Computationally Modeling Dynamic Biological Systems

Katherine Jarvis
University of Maine, katherine.jarvis@maine.edu

Follow this and additional works at: https://digitalcommons.library.umaine.edu/etd

Part of the Biochemistry Commons, Molecular Biology Commons, and the Ordinary Differential Equations and Applied Dynamics Commons

Recommended Citation
https://digitalcommons.library.umaine.edu/etd/3511

This Open-Access Thesis is brought to you for free and open access by DigitalCommons@UMaine. It has been accepted for inclusion in Electronic Theses and Dissertations by an authorized administrator of DigitalCommons@UMaine. For more information, please contact um.library.technical.services@maine.edu.
COMPUTATIONALLY MODELING DYNAMIC BIOLOGICAL SYSTEMS

By

Katherine F. Jarvis

B.S. University of Massachusetts Amherst, 2019

A THESIS

Submitted in Partial Fulfillment of the

Requirements for the Degree of

Master of Science

(in Biochemistry)

The Graduate School

The University of Maine

December 2021

Advisory Committee:

Joshua B. Kelley, Assistant Professor of Biochemistry, Advisor

Andre Khalil, Professor of Biomedical Engineering, Co-Advisor

Robert E. Gundersen, Associate Professor of Biochemistry and Chair of Molecular and Biomedical Sciences

Peter Stechlinski, Assistant Professor of Mathematics
Computationally modeling biological systems helps researchers develop testable hypotheses to enhance our understanding of the system. The simple interplay between individual species in a temporally dynamic environment can be computationally modeled by equation-based or agent-based modeling (ABM). Equation based modeling describes the change in species quantity using ordinary differential equations (ODE) and is dependent on the quantity of other species and predetermined rates of change. These predetermined rates mathematically represent the overall rate of change to mimic dynamic interactions that might not be known in complex unknown systems. ABM tracks each individual agent over time and system progression is driven by equation-based probabilities. Because each agent is kept track of over time, ABM requires a longer computation time than equation-based models but can model multiple heterogeneous processes without the foreknowledge of how they affect the entire system. In one model, we assess the complex interplay between dynamic viral detectability in individuals and testing strategies to find people infected. And in the other model, we examine how post translation modification and protein-protein interactions alter the spatial distribution of signaling receptors. Both models explore developing hypotheses from computational modeling to further understand the dynamics of the system.
DEDICATION

To my loved ones who made sure I stayed sane during my academic career.
ACKNOWLEDGEMENTS

I’d first like to thank my advisor Dr. Joshua Kelley for his guidance and help throughout my graduate studies. He’s always there and excited to answer any question I put to him and has always been there no matter where the research took us. I could not have done this without him.

I’d also like to thank my advisor Dr. Andre Khalil and the whole CompuMaine Lab. Dr. Khalil has always helped me throughout my graduate career to provide helpful advice and discussion. I’d also like to thank the whole CompuMaine Lab for hearing me vent and present my research throughout the years.

I would also like to thank everyone I’ve worked with in the Kelley Lab: Andrew Hart, Will Simke, Sari Mayhue, Cory Johnson, Sudati Shrestha, Nik Hase, Sarah Latario, Phillip Lucus Craig and Loren. This thesis would not have happened without their helpful discussions, insightful comments, and curiosity in the laboratory.

Finally, I would like to thank my family who has always supported my scientific career. Also, I’d like to thank my friends and significant other for supporting me, like visiting me in Maine and helping me practice my scientific presentations.

Text and figures appearing in Chapter 3 have been reprinted from the source
TABLE OF CONTENTS

DEDICATION ......................................................................................................................... iii

ACKNOWLEDGEMENTS ........................................................................................................ iv

LIST OF TABLES ................................................................................................................... viii

LIST OF FIGURES ................................................................................................................ ix

LIST OF EQUATIONS .......................................................................................................... x

LIST OF ABBREVIATIONS .................................................................................................. xi

Chapter

1. INTRODUCTION ............................................................................................................. 1

2. MODEL OF SPATIOTEMPORAL DYNAMICS OF RECEPTOR TRAFFICKING AND MODIFICATION

SUGGESTS ACTIVE RECEPTOR LOCALIZES PERIPHERAL TO THE POLARITY SITE ...................... 5

   Introduction .................................................................................................................... 5

       Gradient Tracking is an essential process commonly modeled .................................

       using the Yeast Pheromone Pathway ................................................................. 5

       Desensitization of GPCR signaling is an essential process to

       amplify G-protein Activation ........................................................................... 5

   Methods ....................................................................................................................... 7

       Agent Based Receptor Model ............................................................................. 7

       Vesicle Trafficking ................................................................................................. 7

       Probability of Binding ......................................................................................... 7

       Yeast Strains ......................................................................................................... 8

       Image Analysis ....................................................................................................... 8

   Results ......................................................................................................................... 9

       Development of a Reaction Diffusion model to examine GPCR dynamics ............ 9
Development of an Agent Based Model of Receptor Dynamics .................................................. 12
ABM Sensitivity Analysis hypothesized RGS effect on Receptor Distribution, verified
Experimentally ........................................................................................................................................ 14
Our Model Suggests that Active Receptor Localize Peripheral to the Site of Secretion
based on GPCR post translational modifications .................................................................................. 19
Discussion .................................................................................................................................................. 23

3. TEMPORAL DYNAMICS OF VIRAL LOAD AND FALSE NEGATIVE RATE INFLUENCE THE LEVELS OF
TESTING NECESSARY TO COMBAT COVID-19 SPREAD ..................................................................... 25

Introduction ............................................................................................................................................. 25
Method ....................................................................................................................................................... 29
Model ......................................................................................................................................................... 29
Symptoms ................................................................................................................................................... 31
Recovery .................................................................................................................................................... 33
Probability of Viral Spread ......................................................................................................................... 33
Testing ......................................................................................................................................................... 34
Contact Tracing ......................................................................................................................................... 35
Results ......................................................................................................................................................... 35
Nonuniform false negative rates can delay detection of infected individuals ........................................ 37
An undetectable period leads to high apparent false negative rates ......................................................... 38
Superspreader events increase the variability in the spread of COVID-19 ............................................. 45
Discussion ................................................................................................................................................... 47

4. CONCLUSIONS AND FUTURE DIRECTIONS ..................................................................................... 50

BIBLIOGRAPHY .......................................................................................................................................... 53
Appendix A. Supplemental Information Tables for MODEL OF SPATIOTEMPORAL DYNAMICS OF RECEPTOR TRAFFICKING AND MODIFICATION SUGGESTS ACTIVE RECEPTOR LOCALIZES PERIPHERAL TO THE POLARITY SITE

Appendix B. Supplemental Information MODEL OF SPATIOTEMPORAL DYNAMICS OF RECEPTOR TRAFFICKING AND MODIFICATION SUGGESTS ACTIVE RECEPTOR LOCALIZES PERIPHERAL TO THE POLARITY SITE

Appendix C. Supplemental Information MODEL OF SPATIOTEMPORAL DYNAMICS OF RECEPTOR TRAFFICKING AND MODIFICATION SUGGESTS ACTIVE RECEPTOR LOCALIZES PERIPHERAL TO THE POLARITY SITE

BIOGRAPHY OF THE AUTHOR
LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Table A.1</td>
<td>Agent Based Receptor Model Parameters</td>
<td>60</td>
</tr>
<tr>
<td>Table A.2</td>
<td>GTP ODE Parameters</td>
<td>61</td>
</tr>
<tr>
<td>Table B.1</td>
<td>Yeast Strains</td>
<td>62</td>
</tr>
<tr>
<td>Table B.2</td>
<td>Plasmids</td>
<td>62</td>
</tr>
<tr>
<td>Table B.3</td>
<td>Primers</td>
<td>62</td>
</tr>
<tr>
<td>Table C.1</td>
<td>ODE Receptor Model Parameters</td>
<td>65</td>
</tr>
</tbody>
</table>
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure 1.1</td>
<td>Mathematical Modeling SIR model vs. Molecular Modeling</td>
<td>3</td>
</tr>
<tr>
<td>Figure 2.1</td>
<td>ODE Receptor Model of RGS’s effect of G-Protein</td>
<td>11</td>
</tr>
<tr>
<td>Figure 2.2</td>
<td>ABM of Receptor Modification</td>
<td>13</td>
</tr>
<tr>
<td>Figure 2.3</td>
<td>Membrane Trafficking Modeling is based on Experimental Actin Localization</td>
<td>15</td>
</tr>
<tr>
<td>Figure 2.4</td>
<td>Sensitivity Analysis of RGS Unbinding Probability on Receptor Dynamics</td>
<td>17</td>
</tr>
<tr>
<td>Figure 2.5</td>
<td>Ste2-pHfluorin (GPCR) Polarization with and without RGS binding Receptor</td>
<td>19</td>
</tr>
<tr>
<td>Figure 2.6</td>
<td>Average Active Pheromone Localizes Peripheral to the Secretion Site</td>
<td></td>
</tr>
<tr>
<td></td>
<td>In Computational ABM</td>
<td>21</td>
</tr>
<tr>
<td>Figure 3.1</td>
<td>Dynamic SARS-CoV-2 Viral Load and False Negative Rate</td>
<td>27</td>
</tr>
<tr>
<td>Figure 3.2</td>
<td>Stochastic Agent Based Epidemiological Model</td>
<td>30</td>
</tr>
<tr>
<td>Figure 3.3</td>
<td>Stochastic Agent Based Epidemiological Model Parameters</td>
<td>31</td>
</tr>
<tr>
<td>Figure 3.4</td>
<td>Effects of False Negative Rates on Detection</td>
<td>36</td>
</tr>
<tr>
<td>Figure 3.5</td>
<td>High asymptomatic transmission and dynamic false negative rate lead</td>
<td></td>
</tr>
<tr>
<td></td>
<td>to more required testing to prevent a COVID-19 outbreak</td>
<td>40</td>
</tr>
<tr>
<td>Figure 3.6</td>
<td>Masking allows fewer tests and lower frequencies of testing</td>
<td>42</td>
</tr>
<tr>
<td>Figure 3.7</td>
<td>Assumptions about Mild COVID Cases Isolating Influence Testing</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Requirement</td>
<td>44</td>
</tr>
<tr>
<td>Figure 3.8</td>
<td>Superspreading leads to more variability between discrete simulations</td>
<td>46</td>
</tr>
</tbody>
</table>
LIST OF ABBREVIATIONS

RGS- Regulator of G-Protein Signaling

GPCR- G-Protein Coupled Receptor

SIR- ‘Susceptible Infectious Recovered’ Epidemiology Model

ABM- Agent Based Modeling

ODE- Ordinary Differential Equations

MAPK- Mitogen-activated protein kinase

Bem1- Bud Emergence Protein 1

Ste2- Pheromone alpha factor Receptor

Gα- Gpa1

Ede1- EH domain-containing and endocytosis protein 1

Exo84- Exocyst complex component EXO84

Rt- Reproduction Number
CHAPTER 1

INTRODUCTION

Biological systems in nature are known to be highly dynamic and complex. Biological experiments are designed to answer hypotheses to understand the mechanisms of the system. Proposed mechanisms are often represented with simple visualizations to enhance our working understanding of the system. Then, the system can be further understood through description of biological observations to mathematic equations or mathematic modeling. Computationally simulating these mathematical models help researchers to propose new hypotheses about the dynamics of the system, also known as computational models (Fischer, 2008). For observed biological properties to emerge in a computational model, it is necessary to include all essential dynamics to mimic the system. A variable removal that doesn’t affect the overall system allows the hypothesis of that variable in biology may contribute to a different process. The system’s hypothesized dynamics based on computational modeling can then be tested using experimental approaches.

In biochemistry, we mathematically model molecular species undergoing a chemical reaction as mass action kinetics. Mass action kinetics of a species can be written as a simple differential equation and a system of mass action kinetics can be described as ordinary differential equations (ODE) (Ingalls, 2018a). ODEs apply a predetermined rate of change to a summed amount of the species through compartmental modeling. Unfortunately, mass action kinetics assumes the reaction has a high concentration of species and is well mixed, assumptions that are often not true in molecular biology due to the discrete number of molecules in the system and the lack of spatial details (Soheilypour & Mofrad, 2018). Another computational modeling method for mass action diffusion molecular modeling is agent-based modeling (ABM). ABM simulates interactions between multiple independent individuals or ‘agents’ and assesses their effect on the overall system (Soheilypour & Mofrad, 2018). ABM gives flexibility because it integrates simple individual based conditions. Unfortunately, iterating through
individuals heightens the computational cost (Glen et al., 2019). Both ODE or ABM methods can be used to model molecular reactions with the help of simulating the basic principles of diffusion and mass action kinetics. Computationally modeling both the mass action and the diffusion is called Reaction-Diffusion Modeling and it reveals emergent properties of the system.

Reaction diffusion models can be applied to cellular biology to relieve the constraints of investigation of molecular dynamics only through experimentation (Ingalls, 2018b). In yeast cellular biology, the MAPK pathways have been extensively modeled to understand their receptor mediated signaling as well as their crosstalk with other molecular signaling pathways (Hughey et al., 2010; Pappalardo et al., 2016; Schoeberl et al., 2002). The MAPK pathway in the yeast pheromone pathway consists of a cascade of MAPK kinases that phosphorylate the subsequent MAPK kinase. This signaling cascade can be modeled spatiotemporally through ODE modeling to quantify the change in concentration of each protein species based on other species concentration. Molecular cascades can also be modeled through ABM by simulating each individual molecule and probability-based reactions. Both ODE and ABM methods are easily implementable for molecular modeling to answer hypotheses of MAPK signaling dynamics.

Computational biological models are useful outside of biochemistry to develop hypotheses in a variety of applications (Fischer, 2008). Epidemiology computational models are a tool to study the spread of disease in a population to potentially make public health decisions. ‘Susceptible infected recovered’ or SIR models describe the transmission of disease through time in a constant population. Due to the often-large population size of epidemiological models, ODE modeling methods are often utilized because compartmental models are computationally efficient. An example of compartmental modeling is determining what percentage of the susceptible population becoming infected using simple nonlinear ODEs and the population transmission rate (Hritonenko et al., 2020). Alternatively, individual-based transmission rates are used for epidemiological ABM to determine the probability of one
individual infecting another. These models can be used to hypothesize the spatial spread of COVID-19 as well as the effects of preventative measures to slow the spread of disease (Hritonenko et al., 2020; Mohamadou et al., 2020). Computational modeling epidemiology dynamics is very similar to molecular based modeling because they can use similar modeling methodologies (Figure 1.1).

![Figure 1.1 Mathematical Modeling SIR model vs. Molecular Modeling](image)

Here we develop computational models for two different biological systems. First, we develop a computational model of G-protein Coupled Receptor (GPCR) signaling and how its desensitization mechanisms alter G-Protein activity with relation to the external gradient. Both ODE and ABM methodologies were used to examine the spatiotemporal effects of desensitization cellular processes in conjunction with experimental approaches to answer hypotheses. Next, we model viral infection outbreak on a college campus over time and the effect of dynamic viral transmissivity on reducing the infection rate on campus. Our agent-based epidemiological model showed dynamic transmissivity and false negative testing influences the amount of testing required on campus to prevent an outbreak. This
thesis takes a deeper look at testing hypotheses proposed by computational models to understand the complex biological system.
CHAPTER 2
MODEL OF SPATIOTEMPORAL DYNAMICS OF RECEPTOR TRAFFICKING AND MODIFICATION SUGGESTS
ACTIVE RECEPTOR LOCALIZES PERIPHERAL TO THE POLARITY SITE

Introduction

Gradient Tracking is an essential process commonly modeled using the Yeast Pheromone Pathway

A cell’s ability to track and polarize in the direction of extracellular signal is an essential cellular process, commonly utilized by leukocytes finding invading pathogens in the human body (Kolaczkowska & Kubes, 2013; Lakhani & Elston, 2017). Detecting the direction of signal is the job of surface receptors such as G-protein coupled receptors (GPCR) that bind the extracellular chemoattractant (Weiner, 2002). Downstream effector proteins of GPCR signaling orient cytoskeletal elements towards the signal to facilitate growth (Bi & Park, 2012). A common model to study GPCR gradient tracking is S cerevisiae or baker’s yeast undergoing the yeast pheromone pathway (Arkowitz, 2009; Segall, 1993).

The yeast pheromone pathway is initiated by extracellular pheromone or alpha factor binding GPCR, STE2 (Jackson & Hartwell, 1990). Heterotrimeric G-Protein consisting of subunit Gα (Gpa1), and Gβγ is bound to the GPCR until activation (Wang & Dohlman, 2004). Active GPCR acts as a GEF to Gα subunit, causing dissociation of the G-Protein complex (Dohlman & Thorner, 2001). Free Gα and Gβγ activate cascades of effector proteins to initiate chemotrophic growth forming a shmoo projection (Arkowitz, 2009; Moore et al., 2013). The spatial properties of GPCR signaling effects downstream signaling processes such as polarization and morphogenesis (Hao et al., 2008; Valtz et al., 1995).

Desensitization of GPCR signaling is an essential process to amplify G-protein Activation

GPCR signaling negative regulation, or desensitization, amplifies the gradient of active G-protein to polarize downstream effector cascades to help gradient tracking accuracy (Henderson et al., 2019a). GPCR signal can be desensitized through either GPCR internalization or RGS-directed G-Protein inactivation.
Removal of active GPCR signal from the membrane through internalization is a receptor desensitization mechanism because it reduces the amount of GPCR signal. GPCR internalization require many time-dependent post translational modifications including phosphorylation and ubiquitination. Pheromone activated GPCR promotes its phosphorylation and thus internalization (Hicke et al., 1998a). Receptor polarization requires receptor internalization and not receptor secretion in the direction of signal (Suchkov et al., 2010a). GPCR internalization is regulated through cargo-dependent clathrin coated pits (Pedersen et al., 2020). The spatiotemporal dynamics of GPCR post translational modifications, endocytic pit distribution and their effect on internalization localization is unknown.

The Regulator of G Protein Signaling (RGS) is another GPCR desensitization mechanisms because it negatively regulates the G-protein complex (Dohlman et al., 1995; Segall, 1993). After activation of Go by pheromone bound receptor, RGS acts as a GAP to Go to deactivate the complex and desensitize the pathway (Ballon et al., 2006a; Hague et al., 2005; Kovoor et al., 2005 Apanovitch 1998). RGS and GPCR interact directly to presumably enhance RGS GAP activity (Hague et al., 2005; Kovoor et al., 2005; Ballon et al., 2006). RGS binding GPCR promotes receptor retention from direct inhibition of GPCR phosphorylation (Venkatapurapu et al., 2015a). The dynamic between RGS binding causing GPCR retention but RGS activity negative regulating G-protein activation has an unknown effect on the spatiotemporal dynamics of receptor mediated signal.

Here, we examine the desensitization mechanisms of GPCR signaling and its effects on G-Protein spatiotemporal activation using computational modeling and experimental approaches. A previous model determined receptor retention due to RGS binding is a required dynamic for modeling the spatial effects of receptor mediated signaling on G-protein activity. Because of the unknown effect of receptor retention on the endocytosis distribution, we used an agent-based model (ABM) to incorporate the spatiotemporal dynamics of individual GPCR post translational modifications to determine internalization location. The model describes receptor undergoing time dependent modifications during
GPCR signaling to determine localization of active G-protein. From this, we found a bifurcated pattern of G-protein activation peripheral to the polar cap, dividing the active signal to possibly enable competition and enhance wandering to improve gradient tracking. This result led us to explore RGS driven polarization based on receptor activity. Experimentally we verify RGS binding to receptor enables receptor retention and regulates receptor activity. These data show RGS regulates GPCR polarization by modulating the internalization of the receptor depending on receptor activity.

**Method**

**Agent Based Receptor Model**

This model is made up of receptor of GPCR proteins that bind pheromone, are phosphorylated and ubiquitinated. Each GPCR is individually tracked over time in a uniform spatially discretized 1D space. Uniform inactive receptors are initialized equidistant on the membrane and ligands bind according to probabilities determined by their $k_{\text{cat}}$ rate listed in Table A.1. New inactive GPCRs are delivered to the predetermined location of polarity and the number of receptors delivers is half the maximum number of receptors per bin. The maximum number of receptors per bin was determined based on the length of a bin and the diameter of Ste2 GPCR (Gurevich & Gurevich, 2008). GPCRs are internalized through endocytic pits. Endocytic pits are deposited where the most about of ligand-bound receptors are located and have a lifetime of 1 to 5 minutes (Pedersen et al., 2020) and density and distribution of endocytic pits described in Results section.

**Vesicle Trafficking**

Vesicle trafficking rates were based on Carrillo et al., 2015 rates and altered to maintain a consistent membrane size and receptor polarization (listed in Table A.1). Vesicle trafficking events occur uniformly according to the rate listed in Table A.1. Vesicle trafficking method in the computational model is described in the Results section.
**Probability of Binding**

All probabilities used are listed in Table A.1. Binding probabilities were converted from either disassociation rates and/or $k_{cat}$ rates. Receptor binding pheromone is known to be $0.185 \text{s}^{-1} \text{uM}^{-1}$ and converted to a probability by multiplying by the timestep and concentration of pheromone per bin and normalizing based on the maximum number of receptors per bin. Phosphorylation probability was based on the $k_{cat}$ rate and normalized by the timestep and assuming one enzyme is available to phosphorylate each receptor for each timestep. Ubiquitination probability was calculated through the half-life of Ste2 during the pheromone response (Hicke et al., 1998b) and converted to probability by normalizing to the timestep.

**G-Protein Activation**

Receptor post translational modification and ligand binding regulate G-Protein activation. How the receptor regulates G-protein spatially was modeled through ODEs and using the distribution of pheromone bound receptors as GEF parameter and RGS bound receptors distribution as GAP parameter. G-protein is initialized in a uniform off state discretely modeled in a uniform distribution. Parameters and listed in Appendix A in Table A.1. Diffusion of G-protein is modeled through Fast Fourier transformation and ODE are modeled though MATLAB function ode23. Equations are listed in Appendix A.

**Yeast Strains**

Yeast Strains in this study are shown in Table A.2 Stain List in Appendix A. All yeast strains used have a background of BY4741 MATa haploid Saccharomyces cerevisiae. Proteins were tagged with eGFP, pHluorin and mRuby2 using through oligonucleotide-directed homologous recombination with using the primers listed in Table B.3. Primers were designed to be used with pFA6a plasmids to tag proteins in yeast strains listed in Table B.2. The yeast strains were transformed using the standard lithium acetate yeast transformation protocol. Individual colonies were isolated by growth on selection plates.
The cells were grown in a rich media (YPD) at 30°C overnight through serial dilution. Cultures were grown to mid-log phase (OD600 = 0.6), the cells were further diluted and treated with pheromone (100nM, 500nM and 1uM) for 90 minutes. The cells were then fixed using 4% PFA, permeabilized with Triton X-100, and mounted using polylysine and Vectashield. Microscopy was performed using an IX83 (Olympus, Waltham MA) microscope with a Prime 95B CMOS Camera (Photometrics). Fluorescence and Differential Interference Contrast (DIC) images were acquired using an Olympus-APON-100X-TIRF oil immersion objective and magnified using 2X magnifier. Z-stacks of GFP and RFP images were acquired using an Xcite 120 LEDBoost (Excilites).

**Image Analysis**

Unless otherwise specified, images were deconvolved using Huygens Software (Scientific Volume Imaging, Hilversum, Netherlands) Classic Maximum Likelihood Estimation (CMLE) Deconvolution Algorithm. Masks of cells of interest were made using ImageJ (Schindelin et al., 2012). Data analysis to quantify the profiles was performed using MATLAB (MathWorks, Natick, MA). Protein fluorescent intensity was calculated using a line width of 5 pixels along the membrane and averaged. Proteins profiles were centered according to itself unless otherwise specified. Profiles that report fraction of protein at each position are normalized to sum to one. The shaded region of normalized fluorescent intensity shows the 95% confident intervals derived by bootstrapping 1000 resampling.

**Results**

**Development of a Reaction Diffusion model to examine GPCR dynamics**

Receptor mediated signaling provides a spatial signature to activate downstream effector proteins of the yeast pheromone response. Yeast cells undergoing the mating pheromone response can accurately detect shallow gradients of pheromone by GPCR desensitization mechanisms steepening the intracellular gradient. We wanted to know how the receptor location on the membrane changes due to desensitization mechanisms and hypothesize how these mechanisms effect the receptor dynamics
spatially and modulate G protein activation to steepen the intracellular gradient. To test this, we
developed computational models to simulate spatiotemporal GPCR signaling during the yeast
pheromone response. This simulation includes the receptor binding ligand (pheromone), then activating
intracellular G-protein polarization on receptor mediated signaling. Desensitization of the pheromone
pathway is mediated by either RGS action to turn off the Gα of the G-protein complex, or receptor
internalization, reducing the amount of membrane-bound signaling receptor (Dohlman et al., 1995;
Segall, 1993; Suchkov et al., 2010a). A computational model of receptor dynamics will allow us to test
hypotheses about the effects of desensitization mechanisms on spatial signaling.

The Kelley lab has previously attempted to model spatial GPCR signaling and G-protein
activation using a deterministic ODE computational model. The goal of the model was to simulate
receptor mediated signaling regulation by desensitization mechanisms and output spatial G-protein
activity (Figure 2.1A). Our biological work shows that RGS phosphorylation effects the location of active
G-protein, causing MAPK signaling to broaden (Simke et al., 2020). We wanted to know whether either
desensitization mechanism caused the observed broadening of G-protein activity in response to RGS
phosphorylation. The mathematical model schematic can be seen in Figure 2.1A, details of the model
are included in Appendix B. The deterministic ODE model was not able to capture the experimental data
profiles and the MAPK asymmetrical polarization (Figure 2.1D and Figure 2.1E). This model did not
consider the dynamic spatial effects of endocytosis localization due to RGS activity modulating GPCR
post translational modification and causing receptor retention. To overcome these obstacles, we
developed a stochastic agent-based model (ABM) of receptors to track GPCR post translational
modifications and each vesicle trafficking event.
Figure 2.1 ODE Receptor Model of RGS’s effect of G-Protein

Diagram of the computational ODE based model simulating GPCR signaling, activating subunits $\alpha$ and $\beta\gamma$ and negatively regulated by RGS (A). (B) Scaled distributions of vesicle trafficking to fit $\alpha$-GTP to the experimental wildtype MAPK distribution (C). From there, to compare to mutant RGS MAPK profiles, RGS activity was altered tenfold increase and decrease (D). A similar method was applied to the frequency of endocytosis events (E). B-E are aligned to the polar cap as site of secretion. Profiles were normalized with minimum subtracted.
Development of an Agent Based Model of Receptor Dynamics

The post translational modifications to receptor are critical for modulation of receptor activity and lifetime. Receptor activity is modulated by RGS-GPCR interaction (Hague et al., 2005; Kovoor et al., 2005; Ballon et al., 2006). In addition to RGS acting as a GAP to the G-Protein complex, RGS-GPCR interaction increases receptor levels on the plasma membrane, presumably through blocking GPCR phosphorylation (Venkatapurapu et al., 2015a). The ability of the RGS to modulate receptor signaling and receptor lifetime led us to hypothesize the localization of active receptor is dependent on post translational modification and RGS-GPCR interaction. Theses modification dynamics to receptor were incorporated into the agent-based receptor model to determine the localization of active receptor.

The model works as follows (diagrammed in Figure 2.2): New receptors are secreted at the polarity site and internalized through endocytic pits (described in Figure 2.3B). The polarity site is always considered to be the center of the model, as we are trying to model relative to the center of the polarity site. Inactive receptor can bind pheromone and become active. Active receptor can be phosphorylated six times (Toshima et al., 2009), then finally ubiquitinated, which will mark it for endocytosis (Figure 2.2A and Figure 2.2B). Receptor that has bound RGS cannot be phosphorylated, while phosphorylated receptor cannot bind the RGS (Figure 2.2A). The downstream signaling activity is modeled through ODEs to simulate the overall activation of G-protein based on the amount of active receptor, the GAP activity of the RGS bound to receptor, and a baseline GAP activity of cytosolic RGS as described in the methods. The computational ABM iterates through each receptor on the membrane to determine if ligand bind or diffuses to a different location. Ligand binding and post translational modification probabilities are based on reaction rates as described in the Methods. Receptor diffusion spatiotemporal dynamics is simulated by a calculating a movement probability that considers the diffusion rate, length of the bin, change in time and space availability (Figure 2.2C) (Azimi et al., 2011).
Figure 2.2 ABM of Receptor Modification

(A) Receptor Lifetime on the membrane simulated in the model with RGS negatively regulating phosphorylation (B) The Receptor Lifetime on the membrane simulated in the model without RGS (C) Receptor diffusion was modeled by calculating a movement probability and choosing a final location bin using the Reduced probability method (Azimi et al., 2011) ensuring diffusion to the least concentrated bin
Endocytosis of the receptor is an important regulator of receptor mediated signaling. GPCR stimulation induces clathrin-mediated endocytosis. Clathrin-mediated endocytosis forms actin endocytic pits to internalize protein (Merrifield, 2005). To determine how to distribute endocytic events in the model, we examined the actin distribution in shmooing cells during the pheromone response. We found about half of the actin pits are localized on the polar cap which accounts for about 20% of the membrane (Figure 2.3A). The computational model incorporated this observation by ensuring 20% of bins closest to the polar cap consistently have at least half of the endocytic pits. The density of endocytic pits was determined to be 0.97 actin pits per µm. The density of endocytic pits in the computational model required slight upscaling to be proportional to endocytic frequency. Endocytic pits also effect GPCR dynamic because actin deposition restricts diffusion (Weinberg & Puthenveedu, 2019). The computational model simulated this dynamic by inhibiting diffusion of 30% of ubiquitin bound receptors in endocytic pits. Internalization of endocytic pits is modeled based on clathrin mediated endocytosis position dependent site maturation and accounts for the polarized actin patches in shmooing yeast. Clathrin mediated endocytosis site maturation is dependent on a cargo sensitive step causing the endocytic pit lifetime to range from 1 minute to 5 minutes (Pedersen et al., 2020). To model this, endocytic events initialize by selecting the mature endocytic pits (> 1 minute membrane bound). The model then internalizes the mature endocytic pit that has the most ubiquitin bound receptor (Figure 2.3B). All protein in the internalized endocytic pit bin is also internalized and removed from the membrane. Membrane trafficking events are simulated by removing or inserting membrane in the appropriate location and realigning the membrane to the secretion site or point of polarization (Figure 2.3C).
Figure 2.3 Membrane Trafficking Modeling is based on Experimental Actin Localization

(A) Representative fluorescent image of Phalloidin 647 stain in the wildtype strain BY4741 treated with 3µM pheromone for 90 minutes and fixed in 4% PFA. Red denotes polar cap membrane and Blue denotes peripheral membrane. Quantification of actin pit localization, bar graphs represent the mean percentage of the membrane ± SD (n = 7). 50% of the endocytic pits localize to the polar cap which takes up 20% of the total membrane. (B) Diagram of how endocytic pits in the computational model effect GPCR spatiotemporal dynamics. In the computational model endocytic pits are deposited in the location with the most modified receptors and have a lifetime from 1 to 5 minutes (Pedersen et al., 2020). The endocytic pit with the most ubiquitin bound receptors is internalized. Endocytic pits in the computational model inhibit diffusion of 30% of proteins in the endocytic pit. (C) Methodology of vesicle trafficking in the computational model to maintain site of secretion alignment by shifting the membrane by bin length to introduce new or remove membrane. Numbers refer to the distance from the polar cap (µm).
ABM Sensitivity Analysis hypothesized RGS effect on Receptor Distribution, verified Experimentally

RGS binding to the GPCR modulates plasma membrane levels of the receptor, presumably through inhibition of receptor phosphorylation. To test the function of our model, we performed a sensitivity analysis of the RGS unbinding probability to analyze its effects on receptor localization. RGS dynamics in the model are illustrated in Figure 2.3B; receptor phosphorylation is inhibited due to RGS binding receptor. To model the increase internalization due to receptor activation, receptor phosphorylation probability is scaled with respect to the amount of G-protein activated (further details are described in the methods). Because RGS is assumed to be bound to receptor until phosphorylation (Ballon et al., 2006), we chose a low unbinding probability of 1e-100 per 10 msec to show receptor retention effects on the receptor profile and model irreversible binding. The sensitivity analysis iteratively increased the unbinding probability ten-fold, therefore increasing the unbinding rate exponentially. A higher unbinding rate represents a higher proportion of receptor without RGS bound, therefore a higher receptor phosphorylation availability. After simulating each condition for 5 iterations, the results show irreversible binding provides a high proportion of receptors with RGS bound and low proportion of phosphorylated receptors (Figure 2.4). In addition, the irreversible RGS binding shows dispersed phosphorylated receptors throughout the periphery. Whereas lower RGS-GPCR binding promotes a high proportion of the phosphorylated receptor and a very low proportion with RGS bound. Additionally, there was an overall increase in membrane-bound receptors with increased RGS-GPCR interaction in the computational model. To compare this output of the model with the biological system, we next looked experimentally at receptor profiles of yeast undergoing the pheromone response using fluorescence microscopy.
[Figure 2.4 Sensitivity Analysis of RGS Unbinding Probability on Receptor Dynamics]

GPCR Model adds RGS to mitigate receptor retention. Sensitivity analysis of RGS unbinding rate to compare RGS and phosphorylated receptor distribution. RGS binding rate was kept constant and unbinding probability was increased 10-fold ($1 \times 10^{-100}, 1 \times 10^{-10}, 0.1$) to decrease RGS binding receptor. Increasing the unbinding probability increases phosphorylated receptor concentration at the polar cap and decreases RGS bound receptors. 5 iterations for each unbinding probability and the shaded region is 95th CI through bootstrapping 1000 iterations.
To assess the function of the computational model, we compared the computational models’ effects of RGS interaction with receptor distribution to experimental receptor distribution, with and without RGS binding. The GPCR (Ste2) was fluorescently tagged with the pH-sensitive GFP, pHluorin to visualize membrane bound GPCR and cells were treated in pheromone (100nM, 500nM and 1uM) for 90 minutes in wildtype yeast cells (BY4741). To test without RGS binding, we repeated this assay with a mutant yeast strain, sst2<sup>Q304N</sup>, whose RGS is unable to bind GPCR but maintains its GAP activity. We modulated receptor activation through the pheromone concentration to see RGS binding’s effect on gradient tracking. Fluorescent microscopy images show GPCR polarization to the polar cap increases generally as receptor activation increases in wildtype yeast cells (Figure 2.5A). Whereas in the sst2<sup>Q304N</sup> strain, the Ste2-pHluorin membrane bound fluorescence was significantly reduced and changed minimally with increasing receptor activation (Figure 2.5A). These experimental results confirm the computational results of RGS- GPCR interaction causing increased receptor retention therefore more overall membrane-bound receptor signal. Next, we quantified the GPCR polarization by aligning the normalized membrane-bound Ste2-pHluorin profile to itself to determine the fraction of the receptors localize to the polar cap. Wildtype receptor polarization increases as receptor activation increase except for high pheromone, receptor polarization decreases significantly from 4.6% to 3.4% in saturating pheromone conditions (Figure 2.5B). In mutant sst2<sup>Q304N</sup> strain, the polarization decreases significantly in all pheromone conditions compared to its wildtype counterpart (Figure 2.5B). The overall increase in membrane bound receptor fluorescence with RGS binding verifies the computational model results of RGS binding modulating receptor retention.
Our Model Suggests that Active Receptor Localize Peripheral to the Site of Secretion based on GPCR post translational modifications

From the microscopy data, we can clearly see receptor retention in wildtype compared to sst2 Q304N through increased membrane fluorescence and increased receptor polarization. This verifies the computational model’s sensitivity analysis that showed RGS binding modulating the receptor profile through receptor retention. With the computational model functionality verified, we can us the model to help us hypothesize active receptor localization to help us understand how GPCR accurately tracks external gradients. RGS binding receptor reduces receptor activity and receptor phosphorylation inhibits RGS binding, therefore we assumed the phosphorylated receptor to be active. The receptor agent-based model without RGS suggests phosphorylated receptor localize peripheral to the polar cap (Figure 2.6).

Figure 2.5 Ste2-pHluorin (GPCR) Polarization with and without RGS binding Receptor

(A) Representative images of Ste2-pHluorin in Wildtype and Q304N in increasing pheromone concentrations. Strains have been treated with pheromone for 90 minutes and fixed in 4% PFA and imaged using Widefield Microscopy. (B) Ste2-pHluorin fluorescent profiles have been normalized and self-aligned.
Inactive receptor is delivered to the polar cap, requiring the receptors who have been on the membrane for longer to diffuse away from the site of polarization. Modifications to GPCR such as pheromone binding and phosphorylation increase as more time passes and diffuse further from the site of secretion. With no RGS present, active receptors localize from 0.5 to 1 µm away from the site of polarized secretion according to the computational model. This result shows receptor activation may peak at a distance away from the point of polarization based on the temporal dynamics of receptor modification. From the sensitivity analysis in Figure 2.4, the phosphorylated receptor distribution is more uniform and dispersed with more RGS binding but exhibit the same bifurcated pattern. This hypothesized bifurcated pattern distribution from the ABM may contribute to polarization competition of intracellular processes. From this we learn there may be competition introduced through receptor mediated signaling to induce wandering and increase gradient tracking accuracy.
Figure 2.6 Average Active Pheromone Localizes Peripheral to the Secretion Site in the Computational ABM

Computational Model was developed to track localization of receptor with modifications through time to determine the localization of active receptor. The receptor first binds pheromone, then phosphorylated 6 times (individually) and finally ubiquitinated. Average endocytic pit distribution is the background, and they localize a micron or less from the polar cap. Endocytic events occur around one micron from the polar cap, internalizing 5-15 receptors each event.
**Discussion**

GPCR signaling relies on desensitization mechanisms to mitigate G-Protein activity and regulate downstream signaling. RGS downregulates G-protein activity and extends the receptor lifetime by inhibiting phosphorylation of the receptor (Venkatapurapu et al., 2015b). Because the receptors undergo desensitization by time-dependent modifications, modeling this system using ODEs would require incorporation of the receptor’s lifetime emergent properties on the system if modeled, an unknown effect we are trying to figure out. Therefore, ABM can easily be used to model the lifetime of the receptor and modification individual dynamics to determine the emergent properties of the system rather than compartment-based modeling.

By using ABM, we found GPCR signaling leads to a bifurcated active receptor pattern using a spatiotemporal computational stochastic model of receptor modification without RGS dynamics. This result suggests receptors without RGS bound are not localized to the point of polarization (aka secretion) and are localized peripherally instead. This receptor mediated signaling localization may signal localization to downstream elements spatiotemporally such as Rho signaling. In addition, rho signaling is known to be enhanced through competition of multiple polarization sites. The bifurcated pattern of phosphorylated receptor suggests possible competition between downstream elements signaling elements to determine the point of polarization. This competition would enhance the accuracy of polarization and promote wondering due to the blind spot at the site of polarization due to secretion.

This blind spot at the secretion site is broadened from adding RGS to the computational model. To test RGS’s ability to retain membrane-bound receptor, we modulated the RGS-GPCR unbinding probability to see its effect on the overall receptor distribution. This sensitivity analysis showed normal RGS-GPCR binding increasing the overall number of receptors on the membrane and most of the receptor being RGS bound. Whereas low RGS-GPCR binding decreases the amount of receptor membrane-bound because of its increased internalization rate due to low RGS interaction. From this, we
learn the computational model suggests RGS-GPCR interactions regulate spatially receptor post
translational modifications and alter the overall number of receptors on the membrane. To confirm the
computational model, we tested the overall amount receptor with and without RGS-GPCR interaction
and RGS localization experimentally.

We wanted to know the effect of RGS-GPCR interaction on the distribution of membrane-bound receptors experimentally to confirm the computational model qualitatively. Therefore, we analyzed the fluorescence of Ste2- pHluorin with and without RGS-GPCR interaction (wildtype and sst2^{Q304N}). Wildtype receptor fluorescence increased, and polarization increased compared to sst2^{Q304N} receptor fluorescence. This confirmed the computational model dynamic of receptor retention due to RGS-GPCR binding and verified the dynamics of the model reflect biological observation. Interestingly, we learn wildtype receptor polarization is not proportional to receptor activation and is inverted compared to sst2^{Q304N} receptor polarization. This dynamic may suggest there are other positive and negative regulators effecting receptor polarization in addition to RGS.

GPCR’s role in maintaining homeostasis in the cell is essential for life specifically for cardiovascular tissue in humans. Cardiovascular tissue relies on a number of GPCRs such as adrenergic receptors to maintain homeostasis to prevent heart failure onset and progress. Because of this, GPCRs have been pharmacologically targeted to treat heart failure through researching GPCR function and post receptor modification to mediate responses. Receptor modifications include desensitization methods such as inhibitory proteins and internalization. Understanding these desensitization mechanisms help our understanding of how the GPCR systems modulates activation and is applicable to human health.

Future directions of this research include exploring how RGS and RGS-GPCR interaction effects gradient tracking. The bifurcated pattern of receptor activation may lead to competition between downstream elements. This competition promotes wandering and tracking accuracy. By altering the gradient steepness and the offset of peak pheromone, we can determine how RGS helps track a gradient
through receptor mediated signaling. This information can give us insight of how shallow extracellular gradients translate to steep intracellular gradients to help track signal.

This research gives further insight into the importance of spatiotemporal receptor modification and desensitizing receptor mediated signaling is for polarization and G-protein activity. This study explored how post activation modifications of GPCR lead to altered vesicle trafficking and how inhibitory proteins mediate GPCR activity to regulate downstream positive feedback mechanisms. This allows a dose dependent response to stimuli to give the cell further insight into how to polarize toward the signal.
CHAPTER 3

TEMPORAL DYNAMICS OF VIRAL LOAD AND FALSE NEGATIVE RATE INFLUENCE THE LEVELS OF TESTING NECESSARY TO COMBAT COVID-19 SPREAD

Text and figures appearing in this chapter have been reprinted from the source


Colleges and other organizations are considering testing plans to return to operation as the COVID-19 pandemic continues. Pre-symptomatic spread and high false negative rates for testing may make it difficult to stop viral spread. Here, we develop a stochastic agent-based model of COVID-19 in a university sized population, considering the dynamics of both viral load and false negative rate of tests on the ability of testing to combat viral spread. Reported dynamics of SARS-CoV-2 can lead to an apparent false negative rate from ~17 to ~48%. Nonuniform distributions of viral load and false negative rate lead to higher requirements for frequency and fraction of population tested in order to bring the apparent Reproduction number (Rt) below 1. Thus, it is important to consider non-uniform dynamics of viral spread and false negative rate in order to model effective testing plans.

Introduction

As schools consider their return to normal classes, they are relying on the use of tests to combat COVID-19 transmission(Carl T. Bergstrom, 2020). With little information about how COVID-19 will spread through schools, decision-makers are turning to models of viral spread to estimate the amount of testing and the testing frequency required to allow a normal return to schools, as well as other interventions(Bradley et al., 2020; Gressman & Peck, 2020; Paltiel et al., 2020).

Central to the efficacy of mathematical models is the choice of the parameters in those models that describe the spread of the disease. In order to model testing, the model must make assumptions
about how long after infection a virus is present at a level that can be detected as well as frequency of the false negative rate. Considerations about the rate of transmission of disease are also important because high levels of transmission prior to symptom onset make it harder to control the outbreak (Hellewell et al., 2020). Both detection of virus by a PCR based test and transmission of disease to another person are processes that should be proportional to viral load in the patient because the presence of virus in the patient serves as the infectious agent and as the template for the test. Transmission probability relative to the date of symptom onset has been estimated by He et al. based on the serial interval of multiple transmission pairs (He et al., 2020a). They found that transmission probability likely starts rising just over two days before symptom onset, and that ~44% of transmission may occur prior to symptom onset (He et al., 2020a) (Fig. 3.1).
Figure 3.1 Dynamic SARS-CoV-2 Viral Load and False Negative Rate. Viral transmission data and test false negative rate data both suggest that SARS-CoV-2 is undetectable until ~2 days prior to symptom onset. Shown in cyan is the viral load data by day from onset of symptoms from He et al. Shown in magenta is the false negative rate of tests by day from Kucirka et al. Transmission probability begins increasing ~2 days before symptom onset, at the same time that the false negative rate of tests begins dropping.

Assessing the efficacy of tests relies upon understanding the false negative rate of testing. The false negative rate of testing can be broken down into two basic types of false negative, one is a
technical failure where the test fails on a sample with detectable levels of virus. Another type is a false negative due to the latent period of the virus, where there is not yet sufficient viral titer in the sample for it to be detected by the test. The transmission and viral load data from He et al. would suggest that prior to 2.4 days before symptom onset, infected people may not have sufficient virus to be detectable by a test. In a study by Kucirka et al., the dynamics of false negative rate over time was determined by examining data on false negative test in patients who were eventually found to be positive (Kucirka et al., 2020). False negative rates were found to be 100% until two days prior to symptom onset and they reached a minimum of approximately 25% two days after symptom onset (Fig. 3.1).

These two studies represent two different data sets that can inform assumptions about viral load, as the ability to transmit disease and detect infection are both likely to be proportional to viral load. While He et al. estimated transmission rates from transmission pairs and compared that to measured viral loads, Kucirka et al. measured the likelihood of a positive test relative to symptom onset and collected data points from presymptomatic patients. The data from both studies predict that detection and viral spread are likely to begin approximately 2 days before symptom onset.

The ability of testing to slow the spread of disease is related to the accuracy and function of the test but also to how fast the disease spreads. In order to stop disease spread, each infected person must, on average, infect less than one other person (an effective Reproduction number (Rt) below 1) (Inglesby, 2020). If a large amount of transmissibility occurs in a small window of time, it is more difficult to identify the infected individuals before they transmit to more than one person (Hellewell et al., 2020). We hypothesize that the interplay between an undetectable period during incubation and a non-uniform distribution of transmissivity leads to different outcomes for the efficacy of tests in combating disease spread compared to simple estimates of a uniform chance of transmission and a uniform false negative rate. To examine this, we developed a stochastic agent-based Susceptible-Exposed-Infectious-Recovered (SEIR) model of 10,000 students, roughly the size of the University of
Maine. We find that the period of undetectable virus leads to a high basal apparent false negative rate, regardless of test sensitivity. When we consider the scenario where only testing is used to combat spread, we find that a simple model that assumes uniform viral spread and perfect tests predicts that testing everyone every 14 days may be sufficient to bring the Rt below 1. However, a model using the combination of disease spread based on the transmissivity data from He et al. and the dynamic false negative rates for tests from Kucirka et al. predict that as much as 100% of the population may need daily testing to bring the Rt below 1 and stop viral spread. While lower levels of testing can be effective in the presence of other interventions such as masking or social distancing, we conclude that the dynamics of an undetectable period, viral transmission that is biased early in the disease, and dynamic false negative rates significantly change the predictions of an SEIR model, and these factors should be considered when developing models to plan for public health interventions to combat COVID-19.

**Methods**

**Model**

We chose to build a stochastic agent-based model for two reasons: (1) it would allow us to easily implement nonuniform probabilities over the course of infection and (2) a stochastic model would capture the inherent noise in a system that is presumed to start with a small number of infected cases. We implemented the model in MATLAB using the indicated probabilities and if–then statements. The test was performed with 10,000 individuals to represent the college student body. The model runs daily for 120 days, approximating a semester. The basic structure of the model is outlined in Fig. 3.2A. Because it is a stochastic model, we perform 100 independent runs (Fig. 3.2B), and report the median and 95th percentile results. The model can be found on GitHub at https://github.com/Kelley-Lab-Computational-Biology/coronamodel.
Figure 3.2. Stochastic Agent Based Epidemiological Model A stochastic agent-based model of COVID-19 transmission. This is a stochastic SEIR model implemented in MATLAB. Each transition in state is based on if–then statements with specific probabilities described in Fig. 3.3. (A) Individuals start as susceptible, and the initial population is seeded with 10 random infected individuals, each starting at a random point of progression through the disease, and with random symptoms. Upon being infected, an individual become exposed (presymptomatic), and is assigned a day for symptom onset. Detectability for testing and infectiousness both begin at 2 days prior to onset of symptoms. Infectious individuals can be either asymptomatic, or symptomatic with mild or severe symptoms. Those with severe symptoms will self-isolate and initiate contact tracing through seeking medical attention. Asymptomatic individuals and those with mild symptoms can be isolated through contact tracing or through detection by a test. Infectious individuals will recover randomly with a median time of 14 days. (B) An example of 100 independent simulations with the model. Shown are susceptible, infected (encompassing exposed, infectious, and isolated individuals), and recovered individuals in simulations where no interventions were implemented. Each individual simulation is represented as semi-transparent points, while the median value of all simulations is plotted as a line.
Symptoms

For the timing of symptom onset, we used the symptom onset distribution calculated by He et al. This distribution has a median onset time of 4.2 days, and 99% of cases experience symptom onset by 14 days (Fig. 3.3A)(He et al., 2020a). An update to the He et al. paper considers an altered time to symptom onset (~ 6 days on average)(He et al., 2020b), but does not lead to significant differences in the model output.

Figure 3.3. Stochastic Agent Based Epidemiological Model Parameters

Probability distribution of onset of symptoms from He et al. (B) Breakdown of symptom groups in the
model. (C) Probability distribution of recovery based on a median time to recovery of 14 days. (D) R0 of 2.5 scaled to a uniform transmission probability distribution. The gray box indicates where the cumulative probability reaches 1. Individuals must be detected prior to this, on average, in order to reduce the apparent R0 below 1. (E) The R0 of 2.5 scaled to the viral load based on He et al. The gray box is the same as above. (D) The R0 of 2.5 scaled to the positive test rate from Kucirka et al. This was done because the changes in positive test rate are likely related to viral load, and so may be an alternative representation of transmission likelihood. The gray box is the same as above.

The CDC reports an overall asymptomatic rate of 40% (COVID-19 Pandemic Planning Scenarios, 2020), but we are concerned about the likely asymptomatic rate among a young population. When the aircraft carrier Theodore Roosevelt had an outbreak of COVID-19, they reported that as many as 350 out of 600 sailors were asymptomatic for an asymptomatic rate of 58% (Correll, D. S., 2020). As the population aboard a navy ship are likely to skew younger and healthier than the population as a whole, we felt they may be more representative of college age students. Thus, we assumed an asymptomatic rate of 50%. People who are symptomatic are then assigned either mild or severe symptoms, based on CDC data that 81% of people experience mild symptoms, 14% severe, and 5% critical (Interim Clinical Guidance for Management of Patients with Confirmed Coronavirus Disease (COVID-19), 2020). We consider severe and critical together, as we expect both to seek medical assistance, and then be isolated from the general population. We also assumed that these numbers represent the percentages of symptomatic people, so ultimately the model assigns 50% asymptomatic, 40.5% mild, and 9.5% severe (Fig. 3.3B). We assume that those experiencing severe symptoms seek medical attention at the beginning of symptom onset and are isolated, and initiate contact tracing. Unless explicitly stated otherwise, we assume that mild cases do not self-isolate, as they may not realize that their symptoms
are COVID-19 related, or they may be reluctant to identify themselves as ill for fear of isolation and removal from their normal college activities (Pagoto, 2020).

**Recovery**

The CDC reports median recovery time as 14 days for mild illness (Interim Clinical Guidance for Management of Patients with Confirmed Coronavirus Disease (COVID-19), 2020). We assume the same recovery period for asymptomatic people. Because severe illness results in medical attention and isolation, we did not consider the extended recovery period for severe illness as it would not change transmission in our model. The daily recovery probability is modeled as a binomial distribution with a mean of 14 days (Fig. 3.3C).

**Probability of viral spread**

The model assumes an R0 of 2.5 (COVID-19 Pandemic Planning Scenarios, 2020). Each individual in the model receives an R0 normally distributed around 2.5 (with a standard deviation of 1) to allow for variability in transmissibility between people. We treat the population as well-mixed, and so when a transmission event occurs, it has an equal probability of infecting any individual in the population. If the randomly selected individual is isolated or recovered, then no infection takes place; only susceptible individuals are infected. For the normally distributed R0, we took three different approaches to viral transmission probability. (1) We assume a uniform daily transmission probability equal to 2.5/14 (R0 / median time of illness) (Fig. 2.3D). (2) We assume that daily transmission rate follows the transmission dynamics estimation from He et al. (He et al., 2020a), where transmission starts 2 days prior to symptom onset (Fig. 2.3E). An update to the transmission estimations from He et al. was published by the authors (He et al., 2020b) that changes the daily transmission probabilities but maintains the 44% presymptomatic transmission. We compared the effect of using the newer profiles in our model and found that they had a minimal impact on the predictions (Supplemental Figure S1) and so we continue to use the original profile published in (He et al., 2020a). (3) A daily transmission probability scaled to the
false negative test rate reported by Kucirka et al (Kucirka et al., 2020), under the assumption that the
dynamics of the false negative rate are related to the viral load (Fig. 3.3F). For each of these
assumptions about viral spread, people must be detected on average before they spread virus to one
other person (Rt below 1). We have indicated in Fig. 3.3D,E,F with a shaded rectangle the time in which
sick individual must be detected to keep the average number of new infections below 1.

We also explored the potential impact of super-spreader events on the outcome of the model, as superspreading is thought to be a strong driver of COVID-19 transmission(Laxminarayan et al., 2020; Majra et al., 2021). In this case, we generated a distribution of R0s that leads to 20% of individuals
causing 80% of transmission, following the 80/20 rule (Fig. 8)(Lloyd-Smith et al., 2005a) Each individual
in the superspreader model has 35 interactions a day to allow for multiple transmission events, while
using 1/35 of their daily transmission probability for each.

**Testing**

Tests can be administered to the entire population, or to randomly selected subsets of the
population either daily or at varied frequencies. Unless test delay is explicitly mentioned, we assumed
tests are resolved on the day they are administered. We consider a few scenarios for false negative
rates: (1) Perfect tests, where there is no chance of a false negative rate, and there is no period of
undetectable infections. (2) Our “simple” scenario where the virus is undetectable until 2 days prior to
symptom onset, after which tests have a uniform 5% false negative rate. (3) Dynamic false negative
rates based on those measured by Kucirka et al.(Kucirka et al., 2020). Like the simple scenario, there is
no chance of detecting an infected individual prior to 2 days before symptom onset. We do not consider
the ramifications of false positive rate. While the false positive rate is important due to the burden that
incorrectly identified cases place on resources(Paltiel et al., 2020), that consideration does not affect the
Rt of the system. Tests work equally well on individuals who are symptomatic or asymptomatic in our
model. The only difference in asymptomatic individuals is that they cannot self-isolate, they otherwise have a disease progression and viral load dynamics as if they were symptomatic. Testing of a fraction of the population was done by random sampling without replacement at the indicated frequency. Thus, the identity of tested individuals is independent of previous tests, and no individual is tested more than once per day of testing.

**Contact tracing**

For each individual in the model, we store the identity of the source of their infection, and the identities of people they transmit to. If someone is identified as sick by self-isolating and seeking medical attention, or if they are identified by a randomly administered test, contact tracing is initiated. We assume a 75% chance to identify each infected contact of the individual, as this is sufficient for contact tracing to work (Eames & Keeling, 2003), but not overly optimistic about the ability to find transient contacts in a University setting. Contacts that do not lead to infection are not explicitly modeled, thus the only contacts recorded in the model are transmission events. For each transmission event, the model determines if the individual is detected by the contact tracing success rate (75%) and then successfully identified contact will be moved to the isolated pool without the requirement for a positive test.

**Results**

**Nonuniform false negative rates can delay detection of infected individuals**

Why are we concerned about uniform versus nonuniform false negative rate? To illustrate the issue, we can examine the first day of disease progression at which an infected individual is likely to be detected when tested daily. To examine the effect of detectability, we explore a hypothetical population where every person is sick and is tested daily (not the full SEIR model). We compare three different false negative rate dynamics over 14 days of disease progression assuming testing everyone, every day, and we assume symptom onset at day 5. The average false negative rate of each is the same (50.42%,
chosen to match the overall average false negative rate of the Kucirka et. al. data (Kucirka et al., 2020)), but the way the rates change over time differs (Fig. 3.4). We have indicated with a gray rectangle the two days prior to symptom onset that may represent as much as 44% of viral transmission capability (He et al., 2020a). (1) A completely uniform false negative rate leads to most infected people being detected by day 3, prior to becoming infectious. (2) An undetectable period followed by a uniform rate of detection catches most individuals by day 5 (it is, after all, just a two-day offset of (1), with the uniform rate rescaled to still average to the same overall false negative rate). These assumptions about the dynamics of viral spread allow more people to spend time in the infectious period prior to being detected than the completely uniform assumption. (3) The dynamic false negative rates of Kucirka et al. means that few individuals are likely to be caught prior to the potential for significant viral spread (Kucirka et al., 2020).

Figure 3.4. Effects of False Negative Rates on Detection Effects of False Negative rates on detection. (A) Non-uniform false negative dynamics can delay detection of infected individuals. Shown is the chance of
first being detected at each day of disease progression based on three scenarios with the same average false negative rate across the 14 days shown, but different temporal dynamics. For this graph, we assume that symptom onset begins at day 5. In yellow is the undetectable period prior to 2 days before symptom onset. The two days before symptom onset is shown in gray. Viral load data suggests that as much as 44% of transmissibility may occur in these two days. The line represents histograms of the first day that an individual would be detected by a daily test with the given false negative rate dynamics. (B) The undetectable period and temporal dynamics of the false negative rate lead to high apparent false negative rates. The first day of the model was run 100 times with 10,000 sick individuals. In cyan we show the model run with the simple assumption that infected individuals were undetectable before viral load begins, (2 days prior to symptom onset, based on the He et al. data), and that after that point the tests will always detect infected individuals. In magenta, the model uses the dynamic false negative rates from Kucirka et al., in which both test error and inability to detect due to low viral load are mixed together. Also included as a comparison is the effect of perfect tests shown in gray.

An undetectable period leads to high apparent false negative rates

The data from He et al. suggests that viral load reaches its maximal level prior to symptom onset and the transmission pair data suggest that virus first rises to a detectable level two days prior to symptom onset. Since viral RNA is the template for PCR based tests, the ability of the tests to detect the virus will be dependent upon the viral load, so we made the simple assumption that virus was undetectable prior to 2 days before symptom onset, and that it was uniformly detectable after this point. The measured false negative rates reported by Kucirka et al. validate this assumption, and provide daily false negative rates after viral load begins increasing. We made a separate model using the Kucirka et al. measured false negative rates.
We used the model to test the effect of these different assumptions on the overall false negative rate that would be encountered during testing for the virus on the first day of classes, where the people who are positive are randomly distributed (i.e. they have contracted the virus through independent events) through the progression of the disease. For example, while the median of symptom onset is between 4 and 5 days, 12% of cases would have a start of symptoms at 9 days or later. In this case, there would be at least 7 days during which there is insufficient virus present to detect an infection, regardless of the efficacy of the test. Simulations of the first day of the semester were run 100 times, and the median value of the false negative rate is reported (Fig. 3.4B). We found that the simple model, which assumes uniformly perfect tests after 2 days prior to symptom onset displays an apparent false negative rate of 17%. In the case of the Kucirka data (Kucirka et al., 2020), which has both the undetectable time period before virus replication begins and the measured daily false negative rates afterward, which reach a minimum of ~ 25% two days after symptom onset, the overall false negative rate of the simulation was 48%. It is worth reiterating that this is the false negative rate one would experience testing a random group of people, not the false negative rate expected for directed testing, such as testing someone who is symptomatic. The Kucirka et al. false negative data is a compilation of both the false negative rate of the test, and the false negative rate due to the viral infection dynamics (Kucirka et al., 2020). The simple model considers only the false negative rate from the viral dynamics and places the lower bound at 17% false negative, which is large but within the realm of consideration (Paltiel et al., 2020).

An undetectable period and high early transmission levels lead to a need for higher levels of testing

If the effective false negative rate ranges from 17% to as high as 48%, it is likely to affect the level of testing required to combat the spread of COVID-19. We set out to examine the effect of testing on the spread of disease by calculating the effective R0 of the virus when different testing regimens are used, while varying the dynamics of detectability and test false negative rate. We varied the fraction of
the population being tested and the frequency of the test for four scenarios. For random testing of a
fraction of the population, each person can only be tested once per testing and the sampling for testing
on different days is independent of previous tests, such that a given individual may be tested multiple
times in a row or may be skipped. The four different viral dynamics scenarios are:

Scenario 1: “Perfect tests, Uniform Spread” where we assume no period of undetectability, no
false negative rate, and a uniform chance of transmission equal to 2.5 / 14.

Scenario 2: “Simple Undetectability, Fast Spread” where we assume that the virus is not
detectable until 2 days prior to symptom onset, and then has a 5% false negative rate after that point
(this 5% false negative rate is a change from the simple assumption above (Figure 3.6), which assumed
perfect tests). This condition uses the He et al. viral load data to scale the R0 (Figures 3.1 and 3.4E),
which results in ~45% of transmissivity prior to symptom onset.

Scenario 3: “Dynamic False negative, Slow Spread.” This uses the day-by-day false negative rates
reported by Kucirka et al. for testing (Figure 3.1). For transmissivity, we use the day-by-day positive rates
from the Kucirka et al. data as a stand-in for viral load (Figure 3.4F). The shape of this profile still biases
spread early in the disease, but not as early as the He et al. viral load data.

Scenario 4: “Dynamic False negative, Fast Spread.” This scenario uses the day-by-day false
negative rates from Kucirka et al. for testing, and the He et al. transmissibility estimates.

These simulations are run with testing and subsequent isolation of positives being the only
intervention used to combat viral spread. We report the median effective Rt as well as the 95th
percentile Rt for each condition because testing regimens that work only half the time may not be useful
when considering public health. We see that perfect tests can be effective while testing as little as 25%
of the populace every other day (Fig. 3.5). All simulations that do not assume perfect tests require a
larger proportion of the population to be sampled under these conditions. Scenario 2 and Scenario 3 result in remarkably similar results for which testing regimens are required for suppression of viral spread. The fast viral-spread and sensitive tests of Scenario 2 are therefore compensated for by the slower viral spread and insensitive testing of Scenario 3. With scenario 4, where the transmission occurs early in the disease and false negative rates are high, only testing of every individual every day was able to bring the Rt below 1. Thus, viral transmission that is biased early in the progression of the disease and higher false negative rates require a more aggressive testing regimen than would be suggested by uniform assumptions.
Figure 3.5. High asymptomatic transmission and dynamic false negative rate lead to more required testing to prevent a COVID-19 outbreak

High asymptomatic transmission and dynamic false negative rate lead to a requirement for more testing to bring the viral spread under control. Heatmaps show the effective Reproduction number (Rt) from 100 simulations run with the given proportion of the population tested at the indicated frequency. The top row of matrices shows the median Rt, while the bottom row of matrices shows the value of the upper 95th percentile (i.e. conditions that will work in 19 out of 20 situations). While the scenario 1 perfect tests suggest testing the entire population every two weeks may work to stop spread of the virus, using scenario 4 parameters predicts that testing the entire population daily was necessary.

While these simulations suggest that testing would have to be very aggressive to bring viral spread under control, they are not assuming any interventions beyond testing followed by contact tracing and isolation. In reality, testing is likely to be a component of a multi-pronged approach to combating viral spread that would include social distancing and masking. We decided to examine the efficacy of testing under a situation where other interventions had brought the viral spread down, but not below an Rt of 1. A recent study of mask efficacy suggests that surgical or cloth mask wearing can reduce the risk of contracting COVID-19 to 33% the risk of those not wearing masks (Chu et al., 2020). Interestingly, this is similar to the percent decrease in particulates that has been described for a cloth mask (van der Sande et al., 2008) (average reduction in particulates to ~ 31% of control over a 3 h experiment). We implemented a model where 70% of the population uses masks that reduce transmission rate by 67%. This results in a median apparent Rt of 1.3, and a 95th percentile value of 1.44. We then performed the simulations using the array of testing regimens as above. For this analysis, we used the Scenario 4 conditions of dynamic false negative rate (Kucirka et al., 2020) and high early viral transmission dynamics (He et al., 2020a), as these conditions are the hardest to reduce and will give
the most conservative results for frequency and amount of testing. We find that under these conditions it would now be possible to bring the Rt below 1 in 95% of cases by testing 25% of the population every day (Fig. 3.6A). Testing every person would now be effective when done once a week.

![Figure 3.6](image)

**Figure 3.6. Masking allows fewer tests and lower frequencies of testing** In the presence of masking, fewer tests and lower frequencies of testing can be successful in driving Rt below 1. (A) Here we implemented 70% of the population using a mask that is 67% effective with the parameters of Scenario 4, early transmission of virus based on the He et al. viral load data, and dynamic false negative rates for tests based on Kucirka et al. The top row of matrices shows the median effective Reproduction number (Rt), while the bottom row of matrices shows the value of the upper 95th percentile. Masking drove the median Rt from 2.5 to ~ 1.3. Tests were then able to drive the 95th percentile Rt below 1 with less aggressive testing schemes than in Fig. 3.5. (B) The same conditions as (A), with an included 1 day turn
around delay in testing results. The magenta line shows the border between an Rt above 1 and an Rt below 1 without a delay. (C) As in (B) with a 2 day turn around delay in testing results.

These previous simulations assume instantaneous turnaround time for the test results. Unfortunately, test results may take a day to several days for results to be available. The delay is typically due to the backlog of samples needed to be tested, lack of testing equipment, and the relatively small number of labs and technicians with proper certification (Barone, 2020). To analyze the effect on delay in receiving the results, we used Scenario 4 conditions with 70% of the population using masks then performed the simulations using the array of testing regimens as above. In the model, students who are tested and found positive start their isolation after they find out their results along with people isolated due to contact tracing. Figure 3.6B shows how implementing a one-day delay has a detrimental effect on the testing requirement in order to prevent an outbreak. A one-day delay in receiving test results leads to a requirement for a two-day increase in frequency, as testing the whole population every 5 days would prevent an outbreak compared to testing every 7 days with no delay. Similarly, a two-day delay in receiving test results leads to a four-day increase in testing frequency necessary to prevent an outbreak (Fig. 3.6C). The delay in test results significantly changes the testing frequency requirements in order to prevent an outbreak.

The degree to which people with mild symptoms will decide that they have COVID-19 instead of dismissing their symptoms as insignificant is unclear. For example, a traveler was originally identified as an asymptomatic spreader early in 2020, when in fact she had symptoms at the time but did not identify them as severe enough to consider herself sick (Rothe et al., 2020). Ultimately, she had severe COVID-19 and was even hospitalized. Presumably, those with moderate symptoms may also have difficulty differentiating between symptoms indicating COVID-19, or something innocuous, such as seasonal allergies. Additionally, a survey of students has suggested that they may be reluctant to remove
themselves from school for mild symptoms (Barone, 2020). For these reasons we assumed that people with mild symptoms did not self-isolate in the above simulations. To examine this issue, we tested how self-isolation by people with mild symptoms would affect the necessary testing regimen to contain spread using Scenario 4 conditions as above. Under circumstances where people are very responsible and self-isolate under any mild symptoms of illness, we see that testing 100% of the population every 4 days becomes sufficient to successfully bring the Rt to 1.00 or below in 95% of simulations (Fig. 3.7). Thus, the degree to which the population of interest takes seriously the instructions to self-isolate if they are feeling ill could influence the efficacy of any given testing strategy and could allow less stringent testing plans to succeed. The frequency of self-isolation due to perceived symptoms is a parameter that this model is very sensitive to (Supplemental Figure S2), and a parameter that may vary significantly from one population to another.
Figure 3.7. Assumptions about Mild COVID Cases Isolating Influence Testing Requirement  
Assumptions about the behavior of people with mild symptoms influence the amount of testing required. Scenario 4 from Fig. 3.6 was repeated with the addition that people experiencing mild symptoms self-isolate on the day of symptom onset. Under these conditions, testing everyone every 4 days is sufficient to bring 95% of situations to an Rt of 1.00 or below.

Superspreader events increase the variability in the spread of COVID-19

Superspreading events, where one individual transmits virus to many more individuals than one would expect on average, seem to play a large role in the spread of COVID-19 (Majra et al., 2021) and it has been estimated that 5% of people may be responsible for as much as 80% of cases (Laxminarayan et al., 2020). In our model, we approach variability in transmission between individuals by assigning each individual their own R0, a constant for that individual’s chance to transmit, rather than giving each person the same R0 and differing numbers of opportunities to transmit. To examine the effect of this phenomenon on the requirement for testing, we generated a superspreading probability distribution where 20% of sick individuals were responsible for 80% of cases, following the general 80–20 rule (Lloyd-Smith et al., 2005b) (Fig. 3.8A), which leads to most people having little to no transmission, and a small number of people being highly transmissible. 500 simulations were run with a first day test of all students and the Scenario 4 transmission and false negative rates for both R0 distributions (Fig. 3.8B). Incorporating superspreaders into the model leads to more variability in the timing of outbreaks and an increase in the number of situations where no disease occurs (lucky testing on day 1, or contagious individuals who do not end up transmitting). In those cases where outbreaks occur, they lead to the same level of illness, as one would expect for models with the same R0. When we examine testing regimen, we find that median Rt values do change a bit, (due to the increased frequency of no transmission) but that the testing required to stop 95% of outbreaks remains the same as that found for
Scenario 4 in Fig. 3.5. Thus, superspreading increases the variability in outcomes for situations. We may have seen superspreading driven variability between discrete situations borne out in the real-world differences that were seen between COVID-19 occurrence on different college campuses in the fall of 2020 (Renner, 2020).
Figure 3.8. Superspreading leads to more variability between discrete simulations

Superspreading leads to more variability between discrete simulations, but requires the same amount of testing as normally distributed transmission probabilities. (A) On the left is the normal distribution of R0s assigned to individuals, while the right distribution is a superspreader distribution where 20% of the population is responsible for 80% of the infections. (B) 500 runs of the model with either the normal distribution or the superspreader distribution of R0s and the Scenario 4 transmission and false negative rate dynamics. Medians of each population are shown as solid lines. Note that the medians of infected and recovered in the superspreader conditions fall on the x-axis line. (C) Effective reproduction number (Rt) for the indicated testing frequency and proportion of population for the superspreader R0 distribution. As in Fig. 3.5, only daily testing of 100% of individuals is sufficient to drive Rt below 1 under these assumptions.

Discussion

Available data on SARS-COV2 viral load over time and on false negative rate of tests over time both suggest that virus may not be detectable prior to ~ 2 days before symptom onset, and transmissibility of the virus is biased towards the beginning of disease progression. Here we have examined the effect of nonuniform viral transmission and nonuniform detectability of disease on the efficacy of testing as a means to stop viral spread. We find that the combination of the non-uniform transmission dynamics and false negative rate predict that tests must cover more of the population and be given more frequently than predicted by a model that assumes uniform distributions. Thus, models that make simple assumptions about viral spread, and false negative rate or underestimate the effect of the undetectable period on the apparent false negative rate may recommend less testing than is necessary to stop viral spread.

The parameters used for these simulations (transmissibility dynamics, false negative rate, voluntary self-isolation, efficacy of masks and level of compliance with masking) are not concrete, and
are likely to vary between institutions, populations, or areas. As the model parameters approach containment of viral spread, the prevalence of virus in the surrounding community, or other sources of introduction into the system will be more important to the considerations for testing amount and frequency as well as quick turnaround time of results. As such, these results should not be seen as concrete recommendations on specific testing strategies, although the results for Scenario 4 are clearly conservative, and others have come to similar conclusions about the need for frequent testing (Chang et al., 2021; Larremore et al., n.d.). Similarly, these studies should not be construed as saying that tests do not work or that tests should not be a part of the public health strategy for combating viral spread. Instead, the takeaway message is that modeling of tests should be done with consideration of the potential for an undetectable period, nonuniform transmission dynamics, and the potential for viral load to influence false negative rate. Each of these considerations alters the conclusions that a model will come to about the number and frequency of tests required to combat viral spread.

There are many reports in the news media of organizations using a negative test result as a prerequisite for engagement in some activity, such as returning to college or attending a summer camp. The Kucirka et al. data on dynamic false negative rate should already give pause to these types of plans, but we show here that testing a population of people who may have a random distribution of progression through disease may have a false negative rate as high as 48%. The possibility of missing ~1/2 of positive individuals by performing a complete testing of the population of interest should be considered when making these plans. This high false negative rate is specific to tests that are performed on a population likely to have a random distribution of viral progression. In situations where the tests are being given because of symptoms, or because of contact tracing, the population being tested would be biased towards later days in the progression of the disease, and the overall false positive rate would be lower than the 48% value. However, even if one were using a test that was 100% sensitive and specific given a sample that contains template, it is likely that they would still experience
the ~17% false negative rate due to the latent period of the virus before it reaches sufficient levels to be detectable. Thus, plans to allow people to participate in activities dependent upon a negative test should be aware of the greater than 1 in 6 likelihoods of missing an infected person in their testing.

In conclusion, many people are resorting to modeling of disease transmission to assist in the formulation of public health plans for the return to schools and economic activities. When designing these models, simple assumptions of uniformity of transmission and uniformity of false negative rate can give overly optimistic views of the efficacy of testing. These nonuniform dynamics are complicated to implement in a deterministic ODE model, but easier to implement in a stochastic agent-based model. The stochastic model, however, is slow compared to an ODE model. Answering questions about tests does not, however, require a population to be so large as to be unmanageable with a stochastic model, as the trends in testing efficacy should remain the same. Thus, we recommend that stochastic models be used to model efficacy of tests so that complex dynamics can be readily accounted for. The results of stochastic models could then be used to parameterize deterministic models for other uses.
CHAPTER 4

CONCLUSIONS AND FUTURE DIRECTIONS

Here we model a molecular system and an epidemiological system using similar methods to mimic biological systems to make hypotheses. Both explore biological systems using uniform dynamics for essential processes in the system and find their change over time is necessary for realistic results.

Dynamic biological systems can be modeled using ODEs to describe the system or individual agents in ABM. These dynamics on individual components in the system can be determined through experimentation or previous literature to enhance our understanding of the system. How individual dynamics translates to dynamics in a whole biological system may be unknown and harder to determine due to the complexity of biological system. This unknown effect on the biological system is a disadvantage to ODE modeling because dynamic rates need to be explicitly modeled. Whereas ABM can incorporate many heterogenous dynamic processes to model the system, giving ABM an advantage in mimicking dynamic biological systems.

The COVID-19 epidemiological model gave insight into how much testing is required to combat the viral outbreak on a college campus. Modeling the dynamic transmission and the false negative rate of the tests based on day of symptom onset impacted the testing requirements in order to prevent an outbreak of COVID-19. We show this by model the system with and without dynamic transmission/false negativity of tests. These results mean transmission is highly dynamic and is required to be modeled in an epidemiological system to predict viral outbreaks. In addition, even though COVID-19 tests have a highly dynamic false negative rate, results consistently show high testing frequency and minimal delay in results reduce the probability of an outbreak. These results matter to us because epidemiological models are used to determine preventative measure to combat viral spread. Epidemiological models not including assuming uniform transmission may therefore lead to overly optimistic results that are unable to combat viral spread. Our qualitative epidemiological model can’t tell us public health protocols for
combatting spread because of our models’ assumptions of a well-mixed population. A well-mixed epidemiological system meaning every student has an equal opportunity to transmit virus to anyone on campus. In order to accurately model a viral outbreak on campus to be used for public health context, the model would need to incorporate classroom data and regulating the amount of contacts each student has per day would be necessary in order to obtain accurate results based on realistic dynamics.

Since writing the COVID-19 epidemiological chapter in August 2020, many developments in the COVID-19 pandemic have occurred including more transmissive strains of SARS-CoV-2, vaccinations and widely available at-home COVID-19 tests. Each of these developments have altered the effective reproductive number of COVID-19 and therefore effects the required surveillance testing of the system. Testing how effective surveillance testing in a system with a variety of SARS-CoV-2 strains, and with a fraction of the population partially or fully vaccinated are interesting further studies with this model. With these new parameters, this model can be further used to understand how preventative measures such as vaccinations and N95 vs cloth masks effects the spread of covid-19 on a college campus.

In the GPCR agent-based model, we looked at receptor post translation modifications in relation to its relationship to its negative regulator, RGS. RGS-GPCR interactions are known to cause receptor retention and increase GAP activity on the membrane. We didn’t know how RGS-GPCR interaction effects receptor activity localization. From our research, we found a bifurcated pattern of receptor activation peripheral to the site of receptor secretion. By adding RGS, we verified RGS-GPCR cause receptor retention in the computational model. We confirmed these dynamics through experimentally looking at receptor (Ste2) in wildtype cells and in mutant SST2Q304N cell whose GPCR cannot bind RGS. Wildtype had a significantly greater fluorescence compared to the mutant strain confirming the computational model. Additionally, the RGS-GPCR interaction was shown to affect the receptor modification distribution, therefore effecting receptor mediated signaling. The computational model suggests post-translational modification localization rely greatly on GPCR-RGS interaction. This dynamic
is seen experimentally with receptor polarization not being proportional to receptor activation suggesting modulated receptor mediated signaling based on receptor activation. Because this dynamic was directly inverted with mutant RGS unable to bind to GPCR suggesting this dynamic is due to RGS-GPCR interaction. From this study, we learned how GPCR-RGS interactions modulate receptor modifications and altered the receptor mediated signaling to promote gradient tracking.

Future directions include continuing to use the GPCR ABM in further research to understand how receptor mediated signaling effects gradient tracking accuracy. This biological system relies on receptor internalization promotion due to receptor activation as well as RGS mitigating receptor internalization. Computationally modeling this system can give insight on receptor activation localization and its effect on downstream effector localization such as cytoskeletal element. This research provides further insight on the predicting the spatiotemporal regulation of GPCR signaling. Coupling the computational model with experimental results provides us to a broader understanding of the biological system.
BIBLIOGRAPHY


Correll, D. S., (2020). navytimes.com


*Imported:YPM/Parameter Fitting.* (2017). OpenWetWare. [https://openwetware.org/ypm/Parameter_Fitting#Ste2_Internalization](https://openwetware.org/ypm/Parameter_Fitting#Ste2_Internalization).


**APPENDIX A:**

Table A.1 Agent Based Receptor Model Parameters

<table>
<thead>
<tr>
<th>Description</th>
<th>Value Model 1</th>
<th>Value from literature</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial Number of Receptors</td>
<td>500</td>
<td>2000</td>
<td></td>
</tr>
<tr>
<td>bins</td>
<td>200</td>
<td></td>
<td></td>
</tr>
<tr>
<td>timestep</td>
<td>0.01 seconds</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Simulation Length</td>
<td>2,000 seconds</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Circumference</td>
<td>10 µm</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Diffusion</td>
<td>0.0025 µm²/s</td>
<td>0.0025 µm²/s</td>
<td>(Lakhani &amp; Elston, 2017)</td>
</tr>
<tr>
<td>Rate of Pheromone binding</td>
<td>0.185 s⁻¹µM⁻¹</td>
<td>0.185 s⁻¹µM⁻¹ based on Kd and koff</td>
<td>(Bajaj et al., 2004a; Chen &amp; Konopka, 1996; Jenness et al., 1986; Lee et al., 2001)</td>
</tr>
<tr>
<td>Rate of Pheromone bound Receptor unbinding</td>
<td>0.001 s⁻¹</td>
<td>0.001 s⁻¹</td>
<td>(Bajaj et al., 2004a; Jenness et al., 1983; Raths et al., 1988)</td>
</tr>
<tr>
<td>kcat of pheromone bound Ste2 Phosphorylation</td>
<td>0.1 s⁻¹</td>
<td>0.1 s⁻¹</td>
<td>(Imported:YPM/Parameter Fitting, 2017)</td>
</tr>
<tr>
<td>Ste2 Ubiquitination rate</td>
<td>3.2e-3 s⁻¹</td>
<td>degradation</td>
<td>(Imported:YPM/Parameter Fitting, 2017)</td>
</tr>
<tr>
<td>Receptor Synthesis Rate</td>
<td>16</td>
<td>4 s⁻¹</td>
<td>(Yi et al., 2003)</td>
</tr>
<tr>
<td>Frequency of Exo (events per minute)</td>
<td>8</td>
<td>24,2.7</td>
<td>(Carrillo et al., 2015; Johnson et al., 2011)</td>
</tr>
<tr>
<td>Frequency of Endo (events per minute)</td>
<td>16</td>
<td>100,1.6</td>
<td>(Carrillo et al., 2015; Johnson et al., 2011)</td>
</tr>
<tr>
<td>Exocytic Vesicle Length</td>
<td>0.1 µm</td>
<td>0.48695 µm</td>
<td>(Carrillo et al., 2015)</td>
</tr>
<tr>
<td>Endocytic Vesicle Length</td>
<td>0.05 µm</td>
<td>0.41469 µm</td>
<td>(Carrillo et al., 2015)</td>
</tr>
<tr>
<td>Endocytic Pit Mature Age</td>
<td>1-5 minutes</td>
<td></td>
<td>(Pedersen et al., 2020)</td>
</tr>
<tr>
<td>Endocytic Pit Density</td>
<td>3 actinpit/µm</td>
<td>0.9719 actinpit/µm</td>
<td>In Results Section</td>
</tr>
</tbody>
</table>
Table A.2 GTP ODE Parameters

<table>
<thead>
<tr>
<th>Description</th>
<th>Parameter Name</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Value Model RGS Disassociation</td>
<td></td>
</tr>
<tr>
<td>Baseline GAP Activity</td>
<td>( k_{\text{baselinegap}} )</td>
<td>0.5</td>
</tr>
<tr>
<td>RGS GAP Activity</td>
<td>( k_{\text{gap}} )</td>
<td>0.5</td>
</tr>
<tr>
<td>Receptor GEF Activity</td>
<td>( k_{\text{gef}} )</td>
<td>0.5</td>
</tr>
<tr>
<td>Timestep</td>
<td>( dt )</td>
<td>0.01</td>
</tr>
<tr>
<td>Diffusion Rate (( \mu m^2/s ))</td>
<td>( Dm )</td>
<td>0.0036</td>
</tr>
</tbody>
</table>

The following are the ODE based equations used in this model:

\[
\frac{d(GTP)}{dt} = (-GTP \cdot k_{\text{gap activity}} \cdot k_{\text{gap}} + GDP \cdot k_{\text{gef activity}} \cdot k_{\text{gef}} - GTP \cdot k_{\text{baselinegap}}); \\
\frac{d(GDP)}{dt} = (-GDP \cdot k_{\text{gef activity}} \cdot k_{\text{gef}} + GTP \cdot k_{\text{gap activity}} \cdot k_{\text{gap}} + GTP \cdot k_{\text{baselinegap}}); 
\]
# APPENDIX B:

## Table B.1 Yeast Strains

<table>
<thead>
<tr>
<th>Stain</th>
<th>Parent</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>STE2-PHUORIN FAR1-RUBY</td>
<td>BY4741</td>
<td>STE2-PHUORIN::KAN FAR1-RUBY::HIS</td>
</tr>
<tr>
<td>SST2^{Q304N} STE2-PHUORIN FAR1-RUBY</td>
<td>BY4741</td>
<td>SST2^{Q304N} STE2-PHUORIN::KAN FAR1-RUBY::HIS</td>
</tr>
<tr>
<td>BEM1-RUBY</td>
<td>BY4741</td>
<td>BEM1-RUBY::LEU</td>
</tr>
<tr>
<td>SST2^{S539A} BEM1-RUBY</td>
<td>BY4741</td>
<td>SST2^{S539A} BEM1-RUBY::LEU</td>
</tr>
<tr>
<td>SST2^{S539D} BEM1-RUBY</td>
<td>BY4741</td>
<td>SST2^{S539D} BEM1-RUBY::LEU</td>
</tr>
<tr>
<td>BY4741</td>
<td></td>
<td>MATA LEU2Δ MET15Δ HIS3Δ URA3Δ</td>
</tr>
</tbody>
</table>

## Table B.2 Plasmid List

<table>
<thead>
<tr>
<th>Plasmid</th>
<th>Vector</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>PFA6A-PHLUORIN-KANMX6</td>
<td>pFA6a</td>
<td>Tagging of pHluorin-Kan</td>
</tr>
<tr>
<td>PFA6A-YOMRUBY2-HIS</td>
<td>pFA6a</td>
<td>Tagging of Ruby-his</td>
</tr>
<tr>
<td>PRSII405-BEM1-RUBY2</td>
<td>pRSII405</td>
<td>Integrating Bem1-Ruby:: LEU2 Vector</td>
</tr>
</tbody>
</table>

## Table B.3 Primers

<table>
<thead>
<tr>
<th>Oligonucleotide Name</th>
<th>Sequence</th>
<th>Gene</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>5' ATATGGCTGCGGGTGAACAAAA 3'</td>
<td>Far1</td>
<td>FP Tagging Far1 with pFA6a plasmid</td>
</tr>
<tr>
<td>-----</td>
<td>------------------------------</td>
<td>------</td>
<td>---------------------------------</td>
</tr>
<tr>
<td>AHM-58</td>
<td>5' TCGGGAAATAGATAAAACGCC 3'</td>
<td>Far1</td>
<td>RP Tagging Far1 with pFA6a plasmid</td>
</tr>
<tr>
<td>WSM-7</td>
<td>GGAAGCCGAAAGGTTCTGAAGATATAATAATAATTAGCATGACGTCGCTGTTTA</td>
<td>Ste2</td>
<td>FP Tagging Ste2 with pFA6a plasmid</td>
</tr>
<tr>
<td>WSM-8</td>
<td>GAAGGTCACGAAATTACTTTTTCAAGCCGTAATTTTGATCGATGAATCGAGCTCG</td>
<td>Ste2</td>
<td>RP Tagging Ste2 with pFA6a plasmid</td>
</tr>
</tbody>
</table>
APPENDIX C:

Below are the Methods and Results of the study referenced in Chapter 2. This study focused on how RGS mutation effects the desensitization mechanisms of GPCR spatial signaling. The broaden MAPK signaling profiles from phospho-mutant RGS strains experimentally led us to hypothesize a RGS phosphorylation may control desensitization mechanisms for GPCR spatial signal. Desensitization mechanisms in the computational consist of internalization of activated receptor and g-protein and RGS which negatively regulates G-protein. By modulating these desensitization mechanisms, we hoped to confirm which desensitization mechanism mimics the profile seen experimentally. Our results show these dynamics alone could not copy the dynamics observed experimentally. Additionally, the equation based computational model could not properly model receptor retention caused by RGS binding due to not enough known about the spatial effects. To remedy this, we created an agent-based model of receptor dynamics to determine how receptor retention contributes to GPCR signaling.

Method

ODE Based Receptor G-Protein Model

We developed a computational model to investigate RGS activity and endocytosis frequency effect on the Ga-GTP distribution during the pheromone response. The model is based on ordinary differential equations. The pheromone initiates the cascading protein pathway of the pheromone response along the membrane, an array of length 63, to find the mean distribution of Ga-GTP on the membrane. The model consists of eight coupled ordinary differential equations, one for each protein, and 10 parameters using ordinary differential equation solver ode23 in MATLAB. The model simulates 10 minutes of the active pheromone response. Endocytosis events were modeled by removing a small percentage of the membrane based on endocytic marker gaussian fit (MATLAB’s curve fitting tool), ede1 distribution. Exocytosis events were modeled by adding inactive proteins to the membrane based on exocytic marker gaussian fit (MATLAB’s curve fitting tool), exo84 distribution. The Monte Carlo method
of parameter fitting adjusted the gaussian distribution, scaling and baseline of the endocytosis and
exocytosis profiles to fit the experimental MAPK profile, normalized and minimum subtracted. The
Monte Carlo to fit the simulation wildtype Gα-GTP to experimental data of MAPK ran for 1500 iterations
with a standard deviation of 0.02. The model parameters with the top score was used to compare the
fold change in the RGS activity and endocytosis frequency on the Gα-GTP distribution.

The following are the ODE based equations used in this model:

\[
\frac{d(\text{Ste2})}{dt} = -k_1 \times [\text{Ste2}] \times [\text{pheromone}] + k_{1r} \times [\text{Ste2-GEF}]
\]

\[
\frac{d(\text{Ste2-GEF})}{dt} = k_1 \times [\text{Ste2}] \times [\text{pheromone}] - k_{1r} \times [\text{Ste2-GEF}]
\]

\[
\frac{d(\text{Gα-GDP-Gβγ})}{dt} = -k_2 \times [\text{Gα-GDP-Gβγ}] \times [\text{Ste2-GEF}] + k_{2r} \times [\text{Gα-GDP-Gβγ}] + k_5 \times [\text{Gβγ}] \times [\text{Gα-GDP}]
\]

\[\quad - k_{5r} \times [\text{Gα-GDP-Gβγ}]\]

\[
\frac{d(\text{Gα-GTP-Gβγ})}{dt} = k_2 \times [\text{Gα-GDP-Gβγ}] \times [\text{pheromone}] - k_{2r} \times [\text{Gα-GDP-Gβγ}] + k_{3r} \times [\text{Gα-GTP}] \times [\text{Gβγ}]
\]

\[\quad - k_3 \times [\text{Gα-GTP-Gβγ}]\]

\[
\frac{d(\text{Gα-GTP})}{dt} = -k_{3r} \times [\text{Gα-GTP}] \times [\text{Gβγ}] + k_3 \times [\text{Gα-GTP-Gβγ}] - k_4 \times [\text{Gα-GTP}] \times [\text{RGS}] + k_{4r}
\]

\[\quad \times [\text{Gα-GDP}] \times [\text{Ste2-GEF}]\]

\[
\frac{d(\text{Gβγ})}{dt} = k_3 \times [\text{Gα-GTP-Gβγ}] - k_{3r} \times [\text{Gα-GTP}] \times [\text{Gβγ}] - k_5 \times [\text{Gβγ}] \times [\text{Gα-GDP}] + k_{5r} \times [\text{Gα-GDP-Gβγ}]
\]

\[
\frac{d(\text{Gα-GDP})}{dt} = k_4 \times [\text{Gα-GTP}] \times [\text{RGS}] - k_{4r} \times [\text{Ste2-GEF}] \times [\text{Gα-GDP}] - k_5 \times [\text{Gβγ}] \times [\text{Gα-GDP}] + k_{5r}
\]

\[\quad \times [\text{Gα-GDP-Gβγ}]\]

Table C.1 ODE Receptor Model Parameters

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
<th>Notes</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>$k_1$</td>
<td>5.6</td>
<td></td>
<td>(Bajaj et al., 2004b)</td>
</tr>
<tr>
<td>$k_{1r}$</td>
<td>0.0392</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Based on $K_d = 33\text{nM}$ (Alves et al., 2003, 2005; Bush et al., 2016) and $k_{2r} = 0.1$ (Bush et al., 2016; Hein et al., 2006)

<table>
<thead>
<tr>
<th>$k_2$</th>
<th>3.3</th>
<th>Based on $K_d = 33\text{nM}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$k_{2r}$</td>
<td>0.1</td>
<td></td>
</tr>
<tr>
<td>$k_3$</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>$k_{3r}$</td>
<td>0.01</td>
<td></td>
</tr>
<tr>
<td>$k_4$</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>$k_{4r}$</td>
<td>3.3</td>
<td>$k_{4r} = k_3$</td>
</tr>
<tr>
<td>$k_5$</td>
<td>32</td>
<td>$K_d=0.001$; used $k_d=0.0001$ (Bush et al., 2016)</td>
</tr>
<tr>
<td>$k_{5r}$</td>
<td>0.0032</td>
<td></td>
</tr>
</tbody>
</table>

Results

RGS Phosphorylation spatially alters MAPK localization

By computationally simulating a wildtype GPCR signaling event and its desensitization mechanisms, we can determine their spatiotemporal effect on receptor signaling. Desensitization is an essential mechanism in the yeast pheromone pathway to maintain polarity in the system. Desensitization mechanisms negatively regulate receptor mediated signaling such as internalizing active GPCRs and inactivating Gα signaling (REDO NO Dixit et al., 2014; Suchkov et al., 2010a). The negative regulator of the GPCR, RGS, is phosphorylated on its 539th amino acid by MAPK in response to pheromone (Garrison et al., 1999). Recent work has shown this dynamic phosphorylation of alters G-protein activity spatiotemporally (Simke et al., 2020). These results suggest RGS phosphorylation may regulate GPCR signaling desensitization. To determine if RGS activity or endocytosis is altered by RGS phosphorylation, we created a computational model of GPCR signaling to determine the effect of desensitization mechanisms on G-protein activation localization.
Receptor G- protein Desensitization Computational Model

The simulation mathematically models the pheromone bound receptors acting as a GEF to G-protein by ODEs (equations and rates are listed in the Methods section Table C.1). The concentration of each species are spatially modeled in discrete bins relative to the polar cap. In the model, receptor binds pheromone and activates its G-Protein Gα subunit (Gα-GTP) causing dissociation from its G- Protein complex. The Gα-GTP is inactivated by RGS and converted to Ga-GDP to then re-form its inactive G-Protein complex (Figure 2.1). Fitting the Computational Model to the Experimental Data

To compare the model with experimental data, we first fit the simulation to experimental MAPK data due to the similarity in Gα-GTP localization and MAPK (Fus3) localization. Associating MAPK localization with Gα-GTP localization allows us to determine the location of active G-protein experimentally so the same distribution of active G-Protein can be simulated computationally. The MAPK experimental data was obtained by quantifying the MAPK fluorescence during a time course experiment aligned to the polar cap (through maximal Bem1-Ruby fluorescence).

For the computational model simulates active G-Protein and is compared to experimental MAPK distribution, we first found the distribution of an endocytic marker, ede1, and exocytic marker, exo84, during the yeast pheromone response. Quantified membrane fluorescence protein profile of both proteins were aligned to the polar cap (maximal Bem1-Ruby fluorescence), normalized and fitted to gaussian profiles. The distributions of endocytic and exocytic markers were used in vesicle trafficking events in the simulation to either internalize active protein from the membrane or secrete inactive protein to the membrane. The gaussian distributions of endocytosis and exocytosis in the model were parameter were fit for the Gα-GTP profile to match experimental MAPK data based on the Monte Carlo method (Figure 2.1B).

Altering Desensitization Mechanisms in the ODE Model Did Not Accurately Model G- protein Signaling
Experimentally we found mutating GPCR negative regulator, RGS alters the MAPK spatiotemporal distribution. We wanted to know if this altered distribution is due to the mutated RGS altering internalization rates or receptor activity. By altering the endocytosis frequency or RGS, we compared those computational model profiles to the MAPK mutant RGS mutants in order to determine the effect of RGS mutants on desensitization mechanisms. To do this, we first altered RGS activity in the system tenfold from the simulation fitting the wildtype profile. Altering RGS activity tenfold decrease caused peripheral active Gα increase, overall decreasing polarity (Figure 2.1D). Fold change increase in RGS activity decreases polarization through widening the width of active Gα localized to the polar cap. Neither of the mutant RGS experimental MAPK profiles mimic the profile of an up or downregulation in RGS activity.

Next, we altered the rate of endocytosis because RGS modulates receptor retention (Henderson et al., 2019b). We altered the endocytosis rate tenfold increase and decrease in frequency and compared the active Gα signaling distribution to wildtype and the mutant MAPK profiles. Increasing endocytosis frequency in the simulation decreased the fraction of active Gα at the polar cap to a degree which there was a higher fraction of active Gα on the periphery that the polar cap. By decreasing endocytosis, the width of the polarized active Gα increased, decreasing polarization, and not mimicking the experimental MAPK results.
Katherine Jarvis was born in Tyngsborough, MA on September 27, 1996. She was raised in Tyngsborough, MA and graduated from Tyngsborough High School in 2015. She attended the University of Massachusetts Amherst and graduated in 2019 with a Bachelor’s degrees in Biochemistry. She moved to Maine to start her Master’s Biochemistry program at The University of Maine in 2019. After receiving her degree, Katherine will be joining the work force to begin her career in Biotech in Boston, MA. Katherine is a candidate for Master of Science degree in Biochemistry from the University of Maine in D