ABSTRACT

Title of dissertation: MATHEMATICAL MODELS OF QUORUM SENSING

Hana Ueda, Doctor of Philosophy, 2016

Dissertation directed by: Professor William E. Bentley
Department of Bioengineering &
Applied Mathematics & Statistics,
and Scientific Computation (AMSC) program

Professor Konstantina Trivisa
Department of Mathematics &
Applied Mathematics & Statistics,
and Scientific Computation (AMSC) program

Mathematical models of biological phenomena are constructed in order to further the understanding of the known and unknown interactions that result in the behaviors of dynamical systems. We present mathematical models dealing with quorum sensing, which is the biological process of communication in bacteria. The density-dependent means of communication are mediated by molecules called autoinducers that are both synthesized and collected by quorum sensing bacteria.

We employ differential equations to investigate and understand the dynamics of underlying signaling processes. The first two models of this study were constructed with the idea of relating flocking movements observed in birds to gene expression in quorum sensing. To this end, modified Cucker-Smale flocking equations which do not require detailed knowledge of signal transductions mechanisms

or regulatory proteins are used to represent quorum sensing and chemotaxis. The dynamical behaviors of these models are analyzed and approximated using asymptotic analysis and simulations. The coupling of quorum sensing and chemotaxis systems results in the formation of two groups of cells during the migration towards the attractant, which is similar to behavior observed in experiments of chemotaxing *E. coli*. This consequence of density influencing the velocity of bacteria suggests the possibility that density (or a density-dependent system such as quorum sensing) affects the chemotaxis system. We also show an application of this coupled model that produces qualitatively similar results with experimental data.

To further analyze collective behavior emerging from the interactions of quorum sensing and chemotaxis, this study uses statistical physics to derive a partial differential equation that tracks the time evolution in phase space of the distribution of these cells. Lastly, this study combines theory and experimental data to present a compartmental model that predicts p-aminophenol (PAP) response to various autoinducer concentrations in quorum sensing cells. The use of compartments allows for the model to be customized for constructs that do not use autoinducer-mediated production of the β-galactosidase enzyme.

MATHEMATICAL MODELS OF QUORUM SENSING

by

Hana Ueda

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2016

Advisory Committee:
Professor William E. Bentley, Advisor
Professor Konstantina Trivisa, Co-Advisor
Professor Maria Cameron
Professor Raymond Adomaitis
Professor Nam Sun Wang (Dean's Representative)

Dedication

To my mother and father for their love and support.

Acknowledgements

I would like to acknowledge the many people who have supported me throughout my years at the University of Maryland.

First of all, I would like to thank Dr. William E. Bentley and Dr. Konstantina Trivisa. I would like to thank Dr. Bentley for his sharp insights, vision, and unwavering support. His sense of humor, patience, and calm optimism imbued in our discussions were very helpful throughout this journey. I am grateful for the guidance, direction, and opportunities he has provided. Thank you Dr. Bentley for allowing me to be a part of the Bentley Lab family as well as the Institute for Bioscience and Biotechnology Research (IBBR). I would like to thank Dr. Konstantina Trivisa for her enthusiasm and effervescent energy. Thank you for being interested in and supportive of the research of all the AMSC students. I am grateful for the spirited discussions, direction, and advice that she has provided. Thank you Dr. Trivisa for inviting me to join the AMSC program.

I would like to thank my committee members, Dr. Maria Cameron, Dr. Raymond Adomaitis, and Dr. Nam Sun Wang for serving on my committee and for their invaluable time and suggestions. I am very much grateful for their insights. I would like to especially thank Dr. Maria Cameron for her support and feedback.

I am so thankful to the members of the Bentley Lab, with whom I have shared such great conversations. I would not have been able to do this without their support, friendship, and help. A special shout-out to Chelsea Virgile, Dr. Pricila Hauk, Dr. Tanya Tschirhart, Dr. Jessica Terrell, Dr. Melissa Rhoads, and Kristina

Stephens for their awesomeness. I would especially like to thank Dr. Jessica Terrell for her images and experimental data from her work in Servinsky et al. (2016). Thank you also to the mathematicians Dr. Matias Delgadino and Jacky Chong for their insightful discussions and help.

I have had the honor of TA-ing for and/or sharing insightful conversations with many of the professors of the Department of Mathematics, especially Dr. Kenneth Berg, Dr. Maria Cameron, Dr. Casey Cremins, Dr. P.M. Fitzpatrick, Dr. Manoussos Grillakis, Dr. P-E Jabin, Dr. C.D. Levermore, Dr. Doron Levy, Dr. Melissa Macasieb, Dr. Kasso Okoudjou, and Dr. Jacqueline Stone. I have also shared some wonderful conversations with the staff members of the Math Department: Linette Berry, Rhyneta Fleming, Haydee Hidalgo, Alverda McCoy, Celeste Regalado, William Schildknecht, and Sharon Welton. Thank you for aiding me throughout my time in Maryland.

Lastly, I would like to thank my parents for always cheering me on. Thank you for all the love. Thank you for everything.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

AI-2	autoinducer-2
BBGKY	Bogoliubov-Born-Green-Kirkwood-Yvon
β-gal	beta-galactosidase
cAMP	cyclic AMP
CRP	cAMP receptor protein
DPD	4,5-dihydroxy-2,3-pentanedione
FMN	flavin mononucleotide
GFP	green fluorescent protein
lsr	*luxS*-regulated
MetK	methionine adenosyl transferase
ONPG	ortho-nitrophenyl-beta-galactoside
ONP	ortho-nitrophenol
PAP	4-aminophenol
PAPG	4-aminophenyl beta-D-galactopyranoside
SAH	S-adenohomocysteine
SAM	S-adenosylmethionine
SRH	S-ribosylhomocysteine

Chapter 1: Introduction

In order to extract more insights from biology, mathematics can be used to elucidate the complexities of biological phenomena. The field of systems biology strives to accomplish this with the construction of mathematical models. The model breaks down the complex interactions within the system into components. Each of these components can be then analyzed separately in order to gain an understanding of how one influences the other and how each contributes to the system as a whole. By building such a dynamical model, one can study the effects of changing various parameters and the long-term behavior of the various processes involved that would otherwise be difficult to deduce in an experimental setting. There are mechanistic models that seek to model the detailed regulatory networks of a system. There are also simple models that generalize the observed phenomenon. In this dissertation, we build mathematical models of quorum sensing, a form of chemical signalling employed by bacteria.

1.1 Quorum sensing

The term 'quorum sensing' describes the system employed by bacteria that allows them to sense the population density in their surroundings and collectively

express certain genes. The means by which each bacterium accomplish this feat are through the use of molecules called autoinducers. These autoinducers are both synthesized and collected by the bacterium. The number of autoinducers in the external environment increases as the population density increases. When a sufficient concentration is attained, signal transduction mechanisms are activated that in turn, facilitate the expression of specific genes. The result of such a system is the collective expression of these autoinducer-regulated genes when enough bacteria accumulate (ie when a quorum is attained).

Quorum sensing systems differ from species to species. The participating regulatory proteins, signalling pathways, and types of autoinducer can be varied. To illustrate, we elaborate on the LuxI/LuxR system in *Vibrio fischeri* and the LuxS/AI-2 system in *Escherichia coli*.

1.1.1 LuxI/LuxR system in *Vibrio fischeri*

The *Vibrio fischeri* is one of many species of bacterium that produce light. They live in the light organ of the Hawaiian bobtail squid, *Euprymna scolopes*. The two have a symbiotic relationship in which the bioluminescence produced by the bacterium enables the squid to escape from bottom-dwelling predators as the light achieves a camouflage effect by resembling moonlight [67] [92].

There are two divergent operons that regulate the quorum-sensing behavior that generates the bioluminescence in *Vibrio fischeri*. One operon contains the *luxICDABEG* genes. *luxI* aids in the synthesis of an acyl-homoserine lactone (AHL),

the autoinducer in this quorum sensing system. *luxCDE* codes for proteins that aid in the synthesis of the aldehyde substrate needed for luminescence. *luxAB* are the genes responsible for the luciferase enzyme that produces the bioluminescence [36] [109]. *luxG* codes for proteins that reduce flavin mononucleotide (FMN) for reactions that produce luminescence [70] [77].

On the other operon is the gene that codes for the LuxR protein. Together, the LuxR and AHL form a complex that activates the transcription of the *luxICDABEG* genes, allowing for the expression of bioluminescence.

When the cell population is low (ie low density of cells), the concentration of autoinducers in the surrounding medium is low. Hence, there is only a basal (background) level of expression of bioluminescence. When there is a high cell density, there are consequently much more autoinducers which accumulate both outside and inside of the cell. After the autoinducers surpass a certain threshold, the AHL/LuxR complex upregulates transcription of *luxICDABEG* which produces the light. On the other hand, the complex negatively regulates *luxR* transcription which in turn controls the formation of the AHL/LuxR complex which in turn controls the transcription of *luxICDABEG*. Hence, this LuxI/LuxR system consists of positive and negative feedback mechanisms that regulate bioluminescence.

1.1.2 LuxS/AI-2 system in *Escherichia coli*

Escherichia coli is a species commonly found in the lower intestine of humans and known to help with digestion. Depending on the strain of the species, however,

E. coli can produce harmful toxins. The difference between a non-pathogenic *E. coli* and a pathogenic one lies in the genes of the strain: When specific genes of the pathogenic *E. coli* are transcribed, toxins are produced. The instructions of whether or not to transcribe these genes is influenced by a furanosyl borate diester called AI-2, the autoinducer that is both synthesized and collected by *E. coli*.

Each of the cells produce the AI-2 which is used to sense the presence of other cells in their surroundings. *E. coli* first synthesizes the amino acid methionine, which is used to produce S-adenosylmethionine (SAM), with the help of the methionine adenosyl transferase (MetK) enzyme. SAM has further roles in other methylation processes within the cell due to its donatable methyl group. One of its roles is the donation of the methyl group via the *cheR* methyltransferase to a Methyl acceptor, which produces S-adenohomocysteine (SAH) and a methylated product. The enzyme, Pfs, degrades SAH by removing adenine to form S-Ribosylhomocysteine (SRH). SRH, assisted by the LuxS enzyme, then becomes homocysteine and 4,5-dihydroxy-2,3-pentanedione (DPD). Homecysteine is used to produce methionine, which is used to start this entire process again. DPD in turn becomes AI-2. However, the process of how DPD turns into AI-2 is still unknown and how AI-2 is exported outside of the cell is unknown [95] [111].

The divergent operons of *lsr* (consisting of *lsrACDBFG*) and *lsrRK* play an integral role in quorum sensing in *E. coli*. The *lsrRK* operons codes for the LsrR and LsrK proteins. The *lsr* operon encodes for Lsr A, Lsr C, LsrD, LsrB, LsrF, LsrG, and Tam. LsrA, LsrC, LsrD, and LsrB assist in the transport of AI-2 from the extracellular medium to inside the cell. LsrF and LsrG are involved in the degra-

dation of phosphorylated AI-2 (AI-2-P). What becomes of AI-2-P after the actions of LsrF and LsrG is still unknown [111] [112]. The cyclic AMP (cAMP) and the cAMP receptor protein (CRP) bind together to form the cAMP-CRP complex. The cAMP-CRP in turn needs to bind to the promoter region of the operon in order for transcription to occur. The regulator LsrR prevents the transcription of the *lsr* operon as well as the transcription of its own *lsrR*. But when AI-2-P binds to LsrR, LsrR no longer represses transcription (ie given the presence of the cAMP-CRP complex, an increase in AI-2-P causes an increase in *lsr* transcription). LsrK transforms the extracted AI-2 into AI-2-P and Tam helps out in AI-2 synthesis [112].

The working model for uptake of AI-2 is as follows. AI-2 is transported back into the cell via Lsr A, Lsr B, Lsr C, and Lsr D. The rate of transport of AI-2 depends on the amount of glucose outside of the cell [111]. When glucose is present, there are low levels of cAMP and CRP intracellularly, which reduces transcription of the *lsr* operon. Since the *lsr* operon codes for the LsrA, LsrB, LsrC, and LsrD, this causes less extracellular AI-2 to be taken up by the cell.

Low levels of cAMP and CRP also cause an increase in LuxS. However, how cAMP and CRP accomplish this is still unknown as the cAMP-CRP complex does not bind to promoter region of *luxS* [111]. Nonetheless, the increase in *luxS* transcription aids in the increase in the synthesis of AI-2. Besides facilitating the conversion of SRH to DPD and Homocysteine, however, the direct process of an increase in LuxS causing an increase in AI-2 is not clear. Most likely, glucose also aids in the production of SAM and other substrates at the beginning of the AI-2 synthesis process [111].

When glucose is not present, the cAMP-CRP complex binds to the *lsr* operon, allowing for transcription of the *lsr* and *lsrRK* operons. Thus, production of LsrA, LsrB, LsrC, and LsrD allows for the uptake of extracellular AI-2. The AI-2 then goes through phosphorylation via LsrK to become AI-2-P. Although transcription of *lsrR* means that the repressor LsrR may stop the transcription of the *lsrACDBFG* operon, AI-2-P binds with LsrR in order for transcription to continue. LsrF and LsrG help regulate this cycle by degrading AI-2-P since low levels of AI-2-P will allow the LsrR to continue repressing transcription [111].

So, different species of bacteria can have different ways of quorum sensing. In *Pseudomonas aeruginosa*, there is the LasI/LasR-RhlI/RhlR system. In *Agrobacterium tumefaciens*, there is the TraI/TraR system and we have only discussed some species of gram-negative bacteria. For more details, several more quorum sensing systems are explained in Miller and Bassler [69].

While bacteria are quorum sensing, many do not remain motionless: there is movement. When there is movement as a result of chemical stimuli, bacteria are said to be chemotaxing. Chemotaxis describes the directed movement of bacteria towards or away from areas of higher concentrations of certain chemicals. For *E. coli*, these chemicals can range from food sources such as glucose (chemoattractant) to amino acids such as leucine (chemorepellent). The movements toward these chemicals take the form of tumbles and runs, in which the duration of the run is longer when an increase in concentration of chemoattractants is detected. Chemoreceptors sense the changes of chemicals in the environment and activate a signal transduction cascade that influences the flagella. By using this system of motility, bacteria can travel to-

ward environments favorable to their survival. More details of the specific signalling pathways involved in chemotaxis can be found in Wadhams and Armitage [110] and in Adler [1] [2], among others [12] [16] [66] [100].

1.2 Quorum sensing models

Mathematical models of quorum sensing began fairly recently, with the deterministic models of James et al. [55], Dockery and Keener [29], and Ward et al. [115]. The first two models focused on the regulatory mechanisms involved in the quorum sensing systems of *Vibrio fischeri* and *Pseudomonas aeruginosa*, respectively. Both papers used ordinary differential equations (with the latter presenting a second model of ordinary and partial differential equations) to analyze the effect of autoinducers on transcription of regulatory proteins. Various simplifying assumptions were used to perform a steady state and stability analysis which showed there exists a switch between the states of quorum sensing-related expression dependent on certain parameter values as well as autoinducer concentration. Ward et al. [115] provided a more generalized model of the regulation of *lux* genes in *Vibrio fischeri* using ordinary differential equations. The model, focusing on density of cell population and autoinducer concentration, was fitted to experimental data to extract parameter values that establish the stable steady state solution of the system.

The models developed since then either build on the complexity of the regulatory processes [34] [33], such as taking into account negative feedback mechanisms [116] and oxygen concentrations in the medium [20], or making generaliza-

tions about the modelled system, such as in Koerber et al. [61] and Ward et al. [114]. Quorum sensing models range from deterministic (ordinary/partial differential equations) [37] [39] [52] [73] [108] to stochastic [40] [63] [89].

Many quorum sensing models, such as those listed above, do not incorporate movement of the bacterium even though the modelled species is motile. One example of a chemotaxis model that incorporates quorum sensing in some way is the Painter and Hillen model [78]. This system of partial differential equations analyzes the density of cell population and chemoattractant concentration. They include an assumption of 'quorum-sensing type behavior' as a means of regulation of the cell aggregation that occurs in their volume-filling model. That is, they have their cells change response to the attractant based on the density of their surroundings. They construct their model in this way to allow for cell migration to a lower density, which reduces the traffic caused by the aggregation in their model.

We model quorum sensing bacteria that are chemotaxing towards some source of attractant. Our goal is to explore how the two systems of quorum sensing and chemotaxis influence one another. In order to do this, we use the concept of flocking.

1.3 Flocking

The study of self-organization of autonomous agents provides for a better understanding of such observed phenomena in nature. These kinds of collective behavior can be reflected in motions of birds, fish, wildebeest, and in certain species of bacteria [6] [15] [41] [79]. By constructing mathematical models, one can analyze how

local interactions of these agents can influence overall, complex behaviors and can explain how such multi-agent interactions can benefit the population biologically. Many of these mathematical models are particle based models, in which the behaviors of each agent are characterized by rules that take the form of differential equations governing position and velocity or a Fokker-Planck equation [18] [106] [119].

One such model is the Cucker-Smale flocking model [24]. This model focuses on the flocking, or collective motion, observed in birds. By using equations of the continuous and discrete variety with respect to time, they showed that if every bird in the flock adjusts its velocity according to the velocities of birds in its surroundings, then depending on certain conditions, the birds will all converge to the same velocity. The continuous time formulation of the Cucker-Smale model expanded on by Ha and Tadmor [44] is shown below.

$$\frac{d}{dt}\mathbf{x}_i(t) = \mathbf{v}_i(t)$$
$$\frac{d}{dt}\mathbf{v}_i(t) = \frac{\lambda}{N}\sum_{1\leq j\leq N} k(\mathbf{x}_i(t), \mathbf{x}_j(t))(\mathbf{v}_j(t) - \mathbf{v}_i(t))$$

Here, each agent i's position and velocity are represented by ordinary differential equations for a dynamical system of N agents. Each agent i's velocity is adjusted based on the surrounding agents j. The differential equation governing cell i's velocity, with constant coefficient λ, contains the function $k(\mathbf{x}_i, \mathbf{x}_j)$ which acts as a weight for $(\mathbf{v}_j - \mathbf{v}_i)$ and is based on the proximity of agents i and j. The weights can be constructed so that the closer agent j is to agent i, the velocity of agent j will have more influence on the velocity of agent i.

1.4 Other methods of modelling

Various other techniques can be used to model biological processes, such as those used in enzyme kinetics and in computational neuroscience. Enzyme kinetics is a field devoted to the analysis of enzymatic reactions in order to understand the substrate-enzyme interactions within the cell as well as the conditions which affect such interactions. The Michaelis-Menten model is used to analyze the fast enzymatic reactions that convert substrates into products. This model takes into account the rate in which the enzyme binds to the substrate, the rate in which the enzyme unbinds from the substrate, and the rate in which the enzyme unbinds from the product. By simplifying this dynamics process into three parts, Michaelis and Menten constructed an equation that gives the rate of product formation, given the concentration of substrate and the initial concentration of the enzyme.

$$\frac{d[P]}{dt} = \frac{k_{cat}[E_0][S]}{K_M + [S]}$$

where k_{cat} is the rate in which the product forms from the enzyme-substrate complex, K_M is the Michaelis constant, E_0 is the initial enzyme concentration, S is the concentration of substrate, and P is the concentration of product.

In computational neuroscience, models are used to find out more about the workings of the brain, such as the ways in which the many neurons transmit information in the form of chemical and electrical signals. The system of nonlinear differential equations that is the Hodgkin-Huxley model [50] showed how a model that takes into account voltage, membrane capacitance, and various currents could

capture the dynamics of action potentials. Developed in the 1950's, Hodgkin and Huxley analyzed the action potentials in the squid giant axon that was used in electrophysiological experiments [75]. The Hodgkin-Huxley equations formed the basis of subsequent compartmental modelling of neuron behavior.

Compartmental models have applications in epidemiology, ecology, and neuroscience [49] [80] [101] [105] to name a few. The principle of breaking down a system into components is what compartmental models literally aim to do. The system is broken down into connected subsystems. Each subsystem is modelled. The subsystem can be analyzed separately. It can also be hooked up to the other subsystems to see how the sum of the parts make up the whole.

To elaborate on how such a construction of models can provide insight into biological processes, one can look at how compartmental models are used in investigating the workings of a neuron. Neurons relay chemical and electrical signals that are conveyed through the dendrites, soma, and axons. The dendrites receive signals from other neurons and relays them to the cell body. The axons take information away from the cell body and relay them to other neurons, muscles, or nerves via synapses. There are several ways in which compartmental models are constructed for such a system, such as in the Traub et al. 19-compartment system [101] or the Pinsky/Rinzel 2-compartment model [80]. But the general idea is that the output of one compartment is used as the input of another compartment. The proximal and distal dendrites and soma can be all represented as compartments. Alternatively, the dendrites can be lumped into one compartment and the soma can be the other compartment.

1.5 Statement of Purpose

The discovery of quorum sensing showed that bacteria can sense the cell density in their surroundings as well as demonstrate group behavior in the form of collective gene expression. Mathematical models seek to further the understanding of the working of complex systems such as these. Many quorum sensing models analyze the regulatory circuits that coordinate this signalling system but do not factor in the effects of movement in species that are motile. In our quorum sensing models, we take movement into account by having the cells chemotax toward a source of attractants. We approximate the two biological processes of quorum sensing and chemotaxis with Cucker-Smale type equations. The purpose of incorporating Cucker-Smale is to construct a generalized model. Developing a mechanistic model would require the modelling of detailed regulatory mechanisms that differ from species to species. By using Cucker-Smale flocking generalizations, the goal is to construct a model that can be applied to multiple species of motile bacteria and to analyze the resulting behaviors of this dynamical system.

This dissertation is an exploration in the modelling of quorum sensing in motile bacteria. Chapter 2 and Chapter 3 introduce flocking as a representation of bacterial movement and communication. These mathematical models present novel, deterministic, non-mechanistic, particle-based ordinary differential equations that explore the connection between the quorum sensing system and chemotaxis system. In our flocking models, we look at each individual cell and we track the following aspects: position, velocity, auto-inducer triggered gene expression, and surrounding

concentration of autoinducers. We analyze the quorum sensing behaviors of each cell while they chemotax towards a specified source and we perform asymptotic analysis, sensitivity analysis and explore any trends in dynamics evoked by the interplay of both quorum sensing and chemotaxis systems.

Chapter 4 focuses on the development of a kinetic partial differential equations model. From our flocking model, we employ statistical physics to derive an N-particle partial differential equation that describes the time evolution in phase space of a distribution of N particles. We reduce the N-particle distribution into a j-particle formulation, referred to as the Bogoliubov-Born-Green-Kirkwood-Yvon (BBGKY) hierarchy.

Finally, Chapter 5 presents a compartmental model for an electrochemical method of measuring autoinducer-triggered gene expression. In that model, there are two compartments which model regulatory mechanisms of quorum sensing and enzyme kinetics, respectively. Each compartment is fitted to the corresponding data set. Then, the compartments are combined together as one system of differential equations. The results of the model are compared to the results of the experiment to see how well the model predicted the observed dynamics.

Chapter 2: Quorum Sensing as a form of flocking

2.1 Introduction

Quorum sensing is a form of communication that bacteria employ in order to coordinate behavior in a collective manner. The means of communication, or collective gene regulation, is accomplished through chemical signals called autoinducers. By producing and collecting autoinducers, bacteria such as *Escherichia coli*, *Vibrio fischeri*, and *Pseudomonas aeruginosa*, can determine the cell density in their surroundings. An increase in cell density through changes in the concentration of autoinducers triggers a series of signal transduction mechanisms. This in turn allows for changes in gene expression. Hence, when there is a sufficient amount of bacteria (a quorum) in an area, bacteria collectively exhibit various phenotypes, such as bioluminescence, biofilm, and virulence factors, to name a few [69] [117].

One approach to modelling these kinds of regulatory systems involves differential equations that represent signal synthesis, uptake, binding, and the transcription and translation from gene to protein as well as the various enzyme interactions (e.g, repression, phosphorylation) that allow for these processes to occur [52] [84] [104]. We propose a different approach, in which we model the general behaviors of cellular dynamics. By not emphasizing the regulatory mechanisms, we solely focus on the

general picture pertaining to quorum sensing and the general behaviors governed by quorum sensing such as chemotaxis. In order to accomplish this, we introduce flocking. To date, there are no reports of quorum sensing behavior attributed to flocking.

Flocking describes the movement of biological entities that prescribes to the information from their surroundings to adjust their actions, forming a collective motion. Collective motion is observed, for example, in birds, fish, wildebeest, and even in *E. coli* [6] [26] [41] [79]. To quantify this emergent behavior in order to further study the self-organizing dynamics of biological systems, Cucker & Smale [24] [25] introduced a dynamical model that analyzes these movements using weighted averages. This representation of flocking was introduced as a method to examine the dynamics behind flocking formation in birds by taking weighted averages of the surrounding velocities as an illustration of how birds adjust their velocity with respect to other birds. When each bird in a flock adjusts its speed in this way, every bird ultimately attains the same velocity. The biological coordination of bird flocking has been attributed to sight [82], wherein mathematically expanded on by Ha and Tadmor [44], the Cucker-Smale model is as follows:

$$\frac{d}{dt}\mathbf{x}_i(t) = \mathbf{v}_i(t)$$
$$\frac{d}{dt}\mathbf{v}_i(t) = \frac{\lambda}{N}\sum_{1\leq j\leq N} k(\mathbf{x}_i(t), \mathbf{x}_j(t))(\mathbf{v}_j(t) - \mathbf{v}_i(t))$$

The model states that the velocity of each bird i, represented by the state variable $\mathbf{v}_i$(t), changes based on the velocities of the surrounding birds j, with the symmetric weights $k(\mathbf{x}_i(t), \mathbf{x}_j(t))$ determining the strength of influence of the neighboring birds.

$\mathbf{x}_i(t)$ and $\mathbf{x}_j(t)$ represent the positions of birds i and j and $k(\mathbf{x}_i(t), \mathbf{x}_j(t))$ is a function of the distances between the birds. That is, $k(\mathbf{x}_i(t), \mathbf{x}_j(t)) = k(|\mathbf{x}_i - \mathbf{x}_j|)$. N is the total number of birds and $\lambda > 0$ is a factor controlling the strength of the flocking term. From this differential equation, Cucker & Smale prove that every bird approaches the same velocity.

The mathematical definition of flocking, provided in [72], is given as:

Definition 2.1.1. *Given a particle model with distances and velocities of each cell i represented by $\{\boldsymbol{x}_i(t), \boldsymbol{v}_i(t)\}_{i=1,\dots,N}$, the system converges to a flock provided that the following properties hold:*

(a) $\lim_{t\to\infty} \max_{i,j} |\boldsymbol{v}_j(t) - \boldsymbol{v}_i(t)| = 0$

(b) $\sup_{t\geq 0} \max_{i,j} |\boldsymbol{x}_j(t) - \boldsymbol{x}_i(t)| < \infty$

In order to qualify as flocking, the distances must remain bounded and the velocities must asymptotically converge to the same value.

As for applications of flocking to cells, Di Costanzo et al. [28] incorporated this in the modelling of motility of cells during the morphogenesis process of the posterior lateral line primordium in zebrafish. Ha & Levy [43] incorporated flocking into their model when modelling phototaxis of cyanobacterium *Synochocystis* sp.. In this paper, we build on the groundwork established by Ha & Levy [43] and Ha & Tadmor [44]. Our model introduces flocking not just as a method of representing velocities but as a way of representing quorum sensing behavior. To do this, we have expanded the number of state variables of interest. By taking the average of autoinducer-triggered cell activity in a similar way to which the velocities are aver-

aged, positive and negative feedback behavior can be encapsulated.

We apply the concept of taking weighted averages to describe the dynamics in bacterial communication in tandem with chemotaxis. The model takes into account the fluctuations in both cell velocity and in cell gene expression as a result of changing autoinducer concentrations in the surrounding medium. The model that we present can be applied to any quorum sensing species that (1) demonstrates positive and/or negative feedback and (2) produces and collects autoinducers via diffusion or potentially even activated transport.

Quorum sensing systems vary from species to species with different kinds of autoinducers, different signal transduction motifs, and different emergent phenotypes or behaviors. There are many variations and importantly, not all mechanisms are known. By incorporating the Cucker-Smale flocking terms, our goal is to create a minimal model that globally describes these quorum sensing behaviors.

2.2 Flocking Model

We introduce a general model that focuses on the following four aspects of each cell:

1. $\mathbf{x}_i$: position of cell i
2. $\mathbf{v}_i$: velocity of cell i
3. p_i: cell activity of cell i
4. A_i: concentration of autoinducers surrounding cell i

The model is a system of $4N$ ordinary differential equations that tracks the dynamics of the above four state variables. N represents the total number of cells.

$$\frac{d\mathbf{x}_i}{dt} = \mathbf{v}_i \tag{2.1}$$

$$\frac{d\mathbf{v}_i}{dt} = \frac{\lambda_1}{N}\sum_{j=1}^{N} k_1(\mathbf{x}_i, \mathbf{x}_j)(\mathbf{v}_j - \mathbf{v}_i) + F_0(v_o\hat{\mathbf{e}}_\mathbf{s} - \mathbf{v}_i) \tag{2.2}$$

$$\frac{dp_i}{dt} = \frac{\lambda_2}{N}\sum_{j=1}^{N} k_2(\mathbf{x}_i, \mathbf{x}_j)(p_j - p_i) + L_0(p_\infty - p_i)\psi_A(A_i) \tag{2.3}$$

$$\frac{dA_i}{dt} = \lambda_3\sum_{j=1}^{N} k_3(\mathbf{x}_i, \mathbf{x}_j)r - k_u(\mathbf{x}_i, \mathbf{x}_j)A_i - k_dA_i \tag{2.4}$$

with the corresponding weights,

$$k_1(\mathbf{x}_i, \mathbf{x}_j) = k_2(\mathbf{x}_i, \mathbf{x}_j) = k_3(\mathbf{x}_i, \mathbf{x}_j) = \frac{1}{(1 + |\mathbf{x}_j - \mathbf{x}_i|^2)^\beta}$$

logistic function,

$$\psi_A(A_i) = \frac{1}{(1 + e^{c_0(c_1 - c_A A_i)})^\alpha}$$

and uptake function

$$k_u(\mathbf{x}_i, \mathbf{x}_j) = exp\left(\frac{-(a(\gamma(\sum_{j=1}^{N}|\mathbf{x}_j - \mathbf{x}_i|^{2\beta})) - b)^2}{2c}\right) \tag{2.5}$$

The weight functions k_1, k_2, k_3 are symmetric, which will prove useful in the subsequent proofs. These functions are decreasing as the distance between the cells increase. The power of the denominator controls how fast the functions decrease from 1. Qualitatively, the β represent the scope of influence that the neighboring cells have on the individual. So, low β values (such as $\beta = 0$) would model a system in which the individual is influenced by all cells.

The logistic function $\psi_A(A_i)$ acts as smoothed piecewise constant functions and take values ranging from 0 to 1 depending on the level of cell activity (p_i) and surrounding autoinducer concentration (A_i) of cell i, respectively. The uptake function is modelled as a truncated, shifted Gaussian. The truncation is enforced solely by the non-negative domain. $k_u(x_i, x_j)$ is constructed in such a way so that the uptake rate increases as the density increases up until a certain point. Then the uptake decreases when the density is high. This is to reflect a decrease in metabolic activity caused by overcrowding (i.e. decreased nutrient availability when cells are crowded). By using this Gaussian, we are assuming that there is a density range in which uptake is highest.

Equation (2.1) is the definition of velocity. Equation (2.2) models the changing dynamics of the velocity of cell i by employing a flocking term and a source term. The flocking term addresses how the velocities of neighboring cells influence the individual's velocity. The weight function k_1 as well as the parameter β determine how much influence the other cells j have on cell i. Next, the source term models the chemotaxing behavior in which the cell runs toward the higher concentrations of chemoattractant at a velocity of $v_o\hat{\mathbf{e}}_\mathbf{s}$, where $\hat{\mathbf{e}}_\mathbf{s}$ represents the unit direction vector. The strength of this term is controlled by the constant F_0. The combination of both the flocking term and the source term generalizes the movements of these cells in response to the chemical stimuli.

Equation (2.3) tracks the dynamics in autoinducer-triggered gene expression which lead to collective expression of such phenotypes as biofilm formation. To motivate the existence of equation (2.3), for the sake of a concrete example, we use

the expression of bioluminescence by *Vibrio fischeri* [70] [117]. In this species, the autoinducer AI-1(also known as acyl-homoserine lactone) is involved in the quorum sensing process. The regulatory LuxI/LuxR system governs the production of both the autoinducers and various proteins. LuxI is responsible for the production of AI-1 and LuxR is the repressor that interacts with the autoinducers, regulating transcription of the *luxCDABEG* operon. The *luxCDABEG* genes code for the regulation of bioluminescence, including the expression of the luciferase enzymes that are responsible for light. The flocking term then "averages" the bioluminescence of the surrounding cells. The additional term represents the bioluminescence triggered when the surrounding concentration of autoinducers passes a certain threshold. Above this threshold, the logistic function ψ_A approaches 1, thereby activating the term. The strength of this term is controlled by the constant L_0.

The changes in the surrounding concentration of autoinducers of a particular cell is represented by Equation (2.4). The equation is constructed as

$$\begin{aligned}\frac{dA_i}{dt} &= \text{synthesis} - \text{uptake} - \text{sloughing off/degradation of autoinducers} \\ &= \lambda_3 \sum_{i=1}^{N} k_3(\mathbf{x}_i, \mathbf{x}_j) r - k_u(\mathbf{x}_i, \mathbf{x}_j) A_i - k_d A_i\end{aligned}$$

2.2.1 Asymptotics and Analysis

In this section, we prove that the system represented by equations (2.1)-(2.4) satisfy the above definition of flocking. We first analyze the asymptotic behaviors of cell velocity and then proceed to the analysis regarding cell position.

2.2.1.1 Velocity

Proposition 1. *Let* $\boldsymbol{v}_{avg}(t) := \frac{1}{N}\sum_{j=1}^{N}\boldsymbol{v}_j(t)$ *represent the average velocity of all the cells. Then,*

$$|\boldsymbol{v}_{avg}(t) - v_o\hat{\mathbf{e}}_\mathbf{s}| \leq |\boldsymbol{v}_{avg}(0) - v_o\hat{\mathbf{e}}_\mathbf{s}|e^{-F_o t}$$

Proof.

First we sum the rate of change in velocities of all the cells. $k_1(\mathbf{x}_i, \mathbf{x}_j)$ is a symmetric function, which allows us to simplify further.

$$\begin{aligned}\frac{d}{dt}\sum_{i=1}^{N}\mathbf{v}_i &= \frac{\lambda_1}{N}\sum_{i=1}^{N}\sum_{j=i}^{N}k_1(\mathbf{x}_i, \mathbf{x}_j)(\mathbf{v}_i - \mathbf{v}_j) + \sum_{i=1}^{N}F_o(v_o\hat{\mathbf{e}}_\mathbf{s} - \mathbf{v}_i) \\ &= NF_o v_o\hat{\mathbf{e}}_\mathbf{s} - F_o\sum_{i=1}^{N}\mathbf{v}_i\end{aligned}$$

For conciseness, let $\mathbf{v}_{avg}(t) := \frac{1}{N}\sum_{j=1}^{N}\mathbf{v}_j(t)$, the average velocity of all the cells. Re-writing the above equation, we have a first order linear differential equation.

$$\frac{d}{dt}\mathbf{v}_{avg}(t) + F_o\mathbf{v}_{avg}(t) = F_o v_o\hat{\mathbf{e}}_\mathbf{s}$$

Solving for $\mathbf{v}_{avg}(t)$,

$$\mathbf{v}_{avg}(t) = v_o\hat{\mathbf{e}}_\mathbf{s}\left(1 + \frac{\mathbf{v}_{avg}(0) - v_o\hat{\mathbf{e}}_\mathbf{s}}{v_o\hat{\mathbf{e}}_\mathbf{s}}e^{-F_o t}\right)$$

$$|\mathbf{v}_{avg}(t) - v_o\hat{\mathbf{e}}_\mathbf{s}| \leq |\mathbf{v}_{avg}(0) - v_o\hat{\mathbf{e}}_\mathbf{s}|e^{-F_o t}$$

□

Hence,

$$\lim_{t\to\infty}\mathbf{v}_{avg}(t) = v_o\hat{\mathbf{e}}_\mathbf{s}$$

(i.e. the average velocity of all the cells approaches $v_o\hat{\mathbf{e}}_\mathbf{s}$).

Proposition 2. *Let* $\mathscr{V}(t) := \sum_{i=1}^{N}|\boldsymbol{v}_i(t) - v_o\hat{\mathbf{e}}_\mathbf{s}|^2$. *Then,* $\mathscr{V}(t) \leq \mathscr{V}(0)e^{-2F_o t}$ *and* $\lim_{t\to\infty}\boldsymbol{v}_i(t) = v_o\hat{\mathbf{e}}_\mathbf{s}$.

Proof.

Taking the derivative and plugging in equation (2.2), we obtain

$$\frac{d}{dt}\sum_{i=1}^{N}|\mathbf{v}_i(t) - v_o\hat{\mathbf{e}}_\mathbf{s}|^2 = 2\sum_{i=1}^{N}(\mathbf{v}_i - v_o\hat{\mathbf{e}}_\mathbf{s})\frac{d\mathbf{v}_i}{dt}$$

$$= \frac{2\lambda_1}{N}\sum_{i,j}k_1(\mathbf{x}_i,\mathbf{x}_j)(\mathbf{v}_i - v_o\hat{\mathbf{e}}_\mathbf{s})(\mathbf{v}_j - \mathbf{v}_i) - 2F_o\sum_{i=1}^{N}(\mathbf{v}_i - v_o\hat{\mathbf{e}}_\mathbf{s})^2$$

Since k_1 is symmetric, $\sum_{i,j}k_1(\mathbf{x}_i,\mathbf{x}_j)(\mathbf{v}_j - \mathbf{v}_i) = 0$, therefore

$$\frac{d}{dt}\sum_{i=1}^{N}|\mathbf{v}_i(t) - v_o\hat{\mathbf{e}}_\mathbf{s}|^2 = \frac{2\lambda_1}{N}\sum_{i,j}k_1(\mathbf{x}_i,\mathbf{x}_j)\mathbf{v}_i(\mathbf{v}_j - \mathbf{v}_i) - 2F_o\sum_{i=1}^{N}(\mathbf{v}_i - v_o\hat{\mathbf{e}}_\mathbf{s})^2$$

$$= -\frac{\lambda_1}{N}\sum_{i,j}k_1(\mathbf{x}_i,\mathbf{x}_j)|\mathbf{v}_j - \mathbf{v}_i|^2 - 2F_o\sum_{i}|\mathbf{v}_i - v_o\hat{\mathbf{e}}_\mathbf{s}|^2$$

The lefthand term is nonpositive for $\forall t$, so

$$\frac{d}{dt}\sum_{i=1}^{N}|\mathbf{v}_i(t) - v_o\hat{\mathbf{e}}_\mathbf{s}|^2 \leq -2F_o\sum_{i}|\mathbf{v}_i - v_o\hat{\mathbf{e}}_\mathbf{s}|^2$$

$$= -2F_o\mathscr{V}(t)$$

$$\frac{d}{dt}\mathscr{V}(t) \leq -2F_o\mathscr{V}(t)$$

By Grönwall's inequality,

$$\mathscr{V}(t) \leq \mathscr{V}(0)e^{-2F_o t}$$

meaning

$$\lim_{t\to\infty} \mathbf{v}_i(t) = v_o \hat{\mathbf{e}}_\mathbf{s}$$

for every cell i.

□

Earlier, we showed that the average velocity $\mathbf{v}_{avg}(t)$ asymptotically approaches the same value as $v_o\hat{\mathbf{e}}_\mathbf{s}$. Following the assumption that velocity of each cell is the weighted average of the velocities of cells surrounding it, the velocity of each cel approaches the average velocity, represented by the value $v_o\hat{\mathbf{e}}_\mathbf{s}$. Another interpretation is that since $\mathscr{V}(t)$ is a measure of variance, the variance about the mean approaches 0 asymptotically. Visually, this is represented in Figure 2.1. The velocities of 100 cells are randomly generated. Using these initial conditions, simulations show that they all converge to a common velocity. That is, there is little to no variance in velocity as time progresses.

2.2.1.2 Position

Proposition 3. *For any cell i, j at any time t, $\sup_{t\geq 0} \max_{i,j} |\boldsymbol{x}_j(t) - \boldsymbol{x}_i(t)| < \infty$. Specifically,*

$$|\boldsymbol{x}_i(t) - \boldsymbol{x}_j(t)| \leq |\boldsymbol{x}_i(0) - \boldsymbol{x}_j(0)| + \frac{2\mathscr{V}(0)^{1/2}}{F_0}$$

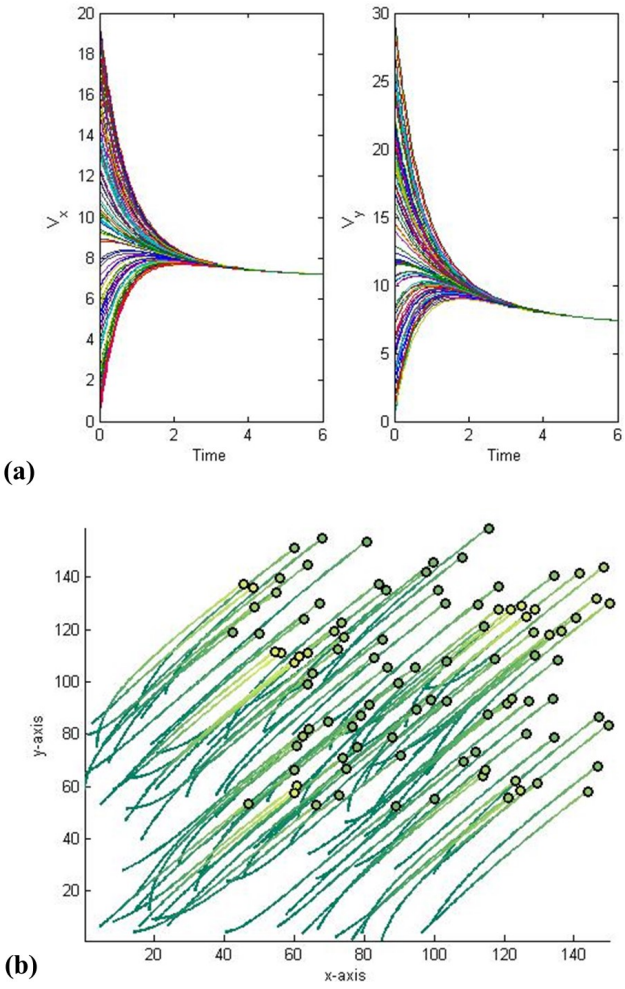


Figure 2.1: (a) Convergence of velocity in the x-direction (V_x) and y-direction (V_y) with $\lambda_1 = 5$ and $\beta = 0.2$ for 100 cells, with initial conditions randomly generated from a uniform distribution on the interval [0,20] and [0,30] for V_x and V_y, respectively. (b) Cells remain bounded (shown here for fixed time). The cells are represented by circles. The trail represents the path travelled by the cell. Initial positions were randomly generated from a uniform distribution on the interval [0,100] for both x and y.

Proof.

For any i, j, using the mean value theorem

$$\begin{aligned}|\mathbf{x}_i(t) - \mathbf{x}_j(t)| &= \left|\mathbf{x}_i(0) + \int_0^t \mathbf{v}_i(s)ds - (\mathbf{x}_j(0) + \int_0^t \mathbf{v}_j(s)ds)\right| \\ &= \left|(\mathbf{x}_i(0) - \mathbf{x}_j(0)) + \int_0^t (\mathbf{v}_i(s) - \mathbf{v}_j(s))ds\right|\end{aligned}$$

Adding and subtracting $(\mathbf{v}_i(s) - \mathbf{v}_j(s))$ as $(\mathbf{v}_i(s) - v_\infty \hat{\mathbf{e}}_\mathbf{s} - (\mathbf{v}_j(s) - v_\infty \hat{\mathbf{e}}_\mathbf{s}))$, we obtain

$$\begin{aligned}|\mathbf{x}_i(t) - \mathbf{x}_j(t)| &\leq |\mathbf{x}_i(0) - \mathbf{x}_j(0)| + \left|\int_0^t (\mathbf{v}_i(s) - v_\infty \hat{\mathbf{e}}_\mathbf{s} - (\mathbf{v}_j(s) - v_\infty \hat{\mathbf{e}}_\mathbf{s}))ds\right| \\ &\leq |(\mathbf{x}_i(0) - \mathbf{x}_j(0))| + \int_0^t |\mathbf{v}_i(s) - v_\infty \hat{\mathbf{e}}_\mathbf{s}| + |\mathbf{v}_j(s) - v_\infty \hat{\mathbf{e}}_\mathbf{s}|ds\end{aligned}$$

Since $\int_0^t |\mathbf{v}_i(t) - v_\infty \hat{\mathbf{e}}_\mathbf{s}| \leq \mathscr{V}(0)^{1/2} \int_0^t e^{-F_0 s} ds$ for any i as shown in Proposition 2,

$$\begin{aligned}|\mathbf{x}_i(t) - \mathbf{x}_j(t)| &\leq |\mathbf{x}_i(0) - \mathbf{x}_j(0)| + 2\mathscr{V}(0)^{1/2} \int_0^t e^{-F_0 s} ds \\ &\leq |\mathbf{x}_i(0) - \mathbf{x}_j(0)| + 2\mathscr{V}(0)^{1/2} \int_0^\infty e^{-F_0 s} ds\end{aligned}$$

Therefore,

$$|\mathbf{x}_i(t) - \mathbf{x}_j(t)| \leq |\mathbf{x}_i(0) - \mathbf{x}_j(0)| + \frac{2\mathscr{V}(0)^{1/2}}{F_0}$$

for all i, j.

□

This means that the distance between any two cells will remain bounded for all t. The boundedness of the set $x_i(t)$ is illustrated in Figure 2.1(b) by showing the paths of bacteria for the time interval $[0, 6]$.

2.2.1.3 Cell Activity

Proposition 4. *Let* $p_{avg}(t) := \frac{1}{N}\sum_{j=1}^{N} p_j(t)$ *represent the average cell activity of all the cells. For the case when* $\psi_A(A_i) = 1$ *at some time* t_o*,*

$$|p_{avg}(t) - p_\infty| \leq |p_{avg}(0) - p_\infty| e^{-L_o t}$$

For the case when $\psi_A(A_i) = 0$ *for all* t*, the differential equation reduces to the standard Cucker-Smale equation and hence,* $\lim_{t\to\infty} \max_{i,j} |p_j(t) - p_i(t)| = 0$ *for* $\beta < 1/2$*.*

Proof.

Similar to the proof of Proposition 1, we start with summing the rate of change of p_i of all the cells. The symmetric property of $k_2(\mathbf{x}_i, \mathbf{x}_j)$ allows for further simplification.

$$\begin{aligned}\frac{d}{dt}\sum_{i=1}^{N} p_i &= \frac{\lambda_2}{N}\sum_{i,j} k_2(\mathbf{x}_i, \mathbf{x}_j)(p_j - p_i) + \sum_{i=1}^{N} L_o(p_\infty - p_i)\psi_A(\mathbf{x}_i, \mathbf{x}_j) \\ &= L_o \sum_{i=1}^{N} (p_\infty - p_i)\psi_A(\mathbf{x}_i, \mathbf{x}_j)\end{aligned}$$

Case: $\psi_A = 1$ *for all* t

For conciseness, let $p_{avg}(t) := \frac{1}{N}\sum_{j=1}^{N} p_j(t)$, the average cell activity of all the cells. Rewriting the above equation, we have a first order linear equation.

$$\frac{d}{dt}\sum_{i=1}^{N} p_i = NL_o p_\infty - L_o N p_{avg}(t)$$

$$\frac{d}{dt} p_{avg}(t) + L_o p_{avg}(t) = L_o p_\infty$$

Solving for $p_{avg}(t)$,

$$p_{avg}(t) = p_\infty \left(1 + \frac{p_{avg}(0) - p_\infty}{p_\infty} e^{-L_o t}\right)$$

$$|p_{avg}(t) - p_\infty| \leq |p_{avg}(0) - p_\infty| e^{-L_o t}$$

Hence,

$$\lim_{t\to\infty} p_{avg}(t) = p_\infty$$

(i.e. the average cell activity of all the cells approaches p_∞).

Case: $\psi_A = 0$ *for all* t

Equation (2.3) simplifies to

$$\frac{dp_i}{dt} = \frac{\lambda_2}{N} \sum_{j\neq i} k_2(\mathbf{x}_i, \mathbf{x}_j)(p_j - p_i)$$

What remains is the Cucker-Smale flocking term. As such, $\lim_{t\to\infty} p_i(t) = \frac{1}{N}\sum_{j=1}^{N} p_j(t^*)$ for some t^*, for $\beta < 1/2$.

□

Proposition 5. *Let* $\mathscr{L}(t) := \sum_{i=1}^{N} |p_i(t) - p_\infty|^2$. *Then,* $\mathscr{L}(t) \leq \mathscr{L}(0)e^{-2L_o t}$ *and* $\lim_{t\to\infty} p_i(t) = p_\infty$.

Proof.

$\mathscr{L}(t) \leq \mathscr{L}(0)e^{-2L_o t}$ with $\mathscr{L}(t) := \sum_{i=1}^{N} |p_i(t) - p_\infty|^2$

$$\begin{aligned}
\frac{d}{dt}\sum_{i=1}^{N}|p_i(t)-p_\infty|^2 &= 2\sum_{i=1}^{N}(p_i-p_\infty)\frac{dp_i}{dt} \\
&= \frac{2\lambda_1}{N}\sum_{i,j}k_1(\mathbf{x}_i,\mathbf{x}_j)(p_i-p_\infty)(p_j-p_i) - 2L_o\sum_{i=1}^{N}(p_i-p_\infty)^2 \\
&= \frac{2\lambda_1}{N}\sum_{i,j}k_1(\mathbf{x}_i,\mathbf{x}_j)p_i(p_j-p_i) - 2L_o\sum_{i=1}^{N}(p_i-p_\infty)^2 \\
&= -\frac{\lambda_1}{N}\sum_{i,j}k_1(\mathbf{x}_i,\mathbf{x}_j)|p_j-p_i|^2 - 2L_o\sum_{i}|p_i-p_\infty|^2 \\
&\leq -2L_o\sum_{i}|p_i-p_\infty|^2 \\
&= -2L_o\mathscr{L}(t)
\end{aligned}$$

$$\frac{d}{dt}\mathscr{L}(t) \leq -2L_o\mathscr{L}(t)$$

By Grönwall's inequality,

$$\mathscr{L}(t) \leq \mathscr{L}(0)e^{-2L_o t}$$

□

Hence, as $t \to \infty$, $\lim_{t\to\infty} p_i(t) = p_\infty$, which we have shown to be the asymptotic mean of the population's cell activity. To put this in the context of *Vibrio fischeri*, each of the cells will collectively express bioluminescence at some average level.

2.2.1.4 Autoinducers

To perform asymptotic analysis on Equation (2.4), we first make assumptions on the form of $k_u(\mathbf{x}_i, \mathbf{x}_j)$. Since the form of Equation (2.5) causes difficulties in

solving the linear differential equation, we focus on the simple case of $k_u(t) = \frac{1}{1+e^{-t}}$. (Note that this logistic function can be generalized but we work with this simple case in order to present a proof less cluttered.) This applies to scenarios in which the rate of uptake is initially low and increases to some stable or maximum uptake rate over time. In other words, this can be observed in scenarios in which the cell density increases up to a certain point and maintains that density.

Proposition 6. *For the case when* $k_u(t) = \frac{1}{1+e^{-t}}$ *(and* $\sum_{j=1}^{N} k_3(\boldsymbol{x}_i, \boldsymbol{x}_j) = M_i$ *for all* $t \geq t^*$*),*

$$\lim_{t\to\infty} A_i(t) = \frac{\lambda_3 r M_i}{k_d + 1}$$

Proof.

$$\frac{dA_i}{dt} + k_u(t)A_i + k_d A_i = \lambda_3 r \sum_{j=1}^{N} k_3(\mathbf{x}_i, \mathbf{x}_j)$$

Multiplying both sides by $e^{\int k_u(s)ds + k_d t}$ in order to perform reverse product rule,

$$\frac{d}{dt}\left(e^{\int k_u(s)ds + k_d t} A_i\right) = \lambda_3 r \sum_{j=1}^{N} k_3(\mathbf{x}_i, \mathbf{x}_j) e^{\int k_u(s)ds + k_d t}$$

$$e^{\int k_u(s)ds + k_d t} A_i = \lambda_3 r \int_0^t \sum_{j=1}^{N} k_3(\mathbf{x}_i, \mathbf{x}_j) e^{\int k_u(s)ds + k_d t} dt$$

Let $G(t) := \int k_u(s)ds$, the antiderivative of $k_u(t)$.

$$A_i = e^{-G(t) - k_d t} \lambda_3 r \int_0^t \sum_{j=1}^{N} k_3(\mathbf{x}_i, \mathbf{x}_j) e^{G(s) + k_d s} ds$$

Since $\sum_{j=1}^{N} k_3(\mathbf{x}_i, \mathbf{x}_j) = M_i$ for all $t \geq t^*$,

$$A_i = e^{-G(t) - k_d t} \lambda_3 r \left(\int_0^{t^*} \sum_{j=1}^{N} k_3(\mathbf{x}_i, \mathbf{x}_j) e^{G(s) + k_d s} ds + \int_{t^*}^{t} M_i e^{G(s) + k_d s} ds \right)$$

The indefinite integral of $k_u(t) = \frac{1}{1+e^{-t}}$ is $ln(1+e^t)$.

$$A_i = \frac{\lambda_3 r e^{-k_d t}}{1+e^t}\left(\int_0^{t^*}\sum_{j=1}^{N} k_3(\mathbf{x}_i,\mathbf{x}_j)e^{G(s)+k_d s}ds + M_i\int_{t^*}^{t}(1+e^s)e^{k_d s}ds\right)$$

Integrating the second integral

$$\begin{aligned}\int_{t^*}^{t}(1+e^s)e^{k_d s}ds &= \int_{t^*}^{t} e^{k_d s} + e^{(k_d+1)s}ds \\ &= \frac{1}{k_d}(e^{k_d t} - e^{k_d t^*}) + \frac{1}{k_d+1}(e^{(k_d+1)t} - e^{(k_d+1)t^*})\end{aligned}$$

and the limit of the second term simplifies to

$$\lim_{t\to\infty}\frac{\lambda_3 r e^{-k_d t}}{1+e^t}M_i\left(\frac{1}{k_d}(e^{k_d t} - e^{k_d t^*}) + \frac{1}{k_d+1}(e^{(k_d+1)t} - e^{(k_d+1)t^*})\right) = \frac{\lambda_3 r M_i}{k_d+1}$$

since the following limits hold:

$$\lim_{t\to\infty}\frac{e^{k_d(t^*-t)}}{1+e^t} = 0, \qquad \lim_{t\to\infty}\frac{e^t}{1+e^t} = 1, \qquad \lim_{t\to\infty}\frac{e^{k_d(t^*-t)+t^*}}{1+e^t} = 0$$

The first integral

$$\lim_{t\to\infty}\frac{\lambda_3 r e^{-k_d t}}{1+e^t}\left(\int_0^{t^*}\sum_{j=1}^{N} k_3(\mathbf{x}_i,\mathbf{x}_j)(1+e^s)e^{k_d s}ds\right)$$

$$\begin{aligned}&\leq \lim_{t\to\infty}\frac{\lambda_3 r N e^{-k_d t}}{1+e^t}\int_0^{t^*}(1+e^s)e^{k_d s}ds \\ &\leq \lim_{t\to\infty}\frac{\lambda_3 r N e^{-k_d t}}{1+e^t}\left(\frac{1}{k_d}(e^{k_d t^*} - 1) + \frac{1}{k_d+1}(e^{(k_d+1)t^*} - 1)\right) \\ &= 0\end{aligned}$$

since the $\lim_{t\to\infty}\frac{e^{-k_d t}}{1+e^t} = 0$. Hence, the limit of the first integral is 0 since the limit of non-negative functions bounded above by 0 signify that the functions converge asymptotically to 0. Putting that all together,

$$\lim_{t\to\infty} A_i(t) = \frac{\lambda_3 r M_i}{k_d+1}$$

□

The surrounding concentration of autoinducers around cell i is ultimately determined by synthesis of autoinducers, the surrounding density of cells, and the rate at which autoinducers are left behind or degraded.

2.3 Simulations and Discussion

The particle-based model presented in this paper was constructed to analyze both movement of cells towards a chemoattractant and autoinducer-triggered gene expression. This system of $4N$ differential equations, in which the four state variables of position, velocity, gene expression, and surrounding autoinducer concentration are assigned to each cell. N represents the total number of cells. The equation governing the surrounding concentration of autoinducer of cell i consists of a weighted synthesis term, an uptake term, and a degradation term. The uptake function can be customized to reflect the system. We initially used a logistic uptake function that was dependent on time to prove that the surrounding concentration of autoinducers converged to a constant value, which differs for each cell. The details are presented in Proposition 2. The assumption here is that the cells grow sufficiently dense enough as time passes that the uptake increases at some point in time and remains at that elevated level. The purpose of this logistic function was to find a simple representation of this general behavior that was with respect to time, instead of position, so that one could run the asymptotic analysis on the differential

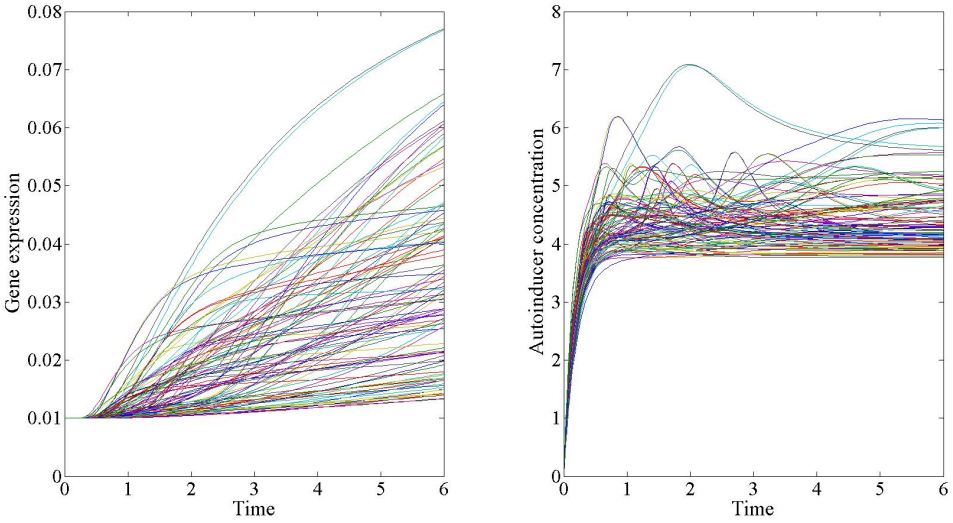


Figure 2.2: Gene expression and surrounding concentration of autoinducers of 100 cells. The combination of the Gaussian uptake and the logistic function $\psi_A(A_i)$ allows for the observed gene expression variation. Parameter values for the Gaussian uptake function: a = 4, b = 4, c = 27.8085 with $\beta = 0.62$ and $\gamma = 1/2000$.

equation.

Simulations were performed using a shifted Gaussian for the uptake function in which at low cell densities, the uptake is low (close to 0). At higher densities, the uptake increases. At even higher densities, the uptake begins to decrease. The reason for this is to account for the effect of overcrowding on metabolic processes. Also, this allows for more interesting dynamics in the surrounding concentration of autoinducers, which leads to a higher variance in gene expression as seen in Figure 2.2. Note that the variation in gene expression here is also attributed to the logistic function, $\psi_A(A_i)$.

The β parameter in the model can be interpreted as the weight placed on the behaviors of the neighbors. When $\beta = 0$, the individual cell places equal weight on all other cells. As β increases, the individual cares less about the behaviors of the cells further away from it. We modified this value to investigate the resulting dynamics. As seen in Figure 2.3, the β values analyzed were $\beta = .5$, $\beta = .58$, $\beta = .66$, and $\beta = .7$. As the β value increases, there is a change in the gene expression dynamics of cells. For lower values of β (i.e. $\beta = .5$, $\beta = .58$), the cells at a higher density are expressing more than the cells at a lower density. In this case, these were the cells that were on the outskirts of the group. As β increases to $\beta = .66$ and $\beta = .7$, the cells at a lower density are expressing more. This is due to the Gaussian uptake function. As the β increases, the difference in positions of the cells (i.e. $|\mathbf{x}_i(t) - \mathbf{x}_j(t)|$) take on a higher value. So the uptake rate approaches 0. Since autoinducers are still being synthesized by our formulation, genes are still being expressed. The cells at a higher density continue to uptake the autoinducers since the distances between cells are smaller than those at a lower density. So at higher values of β, the autoinducer-triggered gene expressions would be lower for cells of higher density than cells of lower density.

Next, we focused on varying the β values on each of the weight functions, k_1, k_2, and k_3. Changing k_1 had the most widespread effect on the dynamics of our system. The changes in velocity governs the positions of the cells. As proved in Proposition 2, the velocity converges for all values of β. In the Cucker-Smale model [24] [44], the differential equation for velocity only consists of the flocking

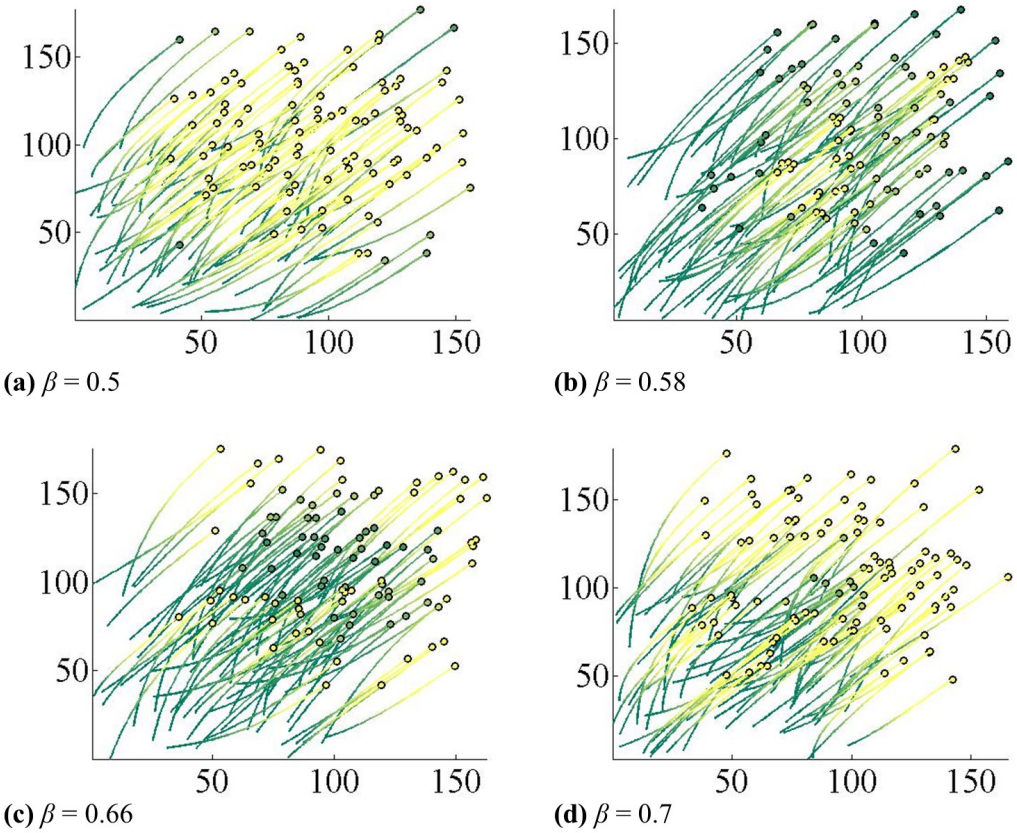


Figure 2.3: (a)-(d) Cell paths for $\beta = .5$, $\beta = .58$, $\beta = .66$, and $\beta = .7$, respectively. The circles represent the cells and the lines represent the path the cells have travelled. Green represents low cell activity and yellow represents higher cell activity. As the β value increases, there is a reversal. The cells which are at a higher density are expressing less than the cells at a lower density. This is due to the Gaussian uptake term, in which the β value increases the value of the sum of the distances between cells. This causes the value of uptake to approach 0 and hence, the higher gene expression.

term and hence, convergence is only guaranteed for $\beta < \frac{1}{2}$ and guaranteed under specific conditions for $\beta \geq \frac{1}{2}$. So it makes sense that changing the strength of the source term in Eq. (2.2) (by changing F_0) and setting the β value to 0.7, for example, could change the rate of convergence in our velocities. This allows more dynamics in the positions of the cells and hence more variation in gene expression and surrounding concentration of autoinducers. In Figure 2.4, we show the effect of a slower convergence in velocity. This is in contrast to the velocity profile shown in Figure 2.1, where the velocities are converging quickly to a common value. The velocities of the cells have not converged to a common value within the time frame shown and the autoinducer profile now has a downward slant, unlike the profile shown in Figure 2.2. If we had increased β_3 to a higher value, that in conjunction with the degradation rate k_d will cause fluctuations that will, in conjunction with the logistic function $\psi_A(A_i)$, influence the variance in gene expression. Altering the β values of k_2 along with the L_0 coefficient will also influence the effects of velocity and autoinducer concentration on gene expression. Hence, the change in β, or the change in the extent to which the neighbors influence the individual, causes variation in the dynamics of our system. Specifically, setting $\beta > \frac{1}{2}$ and weakening the effect of the source term by lowering the values of coefficients F_0 and L_0 promotes more interesting behaviors in our cells.

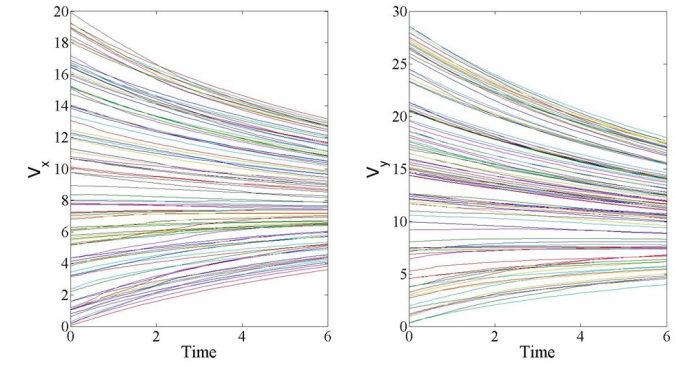


Figure 2.4: (a) Velocity profile for 100 cells in the x-direction and y-direction. The slowness in the rate of convergence causes the downwards slant in the autoinducer profile shown in (b). Here, $F_0 = 0.1$, $\beta_1 = 0.7$, $\beta_2 = 0.7$, $\beta_3 = 0.2$.

2.4 Conclusion

We introduced a method of modelling quorum sensing that does not a priori require knowledge of the intrinsic protein interactions that are necessary to communicate with other bacteria. Instead, the autoinducer-triggered cell activity changes with respect to the cell activity of neighboring cells. That is, the flocking term generalizes quorum sensing in the form of weighted averages.

This keeps the ordinary differential equations simple enough to perform asymptotic analysis while encapsulating the dynamics of bacterial communication in tandem with chemotaxis. This method of modelling also allows us to analyze these dynamics in 2 and 3-dimensions.

Since the cell activity is independent of the known protein interactions that differ from species to species, this kind of flocking representation allows for the generalized modelling of positive feedback and negative feedback mechanisms in a variety of quorum-sensing bacteria.

Future directions include adapting stochastic elements to the system that would account for the variations observed in movement and in behavior. This would also allow for analysis of what kinds of random processes applied to movement, velocity, cell activity, and autoinducer concentration would affect the modelled system that would facilitate a closer representation of bacterial chemotaxis and quorum sensing interactions.

Chapter 3: Modeling the interplay of quorum sensing and chemotaxis

3.1 Introduction

Bacteria exhibit collective gene expression called quorum sensing, which was first discovered when specific phenotypes were collectively expressed by bacteria at a certain density [8]. Further investigations revealed the presence of chemical signals called autoinducers which the bacteria both produce and collect. At some threshold in concentration of autoinducers, signal transduction networks are triggered causing coordinated expression of specific genes. This can be observed in *Vibrio harveyi* through the collective production of bioluminescence. In *Pseudomonas aeruginosa*, quorum sensing enables the production of various virulence factors as well as biofilms [93]. This kind of molecular communication system allows bacteria to collectively exhibit these various phenotypes.

Another form of dynamical behavior observed in bacteria is chemotaxis. Chemotaxis is the movement of bacteria triggered in response to a gradient in concentration of particular substances which are classified as either attractants or repellents. The changes in concentrations of these substances activate a cascade of signal transduc-

tion mechanisms that affect motility [10].

Not much is known about the direct connections between the quorum sensing system and the chemotaxis system. Any connections that exist tend to describe the interactions or effects of one major protein in one system influencing the other system in some way. For example, AI-2 is a known attractant for *E. coli* [7] [32] [48] and the LsrB protein, which is responsible for AI-2 uptake, is necessary for the chemotaxing to AI-2 [48]. Here, the LsrB protein from the quorum sensing system is integral to the chemotaxis system, with regards to the movements toward AI-2. Another example of a connection between quorum sensing and chemotaxis is that there was an up-regulation in motility and production of flagella in *luxS* knockouts in *E. coli*. *luxS* is the gene that contributes to AI-2 production [64] in the quorum sensing system and removing this gene had caused changes to the chemotaxis system.

To explore the cellular behavior resulting from the possible interactions between these two complex systems, we look to differential equations. The regulatory systems in quorum sensing and chemotaxis consist of various transcriptions and translations of genes to proteins that contribute to the dynamical behaviors as well as enzymatic activities (e.g. phosphorylation, induction) that are involved in these processes [110] [111]. Taking into account the various reactions and regulatory mechanisms of these systems can be complicated so instead, we aim to generalize the behaviors of the quorum sensing and chemotaxis system with help from the concept of flocking.

Flocking in nature is the description applied towards the behavior of a group

of birds in flight [31]. The ways in which birds maintain such organized formations while flying has led to several theories and various mathematical models and simulations [24] [45] [82] [88]. One such model is the Cucker-Smale model [24] [25]. This is a dynamical model constructed from ordinary differential equations that demonstrates that the speed of each bird approaches the speed of the flock, under the assumption that each bird adjusts its velocity by weighing the velocities of its neighbors. This model was expanded upon by Ha & Tadmor [44] and the system is described as:

$$\frac{d}{dt}\mathbf{x}_i(t) = \mathbf{v}_i(t)$$
$$\frac{d}{dt}\mathbf{v}_i(t) = \frac{\lambda}{N}\sum_{1\leq j\leq N} k(\mathbf{x}_i(t), \mathbf{x}_j(t))(\mathbf{v}_j(t) - \mathbf{v}_i(t))$$

The state variables are the position of particle i ($x_i(t)$) and the velocity of particle i ($v_i(t)$), with N representing the total number of particles. The velocity of particle i changes according to the neighboring particles j, with the function $k(\mathbf{x}_i(t), \mathbf{x}_j(t))$ acting as weights. That is, the function determines the scope of influence of the surrounding particle j on the individual particle i. The λ acts as a coefficient representing the strength of this flocking term.

The definition of flocking in [72] requires two conditions to be met.

Definition 1. *The N agents of a particle system, with each agent i's distances and velocities represented by $\{\boldsymbol{x}_i(t), \boldsymbol{v}_i(t)\}_{i=1,\dots,N}$, converge to a flock provided the following properties hold:*

(a) $\sup_{t\geq 0} \max_{i,j} |\mathbf{x}_j(t) - \mathbf{x}_i(t)| < \infty$

(b) $\lim_{t\to\infty} \max_{i,j} |\mathbf{v}_j(t) - \mathbf{v}_i(t)| = 0$

In order to classify a group of N particles as flocking, the distances between agent i and agent j remain bounded for all time t and the velocities of each agent approach a common velocity.

This Cucker-Smale flocking method of modelling group dynamics has been applied to cells. Di Costanzo et al. [28] employed flocking to model the process of morphogenesis of cells of the posterior lateral line primordium in zebrafish. Ha & Levy [43] incorporated flocking in the modelling of phototaxis of cyanobacterium *Synochocystis* sp.. In this paper, building on the work of Ha & Levy [43] and Ha & Tadmor [44], we introduce flocking behavior as a representation of quorum sensing behavior.

The goal of this paper is to suggest a minimal model that explores the dynamical behaviors that result from the coupling of the biological processes of quorum sensing and chemotaxis. The regulatory systems involved in quorum sensing differ from species to species. By employing Cucker-Smale flocking terms, we can model positive and negative feedback properties of quorum sensing without accounting for the specific signalling mechanisms involved.

Our general model looks at the dynamics of the following four state variables. Each cell i's position, velocity, autoinducer-triggered gene expression, and surrounding autoinducer concentration is focused on.

1. $\mathbf{x}_i$: position of cell i
2. $\mathbf{v}_i$: velocity of cell i
3. p_i: gene expression of cell i

4. A_i: autoinducer concentration surrounding cell i

The behavior of the state variables are governed by the following system of differential equations.

$$\frac{d\mathbf{x}_i}{dt} = \mathbf{v}_i \tag{3.1}$$

$$\frac{d\mathbf{v}_i}{dt} = \frac{\lambda_1}{N}\sum_{j=1}^{N} k_1(\mathbf{x}_i, \mathbf{x}_j)(\mathbf{v}_j - \mathbf{v}_i) + F_0(v_o\hat{\mathbf{e}}_\mathbf{s} + (v_\infty - v_o)\hat{\mathbf{e}}_\mathbf{s}\psi_p(p_i) - \mathbf{v}_i) \tag{3.2}$$

$$\frac{dp_i}{dt} = \frac{\lambda_2}{N}\sum_{j=1}^{N} k_2(\mathbf{x}_i, \mathbf{x}_j)(p_j - p_i) + L_0(p_\infty - p_i)\psi_A(A_i) \tag{3.3}$$

$$\frac{dA_i}{dt} = \lambda_3\sum_{j=1}^{N} k_3(\mathbf{x}_i, \mathbf{x}_j)r - k_u(\mathbf{x}_i, \mathbf{x}_j)A_i - k_dA_i \tag{3.4}$$

If there are N cells in the modelled system, there are $4N$ ordinary differential equations.

The differential equations for velocity and gene expression (equation (3.2) and equation (3.3)) consist of the flocking term and source term. The flocking term is the weighted adjustment of the velocity (or gene expression) of cell i with respect to the velocities (or gene expression) of all other cell j's. The $k_1(x_i, x_j)$ and $k_2(x_i, x_j)$ are symmetric function with values which depend on distances between cells.

$$k_1(\mathbf{x}_i, \mathbf{x}_j) = \frac{1}{(1 + |\mathbf{x}_j - \mathbf{x}_i|^2)^{\beta_1}}$$

$$k_2(\mathbf{x}_i, \mathbf{x}_j) = \frac{1}{(1 + |\mathbf{x}_j - \mathbf{x}_i|^2)^{\beta_2}}$$

The weight function takes on higher values for cells closer to cell i, meaning that cells would adjust their velocities and gene expression according to the velocities and expression of cells in closer proximity than those further away. The β power represents the scope of influence that the neighbors would have on cell i. For low β

values (e.g. close to 0), the cells further away still have weight values that are closer to the weight values of cells that are nearby. As β increases, the cells further away have much less influence.

The strength of the source term for velocity is governed by constant F_0 and $\hat{\mathbf{e}}_\mathbf{s}$ stands for the unit direction vector. $\psi_p(p_i)$ is a logistic equation that takes values between 0 and 1 depending on the values of gene expression corresponding to that cell.

$$\psi_p(p_i) = \frac{1}{(1 + e^{c_0(c_1 - c_p p_i)})^{\alpha_1}}$$

If $\psi_p = 0$ for all t, the source term reduces to $F_0(v_o \hat{e}_s - v_i)$. If $\psi_p = 1$ for all t, the source term becomes $F_0(v_\infty \hat{\mathbf{e}}_\mathbf{s} - \mathbf{v_i})$. v_o is the magnitude for the basal level of velocity. After the gene expression approaches a certain amount of activity, the cell's velocity increases, approaching v_∞ if gene expression p_i continues to increase.

The source term for the differential equation governing autoinducer-triggered gene expression is not as complicated. The strength of this term is controlled by the constant, L_0. $\psi_A(A_i)$ is a logistic function that takes values between 0 and 1 depending on the value of surrounding concentration of autoinducer for that cell.

$$\psi_A(A_i) = \frac{1}{(1 + e^{c_2(c_3 - c_A A_i)})^{\alpha_2}}$$

If $\psi_A = 0$, the source term disappears altogether. If $\psi_A = 1$, then the gene expression will approach some maximal expression level. Biologically, this ψ_A represents the putative threshold of surrounding autoinducer concentration required for quorum sensing-related expression of, for example, luciferase in *Vibrio harveyi* or biofilm in *Pseudomonas aeruginosa*.

The differential equation governing the surrounding concentration of autoinducer (equation (3.4) does not contain Cucker-Smale terms. However, the first term does use a symmetric weight function $k_3(x_i, x_j)$ that is the same as the Cucker-Smale weights.

$$k_3(\mathbf{x}_i, \mathbf{x}_j) = \frac{1}{(1 + |\mathbf{x}_j - \mathbf{x}_i|^2)^{\beta_3}}$$

This equation consists of three parts: the synthesis term, the uptake term, and the degradation/sloughing-off term. The synthesis term is a weighted term with r representing the rate of synthesis of autoinducers by the cells. The higher the rate of synthesis, the more autoinducers will accumulate in the surroundings. The uptake term represents the concentration of autoinducers that are collected by the cells and thus removed from the surroundings. For the purposes of this model, a shifted Gaussian was used. The reason for the use of this kind of function lies in the ability of the autoinducer uptake rate of each cell depending on the surrounding density. At higher densities of cells, a sufficently higher autoinducer concentration in the surroundings triggers an increase in uptake rate. As the density increases, the cells become crowded enough that there is a reduction in cellular metabolic processes. So at some sufficiently high density, there is a reduction in uptake rate. Hence, the shape of the Gaussian proves useful in this scenario. The last term accounts for the autoinducers that are degraded or left behind as a result of movement towards the chemoattractant.

$$k_u(\mathbf{x}_i, \mathbf{x}_j) = e^{\frac{-(a(\sum_{j=1}^{N} |\mathbf{x}_j - \mathbf{x}_i|^{2\beta_3}) - b)^2}{2c}} \tag{3.5}$$

3.2 Sensitivity Analysis

The resulting dynamics produced by the model is very much dependent on the autoinducer concentration. Biologically, this makes sense as gene expression is powered by the autoinducers in the surrounding medium. We can view the changes in behavior of the model by exploring the sensitivities of various parameter values.

3.2.1 Adjusting the β values

β_1, β_2, and β_3, which we will also refer to as the β values, are the parameters that control the behavior of the weighted values in the Cucker Smale flocking terms in equations (3.2 and 3.3) and in the rate of synthesis term in equation (3.4). Qualitatively, the β values represent the scope of influence of the neighbors on the individual. For low values of β (e.g. $\beta = 0$), every cell influences the individual. As β values increase, the dynamics of the cells furthest away from the individual exert little to no influence. To observe how this scope of influence affects the velocity, gene expression, and surrounding concentration of autoinducers, we adjust the values of β_1, β_2, and β_3.

First, we analyze the dynamics of the state variables by varying β_1 and $\beta_2 = \beta_3$. After cycling through many β values, we focus on the two values of 0.2 and 0.7 to demonstrate the effects of these parameters on our model. These values were chosen because there was a change in the profiles of our state variables when the β values passed 0.5.

When all the β values are 0.2, the autoinducer concentration remains high

so all cells exhibit the same expression profile. The velocity profiles initially show flocking behavior, in which all the different initial conditions start converging to a common velocity. Then, the threshold determined by $\psi_p(p_i)$ kicks in, causing an increase in velocity. As the gene expression values for all cells are the same, all cells increase in velocity at the same time. If $\beta_1 = 0.7$, while β_2 and β_3 remain at 0.2, the velocity profiles of each of the cells show more variation before converging to a common velocity.

When $\beta_1 = 0.2$, while β_2 and β_3 are raised to 0.7, the autoinducer concentration shows much variance. This is a result of the higher β_3 value (as opposed to the 0.2), the uptake function, and the degradation parameter of k_d. Consequently, the gene expression values are different for each cell as seen in Figure 3.1. The velocities of the cells initially take on an assortment of values but the $\beta_1 = 0.2$ ultimately steers the cell velocities to take on the same values. If β_1 is raised to 0.7, the cell velocities don't converge to the same value within the time frame shown in Figure 3.1. Effectually, we see the most variation in the profiles of velocity, gene expression, and autoinducer concentration when all β values are set to 0.7.

If we consider situations when β_2 and β_3 take on different values, most of the variation in autoinducer-activated phenotypic expression is observed when $\beta_3 \geq 0.5$.

3.2.2 Adjusting the α values

Parameters that can also influence our model are those associated with the logistic function of our source terms. Specifically, we focus on varying the α_1 and

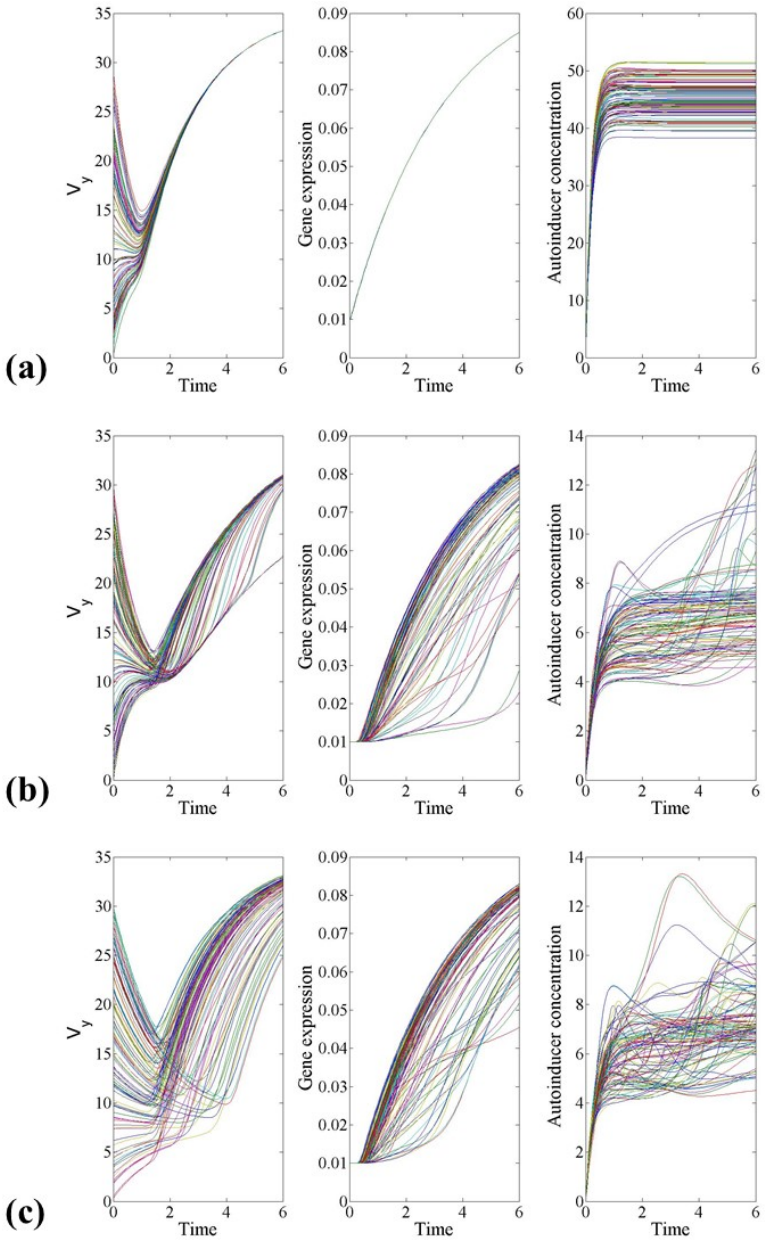


Figure 3.1: Profiles of velocity in y-direction (V_y), gene expression, and autoinducer concentration. (a) $\beta_1 = 0.2$, $\beta_2 = \beta_3 = 0.2$. The autoinducer concentration remains high enough that we obtain convergence in V_y and in gene expression. (b) $\beta_1 = 0.2$, $\beta_2 = \beta_3 = 0.7$. Fluctuations of autoinducer concentration cause variation in velocity and gene expression. (c) $\beta_1 = 0.7$, $\beta_2 = \beta_3 = 0.7$. The variance in velocity causes more fluctuations in autoinducer concentration.

α_2 values (which we will also refer to as the α values) of $\psi_p(p_i)$ and $\psi_A(A_i)$ while fixing all other parameter values. $\psi_p(p_i)$ controls the threshold in which the motility of the cell is upregulated, as a result of gene expression and indirectly, autoinducer concentration. $\psi_A(A_i)$ is the concentration threshold for autoinducers, in which specific genes are upregulated because of an accumulation of autoinducers. Changing the location of the threshold has a greater influence on the dynamics of velocity, gene expression, and autoinducer concentration than from adjusting the β values. Note that modifying the c_p and c_A values will also influence the location of the thresholds. But we fix those values and toggle the α values instead, as the effect of moving the thresholds is the same regardless of the means in which we change the threshold values.

Raising the α values results in the raising of the threshold. That is, more expression is required for an increase in velocity and more autoinducers are necessary for an increase in gene expression. When α values are high (e.g. $\alpha_1 = \alpha_2 = 50$) as shown in Figure 3.2, the velocity of a majority of cells show an alignment in velocity. Gene expression levels remain low for most of the cells. When the threshold for $\psi_A(A_i)$ is lowered, (e.g. $\alpha_1 = 50$, $\alpha_2 = 5$), we see as a result that the cells express a variety of expression, causing one group of cells to converge to one velocity and another group of cells (ie the cells which have expressed high enough gene expression so that $\psi_p \neq 0$) are gravitating towards a higher velocity. Lowering both threshold values (e.g. $\alpha_1 = \alpha_2 = 1$) resulted in all cells approaching a higher velocity and higher expression levels. Since velocity of cells influence their positions, the profiles of autoinducer concentration show more dynamics when the velocities fluctuate.

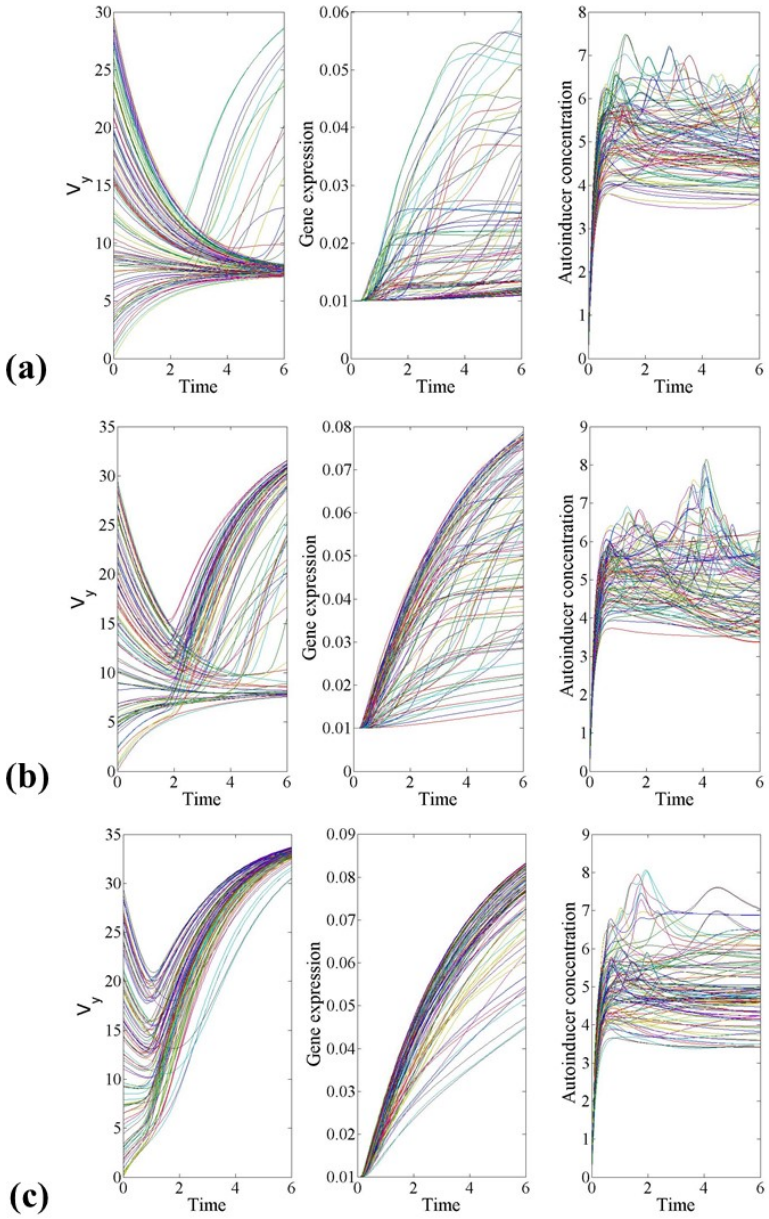


Figure 3.2: Effects of varying α parameters: profiles of velocity in y-direction (V_y), gene expression, and autoinducer concentration. (a) α_1 = 50, α_2 = 50. Majority of cells converge to a lower velocity. Gene expression are varied but mostly stay low. (b) α_1 = 50, α_2 = 5. Two groups of cells approach either a higher or lower velocity. High variance of gene expression. (c) α_1 = 1, α_2 = 1. All cells approach a higher velocity. Gene expression is the least varied with most cells exhibiting higher gene expression than in (a) and (b).

3.3 Discussion and application

The $4N$ system of differential equations can be customized for quorum sensing systems by using different threshold values, which are controlled by changing α_1 and α_2 (or alternatively, c_p and c_A), modifying the scopes of influence (β_1, β_2, β_3), customizing the uptake function in Equation 3.4 (e.g. Gompertz function rather than Gaussian), and toggling the strength of each of the flocking and source terms (λ_1, λ_2, λ_3, F_0, L_0).

When we set our α values as the same, we observed notable dynamics in our cells: Two groups of cells had formed. We conclude that the reason for this behavior in our model is attributed to the coupling of quorum sensing and chemotaxis systems. The rationale for this conclusion is that in our model from Chapter 2, the cells maintain one single group with only a portion of the group expressing more than the rest. Since quorum sensing does not influence velocity in that model, cells with higher expression would not move faster. When we couple the model such that gene expression influences velocity, the portion with higher expression would move faster, leaving the rest of the group behind. Not only would we observe that different cells would have different velocities, we can observe the formation of two groups as shown in Figure 3.3.

The organization of cells into two groups is consistent with experimental observations. In chemotaxis experiments conducted using capillary tubes, where *E. coli* was placed on one end and the chemoattractant (galactose) was placed on the other end, *E. coli* moved towards galactose. What was unexpected was the appearance of

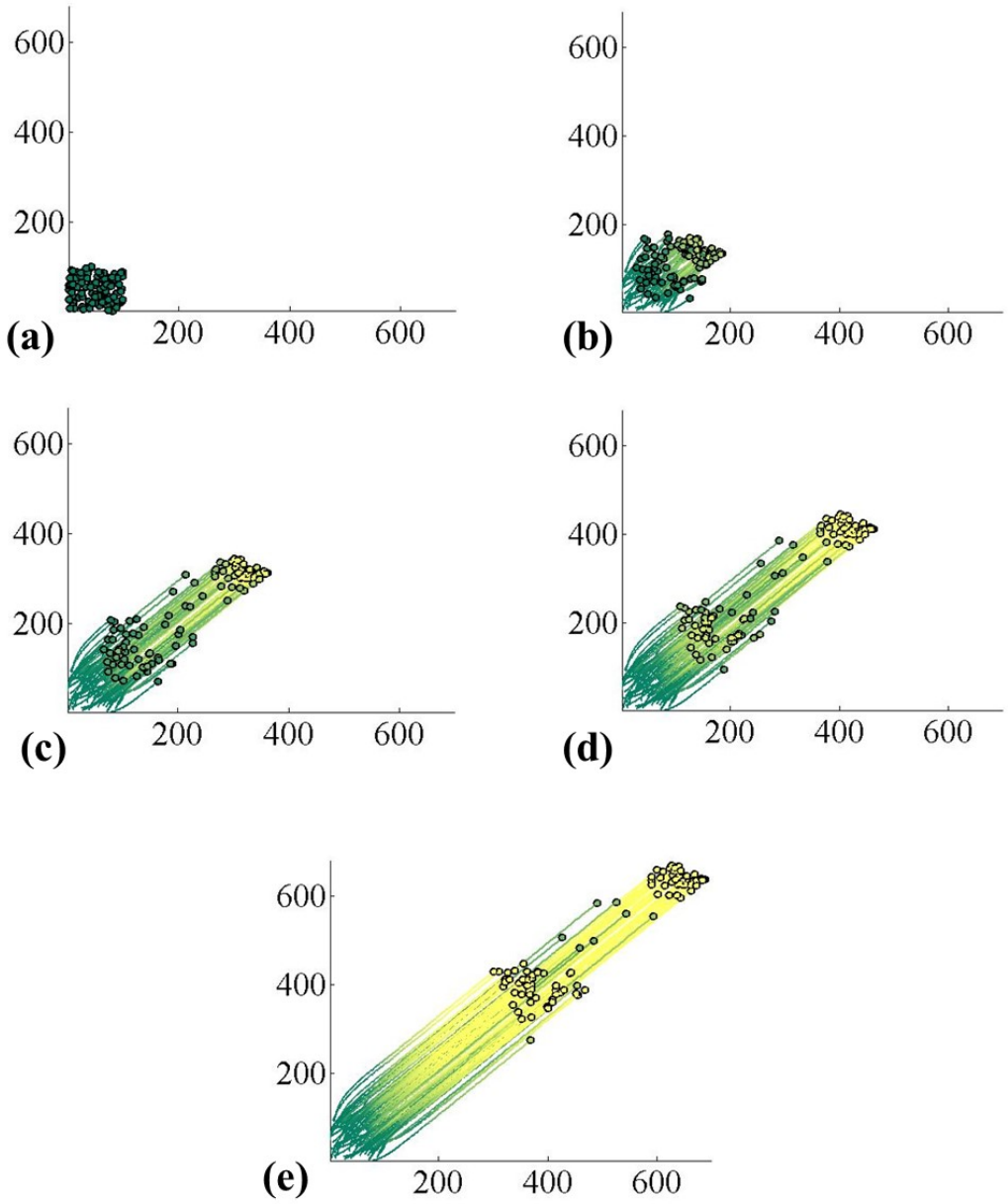
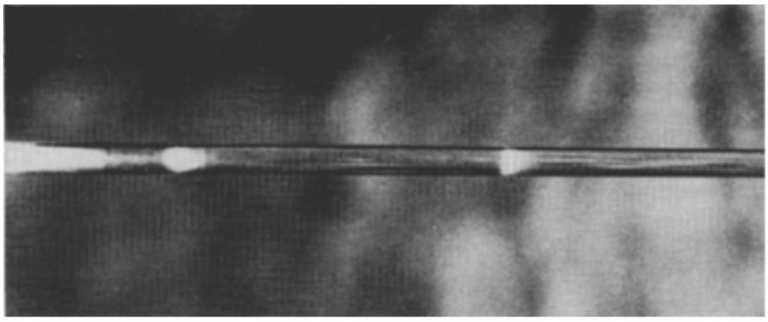


Figure 3.3: Simulation resulting in formation of two groups: (a) Randomly placed cells in [0,100] $\times$ [0,100] move towards source of attractant. (b) The dense portion of cells reach a higher velocity, which (c)-(d) splits away from the rest of the group and (e) forms two groups.



(Adapted from Adler et al. (1966))

Figure 3.4: Chemotaxis experiment in Adler et al. [1] in which *E. coli* (white) were placed at the left end of the capillary tube and galactose (chemoattractant) was placed on the opposite end. The bacteria split into two groups as they chemotaxed, as shown by the two white bands.

two bands, shown in Figure 3.4. To prove that this behavior was not attributed to the heterogeneity in the cells, they scraped the cells from one of the two bands and repeated the experiment with those cells. Two groups of cells formed, yet again.

Our model presents the possibility that quorum sensing systems, or some kind of density-dependent mechanisms, are connected to the chemotaxis system. The increased gene expression caused by the density of cells and the subsequent separation of this group from the original pack would explain the formation of two sets of cells during the process of chemotaxis.

What kind of response would we see when $\alpha_1 \neq \alpha_2$? What kind of experiment would we observe an $\alpha_1 \neq \alpha_2$ scenario? We first respond to the latter question by describing the experiments conducted in a transwell apparatus in Servinsky et al. [97].

The transwell experiments presented the movement of recruitable cells towards AI-2 (also a chemoattractant) and the monitoring of the expression of green fluorescent protein (GFP) as well as constitutively expressed DsRed. There are two sets of cells in this experiment. One set are sentinel cells that secrete AI-2. These are placed in the lower chamber of the transwell. The other set of cells, called recruitable cells, are genetically engineered to produce GFP when transcription is induced by autoinducers. Located in the upper chamber, these cells also constitutively express DsRed as a method of indicating their accumulation in the lower chamber of the transwell. These recruitable cells are not engineered to produce AI-2 so that the only source of AI-2 was the sentinel cells located in the lower chamber. The schematic of this arrangement is visualized in Figure 3.5(a). The results of this experiment showed that signs of DsRed, but not GFP, were detected at the 16 hour point. However, both GFP and DsRed appear at the 40 hour time point. The conclusion was that the motility response is more sensitive than the quorum-sensing-triggered gene expression response with respect to AI-2. In such a scenario, $\alpha_1 < \alpha_2$.

We model a similar system in which the chemotaxis system is more sensitive than the (quorum-sensing-mediated) gene expression system. This can be modelled by lowering the threshold of $\psi_p(p_i)$ (ie the value of p_i for which the logistic function increases to 1) and raising the threshold of $\psi_A(A_i)$. That is, setting $\alpha_1 < \alpha_2$. In this simulation, we mimic the upper and lower chambers of the transwell apparatus by assigning some fixed distance from the origin as the "membrane". We recorded the number of cells that pass this "membrane" as well as their gene expression in our simulation. We classify the cells that express greater than 0.7 as fully expressing

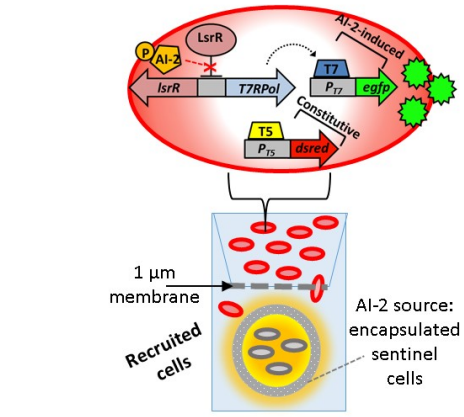
GFP. Figure 3.5(c) shows the increasing cell count in the "lower chamber" of our simulation as well as the increasing fraction of cells fully expressing. We compare that to the results in Servinsky et al. [97], shown in Figure 3.5(b) and we observe qualitatively similar results: As time passes the cell count increases but GFP is not expressed until afterwards.

3.4 Conclusion

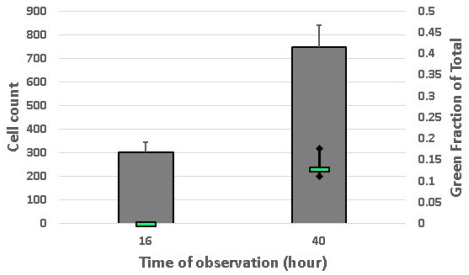
We presented a simple model representing a quorum-sensing, chemotaxing group of bacteria using Cucker-Smale flocking terms. The model requires little knowledge behind the specific signal transduction and other regulatory mechanisms involved in either quorum sensing or chemotaxis. The flocking terms as well as the source terms of our model generalize the complex signalling process at play.

The logistic functions of $\psi_p(p_i)$ and $\psi_A(A_i)$ couple together our system of $4N$ differential equations (Equation (3.1) - Equation (3.4)). Velocity is influenced by gene expression after a certain threshold is attained. Gene expression is influenced by autoinducers after a certain threshold is reached. While there exists a threshold of autoinducers required before triggering the expression of phenotypes such as biofilm or bioluminescence, the connection between gene expression via the quorum sensing system and the chemotaxis system remains to be proven. That is, few research exists on the upregulation of velocity due to autoinducer-related gene expression.

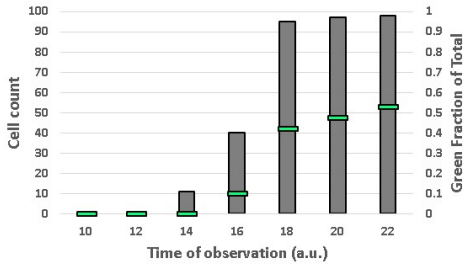
In our model we suggest a possible connection between these two systems. Our simulations show the formation of two groups of cells, as shown in Figure 3.3.



(a) (Adapted from Servinsky et al. (2016))



(b) (Adapted from Servinsky et al. (2016))



(c)

Figure 3.5: Comparison of Servinsky et al. [97] results to our simulation. (a) In the transwell, the recruitable cells in the upper chamber and the sentinel cells in the lower chamber are separated by a membrane. The recruitable cells constitutively express DsRed and express green fluorescent protein (GFP) in the presence of AI-2. (b) The bar chart shows the number of cells in the lower chamber at hour 16 and hour 40. The fraction of cells expressing GFP are shown by the green bars. (c) Our simulation in which we observe qualitatively similar results: The fraction of cells expressing is initially 0 and increases as time passes.

If our system of differential equations were decoupled so that chemotaxis was not dependent on quorum sensing (ie $\psi_p(p_i) = 0$) as was the case in Chapter 2, the cells remain in one group. The two groups that formed in our model was the consequence of autoinducer-triggered gene expression influencing the velocity of our cells. This suggests the possibility that chemotaxis is influenced in some way by another regulatory process that is density dependent. In our case, we demonstrated that a coupling of quorum sensing with chemotaxis can lead to behavior consistent with experimental observations [1]. When we simulated a system in which the chemotaxis response is more sensitive than the quorum-sensing-triggered gene expression response, we were able to obtain trends similar to those produced by transwell experiments in Servinsky et al. [97].

Chapter 4: Application of plasma theory to bacteria

4.1 Introduction

Plasma is a gas consisting of charged particles (ions and electrons) and plasma kinetic equations are differential equations that describe the dynamics of these particles. The goal of this paper is to treat bacteria as these charged particles in order to derive partial differential equations that describe the evolution of these systems over time.

We begin by discussing the Liouville equation, one of the canonical equations of plasma dynamics. This looks at the motion of an ensemble of particles described as a product of delta distributions. The time evolution of each particle is analyzed on the phase space, $(\mathbf{x}, \mathbf{v})$, which represent the random variables of position and velocity of a particle.

Traditionally, differential equations are used to study the movements and expression of bacteria. For a system in which the dynamics of every bacteria are tracked, this would require many initial conditions and we would be working under the assumption that their behaviors are deterministic. Instead, we turn to a probabilistic representation to characterize our ensemble of bacteria. For this paper, we turn to statistical mechanics for assistance [5] [53] [76] [86] [87]. That is, we treat

bacteria like a system of particles in which the density of a particular point in time, position, velocity, and any other state variables is assigned a probability distribution, f^N. Instead of Maxwell's equations and Newton's laws governing the movement of electrons, we have a flocking model governing the various characteristics of bacteria. The flocking model is as follows:

$$\frac{d\mathbf{X}_i}{dt} = \mathbf{V}_i \tag{4.1}$$

$$\frac{d\mathbf{V}_i}{dt} = \frac{\lambda_1}{N}\sum_{j=1}^{N} k_1(\mathbf{X}_i, \mathbf{X}_j)(\mathbf{V}_j - \mathbf{V}_i) + F_0(v_o\hat{\mathbf{e}_\mathbf{s}} - \mathbf{V}_i) \tag{4.2}$$

$$\frac{dP_i}{dt} = \frac{\lambda_2}{N}\sum_{j=1}^{N} k_2(\mathbf{X}_i, \mathbf{X}_j)(P_j - P_i) + L_0(p_\infty - P_i)\psi_A(A_i) \tag{4.3}$$

$$\frac{dA_i}{dt} = \frac{\lambda_3}{N}\sum_{j=1}^{N} k_3(\mathbf{X}_i, \mathbf{X}_j)r - (k_u + k_d)A_i \tag{4.4}$$

where with the corresponding weights:

$$k_1(\mathbf{X}_i, \mathbf{X}_j) = k_2(\mathbf{X}_i, \mathbf{X}_j) = k_3(\mathbf{X}_i, \mathbf{X}_j) = \frac{1}{(1 + |\mathbf{X}_j - \mathbf{X}_i|^2)^\beta}$$

and logistic function:

$$\psi_A(A_i) = \frac{1}{(1 + e^{c0(c_1 - c_A A_i)})^\alpha}$$

The uptake rate of k_u and degradation rate of k_d are both constant here. First, we derive the modified Liouville Equation that will incorporate our flocking model. Next, we show a formal derivation of the BBGKY hierarchy similar to what is done in Ha & Levy [43].

4.2 Derivation of a Modified Liouville Equation

The density of a system of N particles in phase space is written as

$$\begin{aligned} &L(\mathbf{x}_1, \mathbf{v}_1, p_1, a_1, ..., \mathbf{x}_N, \mathbf{v}_N, p_N, a_N, t) \\ &\qquad = \prod_{i=1}^{N} \delta[\mathbf{x}_i - \mathbf{X}_i(t)]\delta[\mathbf{v}_i - \mathbf{V}_i(t)]\delta[p_i - P_i(t)]\delta[a_i - A_i(t)] \\ &\qquad = \prod_{i=1}^{N} \delta_{x,v,p,a,i} \end{aligned}$$

Here, we introduce the shorthand notation of

$$\delta_{\mathbf{x},\mathbf{v},p,a,i} = \delta[\mathbf{x}_i - \mathbf{X}_i(t)]\delta[\mathbf{v}_i - \mathbf{V}_i(t)]\delta[p_i - P_i(t)]\delta[a_i - A_i(t)]$$

to represent the density of particle i. If we were in 3-dimensional space,

$\delta[\mathbf{x}_i - \mathbf{X}_i] = \delta(x_i - X_i)\delta(y_i - Y_i)\delta(z_i - Z_i)$.

We take the derivative of $L(\mathbf{x}_1, \mathbf{v}_1, p_1, a_1, ..., \mathbf{x}_N, \mathbf{v}_N, p_N, a_N, t)$ with respect to time

$$\begin{aligned} \frac{\partial L}{\partial t} = \sum_{i=1}^{N} \Bigg(&\frac{\partial}{\partial t}(\delta[\mathbf{x}_i - \mathbf{X}_i])\delta_{v,p,a,i} \prod_{k\neq i} \delta_{x,v,p,a,k} + \frac{\partial}{\partial t}(\delta[\mathbf{v}_i - \mathbf{V}_i])\delta_{x,p,a,i} \prod_{k\neq i} \delta_{x,v,p,a,k} \\ &+ \frac{\partial}{\partial t}(\delta[p_i - P_i])\delta_{x,v,a,i} \prod_{k\neq i} \delta_{x,v,p,a,k} + \frac{\partial}{\partial t}(\delta[a_i - A_i])\delta_{x,v,p,i} \prod_{k\neq i} \delta_{x,v,p,a,k} \Bigg) \end{aligned}$$

and apply chain rule

$$\begin{aligned} \frac{\partial L}{\partial t} = -\sum_{i=1}^{N} \Bigg(&\nabla_{\mathbf{x}_i} \cdot \left(\frac{d\mathbf{X}_i}{dt} \prod_{k=1}^{N} \delta_{x,v,p,a,k} \right) + \nabla_{\mathbf{v}_i} \cdot \left(\frac{d\mathbf{V}_i}{dt} \prod_{k=1}^{N} \delta_{x,v,p,a,k} \right) + \\ &\frac{\partial}{\partial_{p_i}} \left(\frac{dP_i}{dt} \prod_{k=1}^{N} \delta_{x,v,p,a,k} \right) + \frac{\partial}{\partial_{a_i}} \left(\frac{dA_i}{dt} \prod_{k=1}^{N} \delta_{x,v,p,a,k} \right) \Bigg) \end{aligned}$$

which simplifies to

$$\frac{\partial L}{\partial t} + \sum_{i=1}^{N} \left(\nabla_{\mathbf{x}_i} \cdot \left(\frac{d\mathbf{X}_i}{dt} L \right) + \nabla_{\mathbf{v}_i} \cdot \left(\frac{d\mathbf{V}_i}{dt} L \right) + \frac{\partial}{\partial_{p_i}} \left(\frac{dP_i}{dt} L \right) + \frac{\partial}{\partial_{a_i}} \left(\frac{dA_i}{dt} L \right) \right) = 0$$

Then, we plug in our flocking model. Applying the $f(x)\delta(x-x_i) = f(x_i)\delta(x-x_i)$ property, we obtain the following modified Liouville equation:

$$\begin{aligned}
&\frac{\partial L}{\partial t} + \sum_{i=1}^{N} \nabla_{\mathbf{x}_i} \cdot (\mathbf{v}_i L) \\
&\quad + \sum_{i=1}^{N} \nabla_{\mathbf{v}_i} \cdot \left(\left[\frac{\lambda_1}{N} \sum_{j=1}^{N} k_1(\mathbf{x}_i, \mathbf{x}_j)(\mathbf{v}_j - \mathbf{v}_i) + F_0(v_o \hat{\mathbf{e}}_\mathbf{s} - \mathbf{v}_i) \right] L \right) \\
&\quad + \sum_{i=1}^{N} \frac{\partial}{\partial p_i} \left(\left[\frac{\lambda_2}{N} \sum_{j=1}^{N} k_2(\mathbf{x}_i, \mathbf{x}_j)(p_j - p_i) + L_0(p_\infty - p_i)\psi_A(a_i) \right] L \right) \\
&\quad + \sum_{i=1}^{N} \frac{\partial}{\partial a_i} \left(\left[\frac{\lambda_3}{N} \sum_{j=1}^{N} k_3(\mathbf{x}_i, \mathbf{x}_j) r - k_u a_i - k_d a_i \right] L \right) = 0
\end{aligned}$$

Therefore,

$$\begin{aligned}
&\frac{\partial f^N}{\partial t} + \sum_{i=1}^{N} \nabla_{\mathbf{x}_i} \cdot \left(\mathbf{v}_i f^N\right) \\
&\quad + \sum_{i=1}^{N} \nabla_{\mathbf{v}_i} \cdot \left(\left[\frac{\lambda_1}{N} \sum_{j=1}^{N} k_1(\mathbf{x}_i, \mathbf{x}_j)(\mathbf{v}_j - \mathbf{v}_i) + F_0(v_o \hat{\mathbf{e}}_\mathbf{s} - \mathbf{v}_i) \right] f^N \right) \\
&\quad + \sum_{i=1}^{N} \frac{\partial}{\partial p_i} \left(\left[\frac{\lambda_2}{N} \sum_{j=1}^{N} k_2(\mathbf{x}_i, \mathbf{x}_j)(p_j - p_i) + L_0(p_\infty - p_i)\psi_A(a_i) \right] f^N \right) \\
&\quad + \sum_{i=1}^{N} \frac{\partial}{\partial a_i} \left(\left[\frac{\lambda_3}{N} \sum_{j=1}^{N} k_3(\mathbf{x}_i, \mathbf{x}_j) r - k_u a_i - k_d a_i \right] f^N \right) = 0
\end{aligned} \tag{4.5}$$

where $f^N(\mathbf{x}_1, \mathbf{v}_1, p_1, a_1, ..., \mathbf{x}_N, \mathbf{v}_N, p_N, a_N, t)$ is the N-particle probability distribution. f^N vanishes at infinity in our phase space and is symmetric, which means that $\mathbf{x}_i, \mathbf{v}_i, p_i, a_i$ can be interchanged with $\mathbf{x}_k, \mathbf{v}_k, p_k, a_k$. For example,

$$f^N(\mathbf{x}_1, \mathbf{v}_1, p_1, a_1, \mathbf{x}_2, \mathbf{v}_2, p_2, a_2, ...) = f^N(\mathbf{x}_2, \mathbf{v}_2, p_2, a_2, \mathbf{x}_1, \mathbf{v}_1, p_1, a_1, ...)$$

4.3 Formal derivation of the BBGKY Hierarchy

The Liouville equation describes the time evolution of the N-body probability density, f^N. This can represent a rather large system. Instead, one can

examine a reduced system via integration. By integrating over the space associated with the $N-j$ particles, one can analyze a reduced j-body system with the resulting j-body probability density. This reduced system forms what is called the BBGKY (Bogoliubov-Born-Green-Kirkwood-Yvon) hierarchy, a set of j coupled integro-differential equations. Each describes the time evolution of a reduced probability density. The reduced j-body probability distribution is written as

$$f_j^N = \int_\Omega (f^N) d\mathbf{x}_{j+1} d\mathbf{v}_{j+1} dp_{j+1} da_{j+1} ... d\mathbf{x}_N d\mathbf{v}_N dp_N da_N = \int_\Omega (f^N) d\Omega$$

where $\Omega = \left\{ \mathbb{R}^{2d(N-j)} \times \mathbb{R}_+^{N-j} \times \mathbb{R}_+^{N-j} \right\}$ is the space which we are integrating over.

To obtain the BBGKY hierarchy, we integrate Equation (4.5) over the space associated with the $j+1$,...,N particles and simplify. Since Equation (4.5) consists of 5 terms, we will call each of these terms as 'sections' to avoid confusion later on. When we integrate over Ω, the first section of Equation (4.5) is simply $\frac{\partial f_j^N}{\partial t}$. We focus on the second section.

Second section of Equation (4.5)

The second section is split into two terms.

$$\begin{aligned} \text{Second section} = &\sum_{i=1}^{j} \int_\Omega \nabla_{\mathbf{x}_i} \cdot (\mathbf{v}_i f^N) d\mathbf{x}_{j+1} d\mathbf{v}_{j+1} dp_{j+1} da_{j+1} ... d\mathbf{x}_N d\mathbf{v}_N dp_N da_N + \\ &\sum_{i=j+1}^{N} \int_\Omega \nabla_{\mathbf{x}_i} \cdot (\mathbf{v}_i f^N) d\mathbf{x}_{j+1} d\mathbf{v}_{j+1} dp_{j+1} da_{j+1} ... d\mathbf{x}_N d\mathbf{v}_N dp_N da_N \end{aligned}$$

The first term simplifies to

$$\sum_{i=1}^{j} \nabla_{\mathbf{x}_i} \cdot \left[\mathbf{v}_i \int_\Omega (f^N) d\mathbf{x}_{j+1} d\mathbf{v}_{j+1} dp_{j+1} da_{j+1} ... d\mathbf{x}_N d\mathbf{v}_N dp_N da_N \right]$$

because $\mathbf{x}_i$ and $\mathbf{v}_i$ are independent. The second term simplifies to 0 because in our N-particle phase space, f^N vanishes at infinity (ie $f^N \to 0$ as each component of

$\mathbf{x}_i \to \pm\infty$). Therefore,

$$\text{Second section} = \sum_{i=1}^{j} \nabla_{\mathbf{x}_i} \cdot \left(\mathbf{v}_i f_j^N\right)$$

Third section of Equation (4.5)

The third section is split into three terms.

$$\begin{aligned}
&\text{Third section} = \\
&\sum_{i=1}^{j} \int_{\Omega} \nabla_{\mathbf{v}_i} \cdot \left(\frac{\lambda_1}{N} \sum_{k=1}^{j} k_1(\mathbf{x}_i, \mathbf{x}_k)(\mathbf{v}_k - \mathbf{v}_i) f^N + F_0(v_o \hat{\mathbf{e}}_\mathbf{s} - \mathbf{v}_i) f^N \right) d\Omega + \\
&\sum_{i=1}^{j} \int_{\Omega} \nabla_{\mathbf{v}_i} \cdot \left(\frac{\lambda_1}{N} \sum_{k=j+1}^{N} k_1(\mathbf{x}_i, \mathbf{x}_k)(\mathbf{v}_k - \mathbf{v}_i) f^N \right) d\Omega + \\
&\sum_{i=j+1}^{N} \int_{\Omega} \nabla_{\mathbf{v}_i} \cdot \left(\frac{\lambda_1}{N} \sum_{k=1}^{N} k_1(\mathbf{x}_i, \mathbf{x}_k)(\mathbf{v}_k - \mathbf{v}_i) f^N + F_0(v_o \hat{\mathbf{e}}_\mathbf{s} - \mathbf{v}_i) f^N \right) d\Omega
\end{aligned}$$

The two parts of the first term can be simplified as such

$$\begin{aligned}
&\sum_{i=1}^{j} \int_{\Omega} \nabla_{\mathbf{v}_i} \cdot \left(\frac{\lambda_1}{N} \sum_{k=1}^{j} k_1(\mathbf{x}_i, \mathbf{x}_k)(\mathbf{v}_k - \mathbf{v}_i) f^N \right) d\Omega \\
&\qquad = \sum_{i=1}^{j} \nabla_{\mathbf{v}_i} \cdot \left(\frac{\lambda_1}{N} \sum_{k=1}^{j} k_1(\mathbf{x}_i, \mathbf{x}_k)(\mathbf{v}_k - \mathbf{v}_i) f_j^N \right)
\end{aligned}$$

$$\sum_{i=1}^{j} \int_{\Omega} \nabla_{\mathbf{v}_i} \cdot \left(F_0(v_o \hat{\mathbf{e}}_\mathbf{s} - \mathbf{v}_i) f^N \right) d\Omega = \sum_{i=1}^{j} \nabla_{\mathbf{v}_i} \cdot \left(F_0(v_o \hat{\mathbf{e}}_\mathbf{s} - \mathbf{v}_i) f_j^N \right)$$

because we are integrating over all $\mathbf{x}_{j+1}, \mathbf{v}_{j+1}, p_{j+1}, a_{j+1} ... \mathbf{x}_N, \mathbf{v}_N, p_N, a_N,$ whereas the derivatives are with respect to the velocities of the $i = 1$ to $i = j$ particles.

The second term is simplified using the same reasoning

$$\begin{aligned}
&\frac{\lambda_1}{N} \sum_{i=1}^{j} \int_{\Omega} \nabla_{\mathbf{v}_i} \cdot \left(\sum_{k=j+1}^{N} k_1(\mathbf{x}_i, \mathbf{x}_k)(\mathbf{v}_k - \mathbf{v}_i) f^N \right) d\Omega \\
&= \frac{\lambda_1}{N} \sum_{i=1}^{j} \nabla_{\mathbf{v}_i} \cdot \left(\int_{\Omega} \sum_{k=j+1}^{N} k_1(\mathbf{x}_i, \mathbf{x}_k)(\mathbf{v}_k - \mathbf{v}_i) f^N d\Omega \right)
\end{aligned}$$

and further simplified

$$= \frac{\lambda_1(N-j)}{N} \sum_{i=1}^{j} \nabla_{\mathbf{v}_i} \cdot \left(\int_{\Omega} k_1(\mathbf{x}_i, \mathbf{x}_{j+1})(\mathbf{v}_{j+1} - \mathbf{v}_i) f^N d\Omega \right)$$

$$= \frac{\lambda_1(N-j)}{N} \sum_{i=1}^{j} \nabla_{\mathbf{v}_i} \cdot \left(\int_{\Omega_{j+1}} k_1(\mathbf{x}_i, \mathbf{x}_{j+1})(\mathbf{v}_{j+1} - \mathbf{v}_i) f^N_{j+1} d\Omega_{j+1} \right)$$

where Ω_{j+1} integrates over all $\mathbf{x}_{j+1}, \mathbf{v}_{j+1}, p_{j+1}, a_{j+1}$. The rationale for such simplifications is as follows. First,

$$\int_{\Omega} k_1(\mathbf{x}_i, \mathbf{x}_{j+1})(\mathbf{v}_{j+1} - \mathbf{v}_i) f^N d\Omega = \int_{\Omega} k_1(\mathbf{x}_i, \mathbf{x}_m)(\mathbf{v}_m - \mathbf{v}_i) f^N d\Omega$$

for $\forall m \in \mathbb{Z}$ such that $j+1 \leq m \leq N$. Within each integral, without loss of generality, we can switch labels on the particle m with particle $j+1$. The value of f^N will not change thanks to the symmetry property. Also, if we define the spaces such that Ω_{j+1} integrates over all $\mathbf{x}_{j+1}, \mathbf{v}_{j+1}, p_{j+1}, a_{j+1}$ and Ω' integrates over the rest, then

$$\int_{\Omega_{j+1}} \left(\int_{\Omega'} k_1(\mathbf{x}_i, \mathbf{x}_{j+1})(\mathbf{v}_{j+1} - \mathbf{v}_i) f^N d\Omega' \right) d\Omega_{j+1}$$

$$= \int_{\Omega_{j+1}} k_1(\mathbf{x}_i, \mathbf{x}_{j+1})(\mathbf{v}_{j+1} - \mathbf{v}_i) f^N_{j+1} d\Omega_{j+1}$$

Moving on to the third term, we can apply the divergence theorem to obtain

$$\sum_{i=j+1}^{N} \int_{\Omega} \nabla_{\mathbf{v}_i} \cdot \left(\frac{\lambda_1}{N} \sum_{k=1}^{N} k_1(\mathbf{x}_i, \mathbf{x}_k)(\mathbf{v}_k - \mathbf{v}_i) f^N \right) d\Omega$$

$$+ \sum_{i=j+1}^{N} \int_{\Omega} \nabla_{\mathbf{v}_i} \cdot \left(F_0(v_o \hat{\mathbf{e}}_{\mathbf{s}} - \mathbf{v}_i) f^N \right) d\Omega = 0$$

Hence, we have simplified the third section down to

$$\begin{aligned}\text{Third Section} =& \frac{\lambda}{N}\sum_{i=1}^{j}\nabla_{\mathbf{v}_i}\cdot\left(\sum_{k=1}^{j}k_1(\mathbf{x}_i,\mathbf{x}_j)(\mathbf{v}_k-\mathbf{v}_i)f_j^N\right)\\ &+F_0\sum_{i=1}^{j}\nabla_{\mathbf{v}_i}\cdot((v_0\hat{\mathbf{e}_\mathbf{s}}-\mathbf{v}_i)f_j^N)\\ &+\frac{\lambda_1(N-j)}{N}\sum_{i=1}^{j}\nabla_{\mathbf{v}_i}\cdot\left(\int_{\Omega_{j+1}}k_1(\mathbf{x}_i,\mathbf{x}_{j+1})(\mathbf{v}_{j+1}-\mathbf{v}_i)f_{j+1}^N d\Omega_{j+1}\right)\end{aligned}$$

Fourth section of Equation (4.5)

$$\begin{aligned}&\text{Fourth section} =\\ &\sum_{i=1}^{j}\int_\Omega\partial_{p_i}\left(\frac{\lambda_2}{N}\sum_{k=1}^{j}k_1(\mathbf{x}_i,\mathbf{x}_k)(p_k-p_i)f^N+L_0(p_\infty-p_i)\psi_A f^N\right)d\Omega\\ &+\sum_{i=1}^{j}\int_\Omega\partial_{p_i}\left(\frac{\lambda_2}{N}\sum_{k=j+1}^{N}k_1(\mathbf{x}_i,\mathbf{x}_k)(p_k-p_i)f^N\right)d\Omega\\ &+\sum_{i=j+1}^{N}\int_\Omega\partial_{p_i}\left(\frac{\lambda_2}{N}\sum_{k=1}^{j}k_1(\mathbf{x}_i,\mathbf{x}_k)(p_k-p_i)f^N+L_0(p_\infty-p_i)\psi_A f^N\right)d\Omega\end{aligned}$$

The two parts of the first term simplify using the same reasoning as in the third section.

$$\sum_{i=1}^{j}\int_\Omega\partial_{p_i}\left(\frac{\lambda_2}{N}\sum_{k=1}^{j}k_2(\mathbf{x}_i,\mathbf{x}_k)(p_k-p_i)f^N\right)d\Omega=$$

$$\sum_{i=1}^{j}\partial_{p_i}\left(\frac{\lambda_2}{N}\sum_{k=1}^{j}k_2(\mathbf{x}_i,\mathbf{x}_k)(p_k-p_i)f_j^N\right)$$

$$\sum_{i=1}^{j}\int_\Omega\partial_{p_i}\left(L_0(p_\infty-p_i)\psi_A f^N\right)d\Omega=\sum_{i=1}^{j}\partial_{p_i}\left(L_0(p_\infty-p_i)\psi_A f_j^N\right)$$

The second term also simplifies using the same reasoning as in the third section.

$$\begin{aligned}
&\frac{\lambda_2}{N}\sum_{i=1}^{j}\int_{\Omega}\partial_{p_i}\left(\sum_{k=j+1}^{N}k_2(\mathbf{x}_i,\mathbf{x}_k)(p_k-p_i)f^N\right)d\Omega\\
&=\frac{\lambda_2}{N}\sum_{i=1}^{j}\partial_{p_i}\left(\int_{\Omega}\sum_{k=j+1}^{N}k_2(\mathbf{x}_i,\mathbf{x}_k)(p_k-p_i)f^N d\Omega\right)\\
&=\frac{\lambda_2(N-j)}{N}\sum_{i=1}^{j}\partial_{p_i}\left(\int_{\Omega}k_2(\mathbf{x}_i,\mathbf{x}_{j+1})(p_{j+1}-p_i)f^N d\Omega\right)\\
&=\frac{\lambda_2(N-j)}{N}\sum_{i=1}^{j}\partial_{p_i}\left(\int_{\Omega_{j+1}}k_2(\mathbf{x}_i,\mathbf{x}_{j+1})(p_{j+1}-p_i)f_{j+1}^N d\Omega_{j+1}\right)
\end{aligned}$$

where Ω_{j+1} integrates over all $\mathbf{x}_{j+1}, \mathbf{v}_{j+1}, p_{j+1}, a_{j+1}$.

The third term simplifies to 0 via the divergence theorem,

$$\begin{aligned}
&\sum_{i=j+1}^{N}\int_{\Omega}\partial_{p_i}\left(\frac{\lambda_2}{N}\sum_{k=1}^{N}k_2(\mathbf{x}_i,\mathbf{x}_k)(p_k-p_i)f^N\right)d\Omega\\
&+\sum_{i=j+1}^{N}\int_{\Omega}\partial_{p_i}\left(L_0(p_\infty-p_i)\psi_A f^N\right)d\Omega=0
\end{aligned}$$

Hence, gathering all the simplified terms of the fourth section together, we get

$$\begin{aligned}
\text{Fourth section } =&\frac{\lambda_2}{N}\sum_{i=1}^{j}\partial_{p_i}\left(\sum_{k=1}^{j}k_2(\mathbf{x}_i,\mathbf{x}_j)(p_k-p_i)f_j^N\right)\\
&+L_0\sum_{i=1}^{j}\partial_{p_i}((p_\infty-p_i)\psi_A f_j^N)\\
&+\frac{\lambda_2(N-j)}{N}\sum_{i=1}^{j}\partial_{p_i}\left(\int_{\Omega_{j+1}}k_2(\mathbf{x}_i,\mathbf{x}_{j+1})(p_{j+1}-p_i)f_{j+1}^N d\Omega_{j+1}\right)
\end{aligned}$$

Fifth (and last) section of Equation (4.5)

$$\text{Fifth section} =$$

$$\sum_{i=1}^{j} \int_{\Omega} \partial_{a_i} \left(\frac{\lambda_3 r}{N} \sum_{k=1}^{j} k_3(\mathbf{x}_i, \mathbf{x}_k) f^N - (k_u + k_d) a_i f^N \right) d\Omega +$$

$$\sum_{i=1}^{j} \int_{\Omega} \partial_{a_i} \left(\frac{\lambda_3 r}{N} \sum_{k=j+1}^{N} k_3(\mathbf{x}_i, \mathbf{x}_k) f^N \right) d\Omega +$$

$$\sum_{i=j+1}^{N} \int_{\Omega} \partial_{a_i} \left(\frac{\lambda_3 r}{N} \sum_{k=1}^{N} k_3(\mathbf{x}_i, \mathbf{x}_k) f^N - (k_u + k_d) a_i f^N \right) d\Omega$$

The two parts of the first term simplify.

$$\sum_{i=1}^{j} \int_{\Omega} \partial_{a_i} \left(\frac{\lambda_3 r}{N} \sum_{k=1}^{j} k_3(\mathbf{x}_i, \mathbf{x}_k) f^N \right) d\Omega$$

$$= \sum_{i=1}^{j} \partial_{a_i} \left(\frac{\lambda_3 r}{N} \sum_{k=1}^{j} k_3(\mathbf{x}_i, \mathbf{x}_k) f_j^N \right)$$

$$\sum_{i=1}^{j} \int_{\Omega} \partial_{a_i} \left((k_u + k_d) a_i f^N \right) d\Omega = (k_u + k_d) \sum_{i=1}^{j} \partial_{a_i} \left(a_i f_j^N \right)$$

The second term simplifies using the same reasoning as in the previous sections.

$$\frac{\lambda_3 r}{N} \sum_{i=1}^{j} \int_{\Omega} \partial_{a_i} \left(\sum_{k=j+1}^{N} k_3(\mathbf{x}_i, \mathbf{x}_k) f^N \right) d\Omega$$

$$= \frac{\lambda_3 r}{N} \sum_{i=1}^{j} \partial_{a_i} \left(\int_{\Omega} \sum_{k=j+1}^{N} k_3(\mathbf{x}_i, \mathbf{x}_k) f^N d\Omega \right)$$

$$= \frac{\lambda_3 r (N-j)}{N} \sum_{i=1}^{j} \partial_{a_i} \left(\int_{\Omega} k_3(\mathbf{x}_i, \mathbf{x}_{j+1}) f^N d\Omega \right)$$

$$= \frac{\lambda_3 r (N-j)}{N} \sum_{i=1}^{j} \partial_{a_i} \left(\int_{\Omega_{j+1}} k_3(\mathbf{x}_i, \mathbf{x}_{j+1}) f_{j+1}^N d\Omega_{j+1} \right)$$

where Ω_{j+1} integrates over all $\mathbf{x}_{j+1}, \mathbf{v}_{j+1}, p_{j+1}, a_{j+1}$.

Applying the divergence theorem, the third term simplifies to 0.

$$\sum_{i=j+1}^{N} \int_{\Omega} \partial_{a_i} \left(\frac{\lambda_3 r}{N} \sum_{k=1}^{N} k_3(\mathbf{x}_i, \mathbf{x}_k) f^N \right) d\Omega$$
$$- \sum_{i=j+1}^{N} \int_{\Omega} (k_u + k_d) \partial_{a_i} \left(a_i f^N \right) d\Omega = 0$$

Hence, the fifth section is simplified down to

$$\begin{aligned} \text{Fifth section } =& \frac{\lambda_3 r}{N} \sum_{i=1}^{j} \partial_{a_i} \left(\sum_{k=1}^{j} k_3(\mathbf{x}_i, \mathbf{x}_j) f_j^N \right) \\ &+ \frac{\lambda_3 r (N-j)}{N} \sum_{i=1}^{j} \partial_{a_i} \left(\int_{\Omega_{j+1}} k_3(\mathbf{x}_i, \mathbf{x}_{j+1}) f_{j+1}^N d\Omega_{j+1} \right) \\ &- (k_u + k_d) \sum_{i=1}^{j} \partial_{a_i} (a_i f_j^N) \end{aligned}$$

Putting all five simplified sections together, we have the following BBGKY formulation.

$$
\begin{aligned}
&\frac{\partial f_j^N}{\partial t} + \sum_{i=1}^{j} \nabla_{\mathbf{x}_i} \cdot \left(\mathbf{v}_i f_j^N\right) + F_0 \sum_{i=1}^{j} \nabla_{\mathbf{v}_i} \cdot \left((v_0 \hat{\mathbf{e}}_\mathbf{s} - \mathbf{v}_i) f_j^N\right) \\
&\quad + L_0 \sum_{i=1}^{j} \partial_{p_i}((p_\infty - p_i)\psi_A f_j^N) - (k_u + k_d) \sum_{i=1}^{j} \partial_{a_i}(a_i f_j^N) \\
&\quad + \frac{\lambda_1}{N} \sum_{i=1}^{j} \nabla_{\mathbf{v}_i} \cdot \left(\sum_{k=1}^{j} k_1(\mathbf{x}_i, \mathbf{x}_j)(\mathbf{v}_k - \mathbf{v}_i) f_j^N \right) \\
&\quad + \frac{\lambda_2}{N} \sum_{i=1}^{j} \partial_{p_i} \left(\sum_{k=1}^{j} k_2(\mathbf{x}_i, \mathbf{x}_j)(p_k - p_i) f_j^N \right) \\
&\quad + \frac{\lambda_3 r}{N} \sum_{i=1}^{j} \partial_{a_i} \left(\sum_{k=1}^{j} k_3(\mathbf{x}_i, \mathbf{x}_j) f_j^N \right) \\
&\quad + \frac{\lambda_1 (N-j)}{N} \sum_{i=1}^{j} \nabla_{\mathbf{v}_i} \cdot \left(\int_{\Omega_{j+1}} k_1(\mathbf{x}_i, \mathbf{x}_{j+1})(\mathbf{v}_{j+1} - \mathbf{v}_i) f_{j+1}^N d\Omega_{j+1} \right) \\
&\quad + \frac{\lambda_2 (N-j)}{N} \sum_{i=1}^{j} \partial_{p_i} \left(\int_{\Omega_{j+1}} k_2(\mathbf{x}_i, \mathbf{x}_{j+1})(p_{j+1} - p_i) f_{j+1}^N d\Omega_{j+1} \right) \\
&\quad + \frac{\lambda_3 r (N-j)}{N} \sum_{i=1}^{j} \partial_{a_i} \left(\int_{\Omega_{j+1}} k_3(\mathbf{x}_i, \mathbf{x}_{j+1}) f_{j+1}^N d\Omega_{j+1} \right) = 0
\end{aligned} \tag{4.6}
$$

4.4 Future directions

To go further with the BBGKY hierarchy, we would take the mean field limit. That is, take $N \to \infty$. But this requires many assumptions from our hierarchy that are difficult to prove rigorously. Formally, if we could take the limit and assuming $\lim_{N\to\infty} f_j^N = f_j$, Equation (4.6) becomes

$$\frac{\partial f_j}{\partial t} + \sum_{i=1}^{j} \nabla_{\mathbf{x}_i} \cdot (\mathbf{v}_i f_j) + F_0 \sum_{i=1}^{j} \nabla_{\mathbf{v}_i} \cdot ((v_0 \hat{\mathbf{e}}_\mathbf{s} - \mathbf{v}_i) f_j) + L_0 \sum_{i=1}^{j} \partial_{p_i}((p_\infty - p_i)\psi_A f_j)$$
$$- (k_u + k_d) \sum_{i=1}^{j} \partial_{a_i}(a_i f_j) \tag{4.7}$$
$$+ \lambda_1 \sum_{i=1}^{j} \nabla_{\mathbf{v}_i} \cdot \left(\int_{\Omega_{j+1}} k_1(\mathbf{x}_i, \mathbf{x}_{j+1})(\mathbf{v}_{j+1} - \mathbf{v}_i) f_{j+1} d\Omega_{j+1} \right)$$
$$+ \lambda_2 \sum_{i=1}^{j} \partial_{p_i} \left(\int_{\Omega_{j+1}} k_2(\mathbf{x}_i, \mathbf{x}_{j+1})(p_{j+1} - p_i) f_{j+1} d\Omega_{j+1} \right)$$
$$+ \lambda_3 r \sum_{i=1}^{j} \partial_{a_i} \left(\int_{\Omega_{j+1}} k_3(\mathbf{x}_i, \mathbf{x}_{j+1}) f_{j+1} d\Omega_{j+1} \right) = 0$$

This is still a coupled system of integro-differential equations. In order to close this hierarchy, we can set $j = 1$ and use the molecular chaos assumption of

$$f_2(\mathbf{x}_1, \mathbf{v}_1, p_1, a_1, \mathbf{x}_2, \mathbf{v}_2, p_2, a_2, t) = f_1(\mathbf{x}_1, \mathbf{v}_1, p_1, a_1, t) f_1(\mathbf{x}_1, \mathbf{v}_1, p_1, a_1, t)$$

to obtain a PDE for f_1.

Chapter 5: Electrochemical Measurement of the β-galactosidase Reporter from Live Cells: A Comparison to the Miller Assay

 This is a collaboration in which I developed the compartmental model for the experiments performed by Dr. Tanya Tschirhart and Xinyi Zhou.

5.1 Abstract

In order to match our ability to conceive of and construct cells with enhanced function, we must concomitantly develop facile, real-time, methods for elucidating performance. With these, new designs can be tested *in silico* and steps in construction incrementally validated. Electrochemical monitoring offers the above advantages largely because signal transduction stems from direct electron transfer - allowing for potentially quicker and more integrated measurements. One of the most common genetic reporters, β-galactosidase, can be measured both spectrophotometrically (Miller assay) and electrochemically. However, since the relationship between the two is not well understood, the electrochemical methods have not yet garnered

the attention of biologists. With the aim of demonstrating the utility of an electrochemical measurement to the synthetic biology community, we created a genetic construct that interprets and reports (with β-galactosidase) on the concentration of the bacterial quorum sensing molecule autoinducer-2. In this work, we provide a correlation between electrochemical measurements and Miller Units. We show that the electrochemical assay works with both lysed and whole cells, allowing for the prediction of one from the other, and for continuous monitoring of cell response. We further present a conceptually simple and generalized mathematical model for cell-based β-galactosidase reporter systems that could aid in building and predicting a variety of synthetic biology constructs. This first-ever in-depth comparison and analysis aims to facilitate the use of electrochemical real-time monitoring in the field of synthetic biology as well as to facilitate the creation of constructs that can more easily communicate information to electronic systems.

5.2 Introduction

Synthetic biology motivates the rewiring of biological systems for a myriad of applications, including sensing, directed feedback response to various inputs and the production of valuable products [99]. One of the aims of the field is to uncover the underlying design principles of biological systems through the rational design of gene and protein circuits [17] [74]. Engineered biological systems have already contributed significantly to our understanding of how natural systems function and interact [14] [85] [98]. Complex genetic circuits and feedback loops have given us

quantitative understanding of gene expression and signal transduction [52] [90], insights into the diversity of behaviors that result from various control loops [102], and the ability to rationally control spatial organization and interactions between cells [21] [51] [120].

For ease of analysis and for developing designs, genetically engineered cells generally use a set of well-characterized reporters, typically fluorescent proteins such as green fluorescent protein (GFP) or enzymes such as β-galactosidase (β-gal) [38]. Fluorescent reporters, when engineered for short half-lives, allow nearly continuous tracking of protein expression in live cells over time under varying conditions with intracellular resolution. In contrast, enzymatic reporters offer high sensitivity through signal amplification. Because the enzyme continually generates more detectable product over time, very small amounts of the enzyme can be measured. Requiring lower copy numbers of reporter in order to achieve a desired sensitivity eases the transcriptional/translational burden on the cell [9] so that the process of detection does not independently influence the system's biology.

The β-galactosidase enzyme is one of the most common genetic reporters, and its Miller assay [68] represents a textbook example of a well-quantified, validated and reproducible method for measuring protein expression and enzyme activity. Typically, cell samples are collected and lysed at specified time points, and the freed enzyme is allowed to react with the substrate, ortho-nitrophenyl-β-galactoside (ONPG). The yellow-colored product o-nitrophenyl (ONP) is produced by enzymatic cleavage and is quantified by measuring absorbance. The amount of product, the extent of the enzymatic reaction and the amount of cells in the sample are used to

calculate Miller Units, which can be used to compare gene expression across samples taken under various conditions. Although reliable, the Miller assay has imitations. The assay requires cell lysis, preventing continuous measurement or intracellular localization. Use of a different β-gal substrate, X-gal, provides an optical measure of localization but is not quantitative. These constraints become more problematic as researchers move from traditional flask and well-plate formats towards microfluidic and lab-on-a-chip formats that better support rapid, highly sensitive and selective measurements. As a result, recent influential papers in synthetic biology have mainly relied on fluorescent reporters [30] [71] [83].

Since an aim of synthetic biology is the analysis and control over natural circuits as well as the design and manipulation of synthetic parts, we need measurement techniques that track the dynamics of the synthetic system. Such techniques would enable more predictive models of metabolic and cellular networks. This paper is largely motivated by many reports of the biosensor and biochip community, where researchers have used the redox-active molecule p-aminophenyl β-D-galactopyranoside (PAPG) as a substrate for electrochemically measuring β-gal activity both inside and outside of cells [4] [11] [19] [23] [59] [60] [62] [81] [96]. The focus to-date, has targeted the development of sensitive electrode and biochip systems that provide cost-efficient real-time measurements on the microscale that are easily integrated into more elaborate electronic systems.

Electrochemical sensing uses electron exchange between redox-active molecules and electrodes to generate measurable electrochemical signals. Their fast kinetics

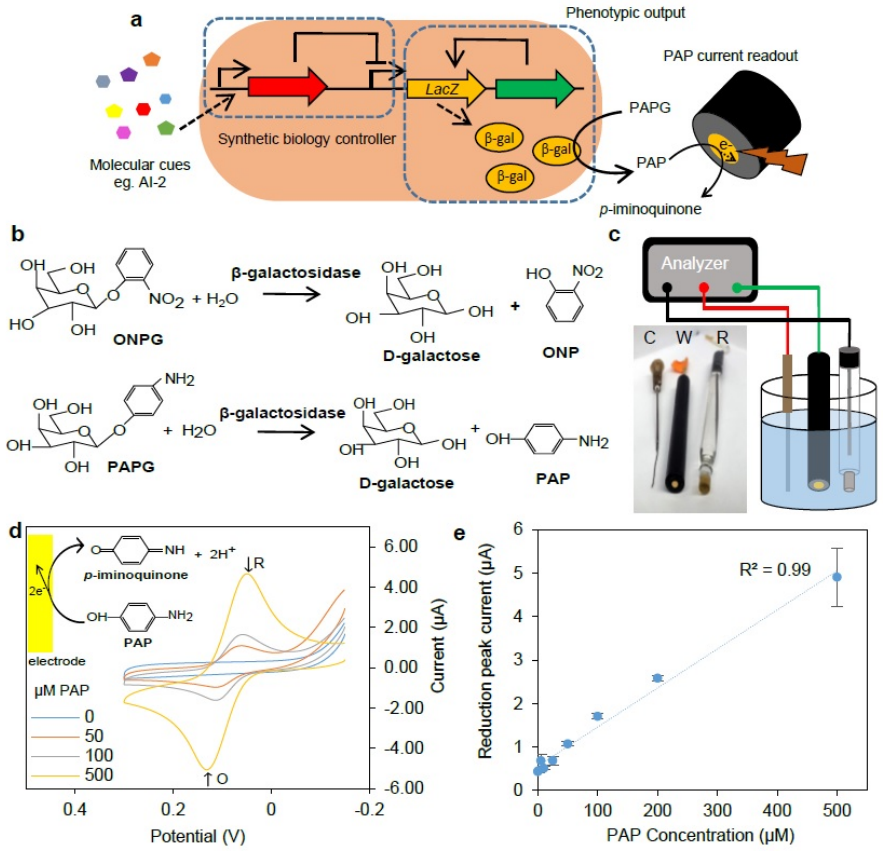


Figure 5.1: (a) A general electrochemical biosensing scheme for a cell producing β-galactosidase as a reporter (phenotypic output) from a promoter of choice or a combination of genetic elements (synthetic biology controller) in response to various molecular cues (AI-2 in this paper). Once β-galactosidase is produced, 4-aminophenyl β-D-galactopyranoside (PAPG) is added. PAPG enters the cell and is cleaved by β-gal into the redox active molecule p-aminophenol (PAP). PAP exits the cell and is quantified through cyclic voltammetry at the electrode. (b) Molecular structures of the cleavage of ONPG to ONP and D-galactose and PAPG to PAP and D-galactose by β-gal.(c) The counter (C), working (W), and reference (R) electrodes used in this work and a schematic of the electrochemical system setup. (d) Sample cyclic voltammograms of different concentrations of PAP using the setup in (c) Inset showing reversible PAP oxidation to *p*-iminoquinone at the electrode. R and O indicate reduction and oxidation peaks, respectively. (e) PAP standard curve shows correlation between PAP concentration and the reduction peak current, taken from (d).

enable dynamic monitoring. Additionally, the electronic output is easily analyzed and used for the development of predictive models. However, most electrochemical devices for β-gal measurement are manufactured in-house, precluding their proliferation among synthetic biology laboratories. Moreover, there remains no direct comparison to the standard that is well understood by biologists - the Miller Unit. This disconnect motivated us to investigate the relationship between the electrochemical (using PAPG) and spectrophotometric (using ONPG) methods for quantifying β-galactosidase activity.

Here, we show for the first time that the electrochemical substrate PAPG can be used to detect β-gal in a manner analogous to ONPG being used as a Miller substrate. We demonstrate the electrochemical method by quantifying the responses of both whole cells expressing β-gal and lysed cells, and correlate both responses to the Miller assay. We created a whole-cell biosensor that detects the quorum-sensing molecule autoinducer-2 (AI-2) and responds by producing β-gal in a concentration-dependent manner. The genetic circuitry is generalizable to any stimuli or molecular cue that evokes a cell response, contingent on a genetic link to a natural or synthetic promoter. We specifically chose to detect AI-2 because it is a molecule secreted by scores of bacteria and facilitates quorum sensing (QS) [95]. QS represents population dependent bacterial communication and response and mediates bacterial virulence and biofilm formation [46]. QS is particularly relevant as the problem of antibiotic resistance escalates [91]. Thus, in our example, we utilize synthetic biology constructs to transduce molecular signals that mediate biomolecular cell-to-cell communication into electrical output signals that can be further quantified, ana-

lyzed, and modeled.

We use both our electrochemical and spectrophotometric data to create a computational model of the cell 'biosensor' under both lysed and whole cell conditions, and correlate the electrochemical sensor measurements to the genetic response through the Miller assay. With this framework, researchers can measure whole cell response with all the advantages of electrochemical detection, without sacrificing understanding of the underlying genetic response. Overall, our analysis of the whole-cell electrochemical AI-2 biosensor and corresponding computational model furthers the use of electrochemical monitoring for synthetic biology.

5.3 Results and Discussion

5.3.1 Electrochemical characterization of *LacZ* expression

A generalized scheme of the electrochemical detection of PAP to characterize cellular *LacZ* expression is presented in Figure 5.1a. Molecular cues signal intracellular synthetic biology controllers (represented in abstract form) to drive *LacZ* gene expression, which produces the β-gal enzyme as a surrogate for phenotypic output. To perform electrochemical enzyme activity measurements, PAPG is added, enters the cell, and is cleaved by β-gal into PAP (Figure 5.1b), exactly analogously to ONPG that is cleaved into ONP for optical measurements. PAP can then exit the cell and be detected electrochemically by oxidation to *p*-iminoquinone using a three electrode system (Figure 5.1b and c) [113]. The current readout, as can be seen in Figures 5.1d (cyclic voltammetry scan) and e (peak current of the cyclic voltammo-

grams), is linearly proportional to PAP concentration over a fairly wide range. See Supplementary Figure 5.11 for additional detail regarding the oxidation currents of PAP. This method allows quantification of β-gal activity whether in cell extracts or within intact cells. Since we foresee this measurement method being especially useful in miniaturized lab-on-a-chip systems, it is important to note that decrease in electrode size should not negatively impact the sensitivity of the system.

5.3.2 Characterization of β-gal produced by biosensor cells

To study the utility of electrochemical β-gal detection in cells, we constructed an *E. coli* sensor strain (*luxS*-, *lsrFG*-, Δ(*argF-lac*)*169*) that relies on native import and phosphorylation of AI-2 to de-repress expression of the *lsr* promoter (Figs. 2a,b). Using a two-plasmid system we amplify the native signal strength (*lsr* promoter) by driving expression of T7 RNA polymerase (from pCT6). This, in turn, amplifies gene expression from a generic pET vector, engineered to produce β-gal. This two-plasmid expression system provides biological amplification that greatly increases the sensitivity of the strain to AI-2, as discussed by Tsao et al. (2010). As seen in Figure 5.2c, the cells respond in a predictable, dose-dependent manner; that is, the β-gal response as measured by the Miller assay correlates with the concentration of added AI-2. These response trends correspond with our previous results for systems involving AI-2 uptake by cells [103]. Because the stoichiometry of the reaction using the electrochemical substrate (PAPG) is the same as that using the colorimetric Miller substrate (ONPG), we expect a direct correlation between the

Miller and electrochemical assays. We selected cells at 3 hours post-AI-2 induction for further experiments since the responses were robust and the Miller response correlated most closely to the AI-2 concentration (Supplementary Figure 5.8).

5.3.3 Real-time electrochemical vs. spectrophotometric measurement of β-gal activity

To establish the electrochemical measurement using our electrodes, we first used purchased β-gal as a reference material. In Figure 5.3, we characterized β-gal activity using both spectrophotometric and electrochemical methods, re-suspending the enzyme in either the Miller assay's Z-buffer or 40 mM Tris-HCl, respectively. These experiments provided a basis for quantifying enzyme (g/L) relative to its activity using both methods and for estimating β-gal concentration within biosensor cells.

The spectrophotometric technique (Fig. 5.3b) allows measurements that are rapid, sensitive, and linear at low β-gal concentrations. However, in less than a minute at the highest tested β-gal concentration - 0.005 U/µl - the ONP generated was above the linear range (>1 AU). Higher concentrations would need to be diluted. The electrochemical technique, however, remained sensitive and linear, better differentiating among higher concentrations of β-gal (Fig. 5.3c). In both cases, rates of product generation were higher with increasing enzyme concentration (as expected), and the two measurements were linearly correlated (Fig. 5.3d). We note also (Fig. 5.3d) that the rate of PAP detection was ~ 5 fold slower than the rate of

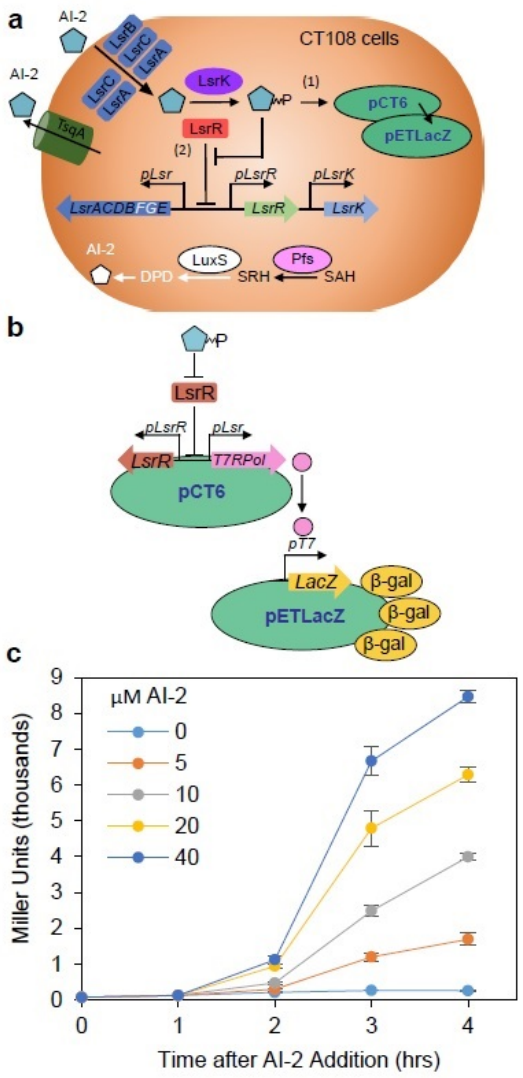


Figure 5.2: (a) The biosensing scheme for a cell producing β-gal as a response to added AI-2. (b) The synthetic construct, in which the Lsr promoter activation induces β-gal production through a two-plasmid system. Plasmid pCT6 responds to AI-2 by producing T7 polymerase and activating the T7 promoter on pETLacZ, which results in β-gal over expression. (c)A time-course response to different added AI-2 concentrations of the above construct, CT108 cells with pCT6 and pETLacZ plasmids. Averages are from 3 samples and error bars indicate S.D.

ONP detection.

5.3.4 Electrochemical response of AI-2 biosensor cells

We then characterized the response of our biosensor cells after a 3 hour AI-2 induction. Since the Miller assay requires cell lysis, we first investigated the electrochemical response of lysed cells. The schematic in Figure 5.4a depicts the relationships between all of the system components. PAP current was measured at 0, 5, 10, and 15 minutes after the addition of PAPG. In Figure 5.4b, the current rose as a function of both time and initial AI-2 concentration. The Miller assay was also performed on the same cells at the 0 time point, and in Figure 5.4c, a linear correlation between the Miller Units and the rate of PAP detection from the same cells was found.

This enabled correlation of the electrochemical measurement to the Miller Units. To our knowledge, this correlation has never been reported for β-gal expressed in any cell type. These data provide direct correlation for the two measurements and demonstrate, at least for this synthetic biology application with *E. coli*, that the assays are easily employed for typical experimental ranges. Finally, these data can be used to approximate the amount of intracellular β-gal in the host cells (Supplementary Figure 5.9).

In Figure 5.5, we continued our studies with whole live cells. PAP current was measured at the 0, 1, and 2 hr after PAPG addition. Recall that whole cell

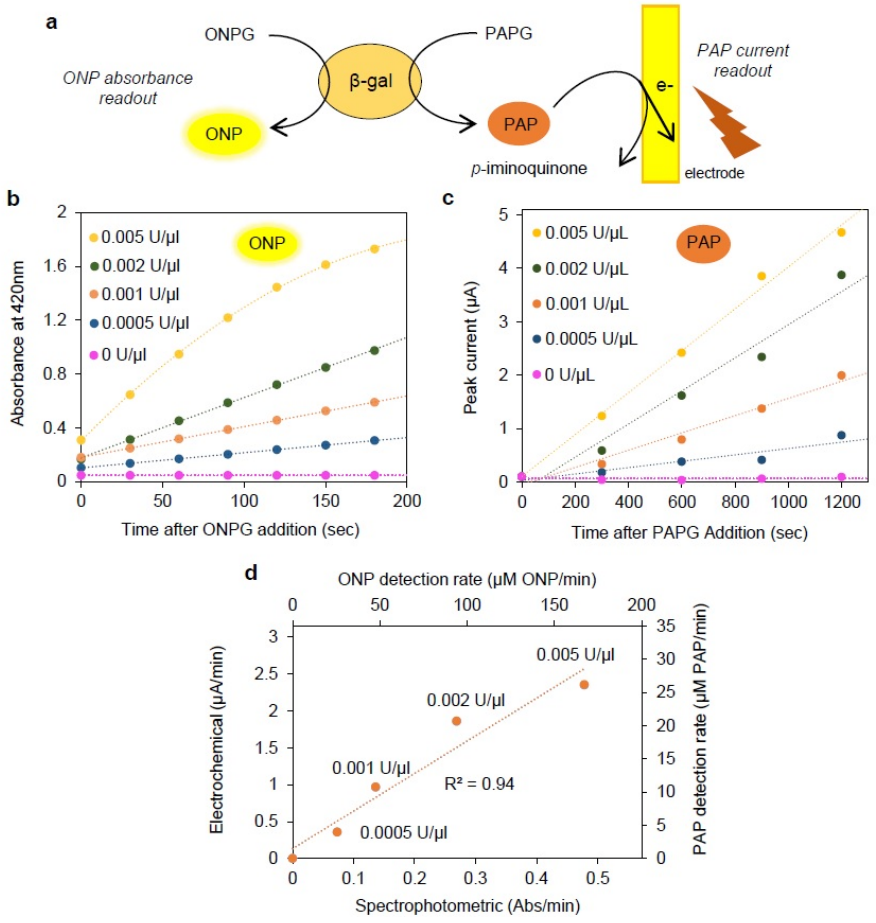


Figure 5.3: (a) Scheme of real-time electrochemical and spectrophotometric detection of PAP or ONP produced by β-gal cleavage of PAPG or ONPG. (b) Spectrophotometric ONP measurement from reactions of various concentrations of β-gal with ONPG over time. (c) The same β-gal concentrations and conditions as in (b), where the activity was instead measured by electrochemical detection of PAPG cleavage to PAP. (d) Correlation between electrochemical measurement rate of PAP detection to spectrophotometric measurement rate of ONP detection from data in (b) and (c), as well as the same data converted to the same units of µM/min of measured PAP or ONP. All lines in B-D indicate linear trendlines except for 0.005 U/µl in B, which indicates a best fit 2nd order polynomial trendline.

measurements are not possible via the Miller protocol, which lyses cells. In order to perform measurements in live cells, PAPG and PAP diffusion through the cell membrane is required, which can be quantified using electrochemical techniques. As seen in Figure 5.5b, the current response correlates with the initial amount of AI-2 added, as was the case with the lysed cells. Multiple experiments are shown as performed on different days. In Figure 5.5c, each point represents both the rate of PAP detection at 2 hours (as in Figure 5.5b) and the corresponding Miller results of the same cell samples. It is important to note that during those two hours following PAPG addition, the quantity of β-gal within the cells did not vary significantly. Figure 5.5c shows correlation, R2 = 0.97, with R2 = 0.92 for all points between the Miller Units and the electrochemical response up to $\sim$ 2000 Miller units. The limit tested for the lysed cells was also in this range. We also note that the current generation rate was much lower in whole cells than in lysed cells (Fig. 5.4). The current obtained in the 40 µM case for the whole cells reached 1.5 µA after 2 hours, but the same output was obtained within $\sim$ 7 min for the lysed cells. Clearly, the intact cells represent a barrier to PAPG and/or PAP transport, in addition to unknown factors that affect the molecule's concentrations or reactions.

In Figure 5.6a, we show a comparison between the lysed and whole cell electrochemical response rates, dependent on the corresponding Miller Unit measurements. This again allows a prediction of the current for either lysed or whole cells based on Miller assay results. The linear relationships from Figures 5.4c & 5.5c were used to construct a line in Figure 5.6b (see Methods for further information and calcu-

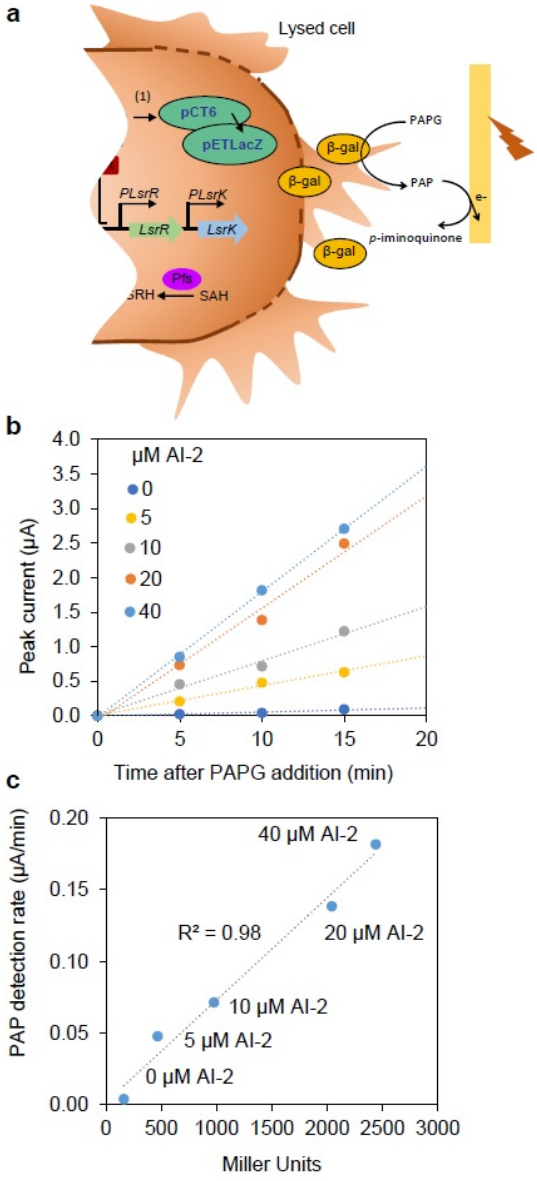


Figure 5.4: (a) Scheme of detection of AI-2 by electrochemical measurement of PAP production by lysed biosensor cells. (b)Detection of PAP from cells induced with the indicated AI-2 concentrations for 3 hours and lysed prior to PAPG incubation for the indicated time. (c) Comparison of the rate of PAP detection from the samples in (b) to the Miller units measured from the same samples, indicating a linear relationship and low background. Dotted lines indicate linear trendlines.

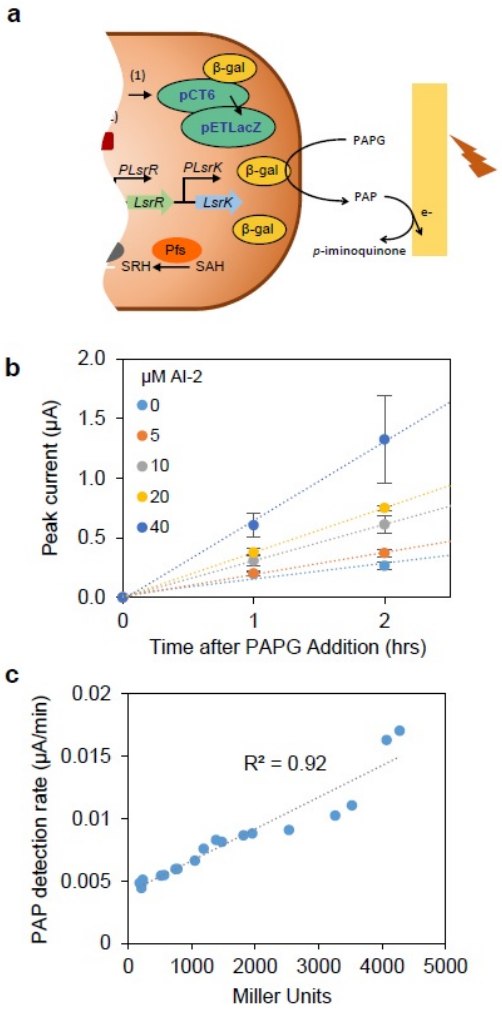


Figure 5.5: (a) Scheme of detection of AI-2 by electrochemical measurement of PAP production by whole biosensor cells. (b) PAP detection from whole cells after incubation with the given AI-2 concentrations for 3 hours and then PAPG for the indicated time. Averages are from 3-4 replicates on separate days and error bars show S.D. (c) The calculated rates of PAP detection over the two hour period versus the Miller Units of the same cell samples shows a linear relationship. Dotted lines indicate trendlines.

lations), which relates the lysed cell assay to the intact cell assay, both measured via electrochemical means. It is evident from the graphs that lysed cell measurements are more than an order of magnitude faster than whole cell measurements for samples with similar Miller Units. Nonetheless, the fact that the detection rates are linear at even low levels of β-gal suggests that relatively rapid analysis of gene expression is possible using this method, especially if more sensitive electrodes were used. Additionally, we performed similar preliminary experiments using constructs with single plasmids and different promoters and were able to confirm similar trends. One aspect that is not obvious from the plot is that the amount of background signal for whole cells ($\sim$ 0.005 µA/min, Fig. 5.6b) was similar to that for lysed cells ($\sim$ 0.004 µA/min, Fig. 5.6a). This background could be due to a combination of factors: physical PAPG degradation to PAP or other uncharacterized cell processes resulting in PAPG cleavage, rather than due to reaction by β-gal. We note, however, that experiments without the addition of PAPG resulted in near zero background (Supplementary Figure 5.10), indicating that the cells are not producing additional electrochemical species that overlap with the signal of PAP.

5.3.5 The two-compartment mathematical model

Finally, we constructed a simple computational model that simulates β-gal levels and PAP concentrations in response to PAPG and AI-2 addition to cells with respect to time. In order to build this model, compartments (i.e., sub-models) were constructed separately (Figure 5.7a). The first compartment modeled cell behavior

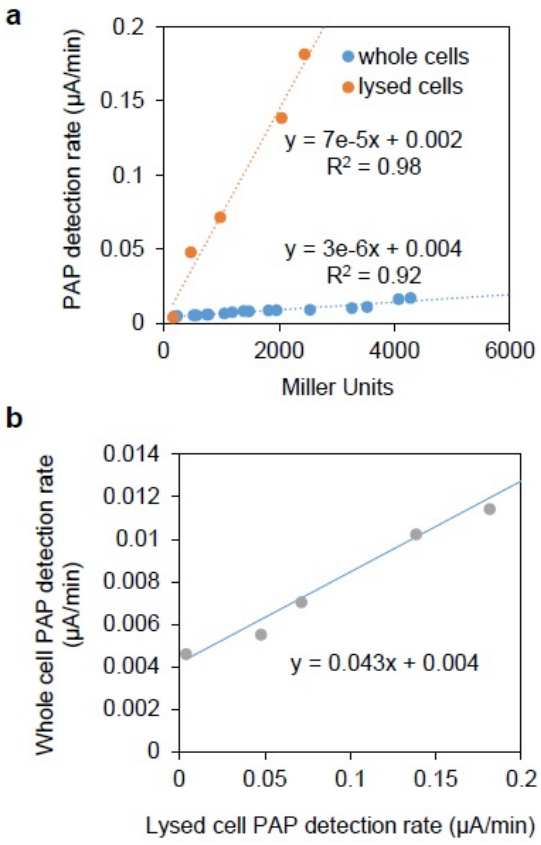


Figure 5.6: (a) Comparison of the rates of PAP detection from whole versus lysed CT108 pCT6/pETLacZ sensor cells based on measured Miller Units in response to different AI-2 concentrations. (b) The line indicating the relationship/conversion of PAP detection rate from lysed cells directly to whole cells. The points on the line indicate calculated whole-cell rates for lysed cells with specific Miller Units. Dotted lines indicate trendlines.

(LacZ expression) in response to AI-2. This model is a surrogate for any cell-based model for the design of a genetic circuit with corresponding elements of gene expression. The second compartment modeled the production of PAP via the substrate PAPG and its response to the enzyme, β-gal. This model represents a calibration 'standard' that can be integrated into the first model to provide connection to the experimental system. These sub-models (Figure 5.7a, 5.7b) were combined into one final model that was then compared to the lysed cell data. Figure 5.7c shows that there exists a good correlation between the model predictions and the PAP measurements (see also Supplementary Figures 5.12-5.14). That is, we have divided the model into two compartments so that the latter compartment could be an electrochemical "plug in" to the first compartment, which in turn, would be constructed for any generic construct wherein the first principles reactions could be conceptualized. In our case, the first compartment models the AI-2 - mediated expression of β-gal.

5.4 Conclusion

In this work, we report the first in-depth comparison of the electrochemical method for measuring β-gal activity in both lysed and intact cells to the gold-standard Miller assay. Our results indicate that the electrochemical method is a feasible alternative that enables added versatility for analysis of synthetic biology constructs, especially in applications exploiting lab-on-a-chip devices or real-time measurements. We demonstrated this by developing a biosensor bacterium that converts the signaling function of quorum sensing molecule AI-2 to electronic sig-

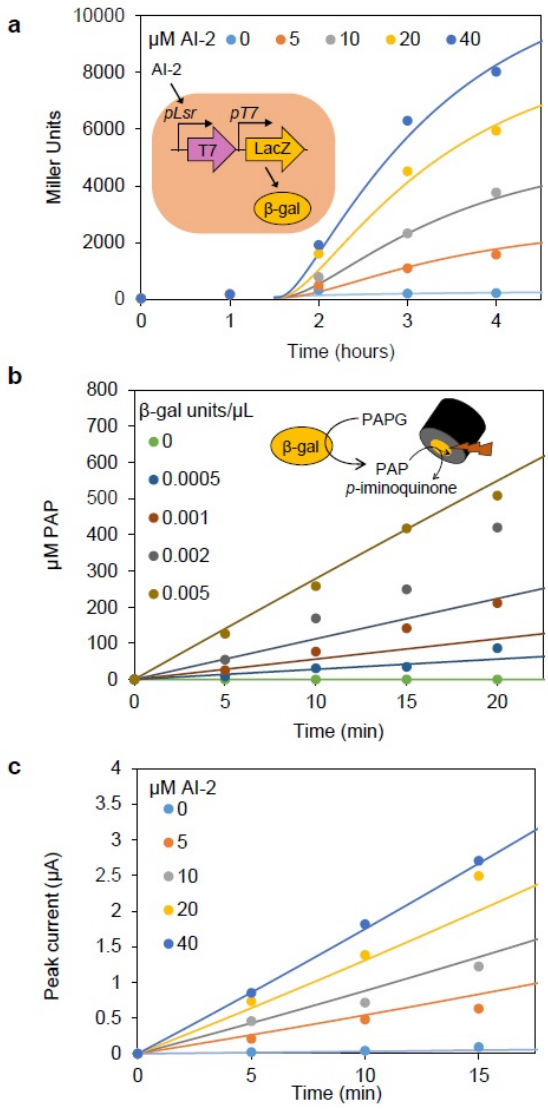


Figure 5.7: (a) Model results of cell production of β-gal depending on AI-2 induction, converted to Miller Units and overlaid with experimental data. Insert indicates modeled processes. (b) Model results of β-gal conversion rate of PAPG to PAP overlaid with experimental results. Inserted scheme indicates modeled processes. (c) Model results of lysed-cell production of β-gal and PAP production converted to current (combination of two modeled compartments in a and b) and overlaid with experimental results.

nals through a genetic (β-gal) and chemical (PAPG) transducer system. Our results showed that the electronic current measured was proportional to Miller Units from the same samples, and can be used to approximate gene expression in a multi-component synthetic circuit. Such electrochemical measurements would allow for *in situ* analyte detection as they integrate seamlessly into electronic systems, and could therefore provide for programmable measurements of synthetic biology constructs. Additionally, we created a two-compartment model that can be used to predict PAP (and thus electrochemical) output from our construct, and can be modified to accommodate other systems driving *LacZ* expression. We believe this work connects, for the first time, the work of the biosensor community in enzymatic reporter detection, with the standard optical measurement techniques biologists use for the same purpose. A better link between synthetic biologists and biochip / microelectronics designers enables parallel use of each other's tools and advances the potential for synergistic outcomes.

5.5 Methods

5.5.1 Chemical Reagents and Biocomponents.

4-aminophenyl β-D-galactopyranoside (PAPG), 4-aminophenol (PAP), β-galactosidase (β-gal) and ortho-Nitrophenyl-β-galactoside (ONPG) were from Sigma-Aldrich. PAPG was dissolved in diH2O and β-gal was dissolved in 40 mM Tris-HCl. PAP and ONPG were dissolved in 0.1 M phosphate buffer (PB). *"In vitro"* autoinducer-2 (AI-2) was produced through previously described biological nanofac-

tories [47]. Briefly, the nanofactory fusion protein HGLPT (His_6-protein G-LuxS-Pfs-Tyr_5) was purified from *E. coli* BL21 $luxS^-$ and incubated with 1 mM of the substrate S-(f'-deoxyadenosin-5')-L-homocysteine (SAH) at 37°C with shaking at 250 rpm, resulting in AI-2 synthesis. The enzymatic reaction product was twice extracted by an equal volume of chloroform.

5.5.2 Electrochemical PAP Detection.

PAP was detected electrochemically through cyclic voltammetry (CV). CVs were performed with a CHI Instruments 600-series electrochemical analyzer (CH instruments, Inc.) using an Au working electrode (2 mm diameter, CH Instruments, Inc.), a 4 cm-long platinum counter electrode (Alfa Aesar), and Ag/AgCl reference electrode (BASi). CVs were run from -0.15 to 0.3 V at a scan rate of 50 mV/s. The current at the reduction peak was used to measure PAP concentration.

5.5.3 Bacterial Strains, Plasmids and Culture Conditions.

In order to enhance the AI-2 responsiveness of *E. coli* ZK126 [22], a double chromosomal knockout of *lsrFG* and *luxS* strain, CT108, was created using one-step inactivation method [27]. Briefly, pKD4 was PCR amplified using primers lsrFGHP1, ATGGCAGATTTAGACGATATTAAAGATGGTAAAGATTTTCGTGTAGGCTGGAGCTGCTTC, and lsrHP2 [111]. The PCR product was electrically transformed into ZK126 pKD46 (to express the Red recombinase). *lsrFG:Kan*^{r} recombinants were screened with kanamycin and pKD46 plasmid was cured by growing at

37°C. After confirming the gene replacement by PCR, the kanamycin resistant gene, Kan^r, was removed by transforming a helper plasmid, pFLPe-Tet (Gene Bridges), to express the FLP recombinase. Then *luxS* was knocked-out from ZK126 $\Delta lsrFG$ following the similar procedure published before [111]. CT108 was created after the removal of the kanamycin resistance gene similarly descripted as above. Plasmids pCT6 [103] and then the commercially-available pET200/D/LacZ (Life Technologies) were introduced via standard heat-shock protocols of chemically-competent cells.

5.5.4 Miller Assay (Colorimetric β-gal Activity Detection).

Miller assay was performed according to standard protocols [68]. Briefly, cells were lysed with chloroform and sodium dodecyl sulfate (SDS) to release β-gal. The substrate ONPG was added and cleaved by β-gal into a yellow molecule, o-nitrophenol. Absorbance at 600 nm, 550nm, and 420 nm was quantified by a Molecular Systems plate reader. The OD at 600nm was measured from 250 µl of cells and the ODs at 420 nm and 550 nm were measured from 200 µl of cells.

5.5.5 AI-2 Detection with Bacterial Biosensor.

E. coli CT108 pCT6/pETLacZ was grown in LB Broth, Miller (Fischer Scientific) at 37°C with aeration by shaking at 250 rpm, then reinoculated at 2% and grown to early log phase (OD600 0.2). Cells were incubated with AI-2 at 37°C with shaking at 250 rpm, then spun down and resuspended in phosphate-buffered saline

(PBS) at an OD600 of 2. 0.5 mg/mL PAPG was added and CVs were performed as described above. To validate the electrochemical results, the commonly used Miller assay was performed on lysed cells to quantify β-galactosidase activity.

5.5.6 Cell Lysis.

10 % chloroform and 0.005 % SDS were added to cells resuspended in PBS. The reaction was vortexed and incubated at room temperature for 5 min. For absorbance measurements, the reaction was spun down for 4 minutes at 8000 rpm and the supernatant was removed for measurement. For electrochemical measurements, the reaction was spun down for 1 minute at 14,000 rpm and the supernatant was removed for measurement.

5.6 Supplemental section

Calculations: In order to convert absorbance (Abs/min) or electrochemical (µA/min) measurements to β-gal (U/µl), the following equations were used, respectively: $x = (y - 0.0016)/134$ and $x = (y + 0.035)/94.8$ (x is the U/µl and y the measurement in either Abs/min or µA/min) . Both equations were calculated from the data in Figure 3, where for each β-gal amount a calculated rate (Abs/min or µA/min) was found using Excel's LINEST function. For the absorbance data, the highest enzyme concentration's data was omitted as it was not in the linear range.

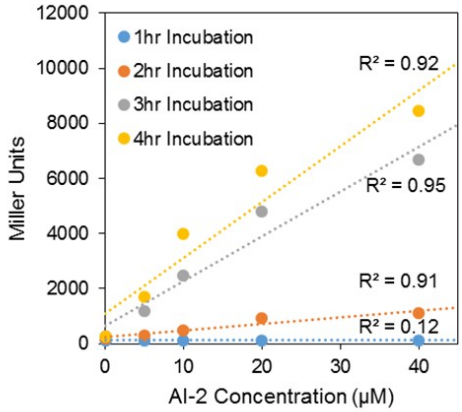


Figure 5.8: Correlation between the concentration of AI-2 used to induce the biosensor cells and the Miller Units, based on the time after induction (incubation).

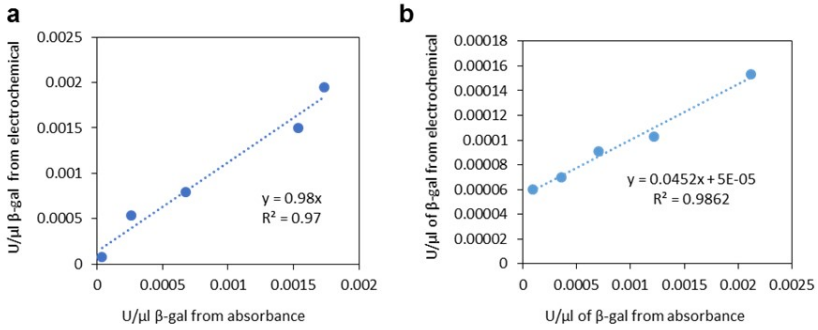


Figure 5.9: (a) The linear correlation between the enzyme units calculated from absorbance or electrochemical measurements for lysed cells. (b) Similar results as in (a) but with electrochemical measurements coming from whole cells.

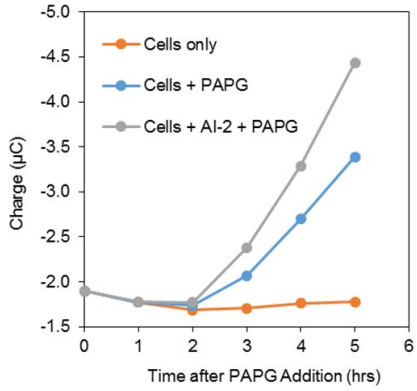


Figure 5.10: Control experiments were performed with biosensor cells different than ones used in the main paper, but which behaved similarly. PAPG concentrations were 0.5 mg/ml and AI-2 was 40 µM. Procedure was similar to that in the paper, and charge measurements were recorded instead of current, but represent similar data.

Discussion: We can see that in Supplementary Figure 2a, the rates correlate in almost a 1-to-1 ratio since in both cases the cells are lysed. In Supplementary Figure 2b, however, the electrochemical measurements show a much lower apparent enzyme amount compared to the absorbance measurements. This is because although the absorbance measurements (as part of the Miller assay) were done on lysed cells, the electrochemical measurements were done on whole cells. The relationship allows us to calculate the underestimation of the electrochemical whole-cell measurement.

Calculations for line in Figure 6b The line was-calculated by equating the x's of the two equations of the lines in 4c and 5c, representative of the Miller Units, and rearranging the combined equation so that the current for whole cells (y1) depended

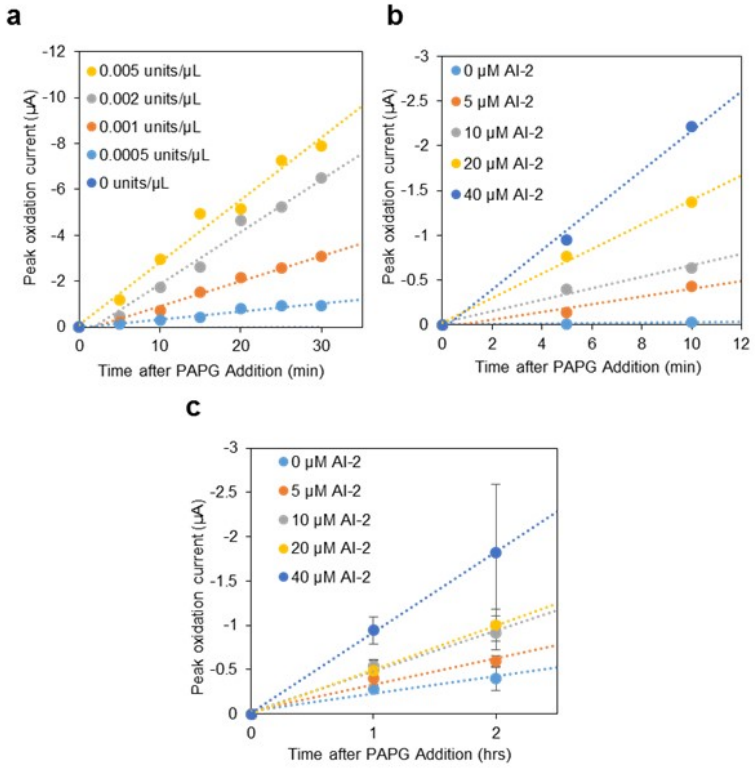


Figure 5.11: (a) Oxidation currents of PAP generated by various β-gal concentrations, same as in Figure 3c. (b) Oxidation current of PAP generated by lysed cells induced with the indicated concentrations of AI-2, same as in Figure 4b. (c) Oxidation current of PAP generated by whole cells induced with the indicated concentrations of AI-2, same as in Figure 5b.

on the current for lysed cells (y2). This allowed for direct conversion of lysed-cell rate to whole-cell rate for samples with the same Miller Units. The points on the graph indicate lysed-cell points [(x, y) being Miller, current] whose Miller Units were taken and plugged into the whole-cell equation to calculate what the whole-cell current would be for cells with the same Miller Units (x's). The results were used as the y's for the points on the graphs, with the x's being the experimentally-measured lysed cell currents.

5.7 The Two-Compartment model

The model for tracking beta-gal (β-gal) activity is a two-compartment model. That is, the main model is a construction of two sub-models. The first sub-model is a system of 3 differential equations that model β-gal activity expressed in cells exposed to different levels of AI-2. The known parameter values are taken from literature and the rest are fitted to the data. The second sub-model uses the two Michaelis-Menten differential equations to model PAP production at various concentrations of β-gal. The Michaelis constant (K_{m_1}) is taken from the literature. The turnover rate (k_{cat_1}) is fitted to the data. The two sub-models are combined to create a model that predicts PAP concentrations at various levels of AI-2 added. To observe the quality of our predictions, the model output is compared to data collected from the lysed cells.

5.7.1 The first compartment

Data Conversion

In order to fit the sub-model, data was converted from Miller units to µM of β-gal. Using Beers Law, with extinction coefficient $e = 4800$ and length = 1 cm, we converted the absorbance at 420nm into µM concentration of ONP. In order to convert µM ONP into µM β-gal, we use the following Michaelis-Menten equations:

$$\frac{d(ONP)}{dt} = \frac{k_{cat_1} E_{o_1}(ONPG)}{K_{m_1} + ONPG} \tag{5.1}$$

$$\frac{d(ONPG)}{dt} = -\frac{k_{cat_1} E_{o_1}(ONPG)}{K_{m_1} + ONPG} \tag{5.2}$$

with known turnover rate $k_{cat_1} = 620/sec$ and Michaelis constant $K_{m_1} = 120$ µM [56] [65]. E_{o_1} stands for the initial enzyme concentration. Assuming that $K_{m_1} \ll$ ONPG, the ONP differential equation simplifies to:

$$\frac{d(ONP)}{dt} = k_{cat_1} E_{o_1} \tag{5.3}$$

The slope is estimated from the linear fit of the data so E_{o_1} can be solved at each time point to extract the µM of β-galactosidase.

Model fitting

$$\text{AI-2}_{\text{out}} \xrightarrow{k_{\text{AI-2}}} \text{AI-2}_{\text{in}} \tag{5.4}$$

$$C_R + \text{AI-2}_{\text{in}} \xrightarrow{k_2} \text{DNA} \tag{5.5}$$

$$4R + \text{DNA} \xrightarrow{k_3} C_R \tag{5.6}$$

$$\text{DNA} \xrightarrow{k_4} \text{DNA} + \beta + R \tag{5.7}$$

For this model, we assume a first order reaction for the influx of AI-2 from the outside. We have simplified the intracellular dynamics and provide the three equations below which describe the β-gal enzyme concentration, the LsrR repressor concentration, and the amount of LsrR bound to DNA [118]. These equations base β-gal expression on the interplay between LsrR bound as a repressor and its freely dissociated form. LsrR repression is known to result from the formation of an LsrR tetramer bound between two distal looped strands of DNA comprising the intergenic region of the lsr regulon [42]. While a significant simplification of the known biochemistry, the topology of the expression kinetics agrees well with our experimental observations. For example, we do not include the phosphorylation of AI-2 by LsrK. We generalize the transcriptions on the pCT6 and pETLacZ into one equation that codes for LacZ and LsrR. Hence, we call k2 a generalized rate as it takes into account these factors. The law of mass action is used to translate these reactions into differential equations. The constraint $C_R + \text{DNA} = D_{total}$, in which D_{total} is a constant, simplifies the system of differential equations into:

$$\frac{d\beta}{dt} = k_4(D_{total} - C_R) - k_{d_2}\beta \tag{5.8}$$

$$\frac{dC_R}{dt} = k_3R^4(D_{total} - C_R) - k_2C_R\left(C_{\text{AI-2}}\left(1 - e^{-k_{\text{AI-2}}(t-1.5)}\right)\right) - k_{d_3}R \tag{5.9}$$

$$\frac{dR}{dt} = k_4(D_{total} - C_R) - 4k_3R^4(D_{total} - C_R) - k_{d_3}R \tag{5.10}$$

The application of this constraint is explained in Alon [3]. However, in our equations we did not apply the assumption that $\frac{dC_R}{dt} = 0$ (steady state amount of bound repressor).

We induced the synthesis of LacZ by the addition of AI-2; typically aiming for an initial OD of X. To facilitate errors in dynamic lag associated with altered OD's for induction and to account for variability in initial growth kinetics, we have set the initial condition so that the initial appearance of β-gal in the zero AI-2 control culture matched with the simulation at that point. This meant that the simulations commenced at t =1.5 hrs, or about a half hour before the initial β-gal measurement. The rest of the initial conditions are $C_R(1.5) = D_{total} - 1.175 \times 10^{-6}$ µM and $R(1.5) = 1.692 \times 10^{-5}$ µM. The differential equations were solved using MATLAB.

Table 2 shows the parameter values used in the model. Note that for k_4 (the transcription rate of LsrR and LacZ), we take into account the rate at which the endogenous RNA polymerase transcribes the T7RNA polymerase (2.65 kbp at 39 nt/s) [13] [35] as well as the rate at which the T7RNA polymerase transcribes LacZ

Name	Description
C_R	LsrR repressor bound to operon (ie acts as a cap over the DNA to block transcription)
AI-2_{out}	Extracellular AI-2
AI-2_{in}	Intracellular AI-2
DNA	represents DNA that codes for β and LsrR
R	LsrR repressor
β	Beta-galactosidase enzyme
$k_{\text{AI-2}}$	rate at which AI-2 enters cell
k_2	generalized rate of de-repression by AI-2
k_3	rate at which LsrR binds to operon
k_4	transcription rate
$C_{\text{AI-2}}$	initial AI-2 concentration
k_{d_2}	LacZ degradation rate
k_{d_3}	LsrR degradation rate

Table 5.1: Variable and Parameter Descriptions

Parameter Value	Source
$k_{\text{AI-2}} = 1.7/\text{hr}$	Li et al. (2006)
$k_2 = 0.73/(\mu M\text{-hr})$	fitted to data
$k_3 = 1.055 \times 10^{20}\ /(\mu M^4\text{-hr})$	fitted to data
$k_4 = (1/83.324)/\text{s}$	Fuerst et al. (1986), Bremer/Dennis (1996), Iost et al.(1991), Kanehisa et al.(2000, 2014)
$D_{total} = 13.1 \times 10^{-5} \mu M$	fitted to data
$k_{d_2} = 0.6/\text{hr}$	Santillan et al. (2007)
$k_{d_3} = 0.6/\text{hr}$	assumed (same as LacZ degradation rate)

Table 5.2: Parameter values and source

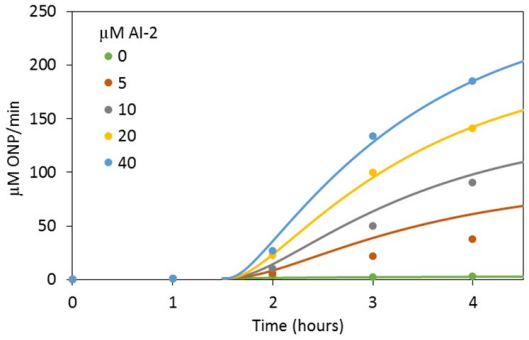


Figure 5.12: Cell model and data overlay of experimental data.

(3.075 kbp at 200 nt/s) [54] [57] [58]. Figure 5 shows the resulting first sub-model.

Converting µM ONP/min data from model to Miller Units for Figure 7a

In order to convert the data outputted by the model, in µM ONP/min, Miller

Units were calculated as follows:

$$\text{Miller Unit} = \frac{1000 \times (OD_{420} - 1.75 \times OD_{550})}{(\text{volume} \times \text{reaction time} \times OD_{600})} \tag{5.11}$$

To convert the model's output to an absorbance value at 420 nm (OD_{420}) we used Beer's Law as described above. Since the model data was not actual data from cells, we used values for OD_{550}, volume, reaction time, and OD_{600} that were similar to those from the experimental data.

For the OD_{550}, since the experimental values were all almost identical, we used the overall mean for all conversions. The volume and time stayed the same as in the experimental data. The OD_{600}, indicative of the amount of cells present, was averaged for all samples treated with the same AI-2 concentration. These values were then used to calculate the corresponding Miller Units for the model data and plotted along with the experimental data in Figure 7a.

5.7.2 The second compartment

Data Conversion

First, the data were converted into µM. The data describe concentrations of PAP in response to different levels of β-gal. We generated a standard curve that converts current (µA) to µM of PAP. Then, we convert to units of β-gal. To convert units of purchased and powdered β-gal into µM of β-gal, we introduced a factor of $\frac{1}{4}$ (ie $\frac{1}{4}$ gram of β-gal for every 1 gram of powdered β-gal). This factor can also compensate for the differences between cell-generated β-gal activity and store-bought β-gal

activity. After fitting the model (explained below), the resulting turnover rate, k_{cat}, is closer (103.26/s) to the value stated in literature (90/s) [107].

Model Fitting

Michaelis-Menten equations were again used for the PAPG substrate and β-gal enzyme. However, the parameter values (turnover rate k_{cat} and Michaelis constant K_m) were not documented as well as the ONPG substrate. Viratelle and Yon [107] provided values of $k_{cat} = 90$/s and $K_m = 330$ µM. Since these values, nor any combination of k_{cat} and K_m, do not produce a reasonable fit that encompasses all our data, we fixed K_m and searched for the k_{cat} that minimized the distance between the model and the data points. Overall, the values we found were within expected variances based on our experience. ($k_{cat} = 103.26$/s and $K_m = 330$ µM). Similar to above in which µM of ONPG was converted into µM of β-gal, we invoked a similar assumption that $K_m \ll$ PAPG (excess PAPG). The PAP differential equation of Michaelis-Menten then simplified to:

$$\frac{d(PAP)}{dt} = k_{cat} E_o \tag{5.12}$$

Since we had PAP (µM) at 5 time points for 4 nonzero levels of β-gal, the slopes were estimated by resolving a linear fit of each of the four data sets. The slopes were applied to the above equation for each level of enzyme to obtain 4 k_{cat} values. This provided a range of k_{cat} values that was used to find the best fit k_{cat} by calculating the least squares fit of the first four time points (as the fifth time point of one of

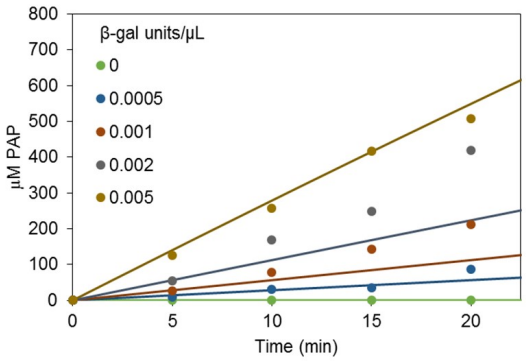


Figure 5.13: PAP model and overlay of experimental data.

the datasets seemed noisy). The resulting turnover rate was $k_{cat} = 103.26/\text{s}$.

In this way, we used the average $k_{cat} = 103.26/\text{s}$ and the literature value of K_m (330 µM) for our Michaelis-Menten model, shown below. We used this model to describe PAP activity after the addition of PAPG to β-gal (E_o).

$$\frac{d(PAP)}{dt} = \frac{k_{cat}E_o(PAPG)}{K_m + PAPG} \tag{5.13}$$

$$\frac{d(PAPG)}{dt} = -\frac{k_{cat}E_o(PAPG)}{K_m + PAPG} \tag{5.14}$$

The model output is shown below.

5.7.3 The combined model

The above models (equations 8-10 and 12-13) were then combined into one system of differential equations. The output of the combined model was compared to the data collected from the lysed cells in which AI-2 was added to the cells and

the levels of PAP were measured. (Note that cell lysate was used so our results would not include any time lag associated with transport of small molecules (AI-2) or proteins (β-gal) passing through the membrane.) The data collected represents the PAPG that was added to the cell extracts after hour 3, with the PAP levels measured in the subsequent 15 minutes, at 5 minute intervals. The model was run accordingly and the predictions are shown in Figure 3 along with the data points. The data support well, the Michaelis-Menten assumptions and constants evaluated here and reported by others [13] [35] [54] [57] [58] [63] [94].

Again, there is reasonable agreement between the models and the corresponding data, suggesting the mechanisms underpinning the system of ODEs represents a reasonable representation of the actual system. We do not, however, claim that because the model simulations and the data match, the mechanistic bases are validated. Instead, we suggest that the mechanistic basis for the *in vitro* electrochemical reactions can be faithfully represented by Michaelis-Menten kinetics, largely supported by constants obtained by previous researchers, and that the simplified cell based model describing the kinetics of β-gal expression is sufficiently accurate so that any future cell-based model describing LacZ kinetics could be added to sub-model 2 described here. While this is conjecture at this time, it is the primary motivation for partitioning the model as described and for including here.

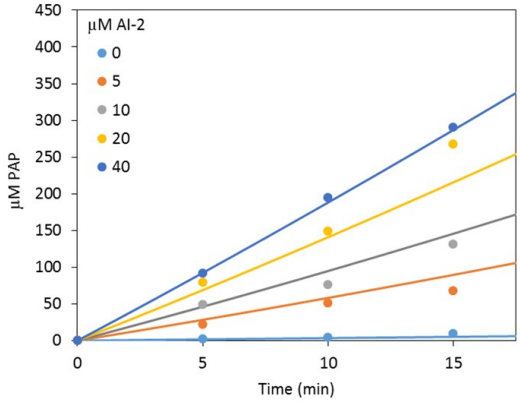


Figure 5.14: Final model predictions and lysed cell overlay of experimental data.

5.8 Determining of a unit for electrochemical measurements analogous to the Miller Unit

A unit analogous to the Miller Unit would ideally be developed for normalizing electrochemical PAP measurements. Below we use our data to demonstrate the concept. Here lysed cells will be used since those measurements are most similar to the way the spectrophotometric measurements are done to get the Miller Units.

$$\text{Miller Unit} = \frac{1000 \times (OD_{420} - 1.75 \times OD_{550})}{(\text{volume} \times \text{reaction time} \times OD_{600})} \tag{5.15}$$

In the Miller Unit the OD_{420} indicates the resulting ONP generated in a volume of cells (of OD_{600}) after the indicated reaction time. The OD_{550} indicates the scatter from cell debris, and when multiplied by 1.75 approximates the scatter at 420nm.

AI-2 concentration (µM)	Peak current at end (µA)	Time (min)	Volume (mL)	Cell OD_{600}	Electrochemical Units
0	0.093	15	1	2	0.003
5	0.631	15	1	2	0.021
10	1.220	15	1	2	0.041
20	2.490	15	1	2	0.083
40	2.710	15	1	2	0.090

Table 5.3: Calculation of Electrochemical units

$$\text{Electrochemical Unit} = \frac{\text{current}}{(\text{volume} \times \text{reaction time} \times OD_{600})}$$

Our analogous unit would take into account the current (in µA) measured after the indicated time (min) of a specific volume (mL) of cells of a certain OD_{600}.

In order to further develop and standardize the electrochemically-measured unit, measurements would need to be performed with various cell amounts and volumes and using different electrode materials, sizes, and setups (the units would likely be different based on the electrodes used). Additionally, any background interference between the cell lysate and the electrode surface would need to be studied and taken into account, especially for smaller electrodes. Below are our sample calculations for lysed cell measurements from Figure 4:

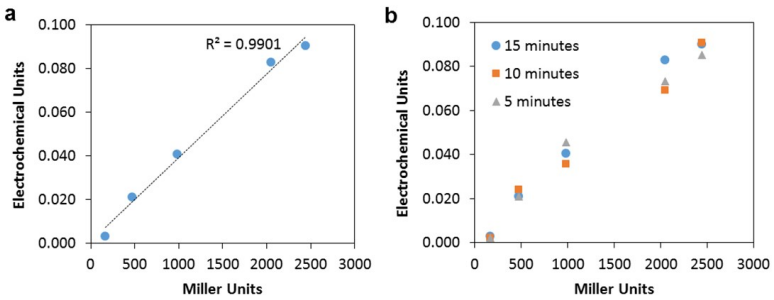


Figure 5.15: a. Linear relationship between the Miller Units measure from the lysed cells in Figure 4c and the electrochemical units calculated as described above from the same cells. b. Linear relationship is also maintained with the 5, and 10 minute measurements.

When plotted against the Miller Units from Figure 4c, there is a very good correlation (Supplementary Figure 5.15 below), though again, to standardize this unit much more data would have to be used.

Chapter 6: Conclusions

Mathematical models were developed in this dissertation to gain a deeper understanding of quorum sensing bacteria. Flocking models provide for a non-mechanistic representation of a dynamical system. By representing gene expression and motility as a form of flocking, we can model the system without the specific knowledge of detailed regulatory mechanisms that vary from species to species.

In our study, we used asymptotic analysis and simulations to explore the complex dynamics that results from quorum sensing, chemotaxing bacteria. Simulations suggest that the chemotaxis system may be influenced by density-dependent mechanisms. Two groups form in our simulations that do not appear when chemotaxis is not influenced by autoinducer-triggered gene expression. The dense portion of our cells increase in velocity and separate from the rest of the group. The remaining cells take longer to catch up, forming two groups of cells. This provides a possible explanation into the phenomenon observed in chemotaxis experiments [1].

The ordinary differential equations modelling velocity and gene expression have a flocking term and source term. Logistic functions are used to model threshold values in which velocity is influenced by gene expression and gene expression is influenced by autoinducer concentrations. By adjusting these threshold values, our

simulations showed similar qualitative behavior to that observed in experiments in Servinsky et al. [97].

We also developed a kinetic model from our flocking model in which we explored the application of statistical physics towards cells. The last model presented was a mechanistic compartmental model of an electrochemical method for measuring quorum sensing activity.

The differential equations in this dissertation aimed to gain insight or predict behavior of the complex system that is quorum sensing. By developing models, mechanistic or nonmechanistic, they help make connections that can be difficult to intuit from experiments.

Bibliography

[1] Julius Adler. Chemotaxis in bacteria. *Science*, 153(3737):708–716, 1966.

[2] Julius Adler. Chemotaxis in bacteria. In *Biochemistry of Sensory Functions*, pages 107–131. Springer, 1974.

[3] Uri Alon. *An introduction to systems biology: design principles of biological circuits*. CRC press, 2006.

[4] Eva Baldrich, Francesc Xavier Muñoz, and Cristina García-Aljaro. Electrochemical detection of quorum sensing signaling molecules by dual signal confirmation at microelectrode arrays. *Analytical chemistry*, 83(6):2097–2103, 2011.

[5] Radu Balescu. *Statistical mechanics of charged particles*, volume 4. Interscience New York, 1963.

[6] Michele Ballerini, Nicola Cabibbo, Raphael Candelier, Andrea Cavagna, Evaristo Cisbani, Irene Giardina, Alberto Orlandi, Giorgio Parisi, Andrea Procaccini, Massimiliano Viale, et al. Empirical investigation of starling flocks: a benchmark study in collective animal behaviour. *Animal behaviour*, 76(1):201–215, 2008.

[7] Tarun Bansal, Palmy Jesudhasan, Suresh Pillai, Thomas K Wood, and Arul Jayaraman. Temporal regulation of enterohemorrhagic *Escherichia coli* virulence mediated by autoinducer-2. *Applied microbiology and biotechnology*, 78(5):811–819, 2008.

[8] Bonnie L Bassler and Richard Losick. Bacterially speaking. *Cell*, 125(2):237–246, 2006.

[9] William E Bentley, Noushin Mirjalili, Dana C Andersen, Robert H Davis, and Dhinakar S Kompala. Plasmid-encoded protein: the principal factor in the metabolic burden associated with recombinant bacteria. *Biotechnology and bioengineering*, 35(7):668–681, 1990.

[10] Howard C Berg and PM Tedesco. Transient response to chemotactic stimuli in *Escherichia coli*. *Proceedings of the National Academy of Sciences*, 72(8):3235–3239, 1975.

[11] Israel Biran, Levcov Klimentiy, Regine Hengge-Aronis, Eliora Z Ron, and Judith Rishpon. On-line monitoring of gene expression. *Microbiology*, 145(8):2129–2133, 1999.

[12] Robert B Bourret and Ann M Stock. Molecular information processing: lessons from bacterial chemotaxis. *Journal of Biological Chemistry*, 277(12):9625–9628, 2002.

[13] Hans Bremer and Patrick P Dennis. *Escherichia coli* and *Salmonella*: cellular and molecular biology. *Washington (DC): American Society for Microbiology. Chapter, Modulation of chemical composition and other parameters of the cell by growth rate*, pages 1553–1569, 1996.

[14] Jamie T Bridgham, Sean M Carroll, and Joseph W Thornton. Evolution of hormone-receptor complexity by molecular exploitation. *Science*, 312(5770):97–101, 2006.

[15] Elena O Budrene and Howard C Berg. Dynamics of formation of symmetrical patterns by chemotactic bacteria. *Nature*, 376:49–53, 1995.

[16] Susan M Butler and Andrew Camilli. Going against the grain: chemotaxis and infection in *Vibrio cholera*. *Nature reviews Microbiology*, 3(8):611–620, 2005.

[17] Jarred M Callura, Charles R Cantor, and James J Collins. Genetic switchboard for synthetic biology applications. *Proceedings of the National Academy of Sciences*, 109(15):5850–5855, 2012.

[18] JA Canizo, JA Carrillo, and J Rosado. Collective behavior of animals: Swarming and complex patterns. *Arbor*, 186:1035–1049, 2010.

[19] I-Jane Chen and Ian M White. High-sensitivity electrochemical enzyme-linked assay on a microfluidic interdigitated microelectrode. *Biosensors and Bioelectronics*, 26(11):4375–4381, 2011.

[20] David L Chopp, Mary Jo Kirisits, Brian Moran, and Matthew R Parsek. A mathematical model of quorum sensing in a growing bacterial biofilm. *Journal of Industrial Microbiology and Biotechnology*, 29(6):339–346, 2002.

[21] John S Chuang, Olivier Rivoire, and Stanislas Leibler. Simpson's paradox in a synthetic microbial system. *Science*, 323(5911):272–275, 2009.

[22] N Connell, Z Han, F Moreno, and R Kolter. An *E. coli* promoter induced by the cessation of growth. *Molecular microbiology*, 1(2):195–201, 1987.

[23] Fernando Cortés-Salazar, Siham Beggah, Jan Roelof van der Meer, and Hubert H Girault. Electrochemical As (iii) whole-cell based biochip sensor. *Biosensors and Bioelectronics*, 47:237–242, 2013.

[24] Felipe Cucker and Steve Smale. Emergent behavior in flocks. *IEEE Transactions on automatic control*, 52(5):852–862, 2007.

[25] Felipe Cucker and Steve Smale. On the mathematics of emergence. *Japanese Journal of Mathematics*, 2(1):197–227, 2007.

[26] Nicholas C Darnton, Linda Turner, Svetlana Rojevsky, and Howard C Berg. Dynamics of bacterial swarming. *Biophysical journal*, 98(10):2082–2090, 2010.

[27] Kirill A Datsenko and Barry L Wanner. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proceedings of the National Academy of Sciences*, 97(12):6640–6645, 2000.

[28] Ezio Di Costanzo, Roberto Natalini, and Luigi Preziosi. A hybrid mathematical model for self-organizing cell migration in the zebrafish lateral line. *Journal of Mathematical Biology*, 71(1):171–214, 2015.

[29] Jack D Dockery and James P Keener. A mathematical model for quorum sensing in *Pseudomonas aeruginosa*. *Bulletin of mathematical biology*, 63(1):95–116, 2001.

[30] Michael B Elowitz, Arnold J Levine, Eric D Siggia, and Peter S Swain. Stochastic gene expression in a single cell. *Science*, 297(5584):1183–1186, 2002.

[31] John T Emlen. Flocking behavior in birds. *The Auk*, pages 160–170, 1952.

[32] Derek L Englert, Michael D Manson, and Arul Jayaraman. Flow-based microfluidic device for quantifying bacterial chemotaxis in stable, competing gradients. *Applied and environmental microbiology*, 75(13):4557–4564, 2009.

[33] Magnus G Fagerlind, Patric Nilsson, Mikael Harlén, Sandra Karlsson, Scott A Rice, and Staffan Kjelleberg. Modeling the effect of acylated homoserine lactone antagonists in *Pseudomonas aeruginosa*. *Biosystems*, 80(2):201–213, 2005.

[34] Magnus G Fagerlind, Scott A Rice, Patric Nilsson, Mikael Harlén, Sally James, Timothy Charlton, and Staffan Kjelleberg. The role of regulators in the expression of quorum-sensing signals in *Pseudomonas aeruginosa*. *Journal of molecular microbiology and biotechnology*, 6(2):88–100, 2004.

[35] Thomas R Fuerst, Edward G Niles, F William Studier, and Bernard Moss. Eukaryotic transient-expression system based on recombinant vaccinia virus that synthesizes bacteriophage T7 RNA polymerase. *Proceedings of the National Academy of Sciences*, 83(21):8122–8126, 1986.

[36] W Claiborne Fuqua, Stephen C Winans, and E Peter Greenberg. Quorum sensing in bacteria: the LuxR-LuxI family of cell density-responsive transcriptional regulators. *Journal of bacteriology*, 176(2):269, 1994.

[37] Jordi Garcia-Ojalvo, Michael B Elowitz, and Steven H Strogatz. Modeling a synthetic multicellular clock: repressilators coupled by quorum sensing. *Proceedings of the National Academy of Sciences of the United States of America*, 101(30):10955–10960, 2004.

[38] Cheol-Min Ghim, Sung-Kuk Lee, Shuichi Takayama, and Robert J Mitchell. The art of reporter proteins in science: past, present and future applications. *BMB reports*, 43(7):451–460, 2010.

[39] AB Goryachev, DJ Toh, and T Lee. Systems analysis of a quorum sensing network: design constraints imposed by the functional requirements, network topology and kinetic constants. *Biosystems*, 83(2):178–187, 2006.

[40] Andrew B Goryachev, Da-Jun Toh, Keng Boon Wee, Travis Lee, Hai-Bao Zhang, and Lian-Hui Zhang. Transition to quorum sensing in an *Agrobacterium* population: A stochastic model. *PLoS Comput Biol*, 1(4):e37, 2005.

[41] Shay Gueron and Simon A Levin. Self-organization of front patterns in large wildebeest herds. *Journal of theoretical Biology*, 165(4):541–552, 1993.

[42] Jung-Hye Ha, Yumi Eo, Alexander Grishaev, Min Guo, Jacqueline AI Smith, Herman O Sintim, Eun-Hee Kim, Hae-Kap Cheong, William E Bentley, and Kyoung-Seok Ryu. Crystal structures of the LsrR proteins complexed with phospho-AI-2 and two signal-interrupting analogues reveal distinct mechanisms for ligand recognition. *Journal of the American Chemical Society*, 135(41):15526–15535, 2013.

[43] Seung-Yeal Ha and Doron Levy. Particle, kinetic and fluid models for phototaxis. *Discret. Contin. Dyn. Syst. B*, 12:77–108, 2009.

[44] Seung-Yeal Ha and Eitan Tadmor. From particle to kinetic and hydrodynamic descriptions of flocking. *arXiv preprint arXiv:0806.2182*, 2008.

[45] William D Hamilton. Geometry for the selfish herd. *Journal of Theoretical Biology*, 31(2):295–311, 1971.

[46] Kim Rachael Hardie and Karin Heurlier. Establishing bacterial communities by 'word of mouth': LuxS and autoinducer 2 in biofilm development. *Nature Reviews Microbiology*, 6(8):635–643, 2008.

[47] Colin G Hebert, Apoorv Gupta, Rohan Fernandes, Chen-Yu Tsao, James J Valdes, and William E Bentley. Biological nanofactories target and activate epithelial cell surfaces for modulating bacterial quorum sensing and interspecies signaling. *ACS nano*, 4(11):6923–6931, 2010.

[48] Manjunath Hegde, Derek L Englert, Shanna Schrock, William B Cohn, Christian Vogt, Thomas K Wood, Michael D Manson, and Arul Jayaraman. Chemotaxis to the quorum-sensing signal AI-2 requires the Tsr chemoreceptor and the periplasmic LsrB AI-2-binding protein. *Journal of bacteriology*, 193(3):768–773, 2011.

[49] Bernard D Hill and G Bruce Schaalje. A two-compartment model for the dissipation of deltamethrin on soil. *Journal of agricultural and food chemistry*, 33(5):1001–1006, 1985.

[50] Alan L Hodgkin and Andrew F Huxley. A quantitative description of membrane current and its application to conduction and excitation in nerve. *The Journal of physiology*, 117(4):500, 1952.

[51] Seok Hoon Hong, Manjunath Hegde, Jeongyun Kim, Xiaoxue Wang, Arul Jayaraman, and Thomas K Wood. Synthetic quorum-sensing circuit to control consortial biofilm formation and dispersal in a microfluidic device. *Nature communications*, 3:613, 2012.

[52] Sara Hooshangi and William E Bentley. LsrR quorum sensing "switch" is revealed by a bottom-up approach. *PLoS Comput Biol*, 7(9):e1002172, 2011.

[53] Kerson Huang. Statistical mechanics. 1963.

[54] I Iost, J Guillerez, and M Dreyfus. Bacteriophage T7 RNA polymerase travels far ahead of ribosomes in vivo. *Journal of bacteriology*, 174(2):619–622, 1992.

[55] Sally James, Patric Nilsson, Geoffrey James, Staffan Kjelleberg, and Torbjörn Fagerström. Luminescence control in the marine bacterium *Vibrio fischeri*: an analysis of the dynamics of *lux* regulation. *Journal of molecular biology*, 296(4):1127–1137, 2000.

[56] Douglas H Juers, Beatrice Rob, Megan L Dugdale, Nastaron Rahimzadeh, Clarence Giang, Michelle Lee, Brian W Matthews, and Reuben E Huber. Direct and indirect roles of his-418 in metal binding and in the activity of β-galactosidase (*E. coli*). *Protein Science*, 18(6):1281–1292, 2009.

[57] Minoru Kanehisa and Susumu Goto. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic acids research*, 28(1):27–30, 2000.

[58] Minoru Kanehisa, Susumu Goto, Yoko Sato, Masayuki Kawashima, Miho Furumichi, and Mao Tanabe. Data, information, knowledge and principle: back to metabolism in KEGG. *Nucleic acids research*, 42(D1):D199–D205, 2014.

[59] Takatoshi Kaya, Kuniaki Nagamine, Nobuto Matsui, Tomoyuki Yasukawa, Hitoshi Shiku, and Tomokazu Matsue. On-chip electrochemical measurement of β-galactosidase expression using a microbial chip. *Chemical communications*, (2):248–249, 2004.

[60] Francis H Ko and Harold G Monbouquette. Photometric and electrochemical enzyme-multiplied assay techniques using β-galactosidase as reporter enzyme. *Biotechnology progress*, 22(3):860–865, 2006.

[61] AJ Koerber, JR King, JP Ward, P Williams, JM Croft, and RE Sockett. A mathematical model of partial-thickness burn-wound infection by *Pseudomonas aeruginosa*: quorum sensing and the build-up to invasion. *Bulletin of mathematical biology*, 64(2):239–259, 2002.

[62] Olivier Laczka, Rosa María Ferraz, Neus Ferrer-Miralles, Antonio Villaverde, Francesc Xavier Muñoz, and F Javier del Campo. Fast electrochemical detection of anti-HIV antibodies: coupling allosteric enzymes and disk microelectrode arrays. *Analytica chimica acta*, 641(1):1–6, 2009.

[63] Jun Li, Liang Wang, Yoshifumi Hashimoto, Chen-Yu Tsao, Thomas K Wood, James J Valdes, Evanghelos Zafiriou, and William E Bentley. A stochastic model of *Escherichia coli* AI-2 quorum signal circuit reveals alternative synthesis pathways. *Molecular systems biology*, 2(1):67, 2006.

[64] Hua Ling, Aram Kang, Mui Hua Tan, Xiaobao Qi, and Matthew Wook Chang. The absence of the *luxS* gene increases swimming motility and flagella synthesis in *Escherichia coli* K12. *Biochemical and biophysical research communications*, 401(4):521–526, 2010.

[65] Summie Lo, Megan L Dugdale, Nisha Jeerh, Tabitha Ku, Nathan J Roth, and Reuben E Huber. Studies of Glu-416 variants of β-galactosidase (*E. coli*) show that the active site Mg2+ is not important for structure and indicate that the main role of Mg2+ is to mediate optimization of active site chemistry. *The protein journal*, 29(1):26–31, 2010.

[66] Michael D Manson, Judith P Armitage, James A Hoch, and Robert M Macnab. Bacterial locomotion and signal transduction. *Journal of bacteriology*, 180(5):1009–1022, 1998.

[67] Margaret McFall-Ngai. Divining the essence of symbiosis: insights from the squid-*Vibrio* model. *PLoS Biol*, 12(2):e1001783, 2014.

[68] Jeffrey H Miller. Assay of β-galactosidase. *Experiments in molecular genetics*, pages 352–355, 1972.

[69] Melissa B Miller and Bonnie L Bassler. Quorum sensing in bacteria. *Annual Reviews in Microbiology*, 55(1):165–199, 2001.

[70] Tim Miyashiro and Edward G Ruby. Shedding light on bioluminescence regulation in *Vibrio fischeri*. *Molecular microbiology*, 84(5):795–806, 2012.

[71] Tae Seok Moon, Chunbo Lou, Alvin Tamsir, Brynne C Stanton, and Christopher A Voigt. Genetic programs constructed from layered logic gates in single cells. *Nature*, 491(7423):249–253, 2012.

[72] Sebastien Motsch and Eitan Tadmor. A new model for self-organized dynamics and its flocking behavior. *Journal of Statistical Physics*, 144(5):923–947, 2011.

[73] Johannes Müller, Christina Kuttler, Burkard A Hense, Michael Rothballer, and Anton Hartmann. Cell–cell communication by quorum sensing and dimension-reduction. *Journal of mathematical biology*, 53(4):672–702, 2006.

[74] Nagarajan Nandagopal and Michael B Elowitz. Synthetic biology: integrated gene circuits. *science*, 333(6047):1244–1248, 2011.

[75] Mark E Nelson. Electrophysiological models of neural processing. *Wiley Interdisciplinary Reviews: Systems Biology and Medicine*, 3(1):74–92, 2011.

[76] Dwight Roy Nicholson and Dwight R Nicholson. *Introduction to plasma theory*. Cambridge Univ Press, 1983.

[77] Sarayut Nijvipakul, Janewit Wongratana, Chutintorn Suadee, Barrie Entsch, David P Ballou, and Pimchai Chaiyen. LuxG is a functioning flavin reductase for bacterial luminescence. *Journal of bacteriology*, 190(5):1531–1538, 2008.

[78] Kevin J Painter and Thomas Hillen. Volume-filling and quorum-sensing in models for chemosensitive movement. *Can. Appl. Math. Quart*, 10(4):501–543, 2002.

[79] Julia K Parrish, Steven V Viscido, and Daniel Grünbaum. Self-organized fish schools: an examination of emergent properties. *The biological bulletin*, 202(3):296–305, 2002.

[80] Paul F Pinsky and John Rinzel. Intrinsic and network rhythmogenesis in a reduced Traub model for CA3 neurons. *Journal of computational neuroscience*, 1(1-2):39–60, 1994.

[81] Rachela Popovtzer, Tova Neufeld, Dvora Biran, Eliora Z Ron, Judith Rishpon, and Yosi Shacham-Diamand. Novel integrated electrochemical nano-biochip for toxicity detection in water. *Nano letters*, 5(6):1023–1027, 2005.

[82] Wayne K Potts. The chorus-line hypothesis of manoeuvre coordination in avian flocks. 1984.

[83] Arthur Prindle, Phillip Samayoa, Ivan Razinkov, Tal Danino, Lev S Tsimring, and Jeff Hasty. A sensing array of radically coupled genetic 'biopixels'. *Nature*, 481(7379):39–44, 2012.

[84] David N Quan, Chen-Yu Tsao, Hsuan-Chen Wu, and William E Bentley. Quorum sensing desynchronization leads to bimodality and patterned behaviors. *PLoS Comput Biol*, 12(4):e1004781, 2016.

[85] Mikaela Rapp, Susanna Seppälä, Erik Granseth, and Gunnar Von Heijne. Emulating membrane protein evolution by rational design. *Science*, 315(5816):1282–1284, 2007.

[86] Linda E Reichl and Ilya Prigogine. *A modern course in statistical physics*, volume 71. University of Texas press Austin, 1980.

[87] Frederick Reif. Fundamentals of statistical and thermal physics. 1965.

[88] Craig W Reynolds. Flocks, herds and schools: A distributed behavioral model. *ACM SIGGRAPH computer graphics*, 21(4):25–34, 1987.

[89] Francisco J Romero-Campero and Mario J Pérez-Jiménez. A model of the quorum sensing system in *Vibrio fischeri* using p systems. *Artificial life*, 14(1):95–109, 2008.

[90] Nitzan Rosenfeld, Jonathan W Young, Uri Alon, Peter S Swain, and Michael B Elowitz. Gene regulation at the single-cell level. *Science*, 307(5717):1962–1965, 2005.

[91] Varnika Roy, Bryn L Adams, and William E Bentley. Developing next generation antimicrobials by intercepting AI-2 mediated quorum sensing. *Enzyme and microbial technology*, 49(2):113–123, 2011.

[92] Edward G Ruby. Lessons from a cooperative, bacterial-animal association: the *Vibrio fischeri-Euprymna scolopes* light organ symbiosis. *Annual Reviews in Microbiology*, 50(1):591–624, 1996.

[93] Steven T Rutherford and Bonnie L Bassler. Bacterial quorum sensing: its role in virulence and possibilities for its control. *Cold Spring Harbor Perspectives in Medicine*, 2(11):a012427, 2012.

[94] M Santillán, MC Mackey, and ES Zeron. Origin of bistability in the *lac* operon. *Biophysical journal*, 92(11):3830–3842, 2007.

[95] Stephan Schauder, Kevan Shokat, Michael G Surette, and Bonnie L Bassler. The LuxS family of bacterial autoinducers: biosynthesis of a novel quorum-sensing signal molecule. *Molecular microbiology*, 41(2):463–476, 2001.

[96] Donna L Scott, Sridhar Ramanathan, Weiping Shi, Barry P Rosen, and Sylvia Daunert. Genetically engineered bacteria: electrochemical sensing systems for antimonite and arsenite. *Analytical chemistry*, 69(1):16–20, 1997.

[97] Matthew D Servinsky, Jessica L Terrell, Chen-Yu Tsao, Hsuan-Chen Wu, David N Quan, Amin Zargar, Patrick C Allen, Christopher M Byrd, Christian J Sund, and William E Bentley. Directed assembly of a bacterial quorum. *The ISME journal*, 10(1):158–169, 2016.

[98] David Sprinzak and Michael B Elowitz. Reconstruction of genetic circuits. *Nature*, 438(7067):443–448, 2005.

[99] Gregory Stephanopoulos. Synthetic biology and metabolic engineering. *ACS synthetic biology*, 1(11):514–525, 2012.

[100] Hendrik Szurmant and George W Ordal. Diversity in chemotaxis mechanisms among the bacteria and archaea. *Microbiology and Molecular Biology Reviews*, 68(2):301–319, 2004.

[101] Roger D Traub, RK Wong, Richard Miles, and Hillary Michelson. A model of a CA3 hippocampal pyramidal neuron incorporating voltage-clamp data on intrinsic conductances. *Journal of Neurophysiology*, 66(2):635–650, 1991.

[102] Tony Yu-Chen Tsai, Yoon Sup Choi, Wenzhe Ma, Joseph R Pomerening, Chao Tang, and James E Ferrell. Robust, tunable biological oscillations from interlinked positive and negative feedback loops. *Science*, 321(5885):126–129, 2008.

[103] Chen-Yu Tsao, Sara Hooshangi, Hsuan-Chen Wu, James J Valdes, and William E Bentley. Autonomous induction of recombinant proteins by minimally rewiring native quorum sensing regulon of *E. coli*. *Metabolic engineering*, 12(3):291–297, 2010.

[104] Tanya Tschirhart, Xinyi Y Zhou, Hana Ueda, Chen-Yu Tsao, Eunkyoung Kim, Gregory F Payne, and William E Bentley. Electrochemical measurement of the β-galactosidase reporter from live cells: a comparison to the Miller assay. *ACS synthetic biology*, 5(1):28–35, 2015.

[105] Pauline Van den Driessche and James Watmough. Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission. *Mathematical biosciences*, 180(1):29–48, 2002.

[106] Tamás Vicsek, András Czirók, Eshel Ben-Jacob, Inon Cohen, and Ofer Shochet. Novel type of phase transition in a system of self-driven particles. *Physical review letters*, 75(6):1226, 1995.

[107] Odile M Viratelle and Jeannine M Yon. Nucleophilic competition in some β-galactosidase-catalyzed reactions. *European Journal of Biochemistry*, 33(1):110–116, 1973.

[108] Alessandro Usseglio Viretta and Martin Fussenegger. Modeling the quorum sensing regulatory network of human-pathogenic *Pseudomonas aeruginosa*. *Biotechnology progress*, 20(3):670–678, 2004.

[109] Karen L Visick, Jamie Foster, Judith Doino, Margaret McFall-Ngai, and Edward G Ruby. *Vibrio fischeri lux* genes play an important role in colonization and development of the host light organ. *Journal of Bacteriology*, 182(16):4578–4586, 2000.

[110] George H Wadhams and Judith P Armitage. Making sense of it all: bacterial chemotaxis. *Nature Reviews Molecular Cell Biology*, 5(12):1024–1037, 2004.

[111] Liang Wang, Yoshifumi Hashimoto, Chen-Yu Tsao, James J Valdes, and William E Bentley. Cyclic AMP (cAMP) and cAMP receptor protein influence both synthesis and uptake of extracellular autoinducer 2 in *Escherichia coli*. *Journal of bacteriology*, 187(6):2066–2076, 2005.

[112] Liang Wang, Jun Li, John C March, James J Valdes, and William E Bentley. *luxS*-dependent gene regulation in *Escherichia coli* K-12 revealed by genomic expression profiling. *Journal of bacteriology*, 187(24):8350–8360, 2005.

[113] Yifeng Wang, Yi Liu, Yi Cheng, Eunkyoung Kim, Gary W Rubloff, William E Bentley, and Gregory F Payne. Coupling electrodeposition with layer-by-layer assembly to address proteins within microfluidic channels. *Advanced Materials*, 23(48):5817–5821, 2011.

[114] John P Ward, John R King, Adrian J Koerber, Julie M Croft, R Elizabeth Sockett, and Paul Williams. Early development and quorum sensing in bacterial biofilms. *Journal of mathematical biology*, 47(1):23–55, 2003.

[115] John P Ward, John R King, AJ Koerber, P Williams, JM Croft, and RE Sockett. Mathematical modelling of quorum sensing in bacteria. *Mathematical Medicine and Biology*, 18(3):263–292, 2001.

[116] John P Ward, JR King, AJ Koerber, JM Croft, RE Sockett, and P Williams. Cell-signalling repression in bacterial quorum sensing. *Mathematical Medicine and Biology*, 21(3):169–204, 2004.

[117] Christopher M Waters and Bonnie L Bassler. Quorum sensing: cell-to-cell communication in bacteria. *Annu. Rev. Cell Dev. Biol.*, 21:319–346, 2005.

[118] Ting Xue, Liping Zhao, Haipeng Sun, Xianxuan Zhou, and Baolin Sun. LsrR-binding site recognition and regulatory characteristics in *Escherichia coli* AI-2 quorum sensing. *Cell research*, 19(11):1258–1268, 2009.

[119] Christian A Yates, Radek Erban, Carlos Escudero, Iain D Couzin, Jerome Buhl, Ioannis G Kevrekidis, Philip K Maini, and David JT Sumpter. Inherent noise can facilitate coherence in collective swarm motion. *Proceedings of the National Academy of Sciences*, 106(14):5464–5469, 2009.

[120] Lingchong You, Robert Sidney Cox, Ron Weiss, and Frances H Arnold. Programmed population control by cell–cell communication and regulated killing. *Nature*, 428(6985):868–871, 2004.