

PARTICLE-BASED STOCHASTIC REACTION-DIFFUSION MODELS TO INVESTIGATE
SPATIOTEMPORAL DYNAMICS IN CELL BIOLOGY

Vinal Vinodrai Lakhani

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 the Curriculum
of Bioinformatics and Computational Biology

Chapel Hill
2015

Approved by:

David Adalsteinsson

Henrik Dohlman

Timothy Elston

Shawn Gomez

Daniel Lew

© 2015
Vinal Vinodrai Lakhani
ALL RIGHTS RESERVED

ABSTRACT

Vinal Vinodrai Lakhani: Particle-Based Stochastic Reaction-Diffusion Models to investigate
Spatiotemporal Dynamics in Cell Biology
(Under the direction of Timothy Elston)

Many facets of mathematics, science and engineering rely on numerical methods to study complex systems, for which analytical methods fail. For example, in particle-based models the positions of individual particles under the influence of various forces are monitored over time. These models are used to study phenomena ranging from the structure and dynamics of galaxies (where particles represent stars) to proteins (where particles represent atoms). In this document, we present two applications of particle-based modeling to study microscopic dynamics in cell biology, which would otherwise be invisible by current experimental methods. In both models, the particles represent protein molecules, and we calculate the stochastic biochemical reactions and diffusion of tens of thousands of proteins over time. We utilize Generally Programmable Graphics Processing Units (GPGPUs) to achieve the high performance computing necessary to simulate these large scale models.

Chapter 1 begins with the observation that the mobility of Rac1 molecules, which are important regulators of the cell cytoskeleton, is spatially regulated in migrating cells. Specifically, Rac1 molecules near the leading edge have less mobility than those in the trailing edge. We create a particle-based stochastic reaction-diffusion model to test the hypothesis that patches of actin, called ‘actin islands’, are responsible for this observation. We find that these islands are capable of producing the spatially-dependent mobility measured by *in vivo* experiments.

Chapters 2 and 3 discuss a more complex model built to study cellular gradient sensing, which is the ability of singular cells to detect external chemical gradients. This fundamental biological process allows cells to move or grow towards favorable environments and away from toxic environments. We study this process in the context of yeast mating; wherein, a haploid yeast cell senses the pheromone emitted by a mating partner. These cells are capable of sensing shallow gradients, in which molecular-level noise from reactions and diffusion is relatively large. Chapter 2 describes the particle-based stochastic reaction-diffusion model we built to quantify noise-reduction mechanisms proposed elsewhere. Chapter 3 shows neither time-averaging nor receptor endocytosis sufficiently reduces noise; however, the pheromone protease Bar1 may improve gradient sensing in certain cases.

To my family, friends, and teachers,
for nurturing, encouraging and enabling me.

ACKNOWLEDGEMENTS

This document is the culmination of the hard work of many people. First and foremost is my fantastic mentor: Tim. On an educational level, Tim opened my eyes to world of Systems Biology, which looks at the world as a network of connections and interactions. And although this field can be incredibly daunting, Tim showed me the power of simplification. He taught me to consider only the minimal set of components capable of answering the question. He taught me that good science starts this way and adds only one variable at a time. But one of the biggest impacts Tim's tutelage had on me was to motivate and excite me. There were times during my project when I felt discouraged about my results or my lack of progress; but each time I met with Tim, he would find exciting questions and opportunities with the results I DID manage to get. His excitement always spurred me forward, feeling undaunted and eager to answer new questions. Lastly, I cannot stress enough the patience Tim had with me. He was patient with me as I debugged my code and as I struggled to understand the confusing results. Most importantly he was patient with me when my progress slowed due to difficulties in my family life.

Not surprisingly, the students, post-docs and collaborators that Tim keeps in his academic company have also been supportive. Professors Beverly Errede, Henrik Dohlman and Danny Lew gave me great feedback on applications of my model. Professors David Adalsteinsson and Shawn Gomez equipped me with tools for better analyzing and understanding my results. Also, I thank the entire Elston Lab (in no particular order): Guilherme Garcia, Richard Allen, Denis Tsygankov, Meng Jin, John Houser, Sai Venkatapurapu, Josh Kelley, Yan Fu, Matthew Peña,

Patrick McCarter, Lior Vered, Maria Minakova, Conner Sandefur, Ellen O'Shaughnessy, and Alex Chen, for their emotional and intellectual encouragement. I can recall conversations with each member wherein they taught me about yeast biology, gave me great ideas about my model, or brainstormed new ideas.

Tim and I also had the pleasure of collaborating with Drs. Enrico Gratton and Elizabeth Hinde from the Laboratory of Fluorescence Dynamics at UC Irvine. Liz performed the experiments and generated the pCF carpets for her experimental data and my simulations. Our work is presented in Chapter 1.

I must also thank my friends: Michael Zytchow, Lawrence Ngo, Cameron Smith, Devin Hubbard, and Zack Cashion, for dreaming big and encouraging me to do the same. I am indebted to the Hubbard family for their unmatched kindness and support. And finally, I thank the Lakhani family, who raised me and instilled in me a great respect for education. I am grateful for every conversation with my cousins, who, even today, remain invaluable sources of knowledge and intellectual discourse. Finally, although we've only been married for three months, Krupal has patiently supported, encouraged, and pushed me through this ultimate phase of my education. As Tim put it, when I told him that I was getting married, he replied "Now you have motivation to finish."

Without each of these people, I would not be the person I am today.

TABLE OF CONTENTS

LIST OF TABLES	xi
LIST OF FIGURES	xii
CHAPTER 1: SPATIO-TEMPORAL REGULATION OF RAC1 MOBILITY BY ACTIN ISLANDS	1
Introduction	1
1.1 Overview	1
1.2 Intracellular Traffic Observed in <i>in silico</i> simulations by Pair Correlation Analysis	5
1.3 Gradient of Molecular Flow Observed in <i>in vivo</i> experiments by Pair Correlation Analysis	7
1.4 Gradient of Rac1 Molecular Flow Produced by Actin Island Simulations	10
1.5 Discussion	14
1.6 Methods	17
1.6.1 Simulation Algorithm	17
1.6.2 Pair Correlation Carpet	22
1.7 Supplemental Methods	24
1.7.1 Analyzing pCF Carpets	24
1.7.2 Simulation Program Details	25
1.7.3 Calculating Diffusion Steps	25
1.7.4 Calculating Reaction Probabilities	29
1.7.5 Full Set of Simulation Parameters	33

1.7.6 Generating an Intensity Carpet	35
1.7.7 Hardware Details	35
References	37
CHAPTER 2: A MODEL OF GRADIENT SENSING IN THE CONTEXT OF YEAST MATING	38
Introduction	38
2.1 Overview	39
2.2 Binding Reactions	41
2.3 Unbinding Reactions	43
2.4 Diffusion of Pheromone	43
2.4.1 Pheromone Gradient – Method 1	44
2.4.2 Pheromone Gradient – Method 2	46
2.5 Diffusion of Receptors	47
2.6 Algorithm Overview	48
2.7 Receptor Cycling Model	49
2.8 Bar1 Model	50
2.9 Calculating Binding Probability	51
2.10 Calculating Pheromone Reflection off the Cell Surface	56
2.11 Injection Boundary Condition: Derivation of Eqs 2.5 – 2.8	58
2.11.1 Gradient Method 1	59
2.11.2 Injection Distance	61
2.11.3 Gradient Method 2	64
2.12 Receptor Diffusion on the Cell Surface: Derivation of Eq 2.9	66
2.13 Measuring the Pheromone Distribution	66

References	68
CHAPTER 3: USING STOCHASTIC SIMULATIONS TO ASSESS NOISE-REDUCTION MECHANISMS DURING GRADIENT SENSING IN YEAST	70
Introduction	70
3.1 Background	71
3.2 Overview	75
3.3 Equilibrium Fluctuations in Uniform Pheromone	76
3.3.1 Equilibrium Fluctuations in Receptor Occupancy	76
3.3.2 Time Averaging	78
3.3.3 The “Perfectly Absorbing” Cell and Receptor Cycling	82
3.4 Cells at Equilibrium in a Pheromone Gradient	85
3.4.1 Gradient Sharpening due to Steric Effects	85
3.4.2 Slow Reaction Rates and Receptor Diffusion add Spatial Noise	87
3.4.3 Time-averaging and the Angle of Estimation	89
3.5 Sensing During Gradient Formation	92
3.5.1 Formation of the Gradient	92
3.5.2 Transient Differences in Receptor Occupancy	93
3.5.3 Bar1 Improves Gradient Sensing for Fast Reaction Rates	96
3.6 Discussion	97
3.6.1 Mechanisms for Noise-Reduction	97
3.6.2 Gradient Sharpening due to Steric Effects	98
References	100

LIST OF TABLES

Table 1.1 – Reaction Rates and Probabilities	34
Table 2.1 – A Typical Parameter Set	41

LIST OF FIGURES

Figure 1.1 – Pair correlation analysis of a fluorescent protein’s diffusive route reveals how the cell’s architecture directs intracellular traffic	4
Figure 1.2 – Pair correlation analysis of a Rac1 FRET biosensor reveals Rac1 activity to be spatiotemporally regulated by a dynamic gradient of protein mobility	9
Figure 1.3 – Simulations with islands of varying binding affinity	11
Figure 1.4 – Investigating the cellular substructure during three stages of EGF stimulation by comparing pCF carpets	13
Figure 1.5 – Simulation Geometry	19
Figure 1.6 – Gaussian Analysis of Pair Correlation Carpets to Extract Average Delay Time(s) Rac1 Takes to Diffuse Along Simulated and <i>in vivo</i> Line Scans	24
Figure 1.7 – Important Vectors for Reflecting Inside Actin Islands	27
Figure 2.1 – Simulations on our Particle-Based Stochastic Reaction-Diffusion Model	40
Figure 2.2 – Rationale for Doubling P_{bind}	56
Figure 2.3 – Reflecting Pheromone off Cell Surface	57
Figure 2.4 – Molecular Reservoir	60
Figure 2.5 – Injection Distance	61
Figure 2.6 – Probability Density of Injection Distance	63
Figure 2.7 – Cylindrical Bins	67
Figure 3.1 – Gradient Sensing during Yeast Mating	72
Figure 3.2 – Receptor Occupancy at Equilibrium	77
Figure 3.3 – Comparing Mathematical Models	80
Figure 3.4 – Uncertainty as a function of Time-Averaging	82
Figure 3.5 – Absorbing Ligand vs Releasing Ligand	84

Figure 3.6 – Gradient Sharpening due to Steric Effects	85
Figure 3.7 – Receptor Occupancy in Linear Gradient vs Sharpened Gradient	87
Figure 3.8 – Receptor Diffusion adds Spatial Noise	88
Figure 3.9 – Estimated Angle away from Gradient	90
Figure 3.10 – Estimated Angle Improves with Time-averaging	91
Figure 3.11 – Cell’s Physical Boundary Sharpens an Emerging Gradient	92
Figure 3.12 – Initial, Transient Difference in Receptor Occupancy	93
Figure 3.13 – Accuracy and Confidence during Gradient Formation	95

CHAPTER 1: SPATIO-TEMPORAL REGULATION OF RAC1 MOBILITY BY ACTIN ISLANDS¹

Introduction

Rho GTPases play important roles in many aspects of cell migration, including polarity establishment and organizing actin cytoskeleton. In particular, the Rho GTPase Rac1 has been associated with the generation of protrusions at leading edge of migrating cells. Previously it was shown that the mobility of Rac1 molecules is not uniform throughout a migrating cell (Hinde E et. al. PNAS 2013). Specifically, the closer a Rac1 molecule is to the leading edge, the slower the molecule diffuses. Because actin-bound Rac1 diffuses slower than unbound Rac1, we hypothesized that regions of high actin concentration, called “actin islands”, act as diffusive traps and are responsible for the non-uniform diffusion observed in vivo. Here, in silico model simulations demonstrate that equally spaced actin islands can regulate the time scale for Rac1 diffusion in a manner consistent with data from live-cell imaging experiments. Additionally, we find this mechanism is robust; different patterns of Rac1 mobility can be achieved by changing the actin islands’ positions or their affinity for Rac1.

1.1 Overview

Rho GTPases play a critical role in regulating many aspects of cell migration including polarity establishment and the actin cytoskeleton. Rac1 is a Rho GTPase associated with membrane protrusions at the leading edge of the cell [1]. Recent work demonstrated that Rac1 activity is closely regulated in space and time during the retraction portion of the protrusion-

¹ This chapter has been accepted for publication as an article in the PLoS ONE. It is titled “Spatio-temporal Regulation of Rac1 Mobility by Actin Islands” and authored by Lakhani V, Hinde E, Gratton E and Elston T.

retraction cycle. Specifically, Rac1 activity peaks 40 seconds after and 2 μm away from a protrusion event [2]. Previously, fluctuation analysis in polarized cells was used to establish that the time scale of Rac1 diffusion varied with its localization within the cell [3]. Using pair correlation function (pCF) analysis, the time taken for a Rac1 molecule to move 1 μm at each position along the axis of the cell was calculated [3]. In particular, a negative correlation between Rac1's mobility and its proximity to the leading cell edge was found; Rac1 molecules took 100 times longer at the front of the cell than at the back to move 1 μm [3]. These observations led to the hypothesis that diffusive barriers, such as those found in neurons for compartmentalizing proteins [4], are responsible for the observed spatial variation in diffusion. Here we use a computational model to demonstrate that diffusive barriers, in the form of "actin islands", can establish gradients of molecular mobility across the cell similar to those observed for Rac1.

We use a new technique called pair correlation function (pCF) analysis [5,6] to determine the spatial dependence of Rac1 mobility along the axis of a polarized cell. *In vivo* data were collected using a combination of Forster Resonance Energy Transfer (FRET) and Fluorescence Lifetime Imaging Microscopy (FLIM) [3,6]. We performed confocal line scans across the axis of a cell expressing a Rac1 dual chain FRET biosensor. The intensity and lifetime data of the donor and acceptor chain of this construct were collected by FLIM. This mode of acquisition provides us with two important data sets. First, we obtain a time series of the FRET biosensor lifetime in each pixel along the line scan, which describes the spatial distribution of Rac1 activity along the axis of the cell with millisecond resolution. Second, we obtain intensity fluctuations of Rac1 localization in each pixel along the line, which is used for pairwise correlation analysis of molecular flow along the axis of the cell. That is, we can calculate the time Rac1 molecules take to traverse a fixed distance along the line [3,6]. As noted above, using this multiplexed approach

it was found that Rac1 mobility decreases near the leading edge of the cell where we also observe, by FRET analysis, Rac1 activity to be the highest [3]. We hypothesized that cells achieve this spatiotemporal control of Rac1 mobility by using patches of dense actin, we call “actin islands”, to which Rac1 reversibly binds. By strategically placing and adjusting the density of the actin in these actin islands, the cell can reduce mobility of Rac1 in the desired location. For example, to slow diffusion towards the leading edge, the actin islands can be denser towards the leading edge.

To test this hypothesis, we created a computational model to study Rac1 mobility within a cell containing actin-islands. Using a particle-based stochastic simulation algorithm, we explicitly simulate the diffusion of individual Rac1 molecules and their binding/unbinding reactions with actin-islands. Unbound Rac1 freely diffuses throughout the cell. The actin-islands behave as diffusive traps, capable of slowing the diffusion rate and restricting the accessible space for an actin-bound Rac1 molecule. During the simulation we tally the number of Rac1 molecules located in bins along the center axis of the cell. Analogously, in the *in vivo* experiments, we measured the fluorescence intensity of pixels along the axis of the cell. In both cases, we tabulate the molecular counts (or fluorescence intensities) for each bin (or pixel) over time into an “intensity carpet” (Fig. 1.1D). We use the intensity carpet to calculate a pCF carpet (Fig. 1.1H) as described in [3]. Although we model Rac1 diffusion as an unbiased, uniform process throughout the cell, we find the presence of actin-islands can create a spatial bias to Rac1 mobility. Hence, comparing the pCF carpet data, we qualitatively reproduce the *in vivo* data with our *in silico* model. These results give credence to our hypothesis that a mechanism based on strategically arranging actin islands can lead to spatially dependent molecular mobility.

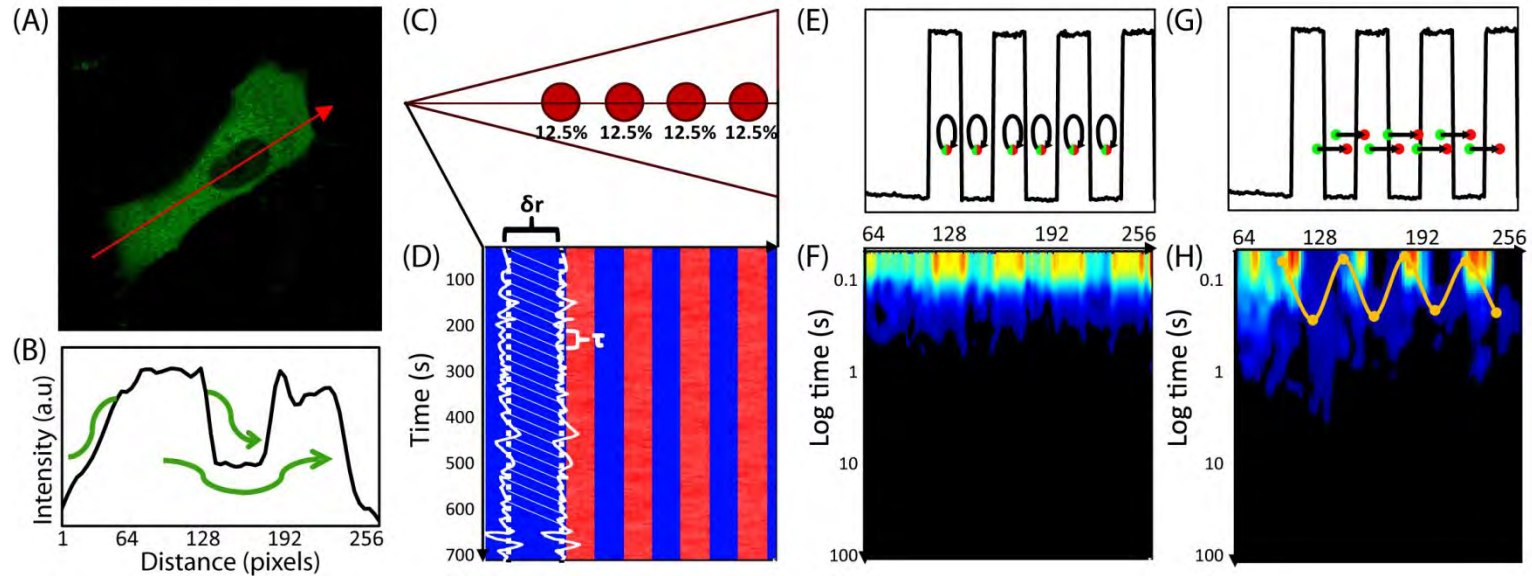


Fig. 1.1: Pair correlation analysis of a fluorescent protein's diffusive route reveals how the cell's architecture directs intracellular traffic.

(A) Intensity image of cell expressing 5-EGFP. Fluorescence data along the arrow is summarized in the next panel. (B) Intensity profile of 5-EGFP across axis of cell shows 5-EGFP exclusion from the nucleus and therefore an obstacle to 5-EGFP free diffusion. Green arrows demonstrate the molecules must diffuse around instead of through the obstacle. (C) Model for the simulation of Rac1 diffusion in a triangular shaped cell ($25.6\mu\text{m}$ by $5\mu\text{m}$) with four, circular traps ($2\mu\text{m}$ wide). Each trap captures, on average, 12.5% of the total Rac1 population. Line scans are taken along the center axis of the cell, as shown by the horizontal line. Each pixel is measured in succession, and one line scan is completed when all 256 pixels have been measured. (D) Many (47000) line scans are combined into an intensity carpet. For this simulation, the intensity carpet shows accumulation of Rac1 colocalized with the four traps. A cartoon of two intensity profiles (intensity vs time) in white is overlaid on the intensity carpet. (E) A representative line scan (intensity vs pixel). The arrows indicate which two pixels are being pair-correlated, from green to red, for the pCF analysis in the next panel. In this case, each pixel is pair-correlated with itself, which is equivalent to an autocorrelation calculation. (F) The pCF(0) carpet reveals that pixels within the traps have higher autocorrelation values for short delay times ($\tau \approx 0\text{s}$) than pixels outside the traps. These values indicate the traps have a higher concentration of Rac1 than elsewhere in the cell. (G) The same representative line scan (intensity vs pixel) as in (E). Here the arrows indicate each pixel (green dot) is pair-correlated with a pixel (red dot) $0.5\mu\text{m}$ to the right ($\delta r = 5$ pixels). (H) The pCF(5) carpet reveals diffusion in and around the traps is slower than elsewhere in the cell. We average the data from every 20 pixels (columns) and smooth this average profile using a Gaussian filter; lastly, we extract the peak time for every 20 columns. We plot a point at each of these peak times. Hence, the yellow highlighted data displays the average time Rac1 takes to diffuse $0.5\mu\text{m}$ to the right. It takes about 0.3s to diffuse $0.5\mu\text{m}$ inside the islands but less than 0.1s to diffuse the same distance elsewhere in the cell.

1.2 Intracellular Traffic Observed in *in silico* simulations by Pair Correlation Analysis

In live cells the default mechanism of motion for many biological molecules is diffusion. Although unregulated diffusion produces a spatially isotropic distribution of molecules, it has been shown that structural features of the cell create intracellular compartments that generate spatially heterogeneous molecule distributions. For example, insights into intra-cellular trafficking have been derived from measuring the effect of the cell nucleus on the diffusion of biologically active and inert molecules [7]. The effect of the nucleus on diffusion is readily apparent, if we scan across the axis of an NIH3T3 cell transiently transfected with the biologically inert fluorescent protein 5-EGFP (Fig. 1.1A). In this case, the fluorescence intensity profile clearly shows the exclusion of 5-EGFP from the nucleus (Fig. 1.1B). From this simple experiment we can deduce that the nuclear envelope behaves as an impenetrable barrier, around which 5-EGFP must diffuse. Given that intracellular trafficking of biologically active molecules is far more complex, the diffusive route traveled by a fluorescently labeled protein is not always evident from simple inspection of the fluorescence intensity distribution, and thus a more dynamic approach is required.

To gain insight into the diffusive motion active proteins, we employ an analysis method that is based on pairwise correlation functions. Using pairwise correlation analysis, it is possible to discern both diffusion rates and particle fluxes along a confocal line scan. These quantities are inferred by measuring temporal cross-correlations in fluorescence intensity between pairs of points a distance δr apart as a function of the time delay τ between measurements. To illustrate this idea consider the following *in silico* example. We simulate the diffusion of individual Rac1 molecules ($D = 10\mu\text{m}^2/\text{s}$) inside a cell containing four diffusive traps (Fig. 1.1C). The cell is $25.6\mu\text{m}$ long and $5\mu\text{m}$ wide; the circular traps are $2\mu\text{m}$ in diameter (see Methods). Rac1

molecules that diffuse into a trapping area can reversibly bind to the trap. When bound in a trap, Rac1 molecules are (1) spatially restricted to remain inside the trap, and (2) the diffusion constant reduces to $1\mu\text{m}^2/\text{s}$. During the simulation we repeatedly take “line scans”, which are measurements along a line that traverses the cell (Fig. 1.1C). Whereas the line scans for the *in vivo* experiments measured fluorescence intensity in each pixel $(0.1\mu\text{m})^2$ along the line, the line scans for our *in silico* simulations measure the number of molecules in square bins $(0.1\mu\text{m})^2$ along the line. In both cases, we summarize the resulting data as an intensity carpet (Fig. 1.1D); wherein, each row is a single line scan (Figs. 1.1E and 1.1G) and each column gives an intensity profile (intensity vs time) for a single pixel (Fig. 1.1D).

The impact on mobility can be measured by pairwise correlation analysis between an intensity profile and a neighboring profile, a distance δr to the right, as a function of the time delay τ . We choose δr such that it is large enough to measure mobility around each trap. For example, we can calculate the correlation between the intensity profile at pixel 64 and the intensity profile at pixel 69 ($\delta r = 5$ or $0.5\mu\text{m}$) τ seconds later. The characteristic delay time, defined as the τ that generates the highest correlation value, is a measure of the time scale for a Rac1 molecule at pixel 64 to diffuse to pixel 69. We repeat this process for all pixels to map the molecular flow pattern of Rac1 along the simulated cell’s axis. We first set $\delta r = 0$ and thus derive an autocorrelation profile (pCF(0)) for each pixel (Figs. 1.1E and 1.1F). For $\tau = 0$, the value of the autocorrelation is equal to the mean squared number of particles in the pixel. Hence, the high value areas in the pCF(0) carpet (Fig. 1.1F) indicate areas of high Rac1 concentration. Unsurprisingly, these areas are co-localized with the actin island traps.

We next introduce a spatial component to the cross correlation function by setting $\delta r = 5$ pixels (pCF(5)) and recalculating the pCF carpet (Figs. 1.1G and 1.1H). Taking $\delta r = 5$ pixels,

corresponds to a distance of $0.5\mu\text{m}$. This distance allows us to cross correlate intensity fluctuations located outside the trap with intensity fluctuations located inside the trap, thus measuring the time taken to enter or exit this environment. To help interpret the pCF carpet, we use the SimFCS software developed at the Laboratory for Fluorescence Dynamics (www.lfd.uci.edu); more details can be found in Section 1.7.1 (Fig. 1.6) as well as the literature [3,5,7]. Briefly, we combine and average every 20 pixels (columns) of the pair correlation values. Each average is smoothed with a Gaussian filter, and we highlight the peak times. Each peak time is the delay time with the maximum pair correlation value (Fig. 1.1H). The highlighted points, plotted every 20 pixels, are connected by interpolation and indicate the time scale for Rac1 to diffuse $0.5\mu\text{m}$ to the right. We find Rac1 takes longer to diffuse in and around the actin islands than elsewhere in the cell. It takes about 0.3s to diffuse $0.5\mu\text{m}$ to the right inside the islands, which is consistent with the diffusional time scale of the islands: $\frac{L^2}{D} = \frac{(0.5\mu\text{m})^2}{1\mu\text{m}^2/\text{s}} = 0.25\text{ s}$. Outside the islands, the same trajectory takes much less than 0.1s, which is as expected: $\frac{L^2}{D} = 0.025\text{ s}$. Hence, pCF analysis is capable of revealing that actin islands act as barriers to mobility.

1.3 Gradient of Molecular Flow Observed in *in vivo* experiments by Pair Correlation Analysis

When probing the spatiotemporal dynamics of signaling molecules like Rac1, it is necessary to measure changes in both position and activity. These measurements are most often achieved by use of a FRET biosensor; wherein, changes in donor emission provide a readout of protein activity. Thus to determine the diffusive route Rac1 adopts upon activation, we recently combined biosensor FRET detection with pair correlation analysis [3]. Here, we repeat the experiment from [3]. We concomitantly measured the fluorescence intensity (Fig. 1.2A) and fluorescence lifetime (Fig. 1.2B) of a Rac1 biosensor in the donor channel along a line that

extended from the rear to the front of a migrating cell. Lifetime analysis of the Rac1 biosensor FRET signal along the line scan revealed that the front of the cell (red time trace) is activated before the back of the cell (green time trace) (Fig. 1.2C). In addition to this spatially dependent Rac1 activation, we found Rac1 mobility was also spatially dependent. We acquired three intensity carpets during the course of the experiment: one before EGF stimulation (Fig. 1.2D), one between 0 – 180s after stimulation and one between 180 – 360s after stimulation. From each intensity carpet data, we performed pair correlation analysis of intensity fluctuations separated by a distance of 800nm ($\delta r = 8$ pixels) along the line scan (Figs. 1.2E – G). As in Fig. 1.1H, we combined the data of every 20 columns into an average pair correlation vs delay time plot; we smooth this plot using a Gaussian filter. The highlighted red data points are the maxima of these smoothed curves (Fig. 1.2E). For the data following EGF stimulation (Figs. 1.2F and 1.2G), some of the smoothed curves have two maxima (see Section 1.7.1). The highlighted red and yellow data points are these maxima (Figs. 1.2F and 1.2G). Before EGF stimulation (Fig. 1.2E), Rac1 diffuses 800nm towards the front of the cell with a characteristic time of 0.03s. Note that this pCF carpet is qualitatively similar to the pCF carpet calculated from a simulation with uniform actin islands (Fig. 1.1H). Three minutes after stimulation (Fig. 1.2F), the time taken to travel 800nm is no longer constant across the axis of the cell (red curve Fig. 1.2F); instead, there is a gradient of molecular flow with slower speeds near the front of the cell. After six minutes (Fig. 1.2G), this gradient of flow is steeper (red curve Fig. 1.2G); the characteristic times range from 0.1 to 1s. Additionally, the second set of highlighted data (yellow curve Fig. 1.2G) is significantly different than the first, which indicates the presence of a second population of Rac1 whose mobility is spatially regulated separately from the first population.

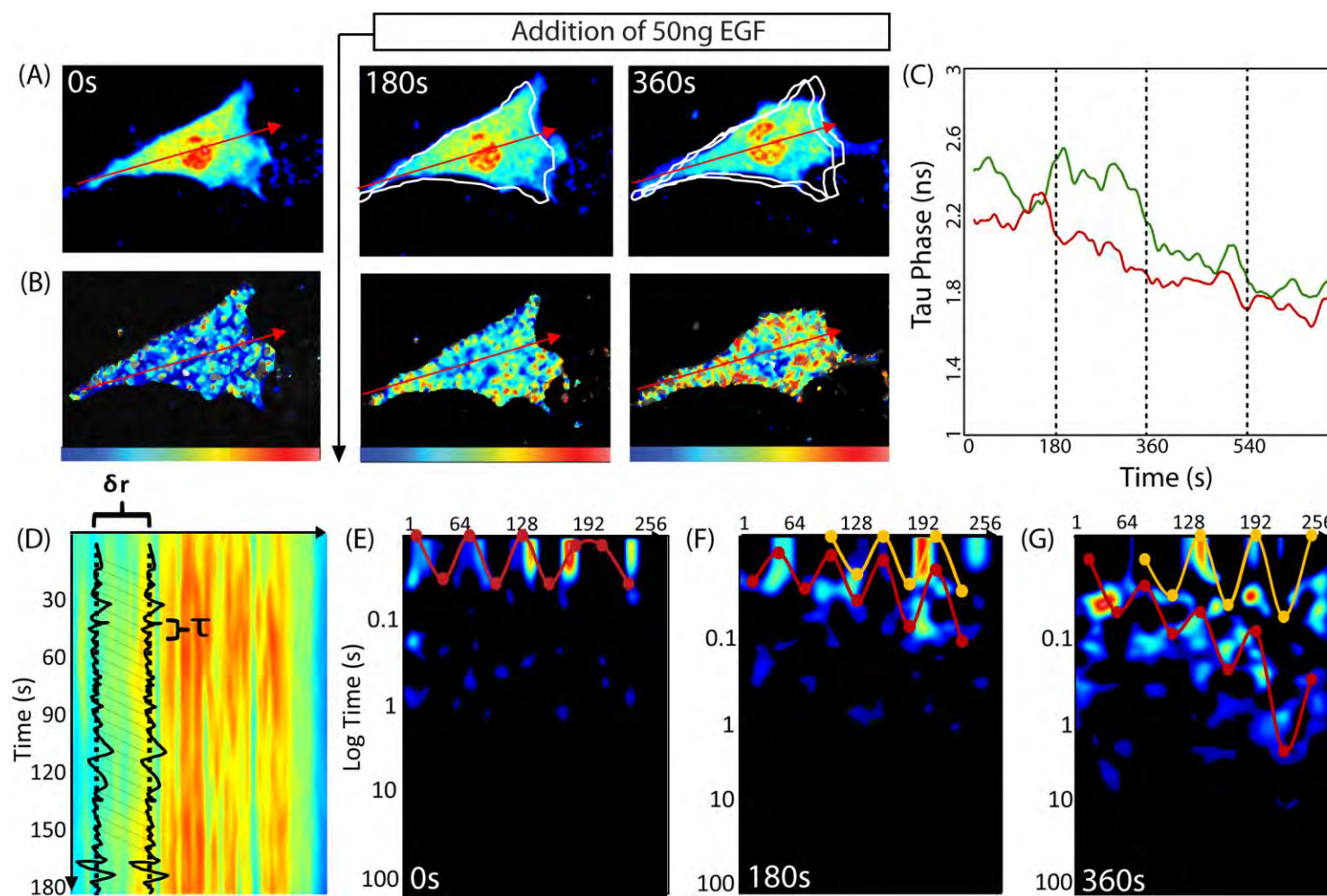


Fig. 1.2: Pair correlation analysis of a Rac1 FRET biosensor reveals Rac1 activity to be spatiotemporally regulated by a dynamic gradient of protein mobility.

Fig. 1.2: Pair correlation analysis of a Rac1 FRET biosensor reveals Rac1 activity to be spatiotemporally regulated by a dynamic gradient of protein mobility.

(A) Intensity image of a NIH3T3 cell expressing the Rac1 dual chain FRET biosensor in the donor channel before and after stimulation with epidermal growth factor (EGF). The white traces outline the cell's position(s) from the previous panel(s). (B) Same cell as in (A) pseudo-colored according to donor lifetime. The blue to red color range corresponds to a change in lifetime from 2 to 3ns and therefore low to high Rac1 activity. The tau phase (in ns) is derived from phasor analysis of the fluorescence decay, as in [3]. Shorter tau-phase times correspond to higher FRET activity. (C) Average lifetime analysis of the first 10 pixels (back of the cell, green time series) and the last 10 pixels (front of the cell, red time series). This comparison reveals that after EGF stimulation Rac1 is activated earlier at the front than the back of the cell. (D) The intensity carpet that is derived from line scans acquired across the axis of the cell in (A). (E) Pair correlation analysis of the intensity carpet acquired before EGF stimulation. The highlighted data shows Rac1 molecular flow is uniform: it takes about 0.03s to traverse 0.8 μ m (pCF(8)). (F) Pair correlation analysis of the intensity carpet acquired 180s after EGF stimulation. The highlighted shows Rac1 molecular flow is non-uniform. At the back of the cell, traversing 0.8 μ m takes about 0.03s; this time becomes gradually longer towards the front of the cell. The second set of highlighted data (yellow curve) is not significantly different from the first (red curve). (G) Pair correlation analysis of data acquired 360s after EGF stimulation. The mobility gradient is steeper (red curve); the delay time ranges from 0.05s at the back to 1.2s at the front. A second gradient emerges (yellow curve); the delay times range from 0.03s to 0.08s.

The simulation in Fig. 1.1 (Fig. 1.1H), wherein each barrier has equivalent affinity for Rac1, qualitatively matches the mobility of Rac1 in an unstimulated cell (Fig. 1.2E); that is, Rac1 mobility is uniform across the cell. However, in a stimulated cell, Rac1 shows variable mobility across the cell (Figs. 1.2F and 1.2G), possibly due to variable density of actin and hence variable Rac1 binding affinity. It may be that Rac1 interacts with different substrates with varying affinities in the membrane, and therefore, Rac1 mobility depends on cell polarization in response to external cues. We next perform simulations guided by this hypothesis.

1.4 Gradient of Rac1 Molecular Flow Produced by Actin Island Simulations

To test our hypothesis that actin-islands can create the spatially dependent mobility observed from *in vivo* experiments, we returned to *in silico* simulations. Instead of assuming all the actin islands bind Rac1 with equal probability, we first tested a scenario in which the affinity of the rear island was taken to be three-times higher than the others (Fig. 1.3A). The resulting intensity carpet showed four regions with higher intensity than the background, corresponding to the islands (Fig. 1.3B). The region with highest intensity corresponds to the island with highest

affinity for Rac1. Calculating the pCF carpet (Fig. 1.3C) reveals that Rac1 molecular flow is slowest in the island with highest affinity, slightly faster in the other three islands and fastest outside the islands. Hence the islands' affinity inversely correlates with local Rac1 mobility.

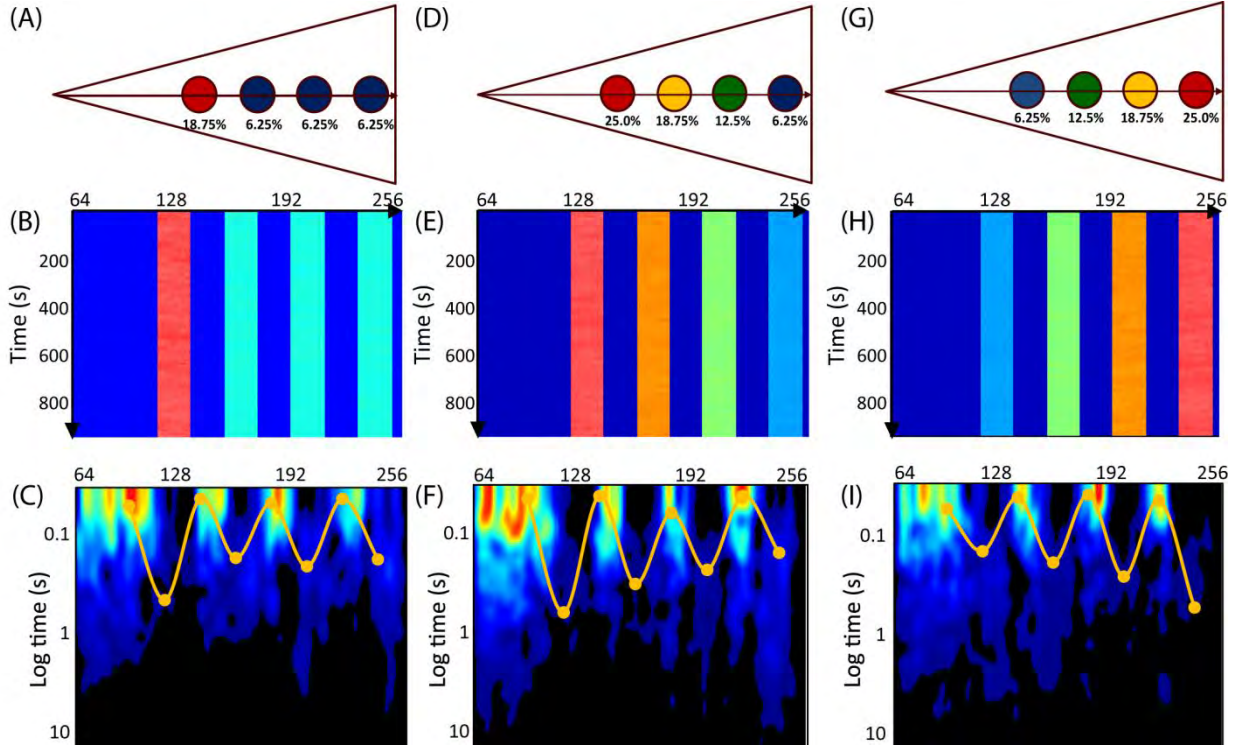


Fig. 1.3: Simulations with islands of varying binding affinity.

(A) A simulation set-up showing the rear (leftmost) island binds on average 18.75% of all Rac1, and each of the other three bind 6.25% on average. Unbound Rac1 diffuses with $D = 10\mu\text{m}^2/\text{s}$. If bound, Rac1 diffuses with $D = 1\mu\text{m}^2/\text{s}$. (B) The resulting intensity carpet shows the highest accumulation at the rear island. (C) The pCF carpet (yellow curve) reveals four arc features, which indicate regions of slow molecular flow, across each island. The time needed to flow $0.5\mu\text{m}$ to the right (pCF(5)) is longer near the island with the highest affinity (0.6s) than the other islands (0.2s). (D) A simulation wherein the islands form a gradient of binding affinities. The affinities range from 37.5-6.25% from the rear to the front of the cell. (E) The resulting intensity carpet shows a gradient of accumulation of Rac1. (F) The pCF carpet reveals a gradient of arc features whose position corresponds to the position of the islands and whose length correlates with the affinity of the islands. The time scale for molecular flow near the islands ranges from 1s, 0.8s, 0.5s and 0.2s (back to front). (G) Simulation of a cell with actin islands forming a gradient of binding affinities. The affinities range from 6.25-37.5% from the back to the front of the cell. (H) The resulting intensity carpet shows a gradient of accumulation of Rac1. (I) The pCF carpet reveals a gradient of arc features whose position corresponds to the position of the islands and whose length correlates with the affinity of the islands. The pair correlation pattern is opposite of the previous simulation (F). The time scale for molecular flow near the islands ranges from 0.2s, 0.5s, 0.8s and 1s (back to front). This gradient is analogous to the gradient calculated for 3 min after EGF stimulation (Fig. 1.2F).

Next, we extend our model by varying the affinity of each actin island. Specifically, islands closer to the leading edge have lower affinity than islands near the trailing edge (Fig. 1.3D). The intensity carpet shows four regions with different intensities (Fig. 1.3E). As before, the average intensity in each of the four regions is directly proportional to the binding affinity. The pCF carpet (Fig. 1.3F) shows a gradient of molecular flow; wherein, Rac1 mobility is faster towards the leading edge. We can reverse this gradient by reversing the island affinities (Fig. 1.3G). This model produces a pCF carpet (Fig. 1.3I) qualitatively similar to the one observed from *in vivo* data (Fig. 1.2F). That is, there is a diffusive gradient such that Rac1 mobility is slower towards the leading edge. Consistent throughout all our simulations, we find the characteristic time to diffuse around or through an actin island is proportional to the island's binding affinity. We find that this mechanism to spatially regulate molecular flow is quite robust; the affinity of the actin island dictates the mobility at that location.

Finally, we aimed to reproduce the molecular flow observed 6 minutes after EGF stimulation (Fig. 1.2G), in which there are two populations of Rac1 whose mobility is regulated separately. It is known that active Rac1 moves slower than inactive Rac1 [8]. In our model, diffusion is regulated by Rac1's affinity for actin (Rac1's primary effector). We postulated that inactive Rac1 also binds actin, but with a significantly reduced affinity. Therefore, keeping all other parameters equal, we consider two populations of Rac1 molecules (inactive and active) that possess different affinities for the actin islands (Fig. 1.4F). In particular, we assumed the active population binds actin with twice the affinity as the inactive population. For computational simplicity, we simulated each population separately (Figs. 1.3G and 1.4D) and combined the resulting intensity carpets. This approach does not affect the results of our simulations, because the two subpopulations are assumed to diffuse and react independently of one another. After

combining the intensity carpets, we perform pCF analysis. The resulting pCF carpet shows two gradients of molecular flow: one for each Rac1 population (Fig. 1.4G). Based on our previous results, we know the red curve corresponds to the high affinity active Rac1 population, and the yellow curve corresponds to the low affinity active Rac1 population. Our results indicate actin islands, which reversibly bind Rac1 and slow its diffusion, are sufficient to produce the spatially dependent molecular flow observed *in vivo*. Additionally, we find that if active and inactive Rac1 have different affinities for actin, then these two subpopulations would show different diffusive behaviors, similar to what has been observed experimentally.

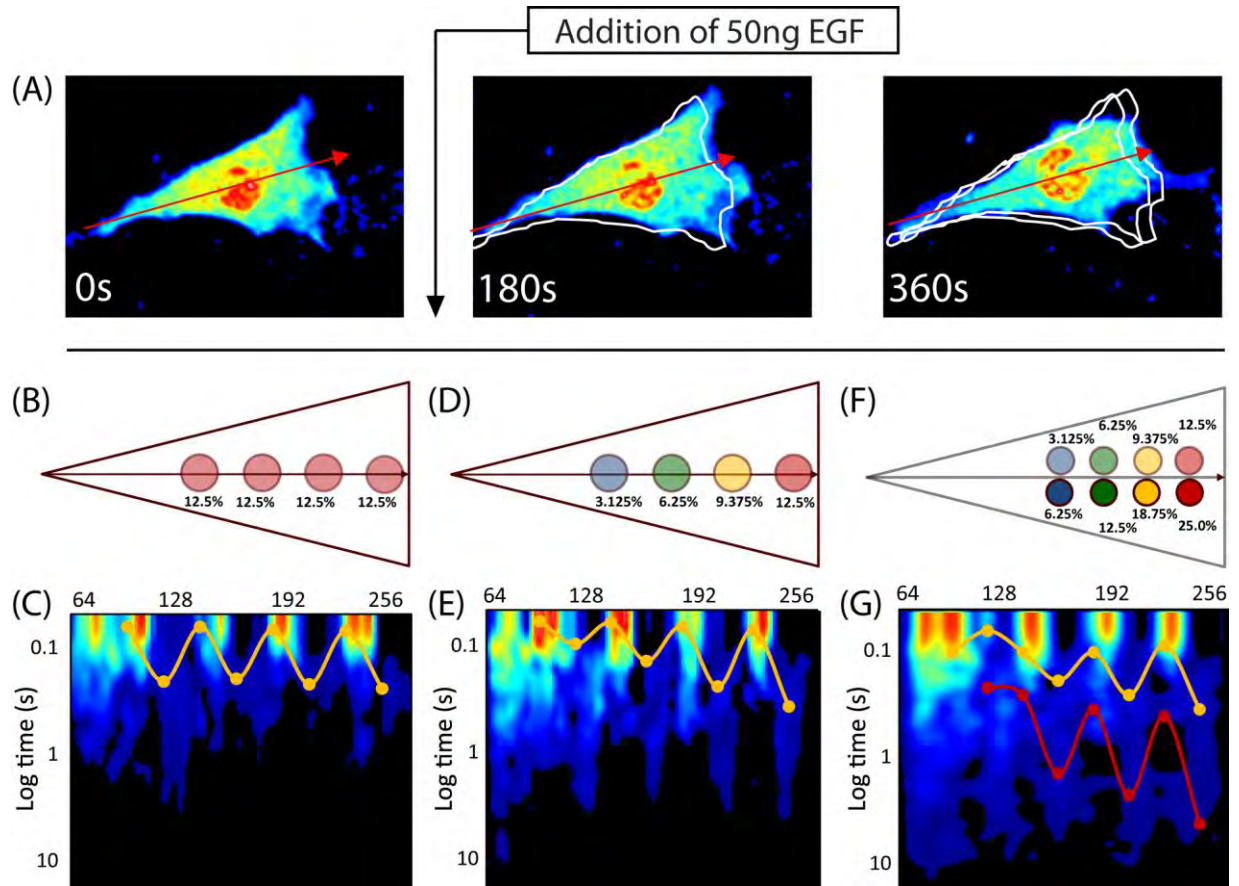


Fig. 1.4: Investigating the cellular substructure during three stages of EGF stimulation by comparing pCF carpets.

(A) Three snapshots of FLIM data before, 3 minutes after and 6 minutes after EGF stimulation. We aim to replicate features in the pCF carpets from the *in vivo* data (Figs. 1.2E – G) using our *in silico* simulations. (B – C) Uniform actin islands, as shown in Figs 1.1C and 1.1H. (D) Actin islands across the axis of the cell bind Rac1 with different affinities. These affinities range from 3.12% to 12.15% from the back to the

front, which is less than the simulation in Fig 1.3G. This arrangement of islands is similar to what we expect is present *in vivo*. (E) pCF analysis reveals four arc features of differing lengths, co-localized with the actin islands. Rac1 mobility is slower towards the front of the cell. (F) Modeling two populations of Rac1 by the superposition of two actin island gradients: one with twice the affinity (Fig. 1.3G) as the other (Fig. 1.4D). We combine the intensity carpets of the two simulations and perform pCF analysis. Although the two sets of islands are superimposed, here, we separate them for illustrative purposes. (G) The resulting pCF carpet for a cell with two populations of Rac1. We find two distinct gradients as highlighted by the curves: one for each population. The red curve corresponds to the Rac1 population with the higher actin affinity.

1.5 Discussion

Here, we developed a computational platform for performing stochastic simulations of intracellular diffusion to study how actin islands are organized to spatially regulate the mobility of signaling molecules. Our simulation platform allows the location, shape, protein binding and unbinding rates and diffusion rate for each island to be varied independently. This flexible computational model allows us to probe what cellular architectures underlie key features of pCF carpets calculated from *in vivo* experiments using a Rac1 biosensor.

As proof of this principle, we considered diffusive regulation of Rac1 at three stages during EGF stimulation (Fig. 1.4A). First, before stimulation, the molecular flow of Rac1 (assessed by the time needed for a Rac1 molecule to diffuse $0.5\mu\text{m}$) is spatially uniform (Fig. 1.2E). Second, for the first 3 minutes after stimulation, Rac1 mobility is inversely correlated with its proximity to the leading edge (Fig. 1.2F). Third, between 3 and 6 minutes after stimulation, this gradient of molecular flow is more pronounced (red curve Fig. 1.2G), and the presence of a distinct second set of arc features (yellow curve Fig. 1.2G) indicates a second population of Rac1 that moves differently than the first. To test if the existence of actin islands can produce arc features in the pCF carpet consistent with our experimental observations, we computationally simulated the diffusion of Rac1 in a cell with islands of equivalent affinity (Figs. 1.1H and 1.4C). The good qualitative agreement between the pCF carpets from *in vivo* (Fig. 1.2F) and *in silico*

(Fig. 1.4C) experiments suggests actin islands are present before EGF stimulation (Figs. 1.4A – C). Next, to test our hypothesis that actin island affinity regulates molecular flow, we performed simulations with actin islands of varying affinity (Figs. 1.3G and 1.4D) and compared these with the *in vivo* observed mobility gradient (Fig. 1.2F). The pCF carpets from these simulations (Figs. 1.3I and 1.4E) show a mobility gradient similar to the one observed *in vivo* (Fig. 1.2F); wherein, the time it takes for Rac1 molecules to flow 0.5 μ m is based on their proximity to the leading edge. Hence, the observed mobility gradient from the *in vivo* experiment is consistent with the presence of actin islands with progressively stronger affinity for Rac1 in moving from the rear to the front of the cell. Finally, we aimed to reproduce the *in vivo* observed dual regulation of Rac1 indicated by two distinct mobility gradients (Fig. 1.2G). We combine the results of two simulations, which are identical in every aspect, except for the actin islands' affinities (Fig. 1.4F). The resulting pCF carpet (Fig. 1.4G) shows two sets of arc features, similar to the ones observed *in vivo* (Fig. 1.2G). Hence, the two sets of mobility gradients observed from the *in vivo* experiments (Fig. 1.2G) may indicate the presence of two forms of Rac1 (e.g. inactive and active) with different affinities to the actin islands.

Our results suggest the Rac1 flow observed *in vivo* (Figs. 1.2F and 1.2G) is produced when actin islands near the leading edge have a higher affinity than islands near the trailing edge (Figs. 1.3G – I). There are many possible explanations for this discrepancy in affinities. Because the affinity in an island is proportional to the actin concentration, one possibility is that the actin concentration is higher in islands near the leading edge than those near the trailing edge. Another possibility is that the conformation of Rac1 near the leading edge is different than Rac1 near the trailing edge; moreover, this difference alters the affinity of Rac1 to actin. Another Rac1 binding

partner may account for the discrepancy in conformation sampling. We discuss the implications of each possible technique for controlling the affinity of actin islands below.

Based on the results of our experimental and computational investigations, we propose that cells can spatially regulate the molecular flow of certain proteins through the use of actin islands. In particular, our results suggest that cells can position the islands in regions where slower flow is desired, e.g. to sequester Rac1 at the leading edge. Because an island regulates molecular flow only locally, the cell can utilize actin islands only where needed. Additionally, the extent to which Rac1 mobility is slowed can be regulated by adjusting the binding affinity. Actin is known to reorganize in seconds [9], and this reorganization may play a role in regulating the position, size, shape and concentration of the islands. In turn, the concentration of actin in each island could affect the affinity for Rac1: denser islands have higher affinity than less dense islands. Organizing actin islands allows cells to spatially regulate molecular flow and therefore establish internal concentration gradients.

Another important feature is molecularly independent regulation; different protein species, and/or different forms of the same protein species, can be regulated separately. As long as each protein population has a different affinity for actin, each population will effectively experience a different set of actin islands (Figs. 1.4F – G). We successfully tested a scenario (data not shown), wherein the two populations had opposing molecular flow gradients (i.e. a combination of Figs. 1.3F and 1.3I) by combining opposing actin island gradients (i.e. Figs. 1.3D and 1.3G). Actin islands can differentially regulate the diffusion of different proteins across the cell; furthermore, pCF carpet analysis is capable of distinguishing these different protein populations and their unique, spatial regulation of molecular flow. Hence, the two sets of features from the *in vivo* experiment (Fig. 1.2G) indicate there are two populations of Rac1

(possibly the inactive and active forms) with different affinities for the actin islands and subsequently separately regulated flow.

1.6 Methods

We developed a simulation platform to test our proposed model of Rac1 behavior. Our program returns intensity carpets that are directly comparable to experimentally measured intensity carpets. As with the *in vivo* experimental data, we perform pair correlation analysis on the *in silico* intensity carpets to determine any spatial dependence on the diffusion constant. Our goal was to determine sufficient conditions for recapitulating the spatial dependence on the diffusion constant observed *in vivo*, by rearranging and varying the binding constant of the actin islands. Our program uses a particle-based stochastic simulation algorithm to simulate reactions and diffusion. Individual Rac1 molecules are modeled as point particles in continuous space and capable of reacting and diffusing discretely in time. This algorithm allows us to closely mimic the stochastic diffusion of Rac1 molecules inside the cell. A computational model that captures these natural fluctuations in concentration is imperative for pair correlation analysis.

1.6.1 Simulation Algorithm

A cell migrating on a 2D substrate in the xy-plane is modeled as an isosceles triangle whose base represents the direction of migration (Fig. 1.5). In this system, diffusion along the z-axis is largely inconsequential, because the molecular counts used to produce the intensity carpets *in vivo* are projected onto the xy-plane, thereby destroying any information from the third dimension. Additionally, the computational complexity is significantly reduced when only 2 dimensions are considered. Rac1 molecules are modeled as non-colliding point particles; their positions are stored in (x,y) Cartesian coordinates. Our hypothesized actin islands, regions of dense actin molecules, are capable of binding freely diffusing Rac1 molecules; hence, these actin

islands behave like diffusive traps, since actin-bound Rac1 has a smaller diffusion constant. The actin islands are modeled as circular patches, which can reversibly bind any Rac1 molecules located inside these regions (Fig. 1.5).

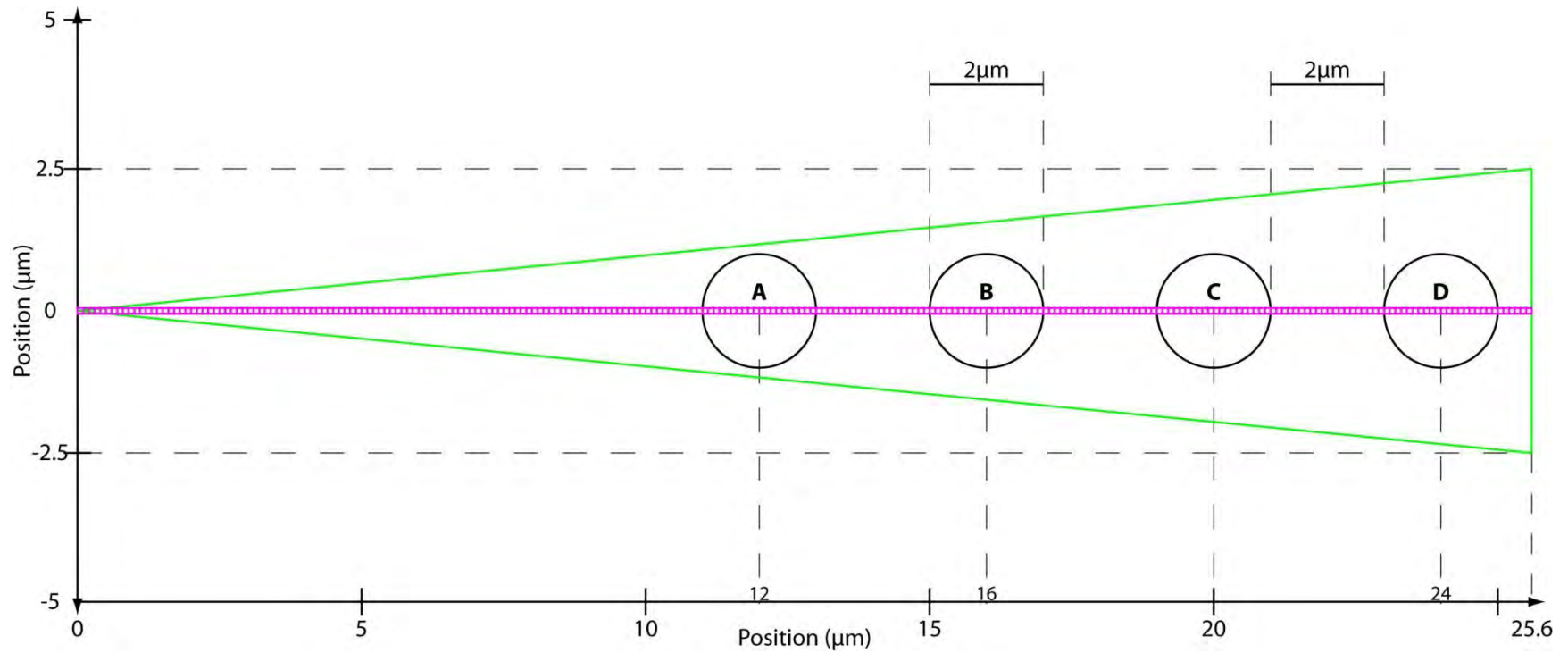


Fig. 1.5: Simulation Geometry.

The cell boundary is shown in green as an isosceles triangle $25.6\mu\text{m}$ wide and $5\mu\text{m}$ tall. The actin islands are shown as black circles with $1\mu\text{m}$ radii. As discussed in the text, unbound Rac1 molecules are free to diffuse throughout the cell and reflect off the cell boundary. Bound Rac1 molecules diffuse slower and are restricted to the inside of the island in which they are bound. We can individually set the K_D of each island thereby setting, on average, the percent of the total Rac1 population bound to each island. The 256 magenta boxes represent the $(0.1\mu\text{m})^2$ bins used to generate the intensity carpet; these bins do not affect the behavior of the molecules.

As suggested by [10], we use the Euler-Maruyama method [11] to perform stochastic simulations of diffusing particles. That is, knowing the current position of a Rac1 molecule, $(x(t), y(t))$, we calculate its position at the next time step, $(x(t+\Delta t), y(t+\Delta t))$, using the following equations:

$$x(t + \Delta t) = x(t) + W_i \sqrt{2D\Delta t} \quad (1.1a)$$

$$y(t + \Delta t) = y(t) + W_j \sqrt{2D\Delta t} \quad (1.1b)$$

Here, D is the diffusion constant, Δt is the time step and the W_n 's are random numbers drawn from a Gaussian distribution with a mean of 0 and a variance of 1. Unbound Rac1 molecules diffuse freely inside the cell with $D_{\text{unbound}} = 10\mu\text{m}^2/\text{s}$. The boundaries of the cell are reflective. Hence, if a molecule attempts to leave the cell (i.e. the diffusion calculation places the particle outside the cell boundaries), then the molecule is elastically reflected back inside the cell. Bound Rac1 molecules diffuse freely inside the actin island to which they are bound with $D_{\text{bound}} = 1\mu\text{m}^2/\text{s}$. The boundary of the island is reflective only to bound molecules. Hence, if a bound molecule attempts to leave the island (i.e. the diffusion calculation places the particle outside the island), then the molecule is elastically reflected back inside the island. Consequently, all bound Rac1 molecules are restricted to an island. The converse, however, is not true; not every Rac1 molecule positioned inside an actin island is bound. Unbound Rac1 molecules, which are positioned inside an island, are considered to be diffusing over/under/through that island without penalty. It is only in this condition that a binding reaction may occur.

In addition to diffusion, binding and unbinding events are also calculated at each time step. If a Rac1 molecule is, unbound and positioned inside an island at time t , then the state of that molecule can be switched to bound with probability: P_{bind} calculated for each island by:

$$P_{\text{bind}} = \left(\frac{v}{V}\right) \left(1 - e^{-\frac{\Delta t}{V} k_{\text{on}}}\right) \quad (1.2)$$

Here, k_{on} is the rate constant for binding to the island; Δt is the time step; v is the volume of the island, and V is the volume of the entire cell (we assume the cell is $1\mu\text{m}$ thick).

Separately, we calculate all unbinding events between time steps. If a Rac1 molecule is bound at time t , then its state can switch to unbound with probability P_{unbind} :

$$P_{\text{unbind}} = 1 - e^{-\Delta t \cdot k_{\text{off}}} \quad (1.3)$$

Here, k_{off} is the dissociation rate constant. When changing the state of a Rac1 molecule, due to either a binding or unbinding reaction, we do not change the position of the molecule.

We calculate the reaction rates based on an average fraction of total Rac1 molecules that we wish to have bound in each island:

$$\frac{1}{V} \frac{k_{\text{on}}}{k_{\text{off}}} = \frac{f}{1 - F} \quad (1.4)$$

Here, f is the average fraction of Rac1 molecules bound in the island, and $1 - F$ is the average fraction of Rac1 molecules not bound in any island. Note k_{on} should be expressed in units of $\mu\text{m}^3/\text{s}$ using a conversion factor such as $0.6022/(\text{nM} \cdot \mu\text{m}^3)$. Additionally, the fraction of molecules bound at each island also relates to the reaction probabilities:

$$\frac{v}{V} \frac{P_{\text{bind}}}{P_{\text{unbind}}} = \frac{f}{1 - F} \quad (1.5)$$

For example, if we desire each of the four islands to bind, on average, 6.25% of all Rac1 molecules, then $F = 0.25$ and for each island $f = 0.0625$. We know the volumes (v and V) from Fig. 1.5. Unfortunately, instead of a value for P_{bind} or P_{unbind} , we are only left with a ratio between these reaction probabilities. We proceed by assigning an off-rate within the biological regime.

We choose the off-rate such that the time-scale for unbinding is equal to the time-scale for a bound molecule to diffuse the length of the island:

$$\Delta x = \sqrt{2D\Delta t}$$

$$k_{\text{off}} = \frac{1}{\Delta t} = \frac{2D_{\text{bound}}}{\Delta x^2} = \frac{2\left(1\frac{\mu m^2}{s}\right)}{(2\mu m)^2} = 0.5\frac{1}{s} \quad (1.6)$$

This rate is substituted into Eq 1.3 to get P_{unbind} , which in turn, is substituted into Eq 1.5 to get P_{bind} . In this case, $P_{\text{unbind}} = 0.5 \times 10^{-6}$ and $P_{\text{bind}} = 0.85 \times 10^{-6}$ for a time-step $\Delta t = 1\mu s$. The reaction probabilities for each simulation are annotated in the Supplemental Methods (Section 1.7.5).

1.6.2 Pair Correlation Carpet

A line-scan technique was used during *in vivo* experiments [3,5,6], which reports the fluorescence intensity of Rac1-GFP molecules. The intensity is sequentially measured from pixel 1 to, about, pixel 300 to complete one line scan. Many line scans are taken and these scans are compiled into an intensity carpet on which pair correlation analysis can be performed. To produce an *in silico* analog to the intensity carpet, we tally the particles by their position as they are simulated (see above). To establish the line along which we will scan, we arrange 256 square bins sized $(0.1\mu m)^2$ along the center of the cell (Fig. 1.5). We cumulatively count the number of Rac1 molecules positioned inside a bin during a $25\mu s$ time window. Each bin is scanned sequentially; that is, we tally the counts from the first bin for the first $25\mu s$; then we tally the counts from the second bin for the second $25\mu s$ etc. After $6.4ms$ ($256 \times 25\mu s$) one full line scan is complete, and the next line scan begins immediately. Each intensity carpet has 47000 *lines*, corresponding to a simulation time of just over 5min. The bin size $(0.1\mu m)^2$ is roughly equivalent to the size of the laser used in the *in vivo* experiments, and the measuring time ($25\mu s$) is equivalent to the length of time a pixel was monitored *in vivo*.

To better understand the flow of particles within the cell, we calculate the pair correlation function (pCF) from the intensity carpet. First, we extract the intensity time trace (intensity versus time) for each pixel (e.g. cartooned white curves in Fig. 1.1D). Second, we calculate the correlation coefficient between two of these traces while sweeping two parameters in space (δr) and time (τ). With regards to space, we consider the traces of two pixels a distance δr apart; note that when $\delta r = 0$, we are computing the autocorrelation value (Figs. 1.1E and 1.1F). With regards to time, we shift the trace from the second pixel by τ seconds. Hence the correlation coefficient for a pixel given a pair of parameters ($\delta r, \tau$) describes the likelihood a particle will diffuse from that pixel to a pixel δr away in τ time. A value of 0 indicates this diffusive trajectory is impossible; while, a value of 1 implies a complete directed flow. Note that there is directionality in our pCF. A left to right calculation (e.g. correlating pixel 0 with pixel 5) describes the diffusion landscape from the back of the cell to the front (line scans are taken as such). A right to left calculation (e.g. correlating pixel 5 with pixel 0) describes the diffusion landscape from the front of the cell to the back.

Our results are independent of the cell's shape, because the cell boundary is sufficiently far from our region of interest. For example, the sharp, leftmost corner produces a slight force to the right; however, this force quickly dissipates and has little effect beyond $x = 5\mu\text{m}$. All the pCF carpets presented here show pCF analysis for $x \geq 6.4\mu\text{m}$.

1.7 Supplemental Methods

1.7.1 Analyzing pCF Carpets

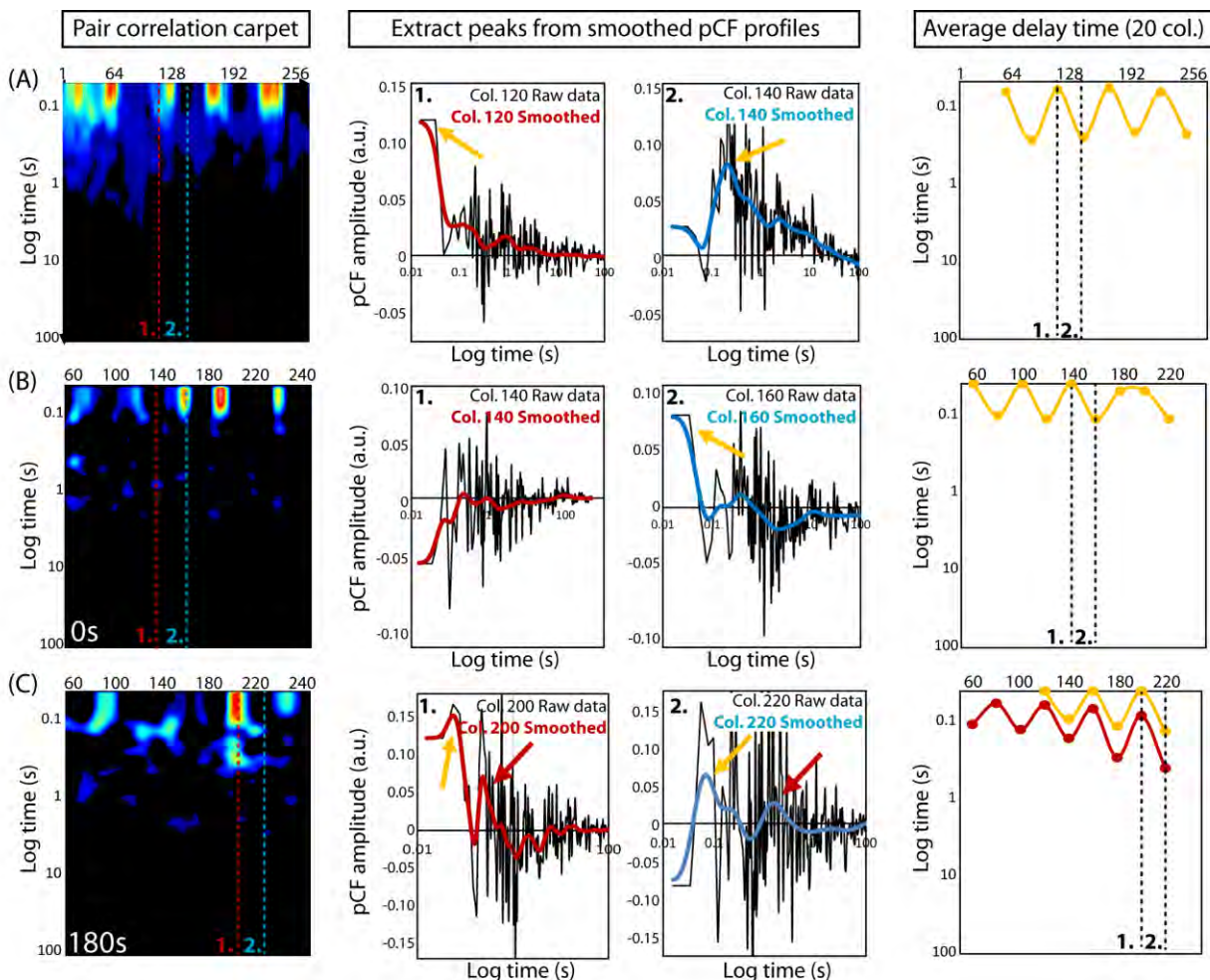


Fig. 1.6: Gaussian analysis of pair correlation carpets to extract average delay time(s) $Rac1$ takes to diffuse along simulated and *in vivo* line scans.

We smoothed every column in the derived pair correlation carpets by use of a Gaussian filter and then extracted the average peak time(s) for every 20 columns to obtain a simplified temporal profile of the molecular flow occurring at each consecutive location along the line scan. We then plotted the extracted peak time(s) as a function of position along the line scan. (A) Pair correlation carpet presented in Fig. 1.1H and an example of how the Gaussian filter smooths the pair correlation profile at two columns (120 and 140). (B) Pair correlation carpet presented in Fig. 1.2E and an example of how the Gaussian filter smooths the pair correlation profile at two columns (140 and 160). (C) Pair correlation carpet presented in Fig. 1.2F and an example of how the Gaussian filter smooths the pair correlation profile at two columns (200 and 220). The analysis presented in (A)-(C) enabled us to extract the average time(s) for molecular flow every 20 columns from pixel 60-220.

1.7.2 Simulation Program Details

We simulate 100,000 molecules, which, given the dimension of the cell (Fig. 1.5) and an assumed height of 1 μ m, corresponds to a concentration of 0.94 μ M. We monitor each molecule's position: (x,y) and state: 'bound' or 'unbound'. All molecules are updated simultaneously during each time step. Depending on the state of the particle, we calculate its new position (Diffusion) then determine if its state changes (Reaction). Consider the following, short pseudo-code.

For each time step:

- I. 'Unbound' molecules: diffuse with $D_{\text{unbound}} = 10 \mu\text{m}^2/\text{s}$
 - a. If outside the cell, then reflect position back inside cell
 - b. If inside an island, then generate a random number to determine if 'bound'
- II. 'Bound' molecules: diffuse with $D_{\text{bound}} = 1 \mu\text{m}^2/\text{s}$
 - a. If outside the island, then reflection position back inside the island
 - b. Generate random number to determine if 'unbound'

1.7.3 Calculating Diffusion Steps

We use Eq 1.1 to calculate the molecules' new position. That is, for each spatial dimension: $x(t + \Delta t) = x(t) + W\sqrt{2D\Delta t}$. Here, W is a random number drawn from a Gaussian Distribution with mean 0 and variance 1. We calculate this random number using the Box-Muller Method, as implemented in the "curand" library of the CUDA 4.0 Toolkit.

Molecules are restricted to certain regions: 'unbound' molecules must remain inside the cell, and 'bound' molecules must remain inside the island to which they are bound. These rules are enforced by imposing reflective boundaries at the cell wall and the edges of the actin islands respectively. In the next two sections, we discuss how the reflections for each boundary are calculated.

Reflections off Cell Boundaries

The cell's boundary can be described as three lines: the 'top' boundary traversing from (0, 0) to (25.6, 2.5), the 'bottom' boundary traversing from (0, 0) to (25.6, -2.5), and the 'right'

boundary traversing from (25.6, -2.5) to (25.6, 2.5) (Fig. 1.5). We describe each line in the form: $ax + by + c = 0$ (restricting $a > 0$). Specifically, the ‘top’ boundary is:

$$2.5x - 25.6y + 0 = 0 \quad (1.7)$$

The ‘bottom’ boundary is:

$$2.5x + 25.6y + 0 = 0 \quad (1.8)$$

The ‘right’ boundary is:

$$2.5x + 0y - 25.6 = 0 \quad (1.9)$$

Let (x', y') be the position that lies outside the cell. Accordingly, the distance from the point (x', y') to a line can be calculated using the expression:

$$d = \frac{|ax' + by' + c|}{\sqrt{a^2 + b^2}} \quad (1.10)$$

Hence, the point (x', y') can be reflected off the ‘top’ and ‘bottom’ boundaries using the equation:

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x' \\ y' \end{bmatrix} + \frac{2d}{\sqrt{a^2 + b^2}} \begin{bmatrix} a \\ b \end{bmatrix} \quad (1.11)$$

To reflect the point (x', y') off the ‘right’ boundary, we calculate:

$$x = x' - 2 * \left| x' + \frac{c}{a} \right| \quad (1.12)$$

During the diffusion of an ‘unbound’ particle, we check if the new position is outside the cell. That is, the position is above the ‘upper’ boundary: $y' > \left(\frac{2.5}{25.6}\right)x'$, below the ‘lower’ boundary: $y' < \left(\frac{-2.5}{25.6}\right)x'$, and/or beyond the ‘right’ boundary: $x' > 25.6$. If the particle is out of bounds, then we reflect the proposed position with respect to the appropriate boundary. For example, if the position is too high, then we reflect the position about the upper boundary. This method of checking and reflecting is repeated until all checks pass (i.e. the position is inside the cell) or twenty-five reflections have been made. This latter condition protects against pathogenic

cases that lead to an infinite check/reflect loop. These rare cases occur when the position is close (within numerical error) to a corner of the cell boundary. These cases are resolved during diffusion in the next time step.

Reflections off Actin Island Boundaries

To reflect a particle inside an actin island (which is a circle in our 2D model), we must know the particle's initial position inside the island, \vec{i} , the proposed position outside the island, \vec{p} , and the center position of the island, \vec{c} (Fig. 1.7).

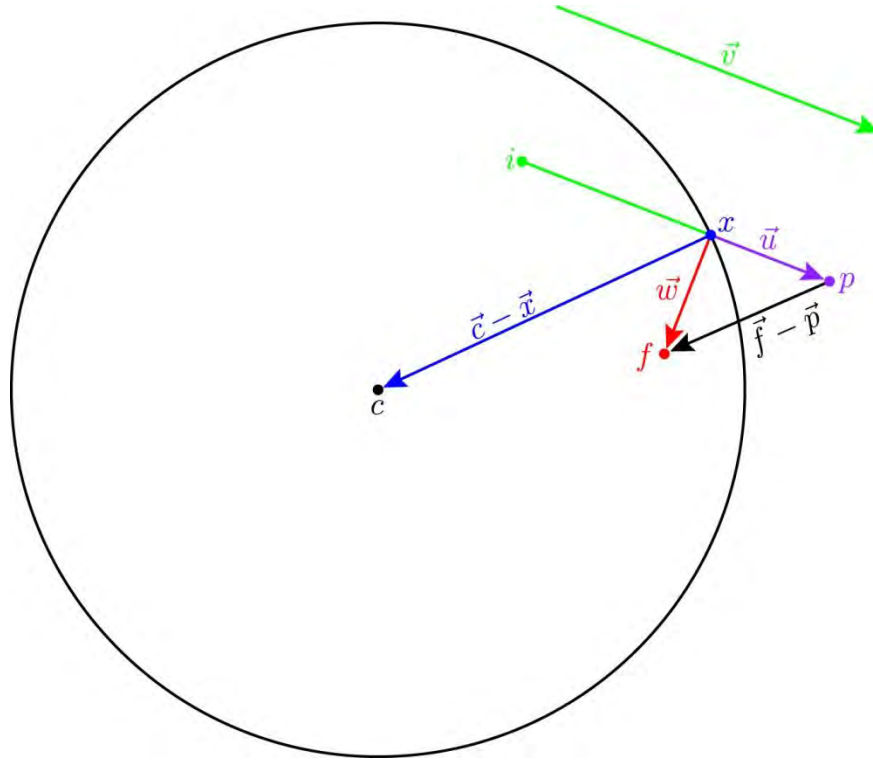


Fig. 1.7: Important Vectors for Reflecting Inside Actin Island.

Here, we diagram an example of a ‘bound’ particle being reflected back inside an actin island while diffusing during a single time step. Initially the particle is at point i ; after freely diffusing for one time step, the proposed final position is at point p . The particle crosses the circular actin island's boundary at point x . For a reflective boundary, the true final position is at point f . Important vectors between these points are indicated. See the text below for details on how to determine the final, reflected position given the points i , p , c and the radius r of the island.

First, we find the intersection point (denoted \vec{x}) between the attempted trajectory ($\vec{v} = \vec{p} - \vec{i}$) and the circle. The attempted trajectory defines a line which can be written $\vec{i} + t\vec{v}$.

$$\vec{x} = \vec{l} + t_x \vec{v} \quad (1.13)$$

Because \vec{x} lies on the boundary, we know the following constraint: $(\vec{x} - \vec{c}) \cdot (\vec{x} - \vec{c}) = r^2$, where r is the radius of the island. After substituting Eq 1.13 into the constraint, we solve for t_x :

$$t_x = \frac{\vec{v} \cdot (\vec{c} - \vec{l}) + \sqrt{\mathbf{Det}}}{\|\vec{v}\|^2} \quad (1.14)$$

$$\mathbf{Det} = (\vec{v} \cdot (\vec{c} - \vec{l}))^2 + (\|\vec{v}\|^2)(r^2 - \|\vec{c} - \vec{l}\|^2)$$

Hence, numerically, \vec{x} is found by substituting known values into Eq 1.14 then Eq 1.13.

The vector \vec{u} ($\vec{u} = \vec{p} - \vec{x}$) is the erroneous portion of the particle's trajectory; that is, the portion in which the particle traveled outside the actin island. To correct the trajectory and reflect the particle back inside the island, we reverse the radial component of \vec{u} . The vector \vec{w} is the corrected continuation of the trajectory. The magnitudes of \vec{u} and \vec{w} are equivalent. They differ by twice the radial component of \vec{u} (Eq 1.15). We denote the direction pointing towards the center of the island as $\widehat{(\vec{c} - \vec{x})} = \frac{\vec{c} - \vec{x}}{\|\vec{c} - \vec{x}\|} = \frac{\vec{c} - \vec{x}}{r}$.

$$\begin{aligned} \vec{w} - \vec{u} &= 2(-(\vec{u} \cdot \widehat{(\vec{c} - \vec{x})}))\widehat{(\vec{c} - \vec{x})} \\ \vec{w} - \vec{u} &= \left(\frac{-2}{r^2}\right)(\vec{u} \cdot (\vec{c} - \vec{x}))(\vec{c} - \vec{x}) \end{aligned} \quad (1.15)$$

And, because $\vec{w} - \vec{u} = \vec{f} - \vec{p}$, we can calculate \vec{f} from Eq 1.15 using the expression:

$$\vec{f} = \vec{p} - \left(\frac{2}{r^2}\right)(\vec{u} \cdot (\vec{c} - \vec{x}))(\vec{c} - \vec{x}) \quad (1.16)$$

We diffuse a 'bound' particle within the circular actin island. The particle is allowed to diffuse freely in space. If the particle lands on the edge of or outside the island, then we reflect the particle back inside. To catch any pathological cases, arising from numerical error & computational limitations, we re-check if the particle is completely inside. In some rare cases,

this re-check will fail; that is, even after reflecting the particle, the new position will still be on the edge of or outside the island. For those scenarios, we repeat the entire process and calculate a different diffusion step. We find these pathological cases occur when the starting position is close (within the numerical error) to the island boundary.

1.7.4 Calculating Reaction Probabilities

After calculating the diffusion step, an ‘unbound’ molecule positioned in an island has a probability to bind, and a ‘bound’ molecule has a probability to unbind. We calculate these probabilities such that the reactions are simulated in a biophysically relevant regime. Our proposed model for the diffusion of Rac1 in the cell does not (nor is it necessary to) speculate on exact values of the reaction rates for the different actin islands. Instead, our model proposes different binding affinities in each island. Below, we derive the relationship between the reaction rates of each island and the probabilities of reaction.

We write out the Master Equation for the binding reaction on the i^{th} island: let A denote the number of ‘unbound’ Rac1 throughout the cell and A_i^* denote the number of ‘bound’ Rac1 in the i^{th} island.

$$\dot{A}_i^* = \left(k_{\text{on}}^i \cdot \frac{1}{V}\right) A - (k_{\text{off}}^i) A_i^* \quad (1.17)$$

Here, k_{on}^i and k_{off}^i are the reaction rates for the i^{th} island, and V is the volume of the cell (we set the cell thickness to $1\mu\text{m}$). The unbinding term, $(k_{\text{off}}^i) A_i^*$, describes the number of molecules which will unbind from the i^{th} island during the infinitesimal time dt . The unbinding term is derived from multiplying the off rate by the number of molecules eligible to unbind: A_i^* . The binding term, $\left(k_{\text{on}}^i \cdot \frac{1}{V}\right) A$, describes the number of molecules that will bind to the i^{th} island

during the infinitesimal time dt . The binding term is derived from multiplying the on-rate by the “concentration” of the island and by the number of molecules eligible to bind to the island.

$$(k_{\text{on}}^i) \left(\frac{1}{v_i} \right) \left(\frac{v_i}{V} A \right) = \left(k_{\text{on}} \cdot \frac{1}{V} \right) A \quad (1.18)$$

Here, v_i is the volume of the i^{th} island; hence, $\frac{1}{v_i}$ is the “concentration” (since the unit for concentration is inverse volume) of actin. The number of molecules eligible to bind is the number of unbound molecules located inside the island; this value is equal to the total number of unbound molecules, A , scaled by the probability of being located in the i^{th} island: $\frac{v_i}{V}$.

At equilibrium, the number of binding and unbinding events is balanced; we set Eq 1.17 equal to 0 and solve for the K_D .

$$K_D^i = \frac{k_{\text{off}}^i}{k_{\text{on}}^i} = \frac{A}{A_i^*} \frac{1}{V} \quad (1.19)$$

We multiply the right-hand-side by $\frac{A_{\text{Tot}}}{A_{\text{Tot}}}$, in order to rewrite the equation in terms of the fraction (with respect to the entire population) of bound and unbound molecules. We denote the fraction of molecules bound to the i^{th} island as $\frac{A_i^*}{A_{\text{Tot}}} = f_i$. We write the fraction of unbound molecules as $\frac{A}{A_{\text{Tot}}} = 1 - \sum f_i = 1 - F$, where F is the total fraction of molecules bound in all islands.

$$\frac{k_{\text{off}}^i}{k_{\text{on}}^i} = \frac{1}{V} \frac{A}{A_i^*} \cdot \frac{A_{\text{Tot}}}{A_{\text{Tot}}} = \frac{1 - F}{f_i} \frac{1}{V} \quad (1.20)$$

We determine the probability a molecule reacts (both binding and unbinding processes) in a time Δt . We choose a sufficiently small time step ($1\mu\text{s}$) such that we may assume a single molecule can perform only one reaction (bind or unbind) during that time step. Hence, we may separately solve the differential equations describing each reaction (i.e. each term in Eq 1.17).

The probability a molecule unbinds during a time step is equal to:

$$\begin{aligned}
P_{\text{unbind}} &= \frac{\# \text{ molecules that unbind in } \Delta t}{\# \text{ bound molecules at time } t} \\
&= \frac{A_i^*(t) - A_i^*(t + \Delta t)}{A_i^*(t)}
\end{aligned} \tag{1.21}$$

Where $A_i^*(t)$ is the number of molecules bound to the i^{th} island as a function of time. We choose a small time step such that only the i^{th} island unbinds.

$$\begin{aligned}
\dot{A}_i^* &= -k_{\text{off}} A_i^* \\
A_i^*(t) &= C e^{-k_{\text{off}} \cdot t}
\end{aligned} \tag{1.22}$$

Substituting Eq 1.22 into Eq 1.21, we find:

$$P_{\text{unbind}} = 1 - e^{-k_{\text{off}} \cdot \Delta t} \tag{1.23}$$

The probability any unbound molecule binds to the i^{th} island during a time step is equal to:

$$\begin{aligned}
P_{\text{on}} &= \frac{\# \text{ molecules that bind in } \Delta t}{\# \text{ unbound molecules at time } t} \\
&= \frac{A(t) - A(t + \Delta t)}{A(t)}
\end{aligned} \tag{1.24}$$

Where $A(t)$ is the number of unbound molecules as a function of time. We choose a small time step such that only the i^{th} island binds.

$$\begin{aligned}
\dot{A} &= -k_{\text{on}}^i \cdot \frac{1}{V} A \\
A(t) &= C e^{-k_{\text{on}}^i \cdot \frac{1}{V} \cdot t}
\end{aligned} \tag{1.25}$$

Substituting Eq 1.25 into Eq 1.24, we find:

$$P_{\text{on}} = 1 - e^{-k_{\text{on}}^i \cdot \frac{1}{V} \cdot \Delta t} \tag{1.26}$$

Eq 1.26 describes the probability that any unbound molecule will bind to the i^{th} island. We rewrite this overall probability as joint probability of, one, the probability an unbound molecule is located inside the i^{th} island and, two, the probability that the same molecule binds:

$$P_{\text{on}} = P_{\text{Island}} P_{\text{bind}} = \left(\frac{v_i}{V}\right) P_{\text{bind}} \quad (1.27)$$

Hence, the probability that an unbound molecule located inside an actin island binds during a time step Δt is given by:

$$P_{\text{bind}} = \left(\frac{V}{v_i}\right) \left(1 - e^{-k_{\text{on}}^i \cdot \frac{1}{V} \cdot \Delta t}\right) \quad (1.28)$$

The exponentials in Eqs 1.23 and 1.28 can be simplified by using the first two terms of their Taylor Expansion. These two expressions for the probability of unbinding and binding can then be written as:

$$P_{\text{unbind}} = k_{\text{off}} \cdot \Delta t \quad (1.29)$$

$$P_{\text{bind}} = k_{\text{on}} \cdot \frac{1}{v_i} \cdot \Delta t \quad (1.30)$$

We combine Eqs 1.20, 1.29 and 1.30 to get three equal expressions for the K_D relating the reaction rates, fraction bound in each island and the probabilities of reacting.

$$K_D^i = \frac{k_{\text{off}}^i}{k_{\text{on}}^i} = \frac{1 - F}{f_i} \frac{1}{V} = \frac{P_{\text{unbind}}}{P_{\text{bind}}} \frac{1}{v_i} \quad (1.31)$$

In every simulation, we customize the affinities for each island; that is, we choose the set of f_i .

We perform two simulations for each set of affinities: in one, the off-rate is assigned a value and kept constant while the on-rate is varied; in the other, the on-rate is assigned a value and kept constant while the off-rate is varied. The affinities of the islands and the macro-chemistry are the same for both cases; however, the mechanism is slightly different between the two cases. An off-rate is assigned such that the time scale for unbinding is equal to the time scale for diffusing

across an actin island in the ‘bound’ state (Eq 1.6). Table 1.1 summarizes the reaction rates, fraction bound and reaction probabilities for each simulation.

1.7.5 Full Set of Simulation Parameters

In all simulations, the volume of the cell ($V = 64\mu\text{m}^3$), the volumes of all islands ($v = 3.14\mu\text{m}^3$), and the time step ($\Delta t = 1\mu\text{s}$), are all kept constant. We vary the fraction bound in each island. In Table 1.1, the values for the islands are reported as $\{A, B, C, D\}$, where ‘A’, ‘B’, ‘C’ & ‘D’ are the labels for each island as shown in Fig 1.5.

Simulation Set-up	\vec{f}	$\overrightarrow{k_{\text{off}}} \left(\frac{1}{s} \right)$	$\overrightarrow{k_{\text{on}}} \left(10^9 \frac{1}{M \cdot s} \right)$	$\overrightarrow{P_{\text{unbind}}}$	$\overrightarrow{P_{\text{bind}}}$
Fig 1.1C & 1.4B	{0.125, 0.125, 0.125, 0.125}	{0.5, 0.5, 0.5, 0.5}	{4.8, 4.8, 4.8, 4.8}	{ 0.5×10^{-6} , 0.5×10^{-6} , 0.5×10^{-6} , 0.5×10^{-6} }	{ 2.55×10^{-6} , 2.55×10^{-6} , 2.55×10^{-6} , 2.55×10^{-6} }
Fig 1.3A	{0.1875, 0.0625, 0.0625, 0.0625}	{0.5, 0.5, 0.5, 0.5}	{5.78, 1.927, 1.927, 1.927}	{ 0.5×10^{-6} , 0.5×10^{-6} , 0.5×10^{-6} , 0.5×10^{-6} }	{ 3.056×10^{-6} , 1.019×10^{-6} , 1.019×10^{-6} , 1.019×10^{-6} }
Fig 1.3D	{0.25, 0.1875, 0.125, 0.0625}	{0.5, 0.5, 0.5, 0.5}	{13, 9.7, 6.5, 3.36}	{ 0.5×10^{-6} , 0.5×10^{-6} , 0.5×10^{-6} , 0.5×10^{-6} }	{ 6.79×10^{-6} , 5.09×10^{-6} , 3.4×10^{-6} , 1.7×10^{-6} }
Fig 1.3G	{0.0625, 0.125, 0.1875, 0.25}	{0.5, 0.5, 0.5, 0.5}	{3.36, 6.5, 9.7, 13}	{ 0.5×10^{-6} , 0.5×10^{-6} , 0.5×10^{-6} , 0.5×10^{-6} }	{ 1.7×10^{-6} , 3.4×10^{-6} , 5.09×10^{-6} , 6.79×10^{-6} }
Fig 1.4D	{0.03125, 0.0625, 0.09375, 0.125}	{0.5, 0.5, 0.5, 0.5}	{0.912, 1.75, 2.6, 3.5}	{ 0.5×10^{-6} , 0.5×10^{-6} , 0.5×10^{-6} , 0.5×10^{-6} }	{ 4.63×10^{-7} , 9.26×10^{-7} , 1.39×10^{-6} , 1.85×10^{-6} }

Table 1.1 Reaction Rates and Probabilities

The reaction rates and corresponding probabilities are shown for each simulation presented in this chapter. Column 1 indicates a figure, which shows a cartoon sketch of a simulation. Column 2 shows the average fraction of molecules bound to each island. Columns 3 and 4 indicate the reaction rates for each island, and columns 5 and 6 show the corresponding reaction probabilities.

1.7.6 Generating an Intensity Carpet

To generate an intensity carpet as done *in vivo* [3,5,6] for our *in silico* data, we tally the number of molecules in each bin sequentially, measuring for 25 μ s at a time. If we label the 256 bins from 0 to 255 and let t be the current simulation time (in μ s), then the current bin being tallied is

$$\left(\text{floor}\left(\frac{t}{25}\right)\right) \bmod 256 \quad (1.32)$$

Where ‘*floor*(x)’ is a function which returns the largest integer that is less than or equal to x . And where ‘*mod*’ indicates the modulo arithmetic operator. The tally is incremented for every particle positioned inside the current bin after diffusing and reacting. For example: for $t \in [0, 24]$, we tally the particles in bin 0; for $t \in [25, 49]$, we tally the particles in bin 1; for $t \in [6375, 6399]$, we tally the particles in bin 255, and for $t \in [6400, 6424]$, we tally the particles in bin 0. We store these tallies as a matrix with 256 columns (one for each bin), and 47000 rows (one for each line scan).

1.7.7 Hardware Details

This program was written in CUDA C using the CUDA 4.0 Toolkit. CUDA is a publicly available parallel programming architecture built for Generally Programmable Graphics Processing Units (GP-GPUs), specifically, those manufactured by the NVIDIA ® Corporation. Utilizing GP-GPUs, we are able to simultaneously access hundreds of computational processors in parallel. Because each particle is independent of all others, we parallelize our algorithm at the particle level. That is, we launch one computational thread in the GPU for each particle in our simulation. Each thread performs all necessary calculations for reacting and diffusing a single particle. All threads and particles are periodically synchronized in time. Every 12.8ms of simulation time, we stop the threads’ calculations and have the CPU download the tallies for the

Intensity Carpet to the RAM. Additionally, the CPU downloads & prints to a file the position & state data of all particles. We run our simulations on the Kill Devil Research Computing Cluster at UNC. Specifically, the program is executed on a computing node whose hardware includes a 12-core 2.67 GHz Intel processor, 48 GB of RAM and an NVIDIA M2070 GPGPU.

REFERENCES

1. Burridge K, Wennerberg K (2004) Rho and Rac Take Center Stage. *Cell* 116: 167–179.
2. Machacek M, Hodgson L, Welch C, Elliott H, Pertz O, et al. (2009) Coordination of Rho GTPase activities during cell protrusion. *Nature* 461: 99–103.
3. Hinde E, Digman M a, Hahn KM, Gratton E (2013) Millisecond spatiotemporal dynamics of FRET biosensors by the pair correlation function and the phasor approach to FLIM. *Proc Natl Acad Sci U S A* 110: 135–140.
4. Katsuki T, Joshi R, Ailani D, Hiromi Y (2011) Compartmentalization within neurites: its mechanisms and implications. *Dev Neurobiol* 71: 458–473.
5. Digman M a, Gratton E (2009) Imaging barriers to diffusion by pair correlation functions. *Biophys J* 97: 665–673.
6. Hinde E, Cardarelli F (2011) Measuring the flow of molecules in cells. *Biophys Rev* 3: 119–129.
7. Hinde E, Cardarelli F, Digman MA, Gratton E (2010) In vivo pair correlation analysis of EGFP intranuclear diffusion reveals DNA-dependent molecular flow. *Proc Natl Acad Sci U S A* 107: 16560–16565.
8. Moissoglu K, Slepchenko BM, Meller N, Horwitz AF, Schwartz MA (2006) In vivo dynamics of Rac-membrane interactions. *Mol Biol Cell* 17: 2770–2779.
9. Andrews NL, Lidke KA, Pfeiffer JR, Burns AR, Wilson BS, et al. (2008) Actin restricts FcepsilonRI diffusion and facilitates antigen-induced receptor immobilization. *Nat Cell Biol* 10: 955–963.
10. Erban R, Chapman SJ (2009) Stochastic modelling of reaction-diffusion processes: algorithms for bimolecular reactions. *Phys Biol* 6: 046001.
11. Higham. DJ (2001) An Algorithmic Introduction to Numerical Simulation of Stochastic Differential Equations. *SIAM Rev* 43: 525–546.

CHAPTER 2: A MODEL OF GRADIENT SENSING IN THE CONTEXT OF YEAST MATING²

Introduction

Sensing an external chemical gradient is a fundamental process in cell biology present in areas such as cancer metastasis, wound healing and embryogenesis. This ability, of decoding spatial information, allows cells to move or grow towards a favorable environment or away from an unfavorable environment. Most eukaryotic cells use a spatial detection mechanism, wherein one portion of the cell membrane has more active receptors than the other regions. This skewed distribution of active receptors reflects the external chemical gradient. Some cells, like yeast (*S. Cerevisiae*), are capable of detecting extremely shallow gradients, wherein the noise from molecular diffusion and stochastic reactions is thought to obscure the desired spatial information. Accordingly, various mechanisms have been proposed to explain how yeast cells overcome this poor signal to noise ratio. To quantitatively assess these mechanisms, we develop a particle-based stochastic reaction-diffusion model. This approach monitors the position of individual molecules and models the two sources of noise present during gradient sensing: diffusion and stochastic binding and unbinding reactions. Here, we first outline the model used to investigate the distribution of active receptors by simulating the reversible binding reaction between pheromone and receptor. Second, to evaluate two potential noise-reduction mechanisms, we discuss adding receptor endocytosis and the pheromone protease Bar1 to our model.

² This chapter is being drafted as part of a publication to be submitted to the journal PLoS Computational Biology. It is authored by Lakhani V and Elston T.

2.1 Overview

We first describe the important biophysical features of gradient sensing that the model should capture. The model must be able to resolve individual signaling molecules in a continuous 3-dimensional space. It also should faithfully capture the stochastic properties of diffusion of both the extracellular signaling molecules and receptors in the cell membrane, and of the biochemical reactions involved in ligand binding and release and receptor internalization. For these reasons, we choose a Particle-Based Stochastic Reaction-Diffusion model. When simulating this model, molecules are modeled as non-colliding point particles which can stochastically diffuse and react, discretely in time and continuously in space.

We simulate molecules in a cubic volume ($10\mu\text{m}$ by $10\mu\text{m}$ by $10\mu\text{m}$). We model the cell as a sphere with a radius of $2.5\mu\text{m}$ at the center of the volume (Fig. 2.1). The state of the system is defined by the position of every pheromone and receptor molecule as well as the state of each receptor (occupied or empty). Pheromone molecules cannot be located inside the cell, and receptor molecules are restricted to the surface of the cell. Given the current state at time t_0 , we determine the subsequent state at time $t_0 + \Delta t$, where Δt is a time step of fixed length, by calculating all binding reactions, unbinding reactions and diffusion of each molecule. A binding event occurs during the time step, Δt , with probability P_{bind} (Table 2.1), if a pheromone molecule is within $r_{\text{bind}} = 4\text{nm}$ (binding radius) of an unbound receptor (Fig. 2.1A). An unbinding event occurs during the time step, Δt , with probability $P_{\text{unbind}} = 1.1 \times 10^{-9}$ (Table 2.1). A bound receptor releases a pheromone molecule 4nm (unbinding radius) away at a randomly chosen angle (Fig. 2.1A). During diffusion, each particle stochastically moves to a new position with appropriate conditions enforced at the boundaries of the cubic volume and the cell (Fig. 2.1B – D). The time step is chosen such that the length scale of pheromone diffusion ($\sqrt{2D_{\alpha}\Delta t}$) is similar to the

binding radius (4nm). Our model captures the stochasticity due to reactions and diffusion, and our model monitors the exact position of all molecules in a 3D space. Below, we explain the microscopic rules that each molecule follows.

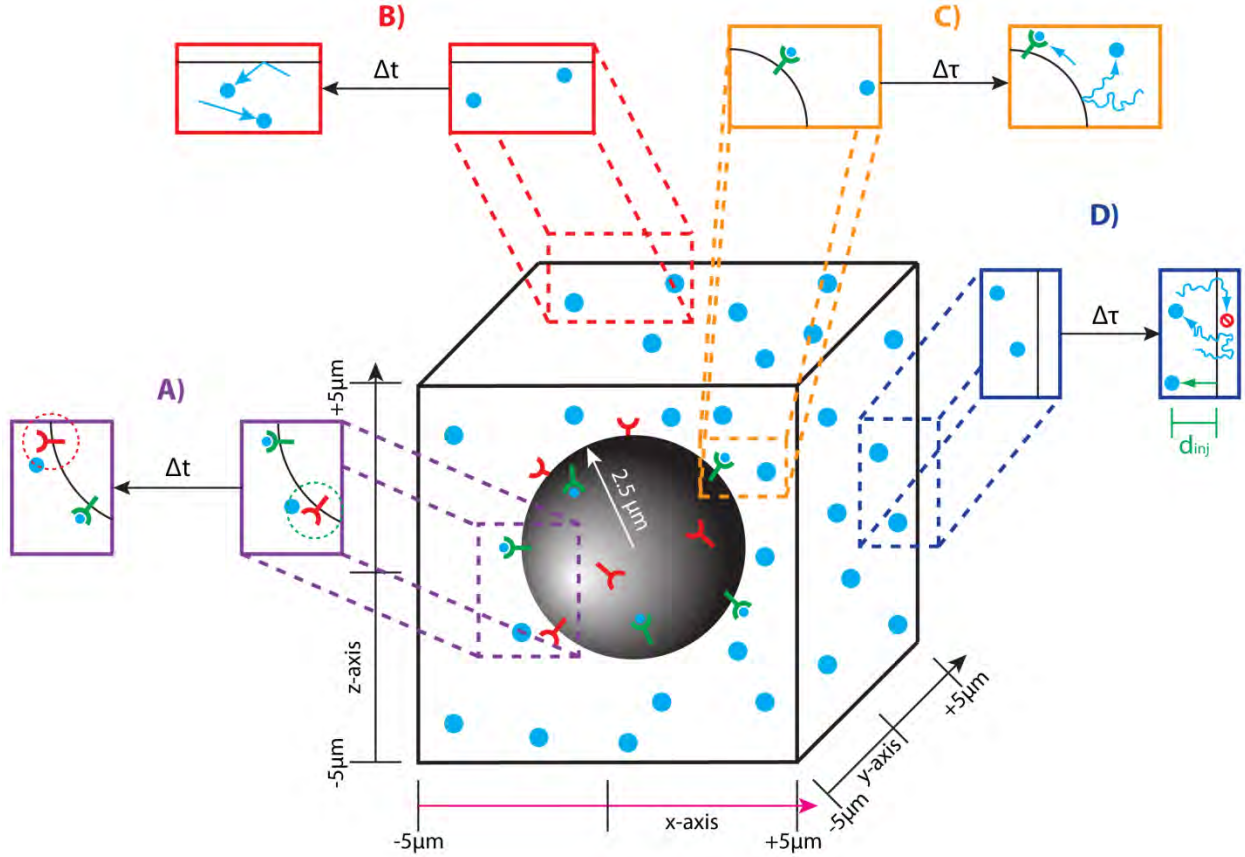


Fig. 2.1: Simulations on our Particle-Based Stochastic Reaction-Diffusion Model

Illustration for microscopic rules. (A) Rules for the binding and unbinding reactions. During each time step (Δt), an 'unbound' receptor (Y) may capture one, nearby pheromone molecule (●). The dashed green circle indicates the binding radius; only pheromone molecules within this region can bind. A 'bound' receptor (Y) may unbind. The new pheromone molecule is released a fixed distance (the unbinding radius) away from the receptor, indicated by the dashed red circle. No pheromone molecules are created inside the cell. (B) Rules for diffusion near the 'y' and 'z' boundaries. Pheromone diffusion is calculated every time step (Δt). Any molecule that diffuses outside the 'y' or 'z' domains are reflected back into the volume of interest. (C) Rules for diffusion near the cell surface. Pheromone molecules reflect off the surface of the cell. Here, the molecule is shown to have taken many diffusion steps, because a longer time has elapsed ($\Delta \tau$) than in (A) or (B). Receptor diffusion on the cell surface is calculated on this longer time scale ($\Delta \tau$). (D) Rules for diffusion near the 'x' boundaries. At these boundaries, pheromone molecules freely diffuse for a time ($\Delta \tau$), shown by the two wiggly tracks. After $\Delta \tau$ time has elapsed, all pheromone molecules outside the 'x' boundaries are removed (fate of top molecule). No consequences are suffered by molecules that exit and re-enter the volume (fate of middle molecule). Lastly, new molecules are created every $\Delta \tau$ near the 'x' boundaries (new bottom molecule). The new molecule is placed a random length (called the 'injection distance') away from the 'x' boundary (shown by the green arrow).

Table 2.1: A Typical Parameter Set

The first portion contains all the parameters necessary to uniquely define a simulation of our model. The second portion contains additional, informative values that are calculable from the first portion.

	Parameter	Value	Description
Customizable Parameters	XDom	10 μm	Length of x-domain
	YDom	10 μm	Length of y-domain
	ZDom	10 μm	Length of z-domain
	Δt	1 μs	Time Step
	$\Delta \tau$	50 μs	Coarse Time Step
	R	2.5 μm	Radius of Cell
	D_a	125 $\mu\text{m}^2/\text{s}$	Pheromone Diffusion Constant [1]
	Grad	0.1 nM/ μm	Pheromone Concentration Gradient along x-axis
	Conc	6.9 nM	Background Pheromone Concentration (equal to K_D)
	D_{Ste2}	0.0025 $\mu\text{m}^2/\text{s}$	Receptor Diffusion Constant [2]
	N	10000	Number of Receptors on Cell Surface [3–6]
	k_{on}	$1.6 \times 10^5 (\text{M} \cdot \text{s})^{-1}$	Binding Rate [7]
	k_{off}	0.0011 s^{-1}	Unbinding Rate [7]
	r_{bind}	4 nm	Binding Radius
	r_{unbind}	4 nm	Unbinding Radius
Additional Values	P_{bind}	0.002	Binding Probability
	P_{unbind}	1.1×10^{-9}	Unbinding Probability
	K_D	6.9 nM	Pheromone/Receptor Dissociation Constant [7]
	c_{high}	7.4 nM	Concentration at High Boundary ($x = 5\mu\text{m}$)
	c_{low}	6.4 nM	Concentration at Low Boundary ($x = -5\mu\text{m}$)
	$n_{\text{inj High}}$	19.88	Average Number of Pheromone to create at High Boundary
	$n_{\text{inj Low}}$	17.19	Average Number of Pheromone to create at Low Boundary
	d_{inj}	Random from Eq 2.6	List of injection distances: random numbers from Eq 2.6

2.2 Binding Reactions

In most particle-based stochastic reaction-diffusion simulations, a binding event is executed as follows. When a ligand molecule is close to (within the binding radius of) an unbound receptor molecule, the ligand molecule is removed from the system, and the receptor molecule is switched to the ‘bound’ state. To match the macroscopic binding kinetics, the

binding radius is calculated from the binding rate and the diffusion constants of the two molecular species:

$$r_{\text{bind}} = \frac{k_{\text{on}}}{4\pi(D_{\alpha} + D_{\text{Ste2}})} \quad (2.1)$$

This expression can be derived from Fick's Law of diffusion and works well when the reaction is diffusion limited; however, α -factor binding to Ste2 is not diffusion limited. Using the rates reported in the literature, this approach requires a binding radius on the order of Angstroms, which is much smaller than the size of Ste2 (GPCRs protrude about 4nm outside the cell membrane [8]). As discussed by Erban and Chapman the unrealistic binding radius results from the assumption that the binding probability is 100% [9]. That is, a ligand molecule within the binding radius of an unbound receptor binds with certainty. The model put forward by Erban and Chapman, removes this assumption and establishes a mathematical framework, in which the binding probability is a function of the binding radius [9,10]. That is, a ligand molecule within a specified binding radius binds with a probability that produces an average binding rate consistent with the macroscopic rate constant k_{on} . We choose the binding radius to be 4nm and calculate the binding probability by numerically solving the system:

$$\frac{k_{\text{on}}\Delta t}{r_{\text{bind}}^3} = P_{\text{bind}} \int_0^1 4\pi z^2 g(z) dz \quad (2.2)$$

where,

$$\begin{aligned} g(\hat{r}) = & (1 - P_{\text{bind}}) \int_0^1 K(\hat{r}, \hat{r}', \gamma) g(\hat{r}') d\hat{r}' \\ & + \int_1^{\infty} K(\hat{r}, \hat{r}', \gamma) g(\hat{r}') d\hat{r}' \\ & + \frac{P_{\text{bind}} K(\hat{r}, \alpha, \gamma)}{\alpha^2} \int_0^1 g(z) z^2 dz \end{aligned}$$

$$K(z, z', \gamma) = \frac{z'}{z\gamma\sqrt{2\pi}} \left(\exp \left[-\frac{(z - z')^2}{2\gamma^2} \right] - \exp \left[-\frac{(z + z')^2}{2\gamma^2} \right] \right)$$

$$\alpha = \frac{r_{\text{unbind}}}{r_{\text{bind}}}$$

$$\gamma = \frac{\sqrt{2(D_\alpha + D_{\text{Ste2}})\Delta t}}{r_{\text{bind}}}$$

derived by Erban and Chapman [10]. For example, using the values for Δt , k_{on} , D_α , D_{Ste2} , r_{bind} and r_{unbind} shown in Table 2.1, a pheromone molecule has a 0.2% chance of binding. We provide a detailed description of how we calculate the probability in Section 2.9. Our customized binding radius and binding probability make our method more accurate and physically realistic than other methods.

2.3 Unbinding Reactions

A ‘bound’ receptor molecule unbinds its ligand molecule in the following way. A new ligand molecule is created a fixed distance away from and in a random direction from the receptor (Fig. 2.1A). The fixed distance is called the unbinding radius (r_{unbind}). We avoid creating the newly released pheromone molecule inside the cell. Lastly, the receptor molecule is switched to the ‘unbound’ state. Given an experimentally measured unbinding rate, we can calculate the unbinding probability, that is, the probability each ‘bound’ receptor has to unbind, using:

$$P_{\text{unbind}} = 1 - \exp[-k_{\text{off}} \Delta t] \quad (2.3)$$

As with the binding radius, we also set the unbinding radius to be 4nm.

2.4 Diffusion of Pheromone

Each pheromone molecule diffuses freely in our volume of interest. To diffuse a pheromone molecule in 3D, let $(x(t), y(t), z(t))$ be the position at time t , then the new position $(x(t + \Delta t), y(t + \Delta t), z(t + \Delta t))$ is found from the following equations:

$$x(t + \Delta t) = x(t) + W_1\sqrt{2D\Delta t} \quad (2.4a)$$

$$y(t + \Delta t) = y(t) + W_2\sqrt{2D\Delta t} \quad (2.4b)$$

$$z(t + \Delta t) = z(t) + W_3\sqrt{2D\Delta t} \quad (2.4c)$$

The factors W_1 , W_2 and W_3 are each random numbers drawn from a Gaussian Distribution with a mean of 0 and a variance of 1. The new position is modified if it is located outside the simulation volume or inside the cell. Reflecting boundary conditions are imposed at four of the boundaries: $y = \pm 5\mu\text{m}$ and $z = \pm 5\mu\text{m}$. Any pheromone molecule that diffuses outside of these boundaries ($y < -5\mu\text{m}$, or $y > 5\mu\text{m}$, or $z < -5\mu\text{m}$, or $z > 5\mu\text{m}$) is reflected back into the volume (Fig. 2.1B).

Additionally, pheromone molecules reflect off the surface of the cell, because the cell membrane is impermeable to pheromone (Fig. 2.1C). This boundary condition prevents pheromone molecules from being located inside the cell. Details for calculating the reflection off the cell surface are provided in Section 2.10.

The last two boundaries, $x = \pm 5\mu\text{m}$, are each uniquely defined, because we wish to establish a linear pheromone gradient along the x-axis. We describe two different methods for treating the $x = \pm 5\mu\text{m}$ boundaries, each of which can establish a gradient. In method 1, each boundary has a fixed concentration. In method 2, one boundary has a fixed concentration while the other is partially absorbing. The next two sections describe the physical interpretation and algorithmic implementation for each method.

2.4.1 Pheromone Gradient – Method 1

In this method, we model a fixed concentration at each boundary. A gradient is formed when we set the concentration at one end of the computational domain higher than at the other. This method is consistent with the design of many microfluidic chambers used to study gradient sensing [1,11–14]. To maintain a fixed concentration at each boundary, pheromone molecules

are added to and removed from the simulation volume in processes called ‘injection’ and ‘ejection’, respectfully (Fig. 2.1D).

For ejection, we remove all pheromone molecules located outside the boundaries ($x < -5\mu\text{m}$, or $x > 5\mu\text{m}$) (Fig. 2.1D). For injection, we create a number of new pheromone molecules and position them near either the $x = 5\mu\text{m}$ or $x = -5\mu\text{m}$ boundary. On average, the number to inject at each time step is calculated using the equation:

$$n_{\text{inj}} = \frac{0.6022}{nM \cdot \mu\text{m}^3} \cdot c \cdot a \sqrt{\frac{D_\alpha \Delta\tau}{\pi}} \quad (2.5)$$

where c is the desired pheromone concentration at the boundary; a is the area of the boundary ($100\mu\text{m}^2$ for most of our simulations); D_α is the diffusion constant for pheromone molecules, and $\Delta\tau$ is the elapsed time between two injection processes. The derivation of Eq 2.5 is found in Section 2.11. Although Eq 2.5 provides the average number of molecules to be injected, due to the stochastic nature of diffusion, the actual number injected can vary for a given time step. During an injection step, the number to inject is a random number drawn from a Poisson distribution with a mean of n_{inj} . The position of a newly injected molecule is also determined randomly. The ‘y’ and ‘z’ positions are determined from a uniform probability distribution across their respective domains (e.g between $-5\mu\text{m}$ and $5\mu\text{m}$, inclusively). The ‘x’ position is calculated as a random distance, called the ‘injection distance’, into the simulation volume from the boundary (Fig. 2.1D). The probability distribution function for the injection distance, d_{inj} , is given by:

$$P(d_{\text{inj}}) = \frac{1}{2} \left[1 - \text{erf} \left(\frac{d_{\text{inj}}}{\sqrt{4D_\alpha \Delta\tau}} \right) \right] \quad (2.6)$$

The derivation of Eq 2.6 and further discussion for implementing Eq 2.6 are provided in Section 2.11. Because it is computationally difficult to generate a random number from the distribution

given by Eq 2.6, we select a random value from a pre-calculated long list. This list has more than 12 million random values whose distribution matches Eq 2.6.

For computational efficiency, injection and ejection of particles are implemented on a slightly coarser time scale, $\Delta\tau$, than the time scale for diffusion Δt . It is important to note that ejection and injection must be calculated on the same time scale. Details and justification for the two time scales are discussed below in Section 2.6: “Algorithm Overview”.

2.4.2 Pheromone Gradient – Method 2

In this method, we model a fixed concentration at one boundary ($x = 5\mu\text{m}$), while the other boundary ($x = -5\mu\text{m}$) is partially absorbing. We use this method for simulations in which pheromone molecules flow toward the cell from one direction ($+\hat{x}$) (Section 3.5).

At the $x = 5\mu\text{m}$ boundary, ejection is the same as method 1. The average number of molecules to inject, n_{inj} , at each time step is given by:

$$n_{\text{inj}} = \frac{0.6022}{nM \cdot \mu\text{m}^3} \cdot \left(c \cdot a \sqrt{\frac{D_\alpha \Delta\tau}{\pi}} + \frac{1}{2} a \cdot g \cdot D_\alpha \Delta\tau \right) \quad (2.7)$$

The derivation of Eq 2.7 is found in Section 2.11. Note that in addition to defining the desired concentration at the boundary, c , we also define the desired gradient at the boundary: g . Eq 2.5 is a special case of Eq 2.7, in which there is no gradient ($g = 0 \text{ nM}/\mu\text{m}$) outside our volume ($x > 5\mu\text{m}$).

At the $x = -5\mu\text{m}$ boundary, no new molecules are injected, and during ejection, not all molecules located outside the boundary ($x < -5\mu\text{m}$) are removed. Instead, each pheromone molecule has a probability of being reflected back inside the volume; otherwise, the molecule is removed. To achieve a steady state gradient of g , the probability of reflection is given by:

$$P_{\text{Ref}} = 1 - \frac{g}{c \sqrt{\frac{1}{\pi D_{\alpha} \Delta \tau}} + \frac{1}{2} g} \quad (2.8)$$

if and only if,

$$\frac{c}{g} \gg \sqrt{4 D_{\alpha} \Delta \tau}$$

The derivation of Eq 2.8 is found in Section 2.11. $\Delta \tau$ is the elapsed time between two ejection processes. The concentration, c , and gradient, g , are the steady state concentration and gradient at the $x = -5\mu\text{m}$ boundary when no cell is present in the computational domain. Because the pheromone molecules coming from the opposite boundary must diffuse around the cell, the resulting concentration will be less than c , and the gradient will be steeper than g .

As in method 1, the injection and ejection processes are implemented on a slightly coarser time scale, $\Delta \tau$, than the primary time scale: Δt . Because these processes are calculated less frequently, our program is more computationally efficient.

2.5 Diffusion of Receptors

Each receptor molecule, ‘bound’ or ‘unbound’, diffuses on the cell membrane (Fig. 2.1C). We approximate diffusion on this surface by first diffusing the receptor in 3-dimensions and then projecting the receptor back onto the surface of the cell. This approach is computationally efficient and accurate for small time steps. The details of these two steps for diffusing a receptor molecule are as follows. First, a new position is calculated using Eqs 2.4 and a diffusion constant appropriate for proteins in the plasma membrane ($D = 0.0025 \mu\text{m}^2/\text{s}$) [2]. Let this new position be $(\tilde{x}, \tilde{y}, \tilde{z})$. Second, we project $(\tilde{x}, \tilde{y}, \tilde{z})$ onto the surface of the cell, which is modeled as a sphere of radius R and centered at the origin $(0, 0, 0)$ using the equations:

$$\tilde{r} = \sqrt{\tilde{x}^2 + \tilde{y}^2 + \tilde{z}^2} \quad (2.9a)$$

$$x(t + \Delta\tau) = \tilde{x} \frac{R}{\tilde{r}} \quad (2.9b)$$

$$y(t + \Delta\tau) = \tilde{y} \frac{R}{\tilde{r}} \quad (2.9c)$$

$$z(t + \Delta\tau) = \tilde{z} \frac{R}{\tilde{r}} \quad (2.9d)$$

The derivation of Eq 2.9 is provided in Section 2.12. For computational efficiency, we diffuse the receptors on a slightly coarser time scale, $\Delta\tau$, than the primary time scale: Δt . We do not sacrifice much accuracy, because the diffusion constant for membrane-bound receptors is small compared to that of extracellular pheromone molecules.

2.6 Algorithm Overview

Our simulation algorithm and order of operations closely follows the general algorithm described by [10]. We modify their algorithm, because of the spatial domains (inside, on or outside the cell) and non-uniform distribution of molecules (the ligands have a linear concentration gradient and the receptors are restricted to the surface of the sphere). Our simulation algorithm is broken into the six processes described above: binding reactions, unbinding reactions, diffusion of pheromone molecules, ejection of pheromone molecules, injection of pheromone molecules and diffusion of receptors. The last three processes are simulated over a more coarse time scale (time step = $\Delta\tau$) than the first three processes (time step = Δt). Below, we outline the pseudo-code of our simulation algorithm.

At each time step

- I. **Binding Reactions** – Each ‘unbound’ receptor has a chance to bind a single, nearby pheromone molecules.
- II. **Unbinding Reactions** – Each ‘bound’ receptor, including those from step 1), has a chance to release its pheromone molecule.
- III. **Pheromone Diffusion** – Each pheromone molecule diffuses. They reflect off the cell’s surface, the $y=\pm 5\mu\text{m}$ and $z=\pm 5\mu\text{m}$ boundaries.

After every ($\Delta\tau/\Delta t$) steps

- IV. **Receptor Diffusion** – Each receptor molecule diffuses on the surface of the cell.
- V. **Ejection** – Remove all pheromone molecules outside the $x=\pm 5\mu\text{m}$ boundaries.

VI. Injection – Add new pheromone molecules near the $x=\pm 5\mu\text{m}$ boundaries.

We choose the coarse time scale, $\Delta\tau$, based on two criteria. One, receptor diffusion should not be many times larger than the binding radius. And two, the average injection distance should be much less than $2.5\mu\text{m}$, which is the distance from the $x = \pm 5\mu\text{m}$ boundaries to the cell. For $\Delta\tau = 50\mu\text{s}$, receptors diffuse about 0.5nm , which is much smaller than 4nm . Also for $\Delta\tau = 50\mu\text{s}$, the average injection distance is about $0.05\mu\text{m}$, which is much smaller than $2.5\mu\text{m}$.

This model, which we derived from physical processes, falls under the class of computational models known as Agent-Based Models. In terms of this class, we simulate many molecules (the “agents”), each of which follows a set of rules. Because many of these rules are independent of other molecules, we can parallelize the algorithm at each process. For example, during pheromone diffusion, process **III**, a new position is calculated for each pheromone molecule. This calculation is independent from all other molecules. Hence, the diffusion of many pheromone molecules can be calculated simultaneously. Ideally, we would calculate the new position of every pheromone molecule in parallel. To achieve massive parallelization with minimal coding effort, we turn to Hardware Acceleration using NVIDIA GPUs. We write the program in CUDA C, which is an extension of the C Programming Language, developed by NVIDIA to facilitate High Performance Computing on their GPGPUs. Each of the six processes is executed on the GPU, one at a time, as arranged above, in order to ensure all reactions are complete before the molecules diffuse. That is to say, there is a global synchronization between processes.

2.7 Receptor Cycling Model

To determine if receptor cycling can reduce noise, we have developed a simplified model of receptor cycling. We compare simulations of the basic model described above, in which

receptors bind and unbind pheromone, to simulations of the simplified receptor cycling model, in which receptors bind pheromone and are endocytosed. Specifically, an active (pheromone-bound) receptor, Ste2*, can be endocytosed and replaced with an unbound receptor, Ste2. The new Ste2 molecule is added to a random position on the cell surface. This method of endocytosis with immediate replacement keeps the total number of receptors constant, and allows us to directly compare results from this model to results from the basic model. Although the endocytosis rate for Ste2* is 0.0021 s^{-1} [15,16], we use a rate of 0.0011 s^{-1} , which is the unbinding reaction rate from the basic model (See Table 1). In our algorithm, process **II** (Unbinding Reactions) is replaced with the following.

- II. Endocytosis Reactions** – Each ‘bound’ receptor, including those from process **I**, has a chance to be endocytosed.

2.8 Bar1 Model

To determine if the pheromone protease Bar1 can reduce noise and improve gradient sensing, we have also developed a simplified model of the cell releasing Bar1. In addition to the basic model described above, in which receptors bind and unbind pheromone, we include the reaction of Bar1 degrading pheromone. In principle this catalytic reaction can be modeled much like the binding reaction; that is, individual Bar1 molecules could be simulated and have a probability of degrading nearby pheromone molecules. However we choose to avoid the computational cost of this method. Instead, we model Bar1 concentration as a static, radial field extending from the surface of the cell:

$$[\text{Bar1}](r) = [\text{Bar1}]_0 \frac{R}{r} \quad (2.10)$$

Where r is the distance from the center of the cell; R is the radius of the cell, and $[\text{Bar1}]_0$ is the Bar1 concentration at the surface of the cell. Similar to previous modeling work, we set $[\text{Bar1}]_0 =$

0.85nM [1]. Based on their distance from the cell surface, pheromone molecules have a probability of being degraded as given by:

$$P_{\text{cat}}(r) \approx k_{\text{cat}} \cdot [\text{Bar1}](r) \cdot \Delta t \quad (2.11)$$

As modeled previously, we set the catalytic reaction rate, k_{cat} , to be $2.5 \times 10^8 \text{ (M} \cdot \text{s)}^{-1}$ [1]. In our algorithm, the process for Bar1-mediated catalytic reaction is inserted between process **II** (Unbinding Reactions) and **III** (Pheromone Diffusion). Here, we label the Bar1 process as **IIB**.

IIB. Catalytic Reactions – Each pheromone molecule has a chance, based on its current position, of being degraded.

2.9 Calculating Binding Probability

To simulate the second-order reaction of a pheromone molecule binding to a Ste2 receptor, we implement the $\lambda - \bar{q}$ method developed by Erban and Chapman [9,10]. The equations presented here, and the technique used to numerically solve them, are taken directly from their paper. We deviate slightly from their technique, in that we double the binding probability calculated from their method (see end of this section).

We use the biophysical parameters defined in Table 2.1. In particular, we use the binding rate (k_{on}), unbinding rate (k_{off}) and diffusion constants for pheromone (D_{α}) and the receptor (D_{Ste2}). Furthermore, we choose a time step (Δt), binding radius (r_{bind}) and unbinding radius (r_{unbind}). The binding and unbinding parameters are chosen based on the size of the receptor. The time step is chosen such that the combined diffusion step size is not much greater than the binding radius. That is, the dimensionless parameter γ (Eq 2.12) is not much greater than 1. In our simulations, $\gamma \approx 4$.

$$\gamma = \frac{\sqrt{2(D_{\alpha} + D_{\text{Ste2}})\Delta t}}{r_{\text{bind}}} \quad (2.12)$$

Ideally, Δt would be much smaller, so that, $\gamma \ll 1$. However, this time step is too small for practical computation, especially because we aim to simulate the system for one hour. Hence, because $\gamma > 1$, our chosen time step is considered “large”. Nonetheless, according to the $\lambda - \bar{q}$ method, we can calculate the binding probability (P_{bind}) based on the parameters listed above.

The dimensionless reaction rate, κ (Eq 2.13), is related to the binding probability by Eq 2.14, which describes the rate of removing particles in one time step.

$$\kappa = \frac{k_{\text{on}}\Delta t}{r_{\text{bind}}^3} \quad (2.13)$$

$$\kappa - P_{\text{bind}} \int_0^1 4\pi z^2 g(z) dz = 0 \quad (2.14)$$

The function $g(\hat{r})$ is defined below (Eq 2.15).

$$\begin{aligned} g(\hat{r}) = & (1 - P_{\text{bind}}) \int_0^1 K(\hat{r}, \hat{r}', \gamma) g(\hat{r}') d\hat{r}' \\ & + \int_1^\infty K(\hat{r}, \hat{r}', \gamma) g(\hat{r}') d\hat{r}' + \frac{P_{\text{bind}} K(\hat{r}, \alpha, \gamma)}{\alpha^2} \int_0^1 g(z) z^2 dz \end{aligned} \quad (2.15)$$

The parameter α is the dimensionless ratio of the binding and unbinding radii (Eq 2.16). The

Kernel $K(\hat{r}, \hat{r}')$ is derived from Green’s function for diffusion (Eq 2.17).

$$\alpha = \frac{r_{\text{unbind}}}{r_{\text{bind}}} \quad (2.16)$$

$$K(z, z', \gamma) = \frac{z'}{z\gamma\sqrt{2\pi}} \left(\exp \left[-\frac{(z - z')^2}{2\gamma^2} \right] - \exp \left[-\frac{(z + z')^2}{2\gamma^2} \right] \right) \quad (2.17)$$

To determine P_{bind} numerically, we use the Secant Root-Finding Method to solve Eq 2.14, which also requires numerically evaluating Eq 2.15. We truncate the semi-infinite integral (second term in Eq 2.15) and approximate $g(\hat{r}) = 1$ for $\hat{r} \geq S$ (Eq 2.18). We choose $S = 5\gamma \approx 20$.

$$\begin{aligned}
g(\hat{r}) = & (1 - P_\lambda) \int_0^1 K(\hat{r}, \hat{r}') g(\hat{r}') d\hat{r}' + \int_1^S K(\hat{r}, \hat{r}') g(\hat{r}') d\hat{r}' \\
& + \int_S^\infty K(\hat{r}, \hat{r}') d\hat{r}' + \frac{P_\lambda K(\hat{r}, \alpha)}{\alpha^2} \int_0^1 g(\hat{r}') \hat{r}'^2 d\hat{r}'
\end{aligned} \tag{2.18}$$

After rearranging Eq 2.18, we get the integral equation:

$$\begin{aligned}
- \int_S^\infty K(\hat{r}, \hat{r}') d\hat{r}' = & (1 - P_{\text{bind}}) \int_0^1 K(\hat{r}, \hat{r}') g(\hat{r}') d\hat{r}' \\
& + \frac{P_\lambda}{\alpha^2} K(\hat{r}, \alpha) \int_0^1 g(\hat{r}') \hat{r}'^2 d\hat{r}' + \int_1^S K(\hat{r}, \hat{r}') g(\hat{r}') d\hat{r}' - g(\hat{r})
\end{aligned} \tag{2.19}$$

The Left-Hand Side of Eq 2.19 can be solved analytically to yield the following function of \hat{r} :

$$\begin{aligned}
F(\hat{r}) = & - \int_S^\infty K(\hat{r}, \hat{r}') d\hat{r}' \\
= & \frac{1}{2} \left[\text{erf} \left(\frac{S - \hat{r}}{\gamma \sqrt{2}} \right) + \text{erf} \left(\frac{S + \hat{r}}{\gamma \sqrt{2}} \right) \right] - 1 - \frac{\gamma^2}{S} K(\hat{r}, S)
\end{aligned} \tag{2.20}$$

For the Right-Hand Side of Eq 2.19, the integrals are numerically evaluated using the quadrature method. That is, the integral is approximated with a polynomial, which can be written as a weighted sum (Eq 2.22). We discretize the domains $\hat{r} \in [0,1]$ and $\hat{r}' \in [0,1]$ at the Chebyshev nodes with $n = 20$ (Eq 2.21). By fitting the integrals at the Chebyshev nodes, we minimize the numerical error near the bounds of the integral.

$$\hat{r}_i = \frac{1}{2} + \frac{1}{2} \cos \left(\frac{2i - 1}{2n} \pi \right) \tag{2.21}$$

$$\int_0^1 f(\hat{r}') d\hat{r}' = \sum_{i=0}^{n-1} a_i f(\hat{r}'_i) \tag{2.22}$$

The coefficients, \vec{a} , are calculated from fitting the integral with the n^{th} degree polynomial:

$$\mathbf{A} * \vec{a} = \vec{h} \quad (2.23)$$

where,

$$\mathbf{A}_{i,j} = (\hat{r}'_j)^i$$

$$h_i = \int_0^1 (x)^i = \frac{1}{i+1}$$

for

$$i, j = 0 \dots n-1$$

Eqs 2.22 & 2.23 solve the first two terms of Eq 2.19. To solve the last integral term (Eq 2.25), the domains $\hat{r} \in [1, S]$ and $\hat{r}' \in [1, S]$ are discretized at the Chebyshev nodes with $m = 13$ (Eq 2.24).

$$\hat{r}_j = \frac{1+S}{2} + \frac{S-1}{2} \cos\left(\frac{2j-1}{2m}\pi\right) \quad (2.24)$$

$$\int_1^S f(\hat{r}') = \sum_{j=0}^{m-1} b_j f(\hat{r}'_j) \quad (2.25)$$

The coefficients, \vec{b} , are calculated by fitting the integral with the m^{th} degree polynomial:

$$\mathbf{B} * \vec{b} = \vec{h} \quad (2.26)$$

where,

$$\mathbf{B}_{i,j} = (\hat{r}'_j)^i$$

$$h_i = \int_1^S (x)^i = \frac{(S)^{i+1} - 1}{i+1}$$

for

$$i, j = 0 \dots m-1$$

Using Eqs 2.22, 2.23, 2.25 & 2.26, the Right-Hand Side of Eq 2.19 is reduced to a function of \hat{r} (Eq 2.27). Because \hat{r} is discretized the same as \hat{r}' , Eq 2.27 produces a system of equations, which can be written in matrix form (Eq 2.28).

$$\begin{aligned}
& (1 - P_\lambda) \int_0^1 K(\hat{r}, \hat{r}') g(\hat{r}') d\hat{r}' + \frac{P_\lambda}{\alpha^2} K(\hat{r}, \alpha) \int_0^1 g(\hat{r}') \hat{r}'^2 d\hat{r}' \\
& + \int_1^S K(\hat{r}, \hat{r}') g(\hat{r}') d\hat{r}' - g(\hat{r}) \\
& = (1 - P_\lambda) \sum_{i=0}^{n-1} a_i K(\hat{r}, \hat{r}'_i) g(\hat{r}'_i) + \frac{P_\lambda}{\alpha^2} K(\hat{r}, \alpha) \sum_{i=0}^{n-1} a_i g(\hat{r}'_i) \hat{r}'_i^2 \\
& + \sum_{j=0}^{m-1} b_j K(\hat{r}, \hat{r}'_j) g(\hat{r}'_j) - g(\hat{r})
\end{aligned} \tag{2.27}$$

$$\vec{F} = \mathbf{M} \vec{g} \tag{2.28}$$

\vec{F} and \vec{g} are, respectively, $(n + m) \times 1$ column vectors of $F(\hat{r})$ (Eq 2.20) and $g(\hat{r})$ evaluated at the discrete \hat{r} points (Eqs 2.21 & 2.24). \mathbf{M} is a $(n + m) \times (n + m)$ square matrix, built from factoring out $g(\hat{r})$ and $g(\hat{r}')$ from Eq 2.27.

Hence the technique is to first guess a value for P_{bind} . Based on this guess, we approximate $g(\hat{r})$ by solving Eq 2.28. This approximation is sufficient to evaluate Eq 2.14 and determine the accuracy of the guess. If Eq 2.14 is incorrect, then we use the Secant Method to update the guess for P_{bind} . This process is repeated until the error falls below a threshold (10^{-15}). We find $P_{\text{bind}} \approx 0.001$.

One important caveat for our system is to double the final value of P_{bind} as obtained using the above technique. The above equations were derived for a system in which both reactant species are in solution. Accordingly, ligand molecules can approach a receptor from any direction. However, in our system, ligand molecules can only approach a receptor from half as many directions (Fig. 2.2). We double the probability and use $P_{\text{bind}} = 0.002$ in our simulations.

In the language of the $\lambda - \bar{\rho}$ method, we adjust Eq 2.14, which integrates over spherical surface areas, to integrate over hemispherical surface areas:

$$\kappa - P_{\text{bind}} \int_0^1 2\pi z^2 g(z) dz = 0 \quad (2.29)$$

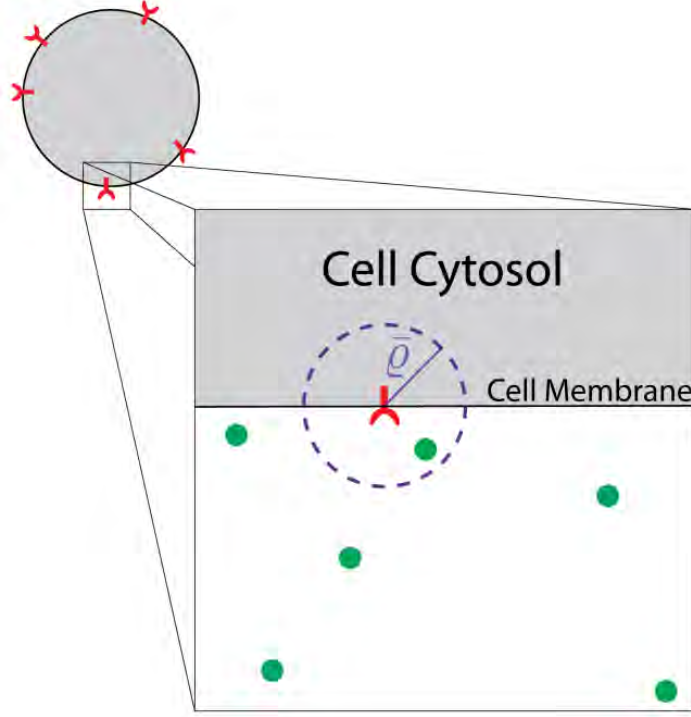


Fig. 2.2: Rationale for Doubling P_{bind}

In the upper left corner, we show a 2D cartoon of the cell in our simulation: a circle with unbound Ste2 receptors on the surface. By focusing closely on a single receptor, we see how the scale of the binding radius, r_{bind} , compares to the scale of the cell. That is, the membrane appears as a straight line, because the cell's radius ($2.5\mu\text{m}$) is three orders of magnitude larger than the binding radius (4nm). The membrane divides the binding volume (purple circle) exactly in half. Pheromone molecules can only enter this volume from one half.

2.10 Calculating Pheromone Reflection off the Cell Surface

The cell membrane is impermeable to pheromone. Our model prevents pheromone from diffusing into the cell, by reflecting molecules off the surface of the cell. In our simulations, the cell is a sphere with a radius of R and center at $(0, 0, 0)$. Figure 2.3 shows how we calculate these reflections.

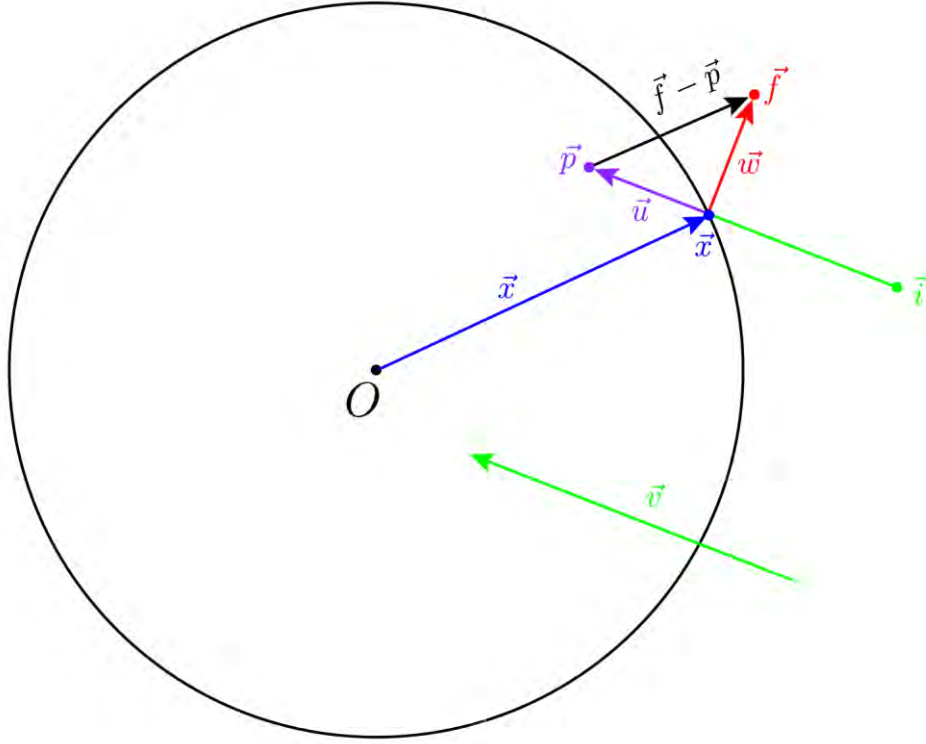


Fig. 2.3: Reflecting Pheromone off Cell Surface

A diagram of important vectors used to calculate how a pheromone molecule reflects off the surface of the cell during a single diffusion step. The molecule is initially at \vec{i} , and after freely diffusing for one step, is positioned inside the cell at \vec{p} . The desired, reflected position is \vec{f} . Although the simulation is in 3D and the cell is modeled as a sphere, all of these points and vectors lie on a single plane (defined by the three points i, p and O). By using vector algebra, we may solve the system on a plane and the resulting equations will be valid in 3D.

Let \vec{i} be the initial position of the pheromone molecule before diffusion. Let \vec{p} be the “proposed” position after diffusion, which is undesirably located inside the cell. The vector \vec{v} is the attempted trajectory of the pheromone molecule for a single time step. That is, $\vec{v} = \vec{p} - \vec{i}$. If the molecule reflected off the cell surface, then its final position would be \vec{f} . Hence, we wish to find a set of equations to calculate \vec{f} , given \vec{i} , \vec{p} and the dimension of the sphere.

First, we calculate the position where the molecule contacts the cell surface: \vec{x} . The attempted trajectory, \vec{v} , defines a line which can be written as $\vec{i} + t\vec{v}$. The intersection, \vec{x} , lies on this line:

$$\vec{x} = \vec{l} + t_x \vec{v} \quad (2.30)$$

Knowing that $\vec{x} \cdot \vec{x} = R^2$, we can solve for t_x :

$$(\vec{l} + t_x \vec{v}) \cdot (\vec{l} + t_x \vec{v}) = R^2 \quad (2.31)$$

$$t_x = \frac{-\vec{v} \cdot \vec{l} - \sqrt{\mathbf{Det}}}{\|\vec{v}\|^2} \quad (2.32)$$

where,

$$\mathbf{Det} = (\vec{v} \cdot \vec{l})^2 + (\|\vec{v}\|^2)(R^2 - \|\vec{l}\|^2)$$

Second, we calculate the vector \vec{u} (Eq 2.33). This vector is the portion of the molecule's trajectory that incorrectly traveled into the cell. Third, to correct the trajectory, we calculate \vec{w} , which is equivalent to \vec{u} except that they have opposite radial components. The direction pointing away from the cell is calculable from the intersection point: $\hat{x} = \frac{\vec{x}}{\|\vec{x}\|} = \frac{\vec{x}}{R}$. Eq 2.34 shows that \vec{w} and \vec{u} differ by twice their radial components.

$$\vec{u} = \vec{p} - \vec{x} \quad (2.33)$$

$$\vec{w} - \vec{u} = 2(-(\vec{u} \cdot \hat{x}))\hat{x}$$

$$\vec{w} - \vec{u} = -\left(\frac{2}{R^2}\right)(\vec{u} \cdot \vec{x})(\vec{x}) \quad (2.34)$$

Finally, because $\vec{w} - \vec{u} = \vec{f} - \vec{p}$, we can calculate \vec{f} :

$$\vec{f} = \vec{p} - \left(\frac{2}{R^2}\right)(\vec{u} \cdot \vec{x})(\vec{x}) \quad (2.35)$$

Numerically, we calculate \vec{x} from Eqs 2.30 & 2.32 using known values for \vec{l} , \vec{p} and R . We then calculate \vec{u} from Eq 2.33. Lastly, we find the final reflected position, \vec{f} , by solving Eq 2.35.

2.11 Injection Boundary Condition: Derivation of Eqs 2.5 – 2.8

We simulate a linear pheromone gradient along the x-axis, which requires special boundary conditions at the x-boundaries defined by the planes: $x = \pm \frac{L}{2}$. We must define how individual molecules behave at these boundaries.

2.11.1 Gradient Method 1

One method to create a pheromone gradient is to fix the pheromone concentration at each x-boundary. The fixed concentrations are molecular reservoirs in the semi-infinite domains ($x < -\frac{L}{2}$ and $x > \frac{L}{2}$) that extend past each planar boundary. A molecular reservoir, analogous to a thermal reservoir, is a large enough volume such that adding or removing molecules does not appreciably change the concentration. Molecules are exchanged between the simulation volume and the reservoir in processes called ‘injection’ (entering the simulation volume) and ‘ejection’ (exiting the simulation volume). These two processes are independent. For ejection, the molecules which leave the simulation volume are considered part of the reservoir and are no longer simulated. For injection, new molecules diffuse into the simulation volume. Here, we derive how many molecules to inject and where to inject them.

For simplicity, we shift the simulation domain such that $x \in [0, L]$ and consider the $x=0$ boundary. Given the molecules’ diffusion constant, D , the concentration of the reservoir, c , the time step Δt , and the area of the boundary, a , we derive how many molecules will enter the simulation volume from the reservoir. The probability for a molecule to diffuse a distance \tilde{x} is:

$$P(\tilde{x}) = \frac{1}{\sqrt{4\pi D\Delta t}} \exp\left[-\frac{\tilde{x}^2}{4D\Delta t}\right] \quad (2.36)$$

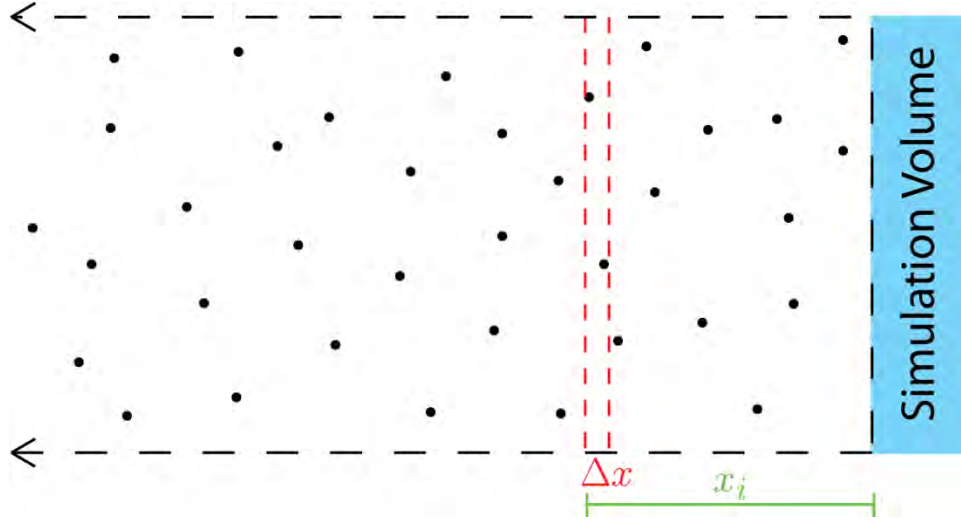


Fig. 2.4: Molecular Reservoir

A cartoon of the molecular reservoir outside one boundary. The simulation volume is colored in blue and continues finitely off the page. The molecular reservoir has a white background and extends to $-\infty$. The border is indicated by the black, vertical dashed line. Also shown is a slice of the reservoir, with thickness Δx , at a distance x_i from the boundary.

Consider a thin slice of the semi-infinite volume, a distance x_i from the boundary (Fig. 2.4). The probability for a molecule in that slice to be injected (diffuse into the simulation volume) is given by Eq 2.37. A molecule must diffuse a distance of at least x_i and less than $L + x_i$.

$$P_{\text{inj}}(x_i) = \int_{x_i}^{L+x_i} P(\tilde{x}) d\tilde{x} = \frac{1}{2} \left[1 - \text{erf} \left(\frac{x_i}{\sqrt{4D\Delta t}} \right) \right] \quad (2.37)$$

The integral's simplification takes advantage of $\text{erf} \left(\frac{L+x_i}{\sqrt{4D\Delta t}} \right) = 1$, because $L \gg \sqrt{4D\Delta t}$ for our values (Table 2.1). Note that for the slice located at the boundary, $x_0 = 0$, $P_{\text{inj}} = 1/2$ as expected; a molecule at $x = 0$ has an equal chance of diffusing right (into the volume) or left (remaining in the reservoir). The entire reservoir can be divided into adjacent, thin slices. For the i^{th} slice, the number of molecules injected is given by $N \cdot P_{\text{inj}}(x_i)$, where N is the number of molecules in this slice. Because the reservoir has a fixed concentration of molecules, $N = c \cdot a \cdot \Delta x$. We sum the contribution from each slice and get a total number of injected molecules:

$$n_{\text{inj}} = c \cdot a \sum_{i=0}^{\infty} P_{\text{inj}}(x_i) \Delta x \quad (2.38)$$

We make the slices infinitely small ($\Delta x \rightarrow 0$), such that Eq 2.38 becomes a solvable integral:

$$n_{\text{inj}} = c \cdot a \int_0^{\infty} P_{\text{inj}}(x) dx = c \cdot a \sqrt{\frac{D\Delta t}{\pi}} \quad (2.39)$$

2.11.2 Injection Distances

We define the injection distance, d_{inj} , as the distance from the boundary that a newly injected molecule is placed. Here, we derive the probability of a molecule being injected a distance d_{inj} . For example, Figure 2.5 shows a molecule can diffuse to the d_{inj} position in many ways. That is, the molecule has many starting positions, from $x = 0$ (at the boundary) to $x \rightarrow -\infty$. Let Δx be the distance the molecule diffuses; therefore, $\Delta x \in [d_{\text{inj}}, \infty)$.

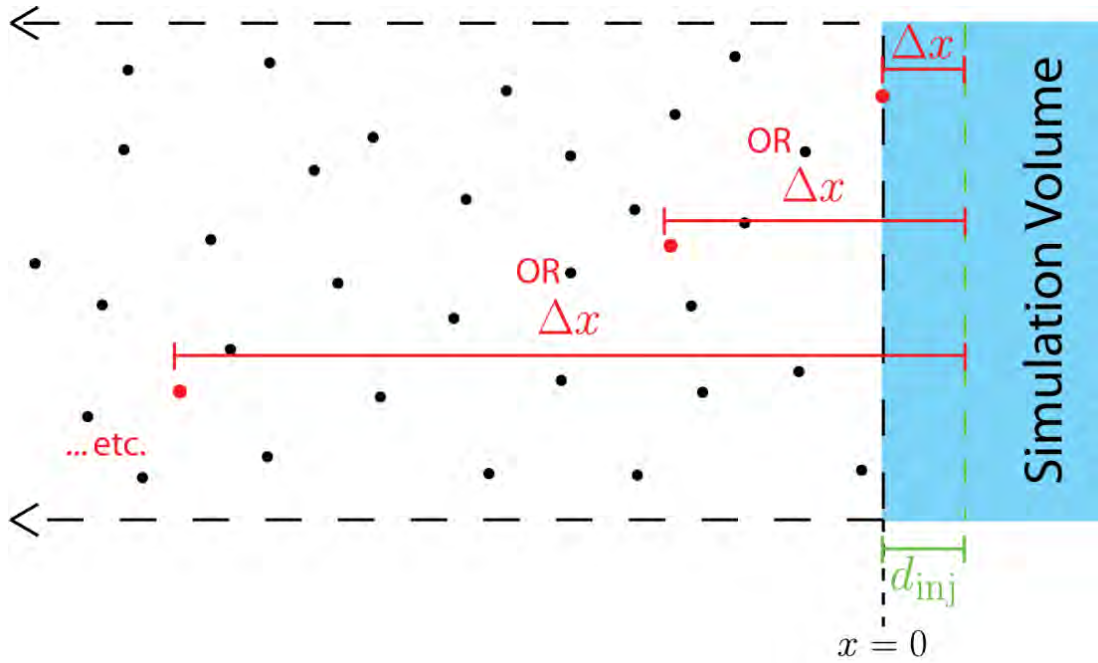


Fig. 2.5: Injection Distance

A cartoon showing the many ways a molecule could diffuse a fixed distance (d_{inj}) into the simulation volume (blue). The red dots show various starting positions; the molecule must travel a distance Δx in order to end at d_{inj} .

We know, as with Eq 2.36, the probability for diffusing a distance Δx is:

$$P(\Delta x) = \frac{1}{\sqrt{4\pi D \Delta t}} \exp \left[\frac{-\Delta x^2}{4D \Delta t} \right] \quad (2.40)$$

The probability of diffusing d_{inj} into the simulation volume is the sum of the probabilities from each potential starting position (Eq 2.41). We then normalize this distribution to get the Probability Density Function (Eq 2.42).

$$P(d_{inj}) = \int_{d_{inj}}^{\infty} P(\Delta x) d(\Delta x) = \frac{1}{2} \left[1 - \operatorname{erf} \left(\frac{d_{inj}}{\sqrt{4D \Delta t}} \right) \right] \quad (2.41)$$

$$PDF(d_{inj}) = \sqrt{\frac{\pi}{4D \Delta t}} \left[1 - \operatorname{erf} \left(\frac{d_{inj}}{\sqrt{4D \Delta t}} \right) \right] \quad (2.42)$$

Note that the distribution depends only on the diffusion constant and the time step. We empirically verify this distribution (Eq 2.42). We simulate diffusion in a cubic volume and record the distance the particles traveled beyond the boundaries in one time step. Figure 2.6 shows the theoretical and empirical distributions. The recorded list has over 2.5 million values, and we draw a d_{inj} for each new molecule during the simulation.

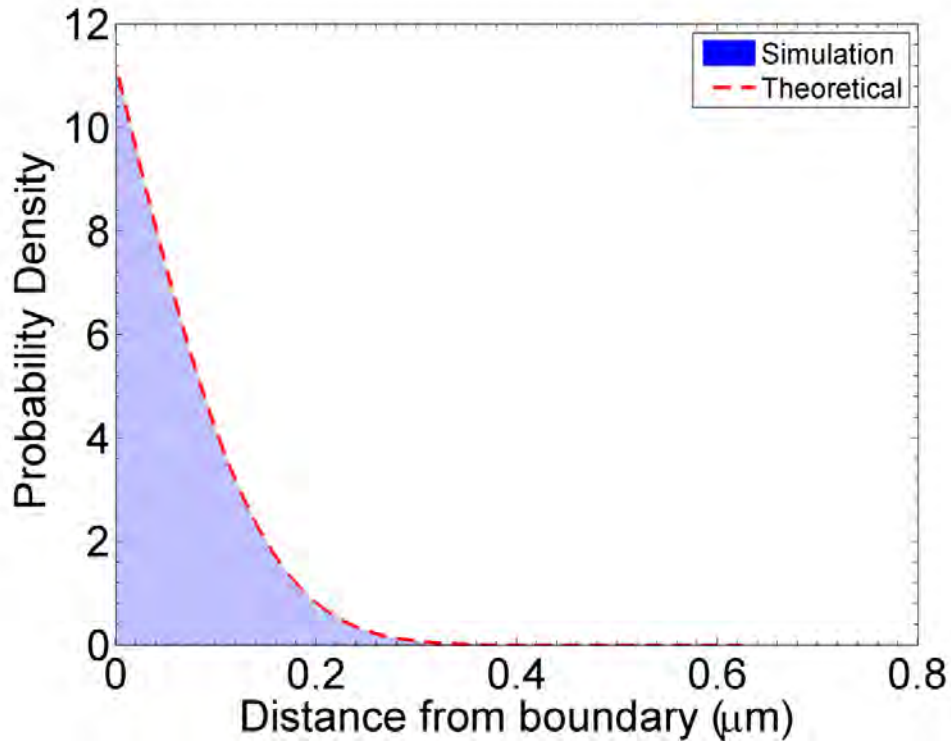


Fig. 2.6: Probability Density of Injection Distance

Here we plot the probability density for an injected molecule to be placed a distance d_{inj} into the simulation volume. The dashed red line is the theoretical, normalized probability density function (Eq 2.42). The blue histogram is from simulated data. In this case, $D = 125 \frac{\mu m^2}{s}$ and $\Delta t = 50 \mu s$.

Ideally, our injection algorithm would generate a random injection distance from this distribution during the simulation (Eq 2.42). However, the corresponding Cumulative Distribution Function:

$$\begin{aligned}
 CDF(d_{inj}) &= \int_0^{d_{inj}} PDF(d_{inj}') d(d_{inj}') \\
 &= \sqrt{\frac{\pi}{4D\Delta t}} d_{inj} \left(1 - \operatorname{erf}\left(\frac{d_{inj}}{\sqrt{4D\Delta t}}\right) \right) + 1 - \exp\left[\frac{-d_{inj}^2}{4D\Delta t}\right] \quad (2.43)
 \end{aligned}$$

cannot be inverted. Instead, we draw randomly from the previously generated list of injection distances. This table-lookup method is more computationally efficient.

2.11.3 Gradient Method 2

In Section 2.11.1, we describe a method to establish a linear gradient by fixing the concentration at the x-boundaries. A second method is to flow molecules from one x-boundary and partially absorb molecules at the opposing boundary. As in Section 2.11.1, we shift the simulation domain such that $x \in [0, L]$. New molecules are added to the system at the $x = L$ boundary, which has a higher concentration than at $x = 0$.

First, we determine the injection rate at the $x = L$ boundary. We define the desired gradient, g , and the concentration at $x = L$ is c_L . Eq 2.44 describes the desired concentration profile at this boundary: $\tilde{x} = x - L$.

$$c(\tilde{x}) = g\tilde{x} + c_L \quad (2.44)$$

The gradient is positive ($g > 0$). We calculate the number to inject by changing $c = c(\tilde{x})$ in Eq 2.38. The resulting discrete sum is converted to an integral and solved:

$$n_{\text{inj}} = a \int_0^\infty c(\tilde{x}) P_{\text{inj}}(\tilde{x}) d\tilde{x} = \frac{1}{2} a \cdot g \cdot D\Delta t + c_L \cdot a \sqrt{\frac{D\Delta t}{\pi}} \quad (2.45)$$

Second, we determine the absorption probability at the $x = 0$ boundary. Eq 2.46 describes the desired concentration profile at this boundary; c_0 is the desired concentration at $x = 0$.

$$c(x) = gx + c_0 \quad (2.46)$$

The absorption probability, P_{Abs} , is the probability a molecule, which has diffused outside the boundary ($x < 0$), is removed from the simulation. If the molecule is not removed, then the molecule is reflected back into the simulation volume. The two quantities are related:

$$P_{\text{Abs}} = 1 - P_{\text{Ref}} \quad (2.47)$$

The reflection probability, P_{Ref} , can be calculated from the expected number of molecules exchanged between the reservoir and simulation volume:

$$P_{\text{Ref}} = \frac{n_{\text{in}}}{n_{\text{out}}} \quad (2.48)$$

The expected number of molecules that exit the simulation volume is n_{out} , given by:

$$n_{\text{out}} = a \int_0^L c(x) P_{\text{inj}}(x) dx = \frac{1}{2} a \cdot g \cdot D \Delta t + c_0 \cdot a \sqrt{\frac{D \Delta t}{\pi}} \quad (2.49)$$

Eq 2.49 is valid for $L \gg \sqrt{4D\Delta t}$, which is the case for our values (Table 2.1). The expected number of molecules that enter the simulation is n_{in} . To calculate n_{in} , a slight change in variables is helpful: $x' = -x$. Accordingly, the desired concentration profile (Eq 2.46) changes:

$$c(x') = -gx' + c_0 \quad (2.50)$$

Because concentration can never be negative ($c(x') \geq 0$), the domain of x' is limited: $x' \leq \frac{c_0}{g}$. The general solution for n_{in} is:

$$\begin{aligned} n_{\text{in}} &= a \int_0^{\frac{c_0}{g}} c(x') P_{\text{inj}}(x') dx' \\ &= \frac{1}{2} a \left[c_0 \sqrt{\frac{D \Delta t}{\pi}} \left(2 - e^{\frac{-c_0^2}{g^2 \cdot 4D \Delta t}} \right) + \frac{1}{2} \frac{c_0^2}{g} - \left(\frac{1}{2} \frac{c_0^2}{g} + g D \Delta t \right) \text{erf} \left(\frac{c_0}{g \sqrt{4D \Delta t}} \right) \right] \end{aligned} \quad (2.51)$$

For our values (Table 2.1), $\frac{c_0}{g} \gg \sqrt{4D\Delta t}$; therefore, Eq 2.51 simplifies:

$$n_{\text{inj}} = c_0 \cdot a \sqrt{\frac{D \Delta t}{\pi}} - \frac{1}{2} a \cdot g \cdot D \Delta t \quad (2.52)$$

Substituting Eqs 2.49 & 2.52 into Eq 2.48, we get the following expression for P_{Ref} :

$$P_{\text{Ref}} = 1 - \frac{g}{c \sqrt{\frac{1}{\pi D \Delta t}} + \frac{1}{2} g} \quad (2.53)$$

2.12 Receptor Diffusion on the Cell Surface: Derivation of Eq 2.9

To calculate receptor diffusion on the surface of the cell, we first diffuse the receptor freely in 3D, and second, we project that new position onto the surface of the cell. Eq 2.9 describes the latter step. Let $(\tilde{x}, \tilde{y}, \tilde{z})$ be the diffused position, which we want to project onto a sphere with a radius of R and center at $(0, 0, 0)$. We can write this position in Spherical coordinates as $(\tilde{r}, \tilde{\theta}, \tilde{\phi})$, where $\tilde{r} = \sqrt{\tilde{x}^2 + \tilde{y}^2 + \tilde{z}^2}$, and $\tilde{\theta}$ & $\tilde{\phi}$ are spherical angles. To convert from Spherical to Cartesian coordinates:

$$\tilde{x} = \tilde{r} \sin \tilde{\theta} \cos \tilde{\phi} \quad (2.54)$$

$$\tilde{y} = \tilde{r} \sin \tilde{\theta} \sin \tilde{\phi} \quad (2.55)$$

$$\tilde{z} = \tilde{r} \cos \tilde{\theta} \quad (2.56)$$

The final projected position, (x, y, z) , has the same spherical angles $\tilde{\theta}$ and $\tilde{\phi}$ but a radius of R .

Because the angles are the same, we can equate the trigonometric factors as follows.

$$\frac{\tilde{x}}{\tilde{r}} = \sin \tilde{\theta} \cos \tilde{\phi} = \frac{x}{R} \quad (2.57)$$

$$\frac{\tilde{y}}{\tilde{r}} = \sin \tilde{\theta} \sin \tilde{\phi} = \frac{y}{R} \quad (2.58)$$

$$\frac{\tilde{z}}{\tilde{r}} = \cos \tilde{\theta} = \frac{z}{R} \quad (2.59)$$

Multiplying Eqs 2.57 – 2.59 by R , yields Eq 2.9.

2.13 Measuring the Pheromone Distribution

To determine the distribution of pheromone molecules, we discretize the simulation volume with thick-walled cylinders (Fig. 2.7). The naïve approach is to discretize the space with small cubes; however, we recognize that our system is symmetric in the ‘y’ and ‘z’ dimensions. Therefore, we reduce these two spatial dimensions into one radial dimension: $\rho = \sqrt{y^2 + z^2}$. We

discretize x-axis using bins of width Δx and discretize the radial dimension using bins of width $\Delta \rho$. Each pheromone molecule is tallied according to its x-position and distance from the x-axis ($\sqrt{y^2 + z^2}$). We determine the average number of pheromone molecules in each bin and subsequently calculate the concentration by dividing by the volume of each cylinder. The pheromone concentration is a function of x and ρ : $c = c(x, \rho)$. Note that for $\rho \leq 2.5 \mu\text{m}$ and $x \in [-2.5, 2.5] \mu\text{m}$, the cylindrical bin intersects with, or lies inside the cell. Because pheromone is excluded from the cell's interior, the effective volume of these bins is reduced. The concentration profiles presented in the main text are $c(x)$ for $\rho \approx 1.01 \mu\text{m}$. Hence, for $-2.28 \mu\text{m} \leq x \leq 2.28 \mu\text{m}$, the expected and measured pheromone concentration is 0 nM.

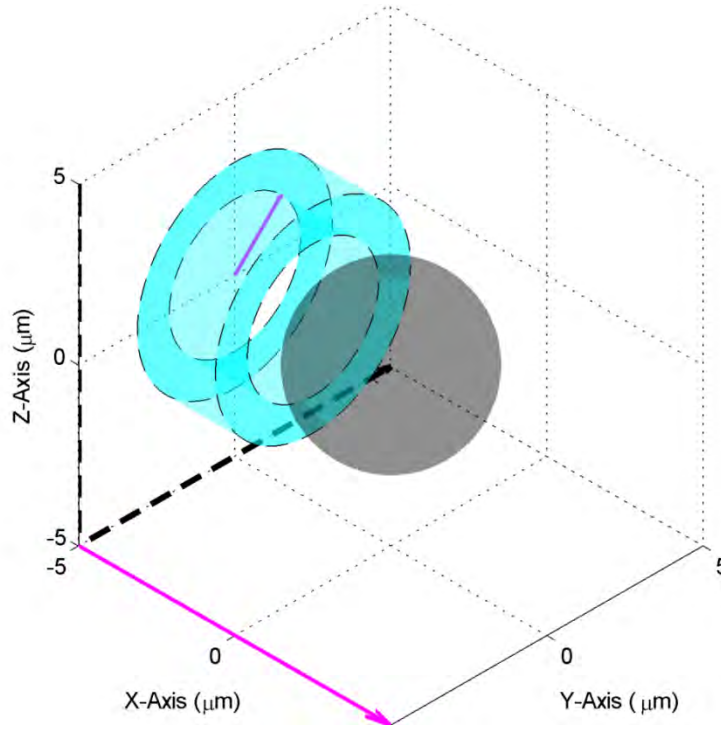


Fig. 2.7: Cylindrical Bins

This cartoon depicts how a single cylindrical bin is oriented with respect to the simulation volume. The cell is shown as a gray sphere in the center of the volume. The height of the cylinder is Δx , and the thickness of the wall is $\Delta \rho$. We count the average number of pheromone molecules between the walls of this cylinder and divide by the volume to estimate the concentration. Specifically, we choose $\Delta x = \frac{10 \mu\text{m}}{64} = 0.15625 \mu\text{m}$ and $\Delta \rho = \frac{5 \mu\text{m}}{32} = 0.15625 \mu\text{m}$. Molecules located near the edges of the simulation volume, $\rho > 5 \mu\text{m}$, are not tallied.

REFERENCES

1. Jin M, Errede B, Behar M, Mather W, Nayak S, et al. (2011) Yeast dynamically modify their environment to achieve better mating efficiency. *Sci Signal* 4: ra54.
2. Valdez-Taubas J, Pelham HRB (2003) Slow Diffusion of Proteins in the Yeast Plasma Membrane Allows Polarity to Be Maintained by Endocytic Cycling. *Curr Biol* 13: 1636–1640.
3. David NE, Gee M, Andersen B, Naider F, Thorner J, et al. (1997) Expression and purification of the *Saccharomyces cerevisiae* alpha-factor receptor (Ste2p), a 7-transmembrane-segment G protein-coupled receptor. *J Biol Chem* 272: 15553–15561.
4. Dube P, Konopka JB (1998) Identification of a polar region in transmembrane domain 6 that regulates the function of the G protein-coupled alpha-factor receptor. *Mol Cell Biol* 18: 7205–7215.
5. Jenness DD, Burkholder AC, Hartwell LH (1986) Binding of alpha-factor pheromone to *Saccharomyces cerevisiae* a cells: dissociation constant and number of binding sites. *Mol Cell Biol* 6: 318–320.
6. Yi T-M, Kitano H, Simon MI (2003) A quantitative characterization of the yeast heterotrimeric G protein cycle. *Proc Natl Acad Sci U S A* 100: 10764–10769.
7. Bajaj A, Celić A, Ding F-X, Naider F, Becker JM, et al. (2004) A fluorescent alpha-factor analogue exhibits multiple steps on binding to its G protein coupled receptor in yeast. *Biochemistry* 43: 13564–13578.
8. Cherezov V, Rosenbaum DM, Hanson M a, Rasmussen SGF, Thian FS, et al. (2007) High-resolution crystal structure of an engineered human beta2-adrenergic G protein-coupled receptor. *Science* 318: 1258–1265.
9. Erban R, Chapman SJ (2009) Stochastic modelling of reaction-diffusion processes: algorithms for bimolecular reactions. *Phys Biol* 6: 046001.
10. Lipková J, Zygalakis KC, Chapman SJ, Erban R (2011) Analysis of Brownian Dynamics Simulations of Reversible Bimolecular Reactions. *SIAM J Appl Math* 71: 714–730.
11. Hao N, Nayak S, Behar M, Shanks RH, Nagiec MJ, et al. (2008) Regulation of Cell Signaling Dynamics by the Protein Kinase-Scaffold Ste5. *Mol Cell* 30: 649–656.
12. Dyer JM, Savage NS, Jin M, Zyla TR, Elston TC, et al. (2013) Tracking shallow chemical gradients by actin-driven wandering of the polarization site. *Curr Biol* 23: 32–41.
13. Dixit G, Kelley JB, Houser JR, Elston TC, Dohlman HG (2014) Cellular Noise Suppression by the Regulator of G Protein Signaling Sst2. *Mol Cell* 55: 85–96.
14. Kelley JB, Dixit G, Sheetz JB, Venkatapurapu SP, Elston TC, et al. (2015) RGS Proteins

and Septins Cooperate to Promote Chemotropism by Regulating Polar Cap Mobility. *Curr Biol* 25: 275–285.

15. Hicke L, Riezman H (1996) Ubiquitination of a yeast plasma membrane receptor signals its ligand-stimulated endocytosis. *Cell* 84: 277–287.
16. Hicke L, Zanolari B, Riezman H (1998) Cytoplasmic tail phosphorylation of the α -factor receptor is required for its ubiquitination and internalization. *J Cell Biol* 141: 349–358.

CHAPTER 3: USING STOCHASTIC SIMULATIONS TO ASSESS NOISE-REDUCTION MECHANISMS DURING GRADIENT SENSING IN YEAST³

Introduction

The ability to detect a chemical gradient is fundamental to many cellular processes. To survive, unicellular organisms may need to move (chemotaxis) or grow (chemotropism) towards a favorable environment. Some cells are capable of detecting extremely shallow gradients, even in the presence of significant molecular-level noise. For example, yeast cells find a mating partner by detecting a pheromone gradient using a spatial sensing mechanism. Reportedly, yeast can detect the direction of gradients as shallow as $0.1 \text{ nM}/\mu\text{m}$. Mechanisms, such as time-averaging and the removal of pheromone molecules, have been proposed to explain how yeast cells filter noise and detect the shallow gradients. Here, we develop a Particle-Based Stochastic Reaction-Diffusion model of pheromone receptor dynamics to determine the limits of gradient sensing and test the efficacy of noise reduction mechanisms. We develop a novel simulation method for establishing chemical gradients that should find applications in other contexts. Our results indicate neither time-averaging nor receptor endocytosis significantly improve the cell's accuracy in detecting gradients over relevant time scales. However, we do find that the pheromone protease Bar1 can improve the cell's ability to detect gradients under certain conditions. Our results demonstrate the physical barrier of the cell membrane sharpens chemical gradients across the cell. This result has implications for interpreting experimental results, because neighboring cells can alter local gradients.

³ This chapter is being drafted as part of a publication to be submitted to the journal PLoS Computational Biology. It is authored by Lakhani V and Elston T.

3.1 Background

Gradient sensing, which is the ability to detect the direction of a chemical gradient, is a fundamental biological process. To survive, single-celled organisms must be able to grow (chemotropism) or move (chemotaxis) towards a favorable chemical, such as nutrients or hormones, or move away from an unfavorable chemical, such as toxins. An ideal system for studying gradient sensing is chemotropism during the mating response in *S. Cerevisiae* (yeast). Yeast cells can exist as one of two haploid types: MATa or MAT α . MATa cells seek a mating partner by sensing a gradient of the pheromone α -factor secreted by MAT α cells (Fig. 3.1). Gradient sensing strategies fall into two major categories: temporal sensing mechanisms and spatial sensing mechanisms. Temporal sensing mechanisms, in which the organism moves around its environment and compares the concentration between its current and previous locations, are commonly utilized by small cells such as *E. Coli* ($\sim 1\mu\text{m}$). Spatial sensing mechanisms, in which the organism compares the concentration difference across the cell body, are commonly used by large cells including most eukaryotes, such as *D. Discoideum* ($\sim 15\mu\text{m}$). The fact yeast cells are not motile suggest they use a spatial sensing mechanism, despite being smaller ($\sim 4\mu\text{m}$ in diameter) than most eukaryotic cells.

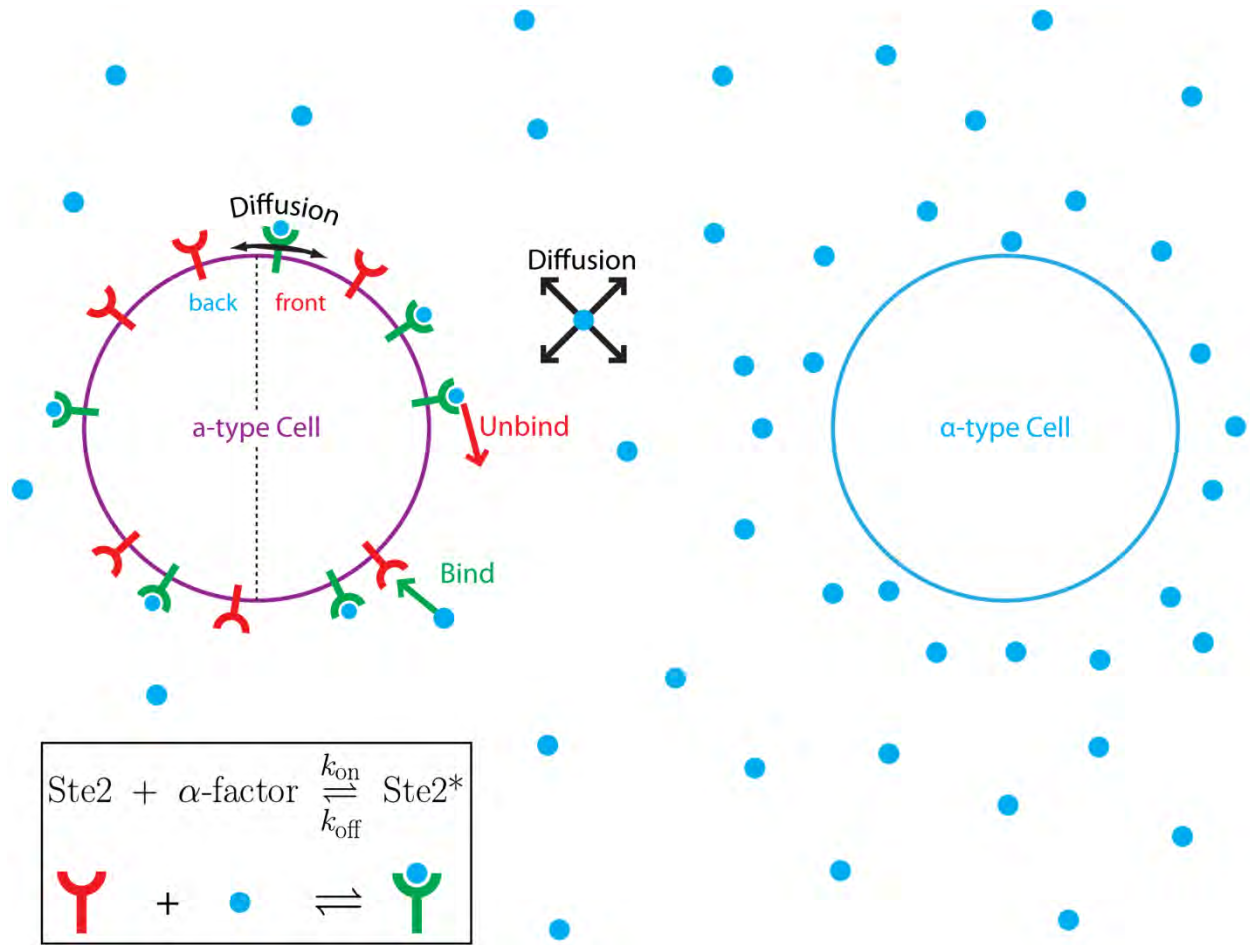


Fig. 3.1: Gradient Sensing during Yeast Mating

The MAT α cell emits the pheromone ' α -factor' (●) to attract a potential mate. Nearby MAT α cells detect the pheromone with a G-Protein Coupled Receptor, called Ste2 (Y='unbound' state and Y='bound' state). The MAT α cell determines its initial direction of growth based on the distribution of the active or 'bound' receptors. Ideally, many more active receptors are located in one part of the membrane (the portion closest to the mate) than another part. However, the stochastic nature of reactions and diffusion act as major sources of noise for a cell sensing a shallow gradient. Consider the number of active receptors in the front half of the cell; we note three basic, biophysical sources of noise for this value. First is the stochastic nature of binding (green arrow) and unbinding (red arrow) reactions. Second, fluctuations in the local pheromone concentration, due to diffusion (black arrows), also contributes. Third, receptor diffusion (black arrows) contributes spatial noise, because active receptors can bind pheromone at regions of high pheromone concentration (i.e. the front half) and diffuse to regions of low pheromone concentration (i.e. the back half). In shallow gradients, these fluctuations can prevent the cell from accurately sensing the gradient.

Experimental studies have reported that yeast cells are capable of sensing linear gradients as shallow as 0.1 nM/ μ m [1,2]. To quantify the challenges faced by a cell in such small gradients, we can estimate the average number of active receptors (receptor occupancy) in the

front half of the cell (pointing up the gradient) versus the back half of the cell (pointing down the gradient) as well as the size of the fluctuations about the mean receptor occupancy (Fig. 3.1).

The average receptor occupancy in each half of the cell is roughly given by $n = N \frac{c_i}{K_D + c_i}$, where c_i is the average concentration in the front (or back) of the cell. This expression is an approximation because it does not correctly take into account the spatial dependence of the gradient across the cell. In a linear pheromone gradient of 0.1 nM/ μm centered at the K_D (about 7 nM) of the receptor (Table 2.1), the difference in receptor occupancy is $\Delta n = n_{\text{front}} - n_{\text{back}} \approx 45$. For a cell with 10000 receptors, this calculation estimates a less than 1% difference in receptor occupancy between the front and back of the cell. Following the work of Lauffenburger, the magnitude of the fluctuations in receptor occupancy is $\sigma_n = \sqrt{N \frac{c K_D}{(K_D + c)^2}} \approx 50$ [3]. Hence for a gradient of 0.1 nM/ μm , the signal is masked by the noise: $\Delta n \approx 45 \pm 50$.

To explain how cells overcome these fluctuations, various noise-reduction mechanisms have been proposed: including time-averaging, gradient sharpening via extracellular degradation of pheromone by the protease Bar1 [4,5] and removal of active receptors via endocytosis to avoid resampling [6,7]. The limits of these mechanisms were estimated from mathematical models. Time-averaging, the first mechanism, requires sufficient time for the cell to sample multiple binding and unbinding events. This sampling frequency is dictated by the reaction rates. In the yeast mating system, the K_D of Ste2 binding to α -factor is known to be around 7nM [8–12]. Reported values for the unbinding rate are extremely slow: on the order of $10^{-4} - 10^{-3} \text{ s}^{-1}$ [12,13]. These values imply a binding rate on the order of $10^4 - 10^5 (\text{M} \cdot \text{s})^{-1}$, which is many orders of magnitude slower than the diffusion limit. With a dissociation rate of 0.0011 s^{-1} [12], changes in receptor occupancy occur on the order of 10s of minutes. Given that yeast cells begin chemotropic growth within 30 minutes of exposure to pheromone, there seems to be insufficient

time for the cell to accurately sense a shallow gradient using the time-averaging mechanism alone. In the second mechanism, MATa cells secrete Bar1: a protease that degrades extracellular α -factor. This cellular process is known to locally sharpen the pheromone gradients between neighboring cells [4,5]. This sharpening effect has been suggested as a mechanism for sensing shallow gradients and ensuring that two or more cells avoid competing for the same mate [4,5]. It is not known if this sharpening sufficiently reduces the noise to enable the cells to gradient sense. In the third mechanism, active receptors are removed from the membrane through endocytosis and newly synthesized receptors are brought to the membrane on vesicles. This receptor cycling has been suggested as a mechanism to improve gradient sensing [6,7], by removing pheromone from the environment thereby preventing re-sampling of the same ligand. Additionally, the endocytosis rate of active receptors, 0.0021 s^{-1} [14,15], is faster than reported unbinding rates. Therefore, receptor endocytosis may improve the sampling frequency.

Most of the estimates and mechanisms discussed above are based on mathematical models in which simplifying assumptions are made to allow analytic tractability. To go beyond these models, we build a computational model based on fundamental biophysical processes that allows us to evaluate noise-reduction mechanisms and study gradient sensing with minimal assumptions. In particular, we develop a Particle-Based Stochastic Reaction-Diffusion Model to study receptor dynamics in a gradient. Our simulations reveal: 1) time-averaging and receptor cycling, wherein Ste2 exocytosis is isotropic (unpolarized), are insufficient for yeast cells to confidently detect the direction of shallow gradients in a realistic time scales, and 2) isotropic exocytosis of Bar1 may improve gradient sensing for cells with fast reaction rates. Additionally, our modeling method reveals that 1) the physical barrier of the cell membrane sharpens the gradient, and 2) diffusion of the receptor reduces the cell's ability to detect the direction of the

gradient. Our approach bridges the theoretical mathematical models and *in vivo* experimental approaches.

3.2 Overview

As outlined in the Introduction, we are motivated by the surprising ability of yeast cells to sense shallow gradients even in the presence of significant amounts of noise. Accordingly, we set the biophysical parameter values in our simulations to match the yeast mating response system (Table 2.1). Although the K_D of the Ste2 receptor is known to be around 7nM [8–12], there is no consensus for the binding and unbinding rates. The diffusion limit for binding is on the order of $10^9 \text{ (M}\cdot\text{s)}^{-1}$. However, one experimental study reported rates as slow as $k_{\text{on}} = 1.6 \times 10^5 \text{ (M}\cdot\text{s)}^{-1}$ and $k_{\text{off}} = 0.0011 \text{ s}^{-1}$ [12], and similar values were measured in [8]. Elsewhere, a computational model used rates ten times faster than these values; although, the experimental sources for these rates are unclear [11]. To compare how these rates affect gradient sensing, we consider both sets of reaction rates.

In Section 3.3, we simulate a cell whose receptors are in equilibrium with a uniform pheromone concentration. From these simulations, we quantify noise levels and investigate noise-reduction through time-averaging. We compare our simulation results to previously published theoretical studies. In Section 3.4 we simulate a cell whose receptors are in equilibrium with a pheromone gradient. We report on how the cell's presence in the gradient generates non-linear effects on the pheromone concentration and the contribution of receptor diffusion to fluctuations in the distribution of active receptors. In Section 3.5 we simulate a cell experiencing the formation of a pheromone gradient and report on the role of the protease Bar1 on the cell's ability to gradient sense.

3.3 Equilibrium Fluctuations in Uniform Pheromone

3.3.1 Equilibrium Fluctuations in Receptor Occupancy

Since the seminal work of Berg & Purcell [16], there have been many theoretical studies on the limits with which cells can measure external ligand concentrations [6,17–19] and strategies for overcoming these limitations [6]. The two main sources of fluctuations considered by these studies are fluctuations in the ligand concentration and stochastic binding and unbinding of the ligand from the receptor. Our simulations accurately capture these sources of noise as well as fluctuations from receptor diffusion. Our model also allows us to investigate non-equilibrium conditions and the effect of the cell in perturbing the ligand concentration profile. We show our simulation results agree most closely with the model from Berezhkovskii and Szabo [17].

Most models assume the cell's receptors are at equilibrium with an external, uniform ligand concentration. Therefore, we begin by performing simulations under equilibrium conditions. First, we initialize the system to have the expected number of active receptors and ligand molecules in the computational domain. Second, we allow the system to equilibrate for 30-minutes to generate a random state. Lastly, this random state is used as the initial condition for a 60-minute simulation, which we then analyze. Figure 3.2 shows the results of sixteen simulations using the parameters listed in Table 2.1, with the exceptions that: 1) there is no ligand gradient ($g = 0 \text{ nM}/\mu\text{m}$) and 2) the reaction rates are taken to be $k_{\text{on}} = 1.6 \times 10^6 \text{ (M}\cdot\text{s)}^{-1}$ and $k_{\text{off}} = 0.011 \text{ s}^{-1}$. Figure 3.2A shows the receptor occupancy time series, $n(t)$, and Figure 3.2B shows a histogram of the data.

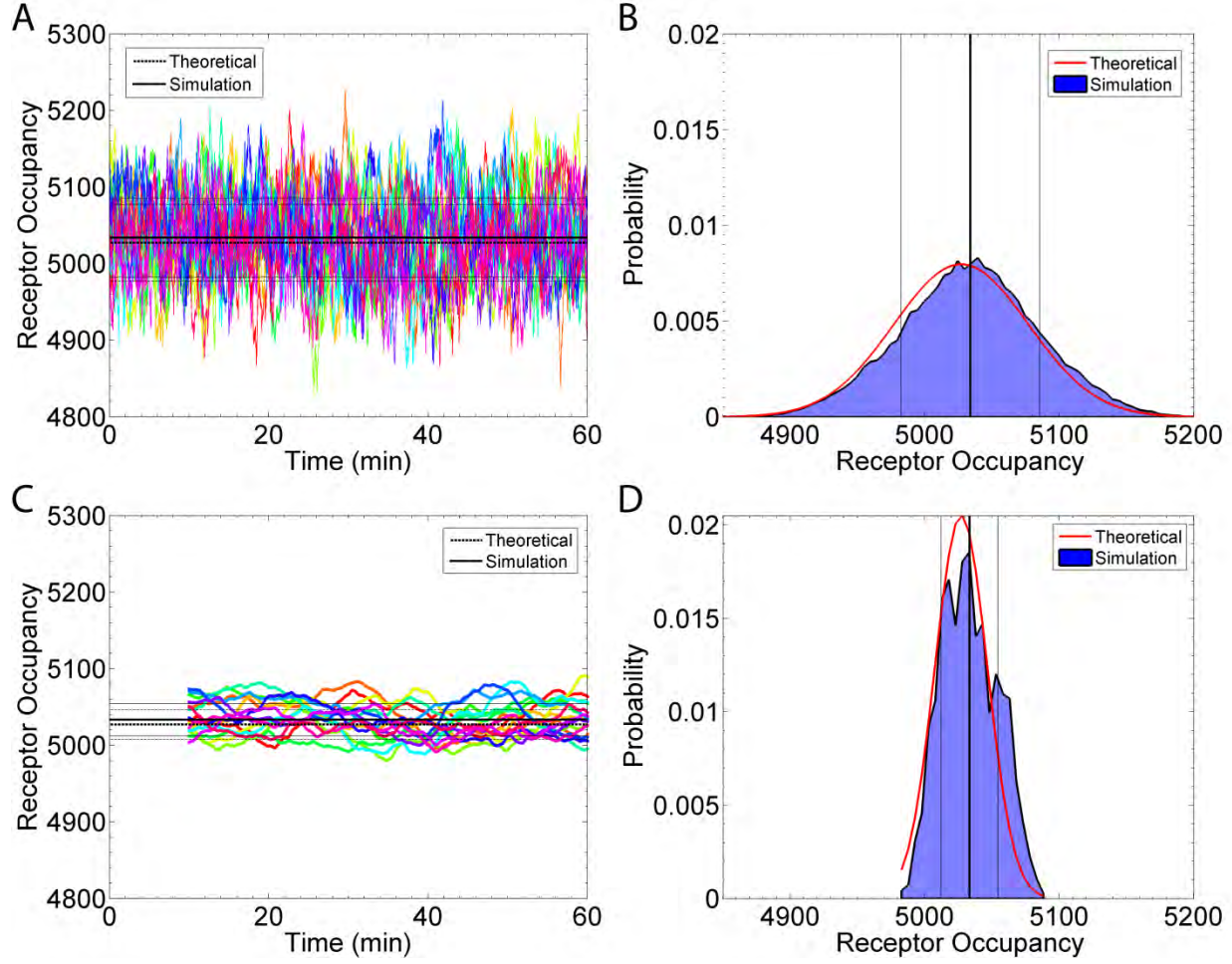


Fig. 3.2: Receptor Occupancy at Equilibrium

Here, we summarize the receptor occupancy data from sixteen simulations. The parameters used for these simulations are reported in Table 2.1, except that the gradient is 0 nM/ μm , the binding rate is $1.6 \times 10^6 \text{ (M} \cdot \text{s)}^{-1}$, and the unbinding rate is 0.011 s^{-1} . (A) A plot of the instantaneous, total number of active receptors against time: $n(t)$. Each simulation is plotted with a different color. The thick, solid, black line is the mean from the data (5035 Ste2*). The thin, solid, black lines represent one standard deviation away from the mean, as calculated from the data (± 52 Ste2*). The thick, dashed, black line is the theoretical mean calculated from Eq 3.1a (5027 Ste2*). The thin, dashed, black lines are one theoretical standard deviation from the mean as calculated from Eq 3.1b (± 50 Ste2*). (B) A histogram of the data in (A). The vertical lines are equivalent to those in (A). The red curve shows the theoretical distribution. (C) A plot of time-averaged (10 min) receptor occupancy: $n_{T=600s}(t)$. Each time point displays the average occupancy of the preceding 10 minutes. No average is available for $t < 10\text{min}$. The simulation mean is 5034 ± 23 Ste2*, and the theoretical, expected value is 5027 ± 19 Ste2*. As in (A), these values are indicated by the black lines. (D) A histogram of the data in (C), displayed in the same style as (B). The time-averaged distribution is much narrower ($\sigma = 23$ Ste2*) than the instantaneous distribution ($\sigma = 52$ Ste2*).

For uniform pheromone concentration, the mean and standard deviation can be calculated from the following equations:

$$\bar{n} = N \frac{c}{K_D + c} \quad (3.1a)$$

$$\sigma_n = \sqrt{N \frac{K_D c}{(K_D + c)^2}} \quad (3.1b)$$

(derived in [3]). In both equations, N is the total number of receptors; K_D is the dissociation constant, and c is the concentration. In Eq 3.1a, \bar{n} is the average receptor occupancy, and in Eq 3.1b, σ_n^2 is the variance of the receptor occupancy. Our simulations match the mean and standard deviation of the receptor occupancy with the theoretical values (Fig. 3.2A – B).

3.3.2 Time Averaging

A common noise reduction technique from signal-processing is time-averaging. Assume the receptor occupancy is averaged for a length of time, T , then starting with the instantaneous occupancy $n(t)$ (Fig. 3.1A), we calculate the time average as $n_T(t) = \frac{1}{T} \sum_{i=1}^{T/\Delta t} n(t - T + \Delta t \cdot i)$, where Δt is sampling interval. Figure 3.1C shows the time-averaged occupancy for $T = 10\text{min}$. The resulting time-averaged standard deviation, which we label as σ_{n_T} , is 23 Ste2* molecules (Fig. 3.1C – D). This time-averaged uncertainty is much smaller than the instantaneous uncertainty of 52 Ste2* molecules (Fig. 3.1A – B). Figure 3.1D shows the corresponding histogram. The theoretical time-averaged values (dashed lines in Figure 3.1C and red curve in Figure 3.1D) are calculated from the theoretical work of Berezhkovskii and Szabo [17]. Below, we discuss the empirical rationale for choosing this model out of the many theoretical models.

In 1977, Berg & Purcell derive an expression for the lower bound on the accuracy a cell can achieve when time-averaging [16]:

$$\text{CV}^2 = \frac{\sigma_c^2}{c^2} = \frac{1}{\pi D R c T} \left(1 + \frac{k_{\text{on}} c}{k_{\text{off}}} \right) \quad (3.2)$$

where $\frac{\sigma_c^2}{c^2}$ is the time-averaged variance in the concentration estimation divided by the average concentration squared; because this ratio is the Coefficient of Variation squared, we label it CV^2 . In Eq 3.2, D is the diffusion constant for the ligand; R is the radius of the cell; c is the concentration of ligand, and T is the length of time-averaging. The Berg & Purcell model assumes the binding rate is diffusion-limited and the cell has an excessive number of receptors ($N \approx \frac{DR}{k_{on}}$) [16]. Hence, their conclusion shows the cell's accuracy is limited by the stochastic arrival of ligand molecules to the cell [16]. In 2005, Bialek & Setayeshgar derived the CV^2 for a binding rate slower than the diffusion-limit [18]:

$$CV^2 = \frac{\sigma_c^2}{c^2} = \frac{2}{k_{on}cT} \left(1 + \frac{k_{on}c}{k_{off}} \right) + \frac{1}{\pi DRcT} \quad (3.3)$$

Eqs 3.2 & 3.3 were reconciled in 2014 by Kaizu et. al. [19]:

$$CV^2 = \frac{\sigma_c^2}{c^2} = \frac{2}{k_{on}cT} \left(1 + \frac{k_{on}c}{k_{off}} \right) + \frac{1}{2\pi DRcT} \left(1 + \frac{k_{on}c}{k_{off}} \right) \quad (3.4)$$

The second term in Eqs 3.3 & 3.4 is the contribution from fluctuations in the arrival of ligand molecules, similar to Eq 3.2. The first term in Eqs 3.3 & 3.4 is the contribution from stochastic binding and unbinding reactions. Eqs 3.3 & 3.4 were derived for a single receptor in solution [20]. In 2013, Berezhkovskii & Szabo derived an expression for the CV^2 that includes both major sources of noise and considers an arbitrary number of receptors [17]:

$$CV^2 = \frac{\sigma_c^2}{c^2} = \frac{2}{Nk_{on}cT} \left(1 + \frac{k_{on}c}{k_{off}} \right) + \frac{1}{2\pi DRcT} \left(1 + \frac{k_{on}c}{Nk_{off}} \right) \quad (3.5)$$

N is the total number of receptors on the cell surface. Note for $N = 1$, Eq 3.5 is equivalent to Eq 3.4. For comparison, we plot Eqs 3.2 & 3.5 along with the results from our simulations (Fig. 3.3). In order to compare our simulation data to these theoretical models, we must express our simulation data as CV^2 . From the receptor occupancy data in Figure 3.2A, we calculate time-

averages of varying lengths ($T = \{10, 20 \dots 3600\}$ sec). For example, the case of $T = 600$ s is shown in Figures 3.2C & 3.2D. Hence, for each value of T , we obtain a dataset (e.g. Fig. 3.2C) and calculate the variance of that dataset: $\sigma_{n_T}^2$, equivalently called the time-averaged variance in receptor occupancy. We also can calculate the instantaneous variance in receptor occupancy, σ_n^2 from the data presented in Figure 3.2A. Therefore, as described in [17], we can calculate an empirical CV^2 using the relationship: $CV^2 = \frac{\sigma_{n_T}^2}{\sigma_n^4}$. By comparing our simulated data to the theoretical models, we find noise from stochastic binding and unbinding events is the biggest source of error when time-averaging (Fig. 3.3).

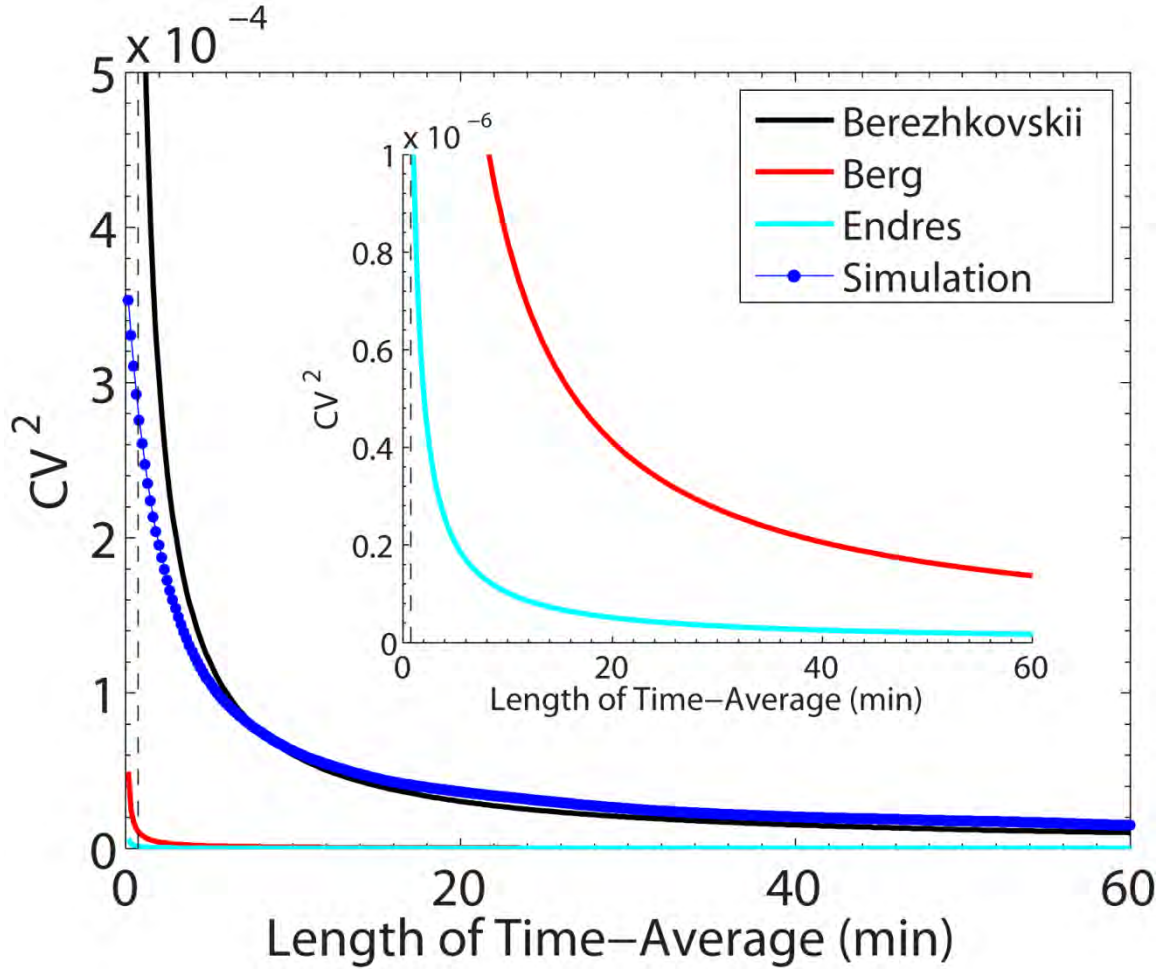


Fig. 3.3: Comparing Mathematical Models

Fig. 3.3: Comparing Mathematical Models

We plot the time-averaged accuracy of a cell measuring an external ligand concentration as predicted by various theoretical models and measured from our simulated data. The inset graph zooms in, along the y-axis. The black curve is a plot of Eq 3.4, taken from Berezhkovskii & Szabo [17]; the black, dashed, vertical line is the relaxation time, τ_N . As discussed in their paper, this model fails for $T \ll \tau_N$ [17]. The red curve is a plot of Eq 3.2, taken from Berg & Purcell [16]. The teal curve is a plot of Eq 3.6, taken from Endres & Wingreen [6]. The blue curve is calculated from the simulation data shown in Figure 3.2A. As discussed in the text, Eqs 3.2 & 3.6 (red & teal curves) are derived from models which predict the cell's accuracy is solely limited by the stochastic arrival of pheromone molecules [6,16]. Eq 3.5 (black curve) is derived from a model which finds the cell's accuracy is additionally limited by the stochasticity of binding and unbinding reactions [17]. Our simulated data agrees with the latter model. Hence, we show the stochastic reaction events play a significant role in limiting the cell's accuracy for measuring an external ligand concentration. We avoid plotting Eqs 3.3 & 3.6, because they were derived for a single receptor in solution, and therefore, their CV^2 values are 10^4 (which is the total number of receptors) times larger than Eq 3.5.

A more intuitive representation of this relationship is to plot the time-averaged standard deviation, σ_{n_T} , as a function of the length of time-averaging, T . We show that our simulations agree with the theoretical predictions from Berezhkovskii and Szabo [17] (Fig. 3.4). In addition, results from our simulations are complimentary to the theoretical model, in that, our simulations can estimate the uncertainty of short time-averaging lengths, for which the theory fails.

Specifically, the theoretical model is valid only for $T \gg \tau_N$, where, τ_N is the equilibration time scale of the system or 'relaxation time' [17]. In the case of slow reaction rates ($k_{on} = 1.6 \times 10^5$ (M·s)⁻¹ and $k_{off} = 0.0011$ s⁻¹) $\tau_N \approx 7.5$ min, which is roughly the inverse of the sum of the rates.

The theoretical predictions diverge from our simulation results for $T < 40$ min (Fig. 3.4B). This time frame is the domain of interest with respect to the yeast mating response, because most yeast cells begin chemotropic growth within 30 minutes after pheromone exposure. Hence our simulated results estimate the time-averaged uncertainty for biologically relevant time-scales, which current theories do not capture. In Figure 3.4, we compare the effectiveness of time-averaging between the two different sets of reaction rates discussed above. For fast rates, as used in [11], there is significant noise-reduction when time-averaging for short lengths of time: 10 min or less (Fig. 3.4A). Whereas, for slow rates, as measured in [12], time-averaging for as long

as 20 min only nominally reduces the noise (Fig. 3.4B). Thus, accurate measurements of the reaction rates are critical for determining how effectively time averaging can reduce noise.

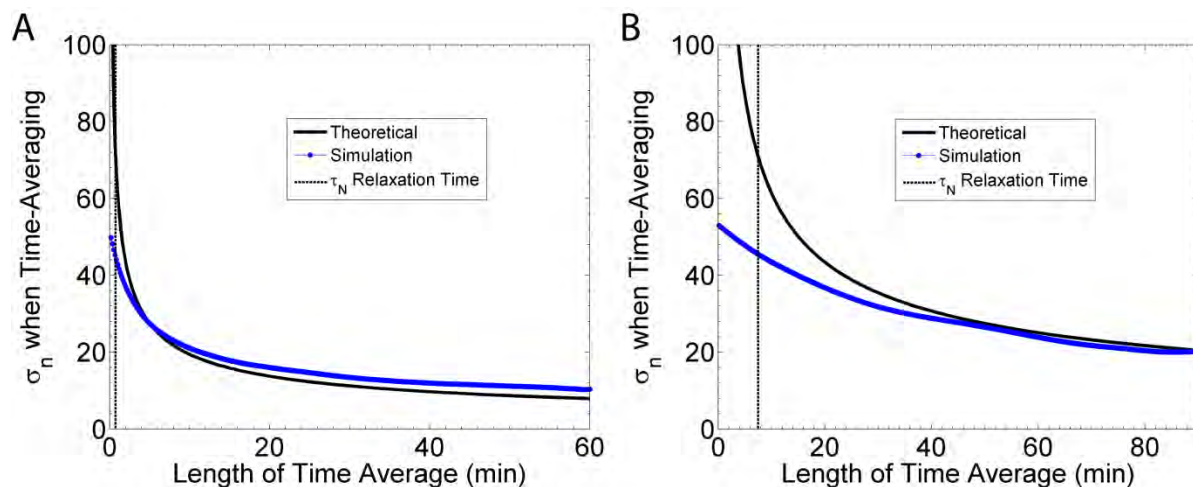


Fig. 3.4: Uncertainty as a function of Time-Averaging

In both panels, the blue line is calculated from simulation data. The solid, black line is the theoretical relationship between the length of time-averaging (T) and the standard deviation of time-averaged receptor occupancy (σ_{n_T}) as derived by Berezhkovskii and Szabo [17]. Importantly, their derived relationship is only valid for $T \gg \tau_N$; where, τ_N is the ‘relaxation time’. Hence, the theoretical curve deviates from the simulation data near this regime. The relaxation time is indicated by the vertical, dashed, black line [17]. (A) This data is calculated from the sixteen simulations used in Figure 3.2A, which use the fast reaction rates: $k_{on} = 1.6 \times 10^6 \text{ (M} \cdot \text{s)}^{-1}$ and $k_{off} = 0.011 \text{ s}^{-1}$. Our simulation data, which includes additional sources of noise, has a slightly larger standard deviation than the theoretical model. (B) This data is calculated from twenty-six simulations, which use the slow reaction rates: $k_{on} = 1.6 \times 10^5 \text{ (M} \cdot \text{s)}^{-1}$ and $k_{off} = 0.0011 \text{ s}^{-1}$.

3.3.3 The “Perfectly Absorbing” Cell and Receptor Cycling

In addition to time-averaging, Endres & Wingreen propose that a cell which “perfectly absorbs” ligand molecules, can more accurately measure an external ligand concentration [6]. A “perfect absorber” is a cell which, after binding a ligand molecule, does not release that ligand molecule back into the environment. Therefore, Endres & Wingreen argue, the cell does not count the same ligand molecule more than once. Similar to Berg & Purcell’s work [16], this model assumes that the binding rate is diffusion-limited and that there are an infinite number of receptors. Endres & Wingreen derive the following expression for the cell’s accuracy [6]:

$$CV^2 = \frac{\sigma_c^2}{c^2} = \frac{1}{4\pi DRcT} \quad (3.6)$$

Eq 3.5 is a more general expression for the CV^2 and reduces to a form similar to Eq 3.6 under certain limiting conditions. If $N \gg \frac{c}{K_D}$, then the second term in Eq 3.5 simplifies: $\frac{1}{2\pi DRcT} \left(1 + \frac{k_{on}c}{N k_{off}}\right) \approx \frac{1}{2\pi DRcT}$. If $k_{on}cN(1 - \bar{n}) = k_{on}cN / \left(1 + \frac{k_{on}c}{k_{off}}\right) \gg 4\pi DRc$, where \bar{n} is the expected fractional occupancy, then the first term in Eq 3.5 is much smaller than the second term, and the expression reduces to the following expression:

$$CV^2 = \frac{\sigma_c^2}{c^2} \approx \frac{1}{2\pi DRcT} \quad (3.7)$$

The two conditions used to derive the above expression, are closely related to the assumptions made for Eqs 3.2 & 3.6. The first condition, $N \gg \frac{c}{K_D}$, indicates the need for many receptors; for typical parameter sets, this condition is easily met. The second condition indicates that every ligand molecule which encounters the cell must be captured by an unbound receptor. This condition can be met by having a high reaction rate (k_{on}), a large number of receptors (N), or both. Hence, if there are enough receptors and the binding rate is fast enough to guarantee the capture of every ligand molecule, then the cell's accuracy is given by Eq 3.7 [17] (see also “The Perfect Instrument” in [16]).

The parameters in our simulations of the yeast system fail to meet these conditions and cannot be considered a “perfect absorber”. The binding rate is 3 – 4 orders of magnitude too slow. Nonetheless, we can simulate a partial absorber to determine if removing ligand from the environment, rather than releasing the ligand back into the environment, can improve the cell's accuracy. A partial absorber does not absorb every ligand molecule that arrives at the surface. Endres & Wingreen suggest receptor cycling is a potential biological mechanism which enables

the cell to absorb ligand [6]. Thus, we modify our simulation algorithm to include a simplified receptor cycling mechanism. In particular, we make the following changes to our algorithm (see Section 2.7 for details). An active receptor, Ste2*, has some rate of endocytosis. Immediately upon endocytosis, an unbound receptor is created in a random position on the cell surface. Thus, in our simplified model, endocytosis is coupled with immediate replacement, which allows us to keep the total number of receptors constant. We also assume, Ste2* cannot unbind a ligand ($k_{\text{off}} = 0 \text{ s}^{-1}$). We find that absorbing ligand molecules does not reduce the fluctuations in time-averaged receptor occupancy as compared to unbinding and releasing ligand molecules (Fig. 3.5).

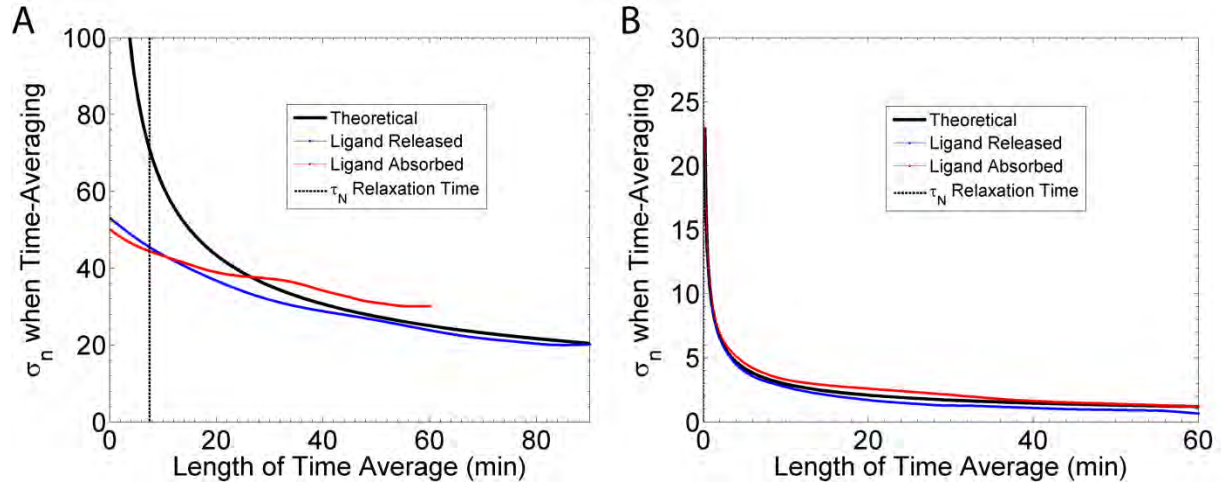


Fig. 3.5: Absorbing Ligand vs Releasing Ligand

As with Figure 3.4, we plot the standard deviation of the time-averaged receptor occupancy (σ_{n_T}) against the length of time-averaging (T). The solid, black curve represents the theoretical values calculated using Eq 3.5 as derived in [17]. The dashed, vertical, black line marks the relaxation time (τ_N). The blue curve is calculated from simulations wherein active receptors, Ste2*, unbind and release their ligand back into the environment. The red curve is calculated from simulations wherein ligand molecules are absorbed into the cell via Ste2* endocytosis. The panels show two sets of reactions rates. (A) These simulations use the measured [12] binding rate of $k_{\text{on}} = 1.6 \times 10^5 \text{ (M}\cdot\text{s)}^{-1}$ and either an unbinding rate of $k_{\text{off}} = 0.0011 \text{ s}^{-1}$ (blue curve) or endocytosis rate of $k_{\text{Endo}} = 0.0011 \text{ s}^{-1}$ (red curve). The blue curve is identical to Figure 3.4B. The red curve is calculated from eight simulations. (B) These simulations have 500 times faster rates than those in (A). These simulations use a binding rate of $k_{\text{on}} = 8 \times 10^7 \text{ (M}\cdot\text{s)}^{-1}$ and either an unbinding rate of $k_{\text{off}} = 0.55 \text{ s}^{-1}$ (blue curve) or endocytosis rate of $k_{\text{Endo}} = 0.55 \text{ s}^{-1}$ (red curve). We calculate the blue and red curves from ten simulations each. For both sets of reaction rates, the fluctuations from the ligand absorbing model are similar to fluctuations from the ligand releasing model.

3.4 Cells at Equilibrium in a Pheromone Gradient

3.4.1 Gradient Sharpening due to Steric Effects

To establish a linear pheromone gradient in our simulations, we fix the pheromone concentration at the boundaries located at $x = 5\mu\text{m}$ and $x = -5\mu\text{m}$ (Section 2.11.1). For example, to create a $0.5\text{ nM}/\mu\text{m}$ gradient with a concentration of 6.9 nM at the midpoint, the concentration at $x = -5\mu\text{m}$ is set to 4.4 nM , and the concentration at $x = 5\mu\text{m}$ is set to 9.4 nM . In addition to these two boundaries, we set the boundary of the cell to be reflective, because the cell membrane is impermeable to pheromone. This additional boundary produces nonlinear steric effects on the pheromone gradient, such that, the resulting gradient is steeper than expected (Fig. 3.6A). Details on calculating the pheromone concentration profiles are provided in Section 2.13.

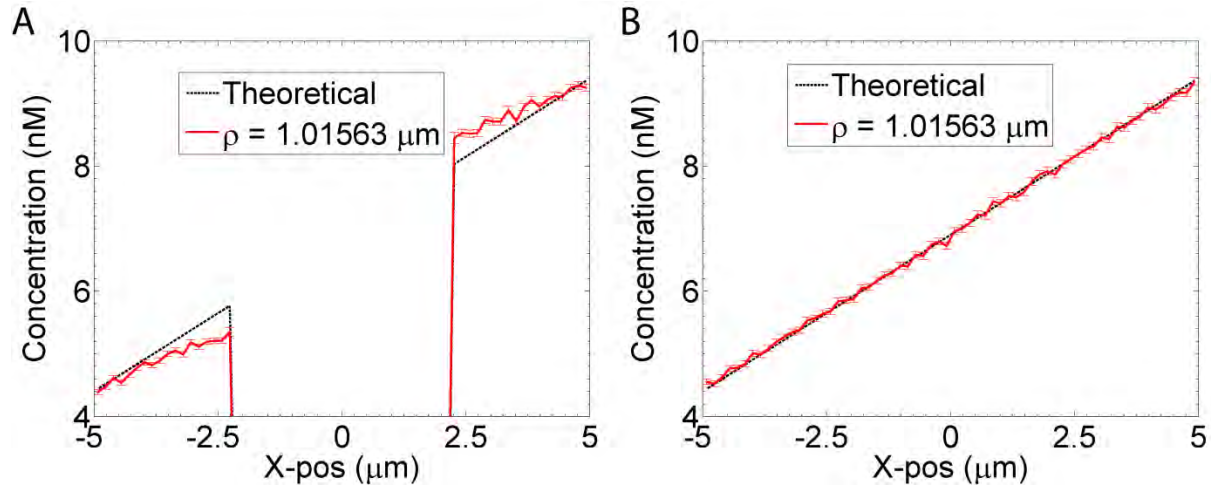


Fig. 3.6: Gradient Sharpening due to Steric Effects

We calculate the pheromone concentration by counting the average number of molecules located at different positions in the simulation space. Specifically, we tally the molecules based on their x -position and distance from the x -axis ($\rho = \sqrt{y^2 + z^2}$). Consequently, the concentration is a function of x and ρ : $c = c(x, \rho)$. Here, we plot $c(x, \rho = 1\text{ }\mu\text{m})$. This concentration calculated from our simulations is shown as a red line. The dashed, black line is the ideal, linear concentration profile. (A) This data is measured from simulations in which pheromone reflects off the cell surface. There is no pheromone between $x = -2.28\text{ }\mu\text{m}$ and $x = 2.28\text{ }\mu\text{m}$, because this region is the interior of the cell. Note that the simulated concentration deviates from the ideal linear gradient near the cell boundary. In fact, $\frac{\partial c}{\partial x} \approx 0$ near the cell boundary. (B) This data is measured from simulations identical to (A), except that the cell boundary has no effect on pheromone diffusion. That is, pheromone molecules can diffuse through the cell membrane and exist inside the cell.

This sharpening effect is solely due to the impermeability of the cell membrane to pheromone. Simulations in which pheromone are free to diffuse through the cell membrane, produce a linear gradient: $\frac{\partial c}{\partial x} = g$ (Fig. 3.6B). However, for an impermeable membrane, there is no flux across the cell boundary. Therefore, according to Fick's Law, $J = -D_\alpha \frac{\partial c}{\partial \hat{n}} = 0$ or $\frac{\partial c}{\partial \hat{n}} = 0$, where \hat{n} is the vector normal to the cell surface. As predicted by this argument, the measured gradient $\frac{\partial c}{\partial x}$ near the boundary is close to 0 (Fig. 3.6A). This constraint produces a higher than expected concentration at the front of the cell and a lower than expected concentration at the back of the cell. We now investigate whether this sharper gradient produces an appreciable difference in the distribution of active receptors.

We first consider the fast set of reaction rates: $k_{\text{on}} = 1.6 \times 10^6 \text{ (M}\cdot\text{s)}^{-1}$ and $k_{\text{off}} = 0.011 \text{ s}^{-1}$. As an initial approach to measuring the distribution of Ste2*, we evaluate the receptor occupancy in the “front half” ($x > 0\mu\text{m}$) and the “back half” ($x < 0\mu\text{m}$). Assuming a true linear gradient, $c(x) = gx + c_0$, we can calculate the mean and standard deviation of receptor occupancy using:

$$n = N \frac{1}{b-a} \int_a^b \frac{c(x)}{K_D + c(x)} dx \quad (3.8a)$$

$$\sigma_n^2 = N \frac{1}{b-a} \int_a^b \frac{K_D c(x)}{(K_D + c(x))^2} dx \quad (3.8b)$$

which are more generalized forms of Eq 3.1. For example, in the front half, $a = 0$ and $b = R$,

which gives the following expression for the mean occupancy: $\bar{n}_{\text{front}} = N + N \frac{K_D}{gR} \ln \left(\frac{K_D + c_0}{K_D + c_0 + gR} \right)$.

Empirically, we know the gradient is sharper and non-linear (Fig. 3.6A). From the simulation data, we count the number of occupied receptors located in each half at a given time. The distribution of the simulated data is plotted (Fig. 3.7A) along with the distribution expected from a true linear gradient (Fig. 3.7B).

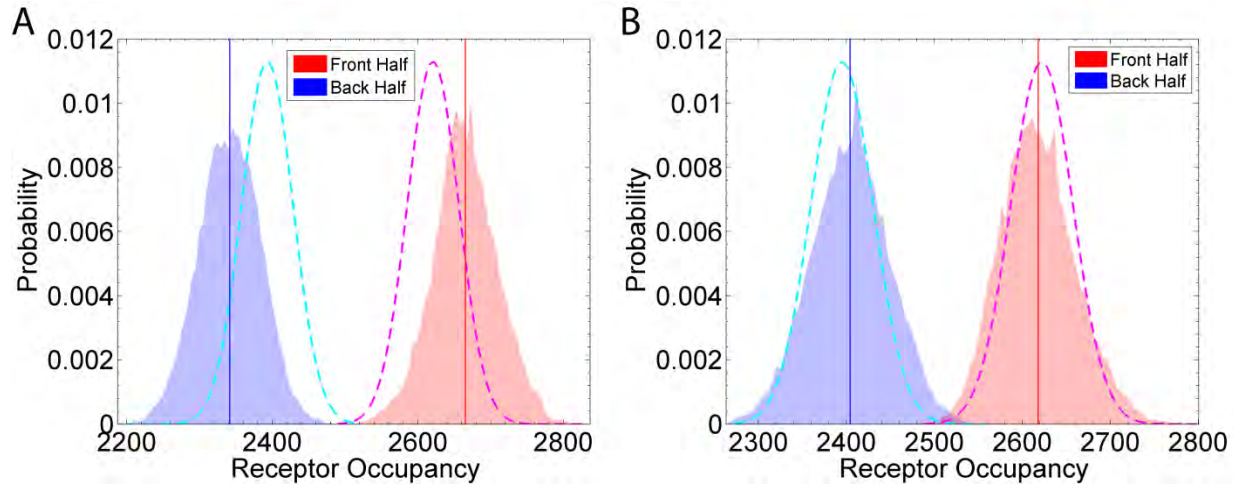


Fig. 3.7: Receptor Occupancy in Linear Gradient vs Sharpened Gradient

The blue and red histograms show the distribution of how many Ste2* are located in the back and front half, respectively, at a given time. The solid, vertical lines indicate the mean for their respective distributions. The dashed curves indicate the theoretical distribution for a cell in a true linear gradient. These theoretical curves are equivalent between the two panels: $n_{\text{back}} = 2395 \pm 35$ Ste2* and $n_{\text{front}} = 2622 \pm 35$ Ste2*. In both panels, the binding rate is $k_{\text{on}} = 1.6 \times 10^6 \text{ (M} \cdot \text{s)}^{-1}$ and the unbinding rate is $k_{\text{off}} = 0.011 \text{ s}^{-1}$. (A) These data correspond to eight simulations with the gradient in Figure 3.6A: a gradient formed with an impermeable cell membrane. From the simulations, we calculate $n_{\text{back}} = 2341 \pm 42$ Ste2* and $n_{\text{front}} = 2665 \pm 44$ Ste2*. Note there are fewer Ste2* in the back than expected, and more Ste2* in the front than expected. (B) These data correspond to eight simulations with the gradient in Figure 3.6B: a true linear gradient. From the simulations, we calculate $n_{\text{back}} = 2404 \pm 45$ Ste2* and $n_{\text{front}} = 2617 \pm 43$ Ste2*. The simulation data agrees with the theoretical prediction.

Based on the distribution of Ste2*, Figure 3.7 shows the cell is measuring the sharper gradient. We find steric effects from the cell membrane can locally sharpen the pheromone gradient. In turn, this sharpening can significantly improve the asymmetry in the distribution of active receptors. For a $0.5 \text{ nM}/\mu\text{m}$ gradient, the difference in active receptors between the front and back ($\Delta n = n_{\text{front}} - n_{\text{back}}$) changes from $\Delta n = 212$ Ste2* without sharpening to $\Delta n = 324$ Ste2* with sharpening, a more than 40% improvement.

3.4.2 Slow Reaction Rates and Receptor Diffusion add Spatial Noise

As shown in Figure 3.7, the sharpened gradient increases the difference in receptor occupancy between the front and back for a system with fast reaction rates. To evaluate the sensitivity of the system with respect to the reaction rates, we repeat the simulations using the

slow set of reaction rates: $k_{\text{on}} = 1.6 \times 10^5 \text{ (M}\cdot\text{s)}^{-1}$ and $k_{\text{off}} = 0.0011 \text{ s}^{-1}$. We find that, although the pheromone gradient is equivalently sharpened (data not shown), the simulations with slow reaction rates do not show a larger than expected difference in receptor occupancy between the front and back (Fig. 3.8A). Simulations with slow kinetics, show $\Delta n = 238 \text{ Ste2}^*$ (Fig. 3.8A); whereas, fast kinetics yield $\Delta n = 324$ (Fig. 3.7A). The cause for this discrepancy is receptor diffusion. Simulations with slow kinetics and no Ste2 diffusion ($D_{\text{Ste2}} = 0 \text{ }\mu\text{m}^2/\text{s}$) show a larger difference in receptor occupancy, $\Delta n = 315$, than simulations with receptor diffusion (Fig. 3.8).

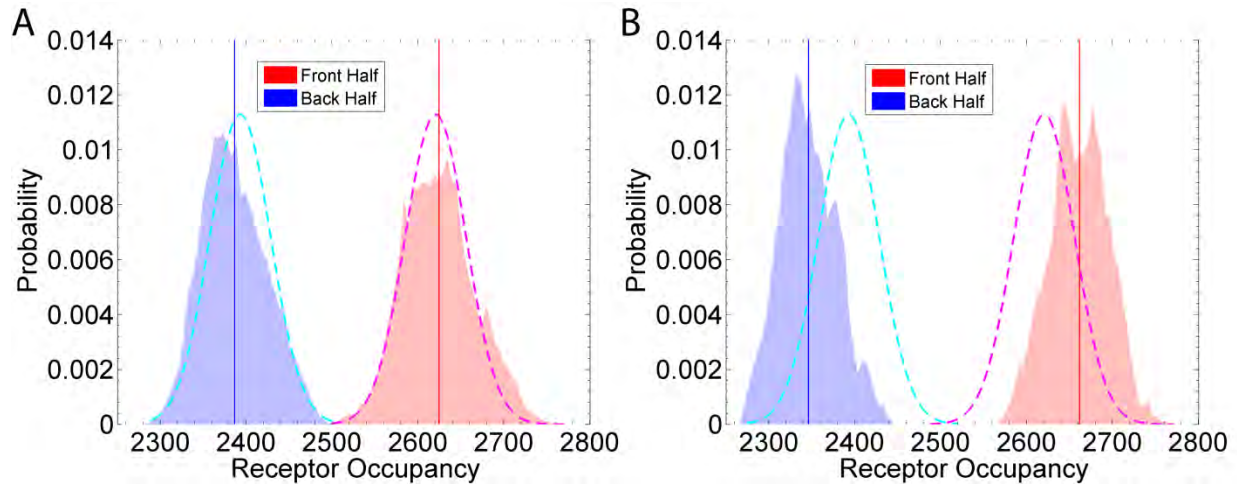


Fig. 3.8: Receptor Diffusion adds Spatial Noise

These curves are similar to those in Figure 3.7. The blue and red histograms show the distribution of the number of Ste2* molecules in the back and front halves respectively. The dashed blue and red curves respectively show the expected back and front distributions for a cell in a true linear gradient. As in Figure 3.7, the expected $n_{\text{back}} = 2393 \pm 35 \text{ Ste2}^*$, and $n_{\text{front}} = 2620 \pm 35 \text{ Ste2}^*$. These dashed curves are equivalent between the two panels. (A) These data correspond to eight simulations with slow reaction rates: $k_{\text{on}} = 1.6 \times 10^5 \text{ (M}\cdot\text{s)}^{-1}$ and $k_{\text{off}} = 0.0011 \text{ s}^{-1}$. The difference in receptor occupancy is $\Delta n = 238 \text{ Ste2}^*$. (B) These data correspond to eight simulations with slow reaction rates and no receptor diffusion; that is, $D_{\text{Ste2}} = 0 \text{ }\mu\text{m}^2/\text{s}$. The difference in receptor occupancy is $\Delta n = 315 \text{ Ste2}^*$.

Compared to a system with a fast unbinding rate, a system with a slow unbinding rate allows activated receptors to diffuse longer distances before reverting to the inactive state. For example, with a slow unbinding rate of $k_{\text{off}} = 0.0011 \text{ s}^{-1}$, a Ste2* molecule typically diffuses

$$\sqrt{2D_{\text{Ste2}} \left(\frac{1}{k_{\text{off}}} \right)} = 2.2 \mu\text{m}, \text{ or } \frac{2.2 \mu\text{m}}{2.5 \mu\text{m}} * \frac{180^\circ}{\pi} \approx 51^\circ, \text{ away from the position at which it bound a}$$

pheromone molecule. Hence, many Ste2* molecules will diffuse from the front half to the back half of the cell or vice-versa. Because there are more Ste2* in the front half than back half, there is a net flux of Ste2* from the front to the back. This flux reduces Δn : the difference in receptor occupancy between the front and back halves.

A second interpretation is that the diffusion of active receptors creates spatial noise. The current position of a Ste2* molecule is unlikely to be the position where that receptor bound a pheromone molecule. The unbinding rate determines how far a Ste2* molecule can diffuse before unbinding the pheromone molecule. A fast unbinding rate reduces this distance; therefore, the current position of each Ste2* molecule is close to where the binding event happened. Although the receptor's diffusion adds spatial noise to the distribution of active receptors, a sufficiently fast unbinding rate minimizes this noise. Alternatively, endocytosis of active receptors would also act to minimize this effect.

3.4.3 Time-averaging and the Angle of Estimation

Up to this point, we have limited our discussion to cells in a gradient of 0.5 nM/ μ m. Figures 3.7A & 3.8A show that for this case, there is a clear distinction between the occupancy in the front and back halves. However, for a shallow gradient, 0.1 nM/ μ m, there is significant overlap in the occupancy distributions for the two halves (data not shown). Additionally, we have thus far limited our analysis of the Ste2* distribution by dividing the cell into two halves. This division artificially introduces spatial information, in that, there are only two options for the direction of the gradient. In reality, the cell has no such spatial information; the complete, 3-Dimensional distribution of Ste2* must be considered in order to determine the direction of the gradient. Therefore, we use the vector that points from the origin to the center of mass of the distribution of active receptors to estimate the direction of the gradient. In Figure 3.9A, we plot

the angle between this estimate and the direction of the gradient (taken to be in the $+\hat{x}$ direction) for each time point. We also plot the distribution of this angle (Fig. 3.9B) and calculate the time-averaged angle (Fig. 3.9C) and the corresponding distribution (Fig. 3.9D). By comparing the distributions (Figs. 3.9B & 3.9D), we find time-averaging improves the likelihood that the average position, or center of mass, of Ste2* correctly points towards the gradient.

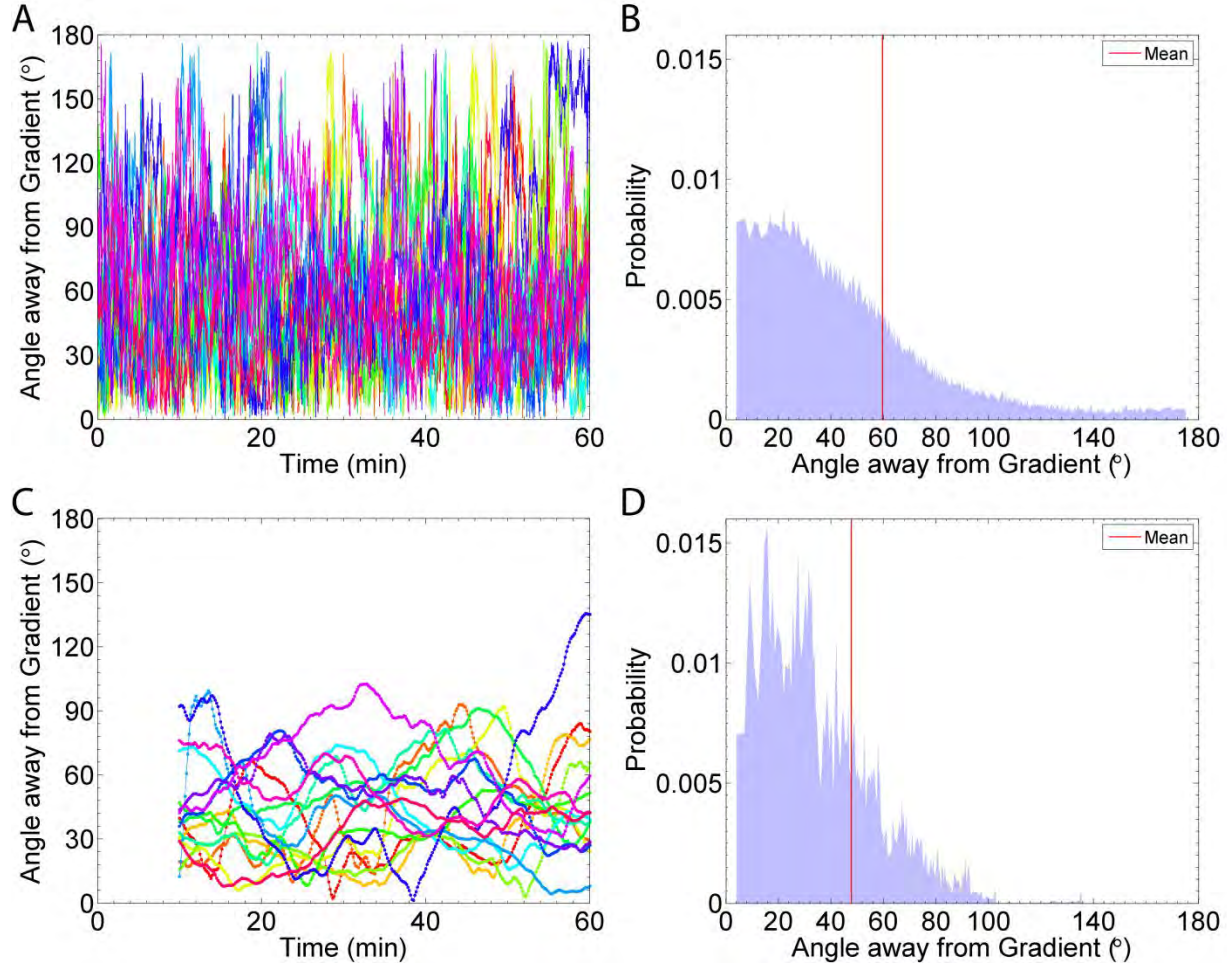


Fig. 3.9: Estimated Angle away from Gradient

The data presented here is calculated from sixteen simulations of cells in a $0.1 \text{ nM}/\mu\text{m}$ gradient, using the fast binding and unbinding rates: $k_{\text{on}} = 1.6 \times 10^6 \text{ (M} \cdot \text{s)}^{-1}$ and $k_{\text{off}} = 0.011 \text{ s}^{-1}$. The “Angle away from Gradient” is the angle between the gradient (pointing to $+\hat{x}$) and the vector of the average Ste2* position. (A) Plot of the instantaneous angle. Each of simulation is plotted in a single color. (B) The distribution of the data in panel (A). The average angle is indicated by the vertical, red line; the mean is 59.7° . (C) A plot of the time-averaged (10 min) angle. (D) The distribution of the data in (C). The average angle is indicated by the vertical, red line; the mean is 47.8° .

From the distributions, we can calculate the probability that the cell's estimate of the gradient is accurate within a given threshold. For example, from the instantaneous distribution (Fig. 3.9B), we find there is an 86% probability, the cell's estimate is within 90° of the gradient's true direction. After time-averaging for 10 minutes (Fig. 3.9D), this probability (that the cell is correct within 90°) increases to 97%. Figure 3.10 shows how this accuracy, for selected thresholds, improves with time-averaging. In the case of fast reaction rates: $k_{\text{on}} = 1.6 \times 10^6 \text{ (M}\cdot\text{s)}^{-1}$ and $k_{\text{off}} = 0.011 \text{ s}^{-1}$, we find that after about 20 minutes of time-averaging, cells know the direction of the gradient within 90° (Fig. 3.10A). In contrast, for the case of slow reaction rates: $k_{\text{on}} = 1.6 \times 10^5 \text{ (M}\cdot\text{s)}^{-1}$ and $k_{\text{off}} = 0.0011 \text{ s}^{-1}$, if a cell time-averages for 90 minutes, then the probability of being correct within 90° is only 88% (Fig. 3.10B). Similarly, we find short lengths of time-averaging (e.g. less than 20 min) are more beneficial in the case of fast reaction rates than slow rates (Fig. 3.10). That is, for 20 min of time-averaging, a system with fast rates improves the probability of being accurate within 60° by 22%; whereas, a system with slow rates improves the probability of being accurate within 60° by only 8%.

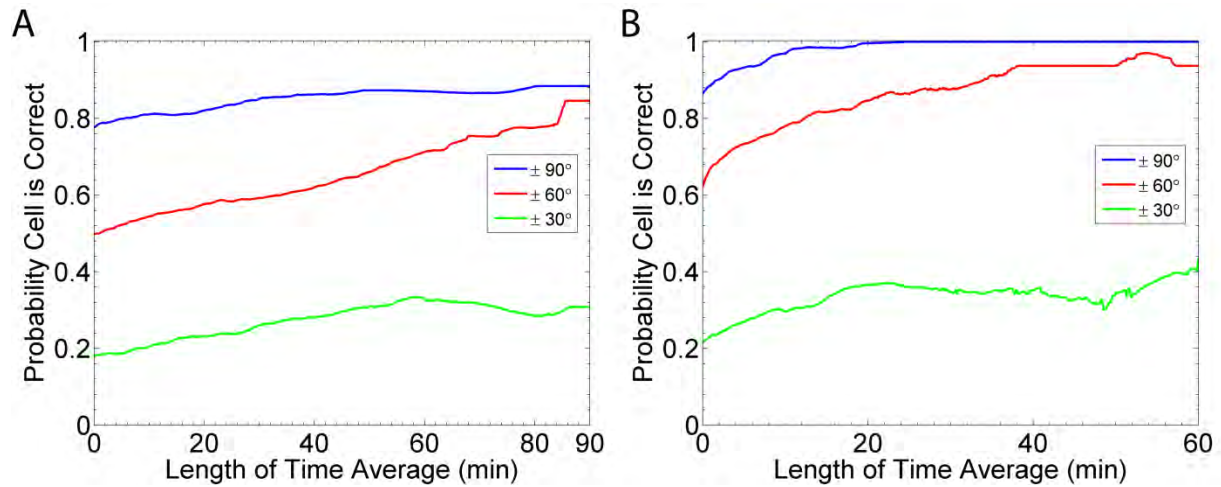


Fig. 3.10: Estimated Angle Improves with Time-averaging

For different lengths of time-averaging, we plot the probability that the cell's estimation of the gradient is accurate within three thresholds: 90° (blue curves), 60° (red curves) and 30° (green curves). In all cases, we simulate cells in a $0.1 \text{ nM}/\mu\text{m}$ gradient. (A) Slow binding kinetics. In these simulations the binding and unbinding reaction rates are $k_{\text{on}} = 1.6 \times 10^5 \text{ (M}\cdot\text{s)}^{-1}$ and $k_{\text{off}} = 0.0011 \text{ s}^{-1}$ respectively. The curves are

calculated from twenty-six simulations, in which, we model the reaction between Ste2 & pheromone. (B) Fast binding kinetics. In these simulations the binding and unbinding reaction rates are $k_{\text{on}} = 1.6 \times 10^6 \text{ (M} \cdot \text{s)}^{-1}$ and $k_{\text{off}} = 0.011 \text{ s}^{-1}$ respectively. The curves are calculated from the same sixteen simulations used to generate Figure 3.9, in which, we model the reaction between Ste2 & pheromone.

3.5 Sensing During Gradient Formation

3.5.1 Formation of the Gradient

We have thus far only considered steady state pheromone gradients. We now study the Ste2* distribution as the gradient forms across the cell. To simulate a developing pheromone gradient, we fix the pheromone concentration at the $x = 5 \mu\text{m}$ boundary and enforce a partially absorbing boundary at $x = -5 \mu\text{m}$ (Section 2.4.2). That is, we do not inject new molecules from the $x = -5 \mu\text{m}$ boundary. Figure 3.11B shows the steady state pheromone concentration profile in the absence of a cell. The inclusion of the cell boundary produces the pheromone concentration profile shown in Figure 3.11A. Again, we find the presence of the cell sharpens the gradient similar to previous cases (Fig. 3.6). However, unlike previous cases, the concentration is not held constant at $x = -5 \mu\text{m}$, and the final concentration at this boundary depends on the presence or absence of the cell (Fig. 3.11).

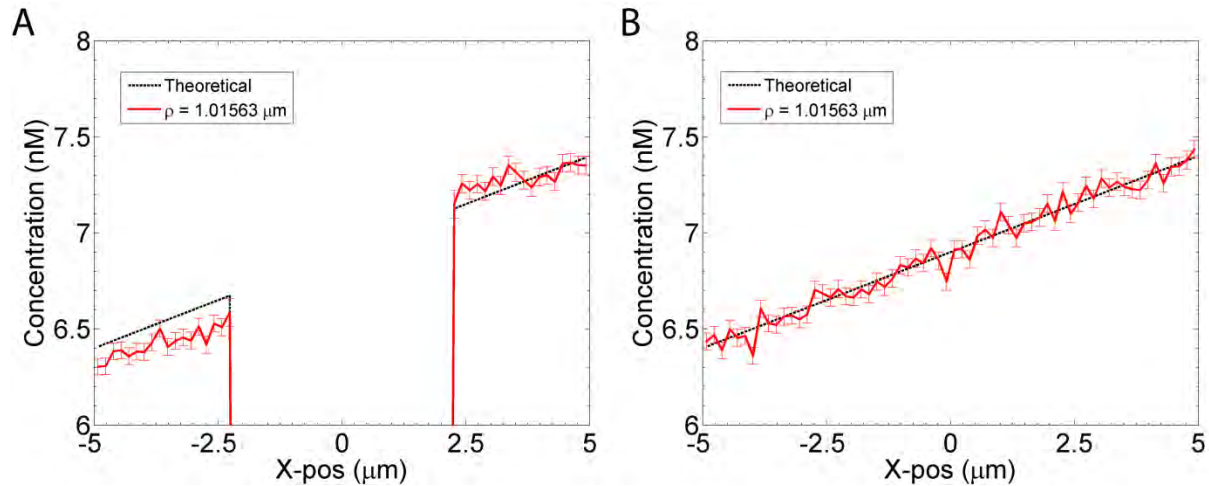


Fig. 3.11: Cell's Physical Boundary Sharpens an Emerging Gradient

We simulate diffusion of an emerging gradient by adding pheromone molecules at the $x = 5 \mu\text{m}$ boundary and partially absorbing molecules at the $x = -5 \mu\text{m}$ boundary (Section 2.4.2). Here, we show the steady state concentration profiles for two scenarios. These simulations are equivalent (same parameter set)

except for one difference. (A) Physical cell boundary present. We model the cell boundary as a sphere, which is reflective with respect to pheromone. We do not simulate any reactions. (B) We simulate neither the physical cell boundary nor any reactions. Note that the physical presence of the cell sharpens the gradient and reduces the overall concentration behind the cell ($x < -2.28\mu\text{m}$).

3.5.2 Transient Differences in Receptor Occupancy

Recent work that evaluated the binding kinetics of the back and front halves of the cell suggested that the largest difference in receptor occupancy occurs transiently, as the receptor occupancy approach equilibrium [13]. Furthermore, the previous work concludes this transient difference in occupancy is significant enough to improve the gradient sensing ability of yeast cells during mating [13]. To investigate these transient effects, we use our method discussed above to simulate a cell whose receptors are all initially unoccupied and flow pheromone from one side. Again, we study the two sets of rates discussed above. The reaction rates determine when the maximum difference in receptor occupancy occurs and the time scale for approaching equilibrium (see also [13]). We find, for both sets of reaction rates, that although the presence of the cell leads to a larger transient difference in receptor occupancy than expected, the maximum difference in receptor occupancy is not significantly larger than at equilibrium (Fig. 3.12).

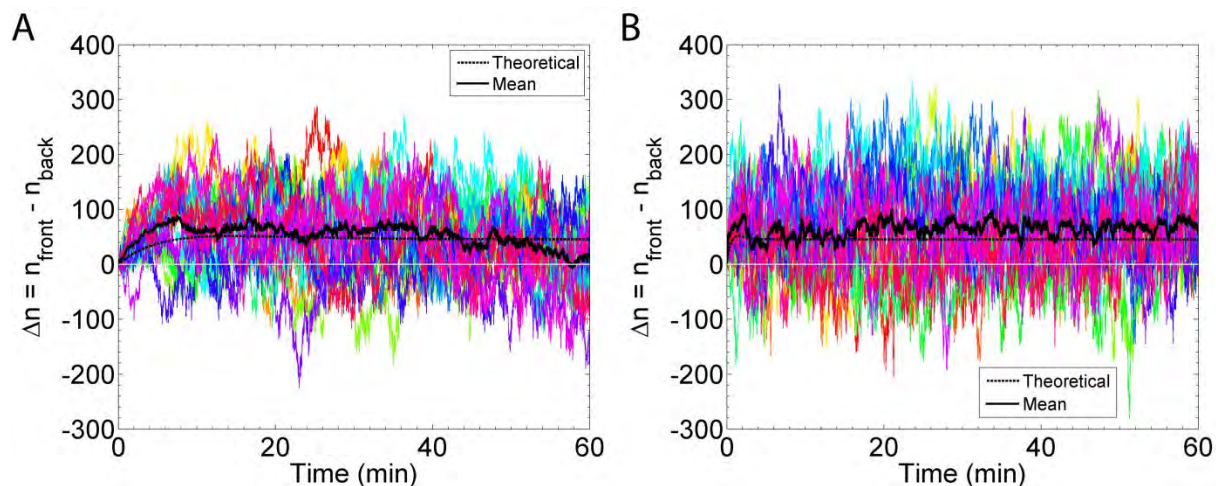


Fig. 3.12: Initial, Transient Difference in Receptor Occupancy

We simulate a cell in an emerging pheromone gradient of $0.1 \text{ nM}/\mu\text{m}$ (whose midpoint is 6.9 nM) and plot the difference in receptor occupancy between the back and front halves: Δn . Each colored time trace shows the result from a single simulation. The solid black lines are the population average. The dashed

black line shows the predicted difference, based on theoretical binding kinetics for an average pheromone concentration of 6.775nM in the back half and 7.025nM in the front half. This theoretical calculation is discussed in [13]. The thin white line demarcates positive from negative Δn values. (A) Slow binding rates. Here, we show the results from twenty simulations, in which, $k_{\text{on}} = 1.6 \times 10^5 \text{ (M} \cdot \text{s)}^{-1}$ and $k_{\text{off}} = 0.0011 \text{ s}^{-1}$. The difference initially peaks around 8 min after pheromone exposure. (B) Fast binding rates. Here, we show the results from eight simulations, in which, $k_{\text{on}} = 1.6 \times 10^6 \text{ (M} \cdot \text{s)}^{-1}$ and $k_{\text{off}} = 0.011 \text{ s}^{-1}$. The difference initially peaks within 2 min of initial pheromone exposure.

As discussed in Section 3.4.3, by analyzing the Ste2* distribution as two halves, we artificially introduce spatial information about the direction of the gradient in our analysis. A more appropriate measure of the Ste2* distribution is to calculate the center of mass of all Ste2* molecules, that is, the average position of all active receptors: $\langle \overrightarrow{\text{Ste2}^*} \rangle$. We interpret the direction of the resulting vector to be the cell's estimate of the gradient's direction. We scale the magnitude of the vector and interpret the value as the cell's "confidence":

$$\text{Confidence} = \frac{1}{R} \frac{n}{N} |\langle \overrightarrow{\text{Ste2}^*} \rangle| \quad (3.9)$$

where, R is the radius of the cell; N is the total number of receptors, and n is the number of Ste2*. The confidence ranges between 0 and 1. A value near 0 indicates that either there are few Ste2* or the Ste2* are close to uniformly distributed. Theoretically, a confidence of 1 indicates all the receptors are bound and located in the same position. For comparison, a cell at equilibrium in a uniform pheromone concentration of 6.9nM, the confidence is $\approx 6.4 \times 10^{-3}$ (data not shown). The cell's confidence is maximum around 8 min for slow rates and 2 min for fast rates. The confidence fluctuates around an average: about 6.7×10^{-3} for slow binding rates (Fig. 3.13A) and about 7.6×10^{-3} for fast binding rates (Fig. 3.13B).

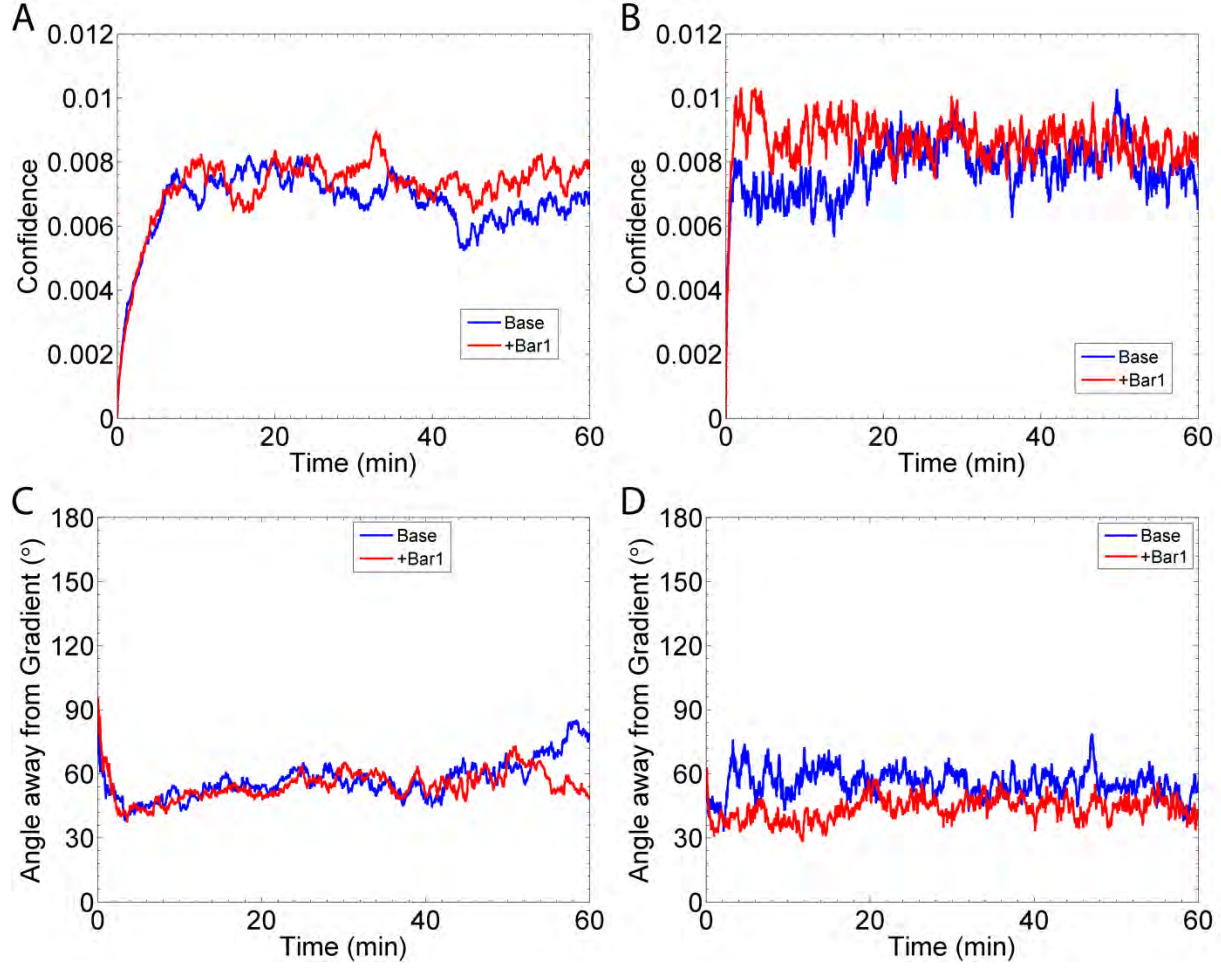


Fig. 3.13: Accuracy and Confidence during Gradient Formation

For the simulations of a cell in an emerging gradient of $0.1 \text{ nM}/\mu\text{m}$, we analyze the Ste2^* distribution over time by calculating the average Ste2^* position: $\langle \text{Ste2}^* \rangle$. The direction of the vector indicates the cell's best estimate of the gradient's direction, and the scaled magnitude determines the cell's "confidence" of that estimate (Eq 3.9). The curves labeled "Base" come from simulations, in which we model only the binding and unbinding reactions between pheromone and Ste2 . The curves labeled "Bar1" come from simulations, in which, we additionally include the Bar1-mediated catalysis of pheromone (Section 2.8). Panels (A) & (B) plot the population-averaged confidence. Panels (C) & (D) plot the population average of the angle between $\langle \text{Ste2}^* \rangle$ and the gradient. (A) Confidence in a system with slow reaction rates ($k_{on} = 1.6 \times 10^5 \text{ (M} \cdot \text{s)}^{-1}$ and $k_{off} = 0.0011 \text{ s}^{-1}$). The Base data comes from the same twenty simulations as in Figure 3.12A, and the Bar1 data comes from sixteen simulations. (B) Confidence in a system with fast reaction rates ($k_{on} = 1.6 \times 10^6 \text{ (M} \cdot \text{s)}^{-1}$ and $k_{off} = 0.011 \text{ s}^{-1}$). The Base model data comes from the same eight simulations as in Figure 3.12B, and the Bar1 data comes from sixteen simulations. (C) Error in Gradient Sensing estimate for slow reaction rates. (D) Error in Gradient Sensing estimate for fast reaction rates.

In the case of slow reaction rates, we find that soon after the cell's confidence is maximized: 8 – 20 min (Fig. 3.13A), an average cell can detect the gradient within 60° of

accuracy (Fig. 3.13C). In the case of fast reaction rates, an average cell can detect the gradient within 90° , between 2 – 20 min after initially binding pheromone (Fig. 3.13D).

3.5.3 Bar1 Improves Gradient Sensing for Fast Reaction Rates

It has been shown that the pheromone protease Bar1 can improve the gradient sensing ability of the cell by locally sharpening the gradient [4,5]. We model the Bar1 concentration as a static field. That is, the Bar1 concentration is a function of the distance from center of the cell (see Section 2.8). Based on previous models, we set the concentration of Bar1 at the cell surface to be 0.85nM [5]. Away from the cell surface, this concentration decreases as $\frac{1}{r}$. At each time step, pheromone molecules have a probability of being catalyzed based on the local concentration of Bar1. We use a catalytic rate of $k_{\text{cat}} = 2.5 \times 10^8 \text{ (M}\cdot\text{s)}^{-1}$, which is based on a previous model [5].

In the case of slow reaction rates, neither the cell's confidence (Fig. 3.13A) nor accuracy (Fig. 3.13C) significantly improves. However, in the case of fast reaction rates, the cell's confidence (Fig. 3.13B) and accuracy (Fig. 3.13D) improve during the first 20 min of gradient formation. The cell's confidence improves to $\approx 8.9 \times 10^{-3}$ with Bar1, as compared to 6.4×10^{-3} without Bar1 (Fig. 3.13B). Furthermore, an average cell can accurately detect the 0.1 nM/ μm gradient within 90° , but with the inclusion of the Bar1 field, the accuracy improves to within 40° (Fig. 3.13D). These results indicate that the true value of the reaction rates may determine how effective Bar1 is in improving gradient sensing. That is, if the true binding and unbinding rates are slow, as measured [8,12], then Bar1 may not be very effective in helping an isolated cell determine the direction of a shallow gradient.

3.6 Discussion

3.6.1 Mechanisms for Noise-Reduction

One potential mechanism for detecting chemical gradients is for cells to use the spatial distribution of active receptors. However, fluctuations in receptor occupancy and receptor diffusion introduce significant uncertainty in receptor occupancy, making this task more difficult. For a cell attempting to sense a shallow gradient, this uncertainty can mask the signal. For example, a cell with 10000 receptors in a shallow gradient of $0.1 \text{ nM}/\mu\text{m}$ centered at the K_D of the receptor, has a difference in occupancy between the front and back of the cell of $\approx 45 \pm 50$ [3]. This estimate does not include the effect of receptor diffusion, which further reduces the difference in receptor occupancy. Recent work predicts the extent to which time-averaging can improve gradient sensing [17]. Simulations of our particle-based stochastic reaction-diffusion model corroborate these predictions (Figs. 3.3 & 3.4). Based on the measured rates, $k_{\text{on}} = 1.6 \times 10^5 \text{ (M}\cdot\text{s)}^{-1}$ and $k_{\text{off}} = 0.0011 \text{ s}^{-1}$ [8,12], we find a yeast cell must wait 40 min or longer to appreciably reduce fluctuations in receptor occupancy. Since yeast cells typically polarize and initiate growth within 20 – 30 min of exposure to pheromone, it is unlikely a cell can sense a gradient as shallow as $0.1 \text{ nM}/\mu\text{m}$ by using only time-averaging. Recent estimates of the reaction rates are an order of magnitude smaller than the rates above [13], which further reduces the likelihood that a yeast cell can sense shallow gradients by time-averaging alone.

Receptor endocytosis has also been suggested as a noise-reduction mechanism, because it removes ligand from the environment thereby eliminating noise from re-binding the same ligand molecule [6]. It may also serve to reduce noise due to receptor diffusion. To investigate the effects of endocytosis on gradient sensing, we adapted our model to absorb rather than release pheromone molecules. Contrary to expectations, we do not find any advantage to removing the

ligand. That is, ligand absorption does not reduce the noise as compared to releasing ligand (Fig. 3.5). In their proposal of this mechanism, the authors assume every ligand molecule that encounters the cell is immediately captured or bound [6]. Whereas, other mathematical models account for non-diffusion-limited binding and unbinding dynamics [18,19]. Our results, which best match the theory of Berezhkovskii and Szabo, indicate these effects from stochastic binding and unbinding are bigger sources of noise than the stochastic arrival of ligand molecules [17].

The pheromone protease, Bar1, has also been suggested as a possible mechanism to improve gradient sensing. We implicitly model Bar1 as a concentration field, which radially decays away from the cell's surface (Section 2.8). Our results indicate that Bar1 can improve the cell's ability to sense an emerging gradient (Figs. 3.13B & 3.13D). Specifically, during the first 20 min that a shallow gradient ($0.1 \text{ nM}/\mu\text{m}$) forms around the cell, the presence of Bar1 increases the cell's accuracy and confidence in estimating the direction of the gradient. Curiously, our results predict this advantage only occurs if the reaction rates for Ste2 and pheromone are fast ($k_{\text{on}} = 1.6 \times 10^6 \text{ (M}\cdot\text{s)}^{-1}$ and $k_{\text{off}} = 0.011 \text{ s}^{-1}$). This advantage may also be dependent upon other parameters, e.g. the concentration of Bar1 near the cell and the catalytic rate of Bar1 on pheromone. Future work will be dedicated to studying how these parameters affect gradient sensing.

3.6.2 Gradient Sharpening due to Steric Effects

Because we directly simulate diffusion, our method captures subtle effects in the distribution of pheromone, e.g. boundary effects from the cell membrane. Because the cell is impermeable to pheromone, it acts to sharpen the gradient (Figs. 3.6 & 3.11). Moreover, this sharpening is reflected in the active receptor distribution (Figs. 3.7 & 3.12). Our simulation methods are analogous to some *in vivo* yeast gradient sensing experiments, in which a

pheromone gradient is established by flowing two different concentrations of pheromone into a microchamber [5,21–24]. In those experiments, it is not uncommon for multiple cells to be adjacent, e.g. as mother-daughter pairs or as multi-cell clusters. Our results suggest these adjacent cells experience a sharper than expected gradient. The effect is strongest if pheromone is flowing into the microchamber from only one side, and the opposite side has no pheromone (i.e. a completely absorbing boundary condition). Additional steric effects may also be present due to the microfluidic chambers themselves. For example, many microfluidic chambers have a height similar to a yeast cell ($\sim 5\mu\text{m}$). Therefore, the presence of a cell severely impedes the flow of pheromone in the chamber and further sharpens the gradient (data not shown). Hence *in vivo* experiments which aim to test the limit of gradient sensing (i.e. the shallowest gradient a cell can detect) must carefully consider steric effects from the cell itself, any adjacent neighbors and the experimental tools. These effects alter the gradient, such that the true gradient experienced by the cells is sharper than expected.

REFERENCES

1. Moore TI, Chou C-S, Nie Q, Jeon NL, Yi T-M (2008) Robust spatial sensing of mating pheromone gradients by yeast cells. *PLoS One* 3: e3865.
2. Chou C-S, Bardwell L, Nie Q, Yi T-M (2011) Noise filtering tradeoffs in spatial gradient sensing and cell polarization response. *BMC Syst Biol* 5: 196.
3. Lauffenburger DA, Linderman JJ (1996) *Receptors: Models for Binding, Trafficking, and Signaling*. Second. Cary, NC, USA: Oxford University Press. 376 p.
4. Andrews SS, Addy NJ, Brent R, Arkin AP (2010) Detailed simulations of cell biology with Smoldyn 2.1. *PLoS Comput Biol* 6: e1000705.
5. Jin M, Errede B, Behar M, Mather W, Nayak S, et al. (2011) Yeast dynamically modify their environment to achieve better mating efficiency. *Sci Signal* 4: ra54.
6. Endres RG, Wingreen NS (2008) Accuracy of direct gradient sensing by single cells. *Proc Natl Acad Sci U S A* 105: 15749–15754.
7. Endres RG, Wingreen NS (2009) Accuracy of direct gradient sensing by cell-surface receptors. *Prog Biophys Mol Biol* 100: 33–39.
8. Jenness DD, Burkholder AC, Hartwell LH (1986) Binding of alpha-factor pheromone to *Saccharomyces cerevisiae* a cells: dissociation constant and number of binding sites. *Mol Cell Biol* 6: 318–320.
9. David NE, Gee M, Andersen B, Naider F, Thorner J, et al. (1997) Expression and purification of the *Saccharomyces cerevisiae* alpha-factor receptor (Ste2p), a 7-transmembrane-segment G protein-coupled receptor. *J Biol Chem* 272: 15553–15561.
10. Dube P, Konopka JB (1998) Identification of a polar region in transmembrane domain 6 that regulates the function of the G protein-coupled alpha-factor receptor. *Mol Cell Biol* 18: 7205–7215.
11. Yi T-M, Kitano H, Simon MI (2003) A quantitative characterization of the yeast heterotrimeric G protein cycle. *Proc Natl Acad Sci U S A* 100: 10764–10769.
12. Bajaj A, Celić A, Ding F-X, Naider F, Becker JM, et al. (2004) A fluorescent alpha-factor analogue exhibits multiple steps on binding to its G protein coupled receptor in yeast. *Biochemistry* 43: 13564–13578.
13. Ventura AC, Bush A, Vasen G, Goldín M a, Burkinshaw B, et al. (2014) Utilization of extracellular information before ligand-receptor binding reaches equilibrium expands and shifts the input dynamic range. *Proc Natl Acad Sci U S A* 111: E3860–E3869.
14. Hicke L, Riezman H (1996) Ubiquitination of a yeast plasma membrane receptor signals its ligand-stimulated endocytosis. *Cell* 84: 277–287.

15. Hicke L, Zanolari B, Riezman H (1998) Cytoplasmic tail phosphorylation of the γ -factor receptor is required for its ubiquitination and internalization. *J Cell Biol* 141: 349–358.
16. Berg HC, Purcell EM (1977) Physics of chemoreception. *Biophys J* 20: 193–219.
17. Berezhkovskii AM, Szabo A (2013) Effect of ligand diffusion on occupancy fluctuations of cell-surface receptors. *J Chem Phys* 139: 121910.
18. Bialek W, Setayeshgar S (2005) Physical limits to biochemical signaling. *Proc Natl Acad Sci U S A* 102: 10040–10045.
19. Kaizu K, de Ronde W, Paijmans J, Takahashi K, Tostevin F, et al. (2014) The Berg-Purcell Limit Revisited. *Biophys J* 106: 976–985.
20. Sun SX (2014) How Accurately Can a Single Receptor Measure Ligand Concentrations? *Biophys J* 106: 778–779.
21. Dixit G, Kelley JB, Houser JR, Elston TC, Dohlman HG (2014) Cellular Noise Suppression by the Regulator of G Protein Signaling Sst2. *Mol Cell* 55: 85–96.
22. Dyer JM, Savage NS, Jin M, Zyla TR, Elston TC, et al. (2013) Tracking shallow chemical gradients by actin-driven wandering of the polarization site. *Curr Biol* 23: 32–41.
23. Hao N, Nayak S, Behar M, Shanks RH, Nagiec MJ, et al. (2008) Regulation of Cell Signaling Dynamics by the Protein Kinase-Scaffold Ste5. *Mol Cell* 30: 649–656.
24. Kelley JB, Dixit G, Sheetz JB, Venkatapurapu SP, Elston TC, et al. (2015) RGS Proteins and Septins Cooperate to Promote Chemotropism by Regulating Polar Cap Mobility. *Curr Biol* 25: 275–285.