

Mechanistic Modeling of Bacterial Nutrient Uptake Strategies

by

Noele Rosalie Norris

B.S., California Institute of Technology (2010)

M.A., Queen Mary, University of London (2011)

M.S., Massachusetts Institute of Technology (2013)

Submitted to the Department of Electrical Engineering and Computer Science
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 2019

©Massachusetts Institute of Technology 2019. All rights reserved.

Author
Department of Electrical Engineering and Computer Science
August 30, 2019

Certified by
Emilio Frazzoli
Professor, Mechanical and Process Engineering, ETH Zürich
Thesis Supervisor

Certified by
Roman Stocker
Professor, Civil, Environmental, and Geomatic Engineering, ETH Zürich
Thesis Supervisor

Accepted by
Leslie A. Kolodziejcki
Professor of Electrical Engineering and Computer Science
Chair, Department Committee on Graduate Students

Mechanistic Modeling of Bacterial Nutrient Uptake Strategies

by

Noele Rosalie Norris

Submitted to the Department of Electrical Engineering and Computer Science on

August 30, 2019

in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Abstract

Bacteria have developed a variety of strategies to find and consume the substrates necessary for both the cell's energy-consuming processes and for the additional biomass needed to replicate. A greater understanding of the diversity and regulation of these strategies can provide us with a number of insights relevant for a variety of applications, from predicting bacterial population dynamics and thus carbon-cycling rates in the ocean to bio-engineering bacteria into microscale robots. Here I use toy, mechanistic models of single-cell metabolism that allow me to quantify the costs and benefits of various nutrient uptake strategies. I find that: *(i)* a sensing-uptake trade-off governs *E. coli*'s regulation of maltose uptake and chemotaxis to maltose; *(ii)* a rate-affinity trade-off in nutrient transport systems governs the speciation of marine oligotrophic and copiotrophic heterotrophs; and *(iii)* an exploration-conservation trade-off governs the prevalence of motility in the marine microbial world. This work thus provides new understanding of how both phenotypic diversity and cellular regulation are governed by trade-offs for maximizing growth rate in different environments.

Thesis Supervisor: Emilio Frazzoli

Title: Professor, Mechanical and Process Engineering, ETH Zürich

Thesis Supervisor: Roman Stocker

Title: Professor, Civil, Environmental, and Geomatic Engineering, ETH Zürich

In memory of my dad, my hero.

“The best thing for being sad,” replied Merlyn, beginning to puff and blow, “is to learn something. That is the only thing that never fails. You may grow old and trembling in your anatomies, you may lie awake at night listening to the disorder of your veins, you may miss your only love, you may see the world about you devastated by evil lunatics, or know your honour trampled in the sewers of baser minds. There is only one thing for it then—to learn. Learn why the world wags and what wags it. That is the only thing which the mind can never exhaust, never alienate, never be tortured by, never fear or distrust, and never dream of regretting. Learning is the thing for you....”

T.H. White, *The Once and Future King*

Acknowledgements

I think I have had an atypical graduate student career. It took me from MIT to ETH Zürich and USC. It provided me with the opportunity to gain different perspectives to research, to pique and satisfy my curiosity in a variety of topics, and even to experience different worldviews. It was additionally a journey of self-discovery. There were moments I was banging my head against the wall, wishing I were just a bit smarter, and others—though I will admit more rare—in which I felt the distinct joy of having brilliantly figured something out. I have learned that this is what doing science is all about, and I would like to acknowledge and thank my three thesis supervisors for giving me the opportunity to learn what it means to be a scientist.

Foremost, I would like to thank Professors Emilio Frazzoli and Roman Stocker. They both ensured that I always had a pay-check each month. They gave me countless resources and opportunities, and they also gave me an incredibly rare gift: the freedom to develop my intellectual curiosity nearly wherever it took me—from taking classes to learn functional analysis to donning a lab coat and learning how to culture *E. coli*. I would like to thank Emilio for teaching me to be open-minded and curious, whether it be about robotic cars, unmanned aerial vehicles, locusts, zebras, or bacteria; and for never once faltering in his support and for ensuring I made it out with a doctoral degree. And I would like to thank Roman for taking me in and integrating a novice experimentalist into his welcoming lab; and for giving me the wonderfully inspiring opportunity to participate in the Simons collaborative grant for Principles of Microbial Ecology.

I would also like to thank my third thesis supervisor, Professor Naomi Levine, for providing me with an intellectual home in Los Angeles. She allowed me to feel part of a close-knit team tackling big problems. She always found the time to talk to me about my work, being both sounding-board and mentor and infusing positivity into my research whenever it was lacking.

I would also like to thank my other two thesis committee members, Professors Munther Dahleh and George Verghese. Munzer convinced me to attend MIT and join Emilio's group and advised me about the practicalities of academia, even though he may just well be the busiest person on campus. And George came to my rescue at a time of great need and took the time to learn about my research and offer useful feedback.

My success is due to the help of many other people. First of all, I would like to thank the following former and current members of the Stocker lab: Vicente Fernandez for his guidance and for ensuring I was on task and on target; Filippo Menolascina for introducing me to the world of the experimentalist and for convincing me I could succeed; Yutaka Yawata for showing me his awe-inspiring experimental wizardry, which finally gave me the insight that allowed me to get my own experiments working; Ela Burmeister for being the most efficient and brilliant lab manager in the universe and also for being a great friend; Jen Nguyen for being a wonderful labmate, for her friendship and guidance

when I was getting started, and for her optimism and cheer that helped me overcome my struggles in the lab; Uria Alcolombri for sharing his enthusiasm and curiosity with me, for spending so much time helping me produce beautiful Western blots, and for showing me that experimental work could be fun; Zachary Landry for opening my eyes to the wonders of the little oligotroph; Mehdi Salek for taking the time and care to help me fabricate my first microfluidic devices; Becky Schilling for being so unbelievably supportive of such a klutzy, experimental novice; Jeanette Wheeler for her camaraderie in the office; Cherry Gao for her help and cheerleading as I reached the finish line; Kwangmin Son for teaching me about the ways of *Vibrio*; Russ Naisbit for being a thoughtful and careful editor; and Joëlle Robinson for her kindness and for ensuring my paychecks made it across the ocean each and every month.

I would also like to acknowledge the help I received from current and former members of the Levine lab, in particular: Emily Zakem for her inspiring work, guidance, and mentorship; and Erin McParland for the sunny disposition that helped me start each day and for continually sending useful papers relevant to my research over my way.

I would also like to thank the staff at LIDS and the EECS headquarters who saved me from much headache, especially: Janet Fischer, the master wizard of the EECS department, without whose tremendous help I would have been unable to make the move to ETH; Alicia Duarte for being always cheerful and ready to help and for so quickly setting everything right; and Jennifer Donovan for her speedy efficiency and for making LIDS a warm and welcoming place to work. I would also like to thank EECS departmental chair Leslie Kolodziejki for taking the time to provide me with both guidance and moral support.

I would also like to thank: Professor Terry Hwa, at UCSD, for his encouragement and feedback on my work; and Professor Cameron Thrash, at USC, for giving me many useful insights into the workings of SAR11.

I am also very grateful to have had the support of the following people, who were there for me when I most needed them: Mitra Osqui for teaching me how the pros do control theory, for her friendship and hugs, and for always taking such good care of me whenever I am in town; Omer Tanovic for being the best officemate ever and for inspiring me with his brilliance and also engaging conversations; Sandhya Ramakrishnan for inspiring me to excel in mathematics and for always checking up on me; Barbara Gee, my Caltech academic grandmother, whose support helped me succeed and whose encouragement keeps me going; and Nancy Sulahian for her understanding and for her encouragement to keep singing.

Finally, I would like to acknowledge and thank my family for their love and support. To my mom, for always being there for me to listen and to advise and for teaching me grit and perseverance and continually reminding me to “just keep swimming”. To my sister, Kelsey Norris, whose hard work and enthusiasm for history has inspired me to be ambitious with my own career and whose thoughtfulness and support is constantly

inspiring me to be a better sister. To Alina, Laz, Ilana, Sisi, Patty, and Tammy, for always being there for my family and for me.

To my fiancé, Edward Chen, for putting up with me for so, so long already—through the ‘good or lean years’; for teaching me optimism and showing me joy; for making me laugh and for recalling bittersweet memories I had nearly forgotten that make me cry; for never once faltering in his support; and for taking such good care of me. I am excited about sharing with him the joys that life brings and am grateful to have him by my side as I face its challenges and disappointments. I would also like to thank my future brother- and sister-in-law, Howard and Angela, and, especially, my future parents-in-law, Sam and Jane, for making me feel a part of the family.

I dedicate my Ph.D. work to the memory of my dad. He sparked and nurtured my love for math and science. He taught me the value of integrity and discipline. And he would never let me give up.

Contents

1	Introduction	19
2	Chemotaxis in <i>E. coli</i>	23
2.1	<i>E. coli</i> 's run-and-tumble chemotaxis strategy	24
2.2	Pioneering experiments that characterized <i>E. coli</i> chemotaxis	25
2.3	The signal transduction pathway	28
2.3.1	Signal amplification	30
2.3.2	A “coarse-grained model” of the chemotactic pathway	31
2.3.2.1	An Ising model	31
2.3.2.2	A MWC model	32
2.4	Multiple signal integration	34
2.4.1	A heterogeneous MWC model	34
2.4.2	Signal integration within the same receptor	35
3	The Effects of Transport on <i>E. coli</i>'s Chemotaxis to Maltose	37
3.1	A transport-and-sensing chemotaxis model	40
3.2	Modeling steady-state population distributions	45
3.3	Chemotaxis assays for single and opposing chemotaxis gradients	46
3.4	Using population-level data to fit a molecular-level model	50
3.5	A sensitivity analysis uncovers a tug-of-war between sensing and transport	52
3.6	Varying growth conditions to modify chemotactic affinity	54
3.7	Conclusions	55
4	The Role of ABC Transport in the Oligotrophic Ocean	57
4.1	Modeling nutrient uptake	60
4.1.1	Michaelis-Menten kinetics	60
4.1.2	Periplasmic uptake	60
4.1.3	Cytoplasmic uptake	62
4.1.4	Modifying the half-saturation concentration of uptake via protein expression	65
4.2	A toy metabolic model	67
4.2.1	Proteomic cost of transport	68
4.2.2	Energetic cost of transport	68

4.2.3	Size constraints	70
4.2.4	A proteome allocation problem	70
4.3	Optimal proteome allocation for ABC versus direct transport systems	74
4.3.1	A rate-affinity trade-off	74
4.3.2	Cell size: a master trait	77
4.3.3	Matching transport to catabolic capacity	79
4.3.4	The cost of high specific affinity	82
4.3.5	Effects of solute size	84
4.4	Conclusions	88
5	To Swim, or Not to Swim	91
5.1	A toy model of motility	93
5.1.1	The benefit of swimming	93
5.1.1.1	Effects of cell's phenotype on particle encounter rate	94
5.1.1.2	Modeling the nutrient landscape	96
5.1.2	The costs of swimming	96
5.1.2.1	The proteomic cost of swimming	96
5.1.2.2	The energetic cost of swimming	97
5.1.3	A dual proteome allocation problem	98
5.2	Results	99
5.2.1	Optimal swimming speeds	99
5.2.2	Varying the particle interaction model	102
5.2.3	The energetic cost of swimming	105
5.2.3.1	The cost of always swimming	105
5.2.3.2	Removing the energetic costs of swimming	107
5.2.4	Implications of chemotaxis	109
5.2.5	Constraining cell size	110
5.3	Conclusions	111
6	Outlook	115
A	Modeling <i>E. coli</i> Chemotaxis to Maltose	117
A.1	An ABC transport model	117
A.2	A chemotaxis model	120
A.2.1	The SPECS model	120
A.2.2	A model for maltose chemotaxis	125
A.2.3	The Heterogeneous MWC model	126
A.3	An analytical approximation of the steady-state distribution	126
A.3.1	Fit using previous indirect binding chemotaxis model	128
A.3.2	Fit using transport-and-sensing chemotaxis model	128
A.3.3	Adding a saturation term	129

A.4	Methods	130
A.5	Supplemental figures	135
B	Proteome Allocation Problem for Nutrient Uptake	141
B.1	Proteome allocation problem	141
B.1.1	Periplasmic uptake	145
B.1.2	Cytoplasmic uptake	146
B.1.2.1	Direct transport	146
B.1.2.2	ABC transport	147
B.1.3	Catabolism	148
B.1.4	Membrane biosynthesis	148
B.1.5	Protein expression	149
B.1.6	Energy generation	150
B.1.7	ADP synthesis	151
B.1.8	Density and surface area constraints	151
B.2	Solving the optimization problem	152
B.2.1	A note on calculating maximal rate of downstream enzymes	152
	Bibliography	155

Science meant for me the most elevating form of revolt against the incoherence of the universe.

Francois Jacob, *The Statue Within*, 1988

One of the fundamental characteristics common to all living beings without exception [is] that of being objects endowed with a purpose or project, which ... they exhibit in their structure and carry out through their performances.

Jacques Monod, *Chance and Necessity*, 1970

This is the ocean, silly. We're not the only two in here.

Dory, the regal tang fish in *Finding Nemo*

Chapter 1

Introduction

Bacteria are among the most successful living organisms on our planet. While some can replicate in less than ten minutes [209], others can survive over a year exposed to cosmic radiation and the vacuum of space [205]. Bacterial cells outnumber human cells in a human body [177], and it has been estimated that there are on the order of 10^{30} bacterial cells on earth [55], outnumbering estimates of the number of stars in the universe by a factor of a million.

Bacteria's ubiquity, diversity, and reproductive success have profound implications for humans and for the other habitats in which they reside. Pathogenic bacteria are the culprits of tuberculosis, pneumonia, the bubonic plague, and cholera, among many other human diseases. Yet, work on the human microbiome suggests that humans live in a symbiotic relationship with countless species of bacteria that is essential to human health [204]. And, in this work, I study both the behavior of *E. coli*, an enteric bacterium that has been genetically modified to produce a variety of compounds, such as life-saving insulin [67]; as well as the behavior of marine heterotrophic bacteria, which via their consumption of nutrients and respiration of carbon dioxide shape the marine chemical landscape and govern the cycling of carbon and other compounds in the ocean [59].

A greater understanding of the behavior of these bacteria is required for a variety of applications. Gaining insight into how *E. coli* forages for nutrients may allow us to take further advantage of our ability to re-engineer it so that we could use it as a microscale robot for the delivery of medical treatments [86, 208]. And gaining insight into the ecology

of marine heterotrophic bacteria may allow us to better predict how climate change will affect levels of carbon in the ocean and the atmosphere [57].

My work is premised on the assumption that underpinning an understanding of microbial behavior and ecology is a theory of microbial metabolism. Evolution’s natural selection has “endowed [bacteria] with a purpose” [139]—that of survival and replication. Bacteria have been evolving for about four billion years, essentially ‘optimizing’ their ability to replicate their “selfish genes” [45]. But what is optimal? How does a species maximize the propagation of its genes? While survival requires, among other traits, successful dispersal in the environment, predator avoidance, and virus avoidance [50], my work focuses on various strategies that bacteria employ to acquire the nutrients needed for energy-consuming processes and for the additional biomass needed to replicate. My work thus assumes that an ‘optimal’ phenotype for a bacterium is one that allows the cell to maximize its nutrient uptake and growth rate and thus minimize the time it needs to double its genes.

The regulation of gene expression during the lifetime of a cell as well as the coexistence of such a huge diversity of bacterial species with various nutrient acquisition strategies clearly demonstrate that no one behavior is optimal for all environments at all times. Therefore, the focus of my work has been to uncover the trade-offs of particular traits that govern bacterial behavior in terms of their optimality for different environments. My general approach throughout this work has been to describe and quantify the trade-offs of a particular bacterial behavior by developing a simple, mechanistic model of that behavior and then incorporating it into a toy model of microbial metabolism, which allows me to assess how variations in that behavior affect growth rate.

My focus on mechanistic (rather than phenomenological) models allows me to hypothesize on the “causality of input-output relationships” [14]. That is, it allows me to hypothesize as to *why* a particular bacterial behavior is being observed in a particular environment and *how* that behavior affects the cell’s growth rate. Thus, my models allow me to develop an understanding of phenotypic trade-offs and to extrapolate from previous experimental research and observations to predict both what phenotypes I expect to find in a particular microbial environment as well as how a bacterium may regulate the expression of its genome to respond to changes in its environment.

I furthermore follow a philosophy of modeling that values the toy model—the model that attempts to achieve the greatest simplicity while still capturing the particular features of interest. I attempted to minimize the number of components and free parameters in my models so that I could gain intuition about which aspects of bacterial behavior were the crucial ones that governed the predicted growth rates.

My thesis work is comprised of two main parts: a study of the chemotaxis of *E. coli* to the sugar, maltose; and a study of the speciation of marine heterotrophic bacteria. In Chapter 2, I review the literature that has provided us with an understanding of *E. coli*'s chemotactic system. In Chapter 3, I present experimental results of chemotaxis assays of *E. coli* in dual gradients of maltose and the amino acid, aspartate, and develop a transport-and-sensing chemotaxis model that incorporates the transport kinetics of maltose to fit the results. This model uncovers a trade-off between the cell's ability to transport maltose at high affinities and its ability to sense maltose over a wider dynamic range. I then broaden my scope to consider trade-offs between various nutrient acquisition strategies in the marine microbial world. In Chapter 4, I differentiate marine heterotrophic bacteria by their abundances in nutrient-poor versus nutrient-rich environments. I develop a metabolic model that considers various nutrient transport models, which suggests that this differentiation of marine bacteria is governed by a rate-affinity trade-off: While one nutrient uptake strategy allows cells to attain high affinities to grow in nutrient-poor environments, an alternative allows cells to attain high growth rates to dominate in nutrient-rich environments. In Chapter 5, I instead differentiate marine heterotrophic bacteria by whether or not they can swim to predict how the energetic costs of motility govern the prevalence of motile bacteria in different environments. Finally, in Chapter 6, I present an outlook and suggest future work that will enable us to gain a better understanding of the ecology of marine bacteria.

Chapter 2

Chemotaxis in *E. coli*

A bacterium is an osmotroph, which consumes nutrients by allowing the nutrients to diffuse across the cell membrane. Because of the thermal motion of water molecules in the cell's microscale water world, diffusive transport is, in fact, faster than transport by flow. Therefore, in an environment with a homogeneous concentration of nutrients, “the fellow who just sits there quietly waiting for stuff to diffuse will collect just as much” [160]. However, many bacteria, including *E. coli*, can swim. Thus, they must have evolved motility because of the prevalence of *heterogeneities* in their environment. They out-swim diffusion to “find greener pastures” [160], environments ideal for survival and replication.

To swim, bacteria have evolved—in a marvelous feat of nature's engineering—ion-driven rotary motors, which whip around spaghetti-like appendages, called *flagella*, that propel the cell forward. Such a complex mechanism for swimming evolved because, since the motile cells are on the order of only one micron long, they live at low Reynolds number, where viscous forces completely dominate inertial forces. As Purcell explained,

If you are at very low Reynolds number, what you are doing at the moment is entirely determined by the forces that are exerted on you *at that moment*, and by nothing in the past.

It helps to imagine under what conditions a man would be swimming at, say, the same Reynolds number as his own sperm. Well, you put him in a swimming pool that is full of molasses, and then you forbid him to move any part of his body faster than 1 cm/min. Now imagine yourself in that

condition: you're under the swimming pool in molasses, and now you can only move like the hands of a clock. If under those ground rules you are able to move a few meters in a couple of weeks, you may qualify as a low Reynolds number swimmer. [160]

Not only does an *E. coli* cell qualify as a low Reynolds number swimmer, it excels at it; it can swim at about ten times its body length in one second. On the other hand, if we were to shrink to the size of an *E. coli* cell, the reciprocal motion of our swimming—the throwing back and forth of our arms and legs—would not propel us forward but instead move us away from and then back to the same position. Bacteria needed to evolve a continuous mechanism with more than one degree of freedom to swim at the microscale. They thus use their cylindrical flagella to “take advantage of the fact that a long cylinder has more resistance to movement perpendicular to its axis than parallel to its axis” [50, p.207]. The turning of the flagella in a periodic, corkscrew motion creates variable drag that allows net forward displacement.

To use these flagella to swim toward ‘greener pastures’, motile bacteria perform *chemotaxis*. For the remainder of this chapter, I review the previous work that helped develop an understanding and mechanistic model of chemotaxis in *E. coli*. This review is partly an expansion of what appeared in [152].

2.1 *E. coli*'s run-and-tumble chemotaxis strategy

To perform chemotaxis, bacterial cells bias their swimming direction in response to spatial gradients of various chemicals, such as nutrients, in their environment. A spatial gradient is indicative of a ‘greener pasture’ ahead because, since diffusion is faster than flow, “the distribution of chemicals is relatively smooth in the micro world, and the local distribution is well represented by a simple gradient” [50, p.222]. To follow this gradient, an *E. coli* cell uses its multiple (on average four) flagella: When all of them spin counter-clockwise (as viewed looking at the end of the flagellum toward the cell), the flagella bundle together and jointly create a corkscrew motion that propels the cell forward and allows it to swim in an approximately straight *run*. When the motor of any of the flagellum switches direction and starts spinning clockwise, the flagella unbundle, causing the cell to *tumble*

and randomly change direction. To achieve chemotaxis, the cell must bias this random walk, and it does so by modulating the rate at which it tumbles. When the cell measures that the concentration of a nutrient has increased, it probabilistically increases the length of its run so that it can travel further in its current, favorable direction [21, Ch.4].

Because this so-called *run-and-tumble* strategy of *E. coli* is probabilistic, a cell is simply more likely to swim up a gradient of nutrients; it may also swim down it. In the field of robotics, we are accustomed to equating randomness to the uncertainty that a robot must overcome to achieve its task. (See, for example, [199].) And, for both the roboticist and the ecologist, a random strategy is ideal only when the environment is unknown [206]. But an *E. coli* cell may swim away from a nutrient even after accurately measuring a gradient. One reason for this randomness is that, because of its tiny size, an *E. coli* cell cannot keep track of where it is going. It is bullied by water molecules that are constantly hitting the cell due to thermal energy. “For a swimming cell, the cumulative effect of this motion over a period of 1 second is displacement in a randomly chosen direction by about 1 μm and rotation about a randomly chosen axis by about 30 degrees. As a consequence, *E. coli* cannot swim in a straight line. After about 10 seconds, it drifts off course by more than 90 degrees, and thus forgets where it is going” [21, p.49].

2.2 Pioneering experiments that characterized *E. coli* chemotaxis

Just as *E. coli* has served as a model organism for the study of biology, its chemotaxis has served as a model biological system because of its relative simplicity and paradigmatic underlying molecular mechanism. (See, for example, how chemotaxis is described in [10] and [157].) There is thus a wealth of information that scientists have accumulated over the decades about *E. coli* chemotaxis.

The term *chemotaxis* was first used by Wilhelm Pfeffer in the 1880s to describe how bacteria seemingly ‘steered’ towards nutrients diffusing out of a capillary tube. However, quantitative experiments on bacterial chemotaxis really only began in earnest in the 1960s. At this time, Julius Adler, using the same capillary assay technique as Pfeffer, quantified the strength of the chemotactic response by culturing onto plates the contents that had

accumulated inside the capillary tube and then counting the colonies that grew onto the plates overnight [5]. Although this approach could not be very precise, it was simple and gave many insights into some of the most important features of chemotaxis. For one, Adler deduced that *E. coli* has dedicated sensors for chemotaxis, *chemoreceptors*, since the cells would respond to various *chemoeffectors* that they could not metabolize. Furthermore, Adler and others recognized the robustness of the chemotactic system, noting that, over a wide range of background chemical concentrations, cells remained very sensitive to small changes in relative concentration [129]. They were thus the first to suggest that *E. coli* chemotaxis may follow Weber's law.

Weber's law is more of an observation than a true law. In the late 1800s, Ernst Weber's work in experimental psychology led him to deduce that our perception to stimuli, such as sound or light, does not depend on the absolute amplitudes of the stimuli but instead on their relative differences from the background or initial amplitudes. Thus, light from a single bulb can seem blindingly bright when suddenly turned on in a dark room but be imperceptible on a sunny day. The founder of psychophysics, Gustav Fechner, named and further developed the law by suggesting that the magnitude of a response to a stimulus is proportional to the logarithm of the stimulus's intensity [77]. Using this *logarithmic sensing*, "organisms adapt to a background stimulus, and are then able to respond to a stronger stimulus, until the background intensity is so high that the receptor becomes saturated" [129].

Another contemporaneous experiment (from the early 1970s) also suggested that bacterial chemotaxis exhibits adaptation by ingeniously studying how *Salmonella* cells respond to temporal gradients instead of spatial ones [120]. Macnab and Koshland developed a "temporal gradient apparatus" that first presented a suspension of cells with an initial uniform concentration and then used rapid mixing to create a step input that quickly presented the cells with a different final uniform concentration. Using a dark-field microscope, they took consecutive photographs of the cells to generate tracks of the cells over time. They concluded that the cells possessed "memory" that allowed them to modify their tumbling rates by comparing past and present measurements of chemical concentrations. Immediately after a positive stimulus, tumbling rates decreased; after a negative stimulus, tumbling rates increased. But, ultimately, the cells adapted to the new uniform concentration and tumbled at the same initial rate.

A couple of years later, Howard Berg and Douglas Brown also confirmed that *E. coli* chemotaxis is based on temporal comparisons with adaptation, using a three-dimensional tracking microscope [24]. This microscope used a servo system to reposition a chamber containing free-swimming cells so that a single cell remains fixed in the reference frame. By keeping track of the three-dimensional coordinates of the chamber over the duration of the experiment, they could recover long tracks of this single cell over time. Computer analysis of these tracks provided very accurate characterizations of the biased random walk that chemotactic cells perform in various spatial gradients. Berg and Brown showed that “*E. coli* extends runs that are favorable (that carry cells up the gradient of an attractant) but fails to shorten runs that are not (that carry cells down such a gradient)” [21]. They quantified this behavior in terms of the cell’s tumbling rate and reorientation bias. They concluded that the run time intervals of a cell are Poisson; that is, the tumble rate is exponentially-distributed so that the probability per unit time that a tumble occurs is constant. They also concluded that reorientation angles were independent of spatial gradients and slightly skewed from a uniform distribution to favor smaller deviations in direction. Berg and Brown were also able to use their 3D tracking microscope to better understand the chemotactic response due to temporal gradients by more accurately creating temporal gradients in isotropic environments. This was achieved by using an enzyme that converts an unperceived chemical into a chemical attractant [37].

Berg, with the help of other colleagues, was also able to quantify the tumbling rate of *E. coli* cells using cells tethered to glass by a single flagellum [28, 29, 176]. In this way, the timing of reversals of the cell’s rotary motor could be accurately determined by observing the cell’s rotation. Pulses of chemicals were delivered to the tethered cells using iontophoretic pipettes. The ejection of a chemical from a micropipette directly in front of the cell created a sudden, saturating but then quickly dissipating stimulus that mimics an impulse. In an article from 2004 [60], de Gennes assumed a particular functional form of the tumbling rate, in which a chemotactic cell responds linearly to concentrations of a chemoeffector. When the system is linear, the impulse response fully describes the response of the cells to any other sequence of stimuli. The form de Gennes chose for the tumbling rate is as given:

$$\lambda(t) = \lambda_0 \left[1 - \int_0^t c(\tau) K(t - \tau) d\tau \right],$$

where $K(t)$ is the impulse response and $c(t)$ is the chemical concentration experienced at time t . This impulse response is thus a bias in the probability of counterclockwise rotation of the rotary motor. This impulse response is “biphasic, increasing the [counterclockwise] bias for the first second and then decreasing it for the next three seconds, with the areas of the two lobes being equal.” Thus, the cell’s ‘memory’ stores information about the concentration of chemicals in the environment over the last four seconds, which “made sense” as it is “a time below the limit set by rotational Brownian motion (~ 10 s)” [23, p. 354].¹

More recent work has suggested that *E. coli*’s response is not only logarithmic with perfect adaptation but, in fact, exhibits fold-change detection (FCD). A FCD is a “response whose entire shape, including amplitude and duration, depends only on fold changes in input and not on absolute values” [179]. Thus, inputs that are proportional to each other create identical responses. In [116], Lazova and others demonstrated that chemotaxis exhibits FCD at both a molecular-level and by observing the population-level dynamics of cells in variously scaled linear gradients of a single attractant. In [179], Shoval and others suggested that bacteria’s chemotaxis evolved to exhibit FCD to more efficiently store and use “information about the position of the [nutrient] source,” which should be “encoded in the shape of the field [and] not in its amplitude.”

2.3 The signal transduction pathway

Previous work has also provided a well-developed understanding of the underlying molecular pathways that drive chemotaxis [53]. The signal transduction pathway for chemotaxis allows a message, in the form of a phosphate group, to be transmitted from the chemoreceptor (the sensor) to the flagellar motor (the actuator). Similar to a number of so-called two-component signaling pathways [191], the pathway involves the transfer of the phosphate group from a histidine protein kinase to an aspartate kinase. (A kinase is an enzyme that catalyzes the transfer of a phosphate group from ATP to a specific molecule.)

¹We can gain better insight by comparing this impulse response to the idealized, simplified form used in [130], where $\lambda(t) = \lambda_0 \left[1 - \int_0^t \log c(\tau) \tilde{K}(t - \tau) d\tau \right]$, with $\tilde{K}(t) = \lim_{\epsilon \rightarrow 0} \frac{1}{M\epsilon} [\delta(t) - \delta(t - \epsilon)]$.

The histidine protein kinase, called CheA, is coupled to the chemoreceptor via the protein CheW. When a chemoreceptor is not bound to an attractant, the chemoreceptor is more likely to be ‘active’ and, via mediation by CheW, activates CheA autophosphorylation. When an attractant binds to the chemoreceptor, the receptor undergoes a conformational modification that deactivates CheA.

The aspartate kinase, called CheY, diffuses freely throughout the cytoplasm and is a so-called response regulator that interacts with the flagellar motor to mediate the cell’s chemotactic response. When CheY binds to phosphorylated CheA, the CheA transfers its phosphate to CheY, forming a CheY-P complex. It is this complex that can bind to the motor to increase the probability of clockwise rotation and thus increase the probability that the cell will tumble. On the other hand, the phosphatase, CheZ, accelerates the dephosphorylation of CheY-P. Therefore, the phosphorylation of CheY by CheA to form CheY-P is offset by the dephosphorylation of CheY-P by CheZ. This allows the level of CheY-P to reach a steady state that depends on the ratio of the rates of the two processes. As CheZ is conserved in the cell, this steady-state value reflects the activity of CheA and determines the cell’s probability of tumbling.

So that the system can adapt to various background concentrations of chemicals, there is also another set of proteins that add or remove methyl groups to four glutamate residues in the chemoreceptor’s cytoplasmic domain. The number of glutamate residues that are methylated acts as the cell’s memory. The methyltransferase, CheR, transfers methyl groups to the residues using S-adenosyl-methionine as the methyl donor. The conformational modifications of the receptor due to methylation increase the rate of CheA phosphorylation, thus increasing activity. The system then adapts by using a negative integral feedback loop [218] in which the rate of CheA phosphorylation increases the rate of demethylation. This is achieved via another response regulator protein, the methylesterase CheB, which, like CheY, is also an aspartate kinase that can be phosphorylated by CheA. Phosphorylated CheB, CheB-P, demethylates the receptor, removing a methyl group via hydrolysis of carboxylic ester bonds. This demethylation decreases the rate of CheA autophosphorylation, thus decreasing activity. Demethylation and methylation are relatively slow in comparison to changes in receptor occupancy. Therefore, when the concentration of a chemoeffector changes steadily throughout time, the corresponding change in the methylation level lags behind. During this lag—that is, before

the system has adapted, the cell's tumbling rate either increases (in the case of a decrease in measured attractant concentration) or decreases (in the case of an increase).

To summarize, I quote Berg:

The receptor complex ... is a remarkable system that acts as a comparator. The output of this comparator (the kinase activity) depends on the difference between the time-average occupancies of the receptor binding sites and the level of methylation. The kinase is activated if the methylation level is relatively high and inactivated if it is relatively low. Changes in the occupancies of the receptor binding sites are very fast, and reflect the present concentrations of ligands. Changes in the levels of receptor methylation, on the other hand, are relatively slow, and reflect the past concentrations of ligands. Thus, the cell is able to make temporal comparisons. [21, p. 84]

2.3.1 Signal amplification

One important aspect of the *E. coli*'s chemotactic response could not be explained by this molecular-level understanding of the signal transduction pathway and puzzled experimentalists for decades. They could not understand why a step increase in attractant concentration of less than 2% from an already high initial concentration could transiently increase the counterclockwise bias of the flagellar motor by 23%. Berg called this puzzle the "gain paradox" [21, Ch. 11], [22].

In [186], Sourjik and Berg showed that the gain was due to a puzzling "front end" amplification of receptor occupancy to CheY-P concentration. They demonstrated that "the fractional change in kinase activity is some 35 times larger than the fractional change in receptor occupancy" [21, p. 97]. To obtain these results, they used a technique that allowed them to directly measure changes in CheY-P concentrations in a single living cell. They used fluorescence resonance energy transfer (FRET) to monitor the amount of CheY-P binding to its phosphatase, CheZ. For the experiments, the cell was genetically modified so that, instead of using CheY and CheZ, the cell used the complex CheZ-CFP (CheZ bound to the cyan fluorescent protein) and CheY-YFP (CheY bound to the yellow fluorescent protein). During dephosphorylation of the CheY-P complex by the CheZ

complex, the two complexes bind together, bringing the CFP and YFP in close proximity to each other. When the fluorescent proteins are this close together and the CFP is excited by light, it transfers some of its energy to the YFP, and the amount of observed YFP fluorescence increases. Since at steady state the rate of phosphorylation of CheY equals the rate of dephosphorylation of CheY-P, this measurement of the dephosphorylation rate can directly provide a measurement of the kinase activity [23, p. 355-6].

Dennis Bray explained why this front end amplification was so puzzling: “Simply put, the issue is as follows. The slowest reaction in the cascade of reactions is the generation of phosphoric groups by CheA.... Subsequent reactions are much faster, so CheY-P should rise and fall more or less in parallel with changes in receptor occupancy” [35]. So what is causing this amplification? It turns out that it is interactions between the chemoreceptors. The chemoreceptors are not evenly distributed throughout the surface of the cell but instead usually cluster near one of the cell’s poles, as was discovered using immunogold labeling of the receptors [121]. It had already been deduced that, since a cell did not preferentially orient this cluster with respect to its swimming direction, the cluster did not act as a sensing ‘nose’. Thus, Bray and colleagues correctly deduced that receptors clustered so that an activated receptor could activate other neighboring receptors [36]. This receptor cooperativity would heighten chemotactic sensitivity and explain the observed signal amplification.

2.3.2 A “coarse-grained model” of the chemotactic pathway

To simulate how an *E. coli* cell would swim in an environment containing gradients of the single chemoeffector aspartate, Yuhai Tu and colleagues developed a “coarse-grained model” for agent-based simulations, SPECS, that takes into account the molecular pathway of chemotaxis as well as receptor cooperativity [202, 203].

2.3.2.1 An Ising model

To model receptor cooperativity, Duke and Bray, in [48], used an Ising model. The Ising model for ferromagnetism uses statistical mechanics to describe the probability that a lattice of magnetic dipole moments, each with either spin up ($\sigma = +1$) or spin down

($\sigma = -1$), takes a specific configuration of spins. The energy of a particular configuration is given by the Hamiltonian function,

$$H(\sigma) = - \sum_{\langle i,j \rangle} J_{ij} \sigma_i \sigma_j - \mu \sum_j h_j \sigma_j,$$

in which the first term accounts for the cooperativity between neighboring dipoles and the second term accounts for the effects of an external magnetic field. The probability of a given configuration is given by the Boltzmann distribution,

$$P(\sigma) = \frac{e^{-\beta H(\sigma)}}{Z},$$

where Z is the normalization partition function.

To describe receptor cooperativity for chemotaxis using the Ising model, they let σ represent whether a receptor is active or inactive and replaced the external magnetic field h_j with the free energy difference between the active state and the inactive state, which is a function of both the methylation level, m , and the ligand concentration, $[L]$:

$$\begin{aligned} h_j &= \Delta f_j(m_j, [L]) = -f_m(m_j) - f_L([L]) \\ &= -f_m(m_j) - \ln \frac{1 + [L]/K_i}{1 + [L]/K_a}, \end{aligned}$$

where K_i and K_a are the dissociation constants between the ligand and the inactive and active receptors, respectively. The Boltzmann distribution is then used to calculate the probability that a lattice of chemoreceptors has a given configuration of activity, and, by computing the expected activity (reflecting the expected concentration of CheY-P in the cell), they could model the probability that the cell would tumble at a particular instant in time.

2.3.2.2 A MWC model

However, the Ising model is difficult to solve analytically, as we must resort to either using Monte Carlo simulations or solving sets of equations for a mean field theory approximation. Therefore, SPECS relies on a simplification of the Ising model that assumes that the

cooperativity interaction strength, J , is infinite. That is, it assumes that all chemoreceptors in a cluster are either all active or all inactive. This gives a nice analytical expression for the expected activity level:

$$\begin{aligned}\langle a \rangle &= (1 + \exp(-N\Delta f))^{-1} \\ &= \frac{1}{1 + \exp[Nf_m(m)] \frac{(1+[L]/K_i)^N}{(1+[L]/K_a)^N}},\end{aligned}$$

where N is the number of receptors in the cluster. This simplified cooperativity model mirrors the “all-or-none” model first proposed by Monod, Wyman, and Changeux (MWC) to describe the allosteric interactions of the hemoglobin subunits when binding oxygen molecules [140].

The MWC model allowed Tu and colleagues to easily develop agent-based simulations for chemotaxis [90]. The simulations take the ligand binding and activity level to be at a quasi-equilibrium and use a simple model for the methylation dynamics that ignores possible saturation of the glutamic residues:

$$\frac{dm}{dt} = k_R(1 - \langle a \rangle) - k_B \langle a \rangle,$$

where k_R and k_B are respectively the rates of methylation and demethylation. (Notice that because, in this model, the receptors are always jointly active or inactive, the methylation level, m , is constant across all receptors.)

They account for the sensitivity of the motor to binding by CheY-P using a Hill coefficient of 10; in their model, the probability of tumbling is a function of the average activity level:

$$P(\langle a \rangle) \propto \langle a \rangle^{10}.$$

This model correctly predicts many of the features of chemotaxis that have been demonstrated in experiments. Not only does it reflect the signal amplification and sensitivity of the chemotaxis system, it also predicts fold-change detection, as derived in [203].

2.4 Multiple signal integration

However, while the MWC model well predicts the response of *E. coli* in environments with *one* chemoeffector, it does not consider the response of cells in environments with *multiple* chemoeffectors.

2.4.1 A heterogeneous MWC model

In [146], Neumann and others reported the results of FRET experiments in which (i) adapted cells were simultaneously stimulated by two different chemoeffectors and (ii) cells adapted to a high concentration of one chemoeffector were stimulated by a second chemoeffector. They reported that “[c]hemotactic response to mixtures of effectors showed that simultaneous stimulation by ligands of different receptors is *additive*, and that adaptation to ligands of one receptor does not interfere with signaling by other receptors.” Furthermore, “computational analysis confirmed that this summation could be accounted for by *additivity of changes in free energy*”. They thus concluded that, since the free energy is dependent on the number of chemoreceptors of a particular type in a cell, signal integration uses “majority voting” to determine where the cell should swim in environments with multiple chemoeffector gradients.

Independently in both [126] and [97], the MWC model was extended to create a heterogeneous MWC (HMWC) model that accounts for this additive response to multiple stimuli. The models predict that the expected activity has the following form in the case of two chemoeffectors:

$$\langle a \rangle = (1 + \exp(f_m(m) + f_L([L]_1, [L]_2)))^{-1},$$

with

$$f_L([L]_1, [L]_2) = N_1 f_{L_1}([L]_1) + N_2 f_{L_2}([L]_2),$$

where N_1 (N_2) is the number of receptors in the cell that bind to attractant 1 (2), with concentration $[L]_1$ ($[L]_2$), and where f_{L_1} and f_{L_2} have the same form as f_L as given above in Section 2.3.2.1.

From this HMWC model, Hu and Tu derived an advection-diffusion equation for the population-level dynamics of chemotactic cells in one-dimension [84]. They used this diffusion equation to solve for the stationary distribution of the cells and predicted that the steady-state cell distribution has the following form:

$$p_s(x) \propto e^{[\eta(\bar{a}(x))f_L([L]_1(x),[L]_2(x))]},$$

where $\bar{a}(x)$ is the expected activity level at location x and $\eta(\bar{a}(x))$ “represents the effective sensitivity of the bacterial population to the environment.”

2.4.2 Signal integration within the same receptor

However, in [75], Hart and colleagues argued that, because various receptor types have different methylation dynamics, the HMWC model may only correctly predict a response to signals that both affect the *same* receptor type. Intriguingly, that is indeed the case for *E. coli*’s sensing of the sugar maltose and the amino acid aspartate, which are both sensed by binding to distinct sites on the so-called aspartate receptor, Tar.

Because this additive signal integration of maltose and aspartate suggested that linear profiles of the two chemoattractants could be used to accumulate *E. coli* cells in the center of environments (Figure A.2), I conducted chemotaxis assays of *E. coli* in opposing gradients of maltose and aspartate. These experiments are presented in the following chapter and show, however, that central accumulation does not occur. Thus, I also develop an extension of the *E. coli* chemotaxis model that is able to fit my data. This new model relies on the HMWC model to describe signal integration and generalizes the chemotaxis model presented above to consider cases in which the periplasmic concentration of a chemoattractant is unequal to the extracellular concentration of it due to variable transport kinetics.

Chapter 3

The Effects of Transport on *E. coli*'s Chemotaxis to Maltose

There is a complex interplay between the sensing of a metabolizable chemoattractant and its uptake by the cell. In this work, I suggest that, as cells regulate the rates at which they consume a nutrient, they may modify their ability to sense it. Therefore, a mechanistic understanding of this interplay is needed to predict chemotactic response in ecologically relevant nutrient landscapes. This understanding will allow greater insight into the ecological impact of chemotaxis as well as into how particular chemotactic bacterial species may be used as “microbots”.

Here I show that existing chemotaxis models (as described in Chapter 2) are unable to predict *E. coli*'s response to maltose. While *E. coli* can respond to concentrations of aspartate over a few orders of magnitude [116], its response to sugars, such as maltose, has a much smaller dynamic range over just one order of magnitude; *E. coli* can transport maltose without saturation over a much larger range than it can sense [76]. This is puzzling, particularly because maltose is sensed by the same receptor as aspartate, the aspartate receptor, Tar. Maltose and aspartate can be sensed independently by Tar because they can bind simultaneously to distinct sites on the receptor [142]. Thus, a Tar molecule effectively acts as two distinct receptors sharing the same methylation state. While *E. coli* uses Tar to sense maltose oligosaccharides, all other known *E. coli* sugar chemoattractants instead bind to the minor receptor Trg [146]. Tar is considered a major receptor because it is present in the cell periplasm at orders of several thousand molecules

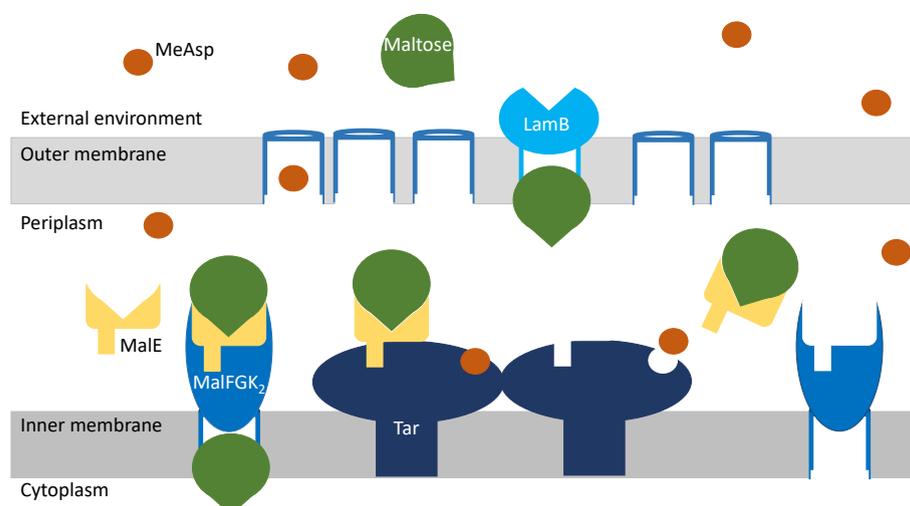


FIGURE 3.1: **Schematic of the transport and sensing of maltose and MeAsp.** MeAsp (α -methyl-DL-aspartate, an analog of the amino acid aspartate) enters and exits the periplasm via general diffusive porins, binds directly to the aspartate receptor, Tar, and is not transported into the cytoplasm. Maltose, in contrast, is larger and enters the periplasm via facilitated diffusion by the maltoporin, LamB. The maltose-binding protein (MalE) both (i) allows Tar to sense maltose via an independent binding site, and (ii) binds to the maltose ABC transporter, MalFGK₂, to allow maltose to be transported into the cytoplasm.

per cell, whereas the minor receptors are present at orders of only a few hundred molecules per cell [185]. A more abundant chemoreceptor should saturate at a higher concentration of attractant, thus allowing the cell to sense over a larger dynamic range. Yet the cell's dynamic range for maltose sensing is just as narrow as its range for other sugars [146].

Neumann and others suggested that the small dynamic range of maltose sensing is due to indirect binding [146]. While aspartate binds directly to Tar, sugars bind indirectly to receptors via binding proteins; maltose binds to Tar only when in complex with the maltose binding protein, MalE [124]. They thus proposed an indirect-binding chemotaxis model to account for the effects of the binding proteins, but this model relied on the assumption that the concentration of free sugars in the periplasm is equal to that of the external environment. This assumption is called into question by experimental results elucidating the variable rates of maltose periplasmic and cytoplasmic uptake, which suggested that, at the micromolar concentrations of maltose that *E. coli* can sense, the concentration of free maltose within the cell periplasm may be nearly zero [80, 106, 148]. I will argue that it is for this reason that this indirect-binding model is unable to fit my experimental data of the response of *E. coli* to maltose.

Transport into the periplasm, and thus the periplasmic concentration of a substrate available for a receptor to sense, depends strongly on the size of the substrate. In steady-state conditions, the periplasmic concentration of a substrate is equal to its external concentration when the substrate's maximal rate of diffusion into the periplasm exceeds the rate of its subsequent uptake into the cytoplasm. However, this is not necessarily the case for saccharides. They are relatively large and thus diffuse through general porins at rates that are orders of magnitude lower than that of smaller compounds, such as amino acids [149]. Because slow rates of diffusion limit uptake and thus growth rates [106], bacteria have evolved a suite of specific porins that facilitate the diffusion of specific sugars into the periplasm [168]. Thus, an *E. coli* cell can, in fact, regulate the abundance of a particular sugar in its periplasm by regulating the expression of the corresponding porin. This in turn allows the cell to regulate its chemotactic sensitivity to a sugar by altering the amount of that sugar available for a receptor to sense [87]. Therefore, to predict the chemotactic response of an *E. coli* cell to maltose, I must explicitly model how the transport of maltose into the periplasm (via the specific porin, the maltoporin LamB [103]) and out of the periplasm (via the ABC maltose transporter, MalFGK₂ [148]) affects the concentration of maltose within the periplasm (Figure 3.1). This work is the first to propose a chemotaxis model that accounts for these transport dynamics.

The development of new chemotaxis models exposes a further obstacle in quantifying and predicting the chemotactic response of various bacteria: we lack simple, reliable, and widely applicable protocols for system identification and model validation of chemotaxis systems. A common technique for fitting molecular-level chemotaxis models uses FRET reporters, which requires fluorescent microscopy on genetically modified cells [187]. On the other hand, microfluidic technology now allows scientists to conduct simpler experiments to accurately quantify the population-level chemotactic responses of wild-type, non-fluorescent cells [6, 94, 217]. Population-level assays have traditionally been described using phenomenological mathematical models, in particular the Keller-Segel model and its variants [82, 128, 173]. However, these models do not consider molecular-level mechanisms and so cannot be used to predict how transcriptional regulation may modify the chemotactic response as a result of changes in the environment. While using agent-based simulators of molecular-level models, such as a modified version of SPECS, would address these concerns, fitting an agent-based simulator to experimental data is computationally expensive and intractable because it requires continually re-running the simulator over a

large parameter space.

In this work, I address both the current inability to model bacterial chemotaxis to certain saccharides and the difficulty in fitting molecular-level chemotaxis models to experimental data on chemotactic migration. I derive a transport-and-sensing chemotaxis model that incorporates the impact of variable porin and transporter expression on the chemotactic response. This model is widely applicable for the chemotaxis of gram-negative bacteria to a variety of sugars and suggests that the puzzlingly narrow dynamic range of maltose sensing is the result of a trade-off between achieving a larger dynamic range of sensing versus a larger range of uptake.

To fit the transport-and-sensing chemotaxis model, I developed a system identification protocol that bridges population-level and molecular-level model fitting techniques. This protocol relies on an intuitive analytical approximation of the steady-state distribution of cells previously derived from the SPECS model [84] and allows us to use population-level experimental data to fit molecular-level parameters. I conducted population-level chemotaxis assays of wild-type *E. coli* cells in a microfluidic device that creates linear gradients of maltose. To better constrain my model fits, I also conducted assays of *E. coli* in various opposing gradients of maltose and α -methyl-DL-aspartate (MeAsp), a non-metabolizable analog of aspartate. I integrated a previous model for MeAsp chemotaxis with my new model for maltose chemotaxis using the Heterogeneous MWC (HMWC) model, which is an extension of the MWC model that describes the integration of multiple chemotactic signals that share the same methylation dynamics [75, 97, 113, 126]. Thus, because both maltose and MeAsp are sensed by Tar, I was able to better constrain my fit for maltose chemotaxis using previous fits of the response to MeAsp. I conducted a sensitivity analysis to test the robustness of the fits and also used previous literature to validate the parameter fits. The robustness and goodness of my fits suggest that my system identification protocol may be used to identify and fit chemotaxis models to chemotaxis systems in other species of bacteria.

3.1 A transport-and-sensing chemotaxis model

Here I outline a transport-and-sensing chemotaxis model that predicts how the transport kinetics of maltose affect the cell's chemotactic response to it. Figure 3.1 shows a cartoon

of the molecular basis of maltose transport and sensing. A detailed derivation of the model can be found in Appendix A.

During chemotaxis, the instantaneous probability that a cell tumbles is a function of the receptor activity level, and the MWC model assumes that, due to receptor cooperativity, receptors form clusters within which receptors are either simultaneously all active or all inactive. The average activity of the clusters, $\langle a \rangle$, is thus the probability that a receptor cluster is active and is a function of the free-energy difference f between the active and inactive states. The HMWC model [97, 126] assumes that the free-energy differences of the different receptor types are additive, so that

$$\langle a \rangle = (1 + e^{f_m(m) + \sum_r n_r f_r})^{-1}, \quad (3.1)$$

where $f_m(m)$ is the free-energy difference between an active and inactive receptor cluster in the absence of chemoattractants and depends on the average methylation level of the entire cluster, m ; n_r is the number of type- r receptors in a cluster; and f_r is the free energy difference between an active and inactive bound receptor of receptor type r . In the direct-binding model that describes aspartate chemotaxis [146, 202],

$$f_r = \log \frac{1 + [L]_r / K_{I,r}}{1 + [L]_r / K_{A,r}}, \quad (3.2)$$

where $[L]_r$ is the concentration of the chemoattractant ligand in the extracellular environment and K_I (K_A) are the dissociation constants between the chemoattractant and the inactive (respectively, active) receptor.

In general, the free energy difference has the following form:

$$f_r = \log \frac{1 + C_I}{1 + C_A}, \quad (3.3)$$

where C_I (C_A) is the ratio of the probabilities of an inactive (respectively, active) receptor being bound versus unbound. Because maltose binds to the aspartate receptor, Tar, only when in complex with the binding protein, MalE, this formulation applies for maltose chemotaxis when C reflects the probability that a maltose-MalE complex is bound to a Tar receptor. Therefore, to derive a more realistic chemotaxis model for maltose sensing, I derive new expressions for the concentrations of active and inactive Tar bound to the maltose-MalE complex. I incorporate not only indirect binding [146] but also the fact

that the cell may modify via its transport kinetics the concentration of maltose within its periplasm and thus the amount of maltose available to bind to MalE and Tar.

To solve for the periplasmic concentrations, I assume a quasi-steady-state equilibrium. At equilibrium, the rate of maltose transport into the periplasm via the maltoporin LamB, v_p , is equal to the rate of its transport out of the periplasm and into the cytoplasm via the ABC transporter MalFGK₂, v_c . I assume that the cytoplasmic uptake rate is proportional to the concentration of MalFGK₂ associated with the sugar-binding protein complex. I simplify a previous model of ABC transport [32] by relying on the fact that, for the transport of maltose in *E. coli*, the abundance of binding proteins greatly exceeds that of the transporters [31, 156]. Therefore, I approximate that the association and dissociation of MalE and maltose proceed without any competing reactions, so that the total amount of free maltose-MalE complex in the periplasm, [L:BP], can be described by the dissociation constant between maltose and MalE, K_{BP} . I furthermore assume that the MalFGK₂ transporter only binds to the maltose-MalE complex, so that I can similarly describe the equilibrium concentration of transporter bound to maltose-MalE complex by the dissociation constant, K_c , between the sugar-binding protein complex (L:BP) and the transporter. Therefore, I approximate the quasi steady-state cytoplasmic uptake rate of maltose as

$$v_c \approx V_c \frac{[\text{L:BP}]}{K_c + [\text{L:BP}]} \quad (3.4)$$

$$\approx V_c \frac{[\text{L}]_p}{\frac{K_c K_{BP}}{[\text{BP}]_{total}} + [\text{L}]_p}, \quad (3.5)$$

where V_c is the maximal cytoplasmic uptake rate, which is a function of the number of expressed transporters; $[\text{L}]_p$ is the concentration of free maltose within the periplasm; and $[\text{BP}]_{total}$ is the total concentration of binding protein in the periplasm. Notice that increasing the total concentration of binding protein in the periplasm decreases the effective half-saturation concentration of cytoplasmic transport, $K_{eff} \approx \frac{K_c K_{BP}}{[\text{BP}]_{total}}$, and thus increases the specific affinity [40] of maltose transport. (See also Chapter 4.) Therefore, this model suggests that the cell expresses levels of binding protein greatly in excess of transporter abundance to achieve high-affinity transport.

While transport into the cytoplasm is active and thus can occur against concentration gradients, specific porins transport sugars into the periplasm via facilitated diffusion,

which can be described by the following Michaelis-Menten equation [33]:

$$v_p = V_p \frac{[L]_{ext} - [L]_p}{K_p + [L]_{ext} + [L]_p}, \quad (3.6)$$

where $[L]_{ext}$ is the extracellular concentration of the ligand; K_p is the half-saturation constant of the specific porin; and V_p is the maximal rate of uptake, which is a function of the number of expressed porins. At steady-state, the periplasmic transport rate must equal the cytoplasmic uptake rate, so that

$$V_p \frac{[L]_{ext} - [L]_p}{K_p + [L]_{ext} + [L]_p} \approx V_c \frac{[L]_p}{\frac{K_c K_{BP}}{[BP]_{total}} + [L]_p}. \quad (3.7)$$

Previous models of chemotaxis have not considered the interplay of the periplasmic and cytoplasmic uptake rates. The indirect-binding chemotaxis model [146] avoided consideration of transport kinetics by assuming that diffusive rates are fast, $V_p \gg V_c$, so that the cell's periplasm saturates with free ligand, $[L]_p \sim [L]_{ext}$. However, I will show that this model cannot fit my data. Furthermore, previous research has suggested that, in the micromolar regime of the experiments, $[L]_p \ll [L]_{ext}$ [58, 80, 148]. Thus, I must obtain $[L]_p$ from Equation 3.7. Although I could directly use Equation 3.7 to solve for $[L]_p$ as a quadratic function of $[L]_{ext}$, I instead limit the number of free parameters in the model by making the simplification that $[L]_p \ll [L]_{ext}$, which yields:

$$[L]_p = \frac{\left(\frac{K_c K_{BP}}{[BP]_{total}}\right) V_p [L]_{ext}}{K_p V_c + (V_c - V_p) [L]_{ext}}. \quad (3.8)$$

Solutions exist for all choices of $[L]_{ext}$ when $V_p \leq V_c$. Therefore, the fact that periplasmic concentrations are much smaller than external concentrations implies that the cell's total uptake rate from the environment into the cytoplasm is diffusion- or porin-limited rather than ABC transporter-limited.

Chemotactic sensing of the periplasmic concentration, $[L]_p$, proceeds via indirect binding, and my transport-and-sensing model makes the same simplifying assumptions as the previous indirect-binding model with respect to receptor sensing—that (i) the receptor only binds to the binding protein when it is in complex with the ligand, and that (ii) receptor binding does not affect binding of the ligand to its binding protein [146]. Using these assumptions, the two ratios of the probabilities of bound versus unbound receptor

are

$$C_{I,A} = \frac{[\text{R:BP:L}]_{I,A}}{[\text{R}]_{total} - [\text{R:BP:L}]_{I,A}}, \quad (3.9)$$

where $[\text{R}]_{total}$ is the total concentration of receptors in the periplasm; and $[\text{R:BP:L}]_{I,A}$ is the total concentration of ligand-binding protein complex bound to an inactive (or active) receptor. This quantity is defined by the dissociation constants, $K_{I,A}$, between the inactive (or active) receptor, R, and the ligand-binding protein complex, L:BP. That is,

$$[\text{R:BP:L}]_{I,A} = \frac{[\text{R}]_{free}[\text{L:BP}]_{free}}{K_{I,A}}. \quad (3.10)$$

Because I do not assume that the ligand-binding protein complex is in much greater excess than the receptors, I make the simplifying assumption that sensing via Tar does not hinder transport by competing for maltose-MalE complex so that I can use the following quadratic equation to solve for this concentration:

$$[\text{R:BP:L}]_{I,A} = \frac{([\text{R}]_{total} - [\text{R:BP:L}]_{I,A})([\text{L:BP}] - [\text{R:BP:L}]_{I,A})}{K_{I,A}}. \quad (3.11)$$

To solve for $[\text{R:BP:L}]_{I,A}$, I use Equation 3.8 to write the total concentration of maltose-MalE complex as a function of the external maltose concentration:

$$[\text{L:BP}] = \frac{[\text{BP}]_{total}[\text{L}]_p}{K_{BP} + [\text{L}]_p} = \beta \frac{[\text{L}]_{ext}}{\alpha + [\text{L}]_{ext}}, \quad (3.12)$$

where

$$\beta \equiv \frac{[\text{BP}]_{total} K_c V_p}{(V_c - V_p) [\text{BP}]_{total} + K_c V_p} \quad \text{and} \quad (3.13)$$

$$\alpha \equiv \frac{[\text{BP}]_{total} K_p V_c}{(V_c - V_p) [\text{BP}]_{total} + K_c V_p}. \quad (3.14)$$

I now have a model for a receptor's activity level (Equations 3.3, 3.9, 3.11 - 3.14) that incorporates the indirect binding of maltose as well as the effects of variable periplasmic and cytoplasmic uptake rates on periplasmic concentrations.

I can include an additional parameter in the model to account for likely steric effects between nearby binding proteins. Because MalE is relatively large, the binding of one molecule to Tar likely inhibits neighboring Tar from binding to a second maltose-MalE

complex [142]. However, while I show in Appendix A.3.3 that this additional parameter improves the fit for higher concentrations of maltose, I leave out this saturation term for the remainder of my analysis and model fitting to limit the number of free parameters in the model.

3.2 Modeling steady-state population distributions

As the purpose of my transport-and-sensing model is to predict the population-level chemotactic response of *E. coli* to various chemoattractants, I developed a protocol that allows us to directly fit my molecular-level model to data from population-level chemotaxis assays. This protocol uses data of steady-state cell distributions from chemotaxis assays on wild-type cells. While simulators like SPECS are capable of predicting steady-state distributions, running an agent-based simulator for each choice of parameters is intractable. Thus, I used analytical approximations of the predicted distributions [84], which assume the validity of the HMWC model, to directly map equations of free-energy differences (Equations 3.2 and 3.3) to a predicted steady-state distribution. Under this analytical approximation, the steady-state distribution of cells in an environment with both MeAsp and maltose is (Appendix A.3):

$$p(x) \propto e^{\eta(x)[f_{MeAsp}(x)+f_{Mal}(x)]}, \quad (3.15)$$

where x is the position along the channel; η “represents the effective sensitivity of the bacterial population to the environment” [84];

$$f_{MeAsp}(x) = \log \left[\frac{1 + \frac{[MeAsp](x)}{K_{I,MeAsp}}}{1 + \frac{[MeAsp](x)}{K_{A,MeAsp}}} \right]; \quad (3.16)$$

and

$$f_{Mal}(x) = \log \left[\frac{1 + C_I(x)}{1 + C_A(x)} \right]. \quad (3.17)$$

I can use equations for C_I and C_A so that they reflect the direct-binding, indirect-binding (Appendix A.3), or transport-and-sensing models (Equation 3.9). However, I am unable to predict the form of $\eta(x)$ because it has no simple molecular-level definition. In its original formulation, $\eta(x)$ varies both between experiments and within a single experiment

and is a nonlinear function of the expected activity level, which reflects adaptation via methylation [84]. However, if I allow η to be a free parameter, I would be overfitting the data. Thus, I made the approximation $\eta(x) = \eta_0$ over all experiments and validated this approximation using SPECS simulations (Figure A.7).

The transport-and-sensing model expands our understanding of chemotaxis and introduces four new parameters to the indirect-binding model [94], which reflect the Michaelis-Menten kinetics of the transport of maltose into (K_p, V_p) and out of (K_c, V_c) the periplasm. However, as seen in Equations 3.11-3.14, we cannot individually distinguish and identify these four parameters when fitting the model to steady-state distributions. Similarly, we also cannot individually distinguish the concentration of MalE in the periplasm and the dissociation constants of Tar to the maltose-MalE complex (Equation 3.11). While this means that I cannot use the model-fitting protocol to directly determine these molecular-level parameters, it does allow us to predict important parameter ratios. Furthermore, under this model-fitting protocol, my transport-and-sensing model has the same number of free parameters as the indirect-binding model; it has just four free parameters: α , $\frac{[R]_{total}}{\beta}$, $\frac{K_{I,Mal}}{\beta}$, and $\frac{K_{A,Mal}}{\beta}$. We will see that the transport-and-sensing model allows for a considerably better fit than the indirect-binding model. Because my model does not increase the number of free parameters, we can thus deduce that the better fits are due to the quadratic term that appears in Equation 3.11, which accounts for the limiting concentrations of both free maltose and the receptors in the periplasm.

3.3 Chemotaxis assays for single and opposing chemotaxis gradients

To test the validity of my model, I conducted chemotaxis experiments using a three-channel microfluidic device that creates steady, linearly varying concentrations of chemoattractants (Figure 3.2, Appendix A.4)[7, 217]. I conducted not only single maltose chemotaxis assays, but also single MeAsp and opposing MeAsp and maltose chemotaxis assays. This allowed us to both test the validity of the HMWC model for chemoattractants that share the same receptor type and to use the parameter fitting for MeAsp to constrain the parameter fitting for maltose.

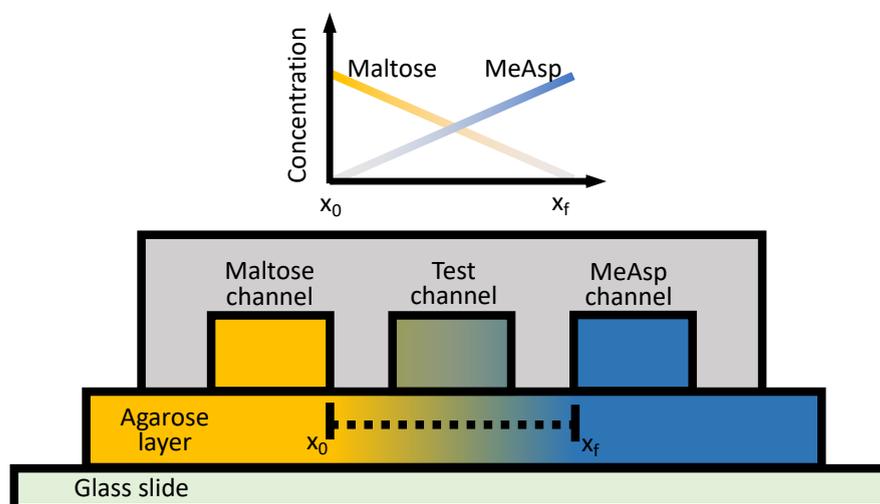


FIGURE 3.2: **Side view of the three-channel microfluidic device used to expose cells to constant chemoattractant gradients.** The channels were 20 mm long, 100 μm deep, 600 μm wide, and with 400 μm spacing between each channel. The agarose layer was 0.5 mm thick. Motility medium with a constant maltose concentration was flowed through the left outer channel, and motility medium with a constant MeAsp concentration was flowed through the right outer channel. For the opposing-gradient experiments, both the maltose and MeAsp concentrations were nonzero; for the single-gradient experiments, one of the two outer channels contained only motility medium. The chemoattractants diffused through the agarose layer and up into the test channel, creating two steady linear concentration profiles in opposite directions. Figure adapted from Yawata, *et al.* [217]

For the MeAsp single gradient experiments, I found the strongest chemotactic response (that is, the steepest steady-state cell distribution in the test channel) when using a MeAsp concentration in the source channel of 46 μM of MeAsp (Figure 3.3B). I was unable to use population-level data to determine the concentration at which Tar's sensing of MeAsp saturates because the cells are, in fact, repelled by very high concentrations of MeAsp (Figure A.1A), likely due to pH taxis [216]. Therefore, I analyzed the response of cells to MeAsp concentrations up to 500 μM , above which increasing the MeAsp concentration caused a drop in pH (Figure A.1B).

Experiments using single gradients of maltose (Figures. 3.3A and 3.4A) corroborate the narrow range of responses previously observed using a FRET reporter [146]. I found that cells show a measurable chemotactic response in gradients created from approximately 0.5 to 20 μM of maltose in the source channel. The response peaked at an intermediate value of 4 μM of maltose, corresponding to a maximum concentration of

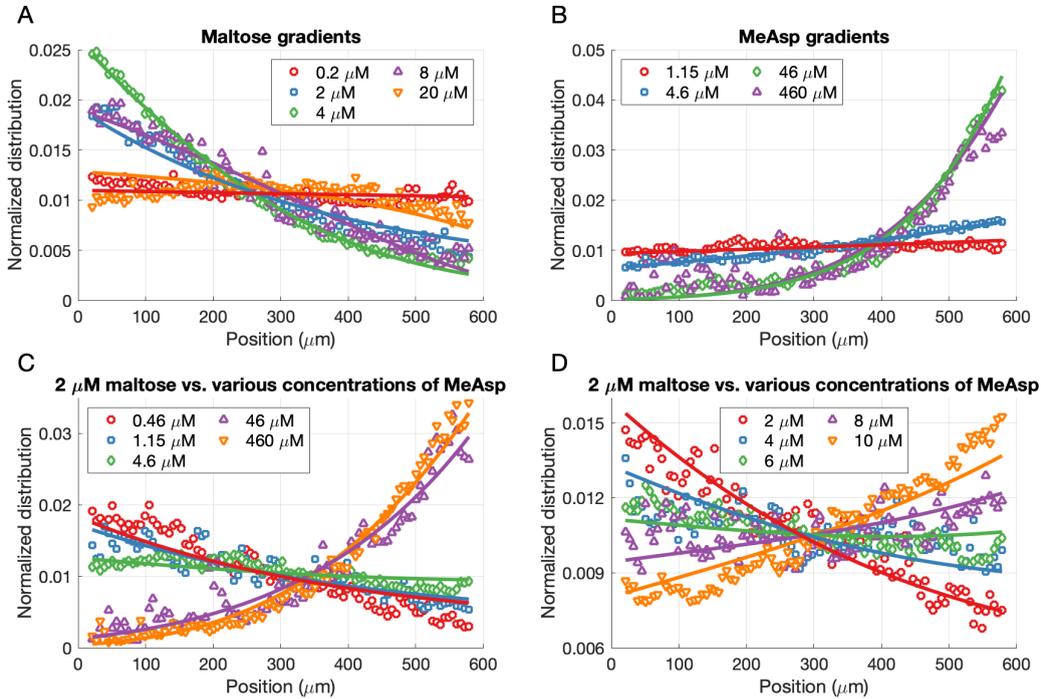


FIGURE 3.3: **Steady-state distributions from experimental chemotaxis assays in single and opposing gradients of MeAsp and maltose along with the best fits obtained using an analytical approximation of the transport-and-sensing model.** In each plot, the maximal maltose concentration is at position $x = 0 \mu\text{m}$ and the maximal MeAsp concentration is at position $x = 600 \mu\text{m}$. Legends give the chemoattractant concentrations in the outer channels. (A) Single-gradient maltose experiments. (B) Single-gradient MeAsp experiments. (C) Opposing-gradient experiments with 2 μM of maltose in the left outer channel and various MeAsp concentrations in the right outer channel. (D) Additional opposing-gradient experiments with 2 μM of maltose in the left outer channel and further intermediate MeAsp concentrations in the right outer channel.

2.86 μM within the test channel (Appendix A.4).

In opposing gradients of maltose and MeAsp, the cells either favored one substrate over the other or had a flat distribution. This response is qualitatively very different from the response predicted from SPECS simulations that used the previous chemotaxis models and fits for aspartate [90] and maltose [146]. SPECS simulations instead predicted that the *E. coli* cells would accumulate in the center of the channel for various intermediate concentrations of maltose and MeAsp (Figure A.2). Under my culture conditions, the cells showed an equal preference for 2 μM of maltose and approximately 7 μM of MeAsp (Figures 3.3 and 3.4). Calculated chemotaxis migration coefficients (CMCs) of the cells

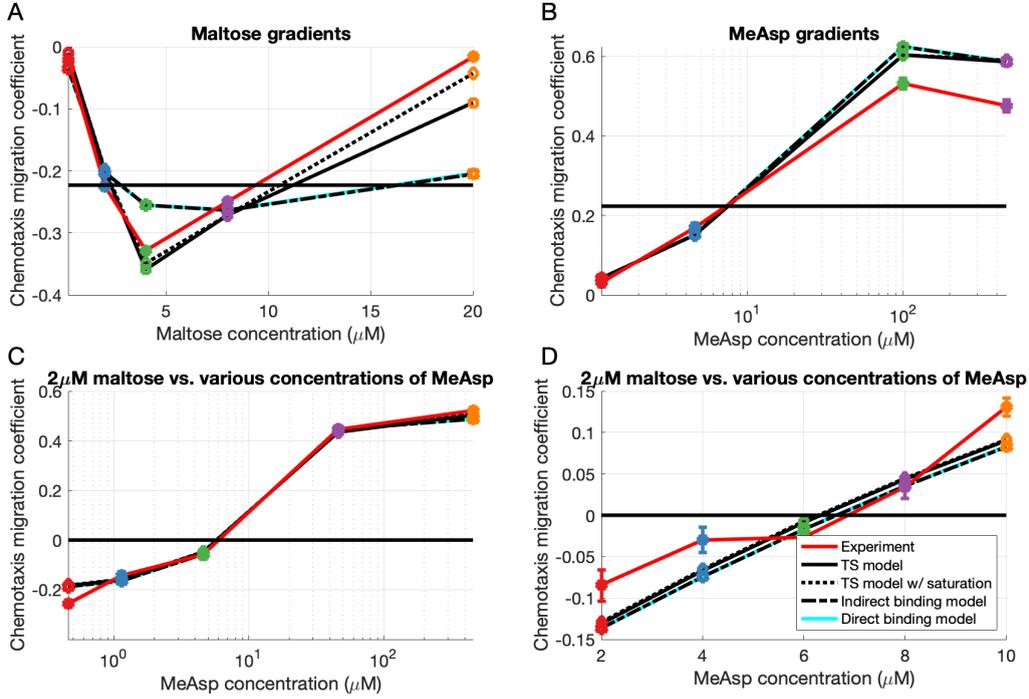


FIGURE 3.4: **Comparison of fits to experimental data among different chemotaxis models.** I compare the experimental distributions as measured by the chemotaxis migration coefficient (CMC) from chemotaxis assays in single and opposing gradients, with values estimated from analytical approximations using: (i) the transport-and-sensing (TS) model, (ii) the transport-and-sensing model with an added saturation term, (iii) the indirect-binding model [146], and (iv) the direct-binding model [90]. The CMC is defined as $CMC = \frac{\langle x \rangle - 300 \mu M}{300 \mu M}$, where $\langle x \rangle$ is the average position of the cells across the 600 μM channel. The CMC is approximately 0 in opposite gradients of 2 μM maltose and 7 μM MeAsp. This is in accordance with the estimates from single gradients: the CMC in a single gradient of 2 μM maltose is approximately -0.225; similarly, my fitting predicts that the CMC in a single gradient of 7 μM MeAsp is approximately 0.225. Panels correspond to the distributions shown in Figure 3.3.

to the single gradients of these two concentrations (Figure 3.4; Appendix A.4) were of equal magnitude (CMC = 0.225).

3.4 Using population-level data to fit a molecular-level model

To fit the approximate analytical expression of the steady-state distributions to direct-binding, indirect-binding, and transport-and-sensing chemotaxis models (Appendix A.3), I used the chemotaxis assay data from single and opposing gradient experiments jointly to fit the parameters for both MeAsp and maltose chemotaxis (Appendix A.4). I used MATLAB's nonlinear optimization function, *fmincon*, to minimize the sum over all of the chemotaxis assays of the maximal differences between the predicted and empirical distributions in each assay.

I was unable to find parameter fits for the direct-binding [90] or indirect-binding [146] chemotaxis models that could capture both the cells' narrow dynamic range of maltose sensing and the population's peak response to 4 μM of maltose (Figure A.3). One assumption of my model fitting protocol, that the concentration of the maltose binding protein, MalE, remains constant over all experimental conditions, is potentially violated by changes in gene expression during the course of the experiments because the *mal* regulon is induced by maltose [172]. However, a Western blot analysis of MalE showed no difference in MalE expression levels between cells acclimated to 10 μM of maltose for the duration of the assay and cells that never had any exposure to maltose (Figure A.6). Therefore, because of the short duration of the chemotaxis experiments, the concentration of maltose binding protein remained constant, so the poor fits of the direct-binding or indirect-binding models cannot be explained by its variation.

In contrast, my transport-and-sensing chemotaxis model was able to capture both the cell's peak response to maltose and the narrow dynamic range of the response to maltose (Figure 3.4). In fact, the predicted steady-state distributions of the model's best fit closely match the measured distributions of all of the corresponding single and opposing gradient experiments that I conducted (Figure 3.3). This best fit uses the following parameter values:

- $\eta_0 = 9.33$,
- $\alpha = 117 \mu\text{M}$,
- $\frac{[\text{R}]_{total}}{\beta} = 0.0146$,

- $\frac{K_{I,Mal}}{\beta} = 0.00334$,
- $\frac{K_{A,Mal}}{\beta} = 0.00635$,
- $K_{I,MeAsp} = 17.0 \mu\text{M}$, and
- $K_{A,MeAsp} = 621 \mu\text{M}$.

Previous studies lend support to the new transport-and-sensing model by validating the fitted parameters. The value of $K_{I,MeAsp} = 17.0 \mu\text{M}$ closely matches the value of $K_{I,MeAsp} = 18.2 \mu\text{M}$ found from FRET analysis [127]. The value of $K_{A,MeAsp} = 621 \mu\text{M}$ is within the range of values reported previously, 500-3,000 μM [127, 146], and its falling at the lower end of that range is in line with the fact that I did not test concentrations of MeAsp greater than 500 μM .

Validating the maltose parameter values is more difficult because we cannot distinguish periplasmic uptake rates from cytoplasmic uptake rates using experimental data of the total uptake rate of maltose from the environment into the cytoplasm [106]. To determine the values of α and β , I rely on previous literature that states that the maltose-MalE complex has a dissociation constant of $K_{BP} \approx 2 \mu\text{M}$ [80] and that the dissociation constants of this complex to Tar, $K_{I,Mal}$ and $K_{A,Mal}$, are in the micromolar range [165]. In my model, I assumed based on previous experimental results that $V_c \geq V_p$. If I furthermore make the assumption that, for greatest efficiency [109], $V_c = V_p$, I obtain for my parameter fits that: $\beta \approx [\text{BP}] \approx 1 \text{ mM}$, which also matches previous literature [31]; $[\text{R}]_{total} \approx 15 \mu\text{M}$; and $\frac{K_c}{K_p} \approx 8.5$. Previous work estimated that *E. coli* has approximately 5,000 Tar receptors [192]; therefore, if I assume a periplasmic volume of approximately $0.3 \mu\text{m}^3$, I obtain an estimate of $[\text{R}] \approx 30 \mu\text{M}$, which reasonably corroborates my prediction given that the concentration of Tar varies extensively over different growth conditions [94, 156]. If I use the effective half-saturation concentration of MalFGK₂ for *B. subtilis*, $K_{eff} = \frac{K_c K_{BP}}{[\text{BP}]_{total}} \approx 50 \text{ nM}$ [174], I obtain $K_c \approx 25 \mu\text{M}$ and $K_p \approx 3 \mu\text{M}$, which matches the micromolar value of the apparent K_m of the total maltose uptake system [31]. Note that for these values and taking $[\text{L}]_{ext} = 2 \mu\text{M}$, my model predicts that $[\text{L:BP}] \approx 17 \mu\text{M}$ and $[\text{L}]_p \approx 34 \text{ nM}$.

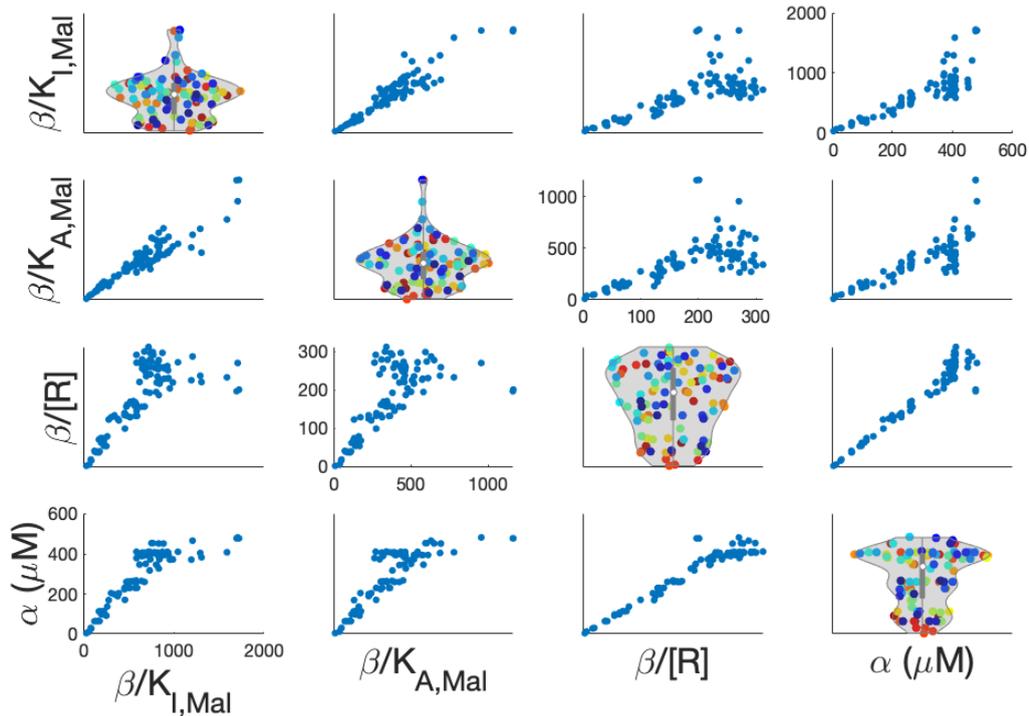


FIGURE 3.5: **Sensitivity analysis for the fitted model parameters that determine the range of maltose sensing.** Plots show the fitted parameter values for 100 “noisy” data sets obtained by randomly selecting 1/1000 of the bacterial position data points without replacement from each experiment. Off-diagonal elements show the correlations between the estimated parameter values. Diagonal elements are violin plots, showing the distribution of estimates (y-axes) obtained for each parameter. For plots of full sensitivity analysis, see Figure A.8.

3.5 A sensitivity analysis uncovers a tug-of-war between sensing and transport

To determine the robustness of the fit, I conducted a sensitivity analysis in which I re-ran the parameter fitting optimization for one hundred different data sets obtained from resampling with replacement all the experimentally obtained bacterial positions (Appendix A.4; Figure A.8). This analysis reveals two striking sets of correlations. Figure A.8 shows strong correlations between η_0 ; the MeAsp dissociation constants, $K_{I,MeAsp}$ and $K_{A,MeAsp}$; and the ratio of the maltose dissociation constants, $\frac{K_{I,Mal}}{K_{A,Mal}}$. Simulations using a modification of SPECS, in which I modified the values of these parameters,

demonstrate that these high correlations exist because each of these four parameters affects the steepness of the steady-state distributions (Figures A.5A,B).

The second striking feature of the sensitivity analysis is the very strong correlation between β and α (Figure 3.5), parameters that incorporate transport kinetics and determine the concentration of maltose bound to MalE (Equations 3.12-3.14). Simulations with modified values of β and α show that increasing the ratio $\frac{\alpha}{\beta}$ increases the range of maltose concentrations that the cell can sense (Figures A.5C,D). Therefore, the analysis is very sensitive to this ratio as it governs the narrow dynamic range of maltose sensing.

When uptake is porin-limited ($V_p \leq V_c$), from Equations 3.13 and 3.14, $\frac{\alpha}{\beta}$ depends on the ratio of ABC transporter units to porins:

$$\frac{\alpha}{\beta}_{V_p \leq V_c} = \left(\frac{K_p}{K_c} \right) \left(\frac{V_c}{V_p} \right). \quad (3.18)$$

However, when uptake is ABC transporter-limited ($V_p > V_c$), the periplasmic concentration of maltose approaches the external concentration ($[L]_p \sim [L]_{ext}$), so that, by Equation 3.12,

$$\frac{\alpha}{\beta}_{V_p > V_c} = \frac{K_{BP}}{[BP]_{total}}. \quad (3.19)$$

The ratio $\frac{\alpha}{\beta}$ may thus provide insight into the cell's narrow sensing range and explain why the cell co-regulates porin and ABC transporter expression so that transport is porin-limited. Because the cell must maintain high levels of binding proteins to achieve high-affinity transport, the ratio is much smaller when the maximal periplasmic uptake rate exceeds the maximal cytoplasmic uptake rate:

$$\left(\frac{\alpha}{\beta} \right)_{V_p > V_c} \ll \left(\frac{\alpha}{\beta} \right)_{V_p \leq V_c}.$$

Thus, the cell would have an even lower dynamic sensing range were it to increase its expression of maltoporins so that maltose uptake is no longer porin-limited. This is because, if maltose uptake were not porin-limited, the periplasm would saturate with maltose at even lower external concentrations, so the receptors would saturate more quickly. Yet, on the other hand, decreasing the permeability of the outer membrane by decreasing porin expression would decrease the total maltose uptake rate. This creates a trade-off between sensing and transport.

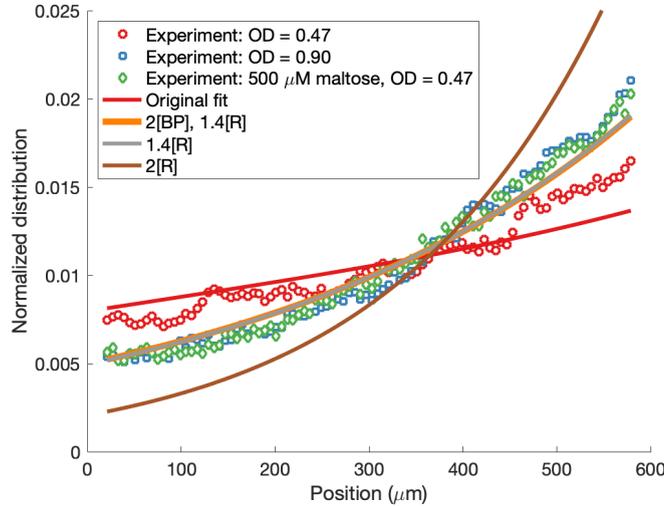


FIGURE 3.6: **Steady-state distributions from chemotaxis assays in opposing gradients created using $2 \mu\text{M}$ maltose and $10 \mu\text{M}$ MeAsp.** Cells were cultured in tryptone broth and harvested as in other experiments at mid-exponential phase ($\text{OD} = 0.47$), harvested at late-exponential phase ($\text{OD} = 0.90$), or cultured in tryptone broth supplemented with $500 \mu\text{M}$ maltose and harvested at $\text{OD} = 0.47$. Curves show fits obtained using an analytical approximation of the transport-and-sensing model, with variants modifying either the concentration of binding proteins, [BP], the concentration of receptors, [R], or both.

3.6 Varying growth conditions to modify chemotactic affinity

An important value of mechanistic chemotaxis models is that they allow one to probe how the cellular migratory response to stimuli may change due to regulation. I performed additional experiments in which I compared the chemotactic response under the original culture conditions, in which cells had no exposure to sugars, with their response when grown in tryptone broth supplemented with $500 \mu\text{M}$ of maltose. Because Tar levels increase over the exponential growth phase of the cells [94], I also performed experiments with cells grown in tryptone broth and harvested later in the exponential phase ($\text{OD} = 0.9$ instead of $\text{OD} = 0.47$). For cells cultured under these three conditions, I conducted an opposing gradient chemotaxis assay with $2 \mu\text{M}$ of maltose and $10 \mu\text{M}$ of MeAsp, concentrations at which neither chemoattractant response was saturated.

As expected from increasing the concentration of Tar in my model, cells grown in tryptone and harvested later in the exponential phase showed a greater relative affinity

to MeAsp. Surprisingly, however, the cells grown in tryptone broth supplemented with maltose showed the same increase in the relative affinity to MeAsp (Figure 3.6). Western blots for MalE confirmed that maltose is indeed an inducer of the mal regulon and that the MalE concentration approximately doubled when maltose was added to the growth medium (Figure A.6). Assuming that the cell maintains a constant $\frac{V_c}{V_p}$ ratio because malK and lamB are on the same operon [30], the model can fit this data if $[R]_{total}$ increases by approximately 40%. Because of the difficulty of distinguishing Tar from Tsr using available immunoblotting methods, I was unable to determine the concentration of Tar over the culture conditions. However, the model prediction closely matches previous research using a different strain of *E. coli* grown on glucose, which found that the cell actively regulated *tar* expression so that the concentration of Tar increased by 40% for similar changes in growth rate [156]. This corroboration gives further support to the transport-and-sensing model.

3.7 Conclusions

In this work, I developed a new transport-and-sensing chemotaxis model that predicts how transport kinetics of maltose into the cell affects the periplasmic concentration of maltose and thus governs the concentration that is available for the cell to sense. Unlike previous molecular-level chemotaxis models, this model does not assume that the periplasmic concentration of a chemoattractant is equal to its extracellular concentration. It introduces parameters that describe the rates of transport of maltose into and out of the periplasm to determine the concentration of maltose-binding protein complex that can bind to receptors. This predictive, mechanistic model thus captures how cells regulate their chemotactic response to maltose.

To fit the molecular-level parameters of my model without requiring genetic modifications or molecular-level experiments, I also developed a simple system identification protocol that uses data from population-level chemotaxis assays, such as those obtained from microfluidic assays. These assays do not require single-cell tracking capabilities but only data of the steady-state distributions of cells. To fit the free parameters of the model to predict the chemotactic response of a species of bacteria to a specific nutrient,

constant-gradient chemotaxis assays should be conducted at various attractant concentrations to determine the approximate dynamic sensing range of the cells (by finding the lowest concentration at which a chemotactic response is observed and the minimal concentration at which the response fully saturates) as well as the approximate concentration at which chemotactic response peaks and the steady-state distribution is steepest. For bacterial species or chemoattractants for which the transport mechanism is unknown, this protocol can be used to compare the fits of the direct-binding model and the new transport-and-sensing model to predict how the cell transports the sensed substrate. In the case of maltose sensing, this approach allowed us to invalidate previous models for chemotaxis to sugars, thus demonstrating the importance of considering porin-limited transport to correctly predict the chemotactic response to saccharides.

The transport-and-sensing chemotaxis model I developed applies more generally to the chemotactic response of gram-negative bacteria to substrates that are transported into the periplasm via facilitated diffusion using specific porins and transported into the cytoplasm via ABC transporters. It thus indicates a general trade-off between sensing and high-affinity transport of sugars. My model predicts that, to achieve high-affinity transport, the cell must express levels of binding protein that greatly exceed the level of ABC transporters. However, if the cell's periplasm were saturated with free sugars as previous work has assumed, the cell's dynamic sensing range would be exceedingly narrow because the high levels of binding protein would quickly saturate the signal. Thus, this work suggests that cells may co-regulate the expression of specific porins and ABC transporters to ensure that total uptake of a specific sugar is porin-limited rather than ABC transporter-limited. With porin-limited transport, the dynamic range of sensing may be greater because it is no longer sensitive to the concentration of binding proteins within the periplasm. Yet, limiting porin expression may decrease total uptake rates. Therefore, the model provides insight into the narrow dynamic range of *E. coli* maltose sensing: this narrow dynamic range is a consequence of the indirect binding of maltose to the Tar receptor and of a trade-off between increasing dynamic sensing range and increasing total uptake rates of maltose into the cytoplasm.

The work in this chapter was done with guidance from Vicente Fernandez, Roman Stocker, and Naomi Levine. Filippo Menolascina conceived the idea to do the sensitivity analysis. Uria Alcolombri performed the Western blot analyses. I would also like to acknowledge the helpful editing of this chapter by Fernandez, Stocker, and Russell Naisbit.

Chapter 4

The Role of ABC Transport in the Oligotrophic Ocean

Marine heterotrophic bacteria play a crucial role in the global cycling of carbon. About half of the global production of organic carbon occurs in the ocean [54], the majority of which is accessible primarily to microbes that thus mediate its flux [13, 38, 65]. Heterotrophic bacteria consume carbon both to build biomass and for energy. The bacteria are thus an important control on the amount of carbon that is sequestered in the ocean versus respired as carbon dioxide back into the atmosphere. Therefore, a mechanistic understanding of their metabolisms is important for predicting how a changing climate will result in further changes in carbon fluxes.

Despite the central role that heterotrophic microbes play in ocean carbon cycling, these dynamics have in general not been included explicitly in global carbon cycling models. This is due to the lack of a mechanistic understanding of marine microbial ecosystem dynamics and links between community composition and function. There are a countless variety of species of heterotrophic bacteria, each of which may specialize in the uptake of particular carbon substrates and have different carbon substrate preferences and utilization rates [62, 79, 159]. Therefore, to predict bacterial carbon fluxes, models must both predict the composition of a microbial community given a particular marine environment as well as how that community's utilization of carbon varies with environmental fluctuations. Yet, we still lack a fundamental understanding of how species should differ given their environmental niche.

Hence, this work attempts to gain insight into the drivers of bacterial speciation in the ocean to better understand how diverse species of heterotrophic bacteria coexist and to better predict how their uptake and growth rates depend on both phenotypic traits and environmental conditions. Previous work suggests that one important driver of speciation is nutrient availability and that there are two main divergent lifestyles: that of copiotrophs, which, by definition, dominate in nutrient-rich environments, and that of oligotrophs, which dominate in nutrient-limited environments [64, 102, 107]. While copiotrophs are unable to persist in starvation conditions, they vastly outcompete oligotrophs in rich environments. Copiotrophs and oligotrophs have very apparent physiological differences. A typical copiotroph, such as a *Vibrio*, exhibits a “feast-and-famine” lifestyle, in which it swims to colonize sporadic, nutrient-rich particles [104, 190]. A copiotroph has a large volume, often greater than $1 \mu\text{m}^3$, and can attain doubling rates greater than 1 per hour [115]. Modeling copiotrophs’ dynamics is thus crucial to predicting degradation versus sequestration rates of particulate organic matter in the ocean.

On the other hand, a typical oligotroph does not attach to particulate organic matter and is thus considered “free-living” [119]. An oligotroph is small, generally with a volume of less than $0.1 \mu\text{m}^3$, and is nonmotile. It is also much slower-growing than a copiotroph, attaining maximal growth rates of less than 0.1 per hour [115]. Although copiotrophs have higher growth rates, oligotrophs are more numerous because the majority of the ocean is nutrient-limited. The most ubiquitous clade of oligotrophs, SAR11, alone makes up about one-third of all bacteria in surface waters [141]. The SAR11 clade is estimated to account for about half of amino acid assimilation in North Atlantic surface waters [123] and to oxidize anywhere from 5-40% of marine primary production [210]. Furthermore, because oligotrophs can consume nutrients found at such low concentrations, their abundances may determine the ambient concentrations of many dissolved organic substrates [200]. Therefore, predicting carbon fluxes in the ocean depends on our ability to model not just copiotrophic dynamics but oligotrophic dynamics as well. However, most of our mechanistic understanding of bacterial metabolism derives from research on copiotrophic-like cells, such as the enteric model organism *E. coli*, that are easier to culture and to experiment with than oligotrophs because they have less auxotrophies and grow faster [61].

However, new expertise is now allowing scientists to culture a variety of slow-growing cells to probe their physiologies and link their characteristics to their proteomes and

genomes. Genomic analyses have yielded important insight into the differentiation of copiotrophic and oligotrophic marine bacteria, suggesting that their divergent traits for growth in nutrient-rich versus nutrient-poor environments are correlated with the presence and absence of particular genes for nutrient transport. Research has found that, while model copiotrophs have a large number of distinct genes for specific phosphotransferase systems (PTS), permeases, and specialized porins, oligotrophs instead rely on a fewer number of “multi-functional” ATP-binding cassette (ABC) transport systems and general porins [43, 115, 150, 189].

In this work, I use this genomic insight to gain mechanistic understanding of the metabolisms of oligotrophs in comparison with that of copiotrophs. I hypothesize that the differences in their nutrient uptake systems are key to understanding their divergent physiologies. I assume that phenotypic variation is the result of trade-offs for maximizing growth rate. I, therefore, quantify trade-offs between PTS or permease (direct) transport systems and ABC transport systems to predict optimal bacterial uptake and growth rates. I develop models for a direct transport system and an ABC transport system. My model for ABC transport demonstrates how cells can tune the effective half-saturation concentration of uptake via regulation of binding protein expression. The transport models furthermore suggest a rate-affinity trade-off: while direct transport systems can achieve higher maximal uptake rates per proteomic cost, ABC transport systems can achieve higher specific affinities [39] per proteomic cost, which allow more efficient uptake at low nutrient concentrations. To quantify these costs, I incorporate the transport models into a simple model of single-cell metabolism that expands a toy model for a self-replicator [138]. This simple model is used in a proteome allocation problem that optimizes cell surface area-to-volume ratio and intracellular concentrations of various metabolites and protein groups to maximize a steady-state growth rate. The optimization results suggest that, to achieve high specific affinities, cells relying on direct transport systems must have very high transport unit to downstream enzyme ratios and thus require efflux systems, while, on the other hand, cells relying on ABC transport systems can more efficiently match their transport and catabolic capacities. Finally, to gain insight into how uptake and growth rates of cells may vary over environmental conditions, I perform a sensitivity analysis on the optimization problem that suggests that the solute size of the growth-limiting carbon source is an important characteristic for determining uptake rates of oligotrophic cells.

4.1 Modeling nutrient uptake

4.1.1 Michaelis-Menten kinetics

Rates of nutrient consumption are commonly described using Michaelis-Menten kinetics [131], in which the uptake rate of a substrate with external concentration $[S]_{ext}$ is described by a saturation function with two parameters:

$$v = V_{max} \frac{[S]_{ext}}{K_m + [S]_{ext}}, \quad (4.1)$$

where V_{max} is the maximal uptake rate and K_m is the half-saturation constant. While the Michaelis-Menten equation is most commonly written in this form, we can more easily interpret the nutrient uptake strategies needed to sustain copiotrophic versus oligotrophic lifestyles if we write the equation in the form pioneered by Button [41]:

$$v = a^\circ V_{max} \frac{[S]_{ext}}{V_{max} + a^\circ [S]_{ext}}, \quad (4.2)$$

where the specific affinity is $a^\circ = V_{max}/K_m$. At high nutrient concentrations, the uptake rate approaches the maximal uptake rate, while, at low nutrient concentrations, the uptake rate is proportional to the specific affinity. Therefore, I expect copiotrophs to have a transport system with a higher maximal uptake rate that allows fast growth in nutrient-rich conditions, while I expect oligotrophs to have a transport system with a higher specific affinity that allows them to persist in nutrient-limited conditions. While previous work has hypothesized the existence of a rate-affinity trade-off [33, 71], a mechanistic explanation for this trade-off has been lacking. Hence, I develop transport models to predict how the maximal uptake rate and the specific affinity are functions of the cell's phenotype and incorporate these transport models into a proteome allocation model to mechanistically explain and quantify the trade-off.

4.1.2 Periplasmic uptake

The majority of observed marine heterotrophic bacteria are gram-negative cells, containing two membranes. Thus, nutrient uptake is a two-step process: (*i*) the substrate passes

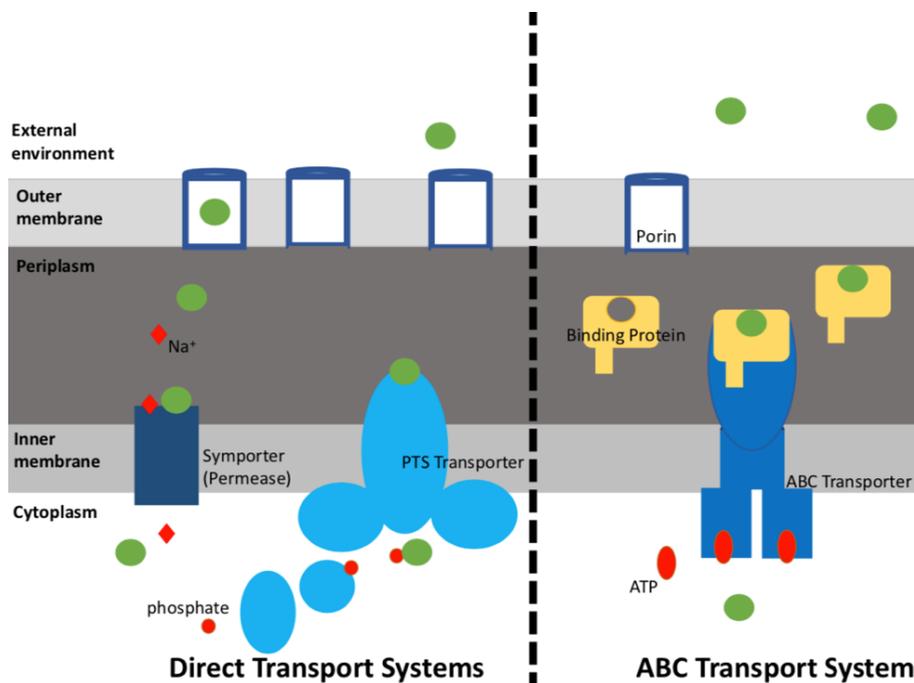


FIGURE 4.1: **A cartoon of nutrient uptake into the cell.** A generic carbon substrate (in green) passively diffuses through the outer membrane into the periplasm via general porins, which are channel-forming proteins. To pass through the inner membrane, membrane-bound transport units expend energy to modify the substrate or translocate the substrate against a concentration gradient. For direct transport systems, the solute binds directly to the membrane-bound transport unit. The two most prevalent types of direct transport systems in marine bacteria are sodium-ion symporters, which co-translocate a sodium ion with the solute, and phosphotransferase systems (PTS) that use a cascade of proteins to phosphorylate particular carbohydrates to draw them across the cytoplasm. On the other hand, ATP-binding cassette (ABC) transport systems use binding proteins in the periplasm to scavenge for and store the solute in the periplasm. These binding proteins bind to a membrane-bound transport unit that uses ATP to translocate the solute.

through the outer membrane to enter the periplasm; and (ii) the substrate passes through the inner membrane to enter the cytoplasm. (See Figure 4.1.)

Heterotrophic marine bacteria are osmotrophs, which rely on the diffusion of nutrients towards the cell surface for their subsequent uptake into the cell. The periplasmic uptake rate is thus potentially limited by the diffusion of the substrate, with diffusivity D , toward the cell. To obtain a simple equation for this rate, I assume that the cell is spherical with radius r to obtain a diffusive uptake rate [12] of:

$$v_{diff} = \frac{3D ([S]_{ext} - [S]_0)}{f_p r^2} \left[\frac{\text{Periplasmic concentration}}{\text{Time}} \right], \quad (4.3)$$

where $[S]_0$ is the concentration of substrate on the surface of the cell and f_p is the periplasmic fraction of the cell volume. Therefore, smaller cells can attain larger periplasmic uptake rates because of their higher surface area-to-volume ratios (that scale with the inverse of r).

To translocate the nutrients from the cell's surface into the periplasm, the cell relies on a concentration gradient so that periplasmic uptake does not require any energy. Oligotrophs use general porins on the outer membrane, which are channel-forming proteins through which the substrates can passively diffuse [20, 144]. While copiotrophs use general porins as well, they also use specific porins to consume larger solutes that diffuse orders of magnitude more slowly than smaller solutes [11, 20, 149]. A specific porin binds to a particular substrate and thus orients the substrate to draw it into the periplasm; this is known as facilitated diffusion.

For simplicity, in this work, I assume that periplasmic uptake is via passive diffusion from a general porin. I thus assume that the periplasmic uptake rate is proportional to the abundance of porins on the cell's outer membrane, N_{porin} , as well as the difference between $[S]_0$ and the periplasmic concentration of that substrate, $[S]_p$:

$$v_p = \frac{k_{porin} N_{porin} ([S]_0 - [S]_p)}{f_p \left(\frac{4\pi r^3}{3}\right)} \left[\frac{\text{Periplasmic concentration}}{\text{Time}} \right], \quad (4.4)$$

where k_{porin} is a flow rate in volume per time that accounts for the permeability of the substrate through the porin. (See Appendix B.1.1 for a derivation of this term.)

Because at steady-state the diffusive uptake rate must equal the periplasmic uptake rate, $v_{diff} = v_p$, I combine Equations 4.3 & 4.4 to remove the surface substrate concentration and obtain a periplasmic uptake rate that depends both on porin abundance and cell radius:

$$v_p = \frac{k_{porin} N_{porin} ([S]_{ext} - [S]_p)}{f_p \left(\frac{4\pi r^3}{3}\right) \left(1 + \frac{k_{porin} N_{porin}}{4\pi D r}\right)} \left[\frac{\text{Periplasmic concentration}}{\text{Time}} \right]. \quad (4.5)$$

4.1.3 Cytoplasmic uptake

For the substrate to enter the cytoplasm from the periplasm, it now needs to go against a concentration gradient, and this requires energy. In this work, I compare two main

types of energy-consuming transport systems: direct transport and ATP-binding cassette (ABC) transport systems.

Direct transport. I define direct transport systems as systems in which the substrate itself binds to the inner membrane-bound transport unit. One type common in marine bacteria is a sodium ion-linked symporter, a protein channel that transports a specific substrate against a concentration gradient by co-transporting it with a sodium ion in the direction of a sodium gradient [92, 145]. Another prevalent type of direct transport used by copiotrophs is a phosphotransferase transport system (PTS) for the uptake of specific carbohydrates [110, 167, 169, 180]. The PTS uses a cascade of proteins that are specific to the sugar to bind a phosphate group to that sugar as well as a general membrane-bound protein carrier that transports the modified sugar into the cytoplasm. Therefore, for PTS, the cell uses energy to modify the sugar so that there is, in fact, no concentration gradient that the cell must work against to translocate it.

I can model direct transport systems as a two-step reaction [32]. First, the periplasmic substrate, S_p , binds to the membrane-bound transport unit, T , to form a bound complex, $T:S$. In this work, I assume that the rate of binding is much faster than its dissociation rate so that I can model it as an irreversible reaction that proceeds with rate $k_{2'}$:



Once the substrate is bound to the transport unit, it is then translocated with rate $k_{3'}$:



I describe the rates using mass-action kinetics and assume a steady state so that the concentration of the substrate-transport complex does not change:

$$\frac{d[T:S]}{dt} = 0 = k_{2'}[S]_p[T] - k_{3'}[T:S], \quad (4.8)$$

where $[T]$ ($[T:S]$) is the number of unbound (bound) transport units divided by the periplasmic volume of the cell. Thus, from Equation 4.8, I obtain the following cytoplasmic uptake rate for direct transport:

$$v_c = k_{3'}[T:S] = k_{3'}[T]_{total} \frac{[S]_p}{K_{T'} + [S]_p} \left[\frac{\text{Periplasmic concentration}}{\text{Time}} \right], \quad (4.9)$$

where $K_{T'} = \frac{k_{3'}}{k_{2'}}$ and $[T]_{total} = [T] + [T:S]$ is the total number of membrane-bound receptor units divided by the volume of the periplasm. (Note that the rate as given describes the periplasmic concentration of substrate leaving the periplasm per unit time. I multiply it by the ratio of the periplasmic volume to the cytoplasmic ratio to obtain the concentration of substrate entering the cytoplasm per unit time.)

ABC transport. For ABC transport, binding proteins scavenge for the incoming nutrients in the periplasm so that they do not escape back into the extracellular environment. These binding proteins, when in complex with the substrate, bind to membrane-bound transport units, which require ATP to translocate the substrate from the periplasm into the cytoplasm [44, 80, 83]. I assume that the periplasmic substrate binds reversibly to the binding protein, BP, to form substrate, binding-protein complex, S:BP, with association rate k_{1f} and dissociation rate k_{1r} :



I make the same simplifying assumption as I did for direct transport that binding of the complex to the membrane-bound transport unit to form substrate, binding protein, transport unit complex, T:S:BP, is irreversible with rate k_2 :



I additionally assume that the dissociation of the binding protein from the transport unit is simultaneous with translocation of the substrate with rate k_3 :



At steady-state, the periplasmic concentrations remain unchanged:

$$\frac{d[BP]}{dt} = -v_1 + v_3, \quad \frac{d[S:BP]}{dt} = v_1 - v_2, \quad \frac{d[T:S:BP]}{dt} = v_2 - v_3, \quad (4.13)$$

where, again using mass action kinetics,

$$v_1 = k_{1f}[S]_p[BP] - k_{1r}[S:BP], \quad (4.14)$$

$$v_2 = k_2[T][S:BP], \quad \text{and} \quad (4.15)$$

$$v_3 = k_3[T:S:BP]. \quad (4.16)$$

Therefore, at steady-state, $v_1 = v_2 = v_3$. I furthermore can define the following periplasmic conservation equations to solve for all periplasmic concentrations:

$$[BP]_{total} = [BP] + [S:BP] + [T:S:BP] \quad (4.17)$$

$$[T]_{total} = [T] + [T:S:BP] \quad (4.18)$$

Because $v_2 = v_3$,

$$[T:S:BP] = [T]_{total} \frac{[S:BP]}{K_T + [S:BP]}, \quad K_T = \frac{k_3}{k_2}. \quad (4.19)$$

Because $v_1 = v_2$,

$$(k_{1r} + k_2[T]) [S:BP] = k_{1f}[S]_p[BP]. \quad (4.20)$$

I use Equations 4.17-4.20 to solve for the four unknown concentrations in Equation 4.20 using Wolfram Mathematica. These exact solutions are used in the proteome model discussed below (Appendix B.1.2.2).

4.1.4 Modifying the half-saturation concentration of uptake via protein expression

We can gain useful insight into ABC transport systems if we consider the case that the concentration of binding proteins greatly exceeds the concentration of transport units, $[BP]_{total} \gg [T]_{total}$, so that $k_2[T] \ll k_{1r}$. (This is the case for the best-studied ABC transport system, the maltose transport system in *E. coli* [31, 47]. See also Chapter 3.) In this case, $[S:BP] \gg [T:S:BP]$, so

$$[S:BP] \approx [BP]_{total} \frac{[S]_p}{K_{BP} + [S]_p}, \quad \text{where } K_{BP} = \frac{k_{1r}}{k_{1f}}, \text{ and} \quad (4.21)$$

$$[T:S:BP] \approx [T]_{total} \left(\frac{[BP]_{total}}{K_T + [BP]_{total}} \right) \left(\frac{[S]_p}{\frac{K_T K_{BP}}{K_T + [BP]_{total}} + [S]_p} \right). \quad (4.22)$$

We can better understand the implications of ABC versus direct transport by comparing the maximal uptake rates and specific affinities that these approximations suggest are attained by these different transport strategies. From Equations 4.9, 4.16, and 4.22, the maximal cytoplasmic uptake rates of direct and ABC transport systems, which are attained when the periplasmic concentration of the substrate saturates, are:

$$V_{max,direct} = k_{3'}[T]_{total} \quad \text{and} \quad (4.23)$$

$$V_{max,ABC} = k_3[T]_{total} \left(\frac{[BP]_{total}}{K_T + [BP]_{total}} \right). \quad (4.24)$$

Therefore, the fact that copiotrophs achieve higher maximal uptake rates than oligotrophs may suggest that direct transport systems have either higher translocation rates ($k_{3'}$ versus k_3) and/or that copiotrophs express higher levels of membrane-bound transport units. Similarly, from the same equations, the cytoplasmic specific affinities and half-saturation concentrations of direct and ABC transport systems are:

$$a_{direct}^{\circ} = k_{2'}[T]_{total}, \quad a_{ABC}^{\circ} = \frac{k_2[T]_{total}[BP]_{total}}{K_{BP}} \quad (4.25)$$

$$K_{M,direct} = K_{T'}, \quad \text{and} \quad K_{M,ABC} = \frac{K_T K_{BP}}{K_T + [BP]_{total}}. \quad (4.26)$$

For direct transport systems, my expression for the cytoplasmic specific affinity is consistent with Button's theory [40, 41], which suggests that the specific affinity is proportional to the number of transport units on the inner membrane as well as the diffusivity of the substrate, which affects both the affinity of periplasmic uptake and the transport rate $k_{2'}$. Thus, because of the lack of a mechanistic understanding of ABC transport, the effective half-saturation concentration of any transport system has traditionally been viewed as a constant that is intrinsic to the transport proteins. In my formulation, the half-saturation constant of direct transport systems is equal to the dissociation constant $K_{T'}$ between the substrate and transport unit, in agreement with Button's theory. Indeed, Button developed his theory for specific affinity because the effective half-saturation concentrations of oligotrophic transport were too small to be explained by changes in the dissociation constant, K_T [39]. More recent literature has conflated the effective half-saturation concentration of oligotrophic ABC transport with the dissociation constant of the substrate and binding protein, K_{BP} . For example, see [150].

However, as my analysis makes clear, the half-saturation concentration of an ABC transport system is not, in fact, a constant intrinsic to the transport proteins. The effective half-saturation concentration for ABC transport depends not only on both dissociation constants but also on the concentration of binding proteins in the periplasm. This matches previous analysis that suggests that binding proteins increase the encounter rate of the substrate and the membrane-bound transport unit [32]. My analysis thus suggests that oligotrophs rely primarily on ABC transport systems because they allow the cell to decrease the half-saturation concentration of uptake via protein expression to achieve high specific affinities.

4.2 A toy metabolic model

In the previous section, I described a model of diffusive periplasmic uptake as well as models of direct and ABC cytoplasmic uptake to understand how a cell can achieve high specific affinities and high maximal uptake rates. However, maximizing either quantity will not necessarily confer an advantage to a cell if it comes with too high a cost. To understand the implications of these transport systems, we need to understand how they affect not just the cell's uptake rate but also the cell's growth rate. Therefore, in this section, I incorporate the transport system models into a simple, mechanistic single-cell metabolic model to predict how phenotypes should vary in oligotrophic versus copiotrophic environments. This metabolic model is an extension of the "self-replicator" proposed by Molenaar and others [138] and, as in their work, is used in a proteome allocation problem that determines the amount of each type of protein that the cell should express in order to maximize its growth rate given a specific external concentration of a generic carbon substrate.

Therefore, the main assumptions of this model are not only that cells have evolved to maximize an instantaneous growth rate but also that oligotrophic and copiotrophic environments can be distinguished by their carbon availability. The implication of these two assumptions is that the phenotypic traits that vary between oligotrophs and copiotrophs must have cost-benefit trade-offs for growth rate maximization in their different environments. Consequently, to understand the evolutionary advantage of a particular phenotypic trait, we must quantify how it affects the cell's growth rate. The benefit of

increasing the expression of transport units and/or binding proteins is clear: it increases the uptake rate of carbon into the cell to increase the cell's growth rate. At saturating nutrient concentrations, this is equivalent to increasing the cell's maximal uptake rate; at very low nutrient concentrations, this is equivalent to increasing the cell's specific affinity.

4.2.1 Proteomic cost of transport

We similarly can think of a cost as causing a reduction in the cell's growth rate. Any trait has associated opportunity costs because some portion of the cell's resources are diverted to implementing that trait instead of being used directly for replication. A phenotype may confer a variety of opportunity costs. A primary opportunity cost is the proteomic cost of a trait [16, 17, 85, 93, 151, 156]. Research has shown that, under a variety of experimental conditions, a cell's growth rate is proportional to the fraction of its proteome devoted to ribosomes [175]. Therefore, expressing higher abundances of membrane-bound transport units or binding proteins to increase uptake rate presents an opportunity cost because the cell could have instead devoted a greater proportion of its proteome to ribosomes. Therefore, in my metabolic model, I quantify the proteomic cost of transport by approximating the number of amino acids needed per transport unit. For my baseline model, I approximate that the membrane-bound ABC transport unit has the same proteomic cost as the direct transport unit and that a binding protein has approximately one fourth of this proteomic cost. (For a specification and explanation of all parameter values used, see Appendix B.1.)

4.2.2 Energetic cost of transport

Similarly, many traits also have an energetic cost. Cells maintain an ATP/ADP ratio about 10^8 times greater than its equilibrium ratio to drive energetically unfavorable processes by coupling them with ATP hydrolysis [73, 108]. Thus, as the cell converts ATP to ADP to carry out energetically costly reactions, it must maintain energy homeostasis by continually using energy to convert ADP back into ATP. Carbon is a crucial substrate for heterotrophic bacteria because it is used not only as a building block for amino acids and other biomass but is also a means for synthesizing ATP. Aerobic heterotrophs can use carbon to produce ATP via two main processes. First, the cells can perform glycolysis,

a substrate-level phosphorylation within the cytoplasm in which a glucose molecule is transformed into pyruvate and generates two ATP molecules from ADP. In addition, by a second process, called oxidative phosphorylation or cellular respiration, the cell uses oxygen as an electron acceptor to further break down the pyruvate into carbon dioxide to create at most a total of 36 ATP molecules per molecule of glucose [154]. Cellular respiration is carried out by a cascade of proteins, called the electron transport chain, and ATP synthase. The electron transport chain units are bound to the inner membrane, as they generate an electrochemical proton gradient across the membrane, known as the proton-motive force. This force, via a translocation of protons, drives ATP synthase to produce ATP from ADP. Thus, the proton-motive force is an additional currency of energy used by the cell [73], and is, in fact, used directly as energy by a number of symporters.

In my metabolic model, I specify the cost of an energy-consuming trait by quantifying the equivalent number of ATP used per second. The ABC transport system uses ATP directly. However, it is not known how tightly linked substrate translocation is to ATP usage [162]. While some studies suggest a ratio of 1-2 ATP units used per substrate translocated [143, 155], others suggest that ABC transport may be more energetically costly and that it may even use ATP in a futile cycle when no substrate is present [158]. On the other hand, research suggests that symporters, a type of direct transport, co-transport between exactly one to three ions per substrate [92, 111]. Because four protons are pumped through ATP synthase per each ATP synthesized [46], a permease therefore has an equivalent energetic cost of 0.25 to 0.75 ATP per substrate translocated. Finally, it has been argued that direct transport by PTS is the most energy efficient method of transport for carbohydrates because its energy-consuming phosphorylation step, in which a phosphoryl group from phosphoenolpyruvate is transferred to the sugar substrate, is also a required step in glycolysis [167]. The energetic cost of this phosphorylation is equivalent to one ATP consumed per substrate because the transformation of the phosphoenolpyruvate into pyruvate produces one ATP molecule [117].

This analysis suggests that ABC transport has both higher proteomic and energetic costs per substrate translocated than both primary types of direct transport systems, symporters and PTS. However, so that I can individually compare the effects of proteomic and energetic costs, I assume in my baseline model that all transport systems have the same energetic cost of one ATP molecule used per substrate translocated. I study the effects of variable energetic costs in a sensitivity analysis.

4.2.3 Size constraints

An additional opportunity cost of a trait is its cell size requirement. As apparent from Equation 4.3 of the diffusive periplasmic uptake rate, small radii are beneficial to the cell because they confer larger surface area-to-volume ratios and thus larger effective specific affinities and diffusive uptake rates [50, 98, 105, 220]. However, proteins take up space so that greater abundances require larger cell sizes. I follow the self-replicator model [138] and assume that the cytoplasmic density of the cell cannot exceed a specified amount so that diffusion of proteins within the cytoplasm can proceed at a reasonable rate. Similarly, I assume that the periplasmic density cannot exceed that same amount and allow the cell to modify the fraction of its volume devoted to the periplasm versus the cytoplasm.

In addition, the membrane-bound transport proteins incur further cell size requirements because they take up space on the membrane and therefore have associated “real estate costs” [197]. Hence, in my model, I additionally specify the surface area of each membrane-bound protein unit.

I assumed spherical cells to derive simple analytical expressions for the rate of diffusion of substrates to the surface of the cell and to constrain the number of membrane-bound units based on membrane surface area. However, it is important to note that it is the ratio of the abundances of the membrane-bound units to the intracellular metabolites that determine rates and thus optimal cell radii. Therefore, my model is not, in fact, determining optimal cell volumes when optimizing over cell radius; it is optimizing surface area-to-volume ratio. In my model, I do not consider any factors such as motility requirements or predation that may favor particular cell volumes [105, 220].

4.2.4 A proteome allocation problem

My abstract model of cell metabolism attempts to minimize the components needed to account for the different opportunity costs discussed above. (See Figure 4.2 and Appendix B.) This toy model tracks the transformation of a generic carbon substrate into seven different intracellular components by eight protein groups. The generic carbon substrate enters the periplasm via a generic porin. The periplasmic substrate is drawn into the cytoplasm via membrane-bound transport units, in complex with binding proteins for

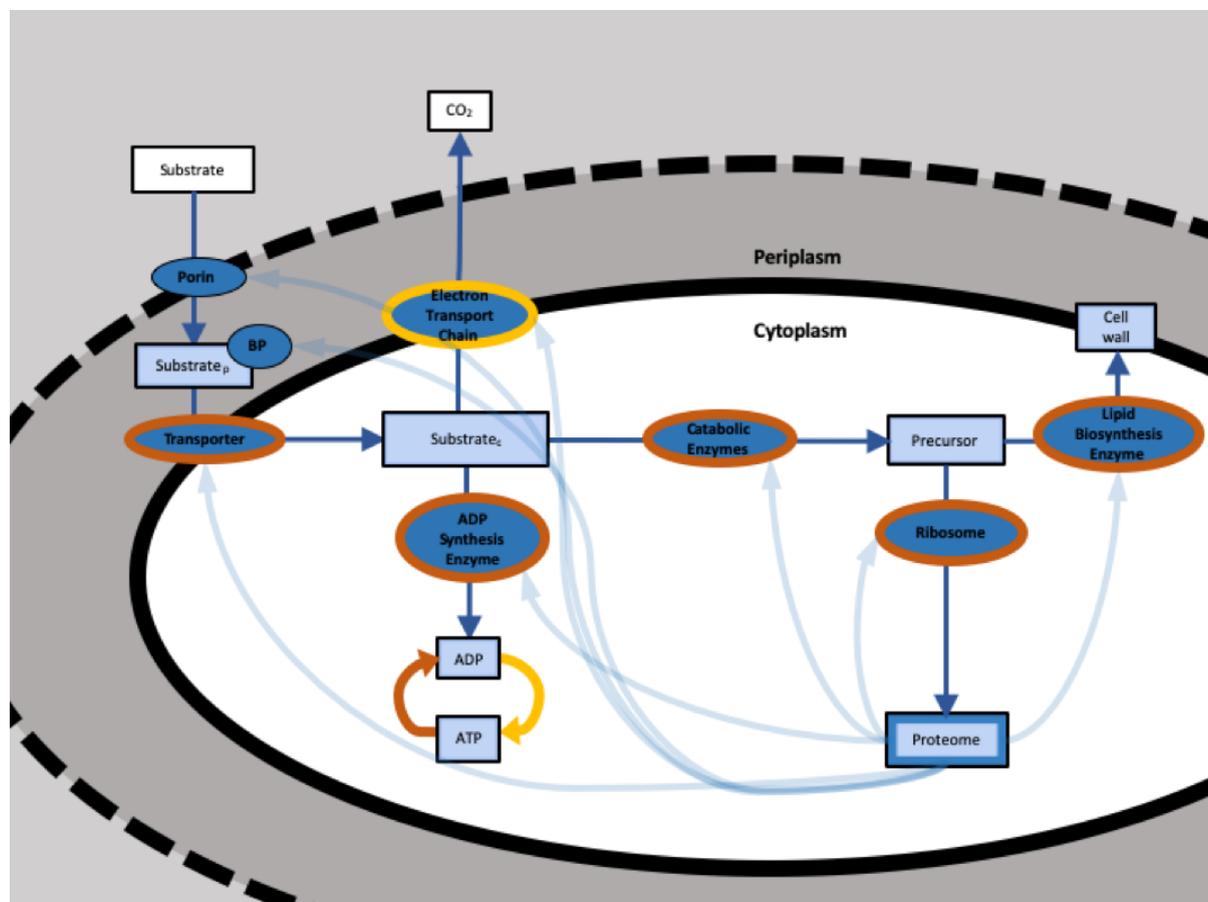


FIGURE 4.2: **A toy model of bacterial metabolism.** The single-cell metabolic model includes seven intracellular metabolites (light blue), one of which is the proteome that is comprised of eight protein groups (dark blue). The model tracks the utilization of carbon by the cell (dark blue arrows). The solute enters the periplasm via porins, may bind to binding proteins in the periplasm, and is then transported into the cytoplasm by membrane-bound transporters. The intracellular carbon substrate is either used to synthesize ADP (via an ADP synthesis protein group), generate energy (via cellular respiration, by the membrane-bound electron transport chain) or is assimilated into biomass. Cellular respiration generates energy by converting ADP to ATP (yellow). The protein groups that use energy result in the conversion of ATP back into ADP (orange). For biomass assimilation, catabolic enzymes transform the intracellular substrate into a precursor (a generic amino acid). This precursor is used by lipid biosynthesis enzymes to make the cell wall and by ribosomes to make proteins. This model is subject to a number of constraints (specified in the text) to determine the proteome allocation that maximizes a steady-state growth rate.

ABC transport systems. This cytoplasmic carbon substrate is transformed into a generic building block via a catabolic protein group. The building block is used both for proteins, expressed by ribosomes, and for the inner and outer membranes, synthesized by a

membrane biosynthesis protein group.

These foregoing components are the minimal required to account for the proteomic cost and size requirements of transport [138]. However, to account for the energetic cost of transport as well, the model additionally tracks the proportion of carbon consumed by the cell that is used for energy instead of for biomass. Thus, the cytoplasmic carbon substrate can also be transformed and excreted from the cell (mimicking the release of carbon dioxide in aerobic respiration) via an energy generation protein group that encapsulates catabolic enzymes, membrane-bound electron transport chains, and ATP synthase. Additionally, the carbon substrate may be used by an ADP synthesis protein group to synthesize ADP, which is converted to ATP to store the generated energy.

I use this model in a nonlinear optimization problem run using MathWorks's MATLAB *fmincon* function, to find the proteome allocation that maximizes the steady-state, balanced growth rate given a particular extracellular concentration of the carbon substrate. To allocate the proteome, the optimization solves both for the total concentration of protein in the cell as well as for the fraction of the proteome devoted to each of the eight protein groups. The code also solves for the optimal concentrations of periplasmic carbon substrate, cytoplasmic carbon substrate, precursor, cell membrane units, and ADP. For energy homeostasis, I assume that the concentration of ATP must be ten times the concentration of ADP. In addition, I include a futile cycle, which simply converts ATP to ADP without requiring any proteome, and allow the code to optimize the rate of this cycle to ensure that a balanced, steady-state solution is found while also ensuring that all nonzero quantities have abundances of at least one. Finally, I also allow the code to optimize over cell radius and periplasmic volume fraction. Thus, including the growth rate as a variable, there are eighteen variables that the code optimizes over.

This optimization problem is subject to a number of constraints, many of which are nonlinear:

1. The cell experiences steady-state exponential growth: $\frac{dx_m}{dt} = Nv - \mu x_m$, where x_m are the seven intracellular substrates; μ is the growth rate that the optimization problem is attempting to maximize; N is a stoichiometry matrix that specifies how many of each input substrate is needed to form an output substrate; and v are the corresponding reaction rates, which are functions of the concentrations of each of

the eight protein groups as well as the concentration of input substrate and are modeled using Michaelis-Menten kinetics.

2. The proteome fractions must sum to one.
3. The sum of the densities of the cytoplasmic components (cytoplasmic carbon substrate, generic precursor, ADP, ATP, catabolic proteins, ribosomes, cell membrane biosynthesis proteins, and ADP synthesis proteins) cannot exceed a threshold cell density.
4. The sum of the densities of the periplasmic components (periplasmic carbon substrate and binding proteins) cannot exceed the same threshold cell density.
5. The abundance of synthesized membrane units must equal the amount needed to cover both the inner and outer membranes.
6. The electron transport chain units as well as transport units must fit on the surface area available on the inner membrane; and
7. The porins must fit on the surface area available on the outer membrane.

The full proteome allocation problem with all equations, constraints, and parameters specified is given in Appendix B. While I use realistic parameter values based on previous knowledge of *E. coli* for all quantities, I cannot expect my toy model of cellular metabolism to predict precise growth and uptake rates. Instead, I use realistic parameter values to substantiate the trends that my model predicts and to evaluate the relative impact of the various uptake mechanisms. Note also that the solutions to the optimizations predict optimal phenotype; that is, they predict the characteristics of the species evolved in a landscape with a specific concentration of a single generic carbon substrate. These solutions do not, therefore, suggest that a single species can attain each phenotype via regulation.

4.3 Optimal proteome allocation for ABC versus direct transport systems

4.3.1 A rate-affinity trade-off

While the turnover rates of direct transport systems have been well studied and are known to be on the order of 100 per second [133, 207], less is known about the achievable rates of ABC transport systems [162]. However, I hypothesize that the maximum achievable turnover rates of ABC transport systems must be intrinsically lower than that of direct transport systems to explain the prevalence of direct transport systems in copiotrophs.

According to my model, the transport systems should have similar proteomic costs per uptake rate at high nutrient concentrations if their turnover rates, k_3 and $k_{3'}$, were the same. And, furthermore, the model suggests that potentially lower energetic costs of direct transport systems do not confer substantial gains at high nutrient concentrations where energy is not limited (Figure 4.3). In addition, previous research suggests that direct transport systems have other costs that my model does not consider because they, in general, have narrower substrate specificity than ABC transport systems [150, 167, 170, 171]. For this reason, copiotrophs must devote additional proteome to tightly regulate the expression of a diverse number of genes for carbon uptake and take advantage of multiple types of carbon nutrient hotspots [42, 78, 156]. Therefore, if the turnover rates of ABC and direct transport systems were the same, ABC transport systems could outperform direct transport systems in all environments that are not limited by electron acceptors for energy generation. In fact, in my model, at the low nutrient concentrations where efficient energy usage is most important, ABC transport still outperforms direct transport when its energetic cost is ten times greater than that of direct transport (Figures 4.6A & 4.4). Therefore, as direct transport systems must have important benefits to explain copiotrophs' reliance on them, I believe that this benefit must be that these systems allow higher maximal uptake rates.

This hypothesis on direct versus ABC transport turnover rates is supported by a molecular understanding of transport. I expect ABC transport systems to have lower turnover rates if their rates are limited by rates of diffusion within the periplasm. Because the binding proteins of ABC transport systems are much bulkier than the carbon substrate

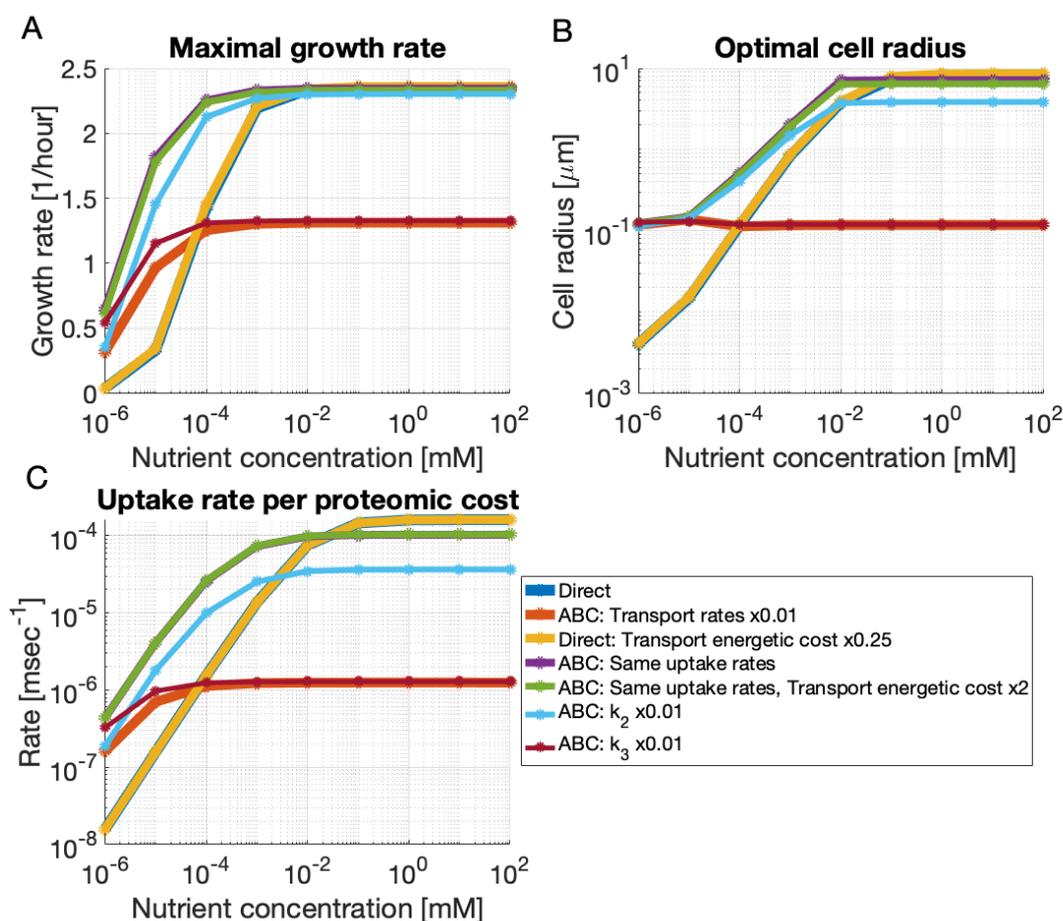


FIGURE 4.3: **Comparing ABC and direct transport.** These plots show results of proteome allocation problems with variations in the direct and ABC transport systems. The proteome allocation problem is solved for different extracellular substrate concentrations (x-axis). (A) shows the maximal growth rates achieved with the optimal proteome allocation, (B) shows the optimal cell radii used to achieve those maximal growth rates, and (C) shows the uptake rate per proteomic cost, which is the rate of carbon consumption divided by the amount of carbon comprising the porins, binding proteins, and membrane-bound transport units. The baseline parameter values for the two transport systems are: ‘Direct’ and ‘ABC: Transport rates x0.01’. I argue that ABC transport rates, k_2 and k_3 , are likely on the order of one hundred times lower than direct transport rates, as in the baseline model, because of the slow diffusivity of the bulky binding proteins and because, when ABC transport has the same transport rates as direct transport (purple), they outcompete direct transport at all concentrations. Direct transport systems may have lower energetic costs (yellow) and ABC transport rates may have higher energetic costs (green), but these modifications do not considerably alter maximal growth rates at high nutrient concentrations. The lower rates of ABC transport (red, maroon) result in smaller optimal cell radii.

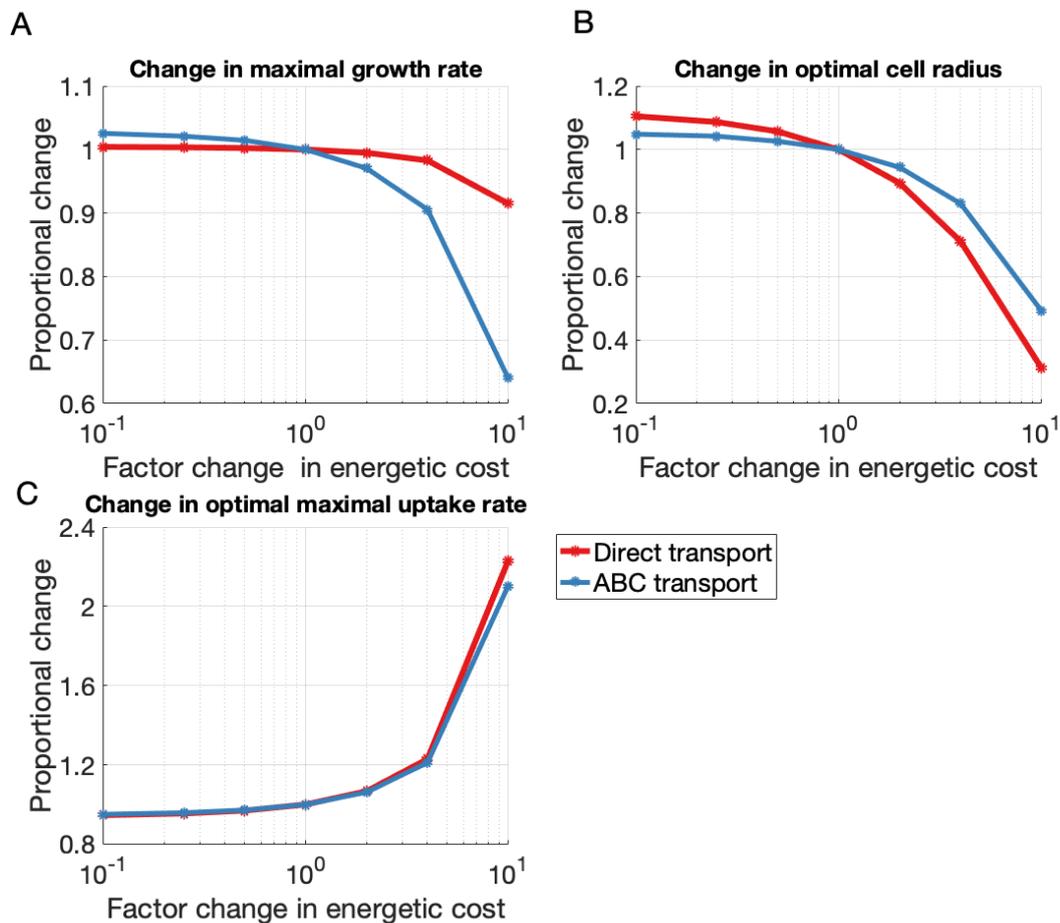


FIGURE 4.4: **Sensitivity analysis of direct and ABC transport systems at saturation, $[S]_{ext} = 100$ mM, over various energetic costs.** The energetic costs were modified by multiplying the number of ATP consumed per substrate translocated by the factor specified on the x-axis.

itself, they diffuse much more slowly. If I assume that the diffusivities of substrates are similar in the periplasm and cytoplasm, a binding protein with a molecular weight of 42.5 kDa (the weight of the maltose binding protein, MalE, in *E. coli*) will have a diffusivity of approximately $2 \mu\text{m}^2/\text{sec}$, while a substrate with a molecular weight of 0.18 kDa (the weight of glucose) will have a diffusivity of approximately $200 \mu\text{m}^2/\text{sec}$ [201]. Therefore, as the association rate k_2 and the turnover rate k_3 may both depend on the diffusivity of the compounds that bind to the membrane-bound transport units, I assume for the remainder of my analysis of the baseline model that these rates are 100 times smaller for ABC transport than for direct transport: $k_2 = 0.01k_2'$ and $k_3 = 0.01k_3'$.

Using these baseline rates, my model predicts that, while direct transport systems

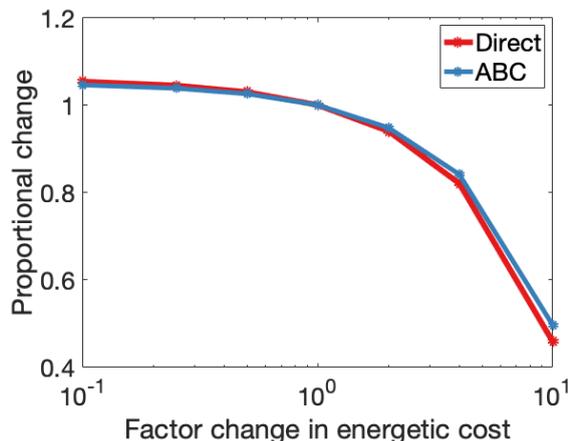


FIGURE 4.5: **Sensitivity analysis of direct and ABC transport systems at $[S]_{ext} = 1$ nM over various energetic costs.** While the optimal cell radii remains nearly constant over factor changes of ten for both ABC and direct transport systems, the energetic costs of transport affect the maximal growth rates.

achieve higher uptake rates per transport proteomic cost at high nutrient concentrations, ABC transport systems achieve higher uptake rates per transport proteomic cost at low nutrient concentrations because they can increase specific affinity at a lower proteomic cost than direct transport (Figure 4.6). Thus, my model suggests that the main opportunity cost between ABC and direct transport systems arises not from variable energetic costs but from a rate-affinity trade-off. While direct transport systems can achieve higher rates per proteomic cost, ABC transport systems can achieve higher affinities per proteomic cost.

4.3.2 Cell size: a master trait

Interestingly, optimization results from my model suggest that the higher turnover rates of direct transport systems support larger cells as they result in smaller surface area-to-volume ratios (Figures 4.6B & 4.7). Therefore, the high turnover rates of direct transport systems may be of particular importance to motile copiotrophs that must be large enough to overcome rotational diffusion and swim effectively toward nutrient hotspots [134].

The effects of cell surface area-to-volume ratio on growth rates is made particularly apparent when I run the optimization problem while constraining cell radii to particular values (Figure 4.8). Smaller cell radii, i.e., higher surface area-to-volume ratios, increase

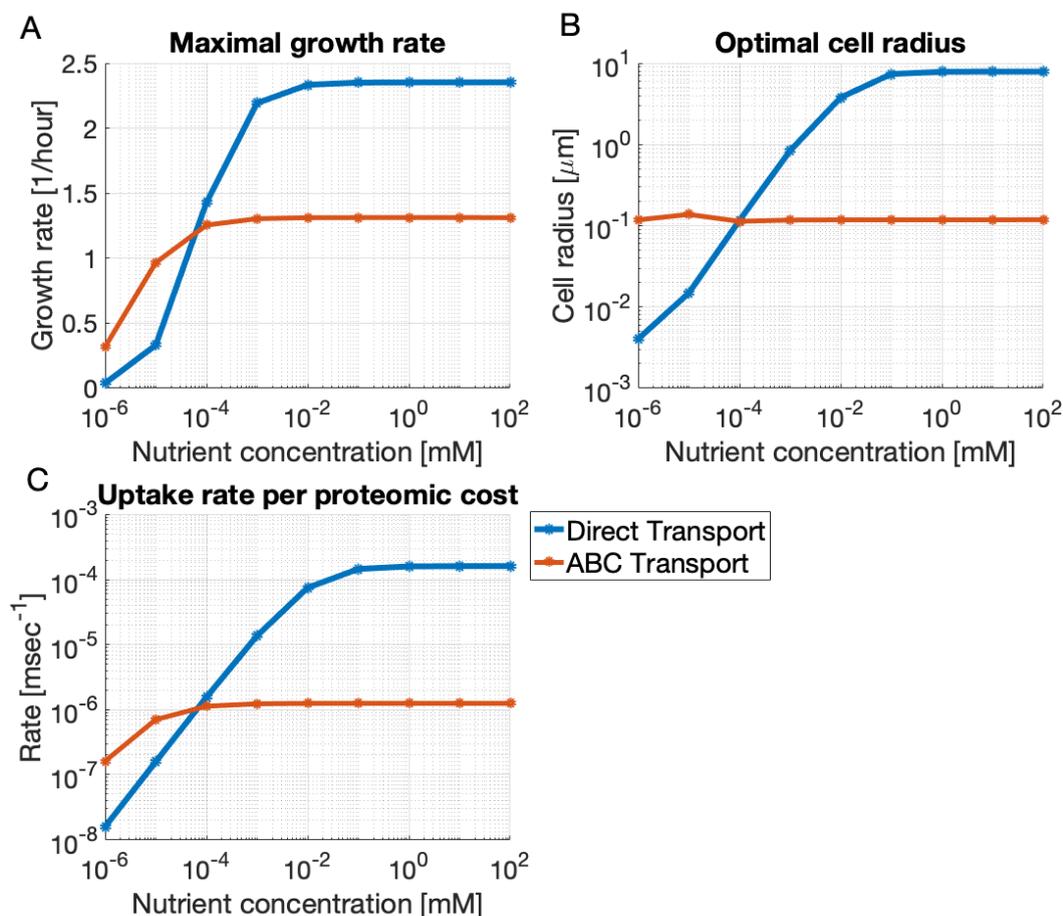


FIGURE 4.6: **A rate-affinity trade-off.** This figure replicates Figure 4.3, focusing on only the baseline models of direct versus ABC transport.

specific affinities and thus growth rates at low nutrient concentrations (Equation 4.5 and Figures 4.8A&B). However, increasing the surface area-to-volume ratio does not increase uptake rates at high nutrient concentrations, where periplasmic uptake rates are not limited by the diffusion of substrates to the cell surface (Figures 4.8C&D). Instead, at high nutrient concentrations, the density and surface area “real estate” constraints limit growth. Therefore, optimal cell radii at substrate saturation scale inversely with maximal allowed cytoplasmic densities; and proportionally with the fraction of the surface area of the inner membrane available for the transport units and electron transport chain units (Figure 4.9).

Although I did not attempt to fine tune my model to ensure realistic cell sizes, I did use realistic parameters for all quantities considered in the above-mentioned constraints.

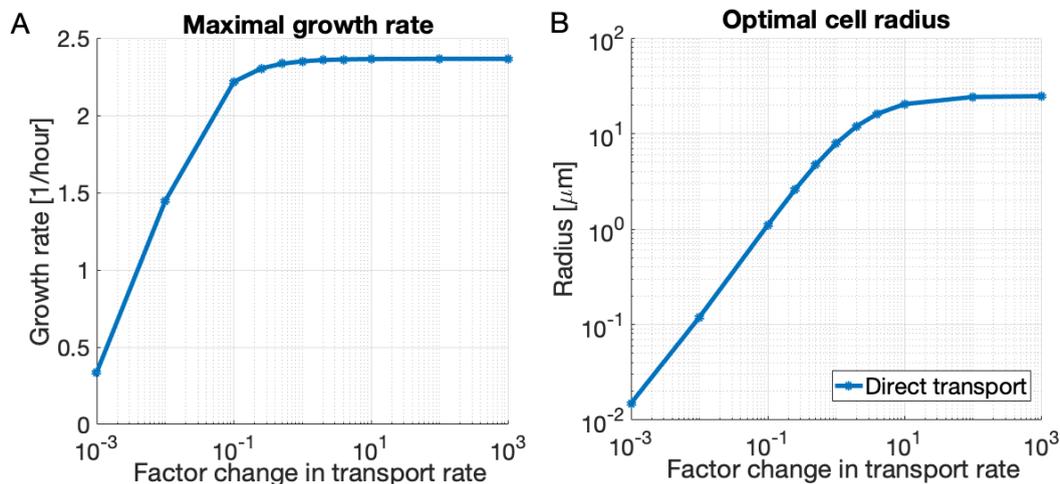


FIGURE 4.7: **Sensitivity analysis of proteome allocation for direct transport to changes in transport rate k_3 at $[S]_{ext} = 100$ mM.** Increases in the transport rate result in (A) higher achievable growth rates and (B) larger optimal cell radii.

Thus, it is worth noting that my model predicts minimal surface area-to-volume ratios that are a factor of 10 smaller than those attained by fast-growing *E. coli* [74]. This may suggest that my model does not consider additional important factors for bacterial survival that constrain cell size. Possible unmodeled factors include high basal energetic requirements [197], predation [220], and virulence [212]. Excitingly, however, further constraining cell size in my model down from a radius of 10 μm to 0.1 μm does not substantially change maximal growth rates and optimal solutions (Figure 4.8C).

4.3.3 Matching transport to catabolic capacity

For cells relying on direct transport, my work replicates results [138] that show that, as extracellular carbon concentration decreases, the optimal transport unit proteome fraction increases and the optimal downstream enzyme proteome fraction (comprising the catabolic enzymes, energy generating proteins, and ADP synthesis proteins) decreases (Figure 4.10D). Therefore, at low nutrient concentrations, the optimal maximal uptake rate is much larger than the optimal maximal rate of utilization of the substrate downstream (Figure 4.10C). Hence, my model predicts a potentially dangerous mismatch between the capacity of the transport system and the capacity of downstream enzymes for

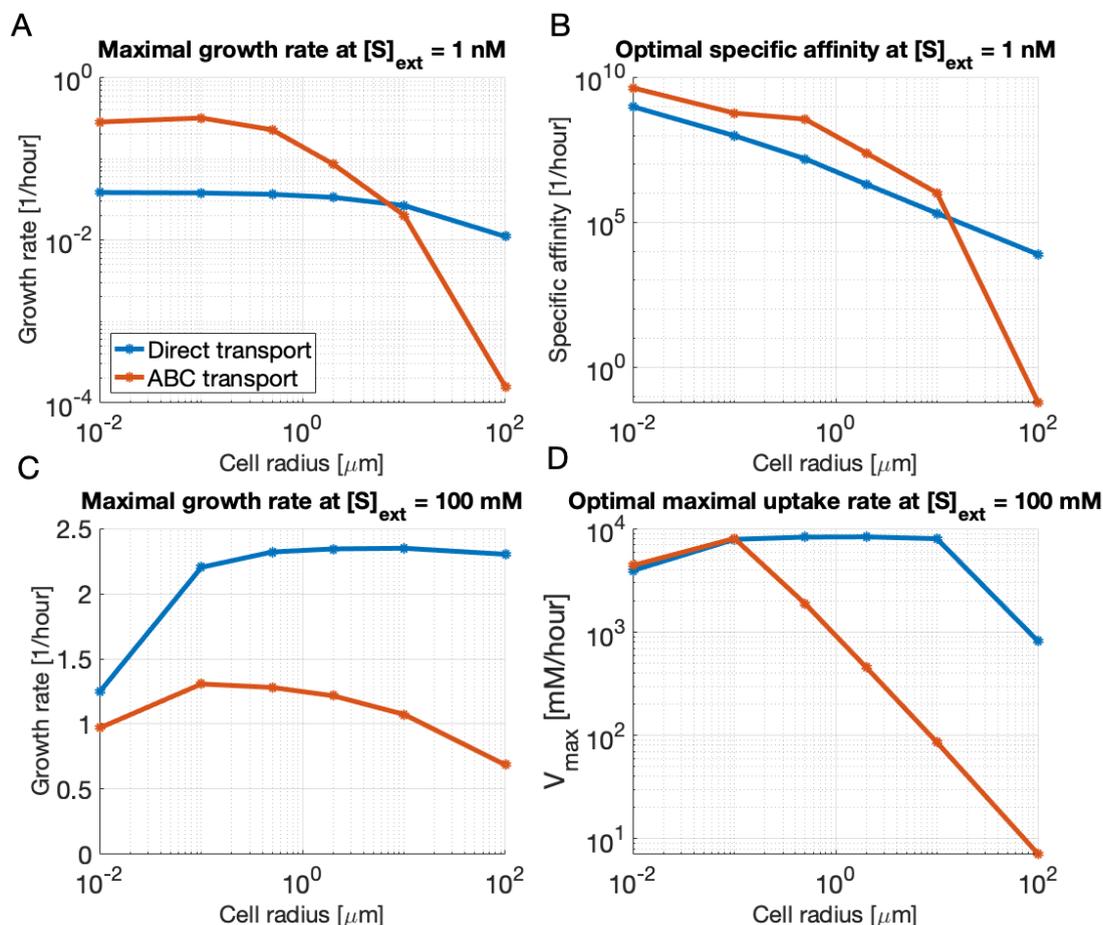


FIGURE 4.8: **Effects of cell radius on maximal growth rate.** (A) and (B): the proteome allocation problem is solved at $[S]_{\text{ext}} = 1 \text{ nM}$ while constraining the cell radius to particular values (x-axis). Because periplasmic uptake is limited by diffusion at low nutrient concentrations, specific affinity increases with increasing surface area-to-volume ratios and thus decreasing cell radii. (C) and (D): the proteome allocation problem is solved at $[S]_{\text{ext}} = 100 \text{ mM}$ while constraining the cell radius to particular values. The maximal growth rates of direct transport systems are not very sensitive to cell radius over a range of $0.1 \mu\text{m}$ to $100 \mu\text{m}$.

cells relying on direct transport systems and optimized for growth at low carbon concentrations. This mismatch is dangerous if the cell cannot excrete consumed substrate because a sudden nutrient up-shift would overwhelm the cell with substrate that it cannot metabolize.

On the other hand, for cells relying on ABC transport, the optimal maximal uptake rate in fact decreases as nutrient concentration decreases (Figure 4.10A). Therefore, my model predicts that the transport capacity of cells relying on ABC transport more closely

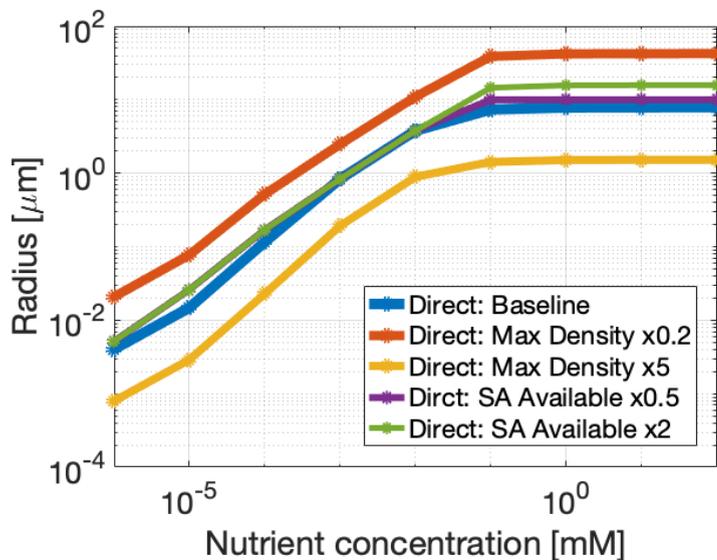


FIGURE 4.9: **Active constraints on cell radius at high nutrient concentrations.** Both the surface area “real estate” constraints and the density constraints are active for the direct transport proteome allocation problem at high nutrient concentrations. Increases in maximal allowed density result in smaller optimal cell radii (red and yellow). Increases in the fraction of the surface area available to the membrane-bound protein groups result in larger optimal cell radii (purple and green). Note that, for the baseline model, I assume that half of the surface area of the inner membrane is available.

matches the capacity of their downstream enzymes (Figure 4.10C). On the other hand, Button’s theory, still prevalently used to explain oligotrophic strategy for nutrient acquisition [63], predicts that oligotrophs have a very high ratio of membrane-bound transport unit to downstream enzyme to achieve high specific affinity [39]. My model suggests that this would only be the case if oligotrophs relied on direct transport systems. Instead, my model suggests that the oligotroph’s success is a result of its ability to achieve high specific affinity without expressing such high levels of membrane-bound transport units.

Therefore, in contrast with Button’s theory, I predict that the model oligotroph does not have any excretion systems because it would rarely have need for them and because they may decrease the efficiency of the cell’s uptake [33]. However, I predict that the model copiotroph is likely to excrete metabolic intermediates when it has acclimated to a low nutrient concentration and is introduced to an environment with higher concentrations.

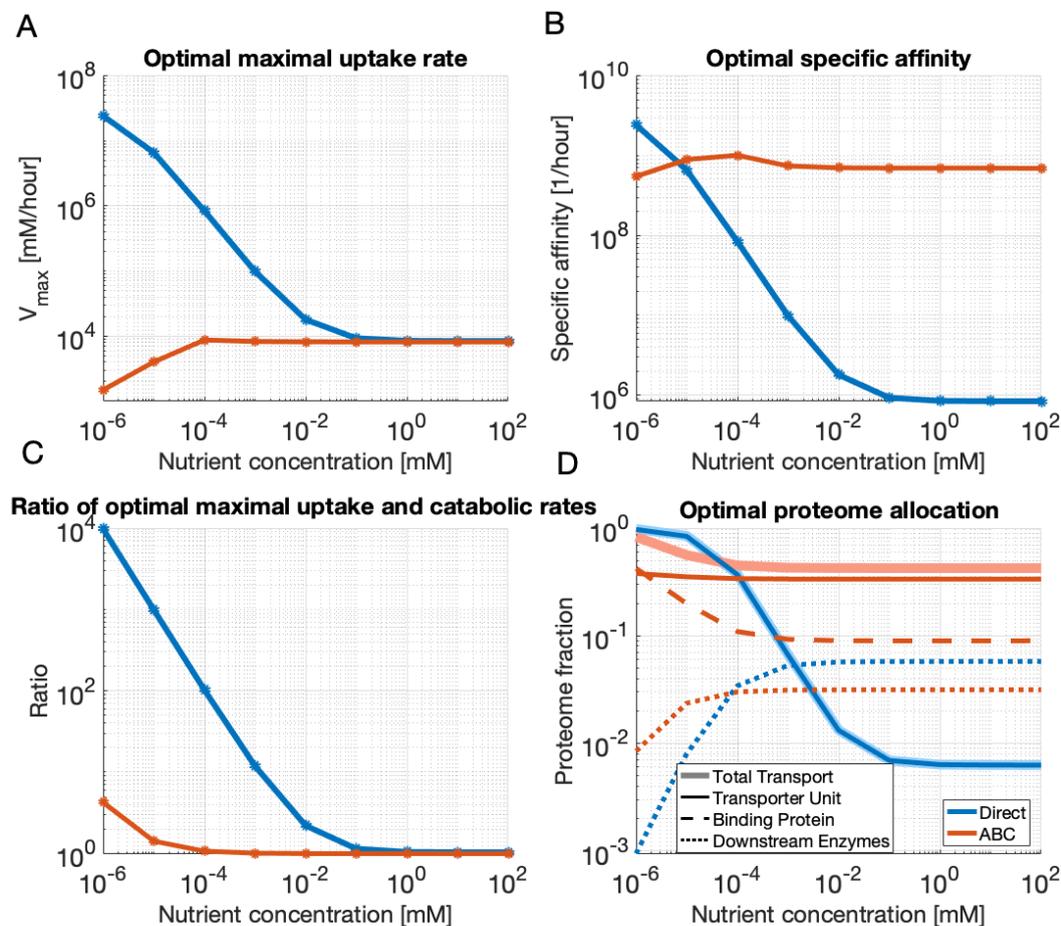


FIGURE 4.10: Optimal proteome allocation for direct versus ABC transport systems. As the extracellular nutrient concentration decreases, the optimal maximal uptake of transport increases for direct transport systems but decreases for ABC transport systems (A), while optimal specific affinities increase for both (B). (C) This results in a large ratio of optimal maximal uptake and maximal downstream enzyme rates for direct transport systems, while it is optimal for ABC transport systems to keep this ratio near one. I calculate maximal downstream enzyme rate by summing the maximal reaction rates of the components that consume the intracellular carbon substrate; see Supplemental Section 2.1. (D) While it is optimal for cells relying on both transport systems to devote nearly all of their proteome to transport at low nutrient concentrations, for ABC transport systems, it is the proteome fraction of the binding proteins that increase and not the fraction of transport units.

4.3.4 The cost of high specific affinity

Although the optimal maximal uptake rate decreases with decreasing nutrient concentration for cells relying on ABC transport, my model still predicts that the cell is using the majority of its proteome for transport at low nutrient concentrations (Figure 4.10D).

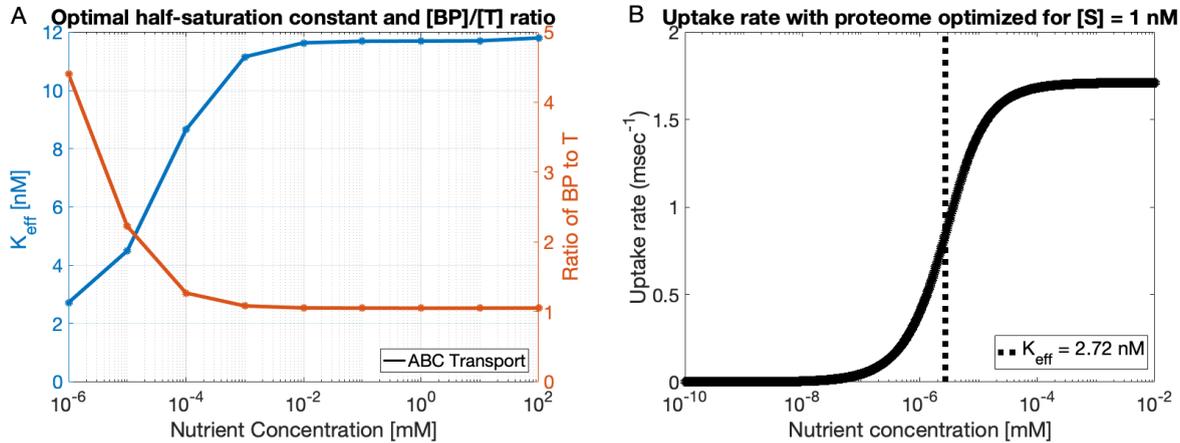


FIGURE 4.11: **The effective half-saturation concentration of ABC transport.** (A) In my model, ABC transport systems achieve lower optimal half-saturation concentrations at low nutrient concentrations by increasing the ratio of the abundance of binding proteins to abundance of membrane-bound transport units. (B) To calculate the effective half-saturation concentration of a solution to a particular proteome allocation, I use Equation 4.5 and Equations 4.16-4.20 to determine the uptake rate over a range of nutrient concentrations (x-axis). Here, I show calculated uptake rates over various nutrient concentrations for the proteome allocation obtained when optimizing the cell for growth at an extracellular concentration of $[S]_{ext} = 1$ nM.

However, instead of the majority of that proteome fraction being devoted to membrane-bound, energy-consuming transport units, it is instead devoted to binding proteins. This is because increasing the concentration of binding proteins in the periplasm increases the specific affinity of transport. Therefore, in my model, when the extracellular carbon concentration is 1 nM, the cell uses a binding protein with dissociation constant $K_{BP} = 1$ μM and a transport unit with dissociation constant $K_T = 10$ μM to achieve an optimal effective half-saturation concentration of $K_m = 2.5$ nM (Figure 4.11), which is four thousand times smaller than the half-saturation constant of direct transport in my model, $K_{eff'} = K_{T'} = 10$ μM .

My model suggests that the proteomic cost of achieving this high specific affinity is substantial. In my model, while a cell relying on direct transport devotes only about 1% of its proteome for transport to attain maximal growth rates at saturation, it devotes 96% of its proteome at an extracellular nutrient concentration of 1 nM to attain even higher optimal specific affinities than are optimal for ABC transport systems (Figure 4.10B). My model thus suggests that the cell relying on ABC transport outperforms cells relying on direct transport at low nutrient concentrations because it can achieve

higher specific affinities per proteomic cost (Figure 4.6C). Yet, in my model, the cell with ABC transport still devotes 84% of its proteome for transport to achieve maximal growth rates in the nanomolar regime (Figure 4.10D). In this regime, about half of the transport proteome (43% of the total proteome) is devoted to the binding proteins. This corresponds to a binding protein to transport unit ratio of five (Figure 4.11A). However, this ratio is estimated to exceed fifty in *E. coli*'s maltose ABC transport system [47, 156]. A sensitivity analysis does show that this ratio is sensitive to the value of the binding protein dissociation constant, and, for example, does reach fifty when the dissociation binding rate, k_{1r} , is a factor of one hundred smaller (Figure 4.13D&E).

As the binding protein to transport unit ratio quantifies the additional proteomic cost that the binding proteins incur to achieve high specific affinities, it would be interesting to obtain proteomic data to determine what this ratio is for model oligotrophs consuming various substrates. Because oligotrophs use ABC transport systems to consume solutes that are smaller than maltose, oligotrophs' binding proteins may in fact have smaller dissociation constants so that perhaps the optimal ratio of binding proteins to transport units is indeed nearer to one. However, it is also possible that oligotrophs can achieve even lower effective half-saturation concentrations that may require higher ratios of binding proteins and thus incur even greater proteomic costs. For this reason, it would also be insightful to determine whether oligotrophs can regulate binding protein abundances to modify this ratio over environmental conditions. If they can, it suggests that the costs of expressing such high levels of binding proteins is greater than the costs of regulating their expression.

4.3.5 Effects of solute size

For carbon-cycling models to better predict how bacterial abundances, uptake rates, and growth rates depend on environmental conditions, these models must quantify the quality of the carbon substrates in the environment. My model suggests that one important characteristic of environmental substrate quality is the size of the solute, and a sensitivity analysis of my model helps us gain insight into how solute size may affect growth rates. My results suggest that solute size is of particular importance for predicting cell sizes and growth rates at low nutrient concentrations.

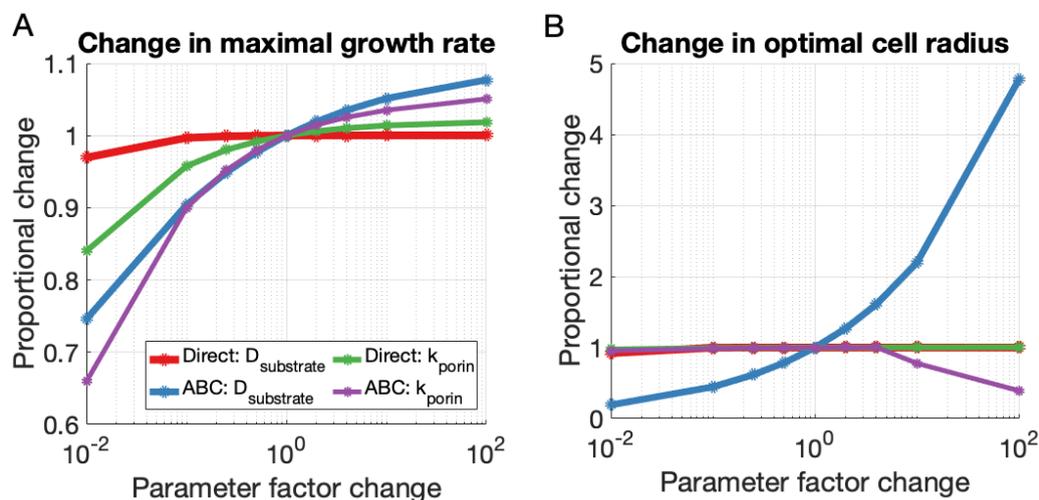


FIGURE 4.12: Sensitivity analysis of direct and ABC transport systems at $[S]_{\text{ext}} = 1$ nM over various periplasmic uptake values.

Periplasmic uptake rates at low nutrient concentrations are dependent on the diffusivity of the solute in the external environment, which scales inversely with size. Thus, as cell surfaces more easily become depleted of solutes with lower diffusivities, larger solutes result in smaller optimal cell radii and smaller maximal growth rates for cells relying on ABC transport (Figure 4.12). (The effects of solute diffusivity on maximal growth rates is much smaller for direct transport using my model’s baseline parameters simply because the optimal cell size is already so small and growth-limiting.) Furthermore, translocation rates of the solutes through general porins depend greatly on the ratio of the effective solute radius to porin radius [147, 164]. Thus, larger solutes also result in smaller values of k_{porin} and further reductions in uptake rates and maximal growth rates (Figure 4.12).

Finally, again due to the smaller diffusion constants of larger solutes, I expect the association and dissociation rates of the solute and binding protein for ABC transport systems (k_{1f} , k_{1r}) as well as the association and dissociation rates of the solute to the membrane-bound transporters for direct transport systems (k_{2f} , k_{3f}) to depend on solute size. And, my sensitivity analysis, in fact, suggests that maximal growth rates and optimal cell sizes are most sensitive to these transport rates at low nutrient concentrations (Figures 4.13A&B). For ABC transport, a reduction in association rate k_{1f} by a factor of ten causes the optimal cell radius to increase by 25% to accommodate greater binding protein abundances; and an 80% reduction in maximal growth rate. Because direct transport systems use even greater fractions of their proteome for transport at low nutrient

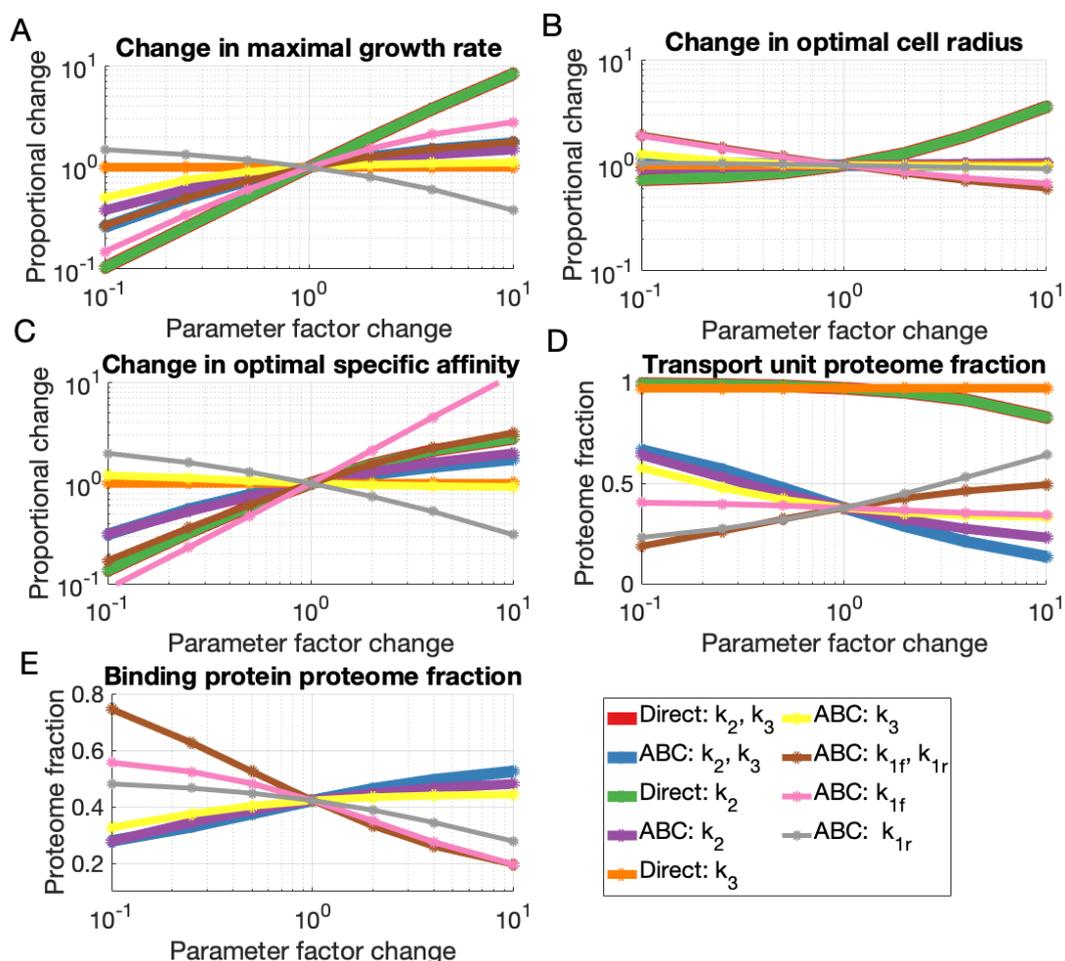


FIGURE 4.13: Sensitivity analysis of direct and ABC transport systems at $[S]_{ext} = 1$ nM over various cytoplasmic transport rates. For ABC transport, I expect the binding-protein association and dissociation rates, k_{1f} and k_{1r} , to be most sensitive to solute size. For direct transport, I expect the association rate of the substrate to the transport unit, k_2 , as well as the translocation rate, k_3 , to be sensitive to solute size. Maximal growth rates for ABC transport are sensitive to both changes in k_{1r} and k_{1f} , while maximal growth rates for direct transport are most sensitive to changes in k_2 .

concentrations, their corresponding maximal growth rates are even more sensitive to changes in association rate k_2 : in my model, a factor of ten reduction in k_2 caused a 90% reduction in maximal growth rate at an external solute concentration of 1 nM.

In contrast, cells relying on direct transport systems are not sensitive to changes in transport rates at saturating levels of nutrients in my model because the limit to maximal growth rate is in the rates of the downstream enzymes so very little of the cell's proteome is devoted to transport (Figure 4.14). Therefore, I do not expect solute size to be an

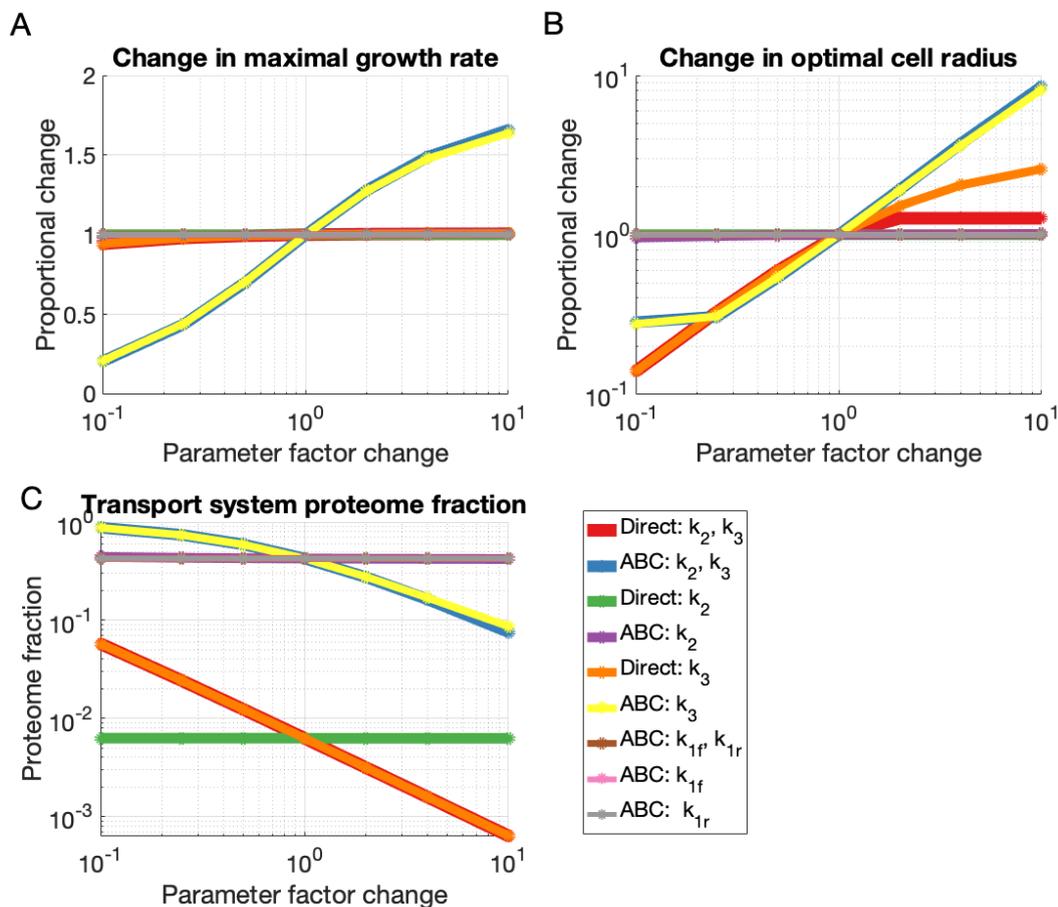


FIGURE 4.14: **Sensitivity analysis of direct and ABC transport systems at saturation over various cytoplasmic transport rates.** Direct transport systems devote so little of their proteome to transport at high nutrient concentrations that modifications in transport rate can be compensated by small shifts in transport proteome fractions. However, the maximal growth rates of ABC transport systems are sensitive to changes in translocation rate, k_3 .

important predictor of growth rate for copiotrophs at high nutrient concentrations. On the other hand, a cell relying on ABC transport is sensitive to changes in the translocation rate of the solute through the inner membrane, k_3 , at saturating nutrient concentrations because, in my model, the cell uses a substantial fraction of its proteome for transport to achieve maximal growth rates at saturation. However, although I expect the rate k_3 to be smaller for larger solutes, I do not expect it to be as sensitive to solute size as the rates k_{1f} and k_{1r} .

My sensitivity analyses suggest that comparing the sizes of consumed solutes may

give insight into the relative growth rates of cells on various compounds. In particular, I hypothesize that solute size is an important factor to consider when predicting growth rates of cells relying on ABC transport. I hypothesize that it may also be important for predicting growth rates of cells relying on direct transport at low nutrient concentrations. I therefore suggest quantifying solute size when considering substrate quality in carbon-cycling models that consider bacterial dynamics. However, one gap in our understanding is of how precisely the transport rates vary with solute size. While previous analyses of periplasmic uptake rates as well as my analysis of the diffusivity of binding proteins versus carbon solutes suggest that these rates could vary substantially with solute size, these differences, to my knowledge, have not been quantified. Therefore, experimental data that relates solute size to oligotrophs' cell sizes and growth rates may be useful for building more predictive ecosystem models.

4.4 Conclusions

This work uses a toy model of cellular metabolism that incorporates the effects of variable surface area-to-volume ratios as well as biophysical transport models into a proteome allocation problem that predicts the cell size and transport phenotype that maximizes growth rate given a specific extracellular carbon substrate concentration. The model's predictions demonstrate how cells can efficiently allocate proteome not only to increase their maximal uptake rates but also to increase their effective specific affinities for transport. While much previous work has considered the effective half-saturation concentration of uptake to be a constant, intrinsic property of the transport proteins, my model of ABC transport shows how a cell can, in fact, regulate binding protein expression to tune its effective half-saturation concentration based on its external environment. Therefore, in my model, a cell using an ABC transport system with a binding protein dissociation constant of $K_{BP} = 1 \mu\text{M}$ and a transporter dissociation constant of $K_T = 10 \mu\text{M}$ can achieve low effective half-saturation concentrations of $K_m \approx 1 \text{ nM}$, while the cell cannot regulate the half-saturation concentration of a direct transport system, $K_{m'} = K_{T'}$.

Therefore, this work suggests that oligotrophs rely on ABC transport systems to efficiently achieve high specific affinities. Button hypothesized that oligotrophs achieve higher specific affinities by having higher permease to downstream enzyme ratios [39].

However, this theory only aligns with my model’s predictions for how cells relying on direct transport may achieve high specific affinity. My results suggest that achieving high specific affinities in this way is very costly for the cell because of both the high proteomic cost as well as the surface area “real estate” requirements of the membrane-bound transport units. Hence, my work instead shows that cells can achieve higher specific affinities per cost using ABC transport systems. Additional binding proteins have lower proteomic costs than additional transport units and also do not incur “real estate” costs on the inner membrane. Instead, in my model, binding protein abundance is subject only to a constraint on the periplasmic density, which the cell can mitigate by modifying the fraction of its volume devoted to its periplasm. My results thus suggest that a majority of an oligotroph’s proteome could be comprised of binding proteins, and this aligns with metaproteomic analysis that found that binding proteins are among the most prevalent bacterial protein found in the oligotrophic ocean [189].

However, I also predict that ABC transport systems have lower maximal rates than direct transport systems because of the slow diffusion of binding proteins in the periplasm. Therefore, I predict that a rate-affinity trade-off is central to the copiotrophic and oligotrophic divide: copiotrophs rely on direct transport systems to efficiently achieve high maximal uptake rates in nutrient-rich environments, while oligotrophs rely on ABC transport systems to efficiently achieve high specific affinities in nutrient-limited environments. Because higher uptake rates support smaller surface area-to-volume ratios, this rate-affinity trade-off is also consistent with observations that copiotrophs are much larger than oligotrophs.

This size difference may also explain why a “superbug” could not exist that dominates in both copiotrophic and oligotrophic environments. This purported “superbug” could rely on direct transport in carbon-rich environments and ABC transport in carbon-limited environments. I hypothesize that this superbug does not exist because of constraints on cell size as well as the regulatory overhead needed to switch between transport systems. Oligotrophs’ high specific affinity is a product not only of their reliance on ABC transport systems but also their small cell size. Furthermore, their higher growth rates in nutrient-limited environments are also due to their highly streamlined genomes, which require not only less space but also less regulation and sensing—each of which add further genomic, proteomic, and energetic costs [64, 78, 178]. I therefore hypothesize that, due to the large genome size and regulatory requirements of the “superbug”, the “superbug” is unable

to achieve the small cell size and low proteomic and energetic requirements necessary to outcompete oligotrophs at low nutrient concentrations.

But how do genome size requirements affect cell size? And what precisely are the proteomic and energetic requirements of sensing and regulation? The ubiquity of cells of the SAR11 clade suggests that the proteomic and energetic costs of copiotrophic plasticity are high. For future work, the mechanistic model of cell metabolism that I developed in this work could be extended to include these additional genomic and regulatory costs to thus provide greater insight into heterotrophic bacterial speciation in the ocean.

The work in this chapter was done with guidance from Naomi Levine, Vicente Fernandez, and Roman Stocker. I would also like to acknowledge helpful feedback from Terry Hwa and Cameron Thrash and helpful edits by Levine and Fernandez.

Chapter 5

To Swim, or Not to Swim

In the previous chapter, I gained insight into the speciation and coexistence of marine heterotrophic bacteria by differentiating oligotrophs and copiotrophs in terms of their transport mechanisms. I hypothesized that central to copiotrophs' and oligotrophs' differentiation is a rate-affinity trade-off: While oligotrophs rely on ABC transporters to increase nutrient affinity and thus uptake rate at low nutrient concentrations, copiotrophs rely on direct transporters to maximize maximal uptake rates at high nutrient concentrations. In this chapter, I instead attempt to gain insight into their differentiation by distinguishing oligotrophs and copiotrophs by their most apparent difference: whether or not they can swim. Genomic analyses show that genes for flagellar synthesis are highly overrepresented in copiotrophs, suggesting that copiotrophs must swim to access the high nutrient concentrations that support their fast growth rates [115].

While previous calculations suggest that motility is a costly trait due to both its high energetic and proteomic requirements [134, 135, 198], previous experimental work has shown motility to be highly beneficial for accessing higher nutrient levels due to spatial heterogeneities in the microbial marine environment. Motile marine bacteria exhibit chemotaxis to better accumulate in regions with elevated levels of dissolved organic matter, for example in the vicinity of phytoplankton [27, 182, 193, 194]. Furthermore, motile cells more easily colonize nutrient-rich particles, such as marine snow and fecal pellets [34, 69, 100, 193]. Motility greatly increases the likelihood that a cell will encounter and colonize sinking particulate organic carbon not only because of the cell's increased diffusivity and ability to perform chemotaxis but because, in addition, nonmotile cells are

unable to cross the streamlines of laminar flow of a sinking particle to reach the particle's surface. Previous work has thus shown that the majority of particle-attached marine bacteria are capable of swimming [99]. In addition, many model copiotrophs that can swim, such as *Vibrio* species, can form biofilms or use pili to attach to particles [95, 125], giving further credence to our understanding that many marine bacteria use motility to find and access *particulate* carbon sources.

In this work, I focus on modeling the benefits of motility for particle colonization because copiotrophs, in contrast to oligotrophs, are able to degrade particulate organic matter [18, 66, 115]. Therefore, gaining insight into the environments that support motility for particle colonization can help us predict the degradation rates of particulate carbon in the ocean to improve marine carbon-cycling models [56, 102, 144].

My assumption in this work that the costs and benefits of motility are due to its use for particle colonization differentiates my work from previous work on the trade-offs of motility in marine bacteria. Previous work has focused on understanding the costs of motility for the uptake of dissolved organic matter and thus considers motile cells to always be swimming; see, for example, [198]. However, in my baseline model of motility, I assume that cells only swim between particle encounters and thus do not expend energy on motility while attached to particles. While very little is known about how marine motile bacteria regulate flagellar synthesis and how cells with expressed flagella regulate motility and speed [188], it is known that cells with expressed flagella can inhibit motility entirely [72]. Experimental results suggest that motility is inhibited on particle surfaces not only because of the energetic costs of swimming but also because its inhibition is a necessary requisite for successful adherence to the surface [72].

In this chapter, I gain insight into the costs and benefits of motility for particle colonization by incorporating a simple model of motility into the metabolic model and proteome allocation problem that I developed in the previous chapter. I model the benefit of motility by quantifying how motility increases particle encounter rates. I assume that motility requires a constant proteome fraction and thus focus on the energetic cost for a cell to swim when not attached to particles. As this work is more exploratory, I study the effects that individual modifications to my model of motility have on my predictions on the optimal cell size and swimming speed given a particular concentration of particles and concentration of nutrients on the particles. However, independent of the modification,

the model consistently predicts a ‘wrong result’: that it is optimal for a motile cell to rotate its flagellum at the maximal rate when it does not sense any nutrient gradients. Thus, I conclude by hypothesizing that it is uncertainty due to temporal heterogeneity that may govern optimal swimming speeds.

5.1 A toy model of motility

5.1.1 The benefit of swimming

Observations suggest that the majority of marine heterotrophic bacteria that swim in the water column have just a single polar flagellum with which they perform a “run-and-reverse” swimming strategy [91]. The cell alternates between rotating its flagellum in one direction to propel the cell body forward (a ‘run’) and rotating it in the opposite direction to pull the cell body backwards (to ‘reverse’). In between a run and a reversal, the cell reorients. Reorientation may be due to Brownian rotational diffusion in some species [193]. However, it is known that various motile marine bacteria actively reorient [15]. In particular, *Vibrio alginolyticus* cells actively reorient by an angle Gaussian distributed about 90 degrees using a ‘flick’, which is due to a mechanical instability in the flagellum that causes it to buckle [183]. To perform chemotaxis, cells bias this run-reverse random walk by increasing the amount of time spent swimming in directions that are up nutrient gradients [153, 214].

Previous work has found that marine motile cells that are capable of active reorientation for performing chemotaxis are better able to colonize particles than motile cells that cannot [99]. This suggests that chemotaxis is crucial for effective motility. Because sinking particles create plumes of dissolved organic matter, reorientations can both help the cell remain in the elevated nutrient concentrations of the plume [194] as well as help the cell follow the plume to the sinking particle, assuming that the cell can swim faster than the particle is sinking. Indeed, experiments have found that many marine motile bacteria increase their swimming speeds when they encounter bands of nutrients, perhaps because they are trying to chase the dissolved nutrients to their particulate source [137].

However, in my baseline toy model, I neither model the plumes of sinking particles nor do I explicitly model run-reverse chemotaxis. Instead, I only implicitly model chemotaxis

by assuming that a cell always successfully attaches to a particle once it encounters the particle. Variations in chemotactic ability can thus be modeled by modifying the effective size of the particle: for cells with chemotactic ability that are more sensitive to the nutrient gradient created by the particle, the effective size of the particle is larger.

I instead assume that the benefit of motility is that it increases encounters with nutrient-rich particles. My model of motility thus captures the benefit of swimming in environments in which the cell does *not* sense any gradients and thus cannot perform chemotaxis.

5.1.1.1 Effects of cell's phenotype on particle encounter rate

I assume that the time a cell spends in the background environment before encountering a particle, τ_{env} , is the reciprocal of the product of an encounter rate kernel, B , and the concentration of particles in the environment, C_p [99]:

$$\tau_{env} = \frac{1}{BC_p}. \quad (5.1)$$

I relate the encounter rate kernel B to the effective translational diffusivity of the cell, D_t , using the equation for the diffusive flux toward a sphere. Thus, I calculate the encounter rate per cell and per particle of an ensemble of uniformly distributed cells diffusing with effective diffusion constant D_t toward uniformly distributed spherical particles [50, 99]. Thus,

$$B = 4\pi aSD_t \left[\frac{\text{length}^3}{\text{time}} \right], \quad (5.2)$$

where a is the effective radius of the particle, which can be larger than the particle itself to model chemotaxis; and S is the Sherwood number, which describes the ratio of mass transfer by convection and mass transfer by diffusion and depends on the size and sinking speed of the particle. (In this work, I assume that the Sherwood number is $S = 2$, corresponding to the Sherwood number estimated for a particle with a radius of 200 μm and sinking at a speed of 60 $\mu\text{m}/\text{sec}$, i.e., 5 m/day, and assuming a diffusivity typical for molecules, $10^{-5} \text{ cm}^2/\text{sec}$ [9, 101].)

For nonmotile cells, the effective translational diffusivity is due solely to Brownian motion and thus decreases with increasing cell size [50]:

$$D_{t,nonmotile} = \frac{k_B T}{6\pi\eta r} \left[\frac{\text{length}^2}{\text{time}} \right], \quad (5.3)$$

where r is the radius of the cell body, which I assume to be spherical; k_B is Boltzmann's constant, T is absolute temperature, and η is dynamic viscosity.

For a motile cell, I assume that the effective diffusivity is [50, p. 92]:

$$D_{t,motile} = \frac{v^2}{6D_R} \left[\frac{\text{length}^2}{\text{time}} \right], \quad (5.4)$$

where v is the swimming speed of the bacterium and D_R is the cell's rotational diffusivity. Thus, a motile cell should attempt to minimize its rotational diffusivity to maximize translational diffusivity and thus particle encounter rates. Because active reorientation can only increase a cell's rotational diffusivity, for the baseline model I assume that cells only reorient due to Brownian motion. This is equivalent to assuming that cells that actively reorient to perform chemotaxis reorient at rates equal to that of or slower than the effective rotational rate due to Brownian motion when they do not sense any nutrient gradients. Therefore, the rotational diffusivity term that I use in my baseline model captures the cell's size requirement for overcoming rotational Brownian motion to swim more straight and efficiently [50, p. 176]:

$$D_r = \frac{k_B T}{8\pi\eta r^3} \left[\frac{\text{radians}}{\text{time}} \right]. \quad (5.5)$$

Hence, in my model, while a nonmotile cell can increase its encounter rate by decreasing its size, a motile cell's encounter rate is proportional to its volume as well as to the square of its swimming speed.

It is important to note that in my metabolic model from the previous chapter, the proteome allocation problem optimized cell radius because of the effects of variable surface area-to-volume ratios on both nutrient uptake and cellular respiration. However, in my encounter rate model above, it is now the length or volume of the cell that is important. Therefore, future work should consider the role of cell shape for mitigating trade-offs between increasing length or volume and decreasing surface area-to-volume ratio [220].

5.1.1.2 Modeling the nutrient landscape

As seen from Equation 5.2, a cell's ability to encounter a particle depends on the particle's radius, a . While the size of organic particles can range over many magnitudes [8, 88], in my baseline model I assume that a particle is a sphere with radius $a = 200 \mu\text{m}$, which is a representative size of a small marine snow particle [100].

It has been estimated that organic particles can contain concentrations of carbon that exceed the background concentration of carbon by 2 to 4 orders of magnitude [8, 193]. In my model, I assume a constant background carbon substrate concentration of $[\text{S}]_{env} = 1 \text{ nM}$ and modify the particle carbon substrate concentration of $[\text{S}]_p > [\text{S}]_{env}$.

I also modify the concentration of particles in the water, C_p . While I in fact calculate an ensemble average encounter rate, I assume that a cell encounters a new particle exactly in time interval, τ_{env} . Thus, there is no stochasticity in this model, which, in effect, assumes both that all particles are located in a precise grid with spacing $d = \frac{1}{C_p^{1/3}}$ and that a cell's diffusive, random swimming in this grid of precisely located particles results in a deterministic first-hitting time [68].

5.1.2 The costs of swimming

5.1.2.1 The proteomic cost of swimming

Previous work suggests that motility may incur a high proteomic cost: the assembly of the polar flagellum by itself requires tens of genes that are carefully regulated in hierarchical cascades [188]. However, in this work, I focus on the energetic cost of swimming and thus keep the proteomic cost of swimming constant. I assume that all the proteins needed for motility require a constant proteome fraction of $\phi_{swim} = 4\%$ [51], which is incurred both while swimming and while on a particle if the cell is a motile one. Thus, I assume that when a cell divides, each of its daughter cells has a single flagellum. Because growth rate is proportional to the fraction of the proteome devoted to replication (Chapter 4), this addition of a proteomic cost causes a reduction in average growth rate by exactly 4%.

5.1.2.2 The energetic cost of swimming

A transmembrane sodium ion potential gradient drives the flagellar rotary motor of marine bacteria [26, 219]. Because ATP synthase also requires a transmembrane gradient, I equate the power needed for a cell of a certain size to swim with a given speed to a required number of ATP molecules used per second.

According to Stokes' law, the power required for motility is [50]:

$$P = 6\pi\eta rv^2 \quad \left[\frac{\text{erg}}{\text{sec}} \right]. \quad (5.6)$$

I assume that the efficiency of the flagellar motor is $\epsilon = 1\%$ [161]. I convert power to a proton motive force, as derived in [133], and use the fact that about 4 protons are used by ATP synthase to generate one ATP molecule to obtain the following energetic cost of swimming:

$$E_{swim} = \left(\frac{P}{\epsilon} \right) \left(\frac{5 \times 10^{12} \text{ H}^+/\text{s}}{1 \text{ erg/s}} \right) \left(\frac{1 \text{ ATP}}{4 \text{ H}^+} \right) \quad \left[\frac{\# \text{ ATP}}{\text{sec}} \right]. \quad (5.7)$$

This power is achieved via rotation of the flagellar motor so that a cell's swimming speed is determined by the cell's size and the rate of rotation of the flagellar motor. Experiments suggest that the mechanics of the flagellum limit the maximal rotation rate of the flagellar motor to approximately 1,700 revolutions per second for *Vibrio* and that, additionally, approximately 1,000 ions are used per revolution of the motor [26, 219]. Because the energy per ion used is about 0.2×10^{-12} ergs [133], I thus assume the following constraint on motility:

$$E_{swim} \leq E_{max}, \quad (5.8)$$

where

$$E_{max} = \left(1700 \frac{\text{revolutions}}{\text{sec}} \right) \left(1000 \frac{\text{ions}}{\text{revolution}} \right) \left(0.2 \times 10^{-12} \frac{\text{erg}}{\text{ion}} \right) \quad (5.9)$$

$$= 3.4 \times 10^{-7} \frac{\text{erg}}{\text{sec}}. \quad (5.10)$$

5.1.3 A dual proteome allocation problem

I incorporate this simple model of motility into the metabolic model developed in the previous chapter. I assume that the cell can only consume substrate using direct transport and optimize its allocation of two proteomes: one with an associated growth rate μ_{env} for when the cell is in the background environment experiencing a carbon substrate concentration of $[S]_{env}$; and one with an associated growth rate μ_p for when the cell is attached to the particle experiencing a nutrient concentration of $[S]_p$. Thus, for simplicity, I assume that the cell can instantaneously modify its proteome upon attaching to a particle or leaving it and that the substrate on the particle is dissolved organic matter. I therefore do not model the additional costs of degrading particulate organic matter via extracellular enzymes [19, 163, 181].

Additionally, I assume that the cell maintains a constant radius r (which I limit to a maximum radius of 10 μm) while in the background environment and while attached to a particle; and is either nonmotile or swims with a constant swimming speed v when not attached to a particle. While I test various functional forms for the amount of time that a cell may spend on a particle before leaving to encounter a new one (τ_p), I assume for the baseline model that it is equal to the doubling time of the cell on the particle. That is,

$$\tau_p = \frac{\ln(2)}{\mu_p}. \quad (5.11)$$

This is equivalent to assuming that the cell leaves the particle after replicating, perhaps to maximize a rate of dispersal. This functional form is a particularly appealing and intuitive form because it matches my model's assumption that the cell maintains a constant volume while swimming and also a constant surface area-to-volume ratio while growing on the particle, as the cell can elongate while on the particle to maintain its original surface area-to-volume ratio while doubling in volume. Once the cell has doubled in volume, this model assumes that the two daughter cells (one having grown a new flagellum) escape from the sinking particle to encounter a new one.

The allocation problem solves for the cell radius, swimming speed, background proteome allocation, and particle proteome allocation that maximizes the average growth rate:

$$\mu_{opt} = \frac{\tau_p}{\tau_p + \tau_{env}} \mu_p + \frac{\tau_{env}}{\tau_p + \tau_{env}} \mu_{env}. \quad (5.12)$$

To incorporate the energetic cost of motility into my metabolic model, I add the following equality constraint to the dual proteome allocation problem:

$$\tau_{env} E_{swim} = \tau_p v_{futile,p} + \tau_{env} v_{futile,env}, \quad (5.13)$$

where $v_{futile,p}$ ($v_{futile,env}$) is the rate of the futile cycle that converts ATP to ADP (Appendix B.1) while the cell is on the particle (respectively, in the background environment). By using this constraint, I am essentially assuming that storage of the energy required for swimming (i.e., glycogen storage) is lossless and does not take up any space. Therefore, I may be underestimating the energetic cost of motility.

5.2 Results

5.2.1 Optimal swimming speeds

In my simple baseline model of cellular metabolism with motility, both the effective benefit and effective energetic cost of motility scale proportionally with the square of the swimming speed (Equations 5.4 and 5.6). Thus, for motile cells, the additional cost of swimming at a faster speed, in essence, cancels out with the additional benefit of swimming at that faster speed. Therefore, for the baseline model, when motility is optimal, the optimal swimming speed is purely a function of the optimal cell size; in particular, it is the maximal speed that the cell with that size can attain (Figure 5.1).

I can reach this conclusion by analyzing the model as presented above. From Equations 5.1 to 5.5, I obtain that, for motile cells in an environment with a given concentration of particles, the time between particle encounters is:

$$\tau_{env} \propto \frac{1}{r^3 v^2} \quad (5.14)$$

The energetic cost of motility is due to the concentration of ATP that the cell must use per second ($\tilde{E}_{swim} = E_{swim}/\text{Volume}$), so, from Equations 5.6 and 5.7,

$$\tilde{E}_{swim} \propto \frac{v^2}{r^2}. \quad (5.15)$$

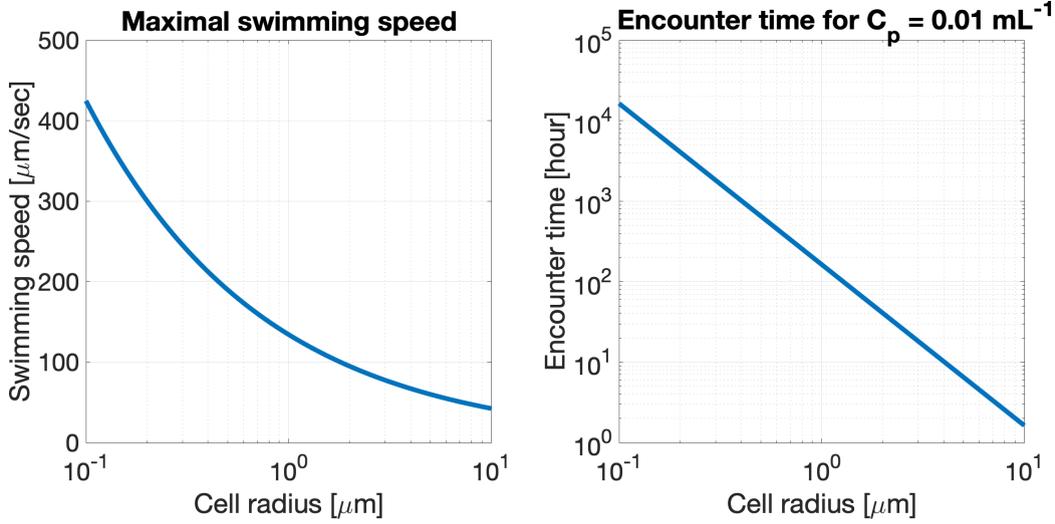


FIGURE 5.1: **Maximal swimming speeds for given cell radii and corresponding encounter times.** The square of the maximal swimming speed a cell of radius r can attain is inversely proportional to r because I assume that the flagellar motor rotates at a maximum frequency of 1,700 Hz and thus has a maximal power output. In the baseline model, this maximal swimming speed is optimal because it minimizes encounter times for a given cell size and the additional energetic cost of swimming at a faster speed is negated by the decrease in time spent swimming. The plot on the right shows the encounter times attained when maximizing flagellar rotation rate and assuming that the concentration of particles is $C_p = 10^{-2} \text{ mL}^{-1}$.

Therefore, if I assume that all of the energy required for swimming is generated while the cell is on the carbon-rich particle, then, from Equation 5.13, the total cytoplasmic concentration of ATP that is used for motility over one period of swimming and replicating is

$$\tau_p \tilde{v}_{futile,p} \propto \frac{1}{r^5}, \quad (5.16)$$

where $\tilde{v}_{futile,p}$ is the cytoplasmic concentration of ATP used by the cell per second for motility. And this quantity does not depend on the swimming speed v .

This analysis is verified by the results of the dual proteome allocation problem (Figure 5.2). In these results, the optimal solution is for either the cell to be nonmotile or for it to be motile and swimming with the flagellar motor attaining its maximum rate of rotation, $E_{swim} = E_{max}$. In these results, maximizing cell volume is optimal at the lowest particle concentrations for which motility is optimal because large cell volumes increase effective diffusivity and thus encounter rates. When encounter rates are too low at the maximal allowed cell volumes because of the scarcity of particles, the optimal solution is to be nonmotile. As the particle concentration increases, maximizing cell volume is no

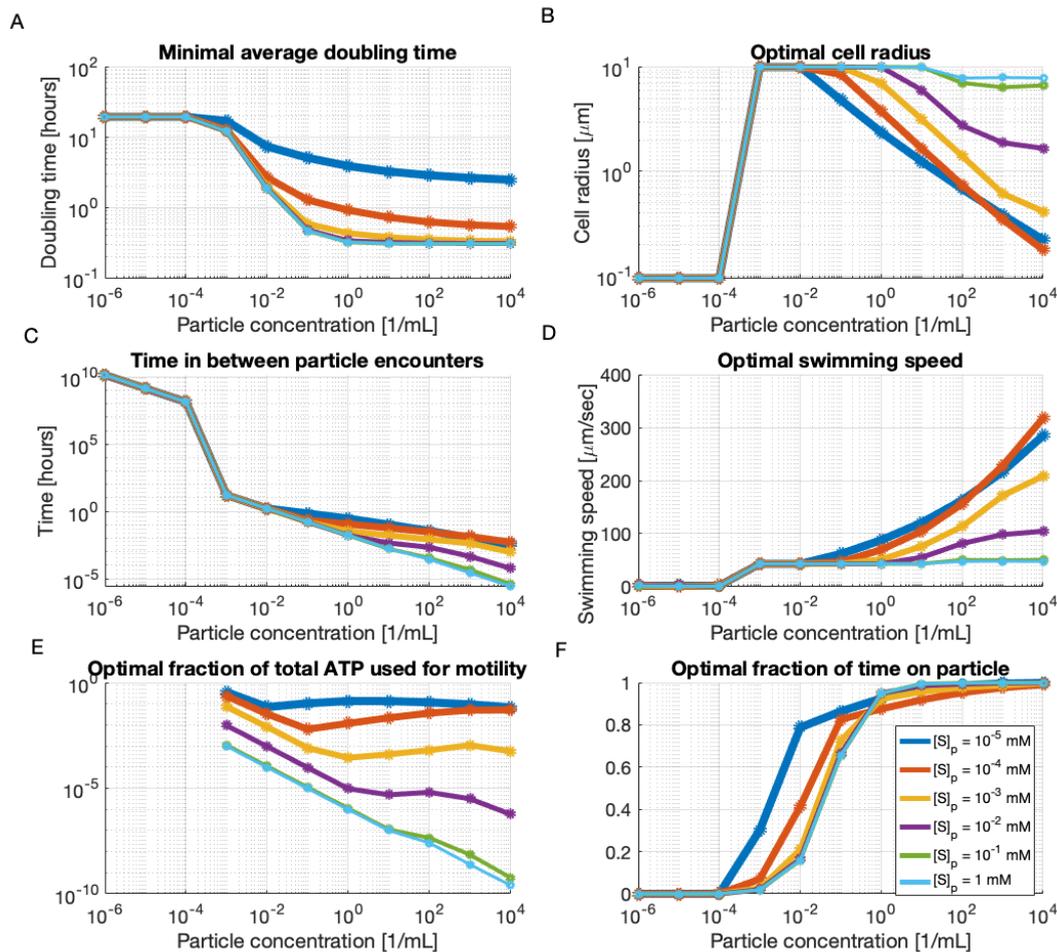


FIGURE 5.2: **Cell doubles on particle.** Proteome allocation problems were run for various values of the substrate concentration on particle, $[S]_p$, where the time spent on a particle is equal to the doubling time of the cell, $\tau_p = \ln(2)/\mu_p$. The plots are of a doubling time, $\ln(2)/\mu_{opt}$, which the dual proteome allocation problem attempted to minimize (A); the corresponding optimal cell radius (B); the corresponding time in between particle encounters, τ_{env} (C); the corresponding optimal swimming speed, which is a function of the optimal cell radius by Equations 5.6 and 5.8 (D); the fraction of total ATP consumed used for motility (E); and the corresponding fraction of time the cell spends on a particle (F).

longer necessary for achieving high encounter rates (Fig. 5.2F). Therefore, as the fraction of time spent on the particle approaches one, the optimal volume of the cell decreases to match the optimal surface area-to-volume ratio for uptake on the particle (Chapter 4). On the other hand, for nonmotile cells, the optimal cell size is the one that optimizes surface area-to-volume ratio for nutrient uptake in the background environment.

Therefore, it is the cell radius that is the master trait. For motile cells, there is a trade-off between increasing cell volume to increase encounter rates and tuning the cell surface area-to-volume ratio to optimize growth for a particular nutrient concentration.

5.2.2 Varying the particle interaction model

Because it is unknown how particle-attached cells determine when to leave a particle, I here modify the model of the time the cell remains on a particle before leaving to encounter a new one, τ_p .

To fully exploit an encountered particle, a cell should remain attached to it until it has depleted it. The cell's and its progeny's combined rate of consumption of the particle at a time t after the cell first attached to the particle is equal to

$$v_{uptake}(t) = V_0 e^{\mu_p t} \tilde{v}_c, \quad (5.17)$$

where $V_0 = 4\pi r^3/3$ is the initial volume of the cell, μ_p is the cell's growth rate on the particle, and \tilde{v}_c is the cellular concentration of carbon substrate consumed per second (Chapter 4.1.3). I assume that the abundance of substrate available to be consumed is $4\pi a^3[S]_p/3$, where $[S]_p$ is the carbon substrate concentration on the particle and a is the radius of the particle. Thus, defining τ_p as the time it takes for the cell and its progeny to deplete the particle,

$$\int_0^{\tau_p} v_{uptake}(t) dt = \frac{4\pi a^3 [S]_p}{3}. \quad (5.18)$$

Thus,

$$\tau_p = \frac{1}{\mu_p} \ln \left(\frac{\mu_p a^3}{\tilde{v}_c r^3} [S]_p + 1 \right). \quad (5.19)$$

I contrast this modification with the baseline case in Figure 5.3. This figure shows two different cases. In the first case, when the particle carbon substrate concentration is $[S]_p = 10^{-4}$ mM, the cell depletes the particle before doubling so that the cell spends less time on the particle than in the baseline case and thus has lower average growth rates. In the second case, when the particle carbon substrate concentration is $[S]_p = 10^{-2}$ mM, the cell remains on the particle with progeny after doubling to deplete the particle. Thus, the cell spends greater fractions of time on the particle and has higher average growth

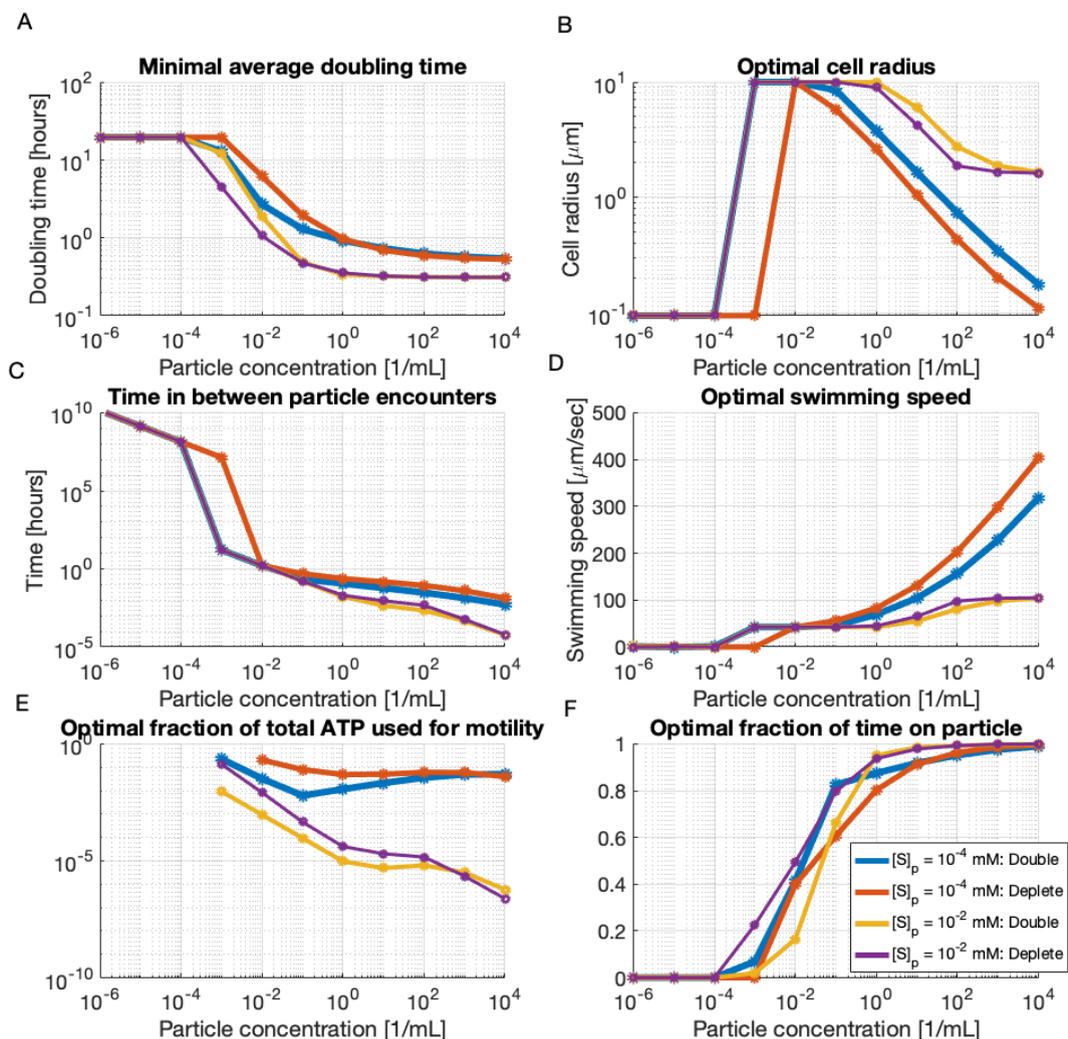


FIGURE 5.3: **Cell depletes particle.** I compare the results of the dual proteome allocation problem for two different particle substrate concentrations in which either the cell leaves the particle immediately after doubling (Double) or the cell (and possibly progeny) remains on the particle until the particle is depleted (Deplete). For the particle depletion interaction model, the time spent on the particle is as given in Equation 5.19.

rates. These differences in the amount of time spent can modify the optimal cell radius from the baseline case by about 50%.

If cells have instead evolved to abandon a particle so that they do not sink to extinction, we might expect cells to remain on a particle until it reaches a certain depth. I can approximate this by assuming that the cell remains on the particle for a constant amount

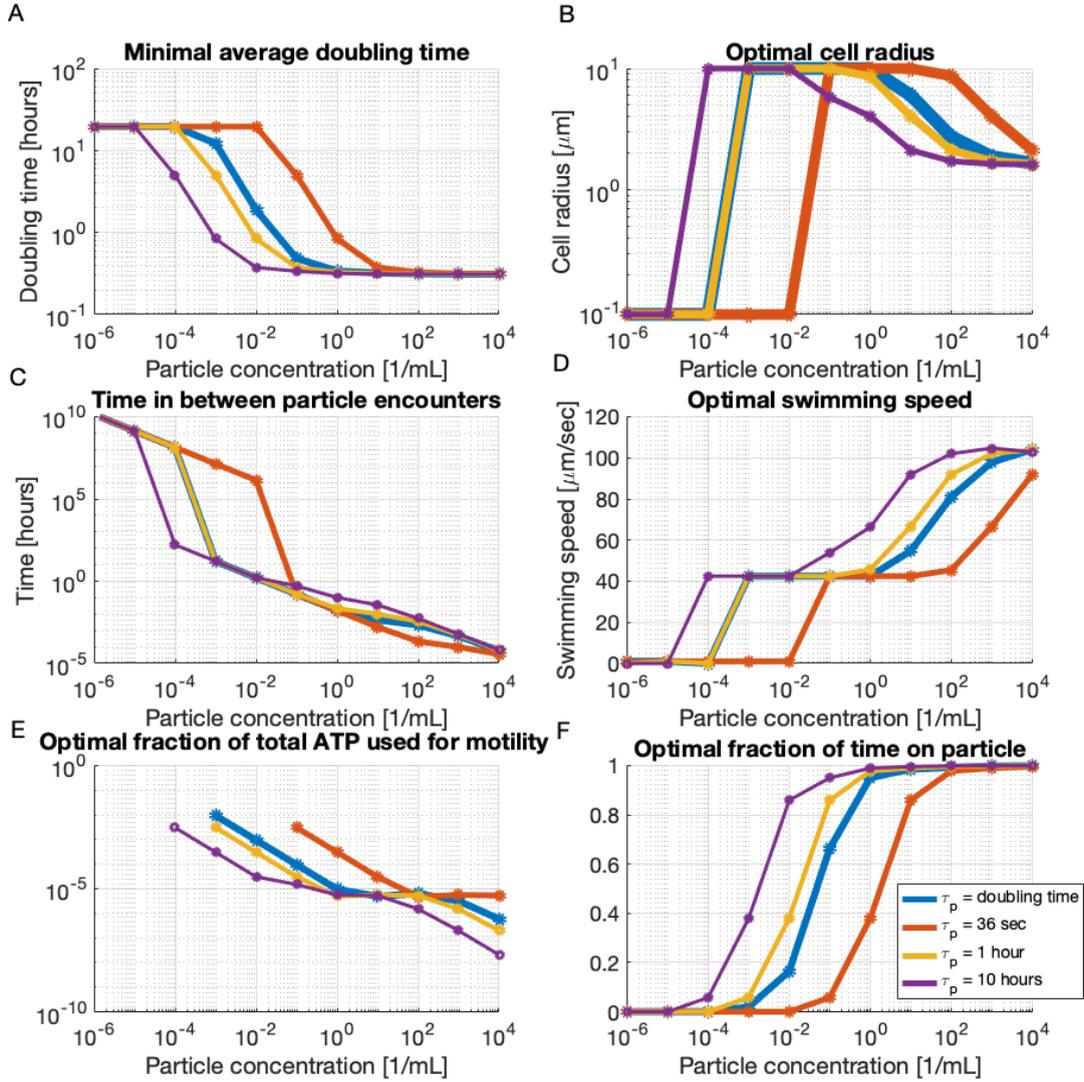


FIGURE 5.4: **Cell escapes sinking particle.** Proteome allocation problems were run as before but assuming $[S]_p = 10^{-2}$ mM and taking τ_p to be either the cell's doubling time on the particle (blue line) or different constant values, specified in the legend.

of time:

$$\tau_p = \text{Constant}. \quad (5.20)$$

I tested this particle interaction model assuming that the concentration of substrate on the particle is $[S]_p = 10^{-2}$ mM (Figure 5.4). The trends of this model are identical to the trends obtained when I instead assumed that the time spent on the particle, τ_p , is equal to the doubling time. In fact, as τ_p increases, all the plots simply shift to the left, as the

cell can better exploit an encountered particle so that motility becomes optimal at lower particle concentrations.

Both of these particle interaction models follow similar trends as the baseline model in which I assumed that the particle cannot be depleted and that the cell remains on it until doubling: The optimal rate of the flagellar motor is its maximum so that the optimal swimming speed is a function of the optimal cell radius, which decreases as the particle concentrations increase and which is smaller for a given particle concentration when the time on particle is greater. Thus, for the remainder of the analysis, I assume that the time spent on the particle is equal to the cell's doubling time because it both limits free parameters and does not depend on the particle size.

5.2.3 The energetic cost of swimming

To gain insight into the energetic cost of swimming, I plot the fraction of ATP generated by the cell that is used for motility (Fig. 5.2C). When particles are sparse and the concentration of substrate on the particle is low, the fraction of ATP used for motility can be substantial; for my parameter choices and with particle substrate concentration $[S]_p = 10^{-5}$ mM and a particle concentration of $C_p = 10^{-3}$ 1/mL (oligotrophic-like conditions), motility is still optimal although it requires 36% of total generated ATP. However, as the substrate concentration of the particle increases or the particle concentration increases, the energetic demands on motility become negligible in the baseline model.

5.2.3.1 The cost of always swimming

In my baseline model, the cell can inhibit swimming while attached to the particle. Thus, the cost of swimming becomes negligible at higher particle concentrations partly because the optimal motile cell is spending nearly its entire lifetime attached to particles on which it does not expend energy swimming (Figure 5.2F). This is the crucial difference between my model of the benefits and costs of motility and previous models, such as the model presented in [198]. It is because the motile cell does not spend all of its time expending energy swimming that the optimal flagellar rotation rate is the maximum allowed. If, on the other hand, I assume that the cell must always expend energy to swim at a particular speed, the maximal flagellar rotation rate is no longer optimal because of its high energetic

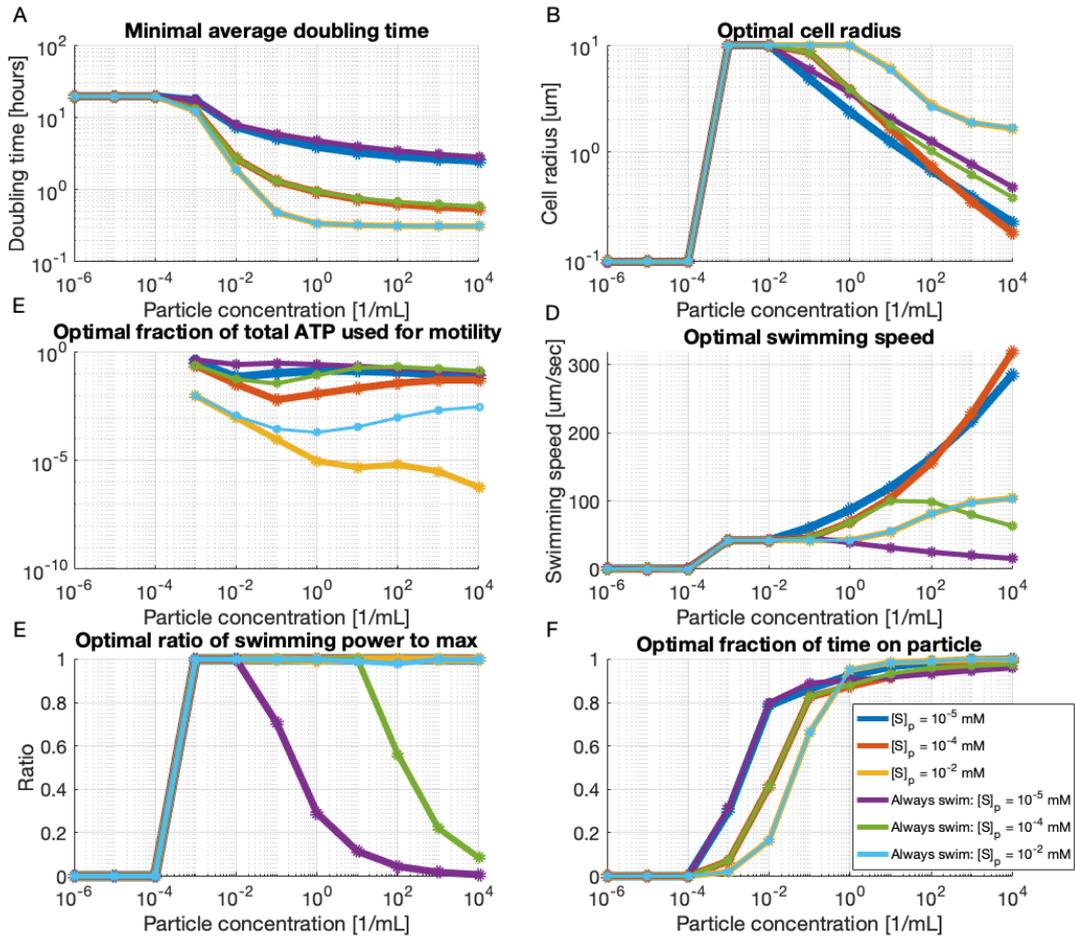


FIGURE 5.5: **Optimal solutions if cell cannot inhibit motility on particles.** For various particle carbon substrate concentrations, $[S]_p$, I compare the baseline case, in which the cell does not expend energy while attached to a particle, to the case in which I instead assume that the cell is always expending energy to swim at a particular speed (‘Always swim’). When the cell is always expending energy to swim, swimming at the maximal flagellar rotation rate is no longer optimal because its effective energetic cost is proportional to the square of the swimming speed even if it spends all time “attached” to a particle. Thus, maximal optimal swimming speeds are lower than in the baseline case, except in cases in which particle carbon concentrations are so high ($[S]_p = 10^{-2}$ mM) that the energetic cost of swimming is negligible compared to its benefit.

cost (Figure 5.5). Thus, as the fraction of time spent on the particle approaches 1, the cellular concentration of ATP consumed per second for motility, $\tilde{v}_{futile,p}$, does not go down to zero, as in the baseline model, but remains proportional to the square of the swimming speed.

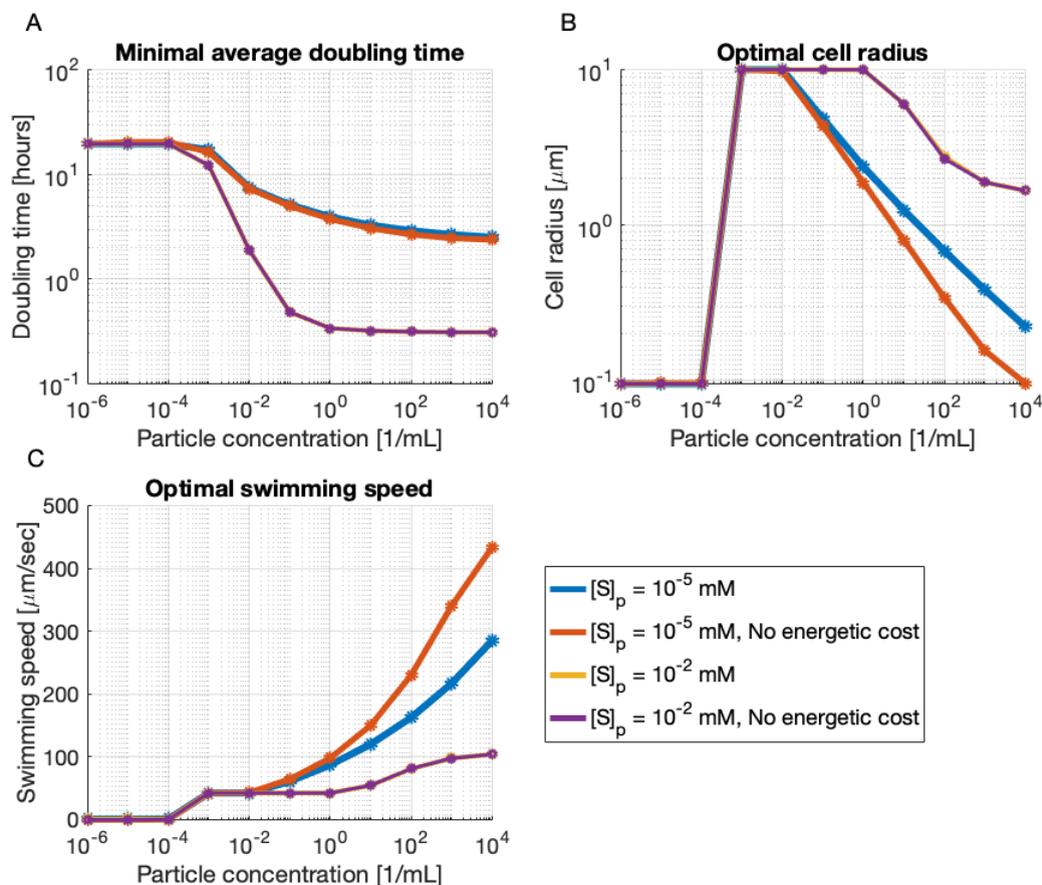


FIGURE 5.6: **Removing the energetic cost of motility.** If I remove the energetic cost of motility from the baseline model, the optimal cell sizes are smaller and swimming speeds correspondingly higher when the concentration of substrate on the particle is low ($[S]_p = 10^{-5}$ mM). For high particle carbon substrate concentrations ($[S]_p = 10^{-2}$ mM), the benefit of motility outweighs its cost so the optimal solution is identical when both accounting for the energetic cost of motility and when not.

5.2.3.2 Removing the energetic costs of swimming

On the other hand, if I remove the energetic costs of motility entirely from the model but still limit the maximal flagellar rotation rate, the optimal cell sizes are smaller and optimal swimming speeds are correspondingly higher when the carbon substrate concentration on the particle is low (Figure 5.6). This clearly demonstrates the trade-off between decreasing cell size to optimize surface area-to-volume ratio for nutrient uptake on the particle and increasing cell size to decrease the effective energetic cost of motility for a given particle encounter rate. That is because, as seen in Equation 5.15, the effective cost of rotating

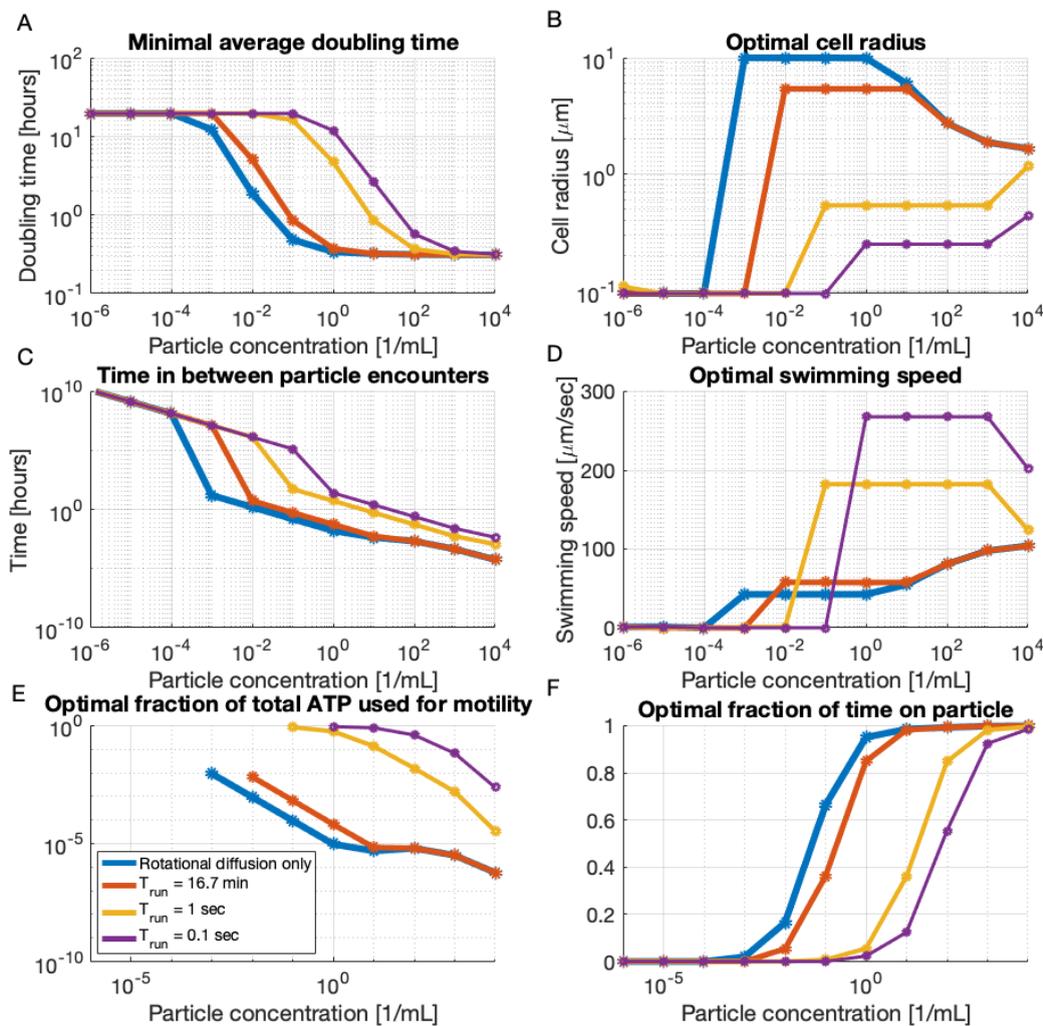


FIGURE 5.7: **Implications of active reorientation.** Proteome allocation problems were run assuming a particle substrate concentration of $[S]_p = 10^{-2}$ mM and different active reorientation rates and are compared with the baseline model (blue line). For active reorientation, I assume that the cell's effective rotational diffusivity is set not by rotational Brownian motion but by a specified average time in between reorientations, T_{run} , giving an effective rotational diffusivity of $\tilde{D}_r = 1/T_{run}$.

the flagellum at the maximal allowed rate increases with decreasing size as it requires the cell to use larger cytoplasmic concentrations of ATP per second.

5.2.4 Implications of chemotaxis

My baseline model assumes that motile cells reorient only due to rotational diffusion because ballistic swimming has maximal translational diffusivity, which maximizes particle encounter rates. However, it is known that many motile cells actively reorient even when in homogeneous nutrient landscapes [193]. Therefore, here I modify the effective translational diffusivity of a motile cell to:

$$\tilde{D}_{t,motile} = \min\left\{D_{t,motile}, \frac{v^2 T_{run}}{6}\right\}, \quad (5.21)$$

where T_{run} is the average amount of time that the cell swims straight before actively reorienting. Therefore, for sufficiently large cells, the effective rotational diffusivity of the cell is now $\tilde{D}_r = 1/T_{run}$.

Because active reorientation may limit diffusivity and thus decrease encounter rates, the lowest particle concentration for which motility is optimal increases with increasing reorientation rates (Figure 5.7). Therefore, fast reorientation rates in homogeneous nutrient landscapes may suggest that motility evolved primarily to facilitate attachment to particles within a concentrated hotspot of nutrient-rich particles.

At the lowest particle concentrations for which motility is optimal, the optimal cell volume is that for which its effective rotational diffusivity due to Brownian motion, $D_{t,motile}$, is equal to that of its diffusivity due to active reorientation, $\tilde{D}_{t,motile}$. This implies that, for very fast active reorientation rates, the optimal cell radius at the sparsest particle concentrations for which motility is optimal is, in fact, smaller (and not larger, as before) than the optimal cell radius for uptake on the particle. This is because, with active reorientation, larger cell volumes do not increase encounter rates. Instead, due to the maximal rate of rotation of the flagellar motor (Equation 5.6), smaller cell volumes allow for faster swimming speeds, which maximize encounter rates.

These results suggest that a cell's rotational diffusivity due to active reorientation in homogeneous nutrient landscapes should not exceed its rotational diffusivity due purely to passive reorientation from rotational Brownian motion. Otherwise, the active reorientation would unnecessarily decrease the cell's translational diffusivity and thus its ability to explore its environment. Previous work indeed suggests that the effective rotational

diffusivity due to active reorientation by the model copiotroph *Vibrio alginolyticus* approximates the effective rotational diffusivity it would have if it simply reoriented passively due to Brownian motion [214].¹ This suggests that, given *Vibrio*'s cell size, its active reorientation strategy does not decrease its effective particle encounter rate, which supports my model's assumption that motility evolved to increase encounters with particles.

5.2.5 Constraining cell size

Because my model does not consider many factors that may constrain cell volume, such as variable predation rates [220], I also solved the dual proteome allocation problem while setting the cell radius to different constants (Figure 5.8). The solutions of this problem make exceptionally clear the high effective costs of motility for small cells. This work thus suggests that small cells do not swim because the environments in which their motility would be optimal (environments with particle concentrations exceeding 100 per milliliter) are rare. On the other hand, my work suggests that motility may be optimal for cells with an effective cell radius of $0.5 \mu\text{m}$ in copiotrophic environments, which is supported by observations of a variety of motile marine bacteria with effective cell radii of $0.4 \mu\text{m}$ (having body lengths of $0.8 \mu\text{m}$) [49].

¹The approximation of a cell's rotational diffusivity in [214] assumed that a *Vibrio* cell is an ellipsoid with semi-minor axis of length $2a = 0.4 \mu\text{m}$ and semi-major axis of length $2b = 2.3 \mu\text{m}$ so that $D_r \approx k_B T (\ln(2b/a) - 0.5) / (8\pi\eta b^3/3)$. Mitchell suggested in an earlier work [134] that this was not an appropriate approximation of the cell's effective rotational diffusivity because it neglected friction from the cell's flagellum. Thus, Mitchell applied this same equation for the rotational diffusivity of a rigid ellipsoid to also calculate the rotational drag of the flagellum, where now b is the length of the helical flagellum and a is the effective radius of the helix. Therefore, assuming that the flagellum is a rigid ellipsoid rigidly attached to the cell body, Mitchell concluded that a flagellum 'stabilizes' the cell and that an arbitrarily long flagellum could decrease the cell's rotation due to Brownian motion effectively to zero. However, more recent work suggests that this rigid body approximation vastly overestimates drag due to the flagellum and thus that the flagellum cannot 'stabilize' the cell [114, 166]. Therefore, I approximate the translational and rotational drag of the cell as being purely due to the cell body. This also differentiates my estimates of the cost of motility from that presented in [198]. Their estimates increased the power requirements of motility by a factor of four compared to mine because they assumed that the flagellum contributed 75% of the total translational drag of the cell. However, the more recent theory suggests that the translational drag of the flagellum should not exceed the translational drag due to the cell body [166].

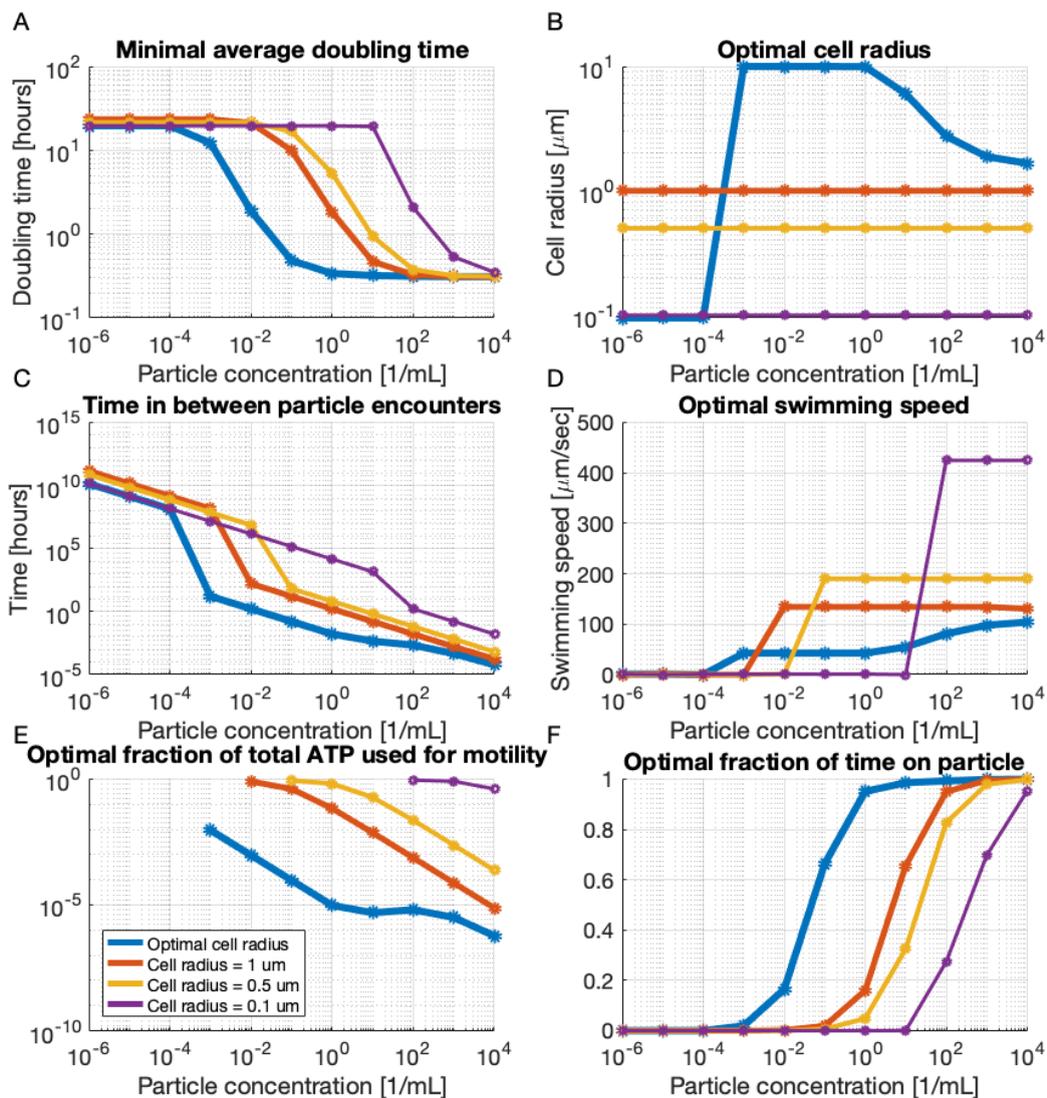


FIGURE 5.8: **Constraining cell size.** Proteome allocation problems were run assuming a particle substrate concentration of $[S]_p = 10^{-2}$ mM but assuming various constant cell radii, though still assuming only passive rotational diffusion while swimming in homogeneous nutrient landscapes.

5.3 Conclusions

In contrast to previous work that assumes that motile cells always swim [198], this work assumes that cells swim to forage for particles, on which they can attach so that they no longer need to swim. Therefore, this work suggests that motility has lower effective

costs than previously estimated. Furthermore, it suggests that motility has tremendous benefits for cells that use it to find and adhere to nutrient-rich particles in environments where the carbon in the background is low but the concentration of nutrient-rich particles is sufficiently high. The costs of motility are effectively low because cells that encounter particles more quickly spend greater fractions of their lifetimes attached to particles on which they no longer need to expend energy for swimming. Thus, in effect, my model suggests that motility may have evolved so that marine bacteria could “island hop” from one nutrient-rich particle to the next. This result follows our intuition that copiotrophs live a “feast or famine” existence [104, 190] and are only motile in feast conditions, where nutrient-rich particles are plentiful. This is supported by experimental evidence showing that much higher percentages of bacterial communities are motile following an algal bloom [70] and also by experimental evidence showing that the majority of motile cells stop swimming soon after the cessation of a nutrient input [122].

Yet, because in my model the carbon concentration in the background environment is so low, motility can still be optimal in cases in which it has substantial total energetic costs. If I assume that a bacterial cell cannot have an effective radius larger than $1\mu\text{m}$, then the critical particle concentration at which the cost of motility is nearly equal to its benefit is approximately 10^{-2} particles per milliliter (Figure 5.8). This concentration is equivalent to an average distance between particles of approximately 5 cm. Because of rotational diffusion, it takes a cell of this size on average 160 hours (nearly a week!) to encounter particles at this distance when swimming at the optimal swimming speed of approximately $130\ \mu\text{m}/\text{sec}$. I have underestimated the cost of motility because I do not consider the cost of storing the energy generated while growing on the particle and used while swimming. Yet, the ability of cells to swim such long periods of time is supported by observations that about 3% of a population of a motile marine bacterium, *Photobacterium angustum*, were still motile after two weeks of starvation [122, 195].

In fact, I may still be overestimating the effective energetic cost of motility over a cell’s lifetime from the standpoint of natural selection because I assumed that the cell always finds a new particle at exactly the estimated ensemble *average* encounter time. However, ‘lucky’ cells can encounter particles much faster than the average, and, as Dawkins convincingly argued [45], “selfish genes” require only lucky cells for their successful propagation. Thus, I may be able to better predict the environments in which

motility has evolved by considering the time it takes for, say, the top 1% or 10% of cells to encounter a particle.

One discrepancy between my model and observations, however, is that my model predicts higher swimming speeds in homogeneous nutrient landscapes than are observed. While my model predicts that a cell with an effective radius of 1 μm will swim at speeds of about 130 $\mu\text{m}/\text{sec}$, many marine cells that size have been observed swimming at much lower speeds, around 30 $\mu\text{m}/\text{sec}$ [15, 136]. In reality, swimming speeds are very variable within a clonal population of marine bacteria, and ‘burst’ velocities exceeding 400 $\mu\text{m}/\text{sec}$ have been observed [137]. Furthermore, swimming speeds are sensitive to both harvesting and experimental conditions: Both larger cells (grown in richer nutrient sources) [137] and starved cells [215] swim slower, while cells swim faster when sensing a gradient [15, 136]. Additionally, swimming speeds depend on environmental conditions such as salinity and pH [136, 184].

Yet, the main discrepancy between my predictions of optimal swimming speeds and observed speeds is that my model incorrectly predicts that a cell swims at the maximal flagellar rotation rate in homogeneous nutrient landscapes. I hypothesize that this is because there is no uncertainty in my model. In my model, the cell is guaranteed to encounter a particle at exactly the calculated time τ_{env} , which depends on its swimming speed. The cell in my model thus does not need to conserve energy in the case it does not encounter a particle within the specified time. A more realistic model would thus not only consider the cost of storing energy but also consider the effects of uncertainty. While spatial heterogeneity causes uncertainty in encounter rates, ‘lucky’ cells will, with very high probability, still be able to encounter particles within the required estimated encounter time, given a specific concentration of particles, if swimming at maximal speeds.

Therefore, temporal heterogeneity in the environment may be a more important determinant of how cells regulate motility. For example, cells may be uncertain about when an algal bloom will occur and may thus conserve energy to survive a period of ‘famine’. I predict that it is this uncertainty and need to conserve energy that causes marine motile cells to swim at reduced speeds when not sensing any nutrients. The cells thus only swim at their maximal speeds when sensing a gradient to attempt to catch up with the gradient’s nutrient-rich source, as experimental observations suggest.

To test this hypothesis, experimental work should be conducted, similar to that done in [217], that links variable cellular regulation of motility from different strains of marine cells to differences in the environments in which the cells evolved.

The work in this chapter was done with guidance from Roman Stocker, Vicente Fernandez, and Naomi Levine.

Chapter 6

Outlook

In the previous chapter, I suggested that the energetic cost of motility is substantial and prohibits motility in environments with lower concentrations of carbon. This energetic cost is even more inhibitive in “energy-limited” environments that lack sufficient concentrations of the electron acceptors that the cells require for respiration [81, 118]. In addition, copiotrophs have a number of specialized traits that are energetically costly and thus may also preclude their ability to survive in a variety of environments.

These many additional traits—for example, copiotrophs’ ability to degrade a variety of complex substrates, to attach to particles, and to glide on particles—enable their “opportunistic” lifestyles [42]. Yet, these traits also require larger genomes. Model oligotrophs, on the other hand, have “streamlined genomes” that are approximately one-fifth the size of the genomes of model copiotrophs [64, 66]. The prevalence of organisms with streamlined genomes suggest that the effective cost of large genomes is high. A large genome clearly requires more metabolites to replicate. Additionally, because of the associated proteomic costs, copiotrophs do not constitutively express their entire arsenal of traits but carefully regulate gene expression. Thus, while oligotrophs lack two-component systems for sensing and rely heavily on the constitutive expression of genes, copiotrophs have higher proteomic and energetic costs due to the complexity of their regulatory networks [78]. But how can we quantify the cost of this phenotypic plasticity?

We unfortunately lack an understanding of how regulatory cellular processes contribute to the basal proteomic and energetic requirements of the cell. In fact, we lack

estimates of the basal energetic requirements of the cell [96]. Yet, predicting the basal requirements of various lifestyles will not only enable us to have a much better understanding of speciation among marine heterotrophic bacteria but will also enable us to develop more predictive models of carbon-cycling in the ocean. To explicitly account for heterotrophic bacterial dynamics and their consumption and respiration of carbon, biogeochemical models can use simple ordinary differential equations of the following form to track the abundance of a bacterial population, B , and the corresponding amount of carbon remineralized by the population, C_{remin} [221]:

$$\begin{aligned}u_B &= V_{max} \frac{C_O}{C_O + K}, \\ \frac{dB}{dt} &= y u_B B - LB, \\ \frac{dC_{remin}}{dt} &= (1 - y) u_B B,\end{aligned}$$

where C_O is the concentration of organic carbon; u_B is the uptake rate per cell described via a maximal uptake rate, V_{max} , and half-saturation concentration, K ; L is a mortality loss term; and y is the population's growth efficiency. This growth efficiency term is thus the fraction of the consumed organic carbon that the population uses for biomass rather than for energy generation.

The work in this thesis provided insight into how both phenotypic diversity and environmental conditions determine the values of the maximal uptake rate, V_{max} , and the half-saturation concentration, K . However, we do not understand how the growth efficiency, y , is a function of both phenotype and environment [13]. Gaining this understanding will require a combination of: field observations to probe uptake and remineralization rates of *in situ* microbial communities; experiments that can demonstrate how various species' uptake rates and growth efficiencies are functions of culture conditions; and modeling work that attempts to extrapolate from these results to make testable predictions. This future work would elucidate bacterial energetics, allowing further insight into how metabolism determines physiology and ecology in the microbial world.

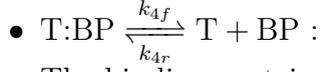
Appendix A

Modeling *E. coli* Chemotaxis to Maltose

A.1 An ABC transport model

I describe maltose transport into the cell by the following reactions and mass-action kinetic rates:

- $L_{\text{ext}} \xrightarrow{v_p} L_p$:
The ligand enters the periplasm from the external environment with periplasmic uptake rate, v_p .
- $L_p + \text{BP} \xrightleftharpoons[k_{1r}]{k_{1f}} \text{L:BP}$:
The periplasmic ligand, L_p , reversibly associates with the binding protein, BP, to form ligand-binding protein complex, L:BP.
- $\text{L:BP} + \text{T} \xrightleftharpoons[k_{2r}]{k_{2f}} \text{T:L:BP}$:
The ligand-binding protein complex reversibly associates with the membrane-bound ABC transporter to form T:L:BP.
- $\text{T:L:BP} \xrightarrow{k_3} \text{T:BP} + L_i$:
The ligand is actively transported into the cytoplasm leaving behind the binding protein in complex with the ABC transporter, T:BP.



The binding protein reversibly associates with the ABC transporter.

The corresponding reaction rates are:

$$v_1 = k_{1f}[L]_p[BP] - k_{1r}[L:BP], \quad (\text{A.1})$$

$$v_2 = k_{2f}[T][L:BP] - k_{2r}[T:L:BP], \quad (\text{A.2})$$

$$v_3 = k_3[T:L:BP], \quad (\text{A.3})$$

$$v_4 = k_{4f}[T:BP] - k_{4r}[T][BP]. \quad (\text{A.4})$$

At steady-state, all of the concentrations remain constant:

$$0 = \frac{d[BP]}{dt} = -v_1 + v_4, \quad (\text{A.5})$$

$$0 = \frac{d[L:BP]}{dt} = v_1 - v_2, \quad (\text{A.6})$$

$$0 = \frac{d[T:L:BP]}{dt} = v_2 - v_3, \quad (\text{A.7})$$

$$0 = \frac{d[T:BP]}{dt} = v_3 - v_4, \quad (\text{A.8})$$

$$0 = \frac{d[T]}{dt} = -v_2 + v_4, \quad (\text{A.9})$$

$$\implies v_1 = v_2 = v_3 = v_4. \quad (\text{A.10})$$

In addition are two conservation equations:

$$[BP]_{total} = [BP] + [L:BP] + [T:L:BP] + [T:BP] \quad (\text{A.11})$$

$$[T]_{total} = [T] + [T:L:BP] + [T:BP] \quad (\text{A.12})$$

In the following, I make a number of simplifying assumptions based on our understanding of maltose transport in *E. coli*. First, I assume that the dissociation of the binding protein from the transporter is very fast so that $[T:BP]$ is much smaller than $[T]$.

Thus, because $v_2 = v_3$, the cytoplasmic uptake rate is

$$v_c = k_3[\text{T:L:BP}] \quad (\text{A.13})$$

$$\approx k_3[\text{T}]_{total} \frac{[\text{L:BP}]}{K_c + [\text{L:BP}]}, \quad K_c = \frac{k_3 + k_{2r}}{k_{2f}}. \quad (\text{A.14})$$

Because $v_1 = v_2$,

$$(k_{1r} + k_{2f}[\text{T}]) [\text{L:BP}] = k_{1f}[\text{L}][\text{BP}] + k_{2r}[\text{T:L:BP}]. \quad (\text{A.15})$$

I assume that, when maltose is in the micromolar range in the external environment, the transporters are working near saturation so that $[\text{BP}]_{total} \geq [\text{T}]_{total}$ and $k_{1r} \gg k_{2f}[\text{T}]$. Thus, defining the dissociation constant of the binding protein complex to be $K_{BP} = k_{1r}/k_{1f}$,

$$[\text{L:BP}] \approx \frac{[\text{L}][\text{BP}]}{K_{BP}} + \frac{k_{2r}}{k_{1r}}[\text{T:L:BP}] \quad (\text{A.16})$$

$$\approx \frac{[\text{L}]}{K_{BP}} ([\text{BP}]_{total} - [\text{L:BP}] - [\text{T:L:BP}]) + \frac{k_{2r}}{k_{1r}}[\text{T:L:BP}] \quad (\text{A.17})$$

$$\approx \frac{[\text{BP}]_{total}[\text{L}]_p + \frac{k_{2r}-k_{1f}[\text{L}]_p}{k_{1f}}[\text{T:L:BP}]}{K_{BP} + [\text{L}]_p}. \quad (\text{A.18})$$

In *E. coli*, it has been shown that the abundance of the maltose binding protein greatly exceeds the abundance of transporters [31]; that is, $[\text{BP}]_{total} \gg [\text{T}]_{total}$. Therefore, the abundance of binding proteins also exceeds the abundance of transporters associated with the maltose-binding protein complex: $[\text{BP}]_{total} \gg [\text{T:L:BP}]$. Additionally, I assume that the dissociation rate of the maltose-binding protein complex from the transporter is small, so that I approximate

$$[\text{L:BP}] \approx [\text{BP}]_{total} \frac{[\text{L}]_p}{K_{BP} + [\text{L}]_p}, \quad \text{and} \quad (\text{A.19})$$

$$[\text{T:L:BP}] \approx [\text{T}]_{total} \frac{[\text{BP}]_{total}}{K_c + [\text{BP}]_{total}} \frac{[\text{L}]_p}{\frac{K_c K_{BP}}{K_c + [\text{BP}]_{total}} + [\text{L}]_p}. \quad (\text{A.20})$$

Therefore, for $[\text{BP}]_{total} \gg K_c$, the cytoplasmic uptake rate is

$$v_c \approx V_c \frac{[\text{L}]_p}{\frac{K_c K_{BP}}{[\text{BP}]_{total}} + [\text{L}]_p}, \quad V_c = k_3[\text{T}]_{total}. \quad (\text{A.21})$$

While transport into the cytoplasm is active and thus can occur against concentration gradients, transport into the periplasm via porins is diffusive. Thus, while the cytoplasmic uptake rate has the above form, the periplasmic uptake rate is better described by the following Michaelis-Menten equation [33]:

$$v_p = V_p \frac{[L]_{ext} - [L]_p}{K_p + [L]_{ext} + [L]_p}, \quad (\text{A.22})$$

where K_p is the half-saturation constant of the specific porin; and V_p is the maximal rate of uptake, which is a function of the number of expressed porins. In environments with micromolar ranges of maltose, *E. coli* has nanomolar periplasmic concentrations of maltose ([58], [106]). Thus, I assume $[L]_p \ll [L]_{ext}$ and make the following simplification:

$$v_p \approx V_p \frac{[L]_{ext}}{K_p + [L]_{ext}}. \quad (\text{A.23})$$

At steady-state, the periplasmic transport rate (v_p) must be equal to the cytoplasmic transport rate (v_c), so that by Equations A.21 and A.23,

$$[L]_p \approx \frac{\left(\frac{K_c K_{BP}}{[BP]_{total}}\right) V_p [L]_{ext}}{K_p V_c + (V_c - V_p) [L]_{ext}}. \quad (\text{A.24})$$

Solutions exist for all choices of K_p and $[L]_{ext}$ when the maximum periplasmic uptake rate is less than the maximum cytoplasmic rate, $V_p \leq V_c$. If instead the periplasmic uptake rate exceeds the cytoplasmic rate ($V_p \geq V_c$), the periplasm would saturate with ligand and reach a steady-state of $[L]_p \sim [L]_{ext}$.

A.2 A chemotaxis model

A.2.1 The SPECS model

In [90], Jiang, Ouyang, and Tu developed a ‘‘coarse-grained’’ quantitative model of *E. coli* chemotaxis to methyl-aspartate (MeAsp), called the Signaling Pathway-based *E. coli* Chemotaxis Simulator (SPECS). This model accounts for both the signal transduction molecular pathway of chemotaxis [53] as well as receptor cooperativity [36]. Tu derives

much of this model in his review [202]. He shows how the model relies on an understanding of the free-energy differences between the different possible states of a chemoreceptor and the Monod-Wyman-Changeux (MWC) allosteric approximation for receptor cooperativity [140]. The model also considers methylation dynamics as well as signal amplification at the flagellar motor. It can be used to run agent-based simulations of *E. coli* in various gradients of methyl-aspartate.

For my work, I extended SPECS to incorporate: the heterogeneous MWC model, as it is proposed in both [126] and [97], to consider the chemotactic response of cells to multiple chemoattractants; and a new model of chemotaxis to maltose that takes into account the transport kinetics of maltose into and out of the periplasm as well as the indirect binding of maltose to the aspartate receptor via the maltose-binding protein.

Below I describe the components of the original SPECS model and derive extensions to it.

Probability of tumbling. In SPECS, the probability that a cell switches from running (swimming approximately straight) to tumbling (changing its direction) is a function of the average activity of all of its chemoreceptors, $\langle a \rangle$:

$$\text{probability of tumbling} = \frac{\langle a \rangle^H}{\tau_1 a_0^H}, \quad (\text{A.25})$$

where: τ_1 is the average run time of the cell; H is the Hill coefficient for the sensitivity of the flagellar motor to binding by CheY-P; and a_0 is the steady-state activity.

The activity level is a function of both the memory (methylation state, m) of the cell as well as the concentrations of chemoattractants that the cell is experiencing. In [202], Tu derives the expected activity level of a receptor cluster for a cell in gradients of MeAsp. For a cluster with n receptors, the average activity is:

$$\langle a \rangle = \frac{1}{1 + e^{[nf_m(m) + nf_{MeAsp}([\text{MeAsp})]]}}. \quad (\text{A.26})$$

Methylation state. The first term in the exponential in Equation A.26, $f_m(m)$, is the free-energy difference between an active and inactive chemoreceptor:

$$f_m(m) = \alpha (m_0 - m), \quad (\text{A.27})$$

where: α is the free-energy difference per added methyl group; and m_0 is the reference methylation level. The methylation state of the cell is $m \in [0, 4]$ because the aspartate receptor has four methylation sites. Its dynamics depends on the activity of the cell:

$$\frac{dm}{dt} = k_R (1 - \langle a \rangle) - k_B \langle a \rangle, \text{ for } m \in (0, 4), \quad (\text{A.28})$$

where: k_R is the rate of the methylation process, and k_B is the rate of the demethylation process. Thus, the steady state activity level is:

$$a_0 = \frac{k_R}{k_B + k_R}. \quad (\text{A.29})$$

The average activity level. The remaining term in the exponential in Equation A.26, $f_{MeAsp}([\text{MeAsp}])$, is the free-energy difference between a bound and vacant chemoreceptor.

In [202], Tu derives this free energy term to be the following:

$$f_{MeAsp}([\text{MeAsp}]) = \log \left[\frac{1 + \frac{[\text{MeAsp}]}{K_{I,MeAsp}}}{1 + \frac{[\text{MeAsp}]}{K_{A,MeAsp}}} \right], \quad (\text{A.30})$$

where $K_{A,MeAsp}$ ($K_{I,MeAsp}$) is the dissociation constant of MeAsp to the active (respectively, inactive) Tar chemoreceptor. Because the receptor is more likely to be active when it is vacant, $K_{A,MeAsp} > K_{I,MeAsp}$. Here I derive the activity level for general chemoattractants so that I can expand the SPECS model to include maltose chemotaxis.

As in the MWC model, I assume that the receptors in a cluster are either all active or all inactive. I follow the derivation presented in [202] to determine the free energy difference between the active state and the inactive state of a receptor cluster.

A single receptor has four possible states (active or inactive: $a = \{0, 1\}$; and bound or vacant: $l = \{0, 1\}$) with the following probabilities, $P(a, l)$:

$$\frac{P(1, 0)}{P(0, 0)} = e^{-f_m(m)} \quad (\text{A.31})$$

$$\frac{P(0, 1)}{P(0, 0)} = C_I \quad (\text{A.32})$$

$$\frac{P(1, 1)}{P(1, 0)} = C_A, \quad (\text{A.33})$$

where C_I and C_A are functions that I will derive below.

I solve for the probabilities to determine the expected activity level of a single receptor:

$$\begin{aligned} P(0, 0) &= 1 - P(0, 1) - P(1, 0) - P(1, 1) \\ &= 1 - P(0, 0) [C_I + e^{-f_m(m)} (1 + C_A)] \\ &= \frac{1}{1 + C_I + e^{-f_m(m)} (1 + C_A)}. \end{aligned}$$

Because the expected activity level of a single receptor is $\langle a \rangle_{receptor} = P(1, 0) + P(1, 1)$,

$$\langle a \rangle_{receptor} = \frac{e^{-f_m(m)} [1 + C_A]}{1 + C_I + e^{-f_m(m)} (1 + C_A)}. \quad (\text{A.34})$$

I define the free energy difference, Δf , such that

$$\langle a \rangle_{receptor} = (1 + e^{-\Delta f})^{-1}. \quad (\text{A.35})$$

Thus,

$$\Delta f = -f_m(m) - \log \left[\frac{1 + C_I}{1 + C_A} \right]. \quad (\text{A.36})$$

Because I assume that all of the n receptors are active or all of them are inactive, the expected activity level of the entire receptor cluster is

$$\langle a \rangle = (1 + e^{-n\Delta f})^{-1}. \quad (\text{A.37})$$

(A derivation of this expression can be found in [157].)

Therefore, a general formulation for the average activity level of a cell sensing chemoattractant L is

$$\langle a \rangle = \frac{1}{1 + e^{[nf_m(m) + nf_L([L])]}}, \quad (\text{A.38})$$

with

$$f_L([L]) = \log \left[\frac{1 + C_I}{1 + C_A} \right], \quad (\text{A.39})$$

where C_I (C_A) is the ratio of the probabilities of a receptor being bound versus ligand-free for an inactive (respectively, active) receptor. I derive the C terms for both MeAsp and maltose below.

The MeAsp free energy term. By definition, the dissociation constant K_{MeAsp} is defined to be the following under equilibrium:

$$K_{MeAsp} = \frac{[R]_{free}[MeAsp]}{[R:MeAsp]}, \quad (\text{A.40})$$

where: $[R]_{free}$ and $[MeAsp]$ are the effective concentrations of free receptor and free ligand in the periplasm; and $[R:MeAsp]$ is the concentration of bound receptor. Defining $[R] = [R]_{free} + [R:MeAsp]$,

$$[R:MeAsp] = \frac{[R][MeAsp]}{K_{MeAsp} + [MeAsp]} \quad (\text{A.41})$$

Therefore, the probability a receptor is bound is

$$P = \frac{[R:MeAsp]}{[R]} = \frac{[MeAsp]}{K_{MeAsp} + [MeAsp]}, \quad (\text{A.42})$$

so

$$C = \frac{P}{1 - P} = \frac{[MeAsp]}{K_{MeAsp}}. \quad (\text{A.43})$$

Therefore, distinguishing an active from an inactive receptor,

$$f_{MeAsp}([MeAsp]) = \log \left[\frac{1 + \frac{[MeAsp]}{K_{I,MeAsp}}}{1 + \frac{[MeAsp]}{K_{A,MeAsp}}} \right]. \quad (\text{A.44})$$

Note that because MeAsp is not metabolized by the cell, the steady-state concentration of MeAsp in the periplasm is equal to the extracellular concentration of MeAsp.

A.2.2 A model for maltose chemotaxis

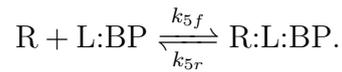
Adding sensing to the ABC transport model complicates an already complicated system. To sense maltose, the receptors must compete with the ABC transporters to bind with the ligand-binding protein complex. Optimally, however, sensing would minimally hinder transport to thus minimally decrease the cell's growth rate. I thus make the simplifying assumption that sensing does not affect transport but simply "reads" the state of the system. This is a reasonable approximation given that the abundance of maltose-binding protein greatly exceeds the abundance of the cognate ABC transporter.

Therefore, in my model, I assume that the ABC transporters and receptors do not compete for the ligand-binding protein complex and likewise assume that the receptors do not affect binding and dissociation of the binding protein with maltose. Therefore, I modify the transport model to incorporate sensing only via a simple modification to Equation A.19:

$$[\text{L:BP}]_0 = [\text{R:L:BP}] + [\text{L:BP}] \approx [\text{BP}]_{total} \frac{[\text{L}]_p}{K_{BP} + [\text{L}]_p}, \quad (\text{A.45})$$

where R:L:BP is the receptor bound to the ligand-binding protein complex.

I assume that the receptor can only bind to the complex and not to the binding protein on its own. Therefore, I describe sensing by the following mass-action kinetics:



At steady-state, the concentration of the bound receptor does not change, so that

$$[\text{R:L:BP}] = \frac{[\text{R}][\text{L:BP}]}{K}, \quad K = \frac{k_{5r}}{k_{5f}}. \quad (\text{A.46})$$

Combining Equations A.45 and A.46, I have that the total concentration of maltose-MBP complex bound to an inactive (I) or active (A) receptor is

$$[\text{R:L:BP}]_{I,A} = \frac{([\text{R}]_{total} - [\text{R:L:BP}]_{I,A})([\text{L:BP}]_0 - [\text{R:L:BP}]_{I,A})}{K_{I,A}}, \quad (\text{A.47})$$

where $K = K_I$ when the receptors are inactive and $K = K_A$ when the receptors are active. Therefore, the terms $[\text{R:L:BP}]_{I,A}$ are the solutions to quadratic equations, and the free energy term for maltose is:

$$f_{Mal}([\text{L}]_{ext}) = \log \left[\frac{1 + C_I}{1 + C_A} \right], \quad (\text{A.48})$$

where

$$C_{I,A} = \frac{[\text{R:L:BP}]_{I,A}}{[\text{R}]_{total} - [\text{R:L:BP}]_{I,A}}. \quad (\text{A.49})$$

A.2.3 The Heterogeneous MWC model

By the Heterogeneous MWC (HMWC) model ([126], [97]), in mixed gradients of ligand 1 (L_1 , which binds to receptor R_1) and ligand 2 (L_2 , which binds to receptor R_2), the average activity a is

$$\langle a \rangle = \frac{1}{1 + e^{[n_{R_1} f_{m,R_1}(m_{R_1}) + n_{R_1} f_{L_1}([L_1]) + n_{R_2} f_{m,R_2}(m_{R_2}) + n_{R_2} f_{L_2}([L_2])]}}. \quad (\text{A.50})$$

MeAsp and the maltose-binding protein complex bind independently to distinct sites of Tar [142]. Thus, in mixed gradients of MeAsp and maltose,

$$\langle a \rangle = \frac{1}{1 + e^{n_{Tar} [f_m(m) + f_{MeAsp}([\text{MeAsp}]) + f_{Mal}([\text{Mal}]_{ext})]}}. \quad (\text{A.51})$$

A.3 An analytical approximation of the steady-state distribution

In [84], Hu and Tu used the HMWC model to derive an advection-diffusion equation for population-level dynamics of chemotactic cells in one dimension. They used this

diffusion equation to predict the stationary distribution of the cells in environments with two chemoattractants. The steady-state cell distribution H_u and T_u derived has the following form for mixed gradients of MeAsp and maltose:

$$p(x) \propto e^{(\eta(x)[f_{MeAsp}(x)+f_{Mal}(x)])}, \quad (\text{A.52})$$

where

$$\eta(x) = H n_{Tar} (1 - \bar{a}(x)) \left[1 - \frac{z_0}{z(\bar{a}(x))} \right], \quad (\text{A.53})$$

where: H is the Hill coefficient for signal amplification at the flagellar motor; $\bar{a}(x) = \frac{a^+(x) - a^-(x)}{2}$ is the activity averaged over left-moving and right-moving cells at location x ; z_0 is the rotational diffusion rate; and $z(a)$ is the total tumbling rate:

$$z(\bar{a}) = z_0 + \frac{\bar{a}^H}{\tau_t K_{1/2}^H}, \quad (\text{A.54})$$

where τ_t is the time it takes the cell to tumble; and $K_{1/2}$ is the activity level at which the clockwise flagellar motor bias is 0.5. As explained in the derivations in [84], $\bar{a}(x) \approx a_0$ in shallow gradients.

This formulation is appealing because it is intuitive and easy to calculate. It directly relates the molecular-level dissociation constants of the aspartate chemoreceptor to the population-level relative affinity of the cells to various concentrations of maltose and MeAsp. However, the difficulty with this approximation is that the parameter $\eta(x)$ is unknown and depends nontrivially on the local gradient of chemoattractants at location x . As explained in the supplemental materials in [84], as the chemoattractant gradients increase, the average activity $\bar{a}(x)$ increases and thus $\eta(x)$ decreases. In my single-gradient experimental setups, the gradients are linear, so I can set $\eta(x) = \eta_j$ for different experimental conditions j , where $\eta_k > \eta_l$ if experiment k used a lower concentration of the chemoattractant than in experiment l . H_u and T_u suggest that η can vary widely over environmental conditions. However, I found instead very good agreement between analytical approximations with $\eta_j = \eta_0$ for all j and the results of SPECS agent-based simulations for a variety of parameter choices of interest and my experimental conditions. (See Figure A.7.) Therefore, I fit the data using $\eta_j = \eta_0$ for all j , which ensured that I was not overfitting the data.

A.3.1 Fit using previous indirect binding chemotaxis model

In [146], Neumann and others derive the following expression for maltose chemotaxis:

$$f_{Mal} = \log \left[\frac{1 + \frac{[BP]}{\tilde{K}_{I,Mal}} \left(\tilde{p}_0 + \frac{[Mal]}{[Mal] + K_{BP}} \right)}{1 + \frac{[BP]}{\tilde{K}_{A,Mal}} \left(\tilde{p}_0 + \frac{[Mal]}{[Mal] + K_{BP}} \right)} \right]. \quad (\text{A.55})$$

The parameters that minimized J for this model of maltose chemotaxis were the following (Methods):

$$\begin{aligned} \eta_0 &= 11.23, \\ K_{BP} &= 4.61 \mu\text{M}, \\ \tilde{K}_{I,Mal}/[BP] &= 0.703, \\ \tilde{K}_{A,Mal}/[BP] &= 47.4, \\ \tilde{p}_0 &= 0.210, \\ K_{I,MeAsp} &= 20.3 \mu\text{M}, \\ K_{A,MeAsp} &= 399 \mu\text{M}. \end{aligned}$$

However, as you can see from Figure A.3, this is not a good fit. The problem with this model is that it assumes that the free concentration of maltose within the periplasm matches the external concentration of maltose. However, as explained in Section A.1, this assumption is incorrect.

A.3.2 Fit using transport-and-sensing chemotaxis model

I get a better fit when I replace f_{Mal} with the newly derived function, Equation A.48, which takes into account variable maltose transport kinetics.

To fit the data, I write (from Equations A.19 and A.24):

$$[L:BP]_0 = \beta \frac{[L]_{ext}}{\alpha + [L]_{ext}}, \quad (\text{A.56})$$

with

$$\alpha = \frac{K_p V_c [\text{BP}]_{total}}{(V_c - V_p) [\text{BP}]_{total} + K_c V_p} \quad (\text{A.57})$$

$$\beta = \frac{K_c V_p [\text{BP}]_{total}}{(V_c - V_p) [\text{BP}]_{total} + K_c V_p}. \quad (\text{A.58})$$

Thus, my model also has seven free parameters:

$$\eta_0, \alpha, [\text{R}]/\beta, K_{I,Mal}/\beta, K_{A,Mal}/\beta, K_{I,MeAsp}, \text{ and } K_{A,MeAsp}.$$

The parameters that minimized J (Methods) are:

$$\begin{aligned} \eta_0 &= 9.33, \\ \alpha &= 117\mu\text{M}, \\ [\text{R}]/\beta &= 0.0146, \\ K_{I,Mal}/\beta &= 0.00334, \\ K_{A,Mal}/\beta &= 0.00635, \\ K_{I,MeAsp} &= 17.0\mu\text{M}, \\ K_{A,MeAsp} &= 621\mu\text{M}. \end{aligned}$$

A.3.3 Adding a saturation term

Previous research [142] has suggested that steric effects limit *E. coli*'s ability to sense maltose. The maltose-binding protein complex is relatively large so that binding of one complex to a receptor may preclude binding of another complex to a neighboring receptor. Thus, I should get a better fit if I introduce a saturation term. One possible saturation model is the following:

$$[\text{L:BP}]_{eff} = \min([\text{L:BP}]_0, C_0) \quad (\text{A.59})$$

$$[\text{R:L:BP}] = \frac{([\text{R}]_{total} - [\text{R:L:BP}]) ([\text{L:BP}]_{eff} - [\text{R:L:BP}])}{K} \quad (\text{A.60})$$

Using this new model with a saturation term to fit the experimental data, I do get a better fit because the saturation term allows the model to explain the kink in the steady state distribution in single gradients of $20\mu\text{M}$ of maltose:

$$\begin{aligned}M &= 9.25, \\ \alpha &= 131\mu\text{M}, \\ [\text{R}]/\beta &= 0.0129, \\ K_{I,\text{Mal}}/\beta &= 0.00325, \\ K_{A,\text{Mal}}/\beta &= 0.00621, \\ C_0/\beta &= 0.0622, \\ K_{I,\text{MeAsp}} &= 16.9\mu\text{M}, \\ K_{A,\text{MeAsp}} &= 663\mu\text{M}.\end{aligned}$$

See Figure A.4.

A.4 Methods

Cell culture. I used *E. coli* strain RP437, obtained from the lab of JS Parkinson. For a couple of controls for the Western blot analyses, I also used the following two strains from the ASKA library: JW1875-AM (del-tar) and JW3994-AM (del-malE). I grew the cultures overnight in tryptone broth (10 g/L Bacto tryptone, 5 g/L NaCl) in a shaking incubator at 30°C and 300 rpm, then diluted 1:100 the following morning into fresh tryptone broth. I harvested the cells, unless otherwise noted, when the culture reached $\text{OD}_{600} = 0.47$ in the mid-exponential growth phase. The cells had a growth rate of approximately 0.24 h^{-1} . In the one experiment using tryptone broth supplemented with $500\mu\text{M}$ of maltose as the growth medium, the cells instead had a growth rate of approximately 0.29 h^{-1} . Before their use in experiments, I washed the cells twice by centrifuging at $2,000g$ for 5 min and diluted to $\text{OD}_{600} = 0.05$ in motility medium (10 mM potassium phosphate, 0.1 mM EDTA, $1\mu\text{M}$ methionine, 10 mM lactic acid, pH 7) and kept the cells in a 4°C fridge for 30 min.

Experimental setup. I performed the chemotaxis assays using a microfluidic device made by sandwiching an agarose gel layer between a glass microscope slide and a

polydimethylsiloxane (PDMS) layer patterned with three parallel channels. I created the channels by molding the PDMS onto a silicon wafer with positive relief features. I fabricated the device with the following specifications [217]: the channels were 20 mm long, 100 μm deep, and 600 μm wide, with 400 μm spacing between each channel, and the agarose layer was 0.5 mm thick and consisted of a 3% (w/v %) solution of agarose in motility medium.

Each of the three channels contained an inlet and outlet port. The outer two channels functioned as feeder channels within which a steady flow of medium, at a rate of 10 μL per minute, was maintained using a syringe pump (Harvard Apparatus PHD 2000). I set the syringe pump to “refill” mode to create a negative pressure that, along with loosely fitting clips, helped create a seal between the PDMS and agarose. I flowed motility medium with 0-20 μM of maltose (D-maltose monohydrate; Sigma-Aldrich PHR1497) in the left channel and motility medium with 0-460 μM of MeAsp (α -methyl-DL-aspartate; Sigma-Aldrich M6001) in the right channel. As molecules diffuse freely through agarose, the flow of these solutions in the outer channels created constant gradients within the central test channel [7]. For example, with 2 μM of maltose in the left channel and motility medium in the right channel, the cells in the test channel experienced a linear gradient with a slope of approximately $1.4 \times 10^{-3} \mu\text{M}/\mu\text{m}$ and minimum and maximum concentrations of 0.57 μM and 1.43 μM . It is important to note that, while MeAsp is non-metabolizable, maltose is metabolizable and potentially is consumed during assays. However, the constant flow of nutrients in the outer channel and the low concentration of cells within the test channel ensured that any changes to maltose concentration within the test channel due to consumption were negligible.

Initially, the central channel was empty of liquid, so that, after establishing flow in the outer feeder channels, a steady gradient formed within the lower agarose layer over a timescale of $\frac{L^2}{D}$, where L is the length between the two feeder channels and D is the diffusivity of the molecules through agarose. The agarose gel layer has a diffusivity very similar to water, so, for small molecules, $D \approx 10^3 \mu\text{m}^2/\text{s}$. Therefore, with a 1,400 μm spacing between the edges of the two outer channels, gradients formed across the agarose layer in about 30 min.

Forty-five minutes after the establishment of flow in the outer channels, I pipetted cells into the test channel and sealed the test channel using glass microscope coverslips. The

PDMS and agarose completely blocked advection so that there was no apparent active flow in the central channel; any flow in the central channel was a result of pressure differences between the two ends of the channel and was negligible compared to the swimming speed of the cells. As the channel is only 100 μm deep, the upward diffusion of the chemoattractants from the agarose layer reached a steady state in approximately thirty seconds. Thus, the cells quickly experienced a steady gradient in a no-flow environment. The run-and-tumble chemotaxis of *E. coli* yields an effective diffusivity of $D \approx 300 \mu\text{m}^2/\text{s}$, so I began data acquisition twenty minutes after the injection of the cells into the test channel, after the bacterial distribution had reached a steady state.

Data acquisition and analysis. I acquired images of the chemotaxis assays using a Nikon Eclipse TE2000E inverted microscope fitted with a CCD camera. I imaged the cells using phase contrast with a 20 \times objective (N.A. = 0.45). For each experiment, I focused the objective mid-height and took 1-min videos (at 10 frames per second) of at least 10 different 1-mm segments across the entire length of the test channel.

To determine the positions of the bacteria, I analyzed the videos using in-house MATLAB image analysis code that subtracted any non-motile cells. I determined the positions of the bacteria in all frames and thus obtained bacterial position data for 600 frames per location along the channel and at least ten locations per experiment. There were about 100 bacteria per frame. I conducted each experiment 1-3 times. Therefore, I obtained each experimental bacterial distribution from a sample set of approximately $N=1.2$ million bacterial positions. Note that, because I caught the same bacterium on multiple frames, these positions are not independent. For the parameter fitting, I first smoothed the data by fitting a power curve to the obtained empirical distributions: for accumulation toward MeAsp, the power fit is of form $f(x) = ax^n + b$; for accumulation toward maltose, the power fit is of form $f(x) = a(600 \mu\text{m} - x)^n + b$. I chose a power fit because it approximates the analytical expression for the empirical distribution used in my model.

A simple measure to quantify the affinity of cells toward a chemoattractant is the chemotactic migration coefficient (CMC), defined for my setup as

$$\text{CMC} = \frac{\langle x \rangle - 300 \mu\text{m}}{300 \mu\text{m}},$$

where $\langle x \rangle$ is the average position of the cells across the 600 μm channel. Therefore, experiments in which the cells accumulated toward MeAsp (maltose) had $\text{CMC} > 0$ (respectively, $\text{CMC} < 0$). I determined CMC values using only a subset of the bacterial position data—from 3 frames per location (1 frame per 30 seconds)—to obtain position data that were approximately independent of each other, giving a sample set of about $N_1 = 6,000$ bacterial positions per experiment. I used a bootstrapping method to approximate the variance between locations along the channel and between experimental replicates. To bootstrap, I resampled the sample sets of size N_1 with replacement 1,000 times to obtain 1,000 different sample sets. For each sample set, I calculated the CMC value, and the error bars in the CMC plots give the 95% confidence intervals.

Agent-based simulations. To run agent-based simulations of *E. coli* chemotaxis in opposing gradients of maltose and MeAsp, I modified the free-energy difference equations in the original SPECS to incorporate both the HMWC model as well as my transport-and-sensing chemotaxis model. I assumed all additional parameters have the same values as provided in SPECS [90]. These parameters include the time discretization, swimming velocity, tumble time, methylation dynamics parameters, Hill coefficient of the motor response, and average directional change due to Brownian rotational diffusion. In my simulations, when a cell hits the boundary, it modifies its orientation so that it faces away from the boundary with a random angle from a uniform distribution. I used a time step of 0.1 s. To obtain steady-state distributions, I simulated 500 cells for forty minutes of simulated time (24,000 iterations) and averaged their locations over the final twenty minutes of the simulated time. Because I cannot determine n_{Tar} from η_0 , I ran SPECS over different values of n_{Tar} to fit the analytical expression to the simulation results.

System identification. To fit the predicted analytical approximation of the steady-state distribution, I ran MATLAB’s *fmincon* function one-thousand times with different randomly chosen initial conditions for each iteration. The optimization program minimized the following measure of fit:

$$J = \sum_k \sup_x |p_k(x) - f_k(x)|, \quad (\text{A.61})$$

where $f_k(x)$ is the best power fit of the empirical distribution from experiment k . To obtain the fits, I summed over the following single and opposing gradient experiments k : five maltose single gradient experiments (0.2, 2, 4, 8, and 20 μM of maltose), four MeAsp

single gradients (1.15, 4.6, 46, and 460 μM of MeAsp), and eight maltose and MeAsp opposing gradient experiments (2 μM of maltose and 0.46, 2, 4, 6, 8, 10, 46, 460 μM of MeAsp).

Sensitivity analysis. To test the robustness of my model fits given my experimental design, I ran my system identification protocol on 100 “noisy” data sets. To obtain these data sets, I used a new sample size of $\tilde{N} = \frac{N}{1000} \approx 1,200$ for each experiment, selecting bacterial positions at random from the original set of size N without replacement.

Semi-quantitative Western blot. For the Western blots, we followed my cell culture protocol and then, instead of following my experimental protocol, placed the cells in motility medium with various concentrations maltose for 30 min. We then froze 3 mL samples at -80°C until we performed immunoblotting.

To lyse the cells, we added 200 μL of lysis buffer (50 mM Tris, 100 mM NaCL, 0.1% Triton X-100, 250 U/mL benzonase nuclease, and 0.4 mg/mL of lyzosome) into each 3 mL frozen sample, vortexed, and shook tubes for 30 min at 37°C . We diluted 1:4 the samples into loading Laemmli buffer and maintained them at 95°C for 5 min. We loaded 10 μL of each sample into 12% SDS gel and placed it into a Bio-Rad Tetra cell buffer tank to electrophorese at 100 V for 1 h. We then placed the gel into a transfer buffer (25 mM Tris, 190 mM glycine, 20% methanol) for 15 min, and returned the gel sandwiched against a PVDF membrane back into the Tetra cell at 100 V for 1 h to transfer the proteins to the membrane. To prepare the blot for primary antibody incubation, we blocked the blot in blocking buffer (3% BSA in PBST) for 1 h at room temperature. For the primary antibody, we used 1 $\mu\text{g}/\text{mL}$ anti-MBP (unconjugated rabbit polyclonal antibody; LS Bio LS-C355688) in blocking buffer. We poured the antibody solution over the blot, rocked it for 1 h at room temperature, and then placed it in a 4°C fridge overnight. The following morning, we washed it three times in PBST. For the secondary antibody, we used 1:5000 of goat anti-rabbit IgG secondary antibody (Thermo Fisher 65-6120) in blocking buffer. We poured the secondary antibody over the blot, rocked it for 2 h at room temperature, and then washed three times in PBST. We used 1-step Ultra TMB blotting solution to stain the blot, leaving the blot covered at room temperature for about 20 min. After uncovering the blot, we acquired images after 10, 30, and 90 minutes of exposure.

A.5 Supplemental figures

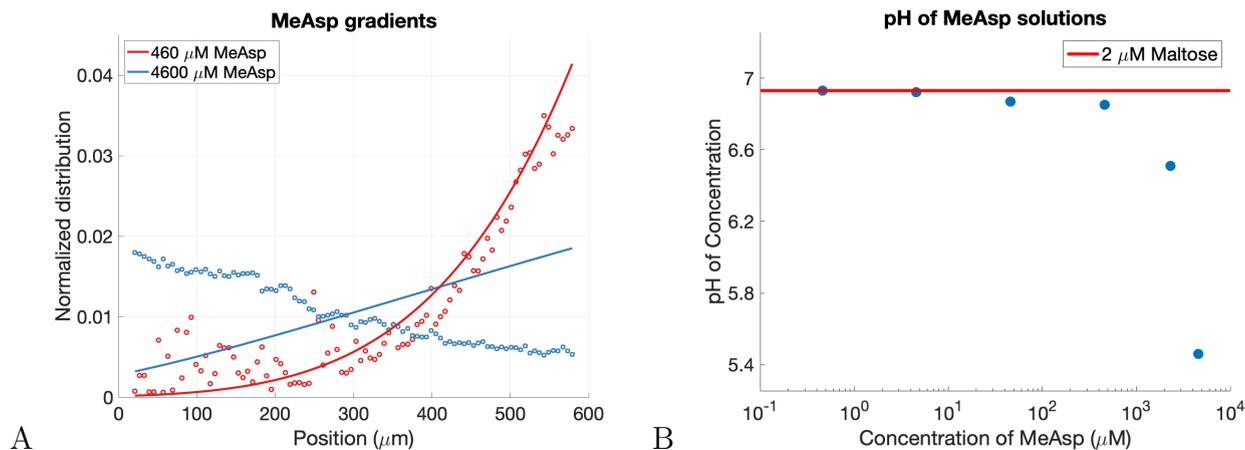


FIGURE A.1: (A) Cells are repelled from high concentrations of MeAsp. Steady-state distributions from experimental chemotaxis assays in gradients of MeAsp along with predicted best-fit using analytical approximation with transport-sensing model. (B) pH of MeAsp solutions used in experiments. Because I suspect pH taxis causes repulsion, I restricted my model fitting to concentrations of MeAsp less than 500 μM .

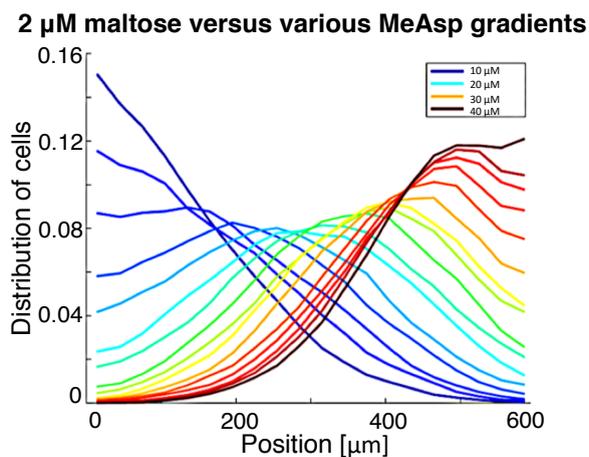


FIGURE A.2: Predicted steady-state distributions of cells using SPECS simulator with HMWC model and Neumann et al. indirect binding model in an opposing gradient of 2 μM of maltose in the left outer channel and various concentrations of MeAsp in the right outer channel.

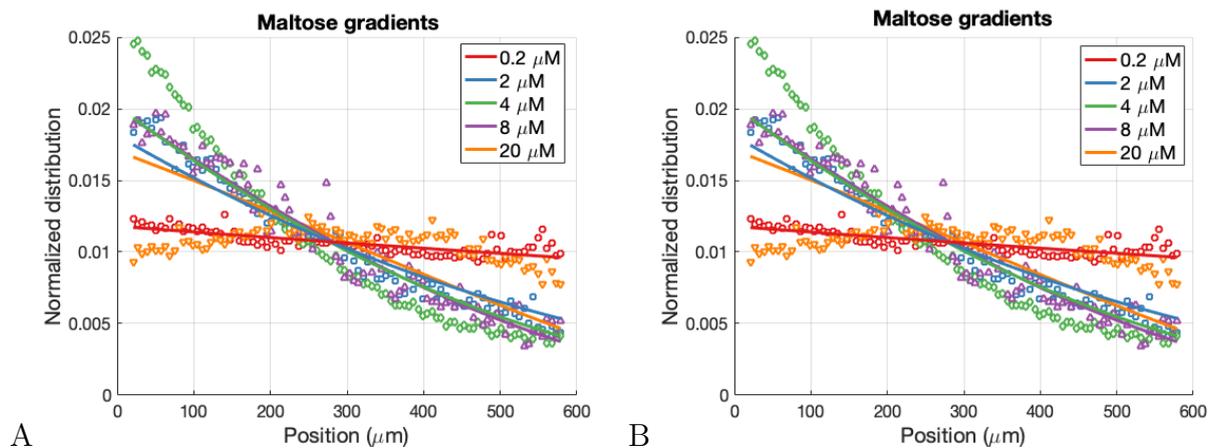


FIGURE A.3: Experimental data and best fit of maltose single gradients using (A) original direct binding model and (B) indirect binding model from Neumann, *et al.*

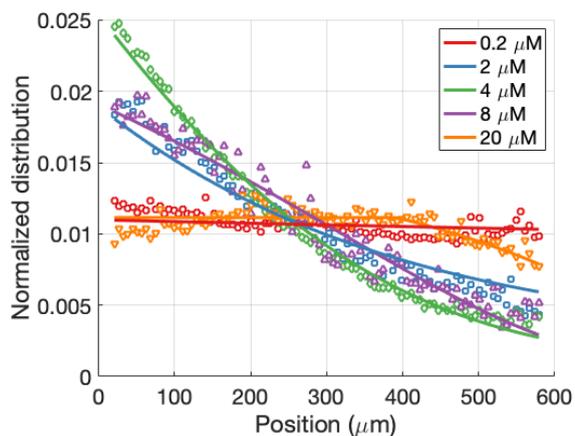


FIGURE A.4: Including a saturation term improves the fit of the steady-state population for 20 μM of maltose.

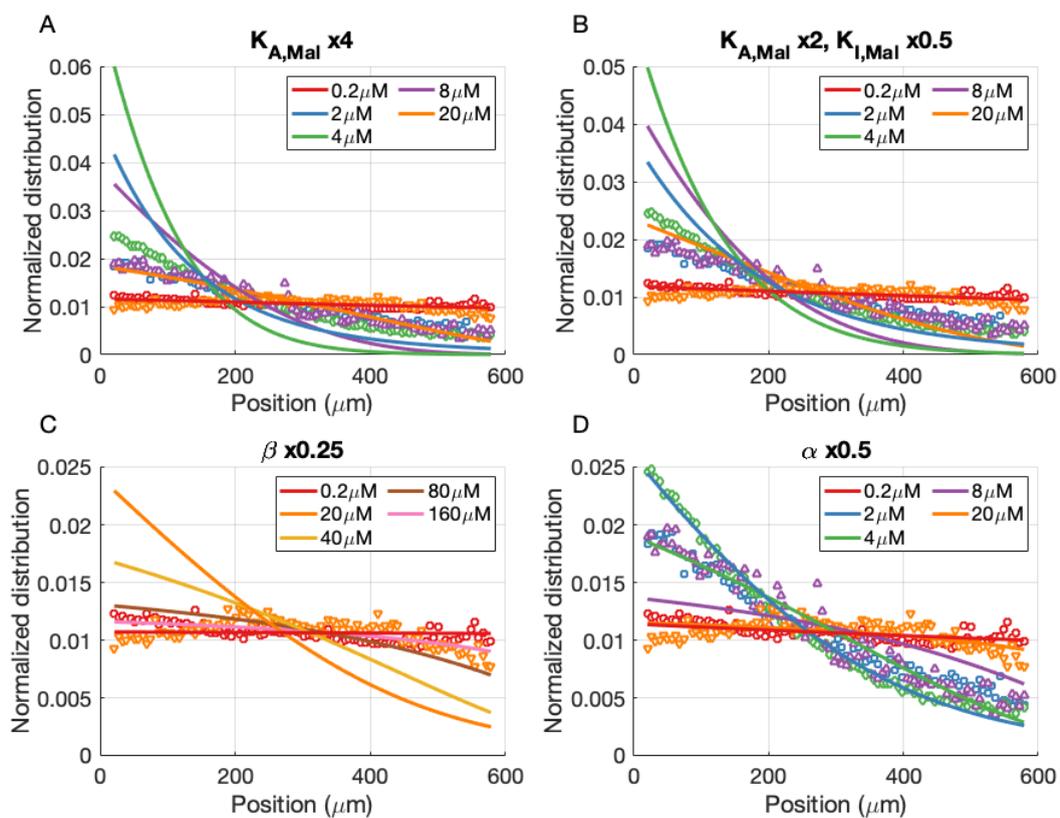


FIGURE A.5: These plots show the effect of modifying the parameter values on the steady-state distribution predictions. Experimental data shown for context.

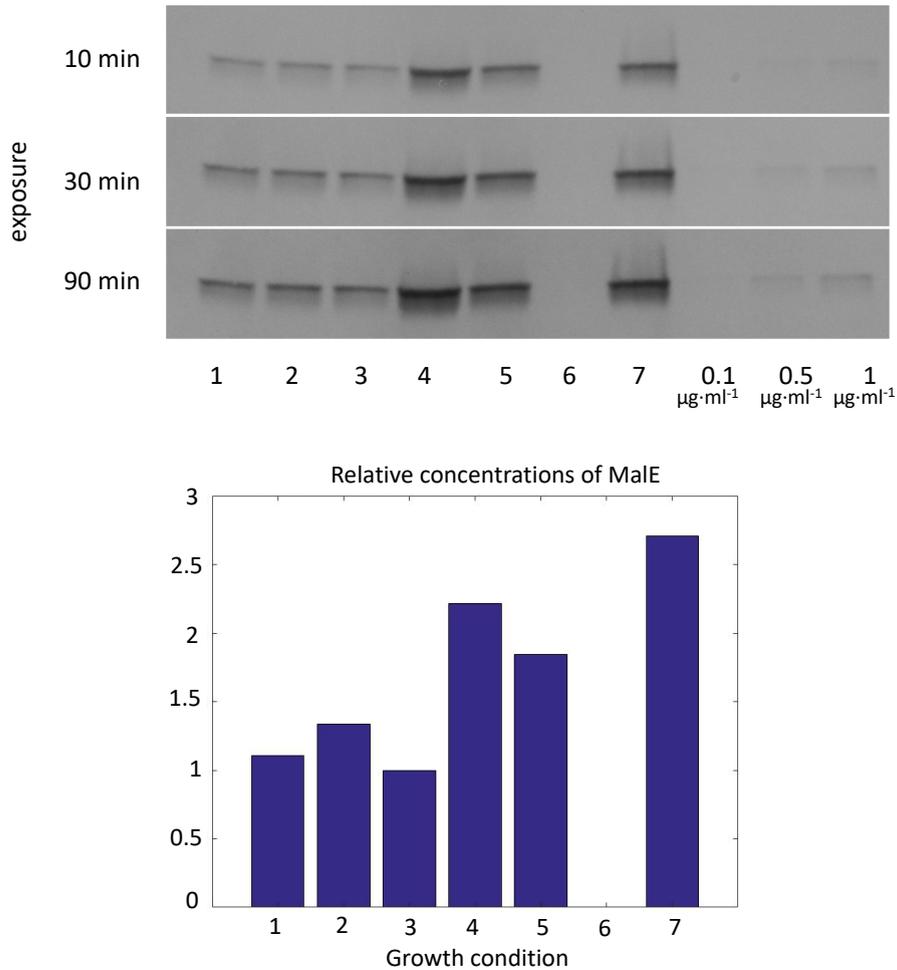


FIGURE A.6: Western blot for MaleE. The bar plot shows the relative concentration of MaleE normalized by total protein concentrations over the following growth and experimental conditions: (1) wild-type cell grown in tryptone, put in 0 maltose, (2) wild-type cell grown in tryptone, put in 1 μM maltose, (3) wild-type cell grown in tryptone, put in 10 μM maltose, (4) wild-type cell grown in tryptone and harvested at $\text{OD}_{600} = 0.9$, put in 1 μM maltose, (5) wild-type cell grown in tryptone with 500 μM maltose, put in 1 μM maltose, (6) del-maleE strain grown in tryptone with 500 μM maltose, put in 1 μM maltose, (7) del-tar strain grown in tryptone with 500 μM maltose, put in 1 μM maltose. (8-10) three MaleE concentrations. The loading volume was 10 μL for all samples.

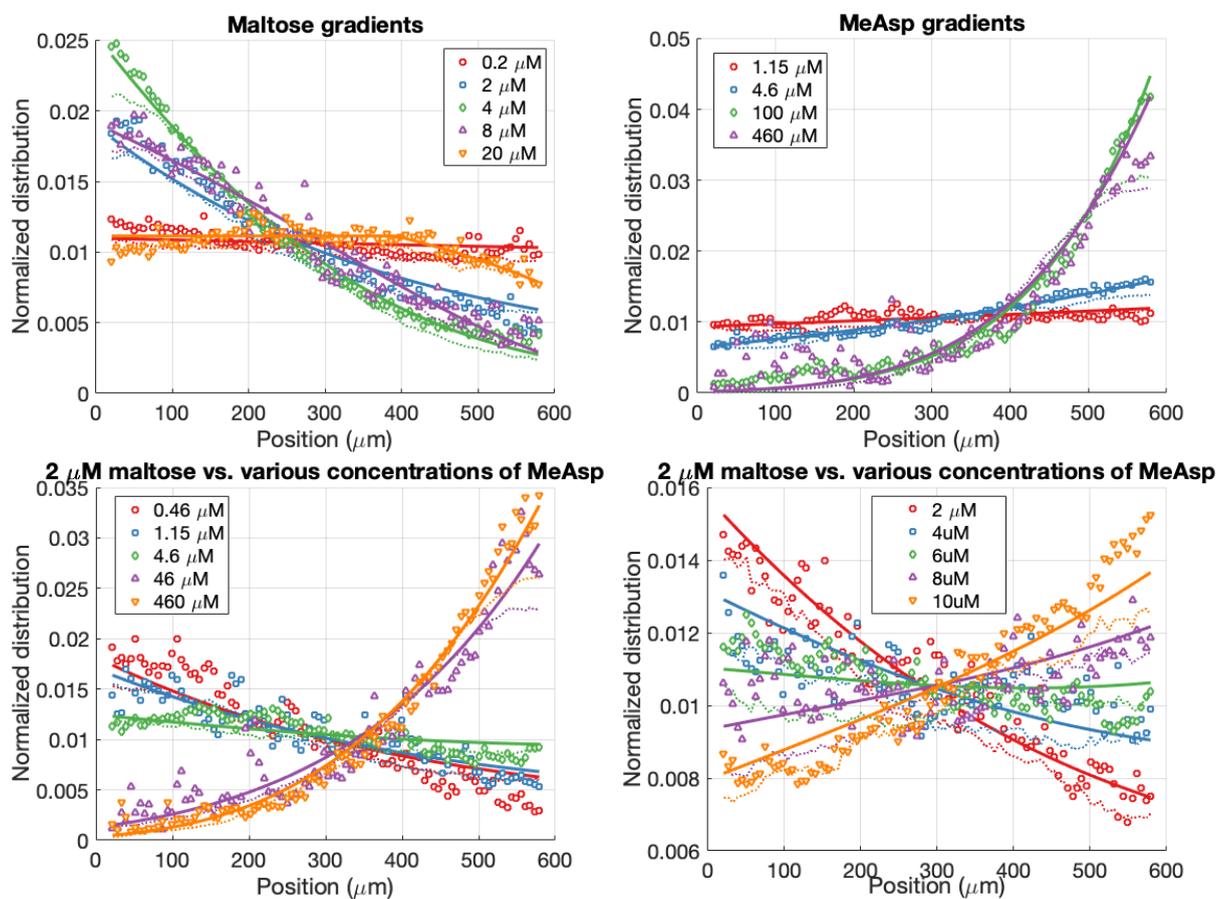


FIGURE A.7: Figures again showing fit with additional saturation term and also showing how analytical equation fits SPEC simulations with $N_{Tar} = 2$. The larger, varied scattered points represent the data; the lines the analytical predictions; and the small dotted lines the results of the SPECS simulations.

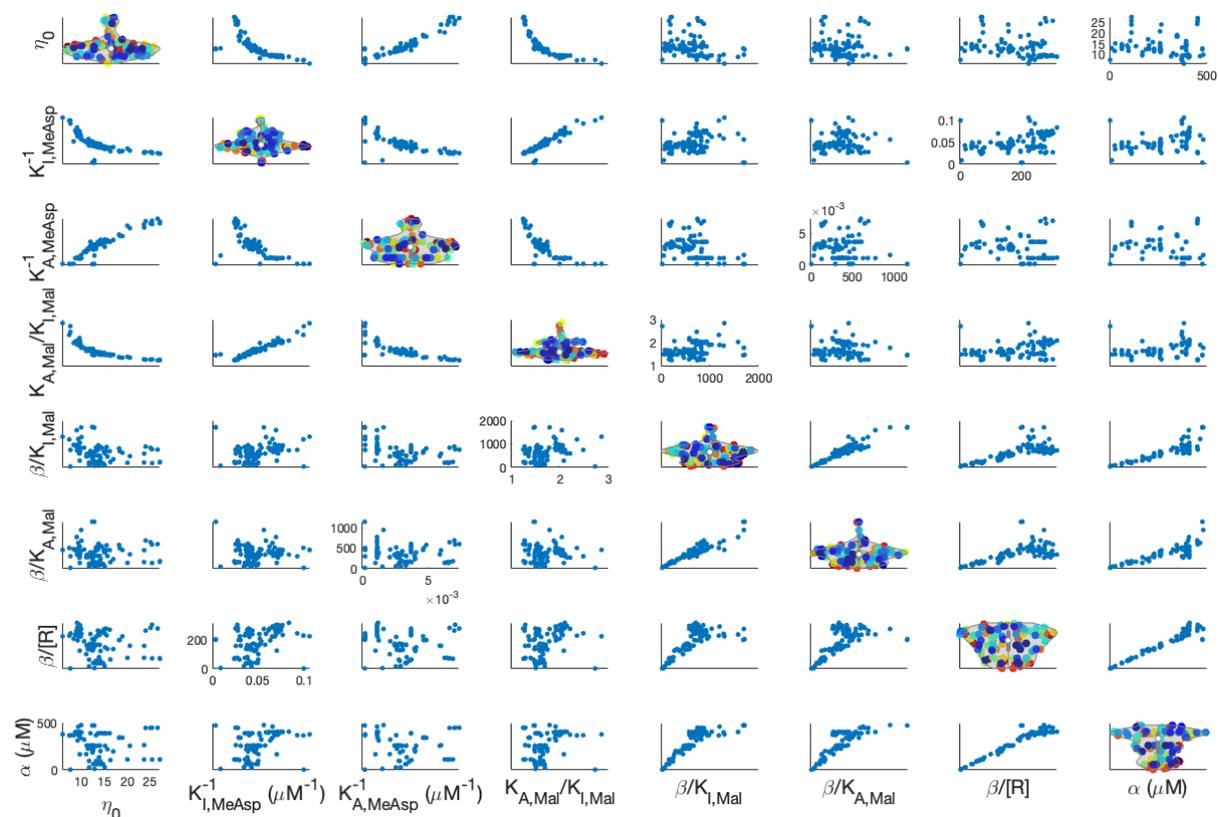


FIGURE A.8: Sensitivity analysis for the fitted model parameters. Plots show the fitted parameter values for 100 “noisy” data sets obtained by randomly selecting 1/1000 of the bacterial position data points without replacement from each experiment. Off-diagonal elements show the correlations between the estimated parameter values. Diagonal elements are violin plots, showing the distribution of estimates (y-axes) obtained for each parameter. Parameters are grouped according to their contribution to the model. The parameters that determine distribution steepness are grouped in the top left portion. The parameters that determine dynamic sensing range are grouped in the bottom right portion.

Appendix B

Proteome Allocation Problem for Nutrient Uptake

B.1 Proteome allocation problem

To compare the trade-offs between direct and ABC transport systems, I incorporate them into a toy, single-cell metabolic model. This model is an extension of the self-replicator model from [138] and is used to solve a proteome allocation problem that maximizes the steady-state growth rate in a homogeneous environment by optimizing the following cell characteristics: cell radius, periplasmic volume fraction, intracellular metabolite concentrations; and fractions of the proteome devoted to particular protein groups. Below I fully specify the problem. The values in blue are the independent variables that the program optimizes over. The values in red are functions of the independent variables. Definitions of each variable are given below.

The optimization problem is to:

maximize μ subject to: Eq.C 1-10, Ineq.C 1-5, and $x_i \geq 0, \forall i$,

where $x = (x_m, \phi, r, f_p, v_{futile}, \mu)$,

$x_m = [[S]_p, [S]_c, [A], [W], [P], [ATP], [ADP]]'$,

$\phi = [\phi_{Porin}, \phi_{BP}, \phi_T, \phi_E, \phi_W, \phi_R, \phi_{ETC}, \phi_{ADP}]'$,

The constraints are:

Eq.C 1-7 (Steady-state growth) : $\frac{d\mathbf{x}_m}{dt} = N\mathbf{v}_r - \boldsymbol{\mu}\mathbf{x}_m = 0,$

$$\mathbf{v}_r = \begin{bmatrix} v_p & (S_{ext} \rightarrow S_p) \\ v_c & (S_p + \text{ATP} \rightarrow S_c + \text{ADP}) \\ k_E[\mathbf{E}] \frac{[S_c]}{K_{M,E} + [S_c]} & (5S_c + 6\text{ATP} \rightarrow 6\text{A} + 6\text{ADP}) \\ k_W[\mathbf{M}] \left(\frac{[\text{A}]}{K_{M,W} + [\text{A}]} \right) & (\text{A} + 4\text{ATP} \rightarrow \text{W} + 4\text{ADP}) \\ k_R[\mathbf{R}] \left(\frac{[\text{A}]}{K_{M,R} + [\text{A}]} \right) \left(\frac{[\text{ATP}]}{K_{M,\text{ATP}} + [\text{ATP}]} \right) & (\text{A} + 4\text{ATP} \rightarrow \text{P} + 4\text{ADP}) \\ k_{ETC}[\mathbf{ETC}] \frac{[S_c]}{K_{M,ETC} + [S_c]} & (2S_c + 35\text{ADP} \rightarrow 35\text{ATP}) \\ k_{ADP}[\mathbf{ADP Syn}] \frac{[S_c]}{K_{M,ADP} + [S_c]} & (10S_c + 60\text{ATP} \rightarrow 66\text{ADP}) \\ v_{futile} & (\text{ATP} \rightarrow \text{ADP}) \end{bmatrix},$$

where the components of the proteome(\mathcal{P}) are: $[\mathbf{P}]_j = \phi_j \alpha_j [\mathbf{P}];$

Eq.C 8 (Proteome allocation) :

$$1 = \sum_{i \in \mathcal{P}} \phi_i;$$

Eq. C 9 (Cell wall) :

$$4\pi (1 + (1 - f_p)^{2/3}) r^2 = [\mathbf{W}] (4/3\pi r^3) a_W;$$

Eq. C 10 (Energy Homeostasis) :

$$[\text{ATP}]/[\text{ADP}] = 10;$$

Ineq. C 1 (Cytoplasmic cell density) :

$$\sum_{j \in \mathcal{M}_{cyto}} m_j x_m(j) \leq \rho;$$

Ineq. C 2 (Periplasmic cell density) :

$$\sum_{j \in \mathcal{M}_{peri}} m_j x_m(j) \leq \rho;$$

Ineq. C 3 (Inner Membrane “Real Estate”) :

$$f_{SA} (4\pi(1 - f_p)^{2/3} r^2) \geq (a_{transporter}[\mathbf{T}] + a_{etc}[\mathbf{ETC}]) (4/3\pi r^3);$$

Ineq. C 4 (Outer membrane “Real Estate”) :

$$f_{SA} (4\pi r^2) \geq (1 - f_p) a_{Porin}[\mathbf{Porin}] (4/3\pi r^3).$$

Note that all components of the proteome are in units of abundance divided by cytoplasmic volume, $[P]_i$. The periplasmic concentration of the binding proteins is thus, $(1 - f_p)[BP]/f_p$, where f_p is the fraction of the cell volume comprised of the periplasm.

This proteome allocation problem optimizes over the following variables to determine the maximal exponential growth rate, μ , given a particular extracellular carbon substrate concentration, $[S]_{ext}$:

- the intracellular metabolite concentrations, x_m , which include
 - $[S]_p$, the periplasmic concentration of generic carbon substrate;
 - $[S]_c$, the cytoplasmic concentration of generic carbon substrate;
 - $[A]$, the cytoplasmic concentration of the precursor building block of both proteins and the cell membranes, presumed to be a generic amino acid;
 - $[W]$, the number of generic cell membrane units divided by the cytoplasmic volume of the cell;
 - $[P]$, the number of amino acids incorporated into proteins divided by the cytoplasmic volume of the cell;
 - $[ATP]$, the cytoplasmic concentration of ATP in the cell; and
 - $[ADP]$, the cytoplasmic concentration of ADP in the cell. I assume, for energy homeostasis, that $[ATP] = 10[ADP]$;
- the proteome fractions, ϕ , which include
 - ϕ_{Porin} , the fraction of the proteome comprised of outer membrane-bound porins, from which the extracellular substrate, S_{ext} , passively diffuses into the cell periplasm;
 - ϕ_{BP} , the fraction of the proteome comprised of periplasmic binding proteins that enable ABC transport (for direct transport, $\phi_{BP} = 0$);
 - ϕ_T , the fraction of the proteome comprised of inner membrane-bound transport units that translocate the periplasmic carbon substrate, S_p into the cytoplasm;
 - ϕ_E , the fraction of the proteome comprised of catabolic enzymes that transform the cytoplasmic substrate, S_c , into a generic amino acid building block, A;

- ϕ_W , the fraction of the proteome comprised of membrane biosynthesis enzymes that build both the inner and outer membranes, W, from the building block A;
 - ϕ_R , the fraction of the proteome comprised of ribosomes that use building block A to elongate peptide chains and thus express the proteome, P;
 - ϕ_{ETC} , the fraction of the proteome comprised of ATP synthase and electron transport chain units that use aerobic respiration to transform carbon, S_c , into energy by converting ADP to ATP; and
 - ϕ_{ADP} , the fraction of the proteome comprised of ADP synthesis enzymes that use the carbon substrate S_c to make ADP;
- the cell radius, r , assuming a spherical cell;
 - the periplasmic volume fraction, f_p ; and
 - a rate of a futile cycle that converts ATP to ADP, v_{futile} , which ensures that a solution exists with balanced rates of energy generation and consumption and which can additionally be used to account for energetic consumption rates that do not scale with growth (which is how I use these rates in Chapter 5).

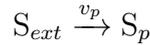
To solve for these optimal parameter values, the optimization problem is subject to a variety of constraints:

- The equality constraints 1-7 are ordinary differential equations that assume a balanced, steady-state exponential growth of each of the cellular components. N is thus a stoichiometry matrix that describes how many of each input metabolite is needed to form an output, and v_r is a vector of Michaelis-Menten reaction rates. Explanations for each parameter choice for the stoichiometry matrix and reaction rates are provided below.
- Equality constraint 8 specifies that the proteome fractions must sum to one.
- Equality constraint 9 specifies that the amount of cell membrane units expressed must cover both the inner and outer membranes of the cell.
- Equality constraint 10 specifies that, for energy homeostasis, the abundance of ATP must be ten times greater than the abundance of ADP.

- Inequality constraint 1 specifies that the sums of the molecular weights of each cytoplasmic metabolite (S_c , A, $(\phi_E + \phi_W + \phi_R + \phi_{ADP})P$, ATP, ADP) multiplied by its corresponding cytoplasmic concentration must be less than or equal to the maximal allowed density, ρ .
- Inequality constraint 2 specifies that the sums of the molecular weights of each periplasmic metabolite (S_p , ϕ_{BPP}) multiplied by its corresponding periplasmic concentration must be less than or equal to the maximal allowed density, ρ .
- Inequality constraint 3 specifies that the inner-membrane bound protein groups (the transport units, T, and the electron transport chains, ETC) must fit on the surface area of the inner membrane.
- Inequality constraint 4 specifies that the porins must fit on the surface area of the outer membrane.

To determine reasonable parameter values, I use glucose consumption by *E. coli* as a model uptake system.

B.1.1 Periplasmic uptake



The rate of the uptake of the carbon substrate in the extracellular environment with concentration $[S]_{ext}$ into the periplasm is derived in Chapter 4.1.2. It is given by the following equation, where $[S]_p$ is the concentration of substrate in the periplasm; r is the cell's radius; $[Porin]$ is the number of porins divided by the cytoplasmic volume of the cell; D is the diffusivity of the substrate in the external environment; k_{Porin} is a porin permeability rate; and f_p is the fraction of the volume of the cell that is comprised of the periplasm.

$$v_p = (1 - f_p)k_{Porin}[Porin] \frac{[S]_{ext} - [S]_p}{f_p \left(1 + \frac{(1-f_p)k_{Porin}[Porin]}{3Dr^{-2}} \right)} \quad \left[\frac{\text{Periplasmic Concentration}}{\text{Time}} \right]$$

In my baseline model, I use the following values:

- $D = 10^{-8} \text{ cm}^2 \text{ msec}^{-1}$, which is approximately the diffusivity of a glucose molecule in water.
- $k_{porin} = 10^{-11} \text{ cm}^3 \text{ msec}^{-1}$. I estimate the permeability rate of a glucose molecule through the general porin OmpF in *E. coli* to be the following [147]:

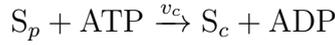
$$k_{Porin} = \left(\frac{D}{d}\right) \left(\frac{a}{a_0}\right) a_{porin}, \quad \left[\frac{\text{cm}^3}{\text{sec}}\right]$$

where the thickness of the outer membrane is $d = 4 \times 10^{-7} \text{ cm}$; (a/a_0) is the Renkin correction factor that accounts for the ratio of the solute size and porin size; and the surface area of a single porin is $a_{porin} = 1 \text{ nm}^2$.

- $\alpha_{Porin} = 1/400$. The *E. coli* general porin OmpF is comprised of 362 amino acids.

I run the optimization problem over a range of external concentrations, from $[S]_{ext} = 10^{-6}$ - 100 mM.

B.1.2 Cytoplasmic uptake



I run the optimization problem using one of two models for cytoplasmic uptake. Both models are explained and derived in Chapter 4.1.3. As explained in Chapter 4.2.2, I assume, in the baseline model, that both ABC and direct transport have the same energetic cost of 1 ATP molecule per substrate translocated.

B.1.2.1 Direct transport

For direct transport, I use the following equation:

$$v_{c,direct} = k_{3'} [\text{T}] \frac{[S]_p}{K_{T'} + [S]_p},$$

where $[\text{T}]$ is the number of membrane-bound transport units divided by the cytoplasmic volume of the cell, k_3 is the turnover rate of the transport-unit, and $K_{T'} (= k_{3'}/k_{2'})$ is the half-saturation constant of the transport-unit.

In the baseline model, I use:

- $k_{3'} = 200 \text{ sec}^{-1}$. The turnover rate of the glucose transport ptsI is approximately 210 molecules/sec. (BNID:103693, [1])
- $K_{T'} = 10 \mu\text{M}$. The effective half-saturation constant for glucose transport was observed to be between 5 - 20 μM [89].
- $\alpha_T = 1/1562$. The number of amino acids of each protein comprising the system are: 831 for PtsI, 85 for PtsH, 477 for PtsG, and 169 for Crr ([3], eco:02060).

B.1.2.2 ABC transport

For ABC transport, I use the following equation

$$v_{c,ABC} = \frac{k_3}{2(k_3 + k_{1f}[S]_p)} \left([k_{1r}K_T + k_{1f}[S]_p ([\text{BP}]_{\text{peri}} + [\text{T}]_{\text{peri}} + K_T) + k_3[\text{T}]_{\text{peri}}] \dots \right. \\ \left. - \sqrt{4[\text{BP}]_{\text{peri}}k_{1f}K_T[S]_p(k_{1r} + k_{1f}[S]_p) + (k_{1r}K_T - [\text{BP}]_{\text{peri}}k_{1f}[S]_p + k_2K_T[\text{T}]_{\text{peri}} + k_{1f}[S]_p(K_T + [\text{T}]))^2} \right),$$

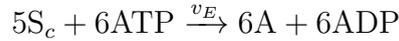
where $[\text{BP}]_{\text{peri}}$ is the concentration of binding proteins in the periplasm; $[\text{T}]_{\text{peri}}$ is the number of membrane-bound transport units divided by the volume of the periplasm; and the transport rates and half-saturation constants are as defined in Chapter 4.1.3.

In the baseline model, I use the following values:

- $k_3 = 2 \text{ sec}^{-1}$. As explained in Chapter 4.3.1, the diffusivity of the binding protein-substrate complex is approximated to be one hundred times slower than the diffusivity of the substrate itself. Thus, I approximate that the transport rate for ABC transport is one hundred times lower than that of direct transport.
- $K_T = 10 \mu\text{M}$. I assume that the transporter dissociation constants for direct and ABC transport are the same.
- $K_{BP} = 1 \mu\text{M}$ ($k_{1f} = 1 \times 10^5 \text{ mM/sec}$ and $k_{1r} = 100 \text{ sec}^{-1}$). The *E. coli* maltose binding protein, MalE, has a dissociation constant of about $K_{BP} = 1 \mu\text{M}$ [80]. I estimated the association and dissociation rates using results from [132].

- $\alpha_{BP} = 1/400$. MalE is comprised of 396 amino acids.
- α_T and a_T are the same as for direct transport.

B.1.3 Catabolism



The set of metabolic enzymes, E, transform the intracellular nutrient, S_c , into amino acids at a rate of

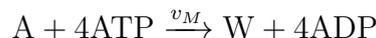
$$v_E = k_E [E] \frac{[S]_c}{K_{M,E} + [S]_c}.$$

For simplicity, I consider proteins to be made of just one type of amino acid and use the following parameter values:

- $k_E = 100 \text{ sec}^{-1}$. The phosphofructokinase reaction is a rate-limiting step of glycolysis ([117]) and has a turnover rate on the order of approximately 100 per second (EC 2.7.1.11, [2]).
- $K_E = 100 \mu\text{M}$. This is the approximate half-saturation constant of phosphofructokinase-2 for fructose 6-phosphate [2].
- $\alpha_E = \frac{1}{40,000}$. There are about 130 genes in glycolysis and the biosynthesis of amino acids ([3], M00002, eco01230).

As shown in [93], energy is created during the breakdown of carbon substrates and used during the biosynthesis of amino acids. The different amino acids have different carbon and ATP consumption costs. For simplicity, I average over all of the amino acids and assume that 5 glucose molecules provide enough carbon to produce 6 amino acids and that there is a net consumption of 1 ATP molecule per amino acid produced.

B.1.4 Membrane biosynthesis



The set of membrane biosynthesis enzymes, M, transform the amino acid building block into the components of both the inner and outer membrane of the cell at a rate of

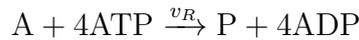
$$v_M = k_M [\text{M}] \frac{[\text{A}]}{K_M + [\text{A}]}.$$

In my model, I use the following baseline parameter values:

- $k_M = 3 \text{ sec}^{-1}$. This is the catalytic rate of penicillin-binding protein 1 [196].
- $K_M = 20 \text{ }\mu\text{M}$. This is the approximate half-saturation constant of MurC for L-alanine [4].
- $\alpha_M = \frac{1}{21,000}$. There are 19 enzymes involved in peptidoglycan biosynthesis ([3], eco:01011).

I take as the unit for the cell membrane a single amino acid and assume that the energetic cost of adding an additional amino acid to the cell wall is equivalent to the energetic cost of adding an amino acid to a polypeptide.

B.1.5 Protein expression



The ribosomes, R, build the proteome, P, from the intracellular pool of the building block amino acid at a rate of

$$v_R = k_R [\text{R}] \frac{[\text{A}]}{K_R + [\text{A}]}.$$

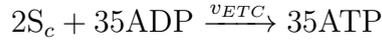
Note that I take [P] to be the concentration of amino acids that comprise the proteome. I use the following baseline parameter values in the model:

- $k_R = 15 \text{ sec}^{-1}$. The peptide chain elongation rate is from about 12 to 20 amino acids per second per ribosome ([1], BNID:107868).

- $K_R = 300 \mu\text{M}$. Estimates within a range of $200 - 400 \mu\text{M}$ have been made for the half-saturation constant of several aminoacyl-tRNA synthetases to various amino acids [213].
- $\alpha_R = \frac{1}{20,000}$. A ribosome is comprised of both ribosomal protein (approximately 7,000 amino acids) and rRNA (approximately 5,000 ribonucleotides) [112].

I take as the unit of proteome a single amino acid. I approximate that the energetic cost of adding one amino acid to a peptide chain is 4 ATP [25].

B.1.6 Energy generation



I assume that the cell is aerobic and generates energy via glycolysis and respiration. The enzymes involved in glycolysis, the electron transport chain and ATP synthase (ETC), are required to generate ATP from the intracellular carbon substrate at the following rate:

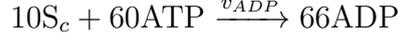
$$v_{ETC} = k_{ETC} [\text{ETC}] \frac{[S]_c}{K_{ETC} + [S]_c}.$$

I use the following baseline parameters:

- $k_{ETC} = 300 \text{ sec}^{-1}$. Previous work estimates that the rate of ATP synthesized by ATP synthesis is approximately 270 per second [52].
- $K_{ETC} = 100 \mu\text{M}$. I approximate the half-saturation constant for energy generation to be the same as for catabolism, K_E .
- $\alpha_{ETC} = \frac{1}{45,000}$. I approximate the proteomic cost of glycolysis as well as of oxidative phosphorylation ([3], M00002, eco00190).

Respiration requires an electron acceptor, such as oxygen, so that the electron transport chain can create a transmembrane ion potential gradient, which ATP synthase utilizes to produce ATP. In glycolysis, 2ATP, 2 NADH, and 2 pyruvate molecules are generated from 1 glucose molecule. I assume that the energy of 1 NADH is equivalent to 1 ATP [154], giving 17.5 ATP per glucose molecule generated under respiration.

B.1.7 ADP synthesis



The ADP synthesis proteins, ADP, transform the intracellular pool of carbon substrate into ADP at the following rate:

$$v_{ADP} = k_{ADP}[\text{ADP Syn}] \frac{[S]_c}{K_{M,ADP} + [S]_c}.$$

I use the following baseline parameter values in the model:

- $k_{ADP} = 100 \text{ sec}^{-1}$. I approximate the turnover rate to be the same as for catabolism.
- $K_{ADP} = 100 \text{ } \mu\text{M}$. I approximate the half-saturation constant for energy generation to be the same as for catabolism, K_E .
- $\alpha_{ADP} = \frac{1}{30,000}$. I approximate the proteomic cost of ADP synthesis using the genes in purine metabolism ([3], ko00230).

Because a glucose molecule contains 6 carbons and an ADP molecule contains 10 carbons, I assume a stoichiometry of an input of $10S_c$ needed for 6 ADP molecules. I approximate that the energetic cost of synthesizing one ADP is identical to that of synthesizing one nucleotide, which was estimated to be about 10 ADP molecules per nucleotide [145].

B.1.8 Density and surface area constraints

The abundances of the intracellular components are constrained by either the maximal density of the cell or, for the membrane-bound components, the surface area of the cell. For Inequality Constraints 2-4, I use the following baseline parameters:

- $\rho = 0.3 \text{ g/mL}$. The density of a cell is approximately 0.34 g/mL ([1], BNID:109049).
- $m_A = 110 \text{ Da}$. The average molecular weight of an amino acid is 110 Da.
- $m_{S_c} = 180 \text{ Da}$. The molecular weight of glucose is 180 Da.

- $m_{ATP} = 500$ Da. The molecular weight of ATP is 507 Da.
- $m_{ADP} = 430$ Da. The molecular weight of ADP is 427 Da.
- $m_W = 220$ Da. The molecular weight of N-acetylglucosamine is 221 Da.
- $a_{transporter} = 50$ nm². The surface area of one transporter equals approximately 50nm² [133].
- $a_{porin} = 1$ nm². The approximate surface area of a single OmpF molecule is 1 nm² [147].
- $a_W = 2$ nm². The surface area of one unit of peptidoglycan is approximately 2 nm² [211].
- $a_{ETC} = 100$ nm². The surface area of one electron transport chain unit is approximately 84 nm² [197].

For the baseline model, I assume that half of the surface areas of both the inner and outer membranes are available for proteins, $f_{SA} = 0.5$.

B.2 Solving the optimization problem

To solve the proteome allocation problem, I use MATLAB's constrained nonlinear multi-variable function solver, *fmincon*. To ensure that the solutions are global and not simply local ones, I run the solver 50 times for each optimization problem, each time using different initial conditions. I additionally transform the units of both the constraints and variables so that their magnitudes are all approximately 1.

B.2.1 A note on calculating maximal rate of downstream enzymes

In my model, direct transport systems have very high optimal ratios of maximal uptake rate to maximal rate of downstream enzymes at low nutrient concentrations, while, for

ABC transport, this ratio is near unity. To calculate this ratio, I use the following:

$$\frac{V_{max,uptake}}{V_{max,downstream}} = \frac{f_p V_{max,c}/(1 - f_p)}{5k_E[E] + 2k_{ETC}[ETC] + 10k_{ADP}[ADP \text{ Syn}]}$$

Bibliography

- [1] *BioNumbers - The Database of Useful Biological Numbers*.
<http://bionumbers.hms.harvard.edu>.
- [2] *Enzyme Database - BRENDA*. <https://www.brenda-enzymes.org>.
- [3] *KEGG PATHWAY Database*. <http://www.genome.jp/kegg>.
- [4] *UniProtKB*. <http://www.uniprot.org/uniprot/P17952>.
- [5] J. ADLER, *Chemotaxis in bacteria*, Science, 153 (1966), pp. 708–716.
- [6] T. AHMED, *Microfluidics for bacterial chemotaxis*, thesis, Massachusetts Institute of Technology, 2011.
- [7] T. AHMED, T. S. SHIMIZU, AND R. STOCKER, *Bacterial Chemotaxis in Linear and Nonlinear Steady Microfluidic Gradients*, Nano Letters, 10 (2010), pp. 3379–3385.
- [8] A. ALLDREDGE, *The carbon, nitrogen and mass content of marine snow as a function of aggregate size*, Deep Sea Research Part I: Oceanographic Research Papers, 45 (1998), pp. 529–541.
- [9] A. L. ALLDREDGE AND C. C. GOTSCHALK, *Direct observations of the mass flocculation of diatom blooms: characteristics, settling velocities and formation of diatom aggregates*, Deep Sea Research Part A. Oceanographic Research Papers, 36 (1989), pp. 159–171.
- [10] U. ALON, *An Introduction to Systems Biology: Design Principles of Biological Circuits*, CRC Press, July 2006.
- [11] C. ANDERSEN, M. JORDY, AND R. BENZ, *Evaluation of the rate constants of sugar transport through maltoporin (LamB) of Escherichia coli from the sugar-induced current noise*, The Journal of General Physiology, 105 (1995), pp. 385–401.
- [12] R. A. ARMSTRONG, *Nutrient uptake rate as a function of cell size and surface transporter density: A Michaelis-like approximation to the model of Pasciak and Gavis*, Deep Sea Research Part I: Oceanographic Research Papers, 55 (2008), pp. 1311–1317.

- [13] F. AZAM AND F. MALFATTI, *Microbial structuring of marine ecosystems*, Nature Reviews Microbiology, 5 (2007), pp. 782–791.
- [14] BAKER RUTH E., PEÑA JOSE-MARIA, JAYAMOHAN JAYARATNAM, AND JÉRUSALEM ANTOINE, *Mechanistic models versus machine learning, a fight worth fighting for the biological community?*, Biology Letters, 14 (2018), p. 20170660.
- [15] G. M. BARBARA AND J. G. MITCHELL, *Bacterial tracking of motile algae*, FEMS Microbiology Ecology, 44 (2003), pp. 79–87.
- [16] M. BASAN, *Resource allocation and metabolism: the search for governing principles*, Current Opinion in Microbiology, 45 (2018), pp. 77–83.
- [17] M. BASAN, S. HUI, H. OKANO, Z. ZHANG, Y. SHEN, J. R. WILLIAMSON, AND T. HWA, *Overflow metabolism in Escherichia coli results from efficient proteome allocation*, Nature, 528 (2015), pp. 99–104.
- [18] B. L. BASSLER, C. YU, Y. C. LEE, AND S. ROSEMAN, *Chitin utilization by marine bacteria. Degradation and catabolism of chitin oligosaccharides by Vibrio furnissii*, Journal of Biological Chemistry, 266 (1991), pp. 24276–24286.
- [19] S. BEIER AND S. BERTILSSON, *Bacterial chitin degradation—mechanisms and eco-physiological strategies*, Frontiers in Microbiology, 4 (2013).
- [20] R. BENZ, ed., *Bacterial and Eukaryotic Porins*, Wiley-VCH Verlag, 2004.
- [21] H. C. BERG, *E. coli in Motion*, Springer-Verlag, New York, 2004.
- [22] —, *The Gain Paradox*, Progress in Biophysics and Molecular Biology, 100 (2009), pp. 2–3.
- [23] —, *Chemotaxis*, in Quantitative Biology: From Molecular to Cellular Systems, CRC Press, Boca Raton, 2013, pp. 341–363.
- [24] H. C. BERG AND D. A. BROWN, *Chemotaxis in Escherichia coli analysed by three-dimensional tracking*, Nature, 239 (1972), pp. 500–504.
- [25] J. M. BERG, J. L. TYMOCZKO, AND L. STRYER, *Biochemistry*, W H Freeman, 5th ed., 2002.
- [26] R. M. BERRY AND J. P. ARMITAGE, *The Bacterial Flagella Motor*, in Advances in Microbial Physiology, R. K. Poole, ed., vol. 41, Academic Press, Jan. 1999, pp. 291–337.
- [27] N. BLACKBURN, T. FENCHEL, AND J. MITCHELL, *Microscale Nutrient Patches in Planktonic Habitats Shown by Chemotactic Bacteria*, Science, 282 (1998), pp. 2254–2256.

- [28] S. M. BLOCK, J. E. SEGALL, AND H. C. BERG, *Impulse responses in bacterial chemotaxis*, *Cell*, 31 (1982), pp. 215–226.
- [29] —, *Adaptation kinetics in bacterial chemotaxis*, *Journal of Bacteriology*, 154 (1983), pp. 312–323.
- [30] W. BOOS AND A. BÖHM, *Learning new tricks from an old dog: MalT of the Escherichia coli maltose system is part of a complex regulatory network*, *Trends in Genetics*, 16 (2000), pp. 404–409.
- [31] W. BOOS AND H. SHUMAN, *Maltose/maltodextrin system of Escherichia coli: transport, metabolism, and regulation*, *Microbiology and Molecular Biology Reviews*, 62 (1998), pp. 204–229.
- [32] E. BOSDRIESZ, S. MAGNÚSDÓTTIR, F. J. BRUGGEMAN, B. TEUSINK, AND D. MOLENAAR, *Binding proteins enhance specific uptake rate by increasing the substrate-transporter encounter rate*, *FEBS Journal*, 282 (2015), pp. 2394–2407.
- [33] E. BOSDRIESZ, M. T. WORTEL, J. R. HAANSTRA, M. J. WAGNER, P. D. L. T. CORTÉS, AND B. TEUSINK, *Low affinity uniporter carrier proteins can increase net substrate uptake rate by reducing efflux*, *Scientific Reports*, 8 (2018), p. 5576.
- [34] P. BOYD AND T. TRULL, *Understanding the export of biogenic particles in oceanic waters: Is there consensus?*, *Progress in Oceanography*, 72 (2007), pp. 276–312.
- [35] D. BRAY, *Bacterial chemotaxis and the question of gain*, *Proceedings of the National Academy of Sciences of the United States of America*, 99 (2002), pp. 7–9.
- [36] D. BRAY, M. D. LEVIN, AND C. J. MORTON-FIRTH, *Receptor clustering as a cellular mechanism to control sensitivity*, *Nature*, 393 (1998), pp. 85–88.
- [37] D. A. BROWN AND H. C. BERG, *Temporal stimulation of chemotaxis in Escherichia coli*, *Proceedings of the National Academy of Sciences*, 71 (1974), pp. 1388–1392.
- [38] A. B. BURD, S. FREY, A. CABRE, T. ITO, N. M. LEVINE, C. LØNBORG, M. LONG, M. MAURITZ, R. Q. THOMAS, B. M. STEPHENS, T. VANWALLEGHEM, AND N. ZENG, *Terrestrial and marine perspectives on modeling organic matter degradation pathways*, *Global Change Biology*, 22 (2016), pp. 121–136.
- [39] D. K. BUTTON, *Biochemical Basis for Whole-Cell Uptake Kinetics: Specific Affinity, Oligotrophic Capacity, and the Meaning of the Michaelis Constant*, *Applied and Environmental Microbiology*, 57 (1991), pp. 2033–2038.
- [40] —, *Nutrient Uptake by Microorganisms according to Kinetic Parameters from Theory as Related to Cytoarchitecture*, *Microbiology and Molecular Biology Reviews*, 62 (1998), pp. 636–645.

- [41] D. K. BUTTON, B. ROBERTSON, E. GUSTAFSON, AND X. ZHAO, *Experimental and Theoretical Bases of Specific Affinity, a Cytoarchitecture-Based Formulation of Nutrient Collection Proposed To Supersede the Michaelis-Menten Paradigm of Microbial Kinetics*, Applied and Environmental Microbiology, 70 (2004), pp. 5511–5521.
- [42] J. A. CHRISTIE-OLEZA, B. FERNANDEZ, B. NOGALES, R. BOSCH, AND J. ARMENGAUD, *Proteomic insights into the lifestyle of an environmentally relevant marine bacterium*, The ISME Journal, 6 (2012), pp. 124–135.
- [43] M. T. COTTRELL AND D. L. KIRCHMAN, *Transcriptional Control in Marine Copiotrophic and Oligotrophic Bacteria with Streamlined Genomes*, Appl. Environ. Microbiol., 82 (2016), pp. 6010–6018.
- [44] A. L. DAVIDSON, E. DASSA, C. ORELLE, AND J. CHEN, *Structure, Function, and Evolution of Bacterial ATP-Binding Cassette Systems*, Microbiology and Molecular Biology Reviews, 72 (2008), pp. 317–364.
- [45] R. DAWKINS, *The Selfish Gene: 30th Anniversary edition*, OUP Oxford, Mar. 2006.
- [46] G. DECKERS-HEBESTREIT, J.-C. GREIE, W.-D. STALZ, AND K. ALTENDORF, *The ATP synthase of Escherichia coli: structure and function of F₀ subunits*, Biochimica et Biophysica Acta (BBA) - Bioenergetics, 1458 (2000), pp. 364–373.
- [47] R. DIPPEL AND W. BOOS, *The Maltodextrin System of Escherichia coli: Metabolism and Transport*, Journal of Bacteriology, 187 (2005), pp. 8322–8331.
- [48] T. A. J. DUKE AND D. BRAY, *Heightened sensitivity of a lattice of membrane receptors*, Proceedings of the National Academy of Sciences, 96 (1999), pp. 10104–10108.
- [49] D. B. DUSENBERY, *Minimum size limit for useful locomotion by free-swimming microbes*, Proceedings of the National Academy of Sciences, 94 (1997), pp. 10949–10954.
- [50] —, *Living at Micro Scale: The Unexpected Physics of Being Small*, Harvard University Press, 2009.
- [51] D. W. ERICKSON, S. J. SCHINK, V. PATSALO, J. R. WILLIAMSON, U. GERLAND, AND T. HWA, *A global resource allocation strategy governs growth transition kinetics of Escherichia coli*, Nature, 551 (2017), pp. 119–123.
- [52] C. ETZOLD, G. DECKERS-HEBESTREIT, AND K. ALTENDORF, *Turnover number of Escherichia coli F₀f₁ ATP synthase for ATP synthesis in membrane vesicles*, European Journal of Biochemistry, 243 (1997), pp. 336–343.

- [53] J. J. FALKE, R. B. BASS, S. L. BUTLER, S. A. CHERVITZ, AND M. A. DANIELSON, *The Two-Component Signaling Pathway of Bacterial Chemotaxis: A Molecular View of Signal Transduction by Receptors, Kinases, and Adaptation Enzymes*, Annual Review of Cell and Developmental Biology, 13 (1997), pp. 457–512.
- [54] P. G. FALKOWSKI, R. T. BARBER, AND V. SMETACEK, *Biogeochemical Controls and Feedbacks on Ocean Primary Production*, Science, 281 (1998), pp. 200–206.
- [55] H.-C. FLEMMING AND S. WUERTZ, *Bacteria and archaea on Earth and their abundance in biofilms*, Nature Reviews Microbiology, 17 (2019), p. 247.
- [56] M. J. FOLLOWS AND S. DUTKIEWICZ, *Modeling Diverse Communities of Marine Microbes*, Annual Review of Marine Science, 3 (2011), pp. 427–451.
- [57] M. J. FOLLOWS, S. DUTKIEWICZ, S. GRANT, AND S. W. CHISHOLM, *Emergent Biogeography of Microbial Communities in a Model Ocean*, Science, 315 (2007), pp. 1843–1846.
- [58] S. FREUNDLIEB, U. EHMANN, AND W. BOOS, *Facilitated diffusion of p-nitrophenyl-alpha-D-maltohexaoside through the outer membrane of Escherichia coli. Characterization of LamB as a specific and saturable channel for maltooligosaccharides.*, Journal of Biological Chemistry, 263 (1988), pp. 314–320.
- [59] J. M. GASOL AND D. L. KIRCHMAN, *Microbial Ecology of the Oceans*, John Wiley & Sons, Mar. 2018.
- [60] P. G. GENNES, *Chemotaxis: the role of internal delays*, European Biophysics Journal, 33 (2004), pp. 691–693.
- [61] H. GEST, *The Modern Myth of “Unculturable” Bacteria/ Scotoma of contemporary microbiology*, (2008).
- [62] S. M. GIFFORD, S. SHARMA, M. BOOTH, AND M. A. MORAN, *Expression patterns reveal niche diversification in a marine microbial assemblage*, The ISME Journal, 7 (2013), pp. 281–298.
- [63] S. J. GIOVANNONI, *SAR11 Bacteria: The Most Abundant Plankton in the Oceans*, Annual Review of Marine Science, 9 (2017), pp. 231–255.
- [64] S. J. GIOVANNONI, J. CAMERON THRASH, AND B. TEMPERTON, *Implications of streamlining theory for microbial ecology*, The ISME Journal, 8 (2014), pp. 1553–1565.
- [65] S. J. GIOVANNONI AND U. STINGL, *Molecular diversity and ecology of microbial plankton*, Nature, 437 (2005), pp. 343–348.

- [66] S. J. GIOVANNONI, H. J. TRIPP, S. GIVAN, M. PODAR, K. L. VERGIN, D. BAPTISTA, L. BIBBS, J. EADS, T. H. RICHARDSON, M. NOORDEWIER, M. S. RAPPÉ, J. M. SHORT, J. C. CARRINGTON, AND E. J. MATHUR, *Genome Streamlining in a Cosmopolitan Oceanic Bacterium*, *Science*, 309 (2005), pp. 1242–1245.
- [67] D. V. GOEDEL, D. G. KLEID, F. BOLIVAR, H. L. HEYNEKER, D. G. YANSURA, R. CREA, T. HIROSE, A. KRASZEWSKI, K. ITAKURA, AND A. D. RIGGS, *Expression in Escherichia coli of chemically synthesized genes for human insulin*, *Proceedings of the National Academy of Sciences*, 76 (1979), pp. 106–110.
- [68] G. GRIMMETT AND D. STIRZAKER, *Probability and Random Processes*, Oxford University Press, Oxford, May 2001.
- [69] H.-P. GROSSART, T. KIØRBOE, K. TANG, AND H. PLOUG, *Bacterial Colonization of Particles: Growth and Interactions*, *Applied and Environmental Microbiology*, 69 (2003), pp. 3500–3509.
- [70] H.-P. GROSSART, L. RIEMANN, AND F. AZAM, *Bacterial motility in the sea and its ecological implications*, *Aquatic Microbial Ecology*, 25 (2001), pp. 247–258.
- [71] I. GUDELJ, R. E. BEARDMORE, S. S. ARKIN, AND R. C. MACLEAN, *Constraints on microbial metabolism drive evolutionary diversification in homogeneous environments*, *Journal of Evolutionary Biology*, 20 (2007), pp. 1882–1889.
- [72] S. B. GUTTENPLAN AND D. B. KEARNS, *Regulation of flagellar motility during biofilm formation*, *FEMS Microbiology Reviews*, 37 (2013), pp. 849–871.
- [73] F. M. HAROLD, *The Vital Force: A Study of Bioenergetics*, W. H. Freeman, 1986.
- [74] L. K. HARRIS AND J. A. THERIOT, *Surface Area to Volume Ratio: A Natural Variable for Bacterial Morphogenesis*, *Trends in Microbiology*, 26 (2018), pp. 815–832.
- [75] Y. HART, A. E. MAYO, O. SHOVAL, AND U. ALON, *Comparing Apples and Oranges: Fold-Change Detection of Multiple Simultaneous Inputs*, *PLoS ONE*, 8 (2013), p. e57455.
- [76] G. L. HAZELBAUER, *Maltose chemoreceptor of Escherichia coli.*, *Journal of bacteriology*, 122 (1975), pp. 206–214.
- [77] M. HEIDELBERGER, *Nature from Within: Gustav Theodor Fechner and his Psychophysical Worldview*, University of Pittsburgh Press, Pittsburgh, 2004.
- [78] N. A. HELD, M. R. MCILVIN, D. M. MORAN, M. T. LAUB, AND M. A. SAITO, *Unique Patterns and Biogeochemical Relevance of Two-Component Sensing in Marine Bacteria*, *mSystems*, 4 (2019).

- [79] F. L. HELLWEGER, *Heterotrophic substrate specificity in the aquatic environment: The role of microscale patchiness investigated using modelling: Substrate specificity and microscale patchiness*, *Environmental Microbiology*, 20 (2018), pp. 3825–3835.
- [80] R. HENGGE AND W. BOOS, *Maltose and lactose transport in Escherichia coli: Examples of two different types of concentrative transport systems*, *Biochimica et Biophysica Acta*, 737 (1983), pp. 443–478.
- [81] T. M. HOEHLER AND B. B. JØRGENSEN, *Microbial life under extreme energy limitation*, *Nature Reviews Microbiology*, 11 (2013), pp. 83–94.
- [82] D. HORSTMANN, *From 1970 until present: the Keller-Segel model in chemotaxis and its consequences*, (2003).
- [83] A. H. F. HOSIE AND P. S. POOLE, *Bacterial ABC transporters of amino acids*, *Research in Microbiology*, 152 (2001), pp. 259–270.
- [84] B. HU AND Y. TU, *Behaviors and strategies of bacterial navigation in chemical and nonchemical gradients*, *PLOS Computational Biology*, 10 (2014).
- [85] S. HUI, J. M. SILVERMAN, S. S. CHEN, D. W. ERICKSON, M. BASAN, J. WANG, T. HWA, AND J. R. WILLIAMSON, *Quantitative proteomic analysis reveals a simple strategy of global resource allocation in bacteria*, *Molecular Systems Biology*, 11 (2015), p. 784.
- [86] I. Y. HWANG, E. KOH, A. WONG, J. C. MARCH, W. E. BENTLEY, Y. S. LEE, AND M. W. CHANG, *Engineered probiotic Escherichia coli can eliminate and prevent Pseudomonas aeruginosa gut infection in animal models*, *Nature Communications*, 8 (2017), p. 15028.
- [87] C. INGHAM, M. BUECHNER, AND J. ADLER, *Effect of outer membrane permeability on chemotaxis in Escherichia coli*, *Journal of Bacteriology*, 172 (1990), pp. 3577–3583.
- [88] G. A. JACKSON, R. MAFFIONE, D. K. COSTELLO, A. L. ALLDREDGE, B. E. LOGAN, AND H. G. DAM, *Particle size spectra between 1 μ m and 1 cm at Monterey Bay determined using multiple instruments*, *Deep Sea Research Part I: Oceanographic Research Papers*, 44 (1997), pp. 1739–1767.
- [89] K. JAHREIS, E. F. PIMENTEL-SCHMITT, R. BRÜCKNER, AND F. TITGEMEYER, *Ins and outs of glucose transport systems in eubacteria*, *FEMS Microbiology Reviews*, 32 (2008), pp. 891–907.
- [90] L. JIANG, Q. OUYANG, AND Y. TU, *Quantitative Modeling of Escherichia coli Chemotactic Motion in Environments Varying in Space and Time*, *PLoS Computational Biology*, 6 (2010), p. e1000735.

- [91] J. E. JOHANSEN, J. PINHASSI, N. BLACKBURN, U. L. ZWEIFEL, AND K. HAGSTRM, *Variability in motility characteristics among marine bacteria*, Aquatic Microbial Ecology, 28 (2002), pp. 229–237.
- [92] H. JUNG, *The sodium/substrate symporter family: structural and functional features*, FEBS Letters, 529 (2002), pp. 73–77.
- [93] C. KALETA, S. SCHÄUBLE, U. RINAS, AND S. SCHUSTER, *Metabolic costs of amino acid and protein production in Escherichia coli*, Biotechnology Journal, 8 (2013), pp. 1105–1114.
- [94] Y. KALININ, S. NEUMANN, V. SOURJIK, AND M. WU, *Responses of Escherichia coli Bacteria to Two Opposing Chemoattractant Gradients Depend on the Chemoreceptor Ratio*, Journal of Bacteriology, 192 (2010), pp. 1796–1800.
- [95] C. KE AND W.-C. B, *Adhesion of bacteria and diatoms to surfaces in the sea: a review*, Aquatic Microbial Ecology, 09 (1995), pp. 87–96.
- [96] C. P. KEMPES, P. M. VAN BODEGOM, D. WOLPERT, E. LIBBY, J. AMEND, AND T. HOEHLER, *Drivers of Bacterial Maintenance and Minimal Energy Requirements*, Frontiers in Microbiology, 8 (2017).
- [97] J. E. KEYMER, R. G. ENDRES, M. SKOGE, Y. MEIR, AND N. S. WINGREEN, *Chemosensing in Escherichia coli: two regimes of two-state receptors*, Proceedings of the National Academy of Sciences of the United States of America, 103 (2006), pp. 1786–1791.
- [98] T. KIØRBOE, *Fluid dynamic constraints on resource acquisition in small pelagic organisms*, The European Physical Journal Special Topics, 225 (2016), pp. 669–683.
- [99] T. KIØRBOE, H.-P. GROSSART, H. PLOUG, AND K. TANG, *Mechanisms and Rates of Bacterial Colonization of Sinking Aggregates*, Applied and Environmental Microbiology, 68 (2002), pp. 3996–4006.
- [100] T. KIØRBOE AND G. A. JACKSON, *Marine snow, organic solute plumes, and optimal chemosensory behavior of bacteria*, Limnology and Oceanography, 46 (2001), pp. 1309–1318.
- [101] T. KIØRBOE, H. PLOUG, AND U. THYGESEN, *Fluid motion and solute distribution around sinking aggregates. I. Small-scale fluxes and heterogeneity of nutrients in the pelagic environment*, Marine Ecology Progress Series, 211 (2001), pp. 1–13.
- [102] D. L. KIRCHMAN, *Growth Rates of Microbes in the Oceans*, Annual Review of Marine Science, 8 (2016), pp. 285–309.
- [103] P. E. KLEBBA, *Mechanism of maltodextrin transport through LamB*, Research in Microbiology, 153 (2002), pp. 417–424.

- [104] A. L. KOCH, *The Adaptive Responses of Escherichia coli to a Feast and Famine Existence*, in *Advances in Microbial Physiology*, A. H. Rose and J. F. Wilkinson, eds., vol. 6, Academic Press, Jan. 1971, pp. 147–217.
- [105] —, *WHAT SIZE SHOULD A BACTERIUM BE? A Question of Scale*, *Annual Review of Microbiology*, 50 (1996), pp. 317–348.
- [106] —, *Microbial physiology and ecology of slow growth*, *Microbiology and Molecular Biology Reviews*, 61 (1997), pp. 305–318.
- [107] —, *Oligotrophs versus copiotrophs*, *BioEssays*, 23 (2001), pp. 657–661.
- [108] B. J. KOEBMANN, H. V. WESTERHOFF, J. L. SNOEP, D. NILSSON, AND P. R. JENSEN, *The Glycolytic Flux in Escherichia coli Is Controlled by the Demand for ATP*, *Journal of Bacteriology*, 184 (2002), pp. 3909–3916.
- [109] W. N. KONINGS, *Microbial transport: Adaptations to natural environments*, *Antonie van Leeuwenhoek*, 90 (2006), pp. 325–342.
- [110] P. KOTRBA, M. INUI, AND H. YUKAWA, *Bacterial phosphotransferase system (PTS) in carbohydrate uptake and control of carbon metabolism*, *Journal of Bioscience and Bioengineering*, 92 (2001), pp. 502–517.
- [111] H. KRISHNAMURTHY, C. L. PISCITELLI, AND E. GOUAUX, *Unlocking the molecular secrets of sodium-coupled transporters*, *Nature*, 459 (2009), pp. 347–355.
- [112] C. G. KURLAND, *Molecular characterization of ribonucleic acid from Escherichia coli ribosomes: I. Isolation and molecular weights*, *Journal of Molecular Biology*, 2 (1960), pp. 83–91.
- [113] G. LAN, S. SCHULMEISTER, V. SOURJIK, AND Y. TU, *Adapt locally and act globally: strategy to maintain high chemoreceptor sensitivity in complex environments*, *Molecular Systems Biology*, 7 (2014), pp. 475–475.
- [114] E. LAUGA AND T. R. POWERS, *The hydrodynamics of swimming microorganisms*, *Reports on Progress in Physics*, 72 (2009), p. 096601.
- [115] F. M. LAURO, D. MCDUGALD, T. THOMAS, T. J. WILLIAMS, S. EGAN, S. RICE, M. Z. DEMAERE, L. TING, H. ERTAN, J. JOHNSON, S. FERRIERA, A. LAPIDUS, I. ANDERSON, N. KYRPIDES, A. C. MUNK, C. DETTER, C. S. HAN, M. V. BROWN, F. T. ROBB, S. KJELLEBERG, AND R. CAVICCHIOLI, *The genomic basis of trophic strategy in marine bacteria*, *Proceedings of the National Academy of Sciences*, 106 (2009), pp. 15527–15533.
- [116] M. D. LAZOVA, T. AHMED, D. BELLOMO, R. STOCKER, AND T. S. SHIMIZU, *Response rescaling in bacterial chemotaxis*, *Proceedings of the National Academy of Sciences*, 108 (2011), pp. 13870–13875.

- [117] A. L. LEHNINGER, D. L. NELSON, AND M. M. COX, *Lehninger Principles of Biochemistry*, Macmillan, 2005.
- [118] M. A. LEVER, K. L. ROGERS, K. G. LLOYD, J. OVERMANN, B. SCHINK, R. K. THAUER, T. M. HOEHLER, AND B. B. JØRGENSEN, *Life under extreme energy limitation: a synthesis of laboratory- and field-based investigations*, FEMS Microbiology Reviews, 39 (2015), pp. 688–728.
- [119] H. LUO AND M. A. MORAN, *How do divergent ecological strategies emerge among marine bacterioplankton lineages?*, Trends in Microbiology, 23 (2015), pp. 577–584.
- [120] R. M. MACNAB AND D. E. KOSHLAND, *The gradient-sensing mechanism in bacterial chemotaxis*, Proceedings of the National Academy of Sciences, 69 (1972), pp. 2509–2512.
- [121] J. R. MADDOCK AND L. SHAPIRO, *Polar Location of the Chemoreceptor Complex in the Escherichia coli Cell*, Science, 259 (1993), pp. 1717–1723.
- [122] K. MALMCRONA-FRIBERG, A. GOODMAN, AND S. KJELLEBERG, *Chemotactic Responses of Marine Vibrio sp. Strain S14 (CCUG 15956) to Low-Molecular-Weight Substances under Starvation and Recovery Conditions*, Applied and Environmental Microbiology, 56 (1990), pp. 3699–3704.
- [123] R. R. MALMSTROM, R. P. KIENE, M. T. COTTRELL, AND D. L. KIRCHMAN, *Contribution of SAR11 Bacteria to Dissolved Dimethylsulfoniopropionate and Amino Acid Uptake in the North Atlantic Ocean*, Applied and Environmental Microbiology, 70 (2004), pp. 4129–4135.
- [124] M. D. MANSON, W. BOOS, P. J. BASSFORD, AND B. A. RASMUSSEN, *Dependence of maltose transport and chemotaxis on the amount of maltose-binding protein*, Journal of Biological Chemistry, 260 (1985), pp. 9727–9733.
- [125] J. S. MATTICK, *Type IV Pili and Twitching Motility*, Annual Review of Microbiology, 56 (2002), pp. 289–314.
- [126] B. A. MELLO AND Y. TU, *An allosteric model for heterogeneous receptor complexes: understanding bacterial chemotaxis responses to multiple stimuli*, Proceedings of the National Academy of Sciences of the United States of America, 102 (2005), pp. 17354–17359.
- [127] —, *Effects of Adaptation in Maintaining High Sensitivity over a Wide Range of Backgrounds for Escherichia coli Chemotaxis*, Biophysical Journal, 92 (2007), pp. 2329–2337.
- [128] F. MENOLASCINA, R. RUSCONI, V. I. FERNANDEZ, S. SMRIGA, Z. AMINZARE, E. D. SONTAG, AND R. STOCKER, *Logarithmic sensing in Bacillus subtilis aerotaxis*, NPJ Systems Biology and Applications, 3 (2017), p. 16036.

- [129] R. MESIBOV, G. W. ORDAL, AND J. ADLER, *The Range of Attractant Concentrations for Bacterial Chemotaxis and the Threshold and Size of Response over This Range: Weber law and related phenomena*, *The Journal of general physiology*, 62 (1973), pp. 203–223.
- [130] A. R. MESQUITA, *Exploiting Stochasticity in Multi-agent Systems*, PhD thesis, University of California, Santa Barbara, 2010.
- [131] L. MICHAELIS, M. L. MENTEN, K. A. JOHNSON, AND R. S. GOODY, *The original Michaelis constant: translation of the 1913 Michaelis-Menten paper*, *Biochemistry*, 50 (2011), pp. 8264–8269.
- [132] D. M. MILLER, J. S. OLSON, J. W. PFLUGRATH, AND F. A. QUIOCHO, *Rates of ligand binding to periplasmic proteins involved in bacterial transport and chemotaxis*, *Journal of Biological Chemistry*, 258 (1983), pp. 13665–13672.
- [133] R. MILO AND R. PHILLIPS, *Cell Biology by the Numbers*, Garland Science, Dec. 2015.
- [134] J. G. MITCHELL, *The influence of cell size on marine bacterial motility and energetics*, *Microbial Ecology*, 22 (1991), pp. 227–238.
- [135] ———, *The Energetics and Scaling of Search Strategies in Bacteria*, *The American Naturalist*, 160 (2002), pp. 727–740.
- [136] J. G. MITCHELL, L. PEARSON, AND S. DILLON, *Clustering of marine bacteria in seawater enrichments*, *Applied and Environmental Microbiology*, 62 (1996), pp. 3716–3721.
- [137] J. G. MITCHELL, L. PEARSON, S. DILLON, AND K. KANTALIS, *Natural assemblages of marine bacteria exhibiting high-speed motility and large accelerations*, *Applied and Environmental Microbiology*, 61 (1995), pp. 4436–4440.
- [138] D. MOLENAAR, R. V. BERLO, D. D. RIDDER, AND B. TEUSINK, *Shifts in growth strategies reflect tradeoffs in cellular economics*, *Molecular Systems Biology*, 5 (2009), p. 323.
- [139] J. MONOD, *Chance and Necessity: An Essay on the Natural Philosophy of Modern Biology*, Knopf, 1971.
- [140] J. MONOD, J. WYMAN, AND J.-P. CHANGEUX, *On the nature of allosteric transitions: A plausible model*, *J. Mol. Biol.*, 12 (1965), pp. 88–118.
- [141] R. M. MORRIS, M. S. RAPPÉ, S. A. CONNOR, K. L. VERGIN, W. A. SIEBOLD, C. A. CARLSON, AND S. J. GIOVANNONI, *SAR11 clade dominates ocean surface bacterioplankton communities*, *Nature*, 420 (2002), p. 806.

- [142] S. L. MOWBRAY AND D. E. KOSHLAND, *Additive and independent responses in a single receptor: Aspartate and maltose stimuli on the tar protein*, *Cell*, 50 (1987), pp. 171–180.
- [143] M. MUIR, L. WILLIAMS, AND T. FERENCI, *Influence of Transport Energization on the Growth Yield of Escherichia coli*, *J. BACTERIOL.*, 163 (1985), p. 6.
- [144] T. NAGATA, *Organic Matter–Bacteria Interactions in Seawater*, in *Microbial Ecology of the Oceans*, D. L. Kirchman, ed., John Wiley & Sons, Inc., Hoboken, NJ, USA, June 2008, pp. 207–241.
- [145] F. C. NEIDHARDT, J. L. INGRAHAM, AND M. SCHAECHTER, *Physiology of the Bacterial Cell: A Molecular Approach*, Sinauer Associates, Jan. 1990.
- [146] S. NEUMANN, C. H. HANSEN, N. S. WINGREEN, AND V. SOURJIK, *Differences in signalling by directly and indirectly binding ligands in bacterial chemotaxis*, *The EMBO Journal*, 29 (2010), pp. 3484–3495.
- [147] H. NIKAIDO, *Effect of solute size on diffusion rates through the transmembrane pores of the outer membrane of Escherichia coli*, *The Journal of General Physiology*, 77 (1981), pp. 121–135.
- [148] ———, *Maltose transport system of Escherichia coli: An ABC-type transporter*, *FEBS letters*, 346 (1994), pp. 55–58.
- [149] ———, *Molecular Basis of Bacterial Outer Membrane Permeability Revisited*, *Microbiol. Mol. Biol. Rev.*, 67 (2003), pp. 593–656.
- [150] S. E. NOELL AND S. J. GIOVANNONI, *SAR11 bacteria have a high affinity and multifunctional glycine betaine transporter*, *Environmental Microbiology*, 0.
- [151] E. NOOR, A. FLAMHOLZ, A. BAR-EVEN, D. DAVIDI, R. MILO, AND W. LIEBERMEISTER, *The Protein Cost of Metabolic Fluxes: Prediction from Enzymatic Rate Laws and Cost Minimization*, *PLOS Computational Biology*, 12 (2016), p. e1005167.
- [152] N. NORRIS, *Exploring the optimality of various bacterial motility strategies: a stochastic hybrid systems approach*, Master’s thesis, Massachusetts Institute of Technology, 2013.
- [153] N. NORRIS, F. MENOLASCINA, E. FRAZZOLI, AND R. STOCKER, *The effect of reversals for a stochastic source-seeking process inspired by bacterial chemotaxis*, in 2014 American Control Conference, June 2014, pp. 3411–3416.
- [154] B. Ø. PALSSON, *Systems Biology: Constraint-based Reconstruction and Analysis*, Cambridge University Press, Jan. 2015.

- [155] J. S. PATZLAFF, T. V. D. HEIDE, AND B. POOLMAN, *The ATP/Substrate Stoichiometry of the ATP-binding Cassette (ABC) Transporter OpuA*, *Journal of Biological Chemistry*, 278 (2003), pp. 29546–29551.
- [156] K. PEEBO, K. VALGEPEA, A. MASER, R. NAHKU, K. ADAMBERG, AND R. VILU, *Proteome reallocation in Escherichia coli with increasing specific growth rate*, *Molecular BioSystems*, 11 (2015), pp. 1184–1193.
- [157] R. PHILLIPS, J. KONDEV, J. THERIOT, AND H. GARCIA, *Physical Biology of the Cell*, Garland Science, 2012.
- [158] B. POOLMAN, M. K. DOEVEN, E. R. GEERTSMA, E. BIEMANS-OLDEHINKEL, W. N. KONINGS, AND D. C. REES, *Functional Analysis of Detergent-Solubilized and Membrane-Reconstituted ATP-Binding Cassette Transporters*, in *Methods in Enzymology*, vol. 400 of Phase II Conjugation Enzymes and Transport Systems, Academic Press, Jan. 2005, pp. 429–459.
- [159] R. S. PORETSKY, S. SUN, X. MOU, AND M. A. MORAN, *Transporter genes expressed by coastal bacterioplankton in response to dissolved organic carbon*, *Environmental Microbiology*, 12 (2010), pp. 616–627.
- [160] E. M. PURCELL, *Life at low Reynolds number*, *Am. J. Phys.*, 45 (1977), pp. 3–11.
- [161] —, *The efficiency of propulsion by a rotating flagellum*, *Proceedings of the National Academy of Sciences*, 94 (1997), pp. 11307–11311.
- [162] D. C. REES, E. JOHNSON, AND O. LEWINSON, *ABC transporters: the power to change*, *Nature Reviews Molecular Cell Biology*, 10 (2009), pp. 218–227.
- [163] G. REINTJES, C. ARNOSTI, B. M. FUCHS, AND R. AMANN, *An alternative polysaccharide uptake mechanism of marine bacteria*, *The ISME Journal*, 11 (2017), pp. 1640–1650.
- [164] E. M. RENKIN, *Filtration, Diffusion, and Molecular Sieving Through Porous Cellulose Membranes*, *The Journal of General Physiology*, 38 (1954), pp. 225–243.
- [165] G. RICHARME, *Interaction of the maltose-binding protein with membrane vesicles of Escherichia coli*, *Journal of Bacteriology*, 149 (1982), pp. 662–667.
- [166] B. RODENBORN, C.-H. CHEN, H. L. SWINNEY, B. LIU, AND H. P. ZHANG, *Propulsion of microorganisms by a helical flagellum*, *Proceedings of the National Academy of Sciences*, 110 (2013), pp. E338–E347.
- [167] A. ROMANO AND M. H. SALER, *Evolution of the Bacterial Phosphoenolpyruvate: Sugar Phosphotransferase System*, in *The Evolution of Metabolic function*, R. Mortlock, ed., CRC Press, 1992.
- [168] M. H. SAIER, *Families of transmembrane sugar transport proteins*, *Molecular Microbiology*, 35 (2000), pp. 699–710.

- [169] M. H. SALER AND J. REIZER, *The bacterial phosphotransferase system: new frontiers 30 years later*, Molecular Microbiology, 13 (1994), pp. 755–764.
- [170] R. I. SARKER, W. OGAWA, T. SHIMAMOTO, T. SHIMAMOTO, AND T. TSUCHIYA, *Primary structure and properties of the Na⁺/glucose symporter (Sg1s) of Vibrio parahaemolyticus.*, Journal of Bacteriology, 179 (1997), pp. 1805–1808.
- [171] R. I. SARKER, W. OGAWA, M. TSUDA, S. TANAKA, AND T. TSUCHIYA, *Properties of a Na⁺/galactose(glucose) symport system in Vibrio parahaemolyticus*, Biochimica et Biophysica Acta (BBA) - Biomembranes, 1279 (1996), pp. 149–156.
- [172] A. SCHLEGEL, A. BOHM, S.-J. LEE, R. PEIST, K. DECKER, AND W. BOOS, *Network regulation of the Escherichia coli maltose system*, Journal of molecular microbiology and biotechnology, 4 (2002), pp. 301–307.
- [173] M. J. SCHNITZER, *Theory of continuum random walks and application to chemotaxis*, Physical Review E, 48 (1993), p. 2553.
- [174] S. SCHÖNERT, S. SEITZ, H. KRAFFT, E.-A. FEUERBAUM, I. ANDERNACH, G. WITZ, AND M. K. DAHL, *Maltose and Maltodextrin Utilization by Bacillus subtilis*, Journal of Bacteriology, 188 (2006), pp. 3911–3922.
- [175] M. SCOTT, S. KLUMPP, E. M. MATEESCU, AND T. HWA, *Emergence of robust growth laws from optimal regulation of ribosome synthesis*, Molecular Systems Biology, 10 (2014), p. 747.
- [176] J. E. SEGALL, S. M. BLOCK, AND H. C. BERG, *Temporal comparisons in bacterial chemotaxis*, Proceedings of the National Academy of Sciences, 83 (1986), pp. 8987–8991.
- [177] R. SENDER, S. FUCHS, AND R. MILO, *Revised Estimates for the Number of Human and Bacteria Cells in the Body*, PLOS Biology, 14 (2016), p. e1002533.
- [178] P. SHARMA, P. P. PANDEY, AND S. JAIN, *Modeling the cost and benefit of proteome regulation in a growing bacterial cell*, Physical Biology, 15 (2018), p. 046005.
- [179] O. SHOVAL, L. GOENTORO, Y. HART, A. MAYO, E. SONTAG, AND U. ALON, *Fold-change detection and scalar symmetry of sensory input fields*, Proceedings of the National Academy of Sciences, 107 (2010), pp. 15995–16000.
- [180] C. SIEBOLD, K. FLÜKIGER, R. BEUTLER, AND B. ERNI, *Carbohydrate transporters of the bacterial phosphoenolpyruvate: sugar phosphotransferase system (PTS)*, FEBS Letters, 504 (2001), pp. 104–111.
- [181] D. R. SMITH AND M. R. CHAPMAN, *Economical Evolution: Microbes Reduce the Synthetic Cost of Extracellular Proteins*, mBio, 1 (2010), pp. e00131–10.

- [182] S. SMRIGA, V. I. FERNANDEZ, J. G. MITCHELL, AND R. STOCKER, *Chemotaxis toward phytoplankton drives organic matter partitioning among marine bacteria*, Proceedings of the National Academy of Sciences, 113 (2016), pp. 1576–1581.
- [183] K. SON, J. S. GUASTO, AND R. STOCKER, *Bacteria can exploit a flagellar buckling instability to change direction*, Nature Physics, 9 (2013), pp. 494–498.
- [184] K. SON, F. MENOLASCINA, AND R. STOCKER, *Speed-dependent chemotactic precision in marine bacteria*, Proceedings of the National Academy of Sciences, (2016), p. 201602307.
- [185] V. SOURJIK, *Receptor clustering and signal processing in E. coli chemotaxis*, Trends in Microbiology, 12 (2004), pp. 569–576.
- [186] V. SOURJIK AND H. C. BERG, *Receptor sensitivity in bacterial chemotaxis*, Proceedings of the National Academy of Sciences, 99 (2002), pp. 123–127.
- [187] V. SOURJIK, A. VAKNIN, T. S. SHIMIZU, AND H. C. BERG, *In Vivo Measurement by FRET of Pathway Activity in Bacterial Chemotaxis*, in Methods in Enzymology, vol. 423, Elsevier, 2007, pp. 365–391.
- [188] O. A. SOUTOURINA AND P. N. BERTIN, *Regulation cascade of flagellar expression in Gram-negative bacteria*, FEMS Microbiology Reviews, 27 (2003), pp. 505–523.
- [189] S. M. SOWELL, L. J. WILHELM, A. D. NORBECK, M. S. LIPTON, C. D. NICORA, D. F. BAROFSKY, C. A. CARLSON, R. D. SMITH, AND S. J. GIOVANONNI, *Transport functions dominate the SAR11 metaproteome at low-nutrient extremes in the Sargasso Sea*, The ISME journal, 3 (2009), pp. 93–105.
- [190] S. SRINIVASAN AND S. KJELLEBERG, *Cycles of famine and feast: the starvation and outgrowth strategies of a marine Vibrio*, Journal of Biosciences, 23 (1998), pp. 501–511.
- [191] A. M. STOCK, V. L. ROBINSON, AND P. N. GOUDREAU, *Two-Component Signal Transduction*, Annual Review of Biochemistry, 69 (2000), pp. 183–215.
- [192] J. B. STOCK, *Adaptive responses in bacterial chemotaxis*, in Regulation of Cellular Signal Transduction Pathways by Desensitization and Amplification, D. Sibley and M. Houslay, eds., Wiley, New York, 1994, pp. 3–24.
- [193] R. STOCKER AND J. R. SEYMOUR, *Ecology and Physics of Bacterial Chemotaxis in the Ocean*, Microbiology and Molecular Biology Reviews, 76 (2012), pp. 792–812.
- [194] R. STOCKER, J. R. SEYMOUR, A. SAMADANI, D. E. HUNT, AND M. F. POLZ, *Rapid chemotactic response enables marine bacteria to exploit ephemeral microscale nutrient patches*, Proceedings of the National Academy of Sciences, 105 (2008), pp. 4209–4214.

- [195] S. STRETTON, S. J. DANON, S. KJELLEBERG, AND A. E. GOODMAN, *Changes in cell morphology and motility in the marine Vibrio sp. strain S14 during conditions of starvation and recovery*, FEMS Microbiology Letters, 146 (1997), pp. 23–29.
- [196] M.-T. SUNG, Y.-T. LAI, C.-Y. HUANG, L.-Y. CHOU, H.-W. SHIH, W.-C. CHENG, C.-H. WONG, AND C. MA, *Crystal structure of the membrane-bound bifunctional transglycosylase PBP1b from Escherichia coli*, Proceedings of the National Academy of Sciences, 106 (2009), pp. 8824–8829.
- [197] M. SZENK, K. A. DILL, AND A. M. R. DE GRAFF, *Why Do Fast-Growing Bacteria Enter Overflow Metabolism? Testing the Membrane Real Estate Hypothesis*, Cell Systems, 5 (2017), pp. 95–104.
- [198] J. R. TAYLOR AND R. STOCKER, *Trade-Offs of Chemotactic Foraging in Turbulent Water*, Science, 338 (2012), pp. 675–679.
- [199] S. THRUN, W. BURGARD, AND D. FOX, *Probabilistic Robotics*, MIT press, 2005.
- [200] D. TILMAN, *Resource Competition and Community Structure*, Princeton University Press, 1982.
- [201] F. TROVATO AND V. TOZZINI, *Diffusion within the Cytoplasm: A Mesoscale Model of Interacting Macromolecules*, Biophysical Journal, 107 (2014), pp. 2579–2591.
- [202] Y. TU, *Quantitative Modeling of Bacterial Chemotaxis: Signal Amplification and Accurate Adaptation*, Annual Review of Biophysics, 42 (2013), pp. 337–359.
- [203] Y. TU, T. S. SHIMIZU, AND H. C. BERG, *Modeling the chemotactic response of Escherichia coli to time-varying stimuli*, Proceedings of the National Academy of Sciences, 105 (2008), pp. 14855–14860.
- [204] P. J. TURNBAUGH, R. E. LEY, M. HAMADY, C. M. FRASER-LIGGETT, R. KNIGHT, AND J. I. GORDON, *The Human Microbiome Project*, Nature, 449 (2007), pp. 804–810.
- [205] P. A. VAISHAMPAYAN, E. RABOW, G. HORNECK, AND K. J. VENKATESWARAN, *Survival of Bacillus pumilus Spores for a Prolonged Period of Time in Real Space Conditions*, Astrobiology, 12 (2012), pp. 487–497.
- [206] G. M. VISWANATHAN, M. G. E. DA LUZ, E. P. RAPOSO, AND H. E. STANLEY, *The Physics of Foraging*, Cambridge University Press, New York, 2011.
- [207] E. B. WAYGOOD AND T. STEEVES, *Enzyme I of the phosphoenolpyruvate: sugar phosphotransferase system of Escherichia coli. Purification to homogeneity and some properties*, Canadian Journal of Biochemistry, 58 (1980), pp. 40–48.
- [208] U. WEGMANN, A. L. CARVALHO, M. STOCKS, AND S. R. CARDING, *Use of genetically modified bacteria for drug delivery in humans: Revisiting the safety aspect*, Scientific Reports, 7 (2017), p. 2294.

- [209] M. T. WEINSTOCK, E. D. HESEK, C. M. WILSON, AND D. G. GIBSON, *Vibrio natriegens as a fast-growing host for molecular biology*, Nature Methods, 13 (2016), pp. 849–851.
- [210] A. E. WHITE, S. J. GIOVANNONI, Y. ZHAO, K. VERGIN, AND C. A. CARLSON, *Elemental content and stoichiometry of SAR11 chemoheterotrophic marine bacteria: SAR11 composition*, Limnology and Oceanography Letters, (2019).
- [211] F. B. WIENTJES, C. L. WOLDRINGH, AND N. NANNINGA, *Amount of peptidoglycan in cell walls of gram-negative bacteria.*, Journal of Bacteriology, 173 (1991), pp. 7684–7691.
- [212] C. WINTER, T. BOUVIER, M. G. WEINBAUER, AND T. F. THINGSTAD, *Trade-Offs between Competition and Defense Specialists among Unicellular Planktonic Organisms: the “Killing the Winner” Hypothesis Revisited*, Microbiology and Molecular Biology Reviews, 74 (2010), pp. 42–57.
- [213] A. D. WOLFSON, J. A. PLEISS, AND O. C. UHLENBECK, *A new assay for tRNA aminoacylation kinetics.*, RNA, 4 (1998), pp. 1019–1023.
- [214] L. XIE, T. ALTINDAL, S. CHATTOPADHYAY, AND X.-L. WU, *Bacterial flagellum as a propeller and as a rudder for efficient chemotaxis*, Proceedings of the National Academy of Sciences, 108 (2011), pp. 2246–2251.
- [215] E. M. YAM AND K. W. TANG, *Effects of starvation on aggregate colonization and motility of marine bacteria*, Aquatic Microbial Ecology, 48 (2007), pp. 207–215.
- [216] Y. YANG AND V. SOURJIK, *Opposite responses by different chemoreceptors set a tunable preference point in Escherichia coli pH taxis*, Molecular Microbiology, 86 (2012), pp. 1482–1489.
- [217] Y. YAWATA, O. X. CORDERO, F. MENOLASCINA, J.-H. HEHEMANN, M. F. POLZ, AND R. STOCKER, *Competition-dispersal tradeoff ecologically differentiates recently speciated marine bacterioplankton populations*, Proceedings of the National Academy of Sciences, 111 (2014), pp. 5622–5627.
- [218] T.-M. YI, Y. HUANG, M. I. SIMON, AND J. DOYLE, *Robust perfect adaptation in bacterial chemotaxis through integral feedback control*, Proceedings of the National Academy of Sciences, 97 (2000), pp. 4649–4653.
- [219] T. YORIMITSU AND M. HOMMA, *Na⁺-driven flagellar motor of Vibrio*, Biochimica et Biophysica Acta (BBA) - Bioenergetics, 1505 (2001), pp. 82–93.
- [220] K. D. YOUNG, *The Selective Value of Bacterial Shape*, Microbiology and Molecular Biology Reviews, 70 (2006), pp. 660–703.

- [221] E. J. ZAKEM, A. AL-HAJ, M. J. CHURCH, G. L. v. DIJKEN, S. DUTKIEWICZ, S. Q. FOSTER, R. W. FULWEILER, M. M. MILLS, AND M. J. FOLLOWS, *Ecological control of nitrite in the upper ocean*, Nature Communications, 9 (2018), pp. 1–13.