ABSTRACT

Title of Dissertation: **THE STANDALONE REGULATOR ROFA OF STREPTOCOCCUS PYOGENES EXHIBITS CHARACTERISTICS OF A PRD-CONTAINING VIRULENCE REGULATOR**

Meaghan T. Hart, Doctor of Philosophy, 2024

Dissertation directed by: Professor Kevin S. McIver, Cell Biology and Molecular Genetics Department

Streptococcus pyogenes (Group A *Streptococcus*; GAS) is a human pathogen estimated to cause nearly 790 million cases of disease annually at diverse tissue sites. To successfully infect these sites, GAS must detect nutrient availability and adapt accordingly. One mechanism employed to detect and import carbohydrates is the phosphoenolpyruvate transferase system (PTS), which mediates both carbohydrate uptake and metabolic gene regulation. Gene regulation by the PTS can occur through phosphorylation of transcriptional regulators at conserved PTS-regulatory domains (PRDs). GAS has several stand-alone regulators that contain PRDs, with corresponding regulons encoding both metabolic genes and important virulence factors. These regulators form a family called PRD-Containing Virulence Regulators (PCVRs). RofA is a putative member of this family and is known to regulate the expression of genes important for virulence. It was hypothesized that RofA is phosphorylated by the PTS in response to carbohydrate levels to coordinate appropriate virulence gene expression. In this dissertation, the RofA regulon

was determined in strain 5448, a representative strain of the globally disseminated M1T1 serotype. The pilus and capsule operons were consistently dysregulated across growth in the absence of RofA. This correlated with increased capsule production and decreased adherence to primary keratinocytes. Purified RofA-His was phosphorylated *in vitro* by the general PTS components EI and HPr, and phosphorylated species of RofA-FLAG were detected *in vivo* late in stationary phase in a glucose-dependent manner. Together, these findings support the hypothesis that RofA is a PCVR that may couple sugar detection and utilization with GAS virulence gene regulation. Additionally, a bioluminescent construct was generated for allelic exchange into any *S. pyogenes* strain. Allelic exchange of this construct into WT 5448 yielded strains that were highly bioluminescent, grew to a similar density as WT, and survived as well as WT when challenged with human neutrophils. This tool could be used to study the contribution of specific proteins on *in vivo* virulence in a non-invasive manner, including RofA and RofA phosphorylation.

THE STANDALONE REGULATOR ROFA OF STREPTOCOCCUS PYOGENES EXHIBITS CHARACTERISTICS OF A PRD-CONTAINING VIRULENCE REGULATOR

by

Meaghan T. Hart

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2024

Advisory Committee:
Professor Kevin S. McIver, Chair
Professor Daniel C. Stein
Professor Vincent T. Lee
Associate Professor Margaret A. Scull
Associate Professor Jason D. Kahn, Dean's Representative

Dedication

This dissertation is dedicated to Betty Lucille Nordaas and James Hoagland French; two individuals who encouraged my intellectual curiosity and academic drive from a very young age. Thank you, Mom-Mom and Grandpa!

Acknowledgements

There are many people to thank and acknowledge for their time, attention, mentorship, and friendship, which have all contributed to me getting here. I owe so much to you all. Thank you!

Thank you to my family: my mom, Holly Hart, my dad, Michael Hart, and my sister, Madison Hart, for the years of support and encouragement in all my pursuits, academic and otherwise. Specifically, thank you Mom, for encouraging me to stay in the biology lab that year I wanted to quit and make beer (ha!). And thank you to all the Harts and Frenchs, I am lucky to have a family like ours.

Thank you to my husband, Connor Sage, who supported me when I wanted to return to graduate school. I sincerely appreciate your patience in these last five years so I could make this happen. This process would have been so much harder if I did not have you to come home to.

Thank you to many of my former colleagues from the Johns Hopkins University Applied Physics Laboratory for all that you taught me in my post-baccalaureate career there. Special thanks to Dr. Audrey Fischer-Hesselbrock, my original scientific mentor, who took the time to teach a 17-year-old how to pipette and encourage me to pursue a path in molecular biology and microbiology.

Thank you to my advisor, Dr. Kevin McIver, for your guidance and wisdom throughout my time in graduate school. I sincerely appreciate the time and effort you put into my training. And thank you to my committee, Dr. Daniel Stein, Dr. Vincent Lee, Dr. Margaret Scull, and Dr. Jason Kahn for your guidance and insights. I've always appreciated our conversations and your feedback on my research.

Finally, thank you to all the members of the McIver Lab, past and present, for your work, conversations, advice, and friendship. I would especially like to thank Dr. Joseph Rom, for your mentorship early in my training. I would also like to thank my fellow graduate students, Aolani Perry, Spencer Lewis, Alejandra Taboada and former lab member Dr. Rezia Era Braza, for being excellent peers and lab mates in the time I was pursuing my Ph. D.

Table of Contents

List of Tables

List of Figures

List of Appendices

List of Abbreviations

Abbreviation	Meaning
1:10	1 to 10, this notation was used for all ratios. (1:2, 1:100 etc)
[^{32}P]EP	^{32}P-radiolabeled phosphoenolpyruvate
~	Approximately; Context dependent
°C	Degrees celcius
µCi	Microcurie
µg	Microgram
µL	Microliter
µm	Micron/Micrometer
^{32}P	Phosphorous-32 radioactive isotope
A	Alanine
A.A.	Ascorbic acid
ABC	ATP-binding cassette
adj-p	Adjusted p-value
Amp	Ampicillin
ANOVA	Analysis of Variance
Ap	Ampicillin
Ap^{r}	Ampicillin resistant
ARF	Acute rheumatic fever
Asc	Ascorbic acid
ATCC	American Type Culture Collection
ATP	Adenosine triphosphate
AtxA	Anthrax toxin activator
AUC	Area under curve
B. cereus	*Bacillus cereus*
B. subtilis	*Bacillus subtilis*
BAP	Blood agar plate
BBI	Brain and Behavior Institute
BBI-AGTC	BBI-Advanced Genomic Technologies Core
bc	because
BLAST	Basic local alignment search tool
BMH	Bis(maleimido)hexane
C	Cysteine
C. difficile	*Clostridum difficle*
cAMP	Cyclic adenosine monophosphate
Carb	Carbenicillin; context dependent
Carb.	Carbohydrate; context dependent
Cat.	Category

Abbreviation	Meaning
CCR	Carbon catabolite repression
CDC	Centers for Disease Control and Prevention
CDD	Conserved domain database
CDM	Chemically defined medium
cDNA	Copy DNA
Cel	Cellobiose
CFU	Colony forming units
ChIP-Seq	Chromatin immunoprecipitation sequencing
COG	Cluster of orthologous gene
Co-IP	Coimmunoprecipitation
COVID-19	Corona Virus Disease of 2019
cre	Catabolite response element
Crp	cAMP receptor protein
C_T	Cycle threshold value
D	Aspartate
Desc.	Description
diH_2O	Deionized water
dL	Deciliter
DNA	Deoxyribonucleic acid
DNQ	Dectected, not quantified
E. coli	*Escherichia coli*
E. faecalis	*Enterococcus faecalis*
ECM	Extracellular matrix
EDTA	Ethylenediaminetetraacetic Acid
EF1	Elution fraction 1; This nomenclature is the same for any abbreviation that is in this form.
EI	Enzyme I of the PTS
EII	Enzyme II of the PTS
EMSA	Electrophoretic mobility shift assay
ENQ	Expected, not quantified
Ery^r	Erythromycin resistant
EST	Early stationary
EX	Exponential
Express.	Expression
FBS	Fetal bovine serum
FC	Fold change
FCT	Fibronectin, collagen-binding, t-antigen
FFluc	Firefly luciferase
Fru	Fructose

Abbreviation	Meaning
Fru-6-P	Fructose-6-Phosphate
FT	Flow through
GAC	Group A Carbohydrate
Gal	Galactose
Gal-1-P	Galactose-1-phosphate
GAS	Group A *Streptococcus*
Gat	Galactitol
gDNA	Genomic DNA
Glc	Glucose
Glu	Glucose
Glu-6-P	Glucose-6-phosphate
H	Histidine
h	Hours
H>A	Substitution of histidine for alanine
H>D	Substitution of histidine for aspartate
H199	Histidine 199. This nomenclature is the same for any abbreviation in this form.
H199A	Alanine substitution mutation at histidine 199. This nomenclature is the same for any abbreviation that is in this form.
HBSS	Hanks Balanced Salt Solution
HCl	Hydrochloric acid
HD	Heat denatured
His	Histidine
His_6	6x-His epitope
HPr	Histidine phosphocarrier protein of the PTS
HPrK/P	Histidine kinase/phosphorylase protein
HRP	Horseradish peroxidase
HTH	Helix-turn-helix
Hyal	Hyaluronidase
IDT	Integrated DNA Technologies
IRB	Insitutional review board
IVIS	In Vivo Imaging System
Kan	Kanamycin
kCPM	Kilo-counts per minute
kDa	Kilodalton
KEGG	Kyoto Encyclopedia of Genes and Genomes
Km	Kanamycin
Km^r	Kanamycin resistant

Abbreviation	Meaning
kV	Kilovolts
L. lactis	*Lactobacillus lactis*
Lac	Lactose
LB	Luria-Bertani broth
log2_FC	Logarithm base2 fold-change
LST	Late stationary
M	Molar (moles/L)
M1	M protein serotype group 1; This nomenclature is the same for any abbreviation that is in this form.
mA	Milliamps
Mal	Maltose
Mal-Tri	Maltotriose
Man	Mannose
Man-6-P	Mannose-6-phosphate
MBP	Maltose-binding protein
Met	Methionine
Metab.	Metabolism
mg	Milligram
Mga	Multiple gene regulator of Group A *Streptococcus*
min	Minutes
mL	Milliliter
mM	Millimolar (millimoles/L)
MOI	Multiplicity of infection
mRNA	Messenger RNA
MW	Molecular weight
MWCO	Molecular weight cut off
NAc -Glu -6-P	N-acetylglucosamine-6-phosphate
NAc-Gal	N-acetylgalactosamine
NAc-Glu	N-acetylglucosamine
NALT	Nasal-associated lymphoid tissue
NAP	Nucleoid associated protein
NaPi	Sodium phosphate buffer
NCBI	National Center for Biotechnology Information
NEB	New England Biolabs
ng	Nanograms
Ni-NTA	Nickel-Nitrilotriacetic acid
Nluc	Nano-Luciferase
nM	Nanomolar (nM)
No.	Number
NS; ns	Not significant

Abbreviation	Meaning
OD_{600}	Optical density measured by absorbance at 600 nm
OE-PCR	Overlap extension PCR
ORF	Open reading frame
P- or P~	Indicates phosphorylation of a protein or residue
P~ylated	Phosphorylated
PANDAS	Pediatric autoimmune neuropsychiatric disorders associated with Streptococcal infection
PBS	Phosphate buffered saline
PBST	phosphate buffered saline plus tween
PCR	Polymerase chain reaction
PCVR	PRD-containing virulence regulator
PEP	Phosphoenolpyruvate
PMNs	Polymorphonuclear lymphocytes
pmol	Picomole
PRD	PTS Regulation Domain
PSI-BLAST	Position specific iterative basic local alignment search tool
PTS	Phosphoenolpyruvate-Dependent Transferase System
PVDF	polyvinylidene difluoride
RALP	RofA-like protein
RBS	Ribosome binding site
Refs.	References
RHD	Rheumatic heart disease
RNA	Ribonucleic acid
RNA-Seq	RNA-Sequencing
RofA	Regulator of F
rRNA	Ribosomal RNA
RT	Reverse transcriptase or room temperature; context dependent
RT-PCR	Reverse-transcriptase PCR
RT-qPCR	Reverse-transcriptase quantitative PCR
S. aureus	*Staphylococcus aureus*
S. enterica	*Salmonella enterica*
S. mutans	*Streptococcus mutans*
S. pneumoniae	*Streptococcus pneumoniae*
S. pyogenes	*Streptococcus pyogenes*
Sal	Salicin
SDS	Sodium dodecyl sulfate
SDS-PAGE	Sodium dodecyl sulfate – polyacrylamide gel electrophoresis
Ser	Serine

Abbreviation	Meaning
SOF	Serum opacity factor
Sp	Spectinomycin
SPE	Streptococcal pyrogenic exotoxin
Spec	Spectinomycin
Sp^r	Spectinomycin resistant
sRNA	Small non-coding RNA
Suc	Sucrose
T. S.	This study
T1	T-antigen serotype group 1; This nomenclature is the same for any abbreviation that is in this form.
TCA	Tricarboxylic acid
TCS	Two-component system
TEAB	triethylammonium bicarbonate buffer
THY	Todd-hewitt plus yeast extract
TLR-2	Toll-Like Receptor 2
Tn	Transposon
Tn-Seq	Transposon mutagenesis sequencing
TR	Transition
Tre	Trehalose
Tre-6-P	Trehalose-6-Phosphate
TSA	Tryptic soy agar
TSS	Toxic shock syndrome
U	Activity units
V	Volts
w/	with
W1	Wash 1; This nomenclature is the same for any abbreviation that is in this form.
WT	Wild type
x g	Gravitational acceleration; g force (unit of relative centrifugal force)
α-FLAG	Anti-FLAG
α-His	Anti-6x-His
α-Lac	alpha-Lactose
β-Lac	beta-Lactose
βME	Beta-mercaptoethanol
Δ	Greek symbol delta was used to indicate the absence of something, most often, indicating a gene deletion

Chapter 1: Introduction and Literature Review

Brief History of *Streptococcus pyogenes* Research

Streptococci were formally discovered in 1879 by Louis Pasteur (1). However, there were earlier written reports of erysipelas and childbed fever by Hippocrates in the 4th century BC, both of which are commonly caused by *Streptococcus pyogenes* (1). The name, *Streptococcus pyogenes*, came from Friedrich Rosenbach in 1884, due to the pus-filled (*pyo*) lesions that were infected with the chained cocci (1). In 1903, Schottmütller observed that certain isolates of *Streptococcus* lysed red blood cells on agar plates (2) and this was the basis for the distinction of *Streptococcus haemoyticus*; a coccoid strain that grew in chains and lysed red blood cells (1). This included the same species that Rosenbach had termed *Streptococcus pyogenes*, and the two terms seem to be used interchangeably in early literature.

Several groups then worked to distinguish Streptococci from one another with a variety of biochemical, phagological, and serological assays (3). One such effort was from Rebecca Lancefield who in 1933 established a serological typing method from Streptococcal surface antigen extracts that successfully distinguished different groups of *Streptococcus* and correlated well with the other established classification methods (4). *Streptococcus pyogenes* was distinguished from other Streptococci by this method due to its distinct cell-wall carbohydrate, designated A (Group A Carbohydrate, GAC), which led to it being termed Group A *Streptococcus* (GAS) (4). Lancefield, in this study, also showed that GAS was the predominant group causing human infections.

Further classification of *S. pyogenes* was achieved through further serotyping. A serotype is a group of organisms that share certain surface antigens that the immune system reacts to. Serotyping is the process of identifying the serotype of a unknown isolate by incubating it with immune sera that has been generated against a known type. For *S. pyogenes*, serotyping was conducted on the basis of surface M-protein, the T-antigen (pilus), and serum opacity factor (SOF) (5). More recently, molecular methods have been used for *emm*-typing, which is based on the hypervariable region at the 5' end of the gene (6). This has led to the classification of over 275 *emm*-types (7) and is the most common way to differentiate *S. pyogenes*. There are fewer serological T-types, with 21 reference strains being used to generate typing sera by the United States Center for Disease Control and Prevention (CDC) (8). Because of cross-reactivity between typing sera and multiple T-types, and the presence of non-T-typable strains, a scheme for pilin genotyping was recently proposed, as pili are the surface component responsible for T-serotyping (8).

S. pyogenes Disease Manifestations and Burden

S. pyogenes is estimated to cause approximately 790 million cases of new and existing disease each year and over 600,000 deaths (9) (**Fig. 1.1**). Most of these are due to pharyngitis (616 million, 78%) (9), an infection of the throat which is often mild and self-limiting (10). Impetigo, also known as pyoderma, is the second most common disease caused by *S. pyogenes* (162 million, 21%) and manifests as a self-limiting skin infection with purulent lesions that quickly crust over (9, 11). Other skin and soft-tissue disease manifestations in order of increasing invasiveness include erysipelas, cellulitis,

necrotizing fasciitis, and myonecrosis (11). Scarlet fever is a systemic complication that arises from an *S. pyogenes* infection and manifests as a skin rash over the whole body including the tongue (10, 12). Scarlet fever is usually associated with pharyngeal infection, though it can arise following infections at other sites, and is believed to be due to the secretion of streptococcal pyrogenic exotoxins (SPEs) (10, 11). Scarlet fever incidence had dramatically reduced in the antibiotic era (10) but has seen a resurgence in several countries, including China, the United Kingdom, South Korea, and Singapore (13-16).

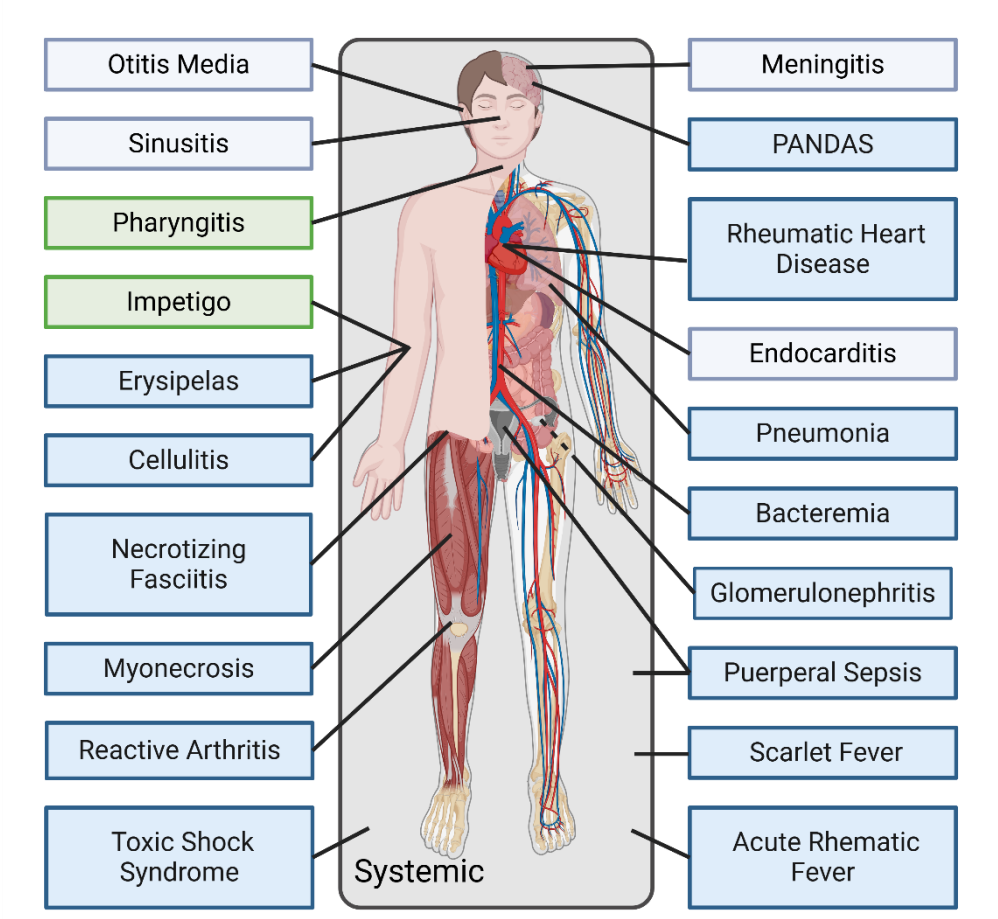


Figure 1.1: Disease Manifestations of *S. pyogenes*
Schematic of the human body and the disease manifestations of *S. pyogenes* at various body sites. The two most common disease manifestations are shown in green. Created in BioRender.com.

Other diseases caused by *S. pyogenes* include puerperal sepsis, bacteremia, pneumonia, streptococcal toxic shock syndrome (TSS) and more rarely endocarditis, sinusitis, otitis media, cervical lymphadenitis, and meningitis (9-11, 17, 18). Several streptococcal disease manifestations, such as bacteremia, puerperal sepsis, pneumonia, necrotizing fasciitis, myonecrosis, and streptococcal TSS, are life-threatening with high mortality rates (9, 18). In the United States, the rate of life-threatening invasive *S. pyogenes* infections have been increasing since 2012, with a brief drop in cases that coincides with the COVID-19 pandemic (19).

Beyond acute disease, there are a variety of immune sequelae that can result from *S. pyogenes* infections, including acute rheumatic fever, rheumatic heart disease, glomerulonephritis, post-streptococcal reactive arthritis, and pediatric autoimmune neuropsychiatric disorders associated with streptococcal infection (PANDAS), though the legitimacy of PANDAS remains controversial (9, 17). Acute rhematic fever (ARF) is the most common cause of acquired heart disease in children worldwide (20). Though the incidence of ARF has dramatically decreased in middle and high-income settings, the incidence in low-resource settings remains high (20). A complication of untreated or recurrent ARF is rhematic heart disease (RHD), which, with its related pathologies is the cause of the majority of deaths attributable to *S. pyogenes* (467,000; 73%) (9).

Research Context

S. pyogenes can cause a diverse array of diseases which range in severity from the self-limiting to the life-threatening. Though there are recognized associations

between specific *emm*-types and tissue preference and certain virulence gene repertoires with specific disease manifestations (17, 21, 22), there is still much to learn about how *S. pyogenes* understands the tissue environment and how it connects this context to virulence regulation at the molecular level. Understanding this connection, ultimately to better develop disease prevention and treatment strategies, inspires much of the work in the McIver Lab and this dissertation.

Nutrient Sensing and Gene Regulation

Infection of a host by a bacterial pathogen requires an essential first step, colonization of the host tissue. To do so, the bacterium must successfully sense the tissue environment it has encountered and respond accordingly by modulating the expression of genes that promote its survival within the niche. Successful infection relies on the ability of the pathogen to modulate its metabolism to access the nutrients needed to survive and replicate within specific host tissues. This is supported by numerous studies that demonstrate the importance of metabolism loci in the virulence of bacterial pathogens. For example, in *Streptococcus pneumoniae*, mutants in two sucrose metabolism loci *scr* and *sus* are attenuated for fitness in nasopharyngeal colonization and virulence in the lung, respectively (23). Additionally, 9 genes for carbohydrate transport and metabolism were found in a screen for genes important for *in vivo* growth of *Mycobacterium tuberculosis* (24). In the Gram-negative pathogen *Salmonella enterica*, 4 promoters for carbohydrate uptake and utilization systems were identified as being activated in murine models of enteritis and typhoid fever (25). Additionally, in *S. enterica* serovar Typhimurium, mutants in several nutrient utilization operons were attenuated in a mouse spleen colonization model (26).

Similarly, mutants of *Salmonella typhimurium* that are unable to utilize fatty acids and citrate are avirulent in peroral infections of mice (27). These studies highlight the importance of metabolism for virulence across diverse bacterial pathogens.

The two major superficial tissue sites for *S. pyogenes* are the skin and the oropharynx, which are relatively low in glucose (28) and contain other carbohydrates from the host diet or host extracellular matrix (ECM) (**Fig. 1.2**). Blood and plasma, which are encountered during more invasive infections, are rich in glucose, proteins, and amino acids compared to the skin and the body fluids found in and near the

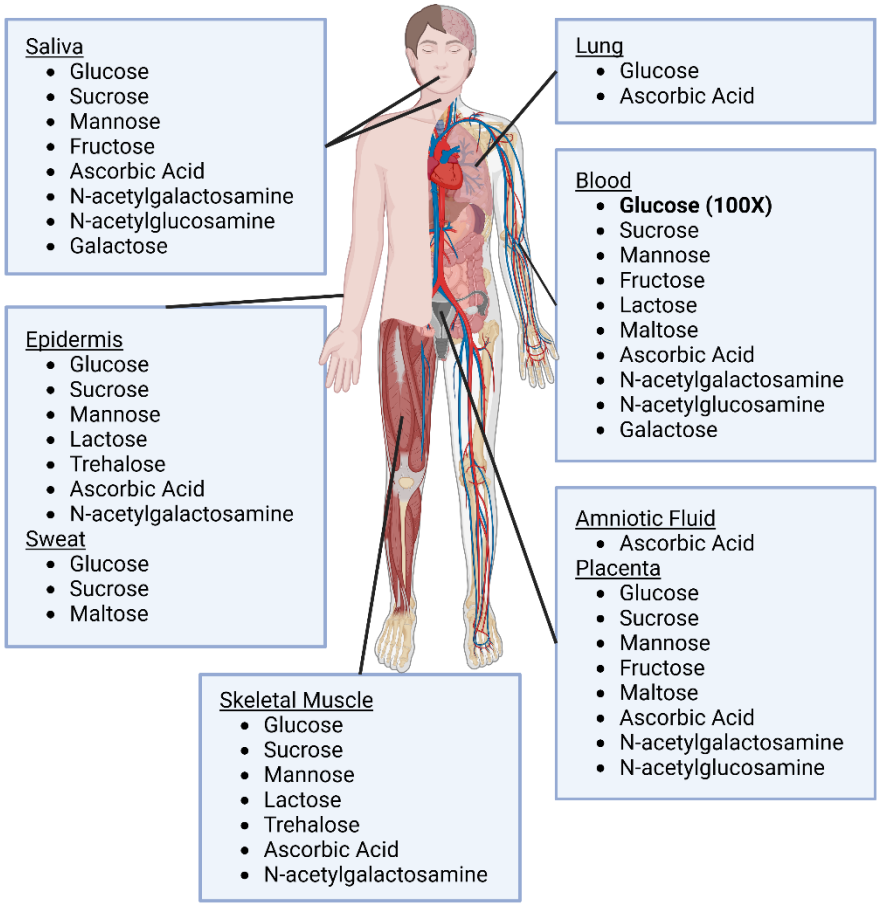


Figure 1.2: Carbohydrates in Body Compartments and Fluids
The carbohydrates present in several relevant human body compartments and body fluids according to the Human Metabolome Database. More detailed information can be found in **Appendix A1.** Created in BioRender.com.

oropharynx (28). In fact, the Human Metabolome Database (29) indicates that glucose is present at approximately 100X the level of other carbohydrates measured in the blood (**Fig. 1.2, Appendix A1**). To successfully infect these niches, *S. pyogenes* must adapt its metabolism as it encounters them.

Adaptation to a niche not only includes genes for products that enable nutrient utilization, but also those that promote infection and aid in evading the host's immune system. These co-existing requirements seem to have led to the evolution of systems that link metabolism to virulence regulation in many pathogens (30, 31). For example, mutation of a key metabolism regulator in *S. pneumoniae* attenuated its virulence in mouse models of nasopharyngeal colonization and lung infection (32). Enterohemorrhagic *Escherichia coli* uses sugar-sensing transcription factors to regulate the expression of virulence factors related to its type III secretion system (33). In *Staphylococcus aureus*, an important metabolism regulator, CcpA, plays a role in regulating biofilm formation and the expression of toxic-shock syndrome toxin 1 (34). In fact, CcpA plays an important role in virulence factor regulation in several pathogens within the phylum Bacillota (formerly Firmicutes) (35-37). These examples highlight the connection between metabolism and virulence factor regulation across diverse bacterial pathogens.

S. pyogenes Carbohydrate Metabolism

Glucose is the primary and preferred carbon source for *S. pyogenes* (28). It is a lactic acid bacterium that relies on glycolysis and fermentation for energy generation (28). Several genes for glycolysis are essential in *S. pyogenes*, highlighting the

importance of this energy generation pathway for this organism (38). *S. pyogenes* lacks a complete tricarboxylic acid (TCA) cycle, cytochromes for respiration, and most genes for the alternative glycolytic pathway (28, 39). Interestingly, it does contain some of the genes for the components of these systems whose functions are not fully understood (28). Glucose seems to primarily be transported by glucose permeases like GlcU, and then is phosphorylated by glucokinase (NagC) as opposed to being imported by the phosphoenolpyruvate transferase system, described below (40).

When glucose is low or not available, *S. pyogenes* will import alternative carbohydrates, many of which are imported using the phosphoenolpyruvate (PEP)-dependent phosphotransferase system (PTS) (28). This system is composed of the general cytosolic components enzyme I (EI) and histidine-containing phosphocarrier protein (HPr); and sugar-specific proteins called EIIs (28, 31, 41). Enzyme II (EII) proteins are either individual proteins or can be a single protein with domains designated EIIA, EIIB, and EIIC. Mannose-family transporters contain an additional protein/domain, EIID (28, 39). The import of a carbohydrate relies on a phosphorelay through the PTS proteins and ultimately onto an incoming sugar (**Fig. 1.3A**) (28). EI is autophosphorylated on a histidyl residue using PEP as a phosphate donor (41, 42). Alignment of the phosphorylated polypeptide identified from *Enterococcus faecalis* (42) to the predicted EI peptide sequence from GAS indicates that the conserved histidine (His) in *S. pyogenes* EI is 191. EI then transfers the phosphate to His-15 on HPr (41). His-15 and the surrounding residues are highly conserved in *S. aureus*, *E. faecalis*, and *Bacillus subtilis* (42). Alignment of *S. aureus* HPr to the *S. pyogenes* HPr peptide sequence demonstrates that it is also conserved in *S. pyogenes*. HPr then

transfers the phosphate onto a histidyl residue of an EIIA, which then transfers it to an EIIB on a cysteinyl residue, in most cases (41). In EIIBs of the mannose family, the EIIB is phosphorylated on a histidyl residue (41). The EIIB then finally transfers the phosphate to the incoming sugar traversing the membrane through permease EIIC/D (41). Interestingly, comparative genomic analysis of the PTS from many different bacteria suggests that the PTS originally evolved as a signal transduction system and that the sugar uptake features evolved later within select groups (30, 43).

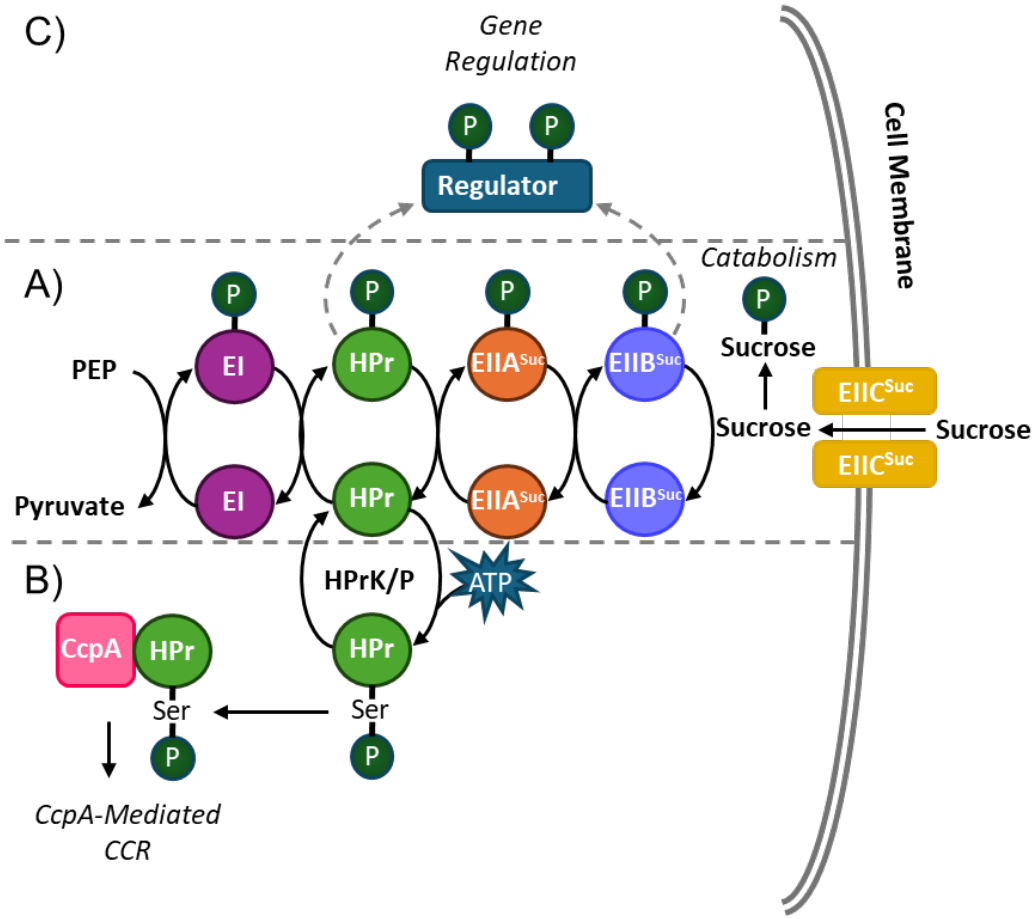


Figure 1.3: Schematic of the PTS Phosphorelay and Regulatory Mechanisms
Sucrose EII proteins are used as an example. EII proteins for other carbohydrates are outlined in Appendix A2. **(A)** Phosphorelay: phosphoenolypyruvate (PEP) is a phosphate donor for the autophosphorylation of EI. Phosphate is transferred to HPr, then sugar-specific EIIA and EIIB, and ultimately onto a carbohydrate moving through the EIIC permease. (B) Ccp-A mediated CCR: HPr is phosphorylated by HPrK at the expense of ATP. P-Ser-HPr then binds with CcpA and the complex modulates gene expression. **(C)** Another mechanism of gene regulation by the PTS is through the phosphorylation of transcriptional regulators which modulates their activity and results in modified gene expression. Adapted from Fig. 1 of reference (95).

It has been reported that *S. pyogenes* has 10-14 sugar-specific PTS systems (28, 39). Currently in the Kyoto Encyclopedia of Genes and Genomes (KEGG) (44-46), there are at least 12 complete PTS systems for maltose, sucrose, β-glucosides, lactose, cellobiose, mannose, galactose, L-ascorbate, and fructose (**Appendix A2**) (47). There are two putative PTS systems for cellobiose, mannose, and L-ascorbate. One of the two PTS systems listed for L-ascorbate, M5005_Spy1662 – Spy1664, was previously annotated as a mannitol PTS (39) and it's not clear which sugar the system transports. Some of these systems, like the trehalose PTS, may not be complete based on the KEGG pathways (47). There is evidence, however, that components from other PTS systems can substitute for the lacking genes (48), and this may be the case for the trehalose PTS based on the trehalose utilization of *S. pyogenes* (39). *S. pyogenes* also has a non-PTS uptake system for maltose/maltodextrin which consists of an ATP-binding cassette (ABC) family transporter (28, 49). The National Center for Biotechnology Information (NCBI) Position Specific Iterative Basic Local Alignment Search Tool (PSI-BLAST) (50) also indicates that *S. pyogenes* has homologs of SusT1 and SusT2, the permease proteins for a sucrose ABC transporter from *S. pneumoniae* (23) with approximately 28% and 33% identity respectively. It's important to note that while the carbohydrate import systems have distinct carbohydrates associated with them, many contribute to the utilization of several carbohydrates (39).

Carbon Catabolite Repression

Among the most well-studied of the systems linking carbohydrate metabolism to virulence regulation is carbon catabolite repression (CCR) (30, 31). CCR is present

in most free-living heterotrophic bacteria, with some exceptions, and is best known for its role in repressing the metabolic operons of secondary carbon sources (30). A well-known mechanism of CCR that regulates diauxic growth in *E. coli* is through the inhibition of adenylate cyclase activity by $EIIA^{Glc}$, which leads to lower levels of cyclic adenosine monophosphate (cAMP) (31). This leaves less cAMP available to bind to the cAMP receptor protein Crp, which when bound to cAMP, then modulates the expression of metabolism operons (31). The mechanism of CCR in the phylum Bacillota is quite different and does not rely on cAMP levels, but instead on the activity of a bifunctional kinase and phosphorylase that responds to glycolytic intermediates as discussed below.

Much of what is known about CCR in the phylum Bacillota is from studies on *B. subtilis*. The current understanding is that all species in the phylum Bacillota, including *S. pyogenes*, use the same mechanism for CCR (28, 30). In Bacillota, CCR is primarily mediated through the HPr protein from the PTS (28, 31). HPr can be phosphorylated by EI at His-15 but can also be phosphorylated by HPr kinase/phosphorylase (HPrK/P) at Serine (Ser)-46 (P-Ser-HPr) (**Fig. 1.3B**) (31, 51). Canonically, P-Ser-HPr interacts with the carbon catabolite control protein A (CcpA) allowing it to bind DNA targets known as catabolite response elements (*cre*) that are upstream of many catabolic operons for alternative carbon sources (30, 31). Interestingly, in *Clostridium difficile* HPr does not need to be phosphorylated for CCR to take place (36). CcpA/(P-Ser)-HPr binds at *cre* sites to act primarily as a repressor by blocking access to the promoter, or by sterically hindering the forward progress of RNA-polymerase (30). However, there is also evidence of carbon catabolite activation,

where P-Ser-HPr/CcpA activates the expression of genes with a *cre* site when the site is located ~56 bp or further upstream of the promoter (52, 53).

The level of P-Ser-HPr within the cell, and thus CcpA-mediated CCR is dependent on the presence and metabolism of carbohydrates. When an efficient carbon source, like glucose, is rapidly metabolized, the levels of ATP and glycolytic intermediates, like fructose-1,6-bisphosphate, are high. This activates the kinase activity of HPrK/P and leads to high levels of P-Ser-HPr (30, 31). In *S. pyogenes* lysates, phosphorylation of HPr at the serine residue was induced by glucose-6-phosphate, 2-phosphoglycerate, and to a lesser degree, fructose-1,6-diphosphate (51). In this way, the regulation activity of CcpA is directly connected to the metabolism of carbohydrates.

As mentioned, CCR also plays an important role in virulence and virulence regulation (32, 35-37, 54, 55). In *Streptococcus mutans*, CcpA was shown to be important for biofilm formation (37, 54) and was shown to play a role in regulating the expression of genes that contribute to other virulence attributes like acid tolerance, acid production, and carbohydrate storage (37). Deletion of the gene for a CcpA homolog, RegM, in *S. pneumoniae* resulted in downregulation of the capsule locus and attenuation in an intraperitoneal mouse model of infection (55) and resulted in attenuation of murine models of pneumonia and nasopharyngeal carriage (32). In *Enterococcus faecium*, *ccpA* deletion attenuated virulence in a rat model of endocarditis (35). In *C. difficile*, inactivation of *ccpA* resulted in derepression of *C. difficile* toxins in the presence of glucose (36). CcpA also plays an important role in *S. pyogenes* virulence, which will be discussed later (53, 56, 57).

CcpA-Independent Carbohydrate Metabolism Regulation

Bacillota species also have CcpA-independent regulation mechanisms for carbohydrate metabolism. For example, while dual phosphorylated P~His/P-Ser-HPr can exist, phosphorylation at P-Ser-HPr makes HPr a poor substrate for phosphorylation by EI (31, 41). In this way, P-Ser-HPr, which forms during rapid and efficient carbohydrate metabolism, acts to inhibit the PTS phosphorelay that allows for alternative sugar uptake. In contrast, under nutrient limitation or when bacteria are taking up a less efficient carbon source, inorganic phosphate accumulates due to the hydrolysis of adenosine triphosphate (ATP) and activates the phosphorylase activity of HPrK/P leading to the dephosphorylation of HPr at Ser-46 (30, 31). Dephosphorylation at Ser-46 makes HPr a better substrate for P~EI phosphorylation (31, 41). Phosphorylation of HPr at His-15 and the subsequent phosphorylation of carbohydrate-specific EII proteins primes the cell for uptake of alternative carbon sources if they become available (41). There is also evidence of inducer exclusion by P-Ser-HPr. An example of this is in *Lactobacillus casei* where P-Ser-HPr interacts with protein MalK and inhibits ATP hydrolysis, thus preventing the transport of alternative carbohydrate maltose (58).

Additionally, P~His-HPr, which accumulates when there is no efficiently utilized carbon source, can directly phosphorylate proteins to modify their function (31) (**Fig. 1.3C**). An example of this is glycerol kinase which is stimulated upon phosphorylation by HPr, allowing for utilization of glycerol as an alternative carbon source (59). P~His-HPr can also phosphorylate proteins that contain a specific domain, known as a PTS regulation domain (PRD). In classic PRD-containing regulators, this

phosphorylation will activate the regulator and induce gene transcription of genes like PTS operons (31, 41). Phosphorylated EIIB (P~EIIB) proteins can also phosphorylate proteins with PRDs and, in contrast to P~His-HPr, this phosphorylation typically has inhibitory function (31, 41). This allows for the negative regulation of carbohydrate-specific EII PTS operons when those carbohydrates are absent since P~EIIB is not transferring its phosphate to the incoming carbohydrate and can therefore phosphorylate regulators (31). In many cases, phosphorylation of the PRDs impacts the ability of the regulator to form dimers, which is important for their function (60). Some proteins also contain EII-like domains that carry conserved histidyl or cysteinyl residues that are subject to phosphorylation by the PTS proteins (41). The same pattern of HPr stimulation and EII inhibition seems to hold for transcriptional activators that contain fused EII domains (41).

There is also evidence in *Bacillus subtilis* that HPr and its homolog Crh directly interact with metabolic proteins and transcriptional regulators to modify their activity (31). An example of this is the interaction of seryl-phosphorylated HPr and Crh interacting with glycolytic enzyme, GapA, to inhibit its activity (61). Therefore, the PTS can modulate the activity of proteins in a variety of different ways to modulate the response of the cell to different metabolic stimuli.

Metabolism and Virulence Connections in *S. pyogenes*

The connection between metabolism and virulence has long been appreciated in *S. pyogenes*. As early as the 1930s Todd and Hewitt found that the specific components in the culture medium impacted the production of the streptococcal

hemolysin (62). In the 1950's, Fox and Krampitz discovered conditions that were nutritionally incomplete, yet still led to the synthesis of the M protein (63), which is a major streptococcal virulence factor (64). This demonstrated that *S. pyogenes* could respond to nutritional conditions of the media and synthesize M protein without robust growth. In the 1970s, it was discovered that the expression of the gene for M protein was activated by the presence of specific carbohydrates in the growth medium (65). Then, in 1997, Chaussee and colleagues found that a critical virulence factor, SpeB, was regulated in response to glucose being depleted in the growth media (66).

More contemporary research has further established the connection between metabolism and virulence. As in other bacterial pathogens, CcpA-mediated CCR is a major mechanism for this connection. It was determined that *S. pyogenes* has 98 potential unique *cre* sites within the genome, including in the promoter region of the gene for the global virulence regulator Mga, and associated with the genes for virulence factors (53). It was found that the CcpA protein and the *cre* site upstream of *mga* together play a role in activation of *mga* transcription (53). Further analysis of the CcpA regulon by two different groups showed that CcpA impacts the expression of not only metabolic genes, as expected, but also the expression of several well-known virulence factors in rich media (56, 57). Interestingly, there were also several virulence factors that were specifically regulated by CcpA in the nutrient limiting conditions of human saliva (57), suggesting that it plays an important role in this host fluid. Additionally, a Δ*ccpA* mutant is less virulent in a mouse model of invasive infection and less fit in an oropharynx colonization model in multiple serotypes (57, 67). In contrast, another study found that a Δ*ccpA* mutant of MGAS5005 was found to be hypervirulent in a

mouse model of invasive infection and produced larger lesions in a subcutaneous soft-tissue infection model (56). However, the findings of Shelburne and colleagues in 2016 featured diverse serotypes and seemed to solidify the role of CcpA in promoting virulence (67).

Beyond CcpA-mediated CCR, other metabolism proteins also play a role in the virulence of *S. pyogenes*. For instance, it was discovered that a paralog of aldolase, LacD.1, from the lactose utilization system functioned as a regulator of virulence factor genes in response to glucose, pH, and NaCl (68). It's also interesting to note that two glycolytic proteins, GAPDH and enolase, have virulence roles when they inexplicably localize to the cell-surface (64). PTS protein EI also plays a role in virulence, as mutation of EI (Δ*ptsI*) leads to the premature expression of streptolysin S (69). This early expression significantly impacts the tissue destruction in a soft tissue infection model in mice (69). The work by Gera and colleagues demonstrated that *ptsI* may contribute to regulating systemic dissemination and inducing the *speB* virulence factor gene (69). *ptsI* deletion also impacts the ability of *S. pyogenes* to survive in human blood (40). Additionally, a two-component regulatory system *sptR*/*sptS* is known to regulate both carbohydrate metabolism genes and virulence factors in human saliva (57). It was also found that specific metabolic pathways for aerobic versus anaerobic fermentation contribute to either growth or tissue destruction during infection (70).

Furthermore, the transcriptome of *S. pyogenes* grown in different *in vitro* and *in vivo* models have demonstrated a close relationship between metabolism and virulence (28, 71-73). For example, when exposed to human blood, 130 genes for carbohydrate metabolism were differentially regulated and 32 genes for virulence

factors were differentially regulated (71). In a murine soft-tissue infection model, carbohydrate metabolism and virulence genes are 7.5% and 5% respectively of the top 200 *S. pyogenes* RNAs detected in infected tissues (72). In a pharyngitis model of cynomolgus macaques, carbohydrate utilization genes and superantigens were upregulated in the early colonization and acute phases of infection (73).

In parallel, major virulence regulators of *S. pyogenes* have been shown to impact the expression of metabolic genes (74-76). For instance, the control of virulence regulator CovR (aka CsrR), which represses the expression of genes for important virulence factors such as capsule (*has)* and streptolysin (*sag*), also impacts the expression of 52 metabolism genes, 17 of which are carbohydrate metabolism genes (74). SptR, a regulator important for the survival in human saliva, plays a role in regulating multiple virulence factors and complex carbohydrate utilization systems (76). Additionally, the virulence regulator Mga has a regulon that includes many genes involved in fermentation or sugar metabolism (36, ~10%) and deletion of *mga* impacted growth in chemically defined medium (CDM) with individual carbohydrates as the sole carbon source (75). Interestingly, the gene for CcpA was downregulated by Mga (75) and there is an activating *cre* site upstream of the *mga* open reading frame (ORF) (53) suggesting there is interplay between these regulatory mechanisms. Overall, the two networks governing metabolism and virulence seem to be highly interconnected.

Virulence Gene Regulation in *S. pyogenes*

Virulence factor regulation is a highly complex interconnected network of different regulators that work directly and indirectly to modulate the expression of

different genes that promote infection (77-79). *Streptoccocus pyogenes* produces a wide array of virulence factors that contribute to infection (17, 80). Some of these include cytolysins like Streptolysin S and Streptolysin O, adhesins like pili and fibronectin binding protein, and immune evasion factors like capsule and C5a peptidase (**Fig. 1.4**). Unlike other organisms, *S. pyogenes* does not rely on the use of alternative sigma factors to regulate growth-phase or stress dependent gene expression (79). The major types of virulence regulators in *S. pyogenes* are two-component systems, stand-alone regulators, and small RNAs, each of which are described in brief below (77).

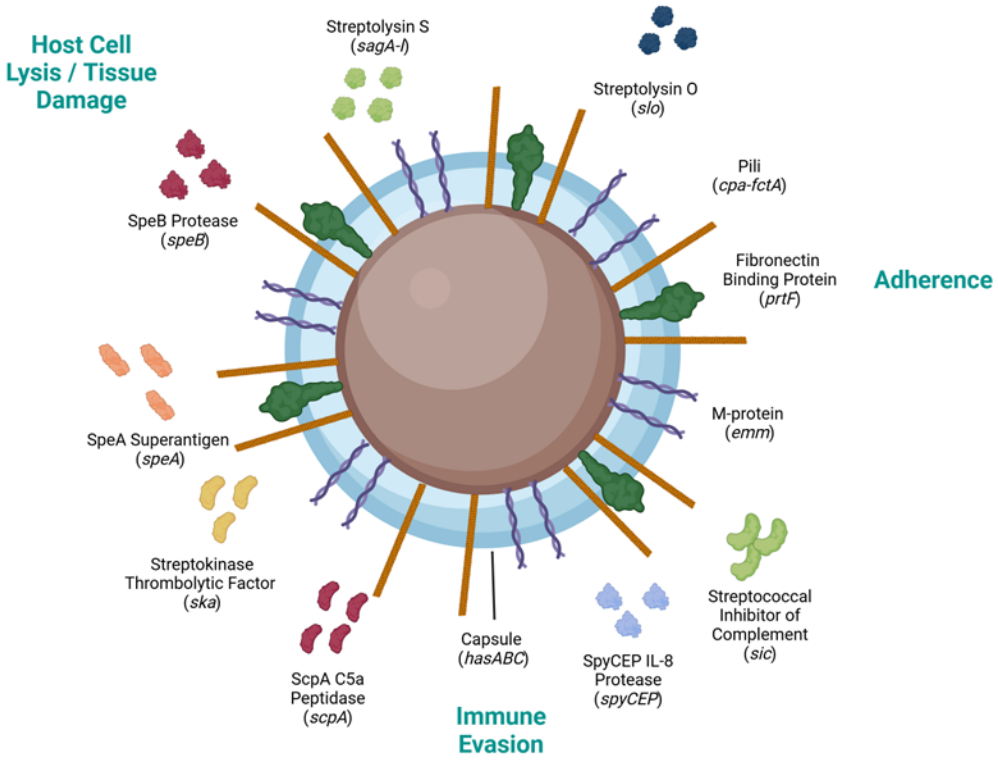


Figure 1.4: Examples of *S. pyogenes* Virulence Factors
Schematic of some important virulence factors that *S. pyogenes* employs to promote infection. This is not comprehensive.

Two-component systems (TCSs) are a well-conserved sensing and response mechanism in bacteria (77, 81). These systems contain two parts: a transmembrane

sensor kinase that recognizes some kind of extracellular signal, and a cytosolic response regulator that activates or represses gene expression (77). For canonical TCSs, upon encountering its stimulating factor the sensor kinase will dimerize and autophosphorylate (81). This autophosphorylation typically occurs at a histidine residue within its cytoplasmic domain (77, 81). This phosphate will then be transferred to an aspartyl residue on the response regulator which modifies its regulatory activity (77, 81). In the case of CovS, the sensor kinase also acts as a phosphorylase on CovR to further regulate CovR-mediated gene expression (77). *S. pyogenes* contains 14 known TCSs, which are distributed heterogeneously across strains (82). Most of the TCSs (9/14) are conserved across >90% of NCBI reference genomes (82). TCSs play diverse roles in modulating gene expression for GAS including metabolism, stress responses, virulence regulation, immune evasion, and quorum sensing (82, 83).

In contrast to two-component systems, stand-alone regulators lack a known sensory component (77). Despite this, stimuli have been identified that impact both the activity and the expression of some of these regulators (77). *S. pyogenes* has 35 known stand-alone transcriptional regulators (84). Many stand-alone regulators are highly conserved, as indicated by the fact 97% of strains carry at least 32 stand-alone regulators. and that the coding sequences of stand-alone regulators are well conserved overall (84). Stand-alone regulators mediate a variety of different activities in *S. pyogenes* that promote infection, including immune evasion, adherence, metabolism, internalization, and host cell-damage (79). Some examples of important stand-alone regulators in *S. pyogenes* include Mga, the RofA-like proteins, and Rgg (79).

Small non-coding RNAs (sRNAs) have become recognized as the most abundant class of post-transcriptional regulators in bacteria (85). sRNAs primarily act to control translation by base-pairing to target messenger RNAs (mRNAs) which can stimulate or inhibit either translation or mRNA degradation (77, 85). sRNAs impact translation by binding RNA in one of three ways. They can bind to and block the ribosome binding site, bind in a manner that induces structural change to the RNA which can either inhibit or promoter translation, or bind at a ribosome standby site to encourage translation (85). They can also act by interacting with RNA-binding proteins to modulate their activity (77). Examples of sRNAs in *S. pyogenes* include the *pel* and *fasX* RNAs which regulates the genes for important virulence factors like M-protein and the streptolysin S (85).

PRD-Containing Virulence Regulators

As mentioned, the PTS proteins play important roles in regulating metabolic activity in the cell, including by phosphorylating PRD-containing regulators (41). Interestingly, there are several virulence regulators in bacterial pathogens that have PRDs and are termed PRD-containing virulence regulators (PCVRs) (reviewed in (86)). The presence of central PRDs with histidine residues suggests that PCVR activity can be modulated by the PTS like the classic PRD-containing metabolism regulators. Of the classic PRD-containing metabolism regulators, PCVRs are most similar to the mannitol operon activator MtlR from *B. subtilis* (41), though all known PCVRs appear to be truncated (86) (**Fig. 1.5**). Their structures include two helix-turn-helix (HTH) domains, two central PRDs, and a C-terminal EIIB-like domain that resembles the EIIB

for galactitol (EIIB^{Gat}) (**Fig. 1.5**). The two most well-studied members of this family are Mga from *S. pyogenes* and AtxA from *B. subtilis*. The RofA-like proteins (RALPs), another family of virulence regulators in *S. pyogenes*, and virulence regulators from other Gram-positive pathogens also have this structure (86, 87). The following sections will discuss the traits of various PCVRs and the similarities and differences between members (**Table 1.1**).

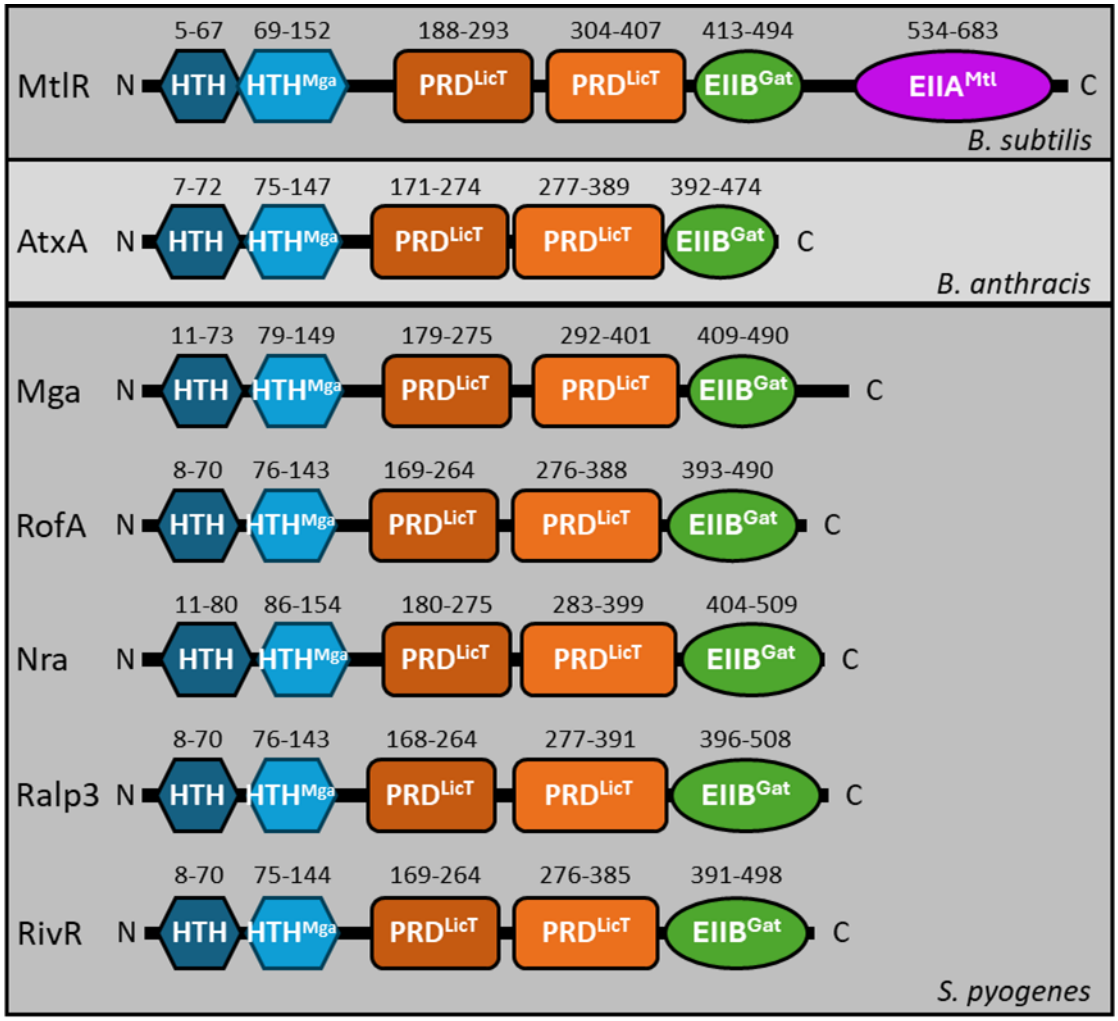


Figure 1.5: Schematic of the Domain Structure of MtlR from *B. subtilis* and PCVRs from *B. anthracis* and *S. pyogenes*.
PCVRs contain two N-terminal helix-turn-helix domains (blue hexagons), two internal PRDs (orange rectangles) and a C-terminal EIIB^{Gat} domain (green circle). This domain structure resembles the N-terminal end of MtlR regulator from *B. subtilis*, but MtlR contains a second EIIA-like domain resembling the EIIA of the mannitol PTS (EIIA^{Mtl}).

Mga

The multiple gene regulator of Group A *Streptococcus* (Mga) is ubiquitous stand-alone regulator in *S. pyogenes* that regulates over 10% of the genome, including genes for metabolism operons and major virulence factors (75). There are two alleles of *mga*, *mga-1* and *mga-2*, which correlate with M-protein class and tissue tropism (75, 88). A strain of *S. pyogenes* will typically either carry a *mga-1* or a *mga-2* allele at the *mga* locus. There is only one known case of a strain carrying both *mga-1* and *mga-2* alleles (84). The expression of *mga* is growth-phase dependent, positively autoregulated, and responds to CO_2 (75, 89-91). Mga is most active in exponential phase, a phase analogous to post-colonization replication *in vivo*, and acts to activate genes that contribute to adhesion and immune evasion, while also modulating metabolism (75, 91, 92). In Mga, the second HTH domain designated HTH^{Mga} is essential for DNA binding, while the first HTH domain only partially contributes to binding (93). There is evidence that Mga has a defined binding box in which Mga can exert direct transcriptional control (75), however, mutational analysis of the Mga binding sequence of 3 Mga-regulated promoters also indicates there is significant variability in how Mga interacts with DNA (94).

Secondary structure predictions of Mga using the Phyre2 software demonstrates that there are 3 histidine residues, H204 and H270 in PRD-1 and H324 in PRD-2, that are closely aligned with the histidines from LicT, LevR, and MtlR (86). In an *in vitro* phosphorylation assay with the general cytosolic proteins EI and HPr, Mga can be phosphorylated and phosphoablative mutations (H>A) at these residues significantly reduced *in vitro* phosphorylation (95). Additionally, phosphoablative and

phosphomimetic (H>D) mutations at these residues impacted the direct transcriptional activity of Mga and Mga-dependent gene expression (92, 95). Interestingly, PRD-1 has a greater impact on Mga activity, but phosphorylation does not follow the canonical activation and repression pattern for most PRD-containing regulators (92, 95). Instead of being activated, Mga is inactivated when phosphomimetic residues are in PRD-1 (H204D and H270D) (92, 95). Importantly, both phosphoablative or phosphomimetic mutations within PRD-1 attenuate virulence of *S. pyogenes* in a mouse model of subcutaneous infection (95), suggesting that these residues are critically important to Mga function.

Mga is also subject to phosphorylation *in vivo* when grown in a high-glucose rich medium (92, 96). This phosphorylation is reduced when *S. pyogenes* is grown in a low-glucose high peptide medium and the Mga regulon is differentially expressed under these conditions (96). These studies demonstrate that Mga can be phosphorylated by EI and HPr, that histidine residues within the PRDs contribute to its activity as a transcriptional regulator, and that phosphorylated Mga is observed *in vivo* in a glucose dependent manner, which may impact its activity. Together, these studies suggest that Mga is subject to PTS regulation in response to carbohydrates.

Mga readily undergoes multimerization *in vitro* due to changes in osmolarity and pH and the presence of divalent cations (97). The $EIIB^{Gat}$ domain is important for this multimerization, as deletion of this domain resulted in a protein that did not significantly oligomerize *in vitro* (97). Elimination of the $EIIB^{Gat}$ domain also eliminated Mga transcriptional activity *in vivo* but did not impact the ability to bind DNA *in vitro* (97), suggesting that multimerization is important to its ability to act as a

regulator. Phosphomimetic mutations of H204 and H270 at PRD-1 prevented multimerization and had decreased levels of transcriptional activity (95), suggesting that phosphorylation of Mga within PRD-1 inhibits Mga by inhibiting its ability to form multimers.

AtxA

AtxA is the anthrax toxin activator from *B. anthracis*. It activates the expression of virulence factors and toxin systems including capsule (*capBCADE*), protective antigen (*pag*), lethal factor (*lef*), and edema factor (*cya*) and is essential for virulence in mouse models of infection (98, 99). AtxA is also found in *Bacillus cereus* strain G9241, which has two alleles of the regulator and loss of both attenuates virulence in mouse models of infection (100).

In AtxA, mutational analysis of histidine residues within the PRDs demonstrated that phosphoablation (H>A) at H199 and H379 completely eliminated the ability to immunoprecipitate AtxA in its phosphorylated form (101). Additionally, phosphomimetic mutations (H>D) at those residues caused activation and repression, respectively (102). Phosphomimetic and phosphoablative mutations in the PRDs also impacted the ability of AtxA to bind the *pag* promoter and impacted *pag* expression (103). McCall and colleagues found that an H199D mutation enhanced binding to the *pag* promoter *in vitro*, while a H199A mutation reduced binding (103). They also found that an H379D mutation decreased the expression of *pag in vivo* and that a stem-loop structure may be important for binding. These findings together demonstrate that AtxA is phosphorylated, and that phosphorylation impacts activity. However, AtxA does not seem to be phosphorylated *in vitro* by purified EI and HPr, and AtxA does not have an

altered ability to express toxin genes in a mutant in which the general PTS proteins EI and HPr are deleted (104). These findings would suggest that while AtxA is phosphorylated *in vivo* at H199 and H379, phosphomimetic and phosphoablative mutations impact DNA binding and AtxA-dependent transcription, the phosphorylation may not be mediated by the PTS.

AtxA can form multimers *in vivo*, but the most abundant multi-protein form is dimers (102). The C402 residue within the EIIB domain was essential for multimer crosslinking, suggesting the EIIB domain is important for oligomerization (102), consistent with Mga. Also consistent with Mga, phosphomimetic mutation in PRD-2 (H379D) reduced the capacity of AtxA to form homodimers (105). The inability to multimerize may contribute to the repressive effects observed by this mutation (102). Glucose has been shown to impact the Atx-dependent expression of *pagA*, however, this was believed to be due to CcpA-mediated regulation of *atxA* transcript and not through PTS modification of AtxA (106). Additionally, *atxA* differential regulation in the presence of glucose was challenged by a report that did not find any glucose-dependent impact on *atxA* transcription (104). Overall, these findings show that AtxA is phosphorylated *in vivo,* and this impacts activity, but the PTS may or may not be playing a role in this phosphorylation.

RofA

Regulator of F (RofA) was identified in serotype M6 strain JRS4 as a trans-acting activator of *prtF*, the gene for fibronectin binding protein, protein F (107). The *rofA* gene resides in the same locus as *prtF* and is known as the fibronectin-binding, collagen-binding, T-antigen (FCT) locus (108) (**Fig. 1.6**). There are nine different

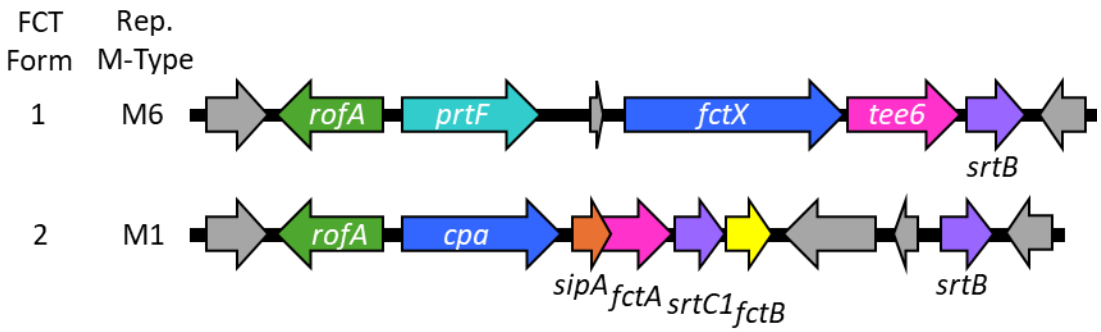


Figure 1.6: Fibronectin-binding, Collagen-binding, T-antigen (FCT) Locus.
Schematic of the genomic organization of the FCT region where *rofA* resides. Figure is adapted from figure 1 of reference 108. The major pilin subunit is shown in pink (*tee6* and *fctA*). The minor pilin subunits are shown in blue and yellow (*fctX*, *fctB*, and *cpa*). Sortases are shown in purple (*srtC1* and *srtB*). Genes and gray are those that are not known to be associated with pilin biogenesis.

forms of the FCT locus in *S. pyogenes*, which are outlined in the 2021 review on *S. pyogenes* pilin by Nakata and Kriekemeyer (109). Figure 1.6 shows the first two of these, FCT-1 and FCT-2. These are shown because *rofA* was discovered in an FCT-1 strain, and because the strain used in this work is M1T1 5448, which is an FCT-2 strain. The *rofA* gene is positively autoregulated (110-112) and *rofA* expression is stimulated by reduced oxygen tension in a serotype M6 strain (110). RofA was found to contribute to the expression of several virulence genes and adhesins, including those for the streptolysin S operon (*sag*), superantigen SpeA (*speA*), M-protein (*emm*), streptokinase (*ska*), the capsule operon (*hasABC*), and pilus operon (*cpa-sipA1-tee1-srtC1-fctB*) in a serotype-dependent manner (112-114). Disruption of *rofA* in an M6 strain resulted in 40% reduced adherence to HEp2 cells and induction of significant amounts of HEp2 cell death compared to the wild-type (WT) strain (112). Disruption of *rofA* in an M2 strain resulted in the opposite, an increase in adherence to host cells (113). These studies demonstrate that RofA impacts the expression of adhesion proteins and

virulence factors that can contribute to colonization of host tissues at the beginning stages of infection, but that this regulation is serotype dependent.

There have been no direct studies investigating the impact of *rofA* on virulence in animal models. However, there have been some interesting findings from transposon mutant library screens in two different *in vivo* infection models (115, 116). In a high throughput transposon-sequencing (Tn-seq) screen for genes important for subcutaneous soft-tissue infection, disruption of *rofA* significantly increased fitness (115). This should be taken with caution, as point mutations in an important virulence factor repressor system, *covRS*, can occur and lead to false positives for genes that increase fitness (115). It is also possible, however, that increased fitness is due to decreased pilus production, which is a toll-like receptor-2 (TLR-2) agonist (117). Another Tn-seq screen identified *rofA* as a gene that contributed to fitness of an M1T1 strain in the primate oropharynx, but not for an M28 strain in the same model (116). It should be noted that these two serotypes have different forms of the FCT locus (109), and that M28 is particularly associated with puerperal sepsis (118), which may impact the role of RofA. The finding that the M1T1 strain is important for pharyngitis is interesting because M1T1 serotype *S. pyogenes* have an A-C *emm* pattern and this *emm*-pattern is associated with throat tissue tropism (119).

No studies were found that have investigated the phosphorylation of RofA within its PRD domains, or its ability to multimerize, two of the shared features of Mga and AtxA. Protein primary sequence shows that there are histidine residues within the predicted PRDs, so these may be subject to PTS phosphorylation. Maltose-binding protein (MBP)-tagged RofA was shown to bind DNA in the intergenic region between

rofA and *prtF* of M6 strain JRS4 through electrophoretic mobility shift assays (EMSAs) (110). Further characterization in M6 strain JRS4 revealed that RofA binds at two sites: a smaller site proximal to *rofA* and a larger site proximal to *prtF* (111). The larger binding site played a major role in activating expression of both *prtF* and *rofA,* while the small site only contributed in a minor role to *rofA* expression (111). Interestingly, similar binding boxes were not found preceding the secreted virulence factors whose expression was impacted by RofA in M6 2/66 (112). This could suggest alternative binding boxes, indirect regulation through a different protein, or sequence-independent binding.

RofA-Like Proteins (RALPs) in *S. pyogenes*

S. pyogenes contains three other Mga paralogs: Nra, Ralp3 and RivR. These paralogs are referred to as RofA-Like Proteins (RALPs). RALPs are on average 52% similar (111) and their presence and combination vary with serotype (84, 113, 120, 121). Nra is found in the same locus as RofA, and strains only carry one of these regulators (84, 109). Expression of *nra* was found to be negatively autoregulated (120), but *rivR* was not autoregulated, at least in the conditions tested (122). In some strains, expression of *nra* is temperature dependent (123). Each of the paralogs have been shown to contribute to the expression of virulence factors and that contribute to fitness, and impact the expression of other virulence regulators, like *mga* (112, 113, 120, 122, 124-127). Like mutants of *rofA*, mutants with disrupted *nra* cause more significant eukaryotic cell death when incubated with HEp2 cells (124), which may suggest that these two limit host cell damage to allow for prolonged persistence. Additionally, deletion of *ralp3* and *rivR* attenuated virulence in mouse models of infection (125, 127).

Interestingly, the *ralp3* coding sequence in the strain used by Kwinn and colleagues has a nonsense mutation halfway through the gene, suggesting this version of Ralp3 is truncated and lacks the second PRD and the $EIIB^{Gat}$ domain.

All the RALPs contain the key structural features of PCVRs including N-terminal DNA-binding domains, central PRDs and a C-terminal EIIB domain (**Fig. 1.5**). Phyre2 analysis shows that each of the RALP PRDs contain histidine residues that may be subject to PTS phosphorylation, including the H392 residue of Nra which aligns with H379 of AtxA (86). No studies were found that have demonstrated phosphorylation of any of the RALPs. The expression of the *nra* gene, however, has been tied to glucose availability in the media, as maximum expression from the *nra* promoter was achieved when high glucose medium was mixed with low glucose medium at a 1:2 ratio (120). Given that Nra is autoregulated, this may suggest that glucose availability, potentially sensed through PTS, plays a role in Nra activity.

The other feature common to Mga and AtxA is the ability to multimerize (97, 102). Multimerization has only been studied for RivR, which showed evidence of multimer formation through cross-linking experiments (128). This cross-linking was dependent upon C470 which resides in the $EIIB^{Gat}$ domain of RivR, suggesting that this domain is important for multimerization (128). Furthermore, mutation at C470 to serine inhibited RivR from regulating its target genes (128). These findings suggest that like Mga and AtxA, RivR forms multimers through its $EIIB^{Gat}$ domain, and that this multimerization is important for its regulatory function. It's reasonable to hypothesize that this is the case for the other RALPs in *S. pyogenes* as well.

PCVRs in Other Organisms

As mentioned, putative PCVRs can be found in several Gram-positive organisms. Within Streptococci, there are many species that contain putative PCVRs and the number of PCVRs seems to correlate with pathogenicity (86). Well-studied putative PCVRs include *S. pneumoniae* Mga homolog, MgaSpn (aka Mgr) (129). MgaSpn regulates the genes within the *rlr* pathogenicity islet (129) and contains histidine residues within its putative PRDs (86). Some of these histidines align to histidine residues that are involved in modulating protein activity in AtxA, Mga, and MtlR (86). Like Mga, MgaSpn forms multimers in solution during purification, and forms multimers on linear DNA (130) . There has been no evidence of MgaSpn phosphorylation, however, the product of a Mga*Spn* targeted gene, phosphorylcholine, increased with increasing glucose concentration in a Mga*Spn*-dependent manner (131), suggesting that MgaSpn may respond to glucose to impact gene expression. Of note, some strains of *S. pneumoniae* also contains a RofA/Nra homolog RlrA (132). Like MgaSpn, RlrA regulates the genes within the *rlr* pathogenicy islet (133) and was found to be important for nasopharyngeal colonization and lung infection (133).

In addition to AtxA, *B. anthracis* contains two additional AtxA homologs AcpA and AcpB. There is evidence that both AcpA and AcpB contribute to capsule regulation (134-136) and that the regulators contribute to *in vivo* virulence (137, 138). AcpA and AcpB contain histidine residues within their PCVRs, some of which align to important histidines in Mga, AtxA and MtlR (86). Like Mga and AtxA, AcpA and AcpB both form homomultimers and this multimerization is dependent upon the $EIIB^{Gat}$ domain

(136). Deletion of the EIIB^{Gat} domain also abolishes AcpB and AcpA-dependent activation of the capsule genes (136).

Table 1.1: Traits of PCVRs

Trait	**Mga**	**AtxA**	**RofA**	**Nra**	**Ralp3**	**RivR**	**Mga-Spn**	**AcpA**	**AcpB**
Auto-regulated	(75)		(111, 112)	(120)		(122)			
Forms Multimers	(97)	(102)				(128)	(130)	(136)	(136)
P~ylated *in vitro* by EI/HPr	(95)								
P~ylated *in vivo*	(92, 96)	(101)							
Regulates Carb. Metab. Genes	(75)	(136)				(122) (127)	(131)	(136)	(136)
Impacts Virulence in Animal Models	(139)	(98, 99)			(125)	(127)	(129)	(138) (137)	(138) (137)
Maximal Express.	EX (89)			EST (120)	EX (125)				

Abbreviations:
"P~ylated" = Phosphorylated
"Carb." = Carbohydrate
"Metab" = Metabolism
"Express." = Expression
"EX" = Mid-exponential
"EST" = Early stationary

Green indicates that the trait has been shown. Yellow indicates 'inconclusive' or 'supportive, but not strong' results. Red indicates that it's been shown that the protein does not have that trait. References are provided in box by number.
MgaSpn was previously referred to as MgrA.

Chapter 2: Materials and Methods

Overview

The goal of this work was to 1) begin characterizing RofA because it is believed to be one mechanism the pathogen *S. pyogenes* utilizes to adapt to host tissue niches, and 2) build a tool to generate bioluminescent bacteria for *in vivo* infection monitoring. All the work was conducted in serotype M1T1 strain 5448 because it is a representative strain of the serotype which is globally disseminated and capable of causing severe invasive disease (see 'Bacterial Strains and Media').

Characterization of RofA was achieved through the investigation of the RofA regulon through RNA-Sequencing and RT-qPCR (See 'RNA Isolation', 'RNA Library Preparation, Sequencing and Data Analysis', 'Endpoint RT-PCR and RT-qPCR'). Then, the phenotypic consequences of differential gene regulation were explored through assays assessing the capsule production and adherence to primary human keratinocytes (See 'Capsule Assay' and 'Adherence Assay'). Additionally, the phosphorylation of RofA was explored through *in vitro* PTS phosphorylation assays with EI and HPr to determine if RofA can be directly phosphorylated by these PTS components (See '*In vitro* PTS Phosphorylation Assays'). *In vivo* phosphorylation of RofA was investigated through Phos-Tag electrophoresis and western blotting (See '*S. pyogenes* Lysis and Phos-Tag Gel Electrophoresis' and 'Western Blotting').

Generation of bioluminescent bacteria was achieved through constructing an allelic exchange plasmid that can be utilized for any *S. pyogenes* strain (See 'pCRS-tsf-lux-pepO' under 'Plasmid Construction'). The use of this construct was demonstrated

by generating luminescent 5448 (See '5448_lux Strains' under 'Mutant Generation by Allelic Exchange'). The growth and luminescence of these strains was evaluated to benchmark the strains against a previously generated strain (See 'Luminescent Growth Curves'). Additionally, the fitness of the strain against human immune cells was assessed (See 'Opsonophagocytic Killing Assay with Primary Human Neutrophils').

Bacterial Strains and Media

All strains utilized in this study are outlined in **Table 2.1**. The *Streptococcus pyogenes* strains 5448 and MGAS5005 were isolated from patients with invasive GAS infections and are representative of the globally disseminated M1T1 serotype (140, 141). 5448 served as the wildtype (WT) parent strain for all mutants. Strain Xen20 was acquired from Caliper Life Sciences (formerly Xenogen). *S. pyogenes* strains were routinely cultured on blood agar plates (BAP) comprised of tryptic soy agar (TSA) + 5% sheep blood (VWR) or Todd-Hewitt Broth (Alpha Biosciences) plus 0.2% yeast extract (THY) plates with appropriate antibiotic. *S. pyogenes* strains were grown in liquid culture in THY broth or C medium (142) in sealed tubes without shaking. Growth was routinely observed by measuring absorbance with a Klett-Summerson Colorimeter (Klett) or a GenSys30 spectrophotometer (Thermo Scientific) at 600 nm (OD_{600})

E. coli strains DH5α, NEB-10beta (New England Biolabs [NEB]), or GB5-alpha (Gold Bio) were used for cloning steps. *E. coli* strain GM48 was used when necessary to isolate plasmid without Dam/Dcm methylation. *E. coli* strains C41[DE3], JM109, and BL21(DE3)-Gold were used for the overexpression of proteins. *E. coli* was

routinely grown in Luria-Bertani (LB) broth (Fisher) for plasmid cloning or autoinduction medium ZYP-5052 (143) for protein expression.

Antibiotics were used as required for plasmid maintenance at the following concentrations: spectinomycin (Spec, Sp) was used at a concentration of 100 µg/mL for both *S. pyogenes* and *E. coli*, ampicillin (Amp, Ap) or carbenicillin (Carb) were used at a concentration of 100 µg/mL for *E. coli*, and kanamycin (Kan, Km) was used at a concentration of 50 µg/mL for *E. coli* and 300 µg/mL for *S. pyogenes*.

Table 2.1: Bacterial Strains

Strain	Relevant genotype or description	Reference or Source
E. coli		
DH5α	*hsdR17 recA1 gyrA endA1 relA1*	(144)
NEB 10-beta	*Δ(ara-leu) 7697 araD139 fhuA ΔlacX74 galK16 galE15 e14 - ϕ80dlacZΔM15 recA1 relA1 endA1 nupG rpsL (Str^R) rph spoT1 Δ(mrr-hsdRMS-mcrBC)*	NEB
GB5-alpha	F$^-$ φ80*lac*ZΔM15 Δ(*lac*ZYA-*arg*F)U169 *rec*A1 *end*A1 *hsd*R17(r_K^-, m_K^+) *poa sup*E44 λ$^-$*thi*-1 *gyr*A96 *rel*A1	Gold Bio
C41[DE3]	*F$^-$ ompT hsdSB* (rB- mB-) *gal dcm* (DE3)	(145)
GM48	*thr leu thi lacY galK galT ara tonA tsx dam dcm supE44*	(146)
JM109	e14-(McrA-) recA1 endA1 gyrA96 thi-1 hsdR17 (rk-mk+) supE44 relA1 Δ(lac-proAB) [F'traD36 proAB lacIqZΔM15]	Promega
BL21[DE3]-Gold	E.coli B: F- ompT hsdS(rB-mB-) tet gal endA hte lambda(DE3) T7 RNA polymerase	(147)
S. pyogenes		
5448	M1T1	(140)
MGAS5005	M1T1, *covS*	(141)
5448 Δ*rofA*_sc	5448 with markerless stop codon mutation in *rofA* coding sequence	This study
5448 *rofA*_R	5448 allelic exchange revertant in *rofA* coding sequence	This study
5448 Δ*ptsI*	Δ*ptsI* mutant in strain 5448. *ptsI* replaced with *aad9* (Sp^r).	(69)
Xen20	M49 591 *spy0535::*Tn4001-*luxABCDE* Km^r	Caliper Life Sciences, (148)

DNA Manipulations

All primer oligos used are listed in **Table 2.2** and were synthesized by Integrated DNA Technologies (IDT). Plasmid DNA was extracted from *E. coli* using either the Wizard Plus SV miniprep system (Promega) or the Qiagen Plasmid Midi Kit (Qiagen). DNA fragments were purified with the Wizard SV gel and PCR Cleanup system (Promega). Polymerase chain reaction (PCR) for cloning was conducted using either Accuprime Pfx (Life Technologies), Phusion polymerase (NEB), or Q5 (NEB) using the manufacturer's recommended protocols. PCR for diagnostic assays was performed with Standard Taq DNA Polymerase (NEB). Genewiz from Azenta Life Sciences performed all DNA sequencing. The Master-Pure complete DNA and RNA purification kit for Gram-positive bacteria (Epicentre) was used to extract genomic DNA (gDNA). All restriction enzymes used are from New England Biolabs unless otherwise indicated. Antarctic phosphatase (NEB) was used for dephosphorylation. Ligations were conducted with T4 DNA Ligase (NEB) using standard protocols.

Table 2.2: Oligos

Name	Sequence (5'→ 3')	Reference
Cloning Primers		
M1_MGAS5005_RofA _NdeI	GTTCATATGTTGATAGAAAAATACTTGGAAT	This study
M1_MGAS5005_RofA _XhoI	TTTCTCGAGTGTTAATTGCTTGGTTAAATCAG CTTGGAATTT	This study
EcoRI_rofA_F	GCGAATTCTTCACCAAATTAGTGTTTTTGAAA ATG	This study
rofA_5'_R	CTCGCAAAAGAATTATTAGCGG	This study
rofA-FLAG_3'_F	CCGCTAATAATTCTTTTGCGAG	This study
OE_rofA-FLAG_R	GACATGGCGCTTACTTGTCGTCATCGTCTTTG	This study
rofA-term_F	CGACAAGTAAGCGCCATGTCACCAC	This study
NcoI_rofA-term_R	ACGAATCCATGGATAGACAGTAAAAAATCCC AGAGAATG	This study
SDM_RofA-FLAG_F2	TTACTTGTCGTCATCGTCTTTGTAGTCG	This study
SDM_SalI_rofA_R	GTCGACGCGCCATGTCACCACATTGCG	This study

Table 2.2 continued…

Name	Sequence (5'→ 3')	Reference
EcoRV_rofA-His_F	CCATCCAAAAAAGGATGATATCTTGCC	This study
SalI_rofA-His_R	ACATGTCGACCAGCCGGATCTCAGTGGTGG	This study
oRofA_SOW1	CCCGAATTCTCTAAGTTAACCCTTAAATCGTG TGC	This study
oRofA_SOW2	GGATCCTCATTATCATGATGATTCC	This study
oRofA_SOW3	TGATAATGAGGATCCCAGTTAATTG	This study
oRofA_SOW4	CCCGAATTCCGAAAATTGATGCCGCTTCC	This study
JRS525_OE-PCR_F1	tcctcctcactattttgattagtacc	This study
JRS525_OE-PCR_R1	AGGGCTTTTCGAATTggtcgattttcgttcgtgaatac	This study
JRS525_OE-PCR_F2	GACAATGATGTTGGAaaagcccgccgaagcggg	This study
JRS525_OE-PCR_R2	AGCTTggatccacgatatcctgcagg	This study
P23_OE-PCR_F	gaacgaaaatcgaccAATTCGAAAAGCCCTGACAAC CC	This study
P23_OE-PCR_R	cttcggcgggcttttCCAACATCATTGTCATTCATATTT TTC	This study
EcoRI_luxA_F2	gtatGAATTCAGGaggactctctatgaaatttggaaac	This study
BamHI_luxE_R2	atgaGGATCCTTAACTATCAAACGCTTCGGTTAa gc	This study
luxA_Seq_F	catgactcaattaaagcgaaagag	This study
luxB_Seq_F	attttagggtttagtgattgcg	This study
luxC_Seq_F	atttcattcattattaacggccag	This study
luxC_Seq_R	ttcttccaggaatcgagtctg	This study
luxD_Seq_R	tcatcttgcttgacccaattatc	This study
M13F	GTAAAACGACGGCCAG	Genewiz
M13R	CAGGAAACAGCTATGAC	Genewiz
oAxtsf.1	tattGGATCCAACTTCTTCGTGAAAAAGG	This study
oAxtsf.2	TATTGGTGTCCGTTTTGAC	This study
oAxpepO.1	AGGAAGCAAAAAACTGATC	This study
oAxpepO.2	attaGGATCCAAAGCAGCCTTTTATTTGG	This study
oAxP23lux.1	GTTGTGTCAAAACGGACACCAATACaatcgaccA ATTCGAAAAG	This study
oAxP23lux.2	CACAGGATCAGTTTTTTGCTTCCTCTTAACTA TCAAACGCTTCG	This study
tsf_seq_F	GGGGTTATCTCAGTTGTCG	This study
tsf_seq_R	CCTTCAGCAGCAAGTTCAG	This study
oAX1781V1	CTGGTGCTGGCGTTATGGAC	This study
oAX1782V2	GCGCTCTTTCAGGAACACCT	This study
Diagnostic Assay Primers		
LicT_RevTrans_F	AAACCATAATGCGGCGATTTC	(149)
LicT_RevTrans_R	TAAGGTGGTAGTTGTACCGC	(149)
RT-qPCR and RT-PCR Primers		
gyrA_RT_L	CGACTTGTCTGAACGCCAAAG	(75)
gyrA_RT_R	ATCACGTTCCAAACCAGTCAAAC	(75)
rofA_M1_RT_L	CGAAGAGTGGATGGCCAAAC	(75)
rofA_M1_RT_R	CTCGACATAGTGGCAAAAAAGATG	(75)
cpa_M1T1_RT_L	CAGACTCTGAAGATGGTGGGA	This study
cpa_M1T1_RT_R	AAGAGGTCACGACCTGCAAT	This study
rofA-FLAG_F2	AAGGAGGUGATATACATATGATAGAAAAATA CTTGGAATCATC	This study

Plasmid Construction

All plasmids used and generated in this study are described below and summarized in **Table 2.3.**

Table 2.3: Plasmids

Plasmid	Description	Reference or Source
pET21a	(*bla*; Ap^r)	Novagen
pJRS525	*E. coli*/*S. pyogenes* shuttle vector. Replicating origin for both with spectinomycin resistance. (*aad9*; Sp^r)	(89)
pCRS	Temperature sensitive *E. coli*/*S. pyogenes* shuttle vector with ColEI origin and thermosensitive pWV01 origin. (*aad9*; Sp^r)	(150)
pET21a_rofA1	Expression vector pET21a with the *rofA* coding sequence from M1T1 strain MGAS5005 in frame with the 6x-His-tag. (*bla*; Ap^r)	This study
pProfA1-FLAGt	*rofA* coding sequence from M1T1 strain 5448 under control of its native promoter with a C-terminal FLAG-tag in pJRS525 backbone. (*aad9*; Sp^r)	This study
pUC19_rofA-HqA	Gene Universal generated *rofA*-HqA allele in pUC19 (*bla*; Ap^r)	Gene Universal/This Study
pET21a_rofA1-HqA	Expression vector pET21a with the *rofA* coding sequence from M1T1 strain 5448 in frame with the 6x-His-tag and alanine substitution mutations at residue 210, 278, 330 and 380. (*bla*; Ap^r)	This study
pProfA1-HqA-FLAGt	Complementation construct with *rofA* coding sequence from M1T1 strain 5448 with alanine substitution mutations at residues 210, 278, 330 and 380 under control of its native promoter with a C-terminal FLAG-tag. (*aad9*; Sp^r)	This study
pKSM357	Expression vector for His_6-EI (*bla*; Ap^r)	(95)
pKSM712	Expression vector for His_6-PtsH a.k.a. HPr (*bla*; Ap^r)	(56)
pKSM879	Expression vector for His_6-PEPCK (*bla*; Ap^r)	(95)
pCRS_rofA_stop	Mutagenic plasmid for allelic exchange to replace the WT *rofA* allele with the stop-codon mutant allele. (*aad9*; Sp^r)	This study
pProfA1-Hist	Complementation construct with *rofA* coding sequence from M1T1 strain 5448 under control of its native promoter with a C-terminal 6x-His-tag. (*aad9*; Sp^r)	This study
pKSM201	*E. coli*/*S. pyogenes* shuttle vector. Replicating origin for both with Kan resistance (*aphA3*; Km^r)	(56)

Table 2.3 continued…

Plasmid	Description	Reference or Source
pProfA1-FLAGt_Km	Complementation construct with *rofA* coding sequence from M1T1 strain 5448 under control of its native promoter with a C-terminal FLAG-tag in pKSM201 backbone (*aphA3*; Km^r)	This study
pProfA1-FLAGt_SalI_ term	Plasmid pProfA1-FLAGt with an added SalI cut site downstream of the stop-codon, and before the predicted terminator (*aad9*; Sp^r)	This study
pOri23	*ermAM ori* ColE1 *P23* (Ery^r)	(151)
pP23-Spec	P23 promoter inserted downstream of lacZα fragment and MCS of pJRS525. (*aad9*; Sp^r)	This study
pP23-luxABCDE-Spec	*luxABCDE* inserted downstream of P23 in pP23-Spec (*aad9*; Sp^r)	This study
pCRS-tsf-lux-pepO	pCRS with allelic exchange cassette to insert *luxABCDE* into the *tsf* - *pepO* intergenic region. (*aad9*; Sp^r)	This study

pCRS_rofA_stop

5448 gDNA was amplified by primer pairs oRofA_SOW1/oRofA_SOW_2 and oRofA_SOW3/oRofA_SOW4. These PCR fragments were joined by overlap-extension PCR (OE-PCR) using oligos oRofA_SOW1 and oRofA_SOW4. The combined fragment was cloned into plasmid pCRS by restriction cloning using EcoRI to form plasmid pCRS_rofA_stop.

pET21-rofA1

The *rofA* coding sequence was amplified from MGAS5005 (NCBI Genome Accession: CP000017) gDNA using primers M1_MGAS5005_RofA_NdeI and M1_MGAS5005_RofA_XhoI. The coding sequence of *rofA* in MGAS5005 is identical to the coding sequence in 5448 (NCBI Genome Accession: NZ_CP008776). This amplicon was digested with XhoI and NdeI and ligated into XhoI/NdeI digested and dephosphorylated pET21a(+) to generate plasmid pET21-rofA1.

pProfA1-FLAGt

The *rofA*-FLAG gene was synthesized by Gene Universal using the coding sequence for *rofA* from 5448. Two codons for glycine followed the last amino acid codon of the *rofA* gene and the sequence for a 1x-FLAG tag (GAC TAC AAA GAC GAT GAC GAC AAG) followed the 2-glycine linker. This coding sequence was amplified from the synthesized gene using primers rofA-FLAG-3'_F and OE_rofA_FLAG_R. The promoter for *rofA* and the RofA binding sites were determined by aligning the *rofA-prtF* intergenic region from JRS4 (111) with the *rofA-cpa* intergenic region from 5448. The RofA binding sites, P_{rofA}, and the 5' end of the gene were amplified from 5448 gDNA using oligos EcoRI_ProfA_F and rofA-5'_R. Finally, the ARNold webserver (152) was used to predict a rho-independent terminator downstream of *rofA*. The terminator region was amplified using primers rofA_term_F and NcoI_rofA_term_R. These three fragments were joined with OE-PCR and the completed P_{rofA}-*rofA*-FLAG-t gene cassette was digested with EcoRI and NcoI and ligated into EcoRI/NcoI digested and dephosphorylated pJRS525, generating plasmid pProfA1-FLAGt.

pProfA1-HqA-FLAGt and pET21a-rofA1-HqA

The *rofA*-HqA allele with alanine substitution mutations at histidines 210, 278, 330 and 380 was synthesized by Gene Universal using site-directed mutagenesis and was sent in vector pUC19 (pUC19_rofA-HqA). To generate plasmid pProfA1-HqA-FLAGt, pUC19_rofA-HqA was digested with EcoRV and ClaI. The 1.3 kb fragment was gel extracted and ligated into EcoRV/ClaI digested pProfA1-FLAGt. To generate

plasmid pET21a_rofA1-HqA, pUC19_rofA-HqA was digested with NsiI and ClaI. The 1.1 kb fragment was gel extracted and ligated into NsiI/ClaI digested and dephosphorylated pET21a-rofA1.

pProfA1-Hist

The pProfA1-FLAGt plasmid was first modified to include a SalI cut site downstream of the stop-codon, but before the ARNold predicted terminator. pProfA1-FLAGt was amplified using the Q5 site-directed mutagenesis kit (NEB) and primers SDM_RofA-FLAG_F2 and SDM_SalI_rofA_R to generate plasmid pProfA1-FLAGt_SalI_term. Then, plasmid pET21a-rofA1 was amplified using EcoRV-rofA-His_F and SalI_RofA-His_R. This fragment was digested with EcoRV and SalI and ligated into EcoRV/SalI digested and dephosphorylated pProfA1-FLAGt_SalI_term generating plasmid pProfA1-Hist.

pProfA1-FLAGt_Km

The pProfA1-FLAGt plasmid was digested with EcoRI and NcoI, and the 1.8 kb P_{rofA}-*rofA1*-FLAG-term cassette was gel extracted. This fragment was then ligated into EcoRI/NcoI digested and dephosphorylated pKSM201 to generate plasmid pProfA1-FLAGt_Km.

pP23-Spec

Plasmid pP23-Spec was generated by adding the P23 lactococcal promoter (151) to the *E. coli/S. pyogenes* shuttle vector pJRS525. The *lacZα* cassette was

preserved to allow for blue-white screening. P23 was joined to fragments upstream and downstream of the desired insertion site by amplifying pJRS525 with oligo pairs JRS525_OE-PCR_F1/JRS525_OE-PCR_R1 and JRS525_OE-PCR_F2/JRS525_OE-PCR_R2. Promoter P23 was amplified from pOri23 using oligo pair P23_OE-PCR_F/P23_OE-PCR_R. These fragments were joined with OE-PCR and the final product was digested with DpnI before PCR purification. The OE-PCR fragment and plasmid pJRS525 were both digested with BamHI and PacI, gel-extracted, ligated together, and transformed into DH5α competent cells. Candidate transformants were screened by restriction digest and confirmed by sequencing. Candidates had 4 SNPs that were different than the reported P23 sequence, but these SNPs were also present in the pOri23 stock.

pP23-*luxABCDE*

The *luxABCDE* was amplified from Xen20 genomic DNA using oligos EcoRI_luxA_F2 and BamHI_luxE_R2 and PCR-purified. pP23-Spec and the *luxABCDE* PCR fragment were digested with BamHI and EcoRI. Digested pP23-Spec was dephosphorylated and PCR purified. *luxABCDE* was ligated into pP23-Spec generating plasmid pP23-luxABCDE-spec, and the ligation product was transformed into NEB 10-beta competent cells. Transformation plates were imaged with an iBright Imaging System (Thermo Fisher) using the 'Chemiluminescence' setting to identify luminescent transformants. Plasmid from luminescent candidates was isolated and confirmed by Sanger sequencing using universal oligos M13F and M13R, as well as luxA_Seq_F, luxB_Seq_F, luxC_Seq_F, luxC_Seq_R, luxD_Seq_R, and luxE_Seq_R.

pCRS-tsf-lux-pepO

To generate the allelic exchange construct, the upstream and downstream homologous flanking DNA from the *tsf-pepO* intergenic region was amplified from 5448 gDNA using oligo pairs oAxtsf.1/oAxtsf.2 and oAxpepO.1/oAxpepO.2. P23-*luxABCDE* was amplified from pP23-*luxABCDE*-Spec with oligo pair oAxP23lux.1/oAsP23lux.2. These regions were adjoined to the P23-*luxABCDE* cassette through OE-PCR. The full 8 kb OE-PCR product was gel-extracted, digested with BamHI, then PCR purified. Plasmid pCRS was digested with BamHI, dephosphorylated, and PCR-purified. The tsf-luxABCDE-pepO OE-PCR product was ligated into pCRS to generate plasmid pCRS-tsf-lux-pepO, and the ligations transformed into GB5-alpha (Gold Bio) competent cells. Transformation plates were imaged with the iBright and luminescent transformants were selected. Plasmid DNA was isolated from luminescent transformants and digested with BamHI to determine if digestion fragments were consistent with the ligation product and these plasmids were sent for sequencing with oAxP23lux.1, luxA_Seq_F, luxB_Seq_F, luxC_Seq_F, luxC_Seq_R, luxD_Seq_R, and luxE_Seq_R, oAxP23lux.2, M13R, tsf_seq_F, and tsf_seq_R. A SNP was found within the *luxA* gene, however this mutation was silent, and the strain remained bioluminescent, so the study continued with this construct.

Bacterial Transformations

E. coli Competent Cell Preparation

To prepare strains DH5α, C41[DE3], and GM48 for electroporation, 500 mL of LB in a 2 L Erlenmeyer flask was inoculated with 5 mL of overnight culture. This flask was incubated at 37°C with 250 rpm shaking until the OD_{600} reached 0.4-0.6. Culture was placed on ice for 30 min, then centrifuged at 7,000 x g for 30 min at 4°C. Supernatant was discarded, then pellet was resuspended in 250 mL ice-cold 10% glycerol. Cells were pelleted again and washed twice more with 100 mL, then 50 mL ice-cold glycerol. Cells were pelleted after final wash, then resuspended in a final volume of 1 mL of 10% glycerol which was aliquoted into 50 µL aliquots. Aliquots were frozen at -80°C until use.

S. pyogenes Competent Cell Preparation

S. pyogenes was inoculated into 10 mL THY + 20 mM glycine and incubated overnight at 37°C. 7.5 mL of this overnight cultures was then used to inoculated 150 mL of THY + 20 mM glycine and incubated at 37°C until cultures reached an OD_{600} of 0.3-0.5. For encapsulated strains, 85 µg/mL hyaluronidase was also added to THY during growth. Cultures were placed on ice, then centrifuged at 7,000 x g for 30 minutes. Supernatant was discarded, and pellets were resuspending in 50 mL of 10% glycerol. Cells were pelleted again and washed twice more in 50 mL of 10% glycerol. Cells were pelleted after final wash and resuspended in a final volume of 1 mL of 10% glycerol which was aliquoted into 200 µL aliquots. Aliquots were frozen at -80°C until use.

Electroporation

10 – 50 ng of intact plasmid DNA was routinely transformed into *E. coli*. 1 – 5 µg of replicating plasmid was routinely transformed into *S. pyogenes*. Plasmid was added up to 20 µg for challenging plasmids. All DNA was drop-dialyzed on a 0.022 µm membrane (Millipore) floating on sterile deionized water (diH_2O) for 30 min to 1 hour prior to electroporation. DNA was then combined with an aliquot of competent cells and flicked gently to mix. This suspension was added to chilled 0.2 cm electroporation cuvettes.

All electroporations were performed with a Gene Pulser XCell (BioRad). NEB 10-beta cells and GB5-alpha cells were transformed following manufacturer's specifications. Lab-prepared competent *E. coli* was electroporated at 2.5 kV, 200 Ω and 25 µF. After pulse, cells were immediately resuspended in 1 mL of LB or SOC and grown at 37°C with shaking for 1 hour. *S. pyogenes* was electroporated with 1.75 kV, 400 Ω, and 25 µF. After pulse, cells were immediately resuspended into 10 mL THY and grown at 37°C stationary for 2 hours. After growth, *E. coli* was serially diluted and plated onto LB agar plus selective antibiotics. *S. pyogenes* was centrifuged at 7,000 x g for 10 minutes, then resuspended into 300 µL, which was then serially diluted and plated onto THY agar plus selective antibiotics.

Mutant Generation by Allelic Exchange

All allelic exchange was performed in 5448 using the protocol described previously (153). Descriptions of the allelic exchange strains and any deviations from the outlined protocol are described below.

5448 Δ*rofA_sc*

5448 Δ*rofA*_sc is an allelic exchange mutant of *rofA* (MGAS5005_Spy0106/ SP5448_RS00675) where codons 10-12 were mutated to three consecutive stop codons to generate strain Δ*rofA_sc* (**Fig. 2.1**). The mutated alleles also contained a BamHI restriction enzyme site downstream of the stop codons to enable efficient screening of candidates. Plasmid pCRS_rofA_stop was transformed into WT 5448 and grown on THY + Spec plates at the replication-permissive temperature of 30°C. Colonies were then inoculated into THY + spec and incubated at the non-permissive temperature of 37°C to select for integrants. Resulting merodiploid strains were again passaged in THY without antibiotic at 30°C to allow for a second crossover event and allow for plasmid curing. Strains were then incubated at 37°C to select for strains cured of the plasmid. Candidates were screened for spectinomycin sensitivity, then, colony PCR was performed with oRofA_SOW_1 and oRofA_SOW_4 to amplify the *rofA* locus. These PCR products were digested with BamHI to determine if they contained mutant or WT allele. 5448 *rofA*_R is an allelic exchange revertant where second crossover event recovered the native *rofA* coding sequence instead of the stop codons (**Fig. 2.1**).

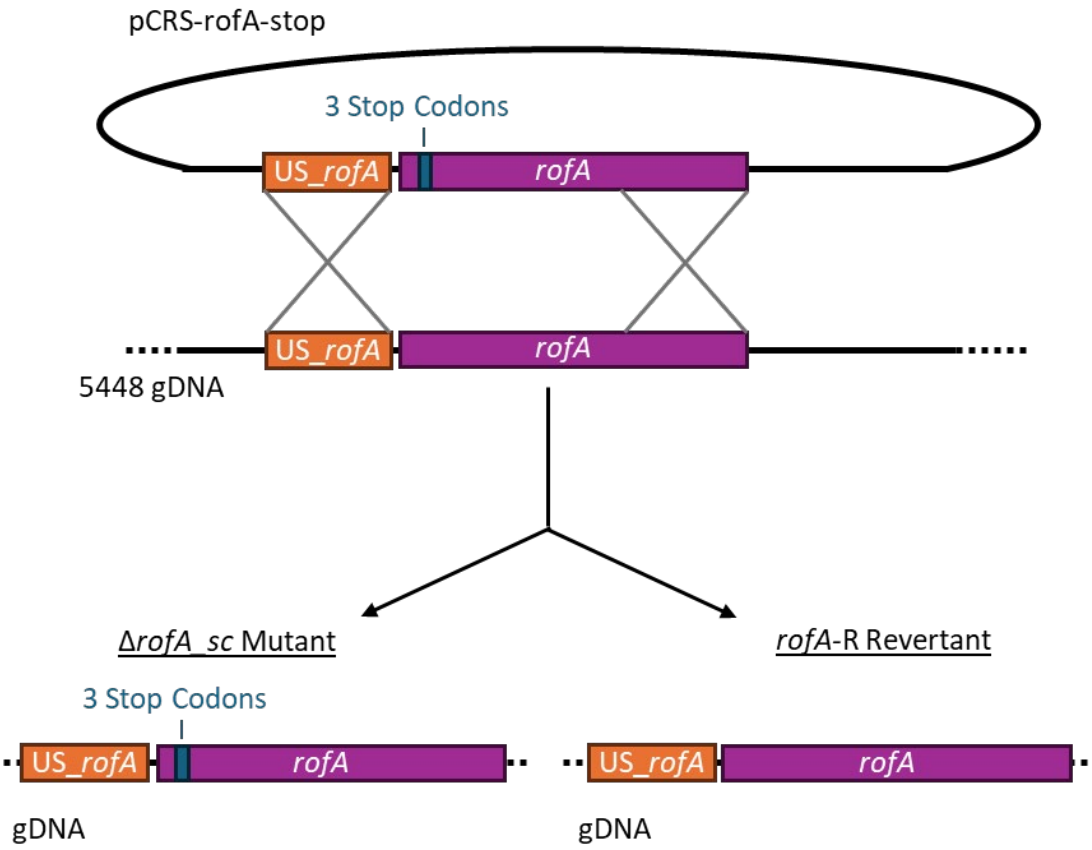


Figure 2.1: Schematic of the Allelic Exchange Process
Plasmid pCRS is used to introduce 3 stop-codon mutations into the *rofA* coding sequence on the 5448 genomic DNA. After plasmid integration, the second crossover event can yield either the mutant (Δ*rofA_sc*) or the revertant (*rofA*-R).

5448_lux Strains

All allelic exchange steps were performed in incubators that maintain 5% CO_2. 20 µg pCRS tsf-lux-pepO was transformed into 5448 and transformation plates were incubated at 30°C for 48 hours. Transformants were patched onto THY + Spec plates and grown at 30°C for an additional 48 hours. Patch plates were then imaged with the iBright and 5 luminescent transformants were identified. Luminescent transformants were inoculated into THY + Spec and grown at 30°C overnight. Cultures were then diluted 1:100 into fresh THY + Spec and grown at the non-permissive temperature of

37°C to select for bacteria with the plasmid integrated into the chromosome. These cultures were then serially diluted and plated for single colonies onto THY + Spec plates and incubated overnight at 37°C. Plates were imaged on the iBright to identify luminescent colonies.

3 luminescent colonies per transformant (15 total isolates) were inoculated into THY + Spec and incubated overnight at 37°C. After incubation cultures were measured for luminescence on a Berthold Centro XS^3 LB 960 Luminometer. Three bright isolates were diluted 1:100 into THY without antibiotic and grown at 30°C overnight. This was repeated for a total of 2 passages to facilitate a second cross-over event to remove the plasmid from the chromosome. Then, strains were diluted 1:100 in THY and grown at 37°C. This was repeated twice more, for a total of 3 passages, to select for strains cured of the plasmid. The cultures from the third passage were then serially diluted, plated on THY, and incubated at 37°C overnight.

300 candidates were patched onto three types of plates, 1) THY only, 2) THY + Spec, and 3) BAPs and grown overnight at 37°C. The 9 isolates that were luminescent, sensitive to spectinomycin, and beta-hemolytic were selected as the candidate 5448_lux strains. 3 isolates that were not luminescent, sensitive to spectinomycin, and beta-hemolytic were selected as candidates that reverted to the wild-type genotype. gDNA was isolated from strains and PCR was used with primer pair oAX1781V1/oAX1782V2 to determine if the P23-*luxABCDE* cassette was present or not.

RNA Isolation

RNA isolation for RNA-sequencing, RT-PCR, or RT-qPCR was performed as described previously (96). Cells were collected at exponential (Klett: 55-60, OD_{600}: ~0.3), transition (Klett: 120-125, OD_{600}: ~0.6), or early stationary (2 h after transition) phase and placed on ice. Cells were harvested by centrifugation and pellets were frozen at -80°C until isolation. Frozen cell pellets were resuspended in 700 µl of TRI Reagent (Zymo Research) and approximately 200 µL of acid-washed glass beads (Sigma) were added. Cells were disrupted by vortexing for 5 min. The RNA was then isolated with the Direct-Zol RNA Mini-Prep Plus kit (Zymo Research) following manufacturers specifications. RNA was then DNase treated with the rigorous two-step treatment of the TURBO DNA-free kit (Invitrogen) plus murine RNase inhibitor (NEB) to remove residual gDNA. To confirm gDNA removal, PCR using standard Taq buffer and polymerase (NEB) with oligos LicT_RevTrans_F and LicT_RevTrans_R (**Table 3.2**) was performed on 250 ng of total RNA for 45 cycles. Samples positive for gDNA were retreated with the Turbo DNA free kit with the standard retreatment plus murine RNase inhibitor. RNA quality was assessed using an Agilent 4150 TapeStation. RNA was quantified on a Nanodrop 8000 spectrophotometer (Thermo Scientific).

RNA Library Preparation, Sequencing, and Data Analysis

130 ng total was used for RNA-Sequencing (RNA-Seq) library preparation using the Stranded Total RNA Prep Ligation with Ribo-Zero Plus (Illumina). The ribosomal RNA (rRNA) depletion step was performed with the addition of 1 µL of a 1 pmol/µL *S. pyogenes* specific oligo probe pool (IDT, **Table 2.4**). Otherwise, the

libraries were prepared following the manufacturer's specifications. Quality and quantity of libraries was assessed by TapeStation (Agilent).

Table 2.4: Custom Oligos for GAS rRNA Depletion

Oligo Number	Sequence
1	ACATGGAATTCCACTCTCCCCTTCTGCACTCAAGTTCTCCAGTTTCCAAA
2	TCTGACTGCCGATGATATCTCGTTGGCATTCGGAGTTTATCTGAGATTGG
3	CCCTCCTGCCGTTTCGCTCGCCGCTACTAAGGCAATCGCGTTTGCTTTCT
4	GTCGTTTGTCACGTCCTTCTTCGGCTTCTAGTGCCAAGGCATCCACCGTG
5	GTTTATTACCGGCAGTCTCGCTAGAGTGCCCAACTTAATGATGGCAACTA
6	TTACTAACATGCGTTAGTCTCTCTTATGCGGTATTAGCTATCGTTTCCAA
7	TAGCTGCTTCTAAGCTAACATCCTAGTTGTCTGTGCAACCCCACATCCTT
8	AACCCAGGGCGGACGAGCCTTCCCCTGGAAACCTTAGTCTTACGGTGGAC
9	GCTAAGCGACTACCTTATCTCACAGGGGGCAACCCCCAACTACATCAGGC
10	TTAGCCGGACTCTTCCAATCGTCCGGTTTAGTTAGCCTACTGCGTCCCTC
11	GTAGAGTATCATTAAACGCTAGAAGCTTTTCTTGGCAGTGTGACATCACT
12	GGTCACATGGTTTCGGGTCTACAACATGATACTAAATCGCCCTATTAAGA
13	TAGGAGGCGACCGCCCCAGTCAAACTGCCCGTCAGACACTGTCTCCGATA
14	GGTTAGAGTAGCCATAACACAAGGGTAGTATCCCAACAACGCCTCATTCG
15	GGGTTGGGAGATGGTCCTCCCAGATTCCAACGAGATTTCACGTGTCTCGC
16	TAGTACAGGAATATCAACCTGTTGGCCATCGGATACACCTTTCGGTCTCT
17	TGGCAAGACAACTGGTACACCAGCGGTAAGTCCACTCTGGTCCTCTCGTA
18	GCTAGGGCTCAAAAGAATTTTAAATACGAGGCTTTTACTCTCTTTGGCTT
19	CAACGCTCGGGACCTACGTATTACCGCGGCTGCTGGCACGTAGTTAGCCG
20	TCTTCTATCCTTTTAAAGTCCCACGTCGCAGTCCTACAACCCCGAAGTGT
21	CTTTCGCTACTCATACCGGCATTCTCACTTCTATGCGTTCCAGCGCTCCT
22	AACTGAGATTGGCTTTAAGAGATTAGCTTGCCGTCACCGGCTTGCGACTC
23	GCCAATGCCTTTAACTTCAGACTTAAAAAACCGCCTGCGCTCGCTTTACG
24	GTCTCAGTCCCAGTGTGGCCGATCACCCTCTCAGGTCGGCTATGTATCGT
25	TTACAGACCAGAGAGCCGCTTTCGCCACCGGTGTTCCTCCATATATCTAC
26	TTACCTCACCAACTAGCTAATACAACGCAGGTCCATCTCATAGTGGAGCA
27	TTAGTGGGTTCCCCCATTCGGACATCTCTGGATCAGCGCTTACTTACAGC
28	ATTCGCTCCCCATCACAGCTCAATGTTATAAAGATAAGCATTTGACTCAT
29	CTATCTACCTGATCATCTCTCAGGGCTCTTACTGATATAAAATCATGGGA
30	CATCTTACGATTTAGCAGAGAGCTGTGTTTTTGATAAACAGTTGCTTGGG
31	GGCGTGCTGATCCGCGATTACTAGCGATTCCGACTTCATGTAGGCGAGTT
32	TCACACATAGAACGCTCTCCTACCATGACACTTTTGTGTCATCCACAGCT
33	AGCCCCGGTACATTTTCGGCGCAGGGTCACTCGACTAGTGAGCTATTACG
34	GGAGAGAACCAGCTATCTCCAAGTTCGTTTGGAATTTCTCCGCTACCCAC
35	TGCTGTGTTCGGCATGGGTACAGGTGTATCTCCTAGGCTATCGTCACTTA
36	AAGAATGCTGTATTGGCACCGGTGATCCACGTTTTTGTATTGAGCCCGCG
37	TTGGATAAGTCCTCGAGCTATTAGTATTAGTCCGCTACATGTATCGCTAC
38	GTTTCCCGAAAATACAAAGAAATAGTCGGTCAACACTTAGATGTGGATTC
39	GGTAACGCGCTCTTTCTTGGAAGTTGCGTCCGCGTGAAACGGCTTGCAGA

Table 2.4 continued…

Oligo Number	Sequence
40	GGGGCTTCGCACTTAGATGCTTTCAGCGCTTATCCCTTCCCTACATAGCT
41	GCTCCGTCTCTTCAACTTAACCTCGCATCATATCGTAACTCGCCGGTTCA
42	CCCTCACCCAAACAGTGCTCTACCTCCAAGAGACTTAACATCGACGCTAG
43	AGTCAATGGCTCCTACCTATCCTGTACATGTGGTACAGATACTCAATATC
44	GATCACTAAGCCCGACTTTCGTCCCTGCTCGAGTTGTAGCTCTCGCAGTC
45	GCTCTCACCCATTAACGGGCTCGAACTTGTTGTAGGCACACGGTTTCAGG
46	TAGATGTTTCAGTTCACTGCGTCTTCCTTCTCATAACCTTAACAGTTATG
47	AACGAGAGTTCTCTCGATCACCTGAGGCTACTCGCCTCGACTACCTGTGT
48	TTTTCAACGTGCCCTGGTTCGGTCCTCCAGTGAGTTTTACCTCACCTTCA
49	GACTTCACCCCAATCATCTATCCCACCTTAGGCGGCTGGCTCCTAATAGG
50	ATGAGGTAGGTTACCTACGCGTTACTCACCCGTTCGCAACTCCTTGAACC
51	CCGATGTACCGAAGTAAAACTCTATCTCTAGAGCGGGCATCGGGATGTCA

Libraries were sequenced on an Illumina NextSeq 1000 in the Brain and Behavior Institute - Advanced Genomic Technologies Core (BBI-AGTC), which is supported by the BBI and the University of Maryland, College Park. Reads were deposited at the Sequence Read Archive under accession PRJNA1037582. Read quality was measured using FastQC, filtered and trimmed using trimmomatic (154), and mapped against the MGAS5005 genome (accession CP000017) using hisat2 (155). Differential expression analyses were performed following visualization of size-factor and quantile normalization of read counts. Potential batch effects were evaluated by comparing the DESeq2 results under three conditions: 1. without any additional parameters to the experimental model, 2. the addition of the most likely batch factor to the model, and 3. after adding sva-derived estimates to the model (156). Differential expression analyses were performed with DESeq2 (157). The resulting metrics of expression were tested for ontology enrichment using goseq (158).

Differentially expressed genes were defined as those that had log2_FC >= 1 or <= -1 and an adjusted p-value <= 0.05. Genes differentially expressed between WT and the Δ*rofA_sc* mutant were compared to those differentially expressed between the *rofA*-

R revertant and the mutant to control for off-target effects of the allelic exchange process. Only genes consistently differentially regulated between the two datasets were considered. Genes were then analyzed for homology to other genes or by known function and categorized into groups. Manual curation of gene ontology was performed using available data from microbesonline.org (159), including the putative operon the gene belonged to, the "cluster of orthologous genes" (COG) designation, and the homology of the gene to others with functional designations. If this information was insufficient, the gene was further investigated by using the 'Paper BLAST' feature of Microbes Online, or using NCBI PubMed and Google Scholar to determine if a functional characterization could be made based on the function of homologs or the literature available.

Endpoint RT-PCR and RT-qPCR

For endpoint reverse transcriptase (RT)-PCR, copy DNA (cDNA) was generated using the Luna Superscript kit (NEB) from 25 ng of total RNA. 5 µL of cDNA was then used as the template for PCR with Standard Taq polymerase (NEB) and oligos rofA_FLAG_F2 and rofA_M1_RT_R. RT-quantitative PCR (RT-qPCR) was conducted on 25 ng of RNA using the Luna Universal One-Step RT-qPCR Kit (NEB) with the manufacturer's recommended protocol. RT-qPCR was performed on a Roche Light Cycler 380. qPCR primers (**Table 2.2**) were designed using Primer-BLAST (160) or Primer3 alone (prev. at: http://biotools.umassmed.edu/bioapps/primer3_www.cgi).

Results for the expression of *rofA* at different time points in growth are reported as the fold-change using the $-\Delta C_T{}'$ method (161) relative to levels of *rofA* at exponential growth phase as the calibrator. Results for the expression of RofA target genes are reported as the fold-change using the $-\Delta\Delta C_T$ method (161) comparing the mutant and revertant to the WT relative to the expression of *gyrA* (Internal control). Each biological replicate (n = 3) was run in technical duplicates whose cycle threshold (C_T) values were averaged after RT-qPCR. Significance was determined using either a Brown-Forsythe and Welch Analysis of Variance (ANOVA) with Dunnett's T3 test for multiple comparisons if only one gene was being assessed or a Two-Way ANOVA with a Tukey's test for multiple comparisons if multiple genes were being assessed.

Capsule Assay

The capsule assay was performed as described previously (162). Briefly, strains were grown to late-logarithmic phase of growth (OD_{600} ~ 0.4-0.5). 2 mL of each culture was harvested and placed on ice until remaining strains reached the desired OD_{600}. Samples were centrifuged at 7,000 x g for 10 minutes, and the medium removed. The pellets were washed twice with 1 mL of diH_2O and finally resuspended in 500 µL of diH_2O. 1 mL of chloroform was added to each sample and samples were placed on the shaker for 1 hour at room-temperature. After incubation, tubes were centrifuged at 14,000 x g for 10 minutes. 200 µL of the aqueous layer was then collected and combined with 600 µL of Stains-All solution (0.2 mg/mL Stains-All (Sigma), 0.06% acetic acid, 50% formamide). The absorbance was measured at 640 nm and compared to a standard curve of known concentrations of hyaluronic acid. Significance was

determined by a Brown-Forsythe and Welch ANOVA with Dunnett's T3 test for multiple comparisons.

Adherence Assay

Adherence assays were performed using a modified versions of protocols described previously (163, 164). Primary keratinocytes (American Type Culture Collection [ATCC]) were thawed at passage 3 and incubated in Dermal Cell Basal Medium (ATCC) supplemented with the Keratinocyte Growth Kit (ATCC) at 37°C + 5% CO_2 until approximately 80% confluent. Once 80% confluency was reached, cells were detached and seeded into a 96-well plate at a seeding density of 2.3 x 10^4 cells per well such that the wells were 100% confluent after approximately 24 hours.

S. pyogenes strains were grown to mid-exponential growth phase (OD_{600} = ~ 0.35) and placed on ice until all strains reached the desired OD. Strains were diluted to approximately 2.3 x 10^6 colony forming units (CFU)/mL in Hanks Balanced Salt Solution (HBSS, Cytiva) plus 1% fetal bovine serum (FBS, Cytiva). Keratinocyte monolayers were washed once in 100 µL of phosphate buffered saline (PBS, ATCC) and 100 µL of the bacterial suspension in HBSS + 1% FBS added to the monolayer to achieve a multiplicity of infection (MOI) of 10. Plates were then spun at 1,600 rpm for 5 minutes. Plates were then incubated for 1 hour at 37°C.

The bacterial suspensions were serially diluted and drop-plated to enumerate the inoculum. After 1 hour, monolayers were washed five times in PBS to remove non-adherent bacteria. Keratinocytes were disrupted by adding 0.25% Trypsin-ethylenediaminetetraacetic acid (EDTA) solution to each well and incubating at 37°C

until the cells rounded. An equal volume of HBSS + 1% FBS was then added with Triton X-100 at a final concentration of 0.025% to lyse keratinocytes. These suspensions were then serially diluted and drop-plated for bacterial enumeration. Adherence was reported as the percentage of the inoculum that were recovered after the wash steps. Significance was determined using a Brown-Forsythe and Welch ANOVA with Dunnett's T3 test for multiple comparisons.

S. pyogenes Lysis and Phos-Tag Gel Electrophoresis

S. pyogenes cultures were centrifuged for 10 minutes at 8,400 x g and the medium removed. Pellets were washed once in cold TN Buffer (50 mM Tris, 150 mM NaCl, pH 8.0) and re-pelleted by centrifugation. Pellets were stored at -20°C until lysis. Frozen pellets were resuspended in cold TN Buffer to a final concentration of 25 OD units/mL. Halt 100x Protease and Phosphatase Inhibitor Cocktail (Thermo Scientific) was added to 1X. DNase I (Zymo) was added to a final concentration of 20 U/mL. 1 µL of PlyC was added per 10 µL of cell suspension. This was incubated for 30 minutes at room temperature. After incubation, 5x cracking buffer (0.312 M Tris pH 6.8, 25% beta-mercaptoethanol (βME), 50% glycerol, 0.1 g/mL SDS, 0.04 % bromophenol blue) was added to 1x, incubated at room temperature for 10 minutes, and then centrifuged at max speed for 10 minutes.

Lysates were loaded onto a Zn^{2+}-Phos-Tag gel using a previously established recipe (165). Phos-Tag gels were run at 30 mA constant current until the loading dye reached the bottom of the gel. Gels were then washed first in methanol-free transfer buffer (0.2 M Tris, 0.2 M Glycine, 0.2% sodium dodecyl sulfate) with 1 mM EDTA

for 10 minutes with rocking, then in methanol-free transfer buffer without EDTA for 20 minutes with rocking. Gels were transferred to polyvinylidene difluoride (PVDF) membranes either overnight at 25 V or for 1.5 H at 125 V at 4°C. Membranes were allowed to dry completely before western blotting.

Western Blotting

PVDF membranes were rehydrated prior to western blotting in methanol, then blocked with 5% non-fat milk in phosphate buffered saline plus tween (PBST) for either 2 h at room temperature or at 4°C overnight. After blocking, membranes were washed briefly in PBST, then primary antibody solutions added in 5% milk-PBST. For anti-FLAG (α-FLAG) western blotting, monoclonal Anti-FLAG M2 antibody (Sigma) was used at a 1:1,250 dilution and for anti-His (α-His) western blotting, monoclonal anti-6x-His (His_6) antibody clone 6AT18 (Sigma) was used at a 1:500 dilution. Membranes incubated with the primary antibody solution for 2 h at room temperature (RT). After incubation, membranes were washed twice briefly in PBST, then washed once for 5 minutes and twice for 10 minutes in PBST. The secondary antibody was an anti-Mouse antibody conjugated with horseradish peroxidase (HRP) (Sigma) diluted 1:5,000 in 5% milk-PBST. Membranes were incubated with the secondary antibody for 1 h at RT and washed as described above. Membranes were visualized with SuperSignal West Pico Plus reagent (Thermo Fisher) and imaged on an iBright Imaging System with the Chemiluminescence Setting (Thermo Fisher).

Protein Expression and Purification

E. coli C41[DE3] carrying pET21a-*rofA1* or pET21a-rofA1-HqA was grown in ZYP-5052 autoinduction medium (143) for ~24 h at RT shaking at 250 rpm. After 24 h, *E. coli* was pelleted by centrifugation and washed in sodium phosphate buffer (NaPi) at pH 7.4 (166). Washed *E. coli* was again pelleted by centrifugation and the pellet was frozen at -80°C until lysis and purification.

To lyse and clarify, pellets were first resuspended in 50 mL of buffer A (20 mM NaPi, 20 mM imidazole, 500 mM NaCl, pH 7.4). Lysozyme was added to the suspension at a concentration of 1 mg/mL and EDTA-free protease inhibitors (Pierce or Sigma) at 1x. The suspension was incubated on ice for 30 min. After incubation, the suspension was sonicated with a Branson Sonifier 450 with a tapered microtip at 50% duty for 1 min 5 times with 3 min rests on ice in between. The insoluble material was pelleted by centrifugation at 20,000 x g for 30 min. The supernatant was clarified further by centrifuging at 20,000 x g for an additional 20 min, then 15 min twice.

For affinity purification, 2 mL of Ni-NTA resin slurry (Invitrogen) was added to a 15 mL conical and the beads were allowed to settle. The storage buffer was removed, and the resin was washed once in 5 mL of buffer A (20 mM Imidazole, 20 mM NaPi, 500 mM NaCl, pH 8.0). The lysate was incubated with the resin for 1 h at 4°C with rocking. After incubation, the lysate/resin mixture was added to an Econo-Column (Bio-Rad). The flow-through (FT) was collected, and the column washed with 25 mL of buffer B1 (50 mM NaPi, 300 mM NaCl, 20 mM imidazole pH 8.0). 25 mL of buffer B2 (50 mM NaPi, 300 mM NaCl, 80 mM imidazole pH 8.0) was then added and the flow valve was adjusted to a rate of 1 mL/min. To elute the protein, 10 mL of

buffer C (50 mM NaPi, 300 mM NaCl, 250 mM imidazole pH 8.0) was added and 1 mL fractions were collected. Immediately after elution, an equal volume of Tris-HCl at pH 8.5, glycerol to 10%, and EDTA to 50 mM was added to each fraction to avoid precipitation.

Fractions (F1-F10) were assessed by SDS-polyacrylamide gel electrophoresis (PAGE) and Coomassie or SimplyBlue SafeStain (Invitrogen) staining (See **Appendix A3**), and fractions containing RofA-His and minimal background protein were pooled. A 10 kDa MWCO filter unit (Pierce) was used for buffer exchanging into Tris-HCl pH 8.5, 10% glycerol, and 50 mM EDTA. After the buffer exchange, the protein was concentrated with another 10 kDa MWCO filter, and concentration was assayed with the Bio-Rad protein assay reagent following manufacturer specifications.

His_6-HPr, His_6-EI and His_6-PEPCK were purified as described previously (95). *E. coli* C41[DE3] carrying plasmids pKSM879 and pKSM357 expressing His_6-PEPCK and His_6-EI, respectively, were grown at RT for 52 h in ZYP-5052. *E. coli* C41[DE3] carrying pKSM712 was grown at 37°C for 24 h in ZYP-5052. Cultures were harvested and lysed following the same methods described above. Purification with NiNTA resin was conducted the same as described above, except that both wash steps were done with Buffer B1.

In vitro PTS Phosphorylation Assays

Assays were carried out as described previously (95). Briefly, Phosphorous-32 (^{32}P)-radiolabeled PEP ([^{32}P]EP) was generated by combining 100 μM oxaloacetic acid (Sigma), 50 μCi of [γ-^{32}P]-ATP (Perkin-Elmer), and 0.1 μM His_6-PEPCK in 50 mM

HEPES, 10 mM $MgCl_2$ at pH 7.5. The reaction was incubated for 20 min at RT, then diluted 10-fold with diH_2O. The reaction was then purified on a 2 mL AG-1-X8-bicarbonate column (Bio-Rad) and washed sequentially with 5 mL of 0.3 and 0.4 M triethylammonium bicarbonate buffer (TEAB) at pH 8.5. The [^{32}P]EP was eluted with sequential additions of 5 mL of 0.6 and 0.7 M TEAB, fractions pooled, and dried in a Savant DNA120 SpeedVac Concentrator without heat. Solid [^{32}P]EP was resuspended into 100 µL of 50 mM HEPES, pH 7.5 and stored at -20°C.

The *in vitro* PTS phosphorylation assay reactions were prepared by combining 250 nM His_6-EI, 1 µM His_6-HPr, and 5 µM of RofA-His or RofA-HqA-His in 50 mM HEPES, 10 mM $MgCl_2$ at pH 7.5. Heat denatured (HD) samples of RofA-His and RofA-HqA-His were incubated at 99°C for 10 min and cooled on ice prior to being added to the reaction. Reactions were initiated by adding 750 – 1,100 kCPM of [^{32}P]EP and allowed to proceed for 20 min at 37°C. After 20 minutes, 5x cracking buffer (0.3M Tris pH 6.8, 25% (v/v) β-mercaptoethanol, 51% (v/v) glycerol, 10% (w/v) SDS, 0.01% (w/v) bromophenol blue) was added to 1x and samples were placed on ice. Reactions were resolved using a 10% SDS-PAGE gel run at 20 mA until the load dye reached the bottom of the gel. The gel was dried in a Model 853 Gel Dryer (Bio-Rad) at 80°C for 1 h, then the gel was exposed to a phosphorimager cassette overnight and imaged on a FLA-7000 imager (GE/Cytiva). The band corresponding to RofA was determined based on the size of the proteins run alone on an SDS-PAGE gel.

Luminescent Growth Curves

Cultures were inoculated from colonies into THY and grown at 37°C overnight. The next morning, the OD_{600} was read and cultures were used to inoculate fresh THY at an OD_{600} of approximately 0.05. OD_{600} was measured every half hour, and luminescence was measured every hour. Luminescence was measured by adding 100 µL of each culture to a 96-well plate and measuring the luminescence on a Berthold Centro XS^3 LB 960 Luminometer with a 5 second exposure. Background autoluminescence of THY was measured in blank wells of 100 µL THY only, and the average of these values was subtracted from the luminescence measurements of the cultures. The correlation between luminescence and OD_{600} was determined by trimming the data to the first 4 hours of growth, and fitting a straight line using the least squares fit in GraphPad Prism.

Area Under the Curve and Doubling Time Calculations

Area under the growth curve was used as a metric to analyze bacterial growth, as it integrates together several contributions to growth (initial population, growth rate, carrying capacity) into a single metric (167) and this metric is included for microbial growth curve analytical programs (167, 168). To determine the area under the curve (AUC), the data was trimmed to the 6.5 hour time point, as growth to this time point was consistent across all growth curves completed. Then, the 'Area Under the Curve' function in prism was applied to the growth curves.

To determine the doubling time, the segment of the growth curve that corresponded to exponential growth was identified, and the data from this region added to a separate spreadsheet. This data was then analyzed using the GraphPad Prism 'Nonlinear fit of Exponential Growth' function and converted from hours to minutes. Statistical significance for both was determined by a Brown-Forsythe and Welch ANOVA with Dunnett's T3 test for multiple comparisons.

Isolation of Human Polymorphonuclear Lymphocytes

Whole human blood was isolated from anonymous donors via venipuncture and collected into sodium-heparin-treated tubes (BD Vacutainer) according to IRB-approved protocol (see Ethics section). Neutrophils were isolated from blood using Polymorphprep^{TM} (Progen), following manufacturer instructions. To remove contaminating red blood cells, neutrophils were treated with red blood cell lysis solution (Thermo Scientific) and washed with Dulbecco's modified PBS (HyClone). Neutrophils were resuspended to 1.25 x 10^6 cells/mL in RPMI 1640 medium supplemented with 2.05 mM L-glutamine (HyClone) until assay.

Opsonophagocytic Killing Assay with Primary Human Neutrophils

Assays were performed as described previously (169) with slight modifications. WT 5448, 5448_lux candidates, and revertant strains were inoculated into THY from colonies and grown overnight in THY at 37°C. Overnight cultures were diluted 1:20 into fresh THY and grown to mid to late-exponential phase (OD_{600} ~0.4) and put on ice

until all strains reached the desired OD_{600}. Cultures were then pelleted by centrifugation at 6,000 x g. THY was removed and cultures were washed once in sterile saline, vortexed for 5 minutes to break up chains, and diluted 1:20 in PBS.

PBS suspensions were then diluted 1:10 into donor plasma for opsonization and incubated at 37°C for 30 minutes. Isolated neutrophils were seeded at a density of $1.0x10^6$ cells/well (800 µL) in 12-well plates with RPMI 1640 medium + 2.05mM L-glutamine (HyClone). 200 µL of bacterial culture in serum was added to seeded neutrophils at an MOI of 0.2 to a final volume of 1 mL. 12-well plates were centrifuged (500 × g for 5 min Sorvall/Heraeus 6445 rotor) to enhance contact between neutrophils and bacteria. Plates were incubated for 2 hours at 37°C + 5% CO_2. As a control, opsonized bacteria in serum were also incubated in RPMI 1640 + 2.05mM L-glutamine without neutrophils.

Following incubation, well supernatants were collected and stored on ice. 1 mL of sterile diH_2O was added to each well to lyse neutrophils and release any intracellular bacteria. These diH_2O suspensions were collected into microfuge tubes and pelleted by centrifugation at 9,000 x g for 4 minutes. Water was removed from the pellets, and pellets were resuspended in the previously collected well supernatants. Tubes were vortexed for 10 minutes to disrupt chains, serially diluted, and drop-plated onto THY agar. Plates were incubated overnight at 37°C + 5% CO_2. The following morning, CFU were enumerated and % survival was determined by the following equation: (Neutrophil CFU/Medium Only CFU) * 100. To combine data gathered from multiple donors on multiple days, each experiment was normalized to the % survival of wild type with the following equation: (% Survival Experimental/% Survival of WT) * 100.

For each donor, assays were completed with 3 biological replicates of bacteria which were assayed in technical duplicate. Statistical significance was determined through a Brown-Forsythe and Welch ANOVA with Dunnett's T3 test for multiple comparison.

Ethics

Whole human blood was acquired by protocol #10-0735, approved by the Institutional Review Board (IRB) at the University of Maryland, College Park.

Chapter 3: The *Streptococcus pyogenes* Stand Alone Regulator RofA Exhibits Characteristics of a PRD-Containing Virulence Regulator

This chapter is adapted from:

Hart MT, Rom JS, Le Breton Y, Hause LL, Belew AT, El-Sayed NM, McIver KS. 2024. The *Streptococcus pyogenes* stand-alone regulator RofA exhibits characteristics of a PRD-containing virulence regulator. Infection and Immunity 92:e00083-24.

DOI: 10.1128/iai.00083-24

Overview

In this study, RofA was investigated to determine if it exhibited characteristics of a PCVR. The RofA regulon was defined with RNA-sequencing (RNA-seq) in 5448, a representative strain of the globally disseminated M1T1 serotype (140) and several novel gene targets of RofA were identified that included PTS operons. RofA was found to be phosphorylated *in vivo* and by the PTS proteins, EI and HPr, *in vitro*. It was also determined that glucose concentration impacts the phosphorylation of RofA *in vivo.* Several histidines were identified as being potential targets of phosphorylation. Alanine substitution mutations at these predicted residues appeared to impact the total amount of RofA that was phosphorylated *in vivo* but did not completely ablate phosphorylation. Taken together, these data support the hypothesis that RofA functions as a PCVR in *S. pyogenes*.

Results

RofA Regulates Virulence Factors and Sugar Metabolism Genes in M1T1 5448

Previous work investigating RofA gene targets was limited to a few genes and was completed in serotypes M6 and M2 (112, 113). The most well-studied targets of RofA regulation are the FCT-region genes involved in pilus expression (112-114). To expand upon these previous findings, the full RofA regulon was explored using RNA sequencing (RNA-seq) in the representative strain 5448 of the globally disseminated M1T1 serotype. To investigate the RofA transcriptome for 5448 (WT), we constructed a mutant strain eliminating the translation of RofA by introducing three consecutive stop-codon mutations at the 5' end of the *rofA* coding sequence (Δ*rofA_sc*). This retained normal transcription and allowed for assessment of *rofA* auto-regulation. An allelic exchange revertant (*rofA*_R) that restored the WT *rofA* allele was generated as a control.

Using RT-PCR and RT-qPCR it was determined that *rofA* transcripts could be detected at exponential, transition, and early stationary phases of growth when WT 5448 was grown in THY (**Fig. 3.1A and C**) and that the expression level remained relatively consistent at each of these growth phases (**Fig. 3.1B**). Therefore, RNA-seq was performed on RNA isolated from WT, the Δ*rofA_sc* mutant, and the *rofA*_R revertant at all three phases of growth (**Fig. 3.2**). RofA was found to impact the expression of the most genes in the exponential phase of growth (**Fig. 3.2B**). In this phase, a total of 219 genes were significantly differentially regulated (**Fig. 3.2A, Table 3.1, Appendix A2**), 153 of which were significantly downregulated and 66 of which

were upregulated. A total of 39 genes and 23 genes were significantly differentially regulated in the transition and stationary phases, respectively (**Figs. 3.2B & 3.3**).

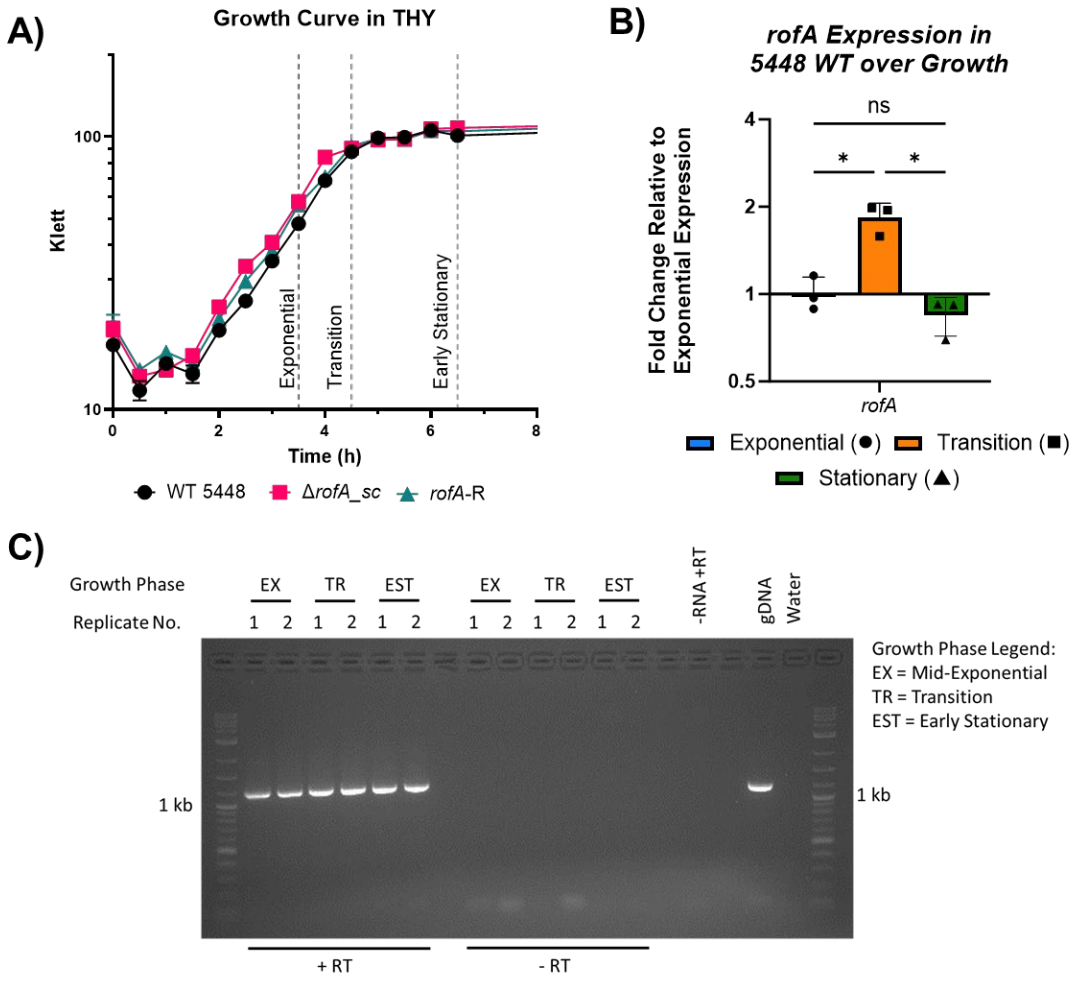


Figure 3.1: *rofA* Transcript Is Detected at All Time Points
(A) Representative growth curve of 5448, mutant Δ*rofA_sc* and revertant, *rofA_R* in THY. Gray dashed lines indicated the collection time points at exponential, transition, and early stationary phases of growth **(B)** RT-qPCR of *rofA* in WT 5448 at each of the growth phases indicated. Analysis was done using the –ΔCt' method. Significance was determined using a Brown-Forsythe and Welch ANOVA and Dunnett's T3 test for multiple comparisons. (*) p-value < 0.05. **(C)** Endpoint RT-PCR for *rofA* transcripts in RNA isolated from M1T1 5448 (WT) at exponential (EX), transition (TR), and early stationary (EST) growth phases. No RT (-RT), no RNA (-RNA), and water were used as negative controls and gDNA was used as a positive control.

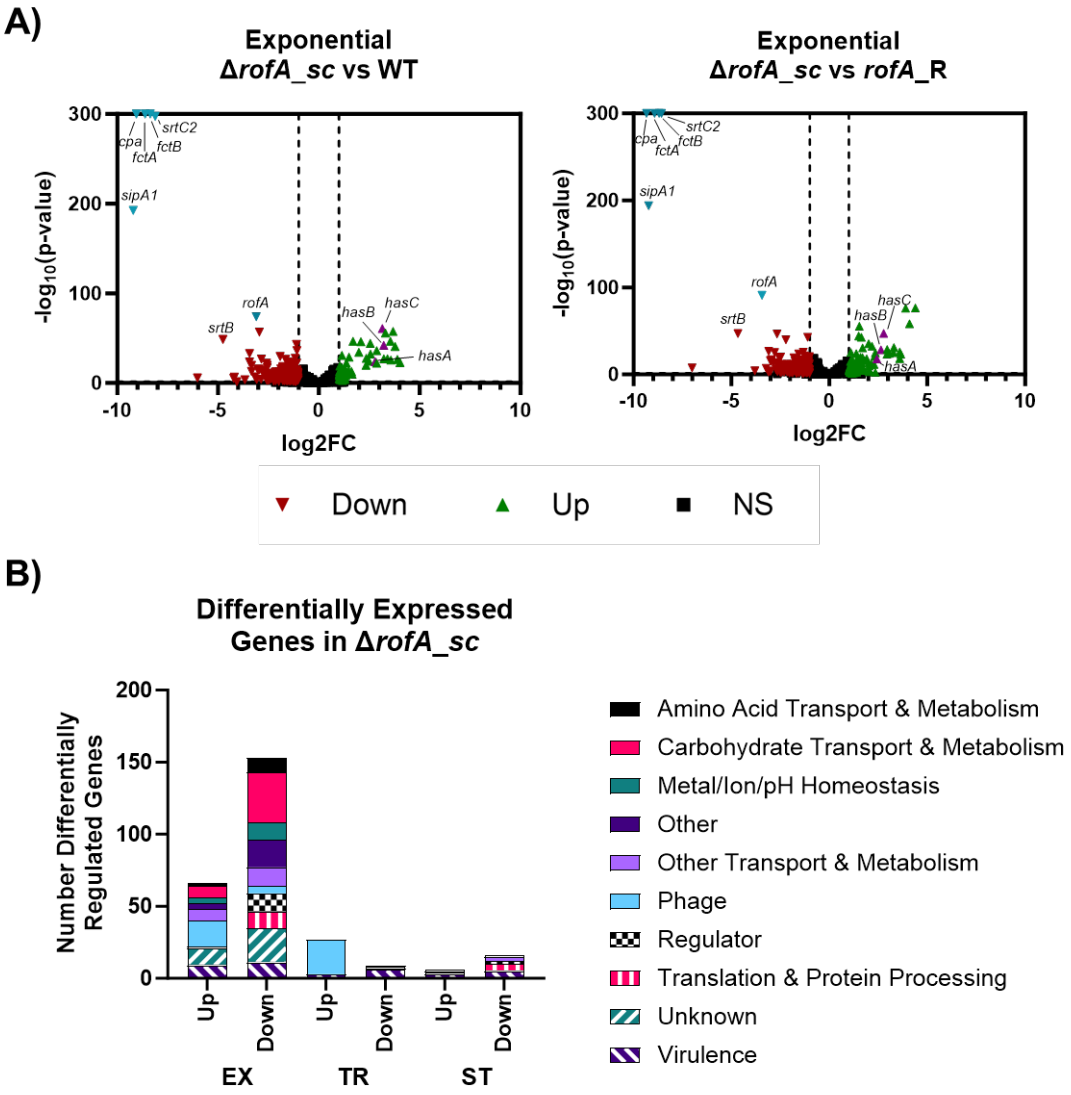


Figure 3.2: Δ*rofA_sc* Mutants Have Differentially Expressed Genes Involved with Sugar Metabolism and Virulence

(A) Volcano plots of the WT versus the mutant and the revertant versus the mutant at exponential growth phase. Pilus biogenesis genes and *rofA* are shown in blue and capsule synthesis genes are shown in purple. Significantly regulated genes were defined as those that were $\log_2$FC of ±1 and p-value <= 0.05. Genes that were significantly differentially upregulated are shown as green triangles, and those that are significantly downregulated are shown as red inverted triangles. Genes that were not significantly (NS) differentially regulated are shown as black squares. **(B)** Summary bar graph of gene groups differentially regulated in Δ*rofA_sc* mutant at each growth phase. Growth curves and collection time points can be seen in Fig. 3.1A.

Table 3.1: Virulence Genes and PTS Systems Differentially Regulated in Δ*rofA_sc* Mutant

			Δ*rofA_sc* v WT		Δ*rofA_sc* v Revert	
Spy#	**Gene**	**Annotation**	**log_2FC**	**adj-p**	**log_2FC**	**adj-p**
Spy0106	*rofA*	transcriptional regulator	-3.104	1.22E-74	-3.434	9.08E-92
Spy0107	*cpa/cbp*	collagen binding protein; minor pilus subunit	-9.055	0.00E+00	-9.336	0.00E+00
Spy0108	*sipA1*	signal peptidase I	-9.21	2.58E-193	-9.232	9.17E-195
Spy0109	*fctA*	major pilus subunit	-8.624	0.00E+00	-8.928	0.00E+00
Spy0110	*srtC2*	srtB family sortase	-8.122	3.59E-298	-8.577	0.00E+00
Spy0111		hypothetical protein	-8.367	0.00E+00	-8.692	0.00E+00
Spy0114	*srtB*	sortase	-4.756	2.32E-49	-4.669	1.11E-47
Spy0148		PTS system, 3-keto-L-gulonate IIC	-1.99	1.05E-02	-1.879	1.68E-02
Spy0149		PTS system, 3-keto-L-gulonate IIB	-2.564	7.41E-04	-2.307	2.74E-03
Spy0150		PTS system, 3-keto-L-gulonate IIA	-2.137	7.38E-06	-1.843	1.47E-04
Spy0216		putative membrane spanning protein	-2.771	6.77E-13	-2.521	8.03E-11
Spy0217	*nanH*	N-acetylneuraminate lyase	-3.441	5.04E-24	-2.964	7.26E-18
Spy0218		N-acetylmannosamine kinase	-2.567	1.38E-17	-2.046	2.35E-11
Spy0404	*stcA*	quorum sensing protein; lysozyme resistance	-1.171	5.54E-05	-1.067	3.04E-04
Spy0477	*scfA*	8-transmembrane protein; subcut. fitness gene A	1.14	1.45E-31	1.528	2.78E-56
Spy0478	*scfB*	2-transmembrane protein; subcut. fitness gene B	1.059	8.94E-07	1.597	4.51E-14
Spy0781	*ptsB*	PTS system, mannose/fructose family IIB	-1.456	5.06E-19	-1.336	4.26E-16
Spy0782	*ptsC*	PTS system, mannose/fructose family IIC	-1.476	5.64E-16	-1.381	4.21E-14
Spy0783	*ptsD*	PTS system, mannose/fructose family IID	-1.207	1.11E-09	-1.267	1.54E-10
Spy0798		IFN-response binding factor 1	-4.222	6.49E-06	-3.111	1.25E-03
Spy1078		hypothetical protein	-1.805	2.58E-05	-1.558	3.91E-04
Spy1079		PTS system, cellobiose IIC	-2.074	8.18E-13	-1.966	1.45E-11
Spy1080		hypothetical protein	-1.851	1.86E-06	-1.443	3.12E-04
Spy1081		PTS system, cellobiose IIA	-1.554	2.66E-04	-1.803	1.91E-05
Spy1082		PTS system, cellobiose IIB	-2.47	1.78E-10	-2.208	2.43E-08
Spy1083		transcription antiterminator, BglG family	-1.789	1.77E-10	-1.835	6.08E-11
Spy1109	*inlA*	internalin protein	-1.614	2.54E-10	-1.309	4.47E-07
Spy1114		protease PrsW	1.472	1.74E-14	1.015	1.95E-07
Spy1407	*sse*	secreted esterase	-1.272	4.28E-07	-1.382	3.64E-08
Spy1540	*endoS*	endo-b-N-acetylglucosaminidase F2 precursor	1.239	1.51E-14	1.447	1.41E-19
Spy1541		hypothetical protein	1.282	9.58E-06	1.692	4.82E-09
Spy1542	*scrA*	PTS system, Sucrose IIABC	1.279	9.87E-08	1.584	2.78E-11

Table 3.1 continued…

			Δ*rofA_sc* v WT		Δ*rofA_sc* v Revert	
Spy#	**Gene**	**Annotation**	**log_2FC**	**adj-p**	**log_2FC**	**adj-p**
Spy1632	*lacG*	6-phospho-beta-galactosidase	3.779	1.30E-41	3.297	6.79E-32
Spy1633	*lacE*	PTS system, lactose IIBC	4.026	1.43E-23	3.596	6.01E-19
Spy1634	*lacF*	PTS system, lactose IIA	3.888	4.03E-27	3.654	4.18E-24
Spy1635	*lacD.2*	tagatose-bisphosphate aldolase	3.575	2.33E-26	3.591	1.81E-26
Spy1636	*lacC.2*	tagatose-6-phosphate kinase	3.404	6.86E-28	3.27	9.05E-26
Spy1637	*lacB.2*	galactose-6-phosphate isomerase lacB subunit	3.233	6.42E-28	3.064	3.61E-25
Spy1638	*lacA.2*	galactose-6-phosphate isomerase lacA subunit	2.869	6.34E-23	2.944	4.70E-24
Spy1662		putative transport protein	-2.156	4.92E-09	-1.609	2.29E-05
Spy1663		PTS system, IIB	-1.604	2.33E-02	-1.669	1.85E-02
Spy1664		transcription antiterminator, BglG family	-1.88	4.97E-13	-1.494	1.91E-08
Spy1702	*smeZ*	mitogenic exotoxin Z	-1.13	4.70E-10	-1.192	4.94E-11
Spy1734	*spi*	streptopain SpeB inhibitor	3.682	3.70E-58	4.388	3.78E-77
Spy1735	*speB*	pyrogenic exotoxin B, cysteine protease	3.31	1.42E-56	3.891	5.10E-77
Spy1783	*dexS*	trehalose-6-phosphate hydrolase	-2.955	1.63E-57	-2.677	3.45E-47
Spy1784		PTS system, trehalose IIBC	-2.565	2.51E-20	-2.59	1.10E-20
Spy1851	*hasA*	hyaluronan synthase; Capsule	2.792	1.79E-24	2.425	1.167E-18
Spy1852	*hasB*	UDP-glucose 6-dehydrogenase; Capsule	3.219	7.61E-43	2.64	4.355E-29
Spy1853	*hasC*	UTP-glucose-1-phosphate uridylyltransferase; Capsule	3.154	1.36E-61	2.773	5.19E-48

Initially, the genes were categorized based on their cluster of orthologous gene (COG) designation (**Fig. 3.4**), however, many genes were labeled as 'Undefined' even though there is literature describing their function and roles in GAS pathogenesis, including the FCT-region genes for pilus synthesis (109). To improve our understanding of the RNA-Seq data, categories were manually curated based upon available data on whether the gene belonged to a putative operon, the COG of the gene, the homology of the gene to others with functional designations, and any functional characterization of homologs in the literature.

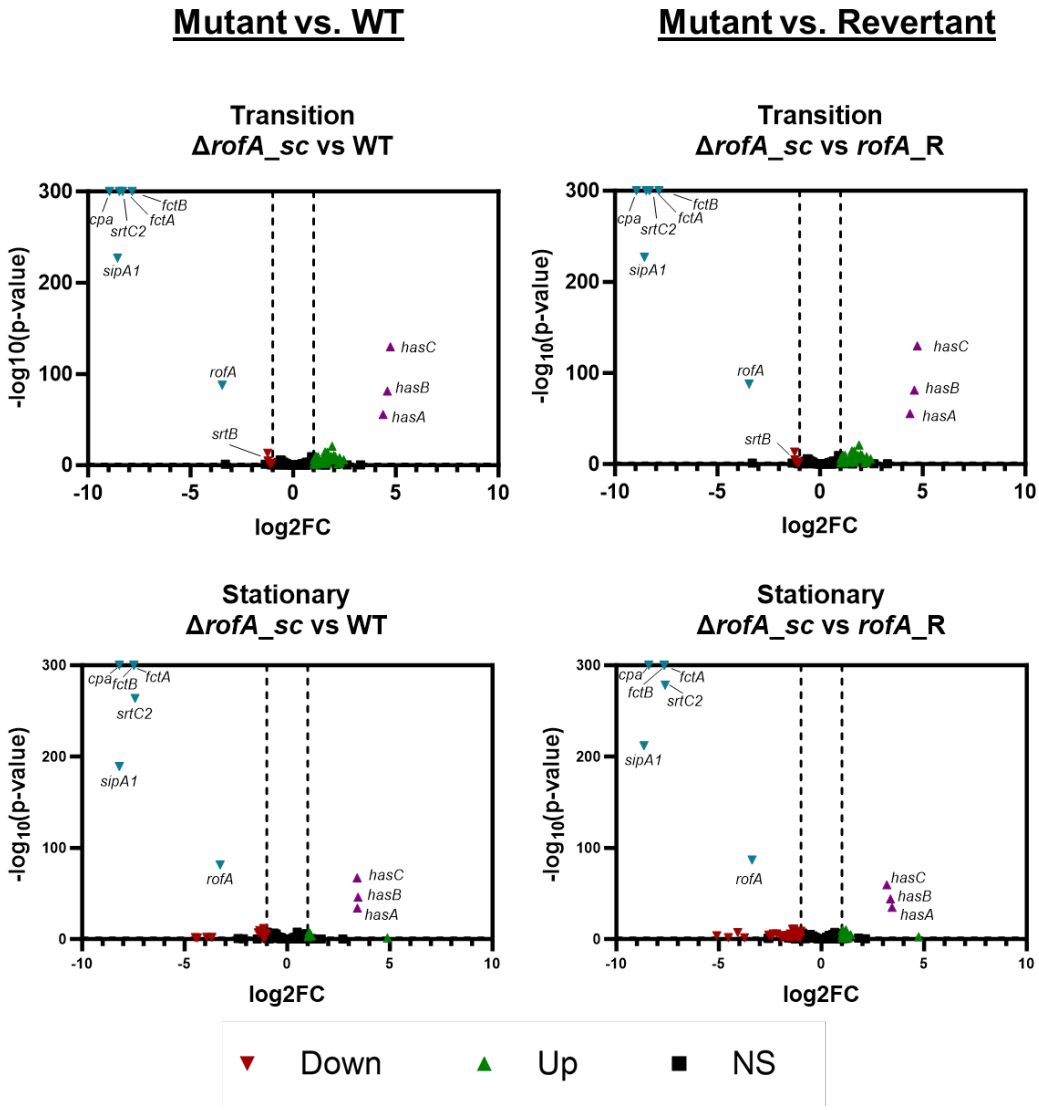


Figure 3.3: Δ*rofA_sc* RNA-Sequencing Volcano Plots from Transition and Early Stationary
Volcano plots of the WT versus the mutant and the revertant versus the mutant at transition and stationary growth phases as determined by DESeq2. Pilus biogenesis genes and *rofA* are shown in blue and capsule synthesis genes are shown in purple. Significantly regulated genes were defined as those that were $\log_2$FC of ±1 and p-value <= 0.05. Genes that were significantly differentially upregulated are shown as green triangles, and those that are significantly downregulated are shown as red inverted triangles. Genes that were not significantly (NS) differentially regulated are shown as black squares.

Using this method, it was found that the biggest group of genes downregulated in the *rofA_sc* mutant were those involved with carbohydrate transport and metabolism and the biggest group of upregulated genes were phage-associated genes (**Fig. 3.2B**). It is notable that there were 20 virulence-associated genes that were differentially

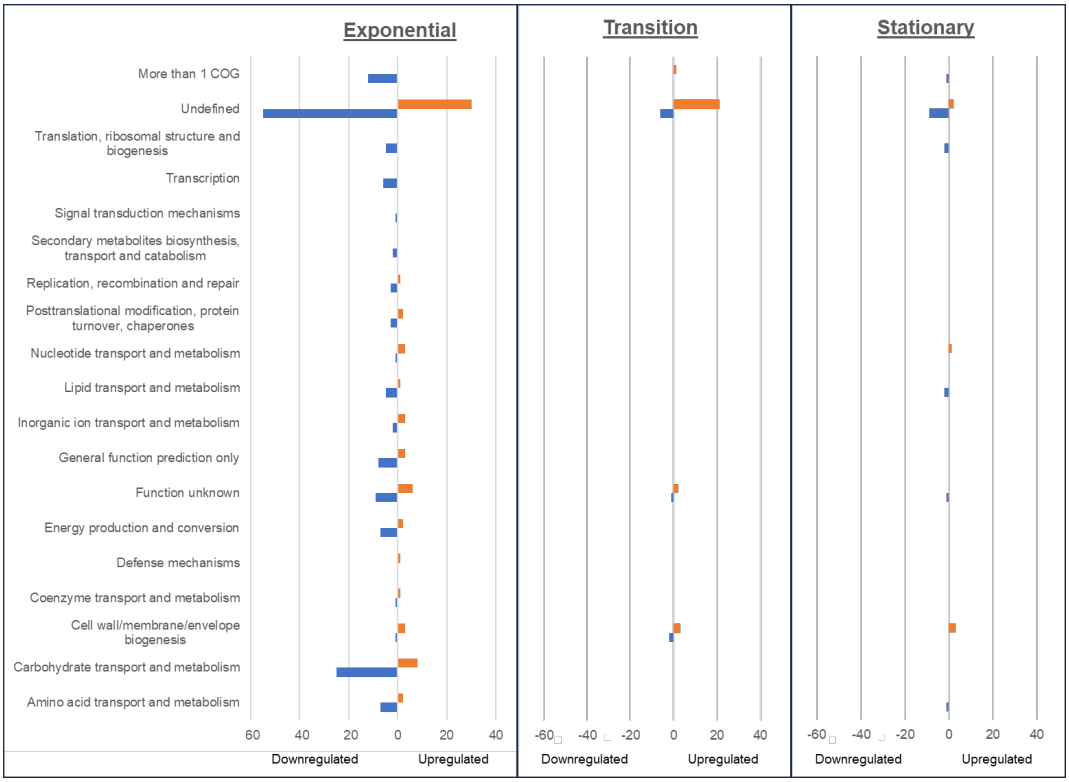


Figure 3.4: RNA-Seq COG Categorization
Bar graph of the number of genes for each category as defined by COG. Any gene without a COG category was listed as undefined. Number of downregulated genes for each category shown in blue. Number of upregulated genes for each category shown in orange.

regulated (11 downregulated and 9 upregulated) in the Δ*rofA_sc* mutant (**Fig. 3.2B**). This demonstrates that RofA plays an important role in regulating both carbohydrate metabolism genes and virulence factors in M1T1 5448 during exponential growth.

Genes impacted in transition growth phase were predominantly phage-associated genes (n=23) all of which were upregulated in the mutant (**Fig. 3.2B**). In the stationary phase, the biggest group of differentially regulated genes were virulence genes (n=8) that were both up and downregulated (**Fig. 3.2B**). Across all phases of growth, the FCT-region genes for pilus biogenesis were consistently downregulated in the mutant (**Fig. 3.2A**, **Fig. 3.3**), which agrees with the findings from previous studies (112, 113). The use of a stop-codon mutant allowed for the observation that *rofA* transcript is also downregulated in the mutant in all growth phases (**Fig. 3.2A**, **Fig. 3.3**),

suggesting RofA positively autoregulates its own expression. Interestingly, in the Δ*rofA*_sc mutant of M1T1 5448, the capsule synthesis operon (*hasABC*) was consistently upregulated in all growth phases (**Fig. 3.2A**, **Fig. 3.3**). This contrasts with the finding that *hasB* is downregulated in a *rofA* mutant in an M2 serotype strain (113). In a recent study from the Sriskandan group, only 1 of the three $M1_{global}$ strains studied (H1499) showed the capsule operon upregulated in the *rofA* knockout (170) demonstrating significant inter-strain variability.

Loss of RofA Impacts Capsule Expression, Adherence, and Aggregation in M1T1 5448

Capsule and pilus are both critical virulence factors for *S. pyogenes* (109, 171). Both biogenesis operons were consistently differentially regulated across growth (**Fig. 3.2A, Fig. 3.3**) and the differential regulation at exponential phase was confirmed by RT-qPCR (**Fig. 3.5**). To determine the impacts of the differential expression of the capsule operon and the pilus FCT region on the physiology of the Δ*rofA_sc* mutant, colony morphology was assessed, and capsule and adherence assays were performed.

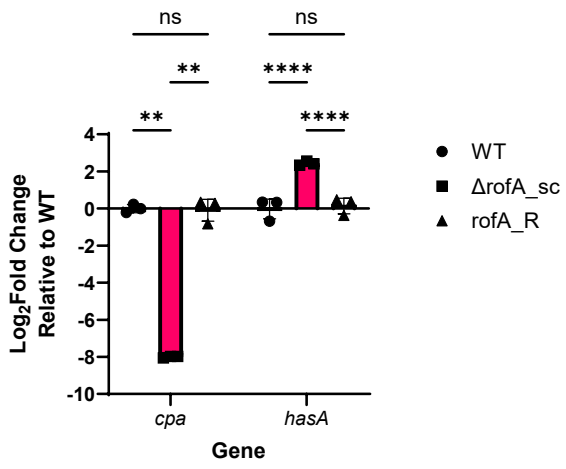


Figure 3.5: Differential Expression of Capsule and Pilus Operons Confirmed by RT-qPCR
The first gene of each operon (capsule, *hasA*; pilus, *cpa*) was used for RT-qPCR. Data represents 3 biological replicates run in technical duplicates. Significance was determined with a Two-Way ANOVA with Tukey's test for multiple comparisons (**) $p < 0.01$, (****) $p < 0.0001$.

On blood agar plates, the Δ*rofA_sc* mutant had an altered colony phenotype compared to both WT 5448 and the *rofA*-R revertant. Colonies of Δ*rofA_sc* were more

matte, consistent with a mucoid phenotype (171) (**Fig. 3.6A**). To quantify the capsule of these strains, each were grown to late-logarithmic phase (OD_{600} ~ 0.4-0.5) in THY and assayed for hyaluronic acid production using a previously established Stains-All method (162). MGAS5005 (*covS*) was included as a positive control given its upregulation of the capsule operon and large increase in hyaluronic acid capsule production (**Fig. 3.6B**) (172). The Δ*rofA_sc* mutant made significantly more hyaluronic acid than WT and the revertant (**Fig. 3.6B**), correlating to the increased transcripts seen by RNA-seq (**Table 3.1**).

Both capsule and pili have been shown to impact the ability of GAS to adhere to human tissues (173-178). Streptococcal pili have been shown to be important adhesins for GAS and impact its ability to adhere to various epithelial cell lines and tissues (173, 174). The effect of capsule on adherence is more complex. In many cases, the presence of capsule on the bacterial surface negatively impacts its ability to adhere to epithelial cells (175-178). However, human receptor CD44 has been shown to bind streptococcal capsule hyaluronic acid (179), and can positively impact adhesion in the absence of another streptococcal adhesin, the M-protein (178). To assess the impact of RofA on the ability of GAS to adhere to host cell tissues, an adhesion assay to primary human keratinocytes was performed. The Δ*rofA_sc* mutant exhibited significantly lower adhesion compared to the WT and the revertant strains (**Fig. 3.6C**). Interestingly, it also showed less adherence than a polar *cpa*::KRMIT transposon mutant that has a disrupted pilus operon (**Fig. 3.6C**), suggesting that additional RofA-regulated factors contribute to the adherence to primary keratinocytes of WT 5448.

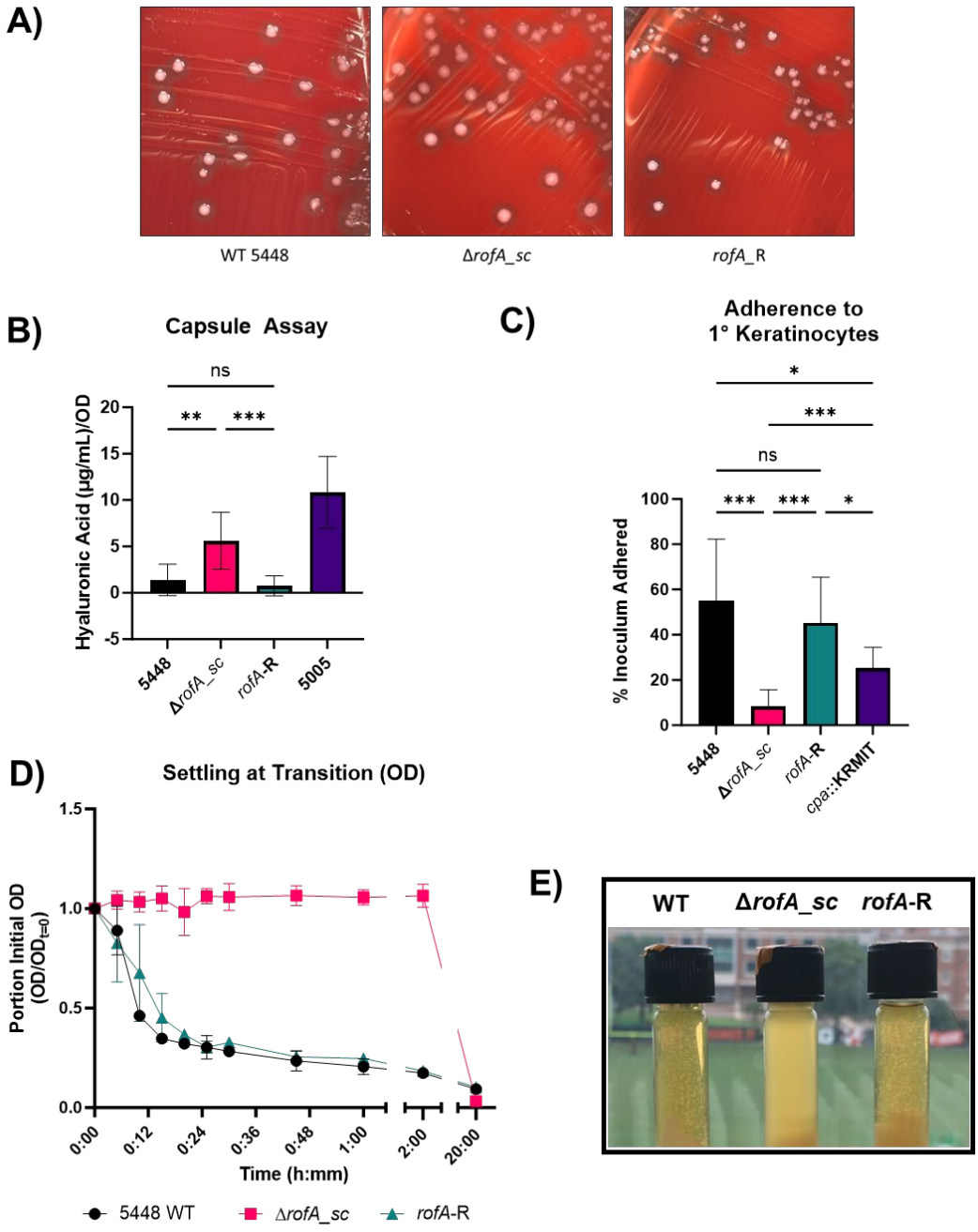


Figure 3.6: Elimination of RofA Impacts Capsule Production and Adhesion
(A) Colony morphology of each strain struck out onto blood agar plates. **(B)** Stains-All capsule assay of each strain with 5005 as a positive control. Data for 5448, Δ*rofA_sc*, and *rofA*-R represents 4 independent experiments, with 3 biological replicates per strain, per experiment. Data for 5005 represents 3 independent experiments with 3 biological replicates per strain. **(C)** Adherence assay with primary keratinocytes. Data represents two independent experiments with 6 biological replicates each. Adherence is reported as the percentage of the initial inoculum that was recovered after washing and disrupting the monolayers to recover the adhered bacteria. **(D)** Settling/precipitation of late exponential/transition phase culture of 3 strains after resuspension was assessed over 20 h time frame. **(E)** Representative image of the cultures at approximately 15 minutes after resuspension. All statistics presented are from Brown-Forsythe and Welch ANOVAs with Dunnett's T3 test for multiple comparisons. (*) p-value < 0.05, (**) p-value < 0.005, (***) p-value < 0.001.

Pilus also impacts the ability of GAS to aggregate in solution and to form microcolonies on epithelial surfaces (173) and capsule impacts the ability of GAS to aggregate in human saliva (180). During growth in THY, it was observed that in the late-exponential growth phase, the Δ*rofA_sc* mutant stayed suspended in the growth medium whereas the wild-type and revertant strains aggregated and dropped to the bottom of the growth tube (**Fig. 3.6E**). Quantification with a settling assay measuring the change in absorbance over time after resuspension (**Fig. 3.6D**) confirmed that WT and revertant strains rapidly aggregate and settle within 15 min, whereas the Δ*rofA_sc* mutant stayed suspended in the growth medium even after 2 h (**Fig. 3.6D**). This lack of aggregation may be due to the downregulation of the pilus operon or upregulation of the capsule operon, but it's not clear which of these, or if other surface components, contribute to this phenotype.

RofA is Phosphorylated *in vitro* and *in vivo*

RofA is homologous to the global GAS virulence regulator, Mga, a paradigm PRD-containing virulence regulator (PCVR) that can be phosphorylated on conserved histidines by the phosphotransferase system (PTS) in response to carbohydrate availability (86, 95, 96). Purified Mga has been shown to be phosphorylated by the general PTS proteins EI and HPr *in vitro* using a reconstituted PTS assay (95). Given that RofA is a paralog of Mga (86), it was hypothesized that RofA would also be phosphorylated in this system. The *in vitro* PTS phosphorylation assay (**Fig. 3.7A**) was performed on purified RofA with a C-terminal 6x-His tag (RofA-His). The additional band not labeled in the phosphorimage is believed to be either a contaminating protein or a RofA degradation product and was observed in multiple experiments. Despite

loading equal amounts of EI to each reaction, the signal of phosphorylated EI qualitatively decreased in reactions lacking HPr (**Fig. 3.7B, lane 1&5**). It is not clear

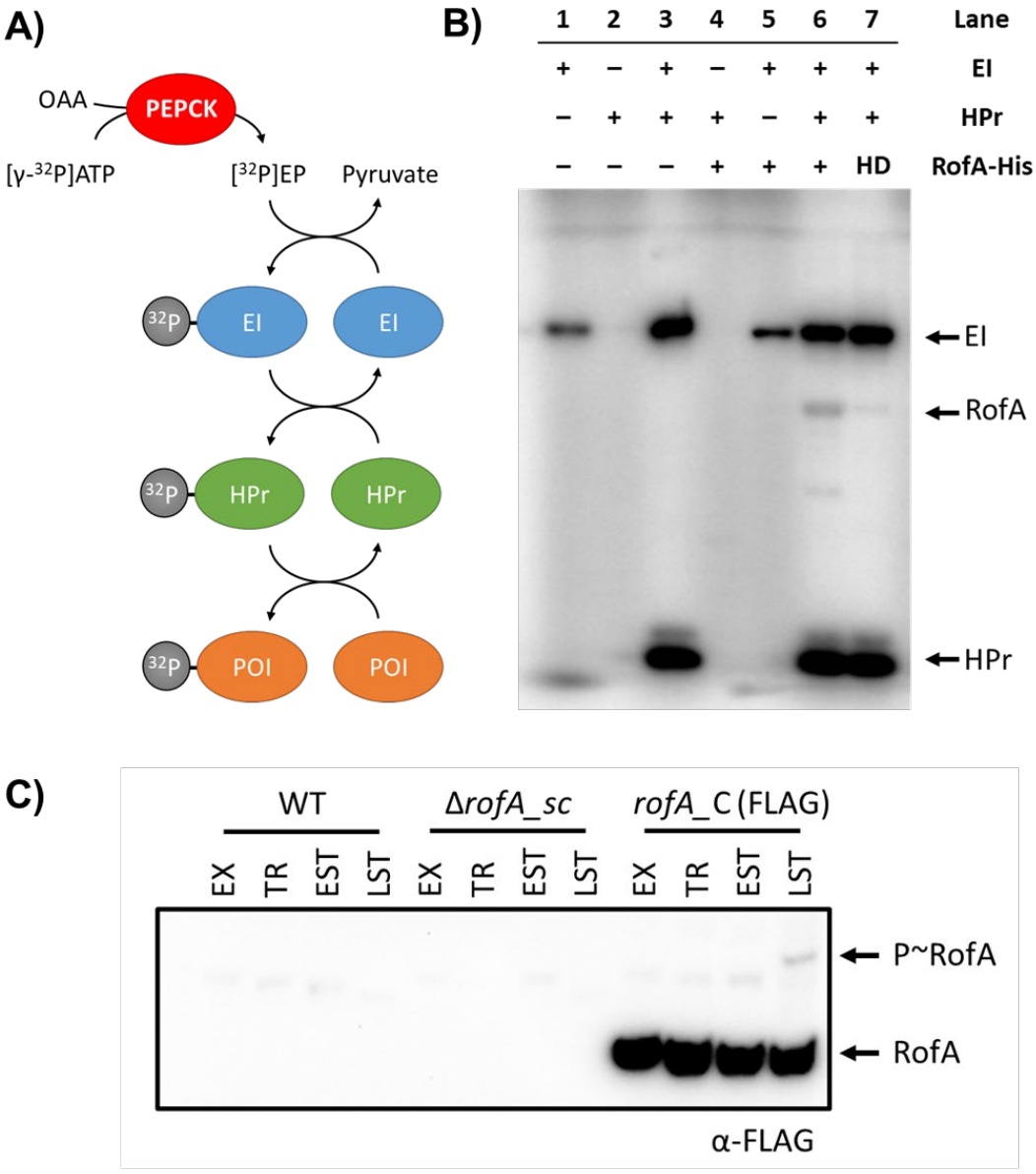


Figure 3.7: RofA Is Phosphorylated *in vitro* and *in vivo*
(A) Schematic of the *in vitro* phosphorylation assay. All proteins were tagged with a 6X-His tag for affinity purification. **(B)** Phosphorimage of the PTS reactions run on an SDS-PAGE gel. HD indicates the protein was heat-denatured by incubating at 99°C for 10 minutes before addition to the reaction. **(C)** Anti-FLAG western blot of a Phos-Tag gel of the whole-cell lysates from WT or Δ*rofA_sc* carrying empty vector pJRS525 as a negative control and *rofA*_C(FLAG) which is Δ*rofA_sc* carrying plasmid pProfA1-FLAGt. Each strain was collected at mid-exponential (EX), transition (TR), early stationary (EST), or late stationary (LST) growth phases. See Fig. 3.9A for growth curve and collection points.

why this occurred, but it is possible that HPr has a stabilizing effect on phosphorylated EI. Phosphorylated RofA-His was observed only when [^{32}P]-EP, EI, and HPr were all present (**Fig. 3.7B, lane 6**) and heat denaturing RofA-His resulted in a significant reduction in signal (**Fig. 3.7B, lane 7**). This result indicates that RofA-His can be phosphorylated *in vitro* by EI and HPr and that the appropriate tertiary structure is important for phosphorylation.

Since the *in vitro* PTS phosphorylation assay does not reflect conditions inside the bacterial cell, the next step was to determine if RofA was phosphorylated *in vivo*. A plasmid was generated with the native promoter region, including the predicted *rofA* promoter and the RofA binding sites (111), the coding sequence for *rofA* with a C-terminal FLAG-tag, and the native 3' intergenic region containing a hypothesized *rho*-independent terminator as predicted by the PePPER webserver (181). The plasmid was transformed into the Δ*rofA_sc* mutant to generate the complementation strain 5448 *rofA*_C (FLAG). RT-qPCR of the first gene of the pilus operon, *cpa*, confirmed the epitope tagged RofA in *rofA*_C (FLAG) was functional to complement the mutant (**Fig. 3.8**). 5448 WT and Δ*rofA_sc* strains were transformed with the empty vector, pJRS525, such that all strains could be maintained with spectinomycin selection. Strains were grown in C

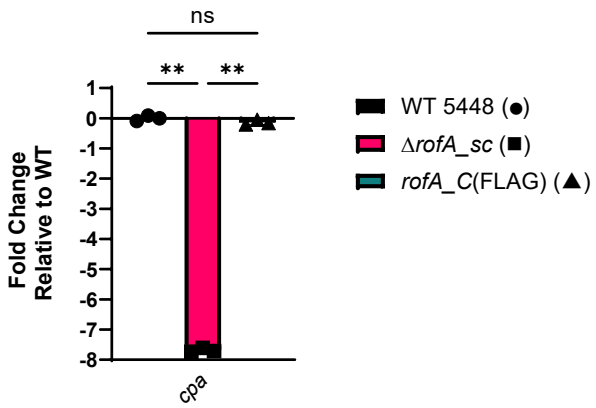


Figure 3.8: Complementation with FLAG-tagged RofA Restores *cpa* Expression
RT-qPCR for *cpa*, the first gene of the pilus operon, with WT and Δ*rofA_sc* carrying empty vector pJRS525 and *rofA_C* (FLAG) with Δ*rofA_sc* carrying pProfA1-FLAGt. Significance determined by Brown-Forsythe and Welch ANOVA with Dunnett's T3 test for multiple comparisons. (**) $p < 0.01$.

medium until they reached exponential, transition, early stationary, or late stationary phase of growth (**Fig. 3.9A**). Whole cell lysates were run on a Phos-Tag polyacrylamide gel and analyzed by western blot with anti-FLAG-tag antibodies (**Fig. 3.7C**). A shifted species of RofA-FLAG was observed in the late stationary phase of growth, suggesting RofA-FLAG was phosphorylated *in vivo* at this time point. No bands were observed in the negative control WT and Δ*rofA_sc* containing an empty vector, indicating the bands were specific to RofA-FLAG. The faint bands seen in the other growth phases were also found in the WT 5448 samples and the 5448 Δ*rofA_sc* samples, suggesting a cross-reactive protein at these timepoints.

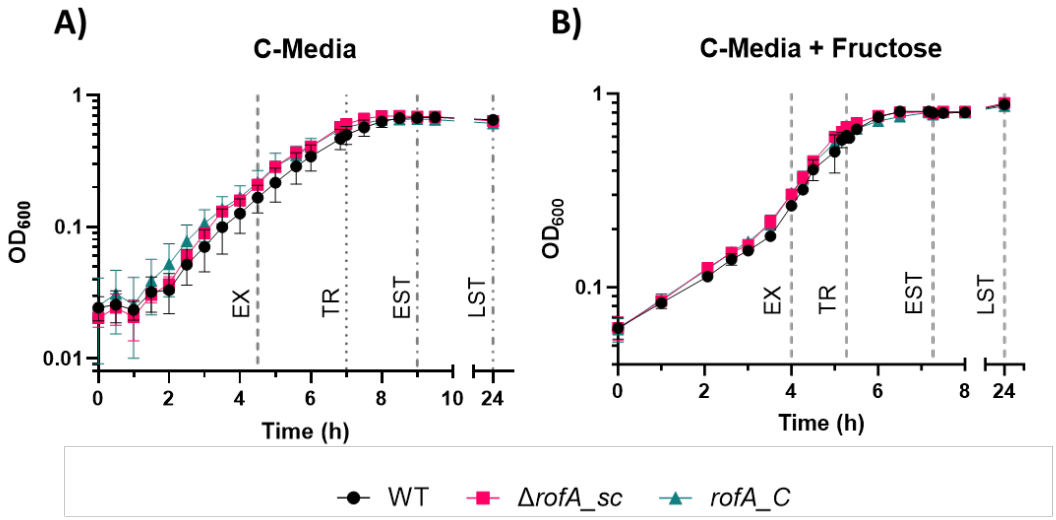


Figure 3.9: Growth Curves in C Medium and C Medium + 1% Fructose
Representative growth curve of strains 5448 (WT) and Δ*rofA_sc* (Δ*rofA*_sc) carrying the empty vector pJRS525, and Δ*rofA_sc* with the plasmid pProfA1-FLAGt (*rofA_C* (FLAG)) grown in **(A)** C medium or **(B)** C medium + 1% fructose. Sample collection points are shown with dashed gray lines. Mid Exponential at OD_{600} = ~0.3) (EX), Transition at OD_{600} = ~0.6 (TR), Early stationary – defined as 2 h after transition reached (EST), and late stationary (LST) at 24 h.

<u>RofA-FLAG Phosphorylation *in vivo* in C Medium is Dependent upon Glucose Concentration</u>

C medium is a low-glucose, high-peptide growth medium with approximately 30 mg/dL of glucose (96). If RofA is differentially phosphorylated depending on sugar

availability and metabolism, then changing the glucose concentration could alter the phosphorylation state. To test this, C medium was supplemented with glucose to a final concentration of 300 mg/dL, which is the approximate amount found in THY (96). Growing 5448 *rofA*_C (FLAG) in this high glucose C medium resulted in the ablation of the phosphorylated RofA-species (**Fig. 3.10A**), demonstrating that glucose

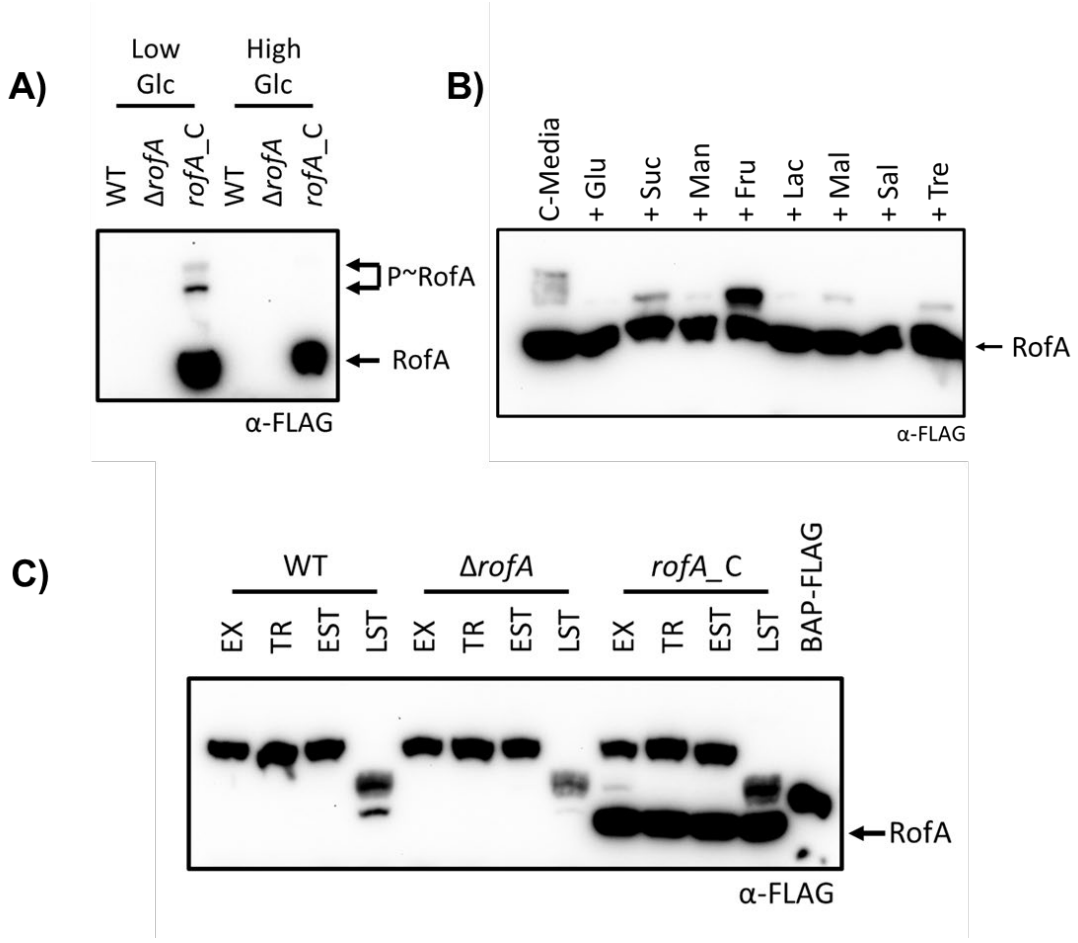


Figure 3.10: Glucose Concentration Impacts RofA Phosphorylation
Anti-FLAG western blot of a Phos-Tag gel of the whole-cell lysates from WT or Δ*rofA_sc* carrying empty vector pJRS525 as negative controls and *rofA_C* (FLAG) which is Δ*rofA_sc* carrying plasmid pProfA1-FLAGt. **(A)** Strains were grown in low glucose (Low Glc, 30 mg/dL) C medium or high glucose (High Glc, 300 mg/dL) C medium to late stationary phase (LST). **(B)** Strains were grown in C medium supplemented with 1% of an additional carbohydrate as indicated to LST phase. Glucose (Glu), Sucrose (Suc), Mannose (Man), Fructose (Fru), Lactose (Lac), Maltose (Mal), Salicin (Sal), or Trehalose (Tre). **(C)** Strains were grown in C medium plus 1% fructose to each of the growth phases indicated: mid-exponential (EX), transition (TR), early stationary (EST), or late stationary (LST). See Fig. 3.9B for growth curves and collection points. BAP-FLAG (Sigma) was used as a FLAG-positive control where indicated.

concentration impacts phosphorylation of RofA and supports the hypothesis that RofA is phosphorylated in response to changes in sugar availability.

To understand how sugars other than glucose impact the phosphorylation of RofA, Phos-Tag analysis was conducted on the lysates of *rofA*_C GAS grown in C medium supplemented with 1% of PTS sugars that are known to allow the growth of M1T1 strain MGAS5005 (39). Interestingly, C medium + 1% ascorbic acid inhibited the growth of 5448 (**Fig. 3.11**), despite the fact that 5448 can grow in C medium without any supplemental sugar (**Fig. 3.9A**) and MGAS5005 can grow in chemically defined medium (CDM) with ascorbic acid as the sole carbon source (39). It was shown, however, that growth was inhibited when ascorbic acid was added to 1.5% in a basal medium comprised of a phosphate buffered tryptone solution (182), which seems similar to C medium.

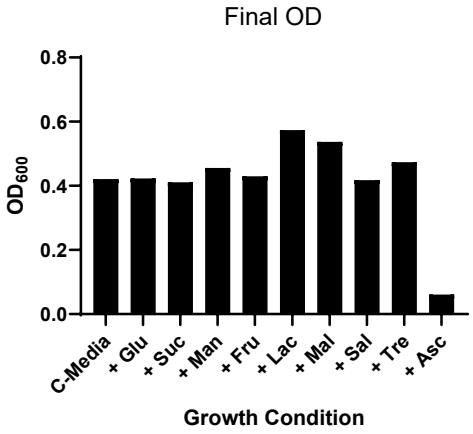


Figure 3.11: Final OD_{600} of *rofA_C* (FLAG) Grown in C Medium + PTS Sugars
OD_{600} measurements for *rofA_C* (FLAG) at 24 h when grown in C medium supplemented with 1% of the sugars indicated. Glucose (Glu), Sucrose (Suc), Mannose (Man), Fructose (Fru), Lactose (Lac), Maltose (Mal), Salicin (Sal), Trehalose (Tre), or ascorbic acid (Asc). Data shown is n=1 and similar findings were seen in multiple experiments.

The PTS sugar screen suggested that there was robust phosphorylation of RofA in C medium + 1% fructose (**Fig. 3.10B**). However, when phosphorylation of RofA was further investigated in C medium + 1% fructose with WT 5448 and Δ*rofA_sc* as negative controls (**Fig. 3.10C**), it showed that the band thought to be phosphorylated RofA (**Fig. 3.10B**) was actually a FLAG cross-reactive protein. Still, it revealed another potential phosphorylation event in exponential growth phase (**Fig. 3.10C**). In the

screen, there were also shifted species observed in C medium supplemented with sucrose, maltose, and trehalose (**Fig. 3.10B**). Further investigation will be needed to determine if these species are phosphorylated RofA or another cross-reactive protein. In contrast, very low or no shifted species were observed when grown in C medium supplemented with mannose, lactose, and salicin, like the phenotype observed when C medium is supplemented with glucose.

Predicted Histidines within PRDs Have an Impact on RofA Phosphorylation

Initial analysis of RofA's amino acid sequence with the Conserved Domain Database (CDD) (183) indicated that residues 171 to 384 made up a PRD in the Mga superfamily (**Fig. 3.12**). Subsequent secondary structure analysis by Phyre2 (184) allowed for more precise analysis of where PRD histidines from paradigm PCVRs Mga and AtxA were in relation to RofA histidines (**Fig. 3.13**) and four histidines were chosen as targets (**Fig. 3.12**, red). H210 was in proximity to H199 and H204 from Mga and AtxA, respectively. H278 was found in a helix adjacent to the corresponding helix

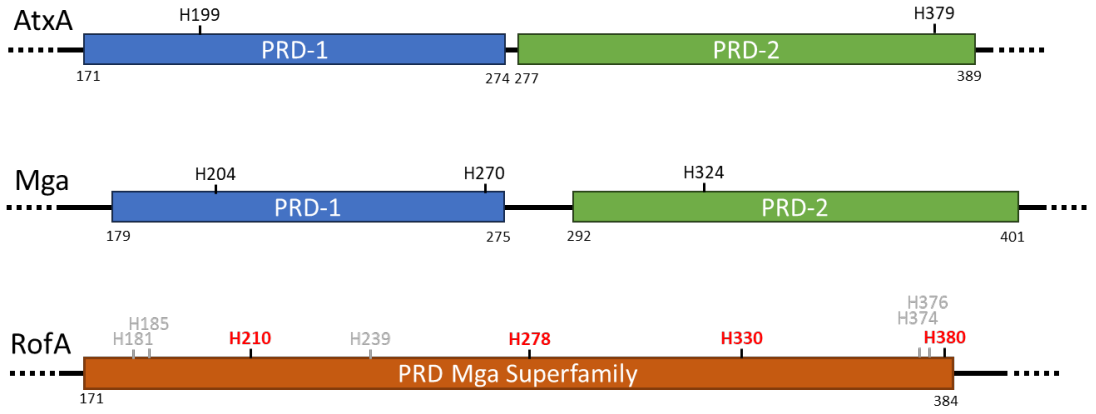


Figure 3.12: Schematic Diagram of PCVR PRDs
Histidines shown to be important for phosphorylation of AtxA and Mga are shown in black. Target histidines chosen for RofA are shown in red. Other potential histidines within the PRDs are shown in gray.

where Mga H270 was found. H330 was found just outside the corresponding helix where Mga H324 was found. And finally, H379 was near H380 from AtxA. All 4

Figure 3.13: Alignments of Phyre2 Predicted Secondary Structure of PCVR PRDs
(A) Alignment of RofA (top) to the solved structure of AtxA (bottom). Predictions for RofA histidines shown on the right. Blue arrows point to the RofA predicted histidines. **(B)** Alignment of Mga (top) to the solved structure of AtxA (bottom). Histidine residues implicated in phosphorylation of AtxA shown on the left. Histidines implicated in phosphorylation of Mga shown on the right and indicated by blue arrows. Regions where Mga histidines were located that lacked a corresponding AtxA histidine were labeled by residues that made up the helix and are shown in blue boxes.

histidines identified (H210, H278, H330 and H380) were mutated to alanine residues in the RofA coding sequence and this allele was cloned into the GAS trans-complementation construct and transformed into 5448 Δ*rofA_sc* to create strain 5448 *rofA*-HqA_C (FLAG). This strain was then analyzed for RofA phosphorylation at stationary phase in C medium by Phos-Tag electrophoresis.

There was still a shifted species observed in the *rofA*-HqA_C (FLAG) lysates, suggesting that RofA-HqA is still phosphorylated *in vivo* (**Fig. 3.14A**). However, there was qualitatively less of the shifted species in the RofA-HqA lysates, and the shifted species ran faster than the shifted species for the WT RofA-FLAG lysates. This suggests that there are less and/or different phosphorylation sites present in the RofA-HqA mutant and that one or more of the targeted histidines are phosphorylated in RofA.

To determine if a similar pattern was observed *in vitro* with purified EI and HPr, the RofA-HqA was expressed, purified, and used in the *in vitro* assay. As before, a contaminating band can be seen in the WT RofA lane but does not seem to be present in RofA-HqA protein preparation. RofA-HqA was phosphorylated to a similar level as the WT RofA (**Fig. 3.14B**), suggesting that the histidines selected are not the targets of EI/HPr phosphorylation. It should be noted, however, that in this *in vitro* assay the Mga protein appeared to have equivalent phosphorylation for single alanine substitutions at each of the residues and it was not until all three residues were mutated that there was any observation of decreased phosphorylation (95). Thus, it is also possible that one or more of the other RofA histidine residues (**Fig. 3.12**; gray) are targets of EI/HPr phosphorylation and their presence in the RofA protein confounds the *in vitro* results.

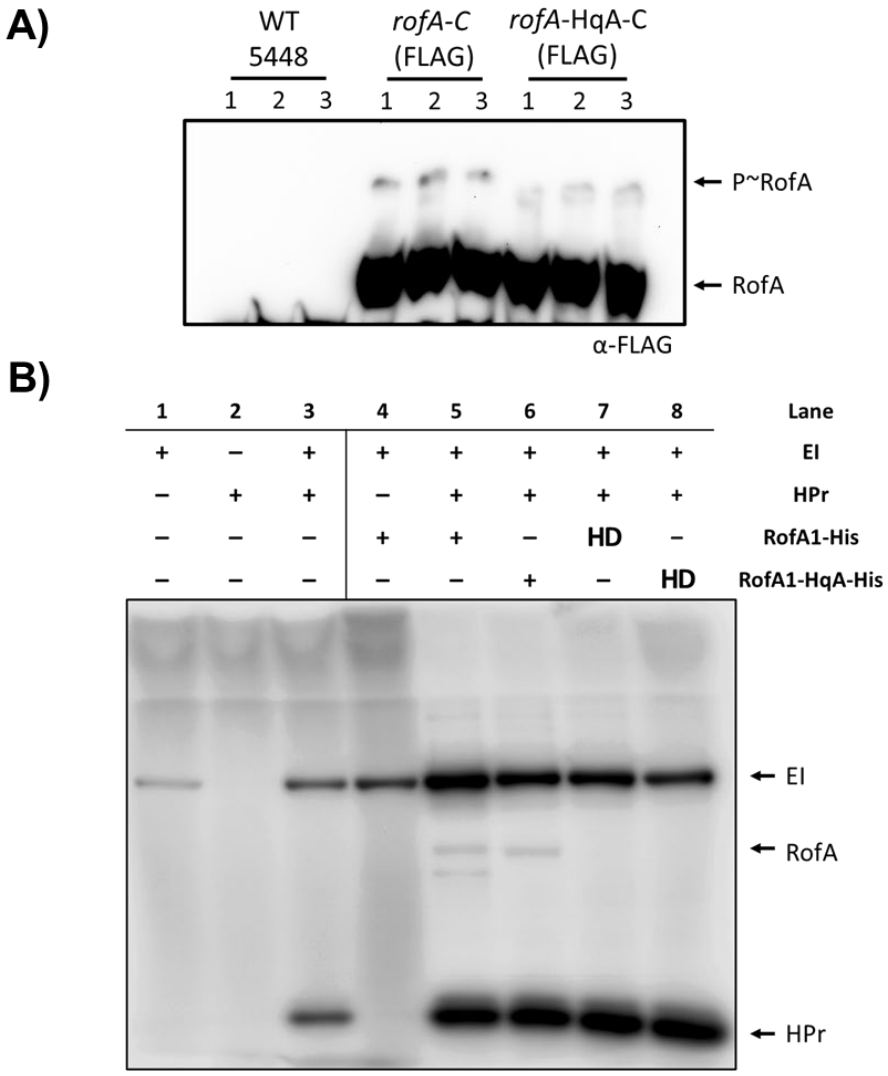


Figure 3.14: Mutation of Predicted PTS Histidines Reduces Phosphorylation *in vivo*
(A) Anti-FLAG western blot of a Phos-Tag gel of the whole-cell lysates from WT carrying empty vector pJRS525 as a negative control, *rofA-C* (FLAG) which is Δ*rofA_sc* carrying plasmid pProfA1-FLAGt, and *rofA*-HqA-C (FLAG) which is Δ*rofA_sc* carrying plasmid pProfA1-HqA-FLAGt in which the RofA coding sequence has mutations to substitute histidines 210, 278, 330, and 380 for alanine residues. All strains were collected at late stationary phase. 3 replicates of each strain are run side by side. **(B)** Phosphorimage of the *in vitro* PTS reactions run on an SDS-PAGE gel. HD indicates the protein was heat-denatured by incubating at 99°C for 10 minutes before addition to the reaction**.**

Discussion

Previous work established that RofA played important roles in virulence regulation in the M6 and M2 serotypes (112, 113), however, these studies were limited to a small set of genes and the role of RofA as a potential PCVR was unexplored. This

work has further established RofA's role as a major virulence regulator in *Streptococcus pyogenes*. RNA-seq showed that in M1T1 5448 RofA acts to regulate expression of virulence factor genes including the pilus operon, capsule operon, IgG cleavage protein *endoS*, secreted esterase *sse,* internalin *inlA, stcA,* superantigen *smeZ,* and cysteine protease *speB* during exponential growth (**Table 3.1**). Two consistently differentially regulated virulence operons in the RNA-seq were capsule and pilus (**Fig 3.2A**, **Fig. 3.3**), and the impacts of this differential gene expression were investigated. It was found that the Δ*rofA_sc* mutant strain produces significantly more capsule than WT and revertant, does not adhere as well to primary human keratinocytes, and has an aggregation defect in culture (**Fig. 3.6**). Both capsule and pilus could contribute to the adherence and aggregation phenotypes (173-178, 180) observed in the Δ*rofA_sc* mutant. It is not clear which of these factors, or a combination of these and/or other unknown factors contribute to these phenotypes.

Capsule production can contribute significantly to virulence of GAS (185-187), however, it can also have detrimental effects on adhesion to host cells (174-176, 178) and this may impact initial colonization. In contrast, pilus is critical for adhering to host cells (173, 174), but acts as a immunostimulant, activating toll-like receptor 2 (TLR-2) (117). This highlights the importance of co-regulation in determining the appropriate temporal expression of pili and capsule during infection. It was also found that the Δ*rofA_sc* mutant adheres less to keratinocytes than the *cpa*::KRMIT transposon mutant with a disrupted pilus operon. This suggests other RofA-regulated factors contribute to adherence beyond the pilus. One possibility is that the observed increase in capsule production in Δ*rofA_sc* obscured other adhesins that are important for keratinocyte

adhesion. This study reports that RofA regulates both *hasABC* and *cpa* genes in M1T1 5448, activating pilus and repressing capsule, which suggests that RofA-mediated virulence regulation is critical during colonization of GAS on epithelial surfaces.

RNA-seq analysis performed in this study showed that RofA plays an important role in regulating carbohydrate metabolism systems (**Fig. 3.2B**), thus opening the possibility that RofA directly connects the regulation of virulence with carbohydrate utilization genes. This work also provides evidence that RofA is phosphorylated by components of the PTS system *in vitro* (**Fig. 3.7B**) and becomes phosphorylated *in vivo* in a carbohydrate-concentration-dependent manner (**Figs. 3.7C & 3.10A**). The combination of these findings: that RofA plays a role in virulence factor expression and that it can be phosphorylated by PTS-proteins, supports the hypothesis that RofA is a PCVR in GAS, acting as an additional regulator to couple carbohydrate metabolism and virulence factor regulation in this pathogen.

Interestingly, this work demonstrated that an M1T1 5448 RofA mutant had the most genes differentially regulated in exponential phase (**Fig. 3.2**). It was expected that RofA would impact expression mostly in the transition growth phase, as RofA and the other RALPs were primarily thought to be transition phase regulators (78). Although, it is possible that the regulation in an infection context is different than *in vitro* growth in a rich medium. As this manuscript was being drafted, another study looking at the RofA regulon was published (170) showing that only 14 genes were found to be differentially regulated consistently across *rofA* mutants in several $M1_{global}$ strains. Of the regulated genes found, the majority were downregulated, which is consistent with the findings from this paper, where 10 of 14 genes observed in that study were also

found here to be differentially regulated in the mid-exponential growth phase. The pilus operon (Spy0106-Spy0111 & Spy0114) was downregulated in the mutants consistent with our findings and from previous literature (112-114). The cellobiose PTS IIA component (Spy1081) was also downregulated, consistent with these findings. Interestingly, the N-acetylmannosamine 6-phosphoate 2-epimerase (Spy0212) and N-acetylneuraminate-binding protein (Spy0213) that were found to be upregulated in their mutants were found to be downregulated in this study with M1T1 5448. Although there were more consistencies between 5448 and individual strains of $M1_{global}$, these findings highlight the large inter-strain variability of the RofA regulon and support published examples of inter-serotype differences (112, 113). Consistent across all studies, however, is the finding that RofA is an important regulator of the pilus operon.

The findings that RofA is phosphorylated *in vitro* and *in vivo* in a glucose concentration-dependent manner are compelling and support the hypothesis that RofA is a PRD-containing Virulence Regulator (PCVR) that responds to sugar metabolism through PTS phosphorylation. RofA is not the only PCVR that appears to be impacted by glucose concentration. Mga was less phosphorylated in the low-glucose C medium compared to the high glucose THY (96). In *Streptococcus pneumoniae*, the Mga homolog Mga*Spn* has not been shown to be phosphorylated, but the product of a Mga*Spn* targeted gene, phosphorylcholine, increased with increasing glucose concentration in a Mga*Spn*-dependent manner (131).

Of note, the changes in phosphorylation of RofA observed in this study were subtle. Only a small portion of RofA is phosphorylated compared to the unphosphorylated RofA both *in vitro* and *in vivo*. Still, the phosphorylation of RofA *in*

vitro was comparable to the levels of phosphorylation seen for Mga (95). The portion of RofA found to be phosphorylated *in vivo* was qualitatively less than that was observed in Mga (92), although phosphorylated Mga was still less abundant than the native protein. It's not clear if there is an unidentified set of conditions that would produce more robust phosphorylation of RofA, or whether the observed ratio of phosphorylated to unphosphorylated protein is a feature of PCVRs. Additionally, while several residues of interest were found within the putative PRDs that seem to reduce the *in vivo* phosphorylation of RofA, this study fell-short of identifying all the residues that are targets for phosphorylation.

Finally, the cross-reactive protein that presented itself in growth of GAS in some of the high-carbohydrate supplemented C medium (**Fig. 3.10B&C**) made it impossible to evaluate the impacts of different PTS sugars on the phosphorylation of RofA. Moving forward, a new system will need to be used to study the phosphorylation of RofA in a growth context where alternative sugars are present as the sole carbon source, instead of being present in addition to glucose and high peptides such as are found in C medium.

Chapter 4: Further Studies of RofA as a PCVR

Overview

Since the publication of the findings presented in Chapter 3, a few experiments have been conducted to further characterize RofA and the Δ*rofA_sc* mutant. The findings from these experiments are presented below. In these experiments, the contribution of capsule to the attenuated aggregation of Δ*rofA_sc* was investigated. Additionally, the *in vivo* phosphorylation of RofA was further interrogated by looking at time points between early stationary and late stationary phases of growth, assessing the phosphorylation of RofA-His_6 tag in C medium supplemented with different carbohydrates, and investigating RofA-FLAG phosphorylation in a Δ*ptsI* mutant. Finally, the tertiary structure of RofA as a monomer and homodimer was predicted by AlphaFold3.

Results

Hyaluronidase Treatment of Δ*rofA_sc* Recovers Aggregation Phenotype

The Δ*rofA_sc* mutant had differentially regulated pilus and capsule across growth (**Figs. 3.2A** & **3.3**). This correlated with a decrease in bacterial aggregation in culture (**Fig. 3.6 D&E**). To begin to understand the individual contributions of pilus and capsule to this phenotype, WT, Δ*rofA_sc*, and *rofA_R* cultures were treated with hyaluronidase to investigate the contribution of the hyaluronic acid capsule. Hyaluronidase was included in each culture at a concentration of 85 µg/mL, which is the concentration used for the growth of encapsulated strains to make competent cells

(153). Cultures with hyaluronidase grew as well as the cultures without hyaluronidase (**Fig. 4.1A**). Δ*rofA_sc* cultures with hyaluronidase aggregated and settled out of solution as quickly as the WT strain with or without hyaluronidase (**Fig. 4.1B**), suggesting that increased hyaluronic acid capsule of Δ*rofA_sc* is the main contributor to the aggregation phenotype.

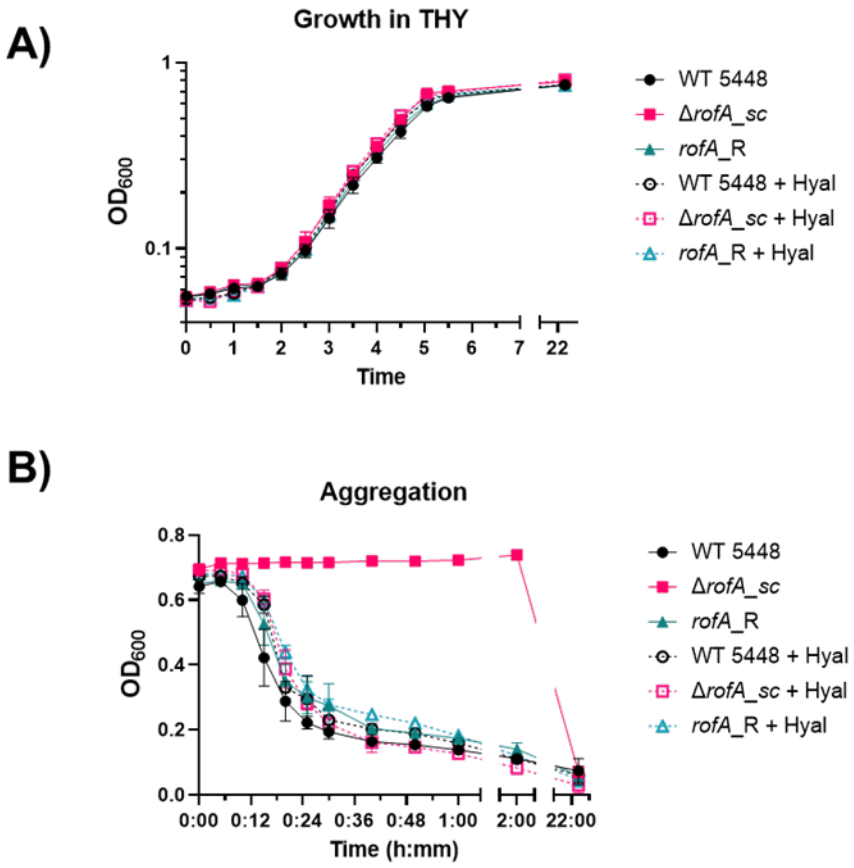


Figure 4.1: Contribution of Capsule to Δ*rofA_sc* Aggregation Phenotype
(A) Representative growth curves of strains in THY and THY plus hyaluronidase (+Hyal). **(B)** Settling/precipitation of transition phase cultures +/- hyaluronidase. Samples without hyaluronidase are indicated with filled symbols and solid lines. Samples with hyaluronidase are indicated with open symbols and dashed lines.

In vivo RofA-FLAG Phosphorylation in Stationary Phase

Previously, we identified a phosphorylated RofA-FLAG species late in stationary (LST) phase (24 h) but not in early stationary (EST) phase (2 h post TR) (**Fig**

3.7C). Therefore, phosphorylation at time points between the EST and LST time points were investigated. The average time point that the strains reached the early stationary phase was approximately 9 hours (**Fig 3.9A**), therefore this was chosen as the first time point for this experiment. Samples were taken every 5 hours thereafter, until 24 hours was reached (**Fig. 4.2A**). The phosphorylated species of RofA appeared at the 14 h time point and was maintained until the 24 h time point (**Fig. 4.2B**).

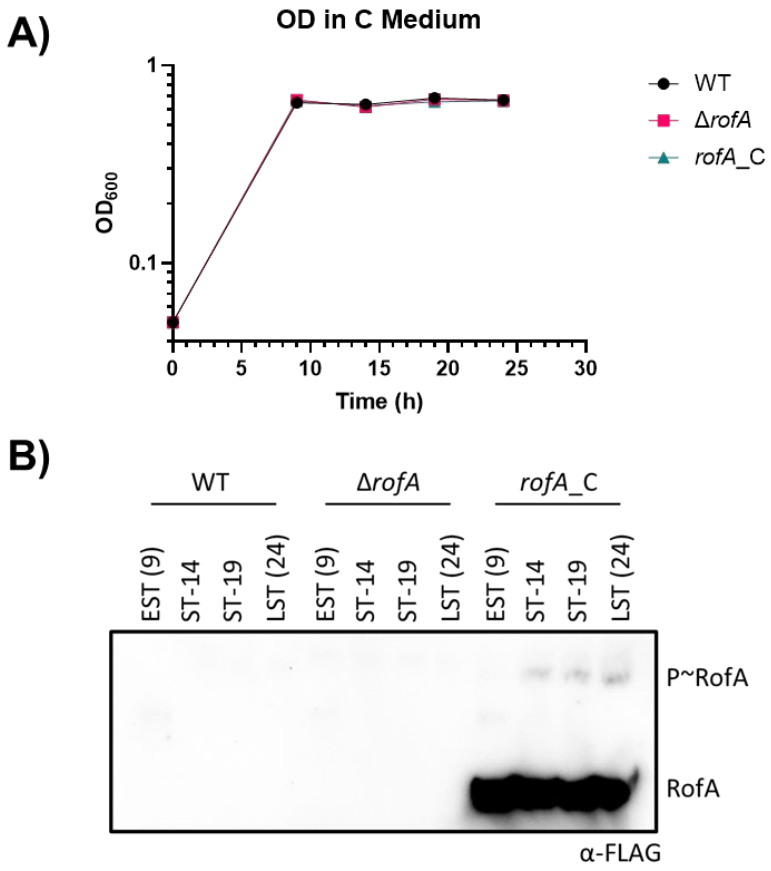


Figure 4.2: RofA Becomes Phosphorylated 14 Hours Post Inoculation in C Medium
(A) OD_{600} of the strains collected at stationary time points 9, 14, 19, and 24 h post inoculation. **(B)** Anti-FLAG western blot of a Phos-Tag gel of the lysates from the samples collected in (A).

RofA-His_6 Phos-Tag Analysis of *S. pyogenes* Grown in C Medium Plus Supplemental Sugar

Due to presence of the FLAG cross-reactive protein in *S. pyogenes* lysates from cultures grown in C medium plus 1% fructose (**Fig. 3.10C**), a 6x-His (His_6) epitope-tagged allele of *rofA* was generated and cloned into the pProfA-FLAGt vector to replace the FLAG tag with a 6x-His tag. This was transformed into Δ*rofA_sc* to generate strain *rofA_C* (His). To confirm that this strain functioned as the *rofA_C*

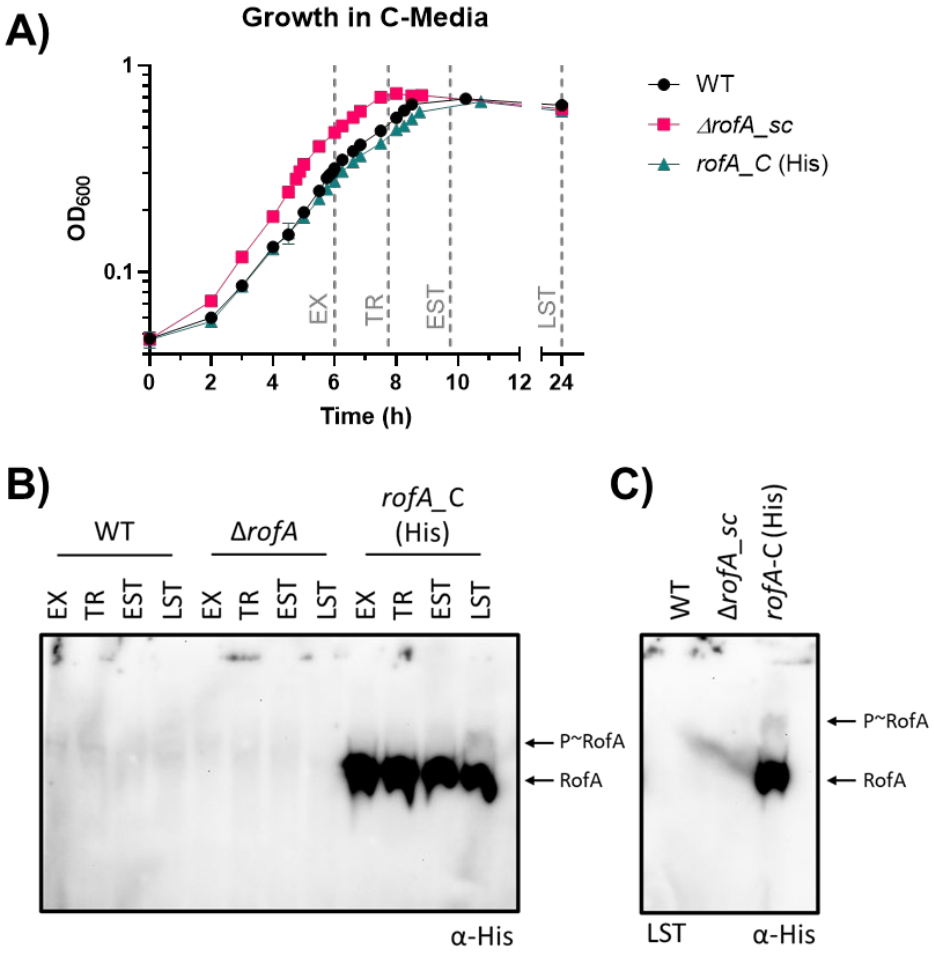


Figure 4.3: Phos-Tag of RofA-His_6 with α-His Western Blotting
(A) Representative growth curve of strains in C medium. Gray dashed lines indicate approximate collection time points. EX = exponential, TR = transition, EST = early stationary, LST = late stationary. **(B)** Anti-His_6 western blots of Phos-Tag gels for lysates of each strain at the indicates time point. **(C)** Anti-His_6 western blot of LST phase run alone.

(FLAG) strain did, it was grown in C medium to exponential, transition, early stationary and late stationary phase (**Fig. 4.3A**). Lysates were then analyzed on Phos-Tag acrylamide gels with anti-His western blotting. Consistent with the RofA-FLAG, the RofA-His was phosphorylated late in stationary phase (**Fig. 4.3B**). The signal for the anti-His western blotting was not as sensitive as anti-FLAG western blotting and resulted in more background signal over the whole blot. This resulted in some concern that the shifted species observed was a cross-reactive protein. To confirm that the band suspected to correspond with P~RofA was not a cross-reactive species, WT, Δ*rofA_sc*, and *rofA_C*(His) lysates from late stationary phase were run alone. Though there was a smear of signal on the blot, it did not seem like there was a cross-reactive species in the WT or Δ*rofA_sc* lanes (**Fig. 4.3C**).

Satisfied that the *rofA_C* (His) strain was behaving as the *rofA_C* (FLAG) strain did, it was grown in C medium supplemented with 1% of each of the carbohydrates presented in figure

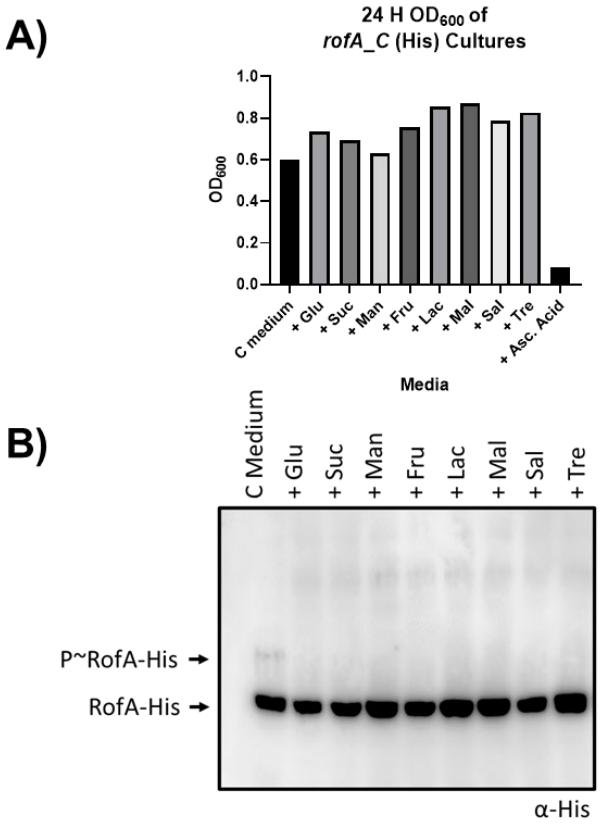


Figure 4.4: *rofA_C* (His) Cultures Grown in C Medium + PTS Sugars
(A) OD_{600} measurements for *rofA_C* (His) at 24 h when grown in C medium + 1% of the sugars indicated. Glucose (Glu), Sucrose (Suc), Mannose (Man), Fructose (Fru), Lactose (Lac), Maltose (Mal), Salicin (Sal), Trehalose (Tre), or ascorbic acid (Asc). n=1, growth pattern was observed in multiple experiments **(B)** Anti-His_6 western blot of Phos-Tag gel of lysates from cultures in (**A**).

3.10B and lysates were analyzed on a Phos-Tag acrylamide gel. Consistent with the findings from RofA-FLAG, the strain did not grow well in C medium + 1% ascorbic acid (**Fig. 4.4A**). No strong phosphorylation event was immediately obvious in the western blot of the Phos-Tag gel (**Fig. 4.4B**). There is a faint band observed in all the supplemented carbohydrate wells, but not evident in the C medium alone well. It is suspected that this is a cross-reactive band, however, because it is the same for all the supplemented carbohydrates. However, follow up with WT cultures grown in C medium + 1% carbohydrate should be completed to provide evidence of that.

In the Phos-Tag gels from the lysates of *S. pyogenes* grown in C medium + 1% fructose, it appeared as though there may have been a phosphorylation in exponential phase of growth (**Fig. 3.10C**). This combined with the finding that mutation of *rofA* impacts the expression of genes mostly in exponential phase (**Fig 3.2**) prompted the investigation of phosphorylation of RofA in exponential growth phase using the RofA-His_6 construct. The *rofA_C* (His) strain grew approximately the same in all conditions tested (**Fig. 4.5A**), and the OD_{600} corresponding to exponential growth phase was identified as ~0.25. In this screen, WT and Δ*rofA_sc* exponential phase lysates from C medium alone were used as a baseline negative control. A lysate of *rofA_C* (His) from late stationary phase was used as a positive control. Unfortunately, Phos-Tag analysis of these lysates had too much background from other proteins to draw any conclusions (**Fig. 4.5B**). It should be noted that this level of background was not observed in the previous blot from growth in C medium alone (**Fig. 4.3B**), therefore it is possible that the blotting conditions could be adjusted to minimize this background. A new tube of

anti-His antibody had been ordered between the two blots, which may explain the differences observed.

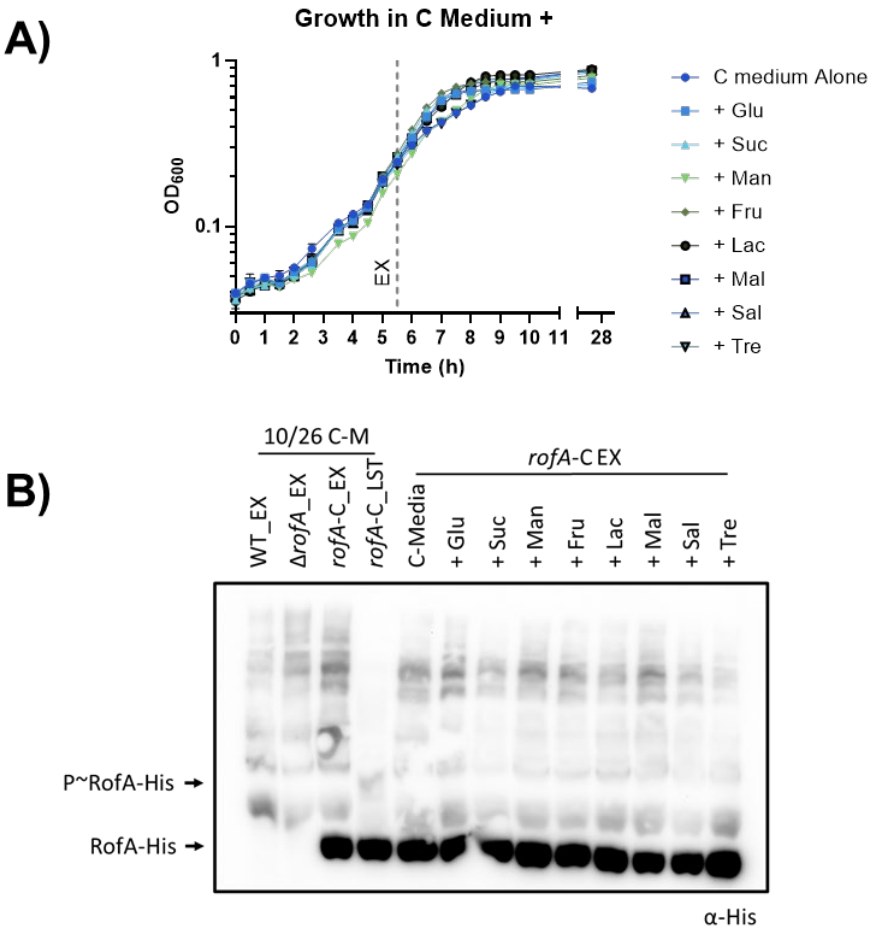


Figure 4.5: Phos-Tag Analysis of *rofA_C* (His) Cultures at Exponential Growth Phase in C Medium + PTS Sugars
(A) Growth curves of *rofA_C* (His) grown in C medium supplemented with 1% of the sugars indicated. Glucose (Glu), Sucrose (Suc), Mannose (Man), Fructose (Fru), Lactose (Lac), Maltose (Mal), Salicin (Sal), Trehalose (Tre). Two replicates were included for each growth condition. Exponential phase collection point is indicated by a dashed gray line. **(B)** Anti-His_6 western blot of Phos-Tag gel of lysates from cultures grown to exponential growth phase

Native RofA Impeded *in vivo* Phosphorylation Analysis in a Δ*ptsI* Mutant

Previous work demonstrated that RofA is phosphorylated *in vitro* by EI and HPr (**Fig. 3.7B**), and that RofA is phosphorylated *in vivo* in a glucose dependent manner (**Figs 3.10A**). To determine if the observed *in vivo* phosphorylation of RofA is

dependent upon the PTS system, the phosphorylation of RofA-FLAG was investigated in a Δ*ptsI* mutant, where the gene for EI is deleted (69). Since the Δ*ptsI* mutant of 5448 was generated by exchange the spectinomycin resistance gene *aad9*, the previously generated pProfA1-FLAGt plasmid could not be used. Therefore, the P_{rofA}-*rofA1*-FLAGt cassette from pProfA1-FLAGt was cloned into the replicating *S. pyogenes* vector pKSM201 (56) with a kanamycin resistance cassette, generating plasmid pProfA1-FLAGt_Km. This plasmid was then transformed into the Δ*ptsI* strain. Because Δ*ptsI* has native RofA, WT 5448 was also transformed with plasmid pProfA1-FLAGt_Km as a control to assess if in the presence of native RofA the phosphorylation of RofA-FLAG could still be observed. pProfA1-FLAGt_Km was transformed into Δ*rofA_sc* to ensure that the RofA-FLAG generated by this new construct behaved as the previous RofA-FLAG did. All strains were also transformed with the pKSM201 empty vector as negative controls. There was evidence of phosphorylation of RofA-FLAG in the Δ*rofA_sc* + pProfA1-FLAGt_Km lysate. Unfortunately, no phosphorylated RofA-FLAG was observed in the WT 5448 + pProfA1-FLAGt_Km lysates (**Fig. 4.6**), therefore no conclusions about RofA phosphorylation absent EI could be drawn.

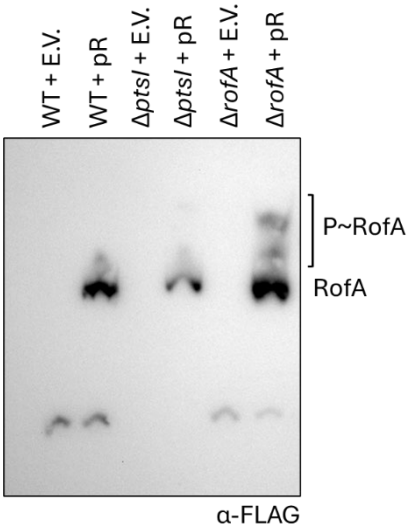


Figure 4.6: Native RofA Prevents Observation of Phosphorylated RofA-FLAG by Phos-Tag
Anti-His_6 western blot of a Phos-Tag gel containing lysates from cultures grown to late stationary phase in C medium. E.V. stands for the empty vector pKSM201. pR stands for pProfA1-FLAGt

RofA AlphaFold3 Predictions Are Similar to Mga

Since publication, AlphaFold3 was released for public use and a user-friendly web-based server was established (188). Previously, the secondary structure of RofA had only been predicted with Phyre2 (184). Therefore, the tertiary structure of RofA from 5448 was predicted with AlphaFold3, then molecular graphics were generated with ChimeraX (189-191). Most of the predicted structure had a "a per-atom confidence estimate" (pLDDT) that fell within the 'Very high' or 'Confident' range. The model also had a "predicted template modeling" (pTM) score of 0.84, well above the 0.5 threshold for a likely accurate model. Cursory inspection of the RofA monomer and the Mga monomer (pTM = 0.82) indicates that the two structures are similar (**Fig. 4.7A & B**). though a more thorough alignment through programs such as DALI (192) should be performed.

The new version of AlphaFold also includes the ability to predict multi-protein structure. Since many PCVRs make dimers and higher order multimers (86) (**Table 1.1**), RofA is expected to also form these structures. Therefore, the RofA homodimer structure was predicted. Again, the pLDDT scores for most of the predicted structure fell within the 'Very high' or 'Confident' range. For interacting domains, AlphaFold3 also provides an "interface predicted template modeling" (ipTM) score. iPTM scores greater than 0.8 indicate high-quality predictions. Scores between 0.6-0.8 indicate that the interactions may or may not be accurate. Then scores less than 0.6 are likely a failed interaction. The ipTM score for the RofA dimer was 0.79; very close to the 0.8 that indicates a high-quality prediction. The protein is predicted to interact mainly with the $EIIB^{Gat}$ domain (**Fig. 4.7C**), which is consistent with the experimental findings for Mga

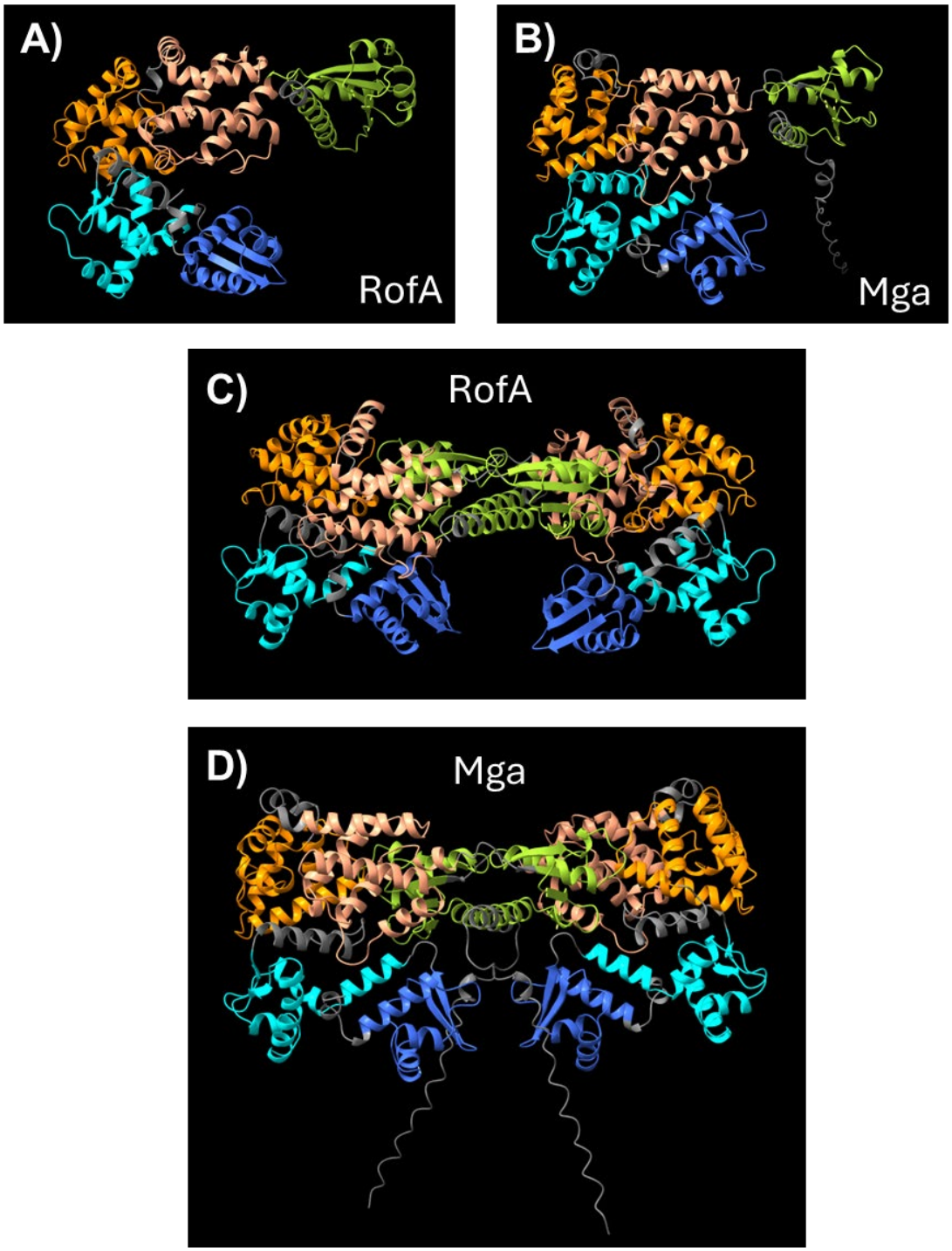


Figure 4.7: AlphaFold3 Structures of RofA Are Similar to Mga
Ribbon diagrams of the AlphaFold 3 predicted structures for **(A)** RofA monomer, **(B)** Mga monomer, **(C)** RofA homodimer, and **(D)** Mga homodimer. The helix-turn-helix domains are shown in blue and cyan, the PRDs are shown in orange and light-orange, and the $EIIB^{Gat}$ domains is shown in green.

and AtxA (97, 102). Additionally, the predicted RofA homodimer structure resembles the predicted Mga homodimer structure (iPTM = 0.63), which also interacts at the EIIB domain (**Fig. 4.7D**). Mga has two C-terminal tails after the EIIB domain that have very low confidence based on the pLDDT values. Considering these are close to the EIIB domains, this may be why the iPTM values are so low.

Discussion

In these additional experiments, it was determined that the higher hyaluronic capsule levels of the Δ*rofA_sc* mutant is the primary contributor to the observed aggregation phenotype (**Fig. 4.1**) and that the phosphorylated RofA-FLAG species observed in lysates of strains grown in C medium appears 14 hours post inoculation (**Fig. 4.2**). Additionally, AlphaFold3 was used to predict the structure of RofA, revealing that its predicted structure is similar to the predicted structure of Mga (**Fig. 4.7**). Finally, predicted structure of the RofA homodimer suggests that two RofA monomers likely interact at the $EIIB^{Gat}$ domain (**Fig. 4.7C**). Unfortunately, there were no carbohydrates identified that induced strong phosphorylation when supplemented in C medium using the 6x-His tag system at late stationary phase (**Fig. 4.4**) and technical limitations inhibited the study of RofA phosphorylation in exponential growth phase (**Fig. 4.5**). Additionally, we were not able to assess the phosphorylation of RofA in a Δ*ptsI* mutant, due to the presence of native RofA (**Fig. 4.6**).

In Chapter 3, the Δ*rofA_sc* mutant was shown to have reduced expression of pili and increased expression of capsule that correlated with reduced aggregation and

reduced adherence to keratinocytes (**Fig. 3.6**). In these experiments, adding hyaluronidase to cultures recovered the aggregation of the Δ*rofA_sc* mutant (**Fig. 4.1**) and suggests that increased hyaluronic acid capsule is the primary contributor to the aggregation defect observed. Most capsule mutants also exhibit reduced adherence to epithelial cells (175-178), though pili mutants also exhibit reduced adherence correlated with reduced aggregation (173). Therefore, it would be interesting to assess if the hyaluronic acid capsule is also the main contributor to the reduced adherence observed for the Δ*rofA_sc* mutant (**Fig. 3.6C**). Hyaluronidase treated cultures could be used for adherence assays to assess whether this is the case.

Use of the standard epitope FLAG and 6x-His tags for analysis of *S. pyogenes* lysates from cultures grown in varying carbohydrates have both had limited utility due to cross-reactive proteins and high background (**Figs. 3.10**, **4.3 - 4.5**). Therefore, it may be necessary to use commercial sources to generate an anti-RofA antibody for western blots and/or immunoprecipitation. This may be of interest considering that there are a variety of potential applications of a native RofA antibody (see Chap. 6: Conclusions and Future Directions).

The phosphorylation of RofA absent a functional PTS system was also investigated. Unfortunately, the presence of native RofA inhibited the observation of phosphorylated RofA-FLAG (**Fig. 4.6**). Therefore, to study the phosphorylation of RofA *in vivo* in the absence of the PTS, a Δ*ptsI*/Δ*rofA_sc* double mutant will need to be generated. Since single mutants of each have already been generated, and the *rofA* mutation is a stop-codon mutation instead of replacement of the *rofA* gene with an

antibiotic resistance cassette, the previous constructs should be able to be used to generate the double mutant.

Finally, AlphaFold3 was used to predict the tertiary structures of Mga and RofA (**Fig. 4.7**). This will enable better predictions of which residues may be important for phosphorylation. By eye, the residues implicated in phosphorylation and activity of Mga can be identified in the tertiary structures, and these compared to the homologous regions on RofA to identify if there are target histidines. The use of structure comparison software, such as DALI, will further enable precise comparison of these two regulators for better predictions of RofA residues that may be targets for phosphorylation.

Chapter 5: Generation of Bioluminescent *S. pyogenes* for RofA Studies *in vivo*

Introduction

RofA is a transcriptional regulator of multiple virulence factor genes including the genes for protein F (*prtF*), M-protein (*emm*), streptolysin S (*sagA*), streptococcal pyrogenic exotoxin A (*speA*), streptokinase (*ska*), capsule (*hasB*), and collagen binding protein (*cpa*) (112-114). The *rofA* gene was also identified as being important to M1T1 fitness in the primate oropharynx (116) and as being highly expressed during soft-tissue infection (72). Therefore, the evaluation of *rofA* on *in vivo* virulence is of interest.

Mice are one of the most common small animal models used for infectious disease research due to their cost-effectiveness, fast breeding time, compact and easy housing, and genetic tractability (193). For *S. pyogenes*, there are several murine models of disease, including models for subcutaneous ulcers, impetigo, systemic disease, septic arthritis, oro- and nasopharyngeal colonization, and vaginal colonization (194). For many of these models, enumeration of bacterial burden is a major infection indicator and often requires sacrificing the animal to excise tissue and enumerate the colony forming units (CFU) within the given tissue (194). If multiple time points over the course of the infection need to be investigated, this requires large numbers of animals. To limit the number of animals required for such experiments, groups have developed bioluminescent and fluorescent pathogens to non-invasively monitor infections (195-197). Using these strains allows investigators to follow a single group of animals over the course of the infection instead of needing several groups of animals for sacrifice at each time point.

Bioluminescence is preferred to fluorescence for *in vivo* imaging in animals because it has a lower signal to noise ratio (195). As such, bioluminescent strains of *S. pyogenes* have been generated for study *in vivo*. The first was generated in 2003 using a plasmid-based transposon system (148). A modified version of the luminescence operon from *Photorhabdus luminescence*, *luxABCDE*, with a Gram-positive ribosomal binding site (RBS) (198) was inserted into the Tn4001 transposon generating pXen5 (formerly pAUL-A Tn4001 *luxABCDE* Km^R) (199). This plasmid was transformed into strain M49 strain 591 and a highly bioluminescent transformant was selected (148). This transformant was found to contain the transposon in the second open reading frame of a gene for a hypothetical protein NP_268809 (148), which is annotated as a sugar phosphate isomerase in NCBI. This strain could be seen on a Xenogen In Vivo Imaging System (IVIS) over the course of 72 hours, and the luminescence observed mirrored the bacterial CFU (148). This strain subsequently became Xen20 and was available commercially first through Xenogen, which was then acquired by Caliper LifeSciences.

Strain Xen20 has been used in subsequent studies investigating *S. pyogenes* host-pathogen interactions and evaluating potential vaccine candidates (200-203). Comparison of infections with bioluminescent strains to non-bioluminescent strains demonstrated that luminescent strains followed similar infection patterns (148, 200). In nasopharyngeal colonization models, Xen20 luminescence was readily observed *in vivo* (148, 200). In a subcutaneous infection model, however, luminescent bacteria were observed at the site of infection, but the bacteria that disseminated to the organs were not emitting a luminescent signal (203). In contrast to these findings, in an

intraperitoneal infection model, luminescence could be detected from infected organs (202). Finally, it was found that the promoter driving the expression of Xen20 luminescence is glucose sensitive (203), which likely impacts the luminescence in different host sites during infection. This emphasizes the risk of using a transposon-based integration system that utilizes native promoters.

The gene identified as the target of the transposon insertion in the Xen20 strain has also been used as the insertion target in generating bioluminescent strains of *emm75* and *emm89 S. pyogenes* (204, 205). These strains differ from Xen20 in that expression of the *luxABCDE* cassette is driven by the P_{help} promoter. Integration of the P_{help}-*luxABCDE* cassette into these strains resulted in attenuated fitness of the *emm75* strain in a nasopharyngeal colonization model (204) and attenuated fitness of the *emm89* strain in a subcutaneous soft-tissue infection model (205). This attenuation was not seen when an empty vector was integrated into the genomic locus (205), suggesting that the expression of the *lux* operon confers a fitness defect on the luminescent strains. Also, luminescence of M89::Lux could not be observed in live animals in a lower respiratory tract infection model, though luminescence was observed in dissected lungs (205), suggesting that the luminescence signal intensity is not sufficient to monitor the lungs of live animals.

Plasmid-based luminescent strains of *S. pyogenes* have also been generated (204-207). In two studies, the use of standard replicative plasmids carrying the *luxABCDE* cassette were able to create bioluminescent strains, however these plasmids were not stable in the population absent antibiotic selective pressure (204, 205). Interestingly, the M1 strain carrying pLux failed to luminesce at all (205). To

circumvent the instability of standard plasmids, another group constructed plasmids which included a toxin-antitoxin system to ensure the luminescent plasmid is maintained during replication (206, 207). In these studies, Loh & Proft found that firefly luciferase operon, FFluc, had a greater bioluminescent signal than the *luxABCDE* operon (206), and the engineered NanoLuc (Nluc) luciferase had an even greater signal (207). However, the FFluc and Nluc luminescent operons both require exogenous substrates to function, whereas the *luxABCDE* operon can generate light without exogenous substrate. This is an important consideration, as the substrate for the Nluc operon is expensive and had not been evaluated for its toxicity in animals at the time of publication (207). Additionally, in M1 strain SF370, the FFluc operon appeared to cause toxicity leading to lower bacterial CFU counts *in vitro* (207). In Loh & Proft's plasmid-based system, the expression of all luciferase operons was driven by the strong constitutive P23 promoter from *Lactococcus lactis* (151). They found that the *luxABCDE* operon in the plasmid system was unstable, and often acquired deletions in the plasmid leading to smaller size (206). Their results indicate that the generation of bioluminescent bacteria that do not need exogenous substrate to function requires the integration of the *luxABCDE* cassette into the chromosome.

Overview

This work sought to improve upon the previously created luminescent systems by generating an integrative construct containing the *luxABCDE* operon under control of the strong and constitutive P23 promoter from *L. lactis* (151). This construct was designed to specifically insert the *lux* cassette into a previously identified and highly

conserved locus that sits between two predicted rho-independent terminators (208). The conserved nature of this locus means the cassette can theoretically by inserted into any strain of *S. pyogenes* without disrupting native genes. Transformation of *S. pyogenes* strain 5448 with this construct allowed for allelic exchange of the P23-*luxABCDE* cassette into the chromosome and generated luminescent clones of varying levels, many of which generated luminescence at higher levels than the commercially available Xen20. The luminescence was found to be stable over 2 passages in THY, and strains grew to similar yields as the WT parent but had longer doubling times. When challenged with freshly isolated human polymorphonuclear lymphocytes (PMNs), one luminescent strain survived as well as wild type, suggesting that the *luxABCDE* cassette is not impairing innate immune evasion functions. Together, this data shows a promising construct for generating bioluminescent strains of *S. pyogenes* for non-invasive monitoring of infection.

Results

Plasmid pP23-*luxABCDE* for Generation of Bioluminescent *S. pyogenes*

The *luxABCDE* operon was amplified from Xen20 gDNA from the Gram-positive RBS and beginning of the *luxA* gene to the end of the *luxE* gene and ligated into plasmid pP23-Spec downstream of the strong constitutive P23 promoter from *L. lactis* (151) to generate plasmid pP23-*luxABCDE*-Spec (**Fig. 5.1A**). To confirm the P23-*luxABCDE* cassette functioned in *S. pyogenes*, pP23-luxABCDE-Spec was transformed into M1T1 5448 (140). Transformants exhibited luminescence when imaged on an iBright Imaging System using the chemiluminescence feature and

cultures of the transformants were highly luminescent as measured on a luminometer (**Fig. 5.1B & C**). These findings demonstrate that the P23-*luxABCDE* cassette is functional in *S. pyogenes*.

After confirming the function of the cassette, a construct for allelic exchange was generated in the temperature-sensitive vector pCRS (150). The allelic exchange construct contains the P23-*luxABCDE* cassette flanked by regions of homology to the *tsf-pepO* intergenic region previously identified when developing the RIVET system (208) (**Fig. 5.2 A & B**). This location was chosen for integration because the region is conserved across *S. pyogenes* genomes, therefore this construct should be able to be used for a variety of different strains. Additionally, it sits between two predicted rho-independent terminators, thus integration should not disrupt any genes. This construct was successfully used to transform the P23-*luxABCDE* cassette into the M1T1 *S. pyogenes* strain 5448 and generated 9 allelic exchange strains (8, 21, 67, 105, 127, 220, 233, 242, and 245), and 3 revertant control strains (R24, R102, R277) that reverted to the WT sequence after the second crossover event. The presence or absence of the cassette was confirmed by amplifying the genomic DNA of each strain within the intergenic region (**Fig. 5.2C**).

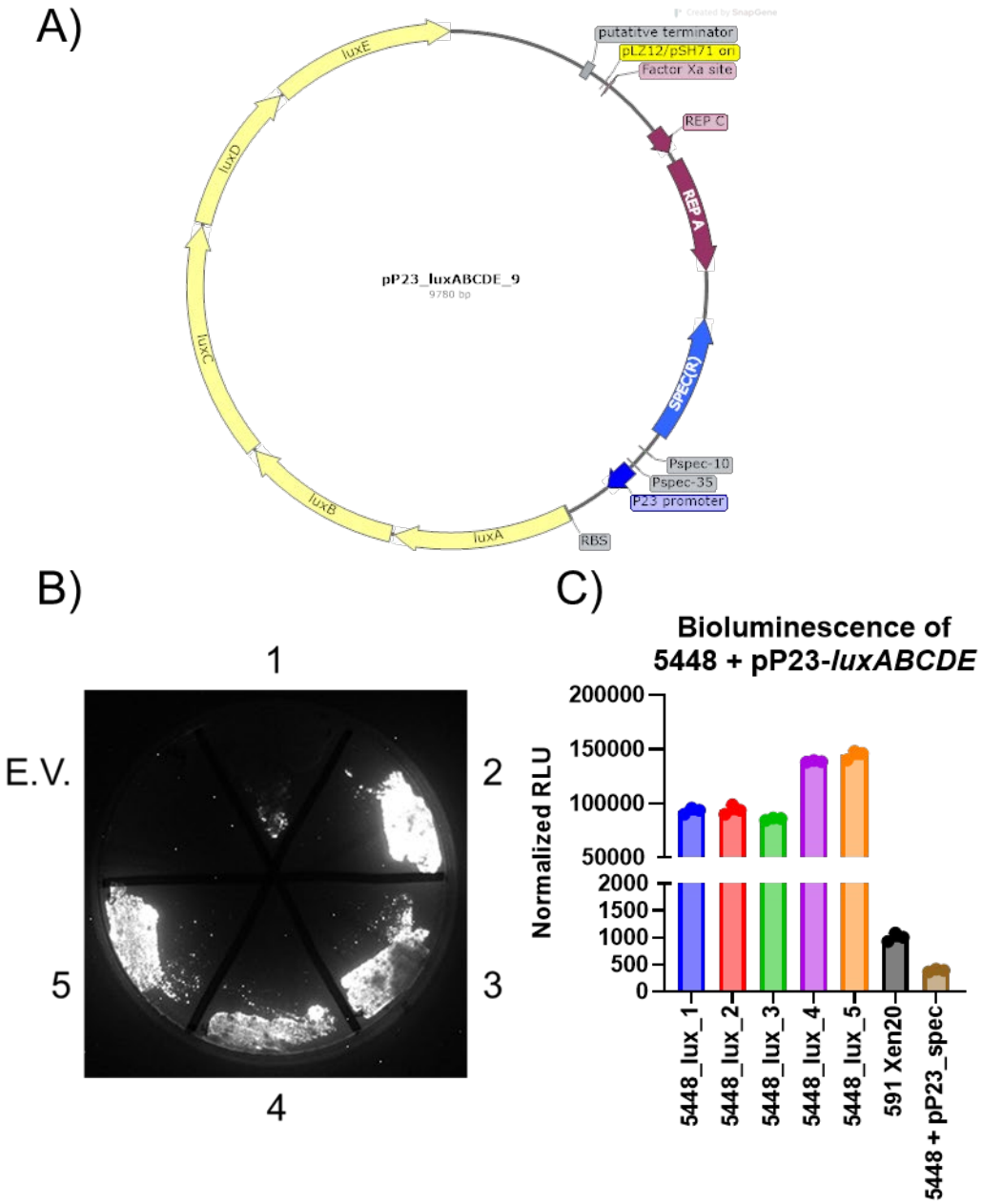


Figure 5.1: P23-luxABCDE Is Functional in *S. pyogenes*
(A) Map of plasmid pP23-luxABCDE which has the *luxABCDE* operon under control of the P23 promoter from pOri23. Vector backbone is pJRS525. **(B)** Luminescence of streaks of *S. pyogenes* + pP23-*luxABCDE* imaged on an iBright Imaging System using the chemiluminescence feature. Labels 1-5 are 5 different transformants and E.V. is *S. pyogenes* transformed with the pP23-Spec empty vector. **(C)** Bar graph of luminescence values measured on a luminometer. Luminescent measurements were normalized to cell density by dividing the measured luminescence by the OD_{600}. Xen20 was included as a positive control and 5448 + pP23_Spec as a negative control.

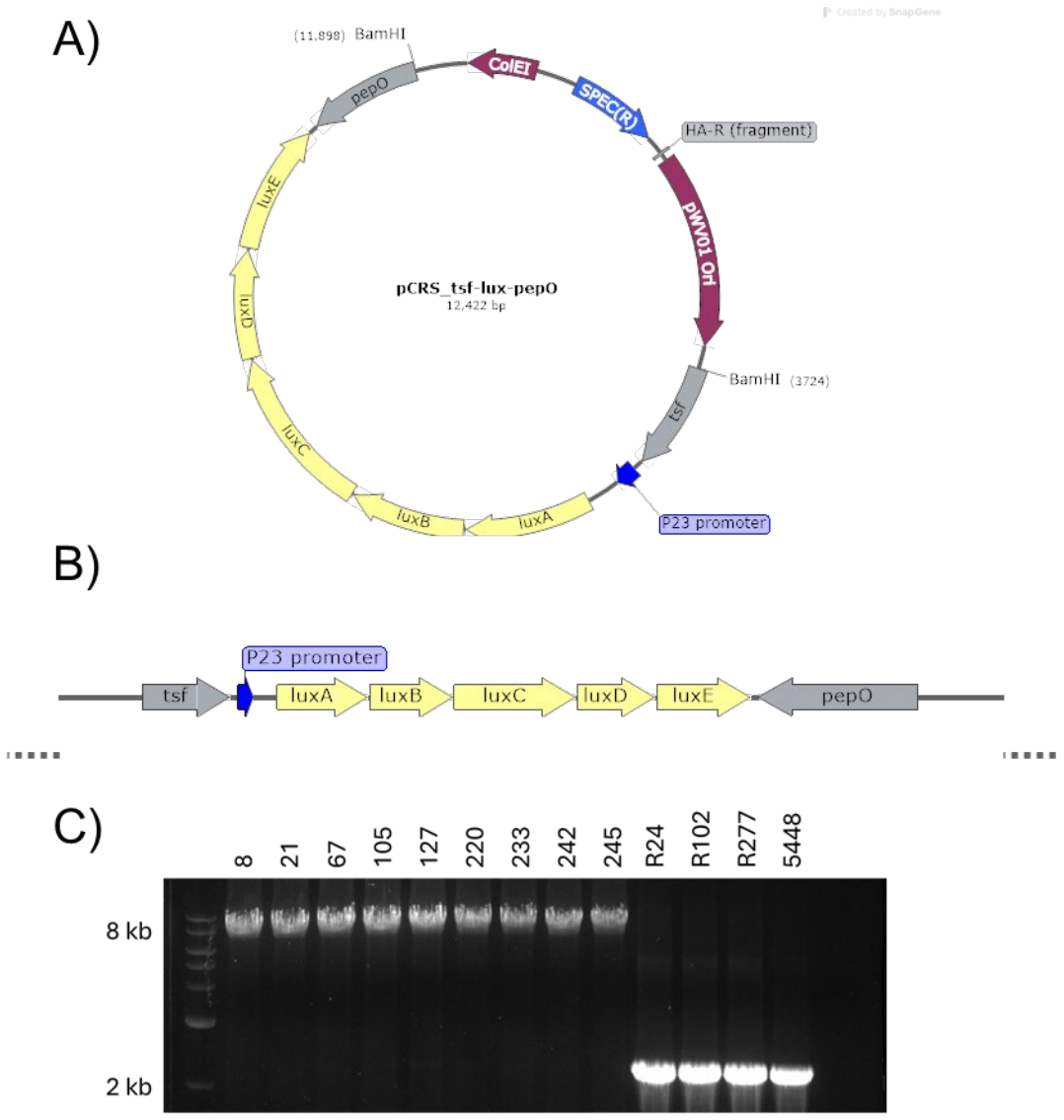


Figure 5.2: Allelic Exchange of P23-*luxABCDE* into *S. pyogenes*
(A) Map of plasmid pCRS_tsf-lux-pepO for allelic exchange of the *luxABCDE* operon into the *tsf-pepO* intergenic region. P23-*luxABCDE* cassette is flanked by regions that are homologous to the *tsf-pepO* intergenic region and are labeled based on which gene they are closer to. Vector backbone is pCRS which contains the thermosensitive pWV01 origin. **(B)** Schematic of P23-*luxABCDE* cassette inserted into the *tsf-pepO* intergenic region. **(C)** Gel of PCR products amplified from *tsf-pepO* intergenic region of candidate gDNA. WT 5448 gDNA included as a control.

Growth and Bioluminescence of 5448_lux Strains in THY

To assess the growth and bioluminescence of the strains, they were grown in Todd Hewitt medium + 0.2% yeast extract (THY) in sealed Klett tubes. Growth was

monitored by measuring absorbance at 600 nm (OD_{600}) every 30 minutes and luminescence was evaluated on a luminometer every hour. Preliminary experiments were conducted with only one replicate because of the number of strains. Growth curves for all the 5448_lux strains and the revertant strains were similar to WT 5448 (**Fig. 5.3A**). The revertant strains exhibited similar luminescent readings to the WT 5448 (**Fig. 5.3B**). The 5448_lux strains exhibited variable luminescence with a group of strains exhibiting high luminescence (8 – 127) and another group exhibiting lower luminescence (220-245). Since the original growth curve was crowded with several strains, the high-luminescent strains and low-luminescent strains were grouped, and their growth curves were compared to WT to assess if high or low output impacted growth. There were no observed differences from the WT growth curve (**Fig. 5.3C**). Because isolates 105 and 127 exhibited the highest bioluminescence, these strains were chosen for further characterization.

Growth Characteristics of lux105 and lux127

The growth and bioluminescence of strains 5448_lux105 and 5448_lux127 were evaluated in triplicate and compared to the previously commercially available Xen20 strain. Overall, the growth curves of both luminescent strains and their corresponding revertant were somewhat similar to the growth of 5448 (**Fig. 5.4A**) though they appeared to have a shorter lag, and more shallow slope in exponential phase. The luminescence of both 5448_lux105 and lux127 strains were considerably higher than the Xen20 strain (**Fig. 5.4B**).

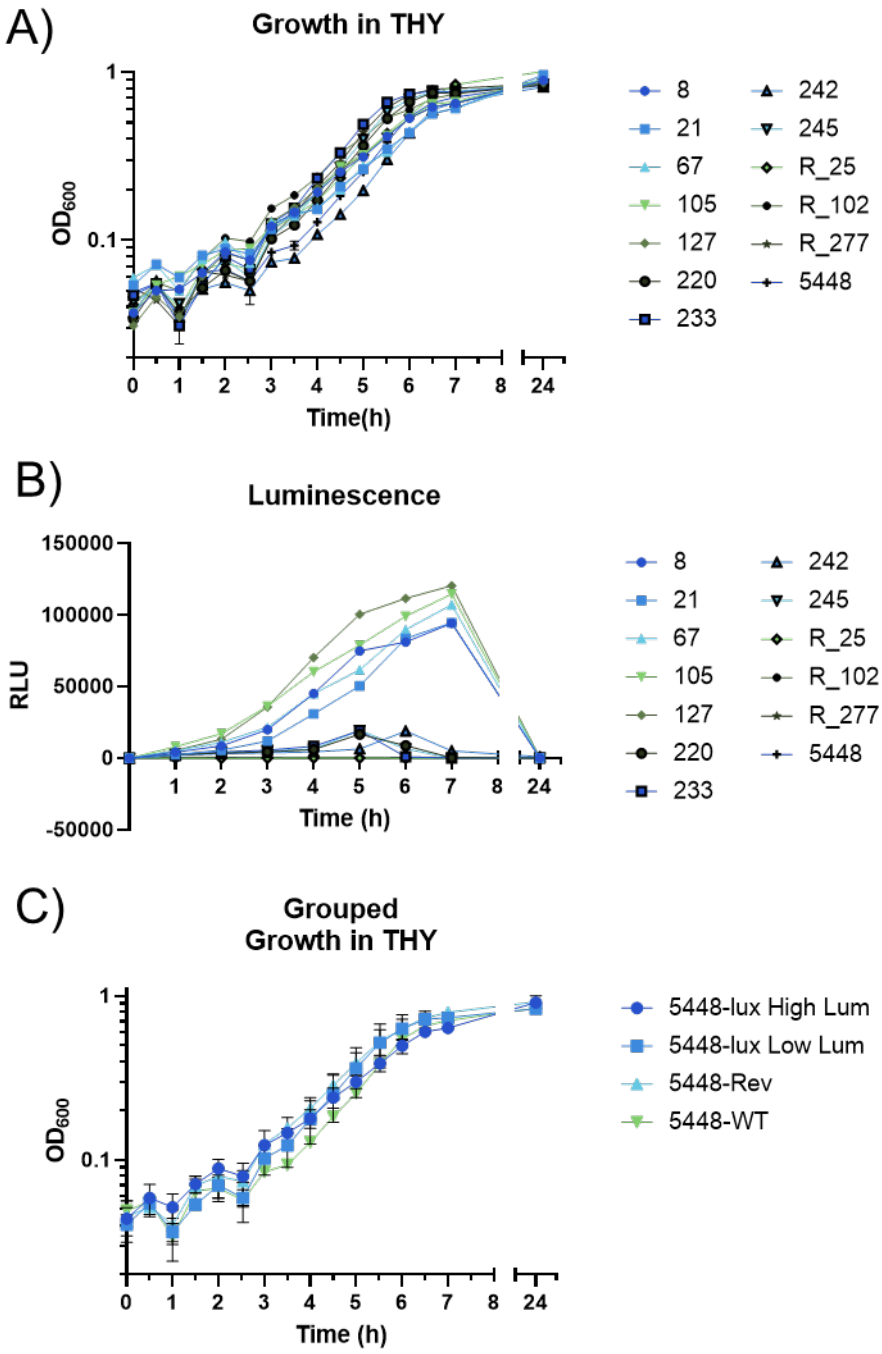


Figure 5.3: Bioluminescence of 5448_lux Strains in THY
(A) Growth curve of strains grown in THY. Each curve is only a single replicate. **(B)** Luminescence of strains across growth in THY. Luminescence was measured on a luminometer with 5 second acquisition time. **(C)** The high luminescent strains (high lum) and low luminescent strains (low lum) were grouped together and plotted over time.

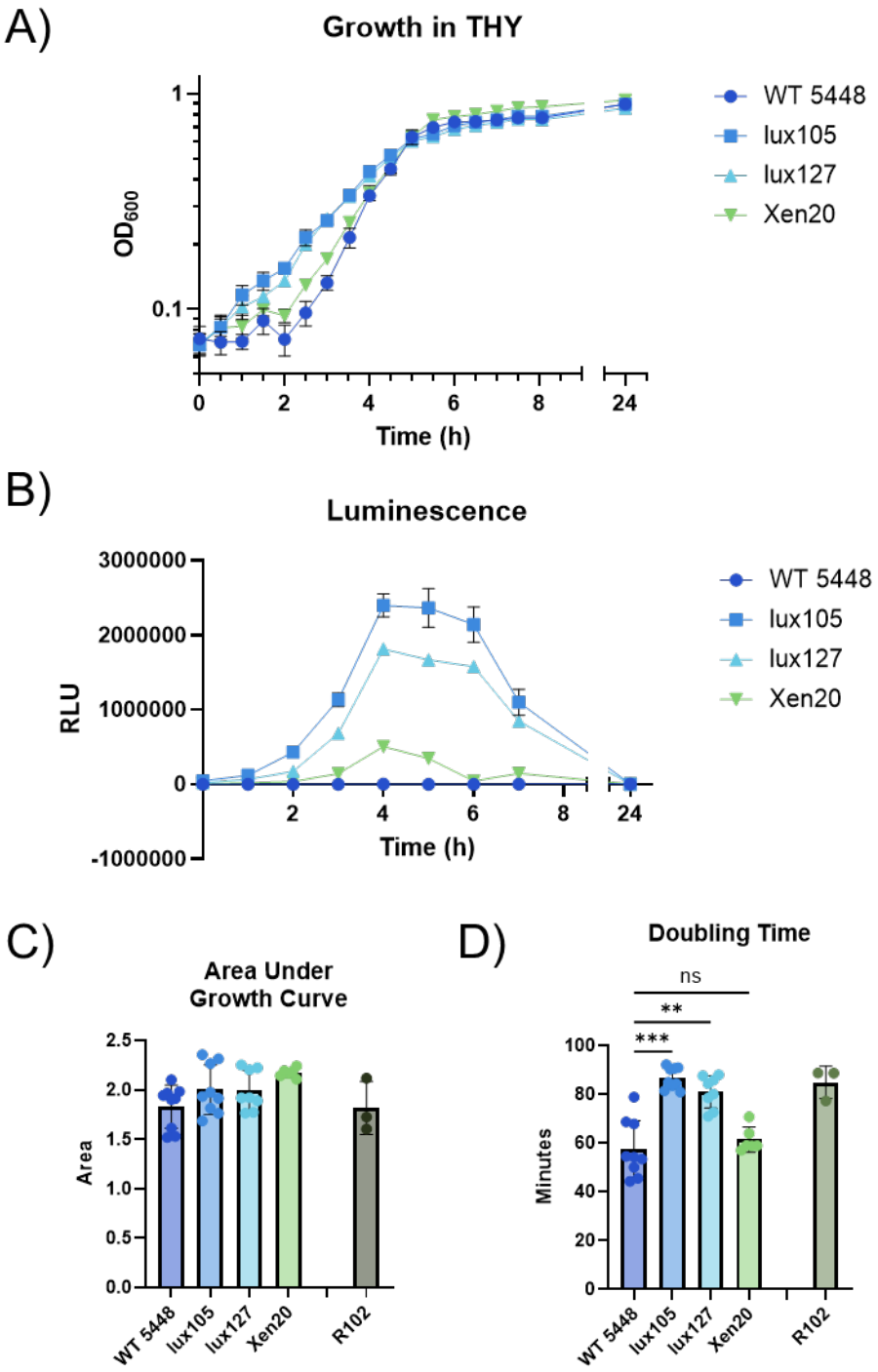


Figure 5.4: Growth Characteristics of 5448_lux105 and lux127 in THY
(A) Representative growth curve of strains grown in THY **(B)** Luminescence of strains across growth in THY. Luminescence was measured on a luminometer with 5 second acquisition. **(C)** AUC measurements for growth from 0 to 6.5 hours. **(D)** Doubling times of each strain during exponential growth phase. Statistics presented are from a Brown-Forsythe and Welch ANOVA with Dunnett's T3 test for multiple comparisons.

To compare the full growth curves, the area under the curve (AUC) was calculated as this metric integrates the contributions of initial population size, growth rate, and carrying capacity (167). The AUC was similar for all strains (**Fig. 5.4C**), indicating that the overall growth of each strain was similar. However, calculation of the doubling time during exponential growth demonstrated that the luminescent strains had longer doubling times than WT 5448 and Xen20 (**Fig. 5.4D**). Growth curve analysis of the corresponding revertant strain demonstrates that the revertant also has a longer doubling time, and that this doubling time is not significantly different from lux105 and lux127 ($p = 0.9998$ and 0.9717 respectively). This may suggest that the strain acquired some kind of off-target mutation during the allelic exchange process that impacts doubling time and that the increase isn't necessarily a result of the luminescent cassette insertion.

An important aspect of creating a bioluminescent strain for *in vivo* monitoring of infection is the ability of the bioluminescent signal strength to indicate the bacterial burden within the tissue. Therefore, it is essential that luminescent intensity correlates with signal strength. As a first step, we plotted the bioluminescent signal as a function of OD_{600} (**Fig 5.5A**) and evaluated the correlation between the two. Like the commercially available Xen20 strain, the luminescence correlates well with the OD_{600} for the first several hours of growth (**Fig 5.5B**) with R^2 values of 0.98, 0.96, and 0.98 for lux105, lux127, and Xen20, respectively.

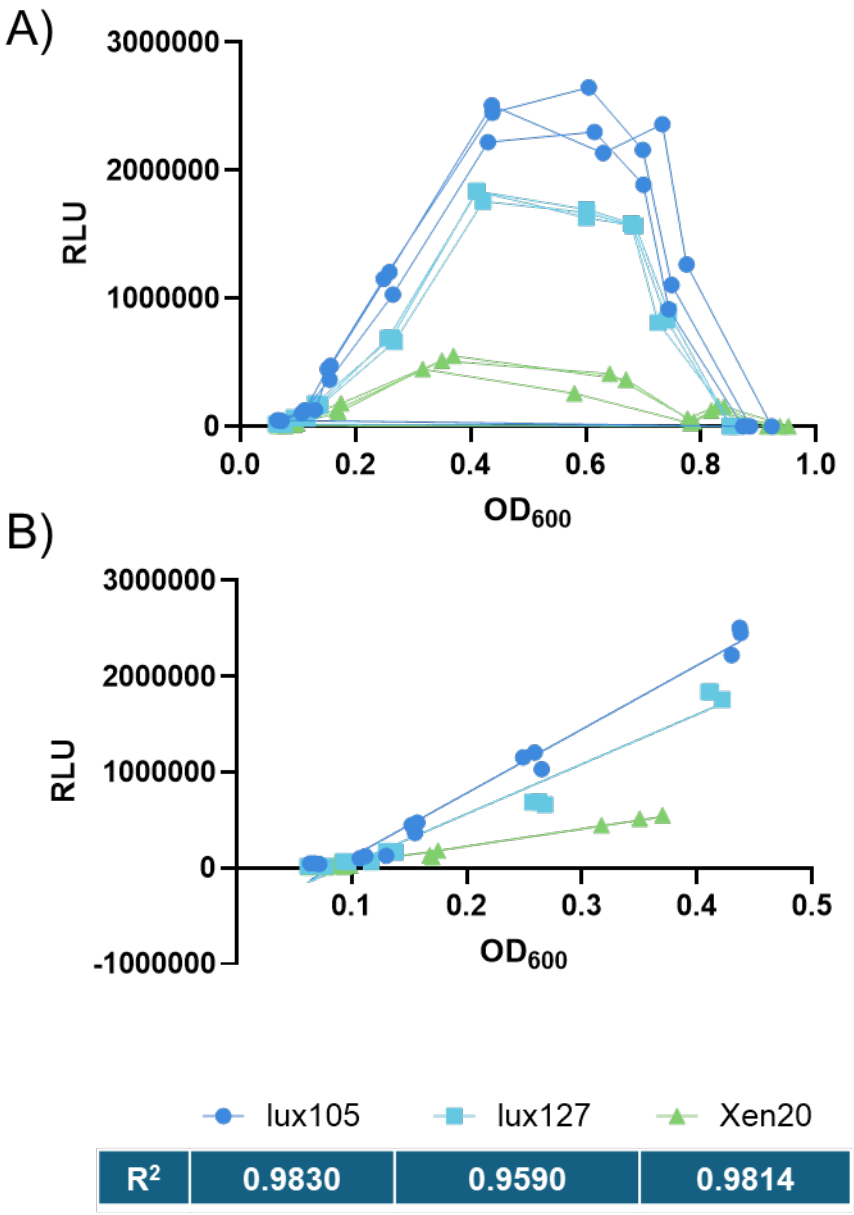


Figure 5.5: Correlation of Luminescence Versus OD600
(A) Luminescence of each replicate plotted versus the OD_{600}. Each strain is represented by a different color and symbol. **(B)** Luminescence versus OD_{600} for the first 4 hours of growth with the line of best fit included. R^2 values are presented below the legend for each strain.

Stability of Bioluminescence

The bioluminescence of the lux105 and lux127 isolates is very strong through much of the growth curve, however, the bioluminescence begins to dissipate around 4 hours and reaches background levels at 24 hours (**Fig 5.4B**). To ensure that this

bioluminescence drop was not the result of mutations to the P23-*luxABCDE* cassette or some other mutation that permanently disrupts the bioluminescence, the 24-hour cultures were re-inoculated into fresh THY, and growth and luminescence were measured over another 24-hour period. In the second passage, the bioluminescence exhibits a similar pattern to passage 1 (**Fig 5.6**), which demonstrates that the bioluminescence cassette is stable over 48 hours.

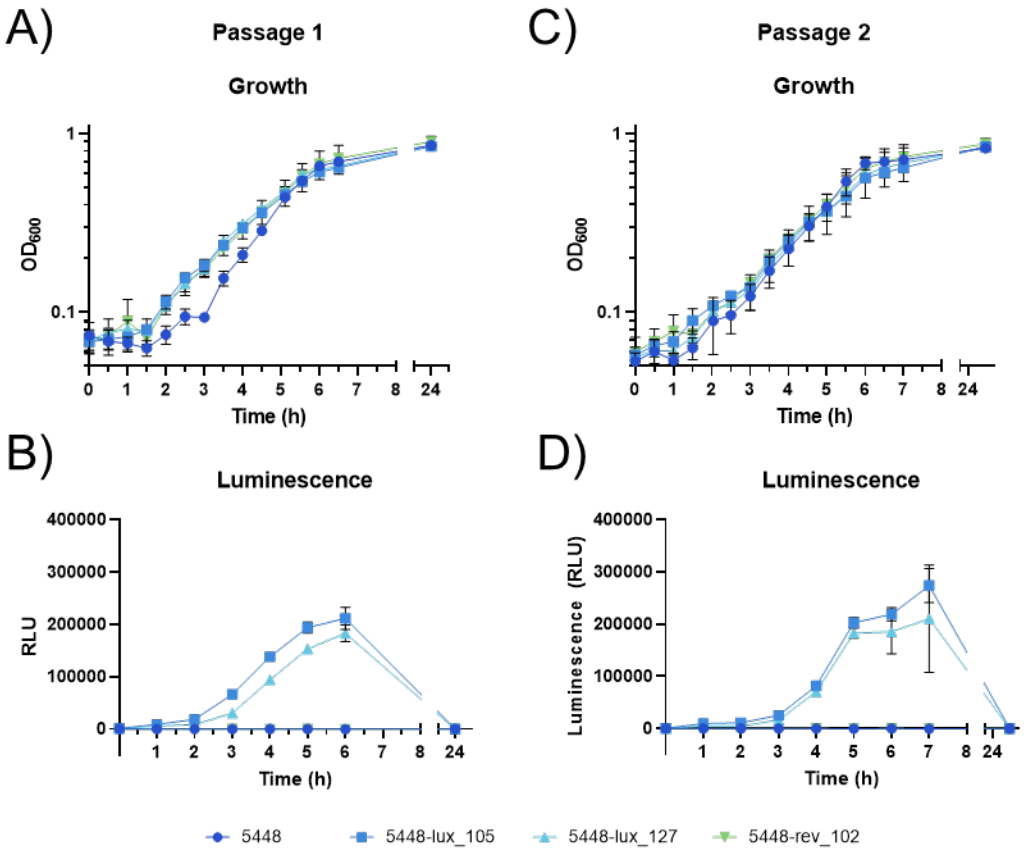


Figure 5.6: Luminescence Is Stable over 48 Hours
(A) Growth and **(B)** luminescence were monitored over 24 hours. Then, cultures from the 24-hour time point in panel A were re-inoculated into fresh THY and **(C)** growth and **(D)** luminescence were monitored over another 24-hour period.

Fitness of 5448_lux Strains in PMN Opsonophagocytic Assay

As a first step toward assessing if this cassette will be useful for studying *in vivo* virulence, the fitness of the bioluminescent strains in an opsonophagocytic assay

with human polymorphonuclear lymphocytes (PMNs) was performed. To combine data gathered from multiple donors on multiple days, each experiment was normalized to the % survival of wild type (see Methods). 5448_lux127 and the revertant R102 both exhibited similar survival to WT 5448 (**Fig. 5.7**). However, lux105 had significantly reduced survival compared to WT. Since the revertant had similar survival to lux127 and WT, this suggests that the luminescent cassette specifically does not impact survival, and that lux105 may have acquired some kind of mutation that has made it less fit.

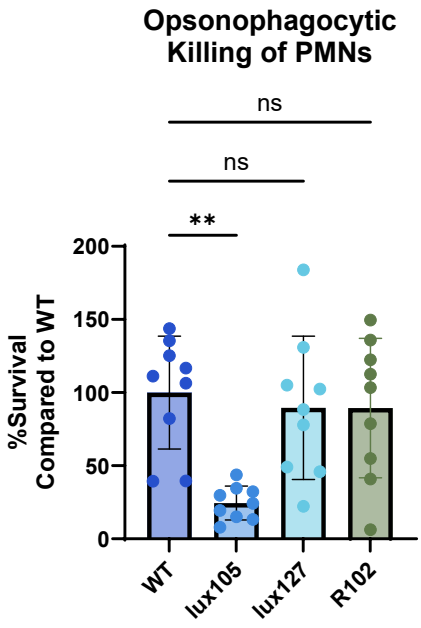


Figure 5.7: 5448 lux127 Is as Fit as WT in Opsonophagocytic PMN Assay
Human PMNs (1 x 10^6 each) were co-incubated with 1 x 10^5 CFU of opsonized *S. pyogenes* WT 5448 (WT), lux105, lux107 and revertant R102. Cells were co-incubated for 2 h at 37°C, mammalian cells were lysed, and bacterial CFU enumerated. Data is presented as the % survival normalized to WT survival (see methods) and represents 3 assays from 3 different donors.

Discussion

In this study, a construct was developed that is theoretically capable of inserting the *luxABCDE* luminescence operon under control of the strong constitutive P23 promoter into any strain, including previously generated mutants, of *S. pyogenes*. This construct was used to generate bioluminescent strains of M1T1 strain 5448. Integration of the cassette yielded strains that grew similarly to WT 5448 (**Fig. 5.3A and C**) and were highly bioluminescent compared to strain Xen20 (**Figs 5.3B and 5.4B**). While the strains have similar growth yields (**Fig. 5.4C**), the luminescent strains and the revertant strain have longer doubling times (**Fig. 5.4D**). The luminescence intensity of the constructed strains correlates well with OD_{600} for the first 4 hours of growth in THY, similar to the Xen20 strain (**Fig. 5.5**) and the luminescence is stable when passaged and grown for 48 h (**Fig. 5.6**). Finally, strain lux127 survived as well as WT 5448 when challenged with human neutrophils, suggesting that the luminescent cassette does not impact fitness in this assay. Together, these findings show that the construct generates strains that produce high amounts of stable luminescence, and that this luminescence does not significantly impact fitness. This construct was also designed to insert into a locus that is uniformly distributed across *S. pyogenes* genomes (208), though integration into other strains and serotypes still needs to be demonstrated.

The bioluminescence of these strains dissipates during stationary phase of growth (**Figs 5.3B and 5.4B**), however this is not without precedent. When this operon was inserted into *S. pneumoniae*, it also had reduced bioluminescence in stationary phase, and this was not due to autolysis, as bacterial CFU counts were still high (199, 209). This was also reported for the *luxABCDE emm3*, *emm89* and *emm75* strains (204,

205) and the plasmid-based FFluc strain (207). Even the luminescence of the ATP-independent Nluc strain exhibited a moderate drop in luminescence in stationary phase, though it maintained luminescence above background levels at 24 hours (207). A similar decrease in luminescence was also observed for the *S. pyogenes* Xen20 strain (**Fig. 5.4B**, **Appendix A5**). Despite the drop in luminescence in stationary phase for most of these strains, many were also demonstrated to be bioluminescent and detectable *in vivo* (148, 199, 200, 204-206). Therefore, it is reasonable to hypothesize that the strains constructed here will also perform adequately *in vivo*.

In this work, it was demonstrated that the luminescence intensity of the strains correlated with OD_{600} in the first 4 hours of growth (**Fig. 5.5**). While this is in line with the luminescence of Xen20, it seems concerning that the luminescence intensity does not correlate across all of growth. The goal of using this system *in vivo* would be to understand the bacterial burden within the tissues from the bioluminescent intensity, regardless of bacterial growth phase. It is worth noting, however, that during the construction of previous bioluminescent *S. pyogenes* strains, the correlation evaluation was determined either for *in vivo* infections (148, 202), or for serial dilutions of exponential phase culture (206, 207). Additionally, as mentioned previously, the luminescence intensity of almost all strains constructed thus far seems to plateau and drop as the bacteria enter the stationary phase (204, 205, 207). Despite this, in most cases, the luminescence intensity was an adequate indicator of the bacterial burden *in vivo*. It is therefore possible that the lack of correlation when comparing the OD_{600} to luminescence is not indicative of the correlation between the two *in vivo*. It will be

important to validate the correlation between CFU and bioluminescence by both methods moving forward.

When *luxABCDE* was introduced into *S. pyogenes* with the toxin-antitoxin plasmid system, the strains lost luminescence within 1-4 days (206). This was due to the plasmid becoming reduced in size. The plasmids presented in this work were also found to be unstable in the *E. coli* strain used for cloning and large deletions were observed within the plasmid for colonies lacking bioluminescence (data not shown). While stable luminescence was demonstrated for the allelic exchange strains over the course of 2 days (**Fig. 5.6**), it will be important to assess the strain when passaged over longer time periods, perhaps even up to the 7 days of the standard subcutaneous infection. Analysis of another bioluminescent *S. pyogenes* strain with an integrated *luxABCDE* was conducted over 7 days and found to be stable (204). Therefore, it is anticipated that this cassette will also be stable over a similar duration.

It will also be important to assess if the bioluminescent strain can be detected by the IVIS, as this will be the device used to image live animals. To do this, serial dilutions will be made of exponentially growing culture in a 96-well plate. This plate will then be imaged with the IVIS. Given that many of the lux strains presented here give greater or equal output of luminescent signal when compared to Xen20 (**Fig. 5.3 and 5.4**), it is anticipated that the strains will be adequate for the IVIS detection.

Finally, variation was observed in the fitness of two luminescent isolates, lux105 and lux127, when challenged with human neutrophils. This underscores the need to validate a luminescent strain in *in vitro* and *ex vivo* assays before utilizing in animal models. Like any strain generated by allelic exchange, it is possible that off

target mutations occur during the necessary passages to generate mutant strains. The difference in fitness may also be due to suppressor mutations that are occurring to accommodate the energy requirements of generating bioluminescence. For both reasons, it will be important to sequence the genomes of generated strains to understand if there are off target mutations. The commercial availability of whole-genome sequencing at relatively low cost should enable these analyses.

Summary

In this work a promising system has been constructed to generate luminescent strains for non-invasive *in vivo* monitoring of infections. The findings here demonstrate the P23-*luxABCDE* cassette can be integrated into 5448 and yield strains with similar growth and fitness to the WT parent with superior luminescent output that should enable highly sensitive *in vivo* monitoring of infection.

Chapter 6: Conclusions and Future Work

Conclusions

S. pyogenes is an important human pathogen that causes significant global morbidity and mortality (9) and is capable of infecting diverse tissue sites. *S. pyogenes* needs to sense its tissue environment and respond accordingly to survive in these niches and promote infection (28). To this end, *S. pyogenes*, like other bacterial pathogens, has seemingly developed systems and networks that couple metabolism and virulence regulation (31, 53, 57, 96). This includes the various PCVRs like Mga (86). In this dissertation, the virulence regulator RofA was investigated to determine if it had characteristics of PCVRs.

This work found that RofA plays a role in regulating several virulence genes and differential regulation has impacts on *in vitro* fitness. However, comparison of the regulon in 5448 and the RofA regulons of other M1 strains from the Sriskandan group (170) demonstrates that there is high intra-serotype variability of the RofA regulon. Additionally, studies on M6 and M2 strains (112, 113) demonstrate that there is a high inter-serotype variability. Combined, it seems that the pilin biogenesis genes are the only genes in the core RofA regulon.

This work also found that RofA was capable of being phosphorylated by EI and HPr *in vitro* and there is evidence of RofA phosphorylation *in vivo* that is dependent on glucose concentration. These findings are important evidence that RofA is phosphorylated by the PTS. However, the timing of the phosphorylation event is curious given that it occurs in stationary phase when transcription slows down. The

consequence of this late phosphorylation event will need to be investigated in future studies.

Overall, these findings support a model where in stationary phase after growth in high glucose conditions, CCR is relieved allowing HPr to become phosphorylated at His-15. Remaining glucose may move through both the PTS transporter and non-PTS transporter, leaving most of the HPr unphosphorylated (**Fig. 6.1**). This leaves no P~HPr available to phosphorylate RofA. After growth in low glucose conditions, glucose is exhausted, and the PTS components are phosphorylated to prime the cells for uptake

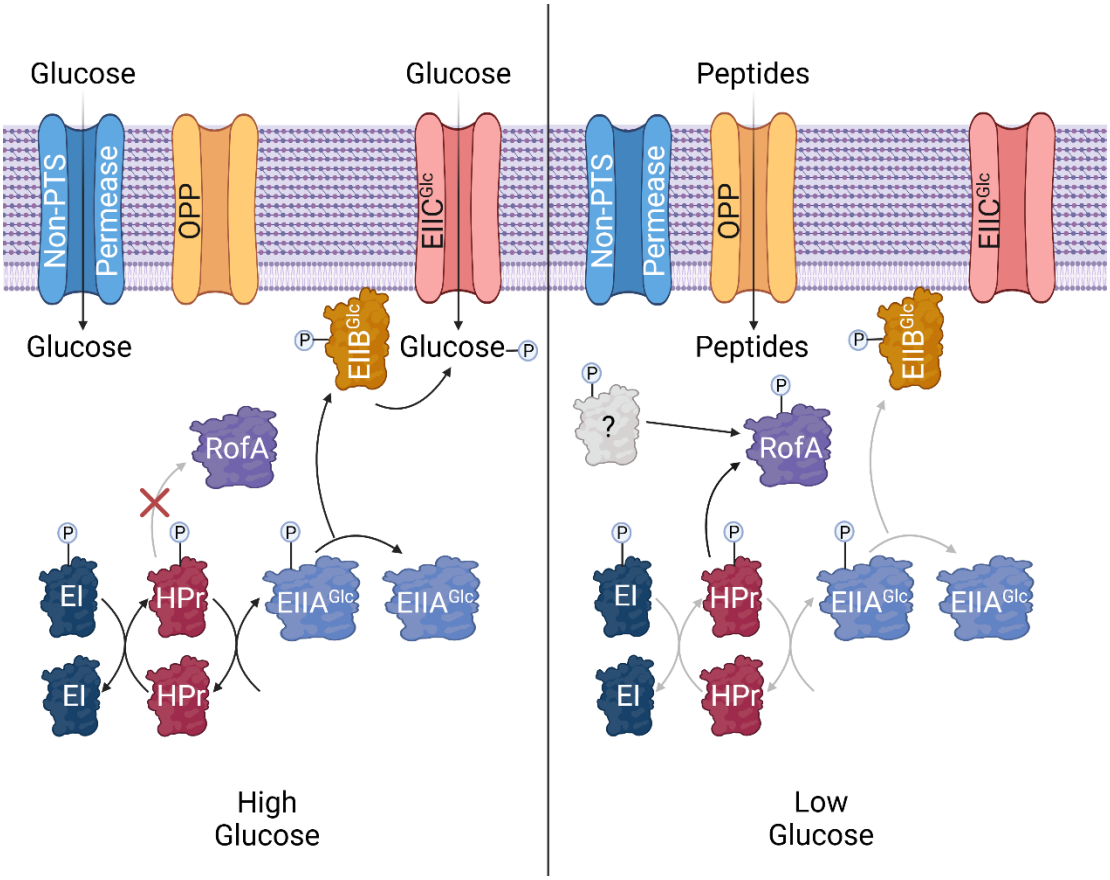


Figure 6.1: Model of RofA Phosphorylation
After growth in high glucose conditions, once CCR is relieved, remaining glucose may still come through both non-PTS transporters and PTS transporters. This leaves HPr largely unphosphorylated. After growth in low glucose conditions, HPr is phosphorylated to prime PTS for uptake of any carbohydrates that become available, and peptides are being transported for energy. This leaves HPr available for phosphorylating RofA. Other proteins (gray) may also be contributing to this phosphorylation. Created in BioRender.com

of carbohydrates should they become available. Cells sustain themselves on peptides, and HPr is available to phosphorylate RofA. It's not clear if there are other proteins also participating in the phosphorylation of RofA, though this would be an interesting avenue for future study.

The findings that RofA has a very small core regulon but impacts high numbers of genes within each strain studied, and that the phosphorylation event occurs late in stationary phase, suggests that RofA is behaving in an interesting way to exert its transcriptional control that is unlike classical transcriptional regulators. Understanding how RofA interacts with the chromosome in different growth phases will enable a better understanding of how it exerts its transcriptional control and may have interesting implications for the behavior of the broader PCVR family.

This dissertation also reports on the development of a bioluminescent construct for generating stable bioluminescent strains for real-time, non-invasive monitoring of *S. pyogenes* infections. Since *rofA* was identified as a gene important for oropharyngeal colonization (116) and as being highly expressed in soft-tissue infection (72), studying the mutant in both subcutaneous infection and pharyngeal colonization model is of interest. The findings that the luminescent 5448 generated has similar growth potential and similar fitness in PMN challenge is an encouraging sign that this system may be useful for *in vivo* studies. However, the increased doubling time is of concern because it suggests the strain has a fitness defect compared to WT. Since the revertant also has this increased doubling time, it seems possible that the strain may have obtained an off-target mutation that is impacting doubling time. It will be important to generate other bioluminescent strains to see if there is a consistent fitness cost associated with

obtaining the bioluminescence cassette. It may be necessary to modify the cassette to reduce the output of luminescence to create strains without that fitness cost.

Table 6.1: Updated Traits of PCVRs Including Data Generated Here

Trait	**Mga**	**AtxA**	**RofA**	**Nra**	**Ralp3**	**RivR**	**Mga-Spn**	**AcpA**	**AcpB**
Auto-regulated	(75)		(111, 112)	(120)		(122)			
Forms Multimers	(97)	(102)				(128)	(130)	(136)	(136)
P~ylated *in vitro* by EI/HPr	(95)		T. S.						
P~ylated *in vivo*	(92, 96)	(101)	T. S.						
Regulates Carb. Metab. Genes	(75)	(136)	T. S.			(122) (127)	(131)	(136)	(136)
Impacts Virulence in Animal Models	(139)	(98, 99)			(125)	(127)	(129)	(138) (137)	(138) (137)
Maximal Express.	EX (89)		TR T. S.	EST (120)	EX (125)				

Abbreviations:
"P~ylated" = Phosphorylated
"Carb." = Carbohydrate
"Metab" = Metabolism
"Express." = Expression
"EX" = Mid-exponential
"EST" = Early stationary
"TR" = Transitions
"T. S." = This study

Green indicates that the trait has been shown. Yellow indicates 'inconclusive' or 'supportive, but not strong' results. Red indicates that it's been shown that the protein does not have that trait. References are provided in box by number. 'EX' is mid-exponential. 'Trans.' is transition. 'Stat' is stationary.
MgaSpn was previously referred to as MgrA.

Overall, this work sought to characterize another member of the PCVR family and establish a tool for efficient *in vivo* infection monitoring. This work further established how RofA is behaving and contributed to the overall understanding of PCVRs (**Table 6.1**) to determine what the common features of this family are. Additionally, a tool was created to study the contribution of RofA to *in vivo* infection that also has broader applications to the studies of other *S. pyogenes* genes and how they contribute to virulence during infection.

Future Work

Determine if *in vivo* RofA Phosphorylation Is Due to the PTS

Though this work found that RofA has behaviors of PCVRs, there remain several outstanding questions. Here we show that EI and HPr can phosphorylate RofA *in vitro*, but it has not been established which systems or individual proteins are responsible for the observed phosphorylation *in vivo*. Attempts to study the phosphorylation in bacteria containing native RofA were not successful (**Fig. 4.6**), therefore, a double Δ*ptsI*/Δ*rofA_sc* mutant is required. Making this mutant would enable phosphorylation studies in the absence of a functional PTS system. It would also be valuable to show interactions between PTS proteins and RofA. This could be achieved using a bacterial two-hybrid screen with HPr and RofA. This system is currently being used by another member of the lab to study interactions between Mga and PTS proteins.

Another approach would be co-immunoprecipitation and western blotting. For this assay, His_6-HPr could be used and a specific antibody for RofA developed by a

commercial source. Alternatively, an antibody for HPr could be developed, and then the existing RofA-FLAG and RofA-His constructs could be used. It may also be interesting to do a pull-down of HPr in various conditions and for the various PCVRs in *S. pyogenes* to see if they are interacting with HPr. This would include Mga, Ralp3, and RivR for an M1T1 strain. Such a pull down could also be done in a strain containing Nra, such as M49 strain 591 or NZ131.

Determine if PTS EIIA and EIIB Proteins Interact with and Phosphorylate RofA

Classic PRD-containing regulators are subject to phosphorylation not just by HPr, but also by sugar-specific EII proteins (41). It is therefore possible that RofA is subject to EII phosphorylation. This may also explain why there was evidence of differential phosphorylation of the quadruple alanine substitution mutant *in vivo*, but not *in vitro* (**Fig. 3.14**). To study the interactions between RofA and the different EII proteins, a bacterial two hybrid approach could be employed. *S. pyogenes* contains 17 PTS EII genes that contain either an EIIA, EIIB, or both. While the time and effort to clone each of these constructs is not trivial, once this system is made it could be used to test the interactions of multiple hypothesized PCVRs.

Alternatively, a simpler approach may be to conduct co-immunoprecipitation experiments for lysates of the RofA-FLAG strain with an anti-FLAG antibody and then identify the co-precipitating proteins through proteomics. The benefit to this approach is that it is unbiased, and we may find interactions between RofA and other proteins that were not expected. If an EIIA or EIIB protein is identified that interacts with RofA,

the phosphorylation of RofA could then be assessed *in vitro* including this EII protein and *in vivo* in deletion mutants of that specific EII to determine if it is phosphorylating RofA.

Impact of Phosphorylation on RofA Activity

This work has shown that RofA is phosphorylated *in vitro* and *in vivo* (**Fig. 3.7**). An important area of future work will be what impact this phosphorylation has on the activity of RofA. The current model for PCVRs is that phosphorylation in response to carbohydrate metabolism modulates the activity of the regulator to modulate both virulence and metabolic gene expression. It is possible, however, that phosphorylation is simply an artifact of the PRD domain structure but has no impact on the protein function.

As a first step, it may be possible to utilize the *in vitro* phosphorylation assay components to phosphorylate RofA using EI and HPr. This phosphoprotein could then be used in *in vitro* transcription assays on the promoter from the *rofA-cpa* intergenic region where the two predicted RofA binding sites are. This method was successfully used to study the impact of phosphorylation on Mga transcriptional activity (95). Though this strategy was successful for Mga, not all RofA becomes phosphorylated in this system, and it seems possible that the presence of non-phosphorylated RofA could obscure differences in transcriptional activity. To address this, the commercially available P~His antibody (Creative Diagnostics, (210)) could be used to enrich for the phosphorylated RofA; then, this enriched sample could be used for the *in vitro* assays. This would require the generation of the pTze phosphohistidine analog to elute the

protein from the pHis antibody, therefore it may require collaboration with a synthetic chemistry group. If EII proteins can be identified that interact with RofA, then the specific purified EIIs could be included in the *in vitro* phosphorylation assay, then this EII-phosphorylated species could also be assayed for its transcriptional activity.

S. pyogenes Phosphoproteome and Phosphorylated Residues of PCVRs

While this work demonstrated that RofA is phosphorylated, the specific amino acid residues that are the targets of phosphorylation are still unknown. Alanine substitutions at H210, H278, H330, and H380 seem to have reduced the phosphorylation *in vivo*, but did not completely ablate it (**Fig. 3.14A**). Additionally, these substitutions did not seem to impact observed *in vitro* phosphorylation at all (**Fig. 3.14B**). It will be important to understand where RofA is being phosphorylated and how this compares to the phosphorylation sites of other PCVRs like Mga and AtxA. The McIver Lab is currently working with a collaborator to establish methods that will successfully enrich phosphohistidine residues so that phosphorylated residues can be identified on PCVRs. This has been challenging historically due to the acidification that is common with most liquid chromatography, tandem mass spectrometry (LC-MS/MS) approaches, as phosphohistidines tend to hydrolyze in acidic conditions (211, 212).

As a first step toward this, a phosphoproteomics experiment could be performed with *S. pyogenes* to establish the full phosphoproteome of WT 5448. Steps should be taken to ensure phosphohistidine residues can be observed. One approach would be to employ a modified version of the sample preparation method developed by Potel and

colleagues that significantly improved the number of phosphohistidine peptides identified (211). In this protocol, cell suspensions would first be incubated with PlyC, a phage lysin that efficiently lyses *S. pyogenes* (213), before adding the remaining denaturing agents from the Potel lysis buffer and proceeding with their outlined digest and clean-up. Two enrichments could be performed on the peptides in sequence to maximize the different types of phosphorylation sites identified. The first enrichment should be done with an Fe^{3+}-IMAC column, which has been shown to enrich for phosphohistidine-containing peptides due to its relatively mild acidic conditions (211). This should be followed by a titanium dioxide (TiO_2) enrichment, which requires acidification, but the phosphohistidine-containing peptides should already have been enriched from the Fe^{3+}-IMAC step. Phosphoproteomics could also be performed on a Δ*ptsI* mutant to understand how the phosphoproteome changes when the PTS is non-functional. If PCVRs function as hypothesized, these should be observed in the WT phosphoproteome, but not in the Δ*ptsI* proteome.

Phosphorylated PCVRs would ideally be identified in the WT enrichments, and the specific phosphorylated residues identified. As an alternative or complementary approach, the commercial P~His antibody can be utilized for immunoprecipitation of specifically histidine phosphorylated peptides (210). These approaches will allow for the interrogation of changes in phosphorylation state with high resolution. Additionally, once the phosphorylated residues have been identified, targeted mutagenesis at these residues can be made to probe how phosphoablative and phosphomimetic residues at these sites impact the activity of RofA.

Multimerization of RofA

Multimerization is a feature that many PCVRs have and this multimerization is important for full transcriptional activation (97, 102, 128, 130, 136) (**Table 6.1**). Multimerization was not specifically assessed in this work, though, like Mga, it was observed that purified RofA forms aggregates in solution. RofA multimerization could be evaluated by native gel electrophoresis of anti-FLAG pull-downs for strains carrying the RofA-FLAG allele, as was done for AtxA (102). As a negative control for cross-reactive proteins, anti-FLAG pull downs will also be performed for a WT 5448 lysate. It should be noted that native gel electrophoresis was not effective for assessing Mga multimerization (97), therefore another strategy may need to be used.

An alternative strategy would be to treat lysates with the crosslinking agent bis(maleimido)hexane (BMH) and then evaluate the treated lysates with SDS-PAGE and western blotting, as was done for RivR and AtxA (102, 128). Two other alternative strategies for assessing multimerization include using gel filtration chromatography, as was done for Mga (97), and *in vitro* Co-IP experiments as was done for AtxA (102). Unfortunately, the coimmunoprecipitation (Co-IP) experiments for *in vivo* multimerization performed for Mga will likely not be possible, as the truncated versions of Mga used to distinguish the Mga-His from native Mga lacked the C-terminal unstructured region. RofA does not have this unstructured region, so C-terminal truncations would impact the EIIB region that we expect participates in multimerization (**Fig. 4.7**). The development of an anti-RofA antibody may also be worthwhile to study the multimerization of native protein lacking an epitope tag.

Impact of RofA Phosphorylation Late in Stationary Phase

The observed phosphorylated RofA species only appeared late in the stationary phase where there is not robust transcriptional activity. It is not clear what role the phosphorylated RofA plays in this context. Given that PCVRs also have properties of nucleoid associated proteins (NAPs) (86), the findings from this study open the question of whether a subset of RofA proteins become phosphorylated in a context of limited transcriptional activity so that they can act as nucleoid scaffolding proteins. Some NAPs, like H-NS, preferentially bind and oligomerize along AT-rich segments of DNA and bind between strands of DNA to organize it (214, 215). Additionally, another putative PCVR, MgaSpn, oligomerizes along DNA in a sequence-independent manner (130). To understand if RofA is forming oligomers along stretches of DNA, EMSAs could be performed with RofA and DNA containing the RofA binding site. The EMSAs could be done with both HPr phosphorylated and unphosphorylated RofA, to determine how phosphorylation impacts DNA binding. Previous EMSA analysis of purified MBP-RofA did not show any higher order multimers being formed (110), however, it seems possible that the conjugated MBP may impact interactions between RofA monomers. Therefore, evaluation of purified proteins with smaller epitope tags, and potentially even of a native RofA protein, would be useful to investigate. The binding of RofA to other DNA probes could also be assessed to determine if RofA binds to DNA in a non-specific manner, like MgaSpn (130). Further, different sequences of different AT concentrations could be assessed to determine if RofA binds preferentially to AT-rich sequences.

Additionally, chromatin immunoprecipitation sequencing (ChIP-Seq) could be employed to determine the regions of DNA that RofA is interacting with and how it is organized along the chromosome. This could be done with samples in context where RofA is unphosphorylated, like in growth in high-glucose C medium and in a context where it is phosphorylated, like in regular C medium. This would be useful to interrogate RofA in its native context, rather than *in vitro*. ChIP-seq could also be done with phosphoablative and phosphomimetic mutants of RofA to determine how these substitutions impact how RofA interacts with DNA. If RofA binds to DNA along long-stretches of chromatin, this may suggest that it is acting as a NAP. However, if RofA is only interacting with the chromosome at discrete sites, this may suggest that it is behaving more like a traditional transcription factor.

In vivo Virulence of Bioluminescence *S. pyogenes*

To ensure that the constitutively expressed bioluminescent reporter does not impact the *in vivo* virulence of *S. pyogenes*, the bioluminescent strains will need to be assessed in infections. To do this, bioluminescent 5448 will be used in the subcutaneous infection model to determine if it has the same infection characteristics as wild type. This would mean similar lesion size, similar lesion severity, and similar lethality. If the strain is attenuated compared to 5448, exploration of different promoters (216) may be pursued to tune the luminescent output to see if less output results in similar virulence. Given that our current 5448_lux candidate is so much more bioluminescent than the previously commercially available Xen20 strain that was shown to work *in vivo*

suggests that some bioluminescent output could be lost and the strain would still be bright enough to be visualized *in vivo*.

An *in vivo* competition assay, as done previously (204), may also be utilized to assess whether the 5448_lux strain is as fit as the WT parent. Even if the luminescent strains are slightly less fit than WT 5448, like the *emm75* luminescent strain was (204), it may not matter if all strains have this attenuation. Virulence could still be compared for luminescent mutants of other genes to luminescent WT, even with some level of attenuation, as the attenuation should be consistent for all strains.

This strain also must be evaluated with the In Vivo Imaging System (IVIS) to take images of the bioluminescent bacteria in a mouse during infection. First, a pilot experiment should be done to determine if the strain remains bioluminescent throughout the 7-day period of our typical subcutaneous infections. Next, the bioluminescence intensity from tissues should be compared to the bacterial burden as determined by CFU enumeration. This would show whether the bioluminescent intensity correlates with bacterial burden.

Use of the Xen20 strain during a subcutaneous infection demonstrated that the bacteria that had disseminated to internal organs were not exhibiting bioluminescence (203). Therefore, it will be important to assess whether or not the 5448_lux strain exhibits bioluminescence in the internal organs. The Xen20 luminescence was shown to be glucose sensitive (203), presumably due to the promoter present at the transposon insertion site. Since the P23-luxABCDE cassette is being driven by a constitutive promoter, it is hypothesized that this strain will not have this issue and will be visible in the internal organs.

It will also be important to assess the same parameters in a pharyngeal colonization model. *S. pyogenes* has been shown to localize to the nasal associated lymphoid tissue (NALT) in a mouse (148), which is analogous to the human tonsil. Bioluminescence would be especially useful for this model as the only metric of this model is bacterial burden, which requires sacrificing the animal, excising the NALT tissues, and enumerating CFU from the homogenates. Assessing these models may require developing a luminescent cassette with an antibiotic resistance gene to enable the specific selection of *S. pyogenes* during CFU enumeration. However, other literature using these models did not mention the use of antibiotics to select for *S. pyogenes* (148, 200, 217), and one group found that no other β-hemolytic streptococci colonized the nasopharyngeal tract of mice (217). Therefore, homogenates could be plated directly onto blood agar as done previously (200, 217), and bacterial burden of *S. pyogenes* specifically enumerated. Alternatively, homogenates can be plated onto Remel Strep Select plates (Thermo-Fisher) to inhibit the growth of other commensals, then the colonies can be replicated plated or patched on blood agar to count the β-hemolytic colonies.

Evaluation of the Impact of RofA on *in vivo* Infections

RofA has been found in transposon mutant library screens to be important for oropharyngeal colonization, and potentially reduces fitness in subcutaneous infections (115, 116). In contrast to the findings that it reduces fitness in a subcutaneous infection model (115), it was also found to be highly expressed during subcutaneous infection (72). Therefore, the contribution of RofA to *in vivo* virulence remains unclear. The

contribution of RofA to *in vivo* virulence should be assessed in each mouse model of infection. The bioluminescent reporter construct will be of great use for these experiments once it is validated. Additionally, it is of interest to determine how phosphorylation of RofA contributes to virulence *in vivo*. If phosphorylated residues can be identified on RofA with LC-MS/MS, then these residues could be targets for phosphomimetic and phosphoablative substitutions and these strains with mutant RofA could be studied in both models.

Another potential avenue of study would be to assess the phosphorylation of RofA during infection. To do this, animals could be inoculated intranasally or subcutaneously on the haunch, to model throat colonization and soft-tissue infection, respectively (194). Then, animals could be sacrificed at different time points for tissue collection and bacterial recovery. The bacteria used for infection would be carrying an epitope-tagged version of RofA at the native locus, which will need to be generated by allelic exchange. RofA could be immunoprecipitated from tissue homogenates, first for Phos-Tag analysis, then for LC-MS/MS. Alternatively, native RofA could be pulled out of these lysates using a commercially generated RofA antibody. It would be wise to include some kind of histidine-phosphorylated positive control to ensure that the methods used for tissue homogenization and lysis do not result in P~His hydrolysis. Since both the His_6-EI and His_6-HPr proteins have been demonstrated to function in the *in vitro* phosphorylation assays, these alleles could be introduced to *S. pyogenes*, then this could be pulled-down and assayed with Phos-Tag and/or LC-MS/MS to ensure that it can be identified in its phosphorylated form.

Appendices

A1: Carbohydrates in Body Compartments and Dietary Sources

A2: Table of Putative PTS Systems in *S. pyogenes*

A3: Protein Purifications of PTS Proteins and RofA for *in vitro* PTS Assay

A4: Differentially Expressed Genes in Δ*rofA_sc* as Identified by RNA-Seq

A5: Caliper LifeSciences Product Literature Page for Xen20

Appendix A1: Carbohydrates in Body Compartments and Dietary Sources

	Ontology			Biological Properties			
Carb.	**Biological Location**	**Exogenous Sources**	**Examples**	**Cellular Locations**	**Biospecimen Locations**	**Tissue Locations**	**Other Notes**
Glucose	Cytoplasm	Food - Animal Origin	Chicken, Turkey, Beef, Pork	Cytoplasm	Blood	Adipose Tissue	
	Saliva	Food - Vegetables	Broccoli, Onion, Carrot, Spinach, Potato	Extracellular	Breast Milk	Adrenal Cortex	
	Feces	Food - Fruits	Pear, Strawberry, Raspberry, Apple, Banana	Lysosome	Cerebrospinal Fluid (CSF)	Adrenal Gland	
		Food - Nuts	Peanut, Walnut, Almond	Endoplasmic Reticulum	Feces	Adrenal Medulla	
		Food - Cereal & Grains	Oat, Wheat, Rice		Saliva	Bladder	
					Sweat	Brain	
					Urine	Epidermis	
						Eye Lens	
						Fibroblasts	
						Intestine	
						Kidney	
						Liver	
						Lung	
						Neuron	

	Ontology			Biological Properties			
Carb.	**Biological Location**	**Exogenous Sources**	**Examples**	**Cellular Locations**	**Biospecimen Locations**	**Tissue Locations**	**Other Notes**
Glucose cont.						Ovary	
						Pancreas	
						Placenta	
						Prostate	
						Skeletal Muscle	
						Testis	
Glucose-6-P	Placenta	Food - Animal Origin	Chicken, Turkey, Beef, Pork	Endoplasmic Reticulum	Blood	Adipose Tissue	
	Kidney	Food - Vegetables	Broccoli, Onion, Carrot, Spinach, Potato		Cellular Cytoplasm	Adrenal Gland	
	Liver	Food - Fruits	Pear, Strawberry, Raspberry, Apple, Banana		Feces	Fibroblasts	
	Adrenal Gland	Food - Nuts	Peanut, Walnut, Almond		Saliva	Kidney	
	Cellular Cytoplasm	Food - Cereal & Grains	Oat, Wheat, Rice		Urine	Liver	
	Blood					Placenta	
	Saliva					Skeletal Muscle	
	Urine						
	Feces						
	Endoplasmic reticulum						
	Adipose Tissue						

	Ontology			Biological Properties			
Carb.	**Biological Location**	**Exogenous Sources**	**Examples**	**Cellular Locations**	**Biospecimen Locations**	**Tissue Locations**	**Other Notes**
Glucose-6-P Cont.	Skeletal Muscle						
	Fibroblasts						
Sucrose	N/A	N/A	N/A	Extracellular	Blood	Adipose Tissue	*Food - Sugar
					Feces	Bladder	*Food - Leaves, Roots
					Saliva	Epidermis	Table sugar (from sugar cane, sugar beetroot, etc)
					Sweat	Fibroblasts	In description lists that its widely found in various plant tissues, suggesting that food plants also contain this sugar, despite that not being listed under exogenous sources.
					Urine	Intestine	
						Kidney	
						Neuron	
						Placenta	
						Platelet	
						Prostate	
						Skeletal Muscle	
						Spleen	
						Testis	
Sucrose-6-P	N/A	Food - Nuts	Walnuts, Almonds, Cashews, Peanut	N/A	N/A	N/A	
		Food - Fruit	Apple, Pear, Banana, Strawberry, Raspberry				

	Ontology			Biological Properties			
Carb.	**Biological Location**	**Exogenous Sources**	**Examples**	**Cellular Locations**	**Biospecimen Locations**	**Tissue Locations**	**Other Notes**
Sucrose-6-P cont.		Food - Vegetables	Broccoli, Onion, Carrot, Spinach, Potato				
		Food - Cereal & Grains	Oat, Wheat, Rice				
Mannose	Urine	Food - Vegetables	Broccoli, Onion, Carrot, Spinach, Potato	Cytoplasm	Blood	Adrenal Medulla	
	Feces	Food - Cereal & Grains	Oat, Wheat, Rice	Extracellular	Breast Milk	Bladder	
	Breast Milk	Food - Fruit	Apple, Pear, Banana, Strawberry, Raspberry	Lysosome	Cerebrospinal Fluid (CSF)	Epidermis	
	Blood	Food - Nuts	Walnuts, Almonds, Cashews, Peanut	Endoplasmic Reticulum	Feces	Fibroblasts	
	Cerebrospinal Fluid (CSF)	Food - Animal Origin	Chicken, Turkey, Beef, Pork	Golgi Apparatus	Saliva	Intestine	
	Saliva				Urine	Kidney	
	Fibroblasts					Neuron	
	Neuron					Pancreas	
	Intestine					Placenta	
	Kidney					Prostate	
	Pancreas					Skeletal Muscle	
	Placenta					Spleen	
	Spleen					Testis	
	Testis						

	Ontology			Biological Properties			
Carb.	**Biological Location**	**Exogenous Sources**	**Examples**	**Cellular Locations**	**Biospecimen Locations**	**Tissue Locations**	**Other Notes**
Mannose cont.	Prostate						
	Bladder						
	Skeletal Muscle						
	Epidermis						
	Adrenal Medulla						
	Cytoplasm						
	Extracellular						
	Golgi apparatus						
	Lysosome						
Mannose-6-P	Testicle	Food - Vegetables	Broccoli, Onion, Carrot, Spinach, Potato	Cytoplasm	Feces	Fibroblasts	
	Fibroblasts	Food - Cereal & Grains	Oat, Wheat, Rice			Intestine	
	Intestine	Food - Fruit	Apple, Pear, Banana, Strawberry, Raspberry			Liver	
	Liver	Food - Nuts	Walnuts, Almonds, Cashews, Peanut			Prostate	
	Spleen	Food - Animal Origin	Chicken, Turkey, Beef, Pork			Spleen	
	Testis					Testis	
	Prostate						

	Ontology			Biological Properties			
Carb.	**Biological Location**	**Exogenous Sources**	**Examples**	**Cellular Locations**	**Biospecimen Locations**	**Tissue Locations**	**Other Notes**
Mannose-6-P cont.	Feces						
	Cytoplasm						
Fructose	N/A	N/A	N/A	Extracellular	Blood	Liver	In the description of "Metabolite Identification" *Food - Honey
					Cerebrospinal Fluid (CSF)	Placenta	*Food - Fruit
					Feces	Prostate	*Food - Root Vegetables
					Saliva	Semen	Tree fruits, berries, melons
					Urine		Beets, sweet potato, parsnip, onion
Fructose-6-P	Cellular Cytoplasm	Food - Animal Origin	Chicken, Turkey, Beef, Pork	Cytoplasm	Blood	Intestine	
	Blood	Food - Vegetables	Broccoli, Onion, Carrot, Spinach, Potato		Cellular Cytoplasm	Prostate	
	Saliva	Food - Fruit	Banana, Apple, Pear, Strawberry, Raspberry		Saliva		
	Intestine	Food - Nuts	Cashew, Peanut, Walnut, Almond				
	Prostate	Food - Cereal & Grains	Oat, Rice, Wheat				
	Cytoplasm						
α-Lactose	N/A	N/A	N/A	Cytoplasm	Blood	Bladder	*Food - Milk
				Extracellular	Breast Milk	Epidermis	"Metabolite Identification"- Desc lists as major sugar in milk.
				Lysosome	Feces	Intestine	

	Ontology			Biological Properties			
Carb.	**Biological Location**	**Exogenous Sources**	**Examples**	**Cellular Locations**	**Biospecimen Locations**	**Tissue Locations**	**Other Notes**
α-Lactose cont…				Golgi Apparatus	Urine	Platelet	
						Skeletal Muscle	
						Spleen	
β-Lactose	Extracellular	Food - Fermented Bevs	Beer, Grape Wine, Liquor	Cytoplasm	Blood	N/A	
	Cytoplasm	Food - Fish	Anchovy, Haddock, Catfish, Crab	Extracellular	Urine		
		Food - Nuts	Cashew, Peanut, Walnut, Almond				
		Food - Fruit	Apple, Pear, Strawberry				
		Food - Vegetables	Broccoli, Carrot, Potato,				
		Food - Animal Origin	Chicken, Turkey, Beef, Pork				
		Food - Cereal & Grains	Oat, Rice, Wheat				
Lactose-6-P	N/A	Food - Animal Origin	Chicken, Turkey, Beef, Pork	N/A	N/A	N/A	
Maltose	N/A	N/A	N/A	Cytoplasm	Blood	Kidney	*Food
				Extracellular	Feces	Liver	Starch is broken down to Maltose
				Lysosome	Sweat	Placenta	Oriental wheats, sweet potato, grape wines, spinach

	Ontology			Biological Properties			
Carb.	**Biological Location**	**Exogenous Sources**	**Examples**	**Cellular Locations**	**Biospecimen Locations**	**Tissue Locations**	**Other Notes**
Maltose cont.					Urine	Platelet	
Maltotriose	N/A	N/A	N/A	Cytoplasm	Blood	Kidney	
				Extracellular	Feces	Liver	
				Lysosome	Urine	Placenta	
						Prostate	
Malto-dextrin	N/A	Food - Fermented Bevs	Beer, Grape Wine, Liquor	N/A	N/A	N/A	
		Food - Fish	Anchovy, Haddock, Catfish, Crab				
		Food - Nuts	Cashew, Peanut, Walnut, Almond				
		Food - Fruit	Banana, Apple, Pear, Strawberry, Raspberry				
		Food - Vegetables	Broccoli, Onion, Carrot, Spinach, Potato				
		Food - Animal Origin	Chicken, Turkey, Beef, Pork				
Salicin	Cytoplasm			N/A	N/A	N/A	
Cellobiose	Cytoplasm	Food - Nuts	Cashew, Peanut, Walnut, Almond	Cytoplasm	Urine	Intestine	

	Ontology			Biological Properties			
Carb.	**Biological Location**	**Exogenous Sources**	**Examples**	**Cellular Locations**	**Biospecimen Locations**	**Tissue Locations**	**Other Notes**
Cellobiose cont…	Intestine	Food - Fruit	Banana, Apple, Pear, Strawberry, Raspberry			Spleen	
	Spleen	Food - Vegetables	Broccoli, Onion, Carrot, Spinach, Potato				
	Urine	Food - Cereal & Grains	Oat, Rice, Wheat				
Trehalose	N/A	N/A	N/A	Extracellular	Feces	Epidermis	
					Urine	Fibroblasts	
						Kidney	
						Platelet	
						Skeletal Muscle	
Trehalose-6-P	Cytoplasm	Food - Animal Origin	Chicken, Turkey, Beef, Pork	Cytoplasm	N/A	N/A	
		Food - Vegetables	Broccoli, Onion, Carrot, Spinach, Potato				
		Food - Nuts	Cashew, Peanut, Walnut, Almond				
		Food - Fruit	Banana, Apple, Pear, Strawberry, Raspberry				
		Food - Cereal & Grains	Oat, Rice, Wheat				

	Ontology			Biological Properties			
Carb.	**Biological Location**	**Exogenous Sources**	**Examples**	**Cellular Locations**	**Biospecimen Locations**	**Tissue Locations**	**Other Notes**
Ascorbic Acid	Adipose Tissue	Food - Fermented Bevs	Grape wine				
	Epidermis	Food - Vegetable	Broccoli, lettuce	Extracellular	Blood	Adrenal Cortex	
	Eye Lens	Food - Milk	Milk (cow)		Breast Milk	Adrenal Medulla	
	Skeletal Muscle				Cellular Cytoplasm	Bladder	
	Blood				Cerebrospinal Fluid (CSF)	Brain	
	Breast Milk				Saliva	Epidermis	
	Cerebrospinal Fluid (CSF)				Urine	Erythrocyte	
	Cellular Cytoplasm					Eye Lens	
	Amniotic Fluid					Fibroblasts	
	Saliva					Heart	
	Adrenal Cortex					Intestine	
	Adrenal Medulla					Leukocyte	
	Bladder					Liver	
	Brain					Lung	
	Testicle					Neuron	
	Heart					Ovary	
	Intestine					Placenta	
	Liver					Platelet	
	Lung					Prostate	

	Ontology			Biological Properties			
Carb.	**Biological Location**	**Exogenous Sources**	**Examples**	**Cellular Locations**	**Biospecimen Locations**	**Tissue Locations**	**Other Notes**
Ascorbic Acid cont.	Placenta					Skeletal Muscle	
	Prostate					Spleen	
	Spleen					Testis	
	Ovary						
	Erythrocyte						
	Fibroblasts						
	Leukocyte						
	Neuron						
	Platelet						
	Urine						
	Cytoplasm						
	Extracellular						
N-acetyl-galactose-amine	N/A	N/A	N/A	Lysosome	Blood	Adipose Tissue	"Metabolite Identification" - Desc: Mucins,
				Golgi Apparatus	Feces	Brain	Brain Glycoproteins
					Saliva	Epidermis	
					Urine	Intestine	
						Liver	
						Neuron	
						Pancreas	
						Placenta	
						Prostate	
						Skeletal Muscle	
						Spleen	
						Testis	

	Ontology			Biological Properties			
Carb.	**Biological Location**	**Exogenous Sources**	**Examples**	**Cellular Locations**	**Biospecimen Locations**	**Tissue Locations**	**Other Notes**
N-acetyl-glucos-amine-6-P	Placenta	Food - Animal Origin	Chicken, Turkey, Beef, Pork	Cytoplasm		Placenta	
	Cytoplasm	Food - Nuts	Cashew, Peanut, Walnut, Almond				
		Food - Fruit	Banana, Apple, Pear, Strawberry, Raspberry				
		Food - Vegetables	Broccoli, Onion, Carrot, Spinach, Potato				
		Food - Cereal & Grains	Oat, Rice, Wheat				
D-Galactose	Membrane	Food - Fruit	Banana, Apple, Pear, Strawberry, Raspberry	Extracellular	Blood	Brain	
	Extracellular	Food - Vegetables	Broccoli, Onion, Carrot, Spinach, Potato	Lysosome	Breast Milk	Liver	
	Lysosome	Food - Milk	Milk (Cow), Cheese, Yogurt		Cellular Cytoplasm	Prostate	
	Prostate	Food - Fish	Cod, Halibut, Trout, Salmon, Crab		Cerebrospinal Fluid (CSF)		
	Brain	Food - Cereal & Grains	Rice, Wheat		Feces		
	Liver	Food - Nuts	Cashew, Peanut, Almond		Saliva		

	Ontology			Biological Properties			
Carb.	**Biological Location**	**Exogenous Sources**	**Examples**	**Cellular Locations**	**Biospecimen Locations**	**Tissue Locations**	**Other Notes**
D-Galactose cont..	Blood	Food - Animal Origin	Beef, Turkey, Pork		Urine		
	Breast Milk						
	Cerebrospinal Fluid (CSF)						
	Cellular Cytoplasm						
	Saliva						
	Urine						
	Feces						
Galactose-1-P	N/A	N/A	N/A	Cytoplasm	Blood	Erythrocyte	
						Platelet	

Carb. = Carbohydrate
Source: Human Metabolome Database (29)

Overview by Compartment

	Glu	Glu-6-P	Suc	Man	Man-6-P	Fru	Fru-6-P	α-Lac	β-Lac	Mal	Mal-Tri	Sal	Cel	Tre	Tre-6-P	A.A.	NAc-Gal	NAc-Glu	NAc-Glu-6-P	Gal	Gal-1-P
Amniotic Fluid																X					
Blood	X	X	X	X		X	X	X	X	X	X					X	X	X		X	X
Breast Milk	X			X				X								X				X	
Cellular Cytoplasm		X					X									X				X	
Cerebrospinal Fluid (CSF)	X			X		X										X				X	
Cytoplasm	X			X	X		X	X	X	X	X	X	X		X	X		X	X		X
Endoplasmic Reticulum	X	X		X																	
Extracellular	X		X	X		X		X	X	X	X			X		X				X	
Lysosome	X			X				X		X	X						X			X	
Feces	X	X	X	X	X	X		X		X	X			X			X	X		X	
Golgi Apparatus				X				X									X				
Membrane																				X	
Saliva	X	X	X	X		X	X									X	X	X		X	

	Glu	Glu-6-P	Suc	Man	Man-6-P	Fru	Fru-6-P	α-Lac	β-Lac	Mal	Mal-Tri	Sal	Cel	Tre	Tre-6-P	A.A.	NAc-Gal	NAc-Glu	NAc-Glu-6-P	Gal	Gal-1-P
Semen						X															
Sweat	X		X							X											
Urine	X	X	X	X		X		X	X	X	X		X	X		X	X	X		X	
Adipose Tissue	X	X	X													X	X				
Adrenal Cortex	X															X					
Adrenal Gland	X	X																			
Adrenal Medulla	X			X												X					
Bladder	X		X	X				X								X					
Brain	X															X	X			X	
Epidermis	X		X	X				X						X		X	X				
Erythrocyte																X					X
Eye Lens	X															X					
Fibroblasts	X	X	X	X	X									X		X					

	Glu	Glu-6-P	Suc	Man	Man-6-P	Fru	Fru-6-P	α-Lac	β-Lac	Mal	Mal-Tri	Sal	Cel	Tre	Tre-6-P	A.A.	NAc-Gal	NAc-Glu	NAc-Glu-6-P	Gal	Gal-1-P
Heart																X					
Intestine	X		X	X	X		X	X					X			X	X				
Kidney	X	X	X	X						X	X			X							
Leukocyte																X					
Liver	X	X			X	X				X	X					X	X			X	
Lung	X															X					
Neuron	X		X	X												X	X				
Ovary	X															X					
Pancreas	X			X													X				
Placenta	X	X	X	X		X				X	X					X	X		X		
Platelet			X					X		X				X		X					X
Prostate	X		X	X	X	X	X				X					X	X			X	
Spleen			X	X	X			X					X			X	X				
Skeletal Muscle	X	X	X	X				X						X		X	X				

	Glu	Glu-6-P	Suc	Man	Man-6-P	Fru	Fru-6-P	α-Lac	β-Lac	Mal	Mal-Tri	Sal	Cel	Tre	Tre-6-P	A.A.	NAc-Gal	NAc-Glu	NAc-Glu-6-P	Gal	Gal-1-P
Testis/ Testicle	X		X	X	X											X	X				

Glu = Glucose; Suc = Sucrose; Man = Mannose; Fru = Fructose; Lac = Lactose; Mal = Maltose; Mal-Tri = Maltotriose; Sal = Salicin; Cel = Cellobiose; Tre = Trehalose; A.A. = Ascorbic Acid; NAc-Gal = N-acetylgalactosamine; NAc-Glu = N-acetylglucosamine; NAc-Glu-6-P = N-acetylglucosamine-6-phosphate; Gal = Galactose

P = phosphate

Source: Human Metabolome Database (29)

Concentrations of Carbohydrates in Different Body Fluids

	Amniotic Fluid (µM)	Blood (µM)	Breast Milk (µM)	Cellular Cytoplasm (µM)	CSF (µM)	Feces (nmol/g wet feces)	Saliva (µM)	Sweat (µM)	Urine (µmol/mmol creatinine)
Glucose		4700	3000		336	9577	1232	180	212
Glu-6-P		16.8		38		D.N.Q.	3.2		D.N.Q.
Sucrose		1.8				D.N.Q.	3361	10	11.75
Mannose		39	66.4		44	1010	D.N.Q.		5.4
Man-6-P						D.N.Q.			
Fructose		39.5			200	D.N.Q.	D.N.Q.		106.1
Fru-6-P		10.1		24.4			0.497		
α-Lactose		E.N.Q.	170000			D.N.Q.			100
β-Lactose		E.N.Q.							D.N.Q.
D-Galacose		95.5	92.3	4.6	166	747	83.5		83.7
Gal-1-P		46.5							
Maltose						D.N.Q.		D.N.Q.	10.7
Maltotriose		D.N.Q.							E.N.Q.
Cellobiose									D.N.Q.
Trehalose						D.N.Q.			5.298
Ascorbic Acid	20.4	50.2	96.6	68	132		3.11		34.3
N-acetyl galactosamine						D.N.Q.	D.N.Q.		
N-acetyl glucosamine		D.N.Q.				D.N.Q.	D.N.Q.		D.N.Q.

If cell is blank, that indicates that there was no data available for that carbohydrate in the body fluid.
E.N.Q. indicates Expected, but not quantified.; D.N.Q. indicates Detected, but not quantified
Source: Human Metabolome Database (29)

Appendix A2: Table of Putative PTS Systems in *S. pyogenes*

Carbohydrate	KEGG Pathway Given in 2016 eBook	KEGG Associated Protein	Component	KEGG MGAS5005 Gene Given (*Gene Name*)	Consistent w Sundar et al 2017?	Notes
Ascorbate	M00283	UlaA	EIIC	M5005_Spy0148/M5005_Spy1662 (*ptxC*/*spy1662*)	Yes/No	Annotated as mannitol
		UlaB	EIIB	M5005_Spy0149/M5005_Spy1663 (*ptxB*/*spy1663*)	Yes/No	Annotated as mannitol
		UlaC	EIIA	M5005_Spy0150/M5005_Spy1664 (*ptxA*/*spy1664*)	Yes/No	Annotated as mannitol
Lactose	M00281	LacE	EIICB	M5005_Spy1633 (*lacE*)	Yes	
		LacF	EIIA	M5005_Spy1634 (*lacF*)	Yes	
Galactitol	M00279	N/A	N/A	N/A	N/A	N/A
Galactose	N/A	EIICGal	EIIC	M5005_Spy1399 (*spy1399*)	Yes	
		EIIBGal	EIIB	M5005_Spy1400 (*spy1400*)	Yes	
		EIIAGal	EIIA	M5005_Spy1401 (*spy1401*)	Yes	
Mannose	M00276	ManY	EIIC	M5005_Spy0782/M5005_Spy1480 (*ptsC*/*manM*)	Yes/Yes	
		ManZ	EIID	M5005_Spy0783/M5005_Spy1481 (*ptsD*/*manN*)	Yes/Yes	
		ManX	EIIA	M5005_Spy0780 (*ptsA*)	Yes and No	Sundar et al. lists Spy0781/*ptsB* as the EIIB also. Though in KEGG this is shown in the fructoselysine PTS as a GfrB homolog (see below)
			EIIAB	M5005_Spy1479 (*manL*)	Yes	

Carbohydrate	KEGG Pathway Given in 2016 eBook	KEGG Associated Protein	Component	KEGG MGAS5005 Spy Number Given (*Gene Name*)	Consistent w Sundar et al 2017?	Notes
Cellobiose	M00275	CelB	EIIC	M5005_Spy1079/M5005_Spy1744 (*spy1079*/*celC*)	Yes/Yes	
		CelC	EIIA	M5005_Spy1081/M5005_Spy1746 (*spy1081*/*celA*)	Yes and No/Yes	Sundar et al. also has Spy1083 showing as an EIIA
		CelA	EIIB	M5005_Spy1082/M5005_Spy1745 (*spy1082*/*celB*)	Yes/Yes	
Fructose	M00273	FruA/B	EIIBC	M5005_Spy0662 (*fruA*)	Yes and no	Sundar et al. listed Spy0667 as EIIABC
Trehalose	M00270	TreB	EIIIBC or EIIBCA	M5005_Spy1784 (*treB*)	Yes and no	Sundar et al. lists as EIIBC.
		TreP/Crr	EIIBC or EIIBCA	Not present	Not present	
Sucrose	M00269	ScrA	EIIBCA or EIIBC	M5005_Spy1542 (*scrA*)	Yes	Sundar et al. lists as the EIIABC
Glucose	M00265	N/A	N/A	N/A	N/A	
N-acetylglucos-amine	None Given	N/A	N/A	N/A	N/A	N/A
N-acetylgalactos-amine	N/A	AgaW	EIIC	M5005_Spy0520 (*agaS*)	Yes	
		AgaE	EIID	M5005_Spy0519 (*agaD*)	Yes	
		AgaV	EIIB	M5005_Spy0521 (*agaV*)	Yes	
		AgaF		Not present	Not present	Not present
Maltose	N/A	MalT	EIICB or EIIBCA	M5005_Spy1692	Yes	Sundar et al. shows it as an EIIABC
β-glucosides	N/A	BglF	EIIBCA	M5005_Spy0475 (*bglP*)	Yes	

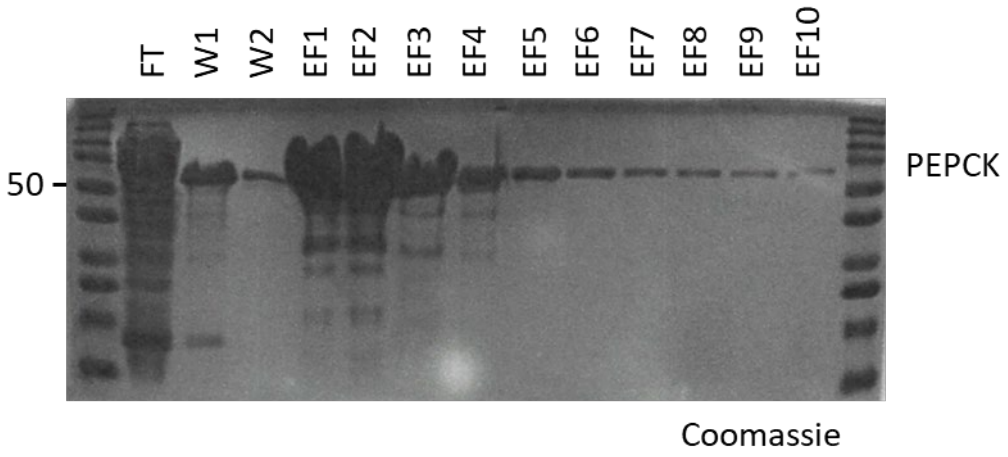
Carbohydrate	KEGG Pathway Given in 2016 eBook	KEGG Associated Protein	Component	KEGG MGAS5005 Spy Number Given (*Gene Name*)	Consistent w Sundar et al 2017?	Notes
Fructoselysine / glucoselysine	N/A	GfrC	N/A	N/A	N/A	N/A
		GfrD	N/A	N/A	N/A	N/A
		GfrA	N/A	N/A	N/A	N/A
		GfrB	EIIB	M5005_0781 (*ptsB*)	No	Same operon as one of the mannose PTS systems. Likely for that and not fructoselysine.

Appendix A3: Protein Purifications of PTS Proteins and RofA for *in vitro* PTS Assay

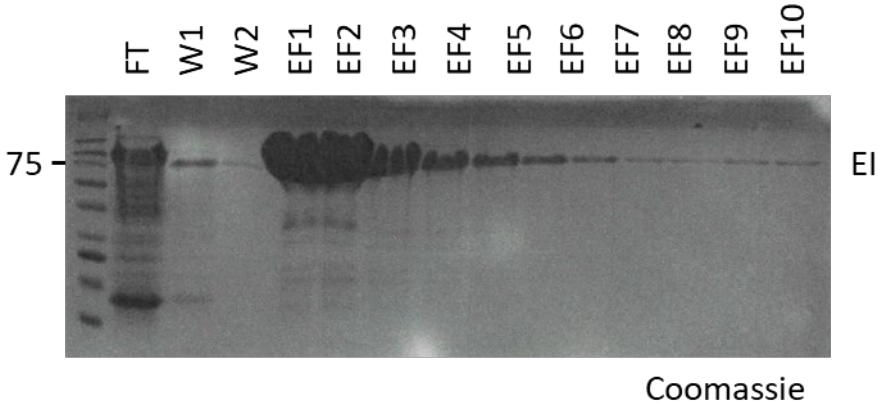
Abbreviations: FT = Flow-through; W# = Wash #; EF# = Elution Fraction #, MW = Molecular Weight; kDa = kilo Dalton

Number on the left hand side indicates relevant band from the molecular weight ladder in kDa

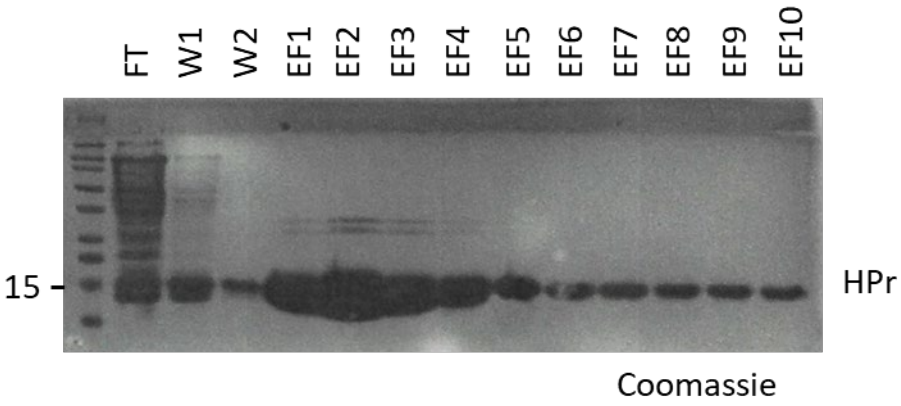
PEPCK: Expected MW: 62.4 kDa; EF 1&2 Pooled



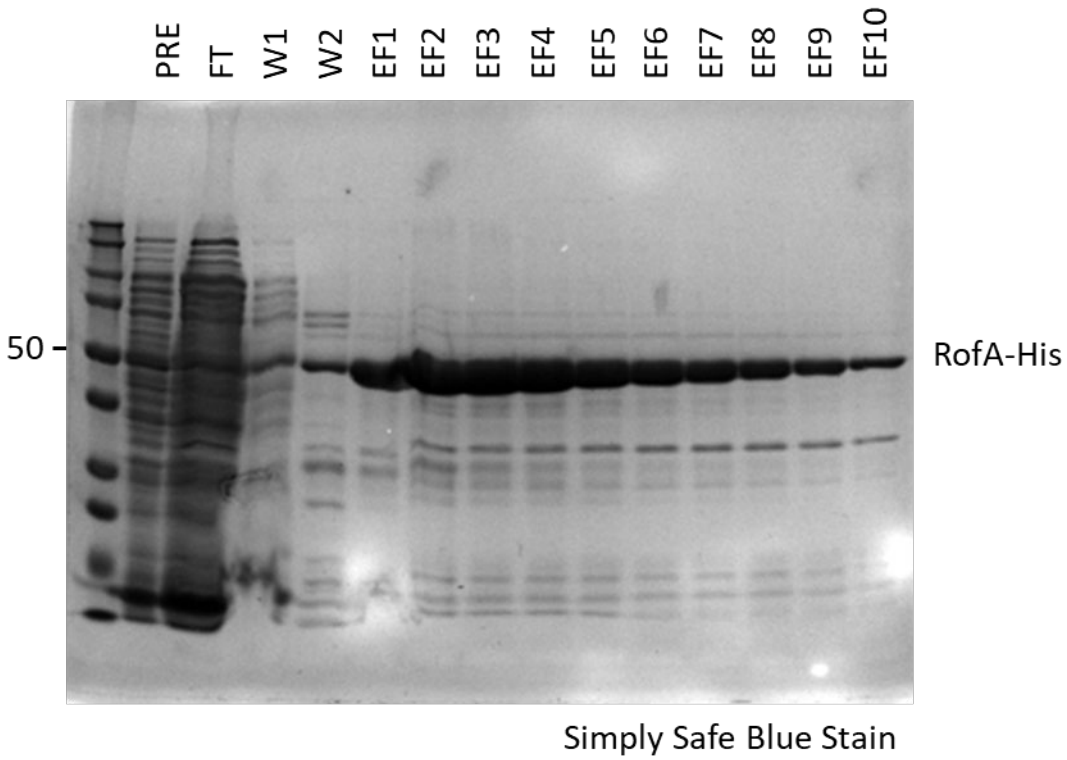
His_6-EI: Expected MW: 65.9 kDa; EF 1&2 Pooled



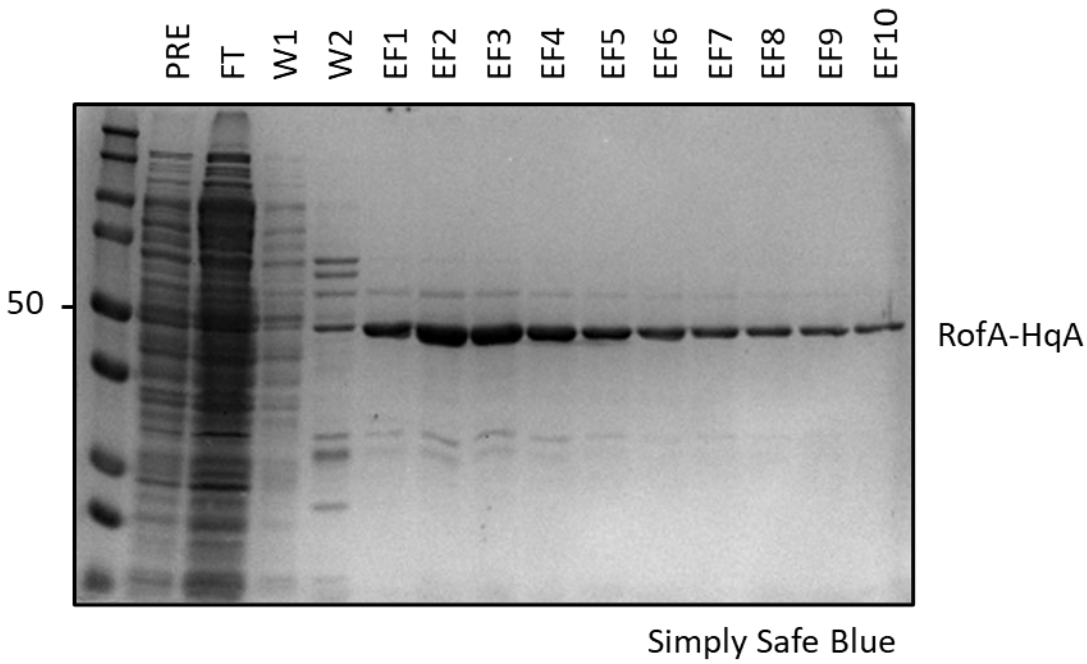
$\underline{His_6\text{-HPr}}$: Expected MW: 11.5 kDa; EF 1- 3 Pooled



$\underline{RofA\text{-}His_6}$: Expected MW: 59.4 kDa; EF 1-5 Pooled



RofA-HqA-His_6: Expected MW: 59.1 kDa; EF 1-5 Pooled



Appendix A4: Differentially Expressed Genes in Δ*rofA sc* as Identified by RNA-Seq

Table Abbreviations:
"Carb.": carbohydrate.; "bc": because. ; "Refs.": references ; "No.": numbers ; "Desc.": description ; "Cat.": category ; "Met": Methionine ; "w/": with;
"*L. monocytogenes*": *Listeria monocytogenes; "S."* of *"S. agalactiae", "S. dysgalactiae", "S. suis",* and "*S. gordonii*": *Streptococcus*

Significantly Down Regulated Genes in Exponential

		Delta v WT		Delta v Revertant		Microbes Online	COG				
5005 Gene No.	**Gene Name**	**deseq_ logfc**	**deseq_ adjp**	**deseq_ logfc**	**deseq_ adjp**	**Desc/Link**	**COG No.**	**COG Cat.**	**Notes**	**My Category**	**Refs.**
Spy0028		-1.072	0.002759	-1.445	3.826E-05	autolysin (NCBI)	COG3942	R	Cell Wall, but bc there was only one entry, shifted to 'Other'.	Other	
Spy0053	*rpsQ*	-1.518	8.348E-13	-1.007	3.543E-06	SSU ribosomal protein S17P (NCBI)	COG186 (RpsQ)	J		Translation & Protein Processing	
Spy0059	*rplF*	-1.428	8.79E-13	-1.055	1.971E-07	LSU ribosomal protein L6P (NCBI)	COG97 (RplF)	J		Translation & Protein Processing	
Spy0062	*rpmD*	-1.38	5.927E-15	-1.148	1.071E-10	LSU ribosomal protein L30P (NCBI)	COG1841 (RpmD)	J		Translation & Protein Processing	
Spy0063	*rplO*	-1.24	3.51E-19	-1.016	3.641E-13	LSU ribosomal protein L15P (NCBI)	COG200 (RplO)	J		Translation & Protein Processing	
Spy0064	*secY*	-1.102	6.819E-23	-1.028	5.913E-20	protein translocase subunit (NCBI)	COG201 (SecY)	U	Originally listed as protein secretion, but that would have been a 1 gene category, so changed 'Translation' to 'Translation & Protein Processing'	Translation & Protein Processing	
Spy0092		-1.045	0.03089	-1.013	0.03936	comG operon protein 6 (NCBI)	N/A	N/A	Aligns with comp. operon from S. mutans & comYG shown to be impt for competence. No conditions have been found in which *S. pyogenes* is naturally competent.	Conjugation & Competence	(218)

		Delta v WT		Delta v Revertant		Microbes Online	COG				
5005 Gene No.	**Gene Name**	**deseq_ logfc**	**deseq_ adjp**	**deseq_ logfc**	**deseq_ adjp**	**Desc/Link**	**COG No.**	**COG Cat.**	**Notes**	**My Category**	**Refs.**
Spy0106	*rofA*	-3.104	1.223E-74	-3.434	9.076E-92	transcriptional regulator (NCBI)	N/A	N/A	Regulated pilus operon in all serotypes tested. Regulates other virulence genes as well.	Regulator	
Spy0107	*cpa /cbp*	-9.055	0	-9.336	0	fibronectin-binding protein (NCBI)	N/A	N/A	Pilus minor subunit ap1. Tip subunit. Could have also been categorized as 'Adhesion'	Virulence	(109)
Spy0108	*sipA1*	-9.21	2.58E-193	-9.232	9.17E-195	signal peptidase I (NCBI)	N/A	N/A	Required chaperone for pilus assembly of FCT type 2 strains. Could have also been categorized as 'Adhesion'	Virulence	(109)
Spy0109	*fctA*	-8.624	0	-8.928	0	fibronectin-binding protein (NCBI)	N/A	N/A	Major pilus subunit. Could have also been categorized as 'Adhesion'.	Virulence	(109)
Spy0110	*srtC2*	-8.122	3.59E-298	-8.577	0	hypothetical protein (NCBI)	COG4509	S	Class B sortase for pilus assembly. Could have also been categorized as 'Adhesion'.	Virulence	(109)
Spy0111		-8.367	0	-8.692	0	hypothetical protein (NCBI)	N/A	N/A	fctB - pilus minor subunit ap2. Base subunit. Could have also been categorized as 'Adhesion'.	Virulence	(109)
Spy0114	*srtB*	-4.756	2.324E-49	-4.669	1.107E-47	sortase (NCBI)	COG3764 (SrtA)	M	Class C sortase for pilus assembly. Could have also been categorized as 'Adhesion'.	Virulence	(109)
Spy0116	*atoE*	-3.167	1.598E-07	-3.223	1.002E-07	short-chain fatty acids transporter (NCBI)	COG2031 (AtoE)	I	Originally categorized as 'Lipid Transport & Metabolism'. To cut down on number of categories for data presentation, recategorized as 'Other Transport & Metabolism'	Lipid Transport & Metabolism	
Spy0119		-2.713	3.156E-10	-2.588	3.057E-09	acetyl-CoA acetyltransferase (NCBI)	COG183 (PaaJ)	I	Same note as Spy0116	Lipid Transport & Metabolism	
Spy0120	*atoD.2*	-1.701	2.923E-05	-1.899	0.0000028	acetate CoA-transferase alpha subunit (NCBI)	COG1788 (AtoD)	I	Same note as Spy0116	Lipid Transport & Metabolism	

		Delta v WT		Delta v Revertant		Microbes Online	COG				
5005 Gene No.	**Gene Name**	**deseq_ logfc**	**deseq_ adjp**	**deseq_ logfc**	**deseq_ adjp**	**Desc/Link**	**COG No.**	**COG Cat.**	**Notes**	**My Category**	**Refs.**
Spy0121		-2.153	2.025E-06	-2.395	1.052E-07	acetyl-CoA:acetoacetyl-CoA transferase beta (NCBI)	COG2057 (AtoA)	I	Same note as Spy0116	Lipid Transport & Metabolism	
Spy0125		-1.276	1.062E-05	-1.208	3.704E-05	hypothetical protein (NCBI)	N/A	N/A		Unknown	
Spy0126	*ntpI*	-1.448	3.765E-08	-1.456	3.44E-08	V-type sodium ATP synthase subunit I (NCBI)	COG1269 (NtpI)	C	KEGG indicated in comments that it pumps H+ & some bacterial V-type sodium ATPases are involved with sodium homeostasis. Further, homologous operon in *E. faecalis* is the primary mechanism for extruding excess sodium (219). *Streptococcus* species MAY use it to overcome pH stress (220). In *Enterococcus hirae*, the homolog pumps Na at high pH to survive at high pHs (221).	Metal/Ion/ pH Homeostasis	(219-221)
Spy0127	*ntpK*	-2.022	1.654E-09	-1.725	4.825E-07	V-type sodium ATP synthase subunit K (NCBI)	N/A	N/A	Same note as Spy0126	Metal/Ion/ pH Homeostasis	(219-221)
Spy0128	*ntpE*	-2.144	3.66E-13	-1.797	1.959E-09	V-type sodium ATP synthase subunit E (NCBI)	N/A	N/A	Same note as Spy0126	Metal/Ion/ pH Homeostasis	(219-221)
Spy0129	*ntpC*	-1.871	1.77E-11	-1.968	1.507E-12	V-type ATP synthase subunit C (NCBI)	COG1527 (NtpC)	C	Same note as Spy0126	Metal/Ion/ pH Homeostasis	(219-221)
Spy0130	*ntpF*	-1.617	1.463E-06	-1.498	0.0000104	V-type sodium ATP synthase subunit F (NCBI)	COG1436 (NtpG)	C	Same note as Spy0126	Metal/Ion/ pH Homeostasis	(219-221)
Spy0131	*ntpA*	-1.889	2.911E-16	-1.312	3.005E-08	V-type sodium ATP synthase subunit A (NCBI)	COG1155 (NtpA)	C	Same note as Spy0126 Subunit has ATP binding	Metal/Ion/ pH Homeostasis	(219-221)

		Delta v WT		Delta v Revertant		Microbes Online	COG				
5005 Gene No.	**Gene Name**	**deseq_ logfc**	**deseq_ adjp**	**deseq_ logfc**	**deseq_ adjp**	**Desc/Link**	**COG No.**	**COG Cat.**	**Notes**	**My Category**	**Refs.**
Spy0132	*ntpB*	-1.785	5.06E-19	-1.47	4.696E-13	V-type sodium ATP synthase subunit B (NCBI)	COG1156 (NtpB)	C	Same note as Spy0126 Subunit has ATP binding	Metal/Ion/ pH Homeostasis	(219-221)
Spy0133	*ntpD*	-1.54	7.983E-12	-1.27	3.187E-08	V-type sodium ATP synthase subunit D (NCBI)	COG1394 (NtpD)	C	Same note as Spy0126.	Metal/Ion/ pH Homeostasis	(219-221)
Spy0148	*ptxC*	-1.99	0.01054	-1.879	0.01681	PTS system, 3-keto-L-gulonate specific IIC component (NCBI)	COG3037 (SgaT)	S	Microbes Online InterPro: Sugar-Specific Permease | EC # would likely be changed to 2.7.1.194. KEGG Comments: This enzyme is a component (known as enzyme II) of a phosphoenolpyruvate (PEP)-dependent, sugar transporting phosphotransferase system (PTS). KEGG Pathways: ec00053: Ascorbate and aldarate metabolism, ec01100: Metabolic pathways, ec01120: Microbial metabolism in diverse environments	Carb. Transport & Metabolism	(39)
Spy0149	*ptxB*	-2.564	0.0007412	-2.307	0.002741	PTS system, 3-keto-L-gulonate specific IIB component (NCBI)	COG3414 (SgaB)	G	Microbes Online InterPro: EIIB component | EC # would likely be changed to 2.7.1.194. KEGG Comments: This enzyme is a component (known as enzyme II) of a phosphoenolpyruvate (PEP)-dependent, sugar transporting phosphotransferase system (PTS). KEGG Pathways: ec00053: Ascorbate and aldarate metabolism, ec01100: Metabolic pathways, ec01120: Microbial metabolism in diverse environments	Carb. Transport & Metabolism	(39)

		Delta v WT		Delta v Revertant		Microbes Online	COG				
5005 Gene No.	**Gene Name**	**deseq_ logfc**	**deseq_ adjp**	**deseq_ logfc**	**deseq_ adjp**	**Desc/Link**	**COG No.**	**COG Cat.**	**Notes**	**My Category**	**Refs.**
Spy0150	*ptxA*	-2.137	7.381E-06	-1.843	0.0001466	PTS system, 3-keto-L-gulonate specific IIA component (NCBI)	COG1762 (PtsN)	G | T	Microbes Online InterPro: EIIA | EC # would likely be changed to 2.7.1.194. KEGG Comments: This enzyme is a component (known as enzyme II) of a phosphoenolpyruvate (PEP)-dependent, sugar transporting phosphotransferase system (PTS). KEGG Pathways: ec00053: Ascorbate and aldarate metabolism, ec01100: Metabolic pathways, ec01120: Microbial metabolism in diverse environments	Carb. Transport & Metabolism	(39)
Spy0151	*ulaD*	-2.065	1.152E-05	-1.534	0.001567	3-keto-L-gulonate-6-phosphate decarboxylase (NCBI)	COG269 (SgbH)	G	KEGG Comments: This enzyme is involved in a pathway for the utilization of L-ascorbate by *E. coli*. | KEGG Pathways: ec00040: Pentose and glucuronate interconversions, ec00053: Ascorbate and aldarate metabolism, ec01100: Metabolic pathways, ec01120: Microbial metabolism in diverse environments	Carb. Transport & Metabolism	

		Delta v WT		Delta v Revertant		Microbes Online	COG				
5005 Gene No.	**Gene Name**	**deseq_ logfc**	**deseq_ adjp**	**deseq_ logfc**	**deseq_ adjp**	**Desc/Link**	**COG No.**	**COG Cat.**	**Notes**	**My Category**	**Refs.**
Spy0152		-1.701	8.224E-05	-1.893	1.061E-05	L-xylulose 5-phosphate 3-epimerase (NCBI)	COG3623 (SgaU)	G	Microbes Online didn't have the EC #, so searched in the KEGG database w/ enzyme name. L-xylulose 5-phosphate 3-epimerase is synonymous with L-ribulose-5-phosphate 3-epimerase. | KEGG Comments: enzyme is involved in a pathway for the utilization of L-ascorbate by *Escherichia coli.* | KEGG Pathways: ec00040: Pentose and glucuronate interconversions, ec00053: Ascorbate and aldarate metabolism, ec01100: Metabolic pathways, ec01120: Microbial metabolism in diverse environments	Carb. Transport & Metabolism	
Spy0153	*araD*	-2.8	1.666E-13	-2.414	3.437E-10	L-ribulose-5-phosphate 4-epimerase (NCBI)	COG235 (AraD)	G	KEGG Pathways: ec00040: Pentose and glucuronate interconversions, ec00053: Ascorbate and aldarate metabolism, ec01100: Metabolic pathways, ec01120: Microbial metabolism in diverse environments	Carb. Transport & Metabolism	
Spy0170	*nadC*	-2.165	4.215E-12	-2.091	2.485E-11	nicotinate-nucleotide pyrophosphorylase (carboxylating) (NCBI)	COG157 (NadC)	H	NAD Metabolism - involved in many cellular processes	Other Transport & Metabolism	(28)
Spy0176		-1.763	6.095E-10	-1.754	8.265E-10	zinc finger protein (NCBI)	COG4357	S		Unknown	
Spy0177		-1.493	8.019E-09	-1.626	2.963E-10	bioY protein (NCBI)	COG1268 (BioY)	R	BLAST: 32% Identity to BioY from *Rhodobacter capsulatus*, 96% query cover. Total score: 73.2. This BioY transporter heterologously expressed in *E. coli* allowed for biotin uptake.	Other Transport & Metabolism	(222)

		Delta v WT		Delta v Revertant		Microbes Online	COG				
5005 Gene No.	**Gene Name**	**deseq_ logfc**	**deseq_ adjp**	**deseq_ logfc**	**deseq_ adjp**	**Desc/Link**	**COG No.**	**COG Cat.**	**Notes**	**My Category**	**Refs.**
Spy0184		-1.106	2.718E-06	-1.148	1.192E-06	hypothetical cytosolic protein (NCBI)	N/A	N/A		Unknown	
Spy0212		-3.341	3.175E-20	-3.022	1.225E-16	N-acetylmannosamine-6-phosphate 2-epimerase (NCBI)	COG3010 (NanE)	G	KEGG Pathways: ec00520: Amino sugar and nucleotide sugar metabolism, ec01100 Metabolism pathways	Carb. Transport & Metabolism	
Spy0213		-2.94	1.949E-15	-2.79	4.993E-14	N-acetylneuraminate-binding protein (NCBI)	COG1653 (UgpB)	G		Carb. Transport & Metabolism	
Spy0214		-3.027	1.272E-15	-2.852	5.379E-14	N-acetylneuraminate transport system permease protein (NCBI)	COG1175 (UgpA)	G		Carb. Transport & Metabolism	
Spy0215		-2.765	6.988E-15	-2.644	1.045E-13	N-acetylneuraminate transport system permease protein (NCBI)	COG395 (UgpE)	G		Carb. Transport & Metabolism	
Spy0216		-2.771	6.772E-13	-2.521	8.033E-11	hypothetical membrane spanning protein (NCBI)	N/A	N/A	InterPro notes it's a DUF624 containing protein. Transcribed w *nanH* and *M5005_Spy0218*	Unknown	
Spy0217	*nanH*	-3.441	5.039E-24	-2.964	7.26E-18	N-acetylneuraminate lyase (NCBI)	COG329 (DapA)	E	KEGG Pathways: ec00520: Amino sugar and nucleotide sugar metabolism; ec01100: Metabolic pathways. No mention of amino acid metabolism, or present in amino acid metabolism pathways.	Carb. Transport & Metabolism	

		Delta v WT		Delta v Revertant		Microbes Online	COG				
5005 Gene No.	**Gene Name**	**deseq_ logfc**	**deseq_ adjp**	**deseq_ logfc**	**deseq_ adjp**	**Desc/Link**	**COG No.**	**COG Cat.**	**Notes**	**My Category**	**Refs.**
Spy0218		-2.567	1.375E-17	-2.046	2.351E-11	N-acetylmannosamine kinase (NCBI)	COG1940 (NagC)	K | G	KEGG Pathways: ec00520: Amino sugar and nucleotide sugar metabolism; ec01100: Metabolic pathways.	Carb. Transport & Metabolism	
Spy0402		-1.394	2.889E-07	-1.235	7.116E-06	hypothetical protein (NCBI)	N/A	N/A	Homolog in NZ131 is spy49_0412 which appears to be quorum sensing regulated (223). Homolog in SF370 is SPy_0412 and might be secreted (224).	Unknown	(223, 224)
Spy0403		-1.542	5.044E-06	-1.016	0.003721	hypothetical protein (NCBI)	N/A	N/A	Homolog to M49 591 protein (SpyoM01000198, 97% Identical) labeled with COG0642: Signal transduction histidine kinase. Homolog is NZ131 (spy49_0413, 97% Identical) and appears to be quorum sensing regulated.	Unknown	(223)
Spy0404	*stcA*	-1.171	5.536E-05	-1.067	0.0003041	hypothetical protein (NCBI)	N/A	N/A	Homolog in NZ131 (spy49_0414) confers resistance to lysozyme and highly positively charged.	Virulence	(223)
Spy0406	*shp2*	-1.109	0.0001892	-1.426	1.062E-06	hypothetical protein (NCBI)	N/A	N/A	Mistake in annotation. Not *shp2* based off DNA seq from Chang et al.	Unknown	(225)
Spy0440	*rgg3*	-1.191	2.811E-07	-1.624	1.092E-12	transcriptional regulator (NCBI)	N/A	N/A	Repressor of both *shp* promoters. Originally categorized as 'Quorum sensing' but this would have made a very specific category with only 1 gene. Could have been recategorized as 'Regulator' or 'Other'. Since 'Regulator' was more descriptive, chose that.	Regulator	(225)
Spy0452		-1.196	9.234E-05	-1.519	4.69E-07	chromosome segregation ATPases (NCBI)	N/A	N/A	Has a DUF2326 domain. Homologs listed as ATPase involved in DNA repair.	Replication and Genome Integrity	

		Delta v WT		Delta v Revertant		Microbes Online	COG				
5005 Gene No.	**Gene Name**	**deseq_ logfc**	**deseq_ adjp**	**deseq_ logfc**	**deseq_ adjp**	**Desc/Link**	**COG No.**	**COG Cat.**	**Notes**	**My Category**	**Refs.**
Spy0459		-2.137	0.00551	-1.685	0.03462	portal protein (NCBI)	N/A	N/A	IPR006944: Phage portal protein. Domains include 'Phage Portal Protein'.	Phage	
Spy0474	*licT*	-1.597	2.772E-17	-1.268	3.76E-11	transcription antiterminator, BglG family (NCBI)	N/A	N/A		Regulator	
Spy0476	*bglA*	-1.326	8.271E-24	-1.277	5.31E-22	6-phospho-beta-glucosidase (NCBI)	COG2723 (BglB)	G	KEGG Pathway: ec00010: Glycolysis / Gluconeogenesis, ec00500: Starch and sucrose metabolism	Carb. Transport & Metabolism	
Spy0523		-1.382	0.003105	-1.335	0.004506	hypothetical protein (NCBI)	N/A	N/A		Unknown	
Spy0650		-1.014	0.005315	-1.583	6.942E-06	hypothetical protein (NCBI)	N/A	N/A		Unknown	
Spy0651	*hupY*	-1.917	2.194E-30	-2.216	1.258E-40	cell surface protein (NCBI)	N/A	N/A	Homolog in NZ131 (Spy49_0661, *hupY*) is important for heme utilization. Binds heme *in vitro*. Deletion results in attenuation in vaginal colonization model.	Metal/Ion/ pH Homeostasis	(226)
Spy0652	*hupZ*	-1.517	3.863E-14	-1.811	6.571E-20	hypothetical cytosolic protein (NCBI)	COG3576	R	Recombinant HupZ from NZ131 (100% identity) degrades heme *in vitro* (227). "HupZ contributes to heme metabolism and host survival, likely as a heme chaperone" (228)	Metal/Ion/ pH Homeostasis	(227, 228)
Spy0754		-1.537	0.03081	-1.553	0.0307	hypothetical protein (NCBI)	N/A	N/A	N/A	Unknown	
Spy0769	*cas9*	-1.181	5.419E-21	-1.326	2.747E-26	hypothetical cytosolic protein (NCBI)	COG3513	S	TIGR01865 CRISPR-associated protein Cas9/Csn1, subtype II/NMEMI [*cas9*] | TIGR01865 CRISPR-associated protein Cas9/Csn1, subtype II/NMEMI [*cas9*]	CRISPR and Other Defense	

		Delta v WT		Delta v Revertant		Microbes Online	COG				
5005 Gene No.	**Gene Name**	**deseq_ logfc**	**deseq_ adjp**	**deseq_ logfc**	**deseq_ adjp**	**Desc/Link**	**COG No.**	**COG Cat.**	**Notes**	**My Category**	**Refs.**
Spy0781	*ptsB*	-1.456	5.06E-19	-1.336	4.257E-16	PTS system, mannose/fructose family IIB component (NCBI)	COG3444	G	EC # likely would transfer to either EC 2.7.1.191 (Mannose phosphotransferase) OR EC 2.7.1.202 (fructose phosphotransferase). Both list ec00051 Fructose and mannose metabolsim under pathways.	Carb. Transport & Metabolism	
Spy0782	*ptsC*	-1.476	5.64E-16	-1.381	4.207E-14	PTS system, mannose/fructose family IIC component (NCBI)	COG3715 (ManY)	G		Carb. Transport & Metabolism	
Spy0783	*ptsD*	-1.207	1.11E-09	-1.267	1.536E-10	PTS system, mannose/fructose family IID component (NCBI)	COG3716 (ManZ)	G		Carb. Transport & Metabolism	
Spy0785	*tcs5R*	-1.155	2.953E-12	-1.283	6.672E-15	two-component response regulator (NCBI)	COG4753	T		Regulator	
Spy0798		-4.222	6.493E-06	-3.111	0.001246	IFN-response binding factor 1 (NCBI)	N/A	N/A	Listed as a virulence-related gene in Valdes *et al*. Homologs in Spn though list it as a replication initiator protein.	Virulence	(229)
Spy0803	*srtI*	-3.669	0.0001829	-3.81	9.808E-05	protein involved in lantibiotic production (NCBI)	N/A	N/A	InterPro: IPR021112: Lantibiotic streptin immunity protein. || Homolog in *S. suis* (AN924_RS09970, 70% identity) is a relaxase involved in plasmid conjugation.	Conjugation & Competence	(230)

		Delta v WT		Delta v Revertant		Microbes Online	COG				
5005 Gene No.	**Gene Name**	**deseq_ logfc**	**deseq_ adjp**	**deseq_ logfc**	**deseq_ adjp**	**Desc/Link**	**COG No.**	**COG Cat.**	**Notes**	**My Category**	**Refs.**
Spy0889		-1.217	0.0009653	-1.507	3.528E-05	regulatory protein (pfoS/R) (NCBI)	N/A	N/A	Homologous to PTS fructpse specific IIBC component in *S. agalactiae* (84%) and EIIC transporter in *S. dysgalactiae* (95%). Homologous to putative toxin regulator in *S. pyogenes* M49 (93%). Since there is conflict with the annotations, I'm inclined to categorize as 'Unknown'.	Unknown	
Spy0890	*ddh*	-1.604	1.317E-05	-1.852	0.0000004	D-lactate dehydrogenase (NCBI)	COG1052 (LdhA)	C| H | R	Kegg Pathway: ec00620 - Pyruvate metabolism; ec01100 - Metabolic pathways; Microbial metabolism in diverse environments	Other Transport & Metabolism	
Spy0891	*satD*	-1.081	8.708E-10	-1.477	1.985E-17	hypothetical protein (NCBI)	N/A	N/A	Homolog in NZ131 (Spy49_0920, 100% Identity) listed as a putative acid tolerance SatD-like protein. So could be Stress Tolerance.	Unknown	
Spy0900		-3.438	1.151E-33	-3.096	2.469E-27	Mg2+/citrate complex secondary transporter (NCBI)	COG2851 (CitM)	C	InterPro: Divalent ion symporter, and citrate transporter.	Other Transport & Metabolism	
Spy0914		-1.035	2.234E-06	-1.281	3.256E-09	phage transcriptional repressor (NCBI)	COG2932	K	Upregulated in response to peroxide stress in a PerR dependent manner (231). 71% identical to *S. uberis* HdiR which is upregulated in response to UV-Damage and may be a part of the SOS response (232).	Regulator	(231, 232)
Spy0964		-1.081	2.018E-15	-1.084	1.603E-15	type I restriction-modification system specificity subunit (NCBI)	N/A	N/A	InterPro entry is IPR000055 and there is an EC # associated with that entry: 3.1.21.4.	CRISPR and Other Defense	
Spy0998		-1.164	0.005854	-1.398	0.0007952	phage protein (NCBI)	N/A	N/A		Phage	

		Delta v WT		Delta v Revertant		Microbes Online	COG				
5005 Gene No.	**Gene Name**	**deseq_ logfc**	**deseq_ adjp**	**deseq_ logfc**	**deseq_ adjp**	**Desc/Link**	**COG No.**	**COG Cat.**	**Notes**	**My Category**	**Refs.**
Spy1078		-1.805	2.575E-05	-1.558	0.0003908	hypothetical protein (NCBI)	N/A	N/A	In operon with 1079 - PTS Cellobiose	Unknown	
Spy1079		-2.074	8.181E-13	-1.966	1.453E-11	PTS system, cellobiose-specific IIC component (NCBI)	COG1455 (CelB)	G	EC # likely would be transferred to 2.7.1.205 for cellobiose phosphotransferase.	Carb. Transport & Metabolism	
Spy1080		-1.851	1.855E-06	-1.443	0.0003116	hypothetical protein (NCBI)	N/A	N/A	InterPro: IPR021701 Domain of Unknown Function (DUF3284). Contains a Bet v1-like domain which bind to hydrophobic ligands.	Unknown	
Spy1081		-1.554	0.0002655	-1.803	1.907E-05	PTS system, cellobiose-specific IIA component (NCBI)	COG1447 (CelC)	G	EC Number is obsolete. Now there is a EC # for cellobiose phosphotransferase: 2.7.1.205.	Carb. Transport & Metabolism	
Spy1082		-2.47	1.784E-10	-2.208	2.43E-08	PTS system, cellobiose-specific IIB component (NCBI)	COG1440 (CelA)	G	EC Number is obsolete. Now there is a EC # for cellobiose phosphotransferase: 2.7.1.205.	Carb. Transport & Metabolism	
Spy1083		-1.789	1.773E-10	-1.835	6.084E-11	PTS system, mannitol (cryptic)-specific IIA component (NCBI)	COG3711 (BglG)	K	EC Number is obsolete. This is a putative transcriptional regulator	Regulator	
Spy1084		-1.94	2.623E-13	-1.841	5.268E-12	outer surface protein (NCBI)	COG3589	S	Peptidyl-prolyl isomerases (PPIs) in *S. aureus* and *S. pneumoniae* aid in biofilm formation and colonization respectively. Generally observed that some play a role in virulence.	Unknown	(233-235)

		Delta v WT		Delta v Revertant		Microbes Online	COG				
5005 Gene No.	**Gene Name**	**deseq_ logfc**	**deseq_ adjp**	**deseq_ logfc**	**deseq_ adjp**	**Desc/Link**	**COG No.**	**COG Cat.**	**Notes**	**My Category**	**Refs.**
Spy1085	*bglA.2*	-2.095	4.638E-11	-2.012	3.205E-10	beta-glucosidase (NCBI)	COG2723 (BglB)	G		Carb. Transport & Metabolism	
Spy1109	*inlA*	-1.614	2.535E-10	-1.309	4.472E-07	internalin protein (NCBI)	COG4886	S	Shares similarity to *L. monocytogenes* internalin family. Mutant lacking this gene was less virulent in I.P. infection and more susceptible to phagocytosis.	Virulence	(236)
Spy1137		-2.169	3.013E-12	-2.167	3.196E-12	putative competence protein/transcription factor (NCBI)	COG4469 (CoiA)	R	InterPro: IPR010330/IPR021176: Competence-induced protein CoiA.	Conjugation & Competence	
Spy1138		-1.094	0.000018	-1.141	7.723E-06	ribosomal small subunit pseudouridine synthase A (NCBI)	COG1187 (RsuA)	J	RsuA modifies the 16S rRNA. Shown to contribute to bacterial survival under stress conditions. Also considered catagorizing as Stress Tolerance, but decided to be consistent with COG.	Ribosomal & tRNA	(237)
Spy1139	*nagB*	-1.458	6.982E-07	-1.44	1.038E-06	glucosamine-6-phosphate isomerase (NCBI)	COG363 (NagB)	G	Pathways: ec00520: Amino sugar and nucleotide sugar metabolism; ec01100: Metabolic pathways	Carb. Transport & Metabolism	
Spy1147	*comEC*	-1.615	1.275E-07	-1.586	2.371E-07	COME operon protein 1 (NCBI)	COG1555 (ComEA)	L		Conjugation & Competence	(238)
Spy1171		-1.651	8.238E-05	-1.687	5.973E-05	phage-associated cell wall hydrolase (NCBI)	COG1705 (FlgJ)	N | U		Phage	
Spy1174		-1.784	0.00229	-1.88	0.0013	phage protein (NCBI)	N/A	N/A	IPR009796: Protein of unknown function DUF1366	Phage	
Spy1194		-1.009	0.001016	-1.316	1.321E-05	phage protein (NCBI)	N/A	N/A	IPR021145: Bacteriophage portal protein, SPP1 Gp6-like	Phage	

		Delta v WT		Delta v Revertant		Microbes Online	COG				
5005 Gene No.	**Gene Name**	**deseq_ logfc**	**deseq_ adjp**	**deseq_ logfc**	**deseq_ adjp**	**Desc/Link**	**COG No.**	**COG Cat.**	**Notes**	**My Category**	**Refs.**
Spy1237	*artP*	-2.597	6.844E-23	-2.484	6.262E-21	arginine transport ATP-binding protein (NCBI)	COG1126 (GlnQ)	E	Obsolete EC #. Transferred to 7.4.2.1. KEGG Comment: The enzyme, found in bacteria, interacts with an extracytoplasmic substrate binding protein and mediates the import of polar amino acids.	Amino Acid Transport & Metabolism	
Spy1238		-2.856	2.29E-27	-2.793	3.61E-26	arginine transport system permease protein (NCBI)	COG765 (HisM)	E		Amino Acid Transport & Metabolism	
Spy1256		-1.177	2.617E-28	-1.017	3.321E-21	rhodanese-related sulfurtransferases (NCBI)	COG607 (PspE)	P		Metal/Ion/ pH Homeostasis	
Spy1257	*glcK*	-1.109	9.92E-44	-1.106	1.798E-43	glucokinase/Xylose repressor (NCBI)	COG1940 (NagC)	K | G	Listed in eBook as the glucokinase involved in glycolysis	Carb. Transport & Metabolism	(28)
Spy1258		-1.8	1.436E-17	-1.528	9.282E-13	hypothetical cytosolic protein (NCBI)	COG4483	S	Homolog in *E. faecalis* V583 EF2789 (60% Identity, 86% coverage) contributes to fitness in the presence of antibiotics. Maybe could be Stress Tolerance, but *S. pyogenes* not known for its antibiotic resistance.	Unknown	(239)
Spy1282	*msrA*	-1.361	0.0004618	-1.147	0.0038	peptide methionine sulfoxide reductase (NCBI)	COG225 (MsrA)	O	KEGG Comments: Enzyme plays a role in preventing oxidative-stress damage caused by ROS by reducing the oxidized form of Met back to Met. Homolog is *S. gordonii* SGO_1176 (69% Identity, 99% Coverage) is needed for oxidative stress tolerance, biofilm formation, and oral colonization.	Stress Tolerance	(240)

		Delta v WT		Delta v Revertant		Microbes Online	COG				
5005 Gene No.	**Gene Name**	**deseq_ logfc**	**deseq_ adjp**	**deseq_ logfc**	**deseq_ adjp**	**Desc/Link**	**COG No.**	**COG Cat.**	**Notes**	**My Category**	**Refs.**
Spy1283	*tlpA*	-1.752	0.0003026	-1.798	0.0002179	thiol:disulfide interchange protein (NCBI)	COG1225 (Bcp)	O	Homolog is *S. gordonii* SGO_1177 (63% identity, 80% coverage) also involved in oxidative stress resistance. Homolog in *S. pneumo* (R6: spr0576, 55% identity, 99% coverage || D39: spd_0572, 54% identity, 99% coverage) invovled in oxidative stress resistance.	Stress Tolerance	(240-242)
Spy1284	*ccdA*	-1.533	0.01761	-1.378	0.03654	cytochrome c-type biogenesis protein (NCBI)	COG785 (CcdA)	O	Homolog in *S. pnuemo* Spd_0885 (60% identity, 98% coverage) is the redox partner of Etrx2 (Spd_0886) and plays a role in oxidative stress resistance.	Stress Tolerance	(28, 241)
Spy1287	*cas4*	-1.318	1.027E-09	-1.41	6.167E-11	hypothetical protein (NCBI)	COG1468	L	IPR013343: CRISPR-associated protein Cas4	CRISPR and Other Defense	
Spy1288	*cds2*	-1.281	2.112E-11	-1.237	1.301E-10	hypothetical cytosolic protein (NCBI)	COG3649	L	IPR006482: CRISPR-associated protein TM1801; IPR013418: CRISPR-associated protein Csd2	CRISPR and Other Defense	
Spy1289	*cds1*	-1.489	4.085E-12	-1.479	6.525E-12	hypothetical protein (NCBI)	N/A	N/A	IPR010144: CRISPR-associated protein, Csd1-type	CRISPR and Other Defense	
Spy1290	*cas5*	-1.599	3.453E-06	-1.542	8.861E-06	hypothetical protein (NCBI)	N/A	N/A	IPR010155: CRISPR-associated protein, Cas5d-type; IPR013422: CRISPR-associated protein Cas5, N-terminal; IPR021124: CRISPR-associated protein, Cas5	CRISPR and Other Defense	
Spy1291	*cas3*	-1.969	3.111E-11	-1.752	4.791E-09	ATP-dependent RNA helicase (NCBI)	COG1203	R	IPR006474: Helicase Cas3, CRISPR-associated, core	CRISPR and Other Defense	

		Delta v WT		Delta v Revertant		Microbes Online	COG				
5005 Gene No.	**Gene Name**	**deseq_ logfc**	**deseq_ adjp**	**deseq_ logfc**	**deseq_ adjp**	**Desc/Link**	**COG No.**	**COG Cat.**	**Notes**	**My Category**	**Refs.**
Spy1308	*ngtS*	-1.415	1.795E-08	-1.06	3.859E-05	sugar-binding protein (NCBI)	COG1653 (UgpB)	G		Carb. Transport & Metabolism	
Spy1309	*ngtP2*	-1.272	0.0001829	-1.073	0.001911	sugar transport system permease protein (NCBI)	COG395 (UgpE)	G		Carb. Transport & Metabolism	
Spy1310	*ngtP1*	-1.542	9.797E-07	-1.52	1.592E-06	sugar transport system permease protein (NCBI)	COG4209 (LplB)	G		Carb. Transport & Metabolism	
Spy1311	*glk*	-1.728	3.356E-05	-1.646	8.965E-05	glucokinase (NCBI)	COG1940 (NagC)	K | G		Carb. Transport & Metabolism	
Spy1374		-1.063	7.875E-14	-1.001	2.335E-12	hypothetical protein (NCBI)	N/A	N/A	IPR015046: Enterocin A Immunity	Other	
Spy1396	*lacC.1*	-1.61	0.002323	-1.303	0.01597	tagatose-6-phosphate kinase (NCBI)	N/A	N/A		Unknown	(28)
Spy1397	*lacB.1*	-1.694	0.004754	-1.316	0.03292	galactose-6-phosphate isomerase lacB subunit (NCBI)	COG698 (RpiB)	G		Carb. Transport & Metabolism	
Spy1401		-1.731	0.002224	-1.269	0.03012	PTS system, galactose-specific IIA component (NCBI)	COG1762 (PtsN)	G | T	EC # for obsolete enzyme. Transferred to 2.7.1.204 - protein-Npi-phosphohistidine--D-galactose phosphotransferase.	Carb. Transport & Metabolism	
Spy1403		-2.483	0.001435	-2.431	0.00182	copper chaperone (NCBI)	N/A	N/A	Role in copper homeostasis (if any) is unknown.	Unknown	(243)
Spy1405	*copA*	-2.065	6.381E-10	-1.13	0.001117	copper-exporting ATPase (NCBI)	COG2217 (ZntA)	P		Metal/Ion/ pH Homeostasis	

		Delta v WT		Delta v Revertant		Microbes Online	COG				
5005 Gene No.	**Gene Name**	**deseq_ logfc**	**deseq_ adjp**	**deseq_ logfc**	**deseq_ adjp**	**Desc/Link**	**COG No.**	**COG Cat.**	**Notes**	**My Category**	**Refs.**
Spy1406	*copY*	-3.337	1.142E-13	-2.566	1.698E-08	copAB ATPases metal-fist type repressor (NCBI)	COG3682	K	Could also be categorized as Metal/Ion/pH Homeostasis	Regulator	
Spy1407	*sse*	-1.272	4.28E-07	-1.382	3.64E-08	esterase (NCBI)	COG657 (Aes)	I	Antibodies against Sse and immunization with Sse protect mice from subcutaneous infection and tissue damage. Authors suggest Sse plays a role in tissue invasion. Also, Sse hydrolyzes platelet activating factor contributing to immune evasion.	Virulence	(244, 245)
Spy1477		-1.123	7.537E-05	-1.262	8.023E-06	guanine-hypoxanthine permease (NCBI)	COG2252	R	IPR006043: Xanthine/uracil/vitamin C permease. KEGG Genome Browser entry lists KO6901 Adenine/Guanine/Hypoxanthine permease.	Other Transport & Metabolism	
Spy1511		-1.631	8.512E-13	-1.344	5.665E-09	pyrazinamidase/nicotinamidase (NCBI)	COG1335 (PncA)	Q	Homolog in *S. pneumo* (SP_1583, 67% identity, 77% coverage) proposed to be involved in NAD salvage pathway.	Other Transport & Metabolism	(246)
Spy1581		-1.177	7.017E-05	-1.133	0.0001407	transcriptional regulator, MerR family (NCBI)	COG789 (SoxR)	K		Regulator	
Spy1582	*dnaQ*	-1.424	3.606E-11	-1.179	6.328E-08	DNA polymerase III, epsilon chain (NCBI)	COG847 (DnaQ)	L		Replication and Genome Integrity	
Spy1604		-1.778	1.671E-06	-1.487	9.705E-05	hypothetical protein (NCBI)	N/A	N/A	Very small peptides gene product. Several paralogs in the 5005 genome. Spy_1832, Spy_1766, Spy_1535, Spy_0259, Spy_0189. C-terminal regions of the listed proteins match with 100% identity to the N-terminal region of Spy_1604.	Unknown	

		Delta v WT		Delta v Revertant		Microbes Online	COG				
5005 Gene No.	**Gene Name**	**deseq_ logfc**	**deseq_ adjp**	**deseq_ logfc**	**deseq_ adjp**	**Desc/Link**	**COG No.**	**COG Cat.**	**Notes**	**My Category**	**Refs.**
Spy1610	*pyrG*	-1.65	4.399E-27	-1.564	2.161E-24	CTP synthase (NCBI)	N/A	N/A		Nucleotide Transport & Metabolism	
Spy1649		-1.005	3.001E-08	-1.115	6.82E-10	hypothetical membrane spanning protein (NCBI)	N/A	N/A		Unknown	
Spy1661		-1.832	3.677E-07	-1.818	5.179E-07	transaldolase (NCBI)	COG176 (MipB)	G		Carb. Transport & Metabolism	
Spy1662		-2.156	4.921E-09	-1.609	2.292E-05	putative transport protein (NCBI)	COG3037 (SgaT)	S	IPR004703: Phosphotransferase system, sugar-specific permease component	Carb. Transport & Metabolism	
Spy1663		-1.604	0.02329	-1.669	0.01852	PTS system, IIB component (NCBI)	COG3414 (SgaB)	G	EC# Obsolete. Transferred to 2.7.1.194	Carb. Transport & Metabolism	
Spy1664		-1.88	4.97E-13	-1.494	1.906E-08	PTS system, mannitol (cryptic)-specific IIA component (NCBI)	COG3711 (BglG)	K		Carb. Transport & Metabolism	
Spy1668		-6.019	2.344E-06	-7.011	2.706E-08	putative transcriptional regulator (NCBI)	N/A	N/A	Adjacent to mannitol PTS	Regulator	
Spy1702	*smeZ*	-1.13	4.698E-10	-1.192	4.94E-11	mitogenic exotoxin Z (NCBI)	N/A	N/A	Streptococcal superantigen. May play a role in invasive disease. IPR008992: Enterotoxin, bacterial; IPR013307: Superantigen, bacterial; IPR016091: Superantigen toxin, C-terminal, Staphylococcal/Streptococcal	Virulence	(247)

		Delta v WT		Delta v Revertant		Microbes Online	COG				
5005 Gene No.	**Gene Name**	**deseq_ logfc**	**deseq_ adjp**	**deseq_ logfc**	**deseq_ adjp**	**Desc/Link**	**COG No.**	**COG Cat.**	**Notes**	**My Category**	**Refs.**
Spy1703		-2.577	0.001704	-3.12	0.0001104	hypothetical cytosolic protein (NCBI)	N/A	N/A		Unknown	
Spy1705	*dppB*	-1.235	2.704E-09	-1.345	8.076E-11	dipeptide transport system permease protein (NCBI)	COG601 (DppB)	E | P		Amino Acid Transport & Metabolism	(28)
Spy1706	*dppC*	-1.092	1.698E-10	-1.177	5.298E-12	dipeptide transport system permease protein (NCBI)	COG1173 (DppC)	E | P		Amino Acid Transport & Metabolism	(28)
Spy1707	*dppD*	-1.363	2.429E-10	-1.211	2.545E-08	dipeptide transport ATP-binding protein (NCBI)	COG444 (DppD)	E | P		Amino Acid Transport & Metabolism	(28)
Spy1739		-2.398	3.209E-05	-2.727	1.88E-06	hypothetical protein (NCBI)	N/A	N/A		Unknown	
Spy1757		-1.439	5.877E-11	-1.289	5.665E-09	hypothetical protein (NCBI)	N/A	N/A	Domains & Families tab indicates a predicted membrane protein. Associated with lipoprotein-rich membrane vesicles.	Unknown	(248)
Spy1759		-2.294	1.117E-11	-2.404	1.092E-12	transcriptional regulator, MutR family (NCBI)	N/A	N/A		Regulator	
Spy1760		-2.222	3.124E-10	-2.344	2.949E-11	transcriptional regulator, MutR family (NCBI)	N/A	N/A		Regulator	
Spy1770	*hutI*	-1.452	0.0005701	-1.616	0.0001225	imidazolonepropionase (NCBI)	COG1228 (HutI)	Q	ec00340 Histidine metabolism	Amino Acid Transport & Metabolism	
Spy1771	*hutU*	-2.571	1.023E-10	-2.634	3.629E-11	urocanate hydratase (NCBI)	COG2987 (HutU)	E	ec00340 Histidine metabolism	Amino Acid Transport & Metabolism	

		Delta v WT		Delta v Revertant		Microbes Online	COG				
5005 Gene No.	**Gene Name**	**deseq_ logfc**	**deseq_ adjp**	**deseq_ logfc**	**deseq_ adjp**	**Desc/Link**	**COG No.**	**COG Cat.**	**Notes**	**My Category**	**Refs.**
Spy1772		-2.858	9.906E-05	-2.649	0.0003657	glutamate formiminotransferase (NCBI)	COG3643	E	ec00340 Histidine metabolism. In KEGG genome browser entry, also associated with 6.3.3.2: 5-formyltetrahydrofolate cyclo-ligase	Amino Acid Transport & Metabolism	
Spy1773		-3.083	0.0001248	-3.007	0.0002049	formiminotetrahydrofolate cyclodeaminase (NCBI)	COG3404	E	KEGG genome browser entry lists EC 3.5.4.9: methenyltetrahydrofolate cyclohydrolase. ec00670 One carbon pool by folate; ec01100 Metabolic pathways.	Other Transport & Metabolism	
Spy1774	*fhs.2*	-2.294	5.09E-12	-1.734	0.0000004	formate--tetrahydrofolate ligase (NCBI)	COG2759 (MIS1)	F	ec00670: One carbon pool by folate; ec00720: Carbon fixation pathways in prokaryotes; ec01100: Metabolic pathways; ec01120: Microbial metabolism in diverse environments	Nucleotide Transport & Metabolism	
Spy1775		-2.632	7.696E-07	-2.694	4.218E-07	hypothetical cytosolic protein (NCBI)	COG3758	S		Unknown	
Spy1776		-2.354	2.747E-06	-2.192	1.482E-05	amino acid permease (NCBI)	COG531 (PotE)	E		Amino Acid Transport & Metabolism	
Spy1777	*hutH*	-2.683	7.758E-12	-2.606	3.586E-11	histidine ammonia-lyase (NCBI)	COG2986 (HutH)	E	ec00340 Histidine metabolism	Amino Acid Transport & Metabolism	
Spy1779		-1.453	3.957E-11	-1.757	7.344E-16	transcriptional regulator, LuxR family (NCBI)	COG3899	R	Homolog in *S. sanguinis* SK36 (SSA_0427, 55% identity, 100% coverage) was identified as a biofilm related transcription factor.	Regulator	(249)
Spy1783	*dexS/ treC*	-2.955	1.628E-57	-2.677	3.454E-47	trehalose-6-phosphate hydrolase (NCBI)	COG366 (AmyA)	G		Carb. Transport & Metabolism	

		Delta v WT		Delta v Revertant		Microbes Online	COG				
5005 Gene No.	**Gene Name**	**deseq_ logfc**	**deseq_ adjp**	**deseq_ logfc**	**deseq_ adjp**	**Desc/Link**	**COG No.**	**COG Cat.**	**Notes**	**My Category**	**Refs.**
Spy1784	*treB*	-2.565	2.506E-20	-2.59	1.096E-20	PTS system, trehalose-specific IIBC component (NCBI)	COG1263 (PtsG)	G	EC # Obsolete. Transferred to 2.7.1.201	Carb. Transport & Metabolism	
Spy1786		-1.466	0.0004538	-1.107	0.009878	transcriptional regulator, MarR family (NCBI)	N/A	N/A		Regulator	
Spy1818	*cadC*	-1.175	2.406E-05	-1.287	3.601E-06	cadmium efflux system accessory protein (NCBI)	COG640 (ArsR)	K		Regulator	
Spy1820		-1.161	3.942E-08	-1.405	2.09E-11	ftsK/SpoIIIE family (NCBI)	N/A	N/A		Replication and Genome Integrity	
SpyT0006		-2.246	0.00898	-2.044	0.01945	tRNA-Thr (NCBI)	N/A	N/A		Translation	
SpyT0048		-1.02	0.001141	-1.026	0.001117	tRNA-Arg (NCBI)	N/A	N/A		Translation	
SpyT0064		-4.193	2.336E-07	-2.506	0.003271	tRNA-Cys (NCBI)	N/A	N/A		Translation	
SpyT0065		-1.516	0.001435	-1.678	0.0003908	tRNA-Asn (NCBI)	N/A	N/A		Translation	
SpyT0067		-2.997	2.86E-16	-2.958	6.455E-16	tRNA-Arg (NCBI)	N/A	N/A		Translation	

Significantly Up Regulated Genes in Exponential

		Delta v WT		Delta v Revertant		Microbes Online	COG				
5005 Gene No.	**Gene Name**	**deseq_ logfc**	**deseq_ adjp**	**deseq_ logfc**	**deseq_ adjp**	**Desc/Link**	**COG No.**	**COG Cat.**	**Notes**	**My Category**	**Refs.**
Spy0039	*adh2*	1.175	0.001582	1.046	0.005391	alcohol dehydrog-enase/ acetaldehyde dehydrogenase (acetylating) (NCBI)	COG1012 (PutA)	C	Multiple pathways in KEGG, most involved with metabolism	Other Transport & Metabolism	
Spy0098		1.377	1E-12	1.372	1.29E-12	hypothetical protein (NCBI)	N/A	N/A		Unknown	
Spy0313		1.981	5.94E-35	2.01	5.56E-36	riboflavin transporter (NCBI)	COG3601	S		Other Transport & Metabolism	
Spy0314		1.701	2.82E-47	1.629	1.7E-43	phosphatidyl-glycerophosphatase B homolog (NCBI)	COG671 (PgpB)	I		Other Transport & Metabolism	
Spy0346		1.278	0.000149	1.598	2.39E-06	hypothetical protein (NCBI)	N/A	N/A		Unknown	
Spy0347	*nrdF.1*	1.301	3.65E-05	2.025	5.97E-11	ribonucleoside-diphosphate reductase beta chain (NCBI)	COG208 (NrdF)	F		Nucleotide Transport & Metabolism	
Spy0348	*nrdI*	1.185	0.00034	2.285	1.88E-12	hypothetical protein (NCBI)	COG1780 (NrdI)	F	IPR020852: Ribonucleotide reductase, class Ib, NrdI, bacterial	Nucleotide Transport & Metabolism	
Spy0349	*nrdE.1*	1.1	0.000564	2.155	3.12E-12	ribonucleoside-diphosphate reductase alpha chain (NCBI)	COG209 (NrdA)	F		Nucleotide Transport & Metabolism	
Spy0417	*pcp*	1.12	5.97E-13	1.123	5.21E-13	pyrrolidone-carboxylate peptidase (NCBI)	COG2039 (Pcp)	O	Protien Turnover could probably also be umbrelled in Amino Acid Metabolism	Amino Acid Transport & Metabolism	

		Delta v WT		Delta v Revertant		Microbes Online	COG				
5005 Gene No.	**Gene Name**	**deseq_ logfc**	**deseq_ adjp**	**deseq_ logfc**	**deseq_ adjp**	**Desc/Link**	**COG No.**	**COG Cat.**	**Notes**	**My Category**	**Refs.**
Spy0421	*gloA*	1.057	8.5E-06	1.158	1.01E-06	lactoylglutathione lyase (NCBI)	COG346 (GloA)	E	COG lists as amino acid metabolism, KEGG orthology places it in carbohydrate metabolism. gloA important for tolerance to methylglyoxal, survival in neutrophils, and survival in whole blood. Stress Tolerance is the category determined, however, could also be virulence.	Stress Tolerance	(250)
Spy0422		1.077	0.000196	1.351	2.48E-06	NAD(P)H-dependent quinone reductase (NCBI)	COG778 (NfnB)	C		Unknown	
Spy0423	*pepQ*	1.151	0.000163	1.615	7.36E-08	Xaa-Pro dipeptidase (NCBI)	COG6 (PepP)	E		Amino Acid Transport & Metabolism	
Spy0477		1.14	1.45E-31	1.528	2.78E-56	hypothetical membrane spanning protein (NCBI)	COG3689	S	Identified as critical for subcutaneous soft-tissue infection.	Virulence	(115)
Spy0478		1.059	8.94E-07	1.597	4.51E-14	hypothetical membrane spanning protein (NCBI)	COG701	R	Identified as critical for subcutaneous soft-tissue infection.	Virulence	(115)
Spy0513		1.359	4.22E-23	1.149	1.02E-16	hypothetical cytosolic protein (NCBI)	N/A	N/A		Unknown	
Spy0514		1.169	2.09E-10	1.231	2.13E-11	dGTP triphospho-hydrolase (NCBI)	COG1078	R		Unknown	
Spy0653	*czcD*	2.564	4.55E-33	1.91	9.23E-19	cobalt-zinc-cadmium resistance protein (NCBI)	COG1230 (CzcD)	P	IPR002524: Cation efflux protein	Metal/Ion/ pH Homeostasis	

		Delta v WT		Delta v Revertant		Microbes Online	COG				
5005 Gene No.	**Gene Name**	**deseq_ logfc**	**deseq_ adjp**	**deseq_ logfc**	**deseq_ adjp**	**Desc/Link**	**COG No.**	**COG Cat.**	**Notes**	**My Category**	**Refs.**
Spy0922	*pdxU*	2.547	4.63E-45	2.177	3.43E-33	hypothetical membrane spanning protein (NCBI)	COG4720	S	Homologous to *S. mutans* SMU_955 (56% identity, 92% coverage) which is a putative pyridoxal transporter. Mislabeled in reference as SMU_865, but oligos used by authors align to SMU_955 in UA159 genome.	Other Transport & Metabolism	(251)
Spy0923	*pdxK*	2.082	9.77E-47	1.7	2.31E-31	pyridoxine kinase (NCBI)	COG2240 (PdxK)	H	KEGG Pathway ec00750: Vitamin B6 metabolism. Homologous to *S. mutans* SMU_954 (44% identity, 72% coverage) which is a putative pyridoxal kinase. Mislabeled in reference as SMU_866, but oligos used by authors align to SMU_954 in UA159 genome.	Other Transport & Metabolism	(251)
Spy1114		1.472	1.74E-14	1.015	1.95E-07	hypothetical membrane spanning protein (NCBI)	COG2339 (prsW)	S	Deletion of homolog is *S. pneumo* (spd_031, 31% identity, 76% coverage) resulted in attenuated virulence in bacteremia and pneumonia models.	Virulence	(252)
Spy1122	*nrdH*	1.615	1.49E-13	1.121	4.63E-07	glutaredoxin (NCBI)	COG695 (GrxC)	O	Listed as a gene involved in oxidative stress resistance in S. pyogenes that is differentially regulated by Ihk-Irr	Stress Tolerance	(253)
Spy1259	*dpr*	1.09	6.88E-06	1.003	3.92E-05	non-specific DNA-binding protein/iron-binding ferritin-like antioxidant protein (NCBI)	COG783 (Dps)	P	KEGG Genome entry listed as nuceloid associated protein.	Stress Tolerance	
Spy1320	*recX*	1.331	2.96E-09	1.081	2.11E-06	regulatory protein (NCBI)	COG2137 (OraA)	R	Could go with Regulator OR Replication and Chromosome Integrity. I have classified all other regulators though as such instead of with the impacts of their gene targets.	Regulator	

		Delta v WT		Delta v Revertant		Microbes Online	COG				
5005 Gene No.	Gene Name	deseq_ logfc	deseq_ adjp	deseq_ logfc	deseq_ adjp	Desc/Link	COG No.	COG Cat.	Notes	My Category	Refs.
Spy1417		1.083	1.4E-09	1.454	3.4E-16	phage protein (NCBI)	N/A	N/A		Phage	
Spy1418		1.32	2.44E-11	1.344	1.24E-11	phage protein (NCBI)	N/A	N/A		Phage	
Spy1419		1.112	8.69E-10	1.374	2.59E-14	phage protein (NCBI)	N/A	N/A	IPR009796: Protein of unknown function DUF1366	Phage	
Spy1421		1.274	5.18E-14	1.399	9.79E-17	phage infection protein (NCBI)	N/A	N/A	IPR012892: Gp58-like	Phage	
Spy1422		1.29	3.85E-12	1.309	1.92E-12	phage protein (NCBI)	N/A	N/A	IPR013320: Concanavalin A-like lectin/glucanase, subgroup	Phage	
Spy1423		1.194	1.86E-11	1.44	3.61E-16	hyaluronoglucos-aminidase (NCBI)	N/A	N/A		Phage	
Spy1424		1.232	1.45E-10	1.616	1.99E-17	phage endopeptidase (NCBI)	N/A	N/A	IPR010572: Domain of unknown function DUF1142	Phage	
Spy1425		1.18	1.17E-10	1.4	1.52E-14	phage protein (NCBI)	N/A	N/A	IPR008841: Siphovirus tail component	Phage	
Spy1426		1.352	1.95E-12	1.261	5.97E-11	phage protein (NCBI)	COG5412	S	IPR011989: Armadillo-like helical; IPR013491: Tape measure domain; IPR016024: Armadillo-type fold	Phage	
Spy1427		1.125	6.93E-07	1.412	3.44E-10	phage protein (NCBI)	N/A	N/A		Phage	
Spy1428		1.059	1.39E-06	1.797	7.2E-17	phage protein (NCBI)	N/A	N/A		Phage	
Spy1429		1.25	6.14E-12	1.354	8.38E-14	phage protein (NCBI)	N/A	N/A	IPR011855: Phage major tail protein TP901-1	Phage	
Spy1430		1.293	8.01E-16	1.454	9.76E-20	phage protein (NCBI)	N/A	N/A		Phage	
Spy1431		1.164	1.36E-10	1.362	4.91E-14	phage protein (NCBI)	N/A	N/A	IPR006967: Protein of unknown function DUF646, phage head/tail component; IPR010064: Phage protein, HK97, gp10	Phage	
Spy1432		1.267	6.42E-12	1.393	3.83E-14	phage protein (NCBI)	N/A	N/A		Phage	

		Delta v WT		Delta v Revertant		Microbes Online	COG				
5005 Gene No.	**Gene Name**	**deseq_ logfc**	**deseq_ adjp**	**deseq_ logfc**	**deseq_ adjp**	**Desc/Link**	**COG No.**	**COG Cat.**	**Notes**	**My Category**	**Refs.**
Spy1433		1.301	4.82E-14	1.591	1.58E-20	phage protein (NCBI)	N/A	N/A	IPR021146: Bacteriophage QLRG family, putative DNA packaging	Phage	
Spy1434		1.281	5.38E-13	1.425	7.46E-16	phage protein (NCBI)	N/A	N/A		Phage	
Spy1436		1.057	2.42E-05	1.219	1.06E-06	phage protein (NCBI)	N/A	N/A		Phage	
Spy1527		1.052	3.94E-05	1.171	4.92E-06	ferrichrome transport system permease protein (NCBI)	COG609 (FepD)	P		Metal/Ion/ pH Homeostasis	(254, 255)
Spy1529	*shp*	1.035	1.57E-05	1.196	5.83E-07	putative heme binding protein (NCBI)	N/A	N/A	IPR020985: Cell surface protein Shp, haem-binding domain	Metal/Ion/ pH Homeostasis	(256)
Spy1530		1.496	1.89E-29	1.37	8.58E-25	putative Fe3+-siderophore transport protein (NCBI)	COG4886	S		Metal/Ion/ pH Homeostasis	(256)
Spy1540	*endoS*	1.239	1.51E-14	1.447	1.41E-19	endo-beta-N-acetylglucos-aminidase F2 precursor (NCBI)	N/A	N/A		Virulence	(257)
Spy1541		1.282	9.58E-06	1.692	4.82E-09	hypothetical protein (NCBI)	N/A	N/A		Unknown	
Spy1542	*scrA*	1.279	9.87E-08	1.584	2.78E-11	PTS system, sucrose-specific IIABC component (NCBI)	COG1263 (PtsG)	G	EC# Obsolete. Transferred to	Carb. Transport & Metabolism	
Spy1628		2.354	2.62E-28	1.618	1.31E-14	ABC transporter ATP-binding protein (NCBI)	COG1136 (SalX)	V	Believed to be responsible for the export of SalA. But M1 GAS does not produce active form of this.	Unknown	(258)
Spy1629	*salX*	2.483	5.73E-25	2.303	1.49E-21	lantibiotic transport ATP-binding protein (NCBI)	N/A	N/A	Believed to be responsible for the export of SalA. But M1 GAS does not produce active form of this.	Unknown	(258)

		Delta v WT		Delta v Revertant		Microbes Online	COG				
5005 Gene No.	**Gene Name**	**deseq_ logfc**	**deseq_ adjp**	**deseq_ logfc**	**deseq_ adjp**	**Desc/Link**	**COG No.**	**COG Cat.**	**Notes**	**My Category**	**Refs.**
Spy1630	*salB*	2.875	7.32E-37	2.348	1.47E-25	serine (threonine) dehydratase (NCBI)	N/A	N/A	IPR017146: Lantibiotic modifying enzyme, RumM. Truncated in *S. pyogenes* SF370.	Unknown	(258)
Spy1631	*salA*	2.346	1.35E-20	2.947	1.69E-29	lantibiotic salivaricin A (NCBI)	N/A	N/A	Only M4 GAS produce an active Salavaricin A peptide.	Unknown	(258)
Spy1632	*lacG*	3.779	1.3E-41	3.297	6.79E-32	6-phospho-beta-galactosidase (RefSeq)	COG2723 (BglB)	G		Carb. Transport & Metabolism	
Spy1633	*lacE*	4.026	1.43E-23	3.596	6.01E-19	PTS system, lactose-specific IIBC component (NCBI)	COG1455 (CelB)	G	EC# obsolete. Transferred to 2.7.1.207	Carb. Transport & Metabolism	
Spy1634	*lacF*	3.888	4.03E-27	3.654	4.18E-24	PTS system, lactose-specific IIA component (NCBI)	COG1447 (CelC)	G	EC# obsolete. Transferred to 2.7.1.207	Carb. Transport & Metabolism	
Spy1635	*lacD.2*	3.575	2.33E-26	3.591	1.81E-26	tagatose-bisphosphate aldolase (NCBI)	COG3684 (LacD)	G		Carb. Transport & Metabolism	
Spy1636	*lacC.2*	3.404	6.86E-28	3.27	9.05E-26	tagatose-6-phosphate kinase (NCBI)	COG1105 (FruK)	G		Carb. Transport & Metabolism	
Spy1637	*lacB.2*	3.233	6.42E-28	3.064	3.61E-25	galactose-6-phosphate isomerase lacB subunit (NCBI)	COG698 (RpiB)	G		Carb. Transport & Metabolism	
Spy1638	*lacA.2*	2.869	6.34E-23	2.944	4.7E-24	galactose-6-phosphate isomerase lacA subunit (NCBI)	COG698 (RpiB)	G		Carb. Transport & Metabolism	
Spy1716		1.37	3.73E-12	1.097	3.68E-08	transposase (NCBI)	COG3666	L		Replication and Genome Integrity	
Spy1733		3.555	2.82E-47	4.093	6.14E-59	hypothetical protein (NCBI)	N/A	N/A		Unknown	

		Delta v WT		Delta v Revertant		Microbes Online	COG				
5005 Gene No.	**Gene Name**	**deseq_ logfc**	**deseq_ adjp**	**deseq_ logfc**	**deseq_ adjp**	**Desc/Link**	**COG No.**	**COG Cat.**	**Notes**	**My Category**	**Refs.**
Spy1734	*spi*	3.682	3.7E-58	4.388	3.78E-77	streptopain fragment (NCBI)	N/A	N/A		Virulence	
Spy1735	*speB*	3.31	1.42E-56	3.891	5.1E-77	streptococcal pyrogenic exotoxin B (NCBI)	N/A	N/A		Virulence	
Spy1750		1.651	1.74E-10	1.482	1.28E-08	hypothetical protein (NCBI)	N/A	N/A		Unknown	
Spy1851	*hasA*	2.792	1.79E-24	2.425	1.17E-18	hyaluronan synthase (NCBI)	COG1215	M		Virulence	
Spy1852	*hasB*	3.219	7.61E-43	2.64	4.36E-29	UDP-glucose 6-dehydrogenase (NCBI)	COG1004 (Ugd)	M		Virulence	
Spy1853	*hasC*	3.154	1.36E-61	2.773	5.19E-48	UTP--glucose-1-phosphate uridylyltransferase (NCBI)	COG1210 (GalU)	M		Virulence	

Significantly Down Regulated Genes in Transition

		Delta v WT		Delta v Revertant		Microbes Online	COG				
5005 Gene No.	**Gene Name**	**deseq_ logfc**	**deseq_ adjp**	**deseq_ logfc2**	**deseq_ adjp2**	**Desc/Link**	**COG No.**	**COG Cat.**	**Notes**	**My Category**	**Refs.**
Spy0106	*rofA*	-3.463	2E-88	-3.547	1.6E-96	transcriptional regulator (NCBI)	N/A	N/A	Regulated pilus operon in all serotypes tested. Regulates other virulence genes as well.	Regulator	
Spy0107	*cpa*	-8.957	0	-9.048	0	fibronectin-binding protein (NCBI)	N/A	N/A	Pilus minor subunit ap1. Tip subunit. Could have also been categorized as 'Adhesion'	Virulence	(109)
Spy0108	*sipA*	-8.559	8.3E-228	-8.579	9.2E-234	signal peptidase I (NCBI)	N/A	N/A	Required chaperone for pilus assembly of FCT type 2 strains. Could have also been categorized as 'Adhesion'	Virulence	(109)
Spy0109	*fctA*	-8.321	0	-8.286	0	fibronectin-binding protein (NCBI)	N/A	N/A	Major pilus subunit. Could have also been categorized as 'Adhesion'	Virulence	(109)
Spy0110	*srtC2*	-8.471	0	-8.409	0	hypothetical protein (NCBI)	COG4509	S	Class B sortase for pilus assembly. Could have also been categorized as 'Adhesion'	Virulence	(109)
Spy0111		-7.855	0	-7.77	0	hypothetical protein (NCBI)	N/A	N/A	fctB - pilus minor subunit ap2. Base subunit. Could have also been categorized as 'Adhesion'	Virulence	(109)
Spy0114	*srtB*	-1.232	2.77E-05	-1.76	6.58E-11	sortase (NCBI)	COG3764 (SrtA)	M	Class C sortase for pilus assembly. Could have also been categorized as 'Adhesion'	Virulence	(109)
Spy1613		-1.246	7.82E-14	-1.124	1.6E-11	mechanosensitive ion channel (NCBI)	COG668 (MscS)	M	Mechanosensitve channels help bacteria deal with osmotic shock.	Stress Tolerance	
SpyT0067		-1.082	0.03122	-1.059	0.0376	tRNA-Arg (NCBI)	N/A	N/A		Translation & Protein Processing	

Significantly Up Regulated Genes in Transition

		Delta v WT		Delta v Revertant		Microbes Online	COG				
5005 Gene No.	**Gene Name**	**deseq_ logfc**	**deseq_ adjp**	**deseq_ logfc2**	**deseq_ adjp2**	**Desc/Link**	**COG No.**	**COG Cat.**	**Notes**	**My Category**	**Refs.**
Spy0995		2.282	0.00121	1.662	0.01886	phage protein (NCBI)	N/A	N/A		Phage	
Spy0998		1.488	0.003738	1.253	0.01772	phage protein (NCBI)	N/A	N/A		Phage	
Spy0999		1.842	2.45E-10	1.41	1.78E-06	phage protein (NCBI)	N/A	N/A		Phage	
Spy1000		2.044	1.36E-10	1.525	1.88E-06	phage protein (NCBI)	N/A	N/A		Phage	
Spy1001		1.531	4.23E-08	1.222	2.1E-05	phage-associated cell wall hydrolase (NCBI)	N/A	N/A	IPR003646: SH3-like domain, bacterial; IPR007921: Cysteine, histidine-dependent amidohydrolase/peptidase; IPR013667: SH3, type-5 bacterial	Phage	
Spy1002		2.241	9.1E-06	1.496	0.003862	N-acetylmuramoyl-L-alanine amidase (NCBI)	N/A	N/A	IPR010026: Phage holin, LL-H	Phage	
Spy1003		1.687	0.00361	1.768	0.001528	phage protein (NCBI)	N/A	N/A		Phage	
Spy1005		1.912	9.93E-08	1.362	0.000162	phage protein (NCBI)	N/A	N/A	IPR009796: Protein of unknown function DUF1366	Phage	
Spy1006		1.551	2.27E-15	1.128	1.48E-08	phage structural protein (NCBI)	N/A	N/A	IPR008577: Protein of unknown function DUF859	Phage	
Spy1007		1.605	1.78E-12	1.373	2.19E-09	phage protein (NCBI)	COG1705 (FlgJ)	N | U	IPR002901: Mannosyl-glycoprotein endo-beta-N-acetylglucosamidase; IPR007921: Cysteine, histidine-dependent amidohydrolase/peptidase; IPR013338: Lysozyme domain, subfamily 2	Phage	

		Delta v WT		Delta v Revertant		Microbes Online	COG				
5005 Gene No.	Gene Name	deseq_ logfc	deseq_ adjp	deseq_ logfc2	deseq_ adjp2	Desc/Link	COG No.	COG Cat.	Notes	My Category	Refs.
Spy1008		1.638	1.54E-13	1.199	1.37E-07	hypothetical protein (NCBI)	COG4722	S	IPR006520: Phage tail component, N-terminal; IPR008841: Siphovirus tail component; IPR013784: Carbohydrate-binding-like fold	Phage	
Spy1009		1.9	1.34E-21	1.417	9.21E-13	phage protein (NCBI)	N/A	N/A		Phage	
Spy1010		1.979	1.89E-11	1.344	5.63E-06	phage protein (NCBI)	N/A	N/A		Phage	
Spy1011		1.515	1.97E-05	1.692	1.78E-06	phage protein (NCBI)	N/A	N/A		Phage	
Spy1012		1.729	7.14E-07	1.327	0.000242	antigen A (NCBI)	COG5437	S	IPR011855: Phage major tail protein TP901-1	Phage	
Spy1013		2.03	1.51E-07	1.59	4.2E-05	antigen B (NCBI)	N/A	N/A		Phage	
Spy1014		1.467	4.19E-06	1.381	1.66E-05	antigen C (NCBI)	N/A	N/A		Phage	
Spy1015		1.542	0.000573	1.203	0.01169	phage protein (NCBI)	N/A	N/A		Phage	
Spy1016		2.263	8.39E-09	1.007	0.01995	phage protein (NCBI)	N/A	N/A	IPR021146: Bacteriophage QLRG family, putative DNA packaging	Phage	
Spy1017		1.637	5.04E-07	1.372	2.93E-05	phage protein (NCBI)	N/A	N/A	IPR003034: DNA-binding SAP; IPR011112: Rho termination factor, N-terminal	Phage	
Spy1018		1.664	4.3E-15	1.02	4.33E-06	phage protein (NCBI)	N/A	N/A		Phage	
Spy1021		1.565	2.76E-08	1.354	1.98E-06	phage protein (NCBI)	N/A	N/A	IPR006528: Phage putative head morphogenesis protein, SPP1 Gp7	Phage	

		Delta v WT		Delta v Revertant		Microbes Online	COG				
5005 Gene No.	Gene Name	deseq_ logfc	deseq_ adjp	deseq_ logfc2	deseq_ adjp2	Desc/Link	COG No.	COG Cat.	Notes	My Category	Refs.
Spy1022		1.459	6.14E-08	1.003	0.000471	portal protein (NCBI)	N/A	N/A	IPR006428: Bacteriophage portal protein, SPP1 family; IPR021145: Bacteriophage portal protein, SPP1 Gp6-like	Phage	
Spy1033		1.242	0.005146	1.029	0.02949	phage protein (NCBI)	N/A	N/A		Phage	
Spy1851	*hasA*	4.387	2.8E-56	4.305	2.03E-56	hyaluronan synthase (NCBI)	COG1215	M		Virulence	
Spy1852	*hasB*	4.588	4.65E-82	4.568	1.32E-84	UDP-glucose 6-dehydrogenase (NCBI)	COG1004 (Ugd)	M		Virulence	
Spy1853	*hasC*	4.739	1.4E-130	4.524	1.2E-124	UTP--glucose-1-phosphate uridylyltransferase (NCBI)	COG1210 (GalU)	M		Virulence	

Significantly Down Regulated Genes in Early Stationary

		Delta v WT		Delta v Revertant		Microbes Online	COG				
5005 Gene No.	**Gene Name**	**deseq_ logfc**	**deseq_ adjp**	**deseq_ logfc2**	**deseq_ adjp2**	**Desc/Link**	**COG No.**	**COG Cat.**	**Notes**	**My Category**	**Refs.**
SpyR0001		-4.379	0.03926	-4.539	0.02187	small subunit 16S ribosomal RNA (NCBI)	N/A	N/A		Translation & Protein Processing	
SpyR0002		-4.484	0.02525	-3.753	0.04861	large subunit 23S ribosomal RNA (NCBI)	N/A	N/A		Translation & Protein Processing	
SpyR0017		-3.671	0.01609	-5.103	0.000298	large subunit 23S ribosomal RNA (NCBI)	N/A	N/A		Translation & Protein Processing	
Spy0057	*rpsN*	-1.155	2.15E-12	-1.004	2.57E-09	SSU ribosomal protein S14P (NCBI)	COG199 (RpsN)	J		Translation & Protein Processing	
Spy0106	*rofA*	-3.277	4.59E-82	-3.38	2.19E-87	transcriptional regulator (NCBI)	N/A	N/A	Regulated pilus operon in all serotypes tested. Regulates other virulence genes as well.	Regulator	
Spy0107	*cpa*	-8.183	0	-8.422	0	fibronectin-binding protein (NCBI)	N/A	N/A	Pilus minor subunit ap1. Tip subunit. Could have also been categorized as 'Adhesion'	Virulence	(109)
Spy0108	*sipA1*	-8.183	1.1E-189	-8.655	1.9E-212	signal peptidase I (NCBI)	N/A	N/A	Required chaperone for pilus assembly of FCT type 2 strains. Could have also been categorized as 'Adhesion'	Virulence	(109)
Spy0109	*fctA*	-7.463	0	-7.657	0	fibronectin-binding protein (NCBI)	N/A	N/A	Major pilus subunit. Could have also been categorized as 'Adhesion'	Virulence	(109)
Spy0110	*srtC2*	-7.426	1.9E-264	-7.623	7.6E-279	hypothetical protein (NCBI)	COG4509	S	Class B sortase for pilus assembly. Could have also been categorized as 'Adhesion'	Virulence	(109)
Spy0111		-7.491	0	-7.681	0	hypothetical protein (NCBI)	N/A	N/A	*fctB* - pilus minor subunit ap2. Base subunit. Could have also been categorized as 'Adhesion'	Virulence	(109)

		Delta v WT		Delta v Revertant		Microbes Online	COG				
5005 Gene No.	**Gene Name**	**deseq_ logfc**	**deseq_ adjp**	**deseq_ logfc2**	**deseq_ adjp2**	**Desc/Link**	**COG No.**	**COG Cat.**	**Notes**	**My Category**	**Refs.**
Spy0123		-1.124	8.5E-06	-1.099	1.43E-05	translation initiation inhibitor (NCBI)	COG251 (TdcF)	J	Name of this gene in SF370 is SPy0145, which is listed in this paper as being homolgous to *L. lactis* AldR which "is a negative regulator of transcription for genes involved in amino acid metabolism in *L. lactis*". IPR006056: YjgF-like protein; IPR006175: Endoribonuclease L-PSP; IPR013813: Endoribonuclease L-PSP/chorismate mutase-like	Regulator	(259)
Spy1487	*fabZ*	-1.337	3.49E-10	-1.411	1.88E-11	(3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase (NCBI)	COG764 (FabA)	I	EC# Incomplete. KEGG genome browser entry lists EC: 4.2.1.59. (3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase. Ec00061: Fatty acid biosynthesis	Lipid Transport & Metabolism	
Spy1488	*accB*	-1.217	7.2E-08	-1.093	2.33E-06	biotin carboxyl carrier protein of acetyl-CoA carboxylase (NCBI)	COG511 (AccB)	I		Lipid Transport & Metabolism	
Spy1490		-1.031	0.000101	-1.146	1.09E-05	3-oxoacyl-[acyl-carrier protein] reductase (NCBI)	COG1028 (FabG)	I | Q | R	KEGG Pathways: ec00061 Fatty acid biosynthesis; ec00780 Biotin metabolism; ec01100 Metabolic pathways ||	Lipid Transport & Metabolism	
Spy1841	*sdhB*	-1.005	1.36E-06	-1.338	1.48E-11	L-serine dehydratase (NCBI)	COG1760 (SdaA)	E	KEGG Pathways: ec00260: Glycine, serine and threonine metabolism; ec00270: Cysteine and methionine metabolism	Amino Acid Transport & Metabolism	
SpyT0067		-1.139	0.01375	-1.087	0.01262	tRNA-Arg (NCBI)	N/A	N/A		Translation & Protein Processing	

Significantly Up Regulated Genes in Early Stationary

		Delta v WT		Delta v Revertant		Microbes Online	COG				
5005 Gene No.	**Gene Name**	**deseq_ logfc**	**deseq_ adjp**	**deseq_ logfc2**	**deseq_ adjp2**	**Desc/Link**	**COG No.**	**COG Cat.**	**Notes**	**My Category**	**Refs.**
Spy0117		1.016	0.00295	1.065	0.001281	transcriptional regulator, LysR family (NCBI)	N/A	N/A		Regulator	
Spy0901		1.171	0.000232	1.185	0.000149	hypothetical protein (NCBI)	N/A	N/A		Unknown	
Spy1165	*pyrD*	1.052	2.92E-08	1.087	7.51E-09	dihydroorotate dehydrogenase (NCBI)	COG167 (PyrD)	F	EC# obsolete. Transferred to 1.3.98.1. KEGG Pathways: ec00240: Pyrimidine metabolism; ec01100: Metabolic pathways	Nucleotide Transport & Metabolism	
Spy1851	*hasA*	3.414	8.42E-35	3.444	1.61E-35	hyaluronan synthase (NCBI)	COG1215	M		Virulence	
Spy1852	*hasB*	3.445	8.33E-47	3.357	9.62E-45	UDP-glucose 6-dehydrogenase (NCBI)	COG1004 (Ugd)	M		Virulence	
Spy1853	*hasC*	3.408	6.13E-68	3.18	3.92E-60	UTP--glucose-1-phosphate uridylyltransferase (NCBI)	COG1210 (GalU)	M		Virulence	

Appendix A5: Caliper LifeSciences Product Literature Page for Xen20

Caliper LifeSciences

Bioware™ Microorganism – *Streptococcus pyogenes* Xen20
In vitro Characteristics

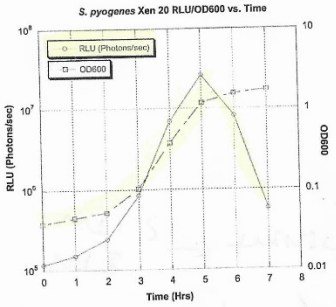
Genetic Characteristics
Streptococcus pyogenes–Xen20 was derived from the parental strain *Streptococcus pyogenes* 591. *S. pyogenes*–Xean20 possesses a stable copy of the modified *Photorhabdus luminescens lux* operon at a single integration site on the bacterial chromosome. Xen8.1 should be stored at -80°C

Growth Characteristics
S. pyogenes–Xen20 grows well in Brain-Heart Infusion (BHI), or Todd-Hewitt broth with 0.5% yeast extract (THY) at 37°C, 5%CO_2, with no aeration. It may also be grown on Trypticase Soy Agar (TSA) supplemented with 5% sheep blood.

Optional Antibiotic Selection
S. pyogenes–Xen20 may be grown selectively on LB agar containing 200 µg/mL kanamycin.
300 µg/ml worked ✓

Colonial Morphology
On TSA + 5% sheep blood plates, *S. pyogenes*–Xen20 appears as small (approx 0.5mm), circular, gray colonies with large zones of beta hemolysis.

Growth Curve
Log-phase growth can be achieved after 4 to 5 hours of subculture in THY broth at 37°C, 5% CO_2, with no aeration in culture tubes. For the above broth culture conditions, an absorbance measurement at 600nm (against a THY blank) of 0.5 is roughly equivalent to $8.0x10^7$ cfu/ml of *S. pyogenes*-Xen20.



Virulence Factors
Hemolysins: β-hemolytic on TSA + 5% sheep blood
M Protein: Serotype M49
Bacitracin: Susceptible

Biochemical Profile
A biochemical profile was obtained for *S. pyogenes*–Xen20 using the api 20 Strep system available from bioMérieux.

Sugar Utilization	
Ribose	-
Arabinose	-
Mannitol	-
Sorbitol	-
Lactose	+
Trehalose	+
Inuline	-
Raffinose	-
Starch	+
Glycogen	-

Other Tests	
Voges Proskauer	-
Hippurate	-
Esculin	-
Pyrrolidone Arylamidase	+
α-galactosidase	-
β-glucuronidase	-
β-galactosidase	-
Alkaline Phosphatase	+
Leucine Arylamidase	+
Arginine Dihydrolase	+

Antibiotic Susceptibility
Disk Diffusion Data
Disk diffusion tests were performed according to methods outlined in the NCCLS Approved Standard M2-A7.

Kirby-Bauer Disk Diffusion Test		
Sensitive	**Intermediate**	**Resistant**
Penicillin 10U	Tetracycline 30	Kanamycin 30
Vancomycin 30	Chloramphenicol 30	SXT
Erythromycin 15		

MIC Data
MIC's were determined using the macro-dilution methods specified in the NCCLS Approved Standard M7-A5.

NCCLS Macrodilution MIC/MBC	
Antibiotic	**MIC (µg/mL)**
Ceftriaxone	1.0
Ciprofloxacin	>4.0
Erythromycin	>16
Gentamicin	>32
Kanamycin	50
Penicillin G	>0.5

- THY Kan.300 plate ✓
- BAP ✓
- THY broth kan300
- THY broth kan30

References:
1. Journal of Immunology, 2003, Vol. 17-1, 2532-2537
2. Infection & Immunity, Aug. 2005, Vol. 73, No. 8, 4512-4521

Contact Information:
If you have any questions regarding these cell lines, please contact Caliper at 508-497-6592 or e-mail: reagents@caliperls.com

References

1. Ferretti J, Köhler W. 2016. History of Streptococcal Research, p 1-26. *In* Vincent A. Fischetti DLS, Joseph J. Ferretti (ed), *Streptococcus pyogenes*: Basic Biology to Clinical Manifestations. University of Oklahoma Health Sciences Center, Oklahoma City (OK).
2. Becker WC. 1916. The Necessity of a Standard Blood-Agar Plate for the Determination of Hemolysis by Streptococci. The Journal of Infectious Diseases 19:754-759.
3. Evans AC. 1936. Studies on Hemolytic Streptococci: II. *Streptococcus pyogenes*. Journal of Bacteriology 31:611-24.
4. Lancefield RC. 1933. A serological differentiation of human and other groups of hemolytic streptococci. Journal of Experimental Medicine 57:571-95.
5. Bessen DE. 2016. Molecular Basis of Serotyping and the Underlying Genetic Organization of *Streptococcus pyogenes*. *In* Ferretti JJ, Stevens DL, Fischetti VA (ed), *Streptococcus pyogenes*: Basic Biology to Clinical Manifestations. University of Oklahoma Health Sciences Center, Oklahoma City (OK).
6. Frost HR, Davies MR, Velusamy S, Delforge V, Erhart A, Darboe S, Steer A, Walker MJ, Beall B, Botteaux A, Smeesters PR. 2020. Updated emm-typing protocol for *Streptococcus pyogenes*. Clinical Microbiology and Infection 26:946 e5-946 e8.
7. Centers for Disease Control and Prevention. . December 2003. Group A streptococcal (GAS) disease. http://www.cdc.gov/ncidod/dbmd/diseaseinfo/groupastreptococcal_t.htm Accessed August 2005.
8. Bessen DE, Beall BW, Hayes A, Huang W, DiChiara JM, Velusamy S, Tettelin H, Jolley KA, Fallon JT, Chochua S, Alobaidallah MSA, Higgs C, Barnett TC, Steemson JT, Proft T, Davies MR. 2024. Recombinational exchange of M-fibril and T-pilus genes generates extensive cell surface diversity in the global group A *Streptococcus* population. mBio 15:e0069324.
9. Hand RM, Snelling TL, Carapetis JR. 2020. Group A *Streptococcus*, p 429-438. *In* Ryan ET, Hill DR, Solomon T, Aronson NE, Endy TP (ed), Hunter's Tropical Medicine and Emerging Infectious Diseases, Tenth ed doi:10.1016/B978-0-323-55512-8.00040-5. Elsevier, London.
10. Wessels MR. 2016. Pharyngitis and Scarlet Fever. *In* Vincent A. Fischetti DLS, Joseph J. Ferretti (ed), *Streptococcus pyogenes*: Basic Biology to Clinical Manifestations. University of Oklahoma Health Sciences Center, Oklahoma City (OK).
11. Stevens DLB, A.E. 2016. Impetigo, Erysipelas, and Cellulitis. *In* Vincent A. Fischetti DLS, Joseph J. Ferretti (ed), *Streptococcus pyogenes*: Basic Biology to Clinical Manifestations. University of Oklahoma Health Sciences Center, Oklahoma City (OK).
12. Centers for Disease Control and Prevention. 2024. Clinical Guidance for Scarlet Fever, *on* Centers for Disease Control. https://www.cdc.gov/group-a-strep/hcp/clinical-guidance/scarlet-fever.html. Accessed Aug 2024.

13. Luk EY, Lo JY, Li AZ, Lau MC, Cheung TK, Wong AY, Wong MM, Wong CW, Chuang SK, Tsang T. 2012. Scarlet fever epidemic, Hong Kong, 2011. Emerging Infectious Diseases 18:1658-61.
14. Turner CE, Pyzio M, Song B, Lamagni T, Meltzer M, Chow JY, Efstratiou A, Curtis S, Sriskandan S. 2016. Scarlet Fever Upsurge in England and Molecular-Genetic Analysis in North-West London, 2014. Emerging Infectious Diseases 22:1075-8.
15. Park DW, Kim SH, Park JW, Kim MJ, Cho SJ, Park HJ, Jung SH, Seo MH, Lee YS, Kim BH, Min H, Lee SY, Ha DR, Kim ES, Hong Y, Chung JK. 2017. Incidence and Characteristics of Scarlet Fever, South Korea, 2008-2015. Emerging Infectious Diseases 23:658-661.
16. Yung CF, Thoon KC. 2018. A 12 year outbreak of scarlet fever in Singapore. The Lancet Infectious Diseases 18:942.
17. Walker MJ, Barnett TC, McArthur JD, Cole JN, Gillen CM, Henningham A, Sriprakash KS, Sanderson-Smith ML, Nizet V. 2014. Disease manifestations and pathogenic mechanisms of Group a *Streptococcus*. Clinical Microbiology Reviews 27:264-301.
18. Stevens DLB, A.E. 2016. Severe Group A Streptococcal Infections. *In* Vincent A. Fischetti DLS, Joseph J. Ferretti (ed), *Streptococcus pyogenes*: Basic Biology to Clinical Manifestations. University of Oklahoma Health Sciences Center, Oklahoma City (OK).
19. Centers for Disease Control and Prevention. 2024. ABCs Bact Facts Interactive Data Dashboard, *on* Centers for Disease Control. https://www.cdc.gov/abcs/bact-facts/data-dashboard.html. Accessed Jul 2024.
20. Sanyahumbi ASC, S.; Wyber, R., Carapetis, J.R. 2016. Global Disease Burden of Group A *Streptococcus*. *In* Vincent A. Fischetti DLS, Joseph J. Ferretti (ed), *Streptococcus pyogenes*: Basic Biology to Clinical Manifestations. University of Oklahoma Health Sciences Center, Oklahoma City (OK).
21. Silva-Costa C, Carrico JA, Ramirez M, Melo-Cristino J. 2014. Scarlet fever is caused by a limited number of *Streptococcus pyogenes* lineages and is associated with the exotoxin genes ssa, speA and speC. The Pediatric Infectious Disease Journal 33:306-10.
22. Roshika R, Jain I, Medicielo J, Wachter J, Danger JL, Sumby P. 2021. The RD2 Pathogenicity Island Modifies the Disease Potential of the Group A *Streptococcus*. Infection and Immunity 89:e0072220.
23. Iyer R, Camilli A. 2007. Sucrose metabolism contributes to in vivo fitness of *Streptococcus pneumoniae*. Molecular Microbiology 66:1-13.
24. Sassetti CM, Rubin EJ. 2003. Genetic requirements for mycobacterial survival during infection. Proceedings of the National Academy of Sciences of the United States of America 100:12989-94.
25. Rollenhagen C, Bumann D. 2006. *Salmonella enterica* highly expressed genes are disease specific. Infection and Immunity 74:1649-60.
26. Steeb B, Claudi B, Burton NA, Tienz P, Schmidt A, Farhan H, Maze A, Bumann D. 2013. Parallel exploitation of diverse host nutrients enhances *Salmonella* virulence. PLoS Pathogens 9:e1003301.

27. Utley M, Franklin DP, Krogfelt KA, Laux DC, Cohen PS. 1998. A *Salmonella typhimurium* mutant unable to utilize fatty acids and citrate is avirulent and immunogenic in mice. FEMS Microbiology Letters 163:129-34.
28. Pancholi V, Caparon M. 2016. *Streptococcus pyogenes* Metabolism. *In* Ferretti JJ, Stevens DL, Fischetti VA (ed), *Streptococcus pyogenes* : Basic Biology to Clinical Manifestations. University of Oklahoma Health Sciences Center, Oklahoma City (OK).
29. Wishart DS, Guo A, Oler E, Wang F, Anjum A, Peters H, Dizon R, Sayeeda Z, Tian S, Lee BL, Berjanskii M, Mah R, Yamamoto M, Jovel J, Torres-Calzada C, Hiebert-Giesbrecht M, Lui VW, Varshavi D, Varshavi D, Allen D, Arndt D, Khetarpal N, Sivakumaran A, Harford K, Sanford S, Yee K, Cao X, Budinski Z, Liigand J, Zhang L, Zheng J, Mandal R, Karu N, Dambrova M, Schioth HB, Greiner R, Gautam V. 2022. HMDB 5.0: the Human Metabolome Database for 2022. Nucleic Acids Reseach 50:D622-D631.
30. Görke B, Stülke J. 2008. Carbon catabolite repression in bacteria: many ways to make the most out of nutrients. Nature Reviews Microbiology 6:613-24.
31. Poncet S, Milohanic E, Mazé A, Abdallah JN, Aké F, Larribe M, Deghmane AE, Taha MK, Dozot M, De Bolle X, Letesson JJ, Deutscher J. 2009. Correlations between carbon metabolism and virulence in bacteria. Contributions to Microbiology 16:88-102.
32. Iyer R, Baliga NS, Camilli A. 2005. Catabolite control protein A (CcpA) contributes to virulence and regulation of sugar metabolism in *Streptococcus pneumoniae*. Journal of Bacteriology 187:8340-9.
33. Carlson-Banning KM, Sperandio V. 2016. Catabolite and Oxygen Regulation of Enterohemorrhagic *Escherichia coli* Virulence. mBio 7:e01852-16.
34. Richardson AR. 2019. Virulence and Metabolism. Microbiology Spectrum 7:10.1128/microbiolspec.gpp3-0011-2018.
35. Somarajan SR, Roh JH, Singh KV, Weinstock GM, Murray BE. 2014. CcpA is important for growth and virulence of *Enterococcus faecium*. Infection and Immunity 82:3580-7.
36. Antunes A, Martin-Verstraete I, Dupuy B. 2011. CcpA-mediated repression of *Clostridium difficile* toxin gene expression. Molecular Microbiology 79:882-99.
37. Abranches J, Nascimento MM, Zeng L, Browngardt CM, Wen ZT, Rivera MF, Burne RA. 2008. CcpA Regulates Central Metabolism and Virulence Gene Expression in *Streptococcus mutans*. Journal of Bacteriology 190:2340-2349.
38. Le Breton Y, Belew AT, Valdes KM, Islam E, Curry P, Tettelin H, Shirtliff ME, El-Sayed NM, McIver KS. 2015. Essential Genes in the Core Genome of the Human Pathogen *Streptococcus pyogenes*. Scientific Reports 5:9838.
39. Sundar GS, Islam E, Gera K, Le Breton Y, McIver KS. 2017. A PTS EII mutant library in Group A *Streptococcus* identifies a promiscuous man-family PTS transporter influencing SLS-mediated hemolysis. Molecular Microbiology 103:518-533.
40. Sundar GS, Islam E, Braza RD, Silver AB, Le Breton Y, McIver KS. 2018. Route of Glucose Uptake in the Group A *Streptococcus* Impacts SLS-Mediated

Hemolysis and Survival in Human Blood. Frontiers in Cellular and Infection Microbiology 8:71.
41. Deutscher J, Aké FM, Derkaoui M, Zébré AC, Cao TN, Bouraoui H, Kentache T, Mokhtari A, Milohanic E, Joyet P. 2014. The bacterial phosphoenolpyruvate:carbohydrate phosphotransferase system: regulation by protein phosphorylation and phosphorylation-dependent protein-protein interactions. Microbiology and Molecular Biology Reviews 78:231-56.
42. Alpert CA, Dorschug M, Saffen D, Frank R, Deutscher J, Hengstenberg W. 1985. The bacterial phosphoenolpyruvate-dependent phosphotransferase system. Isolation of active site peptides by reversed-phase high-performance liquid chromatography and determination of their primary structure. Journal of Chromatography 326:363-71.
43. Cases I, Velazquez F, de Lorenzo V. 2007. The ancestral role of the phosphoenolpyruvate-carbohydrate phosphotransferase system (PTS) as exposed by comparative genomics. Research in Microbiology 158:666-70.
44. Kanehisa M, Goto S. 2000. KEGG: kyoto encyclopedia of genes and genomes. Nucleic Acids Reseach 28:27-30.
45. Kanehisa M. 2019. Toward understanding the origin and evolution of cellular organisms. Protein Science 28:1947-1951.
46. Kanehisa M, Furumichi M, Sato Y, Kawashima M, Ishiguro-Watanabe M. 2023. KEGG for taxonomy-based analysis of pathways and genomes. Nucleic Acids Research 51:D587-D592.
47. Kanehisa Laboratories. Phosphotransferase system (PTS) - *Streptococcus pyogenes* MGAS5005 (serotype M1). https://www.kegg.jp/pathway/spz02060. Accessed Sept 2024.
48. André A, Maccheroni W, Doignon F, Garnier M, Renaudin J. 2003. Glucose and trehalose PTS permeases of *Spiroplasma citri* probably share a single IIA domain, enabling the spiroplasma to adapt quickly to carbohydrate changes in its environment. Microbiology (Reading) 149:2687-2696.
49. Shelburne SA, 3rd, Fang H, Okorafor N, Sumby P, Sitkiewicz I, Keith D, Patel P, Austin C, Graviss EA, Musser JM, Chow DC. 2007. MalE of Group A *Streptococcus* participates in the rapid transport of maltotriose and longer maltodextrins. Journal of Bacteriology 189:2610-7.
50. Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Research 25:3389-402.
51. Deutscher J, Saier MH, Jr. 1983. ATP-dependent protein kinase-catalyzed phosphorylation of a seryl residue in HPr, a phosphate carrier protein of the phosphotransferase system in *Streptococcus pyogenes*. Proceedings of the National Academy of Sciences of the United States of America 80:6790-4.
52. Zomer AL, Buist G, Larsen R, Kok J, Kuipers OP. 2007. Time-resolved determination of the CcpA regulon of *Lactococcus lactis* subsp. *cremoris* MG1363. Journal of Bacteriology 189:1366-81.
53. Almengor AC, Kinkel TL, Day SJ, McIver KS. 2007. The catabolite control protein CcpA binds to P*mga* and influences expression of the virulence

regulator Mga in the group A *Streptococcus*. Journal of Bacteriology 189:8405-8416.
54. Wen ZT, Burne RA. 2002. Functional genomics approach to identifying genes required for biofilm development by *Streptococcus mutans*. Applied Environmental Microbiology 68:1196-203.
55. Giammarinaro P, Paton JC. 2002. Role of RegM, a homologue of the catabolite repressor protein CcpA, in the virulence of *Streptococcus pneumoniae*. Infection and Immunity 70:5454-61.
56. Kinkel TL, McIver KS. 2008. CcpA-mediated repression of streptolysin S expression and virulence in the Group A *Streptococcus*. Infection and Immunity 76:3451-63.
57. Shelburne SA, 3rd, Keith D, Horstmann N, Sumby P, Davenport MT, Graviss EA, Brennan RG, Musser JM. 2008. A direct link between carbohydrate utilization and virulence in the major human pathogen group A *Streptococcus*. Proceedings of the National Academy of Sciences of the United States of America 105:1698-703.
58. Homburg C, Bommer M, Wuttge S, Hobe C, Beck S, Dobbek H, Deutscher J, Licht A, Schneider E. 2017. Inducer exclusion in Firmicutes: insights into the regulation of a carbohydrate ATP binding cassette transporter from *Lactobacillus casei* BL23 by the signal transducing protein P-Ser46-HPr. Molecular Microbiology 105:25-45.
59. Darbon E, Servant P, Poncet S, Deutscher J. 2002. Antitermination by GlpP, catabolite repression via CcpA and inducer exclusion triggered by P-GlpK dephosphorylation control *Bacillus subtilis* glpFK expression. Molecular Microbiology 43:1039-52.
60. Deutscher J, Francke C, Postma PW. 2006. How phosphotransferase system-related protein phosphorylation regulates carbohydrate metabolism in bacteria. Microbiology and Molecular Biology Reviews 70:939-1031.
61. Pompeo F, Luciano J, Galinier A. 2007. Interaction of GapA with HPr and its homologue, Crh: Novel levels of regulation of a key step of glycolysis in *Bacillus subtilis*? Journal of Bacteriology 189:1154-7.
62. Todd EW. 1932. Antigenic streptococcal hemolysin. Journal of Experimental Medicine 55:267-80.
63. Fox EN, Krampitz LO. 1956. Studies on the biosynthesis of the M-protein of group A hemolytic streptococci. Journal of Bacteriology 71:454-61.
64. Fischetti VA. 2016. M Protein and Other Surface Proteins on Streptococci. *In* Ferretti JJ, Stevens DL, Fischetti VA (ed), *Streptococcus pyogenes*: Basic Biology to Clinical Manifestations. University of Oklahoma Health Sciences Center, Oklahoma City (OK).
65. Pine L, Reeves MW. 1978. Regulation of the synthesis of M protein by sugars, Todd Hewitt broth, and horse serum, in growing cells of *Streptococcus pyogenes*. Microbios 21:185-212.
66. Chaussee MS, Phillips ER, Ferretti JJ. 1997. Temporal production of streptococcal erythrogenic toxin B (streptococcal cysteine proteinase) in response to nutrient depletion. Infection and Immunity 65:1956-9.

67. DebRoy S, Saldana M, Travisany D, Montano A, Galloway-Pena J, Horstmann N, Yao H, Gonzalez M, Maass A, Latorre M, Shelburne SA. 2016. A Multi-Serotype Approach Clarifies the Catabolite Control Protein A Regulon in the Major Human Pathogen Group A *Streptococcus*. Scientific Reports 6:32442.
68. Loughman JA, Caparon MG. 2006. A novel adaptation of aldolase regulates virulence in *Streptococcus pyogenes*. EMBO Journal 25:5414-22.
69. Gera K, Le T, Jamin R, Eichenbaum Z, McIver KS. 2014. The Phosphoenolpyruvate Phosphotransferase System in Group A *Streptococcus* Acts To Reduce Streptolysin S Activity and Lesion Severity during Soft Tissue Infection. Infection and Immunity 82:1192-204.
70. Merriman JA, Xu W, Caparon MG. 2023. Central carbon flux controls growth/damage balance for *Streptococcus pyogenes*. PLoS Pathogens 19:e1011481.
71. Graham MR, Virtaneva K, Porcella SF, Barry WT, Gowen BB, Johnson CR, Wright FA, Musser JM. 2005. Group A *Streptococcus* transcriptome dynamics during growth in human blood reveals bacterial adaptive and survival strategies. American Journal of Pathology 166:455-65.
72. Graham MR, Virtaneva K, Porcella SF, Gardner DJ, Long RD, Welty DM, Barry WT, Johnson CA, Parkins LD, Wright FA, Musser JM. 2006. Analysis of the transcriptome of group A *Streptococcus* in mouse soft tissue infection. American Journal of Pathology 169:927-42.
73. Virtaneva K, Porcella SF, Graham MR, Ireland RM, Johnson CA, Ricklefs SM, Babar I, Parkins LD, Romero RA, Corn GJ, Gardner DJ, Bailey JR, Parnell MJ, Musser JM. 2005. Longitudinal analysis of the group A *Streptococcus* transcriptome in experimental pharyngitis in cynomolgus macaques. Proceedings of the National Academy of Sciences of the United States of America 102:9014-9.
74. Graham MR, Smoot LM, Migliaccio CA, Virtaneva K, Sturdevant DE, Porcella SF, Federle MJ, Adams GJ, Scott JR, Musser JM. 2002. Virulence control in group A *Streptococcus* by a two-component gene regulatory system: global expression profiling and in vivo infection modeling. Proceedings of the National Academy of Sciences of the United States of America 99:13855-60.
75. Ribardo DA, McIver KS. 2006. Defining the Mga regulon: comparative transcriptome analysis reveals both direct and indirect regulation by Mga in the group A *Streptococcus*. Molecular Microbiology 62:491-508.
76. Shelburne SA, 3rd, Sumby P, Sitkiewicz I, Granville C, DeLeo FR, Musser JM. 2005. Central role of a bacterial two-component gene regulatory system of previously unknown function in pathogen persistence in human saliva. Proceedings of the National Academy of Sciences of the United States of America 102:16037-42.
77. Vega LA, Malke H, McIver KS. 2022. Virulence-Related Transcriptional Regulators of *Streptococcus pyogenes*. *In* Ferretti JJ, Stevens DL, Fischetti VA (ed), *Streptococcus pyogenes*: Basic Biology to Clinical Manifestations, 2nd ed. University of Oklahoma Health Sciences Center, Oklahoma City (OK).

78. Kreikemeyer B, McIver KS, Podbielski A. 2003. Virulence factor regulation and regulatory networks in *Streptococcus pyogenes* and their impact on pathogen-host interactions. Trends in Microbiology 11:224-32.
79. McIver KS. 2009. Stand-alone response regulators controlling global virulence networks in *Streptococcus pyogenes*. Contributions to Microbiology 16:103-19.
80. Brouwer S, Barnett TC, Rivera-Hernandez T, Rohde M, Walker MJ. 2016. *Streptococcus pyogenes* adhesion and colonization. FEBS Lett 590:3739-3757.
81. Papon N, Stock AM. 2019. Two-component systems. Current Biology 29:R724-R725.
82. Buckley SJ, Timms P, Davies MR, McMillan DJ. 2018. In silico characterisation of the two-component system regulators of *Streptococcus pyogenes*. PLoS One 13:e0199163.
83. Sitkiewicz I, Musser JM. 2006. Expression microarray and mouse virulence analysis of four conserved two-component gene regulatory systems in group A *Streptococcus*. Infection and Immunity 74:1339-51.
84. Buckley SJ, Davies MR, McMillan DJ. 2020. In silico characterisation of stand-alone response regulators of *Streptococcus pyogenes*. PLoS One 15:e0240834.
85. Pitman S, Cho KH. 2015. The Mechanisms of Virulence Regulation by Small Noncoding RNAs in Low GC Gram-Positive Pathogens. International Journal of Molecular Sciences 16:29797-814.
86. Rom JS, Hart MT, McIver KS. 2021. PRD-Containing Virulence Regulators (PCVRs) in Pathogenic Bacteria. Frontiers in Cellular and Infection Microbiology 11:772874.
87. Wu J, McAuliffe O, O'Byrne CP. 2023. A novel RofA-family transcriptional regulator, GadR, controls the development of acid resistance in *Listeria monocytogenes*. mBio 14:e0171623.
88. Bessen DE, Manoharan A, Luo F, Wertz JE, Robinson DA. 2005. Evolution of transcription regulatory genes is linked to niche specialization in the bacterial pathogen *Streptococcus pyogenes*. Journal of Bacteriology 187:4163-72.
89. McIver KS, Scott JR. 1997. Role of *mga* in growth phase regulation of virulence genes of the group A *Streptococcus*. Journal of Bacteriology 179:5178-5187.
90. Okada N, Geist RT, Caparon MG. 1993. Positive transcriptional control of *mry* regulates virulence in the group A *Streptococcus*. Molecular Microbiology 7:893-903.
91. McIver KS, Thurman AS, Scott JR. 1999. Regulation of *mga* transcription in the group A *Streptococcus*: specific binding of Mga within its own promoter and evidence for a negative regulator. Journal of Bacteriology 181:5373-5383.
92. Sanson M, Makthal N, Gavagan M, Cantu C, Olsen RJ, Musser JM, Kumaraswami M. 2015. Phosphorylation events in the multiple gene regulator of group A *Streptococcus* significantly influence global gene expression and virulence. Infection and Immunity 83:2382-95.
93. McIver KS, Myles RL. 2002. Two DNA-binding domains of Mga are required for virulence gene activation in the group A *Streptococcus*. Molecular Microbiology 43:1591-1602.

94. Hause LL, McIver KS. 2012. Nucleotides critical for the interaction of the *Streptococcus pyogenes* Mga virulence regulator with Mga-regulated promoter sequences. Journal of Bacteriology 194:4904-19.
95. Hondorp ER, Hou SC, Hause LL, Gera K, Lee CE, McIver KS. 2013. PTS phosphorylation of Mga modulates regulon expression and virulence in the group A *Streptococcus*. Molecular Microbiology 88:1176-93.
96. Valdes KM, Sundar GS, Belew AT, Islam E, El-Sayed NM, Le Breton Y, McIver KS. 2018. Glucose Levels Alter the Mga Virulence Regulon in the Group A *Streptococcus*. Scientific Reports 8:4971.
97. Hondorp ER, Hou SC, Hempstead AD, Hause LL, Beckett DM, McIver KS. 2012. Characterization of the Group A *Streptococcus* Mga virulence regulator reveals a role for the C-terminal region in oligomerization and transcriptional activation. Molecular Microbiology 83:953-67.
98. Bourgogne A, Drysdale M, Hilsenbeck SG, Peterson SN, Koehler TM. 2003. Global effects of virulence gene regulators in a *Bacillus anthracis* strain with both virulence plasmids. Infection and Immunity 71:2736-43.
99. Dai Z, Sirard JC, Mock M, Koehler TM. 1995. The *atxA* gene product activates transcription of the anthrax toxin genes and is essential for virulence. Molecular Microbiology 16:1171-1181.
100. Scarff JM, Raynor MJ, Seldina YI, Ventura CL, Koehler TM, O'Brien AD. 2016. The roles of AtxA orthologs in virulence of anthrax-like *Bacillus cereus* G9241. Molecular Microbiology 102:545-561.
101. Tsvetanova B, Wilson AC, Bongiorni C, Chiang C, Hoch JA, Perego M. 2007. Opposing effects of histidine phosphorylation regulate the AtxA virulence transcription factor in *Bacillus anthracis*. Molecular Microbiology 63:644-55.
102. Hammerstrom TG, Roh JH, Nikonowicz EP, Koehler TM. 2011. *Bacillus anthracis* virulence regulator AtxA: oligomeric state, function and CO(2) -signalling. Molecular Microbiology 82:634-47.
103. McCall RM, Sievers ME, Fattah R, Ghirlando R, Pomerantsev AP, Leppla SH. 2019. *Bacillus anthracis* Virulence Regulator AtxA Binds Specifically to the *pagA* Promoter Region. Journal of Bacteriology 201:e00569-19.
104. Bier N, Hammerstrom TG, Koehler TM. 2020. Influence of the phosphoenolpyruvate:carbohydrate phosphotransferase system on toxin gene expression and virulence in *Bacillus anthracis*. Molecular Microbiology 113:237-252.
105. Hammerstrom TG, Horton LB, Swick MC, Joachimiak A, Osipiuk J, Koehler TM. 2015. Crystal structure of *Bacillus anthracis* virulence regulator AtxA and effects of phosphorylated histidines on multimerization and activity. Molecular Microbiology 95:426-41.
106. Chiang C, Bongiorni C, Perego M. 2011. Glucose-dependent activation of *Bacillus anthracis* toxin gene expression and virulence requires the carbon catabolite protein CcpA. Journal of Bacteriology 193:52-62.
107. Fogg GC, Gibson CM, Caparon MG. 1994. The identification of *rofA*, a positive-acting regulatory component of *prtF* expression: use of an m-gamma-delta-based shuttle mutagenesis strategy in *Streptococcus pyogene*s. Molecular Microbiology 11:671-684.

108. Bessen DE, Kalia A. 2002. Genomic localization of a T serotype locus to a recombinatorial zone encoding extracellular matrix-binding proteins in *Streptococcus pyogenes*. Infection and Immunity 70:1159-67.
109. Nakata M, Kreikemeyer B. 2021. Genetics, Structure, and Function of Group A Streptococcal Pili. Frontiers in Microbiology 12:616508.
110. Fogg GC, Caparon MG. 1997. Constitutive expression of fibronectin binding in *Streptococcus pyogenes* as a result of anaerobic activation of *rofA*. Journal of Bacteriology 179:6172-6180.
111. Granok AB, Parsonage D, Ross RP, Caparon MG. 2000. The RofA binding site in *Streptococcus pyogenes* is utilized in multiple transcriptional pathways. Journal of Bacteriology 182:1529-1540.
112. Beckert S, Kreikemeyer B, Podbielski A. 2001. Group A streptococcal *rofA* gene is involved in the control of several virulence genes and eukaryotic cell attachment and internalization. Infection and Immunity 69:534-537.
113. Kreikemeyer B, Beckert S, Braun-Kiewnick A, Podbielski A. 2002. Group A streptococcal RofA-type global regulators exhibit a strain-specific genomic presence and regulation pattern. Microbiology 148:1501-11.
114. Calfee G, Danger JL, Jain I, Miller EW, Sarkar P, Tjaden B, Kreikemeyer B, Sumby P. 2018. Identification and Characterization of Serotype-Specific Variation in Group A *Streptococcus* Pilus Expression. Infection and Immunity 86:e00792-17.
115. Le Breton Y, Belew AT, Freiberg JA, Sundar GS, Islam E, Lieberman J, Shirtliff ME, Tettelin H, El-Sayed NM, McIver KS. 2017. Genome-wide discovery of novel M1T1 group A streptococcal determinants important for fitness and virulence during soft-tissue infection. PLoS Pathogens 13:e1006584.
116. Zhu L, Olsen RJ, Beres SB, Saavedra MO, Kubiak SL, Cantu CC, Jenkins L, Waller AS, Sun Z, Palzkill T, Porter AR, DeLeo FR, Musser JM. 2020. *Streptococcus pyogenes* genes that promote pharyngitis in primates. JCI Insight 5:e137686.
117. Takahashi R, Radcliff FJ, Proft T, Tsai CJ. 2022. Pilus proteins from *Streptococcus pyogenes* stimulate innate immune responses through Toll-like receptor 2. Immunology & Cell Biology 100:174-185.
118. Jain I, Sarkar P, Danger JL, Medicielo J, Roshika R, Calfee G, Ramalinga A, Burgess C, Sumby P. 2019. A Mobile Genetic Element Promotes the Association Between Serotype M28 Group A *Streptococcus* Isolates and Cases of Puerperal Sepsis. Journal of Infectious Diseases 220:882-891.
119. Bessen DE, Sotir CM, Readdy TL, Hollingshead SK. 1996. Genetic correlates of throat and skin isolates of group A streptococci. Journal of Infectious Diseases 173:896-900.
120. Podbielski A, Woischnik M, Leonard BA, Schmidt KH. 1999. Characterization of *nra*, a global negative regulator gene in group A streptococci. Molecular Microbiology 31:1051-64.
121. Kreikemeyer B, Nakata M, Köller T, Hildisch H, Kourakos V, Standar K, Kawabata S, Glocker MO, Podbielski A. 2007. The *Streptococcus pyogenes* serotype M49 Nra-Ralp3 transcriptional regulatory network and its control of

virulence factor expression from the novel *eno ralp3 epf sagA* pathogenicity region. Infection and Immunity 75:5698-5710.
122. Roberts SA, Scott JR. 2007. RivR and the small RNA RivX: the missing links between the CovR regulatory cascade and the Mga regulon. Molecular Microbiology 66:1506-22.
123. Nakata M, Sumitomo T, Patenge N, Kreikemeyer B, Kawabata S. 2020. Thermosensitive pilus production by FCT type 3 *Streptococcus pyogenes* controlled by Nra regulator translational efficiency. Molecular Microbiology 113:173-189.
124. Molinari G, Rohde M, Talay SR, Chhatwal GS, Beckert S, Podbielski A. 2001. The role played by the group A streptococcal negative regulator Nra on bacterial interactions with epithelial cells. Molecular Microbiology 40:99-114.
125. Kwinn LA, Khosravi A, Aziz RK, Timmer AM, Doran KS, Kotb M, Nizet V. 2007. Genetic characterization and virulence role of the RALP3/LSA locus upstream of the streptolysin s operon in invasive M1T1 Group A *Streptococcus*. Journal of Bacteriology 189:1322-9.
126. Siemens N, Fiedler T, Normann J, Klein J, Munch R, Patenge N, Kreikemeyer B. 2012. Effects of the ERES Pathogenicity Region Regulator Ralp3 on *Streptococcus pyogenes* serotype M49 Virulence Factor Expression. Journal of Bacteriology 194:3618-26.
127. Treviño J, Liu Z, Cao TN, Ramirez-Peña E, Sumby P. 2013. RivR is a negative regulator of virulence factor expression in group A *Streptococcus*. Infection and Immunity 81:364-72.
128. Ramalinga A, Danger JL, Makthal N, Kumaraswami M, Sumby P. 2017. Multimerization of the Virulence-Enhancing Group A *Streptococcus* Transcription Factor RivR Is Required for Regulatory Activity. Journal of Bacteriology 199:e00452-16.
129. Hemsley C, Joyce E, Hava DL, Kawale A, Camilli A. 2003. MgrA, an orthologue of Mga, Acts as a transcriptional repressor of the genes within the *rlrA* pathogenicity islet in *Streptococcus pneumoniae*. Journal of Bacteriology 185:6640-7.
130. Solano-Collado V, Lurz R, Espinosa M, Bravo A. 2013. The pneumococcal MgaSpn virulence transcriptional regulator generates multimeric complexes on linear double-stranded DNA. Nucleic Acids Research 41:6975-91.
131. Suo W, Guo X, Zhang X, Xiao S, Wang S, Yin Y, Zheng Y. 2023. Glucose levels affect MgaSpn regulation on the virulence and adaptability of *Streptococcus pneumoniae*. Microbial Pathogenesis 174:105896.
132. Hava DL, Camilli A. 2002. Large-scale identification of serotype 4 *Streptococcus pneumoniae* virulence factors. Molecular Microbiology 45:1389-406.
133. Hava DL, Hemsley CJ, Camilli A. 2003. Transcriptional regulation in the *Streptococcus pneumoniae rlrA* pathogenicity islet by RlrA. Journal of Bacteriology 185:413-21.
134. Vietri NJ, Marrero R, Hoover TA, Welkos SL. 1995. Identification and characterization of a trans-activator involved in the regulation of encapsulation by *Bacillus anthracis*. Gene 152:1-9.

135. Drysdale M, Bourgogne A, Hilsenbeck SG, Koehler TM. 2004. *atxA* controls *Bacillus anthracis* capsule synthesis via *acpA* and a newly discovered regulator, *acpB*. Journal of Bacteriology 186:307-15.
136. Raynor MJ, Roh JH, Widen SG, Wood TG, Koehler TM. 2018. Regulons and protein-protein interactions of PRD-containing *Bacillus anthracis* virulence regulators reveal overlapping but distinct functions. Molecular Microbiology 109:1-22.
137. Drysdale M, Heninger S, Hutt J, Chen Y, Lyons CR, Koehler TM. 2005. Capsule synthesis by *Bacillus anthracis* is required for dissemination in murine inhalation anthrax. EMBO Journal 24:221-7.
138. Sittner A, Bar-David E, Glinert I, Ben-Shmuel A, Schlomovitz J, Levy H, Weiss S. 2021. Role of *acpA* and *acpB* in *Bacillus anthracis* capsule accumulation and toxin independent pathogenicity in rabbits. Microbial Pathogenesis 155:104904.
139. Hondorp ER, McIver KS. 2007. The Mga virulence regulon: infection where the grass is greener. Molecular Microbiology 66:1056-65.
140. Chatellier S, Ihendyane N, Kansal RG, Khambaty F, Basma H, Norrby-Teglund A, Low DE, McGeer A, Kotb M. 2000. Genetic relatedness and superantigen expression in group A *Streptococcus* serotype M1 isolates from patients with severe and nonsevere invasive diseases. Infection and Immunity 68:3523-34.
141. Sumby P, Porcella SF, Madrigal AG, Barbian KD, Virtaneva K, Ricklefs SM, Sturdevant DE, Graham MR, Vuopio-Varkila J, Hoe NP, Musser JM. 2005. Evolutionary origin and emergence of a highly successful clone of serotype M1 group A *Streptococcus* involved multiple horizontal gene transfer events. Journal of Infectious Diseases 192:771-82.
142. Lyon WR, Gibson CM, Caparon MG. 1998. A role for trigger factor and an *rgg*-like regulator in the transcription, secretion and processing of the cysteine proteinase of *Streptococcus pyogenes*. EMBO Journal 17:6263-6275.
143. Studier FW. 2005. Protein production by auto-induction in high density shaking cultures. Protein Expression and Purification 41:207-34.
144. Hanahan D. 1983. Studies on transformation of *Escherichia coli* with plasmids. Journal of Molecular Biology 166:557-80.
145. Miroux B, Walker JE. 1996. Over-production of proteins in *Escherichia coli*: mutant hosts that allow synthesis of some membrane proteins and globular proteins at high levels. Journal of Molecular Biology 260:289-98.
146. Marinus MG, Morris NR. 1973. Isolation of deoxyribonucleic acid methylase mutants of *Escherichia coli* K-12. Journal of Bacteriology 114:1143-50.
147. Studier FW, Moffatt BA. 1986. Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. Journal of Molecular Biology 189:113-30.
148. Park HS, Francis KP, Yu J, Cleary PP. 2003. Membranous cells in nasal-associated lymphoid tissue: a portal of entry for the respiratory mucosal pathogen group A *Streptococcus*. The Journal of Immunology 171:2532-7.
149. Braza RE, Silver AB, Sundar GS, Davis SE, Razi A, Islam E, Hart M, Zhu J, Le Breton Y, McIver KS. 2020. Phosphotransferase System Uptake and

Metabolism of the beta-Glucoside Salicin Impact Group A Streptococcal Bloodstream Survival and Soft Tissue Infection. Infect Immun 88.
150. Le Breton Y, Mistry P, Valdes KM, Quigley J, Kumar N, Tettelin H, McIver KS. 2013. Genome-wide identification of genes required for fitness of Group A *Streptococcus* in human blood. Infection and Immunity 81:862-875.
151. Que YA, Haefliger JA, Francioli P, Moreillon P. 2000. Expression of *Staphylococcus aureus* clumping factor A in *Lactococcus lactis* subsp. *cremoris* using a new shuttle vector. Infection and Immunity 68:3516-3522.
152. Naville M, Ghuillot-Gaudeffroy A, Marchais A, Gautheret D. 2011. ARNold: a web tool for the prediction of Rho-independent transcription terminators. RNA Biology 8:11-3.
153. Le Breton Y, McIver KS. 2013. Genetic manipulation of *Streptococcus pyogenes* (the Group A *Streptococcus*, GAS). Current Protocols in Microbiology 30:Unit 9D 3.
154. Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30:2114-2120.
155. Kim D, Paggi JM, Park C, Bennett C, Salzberg SL. 2019. Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. Nature Biotechnology 37:907-915.
156. Leek JT, Johnson WE, Parker HS, Jaffe AE, Storey JD. 2012. The sva package for removing batch effects and other unwanted variation in high-throughput experiments. Bioinformatics 28:882-883.
157. Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biology 15:550.
158. Young MD, Wakefield MJ, Smyth GK, Oshlack A. 2010. Gene ontology analysis for RNA-seq: accounting for selection bias. Genome Biology 11:R14.
159. Dehal PS, Joachimiak MP, Price MN, Bates JT, Baumohl JK, Chivian D, Friedland GD, Huang KH, Keller K, Novichkov PS, Dubchak IL, Alm EJ, Arkin AP. 2010. MicrobesOnline: an integrated portal for comparative and functional genomics. Nucleic Acids Research 38:D396-400.
160. Ye J, Coulouris G, Zaretskaya I, Cutcutache I, Rozen S, Madden TL. 2012. Primer-BLAST: a tool to design target-specific primers for polymerase chain reaction. BMC Bioinformatics 13:134.
161. Livak KJ, Schmittgen TD. 2001. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods 25:402-8.
162. Falaleeva M, Zurek OW, Watkins RL, Reed RW, Ali H, Sumby P, Voyich JM, Korotkova N. 2014. Transcription of the *Streptococcus pyogenes* hyaluronic acid capsule biosynthesis operon is regulated by previously unknown upstream elements. Infection and Immunity 82:5293-307.
163. Khemlani AHJ, Proft T, Loh JMS. 2020. Assays to Analyze Adhesion of Group A *Streptococcus* to Host Cells, p 271-278. *In* Proft T, Loh J (ed), Group A Streptococcus, vol 2136. Humana, New York, NY.
164. Crotty Alexander LE, Maisey HC, Timmer AM, Rooijakkers SH, Gallo RL, von Köckritz-Blickwede M, Nizet V. 2010. M1T1 group A streptococcal pili promote epithelial colonization but diminish systemic virulence through

neutrophil extracellular entrapment. Journal of Molecular Medicine 88:371-381.
165. Nagy Z, Comer S, Smolenski A. 2018. Analysis of Protein Phosphorylation Using Phos-Tag Gels. Current Protocols in Protein Science 93:e64.
166. Cold Spring Harbor. 2006. Sodium phosphate. Cold Spring Harbor Protocols 2006:pdb.rec8303.
167. Sprouffske K, Wagner A. 2016. Growthcurver: an R package for obtaining interpretable metrics from microbial growth curves. BMC Bioinformatics 17:172.
168. Blazanin M. 2024. gcplyr: an R package for microbial growth curve data analysis. BMC Bioinformatics 25:232.
169. Vega LA, Valdes KM, Sundar GS, Belew AT, Islam E, Berge J, Curry P, Chen S, El-Sayed NM, Le Breton Y, McIver KS. 2017. The Transcriptional Regulator CpsY Is Important for Innate Immune Evasion in *Streptococcus pyogenes*. Infection and Immunity 85:e00925-16.
170. Zhi X, Vieira A, Huse KK, Martel PJ, Lobkowicz L, Li HK, Croucher N, Andrew I, Game L, Sriskandan S. 2023. Characterization of the RofA regulon in the pandemic M1(global) and emergent M1(UK) lineages of *Streptococcus pyogenes*. Microbial Genomics 9:001159.
171. Wessels MR. 2016. Cell Wall and Surface Molecules: Capsule. *In* Ferretti JJ, Stevens DL, Fischetti VA (ed), *Streptococcus pyogenes*: Basic Biology to Clinical Manifestations. University of Oklahoma Health Sciences Center, Oklahoma City (OK).
172. Sumby P, Whitney AR, Graviss EA, DeLeo FR, Musser JM. 2006. Genome-wide analysis of group a streptococci reveals a mutation that modulates global phenotype and disease specificity. PLoS Pathogens 2:e5.
173. Manetti AG, Zingaretti C, Falugi F, Capo S, Bombaci M, Bagnoli F, Gambellini G, Bensi G, Mora M, Edwards AM, Musser JM, Graviss EA, Telford JL, Grandi G, Margarit I. 2007. *Streptococcus pyogenes* pili promote pharyngeal cell adhesion and biofilm formation. Molecular Microbiology 64:968-83.
174. Abbot EL, Smith WD, Siou GP, Chiriboga C, Smith RJ, Wilson JA, Hirst BH, Kehoe MA. 2007. Pili mediate specific adhesion of *Streptococcus pyogenes* to human tonsil and skin. Cellular Microbiology 9:1822-33.
175. Bartelt MA, Duncan JL. 1978. Adherence of group A streptococci to human epithelial cells. Infection and Immunity 20:200-8.
176. Ravins M, Jaffe J, Hanski E, Shetzigovski I, Natanson-Yaron S, Moses AE. 2000. Characterization of a mouse-passaged, highly encapsulated variant of group A *Streptococcus* in in vitro and in vivo studies. Journal of Infectious Disease 182:1702-11.
177. Darmstadt GL, Mentele L, Podbielski A, Rubens CE. 2000. Role of group A streptococcal virulence factors in adherence to keratinocytes. Infection and Immunity 68:1215-1221.
178. Schrager HM, Alberti S, Cywes C, Dougherty GJ, Wessels MR. 1998. Hyaluronic acid capsule modulates M protein-mediated adherence and acts as a ligand for attachment of group A *Streptococcus* to CD44 on human keratinocytes. Journal of Clinical Investigation 101:1708-16.

179. Cywes C, Stamenkovic I, Wessels MR. 2000. CD44 as a receptor for colonization of the pharynx by group A *Streptococcus*. Journal of Clinical Investigation 106:995-1002.
180. Courtney HS, Hasty DL. 1991. Aggregation of group A streptococci by human saliva and effect of saliva on streptococcal adherence to host cells. Infection and Immunity 59:1661-6.
181. de Jong A, Pietersma H, Cordes M, Kuipers OP, Kok J. 2012. PePPER: a webserver for prediction of prokaryote promoter elements and regulons. BMC Genomics 13:299.
182. Slade HD, Knox GA. 1950. Nutrition and the role of reducing agents in the formation of streptolysin O by a group A hemolytic *Streptococcus*. Journal of Bacteriology 60:301-10.
183. Marchler-Bauer A, Lu S, Anderson JB, Chitsaz F, Derbyshire MK, DeWeese-Scott C, Fong JH, Geer LY, Geer RC, Gonzales NR, Gwadz M, Hurwitz DI, Jackson JD, Ke Z, Lanczycki CJ, Lu F, Marchler GH, Mullokandov M, Omelchenko MV, Robertson CL, Song JS, Thanki N, Yamashita RA, Zhang D, Zhang N, Zheng C, Bryant SH. 2011. CDD: a Conserved Domain Database for the functional annotation of proteins. Nucleic Acids Research 39:D225-9.
184. Kelley LA, Mezulis S, Yates CM, Wass MN, Sternberg MJ. 2015. The Phyre2 web portal for protein modeling, prediction and analysis. Nature Protocols 10:845-58.
185. Wessels MR, Moses AE, Goldberg JB, DiCesare TJ. 1991. Hyaluronic acid capsule is a virulence factor for mucoid group A streptococci. Proceedings of the National Academy of Sciences of the United States of America 88:8317-8321.
186. Dale JB, Washburn RG, Marques MB, Wessels MR. 1996. Hyaluronate capsule and surface M protein in resistance to opsonization of group A streptococci. Infection and Immunity 64:1495-501.
187. Ashbaugh CD, Warren HB, Carey VJ, Wessels MR. 1998. Molecular analysis of the role of the group A streptococcal cysteine protease, hyaluronic acid capsule, and M protein in a murine model of human invasive soft-tissue infection. Journal of Clinical Investigation 102:550-560.
188. Abramson J, Adler J, Dunger J, Evans R, Green T, Pritzel A, Ronneberger O, Willmore L, Ballard AJ, Bambrick J, Bodenstein SW, Evans DA, Hung CC, O'Neill M, Reiman D, Tunyasuvunakool K, Wu Z, Zemgulyte A, Arvaniti E, Beattie C, Bertolli O, Bridgland A, Cherepanov A, Congreve M, Cowen-Rivers AI, Cowie A, Figurnov M, Fuchs FB, Gladman H, Jain R, Khan YA, Low CMR, Perlin K, Potapenko A, Savy P, Singh S, Stecula A, Thillaisundaram A, Tong C, Yakneen S, Zhong ED, Zielinski M, Zidek A, Bapst V, Kohli P, Jaderberg M, Hassabis D, Jumper JM. 2024. Accurate structure prediction of biomolecular interactions with AlphaFold 3. Nature 630:493-500.
189. Goddard TD, Huang CC, Meng EC, Pettersen EF, Couch GS, Morris JH, Ferrin TE. 2018. UCSF ChimeraX: Meeting modern challenges in visualization and analysis. Protein Science 27:14-25.

190. Meng EC, Goddard TD, Pettersen EF, Couch GS, Pearson ZJ, Morris JH, Ferrin TE. 2023. UCSF ChimeraX: Tools for structure building and analysis. Protein Science 32:e4792.
191. Pettersen EF, Goddard TD, Huang CC, Meng EC, Couch GS, Croll TI, Morris JH, Ferrin TE. 2021. UCSF ChimeraX: Structure visualization for researchers, educators, and developers. Protein Science 30:70-82.
192. Holm L. 2020. DALI and the persistence of protein shape. Protein Science 29:128-140.
193. Sarkar S, Heise MT. 2019. Mouse Models as Resources for Studying Infectious Diseases. Clinical Therapeutics 41:1912-1922.
194. Watson ME, Jr., Neely MN, Caparon MG. 2022. Animal Models of *Streptococcus pyogenes* Infection. *In* Ferretti JJ, Stevens DL, Fischetti VA (ed), *Streptococcus pyogenes*: Basic Biology to Clinical Manifestations, 2nd ed. University of Oklahoma Health Sciences Center, Oklahoma City (OK).
195. Troy T, Jekic-McMullen D, Sambucetti L, Rice B. 2004. Quantitative comparison of the sensitivity of detection of fluorescent and bioluminescent reporters in animal models. Molecular Imaging 3:9-23.
196. Oliveira JC, da Silva AC, Oliveira RA, Pereira VR, Gil LH. 2016. In vivo near-infrared fluorescence imaging of *Leishmania amazonensis* expressing infrared fluorescence protein (iRFP) for real-time monitoring of cutaneous leishmaniasis in mice. Journal of Microbiological Methods 130:189-195.
197. Avci P, Karimi M, Sadasivam M, Antunes-Melo WC, Carrasco E, Hamblin MR. 2018. In-vivo monitoring of infectious diseases in living animals using bioluminescence imaging. Virulence 9:28-63.
198. Francis KP, Joh D, Bellinger-Kawahara C, Hawkinson MJ, Purchio TF, Contag PR. 2000. Monitoring bioluminescent *Staphylococcus aureus* infections in living mice using a novel *luxABCDE* construct. Infection and Immunity 68:3594-600.
199. Francis KP, Yu J, Bellinger-Kawahara C, Joh D, Hawkinson MJ, Xiao G, Purchio TF, Caparon MG, Lipsitch M, Contag PR. 2001. Visualizing pneumococcal infections in the lungs of live mice using bioluminescent *Streptococcus pneumoniae* transformed with a novel gram-positive *lux* transposon. Infection and Immunity 69:3350-8.
200. Park HS, Cleary PP. 2005. Active and passive intranasal immunizations with streptococcal surface protein C5a peptidase prevent infection of murine nasal mucosa-associated lymphoid tissue, a functional homologue of human tonsils. Infection and Immunity 73:7878-86.
201. Georgel P, Crozat K, Lauth X, Makrantonaki E, Seltmann H, Sovath S, Hoebe K, Du X, Rutschmann S, Jiang Z, Bigby T, Nizet V, Zouboulis CC, Beutler B. 2005. A toll-like receptor 2-responsive lipid effector pathway protects mammals against skin infections with gram-positive bacteria. Infection and Immunity 73:4512-21.
202. Sheel M, Pandey M, Good MF, Batzloff MR. 2010. Correlation between bioluminescence and bacterial burden in passively protected mice challenged with a recombinant bioluminescent M49 group A *Streptococcus* Strain. Clinical and Vaccine Immunology 17:127-33.

203. Davis RW, IV, Eggleston H, Johnson F, Nahrendorf M, Bock PE, Peterson T, Panizzi P. 2015. In Vivo Tracking of Streptococcal Infections of Subcutaneous Origin in a Murine Model. Molecular Imaging and Biology 17:793-801.
204. Alam FM, Bateman C, Turner CE, Wiles S, Sriskandan S. 2013. Non-invasive monitoring of *Streptococcus pyogenes* vaccine efficacy using biophotonic imaging. PLoS One 8:e82123.
205. Lamb LE, Zhi X, Alam F, Pyzio M, Scudamore CL, Wiles S, Sriskandan S. 2018. Modelling invasive group A streptococcal disease using bioluminescence. BMC Microbiology 18:60.
206. Loh JM, Proft T. 2013. Toxin-antitoxin-stabilized reporter plasmids for biophotonic imaging of Group A *Streptococcus*. Applied Microbiology and Biotechnology 97:9737-45.
207. Loh JM, Proft T. 2014. Comparison of firefly luciferase and NanoLuc luciferase for biophotonic labeling of group A *Streptococcus*. Biotechnology Letters 36:829-34.
208. Armstrong BD, Herfst CA, Tonial NC, Wakabayashi AT, Zeppa JJ, McCormick JK. 2016. Identification of a two-component Class IIb bacteriocin in *Streptococcus pyogenes* by recombinase-based in vivo expression technology. Scientific Reports 6:36233.
209. Beard SJ, Salisbury V, Lewis RJ, Sharpe JA, MacGowan AP. 2002. Expression of lux genes in a clinical isolate of *Streptococcus pneumoniae*: using bioluminescence to monitor gemifloxacin activity. Antimicrobial Agents and Chemotherapy 46:538-42.
210. Kee JM, Oslund RC, Perlman DH, Muir TW. 2013. A pan-specific antibody for direct detection of protein histidine phosphorylation. Nature Chemical Biology 9:416-21.
211. Potel CM, Lin MH, Heck AJR, Lemeer S. 2018. Widespread bacterial protein histidine phosphorylation revealed by mass spectrometry-based proteomics. Nature Methods 15:187-190.
212. Yagüe P, Gonzalez-Quinoñez N, Fernánez-García G, Alonso-Fernández S, Manteca A. 2019. Goals and Challenges in Bacterial Phosphoproteomics. International Journal of Molecular Sciences 20:5678.
213. Köller T, Nelson D, Nakata M, Kreutzer M, Fischetti VA, Glocker MO, Podbielski A, Kreikemeyer B. 2008. PlyC, a novel bacteriophage lysin for compartment-dependent proteomics of group A streptococci. Proteomics 8:140-8.
214. Dame RT, Rashid FM, Grainger DC. 2020. Chromosome organization in bacteria: mechanistic insights into genome structure and function. Nature Reviews Genetics 21:227-242.
215. Lang B, Blot N, Bouffartigues E, Buckle M, Geertz M, Gualerzi CO, Mavathur R, Muskhelishvili G, Pon CL, Rimsky S, Stella S, Babu MM, Travers A. 2007. High-affinity DNA binding sites for H-NS provide a molecular basis for selective silencing within proteobacterial genomes. Nucleic Acids Research 35:6330-7.
216. Lautenschlager N, Schmidt K, Schiffer C, Wulff TF, Hahnke K, Finstermeier K, Mansour M, Elsholz AKW, Charpentier E. 2024. Expanding the genetic

toolbox for the obligate human pathogen *Streptococcus pyogenes*. Frontiers in Bioengineering and Biotechnology 12:1395659.
217. Alam FM, Turner CE, Smith K, Wiles S, Sriskandan S. 2013. Inactivation of the CovR/S virulence regulator impairs infection in an improved murine model of *Streptococcus pyogenes* naso-pharyngeal infection. PLoS One 8:e61655.
218. Merritt J, Qi F, Shi W. 2005. A unique nine-gene comY operon in *Streptococcus mutans*. Microbiology (Reading) 151:157-166.
219. Solheim M, La Rosa SL, Mathisen T, Snipen LG, Nes IF, Brede DA. 2014. Transcriptomic and functional analysis of NaCl-induced stress in *Enterococcus faecalis*. PLoS One 9:e94571.
220. Tesorero RA, Yu N, Wright JO, Svencionis JP, Cheng Q, Kim JH, Cho KH. 2013. Novel regulatory small RNAs in *Streptococcus pyogenes*. PLoS One 8:e64021.
221. Solioz M, Davies K. 1994. Operon of vacuolar-type Na(+)-ATPase of *Enterococcus hirae*. Journal of Biological Chemistry 269:9453-9.
222. Hebbeln P, Rodionov DA, Alfandega A, Eitinger T. 2007. Biotin uptake in prokaryotes by solute transporters with an optional ATP-binding cassette-containing module. Proceedings of the National Academy of Sciences of the United States of America 104:2909-14.
223. Gogos A, Jimenez JC, Chang JC, Wilkening RV, Federle MJ. 2018. A Quorum Sensing-Regulated Protein Binds Cell Wall Components and Enhances Lysozyme Resistance in *Streptococcus pyogenes*. Journal of Bacteriology 200:e00701-17.
224. Fritzer A, Senn BM, Minh DB, Hanner M, Gelbmann D, Noiges B, Henics T, Schulze K, Guzman CA, Goodacre J, von Gabain A, Nagy E, Meinke AL. 2010. Novel conserved group A streptococcal proteins identified by the antigenome technology as vaccine candidates for a non-M protein-based vaccine. Infection and Immunity 78:4051-67.
225. Chang JC, LaSarre B, Jimenez JC, Aggarwal C, Federle MJ. 2011. Two group A streptococcal peptide pheromones act through opposing Rgg regulators to control biofilm development. PLoS Pathogens 7:e1002190.
226. Cook LCC, Chatterjee N, Li Y, Andrade J, Federle MJ, Eichenbaum Z. 2019. Transcriptomic Analysis of *Streptococcus pyogenes* Colonizing the Vaginal Mucosa Identifies *hupY*, an MtsR-Regulated Adhesin Involved in Heme Utilization. mBio 10:e00848-19.
227. Sachla AJ, Ouattara M, Romero E, Agniswamy J, Weber IT, Gadda G, Eichenbaum Z. 2016. *In vitro* heme biotransformation by the HupZ enzyme from Group A *Streptococcus*. Biometals 29:593-609.
228. Lyles KV, Thomas LS, Ouellette C, Cook LCC, Eichenbaum Z. 2022. HupZ, a Unique Heme-Binding Protein, Enhances Group A *Streptococcus* Fitness During Mucosal Colonization. Frontiers in Cellular and Infection Microbiology 12:867963.
229. Valdes KM, Sundar GS, Vega LA, Belew AT, Islam E, Binet R, El-Sayed NM, Le Breton Y, McIver KS. 2016. The *fruRBA* Operon Is Necessary for Group A Streptococcal Growth in Fructose and for Resistance to Neutrophil Killing during Growth in Whole Human Blood. Infection and Immunity 84:1016-31.

230. Sun Y, Veseli IA, Vaillancourt K, Frenette M, Grenier D, Pombert JF. 2019. The bacteriocin from the prophylactic candidate *Streptococcus suis* 90-1330 is widely distributed across S. suis isolates and appears encoded in an integrative and conjugative element. PLoS One 14:e0216002.
231. Grifantini R, Toukoki C, Colaprico A, Gryllos I. 2011. Peroxide stimulon and role of PerR in group A *Streptococcus*. Journal of Bacteriology 193:6539-6551.
232. Varhimo E, Savijoki K, Jalava J, Kuipers OP, Varmanen P. 2007. Identification of a novel streptococcal gene cassette mediating SOS mutagenesis in *Streptococcus uberis*. Journal of Bacteriology 189:5210-22.
233. Keogh RA, Zapf RL, Frey A, Marino EC, Null GG, Wiemels RE, Holzschu DL, Shaw LN, Carroll RK. 2021. *Staphylococcus aureus* Trigger Factor Is Involved in Biofilm Formation and Cooperates with the Chaperone PpiB. Journal of Bacteriology 203:e00681-20.
234. Hermans PW, Adrian PV, Albert C, Estevao S, Hoogenboezem T, Luijendijk IH, Kamphausen T, Hammerschmidt S. 2006. The streptococcal lipoprotein rotamase A (SlrA) is a functional peptidyl-prolyl isomerase involved in pneumococcal colonization. Journal of Biological Chemistry 281:968-76.
235. Ünal CM, Steinert M. 2014. Microbial peptidyl-prolyl cis/trans isomerases (PPIases): virulence factors and potential alternative drug targets. Microbiology and Molecular Biology Reviews 78:544-71.
236. Reid SD, Montgomery AG, Voyich JM, DeLeo FR, Lei B, Ireland RM, Green NM, Liu M, Lukomski S, Musser JM. 2003. Characterization of an extracellular virulence factor made by group A *Streptococcus* with homology to the *Listeria monocytogenes* internalin family of proteins. Infection and Immunity 71:7043-52.
237. Abedeera SM, Jayalath KS, Xie J, Rauff RM, Abeysirigunawardena SC. 2023. Pseudouridine Synthase RsuA Confers a Survival Advantage to Bacteria under Streptomycin Stress. Antibiotics (Basel) 12:1447.
238. Mashburn-Warren L, Morrison DA, Federle MJ. 2012. The cryptic competence pathway in *Streptococcus pyogenes* is controlled by a peptide pheromone. Journal of Bacteriology 194:4589-600.
239. Gilmore MS, Salamzade R, Selleck E, Bryan N, Mello SS, Manson AL, Earl AM. 2020. Genes Contributing to the Unique Biology and Intrinsic Antibiotic Resistance of *Enterococcus faecalis*. mBio 11:e02962-20.
240. Jalal N, Lee SF. 2020. The MsrAB reducing pathway of *Streptococcus gordonii* is needed for oxidative stress tolerance, biofilm formation, and oral colonization in mice. PLoS One 15:e0229375.
241. Saleh M, Bartual SG, Abdullah MR, Jensch I, Asmat TM, Petruschka L, Pribyl T, Gellert M, Lillig CH, Antelmann H, Hermoso JA, Hammerschmidt S. 2013. Molecular architecture of *Streptococcus pneumoniae* surface thioredoxin-fold lipoproteins crucial for extracellular oxidative stress resistance and maintenance of virulence. EMBO Molecular Medicine 5:1852-70.
242. Farshchi Andisi V, Hinojosa CA, de Jong A, Kuipers OP, Orihuela CJ, Bijlsma JJ. 2012. Pneumococcal gene complex involved in resistance to extracellular oxidative stress. Infection and Immunity 80:1037-49.

243. Stewart LJ, Ong CY, Zhang MM, Brouwer S, McIntyre L, Davies MR, Walker MJ, McEwan AG, Waldron KJ, Djoko KY. 2020. Role of Glutathione in Buffering Excess Intracellular Copper in *Streptococcus pyogenes*. mBio 11:e02804-20.
244. Liu M, Zhu H, Zhang J, Lei B. 2007. Active and passive immunizations with the streptococcal esterase Sse protect mice against subcutaneous infection with group A streptococci. Infection and Immunity 75:3651-7.
245. Liu M, Zhu H, Li J, Garcia CC, Feng W, Kirpotina LN, Hilmer J, Tavares LP, Layton AW, Quinn MT, Bothner B, Teixeira MM, Lei B. 2012. Group A *Streptococcus* secreted esterase hydrolyzes platelet-activating factor to impede neutrophil recruitment and facilitate innate immune evasion. PLoS Pathogens 8:e1002624.
246. Johnson MD, Echlin H, Dao TH, Rosch JW. 2015. Characterization of NAD salvage pathways and their role in virulence in *Streptococcus pneumoniae*. Microbiology (Reading) 161:2127-36.
247. Proft T, Fraser JD. 2016. Streptococcal Superantigens: Biological properties and potential role in disease. *In* Ferretti JJ, Stevens DL, Fischetti VA (ed), *Streptococcus pyogenes*: Basic Biology to Clinical Manifestations. University of Oklahoma Health Sciences Center, Oklahoma City (OK).
248. Biagini M, Garibaldi M, Aprea S, Pezzicoli A, Doro F, Becherelli M, Taddei AR, Tani C, Tavarini S, Mora M, Teti G, D'Oro U, Nuti S, Soriani M, Margarit I, Rappuoli R, Grandi G, Norais N. 2015. The Human Pathogen *Streptococcus pyogenes* Releases Lipoproteins as Lipoprotein-rich Membrane Vesicles. Molecular and Cellular Proteomics 14:2138-49.
249. Zhu B, Song L, Kong X, Macleod LC, Xu P. 2018. A Novel Regulator Modulates Glucan Production, Cell Aggregation and Biofilm Formation in *Streptococcus sanguinis* SK36. Frontiers in Microbiology 9:1154.
250. Zhang MM, Ong CL, Walker MJ, McEwan AG. 2016. Defence against methylglyoxal in Group A *Streptococcus*: a role for Glyoxylase I in bacterial virulence and survival in neutrophils? Pathogens and Disease 74:ftv122.
251. Liao S, Bitoun JP, Nguyen AH, Bozner D, Yao X, Wen ZT. 2015. Deficiency of PdxR in Streptococcus mutans affects vitamin B6 metabolism, acid tolerance response and biofilm formation. Molecular Oral Microbiology 30:255-68.
252. Zhu L, Lin J, Kuang Z, Vidal JE, Lau GW. 2015. Deletion analysis of *Streptococcus pneumoniae* late competence genes distinguishes virulence determinants that are dependent or independent of competence induction. Molecular Microbiology 97:151-65.
253. Henningham A, Dohrmann S, Nizet V, Cole JN. 2015. Mechanisms of group A *Streptococcus* resistance to reactive oxygen species. FEMS Microbiol Reviews 39:488-508.
254. Bates CS, Montañez GE, Woods CR, Vincent RM, Eichenbaum Z. 2003. Identification and characterization of a *Streptococcus pyogenes* operon involved in binding of hemoproteins and acquisition of iron. Infection and Immunity 71:1042-1055.

255. Montañez GE, Neely MN, Eichenbaum Z. 2005. The streptococcal iron uptake (Siu) transporter is required for iron uptake and virulence in a zebrafish infection model. Microbiology (Reading) 151:3749-3757.
256. Zhu H, Liu M, Lei B. 2008. The surface protein Shr of Streptococcus pyogenes binds heme and transfers it to the streptococcal heme-binding protein Shp. BMC Microbiology 8:15.
257. Hynes W, Sloan, M. 2016. Secreted Extracellular Virulence Factors. *In* Vincent A. Fischetti DLS, Joseph J. Ferretti (ed), *Streptococcus pyogenes*: Basic Biology to Clinical Manifestations. University of Oklahoma Health Sciences Center, Oklahoma City (OK).
258. Upton M, Tagg JR, Wescombe P, Jenkinson HF. 2001. Intra- and interspecies signaling between *Streptococcus salivarius* and *Streptococcus pyogenes* mediated by SalA and SalA1 lantibiotic peptides. Journal of Bacteriology 183:3931-8.
259. Wood DN, Weinstein KE, Podbielski A, Kreikemeyer B, Gaughan JP, Valentine S, Buttaro BA. 2009. Generation of metabolically diverse strains of *Streptococcus pyogenes* during survival in stationary phase. Journal of Bacteriology 191:6242-52.