src was enriched within adhesion structures (Figure 4.1C and Figure 4.4), consistent with literature on Src's role in adhesions (Chu et al., 2014; Karginov et al., 2014; Koudelková et al., 2019; Webb et al., 2004; Zaidel-Bar et al., 2007). We used our approach to investigate Src accumulation and activation within adhesions at the single molecule level. To visualize adhesions, we expressed low levels of fluorescently-labeled adhesion proteins (e.g. Paxillin-EGFP) and interweaved TIRF imaging of the adhesion marker and dual-color SPT of tagSrc and Binder (Figure 4.4A). Paxillin-EGFP was used for most of our studies, because Paxillin is present from the beginning of adhesion formation (Gardel et al., 2010), though using FAK-EGFP and Vinculin-EGFP yielded similar results.

Expression of the Paxillin-EGFP construct changed neither the morphology of endogenous adhesions, nor the overall diffusion of tagSrc.

Diffusional analysis of tracks within adhesions highlighted the accumulation of slow-moving Src within adhesions (Figure 4.4A). Quantitative analysis with mean squared displacement curves showed that tagSrc indeed diffused more slowly in adhesions versus out of adhesions (Figure 4.4C). In contrast, the Binder diffused slowly whether it was on adhesions or not (Figure 4.4C). Slowed diffusion is therefore likely caused by additional molecular interactions Src experiences when in the open conformation. Furthermore, once in adhesions,  $73\pm 5\%$  of all tagSrc tracks and  $72\pm 3\%$  of all Binder tracks remained inside before the track disappeared, due to either disassociation from the plasma membrane or photobleaching. These results indicate that in adhesions, Src enters the open conformation and binds slow-diffusing substrates, leading to the accumulation of active Src within the adhesion structure.

Since adhesion-associated tracks persisted on the membrane at least as long as non-adhesion-associated tracks (Figure 4.4B), we were able to measure the density of observed tracks (which are measured over time) as a proxy for tagSrc and Binder concentration. Both tagSrc and Binder are concentrated on larger adhesions, up to about 2-fold versus small adhesions and the plasma membrane, similar to previous reports (Koudelková et al., 2019; Playford and Schaller, 2004) (Figure 4.4C, left).



**Figure 4.4. Interweaving adhesion imaging and dual-color single particle tracking of tagSrc and Binder reveals accumulation of slow-moving active Src.**

(A) Strategy and example of interweaving adhesion imaging and dual-color SPT. Top: cartoon strategy. Bottom: Example images of segmented adhesions and tagSrc track maps (insets, representative snapshots). Right: zoom of cell showing tagSrc tracks colored by diffusion coefficient, and adhesion outlines (black dashed lines) (B) Adhesion-associated tracks persist on the membrane at least as long as non-adhesion-associated tracks. (C) Quantitative mean-squared displacement analysis of diffusion. Error bars show 95% CI. (D) Quantitative analysis of tagSrc and Binder accumulation within adhesions of increasing size. Left, concentrations of tagSrc and Binder. Right, recruitment of tagSrc and Binder. Statistical comparisons were done using ANOVA. Analysis over  $n=16$  cells.

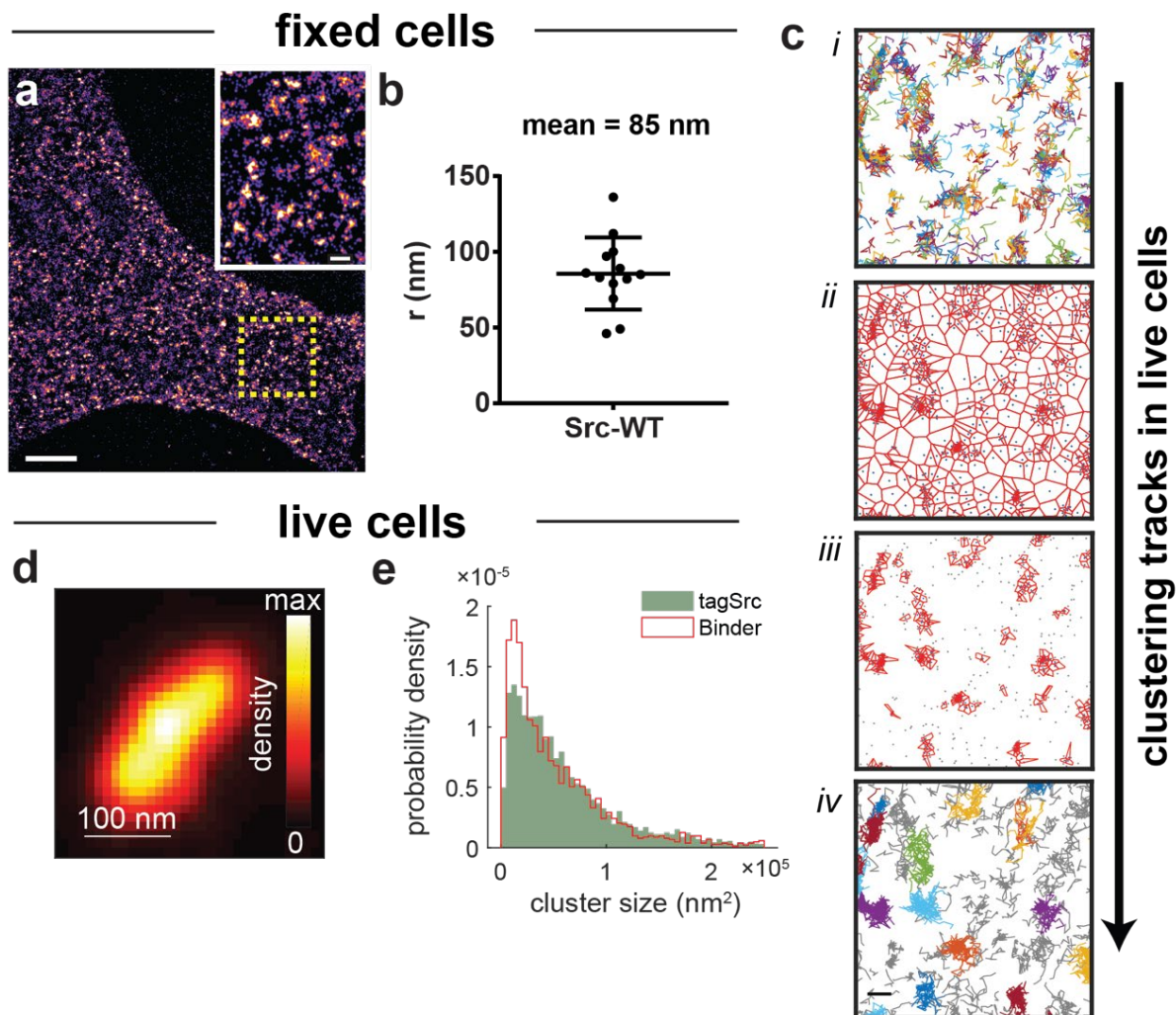
Single particle tracking also allowed us to inspect how tagSrc and Binder entered adhesions. Tracks could either start outside of adhesions and diffuse in, i.e. lateral membrane recruitment, or start inside adhesions, i.e. cytosolic recruitment. A small majority of adhesion-associated tagSrc and Binder tracks started within adhesions ( $57\pm6\%$  tagSrc,  $58\pm4\%$  Binder; 95% CI), indicating that most adhesion-associated Src is directly recruited from the cytoplasm. Furthermore, the cytosolic recruitment of tagSrc and Binder also exhibits a 2-fold increase on large adhesions compared to small adhesions and the plasma membrane (Figure 4.4D). We note that the cytosolic recruitment of Binder cannot distinguish between the cytosolic recruitment of active Src, or simply the cytosolic recruitment of Binder to Src activated after entering adhesions. Overall, these results show that active Src accumulates in adhesions of increasing size, and that the largest adhesions are more effective at recruiting Src. These changes are likely driven by changes in the molecular composition of adhesions as they mature, as Src activity is necessary for the regulation of adhesion maturity (Gardel et al., 2010; Huvneers and Danen, 2009; Webb et al., 2004).

#### **4.4. Nanoscale clusters of Src are dynamic hotspots of activity**

Our SPT maps exhibited clusters of tracks (Figure 4.4A). Studies in fixed cells have found that Src family kinases form clusters (Githaka et al., 2016; Owen et al., 2010; Smith et al., 2016). With fixed-cell PALM imaging, we also observed that wild-type tagSrc forms  $\sim 100$  nm clusters on the plasma membrane (Figure 4.5A, B). Few corresponding studies have been done in living cells, and the dynamics of Src family kinase clusters under physiological conditions are still poorly understood. To address this problem, we studied the clustering dynamics of tagSrc and Binder using dual-color SPT in BT MEFs.

I developed an approach to extract dynamic clustering information from SPT data based on Voronoi tessellation (Figure 4.5C). The method identifies sets of clustered tracks, then isolates each individual cluster for analysis. Voronoi tessellation, which has been used for single molecule cluster analysis (Andronov et al., 2016; Levet et al., 2015), was used to determine local track centroid densities, followed by automated thresholding to propose cluster assignments based on the local density. The proposals were refined spatially by capturing additional tracks based on complete trajectories, and by splitting clusters

when unusually long gaps occurred between new track observations. Additional filters were applied to reject spurious clusters. While other approaches have been developed for live cell single molecule cluster analyses (Griffié et al., 2018), our method combines cluster analysis with single particle tracking data, allowing us to follow the trajectories of individual molecules within clusters.



**Figure 4.5. Observing nanoscale clusters of Src.**

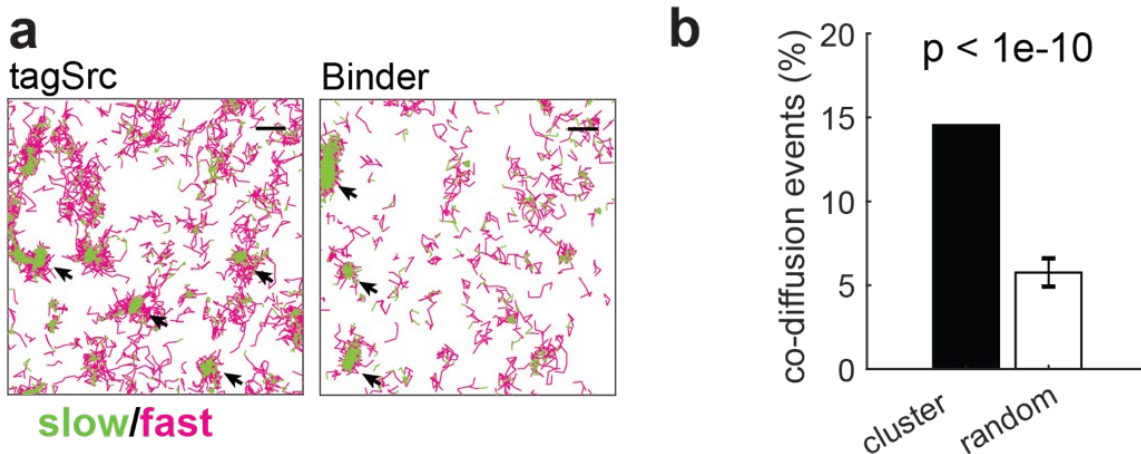
(A) PALM imaging in fixed MEF cells, transiently transfected with Src-WT-mGeosM, scale bar, 5  $\mu\text{m}$ . Inset is a magnification of the yellow box, scale bar, 1  $\mu\text{m}$ . (B) Cluster size estimated using pair correlation PALM, with each dot indicating a measurement from a 3  $\mu\text{m}$  x 3  $\mu\text{m}$  cropped region. (C) Workflow for clustering tracks of tagSrc and Binder in living cells. The SPT maps (i) were transformed into track centroid maps and divided using Voronoi tessellation (ii). A size threshold was used to select for densely packed centroids, leaving behind proposed clusters (iii). Spatial and temporal refinements, and filtering out spurious clusters produced sets of clustered tracks (iv), scale bar, 500 nm. (D) A heat map showing localization densities within an isolated cluster of tracks. (E) Distributions of tagSrc and Binder cluster sizes within living cells. Cluster sizes were defined using regions inside clusters with >50% maximal localization density. Analysis represents  $n=61$  cells with  $n=3972$  tagSrc clusters and  $n=2058$  Binder clusters.

Cluster sizes were measured based on local density maps calculated for isolated clusters (Figure 4.5D). Areas and effective radii were determined by thresholding the density maps at half their maximum density, essentially reporting on the “core” of the cluster. Binder clusters were similar in size to tagSrc clusters, though slightly smaller (Figure 4.5E) (95% CI for median radii: 105-113 and 120-124 nm for Binder and tagSrc, respectively). The “extents” of the cluster, assessed by instead thresholding the density maps at 5% of the maximum density, were more noticeably different (95% CI for median radii, 256-273 nm vs. 289-297 nm for Binder and tagSrc, respectively).

Dynamic analyses of tagSrc and Binder indicate that the clusters are centralized hotspots of activity. Diffusional analysis of the tracks (Figure 4.3B and Figure 4.6A) showed that clusters contain centralized regions of slow diffusion, surrounded by clouds of fast diffusion. The zone of slow diffusion is indicative of a centralized binding site. Compared to the slow tagSrc zone, the slow Binder zone was similar in size, though slightly smaller (95% CI for median radii, 88-94 nm Binder slow vs. 96-100 nm tagSrc slow). In contrast, the fast Binder zone was noticeably smaller than the fast tagSrc zone (95% CI for median radii, 112-119 nm Binder fast vs. 129-134 nm tagSrc fast). These data suggest that the slow zone is primarily active material, while the fast exterior zone is a mixture of active and inactive material.

To further assess activity within the clusters, we identified pairs of co-diffusing tagSrc and Binder tracks. If the clusters were sites of activity, then the co-diffusion events (denoting open-conformation Src) should be enriched within clusters. Clustered tracks constituted approximately 5% of either the tagSrc or Binder data. This result suggests most membrane-associated Src does not participate in clusters. However, almost 15% of all co-diffusion events were cluster-associated. The enrichment of co-diffusion events within clusters is unlikely to happen by chance, based on statistical bootstrapping (Figure 4.6B), which indicates that a clustered Src molecule is approximately 3-fold more likely to be in the open-conformation versus an un-clustered Src molecule.





**Figure 4.6. Src clusters are hotspots of activity.**

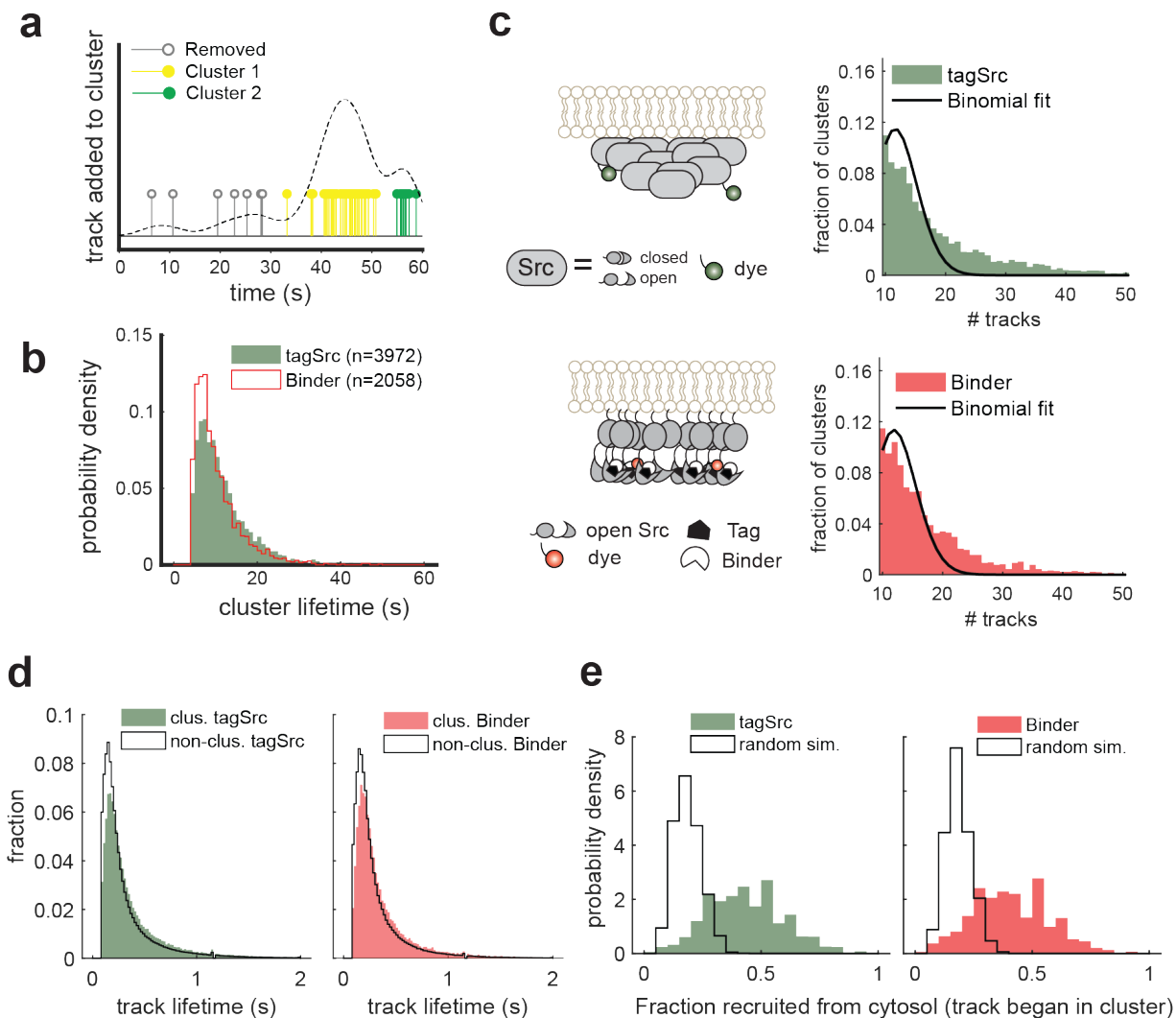
(A) SPT maps colored by slow and fast diffusive state, as defined in Figure 4.3B. Arrows point out enrichment of the slow diffusive state within clusters. Scale bar, 500 nm. (B) Activation within clusters revealed by enrichment of co-diffusion events. Clusters are 3-fold more likely to have co-diffusion events (i.e. active Src) than would be expected by random chance. Error bar obtained from 10 bootstrap selections of tracks. P-value from a two-sided, one-sample t-test. Analysis represents 61 cells with 1094 co-diffusion events, 3972 tagSrc clusters, and 2058 Binder clusters.

The dynamics revealed by our technique also allows us to follow the dynamic birth and death of clusters. We find that molecules arrived at clusters in bursts, a hallmark of dynamic cluster behavior (Cisse et al., 2013) (Figure 4.7A). The clusters existed transiently, with median lifetimes of 10.1-10.5 seconds for tagSrc and 8.5-9.0 seconds for Binder (95% CI of the median) (Figure 4.7B).

The population of clusters we observed was heterogenous, based on comparison to simple models where either a single type of tagSrc or Binder cluster existed. These simple models assume that the cluster contains  $n$  molecules on average, with some probability of observing each molecule  $p$  (related to the sparse dye labeling necessary for SPT, described in Section 4.2). If these simple models were appropriate, then the observed number of molecules within the cluster should follow Binomial distributions. Neither model fit (Figure 4.7C). Heterogeneity could be due to the diversity of Src binding partners for different biological functions, several of which are known to themselves form clusters, such as growth factor receptors (Baumgart et al., 2018; Gao et al., 2015) and integrins (Changede et al., 2015; Rossier et al., 2012). However, we found few differences when we attempted to analyze clusters in vs. out of adhesions. Differences between clusters could also be time-dependent, either on the lifetime of the cluster, or on the



time relative to the start of imaging. However, we again found few differences when we attempted to correlate our measured properties with temporal characteristics.



**Figure 4.7. Src clusters are dynamic, heterogeneous structures.**

(A) Bursts of new molecules arriving at a proposed cluster (see Figure 4.5C, iii). The black dashed line is the temporal density of events, while the colors of the stems denote classification into clusters after refinement. (B) Cluster lifetime distributions. (C) The number of tracks observed per cluster do not follow a Binomial distributions, indicating the simple Binomial models shown (left) are inadequate. (D) Both tagSrc and Binder have longer lifetimes in the plasma membrane when in clusters. (E) Measured recruitment of tracks to the cores of clusters (region of >50% maximal density) versus simulated random controls. Analysis represents 61 cells with 3972 tagSrc clusters, and 2058 Binder clusters.

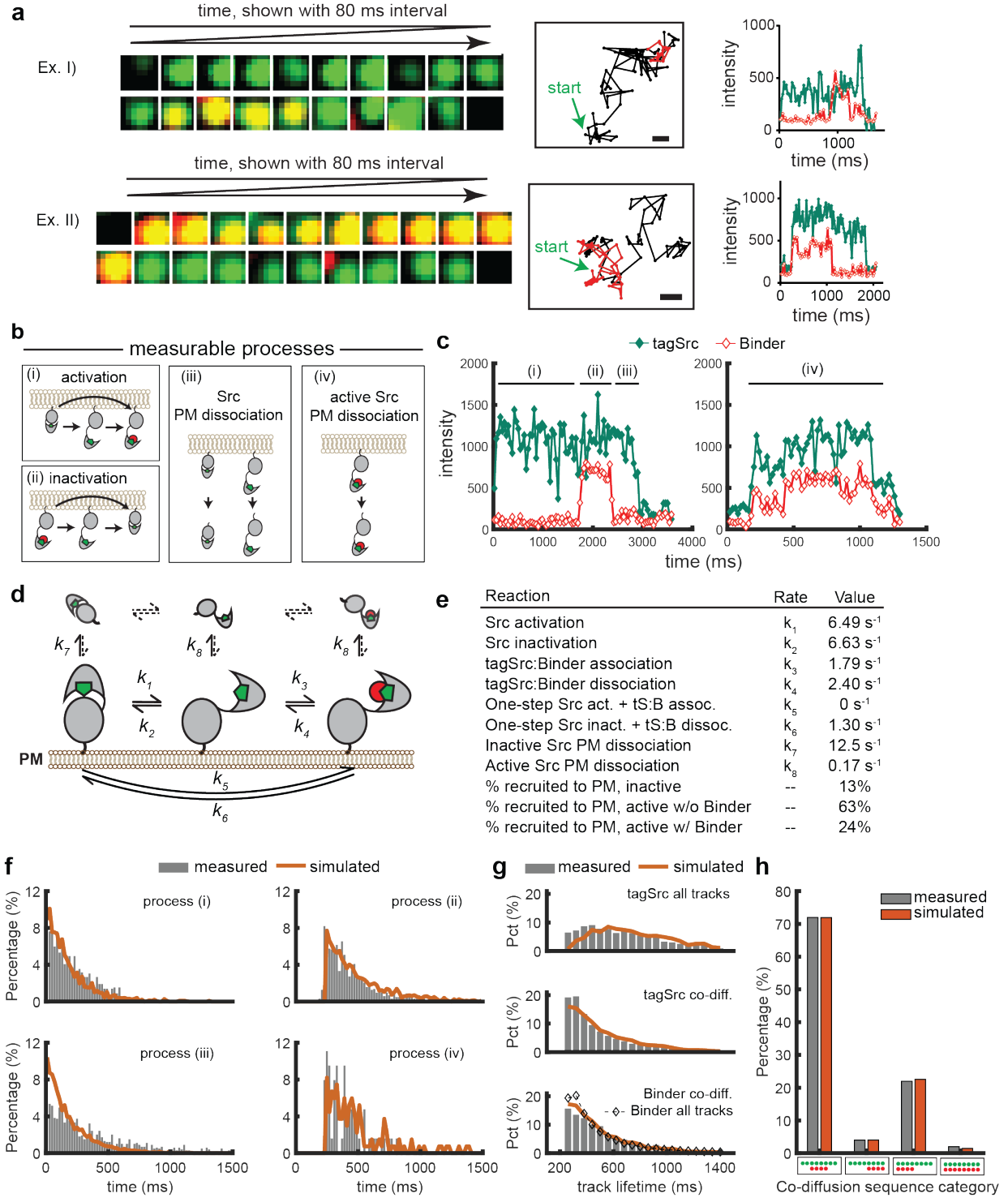
In support of the hypothesis that clusters contain centralized binding sites, clustered tracks persisted longer on the membrane than non-clustered tracks (Figure 4.7D). Our observations apply to both tagSrc

and Binder, which means that the binding sites interact with open-conformation Src, consistent with current models of Src interactions at the plasma membrane (Shvartsman et al., 2007).

Finally, by examining where tracks started, we find that on average only 35% of tracks were recruited directly into the “core” of the cluster, or the zone of half-maximal density (Figure 4.7E). The remaining 65% of tracks entered Src clusters by lateral diffusion along the plasma membrane. In contrast, simulations predict that only 15% of tracks should be recruited from the cytosol by random impingement of Src at the membrane, suggesting that the cores of clusters are enriched in binding sites accessible from the cytoplasm. Altogether, these centralized binding sites represent a heterogeneous population of slow-diffusing molecules that either selectively bind active Src or activate Src, and exist transiently in the plasma membrane.

#### **4.5. Inferring *in vivo* biochemical kinetics**

Src, as well as the ubiquitous Src homology domains SH2 and SH3, has been the subject of many kinetic and thermodynamic analyses (Ladbury and Arold, 2011). Most of these studies have been conducted *in vitro*, which allows rigorous characterization of the pure protein, but may not accurately represent *in vivo* kinetics because of differences in the abundance of native regulators (García-Contreras et al., 2012), the effects of membrane confinement on diffusion (Grecco et al., 2011; Kalay et al., 2012), and macromolecular crowding effects (Acosta et al., 2017; Grima and Schnell, 2006; Ryan et al., 1988; Schnell and Turner, 2004). Here, we inferred Src regulatory kinetics inside live cells using Binder/Tag-based SPT and computational modeling. The kinetic analyses rely upon identifying co-diffusing pairs of tagSrc and Binder tracks (Figure 4.8A). From a model of the co-diffusion events (Figure 4.8B-D) we identified four types of observable processes that report on: i) Src opening, ii) Src closing, iii) the release of Src from the plasma membrane, and iv) the release of open Src from the membrane.

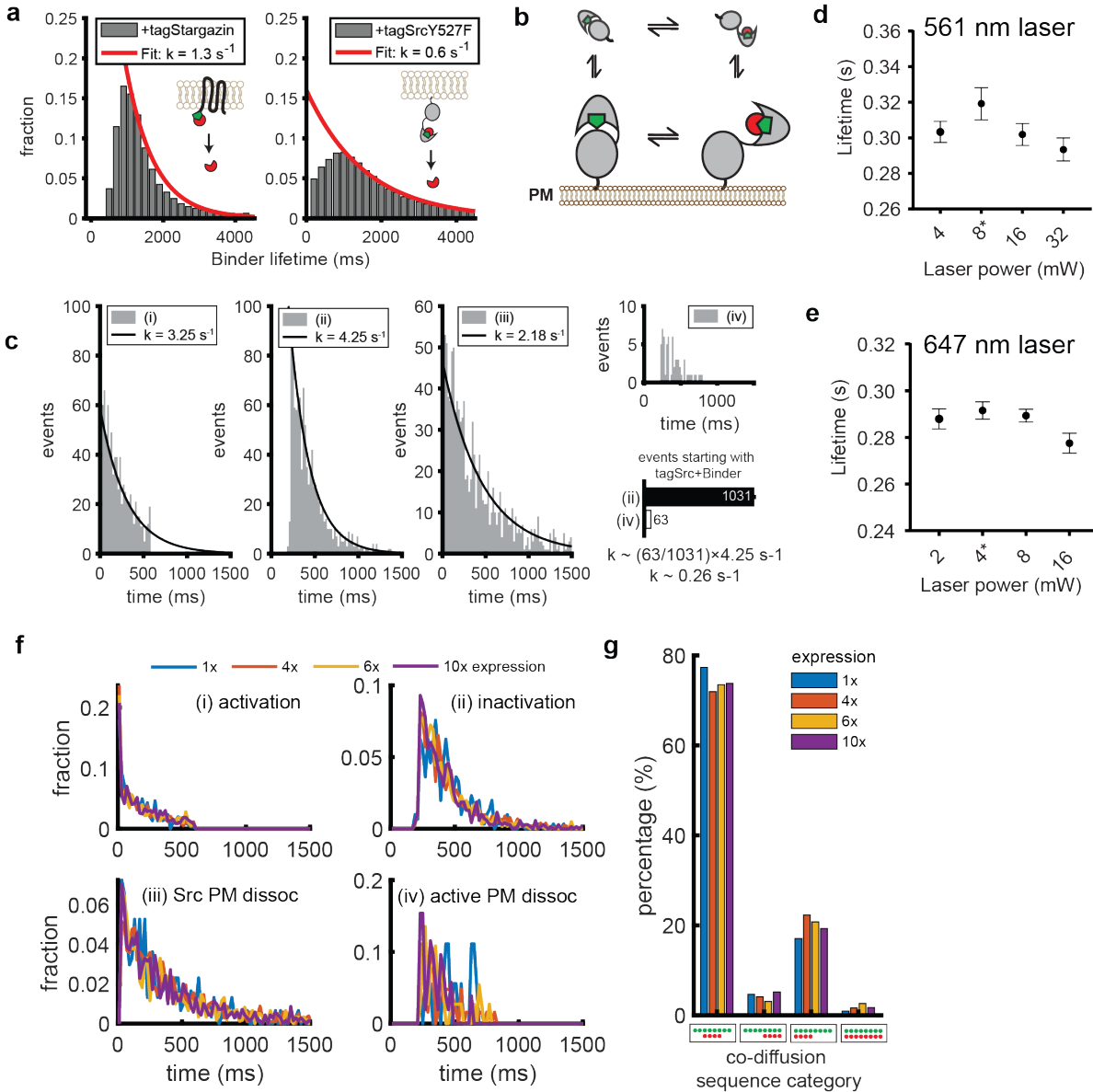


**Figure 4.8. Inferring *in vivo* Src regulatory kinetics through co-diffusion analysis.**

(A) Co-diffusing pairs of tagSrc and Binder tracks. (B,C) Four processes that can be observed from co-diffusion events. Process times (bars, (C)) are a function of individual kinetic steps, denoted with arrows (boxes, (B)). (D) Full kinetic model. (E-H) Stochastic Markov Chain simulations with optimized model parameters yield fits to the data. (F) Process times. (G) Track lifetimes for co-diffusing tracks and all tracks. (H) Sequence of arrival/departure in the co-diffusion event. Results represent  $n=1094$  events.

A complicating factor is that tagSrc/Binder interaction kinetics are entwined with the Src regulatory kinetics of interest. For example, the Src closing process ii might only report spontaneous Binder dissociation ( $k_4$ ). To ensure this is not the case, control experiments to measure spontaneous tag/Binder dissociation in live cells were performed using either constitutively active tagSrc-Y527F ( $k=0.6 \text{ s}^{-1}$ ) or the tagged transmembrane receptor tagStargazin ( $k=1.3 \text{ s}^{-1}$ ) (Figure 4.9A, B). In both cases, spontaneous Binder dissociation was at least 3-fold slower than the measured transition rate for event ii ( $k=4.25 \text{ s}^{-1}$ ) (Figure 4.9C). This implies that inactivating Src accelerates Binder dissociation, most likely driven by structural changes mediated by phosphorylation at Tyr-527. We model this effect as Binder dissociation in concert with tagSrc closing ( $k_6$ ). Another complicating factor is that Binder could force tagSrc into the open conformation. However, Binder does not perturb Src activation or diffusion in BT MEF cells under our working conditions (Figure 4.2). We therefore assume in the model that Binder cannot force tagSrc into the open form ( $k_5 = 0$ ). Photobleaching kinetics were also excluded from the model since the track lifetimes were robust to increases in the laser power (Figure 4.9D, E).

The model represents a continuous time Markov chain, and was simulated stochastically using the three PM-associated states as initial conditions. We used an evolutionary algorithm approach to simultaneously fit the model to all the experimental data (Fortin et al., 2012), and the best parameterization found is shown in Figure 4.8E-H. Altogether, the fitted model is in agreement with the SPT data (Figure 4.8F-H), and the values for the fitted rate constants are given in (Figure 4.8E). There is some mismatch between the model and the data for the process iii waiting times and the lifetimes of co-diffusing tagSrc tracks. This mismatch may be due to the lack of information on the tagSrc conformational state if Binder is not present, as well as the simplifying assumptions made for the model, i.e. treating all steps as first-order reactions.



**Figure 4.9. Controls for analyzing *in vivo* kinetics.**

**(A)** Measuring spontaneous tag/Binder dissociation using tagStargazin and tagSrc-Y527F. **(B,C)** Simplified model and fits treating each observable process as a first-order reaction. Process iv was treated with a frequency-based approach because of the rarity of events and to demonstrate a conditional effect (see text). **(D,E)** Changing the laser power does not significantly change the average track lifetime ( $0.02 \text{ s} = 1 \text{ frame}$ ). The asterisk denotes the working power. **(F,G)** Varying expression between 1-10x endogenous Src does not significantly change the measured kinetics.

The inferred parameters characterize Src regulation at the plasma membrane. For example, open-conformation Src interacts with other proteins in the PM via its SH2 and SH3 domains, whereas closed-conformation Src interacts with the PM primarily via *N*-myristoylation at Gly-2 (Sandilands et al., 2007).

We hypothesized open conformation Src would have slower PM dissociation, as compared to the closed form due to protein-protein interactions mediated by its SH2- and SH3 domains. Consistent with this line of reasoning, we find that  $k_7$  and  $k_8$  (closed and open Src PM dissociation, respectively) are  $12.5 \text{ s}^{-1}$  and  $0.17 \text{ s}^{-1}$ , respectively. In addition to co-diffusion events, the model was fit to track lifetime distributions (Figure 4.8G). These included lifetime distributions for: 1) the subset of tagSrc tracks that at some point underwent co-diffusion with Binder, 2) all tagSrc tracks, and 3) the subset of Binder tracks that at some point underwent co-diffusion with tagSrc. Note that the distribution for all Binder tracks is similar to the subset that underwent co-diffusion. This is expected because presumably all Binder molecules at the plasma membrane are interacting with an open-conformation tagSrc molecule. The difference between the tagSrc lifetime distributions arise because the tracks that underwent co-diffusion must have had open-form Src during this time, whereas the lifetime distribution for all tracks can include tagSrc molecules that were closed during the entire observed event. Therefore, the shift in the tagSrc track lifetime distribution for codiffusion events is likely a result of the increased stability of open tagSrc.

While it is tempting to estimate some of the individual rate constants directly from the distributions in Figure 4.8F, such an approach would lead to incorrect results. For example, we might estimate  $k_8$  by fitting an exponential to the time distribution for process iv. The problem is that in constructing the distribution for process iv, we selected tracks in which the Binder and tagSrc left the membrane together and ignored the tracks in which the Binder dissociated from tagSrc (i.e., the transitions characterized by rates  $k_4$  and  $k_6$ ). By not taking these events into account, the distribution for process iv shown in (Figure 4.8F) represents the distribution for the times that the co-diffusing tagSrc/Binder pair dissociated from the membrane before either transition  $k_4$  or  $k_6$  occurred. Therefore, the rate estimated from this time distribution will be greater than the value for  $k_8$ . In fact, the estimated rate is the sum of the three rates  $k_8+k_6+k_4$ , which is why the measured times for process iv are roughly similar to those of the other processes, despite the rate fitted rate  $k_8$  being slow compared to the other events ( $0.17 \text{ s}^{-1}$ ). A frequency-based analysis, which avoids the conditional effect, provides a similar estimate (Figure 4.9C).

Finally, since changes in expression level could produce changes in tagSrc diffusion as well as changes in Src phosphorylation levels (Figure 4.2), we examined whether Src regulatory kinetics were perturbed by changes in expression. However, we observed no significant changes with expression between 1-10x over endogenous Src (Figure 4.9F), indicating that our kinetic measurements are robust.

## 4.6. Discussion

In collaboration with experimentalists in Klaus Hahn's lab, I studied the conformation, dynamics, and subcellular localization of individual Src molecules. Binder/Tag allowed us to probe the conformational states of Src, while SPT allowed us to follow individual molecules arriving at, diffusing along, and departing from the plasma membrane. Combining these techniques enabled us to monitor the arrival and accumulation of active Src within adhesions at the plasma membrane, the organization of Src into activity-enriched diffraction-limited clusters, and the kinetic regulation of Src via co-diffusion analysis and modeling.

We observed that open conformation Src diffuses more slowly, consistent with studies using mutated and truncated forms of Src (Shvartsman et al., 2007), which is most likely driven by interactions with membrane-bound binding partners (Figure 4.1F). We also observed that open Src is enriched within adhesions (Figure 4.1C and Figure 4.4D). In agreement with biosensor and chemogenetic studies (Chu et al., 2014; Karginov et al., 2014; Koudelková et al., 2019), Src appears to enter the open conformation when it enters adhesions, since tagSrc diffusion slowed within adhesions, while Binder diffusion was slow whether it was in or out of adhesions (Figure 4.4C). We assessed molecules entering and exiting adhesions:  $57 \pm 6\%$  of tagSrc tracks and  $58 \pm 4\%$  of Binder tracks (95% CI) initially started within adhesions, while the remainder diffused in laterally. Once within adhesions,  $73 \pm 5\%$  of tagSrc tracks and  $72 \pm 3\%$  of Binder tracks remained inside adhesions before either dissociation from the plasma membrane or photobleaching. These results indicated that adhesions retain and accumulate open-conformation Src. Src accumulation was a function of adhesion size (Figure 4.4D), which was used to roughly distinguish stages of adhesion maturity (Gardel et al., 2010). The smallest adhesions exhibited statistically identical

accumulation and recruitment of tagSrc and Binder to non-adhesion regions of the plasma membrane. However, medium-sized adhesions (i.e. focal complexes) had elevated Src concentrations, elevated Src activation, and enhanced recruitment of Src. These changes were even more dramatic for the largest adhesions (i.e. focal adhesions), reaching a ~2-fold increase in concentration and recruitment relative to non-adhesion parts of the plasma membrane. We also observed what appeared to be molecular clusters within adhesions, that appeared to accumulate on the edges of adhesions (Figure 4.4A). While statistical analysis of many adhesions across multiple cells revealed little to no preferential localization at adhesion edges, this led us to carefully investigate the clustering behavior of Src in non-adhesion labeled cells.

While the nanoscale organization of Src has been studied in fixed cell systems (Githaka et al., 2016; Owen et al., 2010; Smith et al., 2016), no studies have been conducted to perform single-molecule imaging of Src clusters in living cells. I developed a computational pipeline for detecting, isolating, and characterizing clustered sets of tracks in both time and space (Figure 4.4C). We found that these structures were dynamic, with median lifetimes of 10.1-10.5 seconds for tagSrc and 8.5-9.0 seconds for Binder (95% CI). Furthermore, diffusional analysis highlighted centralized binding sites holding the clusters together (Figure 4.6A), which is supported by prolonged track lifetimes within clusters (Figure 4.7D) and enhanced track recruitment to clusters (Figure 4.7E). These dynamic observations would be inaccessible from fixed cell studies. The clusters were hotspots of Src activity, based on co-diffusion analyses (Figure 4.6B). There are still gaps in our knowledge of Src cluster dynamics and regulation: for example, it is unclear whether Src becomes activated within clusters (i.e. the clusters function as the site of activation) or outside of clusters (i.e. the clusters are where activity is needed). Both may be true, since we anticipate that there should be many different types of Src clusters, for example those driven by integrin clustering and those driven by EGFR clustering (Changade and Sheetz, 2017). Indeed, we observe heterogeneity in the population of observed clusters (Figure 4.7C). It was difficult to analyze adhesion-associated clusters because most tracks were not adhesion-associated.

Finally, we were able to leverage co-diffusion events to infer *in vivo* Src regulatory kinetics (Figure 4.8). These events were rare, totaling 1094 events out of hundreds of thousands of tracks across 61 cells.



We used an evolutionary algorithm approach to iteratively improve the parameters for a stochastic Markov Chain model of the biochemical events. One striking result is the large difference between the predicted inactive Src plasma membrane dissociation rate ( $k_7 = 12.5 \text{ s}^{-1}$ ) and the predicted active Src plasma membrane dissociation rate ( $k_8 = 0.17 \text{ s}^{-1}$ ). The difference is most likely driven by the additional molecular interactions active Src can make in the plasma membrane. Oddly, the difference in these rates is not apparent from looking at the experimentally measured times corresponding to plasma membrane dissociation processes iii and iv (Figure 4.8F). This is because we must select for tracks where Binder and tagSrc left the membrane simultaneously to measure process iv, thus ignoring tracks in which Binder dissociated from tagSrc first (characterized by  $k_4$  and  $k_6$ ). As a result, the measured distribution reflects events where the co-diffusing tagSrc/Binder pair dissociated from the membrane before either the  $k_4$  or  $k_6$  transition occurred. A frequency-based analysis of process iv corrects for this effect and gives a similar estimate for active Src plasma membrane dissociation as our full model fit ( $k = 0.26 \text{ s}^{-1}$ , Figure 4.9C).

These types of *in vivo* kinetic analyses should be complementary to *in vitro* kinetic analysis with purified protein. Implicitly accounting for the native cellular environment is actually a double-edged sword; while it captures the normal signaling context, it also obscures the identity of many measured molecular interactions. For example, a database of protein-protein interactions within adhesions lists 79 proteins Src may interact with (Zaidel-Bar et al., 2007); of course, Src also plays roles in other pathways such as proliferation and growth, so this is a conservative estimate. Comparative analysis of regulatory kinetics in different cellular contexts may help alleviate these issues. We hoped to perform differential kinetic analysis either through changing tagSrc and Binder expression levels or by localizing co-diffusion events within adhesions. However, our kinetic measurements were robust within 1-10x expression of tagSrc over endogenous Src (Figure 4.9F), and there were too few co-diffusion events that occurred within adhesions: roughly 10% of tracks were adhesion-associated in our system, so roughly 10x the data would be needed. Since  $n=61$  cells were used to generate hundreds of thousands of tracks, resulting in only 1094 co-diffusion events, around several hundred adhesion/SPT-correlated cells would need to be collected, which is beyond the scope of this work. Looking forward, other experimental contexts could

better allow comparison of Src regulatory kinetics, such as kinetics at the front vs. back of polarized cells. Furthermore, the Binder/Tag-based kinetic analysis approach shown here should be applicable to characterizing other proteins.

## **4.7. Methods**

### **4.7.1. Single molecule microscopy**

Single molecule tracking was performed on a home-built total internal reflection microscopy microscope based on IX81 (Olympus), equipped with four solid-state lasers (Coherent OBIS 405 nm, 488 nm, 561 nm and 647 nm). A four-band dichroic mirror (DM: Di01-R405/488/561/635, Semrock) was used for multi-color imaging. Fluorescence images were collected by a 150 X TIRF objective (UAPON 150XOTIRF, NA 1.45, Olympus), and projected to an electron-magnified charge-coupled device (EMCCD) camera (Evolve® 512 Delta, Photometrics). A 10-position emission filter wheel (emFW) (Sutter Instrument) and an imaging splitting optics (W-View GEMINI, Hamamatsu) were mounted in front of the camera to enable flexibly switching between different imaging modes. For dual-color SPT, the filters settings were as below: in GEMINI, DM (T647lpxr, Chroma), Long-pass filter (FF01-698/70, semrock) and short-pass filter (FF01-600/52, semrock); emFW was set to empty. For correlative adhesion and dual-color SPT: in GEMINI, the short-pass filter was removed; emFW was set to HQ525/50 M (Chroma) for EGFP-labeled adhesion marker imaging, and was switched to ET570lp (Chroma) for dual-color SPT.

For tracking individual molecules, single B/T MEF cells were imaged for 60 seconds at 50 Hz, generating 3000 frame videos. We first applied à trous-wavelet decomposition to extract individual single molecules. Each identified molecule was fit with a 2D Gaussian function to obtain a precise centroid. Then, we adopted a well-established trajectory linking algorithm to find the corresponding molecules in successive frames (Crocker and Grier, 1996).

For fixed-cell PALM imaging, COS-7 cells were transiently transfected with mGeosM labeled Src for 12~16 hours and plated on fibronectin-coated coverglass for 4 hours. For fixation, cells were treated

with 3.7% PFA for ~20 minutes at room temperature. Each cell was imaged for 10,000–20,000 frames. PALM reconstruction was done by an in-house MATLAB-based software. The cluster size was estimated with pair-correlation PALM analysis as described previously (Sengupta et al., 2011).

#### 4.7.2. Diffusional analysis

Mean square displacement (MSD) is one of the most common approaches to describe the diffusive properties of single molecules. The MSD at different time lags  $\tau = n\Delta t$  was calculated as below:

$$MSD(\tau) = \langle (x_{\tau+k} - x_{\tau})^2 + (y_{\tau+k} - y_{\tau})^2 \rangle$$

where  $x_{i+k}$  and  $y_{i+k}$  describe the position after a time interval,  $k$ , after starting at positions  $x_{\tau}$  and  $y_{\tau}$ .

To get the diffusion coefficient of each trajectory, the MSD curve was fitted to a nonlinear anomalous diffusion model,  $MSD(\tau) = 4D\tau^{\alpha} + 4\sigma^2$ , where  $D$  is the diffusion coefficient,  $\alpha$  is the anomalous exponent, and  $\sigma$  is the localization precision. To get a reliable estimation of the diffusion coefficient, the minimum track length was set to 12 frames and only the first 4 points were used when fit to the diffusion model.

We used moment scaling spectrum (MSS) analysis to categorize trajectories into diffusive modes, such as pure Brownian, confined (sub-diffusive) or directional (super-diffusive) movement (Ferrari et al., 2001). The moments of displacement of order  $\nu$  is defined as:

$$MSS(\tau, \nu) = \langle (x_{\tau+k} - x_{\tau})^{\nu} + (y_{\tau+k} - y_{\tau})^{\nu} \rangle$$

where  $\tau$  is the time lag. The MSD is a special case where  $\nu = 2$ . Moments were calculated for  $\nu = 0\sim6$ .

We also used a Hidden Markov model (HMM) approach to annotate the instantaneous diffusive behavior of trajectories using a previously described method (Persson et al., 2013). Briefly, trajectories with at least 5 frames were analyzed by variational Bayes single particle tracking, producing per-frame classification into discrete diffusive states based on a Brownian motion approximation. Two diffusive states were sufficient to explain most of the dataset based on an elbow analysis of the model scores, after

examination of 1, 2, 3, 4, and 5-state diffusional models. Each cell was analyzed as a separate dataset, and variability in the HMM parameters was reported over cells. The number of iterations were set to 25, bootstrapping to 100, diffusional prior to  $[0.0001, 5] \mu\text{m}^2/\text{s}$ , and dwell time prior to  $[40, 400]$  ms. Occasionally frames were missing within trajectories, so prior to HMM analysis, these frames were interpolated by minimizing the sum of squares of the second derivative at the gap (D'Errico, 2004). After construction of the HMM, the interpolated frames were removed.

#### 4.7.3. Correlating adhesion imaging with single particle tracking

Tracks were considered 'on adhesion' if their centroid was within the adhesion. For each cell, absolute track centroid densities (also 'track densities') were calculated for adhesions binned into three size categories. The track density had units of molecules  $\text{area}^{-1} \text{time}^{-1}$ . Since adhesion-associated tracks had a track lifetime at least as long, if not longer, than non-adhesion-associated tracks, the track density is a reasonable estimate for concentration (molecules  $\text{area}^{-1}$ ). The track density was normalized to the average track density within all adhesions, resulting in a dimensionless quantity we label 'relative track density'. Tracks were considered recruited to adhesions if the first frame of the track was inside an adhesion. The track recruitment density was normalized to the average recruitment density across all adhesions. The adhesion size bins were chosen to reflect previously described estimates for nascent adhesions ( $0.05\text{-}0.25 \mu\text{m}^2$ ), focal complexes ( $0.25\text{-}1 \mu\text{m}^2$ ), and focal adhesions ( $1\text{-}5 \mu\text{m}^2$ ) (Gardel et al., 2010).

#### 4.7.4. Cluster analysis

Tracks were assigned to proposed clusters, subjected to spatial and temporal refinement steps, then filtered to remove spurious clusters. Cluster proposals were generated using Voronoi tessellation (Andronov et al., 2016; Levet et al., 2015) of the track centroids. Track centroids, rather than whole trajectories, were used as the input to the tessellation process to avoid bias for longer-lived tracks and to avoid false-positive cluster detection due to repeated observation of a slow-moving molecule. Reducing

the dataset to track centroids also allowed much more efficient computation. To select clustered Voronoi polygons (islands of small polygons), polygon size thresholds were obtained as described by Andronov et al. Briefly, for each centroid map, 50 realizations of randomly distributed points were subjected to Voronoi tessellation, and probability density functions of the experimental and mean simulated polygon sizes were computed. Their intersection defined the threshold.

Proposed clusters were used as the basis for spatial and temporal refinement. In spatial refinement, tracks left out of the proposal were added to a proposed cluster if their trajectories entered a polygon island. In temporal refinement, proposed clusters were split if enough time passed between new observations of molecules. Since a new molecule was typically observed at a cluster at least once every 4 seconds (>90% of all recruitment events, prior to temporal segregation), any cluster with a longer time gap was split into two. To remove spurious clusters, we ignored clusters that had fewer than 10 tracks, as well as clusters that contained 3 or more tracks at the same instant in time.

The usual correlation analysis methods to assess cluster size for super-resolution microscopy do not account for repeated localization of moving molecules. For example, in PC-PALM (Sengupta et al., 2011), the term  $g(r)^{\text{stoch}}$  is a convolution of the protein pair correlation function at  $r=0$  (i.e. a non-diffusing molecule) and the pair correlation function of the effective point spread function. However, we did apply PC-PALM to analyze tagSrc and Binder clusters in fixed cells for comparison to our live-cell results, and pair-correlation analyses to sets of single-coordinate representations of trajectories (e.g., track centroids). As another approach, we computed kernel density estimates (KDE) using the localizations associated with sets of clustered tracks. Gaussian kernels with optimized bandwidths were used (Bowman and Azzalini, 1997). The domain was padded out by 1  $\mu\text{m}$  in each direction to avoid boundary artifacts, and calculations were performed on 10 nm x 10 nm pixelated grids. After computing a KDE for each cluster, each density map was thresholded at 5% and 50% of its maximum to generate binary masks with areas and effective radii  $r = \sqrt{A/\pi}$ . Simulations were primarily performed using Linux systems (Longleaf

cluster at UNC Chapel Hill, 2.50 GHz and 2.30 GHz Intel Processors), with some subsequent analyses conducted on a Mac (3.4 GHz Intel Processor).

#### 4.7.5. Co-diffusion events and kinetic modeling

To track individual molecules, it was necessary to use low dye concentrations to label a sub-population of the target proteins (tagSrc and Binder), such that not all tagSrc and not all Binder are fluorescent. Whether dye-labeled or not, all Binder events on the membrane should have a corresponding tagSrc. Therefore, when quantifying the activation kinetics of Src, we selected for all tagSrc trajectories that co-diffuse (distance < 100 nm per frame) with another Binder trajectory for at least five consecutive frames (100 ms) during their residence on PM.

The kinetic model of Src regulation has been modeled as a discrete-time Markov chain with four states: (1) closed tagSrc at the plasma membrane, (2) open tagSrc at the plasma membrane, (3) open tagSrc with Binder at the plasma membrane, and (4) an absorbing state representing translocation into the cytosol. The transition probability matrix is:

$$P = \begin{bmatrix} P_{aa} & 1 - \exp(-k_1\Delta t) & 1 - \exp(-k_5\Delta t) & 1 - \exp(-k_7\Delta t) \\ 1 - \exp(-k_2\Delta t) & P_{bb} & 1 - \exp(-k_3\Delta t) & 1 - \exp(-k_8\Delta t) \\ 1 - \exp(-k_6\Delta t) & 1 - \exp(-k_4\Delta t) & P_{cc} & 1 - \exp(-k_8\Delta t) \\ 0 & 0 & 0 & 1 \end{bmatrix},$$

where  $\Delta t$  is the simulated time step, and  $P_{aa}, P_{bb}, P_{cc}$ , are calculated such that the rows sum to one.

The timestep is selected such that  $P_{ij} < 0.1$ , for  $i \neq j$ . Markov chain simulations using a particular parameter vector  $\vec{\theta}$  result in a set of state vectors  $\{\vec{s}_1, \vec{s}_2, \dots, \vec{s}_n\}$  corresponding to the time-evolution of the Markov chain, with each state vector  $\vec{s}_i = \{s_{i1}, s_{i2}, \dots, s_{ij}\}$  a sequence of the four states. Entry to the absorbing state always terminates the vector. The challenge is to find an optimal parameter vector  $\vec{\theta}$  to match the behavior of the state vectors to the experimental data.

Simulations of the Markov chain were performed several thousand times using each of the three non-absorbing states as initial conditions. The relative proportion of events starting from the three states is weighted by the additional parameters  $f_1, f_2$ , and  $f_3$ , which sum to one. Of all co-diffusion events, 24%

started with both tagSrc and Binder together, so  $f_3$  was fixed to 0.24. Furthermore,  $k_5$  was fixed to zero. so the goal is to optimize the parameter vector  $\vec{\theta} = \{k_1, k_2, k_3, k_4, k_5, k_7, k_8, f_2\}$ . Simulated trajectories  $\{\vec{s}_1, \vec{s}_2, \dots, \vec{s}_n\}$  were analyzed to produce simulated distributions with the same constraints as the experimental data: tracks shorter than 12 frames (240 ms) were ignored; tracks where Binder arrived, departed, then arrived again were ignored; and unless Binder was present for >12 consecutive frames, the track was not considered a co-diffusion event. An evolutionary algorithm approach was to fit the data (Fortin et al., 2012). A single evolution consisted of 100 parameter sets (“individuals”) that competed, mated, and mutated over 40 cycles (“generations”). Three-individual tournaments were used for mating, where the two best-fitting individuals mixed parameters to produce a new individual, while the worst-fitting individual was removed. Crossover and mutation rates were set to 0.5 and 0.1, respectively. Fits were quantified using the following score function.

$$\begin{aligned} \text{Score} = & \text{MSE}(\text{process\_i}) + \text{MSE}(\text{process\_ii}) + \text{MSE}(\text{process\_iii}) + \text{MSE}(\text{process\_iii}) \\ & + \text{MSE}(\text{tagSrc\_lifetime\_codiffuse}) + \text{MSE}(\text{tagSrc\_lifetime\_all}) \\ & + \text{MSE}(\text{Binder\_lifetime\_codiffuse}) + \text{MSE}(\text{codiffusion\_sequence\_category}) \end{aligned}$$

Where the  $\text{MSE}(\bullet)$  is the mean-squared error between experimental and simulated distributions. The initial range for parameter optimization was chosen based on exponential fits to a simplified model. Evolutions were conducted using Linux systems (Longleaf cluster at UNC Chapel Hill, 2.50 GHz and 2.30 GHz Intel Processors), each taking at minimum 2-3 days per CPU node, with subsequent analyses conducted on a Mac (3.4 GHz Intel Processor).

#### 4.7.6. Statistical tests

Unless otherwise mentioned, statistical tests were performed in MATLAB. ANOVAs were performed with Tukey-Kramer correction for multiple hypotheses. Multiple t-tests performed within the same plot were subjected to Bonferroni correction. Significance levels: n.s.,  $P > 0.05$ ; \* $P < 0.05$ , \*\* $P < 0.01$ , \*\*\* $P < 0.001$ , \*\*\*\* $P < 0.0001$ .

## REFERENCES

- Acosta, L.C., Perez Goncalves, G.M., Pielak, G.J., and Gorenssek-Benitez, A.H. (2017). Large cosolutes, small cosolutes, and dihydrofolate reductase activity. *Protein Sci.* 26, 2417–2425.
- Andronov, L., Orlov, I., Lutz, Y., Vonesch, J.-L., and Klaholz, B.P. (2016). ClusterViSu, a method for clustering of protein complexes by Voronoi tessellation in super-resolution microscopy. *Sci. Rep.* 6, 24084.
- Baumgart, F., Arnold, A.M., Rossboth, B.K., Brameshuber, M., and Schütz, G.J. (2018). What we talk about when we talk about nanoclusters. *Method Appl Fluoresc* 7, 013001.
- Bowman, A.W., and Azzalini, A. (1997). *Applied Smoothing Techniques for Data Analysis* (New York, NY: Oxford University Press Inc.).
- Changede, R., and Sheetz, M. (2017). Integrin and cadherin clusters: A robust way to organize adhesions for cell mechanics. *Bioessays* 39, 1–12.
- Changede, R., Xu, X., Margadant, F., and Sheetz, M.P. (2015). Nascent Integrin Adhesions Form on All Matrix Rigidities after Integrin Activation. *Dev. Cell* 35, 614–621.
- Chu, P.-H., Tsygankov, D., Berginski, M.E., Dagliyan, O., Gomez, S.M., Elston, T.C., Karginov, A.V., and Hahn, K.M. (2014). Engineered kinase activation reveals unique morphodynamic phenotypes and associated trafficking for Src family isoforms. *Proc. Natl. Acad. Sci. USA* 111, 12420–12425.
- Cooper, J.A., Gould, K.L., Cartwright, C.A., and Hunter, T. (1986). Tyr527 is phosphorylated in pp60c-src: implications for regulation. *Science* 231, 1431–1434.
- Crocker, J.C., and Grier, D.G. (1996). Methods of Digital Video Microscopy for Colloidal Studies. *J. Colloid Interface Sci.* 179, 298–310.
- D’Errico, J. (2004). *Inpaint nans* (MATLAB Central FileExchange: Mathworks).
- Ferrari, R., Manfroi, A.J., and Young, W.R. (2001). Strongly and weakly self-similar diffusion. *Physica D: Nonlinear Phenomena* 154, 111–137.
- Fincham, V.J., Unlu, M., Brunton, V.G., Pitts, J.D., Wyke, J.A., and Frame, M.C. (1996). Translocation of Src kinase to the cell periphery is mediated by the actin cytoskeleton under the control of the Rho family of small G proteins. *J. Cell Biol.* 135, 1551–1564.
- Fortin, F.-A., Rainville, F.-M.D., Gardner, M.-A., Parizeau, M., and Gagné, C. (2012). DEAP: Evolutionary Algorithms Made Easy. *Journal of Machine Learning Research*.
- Freedberg, D.I., and Selenko, P. (2014). Live cell NMR. *Annu. Rev. Biophys.* 43, 171–192.
- Gao, J., Wang, Y., Cai, M., Pan, Y., Xu, H., Jiang, J., Ji, H., and Wang, H. (2015). Mechanistic insights into EGFR membrane clustering revealed by super-resolution imaging. *Nanoscale* 7, 2511–2519.
- García-Contreras, R., Vos, P., Westerhoff, H.V., and Boogerd, F.C. (2012). Why in vivo may not equal in vitro - new effectors revealed by measurement of enzymatic activities under the same in vivo-like assay conditions. *FEBS J.* 279, 4145–4159.



- Gardel, M.L., Schneider, I.C., Aratyn-Schaus, Y., and Waterman, C.M. (2010). Mechanical integration of actin and adhesion dynamics in cell migration. *Annu. Rev. Cell Dev. Biol.* 26, 315–333.
- Githaka, J.M., Vega, A.R., Baird, M.A., Davidson, M.W., Jaqaman, K., and Touret, N. (2016). Ligand-induced growth and compaction of CD36 nanoclusters enriched in Fyn induces Fyn signaling. *J. Cell Sci.* 129, 4175–4189.
- Grecco, H.E., Schmick, M., and Bastiaens, P.I.H. (2011). Signaling from the living plasma membrane. *Cell* 144, 897–909.
- Griffié, J., Burn, G.L., Williamson, D.J., Peters, R., Rubin-Delanchy, P., and Owen, D.M. (2018). Dynamic Bayesian Cluster Analysis of Live-Cell Single Molecule Localization Microscopy Datasets. *Small Methods* 1800008.
- Grima, R., and Schnell, S. (2006). A systematic investigation of the rate laws valid in intracellular environments. *Biophys Chem* 124, 1–10.
- Hodgson, L., Pertz, O., and Hahn, K.M. (2008). Design and optimization of genetically encoded fluorescent biosensors: GTPase biosensors. *Methods Cell Biol.* 85, 63–81.
- Huveneers, S., and Danen, E.H.J. (2009). Adhesion signaling - crosstalk between integrins, Src and Rho. *J. Cell Sci.* 122, 1059–1069.
- Kalay, Z., Fujiwara, T.K., and Kusumi, A. (2012). Confining domains lead to reaction bursts: reaction kinetics in the plasma membrane. *PLoS One* 7, e32948.
- Karginov, A.V., Tsygankov, D., Berginski, M., Chu, P.-H., Trudeau, E.D., Yi, J.J., Gomez, S., Elston, T.C., and Hahn, K.M. (2014). Dissecting motility signaling through activation of specific Src-effector complexes. *Nat. Chem. Biol.* 10, 286–290.
- Kmiecik, T.E., and Shalloway, D. (1987). Activation and suppression of pp60c-src transforming ability by mutation of its primary sites of tyrosine phosphorylation. *Cell* 49, 65–73.
- Koudelková, L., Pataki, A.C., Tolde, O., Pavlik, V., Nobis, M., Gemperle, J., Anderson, K., Brábek, J., and Rosel, D. (2019). Novel FRET-Based Src Biosensor Reveals Mechanisms of Src Activation and Its Dynamics in Focal Adhesions. *Cell Chem. Biol.* 26, 255–268.e4.
- Levet, F., Hosy, E., Kechkar, A., Butler, C., Beghin, A., Choquet, D., and Sibarita, J.-B. (2015). SR-Tesseler: a method to segment and quantify localization-based super-resolution microscopy data. *Nat. Methods* 12, 1065–1071.
- Medalia, O., Weber, I., Frangakis, A.S., Nicastro, D., Gerisch, G., and Baumeister, W. (2002). Macromolecular architecture in eukaryotic cells visualized by cryoelectron tomography. *Science* 298, 1209–1213.
- Murakoshi, H., Iino, R., Kobayashi, T., Fujiwara, T., Ohshima, C., Yoshimura, A., and Kusumi, A. (2004). Single-molecule imaging analysis of Ras activation in living cells. *Proc. Natl. Acad. Sci. USA* 101, 7317–7322.
- Owen, D.M., Rentero, C., Rossy, J., Magenau, A., Williamson, D., Rodriguez, M., and Gaus, K. (2010). PALM imaging and cluster analysis of protein heterogeneity at the cell surface. *J. Biophotonics* 3,

446–454.

- Parsons, S.J., and Parsons, J.T. (2004). Src family kinases, key regulators of signal transduction. *Oncogene* 23, 7906–7909.
- Persson, F., Lindén, M., Unoson, C., and Elf, J. (2013). Extracting intracellular diffusive states and transition rates from single-molecule tracking data. *Nat. Methods* 10, 265–269.
- Playford, M.P., and Schaller, M.D. (2004). The interplay between Src and integrins in normal and tumor biology. *Oncogene* 23, 7928–7946.
- Pufall, M.A., and Graves, B.J. (2002). Autoinhibitory domains: modular effectors of cellular regulation. *Annu. Rev. Cell Dev. Biol.* 18, 421–462.
- Roskoski, R. (2005). Src kinase regulation by phosphorylation and dephosphorylation. *Biochem. Biophys. Res. Commun.* 331, 1–14.
- Roskoski, R. (2015). Src protein-tyrosine kinase structure, mechanism, and small molecule inhibitors. *Pharmacol. Res.* 94, 9–25.
- Rossier, O., Oceau, V., Sibarita, J.-B., Leduc, C., Tessier, B., Nair, D., Gatterdam, V., Destaing, O., Albigès-Rizo, C., Tampé, R., et al. (2012). Integrins  $\beta 1$  and  $\beta 3$  exhibit distinct dynamic nanoscale organizations inside focal adhesions. *Nat. Cell Biol.* 14, 1057–1067.
- Rous, P. (1911). A sarcoma of the fowl transmissible by an agent separable from the tumor cells. *J. Exp. Med.* 13, 397–411.
- Roy, R., Hohng, S., and Ha, T. (2008). A practical guide to single-molecule FRET. *Nat. Methods* 5, 507–516.
- Ryan, T.A., Myers, J., Holowka, D., Baird, B., and Webb, W.W. (1988). Molecular crowding on the cell surface. *Science* 239, 61–64.
- Sakon, J.J., and Weninger, K.R. (2010). Detecting the conformation of individual proteins in live cells. *Nat. Methods* 7, 203–205.
- Sandilands, E., Brunton, V.G., and Frame, M.C. (2007). The membrane targeting and spatial activation of Src, Yes and Fyn is influenced by palmitoylation and distinct RhoB/RhoD endosome requirements. *J. Cell Sci.* 120, 2555–2564.
- Schnell, S., and Turner, T.E. (2004). Reaction kinetics in intracellular environments with macromolecular crowding: simulations and rate laws. *Prog. Biophys. Mol. Biol.* 85, 235–260.
- Schönherr, R., Rudolph, J.M., and Redecke, L. (2018). Protein crystallization in living cells. *Biol. Chem.* 399, 751–772.
- Sengupta, P., Jovanovic-Talisman, T., Skoko, D., Renz, M., Veatch, S.L., and Lippincott-Schwartz, J. (2011). Probing protein heterogeneity in the plasma membrane using PALM and pair correlation analysis. *Nat. Methods* 8, 969–975.
- Shvartsman, D.E., Donaldson, J.C., Diaz, B., Gutman, O., Martin, G.S., and Henis, Y.I. (2007). Src kinase activity and SH2 domain regulate the dynamics of Src association with lipid and protein

- targets. *J. Cell Biol.* 178, 675–686.
- Smith, A.E., Zhang, Z., Pielak, G.J., and Li, C. (2015). NMR studies of protein folding and binding in cells and cell-like environments. *Curr. Opin. Struct. Biol.* 30, 7–16.
- Smith, A.W., Huang, H.H., Endres, N.F., Rhodes, C., and Groves, J.T. (2016). Dynamic organization of myristoylated src in the live cell plasma membrane. *J. Phys. Chem. B* 120, 867–876.
- Stehelin, D., Varmus, H.E., Bishop, J.M., and Vogt, P.K. (1976). DNA related to the transforming gene(s) of avian sarcoma viruses is present in normal avian DNA. *Nature* 260, 170–173.
- Stone, O.J. (2018). *New Approaches for Control and Visualization of Protein Conformation in Live Cells*. Doctoral dissertation. UNC Chapel Hill.
- Ting, A.Y., Kain, K.H., Klemke, R.L., and Tsien, R.Y. (2001). Genetically encoded fluorescent reporters of protein tyrosine kinase activities in living cells. *Proc. Natl. Acad. Sci. USA* 98, 15003–15008.
- Wah, D.A., Levchenko, I., Baker, T.A., and Sauer, R.T. (2002). Characterization of a specificity factor for an AAA<sup>+</sup> ATPase: assembly of SspB dimers with ssrA-tagged proteins and the ClpX hexamer. *Chem. Biol.* 9, 1237–1245.
- Wang, Y., Botvinick, E.L., Zhao, Y., Berns, M.W., Usami, S., Tsien, R.Y., and Chien, S. (2005). Visualizing the mechanical activation of Src. *Nature* 434, 1040–1045.
- Webb, D.J., Donais, K., Whitmore, L.A., Thomas, S.M., Turner, C.E., Parsons, J.T., and Horwitz, A.F. (2004). FAK-Src signalling through paxillin, ERK and MLCK regulates adhesion disassembly. *Nat. Cell Biol.* 6, 154–161.
- Wheeler, D.L., Iida, M., and Dunn, E.F. (2009). The role of Src in solid tumors. *Oncologist* 14, 667–678.
- Zaidel-Bar, R., Itzkovitz, S., Ma'ayan, A., Iyengar, R., and Geiger, B. (2007). Functional atlas of the integrin adhesome. *Nat. Cell Biol.* 9, 858–867.

## **CHAPTER 5: MODELING CLUSTERING-INDUCED ACTIVATION OF RECEPTOR IMMUNOTYROSINE ACTIVATION MOTIFS**

### **Overview**

Clustering is a regulatory feature in signal transduction across diverse organisms, appearing in both unicellular organisms and humans. Clustering, also described as oligomerization and aggregation, is the localized accumulation of many copies of the same molecule, and can effect signaling by inducing a response, broadening the dynamic range of stimulus-induced responses, insulating responses from nearby molecules, and temporally synchronizing responses (Bray et al., 1998; Changede and Sheetz, 2017; Duke and Bray, 1999; Dushek et al., 2011; Goyette et al., 2019; Hartman and Groves, 2011). The simplest and perhaps most classic example of clustering is the dimerization of receptor tyrosine kinases in humans (Lemmon and Schlessinger, 2010; Ullrich and Schlessinger, 1990). Canonically, ligand binding causes monomeric receptors to form signaling-competent dimers. Dimerization can be driven by direct ligand-receptor interactions (Wehrman et al., 2007), receptor-receptor interactions (Garrett et al., 2002; Ogiso et al., 2002), or both (Schlessinger et al., 2000). Higher-order oligomers are also key regulators of signal transduction, and have been observed among receptor tyrosine kinases such as Tie2 and ephrins (Barton et al., 2006; Himanen and Nikolov, 2003), as well as other classes of molecules such as integrins, cadherins (Changede and Sheetz, 2017), and immunoreceptors (Goyette et al., 2019; Holowka and Baird, 2015; Jaumouillé and Grinstein, 2011; Li et al., 2018).

Mammalian immunoreceptors including the T cell receptor (TCR), B cell receptor (BCR), Fc $\epsilon$  receptor, and Fc $\gamma$  receptor undergo ligand-induced clustering, followed by Src family kinase (SFK)-mediated phosphorylation of immunotyrosine activation motifs (ITAMs). ITAMs have two tyrosines that can be phosphorylated in a tandem YxxI/L motif, and are either noncovalently or covalently associated with the receptor (Daëron, 1997; Getahun and Cambier, 2015; Johnson et al., 1995; Love and Hayes,

2010). Dually phosphorylated ITAMs in Fc receptor signaling recruit spleen tyrosine kinase (Syk) for subsequent signal transduction, while the TCR and BCR recruit the homologue ZAP-70 (Mócsai et al., 2010). The number of ITAMs per receptor can vary dramatically across these systems. For example, there is only one ITAM per receptor in the single-chain FcγRIIA isoform (Nimmerjahn and Ravetch, 2008), but ten ITAMs per TCR in TCR:CD3 complexes (Guy and Vignali, 2009). How is receptor clustering relevant across these contexts?

## 5.1. Introduction

To begin, it is helpful to consider FcγR phagocytosis regulated by the FcγRIIA isoform. Fcγ receptor-mediated phagocytosis is the ingestion of large ( $\geq 0.5 \mu\text{m}$ ) particles coated ('opsonized') with immunoglobulin G (IgG) by FcγR-containing cells. The initial step in FcγR-mediated phagocytosis is target recognition, where the receptor binds the Fc domain of surface immobilized IgG molecules. Binding drives FcγR molecules into discrete, sub-micron clusters through an unknown mechanism (Flannagan et al., 2012; Sobota et al., 2005). This process somehow facilitates phosphorylation of the receptor ITAM(s) by Src family kinases (SFKs), initiating the phagocytic signaling cascade. Once phosphorylated, the ITAMs then recruit Syk, which can phosphorylate other ITAMs, and can also initiate downstream signaling through the kinase PI3K, the guanine nucleotide exchange factor Vav, and many adaptor proteins such as Grb2, Gab2, and LAT. The phagocytic receptor FcγRIIA is an excellent system for studying the role of receptor clustering in signal transduction. First, it is one of the best-characterized phagocytic receptors (Flannagan et al., 2012). Second, FcγRIIA is a single polypeptide with a single ITAM, in contrast to the multi-polypeptide, multi-ITAM complexes that comprise the TCR and BCR.

One hypothesis is that cluster formation facilitates the organization of liquid-ordered lipid domains ('lipid rafts' or 'detergent-resistant membranes', DRMs) around the receptors, thereby recruiting SFKs that have *N*-terminal acylated lipid moieties. Experiments in purified neutrophil plasma membrane indicated that clustering FcγRIIA, by crosslinking with 1° and 2° antibodies rather than IgG, caused

receptor enrichment in detergent-resistant membranes. The same set of experiments also indicated that SFKs and Syk were constitutively present in the DRMs, and that cholesterol extraction with methyl- $\beta$ -cyclodextrin caused loss of the SFK Lyn from the DRM, but did not block receptor localization (Rollet-Labelle et al., 2004). Another study showed that methyl- $\beta$ -cyclodextrin caused inhibition of Fc $\gamma$ RIIA phosphorylation without disrupting clustering (Kwiatkowska and Sobota, 2001). The association of Fc $\gamma$ RIIA with the DRMs was dependent upon receptor palmitoylation (García-García et al., 2007). However, no other Fc $\gamma$ Rs are palmitoylated, there are concerns over the side-effects of detergent treatments and cholesterol extraction on cells (Kenworthy, 2008).

A second hypothesis is that cluster formation induces conformational changes in the receptors, resulting in enhanced interactions with SFKs or repressed interactions with phosphatases. However, receptor clustering, phosphorylation, and signaling can be induced not only by exposure to surface-bound IgG (as might be encountered normally on opsonized targets), but also by heat-aggregated IgG, as well as by directly crosslinking the receptor through a combination of 1° anti-Fc $\gamma$ R + 2° antibody (Kwiatkowska and Sobota, 1999; Kwiatkowska et al., 2003). This suggests clustering can activate signaling in a non-specific way. Furthermore, treatment with only the primary anti-Fc $\gamma$ R antibody does not elicit a response, nor does monomeric IgG, which is thought to bind to Fc $\gamma$ R in a 1:1 stoichiometry (Woof and Burton, 2004). Altogether, the variety of ways to trigger a response across different methods of receptor clustering suggests a more generic mechanism than binding-induced conformational changes.

We hypothesized that clustering simply may facilitate repeated encounters with kinases, without altering any affinities for signaling molecules. Diffusion in the membrane is slow relative to diffusion in the cytoplasm. If receptors are sparsely distributed (i.e. unclustered) then a kinase that translocates to the plasma membrane may be unable to diffuse to and phosphorylate many receptors prior to dissociation back to the cytoplasm. However, if receptors are clustered, the search process may be accelerated. This activation must be balanced by inactivation by phosphatases, which will also have an easier time modifying clustered receptors. Furthermore, since fully phosphorylated ITAMs can recruit a kinase that

can also phosphorylate nearby receptors, the cell must somehow balance these interactions to avoid spurious activation due to noise.

Here, I used computational modeling and a simple receptor signaling model to investigate how clustering could cause receptor ITAMs to become phosphorylated without relying upon preferential recruitment of kinases (i.e. via lipid rafts) and without a change in intrinsic biochemical affinities (i.e. a conformational change). Receptors, kinases, and phosphatases were simulated as stochastically diffusing and reacting particles. Clusters were simulated by placing a fixed number of receptors ( $n$ ) into confined zones with  $2 \leq n \leq 20$  receptors per zone, while allowing kinases and phosphatases to diffuse in and out. We used an evolutionary algorithm approach to identify kinetic rate parameters that could generate robust phosphorylation of highly clustered receptors, while suppressing phosphorylation of unclustered receptors. Changing the objective function of the evolutionary algorithm allowed us to tune the clustering-based activation threshold. We find two types of cluster activation in our models: full and committed activation of the cluster, versus partial and indecisive activation of the cluster. Full, committed activation of the cluster was common when the clustering-induced activation threshold was low, while partial and indecisive activation dominated when the threshold was higher. We hope this study will provide a useful framework for studying clustering-induced receptor activation.

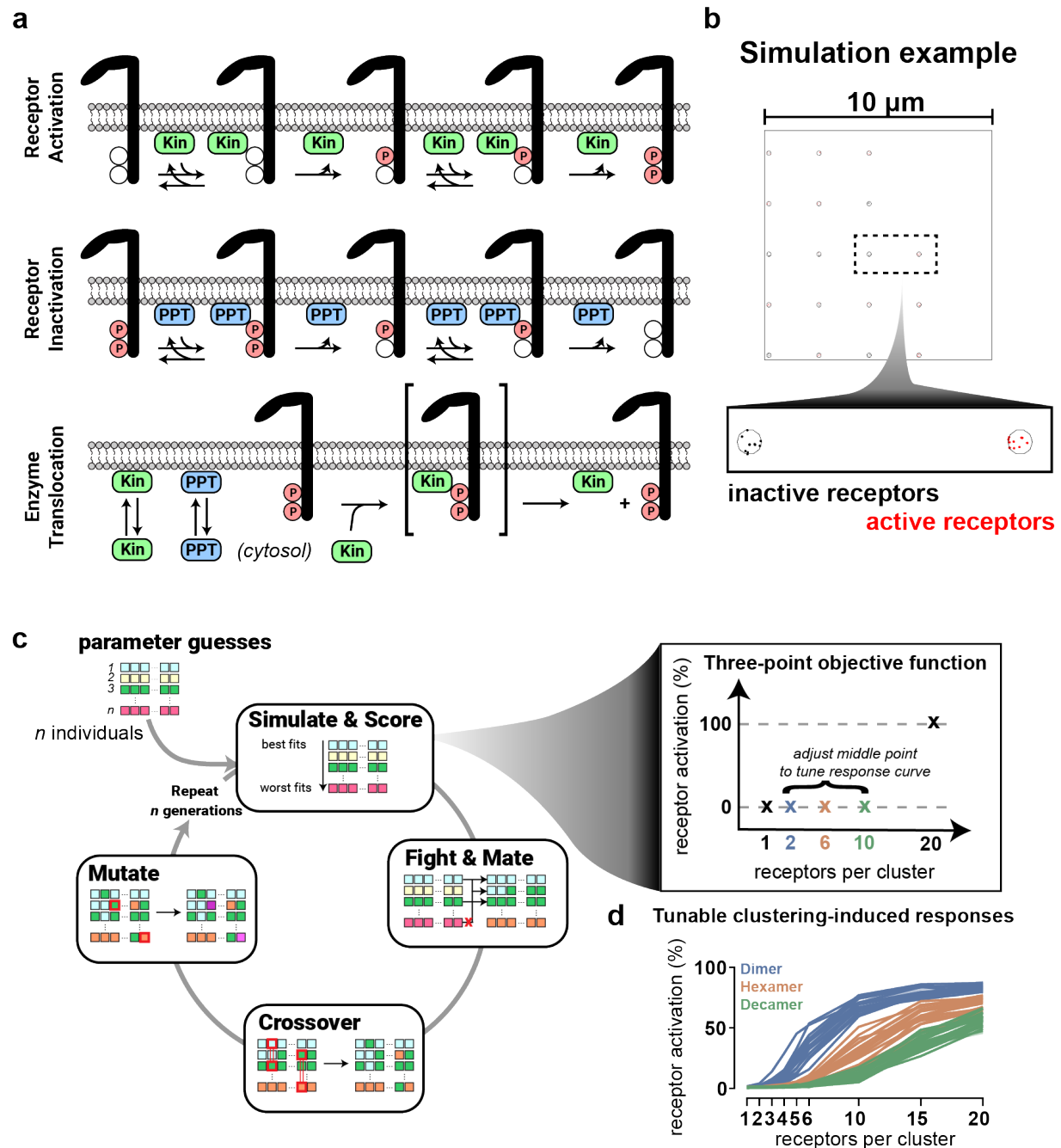
## **5.2. Clustering-mediated receptor activation with tunable response thresholds**

We used a particle-based modeling approach, where each molecule is simulated as a discrete particle subject to Brownian motion and probabilistic chemical reactions. Particle-based modeling is computationally expensive and is often as a follow up to less detailed but more efficient approaches such as reaction-diffusion systems of differential equations (Pablo et al., 2018). However, there are three major advantages to using a particle-based to study receptor clustering. First, we are interested in behaviors ranging from monomers and dimers of receptors up through larger clusters containing approximately 20 receptors. Particle-based approaches are well-suited to modeling stochasticity, and stochasticity is substantial in this low-abundance limit. Second, the clusters are multimers of receptors that can be in

many distinct molecular states, e.g. phosphorylation state and enzyme-binding state. The number of possible states for the overall cluster grows very rapidly with the number of receptors included per cluster. The combinatorial explosion of possibilities makes it impractical to write out all the state variables, and is a common problem in aggregating systems (Sneddon et al., 2011). Third, our hypothesis involves diffusion-limited interactions on a membrane. Diffusion-limited reactions on membranes are difficult to model in the continuum limit (Halatek and Frey, 2018; Yogurtcu and Johnson, 2015).

In our model, each receptor has two phosphorylation sites, similar to FcγRIIA. Cytosolic kinases and phosphatases can associate with the membrane, where they can form encounter complexes with the phosphorylation sites and subsequently catalyze a phosphorylation/dephosphorylation event. These enzymes are assumed to fall off after catalysis, requiring another binding event. Finally, as an approximation of Syk-mediated positive feedback, where Syk can bind to dually-phosphorylated ITAMs and phosphorylate nearby ITAMs (Mócsai et al., 2010; Pao et al., 1997; Turner et al., 2000), our model allows dually-phosphorylated receptors to recruit cytosolic SFK to the membrane. These dually-phosphorylated receptors are thus considered active; non-phosphorylated and mono-phosphorylated receptors are not. All the reactions of our model are shown in Figure 5.1A, and a simulation snapshot is shown in Figure 5.1B. Not all parameter choices for the model will result in clustering-mediated activation. An evolutionary algorithm approach was used to train our model to exhibit clustering-mediated activation (Figure 5.1C). Because the absolute number of receptors per cluster needed to trigger a response is unknown for FcγRIIA, and furthermore is different for other receptors, we checked whether our model could be trained respond with a variety of thresholds. Adjusting the objective function produced sets of parameters that allowed activation at different levels of clustering (Figure 5.1C,D).



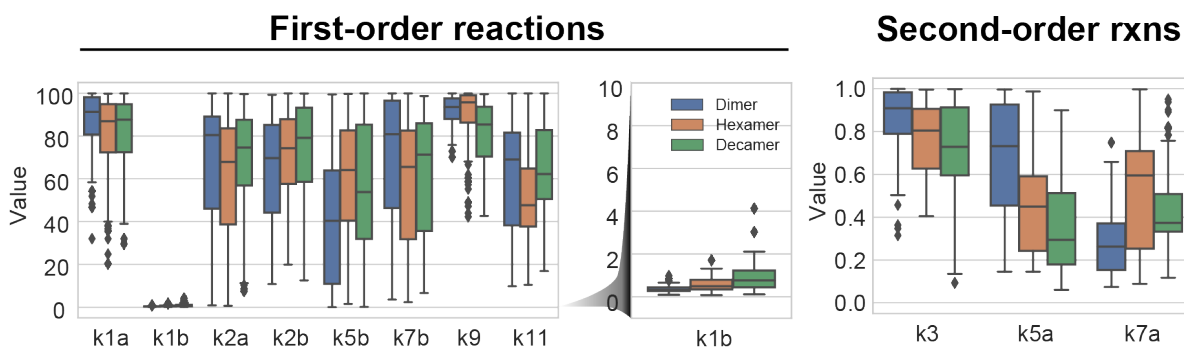


**Figure 5.1. Simulating clustering-mediated activation of receptors.**

(A) Signaling schematic. Kin: kinase; PPT: phosphatase; P: phosphorylated. (B) Simulation snapshot, showing clusters of 10 receptors each. (C) Evolutionary algorithm schematic and adjustable objective function. An optimal simulation has no activation with receptor monomers, full activation with receptor 20-mers, and no activation at some adjustable, intermediate level of clustering. Each evolution contained 200 individuals and was carried out over 50 generations (Methods). (D) Different cluster-response curves obtained by adjusting the objective function, with each line representing the best individual found from an independent evolution. Dimer, Hexamer, and Decamer refer to the adjusted objective shown in (C).

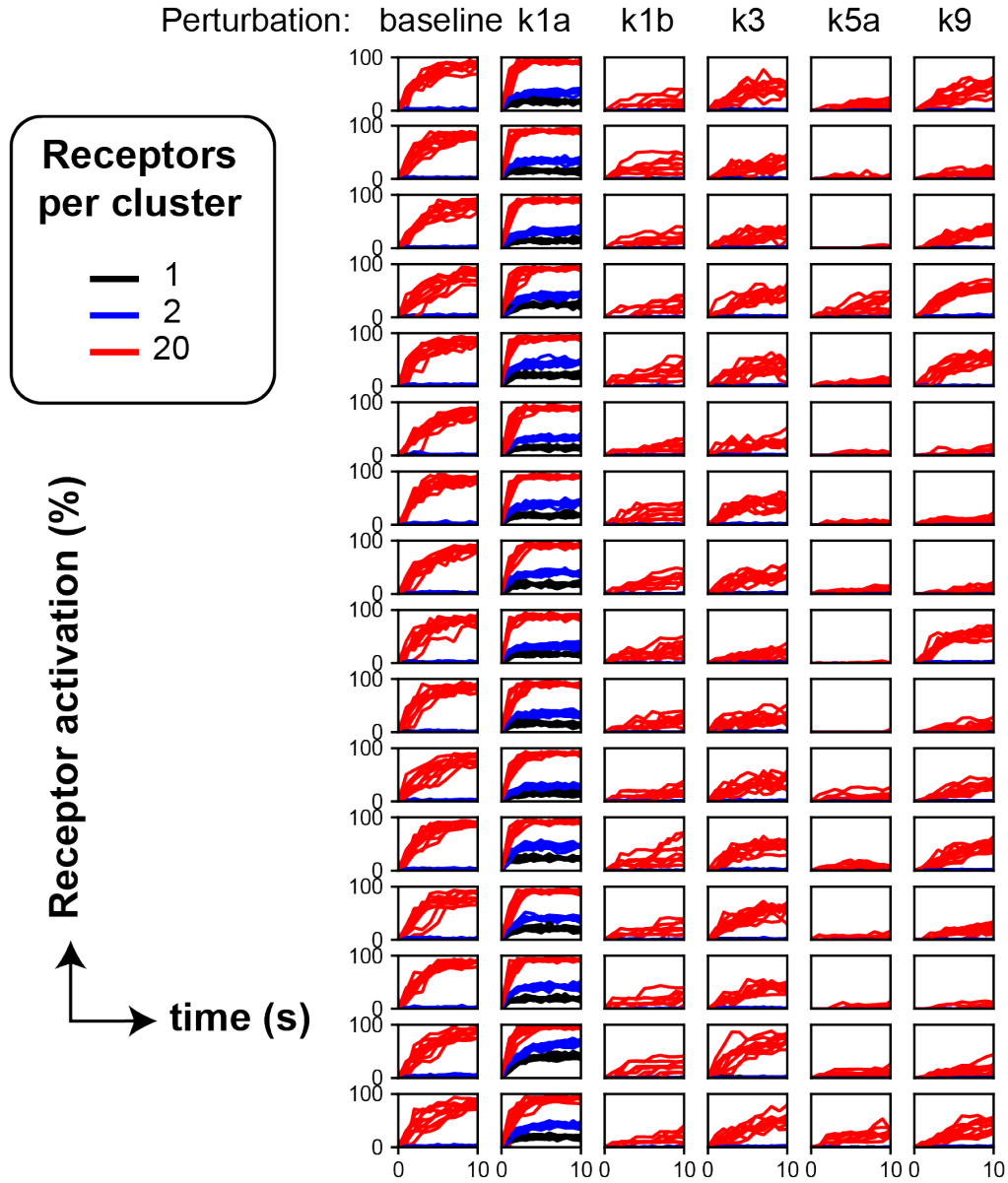
### 5.3. Investigating the role of specific parameters in controlling the clustering threshold

In an attempt to understand the kinetic mechanisms governing changes in the clustering threshold, we visualized the distributions of evolved parameters for each condition (Figure 5.2). Constraints and/or correlations with changing cluster threshold were observed for several parameters. For example, the plasma membrane association rate  $k_{1b}$  was slow, and the dissociation rate for the kinase  $k_{1a}$  was fast; the kinase catalysis  $k_9$  was fast, and phosphatase-receptor association was slow relative to kinases ( $k_{7a} < k_{5a}$ ). The feedback rate  $k_3$  and the kinase catalysis rate  $k_{5a}$  both decreased with the increasing threshold. These parameter constraints are consistent with the proposed mechanism of clustering-induced activation. Because of the limited membrane-cytoplasm exchange, kinases are relatively slow to interact with and explore the membrane, making it difficult to activate sparsely distributed receptors. However, if the kinase reaches a cluster of receptors, it can quickly phosphorylate the group of receptors. The phosphatase interactions must be slow relative to the kinase to limit deactivation through the same mechanism.



**Figure 5.2. Parameter distributions obtained at different levels of clustering-triggered activation.** The first order reaction parameters were allowed to take on values between 0 and 100, while the second order reaction parameters were allowed to take on values between 0 and 1 (see Section 5.6.1). Dimer, Hexamer, and Decamer refer to the variable midpoint constraint in the objective function (Figure 5.1C). Data represent the top 50 individuals from 17, 20, and 17 independent evolutions for the Dimer, Hexamer, and Decamer cases respectively.

We perturbed several of these parameters to confirm their role in controlling the activation response. Parameter sets obtained at the most sensitive response threshold (“Dimer”) were altered to have ten-fold reduced  $k_{1a}$ ,  $k_{1b}$ ,  $k_3$ ,  $k_{5a}$ , or  $k_9$ , then repeated our simulations. We expect all of these perturbations – except for the  $k_{1a}$  reduction – to suppress receptor activation.



**Figure 5.3. Effects of perturbing specific parameters one-by-one.**

The top “dimer”-trained set of parameters from 16 independent evolutions (rows) was perturbed in 5 different ways, reducing  $k_{1a}$ ,  $k_{1b}$ ,  $k_3$ ,  $k_{5a}$ , and  $k_9$  by 10-fold each (columns). Simulations were performed under three levels of clustering: 1, 2, and 20 receptors per cluster (black, blue, and red curves). In the 1 receptor/cluster case, the receptors are freely diffusing. Ten realizations of the simulation were performed to assess variation intrinsic to the parameter set.

As expected, we find that all the perturbations, except reducing  $k_{1a}$ , caused suppression of activity (Figure 5.3). Comparing the effect of the same perturbation across different evolved parameter sets (rows), reveals heterogeneity despite comparable baseline behaviors: reducing the kinase catalysis rate  $k_9$  resulted in little effect in some cases (row 9), but caused complete loss of activity in others (row 6). This

indicates different regulatory features are able to produce similar baseline responses. That is, some evolved parameter sets were less reliant on rapid kinase catalysis because other rate parameters could compensate.

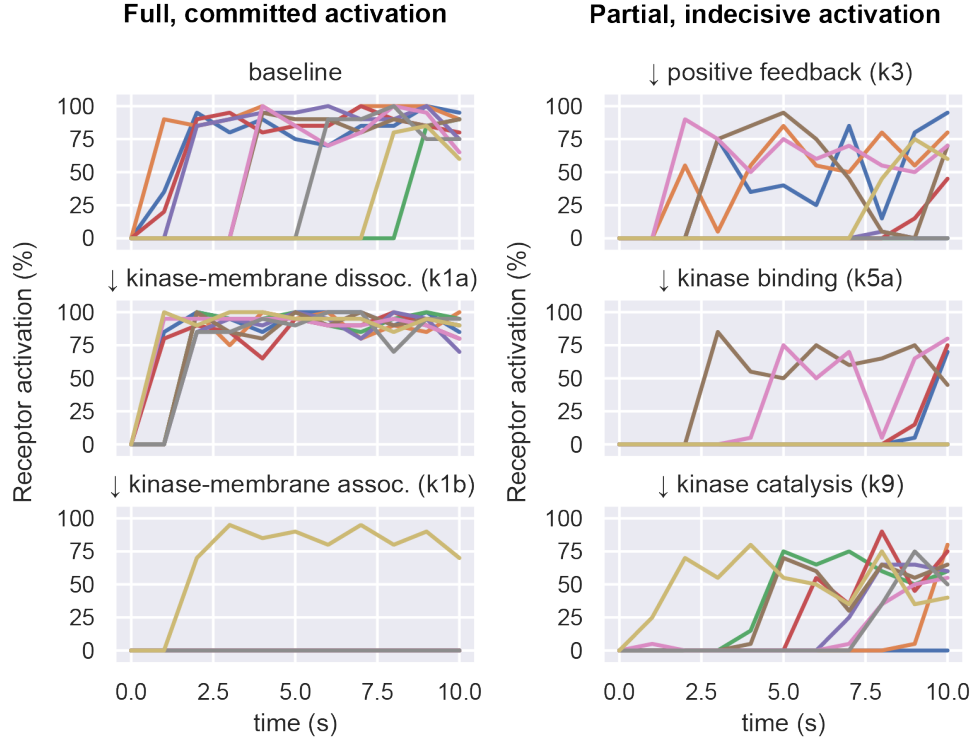
Experimentally, reducing the rate of kinase association to the membrane ( $k_{1b}$ ) might be achieved by perturbing Src family kinase localization in the plasma membrane through either cholesterol extraction with methyl- $\beta$ -cyclodextrin, or by inhibiting *N*-terminal SFK myristoylation/palmitoylation with DL- $\alpha$ -hydroxymyristic acid/2-bromopalmitic acid (Kwiatkowska et al., 2003; Nadler et al., 1993). Weakening the positive feedback ( $k_3$ ) might be approximated through a Syk inhibitor such as piceatannol (Oliver et al., 1994), while slowing the kinase catalysis rate ( $k_9$ ), might be achieved with ATP-competitive Src family kinase inhibitors such as PP2 or Dasatinib (Kong et al., 2011; Milano et al., 2009; Pichot et al., 2009; Roskoski, 2015; Zhu et al., 1999). Luckily, such experiments have been reported. Disrupting Src family kinase localization or kinase activity in U937 cells, a human monocytic cell line, resulted in diminished Fc $\gamma$ RIIA phosphorylation without visually perturbing receptor clustering (Kwiatkowska et al., 2003). In the same system, Syk inhibition with piceatannol does not seem to reduce Fc $\gamma$ RIIA phosphorylation (Kwiatkowska et al., 2003). More recent work performing higher-resolution, quantitative analysis in primary monocyte-derived macrophages and in mouse bone marrow-derived macrophages indicate Syk inhibition does not influence clustering (Lopes et al., 2017), though it seems to reduce receptor diffusion, which was speculated to have an effect on clustering (Jaumouillé et al., 2014). These studies also argue that Fc $\gamma$ RIIA clustering is affected by SFK inhibition. Perturbations of clustering are beyond the current scope of the model, but might indicate downstream feedback regulation of the clustering process.

Returning to the model, we also noticed small, step-like increases in the receptor activity when  $k_{1b}$  was reduced. We hypothesized that individual receptor clusters were becoming activated in rare, switch-like bursts, and therefore investigated the activation state of individual clusters.

#### **5.4. Full and committed vs. partial and indecisive cluster activation**

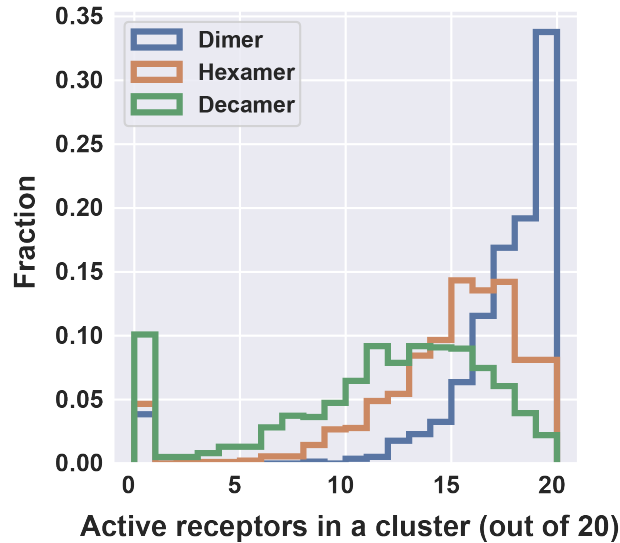
Examining individual cluster responses revealed two different underlying regulatory mechanisms: a “full and committed” cluster response, and a “partial and indecisive” cluster response (Figure 5.4). In the full response, individual clusters tended to become almost entirely active and then stay active. A set of clusters subject to full activation can exhibit intermediate activity on average, because not all clusters become activated at the same time. In the partial and indecisive regime, the degree of receptor activation exhibited large fluctuations around intermediate levels.

To better understand the relevance of the two cluster activation mechanisms, we returned to the three different activation thresholds (Figure 5.1C). We analyzed the number of active receptors at  $t=10$  seconds as a measure of the full versus partial activation behavior. Interestingly, it appears that the full activation mode is predominant when the activation threshold is low (>60% of clusters had receptors fully phosphorylated at our lowest activation threshold), but rare-to-nonexistent as the activation threshold rises (<10% of clusters had 18 of 20 ITAMs fully phosphorylated at our highest activation threshold). This reveals a transition between the mechanisms: full and committed cluster responses are dominant when the activation threshold is low, and partial and indecisive cluster responses are dominant when the activation threshold is high. This is probably because negative regulation (or weakened positive regulation) of clusters is necessary to raise the activation threshold (Figure 5.2). Because of the repression of activity, fully activated clusters are not stable.



**Figure 5.4. Full and committed versus partial and indecisive modes of cluster activation.**

A “dimer” parameter set was subjected to the five perturbations, as in Figure 5.4, and the per-cluster level of activation was visualized (colored lines distinguish individual clusters). Data represent a single evolved parameter set and one realization.



**Figure 5.5. Even under strongly-activating conditions, full activation dominates at the low activation threshold (“dimer”), but disappears at higher activation thresholds (“decamer”).**

Shown are the number of active, dually-phosphorylated receptors within individual clusters at  $t=10s$ . Dimer, Hexamer, and Decamer refer to the variable midpoint constraint in the objective function used to produce distinct thresholds of activation (Figure 5.1C). Simulations were conducted with 20 receptors per cluster, and represent 15 independent evolutions with 10 realizations each.

## 5.5. Discussion and Future Directions

We set out to investigate a simple mechanism that could regulate ITAM phosphorylation upon receptor clustering. The proposed mechanism relied upon a diffusion-limited kinase that could rapidly phosphorylate a cluster upon encounter. Since clustering could enhance the activity of both the kinase and phosphatase, a careful kinetic balance between the two was needed to control the activation threshold. If the kinase was too effective, small clusters would immediately be activated and only weak thresholds could be obtained; if the kinase was not effective enough, it would be outcompeted by the phosphatase and no activation would occur. Our evolutionary algorithm approach (Figure 5.1) facilitated the search for parameter regimes where clustering-induced activation was possible over a range of activation thresholds.

We perturbed several of the evolved parameters to investigate their roles in regulating receptor phosphorylation. As expected, allowing the kinase to stay on the membrane longer heightened the activation profile, while reducing kinase-receptor interactions caused reduced activation (Figure 5.3). Slowing kinase catalysis ( $k_9$ ), we see that some parameter sets lose activation, while others sustain activation, indicating that different parameters may compensate to maintain the same response.

Some simulation parameters exhibited step-like bursts of receptor activation, driven by individual clusters becoming fully activated at different times (Figure 5.3 and Figure 5.4). We find two types of cluster activation: a full, committed response, and a partial, indecisive response. A shift in the response mechanism occurs as the activation threshold is increased: at low thresholds, where few clustered receptors are needed to initiate a response, full activation dominates (Figure 5.5). However, at higher thresholds, where many clustered receptors are needed to initiate a response, partial activation dominates.

Looking forward, it will be interesting to constrain our model with experimentally-measured parameters, such as cluster sizes, molecular abundances, diffusion coefficients, cellular geometry, and rate constants. Since not all the parameters will be readily measurable, fitting the remaining unconstrained parameters should provide insights into the feasibility of our hypothesized mechanism. Pushing the model to fail, then carefully examining the reason for the failure, will highlight important regulatory mechanisms. For example, we may need to look beyond an abundance-based activation threshold, as

receptor clusters might respond instead to the size of the cluster or the concentration within the cluster. These other thresholds could be easily examined by our modeling approach.

#### 5.5.1. Simulation speed and model simplicity

Practically speaking, we kept our model simple to facilitate fast simulations compatible with the evolutionary algorithm approach. Manual exploration of the high-dimensional parameter space was unfruitful, and the evolutionary algorithm approach greatly facilitated unbiased exploration of the parameter space within the provided parameter constraints. Fast simulation times were critical, since the evolutionary algorithm involves many repeated simulations (200 individuals x 50 generations x 3 objective function evaluations = 3000 simulations per evolution), and our simulations tend to take 1-10 minutes, depending on the parameters. After distributing over 200 CPUs and accounting for communication overhead and synchronization, each evolution takes ~2 days (~9600 CPU hours) to complete. Introducing more complexity into the model, altering the timescale of the simulation, or even increasing the number of molecules will all require more compute time or computational optimization.

#### 5.5.2. Receptor co-clustering

Our model ignores the effects of receptor co-clustering. One example where co-clustering may be important is in cells expressing receptors with ITIMs, the immunotyrosine inhibition motif. ITIMs have a single tyrosine that recruits negative regulators when phosphorylated. For example, in macrophage phagocytosis, SIRP $\alpha$  has an ITIM that can recruit the negative regulator SHP-1 (Oldenberg et al., 2001). In primary human monocyte-derived macrophages containing Fc $\gamma$ RI, Fc $\gamma$ RII, and SIRP $\alpha$ , it has been shown that Fc $\gamma$ RI (but not Fc $\gamma$ RII) co-clusters with SIRP $\alpha$  (Lopes et al., 2017). Thus, additional layers of regulation can be obtained through co-clustering.



### 5.5.3. Receptor pre-clustering

Studies of immunoreceptors have shown that the receptors may have a basal level of clustering. For example, in primary human monocyte-derived macrophages, Fc $\gamma$ RIIA forms clusters when not stimulated (seeded onto poly-L-lysine coated coverslips) and when stimulated (seeded onto IgG-coated coverslips) (Lopes et al., 2017). Stimulation allowed the clusters to grow larger over time. In murine B cells, electron microscopy and combined with Monte Carlo simulations demonstrated the existence of B cell receptor pre-clusters (Fiala et al., 2013). Pre-clusters of the T cell receptor have also been described (Alarcón et al., 2006; Castro et al., 2014). Understanding the basal and stimulated levels of receptor clustering will be important for applying our model to specific biological systems.

### 5.5.4. Downstream feedback regulation of clustering and activation

Because our model focuses on the earliest stages of signal transduction during receptor clustering, we have ignored the effects of downstream signaling events. Downstream signaling may exert negative and/or positive feedback on essentially all parts of our pathway. For example, in Fc $\gamma$ R-mediated phagocytosis, reorganization of actin and integrin can cause both enhance clustering (Jaumouillé et al., 2014; Lopes et al., 2017) and exclude negative regulators (Freeman et al., 2016), though at longer timescales than what we have considered. These feedbacks might be approximated in the model by increasing the number of receptors within a cluster proportionally with the number of activated receptors. More complicated approximations may be needed in other systems: for example, in T cells, clusters of the T cell receptor form at the cell periphery, then become transported to the cell center to form larger clusters in the central supramolecular activation cluster (Varma et al., 2006).

Experimentally separating the regulation of the initial clustering events here versus longer-term feedback regulation will require rapid imaging of the first contact between a cell and the cluster-inducing substrate. Furthermore, understanding the oligomerization states of receptors prior to stimulation will be important in establishing how a response is triggered upon further clustering.

### 5.5.5. Noise in receptor abundance and activation

The particle-based simulations identified a mechanism for clustering-induced signaling, but individual simulations were conducted with a fixed number of receptors per cluster. That is, the simulations ignored dynamic changes in the clustered receptor abundance. Preliminary experiments from Bei Liu (not shown) have shown that receptors can diffuse into and away from clusters.

To model the noise in receptor abundance and activation, we set up analytic equations to investigate the fluctuations in the number of each molecular species. The analysis relies on an ordinary differential equation model of receptor activation including mass flux to/from the cluster. For simplicity of the analysis, we omit concentrations of kinase and phosphatase and treat those modifications as simple first-order reactions. Since we are modeling only the individual cluster rather than the whole cell, this should be sufficient. The equations are listed below, along with their rates.

Reaction	Rate
$\emptyset \rightarrow R$	$k_1 + k_2 ppR$
$R \rightarrow \emptyset$	$k_3 R$
$pR \rightarrow \emptyset$	$k_3 pR$
$ppR \rightarrow \emptyset$	$k_3 ppR$
$R \rightarrow pR$	$k_4 R$
$pR \rightarrow R$	$k_5 pR$
$pR \rightarrow ppR$	$k_6 pR$
$ppR \rightarrow pR$	$k_7 ppR$

To couple dynamic changes in the degree of clustering to the activation state of the cluster, we introduced a feedback parameter, such that the rate of receptor influx scaled with the concentration of active receptor ( $k_2$ ). Analytic equations for the fluctuations in receptor abundance and in receptor activity can be calculated as the statistical moments of each of the species  $R$ ,  $pR$ , and  $ppR$ . We used the linear noise approximation to compute statistical moments of all the species (Van Kampen, 2007; Singh and Grima, 2017). The mean-centered second-order moments  $C_{ik}$  between the species  $i$  and  $k$  can be related to the coefficient of variation by normalization to the mean, which is simply computed from the steady-state solution of the system of ordinary differential equations. The moments at steady-state can be approximated using:

$$\sum_{j=1}^M (J_{ij}C_{jk} + C_{ij}J_{kj}) + D_{ik} = 0$$

with

$$J_{ik} = \frac{\partial}{\partial c_k} \sum_{j=1}^M \psi_{ij}(\vec{c})$$

$$D_{ik} = \sum_{j=1}^M \psi_{ij}(\vec{c}) S_{kj}$$

where  $c_k$  is the concentration of species  $k$ ,  $\vec{c}$  is a vector of species concentrations,  $\psi_{ij}$  is the signed rate of change experienced by the  $i$ -th species due to the  $j$ -th reaction, and  $S_{kj}$  is the net change in the number of molecules of the  $k$ -th species  $k$  due to the  $j$ -th reaction. For our system, the matrices  $\mathbf{S}$ ,  $\boldsymbol{\psi}$ ,  $\mathbf{J}$ , and  $\mathbf{D}$  are:

$$\mathbf{S} = \begin{bmatrix} 1 & -1 & 0 & 0 & -1 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 & 1 & -1 & -1 & 1 \\ 0 & 0 & 0 & -1 & 0 & 0 & 1 & -1 \end{bmatrix}$$

$$\boldsymbol{\psi} = \begin{bmatrix} k_1 + k_2\text{ppR} & -k_3\text{R} & 0 & 0 & -k_4\text{R} & k_5\text{pR} & 0 & 0 \\ 0 & 0 & -k_3\text{pR} & 0 & k_4\text{R} & -k_5\text{pR} & -k_6\text{pR} & k_7\text{ppR} \\ 0 & 0 & 0 & -k_3\text{ppR} & 0 & 0 & k_6\text{pR} & -k_7\text{ppR} \end{bmatrix}$$

$$\mathbf{J} = \begin{bmatrix} -k_3 - k_4 & k_5 & k_2 \\ k_4 & -k_3 - k_5 - k_6 & k_7 \\ 0 & k_6 & -k_3 - k_7 \end{bmatrix}$$

$$\mathbf{D} = \begin{bmatrix} k_1 + k_2\text{ppR} + (k_3 + k_4)\text{R} + k_5\text{pR} & -k_4 - k_5\text{pR} & 0 \\ -k_4 - k_5\text{pR} & (k_3 + k_5 + k_6)\text{pR} + k_4\text{R} + k_7\text{ppR} & -k_6\text{pR} - k_7\text{ppR} \\ 0 & -k_6\text{pR} - k_7\text{ppR} & (k_3 + k_7)\text{ppR} + k_6\text{pR} \end{bmatrix}$$

We obtained analytic equations for the moments by substituting these matrices into the equations above. However, the resulting equations were extremely complicated, and did not provide any clarity with respect to the system fluctuations. Using these equations to gain insight into noise regulation may require non-dimensionalization and clever term-grouping to aid analysis.

## 5.6. Methods

### 5.6.1. Particle-based simulation

Simulations were performed with Smoldyn (v2.61) (Andrews and Bray, 2004; Andrews et al., 2010) on Linux systems (2.4 GHz Intel processors on the Dogwood cluster, 2.50 GHz and 2.30 GHz Intel processors on the Longleaf cluster, both at UNC Chapel Hill, Chapel Hill, NC, USA). The Dogwood cluster was used for the evolutionary algorithm approach because it required parallelization across many CPUs (see below) while the Longleaf cluster was used for all other simulations. The lambda-rho algorithm was used for all simulations. The first-order rate constants ( $k_{1a}$ ,  $k_{1b}$ ,  $k_{2a}$ ,  $k_{2b}$ ,  $k_{5b}$ ,  $k_{7b}$ ,  $k_9$ ,  $k_{11}$ ) had units of  $s^{-1}$ . The second order rate constants ( $k_3$ ,  $k_{5a}$ ,  $k_{7a}$ ) are unitless and actually correspond to the probability of reaction upon collision, which is why they are constrained between 0 and 1. The stoichiometry of the phosphorylation sites was considered in binding interactions such that binding was half as likely if one site was phosphorylated, while the other was not. For example, a kinase would associate with a receptor with probability  $k_{5a}$  if both sites were unphosphorylated, but with probability  $0.5 \cdot k_{5a}$  if only one site was unphosphorylated. The simulations were conducted on a square  $10 \mu m \times 10 \mu m$  domain with reflective boundaries. Receptor clusters were constructed by placing receptor molecules within  $0.1 \mu m$  radius circles that act as semi-permeable boundaries: receptors cannot pass through, but kinases and phosphatases can. The simulations contain the species drawn in Figure 5.1A. The simulation domain is 2D, with membrane and cytosolic species distinguished only by their diffusion coefficients and reactivities ( $D_m = 0.1$ ,  $D_c = 1 \mu m^2/s$ ). Molecular radii ( $\rho$ ) were set to 10 nm, and the simulation timestep was set to 0.1 ms. Simulations contained 5000 total kinases, 5000 total phosphatases, and 180 total receptors. All enzymes were initialized in the cytosolic state, placed randomly throughout the domain. Receptors were initialized in the unphosphorylated state, either randomly distributed along the membrane to freely diffuse with no clusters (1 receptor per cluster) or distributed equally into the semi-permeable clusters. The number of clusters was either 90, 60, 45, 36, 30, 18, 12, or 9, corresponding to 2, 3, 4, 5, 6, 10, 15, or 20 receptors per cluster.

### 5.6.2. Parameter optimization via evolutionary algorithm

DEAP (v1.2.2) in python (v2.7.12) was used to set up the evolutionary algorithm (Fortin et al., 2012). Because each simulation took minutes to run, we needed to parallelize each evolution over many CPUs. We used Open MPI (v.3.0.0/intel 17.2) and mpi4py (v3.0.0), which also required cython (v0.29), to parallelize each evolution over 200 CPUs on the Dogwood cluster. Runs were conducted in three-individual tournament mode with 200 individuals over 50 generations per evolution, with a crossover probability of 0.5, a mutation probability of 0.1, a mutation polynomial with parameters  $\eta = 0.1$ , low = 0, high = 1, indpb = 0.2.

## REFERENCES

- Alarcón, B., Swamy, M., van Santen, H.M., and Schamel, W.W.A. (2006). T-cell antigen-receptor stoichiometry: pre-clustering for sensitivity. *EMBO Rep.* 7, 490–495.
- Andrews, S.S., and Bray, D. (2004). Stochastic simulation of chemical reactions with spatial resolution and single molecule detail. *Phys. Biol.* 1, 137–151.
- Andrews, S.S., Addy, N.J., Brent, R., and Arkin, A.P. (2010). Detailed simulations of cell biology with Smoldyn 2.1. *PLoS Comput. Biol.* 6, e1000705.
- Barton, W.A., Tzvetkova-Robev, D., Miranda, E.P., Kolev, M.V., Rajashankar, K.R., Himanen, J.P., and Nikolov, D.B. (2006). Crystal structures of the Tie2 receptor ectodomain and the angiopoietin-2-Tie2 complex. *Nat. Struct. Mol. Biol.* 13, 524–532.
- Bray, D., Levin, M.D., and Morton-Firth, C.J. (1998). Receptor clustering as a cellular mechanism to control sensitivity. *Nature* 393, 85–88.
- Castro, M., van Santen, H.M., Férez, M., Alarcón, B., Lythe, G., and Molina-París, C. (2014). Receptor Pre-Clustering and T cell Responses: Insights into Molecular Mechanisms. *Front. Immunol.* 5, 132.
- Changade, R., and Sheetz, M. (2017). Integrin and cadherin clusters: A robust way to organize adhesions for cell mechanics. *Bioessays* 39, 1–12.
- Daëron, M. (1997). Fc receptor biology. *Annu. Rev. Immunol.* 15, 203–234.
- Duke, T.A., and Bray, D. (1999). Heightened sensitivity of a lattice of membrane receptors. *Proc. Natl. Acad. Sci. USA* 96, 10104–10108.
- Dushek, O., van der Merwe, P.A., and Shahrezaei, V. (2011). Ultrasensitivity in multisite phosphorylation of membrane-anchored proteins. *Biophys. J.* 100, 1189–1197.
- Fiala, G.J., Kaschek, D., Blumenthal, B., Reth, M., Timmer, J., and Schamel, W.W.A. (2013). Pre-clustering of the B cell antigen receptor demonstrated by mathematically extended electron microscopy. *Front. Immunol.* 4, 427.
- Flannagan, R.S., Jaumouillé, V., and Grinstein, S. (2012). The cell biology of phagocytosis. *Annu. Rev. Pathol.* 7, 61–98.
- Fortin, F.-A., Rainville, F.-M.D., Gardner, M.-A., Parizeau, M., and Gagné, C. (2012). DEAP: Evolutionary Algorithms Made Easy. *Journal of Machine Learning Research.*
- Freeman, S.A., Goyette, J., Furuya, W., Woods, E.C., Bertozzi, C.R., Bergmeier, W., Hinz, B., van der Merwe, P.A., Das, R., and Grinstein, S. (2016). Integrins Form an Expanding Diffusional Barrier that Coordinates Phagocytosis. *Cell* 164, 128–140.
- García-García, E., Brown, E.J., and Rosales, C. (2007). Transmembrane mutations to FcγRIIIA alter its association with lipid rafts: implications for receptor signaling. *J. Immunol.* 178, 3048–3058.

- Garrett, T.P.J., McKern, N.M., Lou, M., Elleman, T.C., Adams, T.E., Lovrecz, G.O., Zhu, H.-J., Walker, F., Frenkel, M.J., Hoyne, P.A., et al. (2002). Crystal structure of a truncated epidermal growth factor receptor extracellular domain bound to transforming growth factor alpha. *Cell* 110, 763–773.
- Getahun, A., and Cambier, J.C. (2015). Of ITIMs, ITAMs, and ITAMis: revisiting immunoglobulin Fc receptor signaling. *Immunol. Rev.* 268, 66–73.
- Goyette, J., Nieves, D.J., Ma, Y., and Gaus, K. (2019). How does T cell receptor clustering impact on signal transduction? *J. Cell Sci.* 132.
- Guy, C.S., and Vignali, D.A.A. (2009). Organization of proximal signal initiation at the TCR:CD3 complex. *Immunol. Rev.* 232, 7–21.
- Hartman, N.C., and Groves, J.T. (2011). Signaling clusters in the cell membrane. *Curr. Opin. Cell Biol.* 23, 370–376.
- Himanen, J.-P., and Nikolov, D.B. (2003). Eph signaling: a structural view. *Trends Neurosci.* 26, 46–51.
- Holowka, D., and Baird, B. (2015). Nanodomains in early and later phases of FcεRI signalling. *Essays Biochem* 57, 147–163.
- Jaumouillé, V., and Grinstein, S. (2011). Receptor mobility, the cytoskeleton, and particle binding during phagocytosis. *Curr. Opin. Cell Biol.* 23, 22–29.
- Jaumouillé, V., Farkash, Y., Jaqaman, K., Das, R., Lowell, C.A., and Grinstein, S. (2014). Actin cytoskeleton reorganization by Syk regulates Fcγ receptor responsiveness by increasing its lateral mobility and clustering. *Dev. Cell* 29, 534–546.
- Johnson, S.A., Pleiman, C.M., Pao, L., Schnieringer, J., Hippen, K., and Cambier, J.C. (1995). Phosphorylated immunoreceptor signaling motifs (ITAMs) exhibit unique abilities to bind and activate Lyn and Syk tyrosine kinases. *J. Immunol.* 155, 4596–4603.
- Van Kampen, N.G. (2007). *Stochastic processes in physics and chemistry* (Elsevier).
- Kenworthy, A.K. (2008). Have we become overly reliant on lipid rafts? Talking Point on the involvement of lipid rafts in T-cell activation. *EMBO Rep.* 9, 531–535.
- Kong, L., Deng, Z., Shen, H., and Zhang, Y. (2011). Src family kinase inhibitor PP2 efficiently inhibits cervical cancer cell proliferation through down-regulating phospho-Src-Y416 and phospho-EGFR-Y1173. *Mol. Cell. Biochem.* 348, 11–19.
- Kwiatkowska, K., and Sobota, A. (1999). Tyrosine phosphorylation/dephosphorylation controls capping of Fcγ receptor II in U937 cells. *Cell Motil. Cytoskeleton* 42, 298–314.
- Kwiatkowska, K., and Sobota, A. (2001). The clustered Fcγ receptor II is recruited to Lyn-containing membrane domains and undergoes phosphorylation in a cholesterol-dependent manner. *Eur. J. Immunol.* 31, 989–998.
- Kwiatkowska, K., Frey, J., and Sobota, A. (2003). Phosphorylation of FcγRIIA is required for the receptor-induced actin rearrangement and capping: the role of membrane rafts. *J. Cell Sci.* 116, 537–550.

- Lemmon, M.A., and Schlessinger, J. (2010). Cell signaling by receptor tyrosine kinases. *Cell* 141, 1117–1134.
- Li, J., Yin, W., Jing, Y., Kang, D., Yang, L., Cheng, J., Yu, Z., Peng, Z., Li, X., Wen, Y., et al. (2018). The coordination between B cell receptor signaling and the actin cytoskeleton during B cell activation. *Front. Immunol.* 9, 3096.
- Lopes, F.B., Bálint, Š., Valvo, S., Felce, J.H., Hessel, E.M., Dustin, M.L., and Davis, D.M. (2017). Membrane nanoclusters of FcγRI segregate from inhibitory SIRPα upon activation of human macrophages. *J. Cell Biol.* 216, 1123–1141.
- Love, P.E., and Hayes, S.M. (2010). ITAM-mediated signaling by the T-cell antigen receptor. *Cold Spring Harb. Perspect. Biol.* 2, a002485.
- Milano, V., Piao, Y., LaFortune, T., and de Groot, J. (2009). Dasatinib-induced autophagy is enhanced in combination with temozolomide in glioma. *Mol. Cancer Ther.* 8, 394–406.
- Mócsai, A., Ruland, J., and Tybulewicz, V.L.J. (2010). The SYK tyrosine kinase: a crucial ' ' player in diverse biological functions. *Nat. Rev. Immunol.* 10, 387–402.
- Nadler, M.J., Harrison, M.L., Ashendel, C.L., Cassady, J.M., and Geahlen, R.L. (1993). Treatment of T cells with 2-hydroxymyristic acid inhibits the myristoylation and alters the stability of p56lck. *Biochemistry* 32, 9250–9255.
- Nimmerjahn, F., and Ravetch, J.V. (2008). Fcγ receptors as regulators of immune responses. *Nat. Rev. Immunol.* 8, 34–47.
- Ogiso, H., Ishitani, R., Nureki, O., Fukai, S., Yamanaka, M., Kim, J.-H., Saito, K., Sakamoto, A., Inoue, M., Shirouzu, M., et al. (2002). Crystal structure of the complex of human epidermal growth factor and receptor extracellular domains. *Cell* 110, 775–787.
- Oldenborg, P.A., Gresham, H.D., and Lindberg, F.P. (2001). CD47-signal regulatory protein alpha (SIRPα) regulates Fcγ and complement receptor-mediated phagocytosis. *J. Exp. Med.* 193, 855–862.
- Oliver, J.M., Burg, D.L., Wilson, B.S., McLaughlin, J.L., and Geahlen, R.L. (1994). Inhibition of mast cell Fc εR1-mediated signaling and effector function by the Syk-selective inhibitor, piceatannol. *J. Biol. Chem.* 269, 29697–29703.
- Pablo, M., Ramirez, S.A., and Elston, T.C. (2018). Particle-based simulations of polarity establishment reveal stochastic promotion of Turing pattern formation. *PLoS Comput. Biol.* 14, e1006016.
- Pao, L.I., Bedzyk, W.D., Persin, C., and Cambier, J.C. (1997). Molecular targets of CD45 in B cell antigen receptor signal transduction. *J. Immunol.* 158, 1116–1124.
- Pichot, C.S., Hartig, S.M., Xia, L., Arvanitis, C., Monisvais, D., Lee, F.Y., Frost, J.A., and Corey, S.J. (2009). Dasatinib synergizes with doxorubicin to block growth, migration, and invasion of breast cancer cells. *Br. J. Cancer* 101, 38–47.



- Rollet-Labelle, E., Marois, S., Barbeau, K., Malawista, S.E., and Naccache, P.H. (2004). Recruitment of the cross-linked opsonic receptor CD32A (FcγRIIA) to high-density detergent-resistant membrane domains in human neutrophils. *Biochem. J.* 381, 919–928.
- Roskoski, R. (2015). Src protein-tyrosine kinase structure, mechanism, and small molecule inhibitors. *Pharmacol. Res.* 94, 9–25.
- Schlessinger, J., Plotnikov, A.N., Ibrahimi, O.A., Eliseenkova, A.V., Yeh, B.K., Yayon, A., Linhardt, R.J., and Mohammadi, M. (2000). Crystal structure of a ternary FGF-FGFR-heparin complex reveals a dual role for heparin in FGFR binding and dimerization. *Mol. Cell* 6, 743–750.
- Singh, A., and Grima, R. (2017). The Linear-Noise Approximation and moment-closure approximations for stochastic chemical kinetics. *arXiv*.
- Sneddon, M.W., Faeder, J.R., and Emonet, T. (2011). Efficient modeling, simulation and coarse-graining of biological complexity with NFsim. *Nat. Methods* 8, 177–183.
- Sobota, A., Strzelecka-Kiliszek, A., Gładkowska, E., Yoshida, K., Mrozińska, K., and Kwiatkowska, K. (2005). Binding of IgG-opsonized particles to Fc γ R is an active stage of phagocytosis that involves receptor clustering and phosphorylation. *J. Immunol.* 175, 4450–4457.
- Turner, M., Schweighoffer, E., Colucci, F., Di Santo, J.P., and Tybulewicz, V.L. (2000). Tyrosine kinase SYK: essential functions for immunoreceptor signalling. *Immunol. Today* 21, 148–154.
- Ullrich, A., and Schlessinger, J. (1990). Signal transduction by receptors with tyrosine kinase activity. *Cell* 61, 203–212.
- Varma, R., Campi, G., Yokosuka, T., Saito, T., and Dustin, M.L. (2006). T cell receptor-proximal signals are sustained in peripheral microclusters and terminated in the central supramolecular activation cluster. *Immunity* 25, 117–127.
- Wehrman, T., He, X., Raab, B., Dukipatti, A., Blau, H., and Garcia, K.C. (2007). Structural and mechanistic insights into nerve growth factor interactions with the TrkA and p75 receptors. *Neuron* 53, 25–38.
- Woof, J.M., and Burton, D.R. (2004). Human antibody-Fc receptor interactions illuminated by crystal structures. *Nat. Rev. Immunol.* 4, 89–99.
- Zhu, X., Kim, J.L., Newcomb, J.R., Rose, P.E., Stover, D.R., Toledo, L.M., Zhao, H., and Morgenstern, K.A. (1999). Structural analysis of the lymphocyte-specific kinases Lck in complex with non-selective and Src family selective kinase inhibitors. *Structure* 7, 651–661.