Multi-Scale Modeling of the Innate Immune System: A Dynamic Investigation into Pathogenic Detection

by

Robert W. Gregg

Bachelor of Science in Chemical Engineering, University of Rochester, 2015

Submitted to the Graduate Faculty of

the Swanson School of Engineering in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2020

UNIVERSITY OF PITTSBURGH

SWANSON SCHOOL OF ENGINEERING

This dissertation was presented

by

Robert W. Gregg

It was defended on

July 20, 2020

and approved by

Jason E. Shoemaker, PhD, Assistant Professor, Department of Chemical and Petroleum

Engineering

Saumendra N. Sarkar PhD, Associate Professor, Department of Microbiology and

Molecular Genetics

Morgan V. Fedorchak PhD, Assistant Professor, Department of Chemical and Petroleum

Engineering, Bioengineering

Giannis Mpourmpakis PhD, Associate Professor, Department of Chemical and Petroleum

Engineering

Dissertation Director: Jason E. Shoemaker, PhD, Assistant Professor, Department of

Chemical and Petroleum Engineering

Multi-Scale Modeling of the Innate Immune System: A Dynamic Investigation into Pathogenic Detection

Robert W. Gregg, PhD

University of Pittsburgh, 2020

Having a well-functioning immune system can mean the difference between a mild ailment and a life-threatening infection; however, predicting how a disease will progress has proven to be a significant challenge. The dynamics driving the immune system are governed by a complex web of cell types, signaling proteins, and regulatory genes that have to strike a balance between disease elimination and rampant inflammation. An insufficient immune response will induce a prolonged disease state, but an excessive response will cause unnecessary cell dead and extensive tissue damage. This balance is usually self-regulated, but medical intervention is often necessary to correct imbalances. Unfortunately, these therapies are imperfect and accompanied by mild to debilitating side-effects caused by off-target effects. By developing a detailed understanding of the immune response, the goal of this dissertation is to predict how the immune system will respond to infection and determine how new potential therapies could overcome these threats.

Computational modeling provides an opportunity to synthesize current immunological observations and predict response outcomes to pathogenic infections. When coupled with experimental data, these models can simulate signaling pathway dynamics that drive the immune response, incorporate regulatory feedback mechanisms, and model inherent biological noise. Taken together, computational modeling can explain emergent behavior that cannot be determined from experiment alone. This dissertation will unitize two computational modeling techniques: ordinary differential equations (ODEs) and agent-based modeling (ABMs). Ultimately, they are combined in a novel way to model cellular immune responses across multiple length scales, creating a more accurate representation of the pathogenic response.

TLR4 and cGAS signaling are prominent in a number of diseases and dysregulations including—but not limited to—autoimmunity, cancer, HIV, HSV, tuberculosis, and sepsis. These two signaling pathways are so prevalent because they are activated extremely early

and help drive the downstream immune signaling. Modeling how cells dynamically regulate these pathways is critical for understanding how diseases circumvent feedback mechanisms and how new therapies can restore immune function to combat disease progression. By using ODE and ABM techniques, these studies aim to incrementally expand our knowledge of innate immune signaling and understand how feedback mechanisms control disease severity.
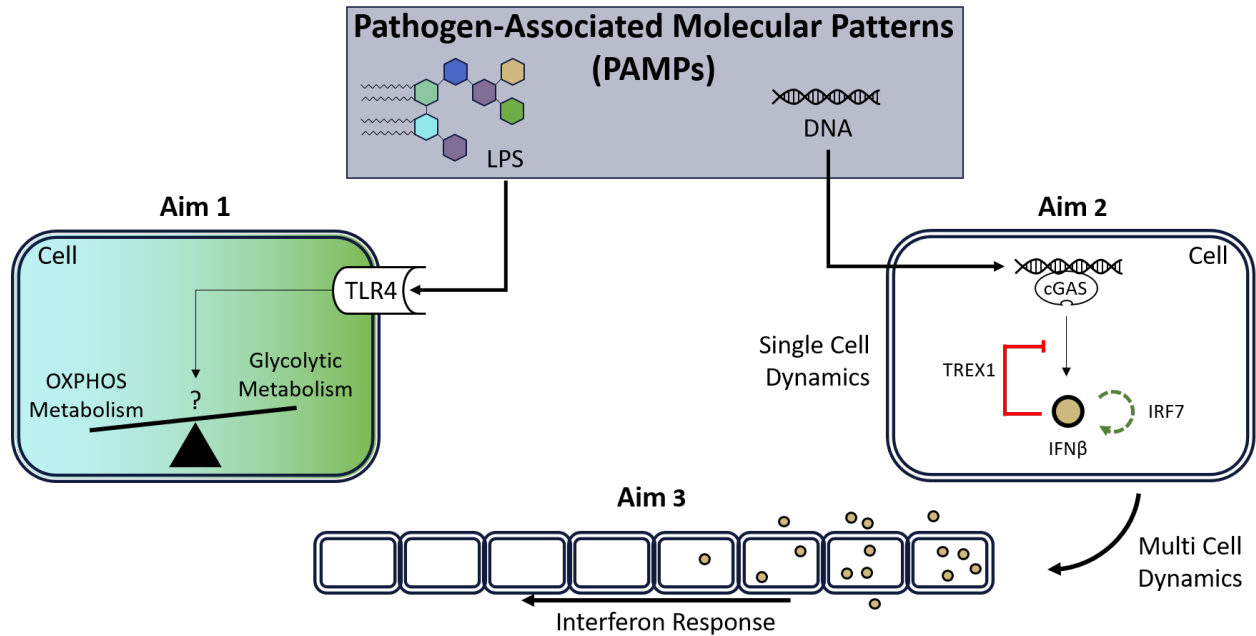

Figure 0.0.1: Visual Summary of Dissertation. Here, a visualization of the three aims in this paper are illustrated. Every aim is built on understanding pathogen associated molecular patterns either in the context of sepsis (Aim 1) or the cGAS pathway (Aim 2). Aim 3 extends the model constructed in the second aim into a full multi-cell, agent-based model.

Table of Contents

List of Tables

List of Figures

Preface

There are so many friends, family, and mentors that have provided support and propelled my education forward. I first want to thank my advisor Jason for guiding me through this PhD and putting in countless hours as we discussed every conceivable method of implementing an agent-based model. He was always willing to listen and provide direction when I lost sight of the larger research goal. Being one of his first students, we both learned and pushed each other to be better researchers and I value learning what it takes to be new PI. I want to thank my lab mates Emily, Muying, Jordan, and Lauren who have made graduate school a joy to experience. They have helped me overcome some major challenges and I hope we continue to support and help one another.

I want to thank my mom and dad for being my advocates during my early education. In the 3rd grade I struggled with reading comprehension but excelled in other subjects like math. The school administration dismissed claims from my teachers that I needed additional help, but they fought tooth-and-nail to get me a reading specialist who eventually diagnosed me with a reading disability. After a few months of painfully slow reading sessions my reading comprehension jumped three grade levels. They continue to help me work toward higher education and opportunities they did not have when growing up and I am grateful for their love and encouragement.

I also want to thank my significant other Megan, who's immense success is something I strive for in my own life. She has loved and supported me throughout this journey and helped me realize the importance/benefits of work-life balance. I want to thank my sister for making me laugh every time we talk long-distance on the phone, my friends both here in Pittsburgh and far away that have stuck with me through this wild experience, my teachers at Durfee and Rochester for believing I could succeed, and the students I've taught for solidifying my passion for teaching.

I finally want to thank my committee for pushing this work to its fullest potential. I am grateful to them and the University of Pittsburgh at large for allowing me to pursue my interest and be a part of the academic community.

1.0   Introduction

1.1   Innate Immune Signaling

1.1.1   Sepsis and Metabolism

When a cell detects a pathogen, a series of signaling events are triggered that lead to the expression of genes that encode for potent immune signaling proteins. Some of these proteins (i.e. cytokines) leave the infected cell to warn neighboring cells about the detected threat, such as IFNβ [1], others will alter cellular function to prevent invading pathogens from utilizing cellular machinery [2], and there are even proteins that are secreted by threatened cells to induce apoptosis (programmed cell death) [3]. Unfortunately, this pathogen detection mechanism is imperfect and can either be deceived (false negative response) or inappropriately activated (false positive response). Cancer is prime example of a false negative response where the immune system should remove malignant cells but is subverted [4]. Cancer cells can accomplish this by expressing cell surface proteins that reduce immune cell potency [5]. Conversely, sepsis is an example of a false positive response where the immune system is activated unnecessarily. Figure 1.1.1 demonstrates how the immune system can be activate far from the local source of infection.

Sepsis typically begins with a local infection isolated to one organ or area in the body. Under normal circumstances, this pathogen can be removed by the immune system with ease. However, if a pathogen or pathogen fragment enters the circulatory system, it can leave the site of infection and relocate to a different organ (often the kidneys because they filter waste from the blood). The pathogenic fragments are recognized by these distant cells and they begin to trigger signaling pathways that upregulate more immune signaling proteins. This process cascades until the entire body is in a state of rampant, uncontrolled inflammation [6, 7].

One potential, and deceptively simple treatment to combat sepsis is to administer anti-inflammatory medication to quell the exacerbated immune response. Unsurprisingly, nothing

1
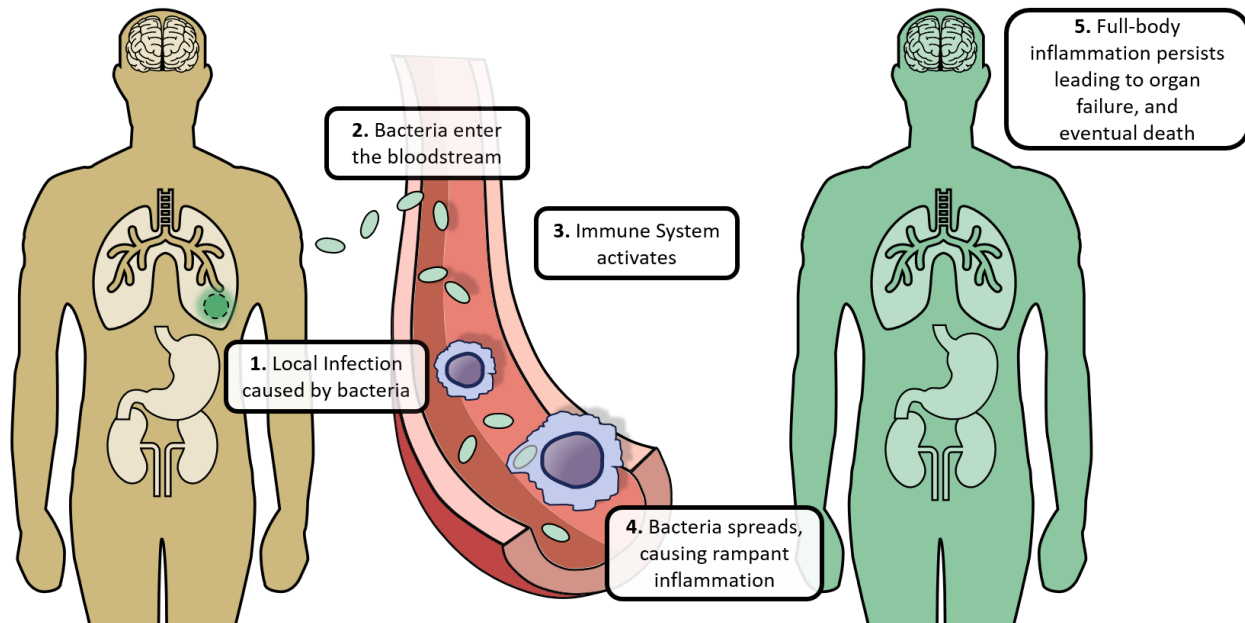
Figure 1.1.1: Sepsis Overview. Sepsis is thought to be caused by the spread of pathogenic signatures away from local infection. Here bacteria from the lungs are shown entering the blood stream and being recognized by leukocytes. Without proper regulation, these processes can lead to lethal, full-body inflammation.

is this simple in medicine. Pharmaceutical intervention aimed to suppress the immune system runs the risk of allowing the original infection to worsen. The treatment would have to be given after the infection is cleared, but before the patient succumbs to organ failure. Paradoxically, even when controlling for this scenario, sepsis clinical trials using anti-inflammatory drugs notoriously fail. Dr. Bone wrote a commentary on this issue over 20 years ago [8] and still today current CDC guidelines to treat sepsis are limited to antibiotics, maintaining blood flow to organs, and treating the source of the infection [9]. Despite these efforts, 30% of the 1.7 million sepsis cases in the United States each year end in death [10]. To improve our chances of developing successful sepsis treatments, we need to look beyond simply reducing inflammation and investigate how else this condition impacts cellular function.

In recent years, there has been increasing evidence that metabolism plays a crucial role during sepsis [11, 12, 13]. Metabolism, generally, is the collection of chemical reactions

that occur throughout the cell to maintain cellular function. Different metabolic pathways are classified as either anabolic, meaning they use energy to synthesize new molecules, or catabolic where they break down molecules to produce energy. The metabolic pathways affected by sepsis are catabolic and are, namely, the oxidative phosphorylation (OXPHOS) and aerobic glycolysis pathways. OXPHOS is considered the default metabolic pathway that occurs in the mitochondria to produce ATP and relies on a continuous supply of oxygen to fuel the electron transport chain, a multi-step redox process the generates an electrochemical gradient to drive the synthesis of ATP. Glycolysis, on the other hand, occurs in the cytosol and is composed of a series of chemical reactions that transforms glucose into pyruvate, generating ATP in the process. This process can occur without oxygen, but will eventually convert pyruvate into lactate, a metabolic dead-end that cannot be further decomposed to produce energy. Typically, cells will use OXPHOS to produce ATP and switch to glycolysis when oxygen is depleted (hypoxia), however, during sepsis, cells are observed to switch to glycolysis despite normoxic conditions (hence aerobic glycolysis) [11, 14, 15]. This is often referred to as the famous Warburg effect which is utilized by cancer cell to meet their high energy requirements [16].

Understanding which metabolic pathway is active during sepsis is important because it determines how a cell's immune system will respond. Cells shifted toward aerobic glycolysis will enter a resistive state, meaning a pathogenic threat is actively being responded to by the immune response, increasing inflammation (see Figure 1.1.2). Cells that remain in OXPHOS exhibit what is known as a tolerant state where cells reduce inflammation to avoid injury caused by an excessive immune response.

Given this idea of resistive and tolerant states associated with metabolism, one would hypothesize that shifting cells away from aerobic glycolysis and toward OXPHOS could improve the outcome of sepsis. However, there is still much to learn about this mechanism as recent studies have suggest that cells undergo biphasic or even cyclic shifts between metabolic states [17, 13]. A single shift toward OXPHOS might not be sufficient. To understand metabolic shifting, computational modeling approaches are used to simulate pathways controlling metabolism. If models can reconstruct this mechanism, they could be used to test different septic states to determine how this shifting results from known signaling pathways.
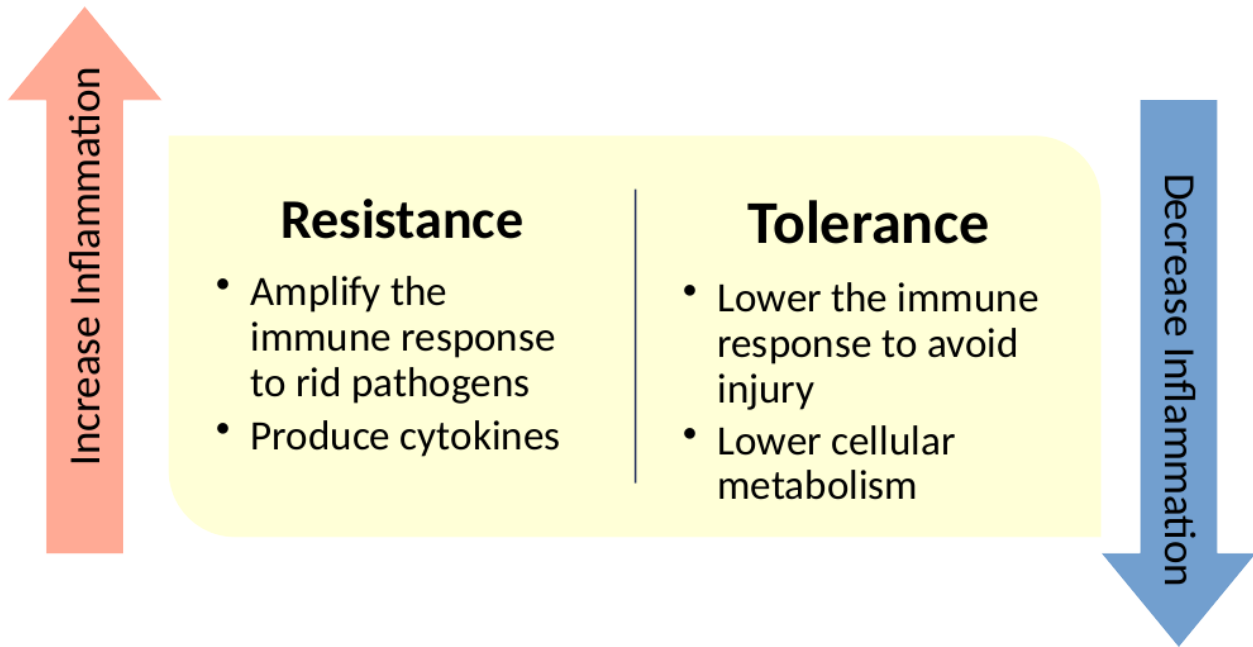
Figure 1.1.2: Balancing Metabolic States. Cells have evolved to leverage one of two immunologic states: Resistive or Tolerant. A resistive state occurs when a cell is actively preventing the advancement of a pathogen. Cells in this state utilize glycolytic metabolism. A tolerant state allows a cell to ignore an infection to reduce the risk of injury caused by the immune system. Cells in this state utilize OXPHOS metabolism.

1.1.2 cGAS Pathway: Sensing DNA

The story behind cGAS begins with a mystery about how cells detect viral invaders. A virus is more or less a package of genetic material surrounded by a capsid, or shell of proteins that protect this genetic material and facilitate its transport. The genetic material can vary in structure depending on the type of virus. Influenza A, for example, is comprised of eight single-stranded RNA molecules that encode for capsid proteins as well as immune suppressing proteins to ensure the success viral proliferation [18]. Herpes Simplex Virus (HSV-1), in contrast, is encoded by a one double stranded DNA molecule that encodes for the entire virus. Still other viruses like human immunodeficiency viruses (HIV) have more complex life cycles because it is encoded by RNA which is then reversed transcribed into DNA when HIV enters a host. The immune system needed to evolve ways to identify this foreign genetic material and differentiate it from its own DNA and RNA molecules. If a cell

could make this distinction, an immune response could be triggered to combat the virus and return to a healthy state.

The method in which cells detect foreign RNA has been established and well understood. It has been shown that RNA is detected through RIG-I like receptors (RLRs, [19]) and toll-like receptors (TLRs, [20]) which promote the activation of transcription factors (e.g. IRF3 and IRF7) and production of type I interferon (IFNα and IFNβ) [21]. Interferon was given this named because it was observed to "interfere" with viral infections. Interferon acts as a master regulator for hundreds of interferon stimulated genes (ISGs) that prevent viruses from proliferating at every step of the infection process [2].

In contrast, the mechanism underlying the recognition of double stranded DNA (dsDNA) in the cytosol remained elusive. Beginning in 2006, there was evidence for dsDNA inducing an IFNβ response, but no definitive signaling pathway was known to explain this observation [22]. Over the next seven years, there were no fewer than seven potential DNA sensors identified but none were met with universal acceptance [23]. This mystery remained unsolved until 2013 when Sun *et. al.* provided the first definitive evidence that cyclic GMP-AMP synthase (cGAS) was indeed the putative molecular sensor of cytosolic dsDNA [24].

They further demonstrated that cGAS binding of dsDNA catalyzes the formation of cyclic guanosine monophosphate adenosine monophosphate (cGAMP), a secondary messenger molecule the binds to a protein bound to the ER called stimulator of interferon genes (STING). As it names suggests, STING undergoes a conformational change after binding to cGAMP, translocates to the perinuclear space to phosphorylate the transcription factor IRF3 [25], which then upregulates interferon transcription (for a more detailed diagram of this signaling pathway, see Figure 3.1.1). This connects cGAS to the more-established RIG-I and TLR pathways, finally creating a clear molecular mechanism linking dsDNA detection to the interferon response.

While the cGAS pathway was identified in the context of viral infections, it also plays pivotal roles in wide range of biological contexts. Many types of cancers, for example, have mutated to specifically downregulate cGAS expression because it bolsters an immune response to prevent tumor growth [26, 27]. Surprisingly, there are instances where the cGAS pathway promotes inflammation-mediated tumor growth [28, 29]. The interplay between

cancer and cGAS is more complex than previously thought and would greatly benefit from modeling techniques derived from systems biology.

Of course, when dysregulation occurs in an immune response pathway, autoimmune diseases are likely to surface. The cGAS pathway is no exception to this and has been implicated in lupus erythematosus [30] as well as rarer diseases such as Aicardi-Goutieres syndrome (AGS) where a primary deoxyribonuclease, TREX1, is functionally inactive [31]. TREX1 is responsible for degrading DNA signals that activate cGAS, which if left untreated, can lead to chronic inflammation.

Pathogens
- Viruses (HPV, HIV)
- Bacteria (Listeria, TB)

Cancer
- Anti-tumor (Prostate Cancer)
- Pro-tumor (Brain Carcinoma)

cGAS Pathway

Aging
- Cellular Senescence
- Stem Cells

Autoimmunity
- Lupus
- Aicardi–Goutières syndrome
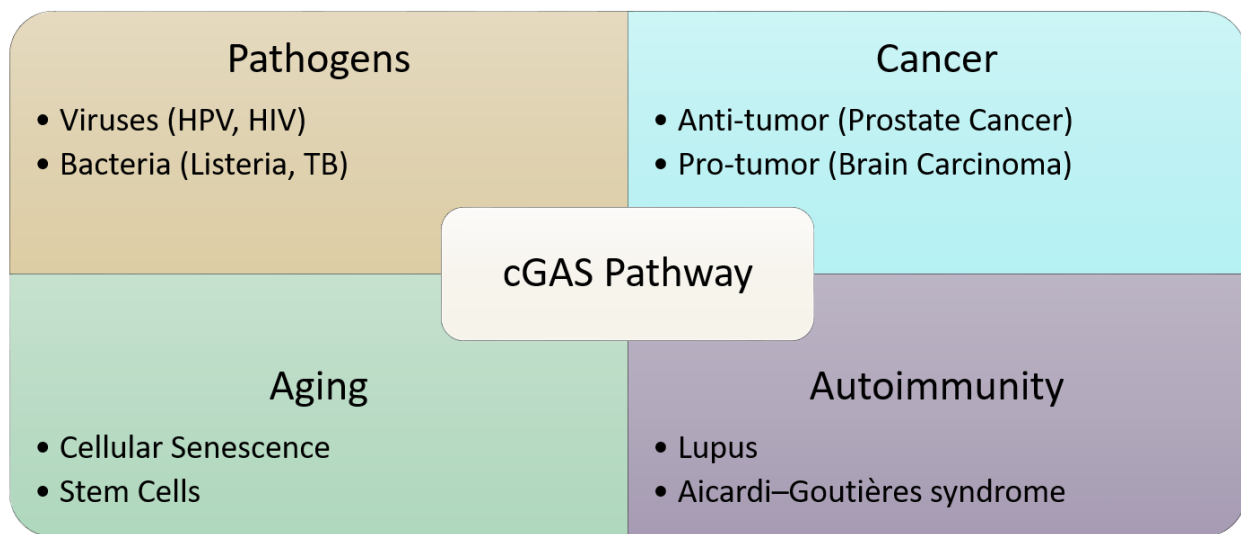
Figure 1.1.3: Roles of cGAS in Disease

Finally, cGAS has been discovered to be a critical component in cellular senescence [32], a process where cells have reached the limit to which they can undergo mitosis. This limitation is to ensure that the accumulation of mutations within the genome is controlled. Usually, after about 50 divisions, signaling events occur to cause the cell to apoptose. cGAS has been shown to detect damaged DNA the accumulates in the cytosol due to errors in cell division [32, 33]. Senescence also have a significant impact on stem cell development as these cells are rapidly dividing and differentiating [34].

Because the success of cGAS signaling is linked to an incredible number of biological processes, it is imperative that this potent immune response be tightly regulated. There are

a few key downstream regulators that control the cGAS pathway and subsequent interferon response, one being the previously mentioned TREX1 protein. cGAS activates TREX1 which then degrades DNA to subdue cGAS activity. Another class of proteins that inhibit the downstream effects of cGAS are Suppressor of cytokine signaling (SOCS) proteins. These class of proteins are upregulated by an interferon response and works to stop incoming interferon signals by inhibiting its surface receptor [35].

cGAS signaling can also be enhanced by positive feedback mechanisms. The transcription factor IRF3 is activated by cGAS to promote the expression of IFNβ. As interferon protein concentrates, another transcription factor (IRF7) is expressed to overtake the role of IRF3 [36]. Because IRF7 is induced by IFNβ which produces more IRF7, strong inflammatory signals can be amplified in this positive feedback loop if left unchecked.

This leads to a need to quantify how the cGAS pathways is regulated and predict how the immune system changes when the pathway is manipulated (either artificially through pharmaceutical intervention or gene knock down). The immune system is a nonlinear dynamic system, which implies that changes in inputs (e.g. the amount of DNA) do not scale with changes in outputs. Reducing the activity of cGAS by 50% will not coincide with an equal reduction in IFNβ expression. In fact, some cells within a population will exhibit no interferon expression despite receiving immunostimulatory DNA (ISD) [37]. To predict how the immune response is going to evolve over time, we need to turn to computational model approaches. In conjunction with experimental observations, these modeling paradigms can help propel our understanding of cGAS pathway regulation and help identify new system behaviors that could not be determined though experiment alone.

1.2 Computational Modeling

1.2.1 Using Ordinary Differential Equations in Systems Biology

There are two main approaches in biological research that are used to understand experimental observation. The first, more traditional method, takes a reductionist approach by breaking down complex biological phenomena into its basic components until the underlining cause is found. Gene knockout studies are an excellent example of this approach. By isolating a primary function of a gene or mechanism within a cell, critical knowledge is gained which can be used to refine a signaling pathway that was previously unresolved or propose a new pharmaceutical target. An insurmountable number of discoveries have been found using this approach and it is usually regarded as the gold-standard for biological research. However, this approach is not infallible due to the complex, non-linear relationships that drive biological mechanisms.

Systems biology takes the opposite approach by considering multiple biological components simultaneously. This more recent, holistic approach benefits from information derived from traditional experiments, and can further identify behavior that only emerges when multiple components come together. Kitano outlines this idea effectively in his seminal paper, *Systems Biology: A Brief Overview*:

> "Because a system is not just an assembly of genes and proteins, its properties cannot be fully understood merely by drawing diagrams of their interconnections. Although such a diagram represents an important first step, it is analogous to a static roadmap, whereas what we really seek to know are the traffic patterns, why such traffic patterns emerge, and how we can control them." [38].

This illustrates an important distinction between the structural and dynamic aspects of biology. Knowing a protein is inhibited by another is important, but it provides minimal information about how that protein interacts with other biological components within the entire cellular system.

A classic example of successfully utilizing systems biology is demonstrated in Hoffmann's work investigating the regulation of NF-κB. This protein is associated with a myriad of biological processes including the innate and adaptive immune response, apoptosis, survival,

and cell growth [39]. NF-κB is perpetually inhibited by IκB proteins through direct binding, which prevents it from acting as a transcription factor to promote gene expression. When specific signals are detected by a cell (e.g. viral RNA), an enzyme phosphorylates IκB, releasing NF-κB, and allowing it to translocate to the nucleus. Interestingly, IκB genes are one of the many genes regulated by NF-κB, meaning its activation eventually leads to auto-inhibition. This probably would have been the endpoint to understanding of NF-κB signaling if the entire system was not further considered. Undeterred, Hoffman continued his examination of the pathway and discovered NF-κB levels can oscillate depending on the dynamics of the input stimulus. This had major implications for which genes NF-κB could activate. Some genes like IP-10 would be expressed after a short impulse (15 minutes) of stimulus, but others like RANTES needed significantly longer stimulus exposures (2 hours) to observe detectable levels of protein. This temporal regulation effectively allows the cell to select which NF-κB genes are expressed despite all being regulated by the same transcription factor.

How were these NF-κB oscillations predicted? There are no traditional biological methods that would have predicted this observation which is why system biology approaches were implemented, specifically in this case ordinary differential equations (ODEs). This branch of mathematics focuses on how a system changes over time and are ubiquitous in science. They can be used to describe simple Newtonian motion like a swinging pendulum or more complicated phenomena like particle wave functions from quantum mechanics. In the case of NF-κB regulation, the following system of ODEs was used:
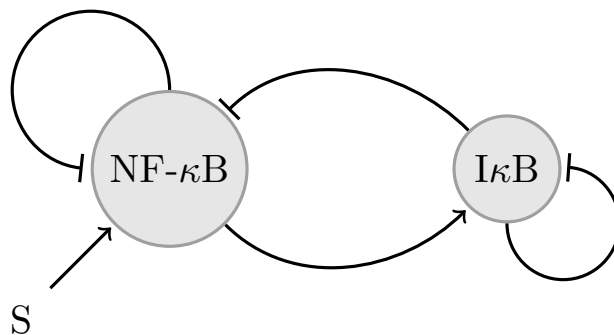
Figure 1.2.1: Two Component NF-κB

9

$$\frac{d[\text{NF-}\kappa\text{B}]}{dt} = S - \alpha[\text{NF-}\kappa\text{B}] - \beta[\text{I}\kappa\text{B}] \tag{1.1}$$

$$\frac{d[\text{I}\kappa\text{B}]}{dt} = \gamma[\text{NF-}\kappa\text{B}] - \delta[\text{I}\kappa\text{B}]. \tag{1.2}$$

These equations relate the change in protein concentration over time to the current protein concentrations in the cell. Here, both proteins have a rate of self degradation (α and δ) which captures the idea of proteolysis. The rate of IκB production is increased by NF-κB (through the γ term), allowing subsequent NF-κB production to be inhibited by the increasing IκB concentration (through the β term). This linear system of ODEs can exhibit a variety of behaviors depending on the values of the rate constants, and because the system is linear, we can easily determine what parameter values correspond with particular system outcomes. By vectorizing the system of equations:

$$\frac{d}{dt}\begin{bmatrix} \text{NF-}\kappa\text{B} \\ \text{I}\kappa\text{B} \end{bmatrix} = \begin{bmatrix} S \\ 0 \end{bmatrix} + \begin{bmatrix} -\alpha & -\beta \\ \gamma & -\delta \end{bmatrix}\begin{bmatrix} \text{NF-}\kappa\text{B} \\ \text{I}\kappa\text{B} \end{bmatrix}, \tag{1.3}$$

it can be shown that the eigenvalues of the rate constant parameter matrix determine the system behavior.

$$p(\lambda) = \det(A - \lambda I) \tag{1.4}$$

$$p(\lambda) = \det\left(\begin{bmatrix} -\alpha & -\beta \\ \gamma & -\delta \end{bmatrix} - \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix}\right) \tag{1.5}$$

$$p(\lambda) = \lambda^2 - (\alpha + \delta)\lambda + (\alpha\delta - \beta\gamma) \tag{1.6}$$

$$p(\lambda) = \lambda^2 + \text{Tr}A\ \lambda + \det A, \tag{1.7}$$

Here, Tr and det are the trace and determinant of the rate constant parameter matrix. Plotting the trace of the rate constant parameter matrix against the determinant, we can easily summarize all the possible behaviors of this system (Figure 1.2.2).

Because the trace of the NF-κB parameter matrix is always negative, and the determinant is always positive, the system will remain stable (staying within the blue region of Figure

10
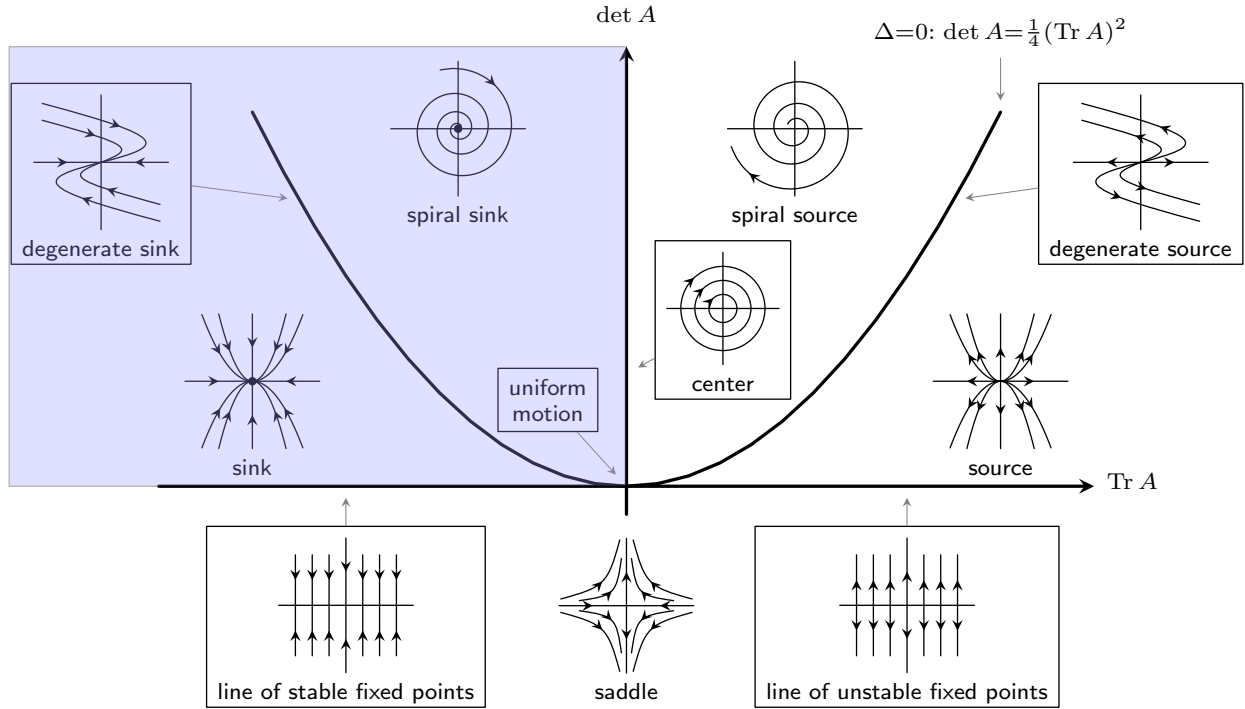
Figure 1.2.2: Characteristic Dynamics of a Coupled Linear System of ODEs. The blue region indicates stable solutions converging toward equilibrium. TrA and $detA$ denote the trace and determinant of the rate constant parameter matrix. Δ represents the discriminant of the characteristic polynomial (Eq 1.4). The different subplots show possible system behaviors where the two axes represent the two components evolving over time.

1.2.2). If the degradation rates for NF-κB and IκB (α and δ) were large and the interaction terms were large (γ and β), the system would smoothly converge to an equilibrium point (sink region). Conversely, if the degradation rates were large and the interaction terms small, the system would experience damped oscillations (spiral sink). This is precisely what was observed experimentally for dynamic NF-κB measurements [40].

Small, linear ODE systems are straightforward to analyze because we can rely on techniques developed in linear algebra. However, these systems become intractable to analyze after the introduction of more components (proteins) and non-linear terms (x^2, xy, $1/(a+x)$). This is the case if we want to model large signaling pathways using mass-action and Michaelis-Menten style kinetics (often used to describe enzymes). When a system describing a cell signaling pathway contains these non-linear terms we can either create a linear approxima-

tion (usually though Taylor series expansions) or—if the system is small enough—use phase portraits as shown in Figure 1.2.2. Although, anything beyond three states becomes too difficult to visualize using this method. One way to overcome these shortcomings is to simulate the system of ordinary differential equation (ODE) numerically.

There is a rich history behind solving ODEs numerically (see [41, 42]), but the algorithms can be largely classified into either explicit or implicit methods. Explicit methods can use the current time point to directly calculate the next, whereas implicit methods require a set of equations to be solved to resolve the next time point. This can be accomplished by estimating the derivative as a finite difference. While implicit methods tend to require additional computations, they also exhibit greater numerical stability. To demonstrate the advantages each class of algorithm brings, we will explore two of the simplest algorithms from each class.

Starting with the explicit Forward Euler Method, if we have a differential equation:

$$\frac{dy}{dt} = f(t, y(t)), \tag{1.8}$$

where $f(t)$ is some well-behaved (i.e. continuous) function, then we can make an approximation of the derivative by its definition:

$$\frac{y(t + \Delta t) - y(t)}{\Delta t} \approx f(t, y(t)). \tag{1.9}$$

We can now solve for the future value of our unknown function $(y(t + \Delta t))$ given the current value at time t

$$y(t + \Delta t) = y(t) + f(t, y(t)) \cdot \Delta t. \tag{1.10}$$

This works well if Δt is taken to be a small value. When the approximation in Eq. 1.9 was formed, a choice for indexing was made. It is just as correct to say:

$$\frac{y(t) - y(t - \Delta t)}{\Delta t} \approx f(t, y(t)), \tag{1.11}$$

which can be simplified to the familiar backward Euler algorithm.

$$y(t) = y(t - \Delta t) + f(t, y(t)) \cdot \Delta t. \tag{1.12}$$

But there is something wrong with Eq. 1.12, the new value for $y(t)$ cannot be isolated directed because it happens to be an argument for the function f. To get the next value for $y(t)$ the system would have to be solved for algebraically or estimated using fixed-point iteration.

A plethora of more advanced ODE algorithm exist. Some utilize the Jacobian to gain higher order derivative information about the system, others have adaptive methods that change Δt depending of the current error estimation. Despite these complexities, almost every ODE solver works by estimating the derivative with a finite difference and taking sufficiently small steps in time to solve the system. By employing these algorithms in a systems biology context, we can simulate how the concentrations of proteins and molecules change over time. This is an invaluable tool because it provides a way to predict everything from how cells will respond to changes in their environment down to how genes are activated over time.

1.2.2 Markov Chain Monte Carlo

Developing ODE models of cellular signaling pathways inevitably leads to a substantial number of unknown parameters that need to be determined. Conventionally, parameterization is performed by measuring all states within the system at a sufficiently dense number of points in time and fitting these measurements to simulation. This optimization would be accomplished by minimizing the sum-of-square differences between the measurements and model simulation, resulting in a best-fit value for each parameter. In addition to being intractable from an experimental perspective, this method is flawed because it ignores two problems in parameter optimization: local minima and identifiability. Local minima are regions within the parameter space that appear to give the best-fit value if you limit your search to a small neighborhood of surrounding points, whereas identifiability occurs when there is many parameterizations all equally minimize the objective function.
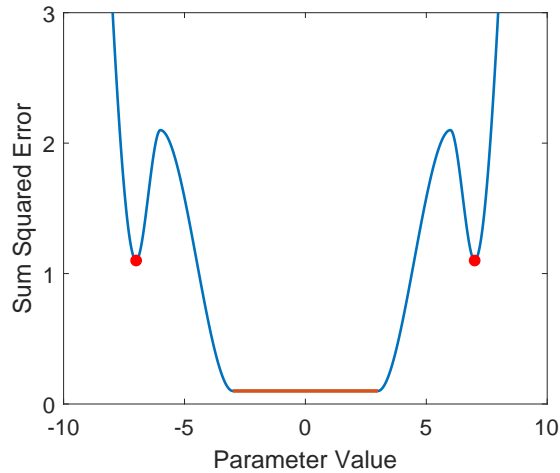
Figure 1.2.3: Example Objective function. Red dots indicate local minima and the orange line demonstrated a region of parameter space that is not identifiable.

Figure 1.2.3 gives a pedagogical demonstration of how tuning one parameter value might lead to either of these problems. Depending on the initialization of the parameter, common optimization methods (i.e. gradient decent algorithms) will either report a value of ± 7 (red dots indicating local minima) or within the range $[-3, 3]$ (orange line representing an identifiability issue). If the optimization was performed only once close to a local minimum, it would lead to an inaccurate conclusion about the best-fit parameter value. Alternatively, choosing an initialization anywhere within the non-identifiable region would lead to ambiguity into exactly what is the best-fit value. All the parameter values between -3 and 3 give the same sum-of-squared error between the measurements and simulation.

Overcoming either scenario requires sufficient topological knowledge of the parameter space, which can be challenging especially in higher dimensional spaces. To combat these problems, we turn to a Bayesian inference approach. Instead of determining what parameter value best fits the data, we instead ask: what is the probability a parameter takes on a particular value, given the experimental evidence? Posing the question in this way allows us to think of the parameters not as a single value, but as a collection of continuous values

14

each with an associated probability. This paradigm is dependent on Bayes Formula:

$$P(\theta|D) = \frac{P(D|\theta) \cdot P(\theta)}{P(D)}. \tag{1.13}$$

Here, we are trying to determine the posterior $P(\theta|D)$, the probability that a particular parameter set θ explains the given data D. The right side of the equation consists of the likelihood $P(D|\theta)$ which tells us how good the data fits to a given parameter set, the prior $P(\theta)$ which summarizes any information we might know about the parameters beforehand, and the evidence $P(D)$ which indicates if the data was generated by the model. The posterior probability cannot be calculated directly, so instead it is numerically estimated by sampling many parameter values using a method called Markov Chain Monte Carlo (MCMC).

The basic MCMC algorithm begins with some initial parameter set. To determine how good this parameter set is, we calculate the likelihood that our data came from this parameter set. If we assume the error in our data is Gaussian (with zero mean and unknown variance σ^2), the likelihood function reduces to the familiar sum-of-square error

$$P(D|\theta) \propto \exp\left(\frac{\sum_{i=1}^{n}(\tilde{y} - y_i)^2}{2\sigma^2}\right), \tag{1.14}$$

where $\tilde{y}$ is our simulated model being subtracted from each to the n total data points (y_i). With this information in mind, MCMC proposes to "jump" to a new location in parameter space and calculates the likelihood of the candidate parameter set. The algorithm now needs to decide if the candidate parameter set will be accepted, which is accomplished by calculating an acceptance ratio between $[0, 1]$ and comparing it to a uniform random number within that same range. The acceptance ratio is calculated by evaluating the ratios of the posterior distributions:

$$\alpha = min\left(1, \frac{P(\theta_{i+1}|D)}{P(\theta_i|D)}\right) = min\left(1, \frac{P(D|\theta_{i+1}) \cdot P(\theta_{i+1})}{P(D|\theta_i) \cdot P(\theta_i)}\right), \tag{1.15}$$

where θ_{i+1} represents the candidate parameter set, and θ_i the current parameter set. The acceptance ratio cleverly uses Bayes formula (Eq. 1.13) to calculate the posterior distributions using just the likelihood and prior distributions, both of which are known. The prior is conventionally given a uniform distribution if no information is known *a priori*. Note that if

the posterior of the candidate parameter set is greater than the current posterior ($\alpha = 1$), it will automatically be accepted when compared to a uniform random number. However, even if the candidate parameter set increases the total error, it could still be accepted with some probability. This ultimately allows the Markov chain to visit all of parameter space, but remain in areas longer if they have a high probability of containing the "best" parameter set, effectively eliminating the problem of local minima. In Figure 1.2.4 we can see how counting the number of times the Markov chain visits a parameter value gives us a distribution we can use to parameterize the model. Areas of high probability are visited more often by the Markov chain which leads to a larger peak on the histogram. Additionally, this algorithm sidesteps some identifiability issues because we can move our focus to investigating different model behaviors the top parameter sets exhibit as opposed to fixating on the best parameterization.
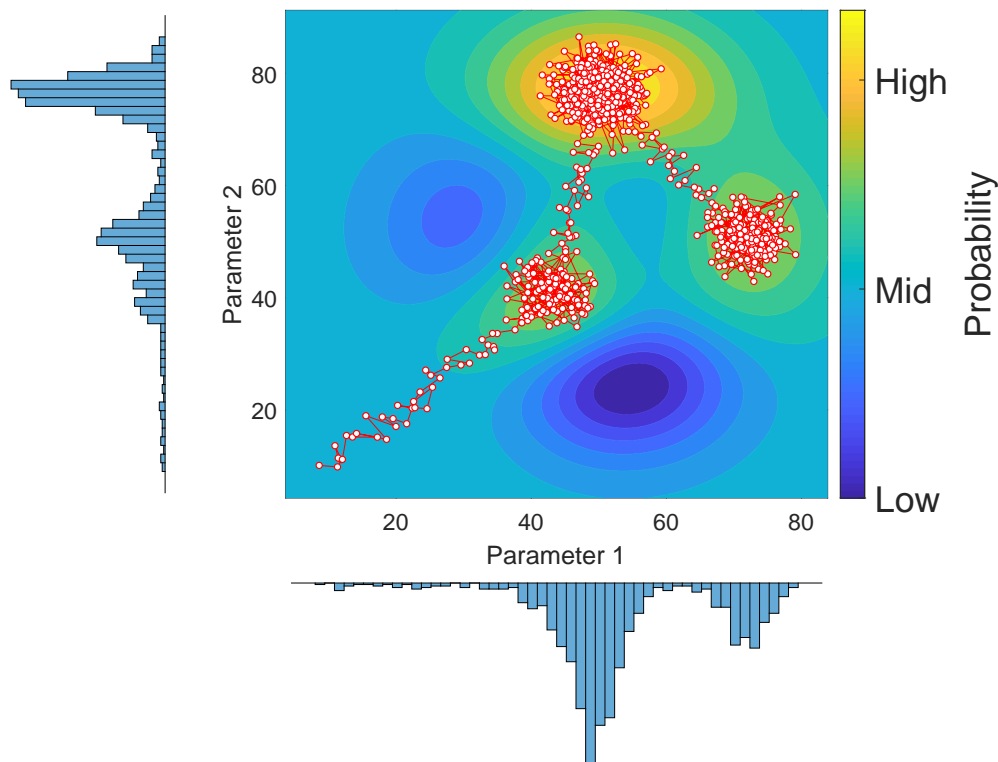
Figure 1.2.4: Demonstration of MCMC Parameterization

Many MCMC methods are available to perform parameter optimization, but based on a recent review by Ballnus *et al.* [43], we chose to use the toolbox PESTO to implement a

parallel tempering (PT)-MCMC algorithm in conjunction with multi-start local optimization. Briefly, PESTO begins by using an interior-point minimization routine in MATLAB (*fmincon*) initialized from 1,000 randomly distributed starting points. These starting points are determined using Latin Hypercube Sampling. (PT)-MCMC then uses these local minima and gradient information to initialize the algorithm. As it executes, multiple chains are run in parallel, each having an associated temperature that dictates how easily a new candidate parameter set is accepted (Figure 1.2.5). Increasing the temperature has the effect of flattening the posterior probably density. Chains are also allowed to swap locations which allows lower temperature chains to further investigate minima found by higher temperature chains.
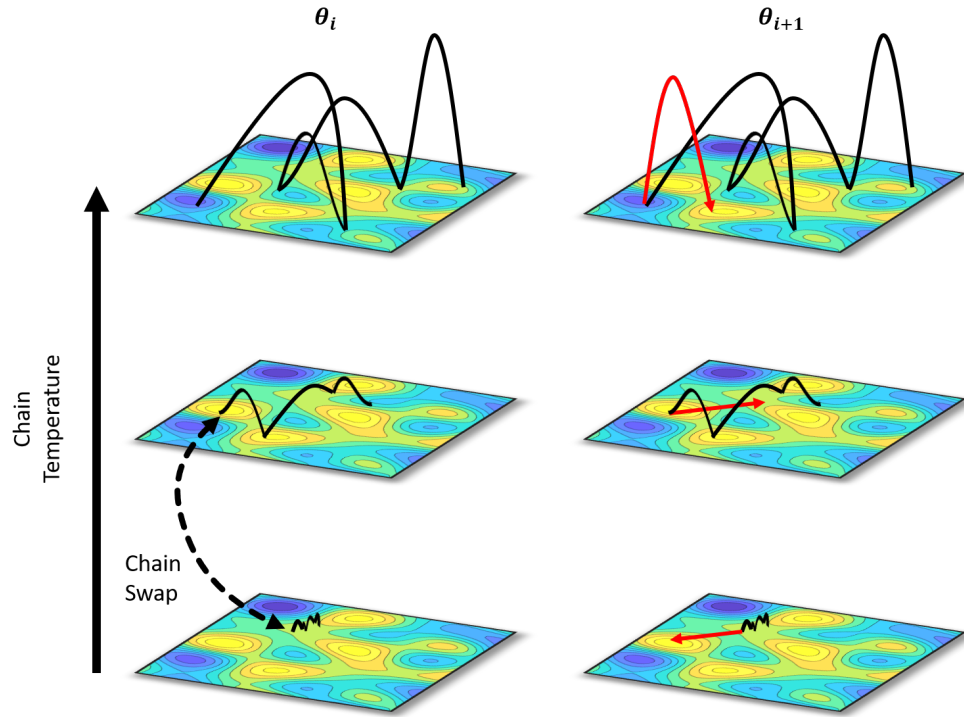
Figure 1.2.5: Combining Parallel Tempering with MCMC. The density plot represents a posterior probability distribution with areas of low (blue) and high (yellow) probability density. MCMC chains are depicted as jumping through parameter space, with higher temperature chains exploring disparate parameterizations more rapidly. Chains will randomly swap positions where, occasionally, lower temperature chains are placed in higher probability regions as demonstrated in the swap shown.

These amendments to the original Metropolis–Hastings algorithm are not necessary—in theory—because MH-MCMC is guaranteed converge to the target distribution. However, the rate of convergence in many circumstances is often intractable, so the addition of multi-start local optimization and parallel tempering routines are used to accelerate this process.

Determining when an MCMC simulation has converged is a difficult task and is still an active area of research. The best tools currently available for this are diagnostics and rules-of-thumb. Generally, we can only determine if a simulation has failed. Because the posterior distribution is unknown, we can never be certain if the entire distribution was sampled. One of the first and easiest ways to determine a successful MCMC simulation is to look at the parameter chain (i.e. the history of the parameter values). If these appear to randomly cover the parameter space, it would be evidence toward a healthy simulation, but if the values appear to depend on one another, this could indicate that the proposal jump is too small. This can be quantified by calculating the auto-correlation between the MCMC samples.

Another test for MCMC simulations is to examine the acceptance rate. If the algorithm is rejecting nearly every proposed jump, the time to convergence will substantially increase. Theoretical estimates put the ideal acceptance ratio at approximately 23% [44], but in practice, the assumptions that go into this estimate are often broken. Nevertheless, this acceptance ratio is often used as a rule-of-thumb.

The last, more computationally intensive diagnostic for MCMC is the Gelman-Rubin diagnostic [45]. This test requires that multiple simulations are run to compare how significantly different posterior estimates are within chains compared to between chains, If one chain happened to explore a unique area in parameter space, it would mean that the MCMC simulations have not yet to converged to the posterior distribution.

By performing all of these diagnostics and checking different rules-of-thumb, we can convince ourselves that the MCMC simulation(s) accurately capture the full posterior distribution.

1.2.3 Multi-Scale Agent-based Modeling

There are many approaches used to model how a cell dynamically evolves over time. ODEs, as described in Section 1.2.1, are useful for modeling intracellular dynamics because they provide a means to easily convert a chemical signaling pathway into a mathematical framework that predicts cell responses to stimuli (e.g. bacteria, cytokines, hypoxia). There are other methods beyond ODEs that, generally, exchange benefits and limitations to better answer specific research question. One example are Petri-Nets where each component in a signaling network is given a discrete number of tokens representing the concentration of that species [46]. Dynamics are simulated by moving tokens around the network. A node can transfer tokens across a connection if it has the required number of tokens matching or exceeding the connections toll (edge weight). This discretization can be useful in modeling small molecular concentrations typical to cells. Another commonly used method in system biology is Boolean modeling [47]. Here, each protein or gene in a network is only considered to be in an "on" or "off" state and logical operators are used to connect this binary states (e.g. if transcription factor 1 *and* transcription factor 2 are on, then gene A is also on). These types of simulations are useful if the signaling networks are large and amount of kinetic data is sparse. All of these methods can, to some degree, be used to model extracellular interactions within a cell population, but invariably make strong assumptions about the spatial layout of a cellular population. ODE modeling, for example, is required to assume no changes in spatial distributions over time. SIR (susceptible, infectious, recovered) models are a good example of this assumption in action. Typical cell-based SIR models [48] simulate cell population dynamics as an infectious agent causes healthy cells to transition infected and recovered states. The distribution of virus across the population is assumed to be equal and any cell within the population can be infected regardless of their neighboring cells' status. To incorporate important spatial information into these modeling schemes, a different approach needs to be utilized.

Agent-Based Modeling (ABM) is one such solution to this dilemma. ABM itself is an umbrella term that covers a wide range of different algorithms and modeling techniques, but all share some common attributes. First, agents (usually in our case cells) are generated and

given a set of intrinsic properties ranging from geometric (like volume or area) to domain specific (like cell type or initial velocity). Second, rules are put into place that instruct the agents how to behave in the simulation. For example, an agent with one attribute might be attracted to an agent with that same attribute. Rules might change the number of agents in a simulation, update specific cell attributes, or even occur randomly. The depth and complexity of these rules is only limited by imagination and computing power. Third, a simulation space for the agents to live in is defined. This could be as simple as a discrete grid of points or contain no grid in continuous 3D space. Agents could be defined as points or span a large region in the space defined. Agents could even be defined in other agents (this would be useful for modeling a mitochondrion inside a cell). Forth, and last, a schedule needs to be created that determines how the rules are enforced as each step in time. This is important because rule order impacts the result of the simulation. If I define two rules: (1) remove a random agent and (2) add two new agents, depending on the execution order, I could end up with a steady population of agents or infinitely many agents.

In addition to providing a foundation for modeling spatial variations in a cell population, ABMs are also useful for simulating the heterogeneous and stochastic properties impacting cellular signaling. Heterogeneity here refers to differences in molecular concentrations across each cell in the population. This can impact cell signaling potency because some cells could have an insufficient amount of protein to recognize the immunologic threats. Stochasticity, with regard to cell signaling, can be broken down into intrinsic and extrinsic sources. Intrinsic sources originate from the cell itself like low concentration levels, whereas extrinsic sources are extracellular events causing different cellular responses. ABMs can incorporate these observed behaviors in contrast to models like reaction-diffusion PDEs which do incorporate spatiotemporal variations but are deterministic in nature. Figure 1.2.6 categorizes the potential for ODE, PDE, and ABM modeling paradigms based on desired attributes. Ideally, models would be able to simulate large cellular populations that incorporate spatiotemporal, heterogeneous, and stochastic behaviors.

In recent years, modelers have found success in hybridizing ABMs with ordinary and partial differential equations [49, 50, 51]. The procedure to accomplish this is difficult because the available software to model advance ABMs (e.g. Netlogo [52]) is typically not integrated
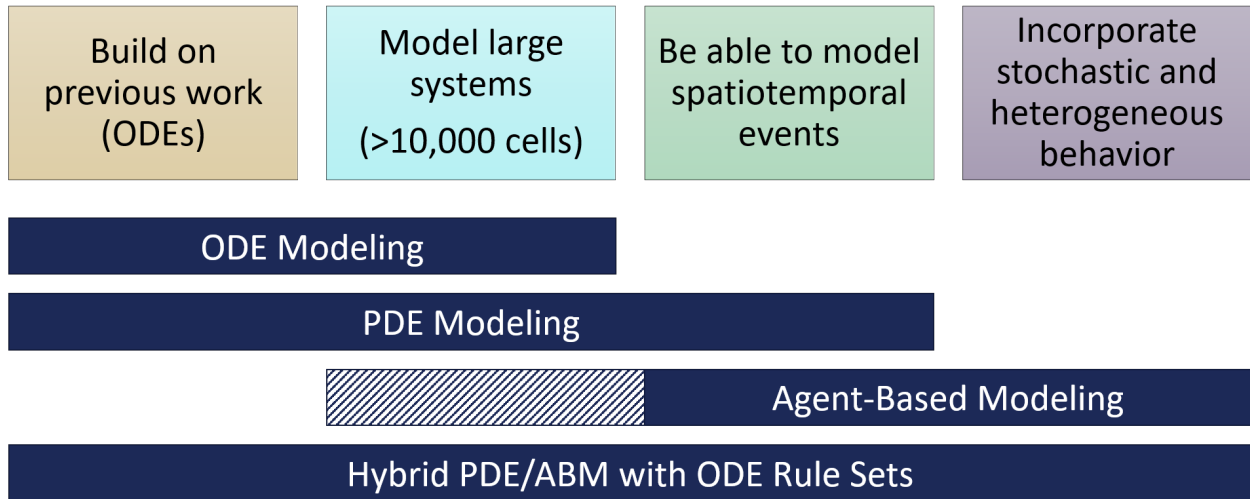
Figure 1.2.6: Comparing Model Strengths and Weaknesses

with advance ODE solver suites (e.g. SUNDIALS [53]). One method formulated in this dissertation to combine ABMs and ODEs works by adapting the DifferentialEquations.jl Julia package [54]. This package provides advanced, adaptive, ODE solvers that can handle large scale systems by leveraging algorithms like automatic sparsity detection, Jacobian coloring, and choices for linear solvers.

When the ODE solver successfully progresses to the next time point, a series of callback events are initiated. Callback events are typically used to simulate discontinuities that arise in a physical system, one example being an ODE system modeling the position and velocity of a bouncing ball. When the ball makes contact with the ground, the velocity abruptly changes sign allowing the ball to bounce away from the ground. This can be modeled with a callback event that checks if the position of the ball matches the ground. Generally, callback events change either the value of the ODE state or parameter. We can adapt these callbacks to implement ABM rules into the ODE system. For example, one callback event that was assigned to the model in Aim 3 was used to transition a cell from an infected to a dead state. After a fixed amount of time after infection, all parameters associated with transcription and translation are set to zero. This *ad hoc* application of callback events is surprisingly flexible, and can be used for more complex ABM rules like adding or moving agents.

A hybridized ODE/ABM model substantially improves the capability of modeling a cell population responding to a pathogenic threat, but still fails to capture intercellular communication. Cellular communication is liaised though the diffusion of cytokines, specialized proteins that bind to neighboring cell surface receptors to initiate signaling pathways. Diffusion can be modeled using partial differential equations (PDEs) through Fick's Law:

$$\frac{\partial f}{\partial t} = \mathcal{D}\nabla^2 f, \tag{1.16}$$

where f is species concentration (a function of space and time), $\mathcal{D}$ is the diffusion coefficient, and ∇^2 is the Laplacian operator. To incorporate this term into the ODE/ABM model, the PDE can be discretized into system of ordinary differential equations and augmented onto the cell signaling dynamics. The time derivative is left unchanged, but the spatial derivatives are approximated by a finite differences. This technique, often referred to as "Method of Lines" provides the major advantage of only needing one ODE solver, as opposed to separate solvers for simulating diffusion and intracellular signaling. The main disadvantage is the resulting size of the system to solve. If a 2-dimensional plane is discretized into 100 segments on each dimension, the accompanying ODE system will have a total 10,000 equations in addition to systems of ODEs controlling cell signaling. Careful consideration needs to be taken to successfully solve this large system without encountering computer time or memory issues.

The Laplacian is discretized using a central order difference method:

$$\nabla^2 f(x_i, t) = \sum_{i=1}^{n} \frac{\partial^2 f(x_i, t)}{\partial x_i^2} \approx \sum_{i=1}^{n} \frac{f(x_i + \Delta x_i, t) - 2f(x_i, t) + f(x_i - \Delta x_i, t)}{\Delta x_i^2}, \tag{1.17}$$

where Δx is the finite difference approximation and n is the number of spatial dimensions. A central difference method is used as opposed to a forward or backward method because the accuracy of the approximation is proportional to the square of the grid spacing (Δx_i^2).

Combining these modeling schemes (see Figure 1.2.7) coalesce into a hybrid model capable of simulating cellular immune responses across multiple spatial scales. The ODEs are used to model changes in intracellular protein concentration over time, PDEs track protein diffusion across a cell population, and ABMs allow cells to probabilistically react to changes to their environment and infectious pathogens encoded by rules.
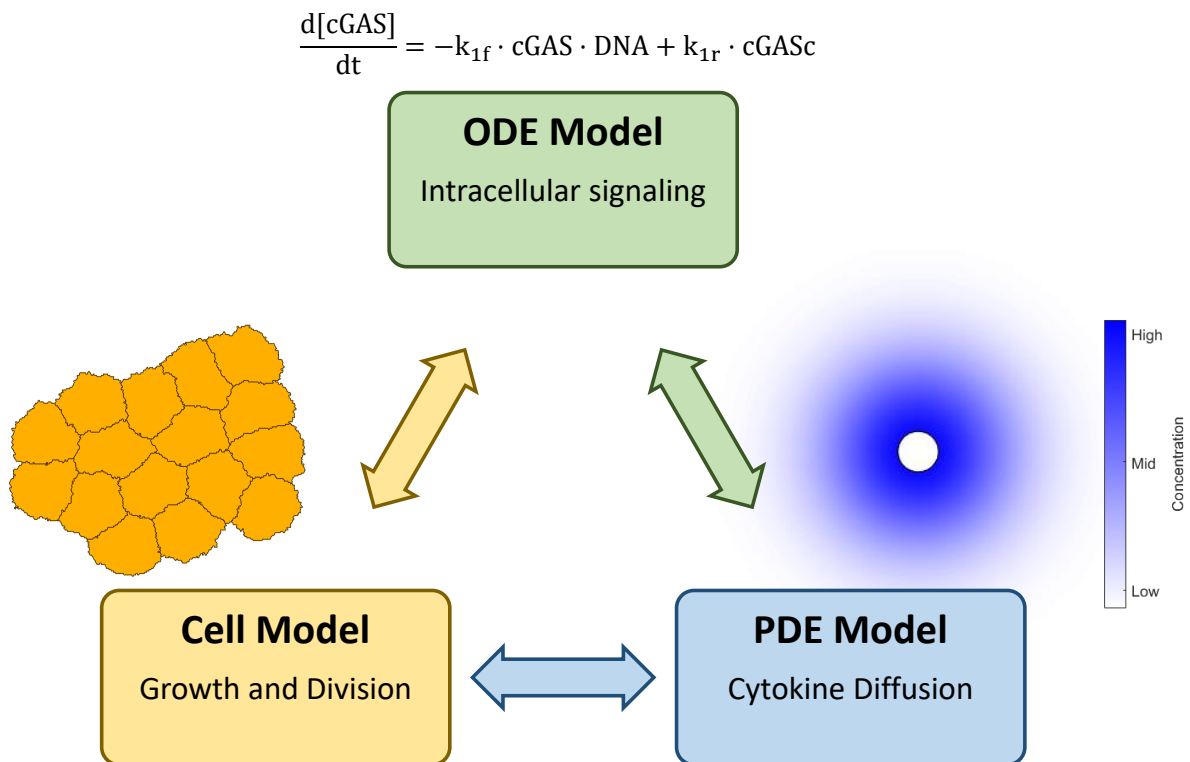
Figure 1.2.7: Hybridizing ODE PDE and ABM Models. Each component of the hybrid model provides a different methodology to simulate a cell population. ODE models are used to track intracellular concentrations over time. PDE modeling allows protein signals to diffuse across the cell population. Cell/ABM Modeling incorporate rules (like the spread of a virus) that influence the current state of the ODE/PDE model.

2.0   Predicting Metabolic Reprogramming in Septic Induced Acute Kidney Injury

2.1   Introduction

Sepsis is a serious medical condition where an aberrant activation of the immune response causes severe, widespread inflammation throughout the body. This activation is initiated by pathogen associated molecular patterns (PAMPs) that enter the circulatory system and disseminate from the origin of infection to induce proinflammatory cytokine production in distal organs despite any sign of local infection. If left untreated, sepsis can quickly devolve into organ dysfunction and, in 30% of the 1.7 million cases in the United States each year, death [10]. The kidney is one of the earliest and most frequently afflicted organs during sepsis, as acute kidney injury (AKI) presents in up to 60% of patients in septic shock [55]. Clinical trials for potential anti-inflammatory therapeutics demonstrate poor efficacy because we lack a mechanistic understanding of the dynamics governing sepsis [56]. To rectify this critical knowledge gap, we turn to computational modeling to identify how sepsis-induced changes in cellular function within the kidney promote AKI.

Sepsis can be experimentally emulated by stimulating renal tubular epithelial cells (TEC) with lipopolysaccharide (LPS) [57], a typical PAMP found within the cell membrane of gram-negative bacteria [58]. TECs detect LPS by binding toll-like receptor 4 (TLR4) protein [59], resulting in a signaling cascade that triggers interferon production [60], releases proinflammatory cytokines, and regulates the infiltration of antigen presenting cells (APCs) and macrophages [61]. Inflammation provides a method to directly resist an infection with the goal of reducing pathogenic loads, despite substantial collateral tissue injury sustained or metabolic resources required.

While several computational studies of LPS-induced signaling have identified mechanisms key to regulating intracellular immunity, most studies do not consider a critically important feature associated with inflammation: metabolic reprogramming [62, 63]. In a process known as tolerance [11], cells change the metabolic pathways used to produce energy, and, in doing

so, prioritize healing associated functions over functions directly targeted at resolving the pathogen. For example, a cell may reprogram itself to use a less efficient metabolic route (aerobic glycolysis) to minimize the production of oxidative byproducts causing cellular damage and tissue injury [64]. The cell can tolerate the infection to deal with the more immediate threat of cell death due to the ongoing immune response. It is also hypothesized that this metabolic reprogramming provides the cell with an opportunity to mount more effective inflammatory responses by allocating energy toward the production of more essential substrates, increasing overall cell survival [65, 66]. Switching from high energetic to lower energetic metabolism to produce metabolites for cell proliferation has been famously observed in tumors, referred to as the Warburg effect [14, 67]. The processes of resisting and tolerating infection work in tandem to maximize pathogenic clearance and minimize injury resulting from the host immune response.

Because metabolism plays a crucial role in regulating immune responses in sepsis [11, 68], it needs to be incorporated into computational models to accurately depict disease progression and outcome. The importance of metabolism suggests that both resistive and tolerant responses to inflammation play a significant role in governing the dynamics driving sepsis. Previous models of the immune response to sepsis focus only the direct inflammatory resistance and do not incorporate the impact of tolerance through metabolic reprogramming [69, 70]. It has been established that exposure to sepsis inducing bacteria signatures like LPS circulating in the blood reach TECs and downregulate ion transporters including ATP pumps which changes how cells produce energy [71]. This metabolism reprogramming induced by sepsis causes TECs to switch between normal efficient oxidative phosphorylation (OXPHOS) to the abnormal inefficient aerobic glycolysis via the Akt/mTOR pathway [11, 72].

Modeling metabolism switching in the context of sepsis has not been investigated, but the signaling pathways governing the metabolism have been modeled extensively (see [73, 74, 75] for the PI3K/Akt/mTOR pathway, [76, 77] for glycolysis, and [78] OXPHOS). By compiling, modifying, and contextualizing these models to LPS induced sepsis, we aim develop an ordinary differential equation model to determine what metabolites or cytokines are responsible for controlling metabolism switching, and show how the dynamics of this biphasic response lead to better cell survival.

2.2 Material and Methods

2.2.1 Establishing Signaling Pathways for Metabolic Reprogramming in Sepsis

Metabolic reprogramming, which has been demonstrated TECs during sepsis, occurs when cellular metabolism shifts from normal oxidative phosphorylation (OXPHOS) to aerobic glycolysis—despite ample oxygen availability. This process is referred to as the Warburg Effect and is controlled by several intersecting signaling pathways that include key regulators such as AMPK, mTOR, Akt, and HIF-1. The first step in building a computational model was to construct a comprehensive map of these regulators to determine the network structure connection these signaling pathways (see Figure 2.2.1).

Both OXPHOS and glycolysis function to convert AMP to ATP, allowing the cell store energy and utilize it maintain cellular mechanisms (e.g. gene expression, molecular transport, etc.). One key regulatory protein that for ATP production is AMPK, which functions as a physiological cellular energy sensor due to its sensitivity to fluctuating AMP levels [79]. When intracellular ATP is reduced to AMP, it binds to the γ subunit of the heterodimer, promoting activation. The phosphorylation of AMPK on Thr172 in the activation loop can also respond to the calcium flux via the CAMKII kinase, so it was included as an additional regulator of AMPK activity [80, 81]. AMPK is an enzyme that plays a vital role in cellular energy homeostasis and thus, in the metabolic shift, acts as a critical control point for both the oxidative phosphorylation and glycolysis pathways.

With evidence that AMPK phosphorylates TSC2 under energy starvation conditions [81], the connection to the mTOR pathway is established through the inverse relationship between the TSC2 and Rheb protein; it was found that TSC2 inhibits Rheb-induced mTOR signaling by reverting Rheb to its inactive GDP-bound state [82]. Rheb, a small GTPase, has been shown to bind to mTORC1 complex and cause a global conformational change that activates the complex allosterically [83, 84]. Experimental observations show that mTORC1 directly increases the stability of HIF-1α [82], a protein that is constitutively expressed, but quickly targeted for ubiquitination [85]. Upon stabilization of HIF-1α, it binds with the HIF-1β to generate the heterodimer HIF-1 complex.
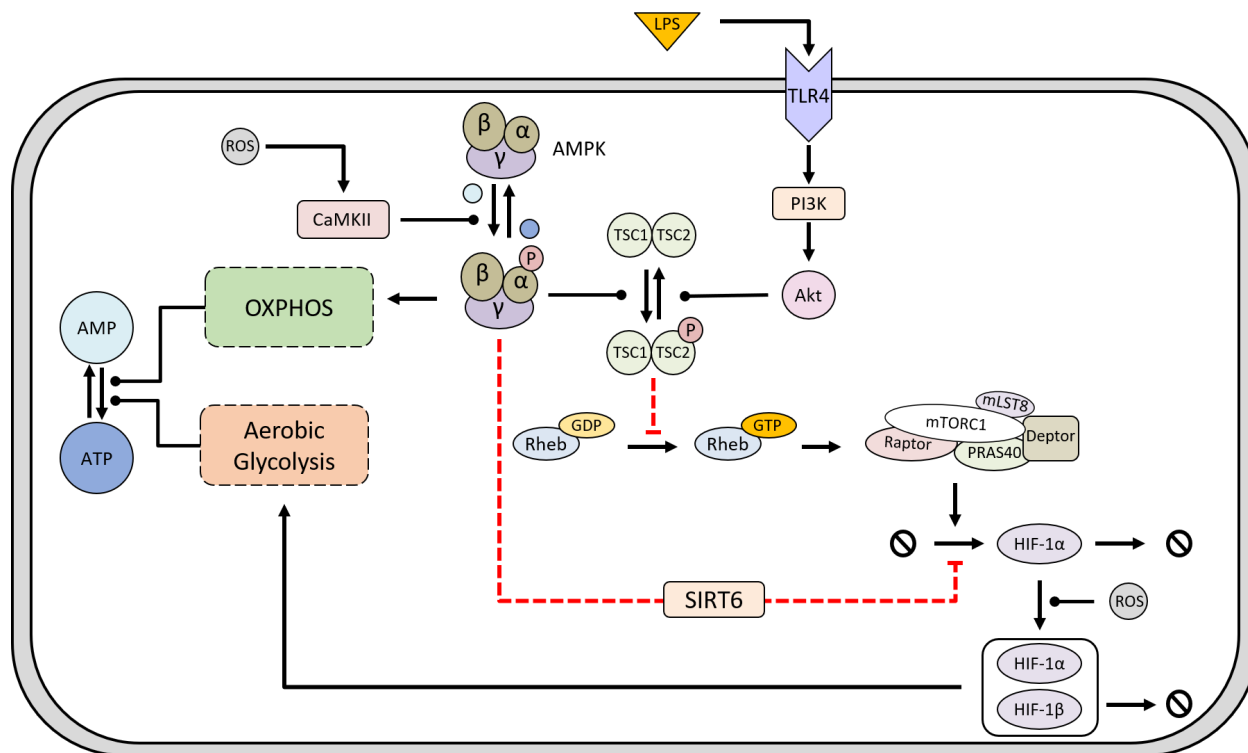
Figure 2.2.1: Sepsis Pathway Diagram. The pathways controlling two major metabolic routes (OXPHOS and glycolysis) are illustrated above. Both catabolic pathways work to enhance ATP production, which is controlled by AMPK and HIF-1 activity. ATP and AMP directly bind to AMPK (shown as light and dark blue circles near AMPK), which acts as a controller to maintain their ratio. AMPK phosphorylates TSC2 which inhibits Rheb activation and binding to mTORC1. This inhibition is reversed by Akt which is activated by LPS stimulation. With LPS stimulus, mTORC1 is activated to stabilizes HIF-1α and enhance aerobic glycolysis. SIRT6 opposes this stabilization when AMPK phosphorylation is pronounced. ROS, a byproduct of OXPHOS, increases HIF-1 complex formation and enhances CaMKII activity to drive further AMPK phosphorylation.

HIF-1 increases the metabolism of glucose using two mechanisms. First, glucose transporters (GLUT1) are upregulated to increase glucose flux into the cell [86]. And second, HIF-1 upregulates glycolytic enzymes like lactate dehydrogenase A (LDH-A) to facilitate the metabolism of glucose into downstream products like pyruvate and lactate. [87].

2.2.2 Integrating Immune Activation into the Metabolic Shift

In this model, LPS is used as a proxy to simulate cellular response to sepsis. LPS, an endotoxin from gram-negative bacteria, is recognized by Toll-like receptor 4 (TLR4) to initiate the innate immune signaling cascade [88]. The activated cell surface receptor can then stimulate the PI3K/Akt pathway, causing tyrosine phosphorylation of surface receptors of PI3K and subsequent molecule phosphorylation of the Akt kinase catalytic region [89]. Akt can completely inhibit TSC2, and as a result, affect the metabolic state of the cell [90, 91]. Furthermore, SIRT6, a NAD$^+$-dependent deacetylase, is activated by AMPK and represses HIF1a transcriptional activity [92]. In addition to affecting metabolism through the Akt/PI3K pathway, LPS has also been shown to inhibit AMPK activation and potentially play a significant role in the downstream metabolic shift from OXPHOS to glycolysis [93].

2.2.3 Developing the ODE System

The combination of this experimental evidence led to the development of the model structure depicted in Figure 2.2.1 with the corresponding ODE model summarized in Appendix B.1. The model consists of 23 states representing protein species in the pathway and 31 parameters that dictate the dynamics driving the evolution of these states. All states are modelled using mass action kinetics except for the conversion between AMP and ATP. The OXPHOS and glycolytic pathways were each lumped into one term which would make this kinetic model inappropriate. Instead, the ATP production rate was modelled to be proportional to AMPKp and HIF-1 activity to represent OXPHOS and glycolysis, respectively. The basal consumption rate of ATP was set to be constant. These rates were informed by experimental studies conducted by Mookerjee *et. al.* where they quantified intracellular glycolytic and oxidative ATP flux [94]. The steady state concentrations for ATP and AMP

are larger in magnitude than other signaling proteins and were taken to be 9600 μM and 280 μM respectively [95].

Proteins that could be present in two different states (e.g. AMPK and AMPKp) were constrained via mass balance, which assume no additional translation or degradation of these proteins. Some species like HIF-1α did have production and degradation rates. Many of these rates were unknown, so they were set to average protein translation and degradation rates based on measurements from hundreds of proteins [96]. The degradation rate of HIF-1α was known and set to value of 5 min^{-1} [85].

To simulate the model, the *DifferentialEquations.jl* package in Julia was used in conjunction with the *DiffEqBiological.jl* package [54]. This method uses a Domain Specific Language (DSL) for writing chemical reactions which we use to define the pathways controlling OX-PHOS and glycolysis metabolism. A system of differential equations is automatically generated with predefined structures like an analytical Jacobian to accelerate solving efficiency. The default ODE solver "AutoTsit5(Rosenbrock23())" was chosen. This solver automatically switches between a Runge-Kutta and Rosenbrock method depending on the stiffness of the solution at the current time point.

2.2.4   Simulating LPS infection and Protein Knockdown

A callback event was used to dose the system with an impulse of LPS. When the simulation reaches a specified time point, the ODE solver is stopped, and the concentration of LPS is changed to a nonzero value. This discontinuous jump is then used as a new initial condition to continue the numerical simulation.

Knockdown simulations were performed in a similar fashion, but instead of changing a state concentration, a parameter value was gradually reduced until it reached a value a zero. At this point, the manipulated rate would be removed the simulation which corresponds to a complete knock out of that reaction.

2.2.5 Global Sensitivity Analysis

The *DiffEqSensitivity.jl* package was used to perform a global sensitivity analysis of the system to determine which states were most impacted by changes in parameters. There were multiple GSA methods available, but the eFAST (extended Fourier amplitude sensitivity testing) method was selected for its robustness and speed. Briefly, eFAST is a variance-based sensitivity analysis that assumes you have some model f with a set of inputs $\mathbf{x} = \{x_1, x_2, x_3, ...x_n\}$ that generates and output Y [97]. The inputs in this case are the different parameters and the output is a scalar value representing the simulation's response to the inputs. Here, the sum of the areas under each curve was used (other metrics like the maximum concentration of a species are also valid, so long as the result is scalar). Without loss of generality, we can treat each input as an independent, uniform random variable on [0,1], and decompose the output Y as a sum of expected input values:

$$Y = f_0 + \sum_{i=1}^{n} f_i(x_i) + \sum_{i<j}^{n} f_{ij}(x_i, x_j) + ... + f_{1,2,...,n}(x_1, x_2, ...x_n), \qquad (2.1)$$

where f_0 is the variance of the output (constant), $f_i(x_i)$ is the variance of one input conditioned on the output, $f_{ij}(x_i, x_j)$ is the variance of two inputs conditioned on the output, etc. By repeatedly sampling these expected values, we can estimate what proportion of the input variances contributes to the total output variance. Differences between variance-based sensitivity methods depend on how variances are estimated. The Sobol method uses Monte Carlo integration to estimate expected values, whereas eFAST uses Fourier series expansion.

2.3 Results

2.3.1 The Metabolic ODE Model Recapitulates Experimental Data

To compare this model against experimental data, a steady state solution needed to be determined. Cells typically operate in a homeostatic range where the generation and consumption of energy and molecules are in equilibrium. There are several ways to determine a steady state solution, but the most straightforward method is to run the ODE simulation until the derivative of each state is below some chosen threshold. Figure 2.3.1A depicts the resulting steady states for each state in the system except for ATP and AMP. These species were removed because their steady states were already known from literature. No LPS is present in this system so the steady state values for LPS and active species in the PI3K. Akt pathway are zero. ROS and SIRT6 possess substantially larger steady states when compared to other signaling protein because of their production are large when compared to their degradation rate. To ensure these values were a steady state solution, the system what solved with this initial condition, and as demonstrated in Figure 2.3.1B every state remained constant across the simulated time frame.

A
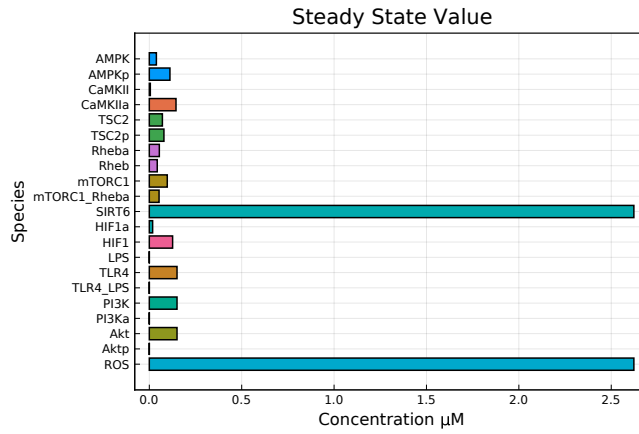
B
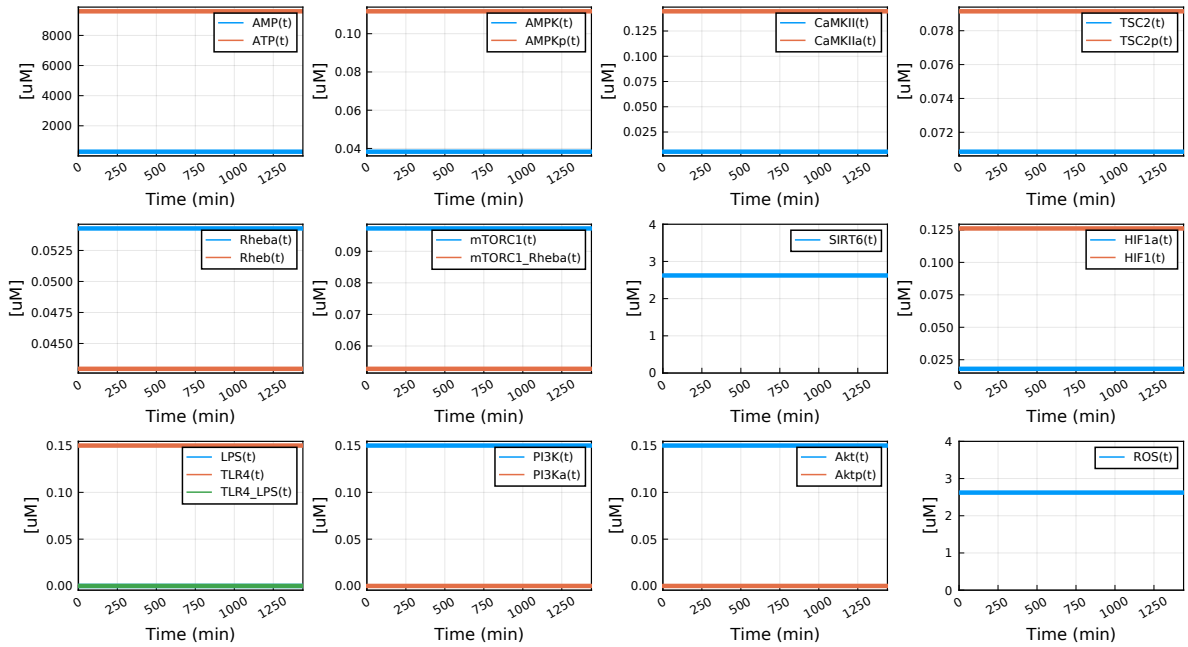
Figure 2.3.1: Determining Metabolic Equilibrium. (A) The ODE is simulated from an arbitrary initial condition until the derivative of each state is below the absolute tolerance of the ODE solver (1×10^{-8}). (B) The determined steady state is feed back into the ODE and simulated out for 24 hours. All states remain constant with respect to the error tolerance.

Figure 2.3.2 shows how the simulation responds to changes in the AMP/ATP ratio. As previous experimental data suggest [79, 98], increasing the proportion of AMP in the cell should trigger the activation of AMPK through phosphorylation. Indeed, at $t = 500$ min, AMPK phosphorylation does increase which triggers a signaling cascade to promote OX-PHOS and restrict glycolysis. HIF-1 concentration decreases because SIRT6 is upregulated and Rheb is deactivated. Interestingly, when the AMP/ATP ratio is decreased, the response from AMPK is strengthened.
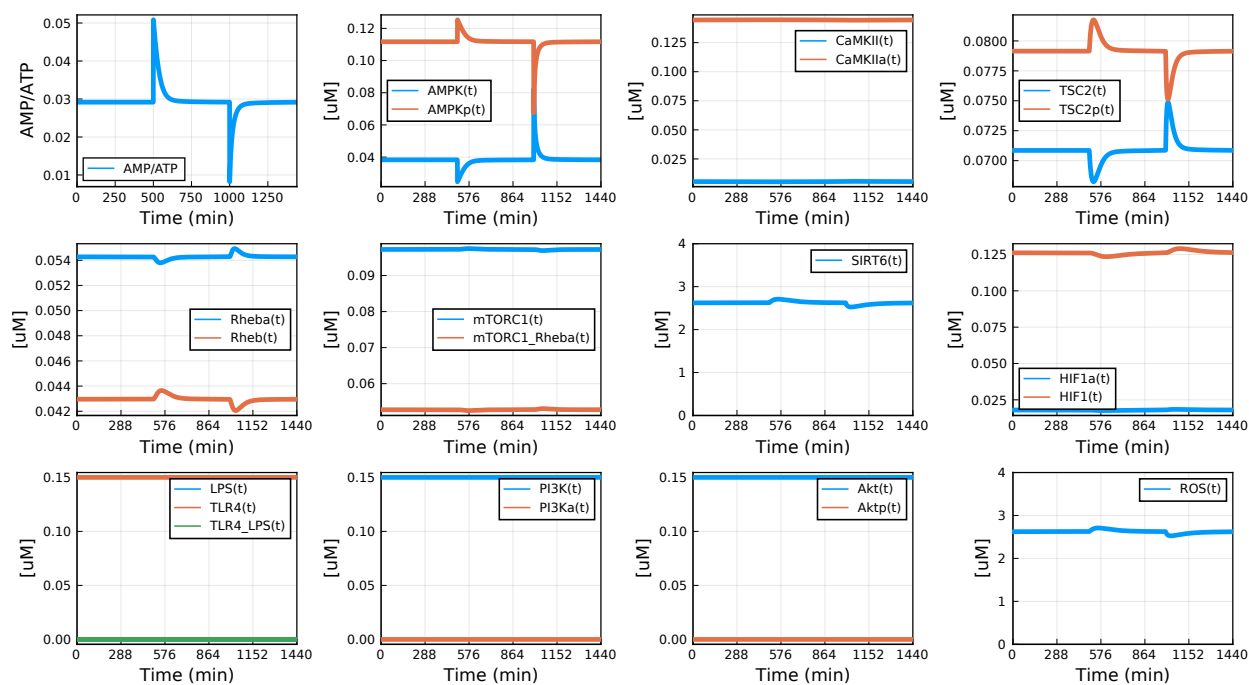
Figure 2.3.2: Modulating AMP/ATP Ratio. Here the AMP/ATP was manipulated to determine if the system would respond similarly to experiment data. At $t = 500$ min, the amount of AMP was increased by 200 μM and ATP was correspondingly decreased by the same amount. The reverse was scenario was performed at $t = 1000$ min by decreasing the amount of AMP in the cell.

2.3.2 Biphasic Metabolic Shifting is Dependent on LPS clearance

To determine how the system behaves when introduce to LPS, the steady state system in Figure 2.3.1 was modified with a dose callback event. At $t = 250$ min, this callback event is trigger and modifies the current LPS concentration from zero to 0.05 μM. This would

be equivalent to a cell receiving an LPS injection *in vitro*, where the concentration of the stimulant suddenly increases.
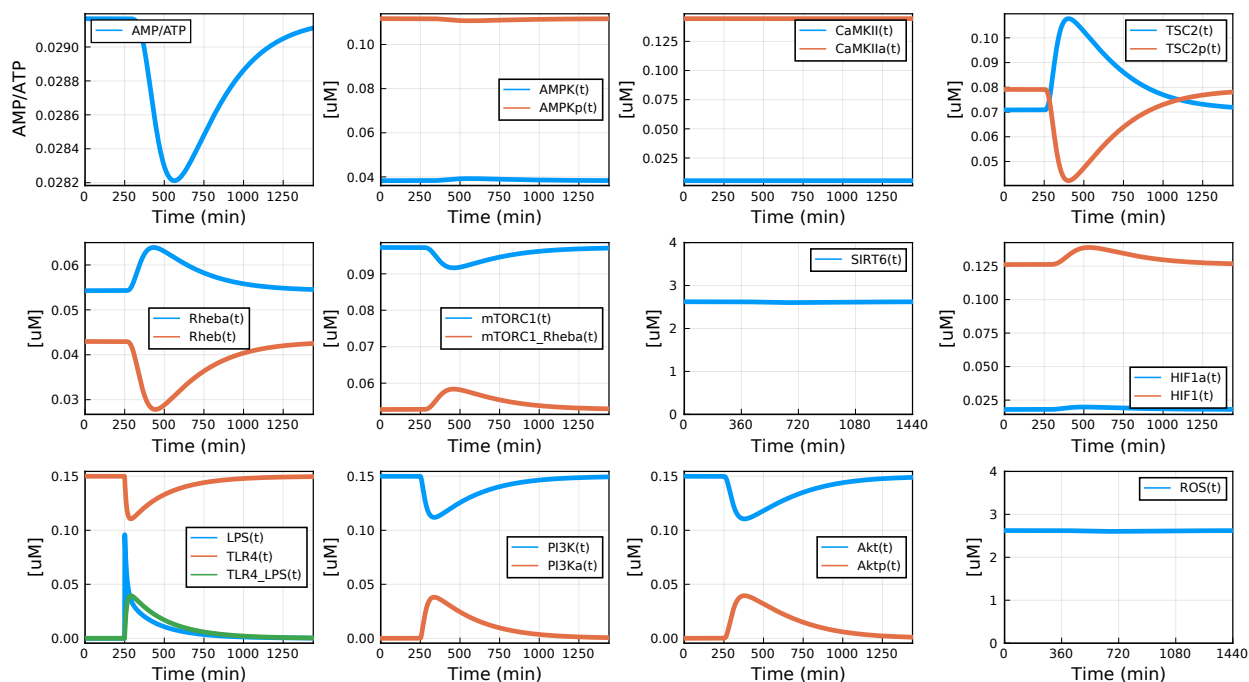
Figure 2.3.3: LPS Impulse Response. The stabilized ODE model was dosed with LPS at $t = 250$ min. LPS quickly binds to TLR4 and degrades, allowing the system to return to steady state.

Figure 2.3.3 illustrates the dynamic response after LPS dosing. Immediately following LPS introduction, the amount of free LPS decreases as it binds to TLR4, forming a complex. This subsequently activates the PI3K/Akt pathway to inhibit TSC2 phosphorylation. This allows mTORC1 activity to rise and finally increase HIF-1. Of note, because HIF-1 is used as an indicator for glycolysis, it slightly increases the amount of ATP in the system (thus decreasing the AMP/ATP ratio). This change is small because the ATP production rate of glycolysis is about 10 times smaller than OXPHOS. This slight increase in ATP has the minuscule effect of lowering AMPKp activity. Nevertheless, the simulation is emulating a cell utilizing OXPHOS, transitioning to a glycolytic state, and reverting back to OXPHOS.
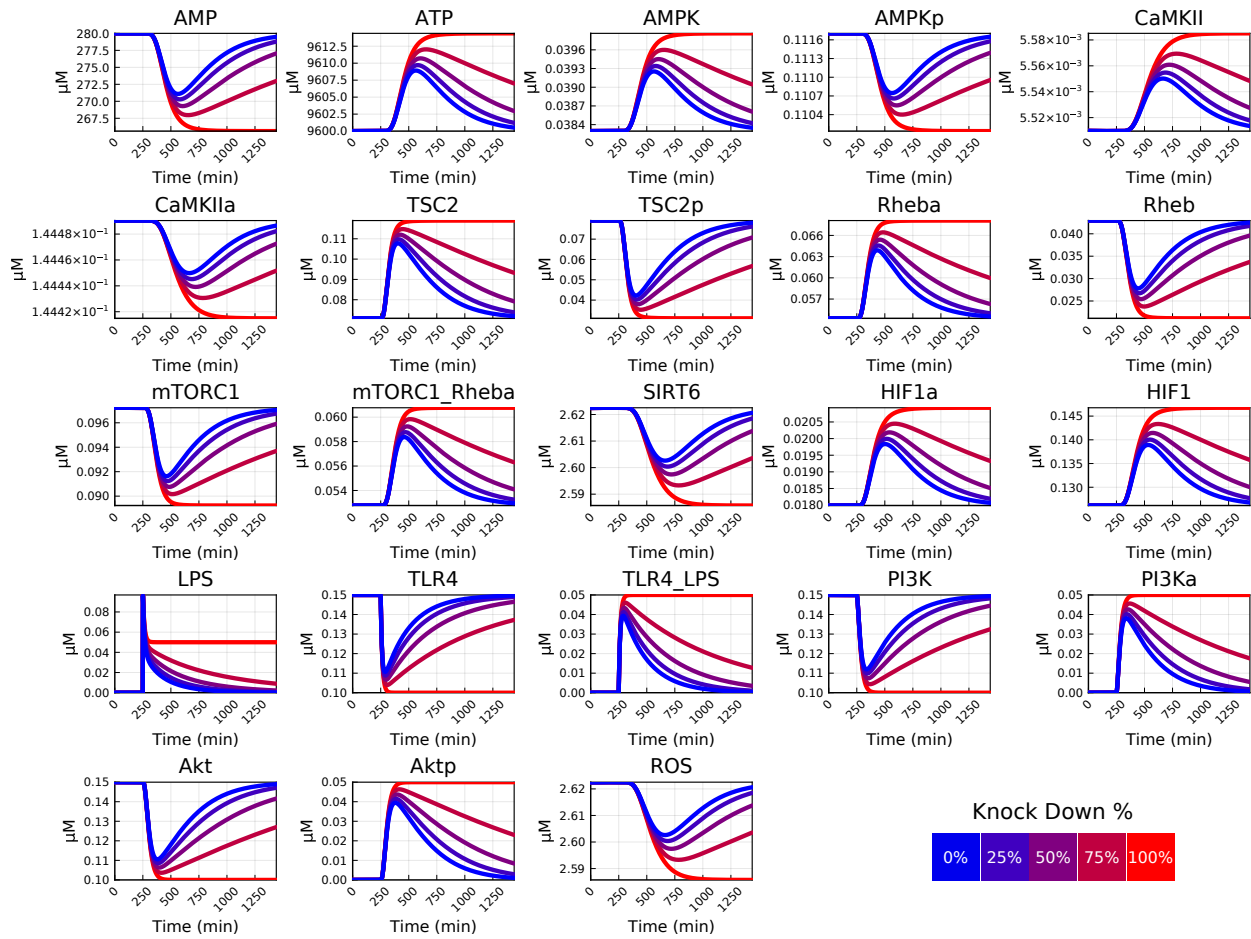
Figure 2.3.4: Effect of LPS Clearance on Metabolic Reprogramming. The ODE system was dosed with LPS at $t = 250$ min, but now with the added effect of varying LPS degradation (τ_{16}). The value of τ_{16} is simulated at its original value (0% knockdown, blue line) all the way to a complete knockout (100% knockdown, red line) where $\tau_{16} = 0$.

.

LPS was given an artificial degradation rate (τ_{16}) to provide the cell a means to clear the stimulus. This would be synonymous to washing cells in culture after stimulation with LPS. If this degradation rate parameter is reduced, the system can be simulated with progressively more prolonged doses of LPS. Figure 2.3.4 shows the same simulation from Figure 2.3.3, but with the addition of progressively lowering the degradation rate by the given percentage. At 50%, for example, the value of τ_{16} is halved and at 100% is set to a value of zero. Without the ability to clear the LPS signal (100% knock-down, red line), HIF-1 and $AMPK_p$ transition to a new equilibrium where HIF-1 levels have increased and $AMPK_p$ levels have decreased. This signifies a transition in metabolism to aerobic glycolysis and away from OXPHOS.

In addition to this permanent trade-off in metabolism, these simulations show a decrease in the AMP/ATP ratio (or equivalently and increase in ATP levels). This accumulation in energy supply could be and indicator for cells preparing to enter a resistive immune state, which would necessitate higher ATP levels.

2.3.3 Global Sensitivity Analysis Reveals AMPK insensitivity

Sensitivity analyses are an excellent tool for characterizing what parameters are contributing most significantly to changes in the ODE states. Figure 2.3.4 illustrates the results from performing an eFAST global sensitivity analysis. This simulation included a callback to trigger an LPS dosing at $t = 250$ min. Without this stimulation, the sensitivity values for the PI3K/Akt pathway would all remain at zero.

One interesting observation from this analysis is how insensitive glycolytic ATP production (k_{1bf}) is compared to OXPHOS (k_{1af}) or the basal consumption rate (k_{1r}). It appears that only SIRT6 exhibits some sensitivity to glycolysis, which is not surprising because it directly inhibits this process. Oddly, activated mTORC1 is insensitive to glycolysis despite directly promoting its activation.

The most sensitive species/parameter combination was LPS and its associated degradation constant. In fact, LPS shows no sensitivity toward any other parameter, which is appropriate as LPS acts as an input to the system.
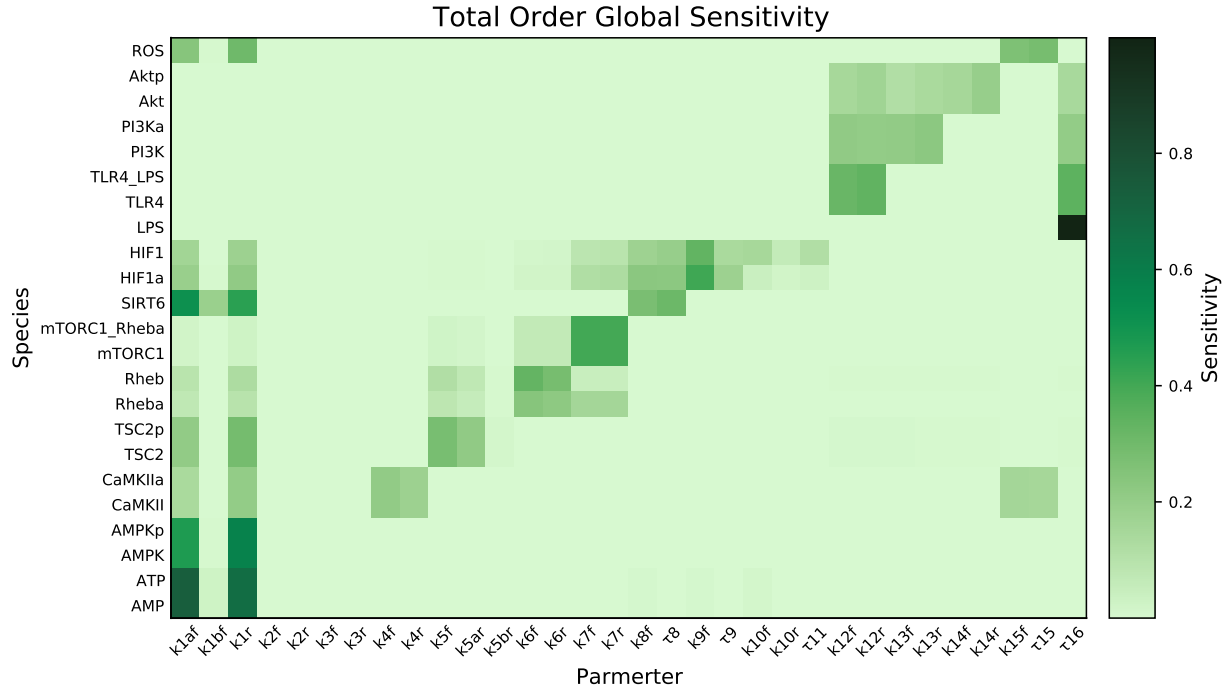
Figure 2.3.5: eFAST Global Sensitivity Analysis. Normalized sensitivities are calculated for each parameter/species combination. A darker color indicates that small perturbations in the parameter lead to large changes in the associated protein/molecule. The parameters were allowed to range between half and double their original values.

The least sensitive parameters are k_{2f}, k_{2r}, k_{3f}, and k_{3r}. These parameters control the AMPK/AMPKp activation level directly or indirectly through CaMKII. As observed in Figure 2.3.3, the magnitude $AMPK_p$ is not significantly impacted by LPS dosing because HIF-1 has a smaller effect on the AMP/ATP ratio. CaMKII activity changes with ROS concentration, a byproduct of OXPHOS and directly affected by $AMPK_p$. CaMKII activity is sensitive to ROS through the k_{15f} and τ_{15}, but this sensitivity is not an adequate to impact AMPK downstream.

2.4 Summary

Sepsis is a complex, multi-organ condition that causes detrimental, and even fatal outcomes. Even with proper treatment and early detection, sepsis still remains a leading cause of death in the United States and is responsible for being the most common cause of death in the ICU [99]. We have spent decades trying to discover potential treatments to combat these morbid outcomes by most fail clinical trial [8]. To investigate therapies that go beyond anti-inflammation strategies, we need to develop a comprehensive, systems level perspective of how sepsis impacts the body. Recent studies have shown that, in addition to elevating the immune response, sepsis is responsible for cells shifting their metabolic state from oxidative phosphorylation to aerobic glycolysis. The goal of this aim was to develop a computational algorithm that demonstrates how this metabolic reprogramming occurs and determine sufficient conditions that cause metabolism to shift back. The model was able to show how LPS stimulation resulted in the deactivation of TSC2 which leads to an increase in glycolysis. It also predicts that biphasic shift is possible if LPS can be cleared from the cell (Figure 2.3.4).

Jin *et. al.* published a study this year investigating the role of AMPK activity in the renal metabolic response to sepsis. In their experiments, cells were given "inflammatory mix" (IM)— a combination LPS and other damage associated molecular patterns to emulate sepsis— and measured for AMPK activity over time. They observed an increase in AMPK activity within the first five minutes of treatment followed by a gradual decline in the next hour. Then, at later time points (2, 6, 12, 18, and 24 hours) under the same condition, AMPK activity began to recover. Our model predicts this gradual decrease and eventual recovery if LPS is being cleared from the cell. It fails to predict the immediate increase in AMPK activity which may indicate some unknown mechanism between LPS and AMPK.

As the model is defined, triggering a septic episode appears to increase the amount of ATP produced by glycolysis, but because its contribution was already low, OXPHOS is not significantly inhibited. What emerges from this simulation is a scenario where both OXPHOS and are glycolysis are utilized, decreasing the AMP/ATP ratio (Figure 2.3.3). This heightened energy state is sustained if LPS remains in the cell (Figure 2.3.4) and is due to a lack of sensitivity in subtle AMP/ATP fluctuations from AMPK (Figure 2.3.4). This

possible co-activation, as opposed to full shift, would need to be experimentally confirmed. One possible experiment design would involve performing OCR and ECAR measurements across mutliple time points to determine the contribution of ATP production from each metabolic pathway.

This computational model comes with some limitations that deserve mention. The first is the lack of dynamic data to inform the model. This is a common problem in systems biological as dynamic information is difficult to measure experimentally (especially without sacrificing the cells to make the measurement). While the predicted trends from these simulation match experiments in a qualitative manner, further work needs to be done to integrate these data into the model. Another limitation related to this is obtaining absolute concentration measurements to inform the initial conditions. This is less concerning though because the ODE model can be easily re-scaled to match these measurements (if there exists an accessible method to determine these values).

Overall, this model demonstrates a simple and user-friendly methodology to generate an ODE model from a given network of interconnected signaling pathways. Advanced model perturbations and analyses were utilized to stabilize the model at a steady state and gain insight into how the system behaved. Using these tools, we were able to confirm the plausibility of a metabolic shift and demonstrate sufficient conditions to cause the (small) shift to reverse. By incrementally building on our knowledge of sepsis and metabolism, we hope to identify and simulate new potential treatments to improve patient outcome.

3.0 Mathematical Modeling of the cGAS Pathway Reveals Robustness of DNA Sensing to TREX1 Feedback

3.1 Introduction

The Cyclic GMP-AMP synthase (cGAS) pathway represents a critical component of the innate immune system that is responsible for detecting dangerous changes in the intracellular environment and maintaining organismal homeostasis. In healthy nondividing cells, DNA is located in the mitochondria and nucleus, so accumulation of DNA in the cytosol typically indicates a breakdown in organelle stability (e.g. tumorigenesis [28, 100] or the presence of a pathogen [101, 102, 103, 104, 105, 106, 107, 108]). While the cGAS pathway must rapidly detect cytosolic DNA and effectively coordinate an immune response, aberrant activation can lead to autoimmune and inflammatory disease [109, 110, 111]. The use of computational tools to model the cGAS pathway can reveal insights about immune regulation and identify unexpected systems-level emergent behaviors that result from changes in the intracellular environment, such as chronic inflammation. Cytosolic DNA is primarily detected by the DNA sensing molecule, cGAS [24]. cGAS and DNA bind to form an activated enzyme complex which triggers a series of molecular interactions culminating in the production of type I interferons (IFNα/β). Studies have confirmed that cGAS is the primary cytosolic DNA sensing molecule for a large variety of pathogens, including DNA-containing viruses (herpes simplex virus [108], adenovirus [105]), retroviruses (HIV [102]), and bacteria (L. monocytogenes [104], M. tuberculosis [101, 107]). There is also evidence of viruses actively inhibiting cGAS pathway activity to promote virus replication. In the case of poxviruses (e.g. smallpox), cGAS is able to bind to viral DNA, but the downstream signaling protein stimulator of interferon genes (STING) is inhibited by the virus, preventing its dimerization and phosphorylation [103]. This example demonstrates the complex relationship between the cGAS pathway and the pathogens it defends against.

cGAS is known to bind to self-DNA that has leaked from the nucleus of damaged cells [112, 113]. This DNA leakage is an early indicator of tumorigenesis in cancer and can be

detected by the cGAS pathway. Interestingly, prolonged activation of the cGAS pathway has been shown to cause inflammation-mediated tumor growth [28], bringing into question the viability of immunomodulatory treatments that promote cGAS activity to regulate tumorigenesis. Dysregulation of the cGAS pathway can lead to impaired immunity or chronic inflammation. A well-established example of this impairment is Aicardi–Goutières syndrome (AGS) [110], an autoimmune disorder where the DNA degradative enzyme three prime repair exonuclease 1 (TREX1) is nonfunctional in its role as a suppressor of the cGAS pathway. This disorder is characterized by elevated interferon levels and chronic inflammation. The cGAS pathway has also been associated with more prevalent autoimmune diseases, such as lupus erythematosus [30], and age-related diseases, such as atherosclerosis and neurodegenerative diseases [32]. With extensive connections to innate and adaptive immunity. as well as various diseases and pathogens, it is critical to develop a systems-level understanding of how the cGAS pathway dynamically regulates the immune response.

To investigate the regulation of the cGAS pathway, mathematical models can be constructed using principles from the fields of reaction kinetics and transport phenomena to capture system-wide dynamic responses. Comprehensive molecular interaction maps of immune signaling pathways have been published [114, 115], however, little is known about the expectant system dynamics under different biological conditions. Elucidating these dynamics can be problematic as data most often only provides a snapshot of a specific event. To mitigate this problem, systems biology approaches are implemented to interpolate the dynamics between static observations [38, 116]. Specifically, ordinary differential equations (ODEs) constructed from mass and energy conservations are employed to track the species' concentrations within a cell over time. This method has demonstrated significant success in describing complex dynamics for a variety of signaling pathways. Key examples include explaining the appearance of robustness within a system [117, 118, 119], oscillatory properties of circadian rhythms [120] and the NF-ϰB pathway [40], and switch mechanisms that give rise to all-or-nothing responses to stimuli [121].

In this study, we introduce an ODE model to quantify cGAS pathway activity and assess the model's robustness to parametric uncertainty in the feedback components that primarily regulate the pathway's dynamics. Conservation of mass principles were used to construct

the ODE description of the cGAS pathway kinetics and enzyme reactions were modeled with Michaelis Menten kinetics. To capture paracrine and autocrine signaling effects, IFNβ responses predicted from the cGAS pathway model were used as inputs into an existing model of the janus kinase/ signal transducer and activator of transcription (JAK/STAT) pathway which simulates IFNβ regulation via the IFN receptor [122]. Additionally, Three Prime Repair Exonuclease 1 (TREX1), a known DNase and interferon stimulated gene (ISG), was recently shown to be essential for degrading cytosolic DNA when the cGAS pathway was activated and was included in the model [110, 123]. A literature search was performed to compile experimental data and fit model parameters to biologically relevant ranges [124, 125, 126, 127, 128, 24, 129]. Applying this model, we test how robust the relative timing and intensity of the IFN response is to perturbations in TREX1 and interferon regulatory transcription factor 7 (IRF7) feedback and compare model predictions to data observed in a drug inhibition study. We find the model to be extremely robust to perturbations in TREX1 activity, but susceptible to changes in IRF7 activity.

3.2   Materials and Methods

3.2.1   Model Construction

As shown in Figure 3.1.1, the cGAS pathway regulates type I interferon production through the JAK/STAT pathway. The cGAS pathway consists of 7 states summarizing how DNA is detected in the cytosol to produce the initial interferon response (see equations 1-7 in Appendix A.1). This pathway begins with cGAS reversibly binding to cytosolic double stranded DNA and enzymatically producing 2'-3'-cyclic guanosine monophosphate–adenosine monophosphate (cGAMP) [130, 24]. The enzymatic activity of cGAS is assumed to be saturated as the basal concentrations of ATP and GTP are orders of magnitude higher relative to cGAS [131, 132]. The produced secondary messenger cGAMP diffuses to the endoplasmic reticulum (ER) and reversibly binds to STING [133, 134], which is modeled using mass action kinetics. This binding induces a conformational change in STING
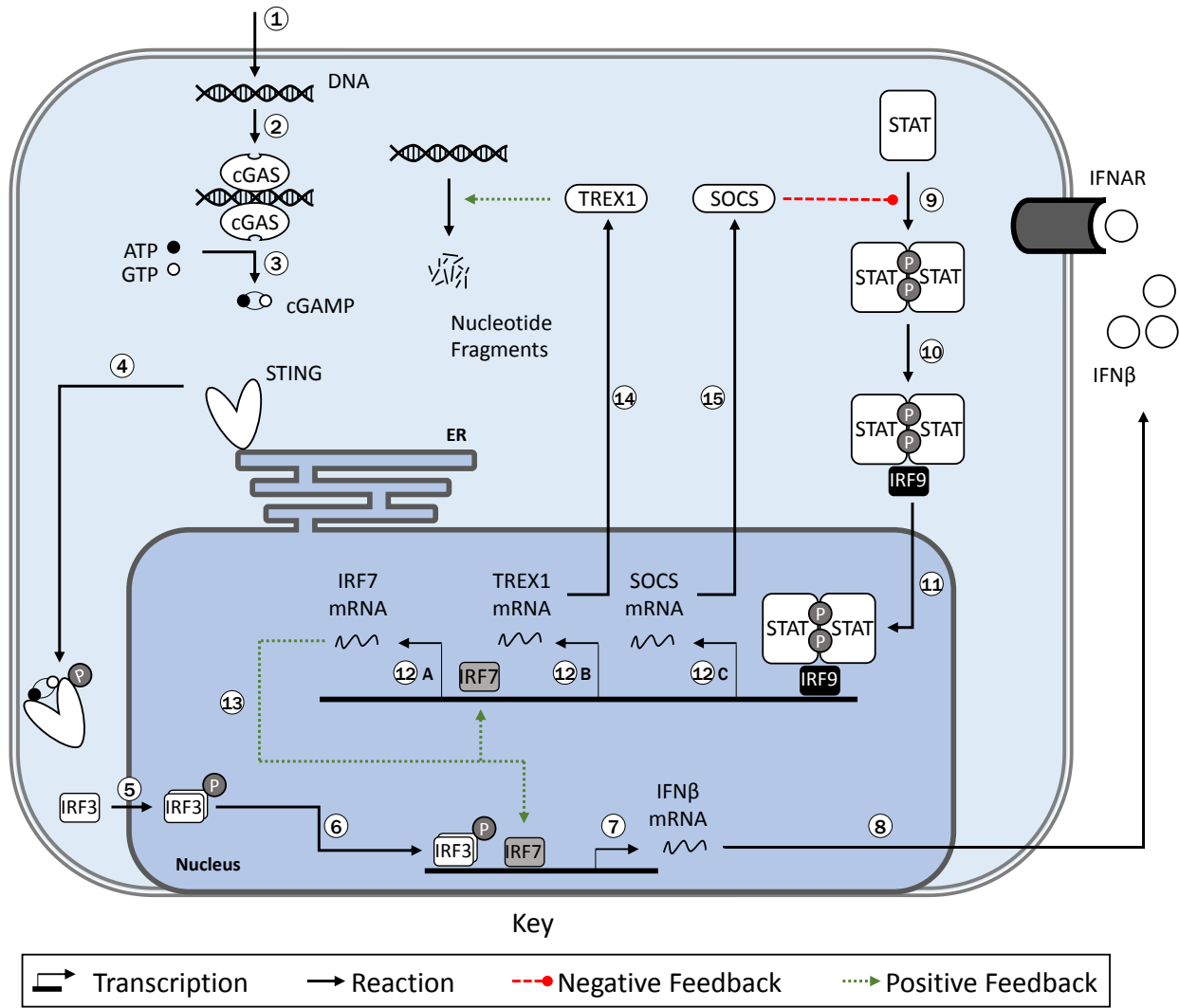
Figure 3.1.1: Diagram illustrating the components of the cGAS and JAK/STAT pathways. Host or pathogenic DNA enters the cytosol and is bound to cGAS, activating the enzymatic conversion of ATP and GTP to cGAMP. The secondary messenger cGAMP binds to STING which undergoes a conformational change releasing it from the ER. STING trafficks to the perinuclear space and acts as a scaffolding protein to allow for the phosphorylation of IRF3. This complex enters the nucleus and acts as a transcription factor for IFNβ. Translation of IFNβ initiates the JAK/STAT pathway through autocrine (shown here) and paracrine signaling. Through the formation of STAT complexes, this pathway activates positive (IRF7) and negative (SOCS, TREX1) feedback mechanisms to regulate the whole system. Systems equations, state descriptions and parameter descriptions are in Appendix A.1

releasing it from the ER and allowing it to migrate to the ERGIC [25]. Here it acts as a scaffolding protein to aid in the phosphorylation of interferon regulatory transcription factor 3 (IRF3). Phosphorylated IRF3 then reversibly dimerizes and enters the nucleus to transcribe IFNβ. The entire process of IRF3 phosphorylation and dimerization was lumped into a Michaels Menten model to minimize the number of parameters necessary to describe the interaction. This method simplifies the model but neglects nonlinear terms associated with IRF3 dimerization. The promoter region of IFNβ is controlled by four positive regulatory domains [36], but it was assumed modeling one (IRF3) would sufficiently recapitulate IFNβ dynamics.

A model describing the JAK/STAT pathway [122] that captured the type I interferon response in influenza infected cells was modified to include the cGAS pathway. The initial JAK/STAT model was unstable. Unstable models cannot capture critical signaling features such as time until system shut down and transitions between steady states. Therefore, the JAK/STAT model was altered and parameterized simultaneously with the complete ODE model system. Specifically, degradation rates were introduced to prevent unbounded production [135, 136, 137] and the piecewise differential equation for phosphorylated Interferon Regulatory Factor 7 (IRF7) was removed and replaced with a mass action rate equation. Additionally, the combined model in this paper makes no distinction between subtypes of type I interferon, so any dynamics involving IFNα were removed.

The modified JAK/STAT model adds an additional 6 states to describe how IFNβ protein promotes the dimerization and phosphorylation of STAT and subsequently induces the transcription of ISGs (see AppendixA.1, EquationsA.8 -A.13). The pathway begins with the activation of interferon-α/β receptor (IFNAR) by IFNβ and leads to the synthesis of STAT complexes which can be inhibited by the presence suppressor of cytokine signaling (SOCS) proteins. Both STAT and SOCS proteins have different subtypes (e.g. SOCS1, STAT2), but no distinction is made between these for the sake of model simplification. Moreover, a quasi-steady state assumption was implemented by Qiao et. al. to directly relate the amount of SOCS mRNA to its protein. The STAT complex enters the nucleus and promotes the transcription of SOCS and IRF7 which in turn activates negative and positive feedback loops, respectively. Other ISGs are transcribed but are assumed to have no effect on the dy-

namics of the pathway. The stability of DNA molecules necessitates a regulatory mechanism for active degradation. A TREX1 double knockout study in mice suggests that TREX1 is the primary DNase that degrades cytosolic DNA in response to cGAS activation [110]. Also, evidence suggests that TREX1 is an interferon stimulated gene and it is therefore regulated by IFNβ in the model [123, 138].

In total, the model is comprised of 13 states whose dynamics are described by 34 parameters separated into three groups: cGAS pathway parameters (13 total), JAK/STAT parameters (11 total), and degradation rate parameters (10 total). It was assumed that none of the cGAS pathway or JAK/STAT parameters were known, and TREX1 negative feedback activation was also unknown. Additionally, one degradation rate was not found in literature, leading to a total of 25 unknown parameters to be optimized.

3.2.2   ODE Simulation

The Advanced MATLAB Interface for CVODES and IDAS (AMICI) toolbox was used to simulate the ODE model. Using symbolic notation in MATLAB, states (species), parameters, and rate equations were written and converted into compiled C code for efficient simulation. AMICI utilizes SUNDIALS to simulate the complied model by implementing a variably stiff ODE solver to numerically estimate the ODE solution.

3.2.3   Markov Chain Monte Carlo (MCMC) Optimization

Many MCMC methods are available to perform parameter optimization but based on a recent review by Ballnus [43] we chose to use the toolbox PESTO to implement a parallel tempering (PT) MCMC algorithm in conjunction with a multi-start local optimization. Briefly, PESTO begins by using an interior-point minimization routine in MATLAB (fmincon) initialized from 1,000 randomly distributed starting points. These starting points are determined using Latin Hypercube Sampling. PT MCMC then uses these local minima and gradient information to initialize the algorithm. Both methods have been shown to greatly improve convergence to a stationary distribution over single chain methods initialized by the prior distribution. Local optimization is also designed to significantly reduce burn-in time.

To further increase the likelihood convergence, PT MCMC ran for 10 million iterations and was repeated in triplicate (labeled as Runs 1-3 in the Results and Discussion).

3.2.4 Data Compilation

To perform the MCMC optimization, data procured from literature was used to inform the unknown model parameters (Supplementary File 1). This data is comprised of 10 unique time course experiments using RT-qPCR and Western blots techniques to measure 7 of the 13 states simulated in the model. Neither of these techniques allow for absolute quantification of concentration, so data were normalized to be bounded on [0,1] and thus provide a measure of the response dynamics but not the magnitude of the responses. Ideally, data would come from a single cell type, but dynamic data for cGAS activation was limited. We collected data from monocytes, macrophages, and fibroblasts. Pulling data from these cell types was justified as (1) monocytes, macrophages, and fibroblasts all are strong interferon producers and (2) demonstrated similar IFNβ and IRF3 dynamics [139].

Table 1: Compilation of Experimental Data

Experiment	Measured State	Label	Reference	Stimulant	Cell Line	Cell Type
1	IFNβ m	Sun	[24]	HT-DNA	L929	Fibroblast
2	IFNβ m	Stanley	[124]	LPS	BMDM	Macrophage
3	IFNβ m	Stanley	[124]	TB	BMDM	Macrophage
4	IFNβ	Guo	[125]	cGAMP	BMDM	Macrophage
5	IRF3c	Sun	[24]	HT-DNA	THP1	Monocyte
6	SOCS1m	Posselt	[127]	LPS	moDC	Monocyte Derived Dendritic Cell
7	IRF3c	Wang	[129]	HT-DNA	L929	Fibroblast
8	STINGc	Jønsson	[126]	dsDNA	THP1	Monocyte
9	IRF3c	Jønsson	[126]	dsDNA	THP1	Monocyte
10	IRF7m	Qing	[128]	poly (I:C)	MEF	Fibroblast

3.2.5 Objective function formulation

To inform the MCMC algorithm, a likelihood function based on a sum-of-squares error (SSE) function was developed:

$$SSE = \sum_{i=1}^{S_m} \sum_{i=1}^{T} \frac{(N_{ij} - O_{ij})^2}{P_j} + \sum_{k=1}^{S} \left(1 - \frac{M_k}{D_k}\right)^2. \tag{3.1}$$

The objective function is separated into two components: the left component evaluating the fit of the systems dynamics when all data and state values are normalized to be [0,1] and the right component evaluating the peak concentrations of each species. The error comparing normalized observations (O_{ij}) and normalized simulations (N_{ij}) is weighted by the number of available data points (P_j) to deprioritize experiments with a large number of observations. The error between simulation magnitude (M_k) and desired magnitude (D_k) allows us to identify solutions in which the maximal concentrations are within or near biologically relevant ranges [96]. Here magnitude refers to the absolute maximum or minimum concentration achieved by a particular species. Most of the species are initiated at a concentration of zero and reach a maximum concentration at some time after DNA stimulation. In contrast, species with non-zero initial conditions achieve a minimum concentration as the simulation progresses. This minimum achieved concentration was optimized to a value 80% of the initial condition. One exception to this was DNA. The model shows DNA concentration monotonically decreasing over time meaning it does not achieve a maximum or minimum concentration. For this reason, it was excluded from this portion of the objective function. The model error is summed over all observed time points (T) for a given measured species (S_m) and the magnitude errors are summed over all observed species (S). Note that S_m represents species with observations and is a subset of S. The negative log likelihood was determined directly from the total sum-of-squares error (SSE) for each parameter set evaluated and was used by the MCMC algorithm to sample the posterior.

3.2.6 MCMC Diagnostics

To determine if the MCMC algorithm was successful in sampling posterior parameter distributions, several diagnostics were performed (see Appendix Figures A.2.1,A.2.2,A.2.3).

None of these tests give definitive proof for convergence but do provide an understanding of the overall health of the MCMC run. Three different diagnostics were performed in total. The first examines the acceptance rate of a new parameter value after each iteration. Ideally, the acceptance rate should be held at 23% [44]. The second diagnostic considers the autocorrelation which ensures PESTO is taking independent parameter samples. With high correlation the resulting effective sample size is severely reduced meaning the MCMC algorithm would have to be run for longer to obtain the same number of independent samples. The last test is the Gelman-Rubin diagnostic [45] which compares multiple independent runs to determine convergence. All diagnostics were performed using the "coda" R package [140].

3.2.7 Sensitivity Analysis

To evaluate parameter sensitivities on different states within the model over the simulated time, the AMICI toolbox was again utilized. The normalized local sensitivity was calculated as follows:

$$Sensitivity = \left(\frac{k}{x(t)} \right) \left(\frac{\partial x(t)}{\partial k} \right), \tag{3.2}$$

here, the sensitivity of state x(t) to the parameter k is shown. The normalization in front of the derivative allows for a better comparison between states [141].

3.2.8 Simulated Drug Inhibition of cGAS

A recent paper by Vincent et. al. [142] demonstrated the viability for small molecule drug interference of cGAS through competitive inhibition of cGAMP formation. To assess the applicability of our ODE model to emulate drug-mediated cGAS dynamics, we introduced two species: Drug and cGASn. Drug is a generic name for the small molecule capable of reversibly binding to the DNA:cGAS complex. That binding event produces cGASn, a drug bound cGAS molecule with no enzymatic activity. The binding event was simulated using first order mass action kinetics requiring both forward (k_{on}) and reverse (k_{off}) rate constants (see AppendixA.1, EquationsA.14 -A.16). These rates constants were unique for each drug and constrained by their experimentally determined disassociation constants

($K_d = k_{off}/k_{on}$). One of the three drugs (RU.332) did not have an experimentally esti-mated K_d value and was excluded from the analysis. The remaining two drugs (RU.365 and RU.521) were utilized in the analysis.

Data from drug dose response curves reported in Vincent et. al. showing IFNβ activation were used to parameterize model simulations. The model was simulated across different drug concentrations and the maximum IFNβ concentration was recorded. To optimize k_{on} and k_{off}, the data and model responses were normalized between zero and one and MATLAB's patternsearch function was used to minimize the sum-of-squares error. Additionally, a non-negative boundary was set on k_{on} and k_{off} to further constrain the solution space.

To compare experimental IC50 values (which were previously determined through non-linear fitting by Vincent et. al.) to our model estimated IC50 values, MATLAB's fzero function was used to calculate what drug concentration resulted in half of the maximum IFNβ expression simulated. Absolute error was then calculated to compare experimental and model determined IC50 values.

3.3   Results

3.3.1   An Experimentally Informed ODE model Emulates the cGAS pathway

We sought to develop a model of the cGAS pathway that could explain previously pub-lished experimental data, and by applying a multi-start MCMC algorithm several times, we observed two distinct parameterizations that fit known pathway responses. The resulting dynamics of those parameterizations can be seen in Figure 3.3.2 where we plot the time profiles for each species for the top 1000 parameterizations (instances) identified in three MCMC parameterizations (labeled as runs 1-3). Overall, the dynamics of each species re-main similar in both timing and magnitude for the different parameterization runs except for DNA. For runs 1 and 3, a delay between DNA binding and TREX1 degradation is observed as DNA concentration stagnates for several hours. In run 2, there is steep decline in DNA concentration—likely due to the lack of knowledge of the rate of DNA degradation. Testing
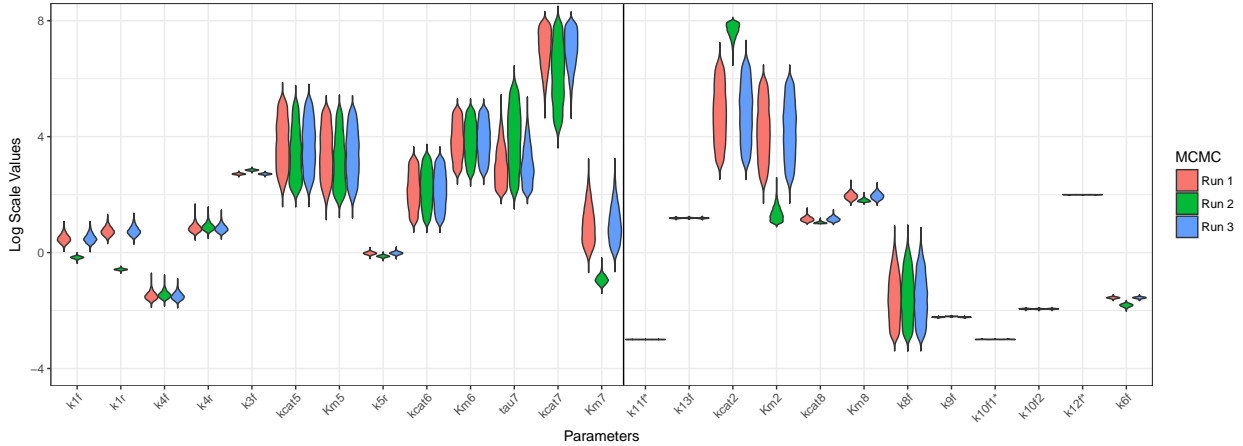
Figure 3.3.1: Model Parameter Distributions

the validity of either prediction experimentally would provide a greater understanding of the dynamics regulating cytosolic DNA.

Despite differences in the apparent dynamics, the top 1000 parameterizations from each run equally minimized the objective function (Equation 3.1). After 10 million iterations of the MCMC algorithm, the absolute difference in error between each run was at most 4×10^{-4} (unitless). As illustrated in Figure 3.3.3, the convergence of the MCMC algorithm occurred after approximately 1 million iterations after which only minor changes in the objective function were observed. To determine the relative error contribution provided by each of the components of the objective function, we separately plotted the dynamic and peak error as defined by Equation 1 (see Appendix Figure A.2.4). Here we observe the error for each component is on the same order of magnitude meaning both offer significant contributions to the overall error.

We next examined the parameter distributions obtained from the MCMC traces (Appendix Figure A.2.5) for the 3 parameterized runs. In Figure 3.3.1 we report summary statistics of the parameter distributions obtained after burn-in for the last 1 million instances per run and confirm differences in parameterization from the second MCMC run. Appendix Figure A.2.6 shows the complete parameter distributions obtained after 1 million MCMC instances post burn-in. Major deviations are observed in k1f and k1r (cGAS/DNA binding)

as well as kcat2 and Km2 (DNA degradation) explaining differences observed in DNA dynamics (Figure 3.3.2). Some parameters (k11f, k10f1, and k12f) were pushed to boundaries set before optimization and shifting these boundaries resulted in integration errors; therefore, conclusions concerning these parameters are limited. For all follow-up analyses of the model (and for the trajectories shown in Figure 3.3.2), we restricted our analyses to the top 1000 parameter sets for each run (3000 sets total). To ensure this parameter subset accurately sampled the total distribution determined by MCMC after burn-in, we compared these distributions against each other and observed sufficient agreement (Appendix Figure A.2.7).
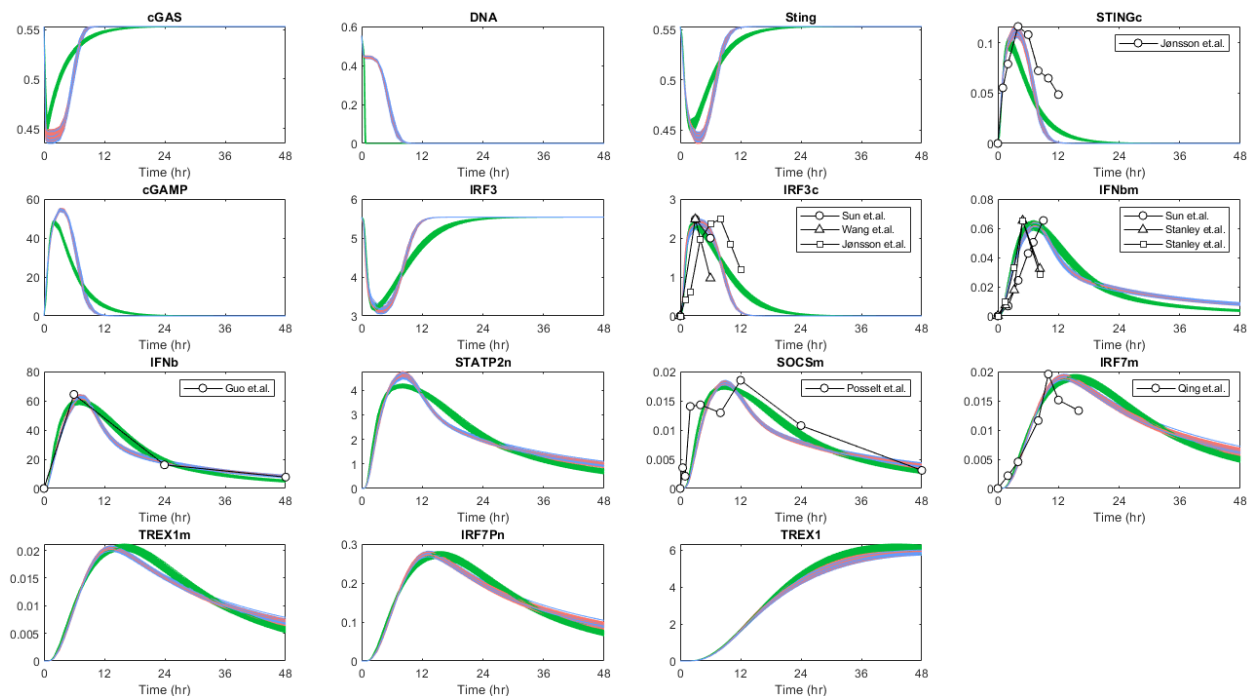
Figure 3.3.2: Trajectories of the top 1000 Parameterizations.

In addition to parameter distributions, we also investigated parameter correlations (Appendix Figure A.2.8) and observed identical correlations for runs 1 and 3, but differences in run 2. Common to all runs was a strong correlation between some Michaels Menten parameters (kcat5/Km5, kcat6/Km6, kcat7/Km7, and kcat8/Km8) and mass action parameters (k4f/k4r) suggesting structural non-identifiability. Looking at differences between the separate runs we note high correlation in the kcat2 and Km2 parameters (associated with TREX1) to k1f, k3f, and k3r (associated with cGAMP) only for run 2. A strong correlation

was also observed exclusive to runs 1 and 3 between k1f and k1r to k4f and k4r. These correlations suggest identifiability issues specific to that parameterization. To alleviate these run specific problems more data would be needed for parametrization.
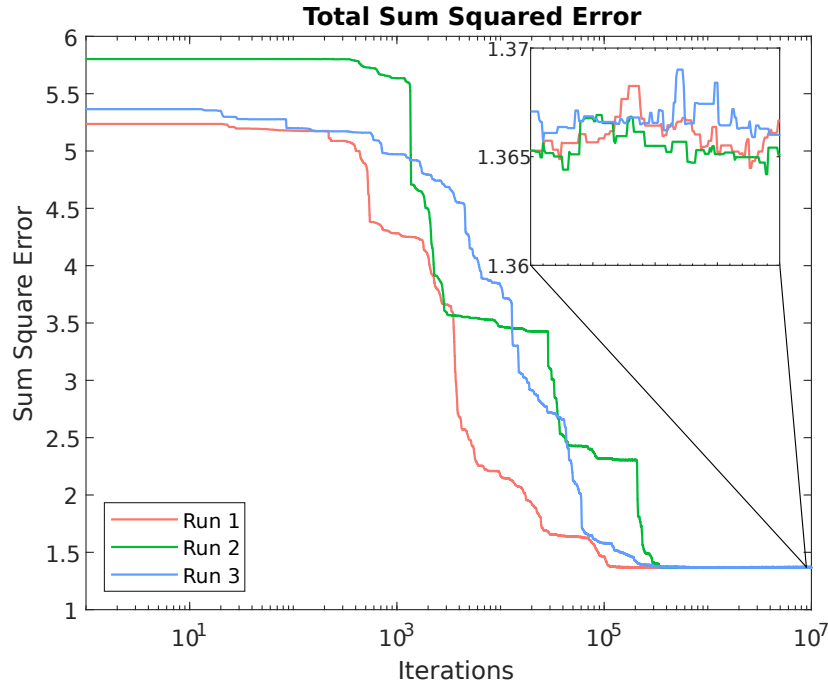
Figure 3.3.3: Error function: MCMC Performance

3.3.2 Sensitivity analysis reveals IFNβ, cGAMP, and TREX1 to be sensitive to parameter perturbations

As MCMC produces ensembles of models capable of fitting relevant data, we selected the top 1000 models identified in each MCMC run (same ensembles as those in Figures 3.3.2) and performed a local sensitivity analysis to identify the most influential processes. To summarize the results of this analysis, we calculated the mean maximum sensitivity achieved for each state in each of the 1000 models across the three independent MCMC runs (3000 model parameterizations in all). Figure 3.3.4 shows the results for all states exhibiting significant parametric sensitivity (see Appendix Figure A.2.9 for all species). Error bars represent the range of maximum sensitivity obtained over all iterations. Comparing the different MCMC

runs we find that all three runs follow similar trends but the absolute magnitude of sensitivity for the second run is lower for most of the parameters. This suggests that differences in sensitivity between parameters are determined by model architecture and independent of the specific parameter value.
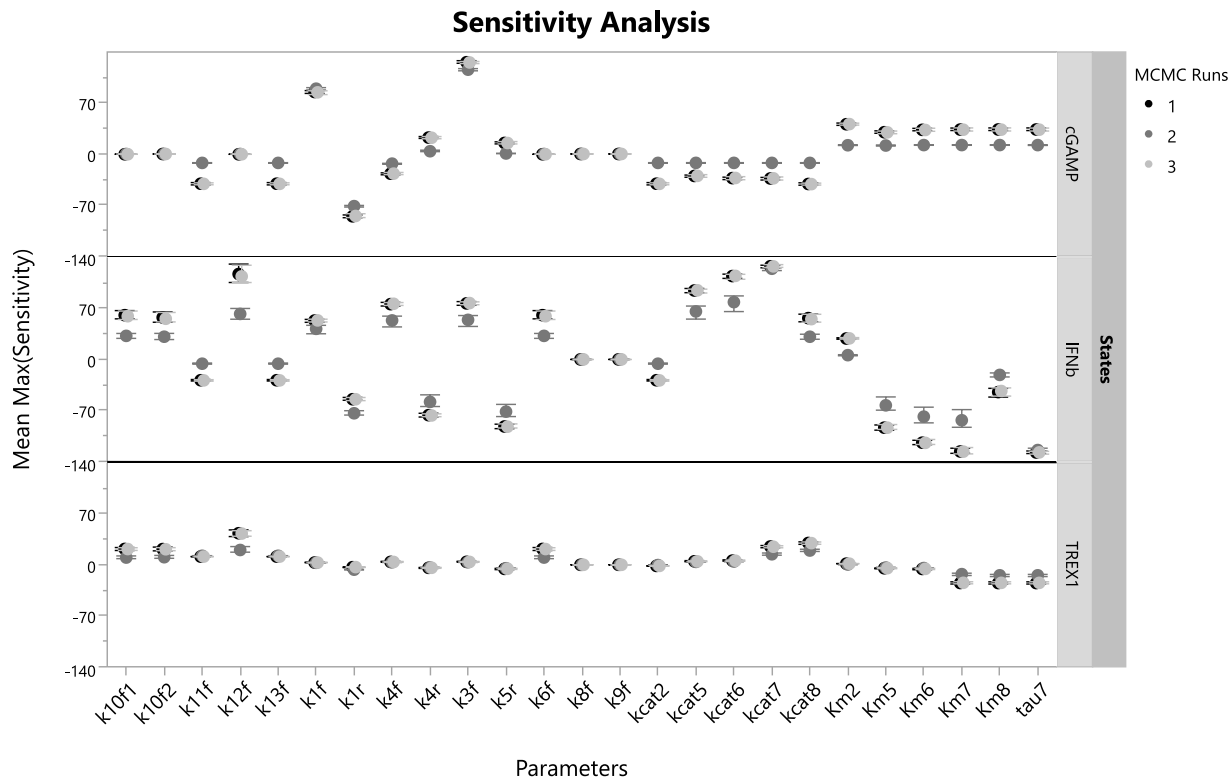
Figure 3.3.4: Sensitivity Analysis. Represented are the three most sensitive species found in the ODE model. Each dot represents the mean maximum sensitivity determined by the AMICI toolbox across the 1000 instances. The error bars show the range of maximum sensitivity found for the top 1000 parameter sets for each run.

IFNβ, TREX1, and cGAMP are most sensitive to parametric uncertainty and are generally affected by processes which directly affect their activity. IFNβ has the largest negative sensitivity with its degradation rate (τ7) and shows the strongest positive sensitivity with the translation of IFNβ mRNA (kcat7). The sensitivity analysis also reveals that IFNβ was insensitive to the SOCS feedback mechanism (k8f and k9f) and was primarily regulated through its degradation rate constant (τ7). This result suggests that the model is insensitive to uncertainty in this feedback mechanism to accurately describe system dynamics.

cGAMP exhibited the second highest overall sensitivity to parametric uncertainty, showing positive sensitivity to STING:cGAMP complex formation (k3f) and negative sensitivity to the unbinding of cGAS and DNA (k1r). Finally, TREX1 is the only other species that showed significant sensitivity to optimized parameters. The highest positive sensitivity was seen from the translation of IRF7 mRNA (k12f) and showed no single significant negative sensitivity. This positive sensitivity suggests that perturbations or uncertainties in the positive feedback mechanism involving IRF7 impacts the negative feedback of TREX1 regulated DNA degradation.

3.3.3 Inhibition of TREX1 activity demonstrates robustness to a chronic inflammatory state

Next, we simulated an *in-silico* knock-down simulation of TREX1 to determine the robustness of IFNβ activation from partial to full inhibition of TREX1 activity. Aberrant activation of the cGAS pathway has been demonstrated to lead to autoimmune diseases like Aicardi-Goutieres syndrome (AGS) and Systemic lupus erythematosus (SLE) and a lack of or inhibited activity of TREX1 has been suggested as the cause [110]. Knock-down was simulated by varying the kcat2 parameter responsible for the degradation reaction between DNA and TREX1. Figure 3.3.5 shows a representative sample of the simulated knock-down for the top 1,000 parameter sets from Runs 1 and 2 (Run 3 results are redundant with Run 1). With both runs, IFNβ levels are strongly robust to partial inhibition but demonstrate sustained IFNβ levels at 100% inhibition. Complete loss of function is necessary to induce a chronically inflamed state. Some major deviations in dynamics are observed between run 1 and run 2. Run 1 (Figure 3.3.5a) shows a higher steady state value for cGAS and DNA than Run 2 (Figure 3.3.5b) by 0.15 nM (25% difference). These differences in steady states are a result of the different parameterizations discussed above.

In addition to simulating TREX1 knock-down, IRF7 knock-down was also performed. Figures 3.3.5c (runs 1 and 3) and 3.3.5d (run 2) show knock-down of IRF7 activity by modifying the k12f parameter responsible for IRF7 translation. Contrary to TREX1 knock-down, IFNβ and TREX1 concentrations were not robust to changes in IRF7 activity. IRF7

knock-down has no impact on upstream species (e.g. DNA and cGAS), but reduction does lead to faster dissipation of IFNβ. This is expected as IRF7 is a transcription factor for IFNβ. In sum, the key observation is that partial inhibition of TREX1 has no impact IFNβ activity, whereas IRF7 inhibition induces changes in IFNβ activity.

(a) Run 1: TREX1 Knock-down

(b) Run 2: TREX1 Knock-down

(c) Run 1: IRF7 Knock-down

(d) Run 2: IRF7 Knock-down
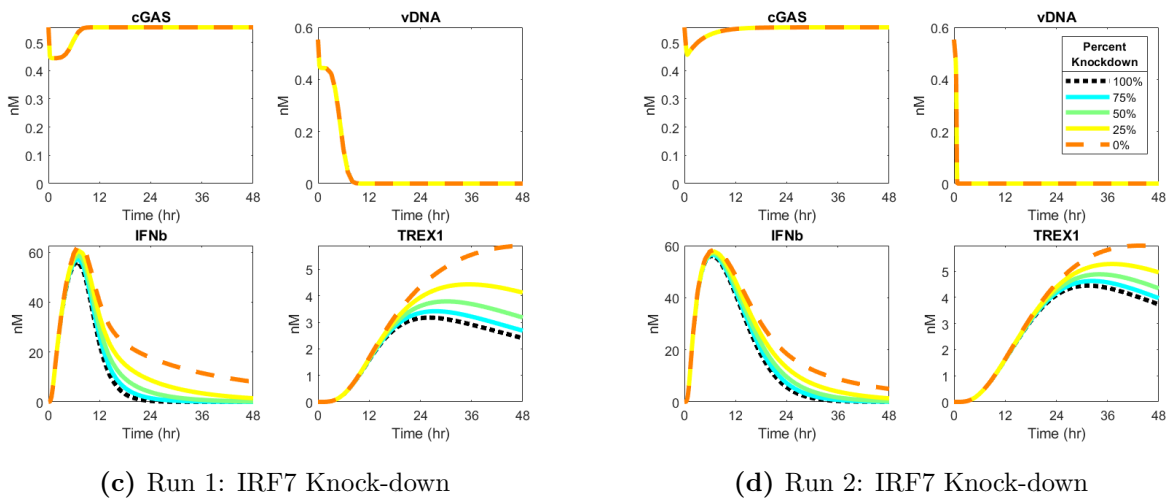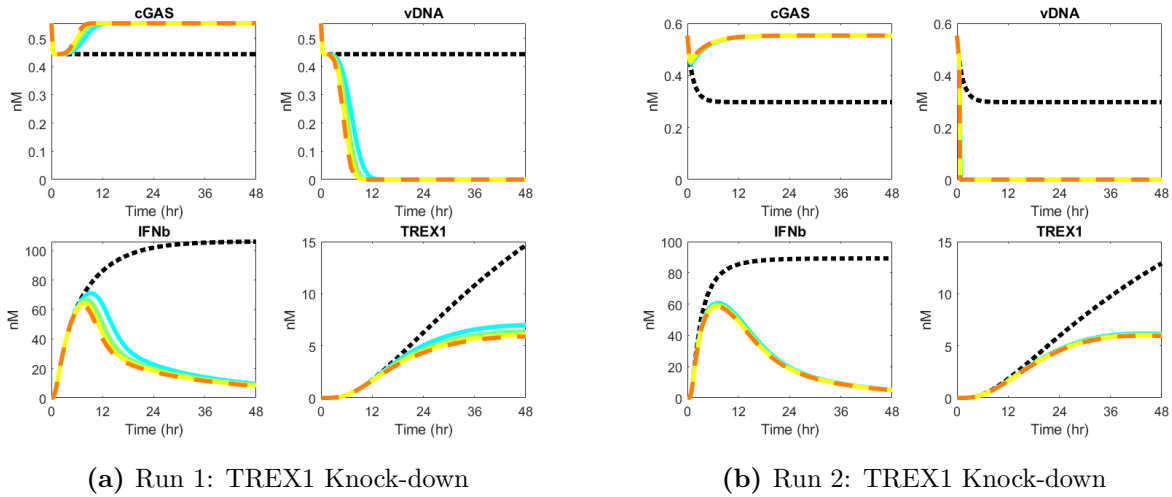
Figure 3.3.5: Comparing Feedback Knock-down of different Parameterizations

3.3.4 Drug inhibition of cGAS activity is recapitulated in simulation

We examined the model's capacity to predict interferon responses in scenarios where cGAS is exposed to an inhibitory drug (Figure 3.3.6). The experimental data originally published by Vincent et. al. [142] confirmed drug inhibition of cGAS using an interferon luciferase reporter assay testing across a wide range of drug concentrations. The two drugs tested (RU.365 and RU.521) all possess a small concentration range within one order of magnitude that induce substantial changes in IFNβ production. If this jump in IFNβ production occurs at lower drug concentrations it suggests a higher drug potency. To model differences in drug potency, we varied the forward rate constant (k_{on}) affecting drug and cGAS binding while simultaneously constraining to the experimentally determined K_d values. One of the three drugs did not have a reported K_d value and was excluded in this analysis. The optimized rate constant for RU.365 ($k_{on} = 2.314 \times 10^{-4}$ $nM^{-1}hr^{-1}$) and RU.521 ($k_{on} = 3.099 \times 10^{-3}$ $nM^{-1}hr^{-1}$) indicate that RU.521 shows the highest potency. The reported IC50 values for each drug follow an inverse trend compared to determined rate constants which is consistent as more potent drugs have lower IC50 values. The IC50 values calculated directly from the model simulations (displayed on Figure 3.3.5b) are in good agreement with the experimentally determined values, with absolute errors of 0.57 μM (RU.365), and 0.16 μM (RU.521).

(a) Mechanism

(b) Dose response curves
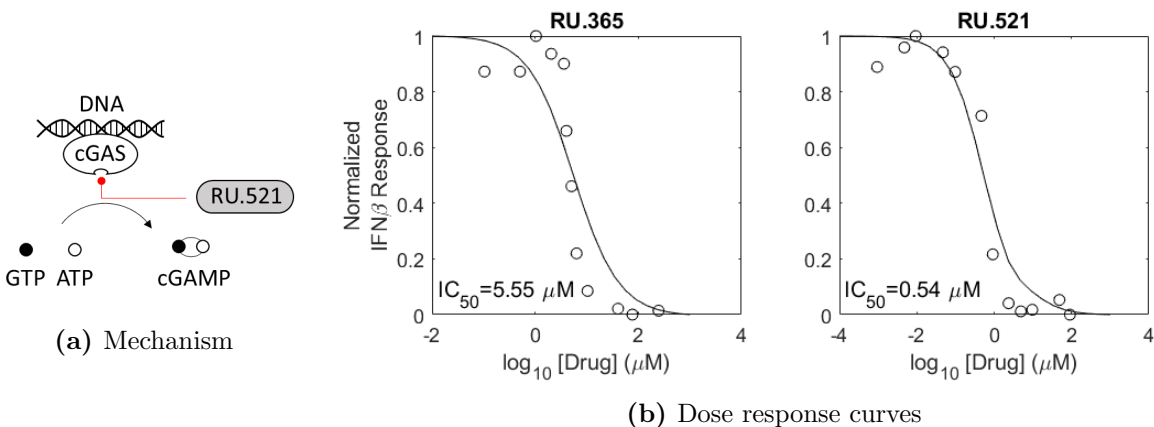
Figure 3.3.6: Simulated Drug Inhibition of cGAS

3.4 Summary

Here, a first principles mathematical model of the cGAS pathway was developed and used to evaluate emergent systems-level properties associated with interferon regulation. The importance of the cGAS pathway is demonstrated in the growing number of pathogens and diseases it is associated with including bacteria [101, 107], DNA viruses [108], retroviruses [143], autoimmunity [109], cancer [29], and aging [32]. There are no existing ODE models for this system, allowing this model to fill a gap in the field's knowledge of cGAS pathway dynamics. In total, working towards a quantitative, systems level understanding of the cGAS pathway can improve insight into interferon regulation and support treatment design.

The data used to inform the model were collected from 10 experimental observations and utilized two different categories of cells: immune cells derived from monocytes and connective tissue cells (Table 1). Both types of cells produce an interferon response through the cGAS pathway, but the magnitude of these responses can vary across cell type with immune cells like plasmacytoid dendritic cells (pDCs) producing more potent responses [144]. Because of this we restricted our data search to cell types whose responses were similar in magnitude. Chessler [139] makes an excellent comparison showing how T. cruzi induces IFNβ expression through the IRF3 transcription factor in both fibroblast (MEF) and macrophage (BMDM). Both cell types produce a 1100-fold change in IFNβ mRNA expression, suggesting our use of these cell types is justifiable. Because we use data from different sources, normalization was performed to equally compare relative responses across the distinct experiments. This ensured the use of different controls within experiments did not influence the model outcome. In addition to differences in magnitude, some cell types are going to respond more rapidly to pathway stimulation [145]. However, the data selected (shown in Figure 3.3.2) comes from cell types which demonstrated similar IFNβ dynamics.

MCMC optimization resulted in two distinct parameterizations determined by three independent runs of the algorithm (Appendix Figure A.2.10 demonstrates the two clusters formed). While these parameterizations all equally minimize the objective function, their emergent behavior is different. The parameterization resulting from run 2 demonstrates behavior that, from a biological standpoint, seems improbable as cytosolic DNA is completely

58

consumed within 1 hour. The dynamics of cytosolic DNA from run 2 parameter sets exhibited no delay between cGAS binding and TREX1 degradation. A time delay should exist because TREX1 is only produced after the cGAS and JAK/STAT pathways are activated.

The overall structuring of the ODE model required several simplifying assumptions to reduce the size of the parameter space and remove nonlinearities, making the optimization more tractable. One major assumption made was neglecting the dimerization of transcription factors like IRF3. The implications of this assumption were tested, and the resulting dynamics were compared (Appendix Figure A.2.11). The only moderate difference between the two models was that more IRF3 was needed to create the activated dimer—leading to a larger drop in IRF3 levels. Because we have limited knowledge of the absolute dynamics of the IRF3 dimerization, and the relative timing of the dynamics are similar, we found this assumption to be reasonable.

Based on calculated correlation values between the optimized parameters, we conclude the model is suspect to issues with structural non-identifiability. The ability to explicitly determine which parameters fall under this category is possible with small linear differential equation models, but by increasing the number the states and introducing non-linear terms (like Michaelis Menten) the task become intractable or even impossible [146]. Having structural non-identifiability means there exist no unique parameter set that best explains the data. Knowing this ubiquitous problem exists in systems biology we used a Bayesian optimization approach to consider families of parameter sets as opposed to one non-unique parameterization. Our focus was not on the individual parameter values, but rather what emergent properties where obtainable across all parametrizations.

The local sensitivity analysis revealed that only 3 states are strongly sensitive to parameter fluctuations (see Appendix Figures A.2.9 and A.2.12) and, as with observations made in other biological pathways [147], the ranking of their sensitivity across parameters is independent of selected parameter set. Sensitivity can guide model refinement and drug targeting, but large uncertainties associated with biological parameters can lead to changing system priorities. We show here that for exceptionally large distributions of parameter values, the ranking of parameter sensitivities is constant; suggesting that sensitivity is the result of the system architecture and not the system parameterization. Supported by this result, the sen-

sitivity analysis suggests that IFNβ is best regulated by manipulating local mechanisms but is still sensitive to several up and downstream kinetics. Thus, there exists multiple methods to manipulate IFNβ concentrations and the models can be used to optimize the proposed manipulation.

The TREX1 regulated DNA degradation appears to be critical to capturing robust interferon signaling. Robustness is an emergent systems property in which a system maintains function (i.e. interferon production) despite large changes in the intracellular environment [117]. As shown in Figure 3.3.5, the presence of the TREX1 feedback is required to degrade cytosolic DNA and ultimately prevent further interferon production. Yet, if TREX1 feedback is present, large changes in parameters associated with the protein's activity do not strongly impact interferon production. Even a 75% reduction in TREX1 activity did not appreciably alter IFNβ/IFNβ mRNA levels, meaning the pathway's activity is highly robust against perturbation in TREX1 feedback. This is important for two reasons: First, it suggests that parameterization does not need to be highly accurate. The model can have relatively large errors in TREX1 associated parameters and still maintain the appropriate dynamic responses (as dictated by the simulation in Figure 3.3.2). And second, this suggests a relatively simple experiment that would provide a strong evaluation of the model.

To show how the mode could replicate experimental measurements performed, we demonstrated its ability to simulate drug inhibition of cGAS. The methods used to simulate this inhibition required several assumptions including: (1) reversible drug binding to the cGAS:DNA complex, (2) first order mass action kinetics, (3) no degradation of drug within the cell, and (4) no interaction between the drug and unbounded cGAS (i.e. no DNA bound). Even with these strong assumptions we were able to successfully recapitulate IFNβ dynamics and determine IC50 values with a maximum absolute error of 0.57 μM. With more sophisticated dynamics, such as the implementation of hill kinetics, drug inhibition of cGAS could be more accurately simulated.

One assumption we did not make while adding the drug interaction into the model was fast equilibration between the drug and cGAS complex. The timescales for DNA or drug binding to cGAS are significantly smaller than transcription and translation events in the model. This can be observed in the steep initial drop-off of cGAS concentration in Figure

2. Using this assumption would be beneficial in analyzing dynamics on long timescales as it simplifies computational complexity but were not included to allow for the potential investigation of small timescale dynamics.

The flexibility of an ODE model provides a fast and effective method to understand how a biological pathway behaves under diverse situations. The model presented in this paper illustrates the dynamics of the cGAS pathway expected from a DNA transfection experiment with and without cGAS inhibition. This is fundamentally different from *in-vivo* activation of the cGAS pathway because it ignores dynamics caused by disease interactions. Many virus strains have evolved to transcribe proteins that target the cGAS pathway [103, 106] and cancer cells have developed mutations to evade the cGAS immune response [28]. The ODE model presented in this paper can be easily modified and extended to predict dynamics specific to these scenarios. Viral proteins, for example, can be included to simulate the effective inhibition of the cGAS pathway. With these additions further hypotheses can be tested to determine if new treatments restore the immune response. Overall predictions garnered by the ODE model will provide crucial insight into how we can manipulate disease outcome.

4.0 Agent-Based Modeling Reveals Benefits of Heterogeneous and Stochastic Cell Populations during cGAS-Mediated IFNβ Production

4.1 Introduction

The advancement of microscopy and single cell analyses have identified IFNβ expression—a critical component of the innate immune response—as an inherently stochastic process that is regulated by interacting, highly heterogeneous cells [148, 145, 37]; however, what advantages stochastic responses and heterogeneous populations may give to signaling in the context of immunity remains an open topic of study. When a cell population encounters a pathogenic threat, a majority of the population will not respond despite being infected, and those that do will vary greatly in their response [148, 149, 150]. Multiple mechanisms have been proposed to explain these observations, with many citing either intrinsic or extrinsic sources [145, 151]. Low protein and mRNA concentrations are common intrinsic sources of stochasticity because they result in probabilistic responses within the cell (e.g. transcriptional busting) [152, 153, 149]. This source of noise can amplify and propagate down signaling pathways eventually affecting gene expression and cellular behavior [154]. In contrast, extrinsic sources implicate heterogeneity in the immune stimulus as well as spatial variation in signaling molecules [155]. Taken together, these layers of noise and variability coalesce to explain the complex dynamics observed in immune cell signaling.

The benefits that stochastic and heterogeneous cell populations bring to innate immune signaling remain unclear. Some evidence suggests that cells are attempting to optimize the potent innate immune response to simultaneously minimize pathogenic loads and maximize cell survival [37]. Overexpressing innate immune cytokines can provoke unnecessary damage to the host [156], leading to cytokine storms and chronic inflammation [157]. More recent evidence suggests that stochasticity allows a cell population to subdivide, where some cells produce substantial immune responses that risk apoptosis while others are preserved to maintain tissue viability [158, 159, 153]. In contrast, hypotheses elucidating the advantages of cellular heterogeneity are less developed. One potential explanation is bet-hedging [160],

where less-fit individuals are maintained as a precaution to drastic changes in the environment. Others suggest that heterogeneity acts as an additional layer of non-genetic variability [161, 162] which can impede pathogenic threats susceptible to noise [163]. Investigating why cell populations employ stochasticity and heterogeneity during the IFNβ immune response allows for the improvement of existing drug therapies [164] and more broadly an increased understanding of host immune regulation.

Multiscale computational modeling can assess how cell stochasticity and heterogeneity impact innate immune signaling across intracellular and cell population scales. However, developing models that incorporate these features are difficult to implement because they are conventionally modeled using separate approaches. Intracellular signaling is typically modeled using ordinary differential equations (ODEs) by assuming spatial differences are negligible within the cell [85, 38]. In comparison, intercellular signaling relies on diffusion gradients described by partial differential equations (PDEs) which need to be accounted for to model cytokine signal propagation. Combining these two modeling paradigms can often lead to stiff systems that become numerically unstable [165]. Fortunately, agent-based modeling (ABM) methods and computational power have greatly advanced, supporting simulation of hybrid ODE, PDE, and rule-based models that easily incorporate cellular behavior like stochasticity and heterogeneity [166, 167]. ABM is a general computational technique that defines individual entities—or agents—containing rulesets which determine how these agents interact with each other and the environment. The rules provided to a cellular agent can be simple, such as basic conditional logic, or complicated, such as using ODEs to track intracellular protein concentrations. The culmination of these interacting agents results in emergent behavior not obtainable from modeling an individual cell [168, 169, 170]. ABMs have been instrumental in revealing the emergence of metastasis [49], angiogenesis [171], structural patterns within the heterogeneous tumor environment [50, 172], as well as simulating granuloma formation in tuberculosis [173], stem cell differentiation [51], and cellular damage in pulmonary fibrosis [174].

Here, we implement an ABM of the cGAS pathway, a critical component of the innate immune system responsible for the detection of foreign DNA and expression of IFNβ [24]. Due to the expansive roles cGAS plays in disease detection (e.g. cancer [28], inflammation

[110], acute kidney injury [175], HSV [108], HIV [102], MTB [101]), it is imperative to further develop an understanding of this signaling pathway to continue improving new drug therapies [176], cancer immunotherapies [177, 178], and vaccine adjuvants [179]. To accomplish this task, the ABM is used to identify the immunologic advantages a cell population may have when the individual cells are heterogeneous and subject to stochastic intracellular IFNβ signaling. We show that cell populations containing heterogeneous protein compositions can mount an IFNβ response that is stronger, weaker, or equal to homogeneous populations. We also find that there exists an optimal level of stochasticity within the cGAS-induced IFNβ response to pathogenic infection. These simulations further suggest that increasing the fraction of cells that produce IFNβ does not lead to an increase in cell survival during infection. In all, these simulations corroborate experimental observations and provide new insight into how cell heterogeneity and stochasticity play a role in optimizing cGAS-mediated IFNβ signaling.

4.2   Materials and Methods

4.2.1   Incorporating ODEs into an Agent-Based Model

The agent-based model is constructed on a 200 by 200 grid where each section represents an epithelial cell (40,000 cells total, see Figure 4.3.1). The agents remain stationary, but paracrine signaling dynamics allow neighboring cells to transmit information about their current state and influence the population as a whole. Grid sections are sized such that the side length is roughly equivalent to the diameter of a mammalian human cell (32 μm), which provides the proper scale for diffusion [61, 62]. An ordinary differential equation model of the cGAS pathway is utilized as a ruleset for each cell in the simulation [55]. This allows an agent to detect immune stimulatory DNA (ISD) or a DNA-encoded virus and respond by upregulating interferon beta (IFNβ). This signaling protein could then diffuse across the cell population and upregulate the JAK/STAT pathway to promote various feedback mechanisms including IRF7, SOCS1/SOCS2, and TREX1. The ODE model remained largely unchanged,

except for the addition of the diffusion term on IFNβ:

$$\frac{\partial[IFN\beta]}{\partial t} = \frac{k_{cat7}[IFN\beta_{mRNA}][DNA]}{K_{m7} + IFN\beta_{mRNA}} - \tau_7[IFN\beta] + \mathcal{D}\nabla^2[IFN\beta] \qquad (4.1)$$

Parameters used to simulate the model are taken from the best (i.e. lowest error) MCMC iteration from the original ODE model. Two independent sources were found for the diffusion coefficient of IFNβ (95 $\mu m^2/s$ [180] [63] and 100 $\mu m^2/s$ [181]) so an average of the two are used in the simulation.

4.2.2 Agent Based Model Simulation

The ABM is simulated using the *DifferentialEquations.jl* package in Julia using a semi-discretized approach [54]. The Laplacian modeling the diffusion process is discretized using a second order central differencing scheme with no flux (Neumann) boundary conditions. This combined with the 13 ODEs encoded in each simulated cell, results in a system of 520,000 ordinary differential equations. The scale of this model requires a well-optimized solver that is efficient in both computing time and memory, as well as being numerically stable. We found that higher order stabilized Runge-Kutta methods (e.g. ESERK5 and ROCK4 [182, 183]) were proficient in some model iterations but led to numerically unstable oscillatory states when simulating viral infections. This is in part caused by jump discontinuities exhibited by some of the species, such as a spike in viral DNA concentration after a cell transitions from a healthy to infected state. We settled on using an implicit ODE solver (backward differentiation formula, BDF [53]) which eliminates numerical instability but requires solving a computationally expensive system of linear equations at each time step. Instead of solving the linear system directly, we use a generalized minimal residual method (GMRES) to accelerate the BDF solver. GMRES can efficiently estimate the solution to the linear system after a few iterations of the algorithm which helps minimizes the computational cost of the model [184].

Two initial conditions are implemented into the model depending on whether an ISD transfection or viral infection was simulated (Figure 4.3.2). The ISD initial condition defines a circular region to specify a boundary between transfected and healthy cells. Cells within

the boundary are considered infected and are given a nonzero concentration of ISD. The viral infection initial condition uses a different approach that assumes each cell within the population has some probability of being initially infected according to a Poisson distribution:

$$P(n) = \frac{(MOI)^n \cdot e^{-MOI}}{n!}, \tag{4.2}$$

here, n refers to the number of times a cell is infected, and MOI is the multiplicity of infection. The fraction of cells initially infected are varied by changing MOI. A low dose simulated infection, for example, would have an MOI of 10^{-3}, which corresponds to an average of 40 cells being initially infected at least once.

4.2.3 Simulating Cellular Heterogeneity and Stochasticity

Two methods are used to simulate differences within individual cells: heterogeneity is modeled by changing initial conditions and stochasticity is simulated by randomly modifying ODE parameters associated with IFNβ production. Initial conditions are varied by sampling from a truncated normal distribution (bounds 0 to infinity) with a mean specified by the original initial condition and a set variance. By surveying different variances, we can determine how sensitive the interferon response is to heterogeneity in cellular composition. The default variance was set to 0.5 nM^2. To simulate stochastic IFNβ signaling, cells are randomly assigned different k_{cat7} parametrizations (Eq. 4.1) This parameter is responsible for IFNβ mRNA translation and is treated as a Bernoulli distributed random variable with the probability of failure equal to the desired percentage of interferon producing cells. If the desired percentage of interferon producing cells was 30%, for example, there would be a 30% chance the k_{cat7} parameter would remain unchanged and a 70% chance the parameter value would be set to zero, effectively preventing the cell from producing IFNβ. The default percentage was set to 20%.

4.2.4 Simulating Viral Infection

To simulate a viral infection, an additional ODE state "Virus" is added to the ABM and the DNA state is modified to incorporate replication.

66

$$\frac{d[Virus]}{dt} = k_{14f}[DNA] - \tau_{14}[Virus] \tag{4.3}$$

$$\begin{aligned}\frac{d[DNA]}{dt} =& k_{1f}[cGAS][DNA] + k_{1r}[cGASc] - \\ & \frac{k_{cat2}[TREX1][DNA]}{K_{m2} + [DNA]} + \frac{[DNA](K_{max} - [DNA])}{K_{max}}\end{aligned} \tag{4.4}$$

When a cell becomes infected by a virus, it releases DNA into the cytosol. This is simulated in the model as a discontinuous jump in DNA concentration subject to a uniform distribution on the interval $[0, K_{max}]$. The modified DNA state can now vary from cGAS binding, TREX1 degradation, or replication specified by the carrying capacity ($K_{max} = 0.55$ nM). As DNA accumulates in the cell, it drives the production of more virions, which can then be used as an indication of how infectious that cell is to neighboring cells.

The probability that a healthy cell transitions into an infected cell is assumed to depend on the surrounding cell's viral concentrations (i.e. Moore's neighbors). Each cell is assigned a viral load threshold value randomly sampled from a normal distribution with a mean of 800 virions and standard deviation of 200 virions. When an infected cell's viral concentration exceeds its given threshold, it attempts to infect each healthy neighboring cell through a Bernoulli trial with the probably of success defaulted to 0.5. Modifying this parameter effects how quickly the infection spreads throughout the population, with low values effectively halting viral spread. When a healthy cell is infected, it is assigned a nonzero concentration of viral DNA and the infected cell is set to no longer be infectious.

In addition to a viral load threshold value, cells are also randomly assigned a length of time they are allowed to be infected. This is also sampled from a normal distribution with a mean 8 hours and standard deviation of 1 hour. When an infected cell surpasses its allowed time to be infected, a callback event is triggered that modifies the cells ODE parameters. All parameters associated with transcription and translation are set to a value of zero. Proteins, mRNA, and other molecules then quickly degrade because the cellular agent has no means to sustain their concentrations. This method is chosen over removing the agent from the simulation because the ODE solver would become unstable as the number of ODEs are changed and discontinuities are introduced.

4.3   Results

4.3.1   Stochasticity and Heterogeneity Modulate IFNβ and Increases Cell Survival
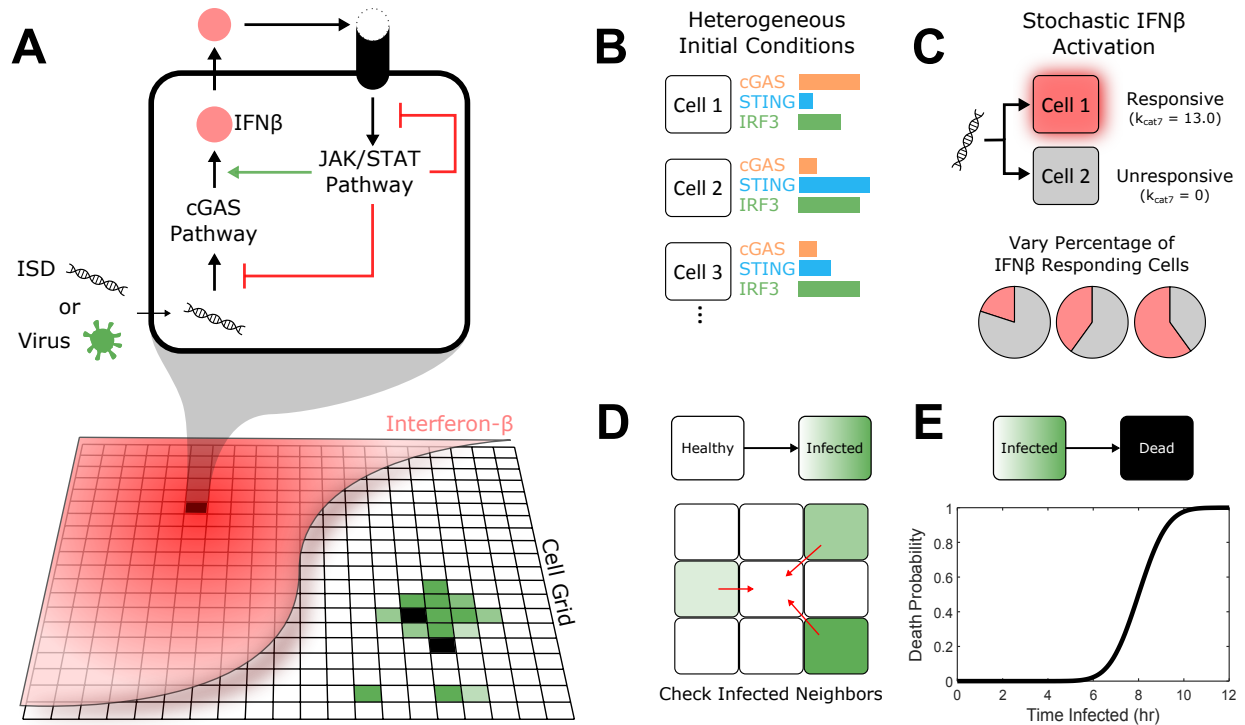
Figure 4.3.1: Overview of the agent-based model construction. **(A)** The model consists of two layers: a bottom grid of cells and a top layer portraying the concentration of IFNβ diffusing across the cells. The bottom grid layer is colored green according to the internal concentration of cytosolic DNA (viral or ISD) with darker colors representing higher concentrations. Healthy cells are colored white and dead cells are colored black. Each cell possesses an ODE model of the cGAS and JAK/STAT signaling pathways which allow for the detection cytosolic DNA, production of IFNβ, and upregulation of feedback mechanisms. **(B)** The ABM can simulate heterogeneous cell populations by manipulating the non-zero initial conditions. Initial conditions are sampled from a normal distribution and assigned to each cell at the beginning of the simulation. **(C)** Cells respond to DNA stimulus stochastically by either producing interferon (responsive) or by producing no interferon (unresponsive). The percentage of responsive cells are varied across simulations. **(D)** Healthy cells transition into an infected state subject to the viral concentrations of neighboring cells (Moore neighborhood). As the concentration of virus increases within an infected cell, the probability of infecting a neighboring cell also increases. **(E)** Infected cells then transition into dead cells over time with increasing probability.

With the development of the ABM (Figure 4.3.1), we evaluated the ability of the model to replicate different population-level experimental results and then analyzed the effects of stochasticity and heterogeneity on cGAS signaling, virus growth, and cell death. To compare across the various model simulations, we choose to monitor IFNβ signaling at the cell population level. Figure 4.3.2 illustrates the IFNβ distribution and cell state (healthy, infected, dead) across the cell population 10 hours after the initial ISD transfection (Figure 4.3.2A and 4.3.2D) or DNA virus infection (Figure 4.3.2B,C and 4.3.2E,F). Several simulations were performed under different cellular and intracellular conditions to compare outcomes when the cells are homogeneous/deterministic versus heterogeneous/stochastic cells. During transfection (Figure 4.3.2A and 4.3.2D), a dashed-lined boundary divides the population into primary cells, which directly receive a DNA transfection, and secondary cells, which receive no DNA stimulus and only respond to paracrine IFNβ signaling. Transfection simulations with homogeneous/deterministic cell populations (Figure 4.3.2A) exhibited higher overall IFNβ production and more secondary cell activation than observed in a heterogeneous/stochastic cell population (Figure 4.3.2D).

Viral infection simulations exhibited a different trend in the IFNβ profiles (Figures 4.3.2B and 4.3.2E), but the number of cells dead at hour 10 is comparable in both viral infection simulations (Figures 4.3.2C and 4.3.2F), at approximately 2%. These results suggest that stochasticity and heterogeneity together can significantly influence the amount of IFNβ a cell population produces (in this case reducing IFNβ); however, this influence does not extend directly to cell survival as it remains unchanged. Stochasticity and heterogeneity appear to benefit the cell population by diminishing the response needed to overcome the same threat. To further develop this hypothesis, the following sections will investigate how stochasticity and heterogeneity affect intracellular signaling and determine how each individually impacts the immune response.
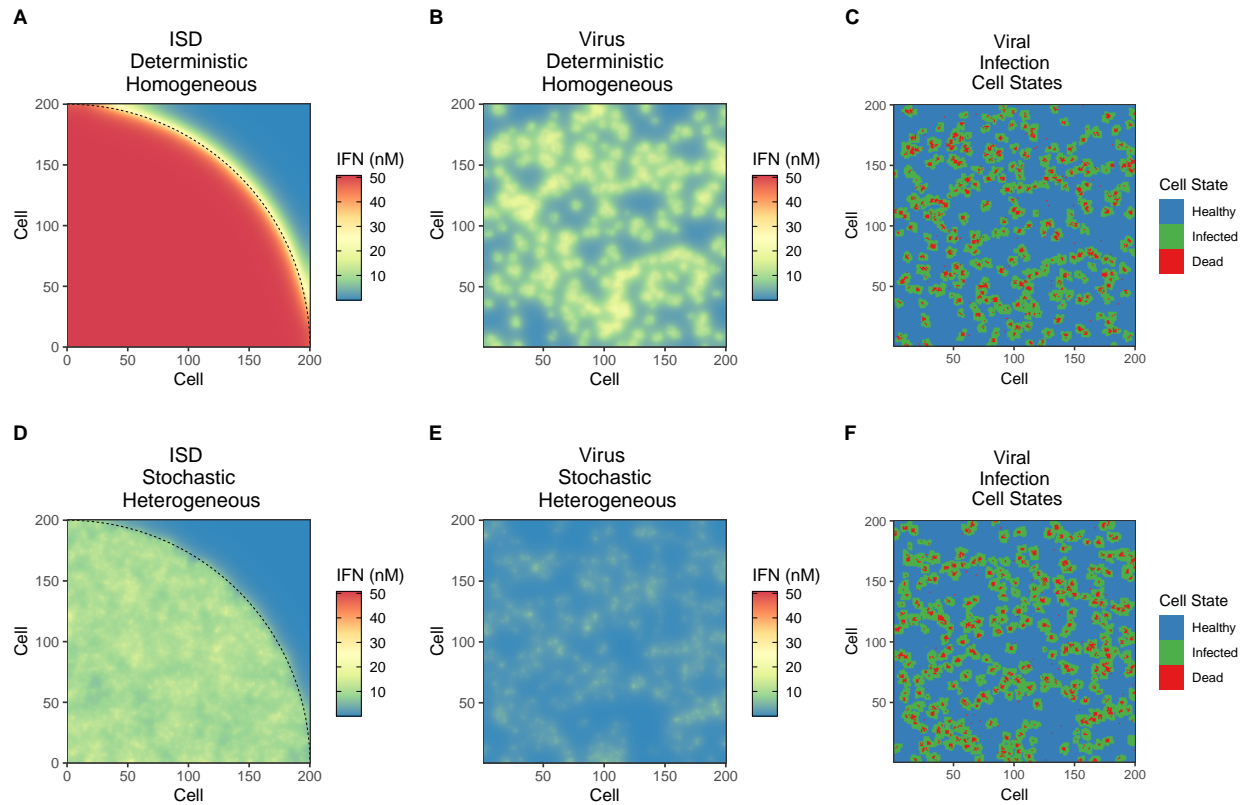
Figure 4.3.2: Simulating IFNβ dynamics across different cell populations. The distribution of IFNβ at 10 hours after the addition of cytosolic DNA. Cell populations were either modeled as deterministic and homogeneous **(A-C)** or stochastic and heterogeneous **(D-F)**. Panels **(A)** and **(D)** show cell populations transfected with immunostimulatory DNA (ISD). The quarter circular region (dotted black line) divides the cells into two populations. Cells within the boundary were transfected with ISD and referred to as primary cells. Cells outside this region were not transfected and referred to as secondary cells. In panels **(B)** and **(D)**, a virus infection is simulated by assigning nonzero initial conditions of DNA according to a Poisson distribution with MOI=0.01 4.2. Panels **(C)** and **(F)** show the same viral simulation but identify cell state (healthy, infected, or dead).

4.3.2 Dynamics of intracellular molecules are consistent with experimental evidence

We next tested the ABM to see if the model could replicate bulk intracellular signaling from cells challenged with ISD and virus (Figure 4.3.3). During ISD transfection (Figure 4.3.3A-D and Figure 4.3.3I-L) we observe a strong IFNβ response in primary cells to the DNA stimulus, followed by the activation of the JAK/STAT pathway to upregulate feedback mechanisms like IRF7 and TREX1. These dynamics are a result from fitting previous experimental data used in the ODE model (Table 1). Secondary cells also respond to IFNβ paracrine signaling, but to a lower degree. Viral infection simulations (Figure 4.3.3E-H and Figure 4.3.3M-P) show an enhanced response from secondary cells, particularly from feedback mechanisms like IRF7 (Figure 4.3.3G and 4.3.3O) and TREX1 (Figure 4.3.3H and 4.3.3P). The variance in these responses is also increased—especially in stochastic/heterogeneous cell populations. This variance is caused by some secondary cells having no response to infection whereas others respond more strongly than even primary cells. Overall, this suggests that primary cells are more important to ISD transfection and that secondary cells play a more central role in regulating viral infections.
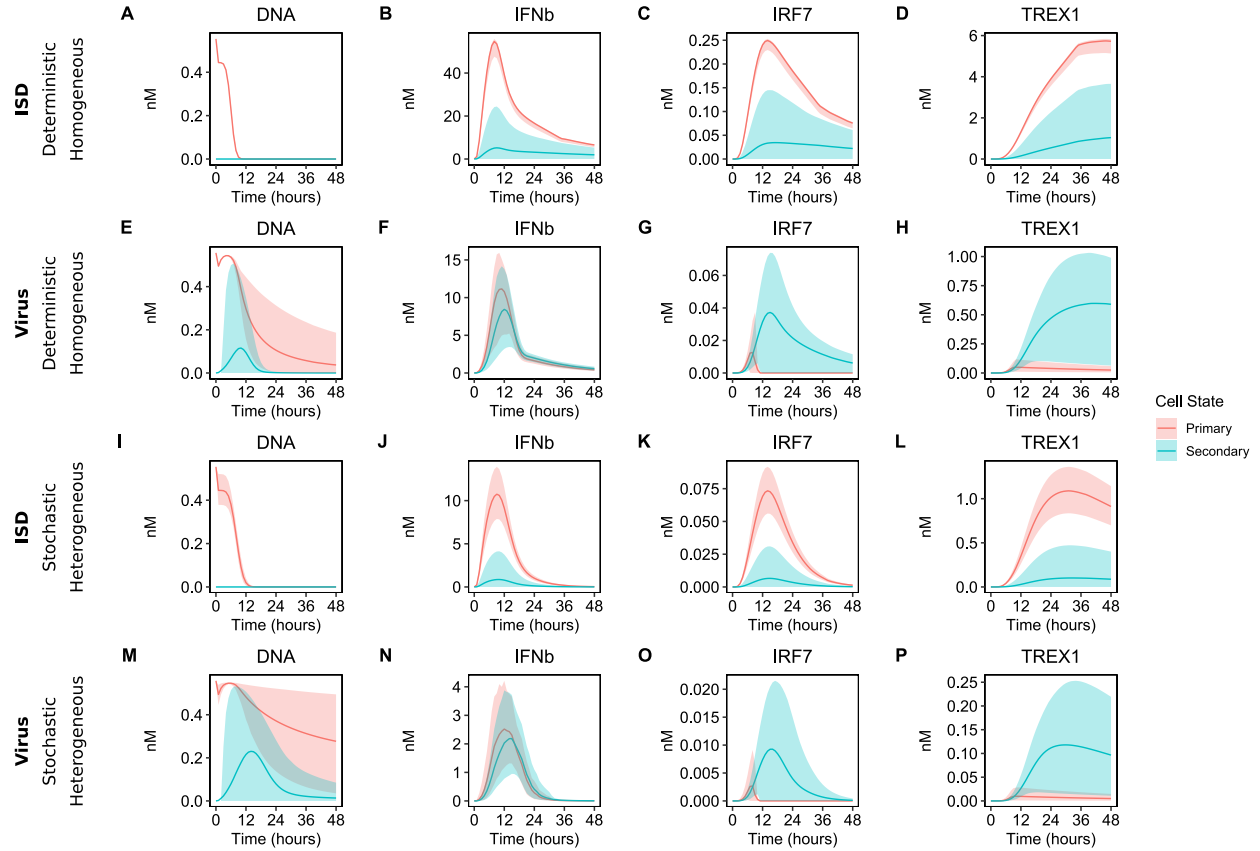
Figure 4.3.3: Simulating intracellular dynamics across different cell populations. The intracellular signaling dynamics in response to ISD transfection **(A-H)** and viral infection **(I-L)**. ISD transfection was further divided into different cell populations including deterministic/homogeneous **(A-D)** and stochastic/heterogeneous **(E-H)**. Average intracellular concentrations are plotted as a function of time and are separated into cells that are transfected/infected (primary) and those that are not (secondary cells). The bold line represents the mean trajectory within the cell population with error bars showing the 5th and 95th percentiles.

To further validate the ABM, we examined the impact heterogeneity and stochasticity had on knockdown simulations. These simulations were compared to simulating just the ODE model under the same knockout conditions. In principle, both models should follow the same trends when the ABM dynamics are averaged over the cell population. Figure 4.3.4A-C show how both the ODE and agent-based models' response to knockdown of the TREX1 negative feedback loop. TREX1 is responsible for cytosolic DNA degradation, and the model shows a robust IFNβ response to its depletion. When TREX1 is completely knocked-out, the ABM and ODE models all exhibit a sustained inflammatory state due to continuous IFNβ signaling.
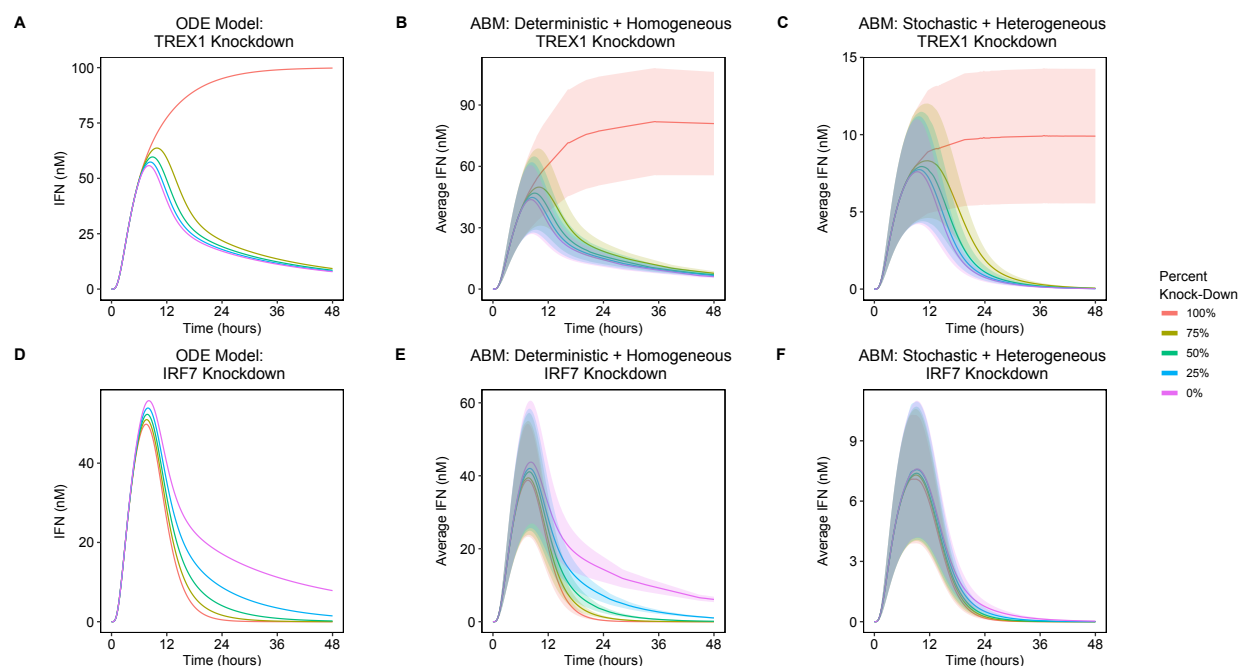
Figure 4.3.4: Comparing ODE and ABM knockdowns. Simulating knock down of feedback mechanisms in three different model types: ODE, ABM with a deterministic/homogeneous cell population, and ABM with a stochastic/heterogeneous cell population. **(A-C)** depict the three simulations were the TREX1 protein was knocked down by reducing the translation rate parameter k_{cat2}. The percentage indicates how reduced the parameter of the respective feedback mechanism was with 0% representing no attenuation and 100% representing total knockout. Average cell trajectories are plotted with the shaded regions marking one standard deviation away from the mean. **(D-F)** depict the same three simulations, but with IRF7 knocked down by reducing the k_{12f} parameter.

IRF7 knockdowns (Figure 4.3.4D-F) show an increase in IFNβ clearance for both models, which is an expected outcome as this protein acts as a transcription factor for IFNβ. IRF3 is the initial transcription factor for IFNβ and is thought to be replaced by IRF7 through a positive feedback loop. With no IRF7 further driving IFNβ transcription, protein levels quickly drop off. Of note, the ABM simulation with a heterogeneous and stochastic cell population still followed this trend, but the separation between the levels of knockdown were smaller because the peak IFNβ concentration was lower. In sum, the ABM captures population-level dynamics observed in experiments, and, unlike the previous ODE model, supports evaluating how cell heterogeneity and stochasticity impact IFNβ signaling.

4.3.3 Cellular heterogeneity can Modulate and Enhance Interferon Production

The concentrations of mRNA and signaling proteins at the initial time of infection will possess some level of variability across the cell population and are suspected to have an impact on immune signaling dynamics [150]. Figure 4.3.5 shows the distribution of peak IFNβ concentrations across the cell population that result when simulating an ISD transfection in a deterministic/heterogeneous cell population (Figure 4.3.5A) or a stochastic/heterogeneous cell population (Figure 4.3.5B).

To distinguish paracrine-induced IFNβ production from production in transfected cells, the cells were divided into two groups: primary and secondary. Primary cells were categorized as being transfected at time zero hours (i.e. had an initial nonzero DNA concentration) whereas secondary cells were not transfected and only produced IFNβ in response to paracrine signaling. As increasing Gaussian noise is added to the initial conditions, we observe a small decrease in peak IFNβ produced (Figure 4.3.5A and 4.3.5B, $\sigma^2 = 0$ to $\sigma^2 = 0.3$) followed by a large increase in IFNβ produced (Figure 4.3.5A and 4.3.5B, $\sigma^2 = 0.4$ to $\sigma^2 = 1.0$) as cellular heterogeneity increases. This trend is more pronounced in the deterministic simulation compared to the stochastic simulation. Secondary cells produce less interferon than their primary counterparts across all levels of variability and model types, which indicates that DNA is a more potent stimulus for interferon production compared to STAT activation in this model. Overall, these simulations suggest that increasing cellular heterogeneity leads to

aberrant levels of IFNβ production. Excessive variability will lead to high levels of interferon, but moderate levels can lower IFNβ production independently of stochasticity.
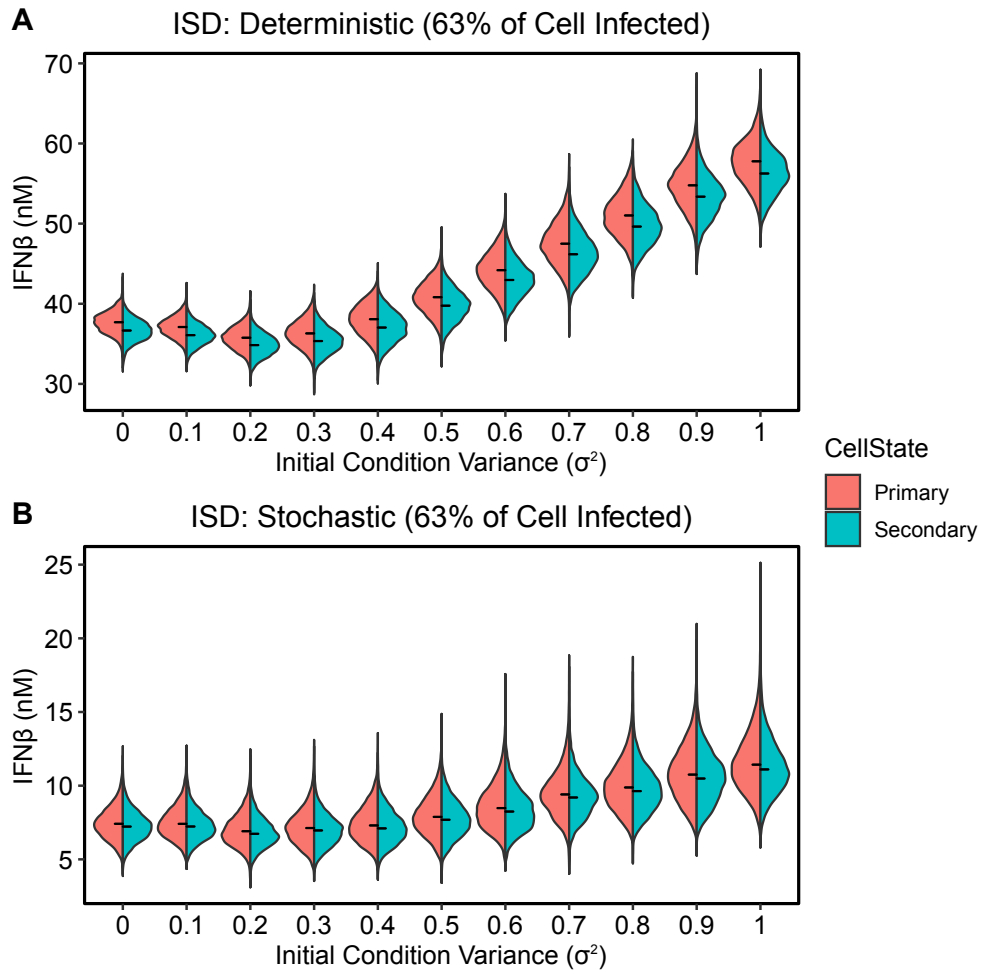
Figure 4.3.5: Effect of heterogeneity on population IFNβ levels. Initial non-zero protein concentrations were sampled from increasingly wider (more variable) normal distributions, as shown on the horizontal axis. With this new initialization, the models were simulated with an ISD infection starting with 63% primary cells. The resulting distribution of peak IFNβ concentrations were taken and partitioned into either primary or secondary cells. The cell population was simulated with **(A)** deterministic IFNβ producing cells and **(B)** stochastic IFNβ producing cells.

4.3.4 Aberrant STAT activation and IFNβ Degradation did not Lead to Chronic Inflammation within Stochastic Cell Populations

The IFNβ response from cGAS activation is typically cleared through negative feedback regulated by SOCS1/SOCS2 proteins inhibiting STAT activation [185, 186] or by TREX1 degrading cytosolic DNA [110]. However, aberrant activation of the interferon response can lead to situations where a cell population can no longer clear the protein and the immune response persists. This chronic inflammatory state can lead to further immune dysfunction and induce potent cytokines including programmed death-ligand 1 (PD-L1), IL-10, and d-indoleamine 2,3 deoxygenase (IDO) [187]. To investigate how heterogeneity and stochasticity impact sustained inflammatory signaling, we ran several simulations in which parameters important to suppressing IFNβ levels are modified across different types of cell populations (homogeneous/deterministic cells, heterogeneous/ deterministic, and homogeneous/stochastic). We selected parameters that describe IFNβ degradation (τ_7) and STAT activation from paracrine signaling (k_{cat8}) because previous work found IFNβ to be sensitive to these parameters [188]. Figure 4.3.6 summarizes these simulations by displaying average IFNβ concentrations simulated one week after the initial infection (this allows the model to reach a steady state). Figure 4.3.6A shows how a deterministic, homogeneous cell population transitions between successfully suppressing IFNβ to having sustained IFNβ levels as the degradation rate decreases and the paracrine activation rate increases. Figure 4.3.6B introduces heterogeneity into the cell population, which increases the parameter ranges that result in non-zero IFNβ steady states, i.e. increases the parameter space in which chronic inflammation is likely. ABM simulations with stochastic cell populations (Figure 4.3.6C) did not result in chronic inflammatory states across the tested parameter ranges. The largest concentrations for IFNβ at steady state never exceeded 0.001 nM (effectively no protein). These simulations suggest that stochasticity is an important signaling feature for limiting chronic inflammation and that features beyond normal systems-level considerations, e.g. negative feedback, should be explored as key regulators of inflammation.
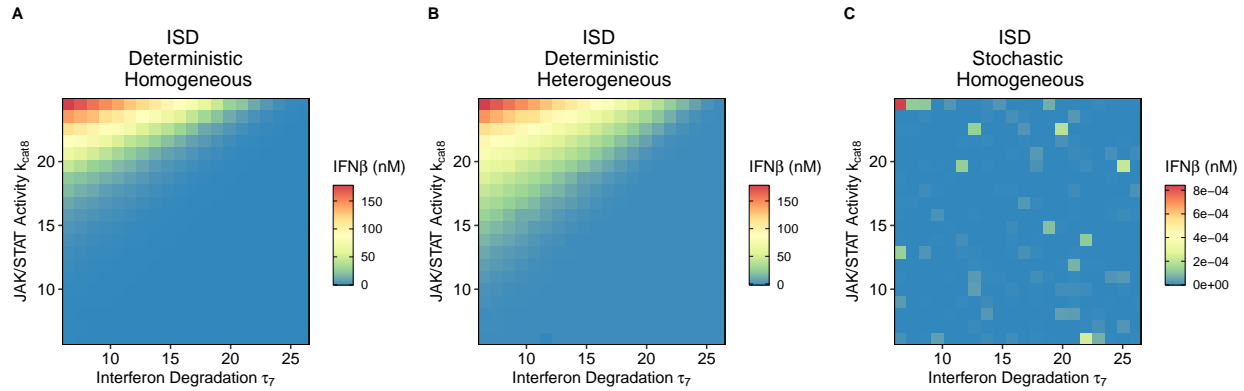
Figure 4.3.6: Driving simulations toward chronic inflammatory states. Each box represents an individual simulation and is colored by the mean concentration of IFNβ at the last simulated time point (168 hours after initial transfection). Two parameters were varied: the degradation rate constant for IFNβ protein (τ_7) and the rate of STAT protein activation by IFNβ (k_{cat8}). Each parameter was varied between half and double its original optimized value. **(A)** Shows simulations with deterministic/homogeneous cell population. **(B)** Simulations with a deterministic/heterogenous cell population. **(C)** Simulations with a stochastic/homogeneous cell population.

4.3.5 There is an optimal level of intrinsic stochasticity to Maximize Cell Survival

Simulating how a cell population would respond to a pathogenic threat provides key insight into the interplay between stochasticity and the success of the innate immune response. We next performed simulations to assess how stochasticity and cell heterogeneity impact immune success by tracking the number of healthy, infected, or dead cells remaining at 48 hours post virus infection (Figure 4.3.7). Figure 4.3.7A and 4.3.7C show the median trajectories and the interquartile ranges of the cell states in a homogeneous/stochastic cell population or heterogeneous/stochastic cell population, respectively. In both cases, the percentage of IFNβ producing capable cells is left at the default value of 20%. Both the homogeneous and heterogeneous cell populations have the largest percentage of cells infected within the first 12 hours simulation and all infected cells transition to dead based on the how long they have been infected (see Section 4.2).
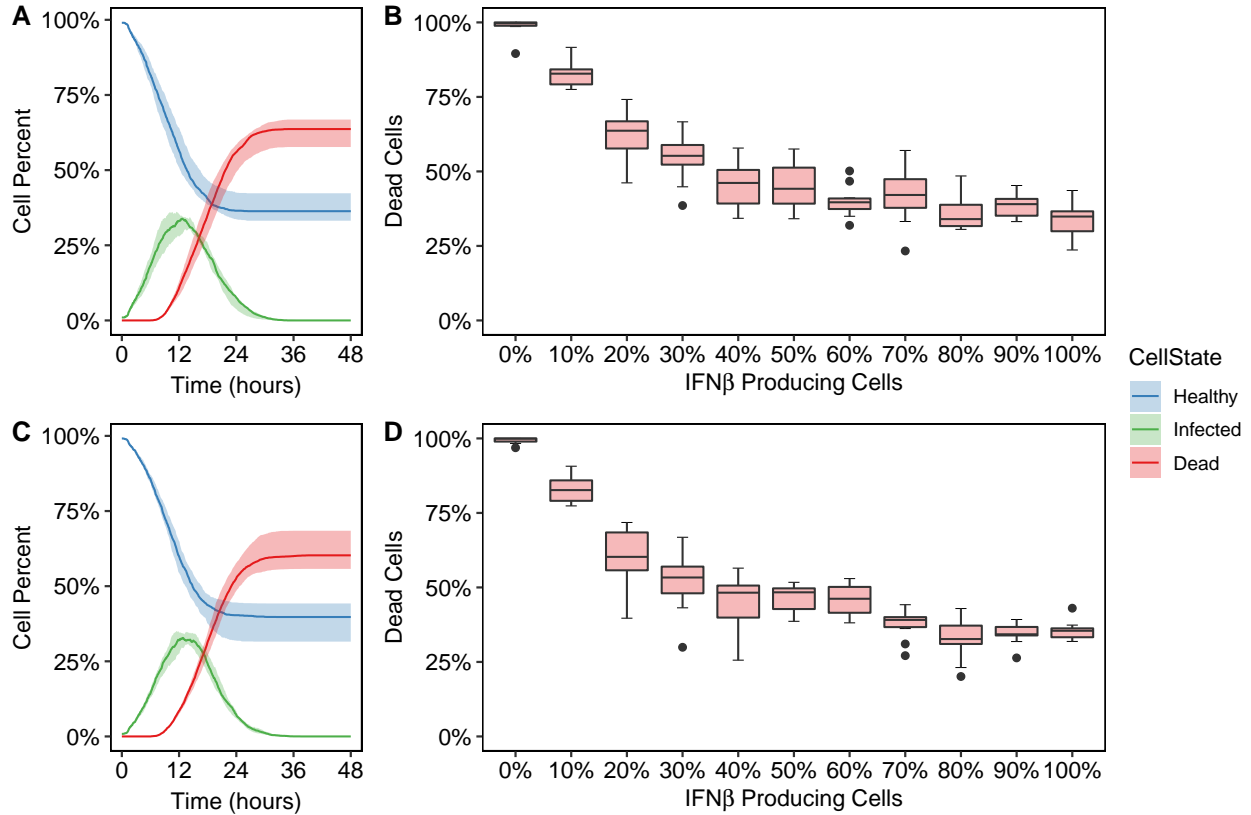
Figure 4.3.7: Stochasticity optimizes cell survival. Stochastic interferon expression of viral infection in homogeneous (A-B) and heterogeneous (C-D) cell populations. **(A)** Cells states (Healthy, Infected, and Dead) were tracked over time in ten simulations with 20% of cells producing IFNβ to obtain median trajectories. The shaded region shows the interquartile range across the simulation. **(B)** Each boxplot summaries the number of accumulated dead cells for ten viral infection simulations (at t=48 hours). The percentage of cells producing IFNβ was varied from 0% (no cells with an IFNβ response) to 100% where all cells mounted a similar IFNβ response. **(C)** The process was repeated with a heterogeneous cell population, tracing cells states with 20% of cells producing IFNβ. **(D)** The percentage of IFNβ producing cells was then again varied and the change in the percentage dead cells was observed.

A successful and optimal host immune response would minimize the number of cell deaths and the amount of IFNβ needed to quench cytosolic DNA, as high levels would likely induce long-term inflammation. With this measure of success defined for the immune response, we evaluated how the cell population responded to changing the number of cells capable of producing an interferon response. This was accomplished by randomly assigning the k_{cat7} parameter (Eq. 4.1) a value of zero based on the desired percentage of IFNβ producing cells (e.g. no IFNβ producing cells would imply that all k_{cat7} parameters were set to zero). Figure 4.3.7B shows how the viral infection simulation was affected by the percentage of IFNβ producing cells in a homogeneous population. When fewer cells are capable of responding to the virus, more cells die. However, there is a wide range (40-100%) where the interferon response has been modulated but the cell survival rate is not significantly affected. A range of 20-30% appears to be the optimal percentage of IFNβ producing cells because it both minimizes the number of cells lost to the infection and the energy needed to bolster an effective immune response. Figure 4.3.7D repeats these simulations, but with a heterogeneous cell population. There were no apparent differences found between the heterogeneous and homogeneous cell populations, indicating the stochastic responses were primarily responsible for the observed trends in cell survival.

4.4 Summary

The ABM method presented in this work was specifically tailored to assess how cellular heterogeneity and stochasticity impact immune signaling. Computational modeling and systems approaches have been used to provide systems-level insight into the regulation of immune signaling systems, however, how cellular heterogeneity and stochastic influence robust, optimal and reliable signaling is an under considered area. Using a customized ABM that is predicated on an experimentally validated ODE model, we performed *in-silico* experiments to determine how cellular heterogeneity and stochasticity impact overall IFNβ production and cell death experienced during the infection. Figure 4.3.5 exemplifies how cell heterogeneity acts as a tuning variable to regulate intracellular immunity. Low heterogeneity

(i.e. small variations in cellular composition) allows a cell population to attain similar IFNβ concentrations without having every cell fully respond. Lowly expressing cells can rely on paracrine IFNβ signaling and not expend the resources necessary to maintain a cellular composition capable of producing a comparable response. Highly heterogeneous cell populations magnify this behavior of having only a few cells produce an excessive amount of IFNβ that overcompensate for the population. By tuning cellular heterogeneity, a cell population can modulate IFNβ response potency without the need for additional cell signaling architecture.

Stochasticity, in comparison, plays an important role in maintaining cell survival. Experimental evidence shows that the IFNβ response is stochastic in nature and results in at most 20% of cells responding to infection [37]. The agent-based model simulates the end result of this stochastic behavior by varying the number of IFNβ producing cells (Figure 4.3.7). We observed that having too few IFNβ producing cells lead to excessive cell death but increasing the fraction of responding cells did not result in improving overall cell survival. We found having 20-30% of cells producing interferon was an ideal balance between these two extremes. The modeled ideal percentage of responding cells does not exactly match Zhao et. al. in Figure 4.3.7, but other experimental studies suggest that this ideal response is virus specific [188]. The ABM virus parameters could easily be modified to fit any viral specific IFNβ response, but more importantly our model provides an explanation for these experimental observations.

Past research in our group explored the incorporation of differential equations into agent-based models using existing software. The user-friendly software NetLogo [52] was first chosen to create an agent-based model of the cGAS pathway with the aim of elucidating the effects of cell-to-cell interactions through paracrine signaling. A reoccurring issue that spanned multiple softwares was the difficulty in incorporating ODE models into the ABM. Following the demo kinetics model in NetLogo, the package allowed for simple kinetics to be added but proved to be challenging for complex differential equations. Additionally, we did not find any established packages to link NetLogo with more general scientific computing languages like Matlab. Coding languages like Matlab were also tested as a platform for ABM simulation but the ODE solvers (e.g. ode15s) were unsuccessful at solving the large systems we desired. For this project, Julia allowed for the incorporation of ordinary differential

equations into an ABM, as well as more control over the ODE solvers needed to simulate the large, stiff system of equations. [189].

Critical tradeoffs were made when developing this ABM, most notably was the number and complexity of the agents. The model assumes cells are static entities on a uniform grid, which works when modeling epithelial cells. Modeling other cell types, conversely, may require agents to possess attributes like shape, movement, and cellular division. These can be more readily implemented in other ABM software like Compucell3D and Morpheus at the cost of reducing agent count [190, 191].

Overall, this paper develops a straightforward method to translate a signaling pathway ODE model into a hybridized agent-based model capable of handling large scale simulations. Common signaling mechanisms like autocrine and paracrine signaling can easily be incorporated into the model as well as additional heterogeneous and stochastic properties that cannot be modelled with ODEs directly. In this example of the cGAS pathway, we were able to take experimental observations like heterogeneity cellular composition and stochasticity in IFNβ expression and show why they would be beneficial to a cell population. With the advent of measuring single cell dynamics in an accurate and precise manner, this modeling framework has the potential to allow researchers to rapidly restructure other ODE models to help explain new observations made at the single cell level.

5.0 Conclusions

The goal of this dissertation was to (1) use systems biology methodologies—like ODE modeling—to examine the regulation of the innate immune system and (2) extend those methods to simulate heterogeneous cell populations across multiple length scales. Two components of the innate immune systems were investigated: metabolic reprogramming during sepsis and interferon beta expression induced by cGAS DNA detection. Both of these immune signaling events are relevant to a number of diseases/conditions (e.g. acute kidney injury [175, 68], viral infection [108, 102], cancer [112, 113, 28]), so building dynamic models to describe their progression is key to developing new pharmaceuticals and improving patient outcome.

5.1 Metabolic Reprogramming

With metabolic reprogramming during sepsis (Aim 1), there was preliminary evidence that tubular epithelial cells (TECs) in the kidney transitioned between normal metabolism (oxidative phosphorylation) in the mitochondria and aerobic glycolysis [11]. This metabolic shift is hypothesized to occur because cells are entering either a resistive or tolerant state to a perceived pathogenic threat. Aerobic glycolysis is triggered early on during infection driving TECs into a more resistive state. If cells remain in this state for a prolonged period of time or fail to transition back to more tolerant oxidative phosphorylation, organ damage like fibrosis can occur, resulting in chronic kidney disease (CKD).

The goal of this first aim was to develop a model of the pathways regulating metabolism in TECs. This involved crosstalk between the AMPK, TLR4/Akt/mTOR, and HIF-1 pathways which controlled oxidative phosphorylation, LPS detection, and glycolytic pathways respectively. With the implementation of such a model, metabolic shifting between OXPHOS and aerobic glycolysis could be simulated. The benefit of this model development is the ability determine what components of the signaling pathways are most sensitive to

promoting metabolic reprogramming. Global sensitivity analysis (GSA) identified AMPK phosphorylation to be highly insensitive to changes in metabolic state (shifting to glycolysis). In other words, small variations in AMP/ATP ratios caused by metabolic shift were compounded with insensitive parameters leading to minimal downregulation in OXPHOS.

This model for simulating metabolic reprogramming in sepsis had some notable limitations. One is that the model focused more on building an understanding of the regulators controlling the two metabolic pathways and was less concerned with simulating dynamics within those metabolic pathways. Effectively, OXHPOS and Glycolysis are lumped into individual terms. This was done because limited kinetic data was available to model the complexities of say the entire electron transport chain. This more minimalist approach is similar to more extreme example of not modeling every quantum mechanical interaction of a box sliding down a ramp. While both of these modeling schemes would be more realistic, they can lead to simulations that are intractable even on modern hardware and difficult to draw conclusions from. The modeler needs to strike a balance between biological detail and model interpretability. That being said, one future direction to improve this model would be the addition of kinetically informed regulations like SIRT1, which is controlled by AMPK and regulates biogenesis of mitochondria through PGC1-α.

A more ambitious direction for this project would be to expand the ODE simulation into a multiscale ABM simulation. This simulation could vary the number of tolerant and resistant cells to determine its impact on acute kidney injury. Additionally, intercellular signaling between TECs and immune cell signaling could be incorporated to develop a more holistic view of the kidney nephron during sepsis.

5.2   Computation Model for the cGAS Pathway

The second aim was focused on model development for the recently discovered cytosolic DNA sensor cGAS. RNA sensors like the RIG-I had been well established at the time [19], but it was not until 2013 that cGAS was identified as the putative DNA sensor [24]. The detection of aberrant DNA in the cytosol is a precursor for several dangerous events including

pathogenic invasion, cancer [29], autoimmune disease [110, 30], and senescence [192, 32]. By developing a computational model of this DNA detection mechanism, we were able to investigate hypotheses explaining dynamic regulation and show emergent system properties *in silico.*

The cGAS model consisted of the titular enzyme, activation of the interferon beta response, and downstream regulators expressed through the JAK/STAT pathway. This included negative feedback loops such as TREX1 degrading cytosolic DNA and positive feedback loops like IRF7 taking over as the primary interferon transcription factor. The model was fit to experimental data using a Bayesian approach (Markov Chain Monte Carlo) which provides significant advantages over traditional parameter fitting. Chief among these are the ability to escape local minimum solutions that gradient descent algorithms fall victim to. If the MCMC algorithm successfully samples the parameter space, distributions for parameter values are obtained which provide additional information about parameter sensitivity. A broad, flat distributions, for example, indicates that parameter can vary greatly with without impacting the overall model fit.

One of the main results of the cGAS ODE model was the prediction of robustness in TREX1 signaling. TREX1 is exonuclease upregulated by a type I interferon response whose function is to degrade double stranded DNA that accumulates in the cytosol. The model in Aim 2 showed that knocking down TREX1 activity had a limited effect on IFNβ signaling and that only complete knockout of TREX1 activity lead to a sustained IFNβ producing state (Figure 3.3.5a). This prediction still needs to be experimentally validated, but looking at previously published studies seems to corroborate this claim. Zhang *et.al.*, for example, was investigating the role of cGAS dependent IFNβ expression in Chlamydia trachomatis infection [193]. In this study, both TREX1 knockouts and siRNA knockdowns were performed to determine the resulting IFNβ mRNA fold change. When polydA:dT (a synthetic DNA immunostimulant) was used as a transfecting agent, the difference in IFNβ mRNA expression for the knockdown was small (control: ∼7 fold, TREX1 KD: ∼11 fold) whereas the difference from the total knockout was large (control: ∼10,000 fold, TREX1 KO: ∼88,000 fold). This does suggest some signaling robustness, but further experimentation would have to be done. These knockout and knockdown measurements were only performed at one time

point and there are no measurements to determine if this IFNβ response if sustained in the cell population. This could be accomplished, for example, by using an shRNA knockdown cell line coupled with a fluorescent reporter for IFNβ mRNA.

5.3   Multi-Scale Agent-Based Modeling

Aim 3 was driven by two key observations exhibited by the interferon response in cells. The first was the apparent failure of some cell in a population to respond to an immunostimulant (whether that be a virus or DNA fragment). The percentage varies depending on context, but up to 85% of stimulated cells will mount no immune response [37, 188]. The model discussed in Aim 2 cannot predict this stochastic observation because ODEs are completely deterministic given an initial condition. If there was cGAS stimulus, the ODE model would always predict an accompanying IFNβ upregulation because this portion of cGAS signaling pathway contains no negative feedback regulation.

The second observation was that the magnitude of interferon produced by cells varied greatly despite be given identical treatments [145, 148]. This was further confirmed when Rand *et. al.* examined "sister" cells (cells that had just divided) and showed that the correlation between pairs of IFNβ expressing cells was very weak ($r^2 = 0.11$) [145]. The ODE model in Aim 2 could not predict this phenomenon in a single execution of the simulation. To model this observation, the ODE model would have to be simulated multiple times while varying the initial conditions (each simulation would then represent a different cell in the population). This works to an extent, but assumes that (1) heterogeneity is caused by differences in intracellular protein concentrations and (2) negligible intercellular signaling.

The main goal of Aim 3 was to build a model that could emulate these observations and explain what benefit—if any—they brought to a cell population. This question is important because it impacts how we modify immune response. If we attempt and stimulate the interferon response by artificially increasing the number of expressing cells, essential regulatory signals that control IFNβ could be disrupted. Excessive tissue damage and organ failure can result from an overactive immune response, as demonstrated in failed IFNβ treatments [194].

To explain these two observations, a new modeling scheme needed to be developed. The ODE cGAS model was already established so it was sensible to try and extend this model as opposed to implementing something entirely new. Extending the model was also advantageous because the process could be generalized to extend other ODE based models like the sepsis model in Aim 1. The resulting model that incorporated both heterogeneity and stochasticity can be best described as a hybrid ODE/PDE/ABM model which benefits from each of the individual modeling types. The ODE system was used to track intracellular concentrations as before, PDEs were used to simulate diffusion across the cell population, and the ABM allowed these previously deterministic models to react to probabilistic events.

The main finding of this aim was the benefit stochasticity provided to a cell population. Simulated cell populations with a reduced number of IFNβ expressing cells experienced a similar number of cell deaths when compared to fully expressing cell populations (Figure 4.3.7). This benefits the cell population in a number of ways. First, mounting an immune response is energy intensive, so only needing some responding cells reduces this burden. Second, the immune response is more challenging to predict, making is more difficult for a pathogenic threat to adapt. And last, by minimizing the interferon producing cells, the risk of overstimulating the immune response and causing tissue damage is reduced.

A future direction for this type of hybrid agent-based modeling scheme would be to replace ODEs with stochastic differential equations (SDEs). With ordinary differential equations, the parameters are usually considered constant. The entire objective of parameter fitting is to determine these constants to make inferences about the ODE model. SDEs break this assumption and allow parameters to be defined as random variables. The result is a system that follows the governing equations used to build the ODE system but can randomly fluctuate away the determined trajectory. Running an SDE will give a different trajectory each time, but the average trajectory will converge to the ODE solution (given small noise limits). This consideration for signal noise can have a substantial effect on a cell population, depending on the extent of noise introduced. Some cells could produce massive IFNβ responses while others could very well produce no detectable signal. Steady states could bifurcate leading to subpopulations of cells responding differently to the same signal. This would be difficult to model using the current hybrid approach as callback rules

for changing parameter values would need to be added which would substantially reduce computational efficiency.

The switch between to SDEs would be straightforward using *DifferentialEquations.jl* because it provides a large library of solvers tailored to different classes of SDEs. One would simply have to define a noise matrix and choose an appropriate solver. Numerical algorithms to solve for SDEs is still a developing field, which means the number of cells simulated in Aim 3 would likely need to be reduced as many of the available ODE solvers were unsuccessful at solving this system. Additionally, the system would have to be simulated multiple times to get average trajectories which would be feasible, but computationally expensive.
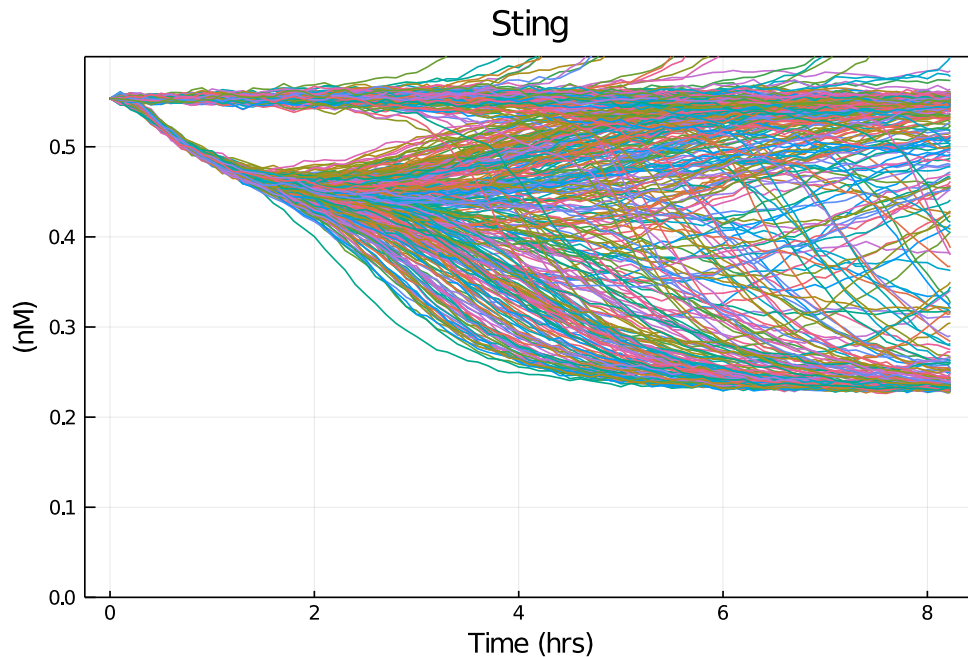
Figure 5.3.1: Preliminary SDE Simulation. Each of the 400 time series represents a cell Sting protein concentration dynamically responding to an ISD stimulation. A small, constant noise term was added to each species in the ABM model and simulated using the SOSRI solver. Some cells are initially transfected with DNA, decreasing their inactive Sting levels while other cells randomly fluctuate at steady due to added noise. Transfected cells can be observed bifurcating, leading to two steady states for Sting.

Overall, these studies were successful in revealing insights into the innate immune system that were difficult to ascertain with traditional biological methods. It puts emphasis on the fact that the innate immunity is a dynamic, complex, and nonlinear system that demands

advanced computational techniques to even begin to understand its intricacies. Intensive collaboration between immunologist and engineers needs to be brought to the forefront of this field because neither discipline can fully decode cell signaling regulation individually. Progress is being made toward the collection of genome wide, live, single cell measurements which will provide invaluable information for computational modeling; but until such data become ubiquitous, systems biology modeling will be a highly iterative, trial-and-error process. The models presented in this work are no exception. New biological discoveries are constantly changing our understanding of this system. For example, there are been recent evidence that cGAS is predominately localized in nucleus and not constitutively activate in the cytosol which would influence how the cGAS pathway is modeled. This work is a snapshot of our current understanding of how the innate immune system detects dangerous situation, what response a cell population takes (changing metabolism, producing cytokines), and how that response is regulated. Ultimately, these insights into immune system dynamics will help explain why certain treatments fail and pave the way for new, more effective therapies.

5.4   Publications Resulting from this Dissertation

- Gregg, Robert W., Saumendra N. Sarkar, and Jason E. Shoemaker. "Mathematical Modeling of the cGAS Pathway Reveals Robustness of DNA Sensing to TREX1 Feedback." Journal of theoretical biology (2018). https://doi.org/10.1016/J.JTBI.2018.11.001

- Gregg, Robert W., Saumendra Sarkar, and Jason E. Shoemaker. "Examining Dynamic Emergent Properties of the DNA Sensing Pathway." IFAC-PapersOnLine 51.19 (2018). https://doi.org/10.1016/J.IFACOL.2018.09.017

- Gregg, Robert W, Shabnam, Fathima, and Jason E. Shoemaker. "Agent-Based Modeling Reveals Benefits of Heterogeneous and Stochastic Cell Populations during cGAS-Mediated IFNβ Production." Bioinformatics. (In review)

- Gregg, Robert W, Shabnam, Fathima, and Jason E. Shoemaker. "Modeling Metabolic Reprogramming during Sepsis" (To be submitted)

6.0   Collaborative Projects

I have collaborated on a number of projects that are peripherally related to this main body of work, but do not fall into the main theme of using systems biology to understand pathogen associated molecular patterns. Here, I will summarize the purpose of these projects and highlight any important results that were determined.

6.1   Vir-ed: An Educational Virtual Reality Application for Teaching Virology

For this project, I led and managed a team of software engineers at Full Sail University in Florida to develop an educational virtual reality (VR) app teaching users about this dissertation research related to cGAS. Because this software utilized VR, users were placed in a 3D virtual cell environment they could explore and use to discover biological concepts related to virology.

The are two campaigns the user can play through. The first follows the process of a virus invading a cell. Users are guided through parts the viral life cycle like surface attachment and replication using mini-games that test for biological knowledge and skill. The second flips this perspective and discusses how the cGAS pathway is used as a cellular defense mechanism against viral infection. The mini-games in this section focus on components of the cGAS pathway, like how cGAMP is formed from ATP and GTP in the enzymatic pocket of cGAS when DNA is bound. Performing well in these mini-games unlocks trophies of the different biological components in the game (e.g. the Sting protein, DNA, etc.) which come with detailed descriptions of their function. After completing both campaigns, users unlock an additional survival-based game were they are tasked with maintaining a cell population and defending them against invading viruses. They have to balance cellular defense, immunity, and production using their knowledge of the given anti-viral interferon stimulated proteins.

To promote this project, it was presented at the STEAM Showcase hosted by Schell Games (the largest full-service education and entertainment game development company in

the United States), and play-tested with students at Plum Senior High School to gather feedback and gauge interest. This project even garnered some publicity when we were interviewed with the Pittsburgh Post-Gazette.

The free app is available on the Google play store and iOS (can also search for "Vir-ed").
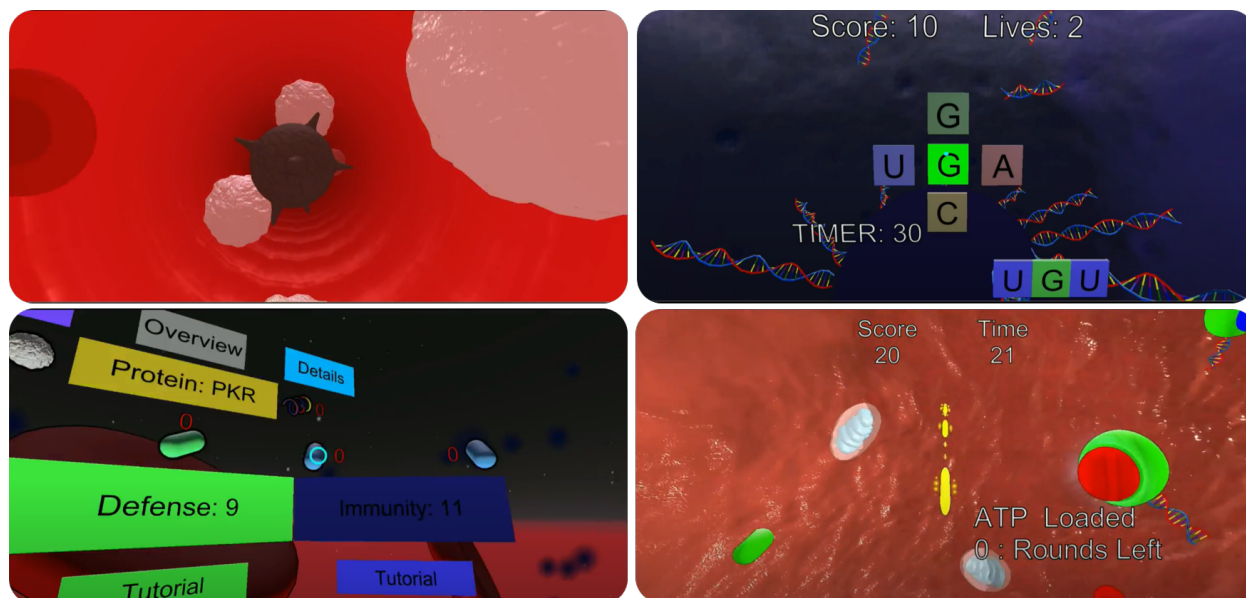
Figure 6.1.1: Screen captures of Vir-ed, the virtual reality app built to educate users about introductory biology, virology, and the cGAS pathway. Pictured above is a virion traveling through the circulatory system, a memory mini-game to explain nucleotide bases, a survival game to defend cells against viral invasion, and an aiming mini-game to place ATP and GTP into DNA bound cGAS.

6.2   Pathogenesis of Influenza A(H7N9) Virus in Aged Nonhuman Primates

In this collaborative project I worked with Yoshihiro Kawaoka's Lab in the Institute of Medical Science at University of Tokyo to understand the impact of aging on influenza A infection. Viral infections, especially influenza A, are observed to preferentially cause higher mortality rates in the elderly, but the exact mechanism for this process is ill-defined. To improve our understanding of this unknown mechanism, Yoshihiro's group infected a non-human primate model with influenza A (H7N9 strain) and measured the pathology, immunology, and genetic variability over time.

My contribution to this project was to perform a gene functional enrichment analysis using the collected micro-array data. This bioinformatic approach takes gene expression data and determines a set of differentially expressed genes (up or down regulated) that contribute to differences between groups. By determining these unique gene signatures, we can use gene ontology to identify matching biological functions that are enriched for this differential expressed genes associated with influenza and aging.

There were two key findings to this study, the first was the identification of IL-6 being down-regulated in aging animals, which likely contributed to the their high pathogenicity of influenza A(H7N9). The second finding showed that while type-1 interferon activated genes were equally expressed in both age groups, protein levels were significantly lower in aged animals. This dissonance between gene expression and protein production was attributed to the association between aging and the down-regulation of protein synthesis found in previous studies [195]. For more information related to this project, one can refer to our recently accepted publication [196].

Appendix A cGAS Model

A.1 Ordinary Differential Equations

$$\frac{d\left[cGAS\right]}{dt} = -k_{1f} \cdot cGAS \cdot DNA + k_{1r} \cdot cGASc \tag{A.1}$$

$$\frac{d\left[DNA\right]}{dt} = -k_{1f} \cdot cGAS \cdot DNA + k_{1r} \cdot cGASc - \frac{k_{cat2} \cdot TREX1 \cdot DNA}{K_{m2} + DNA} \tag{A.2}$$

$$\frac{d\left[cGAMP\right]}{dt} = k_{3f} \cdot (cGAS_{tot} - cGAS) - k_{4f} \cdot cGAMP \cdot STING$$
$$+ k_{4r} \cdot (STING_{tot} - STING) - \tau_3 \cdot cGAMP \tag{A.3}$$

$$\frac{d\left[STING\right]}{dt} = -k_{4f} \cdot cGAMP \cdot STING + k_{4r} \cdot (STING_{tot} - STING) \tag{A.4}$$

$$\frac{d\left[IRF3\right]}{dt} = -\frac{k_{cat5} \cdot IRF3 \cdot (STING_{tot} - STING)}{K_{m5} + IRF3}$$
$$+ k_{5r} \cdot (IRF3_{tot} - IRF3) \tag{A.5}$$

$$\frac{d\left[IFN\beta m\right]}{dt} = \frac{k_{cat6} \cdot (IRF3_{tot} - IRF3)}{K_{m6} + (IRF3_{tot} - IRF3)} + k_{6f} \cdot IRF7Pn - \tau_6 \cdot IFN\beta m \tag{A.6}$$

$$\frac{d\left[IFN\beta\right]}{dt} = \frac{k_{cat7} \cdot IFN\beta m}{K_{m7} + IFN\beta m} - \tau_7 \cdot IFN\beta \tag{A.7}$$

$$\frac{d\left[STATP2n\right]}{dt} = \frac{k_{cat8} \cdot IFN\beta}{Km_8 + IFN\beta} \cdot \frac{1}{1 + k_{8f} \cdot SOCSm} - \tau_8 \cdot SOCSm \tag{A.8}$$

$$\frac{d\left[SOCSm\right]}{dt} = k_{9f} \cdot STATP2n - \tau_9 \cdot SOCSm \tag{A.9}$$

$$\frac{d\left[IRF7m\right]}{dt} = k_{10f1} \cdot STATp2n + k_{10f2} \cdot IRF7Pn - \tau_{10} \cdot IRF7m \tag{A.10}$$

$$\frac{d\left[TREX1m\right]}{dt} = k_{11f} \cdot STATP2n - \tau_{11} \cdot TREX1m \tag{A.11}$$

$$\frac{d\left[IRF7Pn\right]}{dt} = k_{12f} \cdot IRF7m - \tau_{12} \cdot IRF7Pn \tag{A.12}$$

$$\frac{d\left[TREX1\right]}{dt} = k_{13f} \cdot TREX1m - \tau_{13} \cdot TREX1 \tag{A.13}$$

$$\frac{d\left[cGAS\right]}{dt} = -k_{1f} \cdot cGAS \cdot DNA + k_{1r} \cdot cGASc - k_{on} \cdot cGASc \cdot Drug + k_{off} \cdot cGASn \tag{A.14}$$

$$\frac{d[Drug]}{dt} = -k_{on} \cdot cGASc \cdot Drug + k_{off} \cdot cGASn \tag{A.15}$$

$$\frac{d[cGASn]}{dt} = k_{on} \cdot cGASc \cdot Drug - k_{off} \cdot cGASn \tag{A.16}$$

A.2 Parameters

Table 2: Summary of Parameters in ODE Model

Number	Parameter	Units	Minimum Constraint	Maximum Constraint	log_{10} Value (Run 1)	log_{10} Value (Run 2)	log_{10} Value (Run 3)
1	k_{1f}	1/(nM·hour)	-3	1	0.429732	-0.14158	0.571135
2	k_{1r}	1/hour	-2	2	0.685785	-0.60234	0.81712
3	k_{cat2}	1/hour	3	8	5.375736	7.916804	3.922932
4	Km_2	nM	1	6	4.790203	1.236027	3.344868
5	k_{3f}	1/hour	2	6	2.713831	2.853165	2.703932
6	k_{4f}	1/(nM·hour)	-3	1	-1.44837	-1.56758	-1.57522
7	k_{4r}	1/hour	-2	2	0.874307	0.768601	0.762453
8	k_{cat5}	1/hour	2	6	4.348857	2.061381	4.525882
9	K_{m5}	nM	1	5	4.050239	1.740002	4.220628
10	k_{5r}	1/hour	-2	2	-0.02962	-0.0882	-0.02824
11	k_{cat6}	nM/hour	1	5	2.315854	2.670361	2.950268
12	K_{m6}	nM	1	5	4.013067	4.312452	4.643317
13	k_{6f}	1/hour	-2	2	-1.55904	-1.82669	-1.53605
14	k_{cat7}	nM/hour	4	8	7.677969	5.800111	6.377869
15	K_{m7}	nM	-3	5	0.585172	-0.9078	0.996437
16	τ_7	1/hour	2	6	4.087883	3.546575	2.382462
17	k_{cat8}	nM/hour	-2	2	1.114157	1.00444	1.169599
18	K_{m8}	nM	-2	2	1.893234	1.744374	1.97878
19	k_{8f}	1/nM	-2	2	-1.68035	-2.12436	-1.60521
20	k_{9f}	1/hour	-2	2	-2.23089	-2.20471	-2.23297
21	k_{10f1}	1/hour	-2	2	-2.99886	-2.99924	-2.99893
22	k_{10f2}	1/hour	-2	2	-1.95019	-1.9481	-1.95154
23	k_{11f}	1/hour	-4	1	-2.99991	-2.99966	-2.99998
24	k_{12f}	1/hour	-2	2	1.999768	1.998867	1.999795
25	k_{13f}	1/hour	-1	4	1.180231	1.186086	1.187107
Number	Parameter	Units	Value	Reference			
26	τ_3	1/hour	0.96	[197]			
27	τ_6	1/hour	0.347	[137]			
28	τ_8	1/hour	1.238	[135]			
29	τ_9	1/hour	1.507	[137]			
30	τ_{10}	1/hour	0.347	[137]			
31	τ_{11}	1/hour	0.165	[136, 137]			
32	τ_{12}	1/hour	6.93	[198]			
33	τ_{13}	1/hour	0.0178	[135, 136]			
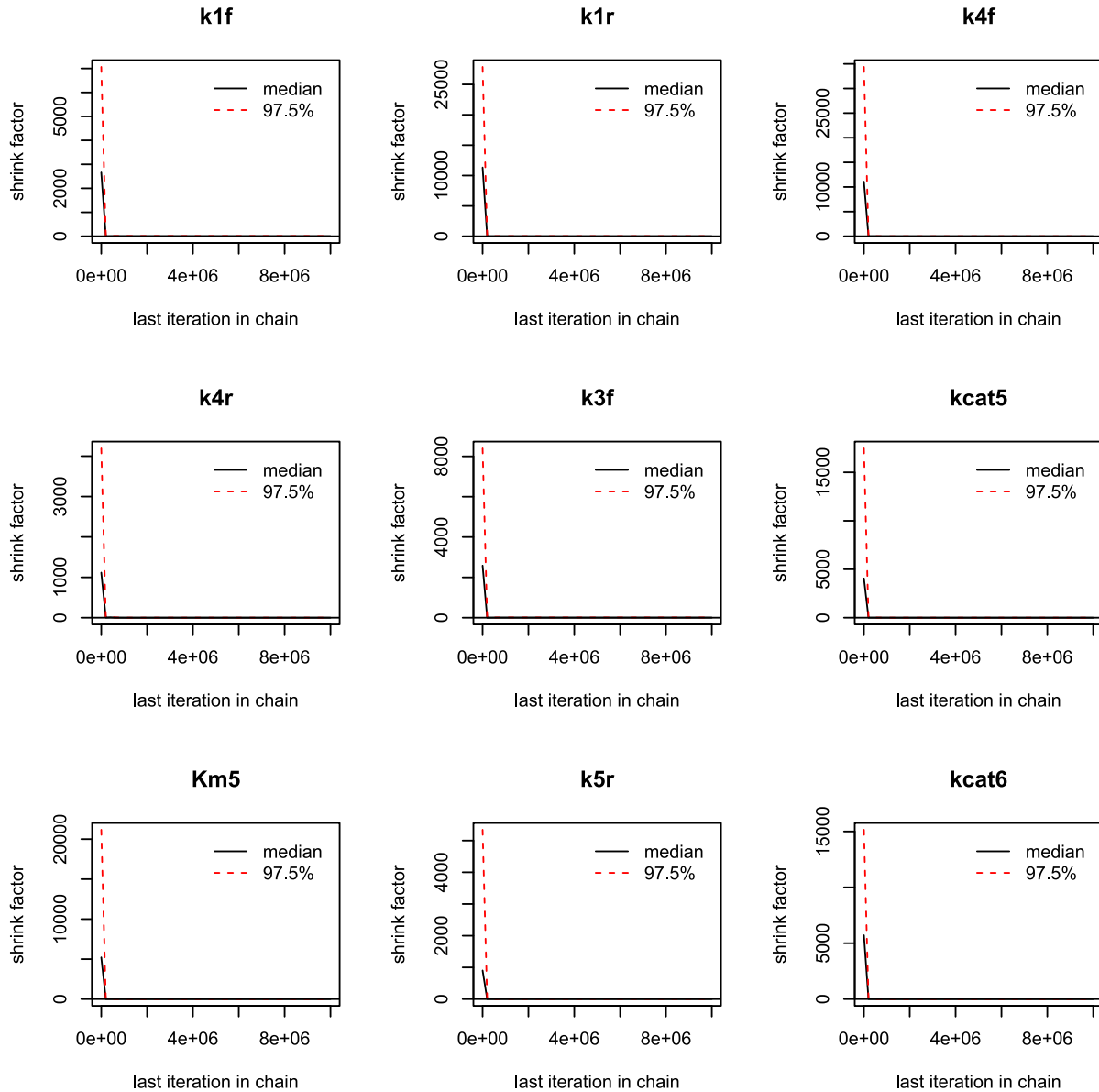
Figure A.2.1: Each plot gives the Gelman-Rubin diagnostic for a given parameter after each iteration. Convergence to a value of one indicates agreement between the different MCMC runs.
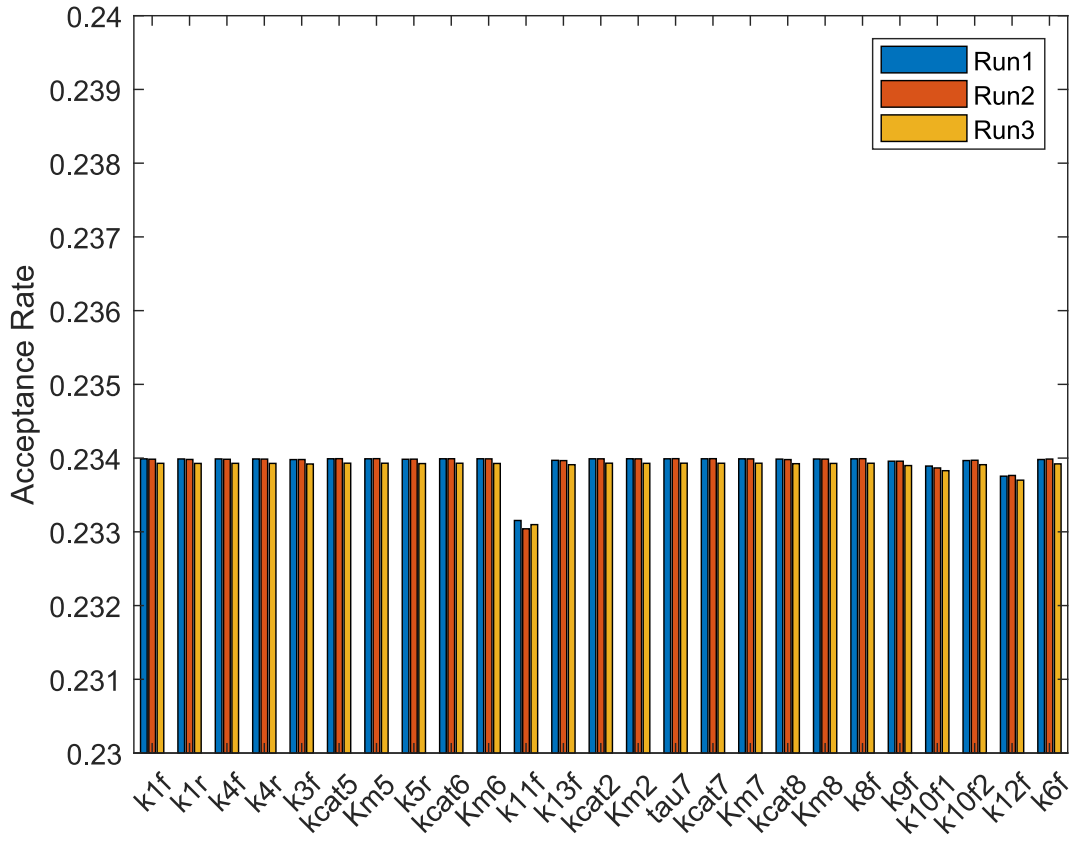
Figure A.2.2: For each run the acceptance rate is shown for all parameters optimized.
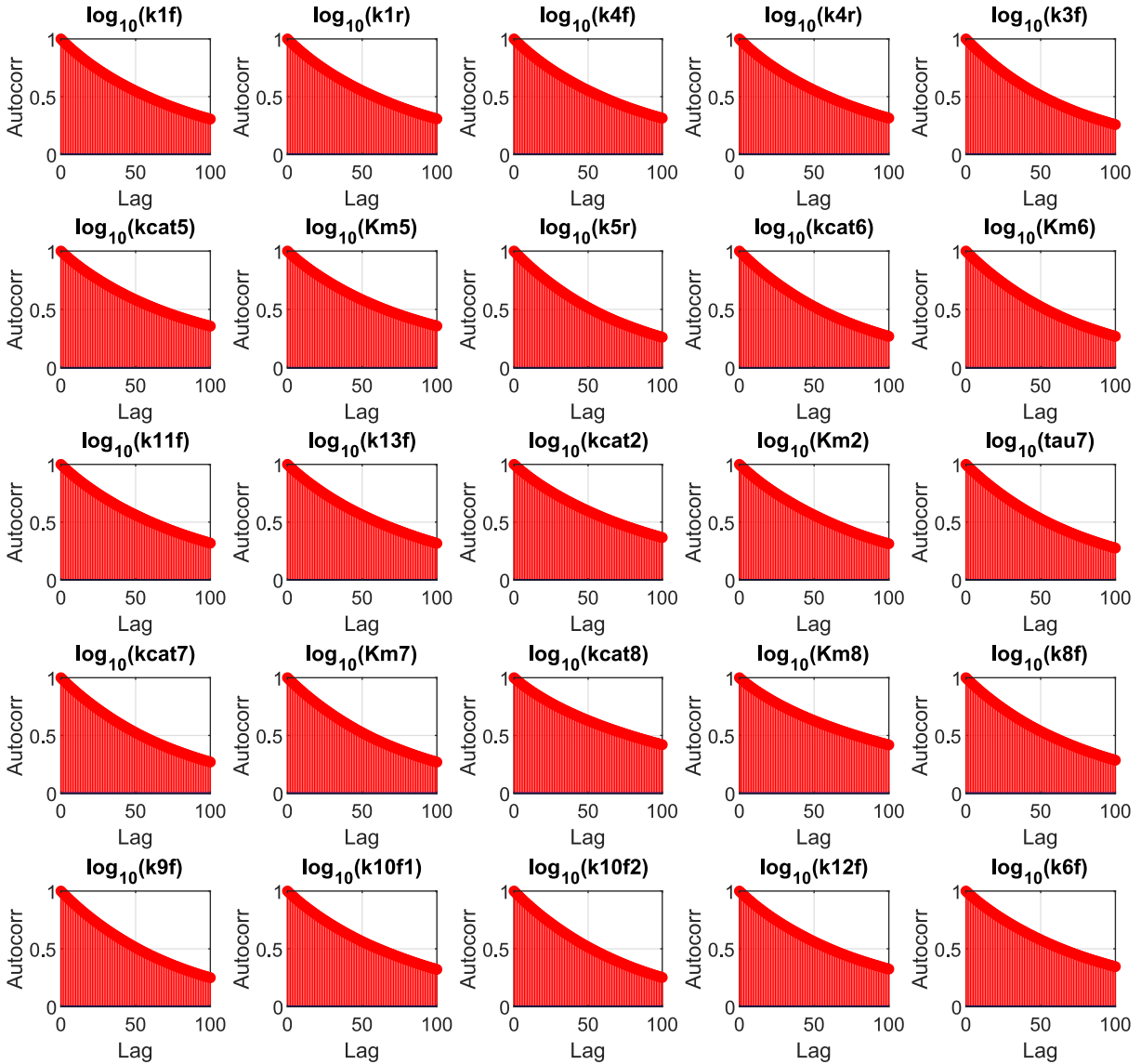
Figure A.2.3: Autocorrelation within chains are tested to determine health of the MCMC runs. Here we show the autocorrelation as a function of lag or distance between tested values for a represented MCMC iteration (Run 1).
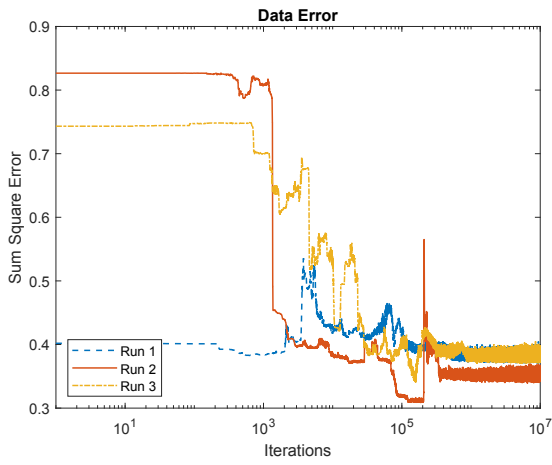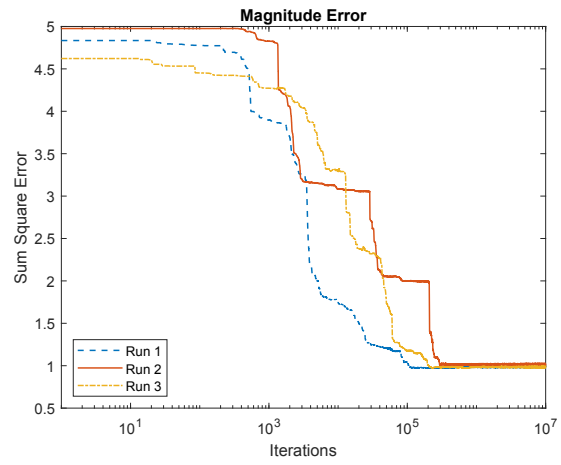
Figure A.2.4: Component errors of the objective function. Here we show the errors associated with (A) the differences between normalized model and observations and (B) differences between expected and peck magnitudes. Each line represents the total error for all species.
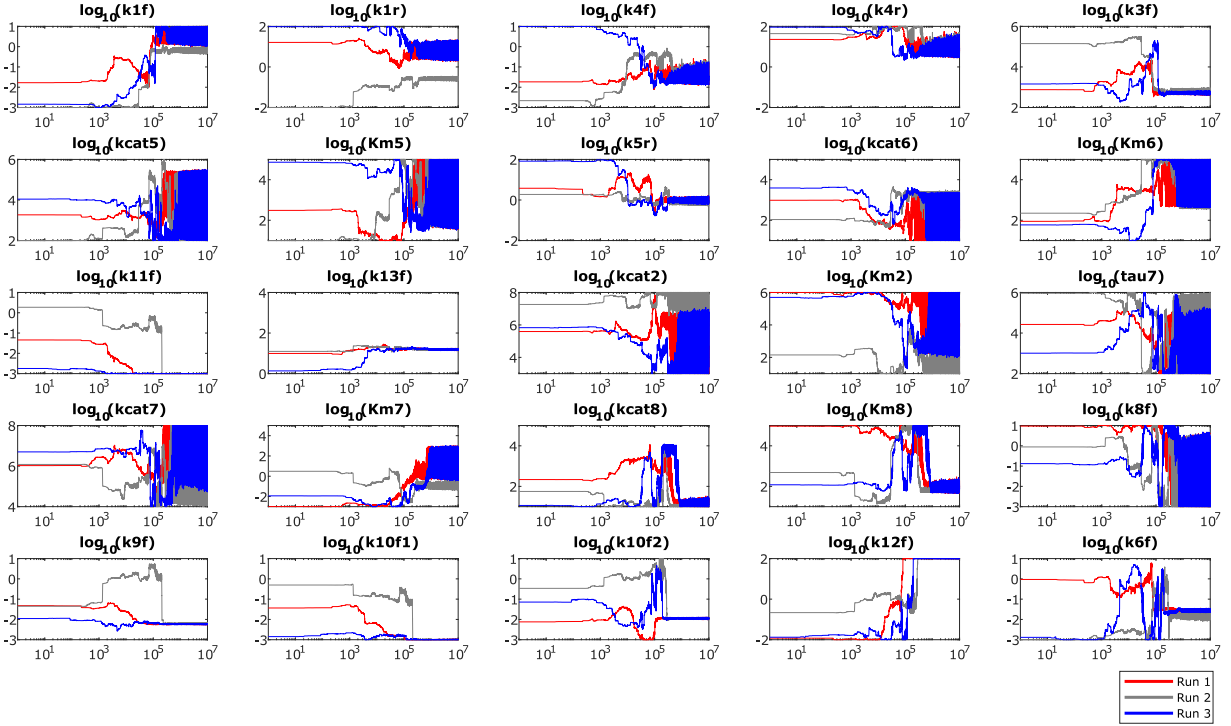
Figure A.2.5: Trace plots for all of the MCMC runs are plotted. The x-axis is the iteration of the MCMC algorithm and the y-axis is the log-scaled value of the parameter. The limits of the y-axis are the scaled to the constraints imposed on the parameters. All of the MCMC reiterations are plotted (run 1 [red], run 2 [gray] and run 3 [blue]) showing how the algorithm sampled the parameter space.
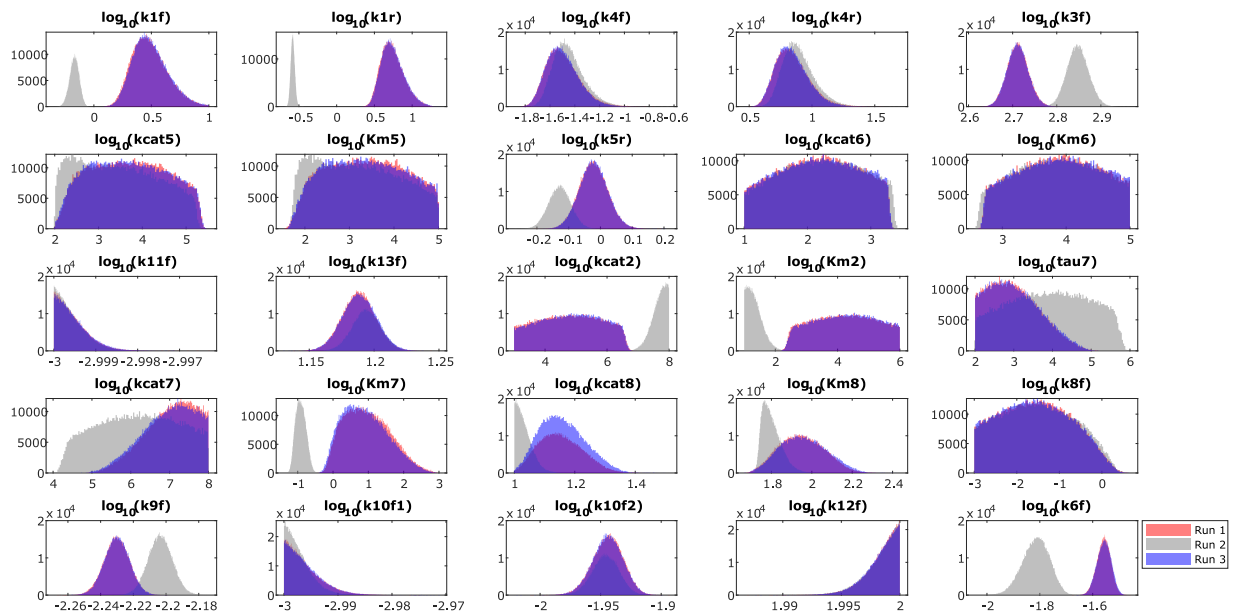
Figure A.2.6: Histograms are given for all the optimized parameters. Each color shows one of the three MCMC runs performed, run 1 (red), run 2 (gray), and run 3 (blue). These histograms represent the posterior probability distributions determined by the MCMC algorithm.
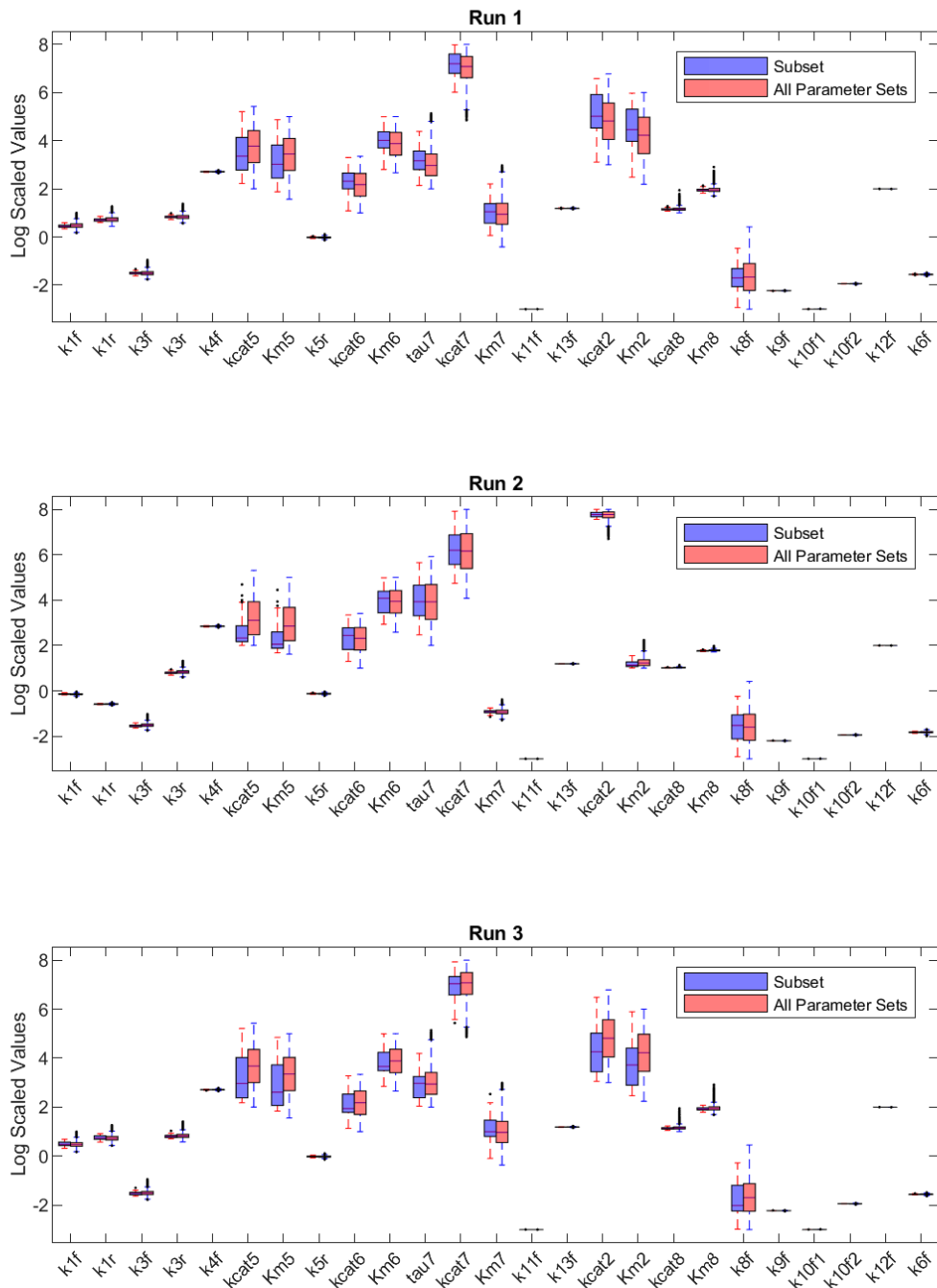
Figure A.2.7: To determine how well the parameter subsets represented the overall parameter sets determined in MCMC we plot the summary statistics for the top 1000 parameter sets (Subset) and the top 1 million parameter sets (All parameter sets). We show all of the comparisons across the different MCMC runs.
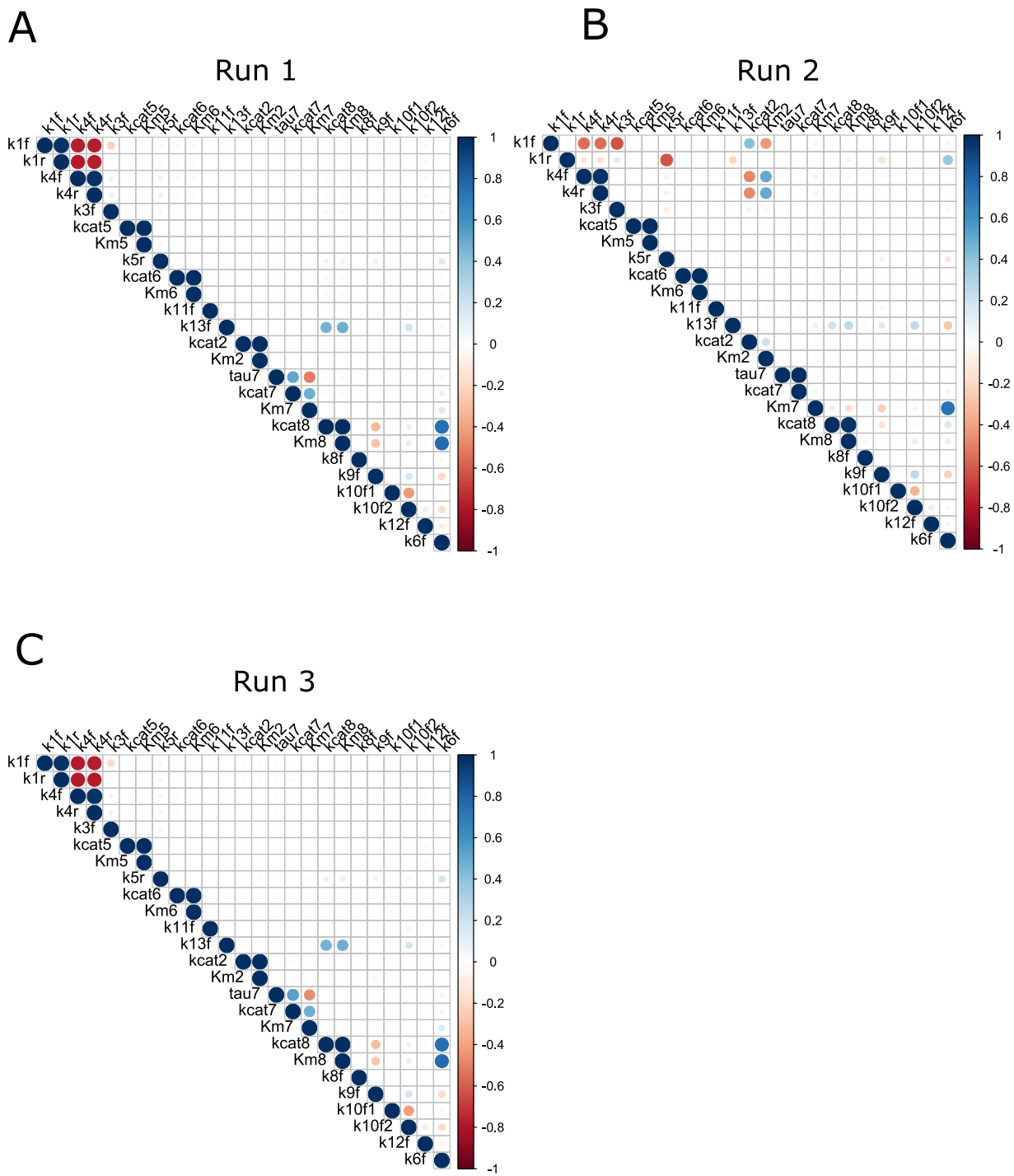
Figure A.2.8: Plotted are the Pearson's correlation between each of the optimized parameters. Blue represents a positive correlation whereas red represents a negative correlation. The last million MCMC iterations are used to examine the correlations.
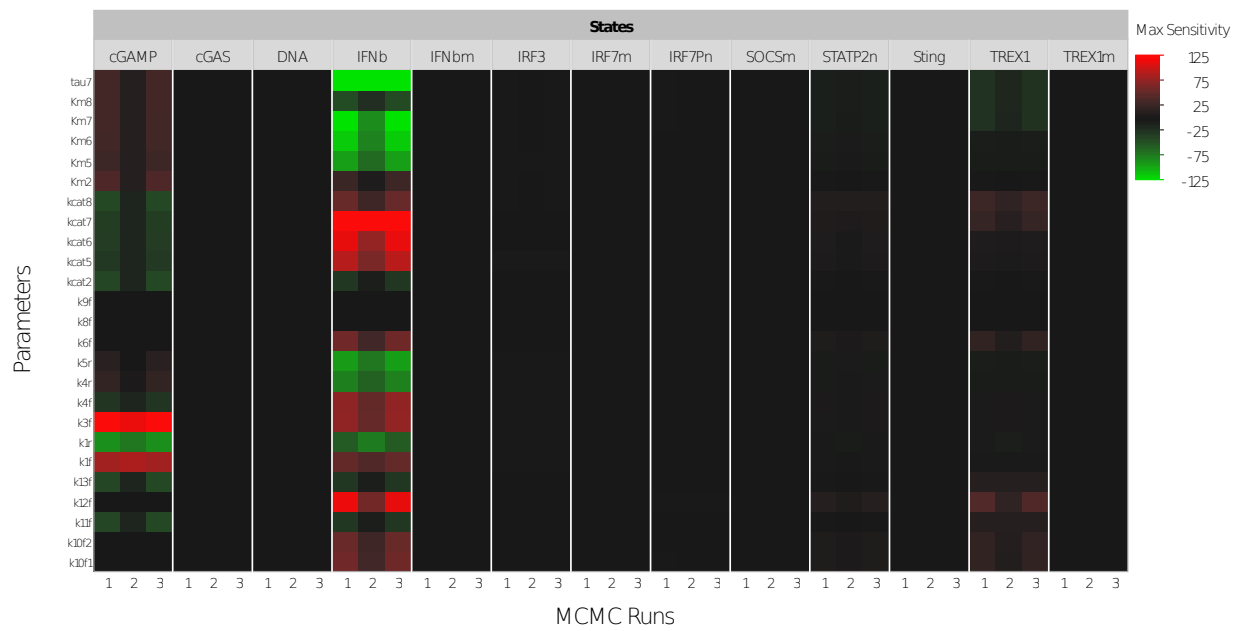
Figure A.2.9: Parameter Sensitivity Analysis (All Species). Here we show another representation of the sensitivity analysis for all of the species in the model. The mean maximum sensitivity is plotted from the top 1000 parameter sets. Positive sensitivity is shown in red and negative sensitivity is shown in green.
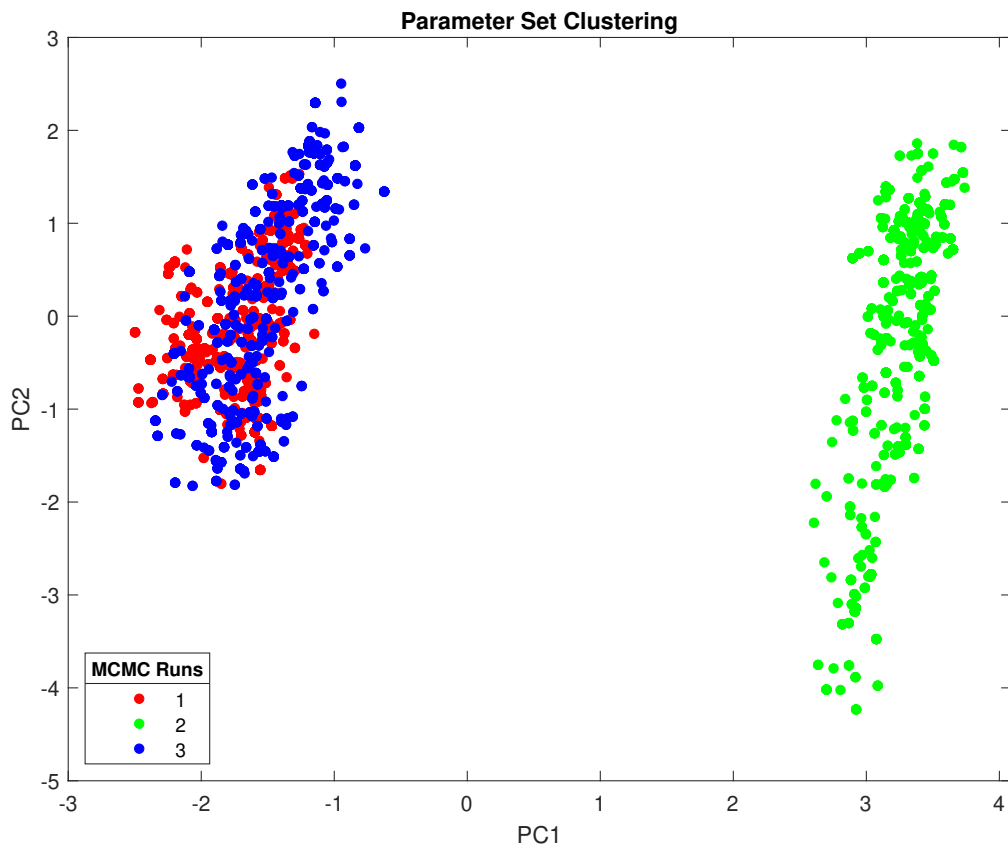
Figure A.2.10: Parameter Clustering. Principle Component Analysis of the top 100 parameter sets for each of the three runs. Plotted are clustered parameter sets as determined by PCA. Each dot represents a column vectors containing the values for the 25 optimized parameters. Colors are assigned based on the MCMC run they originate.
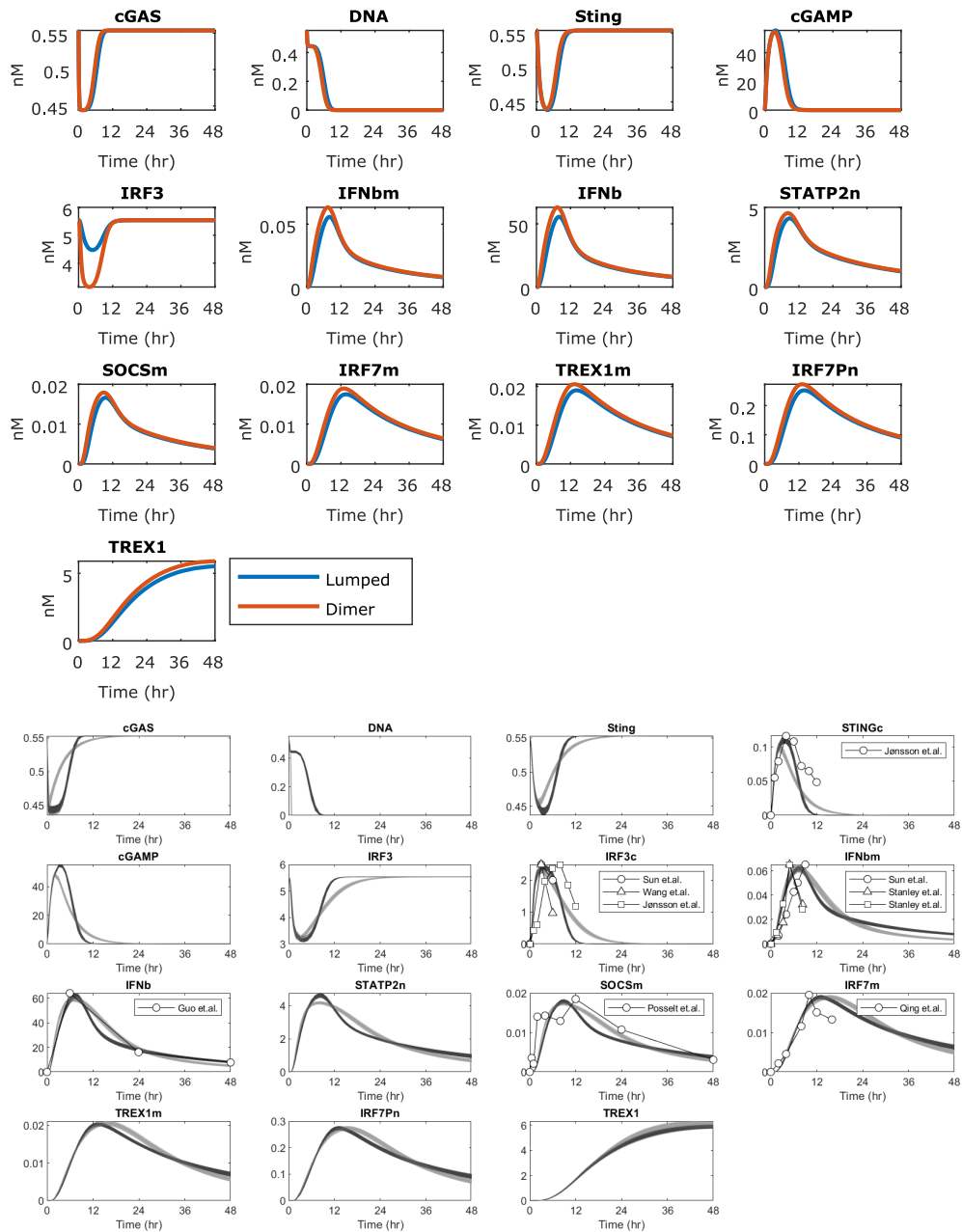
Figure A.2.11: IRF3 Dimerization Assumption. The assumption of lumping dynamics describing IRF3 formation was assessed. Two model structures were simulated: the original (lumped) and a modified model (dimer) that considered dimerization to activate IRF3. (A) Both models were simulated using the best parameterization found in run 1. (B) The same top 1000 parameter ensembles (see figure 2) were simulated showing range of dynamics observed by the modified model.
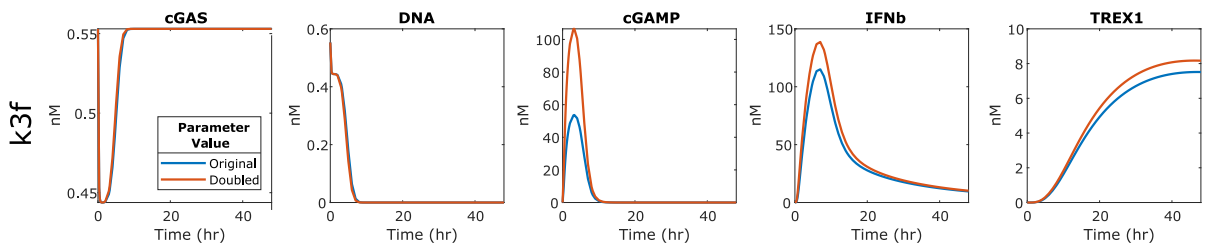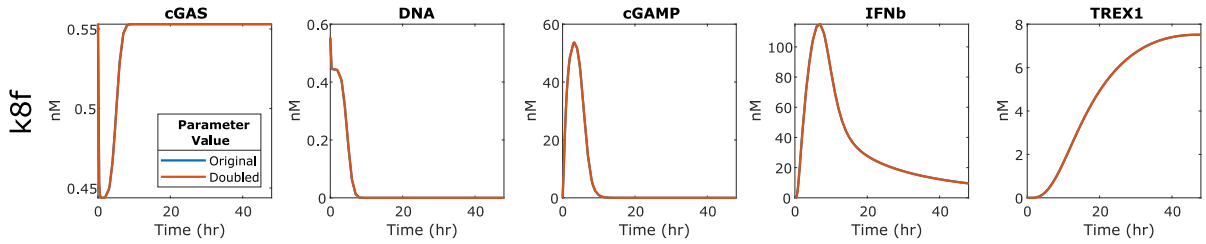
Figure A.2.12: Sensitivity Magnitudes on Parameter Perturbation. Time courses are plotted to demonstrate how the magnitudes of different species change relative to the sensitivity of the parameters. (A) Demonstrates responses for species to a doubling of the sensitive parameter k3f. (B) Demonstrates responses for species to a doubling of the insensitive parameter k8f.

Appendix B Sepsis Model

B.1 Ordinary Differential Equations

$$\frac{dAMP(t)}{dt} = -k1af \cdot AMPKp - k1bf \cdot HIF1 + k1r \tag{B.1}$$

$$\frac{dATP(t)}{dt} = k1af \cdot AMPKp + k1bf \cdot HIF1 - k1r \tag{B.2}$$

$$\frac{dAMPK(t)}{dt} = -AMPK \cdot k2f \cdot AMP + AMPKp \cdot k2r \cdot ATP - AMPK \cdot k3f \cdot CaMKIIa \tag{B.3}$$

$$\frac{dAMPKp(t)}{dt} = AMPK \cdot k2f \cdot AMP - AMPKp \cdot k2r \cdot ATP + AMPK \cdot k3f \cdot CaMKIIa \tag{B.4}$$

$$\frac{dCaMKII(t)}{dt} = -CaMKII \cdot k4f \cdot ROS + k4r \cdot CaMKIIa \tag{B.5}$$

$$\frac{dCaMKIIa(t)}{dt} = CaMKII \cdot k4f \cdot ROS - k4r \cdot CaMKIIa \tag{B.6}$$

$$\frac{dTSC2(t)}{dt} = -TSC2 \cdot k5f \cdot AMPKp + TSC2p \cdot k5r \cdot Aktp \tag{B.7}$$

$$\frac{dTSC2p(t)}{dt} = TSC2 \cdot k5f \cdot AMPKp - TSC2p \cdot k5r \cdot Aktp \tag{B.8}$$

$$\frac{dRheba(t)}{dt} = -Rheba \cdot k6f \cdot TSC2p + k6r \cdot Rheb - k7f \cdot mTORC1 \cdot Rheba + k7r \cdot mTORC1_{Rheba} \tag{B.9}$$

$$\frac{dRheb(t)}{dt} = Rheba \cdot k6f \cdot TSC2p - k6r \cdot Rheb \tag{B.10}$$

$$\frac{dmTORC1(t)}{dt} = -k7f \cdot mTORC1 \cdot Rheba + k7r \cdot mTORC1_{Rheba} \tag{B.11}$$

$$\frac{dmTORC1_{Rheba}(t)}{dt} = k7f \cdot mTORC1 \cdot Rheba - k7r \cdot mTORC1_{Rheba} \tag{B.12}$$

$$\frac{dSIRT6(t)}{dt} = k8f \cdot AMPKp - \tau8 \cdot SIRT6 \tag{B.13}$$

$$\frac{dHIF1a(t)}{dt} = \frac{k9f \cdot mTORC1_{Rheba}}{SIRT6 + 1} - \tau9 \cdot HIF1a - k10f \cdot HIF1a + k10r \cdot HIF1 \tag{B.14}$$

$$\frac{dHIF1(t)}{dt} = k10f \cdot HIF1a - k10r \cdot HIF1 - \tau11 \cdot HIF1 \tag{B.15}$$

$$\frac{dLPS(t)}{dt} = -k12f \cdot LPS \cdot TLR4 + k12r \cdot TLR4_{LPS} \tag{B.16}$$

$$\frac{dTLR4(t)}{dt} = -k12f \cdot LPS \cdot TLR4 + k12r \cdot TLR4_{LPS} \tag{B.17}$$

$$\frac{dTLR4_{LPS}(t)}{dt} = k12f \cdot LPS \cdot TLR4 - k12r \cdot TLR4_{LPS} \tag{B.18}$$

$$\frac{dPI3K(t)}{dt} = -PI3K \cdot k13f \cdot TLR4_{LPS} + k13r \cdot PI3Ka \tag{B.19}$$

$$\frac{dPI3Ka(t)}{dt} = PI3K \cdot k13f \cdot TLR4_{LPS} - k13r \cdot PI3Ka \tag{B.20}$$

$$\frac{dAkt(t)}{dt} = -Akt \cdot k14f \cdot PI3Ka + k14r \cdot Aktp \tag{B.21}$$

$$\frac{dAktp(t)}{dt} = Akt \cdot k14f \cdot PI3Ka - k14r \cdot Aktp \tag{B.22}$$

$$\frac{dROS(t)}{dt} = k15f \cdot AMPKp - \tau15 \cdot ROS \tag{B.23}$$

Appendix C ABM Code

C.1   ProblemGenerator.jl

C.1.1   Import Dependencies

```julia
#Import Packages needed for simulation
using DifferentialEquations, Sundials #For implementing Differential equations
using LinearAlgebra, SparseArrays, Distributions, Statistics, Random #Linear Algebra and Statistics
using CSV, DataFrames #Data handling
using StatsPlots #For graphing
using JLD2, FileIO #Saving simulations
```

C.1.2   Define all the constants that will not change

```julia
#Constants for all cells
const N=200 #number of grid points along one dimensions
const nCells = N^2 #number of cells in the simulation
const cellVol = 3e-12 #Cell Volume (liters)
const Na = 6.02e23 #Avagadro's number
const species = 14 #Number of states within each cell (including virus)
const moi = 1.0e-2 #Multicity of infection
const cellIndicies = CartesianIndices(zeros(N,N)) #set of indices for looping through every cell

#Function that converts molecules to nM
m2c(molecule) = @. 1e9*molecule/(cellVol*Na)

#Constants for all simulations
const tspan = (0.0,48.0) #Time span for simulation
const ΔIFNβ = zeros(N,N) #Define memory space to hold the Laplacian
const statesNames = ["cGAS","DNA","Sting","cGAMP","IRF3","IFNbm","IFNb","STAT",
                     "SOCSm","IRF7m","TREX1m","IRF7","TREX1","Virus"] #for plotting
const θNames = [:k1f, :k1r, :k3f, :k3r, :k4f, :kcat5, :Km5, :k5r, :kcat6, :Km6, :kcat7,
 :Km7, :kcat8, :Km8, :k8f, :k9f, :k10f1, :k10f2, :k11f, :k12f, :k13f, :k6f, :kcat2,
 :Km2, :τ4, :τ6, :τ7, :τ8, :τ9, :τ10, :τ11, :τ12, :τ13, :k14f,:τ14] #Parameter names
 #These species are not starting at zero
 const nonZeroSpeciesIdx = [1,3,5] #cGAS, Sting, IRF3
 const nonZeroSpeciesValues = m2c([1e3, 1e3, 1e4]) #convert to concentration

#This function modifies initial values by adding guassian noise
#Standard deviation is same magnitude as mean
# If p=0.5 then: (u1=0.1 u2=100)  (sigma1=0.05 sigma2=50) instead of (sigma1=0.5 sigma2=0.5)

AddNoise2States(p) = [truncated(Normal(μ,p*μ),0,Inf) for μ in nonZeroSpeciesValues]

 #Often it is useful to pass parameters between functions during the ODE solve,
 #This struct hold all the parameters that we need to keep track of
 mutable struct ParContainer{T}
   par::T #Rate Constants
   mass::Vector{Array{Float64,2}} #Mass balances
   deathParameter::Array{Float64,2} # 0 or 1 indicating cell is dead
   DNAReplicate::Int64 #0 if ISD, 1 if Virus
   cellsInfected::Array{Float64,2} #Time cell was infected
   cellsDead::Array{Float64,2} #Time cell was killed
   infectFirstAttempt::BitArray{2} #Has cell tried to infect neighbors?
 end
```

C.1.3   Helper function that discretizes the Laplacian

```
#u is updated with Laplacian estimate from u
 function (Δu,u)
   #Get dimensions of the input and define some constants
     n1, n2 = size(u)
     Δx = 32.0 #Grid spacing (diameter of cell in μm)
     D=97.5*3600.0 #Diffusion coefficient (μm^2/hr)
     h = D/Δx^2

     # internal nodes
     for j = 2:n2-1
         for i = 2:n1-1
             @inbounds  Δu[i,j] = h*(u[i+1,j] + u[i-1,j] + u[i,j+1] + u[i,j-1] - 4*u[i,j])
         end
     end

     # left/right edges
     for i = 2:n1-1
         @inbounds Δu[i,1] = h*(u[i+1,1] + u[i-1,1] + 2*u[i,2] - 4*u[i,1])
         @inbounds Δu[i,n2] = h*(u[i+1,n2] + u[i-1,n2] + 2*u[i,n2-1] - 4*u[i,n2])
     end

     # top/bottom edges
     for j = 2:n2-1
         @inbounds Δu[1,j] = h*(u[1,j+1] + u[1,j-1] + 2*u[2,j] - 4*u[1,j])
         @inbounds Δu[n1,j] = h*(u[n1,j+1] + u[n1,j-1] + 2*u[n1-1,j] - 4*u[n1,j])
     end

     # corners
     @inbounds Δu[1,1] = h*(2*(u[2,1] + u[1,2]) - 4*u[1,1])
     @inbounds Δu[n1,1] = h*(2*(u[n1-1,1] + u[n1,2]) - 4*u[n1,1])
     @inbounds Δu[1,n2] = h*(2*(u[2,n2] + u[1,n2-1]) - 4*u[1,n2])
     @inbounds Δu[n1,n2] = h*(2*(u[n1-1,n2] + u[n1,n2-1]) - 4*u[n1,n2])

 end
```

C.1.4   This is the ODE model for the cGAS pathway

```
# Define the discretized PDE as an ODE function

 function Model!(du,u,p,t)

   #Species
   cGAS = @view u[:,:,1]
   DNA = @view u[:,:,2]
   Sting = @view u[:,:,3]
   cGAMP = @view u[:,:,4]
   IRF3 = @view u[:,:,5]
   IFNβm = @view u[:,:,6]
   IFNβ = @view u[:,:,7]
   STAT = @view u[:,:,8]
   SOCSm = @view u[:,:,9]
   IRF7m = @view u[:,:,10]
   TREX1m = @view u[:,:,11]
   IRF7 = @view u[:,:,12]
   TREX1 = @view u[:,:,13]
   Virus = @view u[:,:,14]

   #Derivatives
   d_cGAS = @view du[:,:,1]
   d_DNA = @view du[:,:,2]
   d_Sting = @view du[:,:,3]
   d_cGAMP = @view du[:,:,4]
   d_IRF3 = @view du[:,:,5]
```

```
    d_IFNβm = @view du[:,:,6]
    d_IFNβ = @view du[:,:,7]
    d_STAT = @view du[:,:,8]
    d_SOCSm = @view du[:,:,9]
    d_IRF7m = @view du[:,:,10]
    d_TREX1m = @view du[:,:,11]
    d_IRF7 = @view du[:,:,12]
    d_TREX1 = @view du[:,:,13]
    d_Virus = @view du[:,:,14]

    #Parameters
    k1f, k1r, k3f, k3r, k4f, kcat5, Km5, k5r, kcat6, Km6, kcat7, Km7, kcat8, Km8, k8f, k9f, k10f1, k10f2, k11f,
    k12f, k13f, k6f, kcat2, Km2, τ4, τ6, τ7, τ8, τ9, τ10, τ11, τ12, τ13, k14f, τ14 = p.par
    #Constants from the mass balances
    cGAStot, Stingtot, IRF3tot = p.mass
    #Which cells are dead?
    skull = p.deathParameter
    #Should DNA be allowed to replicate (only with virus, not with ISD)
    rep = p.DNAReplicate

    #Calculate the diffusion of IFNβ
    (ΔIFNβ,IFNβ)

    #Update derivatives for each species according to model
    @. d_cGAS = -k1f*cGAS*DNA + k1r*(cGAStot - cGAS)
    @. d_DNA = -k1f*cGAS*DNA + k1r*(cGAStot - cGAS) - kcat2*TREX1*DNA / (Km2 + DNA) + rep*skull*DNA*(0.55-DNA)/0.55
    @. d_Sting = -k3f*cGAMP*Sting + k3r*(Stingtot - Sting)
    @. d_cGAMP = k4f*(cGAStot - cGAS) - k3f*cGAMP*Sting + k3f*(Stingtot - Sting) - τ4*cGAMP
    @. d_IRF3 = -kcat5*IRF3*(Stingtot - Sting) / (Km5 +IRF3) + k5r*(IRF3tot - IRF3)
    @. d_IFNβm = skull*kcat6*(IRF3tot - IRF3) / (Km6 + (IRF3tot - IRF3)) + skull*k6f*IRF7 - τ6*IFNβm
    @. d_IFNβ = skull*kcat7*IFNβm / (Km7 + IFNβm) - τ7*IFNβ + ΔIFNβ #Add the diffusion in here
    @. d_STAT = skull*kcat8*IFNβ / (Km8 + IFNβ) * 1.0/(1.0+k8f*SOCSm) - τ8*STAT
    @. d_SOCSm = skull*k9f*STAT - τ9*SOCSm
    @. d_IRF7m = skull*k10f1*STAT + skull*k10f2*IRF7 - τ10*IRF7m
    @. d_TREX1m = skull*k11f*STAT - τ11*TREX1m
    @. d_IRF7 = skull*k12f*IRF7m - τ12*IRF7
    @. d_TREX1 = skull*k13f*TREX1m - τ13*TREX1
    @. d_Virus = skull*k14f*DNA - τ14*Virus
end
```

C.1.5   Set up function that will return an ODE problem to solve

```
function ModelSetup(infectionMethod,IFNStoch,Hetero)
    #Parameter values for the ODEs
    θVals = [2.6899, 4.8505, 0.0356, 7.487, 517.4056, 22328.3852, 11226.3682,0.9341,
            206.9446, 10305.461, 47639.70295,3.8474, 13.006, 78.2048, 0.0209,
            0.0059, 0.001, 0.0112, 0.001, 99.9466, 15.1436,0.0276, 237539.3249,
            61688.259, 0.96, 0.347, 12.2428736,1.2399, 1.5101, 0.347, 0.165, 6.9295,
            0.0178]
    θVirus = [1.0, 1.0] # k14f τ14 (Virus Parameters)
    append!(θVals,θVirus) #Append the virus parameters to the orginal parameters

    #Should IFN signaling be stochastic?
    if IFNStoch == :Stochastic
        #Keep most parameters the same
        θ = fill.(θVals,N,N)
        #kcat8 produces IFN, make it nonzero ~20% of the time (can be changed later)
        θ[11] .= θVals[11] .* rand(Bernoulli(0.2),N,N)
    else
        θ = θVals #Just keep the parameters as is (same for each cell)
    end

#Now we need to deal with the initial condition
```

```julia
    #Define the initial conditions
    u0 = zeros(N,N,species)

if Hetero == :Hetero
    noiseDistributions = AddNoise2States(0.5)
    for (i,index) in enumerate(nonZeroSpeciesIdx)
        u0[:,:,index] = rand(noiseDistributions[i],N,N)
    end

else
    #Loop through non-zero species and update their concentrations
    for (idx,val) in zip(nonZeroSpeciesIdx,nonZeroSpeciesValues)
        u0[:,:,idx] .= val
    end
end

    #Finally need to set the DNA initial condition
    if infectionMethod == :ISD
        #Define a region on the domain where cells will be infected
        circleOrigin = [0,0] #Where is the center of the drop?
        circleRadiusSquared = N^2 #How big is the drop?
        #Calculate squared distances
        sqDist(x,c) = reduce(+, @. (x-c)^2)
        #Loop though cells and check if they are infected
        for currentCell in cellIndicies
            #Are the cells inside the infected region?
            if sqDist([currentCell[1],currentCell[2]],circleOrigin)  circleRadiusSquared
                u0[currentCell,2] = m2c(1e3)
            end
        end

    elseif infectionMethod == :Virus
        #Assume a poisson ditribution to randomly choose each cell's level of infection
        probDistInfected = Poisson(moi)
        u0[:,:,2] = @. m2c(1e3*rand(probDistInfected,N,N))
    end

#Need to wrap everything up into the ParContainer struct
    mass = [u0[:,:,i] for i in nonZeroSpeciesIdx]
    deathParameter = ones(N,N)
    DNAReplicate = infectionMethod==:ISD ? 0 : 1
    #Keep track of infected cells (save time when infected, Inf means not infected)
    cellsInfected = fill(Inf,N,N) #Make constant when not testing
    cellsInfected[findall(u0[:,:,2] .> 0.0), 1] .= 0.0
    #Keep track of time of death (TOD)
    cellsDead = fill(Inf,N,N) #Inf implies alive
    #Create an array that keeps track of whether or not a cell has tried to infect neighbors
    infectFirstAttempt = trues(N,N)

#Create an instance of the structures
θ = ParContainer(
    θ, #Rate constants
    mass, #Mass balances
    deathParameter, # cell is dead? (1==alive, 0==dead)
    DNAReplicate, #can DNA replicate? (1 if virus)
    cellsInfected, #Time cell was infected
    cellsDead, #Time cell was killed
    infectFirstAttempt) #Has cell tried to infect neighbors?)

return ODEProblem(Model!,u0,tspan,θ)
end
```

C.1.6   Function to count the cell states (healthy,infected,dead)

```
function cellStates(t,θ)
    #Number of healthy cells at time t
    totaHealthy = sum(θ.cellsInfected .> t)
    #Number of dead cells at time t
    totalDead = sum(θ.cellsDead .< t)
    #Number of infected cells at time t
    totalInfected = nCells - totaHealthy - totalDead
    return [totaHealthy,totalInfected,totalDead]
end
```

C.1.7   Display Information about the Parameter Container

```
function Base.show(io::IO,p::ParContainer)
  #Is simulation det/stoch and homo/hetero
  attribute1 = isa(p.par[1],Array) ? "Stochastic" : "Deterministic" #True if stochastic
  attribute2 = all(p.mass[1][1].==p.mass[1]) ? "Homogeneous" : "Heterogeneous" #True is homo
  println("Cell Population: ",attribute1," and ",attribute2)

  #What is the Input for DNA?
  modelInput = p.DNAReplicate==1 ? "Virus" : "ISD"
  println("Model Input: ",modelInput)

  #How many cells initially infected?
  println("Initially Infected: ",100.0*sum(p.cellsInfected.==0.0)/nCells, "%")

  println("Currently Infected: ",100.0*sum(p.cellsInfected.<Inf)/nCells, "%")

  #How many cells are initially dead?
  println("Dead: ",100.0*((nCells - sum(p.cellsDead.==Inf))/nCells), "%")

  #How many cells are not responsive to infection?
  if attribute1 == "Stochastic"
    println("Cells with kcat7=0: ",100.0*sum(p.par[11].==0.0)/nCells, "%")
  end

end
```

C.2   VirusCallBacks.jl

C.2.1   Callback 1: Healthy → Infected

```
#Cells are infected after a certain viral load is met (with μ=800 virons, σ=200 virons)
const maxViralLoad = rand(Normal(m2c(800.0),m2c(200.0)),N,N)

function CheckInfect(u,t,integrator)
    #Reshape the array into 3D
    uReshaped =  reshape(u,N,N,species)

    #Check the whole virus grid for cells with viral conc. greater than 0.4 nM
    return any( @. (uReshaped[:,:,end] > maxViralLoad) & integrator.p.infectFirstAttempt)

end

#These don't need to be constructed every loop
    const Index = CartesianIndices(zeros(N,N))
```

```
        #Get the first and last index
        const Ifirst, Ilast = first(Index), last(Index)

function TryInfect(integrator)
    #Probability that an infected cell will spread infection
    chanceOfInfection = Bernoulli(0.5)
    cellsInfected = integrator.p.cellsInfected
    @show integrator.t

    #Sundials does not presereve problem shape...
    uReshaped = reshape(integrator.u,N,N,species)


    #Loop through all of the infectious cells
     for I in findall(uReshaped[:,:,end] .> maxViralLoad)
        #Loop through all neighbors to see if they get infected
        for J in max(Ifirst, I-Ifirst):min(Ilast, I+Ifirst)
            #If the cell is healthy and not the current cell try to infect
            @inbounds if isinf(cellsInfected[J]) && (I != J) && rand(chanceOfInfection)
                #Add Viral DNA to the cell
                uReshaped[J,2] = rand(Uniform(0.0,m2c(1e3)))
                #Mark that the cell was infected
                cellsInfected[J] = integrator.t
            end
        end
        #This cell has used its one attempt to infect neighboring cells
        integrator.p.infectFirstAttempt[I] = false
    end

    integrator.u = uReshaped[:]
end

cbInfect = DiscreteCallback(CheckInfect,TryInfect, save_positions=(false,false))
```

C.2.2 Callback 2: Infected → Dead

```
#Cells die some time after infection (with μ=8.0 hours, σ=1.0 hours)
const time2Death = rand(Normal(8.0,1.0),N,N)

function DeathCheck(u,t,integrator)
    cellsInfected = integrator.p.cellsInfected
    #Loop through and check if a cell should die
    cellCounter = 1
    while cellCounter < length(cellsInfected)
        #If cell infected and cell past time of death
        pastTimeOfDeath = (t - cellsInfected[cellCounter]) > time2Death[cellCounter]
        isAlive = isinf(integrator.p.cellsDead[cellCounter])
        if pastTimeOfDeath & isAlive
            return true
        end
        cellCounter += 1
    end
    return false
end

function KillCell(integrator)
    cellsInfected = integrator.p.cellsInfected
    #Find all the cells to kill
    pastTimeOfDeath = (integrator.t .- cellsInfected) .> time2Death
    isAlive = isinf.(integrator.p.cellsDead)
    targets = findall(isAlive .& pastTimeOfDeath)
    #println(sum(integrator.p.deathParameter .== 0.0))
    #Set their deathParameter to stop all cell function
    integrator.p.deathParameter[targets] .= 0.0
    #Mark time of death
```

```julia
        integrator.p.cellsDead[targets] .= integrator.t
end

cbDead = DiscreteCallback(DeathCheck,KillCell,save_positions=(false,false))

cb =CallbackSet(cbInfect,cbDead)
```

C.3   Example Code (Figure 4.3.2)

```julia
#Shows what a typical ISD and Viral Infections looks like
using RCall
include("ProblemGenerator.jl")
include("VirusCallBacks.jl")

#########################################################################
#Heatmaps of IFN dynamics (Figure 2)
#########################################################################

#-----------ISD, Not Stochastic, Homo-----------
probISD_nS_Ho = ModelSetup(:ISD,:notStochastic,:Homo)
    solISD_nS_Ho = @time solve(probISD_nS_Ho,CVODE_BDF(linear_solver=:GMRES),saveat=0.1)
ISD_nS_Ho = DataFrame()
ISD_nS_Ho.x = repeat(1:N,N)
ISD_nS_Ho.y = repeat(1:N,inner=N)
ISD_nS_Ho.IFN = vec(solISD_nS_Ho(10.0)[:,:,7])

#-----------Virus, Not Stochastic, Homo-----------
probVirus_nS_Ho = ModelSetup(:Virus,:notStochastic,:Homo)
    solVirus_nS_Ho = @time solve(probVirus_nS_Ho,CVODE_BDF(linear_solver=:GMRES),saveat=0.1,callback=cb)
Virus_nS_Ho = DataFrame()
Virus_nS_Ho.x = repeat(1:N,N)
Virus_nS_Ho.y = repeat(1:N,inner=N)
Virus_nS_Ho.IFN = vec(solVirus_nS_Ho(10.0)[:,:,7])

#-----------Virus, Not Stochastic, Homo, Cell States-----------
function CellStateGrid(θ::ParContainer,t::Float64)
  grid = zeros(Int64,size(θ.cellsInfected))
  grid[θ.cellsInfected .<= t] .= 1
  grid[θ.cellsDead .<= t] .= 2

  return grid
end

CellStateVirus_nS_Ho_raw = CellStateGrid(probVirus_nS_Ho.p,10.0)
CellStateVirus_nS_Ho = DataFrame()
CellStateVirus_nS_Ho.x = repeat(1:N,N)
CellStateVirus_nS_Ho.y = repeat(1:N,inner=N)
CellStateVirus_nS_Ho.Cell = vec(CellStateVirus_nS_Ho_raw)


#-----------ISD, Stochastic, Hetero-----------
probISD_S_He = ModelSetup(:ISD,:Stochastic,:Hetero)
    solISD_S_He = @time solve(probISD_S_He,CVODE_BDF(linear_solver=:GMRES),saveat=0.1)
ISD_S_He = DataFrame()
ISD_S_He.x = repeat(1:N,N)
ISD_S_He.y = repeat(1:N,inner=N)
ISD_S_He.IFN = vec(solISD_S_He(10.0)[:,:,7])

#-----------Virus, Stochastic, Hetero-----------
probVirus_S_He = ModelSetup(:Virus,:Stochastic,:Hetero)
    solVirus_S_He = @time solve(probVirus_S_He,CVODE_BDF(linear_solver=:GMRES),saveat=0.1,callback=cb)
```

```
Virus_S_He = DataFrame()
Virus_S_He.x = repeat(1:N,N)
Virus_S_He.y = repeat(1:N,inner=N)
Virus_S_He.IFN = vec(solVirus_S_He(10.0)[:,:,7])

#-----------Virus, Not Stochastic, Homo, Cell States-----------

CellStateVirus_S_He_raw = CellStateGrid(probVirus_S_He.p,10.0)
CellStateVirus_S_He = DataFrame()
CellStateVirus_S_He.x = repeat(1:N,N)
CellStateVirus_S_He.y = repeat(1:N,inner=N)
CellStateVirus_S_He.Cell = vec(CellStateVirus_S_He_raw)




#-----------R Plot-----------
@rput ISD_nS_Ho
@rput Virus_nS_Ho
@rput CellStateVirus_nS_Ho
@rput ISD_S_He
@rput Virus_S_He
@rput CellStateVirus_S_He

R"""

  library(ggplot2)
  library(ggpubr)

  #To draw an arc on the heatmap to show initial condition
  arc <- data.frame(
  x0 = seq(0, 200, length.out = 1000),
  y0 = sqrt(200^2 - (seq(0, 200, length.out = 1000)^2))
  )

  #Color Range
  clim = c(min(ISD_S_He$IFN), max(ISD_nS_Ho$IFN))
  commonFigureOptions <- list(scale_fill_distiller(palette = "Spectral",
        guide = guide_colorbar(frame.colour = "black", ticks.colour = "black"),
        limits = clim),
    scale_x_continuous(expand=c(0,0)),
    scale_y_continuous(expand=c(0,0)),
    theme_bw(base_size = 12),
    ylab("Cell"),
    xlab("Cell"),
    labs(fill="IFN (nM)"),
    theme(plot.title = element_text(hjust = 0.5),aspect.ratio = 1))

  p1 <- ggplot() +
    geom_raster(ISD_nS_Ho, mapping =aes(x, y, fill=IFN)) +
    geom_line(data=arc, aes(x=x0, y=y0),size=0.2,linetype="dashed") +
    labs(fill="IFN (nM)") +
    ggtitle("ISD \n Deterministic \n Homogeneous") +
    commonFigureOptions

  p2 <- ggplot(Virus_nS_Ho, aes(x, y, fill=IFN)) +
    geom_raster(aes(fill=IFN)) +
    ggtitle("Virus \n Deterministic \n Homogeneous") +
    commonFigureOptions

  p3 <- ggplot(CellStateVirus_nS_Ho, aes(x, y, fill=factor(Cell))) +
    geom_raster(aes(fill=factor(Cell))) +
    scale_fill_manual(values=c("#377EB8","#4DAF4A","#E41A1C"),labels = c("Healthy", "Infected","Dead")) +
    scale_x_continuous(expand=c(0,0)) +
    scale_y_continuous(expand=c(0,0)) +
    theme_bw() +
    labs(title ="Viral \n Infection \n Cell States", x="Cell", y="Cell", fill = "Cell State") +
    theme(plot.title = element_text(hjust = 0.5),aspect.ratio = 1)
```

```
p4 <- ggplot() +
  geom_raster(ISD_S_He, mapping =aes(x, y, fill=IFN)) +
  geom_line(data=arc, aes(x=x0, y=y0),size=0.2,linetype="dashed") +
  ggtitle("ISD \n Stochastic \n Heterogeneous") +
  commonFigureOptions

p5 <- ggplot(Virus_S_He, aes(x, y, fill=IFN)) +
  geom_raster(aes(fill=IFN)) +
  ggtitle("Virus \n Stochastic \n Heterogeneous") +
  commonFigureOptions

p6 <- ggplot(CellStateVirus_S_He, aes(x, y, fill=factor(Cell))) +
  geom_raster(aes(fill=factor(Cell))) +
  scale_fill_manual(values=c("#377EB8","#4DAF4A","#E41A1C"),labels = c("Healthy", "Infected","Dead")) +
  scale_x_continuous(expand=c(0,0)) +
  scale_y_continuous(expand=c(0,0)) +
  theme_bw() +
  labs(title ="Viral \n Infection \n Cell States",x="Cell", y="Cell", fill = "Cell State") +
  theme(plot.title = element_text(hjust = 0.5),aspect.ratio = 1)

figure1 <- ggarrange(p1, p2, p3, p4, p5, p6,
                     labels = "AUTO",
                     align = "h",
                     ncol = 3, nrow = 2)

ggsave("./Figures/Figure2New.pdf",width=12,height=8,units="in")
"""
```

Bibliography

[1] Ernest C Borden, Ganes C Sen, Gilles Uze, Robert H Silverman, Richard M Ransohoff, Graham R Foster, and George R Stark. Interferons at age 50: past, current and future impact on biomedicine. *Nature reviews. Drug discovery*, 6(12):975–90, dec 2007.

[2] William M. Schneider, Meike Dittmann Chevillotte, and Charles M. Rice. Interferon-Stimulated Genes: A Complex Web of Host Defenses. *Annual Review of Immunology*, 32(1):513–545, mar 2014.

[3] José E Belizário, Jennifer M Neyra, and Maria Fernanda Setubal Destro Rodrigues. When and how nk cell-induced programmed cell death benefits immunological protection against intracellular pathogen infection. *Innate immunity*, 24(8):452–465, 2018.

[4] Amira A Barkal, Kipp Weiskopf, Kevin S Kao, Sydney R Gordon, Benyamin Rosental, Ying Y Yiu, Benson M George, Maxim Markovic, Nan G Ring, Jonathan M Tsai, et al. Engagement of mhc class i by the inhibitory receptor lilrb1 suppresses macrophages and is a target of cancer immunotherapy. *Nature immunology*, 19(1):76–84, 2018.

[5] Siddhartha Jaiswal, Catriona HM Jamieson, Wendy W Pang, Christopher Y Park, Mark P Chao, Ravindra Majeti, David Traver, Nico van Rooijen, and Irving L Weissman. Cd47 is upregulated on circulating hematopoietic stem cells and leukemia cells to avoid phagocytosis. *Cell*, 138(2):271–285, 2009.

[6] Xing-hai Chen, Yong-jie Yin, and Jing-xiao Zhang. Sepsis and immune response. *World journal of emergency medicine*, 2(2):88, 2011.

[7] Akinori Takaoka, ZhiChao Wang, Myoung Kwon Choi, Hideyuki Yanai, Hideo Negishi, Tatsuma Ban, Yan Lu, Makoto Miyagishi, Tatsuhiko Kodama, Kenya Honda, Yusuke Ohba, and Tadatsugu Taniguchi. DAI (DLM-1/ZBP1) is a cytosolic DNA sensor and an activator of innate immune response. *Nature*, 448(7152):501–505, jul 2007.

[8] Roger C Bone. Why sepsis trials fail. *Jama*, 276(7):565–566, 1996.

[9] How is sepsis diagnosed and treated? — CDC, https://www.cdc.gov/sepsis/diagnosis/index.html.

[10] Kelly M. Hatfield, Raymund B. Dantes, James Baggs, Mathew R. P. Sapiano, Anthony E. Fiore, John A. Jernigan, and Lauren Epstein. Assessing Variability in Hospital-Level Mortality Among U.S. Medicare Beneficiaries With Hospitalizations for Severe Sepsis and Septic Shock*. *Critical Care Medicine*, 46(11):1753–1760, nov 2018.

[11] Hernando Gómez, John A. Kellum, and Claudio Ronco. Metabolic reprogramming and tolerance during sepsis-induced AKI. *Nature Reviews Nephrology*, 13(3):143–151, mar 2017.

[12] David Bar-Or, Matthew Carrick, Allen Tanner II, Mark J Lieser, Leonard T Rael, and Edward Brody. Overcoming the warburg effect: is it the key to survival in sepsis? *Journal of Critical Care*, 43:197–201, 2018.

[13] Kui Jin, Yujie Ma, Carlos L Manrique-Caballero, Hui Li, David R Emlet, Shengnan Li, Catherine J Baty, Xiaoyan Wen, Nahmah Kim-Campbell, Alicia Frank, et al. Activation of amp-activated protein kinase during sepsis/inflammation improves survival by preserving cellular metabolic fitness. *The FASEB Journal*, 34(5):7036–7057, 2020.

[14] Matthew G Vander Heiden, Lewis C Cantley, and Craig B Thompson. Understanding the Warburg effect: the metabolic requirements of cell proliferation. *Science (New York, N.Y.)*, 324(5930):1029–33, may 2009.

[15] Joshua A Smith, L Jay Stallons, and Rick G Schnellmann. Renal cortical hexokinase and pentose phosphate pathway activation through the egfr/akt signaling pathway in endotoxin-induced acute kidney injury. *American Journal of Physiology-Renal Physiology*, 307(4):F435–F444, 2014.

[16] Otto Warburg, Franz Wind, and Erwin Negelein. The metabolism of tumors in the body. *The Journal of general physiology*, 8(6):519, 1927.

[17] Rob JW Arts, Mark S Gresnigt, Leo AB Joosten, and Mihai G Netea. Cellular metabolism of myeloid cells in sepsis. *Journal of Leukocyte Biology*, 101(1):151–164, 2017.

[18] Patricia Resa-Infante, Núria Jorba, Rocio Coloma, and Juan Ortín. The influenza virus rna synthesis machine: advances in its structure and function. *RNA biology*, 8(2):207–215, 2011.

117

[19] Mitsutoshi Yoneyama, Mika Kikuchi, Takashi Natsukawa, Noriaki Shinobu, Tadaatsu Imaizumi, Makoto Miyagishi, Kazunari Taira, Shizuo Akira, and Takashi Fujita. The RNA helicase RIG-I has an essential function in double-stranded RNA-induced innate antiviral responses. *Nature Immunology*, 5(7):730–737, jul 2004.

[20] Shizuo Akira and Hiroaki Hemmi. Recognition of pathogen-associated molecular patterns by TLR family. *Immunology Letters*, 85(2):85–95, jan 2003.

[21] Alina Baum and Adolfo García-Sastre. Induction of type i interferon by rna viruses: cellular receptors and their substrates. *Amino acids*, 38(5):1283–1299, 2010.

[22] Daniel B. Stetson and Ruslan Medzhitov. Recognition of Cytosolic DNA Activates an IRF3-Dependent Innate Immune Response. *Immunity*, 24(1):93–103, jan 2006.

[23] Leonie Unterholzner. The interferon response to intracellular DNA: Why so many receptors? *Immunobiology*, 218(11):1312–1321, nov 2013.

[24] L. Sun, J. Wu, F. Du, X. Chen, and ZJ J. Chen. Cyclic GMP-AMP synthase is a cytosolic DNA sensor that activates the type I interferon pathway. *Science*, 339(6121):786–791, feb 2013.

[25] Nicole Dobbs, Nikolay Burnaevskiy, Didi Chen, Vijay K. Gonugunta, Neal M. Alto, and Nan Yan. STING Activation by Translocation from the ER Is Associated with Infection and Autoinflammatory Disease. *Cell Host & Microbe*, 18(2):157–168, 2015.

[26] Seng-Ryong Woo, Mercedes B. Fuertes, Leticia Corrales, Stefani Spranger, Michael J. Furdyna, Michael Y.K. Leung, Ryan Duggan, Ying Wang, Glen N. Barber, Katherine A. Fitzgerald, Maria-Luisa Alegre, and Thomas F. Gajewski. STING-Dependent Cytosolic DNA Sensing Mediates Innate Immune Recognition of Immunogenic Tumors. *Immunity*, 41(5):830–842, 2014.

[27] Liufu Deng, Hua Liang, Meng Xu, Xuanming Yang, Byron Burnette, Ainhoa Arina, Xiao-Dong Li, Helena Mauceri, Michael Beckett, Thomas Darga, et al. Sting-dependent cytosolic dna sensing promotes radiation-induced type i interferon-dependent antitumor immunity in immunogenic tumors. *Immunity*, 41(5):843–852, 2014.

[28] Kevin W. Ng, Erin A. Marshall, John C. Bell, and Wan L. Lam. cGAS-STING and Cancer: Dichotomous Roles in Tumor Immunity and Development. *Trends in immunology*, 39(1):44–54, jan 2018.

[29] Qi Chen, Lijun Sun, and Zhijian J Chen. Regulation and function of the cGAS–STING pathway of cytosolic DNA sensing. *Nature Immunology*, 17(10):1142–1149, sep 2016.

[30] Jie An, Laura Durcan, Reynold M. Karr, Tracy A. Briggs, Gillian I. Rice, Thomas H. Teal, Joshua J. Woodward, and Keith B. Elkon. Expression of Cyclic GMP-AMP Synthase in Patients With Systemic Lupus Erythematosus. *Arthritis & Rheumatology*, 69(4):800–807, apr 2017.

[31] Yanick J. Crow et al. Mutations in the gene encoding the 3'-5' dna exonuclease trex1 cause aicardi-goutières syndrome at the ags1 locus. *Nature Genetics*, 38:917, Jul 2006.

[32] Hui Yang, Hanze Wang, Junyao Ren, Qi Chen, and Zhijian J. Chen. cgas is essential for cellular senescence. 114(23):E4612–E4620, jun 2017.

[33] Zhixun Dou, Kanad Ghosh, Maria Grazia Vizioli, Jiajun Zhu, Payel Sen, Kirk J Wangensteen, Johayra Simithy, Yemin Lan, Yanping Lin, Zhuo Zhou, et al. Cytoplasmic chromatin triggers inflammation in senescence and cancer. *Nature*, 550(7676):402–406, 2017.

[34] Qiujing Yu, Yuliya V. Katlinskaya, Christopher J. Carbone, Bin Zhao, Kanstantsin V. Katlinski, Hui Zheng, Manti Guha, Ning Li, Qijun Chen, Ting Yang, Christopher J. Lengner, Roger A. Greenberg, F. Brad Johnson, and Serge Y. Fuchs. DNA-Damage-Induced Type I Interferon Promotes Senescence and Inhibits Stem Cell Function. *Cell Reports*, 11(5):785–797, may 2015.

[35] Nicholas PD Liau, Artem Laktyushin, Isabelle S Lucet, James M Murphy, Shenggen Yao, Eden Whitlock, Kimberley Callaghan, Nicos A Nicola, Nadia J Kershaw, and Jeffrey J Babon. The molecular basis of jak/stat inhibition by socs1. *Nature communications*, 9(1):1–14, 2018.

[36] Kenya Honda, Akinori Takaoka, and Tadatsugu Taniguchi. Type I Inteferon Gene Induction by the Interferon Regulatory Factor Family of Transcription Factors. *Immunity*, 25(3):349–360, sep 2006.

[37] Mingwei Zhao, Jiangwen Zhang, Hemali Phatnani, Stefanie Scheu, and Tom Maniatis. Stochastic Expression of the Interferon-β Gene. *PLoS Biology*, 10(1):e1001249, jan 2012.

[38] Hiroaki Kitano. Computational systems biology. *Nature*, 420(6912):206–210, nov 2002.

[39] Simon Mitchell, Jesse Vargas, and Alexander Hoffmann. Signaling via the nfκb system. *Wiley Interdisciplinary Reviews: Systems Biology and Medicine*, 8(3):227–241, 2016.

[40] Alexander Hoffmann, Andre Levchenko, Martin L Scott, and David Baltimore. The IkappaB-NF-kappaB signaling module: temporal control and selective gene activation. *Science (New York, N.Y.)*, 298(5596):1241–5, nov 2002.

[41] John Charles Butcher. *Numerical methods for ordinary differential equations*. John Wiley & Sons, 2016.

[42] David F Griffiths and Desmond J Higham. *Numerical methods for ordinary differential equations: initial value problems*. Springer Science & Business Media, 2010.

[43] Benjamin Ballnus, Sabine Hug, Kathrin Hatz, Linus Görlitz, Jan Hasenauer, and Fabian J Theis. Comprehensive benchmarking of Markov chain Monte Carlo methods for dynamical systems. *BMC Systems Biology*, 1163(11), 2017.

[44] G. O. Roberts, A. Gelman, and W. R. Gilks. Weak convergence and optimal scaling of random walk Metropolis algorithms. *The Annals of Applied Probability*, 7(1):110–120, feb 1997.

[45] Andrew Gelman and Donald B. Rubin. Inference from Iterative Simulation Using Multiple Sequences. *Statistical Science*, 7(4):457–472, nov 1992.

[46] René David and Hassane Alla. *Discrete, continuous, and hybrid Petri nets*, volume 1. Springer, 2010.

[47] René Thomas and Richard d'Ari. *Biological feedback*. CRC press, 1990.

[48] Roy M Anderson. Discussion: the kermack-mckendrick epidemic threshold theorem. *Bulletin of mathematical biology*, 53(1-2):1, 1991.

[49] Zhihui Wang, Joseph D. Butner, Romica Kerketta, Vittorio Cristini, and Thomas S. Deisboeck. Simulating cancer growth with multiscale agent-based modeling. *Seminars in Cancer Biology*, 30:70–78, feb 2015.

[50] Mahua Roy and Stacey D. Finley. Metabolic reprogramming dynamics in tumor spheroids: Insights from a multicellular, multiscale model. *PLOS Computational Biology*, 15(6):e1007053, jun 2019.

[51] Axel Krinner, Ingo Roeder, Markus Loeffler, and Markus Scholz. Merging concepts - coupling an agent-based model of hematopoietic stem cells with an ODE model of granulopoiesis. *BMC Systems Biology*, 7(1):117, nov 2013.

[52] U. WILENSKY. NetLogo. Center for Connected Learning and Computer-Based Modeling. *Northwestern University, Evanston, IL.*, 1999.

[53] Alan C. Hindmarsh, Peter N. Brown, Keith E. Grant, Steven L. Lee, Radu Serban, Dan E. Shumaker, and Carol S. Woodward. SUNDIALS: Suite of nonlinear and differential/algebraic equation solvers. *ACM Transactions on Mathematical Software*, 31(3):363–396, 2005.

[54] Christopher Rackauckas and Qing Nie. DifferentialEquations.jl – A Performant and Feature-Rich Ecosystem for Solving Differential Equations in Julia. *Journal of Open Research Software*, 5, may 2017.

[55] Osamu Takasu, Joseph P. Gaut, Eizo Watanabe, Kathleen To, R. Eliot Fagley, Brian Sato, Steve Jarman, Igor R. Efimov, Deborah L. Janks, Anil Srivastava, Sam B. Bhayani, Anne Drewry, Paul E. Swanson, and Richard S. Hotchkiss. Mechanisms of Cardiac and Renal Dysfunction in Patients Dying of Sepsis. *American Journal of Respiratory and Critical Care Medicine*, 187(5):509–517, mar 2013.

[56] Daniel Rittirsch, L. Marco Hoesel, and Peter A. Ward. The disconnect between animal models of sepsis and human sepsis. *Journal of Leukocyte Biology*, 81(1):137–143, jan 2007.

[57] Luiz F. Poli-de Figueiredo, Alejandra G. Garrido, Naomi Nakagawa, and Paulina Sannomiya. EXPERIMENTAL MODELS OF SEPSIS AND THEIR CLINICAL RELEVANCE. *Shock*, 30(Suppl 1):53–59, oct 2008.

[58] H. Mayer, R.N. Tharanathan, and J. Weckesser. 6 Analysis of Lipopolysaccharides of Gram-Negative Bacteria. *Methods in Microbiology*, 18:157–207, jan 1985.

[59] A Poltorak, X He, I Smirnova, M Y Liu, C Van Huffel, X Du, D Birdwell, E Alejos, M Silva, C Galanos, M Freudenberg, P Ricciardi-Castagnoli, B Layton, and B Beutler. Defective LPS signaling in C3H/HeJ and C57BL/10ScCr mice: mutations in Tlr4 gene. *Science (New York, N.Y.)*, 282(5396):2085–8, dec 1998.

[60] S Gessani, F Belardelli, A Pecorelli, P Puddu, and C Baglioni. Bacterial lipopolysaccharide and gamma interferon induce transcription of beta interferon mRNA and

interferon secretion in murine macrophages. *Journal of Virology*, 63(6):2785–2789, jun 1989.

[61] N M Zirk, S F Hashmi, and H K Ziegler. The polysaccharide portion of lipopolysaccharide regulates antigen-specific T-cell activation via effects on macrophage-mediated antigen processing. *Infection and immunity*, 67(1):319–26, jan 1999.

[62] Alex J Freemerman, Amy R Johnson, Gina N Sacks, J Justin Milner, Erin L Kirk, Melissa A Troester, Andrew N Macintyre, Pankuri Goraksha-Hicks, Jeffery C Rathmell, and Liza Makowski. Metabolic reprogramming of macrophages: glucose transporter 1 (GLUT1)-mediated glucose metabolism drives a proinflammatory phenotype. *The Journal of biological chemistry*, 289(11):7884–96, mar 2014.

[63] Connie M Krawczyk, Thomas Holowka, Jie Sun, Julianna Blagih, Eyal Amiel, Ralph J DeBerardinis, Justin R Cross, Euihye Jung, Craig B Thompson, Russell G Jones, and Edward J Pearce. Toll-like receptor-induced changes in glycolytic metabolism regulate dendritic cell activation. *Blood*, 115(23):4742–9, jun 2010.

[64] Lise Van Wyngene, Jolien Vandewalle, and Claude Libert. Reprogramming of basic metabolic pathways in microbial sepsis: therapeutic targets at last? *EMBO molecular medicine*, 10(8):e8712, jul 2018.

[65] Róisín M Loftus and David K Finlay. Immunometabolism: Cellular Metabolism Turns Immune Regulator. *The Journal of biological chemistry*, 291(1):1–10, jan 2016.

[66] Luke A. J. O'Neill, Rigel J. Kishton, and Jeff Rathmell. A guide to immunometabolism for immunologists. *Nature Reviews Immunology*, 16(9):553–565, sep 2016.

[67] Maria V. Liberti and Jason W. Locasale. The Warburg Effect: How Does it Benefit Cancer Cells?, 2016.

[68] Paul Waltz, Evie Carchman, Hernando Gomez, and Brian Zuckerbraun. Sepsis results in an altered renal metabolic and osmolyte profile. *Journal of Surgical Research*, 202(1):8–12, may 2016.

[69] Zhenzhen Shi, Chih-Hang J. Wu, David Ben-Arieh, and Steven Q. Simpson. Mathematical Model of Innate and Adaptive Immunity of Sepsis: A Modeling and Simulation Study of Infectious Disease. *BioMed Research International*, 2015:1–31, sep 2015.

[70] Yukihiro Yamanaka, Kenko Uchida, Momoka Akashi, Yuta Watanabe, Arino Yaguchi, Shuji Shimamoto, Shingo Shimoda, Hitoshi Yamada, Masashi Yamashita, and Hidenori Kimura. Mathematical modeling of septic shock based on clinical data. *Theoretical Biology and Medical Modelling*, 16(1):5, dec 2019.

[71] Filippo Mariano, Vincenzo Cantaluppi, Maurizio Stella, Giuseppe Romanazzi, Barbara Assenzio, Monica Cairo, Luigi Biancone, Giorgio Triolo, V Marco Ranieri, and Giovanni Camussi. Circulating plasma factors induce tubular and glomerular alterations in septic burns patients. *Critical Care*, 12(2):R42, mar 2008.

[72] Tie Fu Liu, Vidula T Vachharajani, Barbara K Yoza, and Charles E McCall. NAD+-dependent sirtuin 1 and 6 proteins coordinate a switch from glucose to fatty acid oxidation during the acute inflammatory response. *The Journal of biological chemistry*, 287(31):25758–69, jul 2012.

[73] Anisur Rahman and Jason M Haugh. Kinetic Modeling and Analysis of the Akt/Mechanistic Target of Rapamycin Complex 1 (mTORC1) Signaling Axis Reveals Cooperative, Feedforward Regulation. *The Journal of biological chemistry*, 292(7):2866–2872, feb 2017.

[74] Nurgazy Sulaimanov, Martin Klose, Hauke Busch, and Melanie Boerries. Understanding the mTOR signaling pathway via mathematical modeling. *Wiley interdisciplinary reviews. Systems biology and medicine*, 9(4), 2017.

[75] Thawfeek M Varusai and Lan K Nguyen. Dynamic modelling of the mTOR signalling network reveals complex emergent behaviours conferred by DEPTOR. *Scientific reports*, 8(1):643, 2018.

[76] Yangjin Kim, Gibin Powathil, Hyunji Kang, Dumitru Trucu, Hyeongi Kim, Sean Lawler, and Mark Chaplain. Strategies of Eradicating Glioma Cells: A Multi-Scale Mathematical Model with miR-451-AMPK-mTOR Control. 2015.

[77] Bhanu Chandra Mulukutla, Andrew Yongky, Prodromos Daoutidis, and Wei-Shou Hu. Bistability in Glycolysis Pathway as a Physiological Switch in Energy Metabolism. *PLoS ONE*, 9(6):e98756, jun 2014.

[78] Margit Heiske, Thierry Letellier, and Edda Klipp. Comprehensive mathematical model of oxidative phosphorylation valid for physiological and pathological conditions. *FEBS Journal*, 284(17):2802–2828, sep 2017.

[79] Rong Ke, Qicao Xu, Cong Li, Lingyu Luo, and Deqiang Huang. Mechanisms of AMPK in the maintenance of ATP balance during energy metabolism, apr 2018.

[80] Maria M Mihaylova and Reuben J Shaw. The ampk signalling pathway coordinates cell growth, autophagy and metabolism. *Nature cell biology*, 13(9):1016–1023, 2011.

[81] Ken Inoki, Tianqing Zhu, and Kun-Liang Guan. Tsc2 mediates cellular energy response to control cell growth and survival. *Cell*, 115(5):577–590, 2003.

[82] Stephen C. Land and Andrew R. Tee. Hypoxia-inducible factor 1α is regulated by the mammalian target of rapamycin (mTOR) via an mTOR signaling motif. *Journal of Biological Chemistry*, 282(28):20534–20543, jul 2007.

[83] Haijuan Yang, Xiaolu Jiang, Buren Li, Hyo J Yang, Meredith Miller, Angela Yang, Ankita Dhar, and Nikola P Pavletich. Mechanisms of mtorc1 activation by rheb and inhibition by pras40. *Nature*, 552(7685):368–373, 2017.

[84] Joungmok Kim and Kun-Liang Guan. mtor as a central hub of nutrient signalling and cell growth. *Nature Cell Biology*, 21(1):63–71, 2019.

[85] Chi Ying F. Huang and James E. Ferrell. Ultrasensitivity in the mitogen-activated protein kinase cascade. *Proceedings of the National Academy of Sciences of the United States of America*, 93(19):10078–10083, sep 1996.

[86] Ayako Nagao, Minoru Kobayashi, Sho Koyasu, Christalle CT Chow, and Hiroshi Harada. Hif-1-dependent reprogramming of glucose metabolic pathway of cancer cells and its therapeutic significance. *International journal of molecular sciences*, 20(2):238, 2019.

[87] Gregg L Semenza, Peter H Roth, Hon-Ming Fang, and Guang L Wang. Transcriptional regulation of genes encoding glycolytic enzymes by hypoxia-inducible factor 1. *Journal of biological chemistry*, 269(38):23757–23763, 1994.

[88] Qi Li, Chenyi Wu, Zhenlong Liu, Huiqing Zhang, Yuna Du, Yuxiang Liu, Kuangyu Song, Qiaofa Shi, and Rong Li. Increased tlr4 expression aggravates sepsis by promoting ifn-γ expression in cd38-/- mice. *Journal of Immunology Research*, 2019, 2019.

[89] James P Luyendyk, Gernot A Schabbauer, Michael Tencati, Todd Holscher, Rafal Pawlinski, and Nigel Mackman. Genetic analysis of the role of the pi3k-akt pathway

in lipopolysaccharide-induced cytokine and tissue factor gene expression in mono-cytes/macrophages. *The Journal of Immunology*, 180(6):4218–4226, 2008.

[90] Hung Q Doan, Kanika A Bowen, Lindsey A Jackson, and B Mark Evers. Toll-like receptor 4 activation increases akt phosphorylation in colon cancer cells. *Anticancer research*, 29(7):2473–2478, 2009.

[91] Fei Xu, Lixin Na, Yanfei Li, and Linjun Chen. Roles of the pi3k/akt/mtor signalling pathways in neurodegenerative diseases and tumours. *Cell & Bioscience*, 10:1–12, 2020.

[92] Lei Zhong, Agustina D'Urso, Debra Toiber, Carlos Sebastian, Ryan E Henry, Douang-sone D Vadysirisack, Alexander Guimaraes, Brett Marinelli, Jakob D Wikstrom, Tomer Nir, et al. The histone deacetylase sirt6 regulates glucose homeostasis via hif1α. *Cell*, 140(2):280–293, 2010.

[93] Jean-Marc Tadie, Hong-Beom Bae, Jessy S Deshane, Celeste P Bell, Eduardo R Lazarowski, David D Chaplin, Victor J Thannickal, Edward Abraham, and Jaroslaw W Zmijewski. Toll-like receptor 4 engagement inhibits adenosine 5-monophosphate-activated protein kinase activation through a high mobility group box 1 protein-dependent mechanism. *Molecular medicine*, 18(4):659–668, 2012.

[94] Shona A Mookerjee, Akos A Gerencser, David G Nicholls, and Martin D Brand. Quantifying intracellular rates of glycolytic and oxidative atp production and con-sumption using extracellular flux measurements. *Journal of Biological Chemistry*, 292(17):7189–7207, 2017.

[95] Bryson D Bennett, Elizabeth H Kimball, Melissa Gao, Robin Osterhout, Stephen J Van Dien, and Joshua D Rabinowitz. Absolute metabolite concentrations and implied enzyme active site occupancy in escherichia coli. *Nature chemical biology*, 5(8):593–599, 2009.

[96] Ron Milo and Rob Phillips. *Cell Biology by the Numbers*. Garland Science, 1st edition, 2015.

[97] X-Y Zhang, MN Trame, LJ Lesko, and S Schmidt. Sobol sensitivity analysis: a tool to guide the development and evaluation of systems pharmacology models. *CPT: pharmacometrics & systems pharmacology*, 4(2):69–79, 2015.

[98] Julia M. Corton, John G. Gillespie, and D.Grahame Hardie. Role of the AMP-activated protein kinase in the cellular stress response. *Current Biology*, 4(4):315–324, apr 1994.

[99] Derek C Angus, Walter T Linde-Zwirble, Jeffrey Lidicker, Gilles Clermont, Joseph Carcillo, and Michael R Pinsky. Epidemiology of severe sepsis in the united states: analysis of incidence, outcome, and associated costs of care. *Read Online: Critical Care Medicine— Society of Critical Care Medicine*, 29(7):1303–1310, 2001.

[100] Pengyan Xia, Buqing Ye, Shuo Wang, Xiaoxiao Zhu, Ying Du, Zhen Xiong, Yong Tian, and Zusen Fan. Glutamylation of the DNA sensor cGAS regulates its binding and synthase activity in antiviral immunity. *Nature Immunology*, 17(4):369–378, feb 2016.

[101] Angela C. Collins, Haocheng Cai, Tuo Li, Luis H. Franco, Xiao-Dong Li, Vidhya R. Nair, Caitlyn R. Scharn, Chelsea E. Stamm, Beth Levine, Zhijian J. Chen, and Michael U. Shiloh. Cyclic GMP-AMP Synthase Is an Innate Immune DNA Sensor for Mycobacterium tuberculosis. *Cell Host & Microbe*, 17(6):820–828, jun 2015.

[102] Pu Gao, Manuel Ascano, Yang Wu, Winfried Barchet, Barbara L. Gaffney, Thomas Zillinger, Artem A. Serganov, Yizhou Liu, Roger A. Jones, Gunther Hartmann, Thomas Tuschl, and Dinshaw J. Patel. Cyclic [G(2,5)pA(3,5)p] Is the Metazoan Second Messenger Produced by DNA-Activated Cyclic GMP-AMP Synthase. *Cell*, 153(5):1094–1107, may 2013.

[103] Iliana Georgana, Rebecca P Sumner, Greg J Towers, and Carlos Maluquer de Motes. Virulent poxviruses inhibit DNA sensing by preventing STING activation. *Journal of virology*, pages JVI.02145–17, feb 2018.

[104] K. Hansen, T. Prabakaran, A. Laustsen, S. E. Jorgensen, S. H. Rahbaek, S. B. Jensen, R. Nielsen, J. H. Leber, T. Decker, K. A. Horan, M. R. Jakobsen, and S. R. Paludan. Listeria monocytogenes induces IFN expression through an IFI16-, cGAS- and STING-dependent pathway. *The EMBO Journal*, 33(15):1654–1666, aug 2014.

[105] Eric Lam, Saskia Stein, and Erik Falck-Pedersen. Adenovirus detection by the cGAS/STING/TBK1 DNA sensing cascade. *Journal of virology*, 88(2):974–81, jan 2014.

[106] Zhe Ma and Blossom Damania. The cGAS-STING Defense Pathway and Its Counteraction by Viruses. *Cell host & microbe*, 19(2):150–8, feb 2016.

[107] Aleksey Novikov, Marco Cardone, Robert Thompson, Kevin Shenderov, Kevin D. Kirschman, Katrin D. Mayer-Barber, Timothy G. Myers, Ronald L. Rabin, Giorgio Trinchieri, Alan Sher, and Carl G. Feng. Mycobacterium tuberculosis Triggers Host Type I IFN Signaling To Regulate IL-1β Production in Human Macrophages. *The Journal of Immunology*, 187(5), 2011.

[108] Chenhe Su and Chunfu Zheng. Herpes Simplex Virus 1 Abrogates the cGAS/STING-Mediated Cytosolic DNA-Sensing Pathway via Its Virion Host Shutoff Protein, UL41. *Journal of virology*, 91(6):e02414–16, mar 2017.

[109] Glen N Barber. STING: infection, inflammation and cancer. *Nature Publishing Group*, 15, 2015.

[110] Elizabeth E. Gray, Piper M. Treuting, Joshua J. Woodward, and Daniel B. Stetson. Cutting Edge: cGAS Is Required for Lethal Autoimmune Disease in the Trex1-Deficient Mouse Model of Aicardi–Goutières Syndrome. *The Journal of Immunology*, 195(5), 2015.

[111] Shigekazu Nagata and Kohki Kawane. Autoinflammation by Endogenous DNA. *Advances in Immunology*, 110:139–161, jan 2011.

[112] Sabrina Brzostek-Racine, Chris Gordon, Sarah Van Scoy, and Nancy C Reich. The DNA damage response induces IFN. *Journal of immunology (Baltimore, Md. : 1950)*, 187(10):5336–45, nov 2011.

[113] Anetta Härtlova, Saskia F. Erttmann, Faizal AM. Raffi, Anja M. Schmalz, Ulrike Resch, Sharath Anugula, Stefan Lienenklaus, Lisa M. Nilsson, Andrea Kröger, Jonas A. Nilsson, Torben Ek, Siegfried Weiss, and Nelson O. Gekara. DNA Damage Primes the Type I Interferon System via the Cytosolic DNA Sensor STING to Promote Anti-Microbial Innate Immunity. *Immunity*, 42(2):332–343, feb 2015.

[114] Yukiko Matsuoka, Hiromi Matsumae, Manami Katoh, Amie J Eisfeld, Gabriele Neumann, Takeshi Hase, Samik Ghosh, Jason E Shoemaker, Tiago J S Lopes, Tokiko Watanabe, Shinji Watanabe, Satoshi Fukuyama, Hiroaki Kitano, and Yoshihiro Kawaoka. A comprehensive map of the influenza A virus replication cycle. *BMC systems biology*, 7(1):97, jan 2013.

[115] K Oda, T Kimura, Y Matsuoka, A Funahashi, M Muramatsu, and H Kitano. Molecular interaction map of a macrophage. *AfCS Research Reports*, 2(14), 2004.

[116] Gabriele Lillacci and Mustafa Khammash. Parameter estimation and model selection in computational biology. *PLoS computational biology*, 6(3):e1000696, mar 2010.

[117] Hiroaki Kitano. Biological robustness. *Nature Reviews Genetics*, 5(11):826–837, nov 2004.

[118] Birgit Schoeberl, Claudia Eichler-Jonsson, Ernst Dieter Gilles, and Gertraud Müller. Computational modeling of the dynamics of the MAP kinase cascade activated by surface and internalized EGF receptors. *Nature Biotechnology*, 20(4):370–375, apr 2002.

[119] Jason E Shoemaker, Kalyan Gayen, Natàlia Garcia-Reyero, Edward J Perkins, Daniel L Villeneuve, Li Liu, and Francis J Doyle. Fathead minnow steroidogenesis: in silico analyses reveals tradeoffs between nominal target efficacy and robustness to cross-talk. *BMC Systems Biology*, 4(1):89, jun 2010.

[120] Neda Bagheri, Jörg Stelling, and Francis J. Doyle. Quantitative performance metrics for robustness in circadian rhythms. *Bioinformatics*, 23(3):358–364, feb 2007.

[121] Hisaaki Shinohara, Marcelo Behar, Kentaro Inoue, Michio Hiroshima, Tomoharu Yasuda, Takeshi Nagashima, Shuhei Kimura, Hideki Sanjo, Shiori Maeda, Noriko Yumoto, Sewon Ki, Shizuo Akira, Yasushi Sako, Alexander Hoffmann, Tomohiro Kurosaki, and Mariko Okada-Hatakeyama. Positive Feedback Within a Kinase Signaling Complex Functions as a Switch Mechanism for NF-κB Activation. *Science*, 344(6185), 2014.

[122] Liang Qiao, Hannah Phipps-Yonas, Boris Hartmann, Thomas M Moran, Stuart C Sealfon, and Fernand Hayot. Immune response modeling of interferon beta-pretreated influenza virus-infected human dendritic cells. *Biophysical journal*, 98(4):505–14, feb 2010.

[123] Selma Pereira-Lopes, Teja Celhar, Gloria Sans-Fons, Maria Serra, Anna-Marie Fairhurst, Jorge Lloberas, and Antonio Celada. The exonuclease Trex1 restrains macrophage proinflammatory activation. *Journal of immunology (Baltimore, Md. : 1950)*, 191(12):6128–35, dec 2013.

[124] Jeffery S Cox, Sarah A Stanley, James E Johndrow, and Paolo Manzanillo. The Type I IFN Response to Infection with The Type I IFN Response to Infection with Mycobacterium tuberculosis Requires ESX-1-Mediated Secretion and Contributes to Pathogenesis. *J Immunol References The Journal of Immunology*, 178:3143–3152, 2017.

[125] Xin Guo, Chang Shu, Honggui Li, Ya Pei, Shih-Lung Woo, Juan Zheng, Mengyang Liu, Hang Xu, Rachel Botchlett, Ting Guo, Yuli Cai, Xinsheng Gao, Jing Zhou, Lu Chen, Qifu Li, Xiaoqiu Xiao, Linglin Xie, Ke K. Zhang, Jun-Yuan Ji, Yuqing Huo, Fanyin Meng, Gianfranco Alpini, Pingwei Li, and Chaodong Wu. Cyclic GMP-AMP Ameliorates Diet-induced Metabolic Dysregulation and Regulates Proinflammatory Responses Distinctly from STING Activation. *Scientific Reports*, 7(1):6355, dec 2017.

[126] K. L. Jønsson, A. Laustsen, C. Krapp, K. A. Skipper, K. Thavachelvam, D. Hotter, J. H. Egedal, M. Kjolby, P. Mohammadi, T. Prabakaran, L. K. Sørensen, C. Sun, S. B. Jensen, C. K. Holm, R. J. Lebbink, M. Johannsen, M. Nyegaard, J. G. Mikkelsen, F. Kirchhoff, S. R. Paludan, and M. R. Jakobsen. IFI16 is required for DNA sensing in human macrophages by promoting production and function of cGAMP. *Nature Communications*, 8:14391, feb 2017.

[127] Gernot Posselt, Harald Schwarz, Albert Duschl, and Jutta Horejs-Hoeck. Suppressor of cytokine signaling 2 is a feedback inhibitor of TLR-induced activation in human monocyte-derived dendritic cells. *Journal of immunology (Baltimore, Md. : 1950)*, 187(6):2875–84, sep 2011.

[128] Jing Qing, Cheng Liu, Lisa Choy, Rui-Yun Wu, Joseph S Pagano, and Rik Derynck. Transforming growth factor beta/Smad3 signaling regulates IRF-7 function and transcriptional activation of the beta interferon promoter. *Molecular and cellular biology*, 24(3):1411–25, feb 2004.

[129] Qiang Wang, Liyuan Huang, Ze Hong, Zhongshi Lv, Zhaomin Mao, Yijun Tang, Xiufang Kong, Senlin Li, Ye Cui, Heng Liu, Lele Zhang, Xiaojie Zhang, Lindi Jiang, Chen Wang, and Qin Zhou. The E3 ubiquitin ligase RNF185 facilitates the cGAS-mediated innate immune response. *PLOS Pathogens*, 13(3):e1006264, mar 2017.

[130] Liudmila Andreeva, Björn Hiller, Dirk Kostrewa, Charlotte Lässig, Carina C. de Oliveira Mann, David Jan Drexler, Andreas Maiser, Moritz Gaidt, Heinrich Leonhardt, Veit Hornung, and Karl-Peter Hopfner. cGAS senses long and HMGB/TFAM-bound U-turn DNA by forming protein–DNA ladders. *Nature*, 549(7672):394–398, sep 2017.

[131] Philip J. Kranzusch, Amy Si-Ying Lee, James M. Berger, and Jennifer A. Doudna. Structure of Human cGAS Reveals a Conserved Family of Second-Messenger Enzymes in Innate Immunity. may 2013.

[132] Tomoki Yoshida, Akira Kakizuka, and Hiromi Imamura. BTeam, a Novel BRET-based Biosensor for the Accurate Quantification of ATP Concentration within Living Cells. *Scientific Reports*, 6(1):39618, dec 2016.

[133] Jiaxi Wu, Lijun Sun, Xiang Chen, Fenghe Du, Heping Shi, Chuo Chen, and Zhi-jian J Chen. Cyclic GMP-AMP is an endogenous second messenger in innate immune signaling by cytosolic DNA. *Science (New York, N.Y.)*, 339(6121):826–30, feb 2013.

[134] Wei Zhang and Xiufen Zou. Systematic Analysis of the Mechanisms of Virus-Triggered Type I IFN Signaling Pathways through Mathematical Modeling. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, 10(3):771–779, may 2013.

[135] Sidney B. Cambridge, Florian Gnad, Chuong Nguyen, Justo Lorenzo Bermejo, Marcus Krüger, and Matthias Mann. Systems-wide Proteomic Analysis in Mammalian Cells Reveals Conserved, Functional Protein Turnover. *Journal of Proteome Research*, 10(12):5275–5284, dec 2011.

[136] Björn Schwanhäusser, Dorothea Busse, Na Li, Gunnar Dittmar, Johannes Schuchardt, Jana Wolf, Wei Chen, and Matthias Selbach. Global quantification of mammalian gene expression control. *Nature*, 473(7347):337–342, may 2011.

[137] L. V. Sharova, A. A. Sharov, T. Nedorezov, Y. Piao, N. Shaik, and M. S.H. Ko. Database for mRNA Half-Life of 19 977 Genes Obtained by DNA Microarray Analysis of Pluripotent and Differentiating Mouse Embryonic Stem Cells. *DNA Research*, 16(1):45–58, jan 2009.

[138] Katrin Peschke, Martin Achleitner, Kathrin Frenzel, Alexander Gerbaulet, Servi Remzi Ada, Nicolas Zeller, Stefan Lienenklaus, Mathias Lesche, Claire Poulet, Ronald Naumann, Andreas Dahl, Ursula Ravens, Claudia Günther, Werner Müller, Klaus-Peter Knobeloch, Marco Prinz, Axel Roers, and Rayk Behrendt. Loss of Trex1 in Dendritic Cells Is Sufficient To Trigger Systemic Autoimmunity. *Journal of immunology (Baltimore, Md. : 1950)*, 197(6):2157–66, sep 2016.

[139] A.-D. C. Chessler, L. R. P. Ferreira, T.-H. Chang, K. A. Fitzgerald, and B. A. Burleigh. A Novel IFN Regulatory Factor 3-Dependent Pathway Activated by Trypanosomes Triggers IFN- in Macrophages and Fibroblasts. *The Journal of Immunology*, 181(11):7917–7924, nov 2008.

[140] Martyn Plummer, Nicky Best, Kate Cowles, Karen Vines, Martyn Plummer, Nicky Best, Kate Cowles, and Karen Vines. CODA: Convergence Diagnosis and Output Analysis for MCMC. *R News*, 6:7–11, 2006.

[141] Kansuporn Sriyudthsak, Harumi Uno, Rudiyanto Gunawan, and Fumihide Shiraishi. Using dynamic sensitivities to characterize metabolic reaction systems. *Mathematical Biosciences*, 269:153–163, nov 2015.

[142] Jessica Vincent, Carolina Adura, Pu Gao, Antonio Luz, Lodoe Lama, Yasutomi Asano, Rei Okamoto, Toshihiro Imaeda, Jumpei Aida, Katherine Rothamel, Tasos Gogakos, Joshua Steinberg, Seth Reasoner, Kazuyoshi Aso, Thomas Tuschl, Dinshaw J. Patel, J. Fraser Glickman, and Manuel Ascano. Small molecule inhibition of cGAS reduces interferon expression in primary macrophages from autoimmune mice. *Nature Communications*, 8(1):750, dec 2017.

[143] Daxing Gao, Jiaxi Wu, You-Tong Y.-T. Wu, Fenghe Du, Chukwuemika Aroh, Nan Yan, Lijun Sun, and Zhijian J. Chen. Cyclic GMP-AMP synthase is an innate immune sensor of HIV and other retroviruses. *Science (New York, N.Y.)*, 341(6148):903–6, aug 2013.

[144] Sandra S. Diebold, Maria Montoya, Hermann Unger, Lena Alexopoulou, Polly Roy, Linsey E. Haswell, Aymen Al-Shamkhani, Richard Flavell, Persephone Borrow, and Caetano Reis e Sousa. Viral infection switches non-plasmacytoid dendritic cells into high interferon producers. *Nature*, 424(6946):324–328, jul 2003.

[145] Ulfert Rand, Melanie Rinas, Johannes Schwerk, Gesa Nö Hren, Melanie Linnes, Andrea Krö Ger, Michael Flossdorf, Kristó F Kály-Kullai, Hansjö Rg Hauser, Thomas Hö Fer, and Mario Kö Ster. Multi-layered stochasticity and paracrine signal propagation shape the type-I interferon response. *Molecular Systems Biology*, 8(10), 2012.

[146] Oana-Teodora Chis, Julio R. Banga, and Eva Balsa-Canto. Structural Identifiability of Systems Biology Models: A Critical Comparison of Methods. *PLoS ONE*, 6(11):e27755, nov 2011.

[147] Jörg Stelling, Ernst Dieter Gilles, and Francis J Doyle. Robustness properties of circadian clock architectures. *Proceedings of the National Academy of Sciences of the United States of America*, 101(36):13210–5, sep 2004.

[148] Sivan Leviyang and Igor Griva. Investigating functional roles for positive feedback and cellular heterogeneity in the type I interferon response to viral infection. *Viruses*, 10(10), oct 2018.

[149] Rahul Satija and Alex K. Shalek. Heterogeneity in immune responses: from populations to single cells. *Trends in Immunology*, 35(5):219–229, may 2014.

[150] Florian Wimmers, Nikita Subedi, Nicole van Buuringen, Daan Heister, Judith Vivié, Inge Beeren-Reinieren, Rob Woestenenk, Harry Dolstra, Aigars Piruska, Joannes F. M. Jacobs, Alexander van Oudenaarden, Carl G. Figdor, Wilhelm T. S. Huck, I. Jolanda M. de Vries, and Jurjen Tel. Single-cell analysis reveals that stochasticity and paracrine signaling control interferon-alpha production by plasmacytoid dendritic cells. *Nature Communications*, 9(1):3317, dec 2018.

[151] Peter S. Swain, Michael B. Elowitz, and Eric D. Siggia. Intrinsic and extrinsic contributions to stochasticity in gene expression. *Proceedings of the National Academy of Sciences*, 99(20):12795–12800, oct 2002.

[152] Michael B Elowitz, Arnold J Levine, Eric D Siggia, and Peter S Swain. Stochastic gene expression in a single cell. *Science (New York, N.Y.)*, 297(5584):1183–6, aug 2002.

[153] Arjun Raj and Alexander van Oudenaarden. Nature, Nurture, or Chance: Stochastic Gene Expression and Its Consequences, oct 2008.

[154] Arren Bar-Even, Johan Paulsson, Narendra Maheshri, Miri Carmi, Erin O'Shea, Yitzhak Pilpel, and Naama Barkai. Noise in protein expression scales with natural protein abundance. *Nature Genetics*, 38(6):636–643, jun 2006.

[155] M. B. Schulte and R. Andino. Single-Cell Analysis Uncovers Extensive Biological Noise in Poliovirus Replication. *Journal of Virology*, 88(11):6205–6212, jun 2014.

[156] Sun-Young Hwang, Kye-Yeon Hur, Jeong-Rae Kim, Kwang-Hyun Cho, Seung-Hwan Kim, and Joo-Yeon Yoo. Biphasic RLR–IFN-β Response Controls the Balance between Antiviral Immunity and Cell Damage. *The Journal of Immunology*, 190(3):1192–1200, feb 2013.

[157] Marco Prinz and Klaus-Peter Knobeloch. Type I Interferons as Ambiguous Modulators of Chronic Inflammation in the Central Nervous System. *Frontiers in Immunology*, 3(APR):67, apr 2012.

[158] Nir Drayman, Parthiv Patel, and Luke Vistain. HSV-1 single-cell analysis reveals the activation of anti-viral and developmental programs in distinct sub-populations. *eLife*, 8, may 2019.

[159] Lionel B. Ivashkiv and Laura T. Donlin. Regulation of type I interferon responses. *Nature reviews. Immunology*, 14(1):36–49, jan 2014.

[160] Samantha L. Goldman, Matthew MacKay, Ebrahim Afshinnekoo, Ari M. Melnick, Shuxiu Wu, and Christopher E. Mason. The Impact of Heterogeneity on Single-Cell Sequencing. *Frontiers in Genetics*, 10(MAR):8, mar 2019.

[161] Amy Brock, Hannah Chang, and Sui Huang. Non-genetic heterogeneity a mutation-independent driving force for the somatic evolution of tumours, may 2009.

[162] Eunjung Kim, Jae Young Kim, Matthew A. Smith, Eric B. Haura, and Alexander R.A. Anderson. Cell signaling heterogeneity is modulated by both cell-intrinsic and -extrinsic mechanisms: An integrated approach to understanding targeted therapy. *PLoS Biology*, 16(3), mar 2018.

[163] Frank S. Heldt, Sascha Y. Kupke, Sebastian Dorl, Udo Reichl, and Timo Frensing. Single-cell analysis and stochastic modelling unveil large cell-to-cell variability in influenza A virus infection. *Nature Communications*, 6(1):1–12, nov 2015.

[164] Volkan Özenci, Mathilde Kouwenhoven, Natalia Teleshova, Mikhail Pashenkov, Sten Fredrikson, and Hans Link. Multiple sclerosis: Pro- and anti-inflammatory cytokines and metalloproteinases are affected differentially by treatment with IFN-β. *Journal of Neuroimmunology*, 108(1-2):236–243, aug 2000.

[165] Yohei Motomura, Hiroyuki Hamada, Masahiro Okamoto, Yohei Motomura, Hiroyuki Hamada, and Masahiro Okamoto. An Effective Numerical Calculation Method for Multi-Time-Scale Mathematical Models in Systems Biology. *Applied Mathematics*, 07(17):2241–2268, nov 2016.

[166] Katarzyna A. Rejniak and Alexander R. A. Anderson. Hybrid models of tumor growth. *Wiley Interdisciplinary Reviews: Systems Biology and Medicine*, 3(1):115–125, jan 2011.

[167] Jessica S Yu and Neda Bagheri. Multi-class and multi-scale models of complex biological phenomena. *Current Opinion in Biotechnology*, 39:167–173, jun 2016.

[168] Alexander M. Bailey, Michael B. Lawrence, Hulan Shang, Adam J. Katz, and Shayn M. Peirce. Agent-based model of therapeutic adipose-derived stromal cell trafficking during ischemia predicts ability to roll on p-selectin. *PLoS Computational Biology*, 5(2), feb 2009.

[169] Chad M. Glen, Melissa L. Kemp, and Eberhard O. Voit. Agent-based modeling of morphogenetic systems: Advantages and challenges. 15(3):e1006577, mar 2019.

[170] Craig W. Reynolds. Flocks, herds, and schools: A distributed behavioral model. In *Proceedings of the 14th Annual Conference on Computer Graphics and Interactive Techniques, SIGGRAPH 1987*, pages 25–34. Association for Computing Machinery, Inc, aug 1987.

[171] Zhihui Wang and Thomas S. Deisboeck. Computational modeling of brain tumors: discrete, continuum or hybrid? pages 381–393. Springer, Dordrecht, 2008.

[172] Mohammad Hossein Zangooei and Jafar Habibi. Hybrid multiscale modeling and prediction of cancer cell behavior. *PLOS ONE*, 12(8):e0183810, aug 2017.

[173] Elsje Pienaar, William M Matern, Jennifer J Linderman, Joel S Bader, and Denise E Kirschner. Multiscale Model of Mycobacterium tuberculosis Infection Maps Metabolite and Gene Perturbations to Granuloma Sterilization Predictions. *Infection and immunity*, 84(5):1650–1669, may 2016.

[174] Hayley C. Warsinske, Amanda K. Wheaton, Kevin K. Kim, Jennifer J. Linderman, Bethany B. Moore, and Denise E. Kirschner. Computational Modeling Predicts Simultaneous Targeting of Fibroblasts and Epithelial Cells Is Necessary for Treatment of Pulmonary Fibrosis. *Frontiers in Pharmacology*, 7:183, jun 2016.

[175] Hiroshi Maekawa, Tsuyoshi Inoue, Haruki Ouchi, Tzu Ming Jao, Reiko Inoue, Hiroshi Nishi, Rie Fujii, Fumiyoshi Ishidate, Tetsuhiro Tanaka, Yosuke Tanaka, Nobutaka Hirokawa, Masaomi Nangaku, and Reiko Inagi. Mitochondrial Damage Causes Inflammation via cGAS-STING Signaling in Acute Kidney Injury. *Cell Reports*, 29(5):1261–1273.e6, oct 2019.

[176] Lodoe Lama, Carolina Adura, Wei Xie, Daisuke Tomita, Taku Kamei, Vitaly Kuryavyi, Tasos Gogakos, Joshua I. Steinberg, Michael Miller, Lavoisier Ramos-Espiritu, Yasutomi Asano, Shogo Hashizume, Jumpei Aida, Toshihiro Imaeda, Rei Okamoto, Andy J. Jennings, Mayako Michino, Takanobu Kuroita, Andrew Stamford, Pu Gao, Peter Meinke, J. Fraser Glickman, Dinshaw J. Patel, and Thomas Tuschl. Development of human cGAS-specific small-molecule inhibitors for repression of dsDNA-triggered interferon expression. *Nature Communications*, 10(1), dec 2019.

[177] Anping Li, Ming Yi, Shuang Qin, Yongping Song, Qian Chu, and Kongming Wu. Activating cGAS-STING pathway for the optimal effect of cancer immunotherapy, apr 2019.

[178] Ting Su, Yu Zhang, Kristoffer Valerie, Xiang Yang Wang, Shuibin Lin, and Guizhi Zhu. STING activation in cancer immunotherapy, 2019.

134

[179] Ji Wang, Peiyu Li, and Mei X. Wu. Natural STING Agonist as an "Ideal" Adjuvant for Cutaneous Vaccination. *Journal of Investigative Dermatology*, 136(11):2183–2191, nov 2016.

[180] L E KREUZ and A H LEVY. PHYSICAL PROPERTIES OF CHICK INTERFERON. *Journal of bacteriology*, 89(2):462–9, feb 1965.

[181] Mathieu Coppey, Alexander M. Berezhkovskii, Stuart C. Sealfon, and Stanislav Y. Shvartsman. Time and Length Scales of Autocrine Signals in Three Dimensions. *Biophysical Journal*, 93(6):1917–1922, sep 2007.

[182] J. Martín-Vaquero and A. Kleefeld. ESERK5: A fifth-order extrapolated stabilized explicit Runge–Kutta method. *Journal of Computational and Applied Mathematics*, 356:22–36, aug 2019.

[183] Assyr Abdulle. Fourth order Chebyshev methods with recurrence relation. *SIAM Journal on Scientific Computing*, 23(6):2041–2054, 2002.

[184] Youcef Saad and Martin H. Schultz. GMRES: A Generalized Minimal Residual Algorithm for Solving Nonsymmetric Linear Systems. *SIAM Journal on Scientific and Statistical Computing*, 1986.

[185] Satoshi Yamada, Satoru Shiono, Akiko Joo, and Akihiko Yoshimura. Control mechanism of JAK/STAT signal transduction pathway. *FEBS Letters*, 534(1-3):190–196, jan 2003.

[186] Danielle L. Krebs and Douglas J. Hilton. SOCS Proteins: Negative Regulators of Cytokine Signaling. *Stem Cells*, 19(5):378–387, sep 2001.

[187] Laura M. Snell, Tracy L. McGaha, and David G. Brooks. Type I Interferon in Chronic Virus Infection and Cancer, aug 2017.

[188] Sonali Patil, Miguel Fribourg, Yongchao Ge, Mona Batish, Sanjay Tyagi, Fernand Hayot, and Stuart C Sealfon. Single-cell analysis shows that paracrine signaling by first responder cells shapes the interferon-b response to viral infection.

[189] Jeffrey M. Perkel. Julia: come for the syntax, stay for the speed. *Nature*, 572(7767):141–142, aug 2019.

[190] Maciej H. Swat, Gilberto L. Thomas, Julio M. Belmonte, Abbas Shirinifard, Dimitrij Hmeljak, and James A. Glazier. Multi-Scale Modeling of Tissues Using CompuCell3D. In *Methods in Cell Biology*, volume 110, pages 325–366. Academic Press Inc., 2012.

[191] Jorn Starruß, Walter De Back, Lutz Brusch, and Andreas Deutsch. Morpheus: A user-friendly modeling environment for multiscale and multicellular systems biology. *Bioinformatics*, 30(9):1331–1332, may 2014.

[192] Andrea Ablasser and Muhammed F. Gulen. The role of cGAS in innate immunity and beyond. *Journal of Molecular Medicine*, 94(10):1085–1093, oct 2016.

[193] Yugen Zhang, Laxmi Yeruva, Anthony Marinov, Daniel Prantner, Priscilla B Wyrick, Vladimir Lupashin, and Uma M Nagarajan. The dna sensor, cyclic gmp–amp synthase, is essential for induction of ifn-β during chlamydia trachomatis infection. *The Journal of Immunology*, 193(5):2394–2404, 2014.

[194] Armstrong Murira and Alain Lamarre. Type-i interferon responses: from friend to foe in the battle against chronic viral infection. *Frontiers in immunology*, 7:609, 2016.

[195] Suresh I Rattan, Anastasia Derventzi, and BF Clark. Protein synthesis, posttranslational modifications, and aging. *Annals of the New York Academy of Sciences*, 663:48–62, 1992.

[196] Satoshi Fukuyama, Kiyoko Iwatsuki-Horimoto, Maki Kiso, Noriko Nakajima, Robert W Gregg, Hiroaki Katsura, Yuriko Tomita, Tadashi Maemura, Tiago Jose da Silva Lopes, Tokiko Watanabe, et al. Pathogenesis of influenza a (h7n9) virus in aged non-human primates. *The Journal of Infectious Diseases*, 2020.

[197] Lingyin Li, Qian Yin, Pia Kuss, Zoltan Maliga, José L Millán, Hao Wu, and Timothy J Mitchison. Hydrolysis of 2'3'-cGAMP by ENPP1 and design of nonhydrolyzable analogs. *Nature chemical biology*, 10(12):1043–8, dec 2014.

[198] Arun Prakash and David E. Levy. Regulation of IRF7 through cell type-specific protein stability. *Biochemical and Biophysical Research Communications*, 342(1):50–56, mar 2006.