Summer 8-5-2023

Spatial structure formation by the post-transcriptional gene regulator RsmE in Pseudomonas fluorescens Pf0-1

Anton Evans

SPATIAL STRUCTURE FORMATION BY THE POST-TRANSCRIPTIONAL GENE REGULATOR RSME IN *PSEUDOMONAS FLUORESCENS* PF0-1

A Dissertation

Submitted to the School of Science and Engineering

Duquesne University

In partial fulfillment of the requirements for

the degree of Doctor of Philosophy

By

Anton Fredrick Evans Jr.

August 2023

SPATIAL STRUCTURE FORMATION BY THE POST-TRANSCRIPTIONAL GENE REGULATOR RSME IN *PSEUDOMONAS FLUORESCENS* PF0-1

By

Anton Fredrick Evans Jr.

Approved June 15th, 2023

Wook Kim Associate Professor of Biological Sciences (Research Advisor and Committee Chair)	Michael Jensen-Seaman Associate Professor of Biological Sciences (Committee Member)
Dannie Durand Associate Professor of Biological Sciences, Carnegie Mellon University (Committee Member)	Jana Patton-Vogt Chair and Professor, Department of Biological Sciences (Committee Member)

Ellen Gawalt
Dean, School of Science and Engineering

ABSTRACT

SPATIAL STRUCTURE FORMATION BY THE POST-TRANSCRIPTIONAL GENE REGULATOR RSME IN *PSEUDOMONAS FLUORESCENS* PF0-1

By

Anton Fredrick Evans Jr.

August 2023

Dissertation supervised by Dr. Wook Kim

Microorganisms are often found in microbial communities we call biofilms. Organisms living in these crowded environments have significant evolutionary pressure to retain access to the resources necessary to sustain life. My research uses the bacterium *Pseudomonas fluorescens* Pf0-1 to study how organisms evolve strategies to solve this crowding problem as aging colonies repeatedly generate mutant patches. These mutants expand the reach of the colony resulting in decreased local density as they push themselves up to the resource rich surface. These spatial structures result from social interactions between the mutant and the parental cells mediated through extracellular secretions, resulting in the mutant progeny displaying increased fitness compared to the parent. Loss of function mutations in *rsmE* are exclusively responsible for every mutant patch observed. RsmE and its two paralogs (RsmA and RsmI) are post-transcriptional gene regulators, described to redundantly repress multiple secretions by sequestering associated mRNA, which contradicts our large mutational data set. With the use of

genetic engineering, fluorescent microscopy, competitional, transcriptomic, and proteomic analyses in this dissertation, I characterize the spatiogenetic impact of RsmE regulation. In the following chapters I examine the major RsmE regulated extracellular secretions contributing to the increased fitness of the mutant patches, I explore the impact each Rsm paralog has on the transcriptome, and I use a systematic approach combining proteomics and transcriptomics to identify additional intercellular interactions. Overall, this work provides insight into strategies evolved to survive in dense microbial communities and provides a platform to further explore questions of structure and function at both the protein and cellular levels.

DEDICATION

The amount of support I have received throughout this journey has been immense. Firstly, I would like to thank my advisor, Dr. Wook Kim, for always pushing me outside of my comfort zone and supporting me through every step I needed to take to accomplish everything documented in this dissertation. I would also like to thank my committee members for their wisdom and expertise, which helped me grow into a better scientist.

I would like to thank my previous undergraduate students, specifically Amber, Razi, and Bill, for their hard work and dedication, which allowed me to achieve more than I could have alone in the lab. I would also like to thank my fellow PhD students, specifically Collin, Will, Brooke, Brianna, David, and many others, for always being there for me throughout all of our common challenges.

I want to thank my wife Megan for her support and endless optimism, which carried me through all the struggles and pressures of the daily grind to finish this degree.

Finally, I would like to dedicate this accomplishment to my mother Yolanda, father Anton, and aunt MaSheila for always answering and never discouraging my “what if” and “how come” questions as a child. I also want to thank them and the rest of my family for the countless sacrifices they have made which have allowed me to pursue my dreams.

ACKNOWLEDGEMENT

The work described in this dissertation was conducted by Anton Evans unless stated otherwise with the following exceptions:

In Chapter 2 general experimental design, the Identification of biosurfactant biosynthesis genes by transposon mutagenesis, and the construction E. coli S17λpir strains for knocking out Pfl01_3834 and Pfl01_2211 was previously done by Dr. Wook Kim. Initial phenotype screens and competitional analysis was done by undergraduate student Jordan Denk. Final optimized competitional analysis was done by undergraduate student William Mazza under my supervision.

Undergraduate students Raziel Santos and Amber Delprince assisted me in cell culturing, screening mutants, purifying RNA, PCR, qRT-PCR and countless other tasks necessary to complete the work in both chapter 2 and 3. Graduate student Megan Wells assisted me in data analysis and editing for chapter 2 and 3.

Whole genome sequencing and RNA sequencing was performed by the Microbial Genome Sequencing Center (MiGS) located in Pittsburgh, PA. Protein identification by Mass Spectrometry was outsourced to Michigan State University Proteomics Core in East Lansing, MI.

The work was funded by the National Institute of General Medical Sciences of the NIH (1R15GM132856 to Dr. Wook Kim), Charles Henry Leach II Fund (Dr. Wook Kim) and the Duquesne University BSNES Faculty Development Fund (Dr. Wook Kim).

TABLE OF CONTENTS

LIST OF FIGURES

Chapter 1

Chapter 2

Chapter 3

noticeable gap between the interior dead cells and the exterior border. Dead cells are not observed within the patch formed by M^S*, but they are visible at the boundary against WT. In contrast, dead cells are rarely associated with the patches formed by Δ*rsmE* Δ*Pfl01_5574 and* Δ*rsmE* Δ*Pfl01_5594*, that lack the T6SS. Scale bar represents 50μm..64

Chapter 1 Introduction

Pseudomonads

The pseudomonads are Gamma-proteobacteria that display a broad range of metabolic activities for adapting to diverse ecological niches. Many *Pseudomonas* species have long been studied due to their opportunistic nature as pathogens and also their potential as biocontrol and remediation tools. *P. aeruginosa* is the most clinically significant member of the pseudomonads, which frequently causes eye infections and skin infections in immune compromised individuals with lacerations and burns. In particular, *P. aeruginosa* infections of the lung is the major contributing factor to the mortality in cystic fibrosis patients(3). The combination of their genomic plasticity and propensity for biofilm formation poses tremendous challenges in antibiotic therapy.

P. aeruginosa and nonpathogenic *Pseudomonas* species are readily found in water and soil, including those contaminated with oil(3). Pseudomonads are well known for their production of biosurfactants and the ability to utilize diverse carbon sources, which have made them potential targets for use in bioremediation of industrial contamination of the environment(3, 4). In agriculture, plant-bacterial interactions have been widely studied for increasing crop yield. One of the most well-known interactions is the symbiotic relationship between soybean plants and nitrogen fixing rhizobia, which has led to reduced dependence on fertilizer and implementation of soybean – corn crop rotations. *Pseudomonas putida, Pseudomonas entomophila* and *Pseudomonas fluorescens* have been studied for their potential plant biocontrol applications due to their ecological role as plant pathogens or commensals(3, 5, 6). Most notably, the production of secreted

secondary metabolites has been associated with antibacterial, antifungal, insecticidal, and plant hormetic effects(3, 5, 7-9).

The genetic and metabolic plasticity, along with the propensity to interact with the environment through diverse extracellular secretions, are the root of much of the scientific interest in the pseudomonads. However, our understanding of the *in situ* functional genomics of many *Pseudomonas* species remains limited as the majority of our empirical knowledge stems from studying microbes in liquid culture. At the same time, relatively little about the mechanistic details of microbial social interaction, the importance of spatial structures they form, and signals that drive these interactions.

Intercellular interactions

What we do know is that in the environment pseudomonads typically live in microbial communities. Such an environment creates intense competition between different species, and between individuals within a species for access to resources needed for survival. Interestingly emergent properties of these communities include increased adaptation to environmental fluctuations, antibiotic resistance, and increased overall resiliency compared to free-living systems. Understanding these mechanisms is becoming increasingly important as our tools to control infections and colonization dwindle. It is thus critical to build a better understanding of how microbes live, interact, and evolve within structured communities.

Pseudomonads have the ability to adapt to many environments and have been shown to switch from a planktonic lifecycle to a sedentary one (3, 10, 11). The propensity to make this switch is due to the cell's interpretation of extracellular signals present in

their microenvironment. This can result in quorum sensing, leading to the formation of dense communities, as the production and sensing of extracellular messages signal a positive feedback loop for more cellular adherence(6, 12).

A major component of the extracellular signaling cascade pathway in pseudomonads is the GacS/GacA network(13, 14). This regulatory network has been shown to be involved in governing virulence factors, motility, and the formation of biofilms(5, 14). In this system, the transmembrane protein GacS is stimulated by currently unknown extracellular signals(5, 15). GacS relays the extracellular signal across the membrane by activating the cytosolic protein GacA through phosphorylation (5, 12, 14). Activated GacA then binds the GacA box, and stimulates the transcription of small regulatory RNAs (sRNA)(12). These sRNAs competitively bind a class of post transcriptional regulators known as regulators of secondary metabolism (Rsm) proteins to disrupt their function(12, 16, 17).

Rsm proteins

Known as Rsm proteins in Pseudomonads but first described as Csr (carbon storage regulator) proteins in *E. coli,* these proteins have been found in over 2900 bacterial species, with majority of the distribution found in gammaproteobacteria (18, 19). The Rsm family of proteins function as post-transcriptional regulators by binding mRNA and blocking translation by the ribosome (Figure 1) (1, 12, 20, 21). These proteins have long been known to be involved in the highly conserved GacS/GacA signal transduction pathway, and have been shown to be crucial in the regulation of extracellular secretions, virulence factors, and biofilm formation(3, 12). However, Csr/Rsm proteins

have been recently described as global regulators due to their interaction with over 15% of all mRNA in some species (22-24). Since their discovery, it has been shown that Csr/Rsm proteins share very similar sequences, secondary structure, and function to be categorized in the same class of protein. There have been recent proposals to reclassify all Csr/Rsm proteins as Rsm and their subtype based on the clustering with *Pseudomonas* paralogs(19).

The majority of these Csr/Rsm paralogs, are approximately 70 amino acids long with a well-conserved secondary structure consisting of four beta sheets and an alpha helix, N terminus to C terminus, with diversity arising mainly in the C terminus post alpha-helix (Appendix I) (19, 21, 25). In the minority is a structural inversion where an alpha helix is placed between the second and third beta sheets, as observed in RsmN and RsmF. The Csr/Rsm proteins have been shown to be functionally active as a homodimer with two binding sites per dimer (Figure 1) (19, 21, 26). Classical descriptions of Csr/Rsm proteins recognize and bind to the Shine Dalgarno (SD) site of mRNA and block the docking of the ribosome, inhibiting translation of that mRNA transcript(Figure 1)(27). However, recent studies have proposed and demonstrated more nuanced mechanisms.

Rsm activity

Rsm proteins have been shown to inhibit translation through indirect blocking of the SD site(25). For example, it was reported that CsrA in *E. coli* was shown to repress Hfq, another major bacterial post-transcriptional regulator(28), by binding to a site that overlaps with the SD sequence of the gene, while RsmA in *P. aeruginosa* was shown to

block the SD through stabilization of RNA secondary structure, making it inaccessible to the ribosome(29-31). CsrA was also shown to modulate Rho-dependent transcription termination in which the bound protein inhibits the formation of secondary structure on nascent mRNA that would block a Rho binding site resulting in early termination(29). It has also been proposed that this mechanism stabilizing secondary structure could also promote RNA interference, thus further decreasing the transcripts' stability(19).

Conversely, Rsm proteins have been shown to promote translation either by stabilizing access to the SD through binding, which prevents the formation of the SD hairpin structure, or by binding RNase recognition sites and extending the life of the mRNA transcript(29). All Rsm activity is mediated by different interpretations of the canonical stem loop stabilization interaction with RNA. This interaction is facilitated through the Csr domain of the proteins, which is typically the beta sheets 1 through 4. Alanine scanning of the entire peptide revealed several amino acids critical for facilitating the proteins interaction with RNA, and almost all of these residues are in the Csr domain, with the first 5 amino acids all being critical(25). Genomic alignments across the gene family show that these amino acids are very conserved with most diversity arising in the C terminus post Csr defined domain(19, 25). NMR analysis of the RsmE-*hcnA* RNA complex revealed that the homodimer was able to bind and stabilize two stem-loop RNA structures. This interaction was facilitated primarily through interactions with residues in beta sheets 1, 3, and 4 within the Csr domain(25).

Furthermore, this binding was facilitated through interactions with the peptide backbone and not the side chains, suggesting that RNA specificity is an emergent

property of the three dimensional shape of the dimerized protein(25). However, Rsm proteins are very conserved in all regions shown to interact with RNA, with typical substitutions within the same amino acid class. Rsm proteins vary widely post Csr domain, but this region hasn't been shown or predicted to interact with RNA or be involved with dimerization(27). Because of the high conservation among Csr/Rsm homologs, Rsm paralogs are typically considered functionally redundant(26, 32, 33). This, however, poses the question of why multiple paralogs exist if they carry out the same functions. The presence of paralogs within a species can be attributed to regulatory redundancy or backup system of sort that performs the same function(20) . Multiple paralogs may also have overlapping contributions(32) or a cumulative effect(20) on function.

It has also been shown that Csr/Rsm paralogs are differentially expressed during the cell cycle(19) . In *P. putida*, the three Rsm paralogs, RsmE, RsmA, and RsmI, were found to be sequentially turned on and off, with RsmE and RsmA being expressed earlier in the cell cycle than RsmI(20). While differential expression can support functional redundancy, it can also suggest differential function of the Rsm paralogs. Our model species, *P. fluorescens* Pf0-1, is uniquely equipped to allow us to study Rsm regulation and to tackle this question of functional redundancy between the paralogs.

P. fluorescens and our experimental model

P. fluorescens Pf0-1 was initially isolated from soil and was described as having strong adhesion to soil particles with lophotrichous flagella (34, 35). With a point mutation in the GacA protein, the GacA/GacS regulation system is nonfunctional in this

particular strain(36). This simplifies the regulatory networks and promotes the Rsm proteins to an even more central role in regulation.

Like *P. putida*, *P. fluorescens* contains three Rsm paralogs, RsmA, RsmE, and RsmI. High sequence similarity between these paralogs suggests that they are functionally redundant(14, 37). Aging *P. fluorescens* colonies repeatedly select for mutants that expand space and push themselves up to the surface of the colony, which is rich in oxygen supply and less restrictive in space for further expansion(1). A loss-of-function mutation of a single post-transcriptional regulator gene, *rsmE*, was responsible in every mutant screened (Figure 2). These mutants then spatially outcompete the parental strain when grown together, but the individual mutants exhibit no change in maximal growth rate compared to the parent when cultured in isolation(1). However, these mutants produce extracellular secretions not observed in the parental strain. This suggests that the fitness benefit conferred is a direct result of the spatial structures they form, likely mediated by these secretions.

These observations present an exciting opportunity for which this dissertation examines the underlying molecular mechanisms driving the interplay between the metabolic changes, spatial positioning, and the molecular functions governed by the Rsm system.

Specific Aims

Pseudomonas fluorescens Pf0-1 with its three Rsm paralogs raises several questions pertaining to the functional nature of these proteins. Most centrally is the question of why mutations in *rsmE* are exclusively selected for in these dense

communities, resulting in mutant patches with increased fitness over the parental strain. By using a systematic approach, combining the power of competitional analysis, fluorescent microscopy, genetic engineering, protein biochemistry, and transcriptomics, I have started the unveiling of the Rsm regulation network and its downstream ramifications in *Pseudomonas fluorescens* Pf0-1. My work was accomplished through the following aims and the studies linked to them as presented in this dissertation.

Aim 1. Characterize impact of extracellular secretions produced by *rsmE* mutants on relative fitness and spatial structuring. This aim addresses the role of the extracellular secretions and was accomplished by disrupting biosynthetic pathways associated with the mucoid phenotype identified through transposon mutagenesis and characterizing their influence on relative fitness using competitional analysis and spatial structuring using fluorescent and confocal microscopy, as detailed in chapter 2.

Aim 2. Characterize the changes made to the secretome of *Pseudomonas fluorescens* with the evolution of *rsmE* mutants. This aim addresses the question of how does the derepression of the RsmE-secretome change fitness at the molecular level. This aim was addressed through comparative proteomics, and RNA profiling using mass spectrometry and RNA-seq, as detailed in chapter 3.

Aim 3. Characterize regulation network of RsmE in relation to each Rsm paralog. This aim addresses the question of why only mutation in RsmE result in the mucoid phenotype. This was achieved through quantitative PCR in chapter 2 and through comparative transcriptomics using RNA-seq, and the work is detailed in chapter 3 and discussed in chapter 4.

Significance

This work provides insights into the regulation of Rsm proteins, the pathways they regulate, and identifies cellular mechanisms which lead to the creation and protection of physical space in a crowded colony. It provides clear insight into the consequences of derepression of RsmE regulation using functional genomics to identify key mechanisms which contribute to increased fitness. This work distinguishes the discrete functions of the Rsm paralogs adding to the counter narrative against functional redundancy in this class of proteins.

Figures

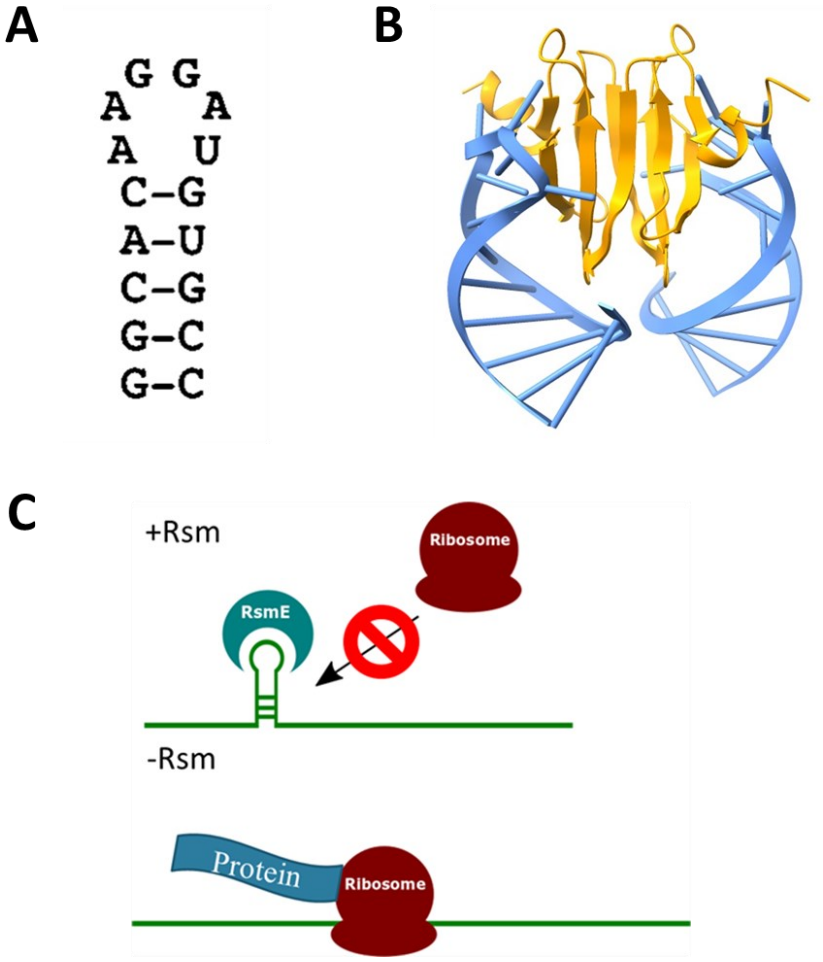


Fig. 1. The canonical Rsm regulation mechanism starts with the protein binding and stabilizing a pentaloop structure within the Shine Dalgarno sequence of the target mRNA (**A**). Each Rsm functions as a homodimer able to bind two discrete mRNA molecules at this site (**B**). Binding the Shine Dalgarno sequence prevents ribosomal docking and translation of the Rsm-bound mRNA (**C**).

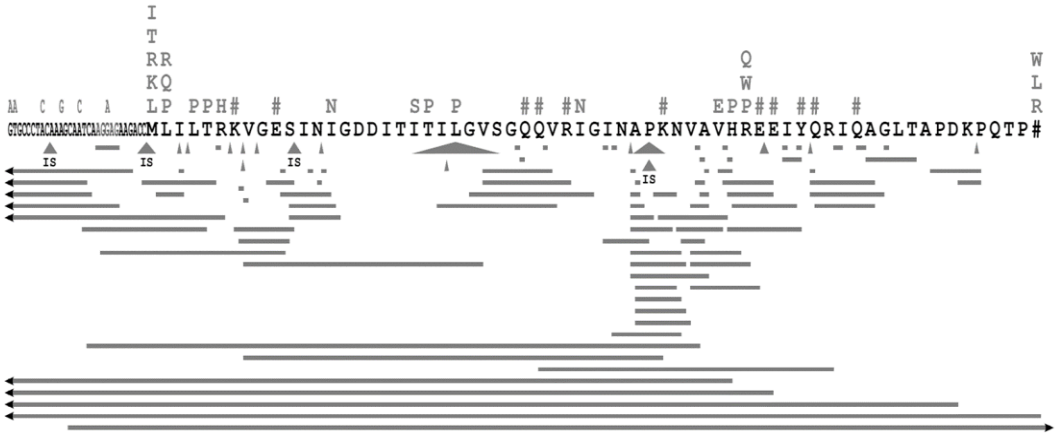


Fig. 2. This figure was adapted with permission from Kim et al. 2014, and depicts a summery the mutations found in denovo M strains independently isolated from WT colonies(1). Nonsense mutations noted by # and missense mutation noted by the substituted amino acid above the location with mutations in the 5' UTR are noted with the base pair change. Triangles indicate the location and size of insertions sequences (IS). Bars represent the size and location of deletions, with arrows indicating deletions extending beyond the shown sequence.

Chapter 2: Spatial structure formation by RsmE-regulated extracellular secretions in *Pseudomonas fluorescens* Pf0-1

Preface

The contents of this chapter were published in the Journal of Bacteriology on September 27, 2022, under the same title. Excluding myself, this manuscript was Authored by Meghan Wells, Jordan Denk, William Mazza, Raziel Santos,Amber Delprince, and Wook Kim, with their contributions noted in the Acknowledgement.

INTRODUCTION

Central to the architecture of microbial communities is the extracellular matrix(38-45), a dynamic cumulus of compounds produced by individual cells that physically define both the spatial arrangements within and the three-dimensional boundaries. Micron-scale spatiogenetic structures readily emerge within surface-grown communities as individual cells produce identical copies of themselves in a given area(46-48). Competition between different genotypes lead to the spatial enrichment of a particular genotype, producing macroscopic regions that stem from a recent common ancestor(47, 49-52). Individual phenotypes could positively or negatively impact the fitness of neighboring cells, including the consumption of limiting nutrients and the secretion of enzymes and toxins that promote or discourage the growth of neighboring cells(45, 53-55). Mechanistic understanding of how individual phenotypes antagonize or synergize with another clearly carries both fundamental and clinical significance.

Researchers employ various experimental approaches to study the interactive dynamics of microbial cells within a community, whether it be computationally(46, 47, 56) or empirically on a variety of abiotic surfaces(42, 57-59). We have previously described a model system based on *Pseudomonas fluorescens* colonies, which shows how spatial structures rapidly evolve within clonal aggregates(1). Mucoid patches repeatedly emerge on the surface of aging colonies due to the activity of specific mutants where they expand space and decrease local density. Remarkably, a mutation in a single gene, *rsmE*, was responsible for each and every case of over 500 independently derived mucoid patches. Importantly, *rsmE* mutants share the same growth rate in isolation compared to the parent cells, and the evolutionary advantage specifically requires the proximal presence of the parent cells. These observations collectively suggest that RsmE-regulated phenotypes physically act to create dominant spatial structures in a densely populated bacterial colony.

RsmE belongs to the CsrA/Rsm family and its homologs are a regulator of social and virulence phenotypes in gamma-proteobacteria(60, 61). CsrA was the first member of the family to be discovered three decades ago in *Escherichia coli* (18), and its homologs are now known to be present in over 2900 species(19). CsrA/Rsm proteins interact with diverse mRNA(22-24, 61) and primarily function as a translation repressor by either directly or indirectly blocking their respective Shine-Dalgarno (SD) sequence(30, 31, 62, 63). CsrA also possesses additional regulatory functions that impact Rho-dependent transcription attenuation, mRNA stabilization and destabilization, and even activation of translation(63). In contrast to CsrA in Enterobacteriaceae, *Pseudomonas* spp. possess varying numbers of Rsm paralogs(19). Rsm paralogs were

initially characterized to repress the production of secondary metabolites and are generally described to overlap or cumulate in function(14, 21, 26, 32, 33, 37, 64).

Although the three paralogs in *P. fluorescens* (RsmE, RsmA, and RsmI) share high sequence similarity, the exclusive selection of mutations in the *rsmE* locus(1) suggests functional specificity of RsmE from its paralogs. Here, we show that all three Rsm paralogs are expressed, but RsmE uniquely governs the production of both a mucoid polymer and a biosurfactant. The biosynthetic genes of the mucoid polymer were previously described(65), and we identify the biosynthetic genes of the biosurfactant in this study. Competition and microscopy analyses of the extracellular polysaccharide and biosurfactant mutants reveal that these extracellular secretions function collectively to confer fitness benefit as a direct result of the spatial structures they form.

METHODS

Strains and culture conditions

Liquid and solid Lennox LB media (Fisher) were used for general overnight cultures. *Pseudomonas* Agar F (PAF) media (Difco) was used for all phenotypic screens, competitions, and microscopy. *Pseudomonas* minimum medium (PMM: 3.5mM Potassium phosphate dibasic trihydrate,2.2mM potassium phosphate monobasic, 0.8mM ammonium sulfate, 100mM Magnesium sulfate, 100mM sodium succinate) was used to selectively grow *P. fluorescens* isolates from conjugations with *Escherichia coli* donors. Routine cloning was carried out in *E. coli* 10B (Invitrogen) or *E. coli* JM109 (Promega), and *E. coli* S17.1λpir(66) was used as the donor strain in conjugations. When required, antibiotics were added to the media at the following final concentrations: kanamycin

(50µg/mL), streptomycin (50µg/mL), ampicillin (100µg/mL), and gentamicin (20µg/mL). *P. fluorescens* was cultured at 30°C or at room temperature (~22°C), and *E. coli* strains were cultured at 37°C. Liquid cultures were incubated while shaking at 250 rpm. All *P. fluorescens* strains used in this study are listed in Table 1.

Quantitative PCR

Total RNA was isolated from colonies grown for 3 days at room temperature on PAF plates using the TRIzol® Reagent (ThermoFisher) under the manufacturer's protocol. Total RNA quality and concentration were assessed using the NanoDrop spectrometer. First strand cDNA synthesis was carried out using the High-Capacity RNA-to-cDNA kit (Applied Biosystems) with 1µg of RNA and random hexamers following the manufacturers protocol. qPCR optimized primers were obtained from Integrated DNA Technologies (Table 2) and their quality was assessed through PCR with gDNA, cDNA, and no-RT cDNA reactions. qPCR was performed using SYBR Green (ThermoFisher) on the StepOnePlus™ instrument (Applied Biosystems). Each reaction was analyzed to ensure only one amplicon was amplified using dissociation curves. Gene expression was calculated using the $2^{-\Delta\Delta CT}$ method with the 16S rRNA gene as an internal reference and quantified relative to *rsmI* expression(67).

Biosurfactant assay

Nuclepore Track-Etch polycarbonate membranes (Whatman: 0.4µM pore size, 90mm diameter) were used for assessing biosurfactant production. As previously described(1), one side of the membrane is shiny and the other is dull due to the manufacturing process. The dull side's surface contains gaps and ridges that physically

trap cells, but the biosurfactant permeates to produce the visible ring around colonies. The shiny side's surface is smooth, which allows biosurfactant producing cells to spread out through growth. Sterile forceps were used to overlay the membrane on the PAF agar surface, and 20µl of overnight culture was spotted directly on the membrane and allowed to fully dry before the plates were inverted and incubated over night at room temperature.

Identification of the biosurfactant biosynthesis genes by transposon mutagenesis

Random transposon mutagenesis, using the plasmid pUT-mini*T*n5-Km*lacZ2* (68, 69) in *E. coli* S17.1λpir as the donor, was carried to identify the biosynthesis genes of the biosurfactant as previously described to identify the biosynthesis genes of the mucoid polymer(65). Briefly, overnight cultures of the donor and M* strains were washed in PMM, mixed at the relative ratio of 1:6, spotted on solid LB to conjugate, incubated at 30°C for three hours, harvested and plated out on solid PMM supplemented with kanamycin. Over 20,000 transconjugant colonies were picked and rearrayed using the QBot (Genetix) into 384-well plates containing kanamycin supplemented PMM, then incubated at 30°C. Surfactant assays on overnight cultures were conducted on PMM plates overlayed with the dull side of the polycarbonate membrane as described above with a disposable 384 pin replicator (Scinomix). Mutants that were defective in biosurfactant production (dull side) were rearrayed into 96-well plates containing kanamycin supplemented PMM, then incubated at 30°C. Overnight cultures were retested for biosurfactant production as described above using a disposable 96 pin replicator (Scinomix). Mutants that failed to produce the biosurfactant ring were selected, ignoring ones that had obvious growth defects. The transposon insertion sites were identified by arbitrary primed PCR as previously described(65).

Mutant construction and tagging

Gene deletion mutants were constructed by the gene splicing by overlap extension method (70), using the plasmid pMQ30(71) or pSR47s(72) as previously outlined(1, 65). PCR primers used to construct and confirm each mutation are listed in Table 2. Briefly, for each targeted gene, approximately 500bp of its flanking upstream and downstream regions were individually amplified, joined together, first cloned into the pGEM-T Easy Vector system (Promega) then sub-cloned into pMQ30 or pSR47s, and transformed into *E. coli* S17.1λpir as the donor strain. Overnight cultures of the donor and target strains were washed in PMM and mixed at an equal ratio, spotted on solid LB, incubated at 30°C overnight, harvested and plated out on solid PMM supplemented with gentamicin (pMQ30) or kanamycin (pSR47s). Transformants were grown on solid LB supplemented with sucrose (5%, w/v) overnight, and the resulting colonies were screened using primers that bind outside the two flanking fragments for expected reduction in the amplicon size. To confirm the gene deletions, we isolated genomic DNA from overnight cultures using the DNeasy UltraClean Microbial Kit (Qiagen) following the manufacturer's protocol and whole genome sequencing was conducted at the Microbial Genome Sequencing Center (MiGS; Pittsburgh, PA). Kanamycin-resistant and streptomycin-resistant strains used in competitions and GFP-tagged and DsRed-Express-tagged strains used in microscopy were constructed using the mini-*Tn7* chromosomal insertion system (73) as previously described (1, 65).

Measurement of monoculture growth

For the measurement of growth in colonies, overnight cultures were resuspended in PMM and 20µL was spotted on PAF plates and incubated at room temperature. To enumerate the initial population size, each cell suspension in PMM was serially diluted and plated out on LB plates, and resulting colonies were counted on the following day. Three spotted colonies were scraped on each day over seven days and resuspended in 5 mL of PMM using a sterilized bent glass Pasteur pipette. Cell suspensions were vortexed until clumps were no longer visible then serially diluted and enumerated as above. For the measurement of growth in liquid, overnight cultures were diluted into a manually formulated PAF without agar (74) in six replicates, and optical density at 600 nm was measured every 30 minutes over 48 hours (30°C, constant shaking) in the Bioscreen C MBR (Oy Growth Curves Ab Ltd.).

Competition assay

Competitions between kanamycin-resistant mutant strains and streptomycin-resistant WT strain were conducted as previously described (1). Briefly, overnight cultures (1.5 mL) were washed in fresh PMM and re-suspended in 1.0 mL PMM, and the mutant strain suspension was serially diluted to 10^{-3} in PMM and mixed with equal volumes of the undiluted WT strain suspension. 20µL of each competition mixture was spotted in triplicate on a PAF plate and incubated at room temperature. To enumerate the initial population size of the competing strains, each competition mixture was serially diluted and plated out on LB plates supplemented with either kanamycin or streptomycin, and resulting colonies were counted on the following day. Four or seven days later, the spotted colonies were scraped and resuspended in 5mL of PMM, serially diluted, plated, and counted as for the initial competition mixture. The results of the competitions were

analyzed by calculating the relative fitness (*W*) of each competing strain against the WT(75).

Statistical analysis

Competition experiments were conducted with at least three biological replicates and two technical replicates for each biological replicate. The data were first analyzed with ANOVA to evaluate if the means of the biological replicates differ significantly, then Tukey's honest significant difference test ($p < 0.05$) was applied to make multiple pairwise comparisons within the dataset. All comparisons were found to be statistically different or noted as n.s. otherwise. Statistical tests were conducted using GraphPad Prism.

Microscopy

Overnight cultures of GFP-labeled strains and DsRed-Express-labeled WT were washed and resuspended in PMM. All GFP-labeled cell suspensions were serially diluted to 10^{-5} in PMM and mixed with equal volumes of the undiluted DsRed-Express-labeled WT suspension. 20µL of each competition mixture was spotted in triplicate on PAF plates and incubated at room temperature. Epifluorescence microscopy was conducted using the Nikon SMZ25 stereo-compound microscope with the 0.5 X SHR Plan Apo objective and the NIS Elements software. For confocal microscopy, an agar slice containing the entire colony was placed on a microscope slide and visualized without a coverslip. Confocal microscopy was conducted using the Nikon Ti2 microscope with the 20X TU Plan Fluor objective or the air-corrected 100X TU Plan Apo objective and the NIS Elements software. Non-fluorescent imaging of colonies was carried out using the

Hayear overhead microscope (HY-2307) or the Canon Rebel EOS T3 DSLR camera. Images were rendered using the NIS Elements and ImageJ software.

RESULTS

RsmE, RsmA, and RsmI in *Pseudomonas fluorescens* Pf0-1 are highly conserved in sequence and all three respective genes are simultaneously expressed

Pseudomonas fluorescens Pf0-1 possesses three Rsm paralogs – RsmA, RsmE, and RsmI – that share high sequence similarity (Fig. 1A). We sought to first determine whether or not all three respective genes are expressed. Quantitative PCR confirmed that all three genes are indeed simultaneously expressed, with *rsmA* and *rsmI* transcripts being the most and least abundant, respectively (Fig. 1B). These results show that the exclusive selection of *rsmE* mutations in our previous experimental evolution study(1) was not simply due to the absence of *rsmA* and *rsmI* expression under the same experimental conditions.

RsmE specifically regulates the production of a mucoid polymer and biosurfactant

Experimentally selected *rsmE* mutants visibly produce a mucoid polymer and/or a biosurfactant(1), which suggests that specific mutations differentially impact RsmE's function. To determine if these extracellular secretions are commonly regulated by the three Rsm homologs, we constructed deletion mutants of the respective genes. Comparison of colony morphologies show that only the *rsmE* mutant exhibits mucoidy (Fig. 2A). In addition, mucoid patches consistently emerge in colonies of WT, *rsmA* mutant, and *rsmI* mutant (Fig. 2A), which are characteristic of naturally mutated *rsmE*

(1). These results confirm that the production of the mucoid polymer is specifically regulated by RsmE. We next compared biosurfactant production on a polycarbonate membrane overlaid on the agar surface. Production of the biosurfactant on the shiny side of the membrane allows the colony to spread out radially, but the cells remain trapped on the dull side of the membrane while the biosurfactant spreads out unhindered(1). Only the *rsmE* mutant produced a visible ring on the dull side of the membrane and also spread out on the shiny side of the membrane (Fig. 2B). These results confirm that the production of both the mucoid polymer and the biosurfactant is uniquely governed by RsmE from its paralogs.

Identification of the biosurfactant as gacamide A

The biosynthetic genes of the mucoid polymer were previously characterized to encode a glucose-rich extracellular polysaccharide, and a corresponding gene was deleted in a mucoid (M) strain with a frameshift mutation in *rsmE* (1) to produce the non-mucoid M* strain(65). To identify the biosynthetic genes of the biosurfactant, we carried out a random transposon mutagenesis in the M* strain background. Seven mutants were independently isolated that no longer produced the secretion on the dull side of the polycarbonate membrane and failed to spread out on the shiny side of the membrane. All transposon insertion sites were mapped to three contiguous loci (annotated as Pfl01_2211, Pfl01_2212, and Pfl-1_2213), which were recently demonstrated to encode non-ribosomal peptide synthetases(76) that produce the cyclic lipopeptide gacamide A (77). Cyclic lipopeptides are indeed classified as a surfactant, and they contribute to surface motility and biofilm formation in many *Pseudomonas* spp.(78, 79). Given that four independent transposon insertions occurred in the Pfl01_2211 locus, we constructed

a corresponding in-frame deletion mutant in the M strain to produce the M^s strain, and the same mutation was also introduced in the non-mucoid M* strain to produce the M^S* strain. Neither M^S nor M^S* produce the biosurfactant ring on the dull side of the membrane and the spreading phenotype on the shiny side of the membrane (Fig. 3), confirming that the Pfl01_2211-Pfl-1_2213 cluster encodes the production of the biosurfactant. Importantly, M* maintains the production of the biosurfactant and M^S maintains the production of the mucoid polymer (Fig. 3), which shows that the biosynthesis of these two secreted products are not genetically linked to one another, but are both regulated by RsmE.

Both the mucoid polymer and biosurfactant confer competitive advantage

All experimentally selected *rsmE* mutants outcompete the WT strain in co-cultured colonies(1). To assess the contributions of the RsmE-regulated mucoid polymer and the biosurfactant, we independently competed M, M^S, M*, and M^S* against the WT in co-cultured colonies and assessed their fitness relative to the WT. All four strains outcompeted the WT throughout the duration of the experiments (Fig. 4), with M and M^S being nearly equal in fitness and M* and M^S* exhibiting decreased fitness at day 4. However, we observed reduced fitness in all secretion mutants compared to M by day 7, with M^S and M* being comparable and M^S* exhibiting further reduction. Such step-wise decreases in fitness indicates that each secreted product independently confers competitive advantage and the two secretions also likely function in an additive manner. Furthermore, the fact that M^S* retains the ability to outcompete the WT indicates that there are additional RsmE regulated genes that contribute to M's dominance over the WT.

The mucoid polymer creates space and the biosurfactant prevents the diffusion of the mucoid polymer at the colony surface

The temporal differences in the relative fitness between M^S and M* (Fig. 4) suggests that the mucoid polymer plays a more significant role early in the competition. Importantly, our secretion mutants exhibit equal growth profiles compared to the WT as monoculture in both liquid and colonies (Fig. 5). The M data here recapitulates the results from our previous study, which also demonstrated that the competitive advantage of *rsmE* mutants specifically requires the formation of spatial structures that decreases local density and provides greater access to oxygen(1).

To explore the functional role of the RsmE-regulated mucoid polymer and biosurfactant in spatial structure formation, we carried out epifluorescence and confocal microscopy analyses of our collection of secretion mutants against the WT. We first introduced a constitutively expressed *gfp* gene into the chromosome of WT, M, M^S, M*, and M^S* strains. Each green fluorescent strain was mixed with red fluorescence labeled WT at the respective ratio of 10^{-5}:1 to best visualize isolated spatiogenetic structures in colonies after 5 days. Epifluorescence imaging of entire colonies shows isolated green fluorescent patches emerging from mostly red fluorescent WT colonies, with M and M^S producing consistently bigger patches compared to M* and M^S* (Fig. 6A). Each co-culture also produced red fluorescent mucoid patches, which represents *de novo rsmE* mutants naturally emerging from the red fluorescent WT cells(1), however, no green fluorescent patches were observed in the WT:WT colonies. With confocal imaging at a low magnification, the green fluorescence signal in the smaller patches formed by M* and M^S* is much more intense compared to those formed by M, and M^S patches

produced the least intense fluorescence signal (Fig. 6B). Individual patches formed by both M and M^S typically merged together with nearby patches through continuous expansion over time, but we consistently observed M^S patches to be much more amorphous in structure with less defined individual boundaries. In contrast, green fluorescent WT patches were rarely observed and appeared to comprise only few cells.

Confocal imaging using an air-corrected 100X Plan Apo objective provided a clear view of individual green fluorescent cells and their spatial arrangement within a given patch surrounded by red fluorescent WT cells (Fig. 6C). M cells were present at a strikingly lower density compared to the neighboring WT cells, with the characteristic black space that is devoid of cells(1). In addition, M patches are defined by a clear boundary formed with a thin layer of M cells which appears to exclude the encroachment of WT cells into the black space. In contrast, M^S patches lacked a clear exclusionary boundary, with M^S cells appearing to flow over the WT cells. This interpretation is also reflected in the lower magnification observations of M^S patches being more mucoid and amorphous (Fig. 6A) and producing much less intense fluorescent signal (Fig. 6B) compared to patches formed by M. M* and M^S* both formed much densely packed patches with clear boundaries against the WT cells, but M^S* cells appear to be even more packed as indicated by the uniquely vertical arrangement of cells (Fig. 6C) and much smaller size of individual patches (Fig. 6A). These observations collectively suggest that the mucoid polymer is the primary driver of creating space while the biosurfactant spatially sequesters the mucoid polymer to prevent their diffusion. However, M^S* retains the ability to produce a spatiogenetic structure that contrasts greatly from the green fluorescent WT cells that form small clusters of only few cells without any organized

structure (Fig. 6C), likely representing daughter cells stemming from initially a single mother cell. As already reflected in our relative fitness data (Fig. 4), there appears to be additional RsmE-regulated genes that specifically promote spatial competition in a densely populated colony.

DISCUSSION

Several members of the gamma-proteobacteria, including *Pseudomonas* spp., possess multiple paralogs of CsrA/Rsm proteins, and their corresponding genes are also present in diverse plasmids and bacteriophages(19). We had previously shown that mutations in *rsmE* are repeatedly selected as mucoid patches in colonies of *P. fluorescens* Pf0-1 by creating space and capturing optimal positioning within a crowded environment(1). The exclusive association of *rsmE* mutations with this striking phenotype suggests that RsmE's function is not entirely redundant from that of its paralogs, RsmA and RsmI. In this study, we show that all three paralogs are accessible to evolutionary selection, since their respective genes are simultaneously expressed under the same experimental conditions. Furthermore, mucoid patches consistently emerged in both *rsmA* and *rsmI* knockout colonies, much like those that emerge from WT colonies through *de novo* mutations in *rsmE*. These observations strongly support our prediction that the formation of beneficial spatial structures occurs specifically through mutations that deregulate RsmE's native function.

We have shown that knocking out *rsmE* results in the production of two visible extracellular secretions, a mucoid polymer and a biosurfactant, but neither are produced

in *rsmA* nor *rsmI* knockouts. Thus, RsmE appears to either directly repress the production of these secretions or modulate the activity of other regulators that directly govern their production. Genetically removing the production of either or both secretions in the *rsmE* mutant significantly reduced competitive advantage against WT in co-cultured colonies. However, all engineered secretion mutants shared the same growth profiles compared to WT in liquid and colony monocultures. These observations collectively suggest that both secretions contribute to the spatial structure formation by the *rsmE* mutant, and we confirmed this prediction through epifluorescence and confocal microscopy.

The two key characteristics associated with the dominant spatial structure formed by the *rsmE* mutant are creation of space with low cellular density and exclusion of the neighboring WT cells from this local environment(1). Here, we have demonstrated that the mucoid polymer is solely responsible for creating the space. We had initially interpreted that the biosurfactant forms the exclusionary boundary due to the mixed presence of the biosurfactant knockout and WT cells. However, we consistently observed that the WT cells rarely invade deeply into the areas of low cellular density at high optical magnification. In addition, the borders of individual patches formed by the biosurfactant mutant were less defined and the mutant cells appeared to flow out on top of the neighboring WT cells, akin to outflowing lava from a volcano. However, these observations indirectly contradict the results of our membrane assay, which showed that the same biosurfactant promotes the spreading of cells on the membrane surface. In fact, we initially referred to the corresponding secretion as a biosurfactant, due to the well-known function of bacterial surfactants that reduce surface tension to promote swarming on semi-solid agar surfaces(80).

We identified the biosynthetic genes of the biosurfactant in this study, which were recently characterized by an independent group to produce a cyclic lipopeptide named gacamide A that promotes swarming(77). *Pseudomonas* spp. produce numerous cyclic lipopeptides that variably contribute to surface-spreading and biofilm formation, and this variability potentially depends on discrete interactions with diverse extracellular or cell membrane-associated products(79, 81). The amphiphilic structure of gacamide A likely promotes its interaction with both hydrophilic compounds, like the mucoid polymer, and hydrophobic compounds that co-accumulate within the patches formed by the *rsmE* mutant. Importantly, removing the production of both the mucoid polymer and gacamide A maintained the respective *rsmE* mutant's ability to outcompete the WT, albeit with much reduced spatial dominance. These observations suggest that there are additional RsmE-regulated products that contribute to the competitive advantage of the *rsmE* mutant, which clearly manifests through beneficial structures(1). Pressure likely builds up internally within a localized patch as the accumulating mucoid polymer constantly pushes away the surrounding WT cells to expand space. We thus speculate that gacamide A physically stabilizes the mucoid polymer and additional RsmE-regulated products to prevent their diffusion at the surface of the colony, which is uniquely devoid of neighboring cells and provides much less resistance.

A potential criticism of this study is the utilization of bacterial colonies to explore spatial structure formation, which lack important mechanical properties that manifest in natural microbial communities(44). However, resolving the problem of space and resource constraints in a densely populated colony likely shares common principles with other organisms in different experimental systems. Extracellular polysaccharides

produced by *Vibrio cholerae* growing in microfluidic device biofilms promotes the formation of isogenic structures that exclude the neighboring non-producers(82), and glycolipid biosurfactants produced by *Streptococcus* spp. selectively displace competing genotypes on the tooth surface (83). Cyclic lipopeptide production in *Bacillus subtilis* is essential for fruiting body formation on an agar surface, and mutants that lack this biosurfactant initially form projecting columns, but they grow laterally and subsequently fuse together (84) much like our biosurfactant knockout cells. Our study also establishes a highly tractable experimental pipeline to identify and characterize additional RsmE-regulated products, and to explore why RsmA and RsmI are functionally excluded from the formation of spatial structures.

Figures

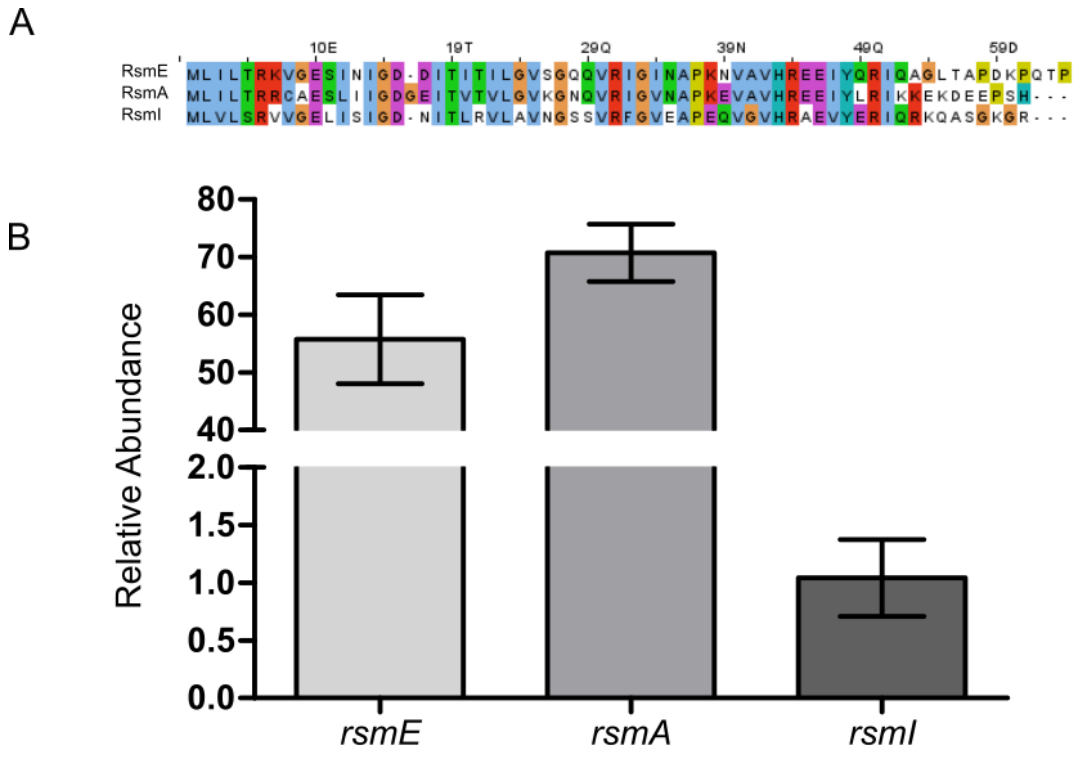


Fig. 1. Rsm paralogs in *P. fluorescens* Pf0-1 share a highly conserved sequence and their respective genes are simultaneously expressed. (A) Sequence alignment of Rsm-paralogs in *Pseudomonas fluorescens* Pf0-1 labeled using the ClustalX to show similarities in amino acid chemical properties(2). (B) Expression of *rsmE*, *rsmA*, and *rsmI* genes assessed in WT by qPCR. Transcripts of all three genes were detected and shown here is the relative abundance of each transcript using the $2^{-\Delta\Delta CT}$ method compared to that of the least abundantly expressed *rsmI*. Plotted are the mean of three biological replicates with three technical replicates for each biological replicate and the error bars represent the standard deviation of the mean.

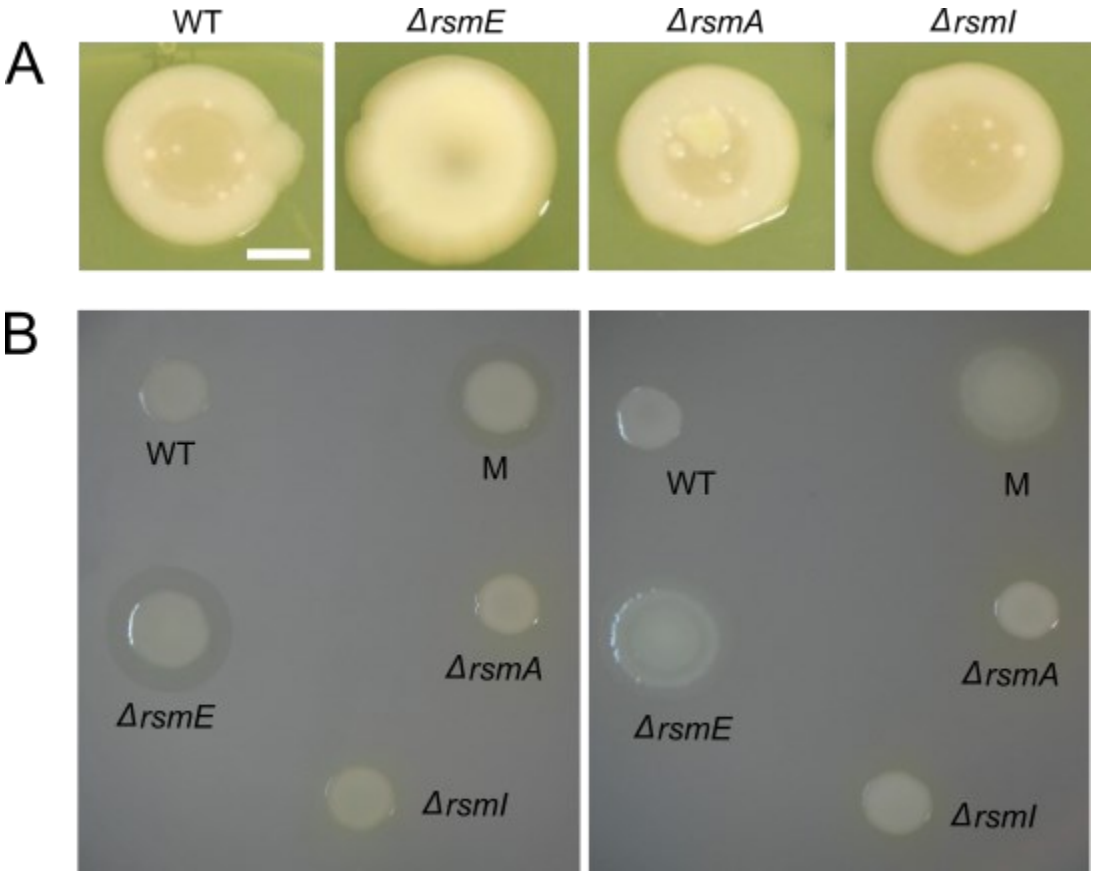


Fig. 2. Both the mucoid polymer and the biosurfactant are regulated by RsmE, but not by RsmA or RsmI. (A) Colony morphology comparisons of WT and deletion mutants of *rsmE*, *rsmA*, and *rsmI*. Liquid cultures were spotted on PAF plates seven days prior to capturing the images. Only the Δ*rsmE* strain is mucoid in appearance and new mucoid patches naturally emerge in WT, Δ*rsmA*, and Δ*rsmI* colonies that characteristically represent de novo *rsmE* mutations. The scale bar represents 10 mm. (B) Comparison of biosurfactant production on the dull (left) and shiny (right) sides of the polycarbonate membrane overlaid on PAF. The M strain is a naturally selected mutant from a WT colony harboring a frameshift mutation in *rsmE*. Only the Δ*rsmE* and M strains produce the biosurfactant ring on the dull side that promotes spreading of cells on the shiny side of the membrane.

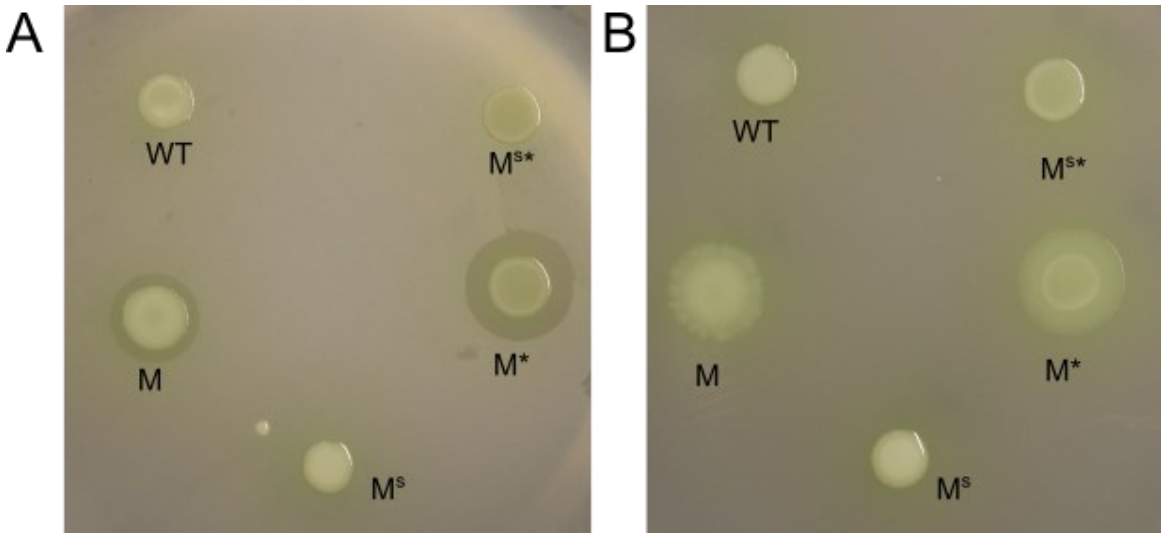


Fig. 3. Deletion of the Pfl01_2211 locus abolishes biosurfactant production. Shown are the results from the dull side (A) and the shiny side (B) of the polycarbonate membrane. M (*rsmE* mutant) and M* (M with the mucoid polymer biosynthesis gene (Pfl01_3834) deleted) produce the biosurfactant and spread on the surface, but M^S (M with Pfl01_2211 deleted) and M^{*S} (M* with Pfl01_2211 deleted) fail to do so like WT with an unaltered *rsmE* gene. These results confirm that the Pfl01_2211-2213 cluster encodes the biosynthetic genes of the biosurfactant, which is now known to be gacamide A.

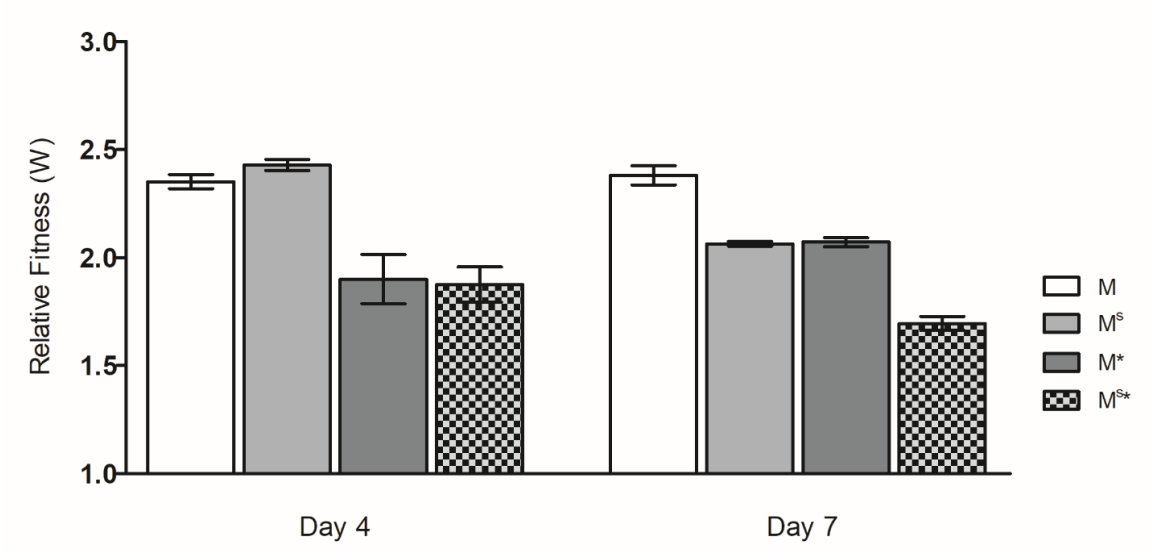


Fig. 4. Competitions of M, with or without mucoid polymer and/or biosurfactant production, against WT show varying levels of relative fitness over time. WT was chromosomally tagged with streptomycin resistance and all mutants were tagged with kanamycin resistance, and these resistance markers were previously shown to produce neutral relative fitness in *P. fluorescens* Pf0-1. Error bars represent the standard deviation of the mean relative fitness (mutant over WT) calculated from three independent populations after four and seven days of incubation. Dataset from each time point was analyzed by ANOVA ($p < 0.0001$) and Tukey's honest significant difference test showed that all pairwise comparisons were significantly different ($p < 0.05$) except for those indicated as nonsignificant (n. s.). Relative fitness (W) of 1 indicates equal fitness of the mutant and WT and a W of greater than 1 indicates that the mutant outcompeted the WT. Both the mucoid polymer and the biosurfactant provide competitive advantage. However, M^{S}* (*rsmE* mutant with biosynthesis genes of both secretions deleted) still outcompetes the WT, suggesting that additional RsmE-regulated products contribute to the competitive advantage of M (*rsmE* mutant) against the WT.

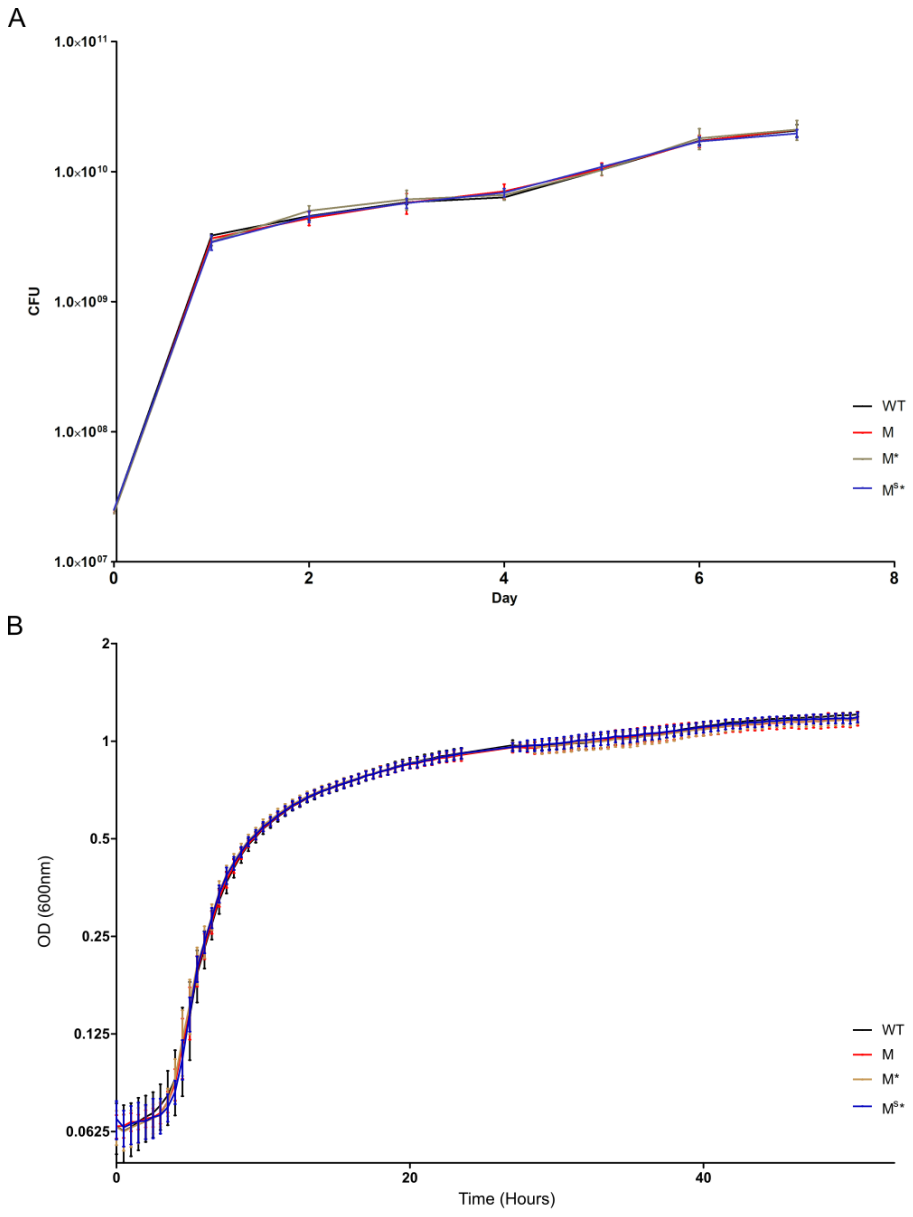


Fig. 5. Production of RsmE-regulated extracellular secretions does not impact growth in monoculture. (A) Growth profiles of single genotype colonies of WT, M, M*, and M^{S}* on solid PAF. Each data point represents the mean colony forming units (CFU) of three populations and the error bars represent the standard deviation of the mean. (B) Growth profiles of single genotypes in liquid PAF as measured by optical density at 600 nm.

Shown are the mean of six independent cultures for each strain, and the error bars represent the 95% confidence interval.

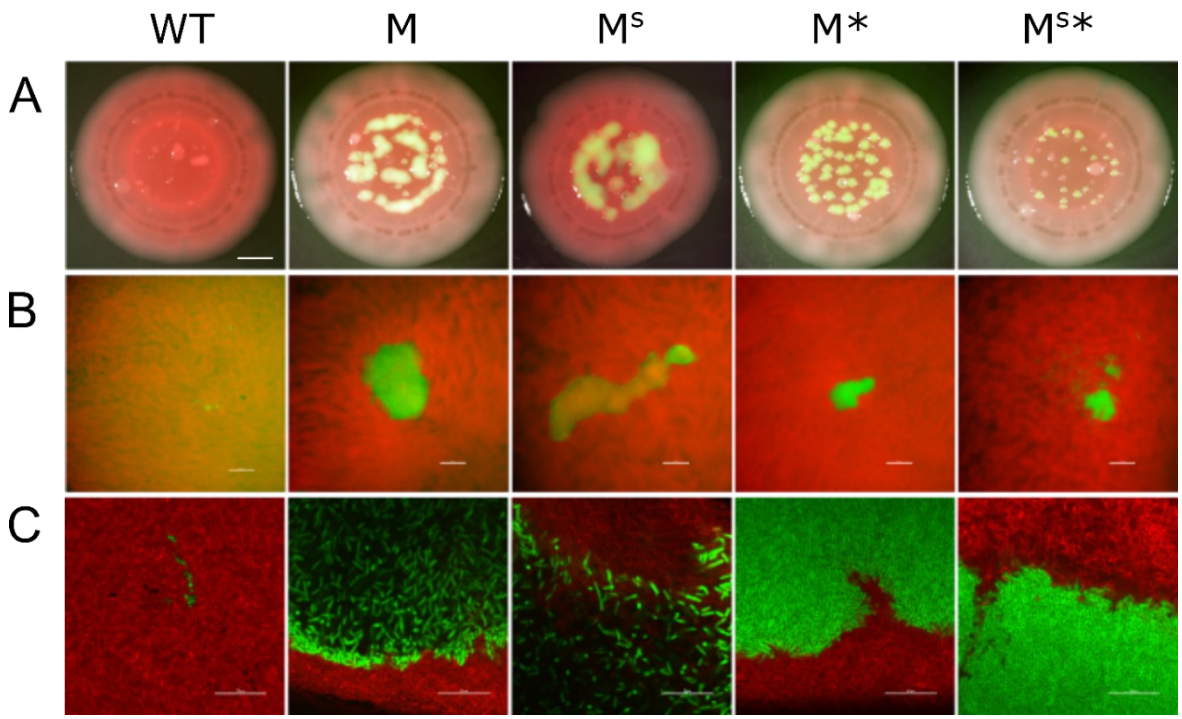


Fig. 6. The mucoid polymer and biosurfactant function together in the formation of dominant spatial structures. Each indicated strain was chromosomally tagged with GFP, heavily under-represented in a mixture with DsRed-Express-tagged WT, and representative co-cultured colonies were imaged five days later. (A) Epifluorescence microscopy images that capture the entire colony (A, scale bar represents 10 mm). Each sample shows the natural emergence of red mucoid patches that are characteristic of *de novo rsmE* mutants stemming from the red-fluorescent WT cells. (B) Confocal microscopy images focusing on the surface of individual patches at a low magnification (scale bar represents 50 um). M^S* produces unique patches that appear to be mixed with red WT cells. (C) Confocal microscopy images at a higher resolution focusing on the boundaries formed between the mutant and WT (scale bar represents 10 um). The mucoid polymer is solely responsible for creating the space of low cell density (black space is devoid of cells) and the biosurfactant appears to physically hold the mucoid polymer and

producing cells from flowing out from the newly created space. M^S* produces the smallest patches that are densely filled, as reflected by vertically aligned cells (spheres) similar to the WT:WT spatial organization (left panel). However, M^S* maintains the ability to form an organized structure that excludes WT cells, suggesting that additional RsmE-regulated products contribute to the spatial dominance of M.

Strain	Relevant genotype	Relevant phenotypes	Source
Pf0-1	WT	Non-mucoid, no biosurfactant	(35)
Pf0-1S	WT (*Tn7*-Sm^R)	Streptomycin resistance	(1)
Pf0-1R	WT (*Tn7*-DsRed-Express)	Red fluorescence	(1)
Δ*rsmE*	WT (Δ*Pfl01_1912*)	Mucoid, biosurfactant	This study
Δ*rsmA*	WT (Δ*Pfl01_4273*)	Non-mucoid, no biosurfactant	This study
Δ*rsmI*	WT (Δ*Pfl01_4104*)	Non-mucoid, no biosurfactant	This study
M	WT (126th nucleotide deleted in *rsmE*)	Mucoid, biosurfactant	(1)
MK	M (*Tn7*-Km^R)	Kanamycin resistance	(1)
MG	M (*Tn7*-Gfpmut2)	Green fluorescence	(1)
M*	M (Δ*Pfl01_3834*)	Non-mucoid, biosurfactant	(65)
M*K	M* (*Tn7*-Km^R)	Kanamycin resistance	This study
M*G	M* (*Tn*7-Gfpmut2)	Green fluorescence	(65)
M^S	M (Δ*Pfl01_2211*)	Mucoid, no biosurfactant	This study
M^{S}K	M^S (*Tn7*-Km^R)	Kanamycin resistance	This study
M^{S}G	M^S (*Tn*7-Gfpmut2)	Green fluorescence	This study
M^S*	M (Δ*Pfl01_2211* Δ*Pfl01_3834*)	Non-mucoid, no biosurfactant	This study
M^S*K	M^S* (*Tn7*-Km^R)	Kanamycin resistance	This study
M^S*G	M^S* (*Tn*7-Gfpmut2)	Green fluorescence	This study

Table 1. *P. fluorescens* strains used in this study

Name	Function[a]	Sequence
M^S-1	Outside F	CCCAACTGCCAATAAAAGTGCGG
M^S-2	Outside R	TGGGCAAATGCATCGAGGCG
M^S-3	Up F	TCCATGTTGCGAAGACTGCCG
M^S-4	Up R	CCCGTGACCGATCAGTCATGGAAGATCCATCT CACGTGATTTGGCG
M^S-5	Down F	CGCCAAATCACGTGAGATGGATCTTCCATGAC TGATCGGTCACGGG
M^S-6	Down R	CAGATCGCCTTCGCCAGCCG
M*-1	Outside F	TGACGATGGCCATGTGTTGCAGG
M*-2	Outside R	GCCGGGGGAGAACATGCAAC
M*-3	Up F	CGTCCCGTGCTTCTTCAACG
M*-4	Up R	CTAACATTGCACTAAAACGCCATCAAGTGGG AA
M*-5	Down F	CGAGTTGGCACTTTTCCCACTTGATGGCGTTTT A
M*-6	Down R	TCCAAAAACAAAGTCACCCG
rsmA-1	Outside F	GTGTCTACACCACTGACCCG
rsmA-2	Outside R	CAGAAAGCGAGAAATGGCCG
rsmA-3	Up F	CTACCAGTCCGCGCAAACC
rsmA-4	Up R	TCTTCCCCGTTTGCAAACATACCTTTCTCCTCA CGCGAATC
rsmA-5	Down F	GATTCGCGTGAGGAGAAAGGTATGTTTGCAA ACGGGGAAGA
rsmA-6	Down R	GGATTCGAACCTGTGACCGC
rsmI-1	Outside F	GCAACAGTCACCCTGACG
rsmI-2	Outside R	GAGCGCACGCACATCGAC
rsmI-3	Up F	GATGTCTTGCATGCTCAATTACC
rsmI-4	Up R	CCTGTTCGACTGAAAACGGGGGAGGTGTCCTT CAAGG
rsmI-5	Down F	CCTTGAAGGACACCTCCCCCGTTTTCAGTCGA ACAGG
rsmI-6	Down R	CTGGCGGCGCTCAGTCG
Pfl01_F	*16S* qPCR F	GCGTAGGTGGTTCGTTAAGT
Pfl01_R	*16S* qPCR R	CACCACCCTCTACCATACTCTA
*rsmE*_F	*rsmE* qPCR F	TGACATCACGATCACCATTCTC
*rsmE*_R	*rsmE* qPCR R	GTAGATCTCTTCCCGGTGTACT
*rsmA*_F	*rsmA* qPCR F	TGCTCGGCGTTAAAGGAAA
*rsmA*_R	*rsmA* qPCR R	CGCAGATAGATTTCTTCACGGT
*rsmI*_F	*rsmI* qPCR F	CATCACTCTGCGCGTTCT
*rsmI*_R	*rsmI* qPCR R	TTTGCGCTGGATCCGTT

Table 2. Primers used in this study

[a] For each gene target, Outside F/Outside R primers were used to screen for deletions, Up F/Up R primers were used to amplify the upstream fragment, and Down F/ Down R primers were used to amplify the downstream fragment. qPCR F/R primers were used to conduct qPCR of the indicated gene.

Chapter 3: Global analyses of RsmE-associated extracellular secretions that function in spatial structure formation by *Pseudomonas fluorescens*

Preface

The contents of this chapter represent a manuscript in preparation for publication. Excluding myself, additional authors for this manuscript will include but are not limited to Megan Wells, Amber Delprince, Raziel Santos, and Wook Kim with contributions as stated in the Acknowledgement.

INTRODUCTION

Bacteria form multicellular communities known as biofilms on virtually any biotic or abiotic surfaces, wherein individual cells of the same and different genotypes interact chemically and physically in close proximity. A universal feature of biofilms is that the entire community is encased within an extracellular matrix, which forms dynamically through the localized accumulation of diverse compounds secreted by individual cells(30, 38, 40, 85, 86). Given the protective nature of the extracellular matrix against predation, dehydration, and lethal chemical infiltration, biofilms have long been generalized to be a conserved genetic differentiation program rooted by cooperating individuals(39, 42, 46). However, recent studies have demonstrated that the production of matrix components are also stimulated in different genotypes to outcompete one another within the common space(4, 33, 86, 87). Although the matrix of individual biofilms comprises similar components in general – carbohydrates, proteins, nucleic acids, and

lipids - specific molecular compositions vary significantly across different species and environmental conditions(4, 38, 85, 86). Furthermore, many questions remain on the functional role of each matrix component, in particular, how they function together, if at all.

Experimental evolution studies of *Pseudomonas fluorescens* biofilms demonstrate that altered production of extracellular secretions lead to the emergence of striking social phenotypes. For example, mutations that elevate the production of aggregative extracellular secretions drive niche separation or cooperative colony expansion, while those that elevate the production of mucoid extracellular secretions drive spatial competition(1, 57, 65, 74, 88-90). In the latter case, diverse mutations in the *rsmE* gene produce a highly advantageous spatial phenotype that is specific to an overcrowded colony(1). RsmE is an RNA-binding post-transcriptional regulator, and the observed mutations appear to de-repress the production of at least two visible extracellular secretions, a polysaccharide composed primarily of glucose and a biosurfactant known to be the cyclic lipopeptide gacamide A(77). Depending on the specific mutation in *rsmE*, the production of the biosurfactant and the degree of fitness vary relative to the parent strain (88). The two secretions function collectively to create and protect the spatiogenetic structure, where the polysaccharide functions to push away neighboring cells to create space and the biosurfactant appears to physically sequester other secretions to maintain a defined genotypic boundary(88). However, knocking out both secretions in a Δ*rsmE* mutant retained its ability to outcompete the parent strain, albeit at a significantly reduced level, indicating that additional secretions contribute to the spatial structure formation(88).

P. fluorescens possesses two paralogs of RsmE – RsmA and RsmI – that do not influence spatial structure formation (88). In addition, the ability of the Δ*rsmE* mutant to outcompete the parent strain manifests exclusively when co-cultured in a structured population, indicating that RsmE-associated extracellular secretions specifically function to increase fitness through spatial interaction(1, 88). Here, we employ three global profiling approaches to identify RsmE-associated secretions that contribute to spatial structure formation beyond the already characterized polysaccharide and biosurfactant. We first characterize mRNA that bind directly to RsmE, then compare genes that are differentially regulated in Δ*rsmE* relative to Δ*rsmA*, Δ*rsmI*, and WT. We next identify extracellular proteins that are unique to Δ*rsmE* compared to WT. Lastly, we focus on genes that overlap in our transcriptomics and proteomics dataset to knockout four select secretion genes for microscopy analysis. Our systematic approach reveals RsmE-associated secretions that explain how Δ*rsmE* protects the spatial structure from the encroachment of neighboring WT cells and additional candidates for future studies.

METHODS

Bacterial strains and culture conditions

All *P. fluorescens* strains used in this study are listed in Table S1. *Escherichia coli* Jm109 (Promega) was used for routine cloning, *E. coli* strains S17-λ*pir* and HB101 were used for conjugations(34, 66). Liquid and solid LB (Fisher) were used for routine growth. Difco Pseudomonas Agar F (PAF; Fisher) was used for all assays, phenotypic screens, competitions, and microscopy. *Pseudomonas* minimum medium (PMM; 3.5mM

Potassium phosphate dibasic trihydrate, 2.2mM potassium phosphate monobasic, 0.8mM ammonium sulfate, 100mM Magnesium sulfate, 100mM sodium succinate) was used to selectively isolate *P. fluorescens* from conjugations with *Escherichia coli*. When antibiotics were added to the media when needed at the following final concentrations: tetracycline (12µg/mL), kanamycin (50µg/mL), gentamicin (20µg/mL), ampicillin (100µg/mL), and streptomycin (50µg/mL). *P. fluorescens* was cultured at 30°C or ambient room temperatures (~22°C) as indicated, while *E. coli* was grown at 37°C.Liquid cultures were shaken at 250 rpm while incubating.

Recombinant protein construction and expression

All PCR primers used in this study are listed in Table 4. The *rsmE* gene was amplified from WT genomic DNA using forward primers containing a 6x his tag. The N-terminally his tagged *rsmE* was cloned into pGEM-T Easy vector system (Promega), then subsequently cloned into pME6000, and transformed into *E. coli* S17-λ*pir(7)*. Multiparental mating was performed as previously to transform Δ*rsmE(88)*. Overnight 5 mL cultures of Δ*rsmE*-and Δ*rsmE* containing pME6000-His_6-*rsmE* was inoculated in 100 mL LB. Cultures were then grown for 2-3 hours or until they reached logarithmic growth, where protein expression was then induced with the addition of IPTG to a final concentration of 1 mM. Cultures were then incubated for an additional 16 hrs.

Protein purification

Induced cultures were collected after incubation and centrifuged at 5000 x *g* to pellet the cells. Cells were then resuspended in a cell lysis buffer (50mM

KH_2PO_4,300mM NaCl,10mM imidazole, 0.5% Tween-40, pH 8.0), and two freeze thaw cycles to lyse the cells. The lysed cell solution was then centrifuged to pellet the insoluble material, and the soluble fraction was transferred to a new tube. 0.5mL of HisPur Cobalt Resin (Thermo Scientific) suspended in storage buffer (20mM Tris, 150mM NaCl, pH 7.0) to added to the soluble cell lysate and incubated at 4°C for 1 hr. The entire solution was then passed through a gravity flow column (product number 29924, Thermo Scientific) where the cobalt resin collected at the bottom and the aqueous solution could freely pass through. The immobilized cobalt resin was then washed with five bed volumes of binding buffer (20mM Tris, 300mM NaCl, pH 8.0), five bed volumes of wash buffer (20mM Tris, 300mM NaCl, 75mM imidazole pH 8.0), and five bed volumes of elution buffer (20mM Tris, 300mM NaCl, 500mM imidazole pH 8.0). The elution was then buffer exchanged into storage buffer and concentrated using 3kDa MWCO concentrator tube (product number 88526, Thermo Scientific). Protein concentration was then calculated using a Nanodrop one (Thermo Scientific) spectrophotometer using bovine serum album (BSA) as a standard with measurements taken at A280 using a mass extinction coefficient of 6.7. To assess protein purity, samples were analyzed using sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) following the standard method using a 14% acrylamide resolving gel (91). Proteins were then visualized by silver stain following the manufactures protocol (Product number 24612, Thermo Scientific).

RNA purification and RNA-seq

For total RNA purification, WT, ΔrsmE, ΔrsmI, and ΔrsmA isolates were grown overnight in 5 mL cultures to saturation. Then, 20µL of each culture was spotted in triplicate on PAF plates and grown at room temperature for 3 days. After incubation, the resulting colonies were harvested, and RNA was extracted using Trizol reagent (Invitrogen) using the manufacturer's protocol. For isolation of Rsm-bound RNA, purified protein was denatured and partitioned from bound RNA using Trizol reagent (Invitrogen) again following the manufacture's protocol. Resulting RNA was quantified using a Nanodrop one (Thermo Scientific) spectrophotometer and the A260/A280 ratio calculated to assess purity. RNA samples were then directly submitted to Microbial Genome Sequencing Center (MiGS; Pittsburgh, PA) for sequencing.

Raw sequencing reads were processed using FastQC V0.11.5 (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/) to assess the quality of each data set. Reads were then aligned with the *P. fluorescens* Pf0-1 genome using HISAT-2 v2.1.0(92) and Bowtie2 V2.3.2(93). Cufflinks v2.2.1(94) was then used to assemble the transcripts, and Cuffdiff v2.2.1(94) was used to identify differentially expressed genes with Log2 fold change of 1 and a false discovery rate adjusted p-value (FDR) of 0.05.

Identification of extracellular proteins

20µL of overnight cultures of WT and Δ*rsmE* were spotted in triplicate on PAF and incubated at room temperature (~25°C) for 3 days. After incubation, three entire colonies and the medium immediately surrounding the colony (~20mm^2) was excised from the plate and placed in a 50 mL conical tube. Samples were immersed in protein extraction buffer (50mM Tris, 150mM NaCl, 1mM PMSF, 1mM EDTA, pH 7.6) and

vortexed until the cells were separated from the agar slices. The buffer was then transferred to a new tube leaving the agar slices behind. The contents of the tube were then passed through a 0.2 μm filter and collected in a new tube. The filtered supernatant was then concentrated using a 3kDa MWCO PES concentrator tube, and 400 μg of each sample was loaded onto a SDS-PAGE gel and run at 100V until the entire sample entered into the stacking gel (4% acrylamide). The gel was then fixed with 25% acetic acid. Samples were then excised and submitted to Michigan State University Proteomics Core (East Lansing, MI) for LC-MS/MS analysis. LC-MS/MS data was analyzed with scaffold viewer (Proteome Software, Inc). Spectra reads were aligned with the Pf0-1 genome. Protein threshold was set at 1.0% FDR, peptide threshold was set at 0.1% FDR with the minimal number of peptides of 3.

Knockout Mutant Construction

Knockout mutants were generated as previously described using the gene splice by overlap extension method(70, 88). Briefly for each targeted gene, ~500bp directly upstream and downstream of the gene was amplified using the up set and down set of primers (Table 4.) respectively and subsequently joined together by PCR uing the forward up primer and the reverse down primer. The ~1000bp joined flanking regions were cloned into pGEM-T, then subcloned into pMQ30 and transformed into *E. coli* S17.1λpir. Mating was then used to transform Δ*rsmE* with pMQ30 as previously described(88). Resulting colonies were then screened using the outside primer set (table s2) for the expected reduction in amplicon size. Gene deletions were further confirmed

using whole genome sequencing at the Microbial Genome Sequencing Center (MiGS; Pittsburgh, PA).

Microscopy

Overnight cultures of WT and each mutant were washed and resuspended in PMM. All mutant strain cell suspensions were serially diluted to 10^{-5} in PMM and mixed with equal volumes of the undiluted WT suspension. All strains were unlabeled except for the GFP-labeled strains as noted. Propidium iodide (Invitrogen) was added to the mixture at a concentration of 300μg/mL. 20μL of each competition mixture was spotted in triplicate on PAF plates and incubated at room temperature. Epifluorescence microscopy was conducted using the Nikon SMZ25 stereo-compound microscope with the 0.5X and 2X SHR Plan Apo objectives and the NIS Elements software. Confocal microscopy was conducted as previously described using and an agar slice containing one entire colony that was placed on a microscope slide and visualized without a coverslip(88). Confocal microscopy was conducted using the Nikon Ti2 microscope with 20x Plan Apo and 100X TU Plan Apo objective and the NIS Elements software. Images were rendered using the NIS Elements and ImageJ software.

RESULTS

RsmE directly binds to a limited number of mRNAs

The Rsm family of proteins are canonically known to function as post transcriptional regulators that bind mRNA to prevent their translation by the ribosome

(18, 25, 27). To determine whether RsmE binds directly to mRNA associated with extracellular secretions, we used a recombinant N-terminally his-tagged *rsmE* expressed in the Δ*rsmE* background for affinity purification of the RNA-Protein complex followed by sequencing (RAP-seq) similar to other recent studies in other pseudomonads(24). Non-denaturing purification of the His_6-RsmE-RNA complex through cobalt affinity chromatography was able to remove nearly all contaminants except for several high molecular weight proteins (Fig. 1A). Due to the partially pure nature of the sample, the contaminating proteins from Δ*rsmE* without His_6-RsmE was separately purified as a control (Fig. 1B).

When compared to the WT transcriptome, 788 transcripts were identified in the control and 816 in the His_6-RsmE sample with 604 transcripts common to both samples. Comparing the His_6-RsmE sample to the control, only four genes were deemed to be specifically enriched in the His6-RsmE sample (Table 1). The four hypothetical genes were *Pfl01_1873*, encoding a putative gene regulator, *Pfl01_2511*, a putative transferase*, Pfl01_4832*, a putative signaling protein or transferase, and *Pfl01_5645,* an uncharacterized membrane protein. The large quantity of background transcripts suggests that our control contained other RNA binding proteins that also copurify with our protein.

With the caveat that the candidate list may not be comprehensive due to the possibility of shared affinity for discrete RNAs between RsmE and the copurified contaminates, these results suggest that RsmE does not bind directly to mRNA that produce extracellular secretions, but rather that the identified putative proteins or other secondary regulators could bridge RsmE's influence on the observed production of extracellular secretions.

Many secretion genes are specifically upregulated in Δ*rsmE*

Given our observation that RsmE could directly binds to only few mRNAs and the potential for hierarchical regulation of RsmE, we sought to identify genes that are specifically upregulated in Δ*rsmE* compared to Δ*rsmA*, Δ*rsmI*, or WT through RNA-seq. We extracted RNA from colonies after three days of incubation, where the mucoid colony phenotype of Δ*rsmE* is clearly visible, but no significant emergence of *de novo* mucoid variants (naturally occurring *rsmE* mutants) in Δ*rsmA*, Δ*rsmI*, or WT colonies is detected(88). The list of differentially regulated genes are summarized in Table 2 and visually presented in Fig. 2 and Appendix II. None of the genes whose mRNA was found to bind directly to RsmE were differentially regulated in any of the samples, indicating that their transcription is independent of RsmE or the paralogs. As expected, we observed the absence of the *rsmE* transcript and the upregulation of the biosurfactant biosynthesis gene (*Pfl01_2211*) only in Δ*rsmE*. There are nearly 20 genes that are predicted to be involved in the biosynthesis or modification of the mucoid polysaccharide(65). However, none of these genes were upregulated in Δ*rsmE*, but among them a gene predicted to encode a UDP-glucose 4-epimerase (*Pfl01_3834*), which converts UDP-galactose (UDP-Gal) to UDP-glucose (UDP-Glc) in *E. coli*, was downregulated(65). This observation was unexpected, since knocking out this particular gene in Δ*rsmE* abolishes its mucoid phenotype(65). Nevertheless, none of the other biosynthesis genes in this pathway were either up- or downregulated in Δ*rsmE*, which suggests that the associated mucoid phenotype may be caused by chemical modifications or intra-membrane transport rather than the overproduction of the polysaccharide.

Comparing the global transcription profiles of Δ*rsmE*, Δ*rsmA*, and Δ*rsmI* against the WT reveals that greatest differential expression occurs in Δ*rsmE* (Figs. 2 and Appendix II). As expected, the expression of each *rsm* gene was not detected in their respective knockout background. Overall, Δ*rsmE* displayed the highest number of genes with altered expression at 48 (31 upregulated and 17 downregulated), followed by Δ*rsmA* at 10 (3 upregulated and 7 downregulated), and Δ*rsmI* at 7 (none upregulated and 7 downregulated). All downregulated genes in Δ*rsmA* and Δ*rsmI* were associated with the 5S ribosomal RNA, while they were entirely absent in the Δ*rsmE* dataset. Instead, downregulated genes in Δ*rsmE* consisted of those encoding nine enzymes including *Pfl01_3834*, three structural proteins, one MSF transporter, one lipoprotein, and four hypothetical proteins. Among the three upregulated genes in Δ*rsmA*, two are annotated to encode type VI secretion system (T6SS) effectors, and the remaining one is annotated to encode a T6SS contractile sheath small subunit. T6SS is best described as a bacterial weapon utilized in ecological warfare, comprising a spear-like apparatus loaded with diverse degradative effector proteins (e.g. proteases, nucleases, and lipases) which fires from the cytoplasm of the attacker to physically stab and deliver the effectors into the nearby competitors(95, 96). Unsurprisingly, producers of T6SS also produce appropriate neutralizing anti-effector molecules for protection. The same set of the three T6SS genes were also upregulated in Δ*rsmE*, in addition to 28 additional genes encoding 9 T6SS structural or effector proteins, 2 transcriptional regulators, 7 enzymes including the biosynthesis of the biosurfactant, 5 structural proteins, and 5 hypothetical proteins. Many of the genes that are uniquely upregulated in *rsmE* are predicted to encode extracellular secretions, some of which are likely to contribute to spatial structure formation.

Knocking out RsmE's function dramatically increases extracellular proteins

Although we have identified a group of upregulated genes in Δ*rsmE*, there is an obvious functional gap toward identifying those that contribute to spatial structure formation. We thus sought to identify extracellular proteins that are specifically produced in Δ*rsmE* by comparing the extracellular proteomes of Δ*rsmE* and WT through LC-MS/MS. WT generated a total of 36,559 MS spectra reads mapping to 191 proteins in the genome and Δ*rsmE* produced 36,138 MS spectra reads mapping to 310 proteins (Appendix IV). All 191 proteins found in WT were also found in Δ*rsmE*, along with 119 unique proteins (Fig. 3).

In both the WT and Δ*rsmE* samples, elongation factor Tu was the most abundant protein identified followed by outer membrane protein W, flagellin and several signaling proteins. The function of elongation factor Tu is best known as a component of the translational machinery in the cytoplasm, but its moonlighting function in the extracellular space as a mediator of diverse matrix interactions is increasingly acknowledged(97). Among the proteins unique to Δ*rsmE*, six overlap with the upregulated genes in our RNA-seq dataset for Δ*rsmE* (Table 2). One of these overlapping targets is the second most abundant protein in Δ*rsmE*, Pfl01_2678, which is annotated as the protease epralysin. The remaining overlapping targets include 3 T6SS proteins, 2 metallopeptidases, and 1 triacylglycerol lipase. An additional overlapping target is Pfl01_2270 (annotated as a dioxygenase), where only one spectral hit was found in WT while Δ*rsmE* produced 11 spectra hits across 5 unique peptides. None of the identified

proteins match the down regulated genes in Δ*rsmE* or those associated with the RsmE-bound mRNA.

T6SS kills WT cells at the genotypic boundary and those that infiltrate Δ*rsmE*'s spatial structure

The overlap across our transcriptomic and proteomic datasets effectively provides a short list of RsmE-associated candidates (Table 2) for functional analysis via engineering knockouts in Δ*rsmE*. Due to the significant proportion of the candidates being associated with the T6SS, we selected two of the respective structural genes, *Pfl01_5574* and *Pfl01_5594*. We also selected two genes associated with abundantly produced enzymes unique to Δ*rsmE*: *Pfl01_2678*, encoding a metallopeptidase, and *Pfl01_2685*, encoding a putative triacylglycerol lipase. Each engineered knockout isolate produced the same colony morphology as Δ*rsmE* (Fig.5). To assess the role of the T6SS in spatial structure formation, unlabeled mutants were seeded in low relative frequency with unlabeled WT in colonies in the presence of propidium iodide. Emerging mucoid patches were imaged by confocal microscopy after 2 days, since mucoid patches formed by *de novo rsmE* mutants typically emerge after 3 days. As indicated by the fluorescent signal in the red channel caused by the cellular accumulation of the propidium iodide, we observed evidence of cell death in the patches formed by Δ*rsmE* and the engineered isolates except for the T6SS mutants, with a striking pattern of T6SS-induced cell death as an outer ring surrounding each patch (Fig. 6)(98). When patches are formed by GFP-labeled Δ*rsmE* surrounded by unlabeled WT cells, this ring of death occurs immediately proximal to the space occupied by the green-fluorescent Δ*rsmE* (Fig. 5). This observation

leads to the interpretation that the WT cells are actively killed by Δ*rsmE* at the genotypic boundary through the T6SS.

A clear difference in the spatial structure formed by *rsmE* mutant without polysaccharide and biosurfactant production (M^S*) compared to that of Δ*rsmE* is the absence of cell-free space, wherein the M^S* cells are tightly packed(88). Coincidentally, cell death was not visible within the patches formed by M^S* in contrast to Δ*rsmE* (Fig. 6). These observations suggest that the dead cells within the patches likely represent the WT cells that had infiltrated the low cell-density space created by Δ*rsmE*. When patches are formed by unlabeled Δ*rsmE* surrounded by GFP-labeled WT cells, the encroachment of green-fluorescent WT cells into Δ*rsmE*'s space is clearly visible when the T6SS is knocked out, which also coincides with the absence of death (Appendix III). The functional roles of the metalloprotease and the putative triacylglycerol lipase in spatial structure formation remains unclear, if any.

DISCUSSION

Our three global profiling approaches collectively reveal that RsmE in *P. fluorescens* Pf0-1 likely exerts its influence on the production of extracellular secretions through secondary regulators. Transcripts that correspond to only four unique genes were found to directly bind RsmE, which contrasts greatly to the outcome of a recent study in *P. putida* KT2440, which showed that RsmE binds to the transcripts of 241 genes(24). The *P. putida* study utilized a similar affinity purification technique using his-RsmE, but the recombinant protein was expressed in WT and RsmE-binding was determined

through >2.15-fold enrichment compared to the total RNA in WT without the recombinant protein. In contrast, we utilized our his-RsmE in Δ*rsmE* to pull down transcripts that were entirely absent in Δ*rsmE* without the recombinant construct. We found hundreds of transcripts that bound the affinity resin in the absence of the his-RsmE construct. Possibly, some of the transcripts that were bound to the resin could have also bound to his-RsmE, which would have been filtered out as false-positives under our stringency rules. We are currently processing additional samples to increase statistical power, which could identify additional targets that bind RsmE. Nevertheless, our results suggest that RsmE does not bind to mRNA that are directly linked to the production of extracellular secretions. The four RsmE-binding targets we discovered correspond to hypothetical proteins with putative functions, including a regulator and a component of signal transduction, that could function to bridge RsmE's regulatory influence on the observed production of extracellular secretions.

The canonical mechanism by which Rsm proteins function is through competitive inhibition of the ribosome's ability to translate Rsm-binding mRNA (25). Under this model, one would expect to observe changes primarily in the proteome in the absence of RsmE, but changes in the transcriptome should also be expected if RsmE alters the translation of other regulatory systems. We thus employed an RNA-seq approach, which revealed numerous upregulated secretion genes unique to Δ*rsmE*, including the biosurfactant. The expression of these secretion genes is likely regulated directly or indirectly through RsmE-associated secondary regulators. A clear outcome, however, is that altered expression of secretion genes are specific to Δ*rsmE*, and not to Δ*rsmA* or Δ*rsmI*. None of the four genes whose transcripts were found to directly bind

RsmE was differentially regulated, suggesting that RsmE's sequestration of their transcripts could be primarily responsible for the observed shift in the global transcription profile. Moreover, we observed dramatic changes to the extracellular proteome in the absence of RsmE, which provides further support to the functional role of RsmE as an important modulator of extracellular secretions.

We had previously demonstrated that RsmE-associated polysaccharide and biosurfactant function together to create and maintain structured space of low cell density in a crowded colony(88). The polysaccharide is solely responsible for creating the space and the biosurfactant appears to physically sequester other extracellular secretions to help physically maintain the genotypic boundary. In this study, we have demonstrated that the production of a T6SS is associated with RsmE, which kills neighboring and encroaching WT cells to protect the spatiogenetic structure. The T6SS is a well-known bacterial weapon in ecological warfare and RsmA of *P. aeruginosa* (lacks RsmE) has been demonstrated to regulate its production(8). The tip of the T6SS, which physically punctures through the membrane of a nearby cell, is loaded with diverse degradative effector proteins that function to kill the competitor(95, 99). The producer of a T6SS also produces immunity proteins that neutralize the activity of each effector protein(95, 99). Therefore, a Δ*rsmE* cell is expected to be immune to the attack from another Δ*rsmE* cell, but a WT cell should succumb to Δ*rsmE*'s attack as we had observed. There is increasing evidence that a T6SS attack could be also neutralized through non-specific means, including the production of protective extracellular polysaccharides to cause steric hinderance or relying on the activation of envelope stress pathways as a general protective measure(99). Such immunity-independent mechanisms could be uniquely

employed by Δ*rsmE*, particularly through multiple extracellular secretions that are lacking in WT.

This study provides new insight on the regulatory function of RsmE, which is clearly distinct from that of its paralogs. Understanding the molecular mechanisms underlying the functional specificity of the paralogs, that harbor extremely little sequence or secondary structure variation, remains a question of significance. Spatial structure formation in *P. fluorescens* appears to be regulated exclusively through RsmE via the production of extracellular polysaccharide, biosurfactant, T6SS, and other secreted products. Our analysis of RsmE-specific metalloprotease and triglycerol lipase suggests a potential role in limiting the infiltration of WT cells into the spatial structure. It remains to be determined whether these two enzymes act as effectors of the T6SS or function independently. The T6SS has been shown to play an important role in self vs non-self recognition in other species(95). Additional RsmE-specific extracellular proteins identified in this study represent a rich resource for exploring novel mechanisms of social interaction, characterizing the hierarchical regulation surrounding RsmE, and understanding the evolution of functional specificity among the paralogs.

Figures

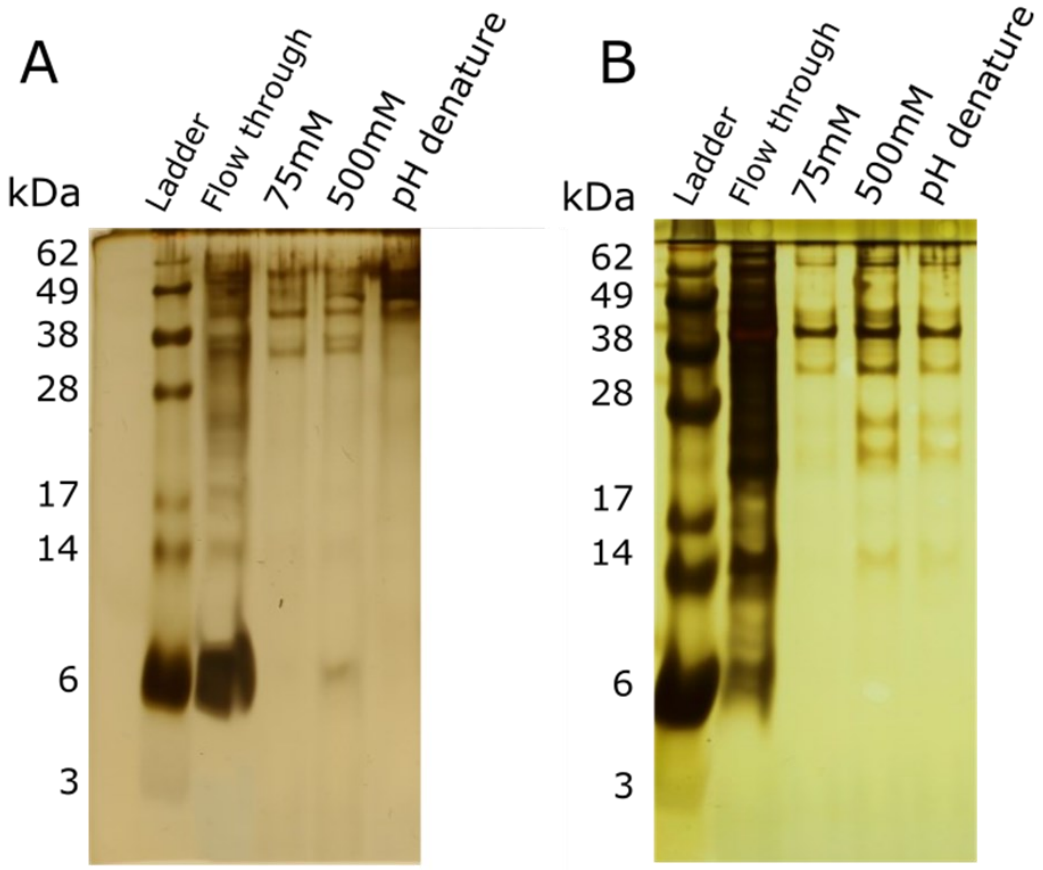


Fig. 1. Purification of his-RsmE. Silver stained SDS-PAGE gel of purification of his-RsmE from Δ*rsmE* (A) and purification of non-specific proteins from Δ*rsmE* background (B). The 500mM imidazole faction from both samples was used to isolate RNA due to the presence of ~7 kDa band corresponding to the size of a RsmE monomer.

Feature Id	Function	GO annotation	Phyre annotation
Pfl01_1873	hypothetical protein	#N/A	gene regulation
Pfl01_2511	hypothetical protein	#N/A	transferase
Pfl01_4832	hypothetical protein	calcium ion binding	signaling protein/transferase
Pfl01_5645	EamA family transporter	membrane	membrane protein

Table 1. Genes whose transcripts were found to directly bind RsmE[a]

[a] RNA was isolated and sequenced from affinity-purified his-tagged RsmE expressed in Δ*rsmE*, and passed through the selective filter (log_2-fold change of 1 and FDR of 0.05) when compared to Δ*rsmE* with no vector.

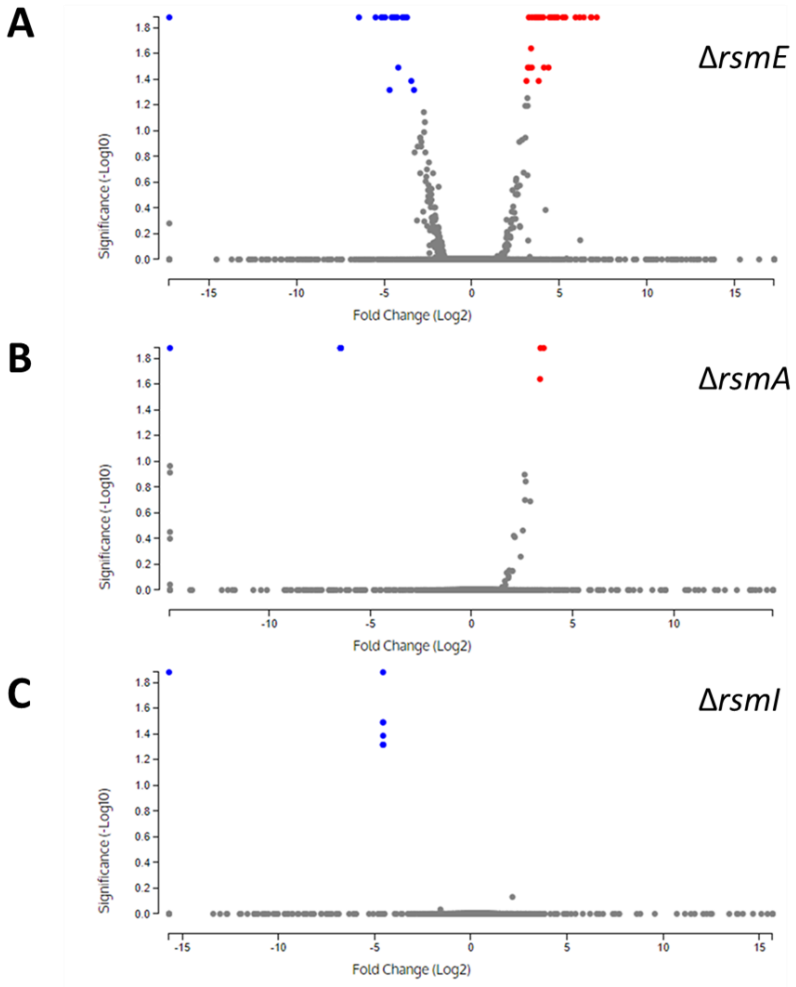


Fig. 2. Among the paralogs, Δ*rsmE* has the greatest influence on differential gene regulation relative to WT. Volcano plots comparing the relative expression of genes between each paralog knockout to WT. Genes with a fold change of 1-Log_2 and FDR of 0.05 (shown as Log_{10}) are labeled in blue for downregulation and red for upregulation, and genes with non-significant changes are labeled in grey. (A) 31 genes are upregulated and 17 are downregulated in Δ*rsmE*. The Δ*rsmA* strain (B) 3 genes are upregulated and 7 are downregulated in Δ*rsmA*. (C) 7 genes are downregulated and none are upregulated in Δ*rsmI*.

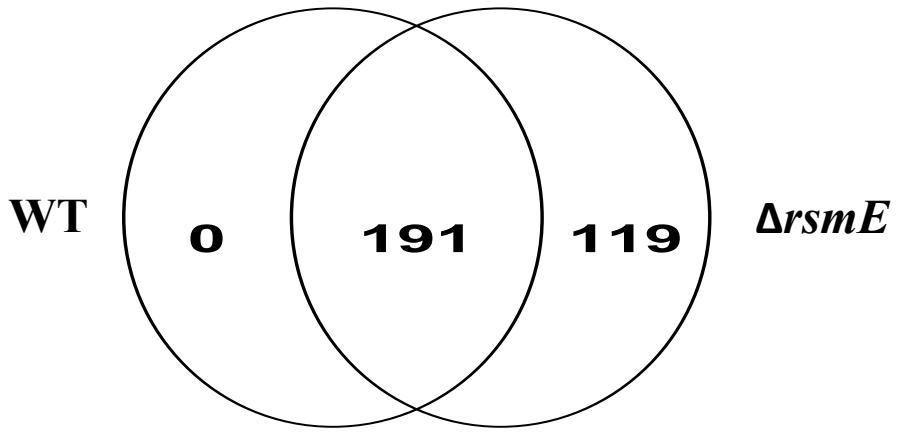


Fig. 3. Absence of RsmE dramatically alters the extracellular proteome. A Venn diagram showing that among the 310 extracellular proteins identified, 191 are common to both WT and Δ*rsmE*, while 119 are unique to Δ*rsmE*.

Upregulated Genes		Downregulated Genes	
ΔrsmE		***ΔrsmE***	
Feature Id	Function	Feature Id	Function
Pfl01_0125	SRPBCC family protein	Pfl01_1316	MFS transporter
Pfl01_1746	hypothetical protein	Pfl01_1912	carbon storage regulator CsrA
Pfl01_2038	hypothetical protein	Pfl01_0292	hypothetical protein
Pfl01_2045[+]*	Hcp family type VI secretion system effector	Pfl01_RS30095	hypothetical protein
Pfl01_2210	helix-turn-helix transcriptional regulator	Pfl01_3827	winged helix-turn-helix domain-containing protein
Pfl01_2211^	non-ribosomal peptide synthetase	Pfl01_3829	undecaprenyl-phosphate glucose phosphotransferase
Pfl01_2214	macrolide transporter subunit MacA	Pfl01_3832	YjbF family lipoprotein
Pfl01_2215	MacB family efflux pump subunit	Pfl01_3833	polysaccharide biosynthesis/export family protein
Pfl01_2270[+]	TauD/TfdA family dioxygenase	Pfl01_RS19235	hypothetical protein
Pfl01_2328*	Hcp family type VI secretion system effector	Pfl01_3834^	UDP-glucose 4-epimerase GalE
Pfl01_2333	hypothetical protein	Pfl01_RS19250	mannose-1-phosphate guanylyltransferase
Pfl01_2427	hypothetical protein	Pfl01_3837	lysophospholipase
Pfl01_2626	sel1 repeat family protein	Pfl01_3838	acyltransferase
Pfl01_2627	sel1 repeat family protein	Pfl01_3841	phosphoribosylaminoimidazole carboxylase
Pfl01_2678[+§]	M10 family metallopeptidase	Pfl01_3843	glycosyltransferase
Pfl01_2680	type I secretion system permease/ATPase	Pfl01_3844	polysaccharide biosynthesis tyrosine autokinase
Pfl01_2682	TolC family outer membrane protein	Pfl01_3845	low molecular weight phosphotyrosine protein phosphatase
Pfl01_2684	autotransporter serine protease	Pfl01_3846	hypothetical protein
Pfl01_2685[+§]	triacylglycerol lipase		
Pfl01_3190	ParA family protein	***ΔrsmA***	
Pfl01_3363[+]	M10 family metallopeptidase	Feature Id	Function
Pfl01_4819	pilin	Pfl01_R0016	5S ribosomal RNA
Pfl01_5573	hypothetical protein	Pfl01_R0005	5S ribosomal RNA
Pfl01_5574[+§]	type VI secretion system tip protein VgrG	Pfl01_4273	carbon storage regulator CsrA
Pfl01_5578	DotU family type IV/VI secretion system protein	Pfl01_R0045	5S ribosomal RNA
Pfl01_5580	type VI secretion system lipoprotein TssJ	Pfl01_R0006	5S ribosomal RNA
Pfl01_5581	type VI secretion system-associated FHA domain protein TagH	Pfl01_R0056	5S ribosomal RNA
Pfl01_5583	sigma-54-dependent Fis family transcriptional regulator	Pfl01_R0077	5S ribosomal RNA
Pfl01_5584	type VI secretion system ATPase TssH	Pfl01_R0090	5S ribosomal RNA
Pfl01_5594[+§]	type VI secretion system contractile sheath large subunit		
Pfl01_5595*	type VI secretion system contractile sheath small subunit	***ΔrsmI***	
		Feature Id	Function
ΔrsmA		Pfl01_R0016	5S ribosomal RNA
Feature Id	Function	Pfl01_R0005	5S ribosomal RNA
Pfl01_2045*	Hcp family type VI secretion system effector	Pfl01_4104	carbon storage regulator CsrA
Pfl01_2328*	Hcp family type VI secretion system effector	Pfl01_R0045	5S ribosomal RNA
Pfl01_5595*	type VI secretion system contractile sheath small subunit	Pfl01_R0006	5S ribosomal RNA
		Pfl01_R0056	5S ribosomal RNA
ΔrsmI		Pfl01_R0077	5S ribosomal RNA
Feature Id	Function	Pfl01_R0090	5S ribosomal RNA
N/A	N/A		

Table 2. Differently expressed genes and selected targets of RsmE regulation. Compared to WT, all differentially expressed genes in Δ*rsmE,* Δ*rsmA,* and Δ*rsmI* are shown with upregulated genes on the left and down regulated genes on the left. Knockouts were constructed for genes matching secreted proteins that were exclusively upregulated in Δ*rsmE.*

[+] Genes for secreted proteins identified by LC-MS/MS and upregulated in Δ*rsmE*

*****Genes upregulated in Δ*rsmE* and Δ*rsmA*

^Genes identified by transposon mutagenesis

§Genes selected for knockout construction

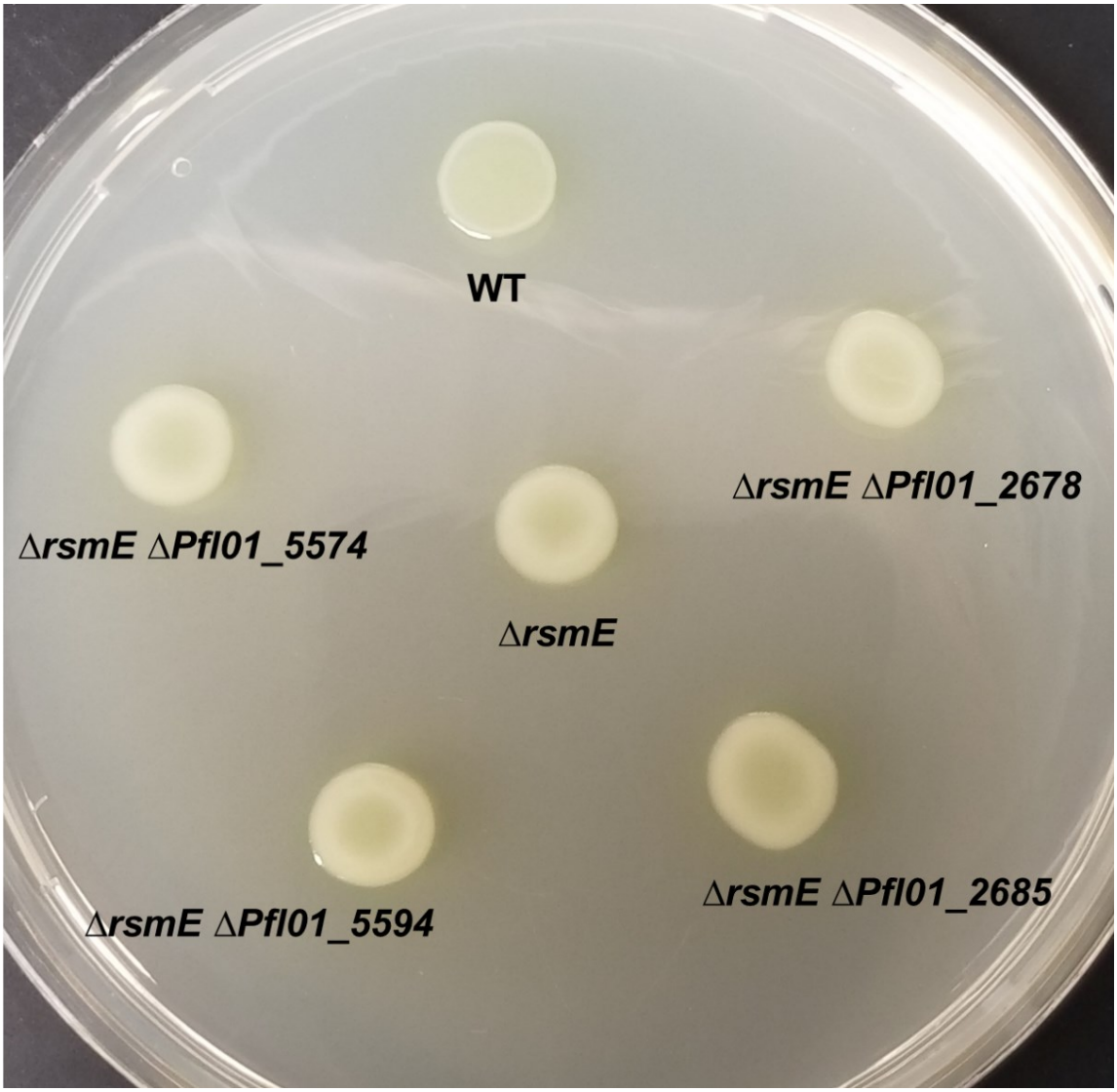


Fig 4. The knockouts of selected targets of RsmE regulation show no change in phenotype. Day 2 Phenotypic screen, displays the typical mucoid phenotype of Δ*rsmE* compared to WT. The knockouts in the Δ*rsmE* background displays no change in phenotype.

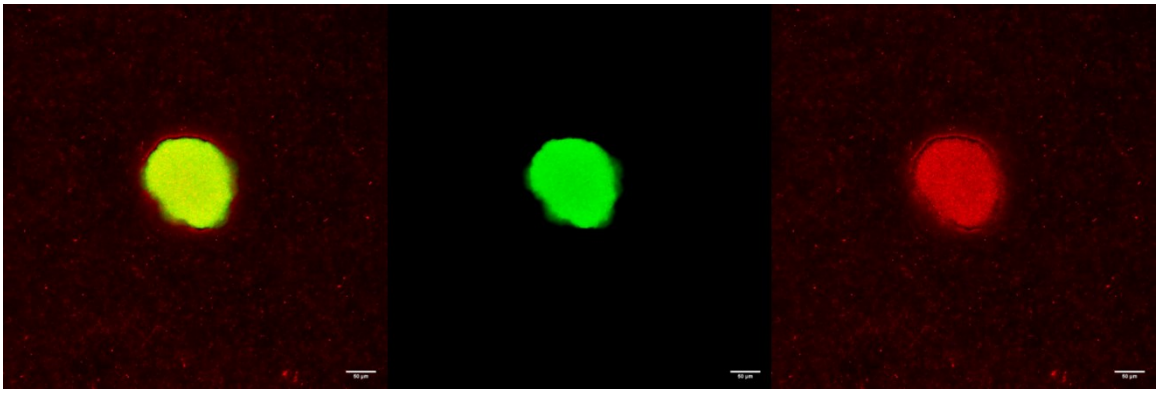


Fig 5. T6SS kills WT cells at the genotypic boundary. GFP-labeled Δ*rsmE* was mixed with unlabeled WT at a 10^{-5} relative frequency, and incubated for 2 days with propidium iodide supplemented to each initial inoculum. Confocal images of the emergent patches were generated using the 20X PLAN APO objective, shown with the composite image (left), green channel (middle), red channel (right). Dead cells (red) can be seen sporadically spread out throughout the WT population away from the green-fluorescent Δ*rsmE* patch. In contrast, a higher concentration of dead WT cells is seen as an outer ring situated outside the genotypic boundary, with a distinct gap separating the dead cells from the border.

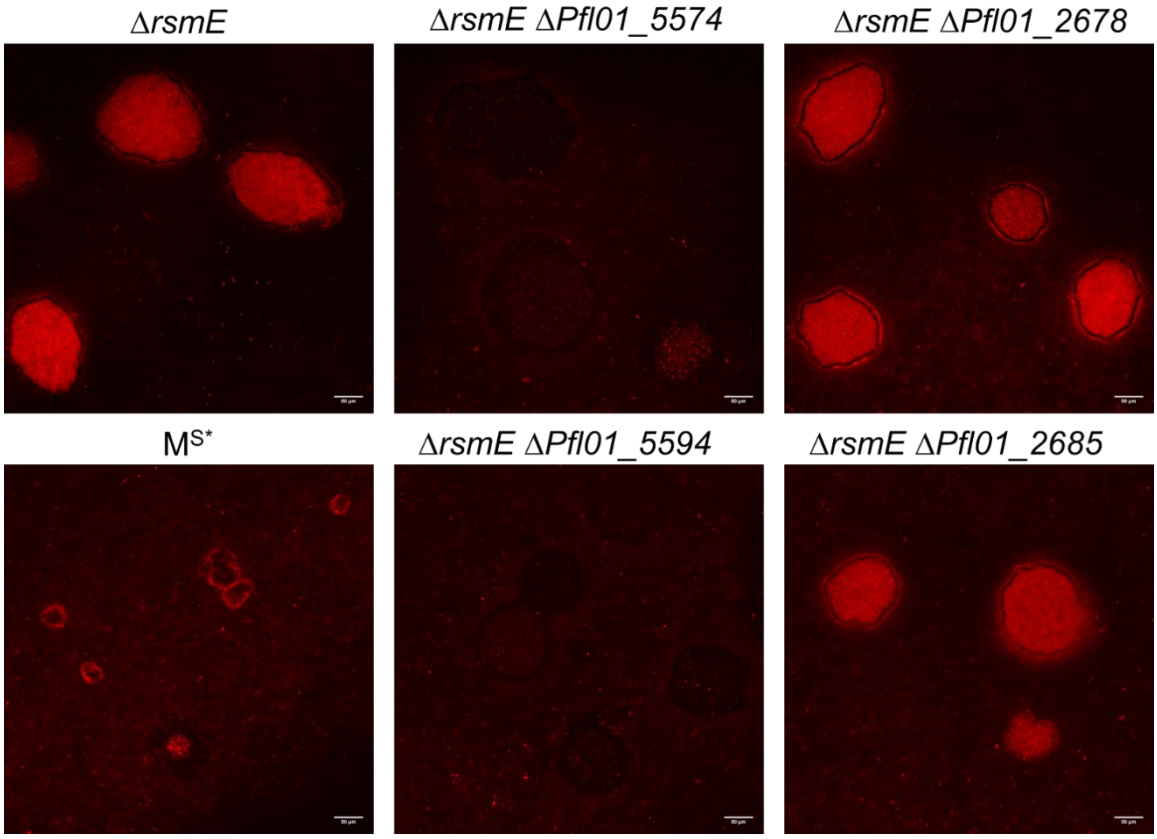


Fig 6. T6SS kills cells at the genotypic boundary and within Δ*rsmE*'s spatial structure. Unlabeled mutants, as noted above each panel, were mixed with unlabeled WT at a 10^{-5} relative frequency and incubated for 2 days with propidium iodide supplemented to each initial inoculum. Confocal images of the emergent patches were generated using the 20X PLAN APO objective. Dead cells (red) are seen in high abundance throughout the patches formed by Δ*rsmE,* Δ*rsmE* Δ*Pfl01_2678* (metalloprotease), and Δ*rsmE* Δ*Pfl01_2685* (triglycerol lipase), and at the respective boundary against WT, with a noticeable gap between the interior dead cells and the exterior border. Dead cells are not observed within the patch formed by M^S*, but they are visible at the boundary against

WT. In contrast, dead cells are rarely associated with the patches formed by Δ*rsmE* Δ*Pfl01_5574 and* Δ*rsmE* Δ*Pfl01_5594*, that lack the T6SS. Scale bar represents 50µm.

Strain	Genotype	Relevant Phenotype	Species	Source
Pf0-1	Wildtype (WT)	WT	*P. fluorescens*	(35)
Pf0-1 GFP	WT (Tn7-Gfpmut2)	Expresses GFP	*P. fluorescens*	(1)
Pf0-1 RFP	WT (Tn7-dsRedExpress)	Expresses RFP	*P. fluorescens*	(1)
M	WT (frameshift mutation in *rsmE*)	Mucoid	*P. fluorescens*	(1)
M GFP	M (Tn7-Gfpmut2)	Expresses GFP	*P. fluorescens*	(1)
M RFP	M (Tn7-dsRedExpress)	Expresses RFP	*P. fluorescens*	(1)
M^S*	M (Δ*Pfl01_2211* Δ*Pfl01_3834*)	No polysaccharide or surfactant	*P. fluorescens*	(88)
M^S* GFP	M^S* (Tn7-Gfpmut2)	Expresses GFP	*P. fluorescens*	(88)
M^S* RFP	M^S* (Tn7-dsRedExpress)	Expresses RFP	*P. fluorescens*	(88)
Δ*rsmE*	WT (Δ*Pfl01_1912*)	Mucoid	*P. fluorescens*	(1)
Δ*rsmE* GFP	Δ*rsmE* (Tn7-Gfpmut2)	Expresses GFP	*P. fluorescens*	(1)
Δ*rsmE*/his-RsmE	Δ*rsmE* (pME6000-his-*rsmE*)	Similar to WT	*P. fluorescens*	This study
Δ*rsmA*	WT (Δ*Pfl01_4273*)	Similar to WT	*P. fluorescens*	(88)
Δ*rsmI*	WT (Δ*Pfl01_4104*)	Similar to WT	*P. fluorescens*	(88)
Δ*rsmE* Δ*Pfl01_2678*	Δ*rsmE* (Δ*Pfl01_2678*)	Mucoid	*P. fluorescens*	This study
Δ*rsmE* Δ*Pfl01_2685*	Δ*rsmE* (Δ*Pfl01_2685*)	Mucoid	*P. fluorescens*	This study
Δ*rsmE* Δ*Pfl01_5574*	Δ*rsmE* (Δ*Pfl01_5574*)	Mucoid	*P. fluorescens*	This study
Δ*rsmE* Δ*Pfl01_5594*	Δ*rsmE* (Δ*Pfl01_5594*)	Mucoid	*P. fluorescens*	This study

Table 3. Pseudomonas Strains used in this study

Name	Function	Sequence
Δ2678-1	Outside F	GAAATGATGGGCAAGCTGTTC
Δ2678-2	Outside R	CCGACTTCAAAGTGCTCAGATA
Δ2678-3	Up F	CGATGCCGCTGAACATCGACC
Δ2678-4	Up R	CGCCGCTTCCCGTGTTACAAAACACACTTCCTTGTTTAGC
Δ2678-5	Down F	GCTAAACAAGGAAGTGTGTTTTGTAACACGGGAAGCGGCG
Δ2678-6	Down R	CTGGATTCTTGAGCCTTCCTTGCCGG
Δ2685-1	Outside F	GAGCCGTTCGCCAATCT
Δ2685-2	Outside R	GCCCTGAGCGTCGTATTT
Δ2685-3	Up F	TTCACCGAGAAGGGTGATG
Δ2685-4	Up R	GAGCATTTCAGTACAACTTTCGTTTGGTTGTGCTCTCTCTTCC
Δ2685-5	Down F	GGAAGAGAGAGCACAACCAAACGAAAGTTGTACTGAAATGCTC
Δ2685-6	Down R	CCTGAGCGCTGGTGTAG
Δ5574-1	Outside F	GCTGCACGGTGACATGAA
Δ5574-2	Outside R	GGGCACAGTCATTAAATTGCTC
Δ5574-3	Up F	GAGCAACGTCATGCTCAG
Δ5574-4	Up R	GAATGTAGCAACGGGTGCGCAAAGTGCTCCCTGTTCGAATG
Δ5574-5	Down F	CATTCGAACAGGGAGCACTTTGCGCACCCGTTGCTACATTC
Δ5574-6	Down R	ATAGAACCAATGCCAGACAAA
Δ5594-1	Outside F	CGGTGCTCAGGAAGAGATTG
Δ5594-2	Outside R	GTCGGATACGCGTGCTG
Δ5594-3	Up F	TGGCAATCGGTGACTACAC
Δ5594-4	Up R	CCGTATCCGTCCATGAGTCGCTCTGTGTTGGCTTCCTTGGC
Δ5594-5	Down F	GCCAAGGAAGCCAACACAGAGCGACTCATGGACGGATACGG
Δ5594-6	Down R	GGGTTCTCCTTAGTTGACCTTC
rsmE N-his F	N-terminal His Tag	ATGCATCATCACCACCACCATCTGATACTCACCCGCAAAG
rsmE N-his R	N-terminal His Tag	TCAGGGCGTTTGTGGCTTG

Table 4. Primers used in this study

Chapter 4: Conclusions

Summary

Bacteria living within dense microbial communities are subject to significant evolutionary pressure. This pressure manifests as competition to maintain access to the resources needed to grow and produce progeny. In this dissertation, I have utilized *Pseudomonas fluorescens* Pf0-1 and its propensity to solve this space problem through mutations in RsmE and as a model to unveil the underlying regulatory network. Previous studies have shown that these *rsmE* mutants uniquely produce extracellular secretions that likely contribute to increased fitness(1, 65). I then characterized two of these extracellular secretions, a biosurfactant and an extracellular polymer, by competition analyses. Competitions between the WT and single and double knockouts of the secretions showed that each secretion discretely and collectively contribute to the increased fitness observed in the *rsmE* mutant strain, each single knockout reduced fitness and the double knockout further reduced fitness.

Epifluorescence and confocal microscopy revealed different contributions of each secretion to spatial structuring. In patches formed by *rsmE* mutants, cells are able to push up and outward to reach the colony surface by creating space and excluding the parental cells from the newly created space. The knockout of the biosurfactant show patches that don't exclude the parental cell once they reach the colony surface, while the knockout of the polymer excludes the parental cells but do not make space resulting in small densely packed patches. The patches formed by the double mutant are similar in phenotype to the polymer knockout, but much smaller in scale. Overall, the mucoid polymer creates space

and the biosurfactant prevents its diffusion, but even the double knockout still had increased fitness over the WT strain, suggesting additional contributing mechanisms.

To identify these additional contributing factors, I explored RsmE's direct role in gene regulation since mutations disrupting the function of RsmE are central to the spatial structure formation. The two major RsmE-secretions had been previously identified by transposon mutagenesis; however further use of this method would be problematic for identifying non-visible secretions. To overcome this challenge, I employed a systematic approach to characterizing RsmE's impact on gene regulation: RIP-seq, RNA-seq, and proteomics. Using recombinant his-tagged RsmE, I successfully purified the RNA directly bound to the protein. Sequencing the copurified RNA revealed no direct connection to any of the previously identified genes, but rather that RsmE likely regulates other potential regulators. I had applied stringent filtering parameters to rule out false-positives, which likely also removed true-positives. While this discovery likely places RsmE higher up in the hierarchy of gene regulation, I decided to apply a more direct approach to identify the key secretions that contribute to spatial structure formation beyond the mucoid polymer and biosurfactant.

Based on the premise that RsmE likely regulates other gene regulators, I compared the transcriptome between WT and Δ*rsmE* and observed 48 differently regulated genes in the knockout. With two additional Rsm paralogs expressed in the WT, and no obvious phenotype in their knockouts, I explored whether the other Rsm paralogs differentially altered gene expression. I independently sequenced the transcriptomes of Δ*rsmA* and Δ*rsmI* and compared them to WT and Δ*rsmE*. Δ*rsmA* and Δ*rsmI* had a much smaller set of altered gene expression compared to WT (10 and 7 respectively). Three

T6SS genes were found upregulated in Δ*rsmA* but these were also found upregulated in Δ*rsmE*. All downregulated genes were annotated as 5S r-RNA shared between Δ*rsmA* and Δ*rsmI*.

To further narrow down the list of secretion targets I then focused on identifying secreted proteins through LC-MS/MS. Between WT and Δ*rsmE*, I found 191 proteins common to both isolates and 119 unique to Δ*rsmE.* Comparing the secreted proteins unique to Δ*rsmE* to the upregulated genes from RNA-seq, I narrowed down the list to 7 secretion candidates and chose 4 to engineer knockouts for further study. These 4 genes consisted of 2 putative extracellular enzymes and 2 structural components of the T6SS. With no obvious colony phenotype, I used a propidium iodine assay to study cellular death associated with the T6SS. I found that Δ*rsmE* does indeed kill the parental strain, and observed a striking pattern where dead cells are concentrated at the genotypic boundaries of the expanding spatial patch. This observation was consistent in all of the engineered strains except the ones with the T6SS components knocked out. This work opens up a multitude of new questions about the nature of Rsm regulation and the evolution of spatial structure formation via RsmE.

T6SS in future studies

With T6SS being identified as an additional mechanism of spatial competition through RsmE, several questions pertaining to this new mechanism still need to be answered. First, does the T6SS contribute to the increased fitness of Δ*rsmE* against the WT, and how does it function alongside other identified extracellular secretions? My propidium iodide studies in the M and M^{s*} strain show that cellular death is occurring at the margins of the expanding patch. However, there is a noticeable difference between

the two, with the M^{s*} strain appearing to produce almost all of the dead cells at the borders of the patch with little to none within, while with the M strain, dead cells are also observed throughout the patch. This difference could suggest that some parent cells do infiltrate the newly created space but they are rapidly killed off through the T6SS. An alternative explanation could be that the mucoid polymer and/or biosurfactant and/or additional secretions play some role in self versus non-self recognition as individual patches of M* or M^{s*} within close proximity appear to not merge together while those of M or M^{s} tend to merge.

To address these questions, competition analyses of T6SS, mucoid polymer, the biosurfactant knockouts could determine to which degree each component is responsible for the increased fitness and disentangle their functional interplay, as the fitness contribution of the mucoid polymer and biosurfactant are discrete from each other. Furthermore, I have established an extensive set of additional extracellular secretions that are uniquely produced by Δ*rsmE*. Additionally, a full metabolomic profile of internal and external proteins, lipids, and small molecules comparing WT to Δ*rsmE* would provide even a richer dataset to mine for such additional undiscovered factors.

The molecular function of Rsm likely extends beyond the canonical mechanism

I have shown conclusively that *P. fluorescens* Pf0-1 Rsm paralogs are not functionally redundant as mutations in *rsmE* alone are responsible the formation of competitive spatial structure(88). Furthermore, I have shown that all three paralogs are expressed with transcripts for rsmA being most abundant followed by rsmE and rsmI with all the paralogs having transcripts at biologically relevant levels. The knockouts of each Rsm paralog produce different transcription profiles, with RsmE being responsible

for the most altered profiles compared to WT. The canonical mechanism of Rsm proteins regulating gene expression by stabilizing a pentaloop structure in the Shine Dalgarno sequence of mRNA is a simple and convenient mechanism to explain turning translation on and off. But digging further into this mechanism, the proteins' activity is likely conferred through interactions with the peptide backbone and not the side chains, with specificity resulting from the overall 3D shape. Such a model suggests that changes in ligand specificity likely require major structural changes to the protein rather than a single amino acid substitution. This hypothesis is problematic in our system because all three paralogs have very similar amino acid sequences and are almost identical within the Csr/Rsm domain containing the residues known to interface with RNA, containing only conservative substitutions. CHIP-seq studies where researchers pulled down Rsm and Csr proteins mostly revealed an overlapping consensus Shine Dalgarno sequence(8, 16, 26, 28, 100, 101). Such an observation led to the generalization that Rsm paralogs are functionally redundant(26, 32, 33). However, multiple other Rsm-RNA interaction have been postulated, ranging from gene activation to protection from nuclease activity resulting in mRNA with a longer half-life(29, 102, 103).

The secondary structure of most Rsm homologs is 4 beta sheets and an alpha helix making up the Csr domain(24, 25, 88). This core Csr domain is conserved across over 2900 species, and involved with both protein dimerization and RNA recognition(19, 25). Additionally, there is a C-terminal tail region post alpha helix that is highly variable in sequence length and structure, and this is where the major differences between our paralogs occur. We observed such patterns while exploring the evolutionary history of Rsm proteins. We built a customized dataset of over 700 Rsm/CsrA homologs from the

RCSB Protein Data Bank and PSI BlastP based searches using each Rsm from the Pf0-1 strain and used the motif-based sequence analysis tools (MEME) suite to identify unique amino acid motifs (Appendix I). This unpublished work showed that the highly variable tail section contained unique but conserved motifs where each Rsm paralog fit a different motif. We were then able to roughly categorize the entire data set into *rsmA*-like, *rsmE*-like and *rsmI*-like bins based on these motifs. Additional research is currently being carried out in the lab to directly assess the molecular role of the C-terminus in paralogs with respect to functional specificity.

Future studies pertaining to Rsm Paralogs

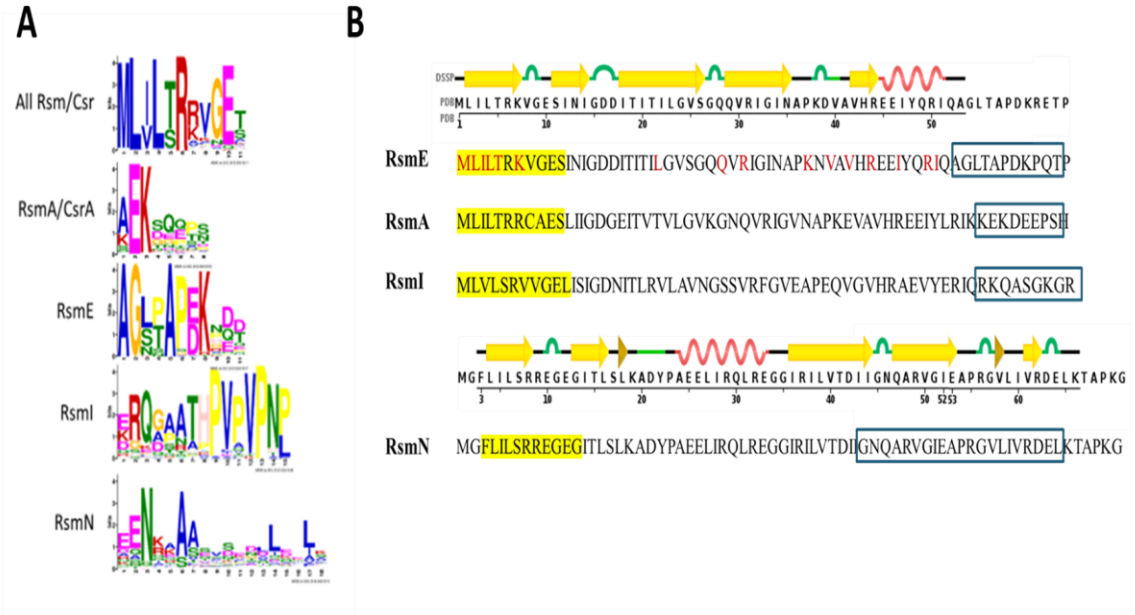
As we were working towards a manuscript of our computational work and attempting to improve statistical analyses, an article entitled *Comparative Genomics and Evolutionary Analysis of RNA-Binding Proteins of the CsrA Family in the Genus Pseudomonas* was published by another group(19). This publication used a much larger dataset containing every CsrA entry in the Pfam and InterPro database to build an evolutionary tree using the neighbor-joining method, resulting in extensive categorization of the Rsm subfamilies. While it was disappointing to be beaten to print, their results support our initial finding as outside of structural rearrangements, such as in RsmN and RsmF, the a C-terminal tail region determined the subcategory. However, the tails still vary wildly while looking at just amino acid sequence, but they seem to share cluster of similar properties such as charge. If this is the case, the tails may play a critical role in determining RNA specification, not necessarily through binding specification, but rather through subcellular localization.

The internal organization of bacterial cells, specifically the cytosol, is an emerging area of study. Long thought to have limited structure due to the lack of membrane bound organelles, recent studies have shown that biomolecular condensates exist with these cells and may function to increase enzymatic kinetics within the cells(104). Recent studies in *Caulobacter crescentus* and *Bacillus subtilis* have described structures coined Bacterial RNP-bodies (BR-bodies)(104, 105). These BR-bodies are the first bacterial biomolecular condensates described that assemble through liquid-liquid phase separation, in which RNaseE forms a scaffold with RNA and other associated proteins(104, 105). These BR-bodies are most analogous to stress granules and P-bodies in eukaryotic cells, in that they concentrate translationally repressed mRNA and exclude ribosomes(104).

In our system, it is possible that the tails determine subcellular localization, with some paralog sequestered in different regions of the cell. If BR-bodies or some analogous structures indeed exist with in *P. fluorescens* Pf0-1, the Rsm tail could mediate the localization inside or outside BR-bodies. It is also possible that with the Rsm dimer structure containing two binding sites, they may serve as linker proteins in subcellular scaffolding similar to RNaseE. This may be possible with RIP-seq studies showing Rsm interacting with essentially every mRNA in the cell. Future studies can potentially attempt to address this question of tail determined function by first looking at subcellular locations of each paralog, potentially with fluorescent tags on the protein. Then it can be coupled with differential staining of nucleic acids or other polymers to see if any lattice network or any biomolecular condensate associations are found.

Overall, our thoroughly characterized model in which the dense and crowed conditions of a *P. fluorescens* colonies selects for mutations in *rsmE* is observed on rich media. To elucidate the full roll of each Rsm paralog it may be necessary to examine the *P. fluorescens* undergrowth conditions replicating its native ecological niche. Previous studies have shown that pseudomonads behavior with regards to root colonization(106), biofilm formation(107), and expression of virulence factors can differ based on the particular carbon source available(5). It could be that the regulation network of each paralog is only biologically relevant for a particular carbon source and we wouldn't see any differential regulation unless these metabolic pathways are activated.

Appendix I: Discovery of amino acid motifs in Rsm Orthologs.

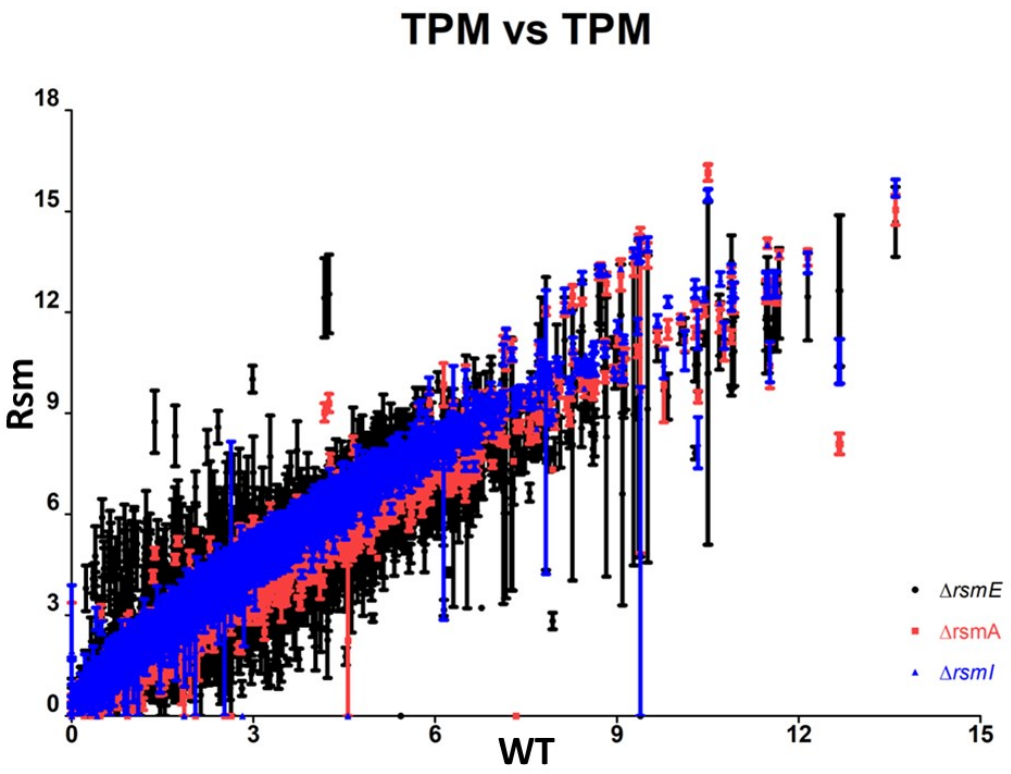


The amino acid sequences of CsrA from *E. coli* K12, RsmE, RsmA, and RsmI from *P. fluorescens* Pf0-1, and RsmN from *P. aeruginosa* PAO1 was used to query the Refseq database(106) using PSI Blastp and the RCSB Protein Data bank (RCSB.org)(107) using mmseq2. The resulting hits were combined, and duplicates were removed, resulting in a dataset of 761 sequences.

Amino acid motifs for this dataset were generated using the Multiple Em for Motif Elicitation (MEME)(108) tool using classic motif discovery mode, zero or one occurrence for each motif per sequence, with the maximum of 7 discoverable motifs. Consensus sequence logos of the resulting amino acid motifs are shown (A). 742 of the 761 sequences contained the first motif located in the N-terminal of the peptide sequences and was subsequently labeled all Rsm/Csr. Each of the query sequences contained different C-terminal motifs except RsmA and CsrA shared the same motif. Within our dataset 149 sequences contained the RsmA/CsrA motif, 53 with the RsmE

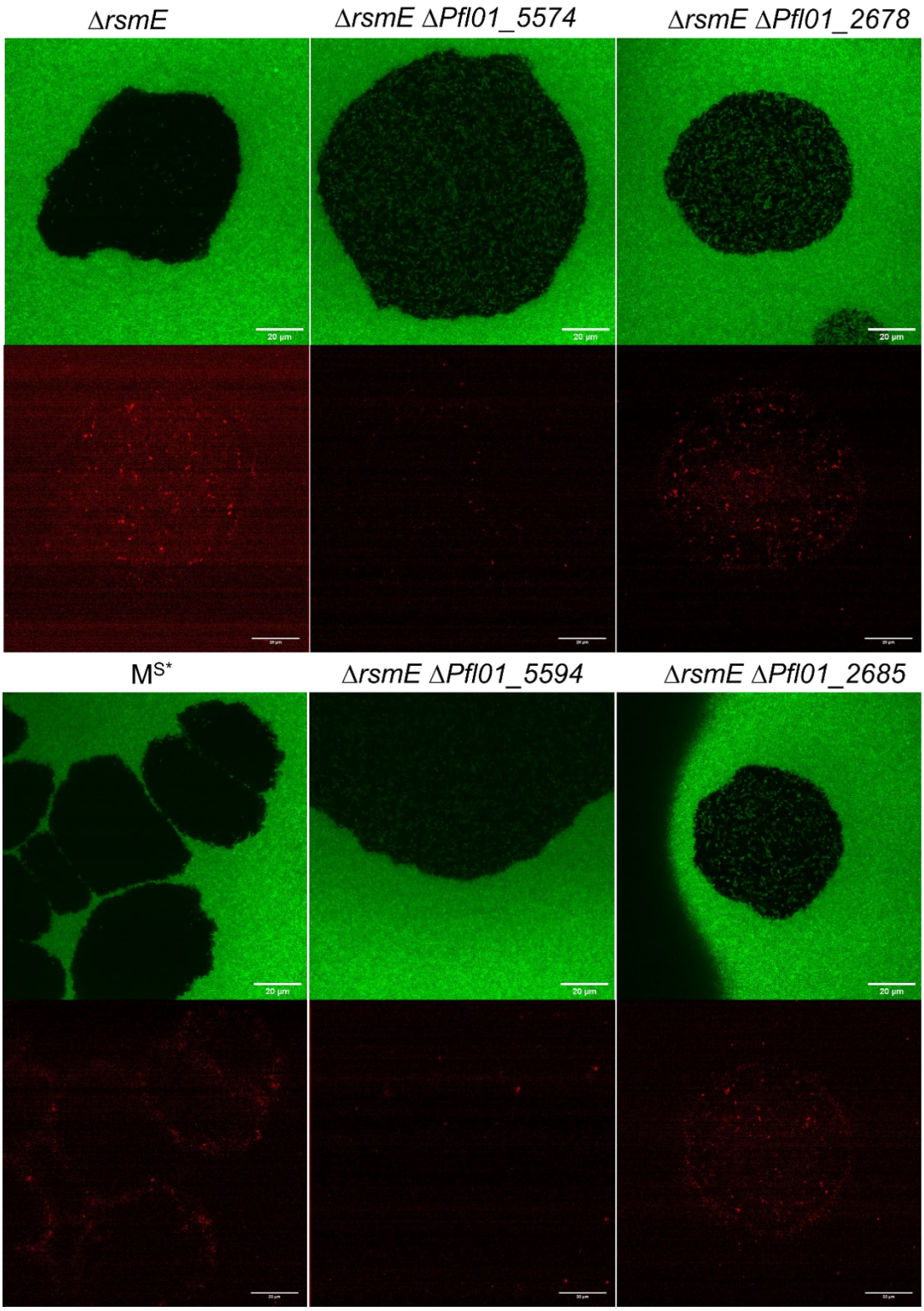
motif, 36 with the RsmI motif, and 87 with the RsmN motif. The location of the resulting motifs is shown (B) for the three Rsm paralogs found in *P. fluorescens* Pf0-1 and the structurally rearranged RsmN from *P. aeruginosa* along with their corresponding secondary structures. The location of the motif common to all Rsm proteins is highlighted in Yellow in and a box is drawn around the location of the motif unique to each subclass.

Appendix II: TPM comparison to WT visually captures that most differentially regulated genes are associated with Δ*rsmE*.



Transcripts Per Kilobase Million (TPM) was calculated for every Pf0-1 gene for each RNA-seq sample using Cufflinks V2.2.1. The scatter plot was generated using the average TPM from the WT sample as the *x*-coordinate, and the average TPM from each knockout as the *y*-coordinate for each gene with the standard error shown. The dataset for Δ*rsmE* is shown in black, Δ*rsmA* in red, and Δ*rsmI* in blue.

Appendix III: The absence of the T6SS, metalloprotease, or triglycerol lipase results in increased encroachment of WT cells into Δ*rsmE*'s patch.



Unlabeled mutants, as noted above each panel, were mixed with GFP-labeled WT at a 10^{-5} relative frequency, and incubated for 2 days with propidium iodide supplemented to each initial inoculum. Confocal images of the emergent patches were generated using the 100X TU PLAN APO objective. Dead cells (red) are seen throughout the patches formed by Δ*rsmE,* Δ*rsmE* Δ*Pfl01_2678* (metalloprotease)*,* and Δ*rsmE* Δ*Pfl01_2685* (triglycerol lipase), and at the respective boundary against WT. Dead cells are not observed within the patch formed by M^S*, but they are visible at the boundary against WT. In contrast, dead cells are rarely associated with the patches formed by Δ*rsmE* Δ*Pfl01_5574 and* Δ*rsmE* Δ*Pfl01_5594*, that lack the T6SS. Increased presence of green WT cells are visible within the patches formed by the engineered mutants.

Appendix IV: Proteins Identified by LC-MS/MS

#	Identified Protein	Accession Number	Molecular Weight	Protein Identification Probability	
				ΔrsmE	WT
1	ref|YP_350809.1|gi|77461302| elongation factor Tu	Pfl01_5081 (+1)	43 kDa	100%	100%
2	ref|YP_348409.1|gi|77458903| cpralysin	Pfl01_2678	50 kDa	100%	0
3	ref|YP_350810.1|gi|77461303| elongation factor G	Pfl01_5082	77 kDa	100%	100%
4	ref|YP_350229.1|gi|77460722| chaperonin GroEL	Pfl01_4501	57 kDa	100%	100%
5	ref|YP_351456.1|gi|77461949| FOF1 ATP synthase subunit alpha	Pfl01_5732	55 kDa	100%	100%
6	ref|YP_347348.1|gi|77457843| dihydrolipoamide dehydrogenase	Pfl01_1616	50 kDa	100%	38%
7	ref|YP_347577.1|gi|77458072| peptide synthase	Pfl01_1845	622 kDa	100%	72%
8	ref|YP_347349.1|gi|77457844| succinyl CoA synthetase subunit beta	Pfl01_1617	41 kDa	100%	100%
9	Cluster of ref|YP_350605.1|gi|77461098| serine hydroxymethyltransferase (Pfl01_4877)	Pfl01_4877 (2)	45 kDa	100%	100%
9.1	ref|YP_350605.1|gi|77461098| serine hydroxymethyltransferase	Pfl01_4877	45 kDa	100%	100%
9.2	ref|YP_350936.1|gi|77461429| serine hydroxymethyltransferase	Pfl01_5208	45 kDa	8%	0
10	ref|YP_351454.1|gi|77461947| FOF1 ATP synthase subunit beta	Pfl01_5730	49 kDa	100%	8%
11	ref|YP_346196.1|gi|77456691| pyruvate dehydrogenase subunit E1	Pfl01_0463	100 kDa	100%	77%
12	ref|YP_346136.1|gi|77456631| malic enzyme	Pfl01_0403	48 kDa	100%	98%
13	ref|YP_347238.1|gi|77457733| flagellar hook associated protein FlgK	Pfl01_1506	70 kDa	100%	65%
14	ref|YP_346877.1|gi|77457372| Phage tail sheath protein FI like	Pfl01_1145	42 kDa	100%	37%
15	ref|YP_349322.1|gi|77459815| isocitrate dehydrogenase (NADP)	Pfl01_3593	46 kDa	100%	85%
16	ref|YP_350216.1|gi|77460709| hypothetical protein	Pfl01_4488	47 kDa	100%	59%
17	ref|YP_346875.1|gi|77457370| Phage tail Collar	Pfl01_1143	86 kDa	100%	89%
18	ref|YP_350120.1|gi|77460613| glycine dehydrogenase	Pfl01_4392	103 kDa	100%	97%
19	ref|YP_346853.1|gi|77457348| phosphopyruvate hydratase	Pfl01_1121	46 kDa	100%	93%
20	ref|YP_350129.1|gi|77460622| OmpA/MotB	Pfl01_4401	18 kDa	100%	49%
21	ref|YP_350263.1|gi|77460756| extracellular solute binding protein	Pfl01_4535	33 kDa	100%	92%
22	ref|YP_347579.1|gi|77458074| amino acid adenylation protein	Pfl01_1847	116 kDa	100%	10%
23	ref|YP_347259.1|gi|77457754| flagellin	Pfl01_1527	29 kDa	100%	100%
24	ref|YP_350249.1|gi|77460742| alkyl hydroperoxide reductase	Pfl01_4521	22 kDa	100%	95%
25	ref|YP_347649.1|gi|77458144| L-asparaginase, type II	Pfl01_1917	39 kDa	100%	0
26	ref|YP_350115.1|gi|77460608| ornithine carbamoyltransferase	Pfl01_4387	38 kDa	100%	78%
27	ref|YP_346495.1|gi|77456990| molecular chaperone DnaK	Pfl01_0763	68 kDa	100%	100%
28	ref|YP_350813.1|gi|77461306| DNA directed RNA polymerase subunit beta'	Pfl01_5085	155 kDa	100%	92%
29	ref|YP_347502.1|gi|77457997| phosphoenolpyruvate synthase	Pfl01_1770	86 kDa	100%	74%
30	ref|YP_350114.1|gi|77460607| arginine deiminase	Pfl01_4386	46 kDa	100%	41%
31	ref|YP_351133.1|gi|77461626| extracellular solute binding protein	Pfl01_5405	40 kDa	100%	71%
32	ref|YP_350260.1|gi|77460753| glycerol kinase	Pfl01_4532	56 kDa	100%	44%
33	ref|YP_350811.1|gi|77461304| 30S ribosomal protein S7	Pfl01_5083	18 kDa	100%	98%
34	ref|YP_346581.3|gi|255961280| fumarate hydratase	Pfl01_0849	48 kDa	100%	0
35	ref|YP_347796.1|gi|77458291| lysine arginine ornithine binding periplasmic protein	Pfl01_2064	28 kDa	100%	0
36	ref|YP_349668.1|gi|77460161| peptide synthase	Pfl01_3940	481 kDa	100%	81%
37	ref|YP_347229.1|gi|77457724| aromatic amino acid aminotransferase	Pfl01_1497	43 kDa	100%	18%
38	ref|YP_350814.1|gi|77461307| DNA directed RNA polymerase subunit beta	Pfl01_5086	151 kDa	100%	84%
39	ref|YP_347509.1|gi|77458004| OmpF	Pfl01_1777	37 kDa	100%	90%
40	ref|YP_346886.1|gi|77457381| bacteriophage Mu tail sheath	Pfl01_1154	53 kDa	100%	16%
41	ref|YP_347580.1|gi|77458075| TonB dependent siderophore receptor	Pfl01_1848	90 kDa	100%	100%
42	ref|YP_346515.1|gi|77457010| polynucleotide phosphorylase	Pfl01_0783	75 kDa	100%	27%
43	ref|YP_347578.1|gi|77458073| peptide synthase	Pfl01_1846	498 kDa	100%	64%
44	ref|YP_350807.1|gi|77461300| 50S ribosomal protein L3	Pfl01_5079	23 kDa	100%	99%
45	ref|YP_350816.1|gi|77461309| 50S ribosomal protein L10	Pfl01_5088	18 kDa	100%	98%
46	ref|YP_351008.1|gi|77461501| S-adenosyl-L-homocysteine hydrolase	Pfl01_5280	51 kDa	100%	47%
47	ref|YP_350273.1|gi|77460766| 5-methyltetrahydropteroyltriglutamate--homocysteine S-methyltransferase	Pfl01_4545	89 kDa	100%	0
48	ref|YP_347350.1|gi|77457845| succinyl CoA synthetase subunit alpha	Pfl01_1618	30 kDa	100%	99%
49	ref|YP_345977.1|gi|77456472| cystine transporter subunit	Pfl01_0244	29 kDa	100%	89%
50	ref|YP_346279.1|gi|77456774| azurin	Pfl01_0546	16 kDa	100%	83%
51	ref|YP_349092.1|gi|77459585| cpralysin	Pfl01_3363	49 kDa	100%	0
52	ref|YP_349845.1|gi|77460338| electron transfer flavoprotein subunit beta	Pfl01_4117	26 kDa	100%	62%
53	ref|YP_351353.1|gi|77461846| aspartate ammonia lyase	Pfl01_5625	51 kDa	100%	99%
54	ref|YP_347909.1|gi|77458404| enoyl ACP reductase	Pfl01_2177	28 kDa	100%	91%
55	ref|YP_346802.1|gi|77457297| adenylate kinase	Pfl01_1070	23 kDa	100%	69%
56	ref|YP_345921.1|gi|77456416| 4-aminobutyrate aminotransferase	Pfl01_0188	45 kDa	100%	53%
57	ref|YP_350795.1|gi|77461288| 50S ribosomal protein L5	Pfl01_5067	20 kDa	100%	94%
58	ref|YP_348294.1|gi|77458788| peptidase S45, penicillin amidase	Pfl01_2563	86 kDa	100%	0
59	ref|YP_350333.1|gi|77460826| nucleoside diphosphate kinase	Pfl01_4605	15 kDa	100%	100%
60	ref|YP_348599.1|gi|77459093| thiolase	Pfl01_2868	41 kDa	100%	0

#	Identified Protein	Accession Number	Molecular Weight	Protein Identification Probability	
				ΔrsmE	WT
61	ref|YP_345956.1|gi|77456451| NLPA lipoprotein	Pfl01_0223	29 kDa	100%	99%
62	ref|YP_350782.1|gi|77461275| 50S ribosomal protein L17	Pfl01_5054	14 kDa	100%	75%
63	ref|YP_351185.1|gi|77461678| thioredoxin	Pfl01_5457	12 kDa	100%	58%
64	ref|YP_346583.1|gi|77457078| manganese and iron superoxide dismutase	Pfl01_0851	23 kDa	100%	0
65	ref|YP_347806.1|gi|77458301| acetyl CoA acetyltransferase	Pfl01_2074	41 kDa	100%	0
66	ref|YP_350800.1|gi|77461293| 50S ribosomal protein L16	Pfl01_5072	15 kDa	100%	0
67	ref|YP_351366.1|gi|77461859| pyruvate carboxylase subunit B	Pfl01_5639	65 kDa	100%	84%
68	ref|YP_346896.1|gi|77457391| hypothetical protein	Pfl01_1164	40 kDa	100%	51%
69	ref|YP_346261.1|gi|77456756| adenylosuccinate synthetase	Pfl01_0528	47 kDa	100%	67%
70	ref|YP_350107.1|gi|77460600| TonB dependent hemoglobin/transferrin/lactoferrin receptor	Pfl01_4379	94 kDa	100%	99%
71	ref|YP_350791.1|gi|77461284| 50S ribosomal protein L18	Pfl01_5063	13 kDa	100%	96%
72	ref|YP_346545.1|gi|77457040| extracellular solute binding protein	Pfl01_0813	59 kDa	100%	71%
73	ref|YP_350993.1|gi|77461486| phosphoglycerate kinase	Pfl01_5265	40 kDa	100%	0
74	ref|YP_350797.1|gi|77461290| 50S ribosomal protein L14	Pfl01_5069	13 kDa	100%	87%
75	ref|YP_346069.1|gi|77456564| preprotein translocase subunit SecB	Pfl01_0336	18 kDa	100%	0
76	ref|YP_347666.1|gi|77458161| 50S ribosomal protein L20	Pfl01_1934	13 kDa	100%	80%
77	ref|YP_349790.1|gi|77460283| DegT/DnrJ/EryC1/StrS aminotransferase	Pfl01_4062	42 kDa	100%	0
78	ref|YP_347346.1|gi|77457841| 2-oxoglutarate dehydrogenase E1	Pfl01_1614	106 kDa	100%	97%
79	ref|YP_350783.1|gi|77461276| DNA directed RNA polymerase subunit alpha	Pfl01_5055	37 kDa	100%	91%
80	ref|YP_350122.1|gi|77460615| glycine cleavage system T protein	Pfl01_4394	40 kDa	100%	33%
81	ref|YP_347777.1|gi|77458272| hypothetical protein	Pfl01_2045	19 kDa	100%	16%
82	ref|YP_350176.1|gi|77460669| pyruvate kinase	Pfl01_4448	52 kDa	100%	90%
83	ref|YP_349124.3|gi|255961273| bifunctional aconitate hydratase 2/2-methylisocitrate dehydratase	Pfl01_3395	95 kDa	100%	61%
84	ref|YP_349368.1|gi|77459861| glutaminyl tRNA synthetase	Pfl01_3639	65 kDa	100%	47%
85	ref|YP_350514.1|gi|77461007| ketol-acid reductoisomerase	Pfl01_4786	36 kDa	100%	11%
86	ref|YP_349323.1|gi|77459816| isocitrate dehydrogenase (NADP+)	Pfl01_3594	80 kDa	100%	83%
87	ref|YP_346643.1|gi|77457138| fumarylacetoacetate hydrolase	Pfl01_0911	47 kDa	100%	11%
88	ref|YP_346834.1|gi|77457329| elongation factor Ts	Pfl01_1102	30 kDa	100%	100%
89	ref|YP_350147.1|gi|77460640| outer membrane porin	Pfl01_4419	48 kDa	100%	86%
90	ref|YP_350142.1|gi|77460635| Ferritin and Dps	Pfl01_4414	18 kDa	100%	42%
91	ref|YP_347625.1|gi|77458120| aspartate semialdehyde dehydrogenase	Pfl01_1893	41 kDa	100%	0
92	ref|YP_349727.1|gi|77460220| elongation factor P	Pfl01_3999	21 kDa	100%	89%
93	ref|YP_350803.1|gi|77461296| 30S ribosomal protein S19	Pfl01_5075	10 kDa	100%	85%
94	ref|YP_350123.1|gi|77460616| cold shock DNA binding protein	Pfl01_4395	8 kDa	100%	89%
95	ref|YP_346346.1|gi|77456841| bifunctional phosphoribosylaminoimidazolecarboxamide formyltransferase	Pfl01_0613	58 kDa	100%	60%
96	ref|YP_346329.1|gi|77456824| TonB dependent copper receptor	Pfl01_0596	77 kDa	100%	57%
97	ref|YP_350139.1|gi|77460632| aspartyl tRNA synthetase	Pfl01_4411	66 kDa	100%	52%
98	ref|YP_349575.1|gi|77460068| transaldolase B	Pfl01_3847	34 kDa	100%	11%
99	ref|YP_346001.1|gi|77456496| phosphoenolpyruvate carboxykinase	Pfl01_0268	55 kDa	100%	0
100	ref|YP_346066.1|gi|77456561| phosphoglyceromutase	Pfl01_0333	55 kDa	100%	0
101	ref|YP_346351.1|gi|77456846| acetyl CoA carboxylase biotin carboxylase subunit	Pfl01_0618	49 kDa	100%	0
102	ref|YP_349607.1|gi|77460100| multifunctional fatty acid oxidation complex subunit alpha	Pfl01_3879	77 kDa	100%	0
103	ref|YP_347661.1|gi|77458156| cold shock DNA binding protein family protein	Pfl01_1929	8 kDa	100%	96%
104	ref|YP_349639.1|gi|77460132| peptidase M24	Pfl01_3911	66 kDa	100%	15%
105	ref|YP_348362.1|gi|77458856| Rhs element Vgr protein	Pfl01_2631	76 kDa	100%	13%
106	ref|YP_350119.1|gi|77460612| glycine cleavage system protein H	Pfl01_4391	14 kDa	100%	7%
107	ref|YP_346980.1|gi|77457475| extracellular ligand binding receptor	Pfl01_1248	40 kDa	100%	0
108	ref|YP_350423.1|gi|77460916| aldo/keto reductase	Pfl01_4695	39 kDa	100%	0
109	ref|YP_350588.1|gi|77461081| 50S ribosomal protein L21	Pfl01_4860	12 kDa	100%	22%
110	ref|YP_350580.1|gi|77461073| isoleucyl tRNA synthetase	Pfl01_4852	106 kDa	100%	17%
111	ref|YP_347357.1|gi|77457852| heat shock protein 90	Pfl01_1625	71 kDa	100%	0
112	ref|YP_350786.1|gi|77461279| 30S ribosomal protein S13	Pfl01_5058	13 kDa	100%	58%
113	ref|YP_349331.1|gi|77459824| isocitrate lyase	Pfl01_3602	49 kDa	100%	14%
114	ref|YP_347145.1|gi|77457640| cysteine synthase	Pfl01_1413	34 kDa	100%	0
115	ref|YP_350089.1|gi|77460582| keto-hydroxyglutarate-aldolase/keto-deoxy-phosphogluconate aldolase	Pfl01_4361	23 kDa	100%	0
116	ref|YP_347261.1|gi|77457756| flagellar hook-associated 2 like	Pfl01_1529	49 kDa	100%	100%
117	ref|YP_346873.1|gi|77457368| baseplate J-like protein	Pfl01_1141	36 kDa	100%	35%
118	ref|YP_347808.1|gi|77458303| 3-oxoacid CoA-transferase	Pfl01_2076	25 kDa	100%	7%
119	ref|YP_350315.1|gi|77460808| GMP synthase	Pfl01_4587	58 kDa	100%	0
120	ref|YP_348656.1|gi|77459150| alkyl hydroperoxide reductase	Pfl01_2925	20 kDa	100%	0

#	Identified Protein	Accession Number	Molecular Weight	Protein Identification Probability	
				ΔrsmE	WT
121	ref|YP_350784.1|gi|77461277| 30S ribosomal protein S4	Pfl01_5056	23 kDa	100%	90%
122	ref|YP_351262.1|gi|77461755| aldehyde dehydrogenase	Pfl01_5534	53 kDa	100%	64%
123	ref|YP_346898.1|gi|77457393| hypothetical protein	Pfl01_1166	38 kDa	100%	47%
124	ref|YP_346754.1|gi|77457249| 50S ribosomal protein L19	Pfl01_1022	13 kDa	100%	0
125	ref|YP_347043.1|gi|77457538| iron-containing alcohol dehydrogenase	Pfl01_1311	40 kDa	100%	0
126	ref|YP_347584.1|gi|77458079| twin-arginine translocation pathway signal protein	Pfl01_1852	48 kDa	100%	0
127	ref|YP_349918.1|gi|77460411| hypothetical protein	Pfl01_4190	63 kDa	100%	0
128	ref|YP_350042.1|gi|77460535| glycerophosphodiester phosphodiesterase	Pfl01_4314	41 kDa	100%	0
129	ref|YP_350817.1|gi|77461310| 50S ribosomal protein L1	Pfl01_5089	24 kDa	100%	98%
130	ref|YP_346225.1|gi|77456720| thiamine biosynthesis protein ThiC	Pfl01_0492	70 kDa	100%	75%
131	ref|YP_348401.1|gi|77458895| phosphoglucomutase	Pfl01_2670	59 kDa	100%	64%
132	ref|YP_351158.1|gi|77461651| extracellular solute-binding protein	Pfl01_5430	37 kDa	100%	33%
133	ref|YP_349802.1|gi|77460295| bifunctional cyclohexadienyl dehydrogenase	Pfl01_4074	78 kDa	100%	30%
134	ref|YP_351302.1|gi|77461795| Rhs element Vgr protein	Pfl01_5574	83 kDa	100%	0
135	ref|YP_345823.1|gi|77456318| oligopeptidase A	Pfl01_0090	76 kDa	100%	0
136	ref|YP_348061.1|gi|77458556| Rhs element Vgr protein	Pfl01_2329	105 kDa	100%	0
137	ref|YP_349655.1|gi|77460148| periplasmic solute binding protein	Pfl01_3927	34 kDa	100%	0
138	ref|YP_351239.1|gi|77461732| methyltransferase	Pfl01_5511	28 kDa	100%	0
139	ref|YP_350104.1|gi|77460597| glyceraldehyde-3-phosphate dehydrogenase	Pfl01_4376	36 kDa	100%	0
140	ref|YP_348597.1|gi|77459091| short chain enoyl-CoA hydratase	Pfl01_2866	28 kDa	100%	0
141	ref|YP_350806.1|gi|77461299| 50S ribosomal protein L4	Pfl01_5078	22 kDa	100%	0
142	ref|YP_346546.1|gi|77457041| outer membrane porin	Pfl01_0814	51 kDa	100%	97%
143	ref|YP_350995.1|gi|77461488| transketolase	Pfl01_5267	74 kDa	100%	63%
144	ref|YP_350421.1|gi|77460914| 30S ribosomal protein S9	Pfl01_4693	15 kDa	100%	10%
145	ref|YP_351280.1|gi|77461773| endoribonuclease L-PSP	Pfl01_5552	13 kDa	100%	7%
146	ref|YP_346832.1|gi|77457327| methionine aminopeptidase	Pfl01_1100	29 kDa	100%	0
147	ref|YP_350757.1|gi|77461250| OmpW	Pfl01_5029	25 kDa	100%	100%
148	ref|YP_346900.1|gi|77457395| tail fiber protein	Pfl01_1168	31 kDa	100%	98%
149	ref|YP_346267.1|gi|77456762| 30S ribosomal protein S6	Pfl01_0534	16 kDa	100%	97%
150	ref|YP_346290.1|gi|77456785| Short chain dehydrogenase/reductase SDR	Pfl01_0557	28 kDa	100%	77%
151	ref|YP_347103.1|gi|77457598| dihydrodipicolinate synthase	Pfl01_1371	31 kDa	100%	62%
152	ref|YP_350802.1|gi|77461295| 50S ribosomal protein L22	Pfl01_5074	12 kDa	100%	44%
153	ref|YP_350541.1|gi|77461034| glucose-6-phosphate isomerase	Pfl01_4813	61 kDa	100%	0
154	ref|YP_348600.1|gi|77459094| Short chain dehydrogenase/reductase SDR	Pfl01_2869	26 kDa	100%	0
155	ref|YP_350805.1|gi|77461298| 50S ribosomal protein L23	Pfl01_5077	11 kDa	100%	0
156	Cluster of ref|YP_350020.1|gi|77460513| lysine-arginine-ornithine-binding periplasmic protein (Pfl01_4292)	Pfl01_4292 (2)	28 kDa	100%	0
156.1	ref|YP_350020.1|gi|77460513| lysine-arginine-ornithine-binding periplasmic protein	Pfl01_4292	28 kDa	100%	0
156.2	ref|YP_346699.1|gi|77457194| lysine-arginine-ornithine-binding periplasmic protein	Pfl01_0967	28 kDa	0	0
157	ref|YP_349424.1|gi|77459917| Lon A peptidase	Pfl01_3695	89 kDa	100%	0
158	ref|YP_346894.1|gi|77457389| baseplate J-like protein	Pfl01_1162	37 kDa	100%	99%
159	ref|YP_348935.1|gi|77459428| glutamate dehydrogenase (NAD)	Pfl01_3206	183 kDa	100%	82%
160	ref|YP_350812.1|gi|77461305| 30S ribosomal protein S12	Pfl01_5084	14 kDa	100%	97%
161	ref|YP_350792.1|gi|77461285| 50S ribosomal protein L6	Pfl01_5064	19 kDa	100%	96%
162	ref|YP_350013.1|gi|77460506| bifunctional N-succinyldiaminopimelate aminotransferase	Pfl01_4285	43 kDa	100%	74%
163	ref|YP_348002.1|gi|77458497| taurine dioxygenase	Pfl01_2270	32 kDa	100%	67%
164	ref|YP_346132.1|gi|77456627| arginyl-tRNA synthetase	Pfl01_0399	64 kDa	100%	38%
165	ref|YP_346567.1|gi|77457062| aspartyl/glutamyl-tRNA amidotransferase subunit B	Pfl01_0835	53 kDa	100%	32%
166	ref|YP_349886.1|gi|77460379| ACP S-malonyltransferase	Pfl01_4158	32 kDa	100%	18%
167	ref|YP_347341.1|gi|77457836| type II citrate synthase	Pfl01_1609	48 kDa	100%	17%
168	ref|YP_346499.1|gi|77456994| carbamoyl phosphate synthase large subunit	Pfl01_0767	118 kDa	100%	8%
169	ref|YP_346077.1|gi|77456572| L-glutamine synthetase	Pfl01_0344	52 kDa	100%	0
170	ref|YP_347657.1|gi|77458152| inosine/uridine-preferring nucleoside hydrolase	Pfl01_1925	37 kDa	100%	0
171	ref|YP_349369.1|gi|77459862| cyclophilin type peptidyl-prolyl cis-trans isomerase	Pfl01_3640	18 kDa	100%	0
172	ref|YP_350830.1|gi|77461323| N-acetyl-gamma-glutamyl-phosphate reductase	Pfl01_5102	37 kDa	100%	0
173	ref|YP_350818.1|gi|77461311| 50S ribosomal protein L11	Pfl01_5090	15 kDa	100%	90%
174	ref|YP_346751.1|gi|77457246| 30S ribosomal protein S16	Pfl01_1019	9 kDa	100%	87%
175	ref|YP_350322.1|gi|77460815| 2-isopropylmalate synthase	Pfl01_4594	62 kDa	100%	82%
176	ref|YP_347651.1|gi|77458146| periplasmic binding protein/LacI transcriptional regulator	Pfl01_1919	33 kDa	100%	25%
177	ref|YP_347932.1|gi|77458427| aminopeptidase	Pfl01_2200	99 kDa	100%	22%
178	ref|YP_348648.1|gi|77459142| 4-hydroxyphenylpyruvate dioxygenase	Pfl01_2917	40 kDa	100%	20%
179	ref|YP_347344.1|gi|77457839| succinate dehydrogenase flavoprotein subunit	Pfl01_1612	64 kDa	100%	14%
180	ref|YP_349385.1|gi|77459878| isovaleryl-CoA dehydrogenase	Pfl01_3656	42 kDa	100%	11%

#	Identified Protein	Accession Number	Molecular Weight	Protein Identification Probability ΔrsmE	WT
181	ref|YP_347583.1|gi|77458078| hypothetical protein	Pfl01_1851	33 kDa	100%	8%
182	ref|YP_347624.1|gi|77458119| 3-isopropylmalate dehydrogenase	Pfl01_1892	39 kDa	100%	6%
183	ref|YP_347581.1|gi|77458076| amino acid adenylation protein	Pfl01_1849	154 kDa	100%	0
184	ref|YP_346345.1|gi|77456840| phosphoribosylamine-glycine ligase	Pfl01_0612	45 kDa	100%	0
185	ref|YP_346833.1|gi|77457328| 30S ribosomal protein S2	Pfl01_1101	27 kDa	100%	99%
186	ref|YP_351145.1|gi|77461638| outer membrane porin	Pfl01_5417	48 kDa	100%	99%
187	ref|YP_346746.1|gi|77457241| phosphoribosylglycinamide formyltransferase 2	Pfl01_1014	42 kDa	100%	66%
188	ref|YP_350685.1|gi|77461178| leucyl-tRNA synthetase	Pfl01_4957	97 kDa	100%	38%
189	Cluster of ref|YP_347944.1|gi|77458439| amino acid adenylation protein (Pfl01_2212)	Pfl01_2212 (2)	593 kDa	100%	36%
189.1	ref|YP_347944.1|gi|77458439| amino acid adenylation protein	Pfl01_2212	593 kDa	100%	0
189.2	ref|YP_347945.1|gi|77458440| amino acid adenylation protein	Pfl01_2213	535 kDa	100%	26%
190	ref|YP_350003.1|gi|77460496| alanyl-tRNA synthetase	Pfl01_4275	94 kDa	100%	22%
191	ref|YP_350841.1|gi|77461334| indole-3-glycerol phosphate synthase	Pfl01_5113	30 kDa	100%	16%
192	ref|YP_347014.1|gi|77457509| NADH:flavin oxidoreductase/NADH oxidase	Pfl01_1282	38 kDa	100%	7%
193	ref|YP_349820.1|gi|77460313| hypothetical protein	Pfl01_4092	19 kDa	100%	6%
194	ref|YP_351455.1|gi|77461948| FOF1 ATP synthase subunit gamma	Pfl01_5731	31 kDa	100%	0
195	ref|YP_349667.1|gi|77460160| alpha/beta hydrolase	Pfl01_3939	30 kDa	100%	0
196	ref|YP_348596.1|gi|77459090| acyl-CoA dehydrogenase	Pfl01_2865	41 kDa	100%	0
197	ref|YP_346892.1|gi|77457387| Phage baseplate assembly protein V	Pfl01_1160	18 kDa	100%	0
198	ref|YP_349773.1|gi|77460266| hexapaptide repeat-containing transferase	Pfl01_4045	23 kDa	100%	0
199	ref|YP_347231.1|gi|77457726| phenylalanine 4-monooxygenase	Pfl01_1499	30 kDa	100%	0
200	ref|YP_350777.1|gi|77461270| single-stranded DNA-binding protein	Pfl01_5049	19 kDa	100%	0
201	ref|YP_349739.1|gi|77460232| outer membrane lipoprotein OprI	Pfl01_4011	9 kDa	100%	100%
202	ref|YP_345785.1|gi|77456280| DSBA oxidoreductase	Pfl01_0052	23 kDa	100%	41%
203	ref|YP_346028.1|gi|77456523| outer membrane porin	Pfl01_0295	46 kDa	100%	41%
204	ref|YP_346085.1|gi|77456580| fructose-1,6-bisphosphatase	Pfl01_0352	37 kDa	100%	19%
205	ref|YP_348416.1|gi|77458910| triacylglycerol lipase	Pfl01_2685	59 kDa	100%	13%
206	ref|YP_349837.1|gi|77460330| hypothetical protein	Pfl01_4109	22 kDa	100%	8%
207	ref|YP_350108.1|gi|77460601| Heme oxygenase	Pfl01_4380	22 kDa	100%	0
208	ref|YP_347304.1|gi|77457799| ChcW protein	Pfl01_1572	18 kDa	100%	0
209	ref|YP_351349.1|gi|77461842| phosphoribosylaminoimidazole carboxylase ATPase subunit	Pfl01_5621	39 kDa	100%	0
210	ref|YP_347776.1|gi|77458271| Rhs element Vgr protein	Pfl01_2044	76 kDa	100%	0
211	ref|YP_347924.1|gi|77458419| NADPH dependent FMN reductase	Pfl01_2192	20 kDa	100%	0
212	ref|YP_351300.1|gi|77461793| hypothetical protein	Pfl01_5572	33 kDa	100%	0
213	ref|YP_349800.1|gi|77460293| 30S ribosomal protein S1	Pfl01_4072	62 kDa	100%	98%
214	ref|YP_349427.1|gi|77459920| trigger factor	Pfl01_3698	48 kDa	100%	82%
215	ref|YP_351012.1|gi|77461505| hypothetical protein	Pfl01_5284	22 kDa	100%	66%
216	ref|YP_350911.1|gi|77461404| malate synthase G	Pfl01_5183	79 kDa	100%	58%
217	ref|YP_349590.1|gi|77460083| soluble pyridine nucleotide transhydrogenase	Pfl01_3862	51 kDa	100%	38%
218	ref|YP_349799.1|gi|77460292| NAD dependent epimerase/dehydratase	Pfl01_4071	38 kDa	100%	18%
219	Cluster of ref|YP_350990.1|gi|77461483| fructose-1,6-bisphosphate aldolase (Pfl01_5262)	Pfl01_5262 (2)	38 kDa	100%	10%
219.1	ref|YP_350990.1|gi|77461483| fructose-1,6-bisphosphate aldolase	Pfl01_5262	38 kDa	100%	7%
219.2	ref|YP_350987.1|gi|77461480| fructose-bisphosphate aldolase	Pfl01_5259	20 kDa	0	0
220	ref|YP_351234.1|gi|77461727| nitrogen regulatory protein P-II	Pfl01_5506	12 kDa	100%	7%
221	ref|YP_346260.1|gi|77456755| ATP phosphoribosyltransferase	Pfl01_0527	43 kDa	100%	7%
222	ref|YP_349194.1|gi|77459687| branched-chain alpha-keto acid dehydrogenase subunit E2	Pfl01_3465	45 kDa	100%	6%
223	ref|YP_347496.1|gi|77457991| 2-methylisocitrate lyase	Pfl01_1764	32 kDa	100%	0
224	ref|YP_347500.1|gi|77457995| 2-methylcitrate dehydratase	Pfl01_1768	55 kDa	100%	0
225	ref|YP_351109.1|gi|77461602| threonine dehydratase	Pfl01_5381	55 kDa	100%	0
226	ref|YP_351322.1|gi|77461815| hypothetical protein	Pfl01_5594	56 kDa	100%	0
227	ref|YP_350771.1|gi|77461264| hypothetical protein	Pfl01_5043	25 kDa	100%	0
228	ref|YP_351252.1|gi|77461745| alanine racemase	Pfl01_5524	40 kDa	100%	0
229	ref|YP_346568.1|gi|77457063| aspartyl/glutamyl-tRNA amidotransferase subunit A	Pfl01_0836	52 kDa	100%	0
230	ref|YP_351115.1|gi|77461608| D-3-phosphoglycerate dehydrogenase	Pfl01_5387	44 kDa	100%	89%
231	ref|YP_350801.1|gi|77461294| 30S ribosomal protein S3	Pfl01_5073	26 kDa	100%	78%
232	ref|YP_350247.1|gi|77460740| dihydroorotase	Pfl01_4519	38 kDa	100%	44%
233	ref|YP_346718.1|gi|77457213| valyl-tRNA synthetase	Pfl01_0986	107 kDa	100%	22%
234	ref|YP_346842.1|gi|77457337| Outer membrane chaperone Skp	Pfl01_1110	19 kDa	100%	13%
235	ref|YP_348359.1|gi|77458853| hypothetical protein	Pfl01_2628	36 kDa	100%	0
236	ref|YP_351226.1|gi|77461719| diaminopimelate decarboxylase	Pfl01_5498	45 kDa	100%	0
237	ref|YP_346603.1|gi|77457098| histidinol dehydrogenase	Pfl01_0871	47 kDa	100%	0
238	ref|YP_347230.1|gi|77457725| pterin-4-alpha-carbinolamine dehydratase	Pfl01_1498	13 kDa	100%	0
239	ref|YP_351086.1|gi|77461579| dihydroxy-acid dehydratase	Pfl01_5358	66 kDa	100%	0
240	ref|YP_346044.1|gi|77456539| extracellular solute-binding protein	Pfl01_0311	27 kDa	100%	0

#	Identified Protein		Accession Number	Molecular Weight	Protein Identification Probability	
					ΔrsmE	WT
241	ref|YP_346836.1|gi|77457331|	ribosome recycling factor	Pfl01_1104	20 kDa	100%	0
242	ref|YP_346514.1|gi|77457009|	30S ribosomal protein S15	Pfl01_0782	10 kDa	100%	96%
243	ref|YP_350316.1|gi|77460809|	inosine 5' monophosphate dehydrogenase	Pfl01_4588	52 kDa	100%	47%
244	ref|YP_350427.1|gi|77460920|	tryptophanyl tRNA synthetase	Pfl01_4699	50 kDa	100%	30%
245	ref|YP_350235.1|gi|77460728|	methionyl tRNA synthetase	Pfl01_4507	76 kDa	100%	0
246	ref|YP_350009.1|gi|77460502|	succinylarginine dihydrolase	Pfl01_4281	49 kDa	100%	0
247	ref|YP_350090.1|gi|77460583|	6 phosphogluconolactonase	Pfl01_4362	25 kDa	100%	0
248	ref|YP_346917.1|gi|77457412|	oxidoreductase FAD/NAD(P) binding	Pfl01_1185	30 kDa	100%	0
249	ref|YP_346099.1|gi|77456594|	histidine ammonia lyase	Pfl01_0366	54 kDa	100%	0
250	ref|YP_347323.1|gi|77457818|	thiosulfate binding protein	Pfl01_1591	37 kDa	100%	59%
251	ref|YP_349428.1|gi|77459921|	bifunctional 5,10 methylene tetrahydrofolate dehydrogenase	Pfl01_3699	31 kDa	100%	56%
252	ref|YP_346195.1|gi|77456690|	dihydrolipoamide acetyltransferase	Pfl01_0462	66 kDa	100%	38%
253	ref|YP_351266.1|gi|77461759|	extracellular solute binding protein	Pfl01_5538	58 kDa	100%	35%
254	ref|YP_346880.1|gi|77457375|	Phage tail tape measure protein TP901, core region	Pfl01_1148	83 kDa	100%	15%
255	ref|YP_351114.1|gi|77461607|	FAD linked oxidase like protein	Pfl01_5386	52 kDa	100%	14%
256	ref|YP_346093.1|gi|77456588|	urocanate hydratase	Pfl01_0360	61 kDa	100%	13%
257	ref|YP_346759.1|gi|77457254|	homoserine dehydrogenase	Pfl01_1027	46 kDa	100%	11%
258	ref|YP_347248.1|gi|77457743|	motility/biofilm formation protein FlhH	Pfl01_1516	28 kDa	100%	10%
259	ref|YP_350326.1|gi|77460819|	Pyrrolo quinoline quinone	Pfl01_4598	41 kDa	100%	10%
260	ref|YP_346890.1|gi|77457385|	DNA circulation like	Pfl01_1158	45 kDa	100%	7%
261	ref|YP_349386.1|gi|77459879|	propionyl CoA carboxylase	Pfl01_3657	57 kDa	100%	6%
262	ref|YP_347610.1|gi|77458105|	glutamyl tRNA synthetase	Pfl01_1878	59 kDa	100%	0
263	ref|YP_350344.1|gi|77460837|	inositol monophosphatase	Pfl01_4616	30 kDa	100%	0
264	ref|YP_349821.1|gi|77460314|	fumarate hydratase	Pfl01_4093	49 kDa	100%	0
265	ref|YP_349195.1|gi|77459688|	dihydrolipoamide dehydrogenase	Pfl01_3466	48 kDa	100%	0
266	ref|YP_349774.1|gi|77460267|	DegT/DnrJ/EryC1/StrS aminotransferase	Pfl01_4046	41 kDa	100%	0
267	ref|YP_346792.1|gi|77457287|	lysyl tRNA synthetase	Pfl01_1060	57 kDa	100%	0
268	ref|YP_350974.1|gi|77461467|	hypothetical protein	Pfl01_5246	48 kDa	100%	0
269	ref|YP_350155.1|gi|77460648|	bifunctional ornithine acetyltransferase/N acetylglutamate synthase	Pfl01_4427	43 kDa	100%	0
270	ref|YP_348836.1|gi|77459329|	B12 dependent methionine synthase	Pfl01_3107	136 kDa	100%	0
271	ref|YP_349367.1|gi|77459860|	cysteinyl tRNA synthetase	Pfl01_3638	52 kDa	100%	0
272	ref|YP_345866.1|gi|77456361|	von Willebrand factor A	Pfl01_0133	519 kDa	100%	19%
273	ref|YP_347497.1|gi|77457992|	methylcitrate synthase	Pfl01_1765	42 kDa	100%	9%
274	ref|YP_346325.1|gi|77456820|	periplasmic binding protein	Pfl01_0592	35 kDa	100%	8%
275	ref|YP_346770.1|gi|77457265|	glucan biosynthesis protein D	Pfl01_1038	61 kDa	100%	7%
276	ref|YP_347029.1|gi|77457524|	glycoside hydrolase family protein	Pfl01_1297	83 kDa	100%	0
277	ref|YP_350826.1|gi|77461319|	tyrosyl tRNA synthetase	Pfl01_5098	44 kDa	100%	0
278	ref|YP_349387.1|gi|77459880|	gamma carboxygeranoyl CoA hydratase	Pfl01_3658	30 kDa	100%	0
279	ref|YP_351217.1|gi|77461710|	argininosuccinate lyase	Pfl01_5489	52 kDa	100%	0
280	ref|YP_347999.1|gi|77458494|	amino acid adenylation protein	Pfl01_2267	333 kDa	100%	0
281	ref|YP_350717.1|gi|77461210|	aldehyde dehydrogenase	Pfl01_4989	56 kDa	100%	0
282	ref|YP_350002.1|gi|77460495|	aspartate kinase	Pfl01_4274	45 kDa	100%	0
283	ref|YP_350245.1|gi|77460738|	argininosuccinate synthase	Pfl01_4517	45 kDa	100%	0
284	ref|YP_347784.1|gi|77458279|	2,3 dimethylmalate lyase	Pfl01_2052	32 kDa	100%	0
285	ref|YP_350007.1|gi|77460500|	succinylglutamate desuccinylase	Pfl01_4279	38 kDa	100%	0
286	ref|YP_349047.1|gi|77459540|	TonB dependent siderophore receptor	Pfl01_3318	78 kDa	100%	0
287	ref|YP_350023.1|gi|77460516|	ribonucleotide diphosphate reductase subunit beta	Pfl01_4295	47 kDa	100%	38%
288	ref|YP_350098.1|gi|77460591|	extracellular solute binding protein	Pfl01_4370	46 kDa	100%	32%
289	ref|YP_351450.1|gi|77461943|	glucosamine fructose 6 phosphate aminotransferase	Pfl01_5726	66 kDa	100%	25%
290	ref|YP_348613.1|gi|77459107|	enoyl CoA hydratase	Pfl01_2882	30 kDa	100%	16%
291	ref|YP_348062.1|gi|77458557|	RHS protein	Pfl01_2330	177 kDa	100%	14%
292	ref|YP_349520.1|gi|77460013|	filamentous hemagglutinin like protein	Pfl01_3791	153 kDa	100%	8%
293	ref|YP_348986.1|gi|77459479|	amidase	Pfl01_3257	61 kDa	100%	5%
294	ref|YP_347411.1|gi|77457906|	C terminal processing peptidase 1	Pfl01_1679	79 kDa	100%	0
295	ref|YP_351346.1|gi|77461839|	aldose 1 epimerase	Pfl01_5618	33 kDa	100%	0
296	ref|YP_346798.1|gi|77457293|	OmpA/MotB	Pfl01_1066	29 kDa	100%	0
297	ref|YP_350859.1|gi|77461352|	aminoglycoside phosphotransferase	Pfl01_5131	38 kDa	100%	0
298	ref|YP_349614.1|gi|77460107|	erythronate 4 phosphate dehydrogenase	Pfl01_3886	41 kDa	100%	0
299	ref|YP_351270.1|gi|77461763|	phosphomannomutase	Pfl01_5542	50 kDa	100%	0
300	ref|YP_349593.1|gi|77460086|	glyceraldehyde 3 phosphate dehydrogenase	Pfl01_3865	53 kDa	99%	74%
301	ref|YP_350730.1|gi|77461223|	hypothetical protein	Pfl01_5002	63 kDa	100%	16%
302	ref|YP_350702.1|gi|77461195|	gamma glutamyl phosphate reductase	Pfl01_4974	45 kDa	100%	0
303	ref|YP_348659.1|gi|77459153|	glutathione reductase	Pfl01_2928	49 kDa	100%	0
304	ref|YP_350781.1|gi|77461274|	catalase like protein	Pfl01_5053	54 kDa	100%	0
305	ref|YP_345743.1|gi|77456238|	glycyl tRNA synthetase subunit alpha	Pfl01_0010	36 kDa	100%	0
306	ref|YP_351058.1|gi|77461551|	thiazole synthase	Pfl01_5330	28 kDa	100%	0
307	ref|YP_348314.1|gi|77458808|	TonB dependent siderophore receptor	Pfl01_2583	82 kDa	100%	0
308	ref|YP_350897.1|gi|77461390|	hypothetical protein	Pfl01_5169	76 kDa	100%	0

References

1. W. Kim, F. Racimo, J. Schluter, S. B. Levy, K. R. Foster, Importance of positioning for microbial evolution. *Proceedings of the National Academy of Sciences of the United States of America* 111, E1639-1647 (2014).
2. M. A. Larkin *et al.*, Clustal W and Clustal X version 2.0. *Bioinformatics* 23, 2947-2948 (2007).
3. M. W. Silby, C. Winstanley, S. A. C. Godfrey, S. B. Levy, R. W. Jackson, Pseudomonas genomes: diverse and adaptable. *FEMS Microbiology Reviews* 35, 652-680 (2011).
4. M. Avinash, S. Aishwarya, F. Zameer, S. Gopal, Pseudomonas aeruginosa biofilm and their molecular escape strategies. *Journal of Applied Biology and Biotechnology* 11, 28-37 (2023).
5. A. Anderson, Y. C. Kim, The Gac/Rsm Signaling Pathway of a Biocontrol Bacterium, *Pseudomonas chlororaphis* O6 Visible light during growth enhances conidial tolerance to different stress conditions in fungi View project Nanoparticles prime crop defenses for abiotic stress View project. 10.5423/RPD.2017.23.3.212.
6. F. Liang *et al.*, Cyclic-di-GMP Regulates the Quorum-Sensing System and Biocontrol Activity of *Pseudomonas fluorescens* 2P24 through the RsmA and RsmE Proteins. *Applied and Environmental Microbiology* 86 (2020).
7. M. Maurhofer *et al.*, Salicylic Acid Biosynthetic Genes Expressed in *Pseudomonas fluorescens* Strain P3 Improve the Induction of Systemic

Resistance in Tobacco Against Tobacco Necrosis Virus. *Phytopathology*® 88, 678-684 (1998).

8. L. P. Allsopp *et al.*, RsmA and AmrZ orchestrate the assembly of all three type VI secretion systems in *Pseudomonas aeruginosa*. *Proceedings of the National Academy of Sciences* 114, 7707-7712 (2017).

9. M. W. Silby *et al.*, Genomic and genetic analyses of diversity and plant interactions of *Pseudomonas fluorescens*. *Genome Biology* 10, R51-R51 (2009).

10. C. L. Miller *et al.*, RsmW, *Pseudomonas aeruginosa* small non-coding RsmA-binding RNA upregulated in biofilm versus planktonic growth conditions. *BMC Microbiology* 16, 1-16 (2016).

11. K. Takeuchi, GABA, A Primary Metabolite Controlled by the Gac/Rsm Regulatory Pathway, Favors a Planktonic Over a Biofilm Lifestyle in *Pseudomonas protegens* CHA0. *Molecular Plant-Microbe Interactions* 31, 274-282 (2018).

12. K. Lapouge, M. Schubert, F. H. T. Allain, D. Haas, Gac/Rsm signal transduction pathway of γ-proteobacteria: From RNA recognition to regulation of social behaviour. *Molecular Microbiology* 67, 241-253 (2008).

13. K. M. Sall *et al.*, A gacS deletion in *Pseudomonas aeruginosa* cystic fibrosis isolate CHA shapes its virulence. *PLoS ONE* 9, 1-10 (2014).

14. C. Reimmann, C. Valverde, E. Kay, D. Haas, Posttranscriptional Repression of GacS/GacA-Controlled Genes by the RNA-Binding Protein RsmE Acting Together with RsmA in the Biocontrol Strain Pseudomonas fluorescens CHA0. *JOURNAL OF BACTERIOLOGY* 187, 276-285 (2005).

15. X. Latour, The Evanescent GacS Signal. *Microorganisms* 8, 1746 (2020).

16. O. Duss *et al.*, Structural basis of the non-coding RNA RsmZ acting as a protein sponge. *Nature* 509, 588-592 (2014).

17. A. Ormazábal, J. Palma, G. Pierdominici-Sottile, Molecular Dynamics Simulations Unveil the Basis of the Sequential Binding of RsmE to the Noncoding RNA RsmZ. *The Journal of Physical Chemistry B* 125, 3045-3056 (2021).

18. T. Romeo, M. Gong, M. Y. Liu, A. M. Brun-Zinkernagel, Identification and molecular characterization of csrA, a pleiotropic gene from *Escherichia coli* that affects glycogen biosynthesis, gluconeogenesis, cell size, and surface properties. *Journal of bacteriology* 175, 4744-4755 (1993).

19. P. M. Sobrero, C. Valverde, Comparative Genomics and Evolutionary Analysis of RNA-Binding Proteins of the CsrA Family in the Genus *Pseudomonas*. *Frontiers in Molecular Biosciences* 7, 127-127 (2020).

20. Ó. Huertas-Rosales, M. I. Ramos-González, M. Espinosa-Urgel, Self-Regulation and Interplay of Rsm Family Proteins Modulate the Lifestyle of *Pseudomonas putida*. *Applied and environmental microbiology* 82, 5673-5686 (2016).

21. E. R. Morris *et al.*, Structural rearrangement in an RsmA/CsrA Ortholog of *Pseudomonas aeruginosa* creates a dimeric RNA-binding protein, RsmN. *Structure* 21, 1659-1671 (2013).

22. F. Ye *et al.*, Molecular basis of binding between the global post-transcriptional regulator CsrA and the T3SS chaperone CesT. *Nature Communications* 9, 1196-1196 (2018).

23. A. N. Edwards *et al.*, Circuitry Linking the Csr and Stringent Response Global Regulatory Systems. *Molecular microbiology* 80, 1561-1561 (2011).

24. Ó. Huertas-Rosales *et al.*, Genome-Wide Analysis of Targets for Post-Transcriptional Regulation by Rsm Proteins in *Pseudomonas putida*. *Frontiers in Molecular Biosciences* 0, 50-50 (2021).

25. M. Schubert *et al.*, Molecular basis of messenger RNA recognition by the specific bacterial repressing clamp RsmA/CsrA. *Nature Structural and Molecular Biology* 14, 807-813 (2007).

26. M. Romero *et al.*, Genome-wide mapping of the RNA targets of the *Pseudomonas aeruginosa* riboregulatory protein RsmN. *Nucleic Acids Research* 46, 6823-6840 (2018).

27. O. Duss, E. Michel, N. Diarra dit Konté, M. Schubert, F. H. T. Allain, Molecular basis for the wide range of affinity found in Csr/Rsm protein-RNA recognition. *Nucleic acids research* 42, 5332-5346 (2014).

28. E. Holmqvist *et al.*, Global RNA recognition patterns of post-transcriptional regulators Hfq and CsrA revealed by UV crosslinking *in vivo*. *The EMBO Journal* 35, 991-1011 (2016).

29. T. Romeo, P. Babitzke, Global Regulation by CsrA and Its RNA Antagonists. 6, 341-354 (2018).

30. Y. Irie *et al.*, *Pseudomonas aeruginosa* biofilm matrix polysaccharide Psl is regulated transcriptionally by RpoS and post-transcriptionally by RsmA. *Molecular Microbiology* 78, 158-158 (2010).

31. C. S. Baker *et al.*, CsrA Inhibits Translation Initiation of *Escherichia coli* hfq by Binding to a Single Site Overlapping the Shine-Dalgarno Sequence. *Journal of Bacteriology* 189, 5472-5472 (2007).

32. D. Ramírez-Zapata *et al.*, Two Homologues of the Global Regulator Csr/Rsm Redundantly Control Phaseolotoxin Biosynthesis and Virulence in the Plant Pathogen *Pseudomonas amygdali* pv. phaseolicola 1448A. *Microorganisms 2020, Vol. 8, Page 1536* 8, 1536-1536 (2020).

33. C. A. Vakulskas, A. H. Potts, P. Babitzke, B. M. M. Ahmer, T. Romeo, Regulation of Bacterial Virulence by Csr (Rsm) Systems. *Microbiology and Molecular Biology Reviews* 79, 193-224 (2015).

34. M. F. DeFlaun, A. S. Tanzer, A. L. McAteer, B. Marshall, S. B. Levy, Development of an Adhesion Assay and Characterization of an Adhesion-Deficient Mutant of *Pseudomonas fluorescens*. *Applied and Environmental Microbiology* 56, 112-119 (1990).

35. G. Compeau, B. J. Al-Achi, E. Platsouka, S. B. Levy, Survival of rifampin-resistant mutants of Pseudomonas fluorescens and Pseudomonas putida in soil systems. *Applied and Environmental Microbiology* 54, 2432-2438 (1988).

36. S. C. Seaton, M. W. Silby, S. B. Levy, Pleiotropic Effects of GacA on Pseudomonas fluorescens Pf0-1 *In Vitro* and in Soil. *Applied and Environmental Microbiology* 79, 5405-5410 (2013).

37. K. H. Schulmeyer *et al.*, Primary and secondary sequence structure requirements for recognition and discrimination of target RNAs by *Pseudomonas aeruginosa* RsmA and RsmF. *Journal of Bacteriology* 198, 2458-2469 (2016).

38. H. C. Flemming, J. Wingender, The biofilm matrix. *Nat Rev Microbiol* 8, 623-633 (2010).

39. I. W. Sutherland, The biofilm matrix- an immobilized but dynamic microbial environment. *Trends Microbiol* 9, 222-227 (2001).

40. J. N. C. Fong, F. H. Yildiz, Biofilm Matrix Proteins. *Microbiology Spectrum* 3 (2015).

41. L. Hobley, C. Harkins, C. E. Macphee, N. R. Stanley-Wall, Giving structure to the biofilm matrix: an overview of individual strategies and emerging common themes. *FEMS Microbiology Reviews* 39, 649-669 (2015).

42. S. S. Branda, Å. Vik, L. Friedman, R. Kolter, Biofilms: the matrix revisited. *Trends in Microbiology* 13, 20-26 (2005).

43. L. Karygianni, Z. Ren, H. Koo, T. Thurnheer, Biofilm Matrixome: Extracellular Components in Structured Microbial Communities. *Trends in Microbiology* 28, 668-681 (2020).

44. A. Persat *et al.*, The mechanical world of bacteria. *Cell* 161, 988-997 (2015).

45. C. D. Nadell, J. B. Xavier, K. R. Foster, The sociobiology of biofilms. *FEMS Microbiology Reviews* 33, 206-224 (2009).

46. J. B. Xavier, K. R. Foster, Cooperation and conflict in microbial biofilms. *Proceedings of the National Academy of Sciences* 104, 876-881 (2007).

47. C. D. Nadell, K. R. Foster, J. B. Xavier, Emergence of Spatial Structure in Cell Groups and the Evolution of Cooperation. *PLoS Computational Biology* 6, e1000716 (2010).

48. S. Mitri, J. B. Xavier, K. R. Foster, Social evolution in multispecies biofilms. *Proceedings of the National Academy of Sciences* 108, 10839-10846 (2011).

49. J. L. Mark Welch, B. J. Rossetti, C. W. Rieken, F. E. Dewhirst, G. G. Borisy, Biogeography of a human oral microbiome at the micron scale. *Proceedings of the National Academy of Sciences* 113, E791-E800 (2016).

50. O. Hallatschek, The noisy edge of traveling waves. *Proceedings of the National Academy of Sciences* 108, 1783-1787 (2011).

51. O. Hallatschek, P. Hersen, S. Ramanathan, D. R. Nelson, Genetic drift at expanding frontiers promotes gene segregation. *Proceedings of the National Academy of Sciences* 104, 19926-19930 (2007).

52. K. S. Korolev, J. B. Xavier, D. R. Nelson, K. R. Foster, A Quantitative Test of Population Genetics Using Spatiogenetic Patterns in Bacterial Colonies. *The American Naturalist* 178, 538-552 (2011).

53. D. A. I. Mavridou, D. Gonzalez, W. Kim, S. A. West, K. R. Foster, Bacteria Use Collective Behavior to Generate Diverse Combat Strategies. *Current Biology* 28, 345-355.e344 (2018).

54. B. J. Crespi, The evolution of social behavior in microorganisms. *Trends in Ecology & Evolution* 16, 178-183 (2001).

55. S. A. West, A. S. Griffin, A. Gardner, S. P. Diggle, Social evolution theory for microorganisms. *Nature reviews. Microbiology* 4, 597-607 (2006).

56. W. P. J. Smith *et al.*, Cell morphology drives spatial patterning in microbial communities. *Proceedings of the National Academy of Sciences* 114, E280-E286 (2017).

57. S. K. Hansen, P. B. Rainey, J. A. J. Haagensen, S. Molin, Evolution of species interactions in a biofilm community. *Nature* 445, 533-536 (2007).

58. D. O. Serra, A. M. Richter, G. Klauck, F. Mika, R. Hengge, Microanatomy at cellular resolution and spatial order of physiological differentiation in a bacterial biofilm. *mBio* 4, e00103-00113 (2013).

59. R. Kolter, E. P. Greenberg, Microbial sciences: the superficial life of microbes. *Nature* 441, 300-302 (2006).

60. L. K, S. M, A. Fh, H. D, Gac/Rsm signal transduction pathway of gamma-proteobacteria: from RNA recognition to regulation of social behaviour. *Molecular microbiology* 67, 241-253 (2008).

61. E. Holmqvist *et al.*, Global RNA recognition patterns of post-transcriptional regulators Hfq and CsrA revealed by UV crosslinking in vivo. *The EMBO journal* 35, 991-1011 (2016).

62. O. Duss, E. Michel, N. D. Dit Konté, M. Schubert, F. H. T. Allain, Molecular basis for the wide range of affinity found in Csr/Rsm protein-RNA recognition. *Nucleic Acids Research* 42, 5332-5346 (2014).

63. T. Romeo, P. Babitzke, Global Regulation by CsrA and Its RNA Antagonists. *Microbiology Spectrum* 6 (2018).

64. Ó. Huertas-Rosales, M. I. Ramos-González, M. Espinosa-Urgel, Selfregulation and interplay of Rsm family proteins modulate the lifestyle of Pseudomonas putida. *Applied and Environmental Microbiology* 82, 5673-5686 (2016).

65. W. Kim, S. B. Levy, K. R. Foster, Rapid radiation in bacteria leads to a division of labour. *Nature Communications* 7, 10508-10508 (2016).

66. R. Simon, U. Priefer, A. Pühler, A Broad Host Range Mobilization System for In Vivo Genetic Engineering: Transposon Mutagenesis in Gram Negative Bacteria. *Bio/Technology* 1, 784-791 (1983).

67. K. J. Livak, T. D. Schmittgen, Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods (San Diego, Calif.)* 25, 402-408 (2001).

68. V. d. Lorenzo, M. Herrero, U. Jakubzik, K. N. Timmis, Mini-Tn5 transposon derivatives for insertion mutagenesis, promoter probing, and chromosomal insertion of cloned DNA in gram-negative eubacteria. *Journal of Bacteriology* 172, 6568-6572 (1990).

69. M. Gonin, E. M. Quardokus, D. O'Donnol, J. Maddock, Y. V. Brun, Regulation of Stalk Elongation by Phosphate in *Caulobacter crescentus*. *Journal of Bacteriology* 182, 337-347 (2000).

70. R. M. Horton, H. D. Hunt, S. N. Ho, J. K. Pullen, L. R. Pease, Engineering hybrid genes without the use of restriction enzymes: gene splicing by overlap extension. *Gene* 77, 61-68 (1989).

71. R. M. Q. Shanks, N. C. Caiazza, S. M. Hinsa, C. M. Toutain, G. A. O'Toole, *Saccharomyces cerevisiae*-Based Molecular Tool Kit for Manipulation of Genes from Gram-Negative Bacteria. *Applied and Environmental Microbiology* 72, 5027-5036 (2006).

72. M. Matthews, C. R. Roy, Identification and Subcellular Localization of the *Legionella pneumophila* IcmX Protein: a Factor Essential for Establishment of a

Replicative Organelle in Eukaryotic Host Cells. *Infection and Immunity* 68, 3971-3982 (2000).

73. L. Lambertsen, C. Sternberg, S. Molin, Mini-Tn7 transposons for site-specific tagging of bacteria with fluorescent proteins. *Environmental Microbiology* 6, 726-732 (2004).

74. C. Kessler, E. Mhatre, V. Cooper, W. Kim, I. S. Druzhinina, Evolutionary Divergence of the Wsp Signal Transduction Systems in Beta- and Gammaproteobacteria. *Applied and Environmental Microbiology* 87, e01306-01321 (2021).

75. M. J. Wiser, R. E. Lenski, A comparison of methods to measure fitness in Escherichia coli. *PLoS ONE* 10 (2015).

76. S. A. Sieber, M. A. Marahiel, Learning from Nature's Drug Factories: Nonribosomal Synthesisof MacrocyclicPeptides. *Journal of Bacteriology* 185, 7036-7043 (2003).

77. G. Jahanshah *et al.*, Discovery of the Cyclic Lipopeptide Gacamide A by Genome Mining and Repair of the Defective GacA Regulator in *Pseudomonas fluorescens* Pf0-1. *Journal of Natural Products* 82, 301-308 (2019).

78. I. De Bruijn, M. J. D. De Kock, P. De Waard, T. A. Van Beek, J. M. Raaijmakers, Massetolide A Biosynthesis in *Pseudomonas fluorescens*. *Journal of Bacteriology* 190, 2777-2789 (2008).

79. I. De Bruijn *et al.*, Genome-based discovery, structure prediction and functional analysis of cyclic lipopeptide antibiotics in *Pseudomonas* species. *Molecular Microbiology* 63, 417-428 (2007).

80. D. B. Kearns, A field guide to bacterial swarming motility. *Nature Reviews Microbiology* 8, 634-644 (2010).

81. J. M. Raaijmakers, I. De Bruijn, O. Nybroe, M. Ongena, Natural functions of lipopeptides from *Bacillus* and *Pseudomonas*: more than surfactants and antibiotics. *FEMS Microbiology Reviews* 34, 1037-1062 (2010).

82. C. D. Nadell, B. L. Bassler, A fitness trade-off between local competition and dispersal in *Vibrio cholerae* biofilms. *Proceedings of the National Academy of Sciences* 108, 14181-14185 (2011).

83. G. Reid *et al.*, Microbiota restoration: natural and supplemented recovery of human microbial communities. *Nature Reviews Microbiology* 9, 27-38 (2011).

84. S. S. Branda, J. E. González-Pastor, S. Ben-Yehuda, R. Losick, R. Kolter, Fruiting body formation by *Bacillus subtilis*. *Proceedings of the National Academy of Sciences* 98, 11621-11626 (2001).

85. H. C. Flemming *et al.*, Biofilms: An emergent form of bacterial life. *Nature Reviews Microbiology* 14, 563-575 (2016).

86. L. Karygianni, Z. Ren, H. Koo, T. Thurnheer, Biofilm Matrixome: Extracellular Components in Structured Microbial Communities. 28, 668-681 (2020).

87. M. Thérien *et al.*, Surfactin production is not essential for pellicle and root-associated biofilm development of *Bacillus subtilis*. *Biofilm* 2, 100021 (2020).

88. A. F. Evans *et al.*, Spatial Structure Formation by RsmE-Regulated Extracellular Secretions in Pseudomonas fluorescens Pf0-1. *Journal of Bacteriology* 204, e00285-00222 (2022).

89. P. A. Lind, E. Libby, J. Herzog, P. B. Rainey, Predicting mutational routes to new adaptive phenotypes. *eLife* 8 (2019).

90. P. A. Lind, A. D. Farr, P. B. Rainey, Evolutionary convergence in experimental Pseudomonas populations. *The ISME Journal* 11, 589-600 (2017).

91. U. K. Laemmli, Cleavage of Structural Proteins during the Assembly of the Head of Bacteriophage T4. *Nature* 227, 680-685 (1970).

92. D. Kim, J. M. Paggi, C. Park, C. Bennett, S. L. Salzberg, Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nature Biotechnology* 37, 907-915 (2019).

93. B. Langmead, S. L. Salzberg, Fast gapped-read alignment with Bowtie 2. *Nat Methods* 9, 357-359 (2012).

94. C. Trapnell *et al.*, Differential analysis of gene regulation at transcript resolution with RNA-seq. *Nature Biotechnology* 31, 46-53 (2013).

95. M. Gallique, M. Bouteiller, A. Merieau, The Type VI Secretion System: A Dynamic System for Bacterial Communication? *Frontiers in Microbiology* 8 (2017).

96. M. S. Dueholm *et al.*, Functional amyloid in Pseudomonas. *Molecular Microbiology* 10.1111/j.1365-2958.2010.07269.x, no-no (2010).

97. K. L. Harvey, V. M. Jarocki, I. G. Charles, S. P. Djordjevic, The Diverse Functional Roles of Elongation Factor Tu (EF-Tu) in Microbial Pathogenesis. *Frontiers in Microbiology* 10 (2019).

98. L. Boulos, M. Prévost, B. Barbeau, J. Coallier, R. Desjardins, LIVE/DEAD® BacLight™: application of a new rapid staining method for direct enumeration of

viable and total bacteria in drinking water. *Journal of Microbiological Methods* 37, 77-86 (1999).

99. S. J. Hersch, K. Manera, T. G. Dong, Defending against the Type Six Secretion System: beyond Immunity Genes. *Cell Reports* 33, 108259 (2020).

100. P. R. Kulkarni *et al.*, A sequence-based approach for prediction of CsrA/RsmA targets in bacteria with experimental validation in *Pseudomonas aeruginosa*. *Nucleic Acids Research* 42, 6811-6825 (2014).

101. K. Lapouge *et al.*, RNA pentaloop structures as effective targets of regulators belonging to the RsmA/CsrA protein family. *RNA Biology* 10, 1030-1041 (2013).

102. P. Müller, M. Gimpel, T. Wildenhain, S. Brantl, RNA Biology A new role for CsrA: promotion of complex formation between an sRNA and its mRNA target in *Bacillus subtilis* A new role for CsrA: promotion of complex formation between an sRNA and its mRNA target in *Bacillus subtilis*. 10.1080/15476286.2019.1605811 (2019).

103. F. Altegoera, S. A. Rensingb, G. Bangea, Structural basis for the CsrA-dependent modulation of translation initiation by an ancient regulatory protein. *Proceedings of the National Academy of Sciences of the United States of America* 113, 10168-10173 (2016).

104. N. S. Muthunayake, D. T. Tomares, W. S. Childers, J. M. Schrader, Phase-separated bacterial ribonucleoprotein bodies organize mRNA decay. *Wiley Interdisciplinary Reviews: RNA* 11 (2020).

105. N. Al-Husini *et al.*, BR-Bodies Provide Selectively Permeable Condensates that Stimulate mRNA Decay and Prevent Release of Decay Intermediates. *Molecular Cell* 78, 670-682.e678 (2020).

106. J. S. Kim, Y. H. Kim, J. Y. Park, A. J. Anderson, Y. C. Kim, The global regulator GacS regulates biofilm formation in Pseudomonas chlororaphis O6 differently with carbon source. *Canadian Journal of Microbiology* 60, 133-138 (2014).

107. R. E. Steinberger, P. A. Holden, Extracellular DNA in Single- and Multiple-Species Unsaturated Biofilms. *Applied and Environmental Microbiology* 71, 5404-5410 (2005).