ABSTRACT

Title of Dissertation: FUNCTIONAL CHARACTERIZATION OF THE LINK BETWEEN CARBOHYDRATE METABOLISM AND THE PATHOGENESIS OF THE INVASIVE M1T1 GROUP A STREPTOCOCCUS

Kayla Maureen Valdes, Doctor of Philosophy, 2016

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

The Group A Streptococcus (GAS), or *Streptococcus pyogenes*, is a strict human pathogen that colonizes a variety of sites within the host. Infections can vary from minor and easily treatable, to life-threatening, invasive forms of disease. In order to adapt to niches, GAS utilizes environmental cues, such as carbohydrates, to coordinate the expression of virulence factors. Research efforts to date have focused on identifying how either components of the phosphoenolpyruvate-phosphotransferase system (PTS) or global transcriptional networks affect the regulation of virulence factors, but not the synergistic relationship between the two. The present study investigates the role of a putative PTS-fructose operon encoded by *fruRBA* and its role in virulence in the M1T1 strain 5448. Growth in fructose resulted in induction of *fruRBA*. RT-PCR showed that *fruRBA* formed an operon, which was repressed by FruR in the absence of fructose. Growth and carbon

utilization profiles revealed that although the entire *fruRBA* operon was required for growth in fructose, FruA was the main fructose transporter. The ability of both Δ*fruR* and Δ*fruB* mutants to survive in whole human blood or neutrophils was impaired. However, the phenotypes were not reproduced in murine whole blood or in a mouse intraperitoneal infection, indicating a human-specific mechanism. While it is known that the PTS can affect activity of the Mga virulence regulator, further characterization of the mechanism by which sugars and its protein domains affect activity have not been studied. Transcriptional studies revealed that the core Mga regulon is activated more in a glucose-rich than a glucose-poor environment. This activation correlates with the differential phosphorylation of Mga at its PTS regulatory domains (PRDs). Using a 5448 *mga* mutant, transcriptome studies in THY or C media established that the Mga regulon reflects the media used. Interestingly, Mga regulates phage-encoded DNases in a low glucose environment. We also show that Mga activity is dependent on C-terminal amino acid interactions that aid in the formation of homodimers. Overall, the studies presented sought to define how external environmental cues, specifically carbohydrates, control complex regulatory networks used by GAS, contribute to pathogenesis, and aid in adaptation to various nutrient conditions encountered.

FUNCTIONAL CHARACTERIZATION OF THE LINK BETWEEN CARBOHYDRATE METABOLISM AND THE PATHOGENESIS OF THE INVASIVE M1T1 GROUP A STREPTOCOCCUS

by

Kayla Maureen Valdes

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

Advisory Committee:
Professor Kevin S. McIver, Chair
Professor Najib M. El-Sayed
Associate Professor Wade C. Winkler
Associate Professor Vincent T. Lee
Associate Professor Daniel C. Nelson, Dean's Representative

Dedication

This thesis is dedicated to my sister, Dr. Thelma Valdes. You were the first person to show me what doing research was really like. I cannot thank you enough for bringing me into lab when I was in high school and teaching me how to pipet, autoclave, and passage cells. Thank you for all your constant guidance throughout my whole life. I would not be the person I am today both personally and professionally, without you. You have always been there whenever I need your wisdom and a teaching moment. Thank you for all that you are to my life. You are my mentor, my sister, and friend. I love you!

Acknowledgements

Quiero agradeser a mis padres por el apoyo, ánimo y motivación que me han siempre dado para lograr mis objetivos en la vida. Ustedes siempre han promovido la educación y me han inculcado la motivación para seguir adelante y buscar éxito en la vida. Siempre han estado a mi lado cuando estoy en situaciónes difíciles y me recuerdan que debo ser paciente. Me acuerdan que lo que pasa es siempre para mejor. Gracias por venir a visitarme tan frecuentemente, que simpre ha sido importante para mi tenerlos cerca. No podría estar más orgullosa de donde vengo y tenerlos como mis padres.

I would like to thank Alan, my best friend, partner and fiancée. Without you, I'm not sure how I would have emerged from this degree a sane person. You have been my rock throughout this whole process, and for always being there to pull me out of the darkness and be my remedy at the end of a tough day. Thank you for having the patience to listen to me vent about the struggles of science. Also, thank you for always being inquisitive about my research when I want to discuss a science topic with you, it really means a lot to me. I am so happy to have found you, and I can't wait to begin our life together in a few months.

I would like to thank my advisor, Dr. Kevin McIver, for his guidance and training to become a scientist. Thank you for your willingness to discuss ideas and helping me turn the downsides of research into learning experiences. I would like to thank my graduate committee Dr. Najib El-Sayed, Dr. Vincent Lee, Dr. Wade Winkler, and Dr. Daniel Nelson for their thoughtful insight on my project.

I would like to thank all the members of the McIver lab, past and present. I would like to especially thank Leigh, Yoann, and Surya. Leigh thank you for being a great undergrad to work with, I appreciated your flexibility when it came to your constantly changing project. Yoann, you have been a great mentor to me, and it's been wonderful being able to work with you on a daily basis. You have taught me the foundations of molecular biology and bacterial genetics. I appreciate your enthusiasm for science, and your willingness to always listen to my ideas and give me suggestions on my projects. Surya, you were a great addition to our lab and I could not have asked for a better peer to work alongside for the past three years. I enjoyed our lab-DJing sessions, always keeping up on the latest rap music and talking about wedding planning struggles. Thank you both for always listening to me complain and, of course, for the daily trips to Stamp; sometimes they were the best moments of the day and were always a much needed break!

Table of Contents

List of Tables

List of Figures

List of Appendicies

List of Abbreviations

α	alpha
aa	amino acid
Ala, A	alanine
Asn, N	asparagine
Asp	aspartic acid
ASPGN	acute post-streptococcal glomerulonephritis
ARF	acute rheumatic fever
ADP	adenosine diphosphate
Asp	aspartate
ATP	adenosine triphosphate
β	beta
B. anthracis	*Bacillus antrhacis*
bp	base pair
BSA	bovine serum albumin
B. subtilis	*Bacillus subtilis*
cAMP	cyclic AMP
CcpA	catabolite control protein A
CCR	carbon catabolite repression
CDM	chemically defined media
cDNA	complementary DNA
CFU	colony forming units
CMD	conserved Mga-like domain
CovRS	control of virulence two-component system
C. perfringens	*Clostridium perfringens*
CRP	cAMP receptor protein
cre	catabolite response element
CTD	C-terminal domain
DEPC	diethyl pyrocarbonate
dH_2O	distilled water

DNA	deoxyribonucleic acid
DNase	deoxyribonuclease
dNTPs	deoxyribonucleic acid triphosphate monomers
DTT	dithiothreiotol
EI	enzyme I
EII	enzyme II
E. coli	*Escherichia coli*
E. faecalis	*Enterococcus faecalis*
emm	gene encoding M protein
EMSA	electrophoretic mobility shift assay
Fba	fibronectin binding protein
FBP	fructose-1,6-bisphosphate
g	gravity
g	gram
GAS	group A streptococcus
gDNA	genomic DNA
Glu, E	glutamine
GRAB	protein G-related α_2-macroglobulin-binding protein
grm	gene regulated by Mga
H_2O_2	hydrogen peroxide
HA	hyaluronic acid
His, H	histidine
HK	histidine kinase
Hpr	histidine phosphocarrier
HprK/P	Hpr kinase/phosphatase
HTH	helix-turn-helix
Ig	immunoglobulin
i.p.	intraperitoneal
Ile	isoleucine
kb	kilobase
kDa	kilo-Dalton

L	liter
LB	Luria-Bertani medium
Leu, L	leucine
L. monocytogenes	*Listeria monocytogenes*
LTA	lipotechoic acid
Lys, K	lysine
M	molar
m	milli
µ	micro
MAC	membrane attack complex
MBS	Mga binding site
MHC	major histocompatability complex
Met, M	methionine
MF	multiplication factor
Mga	multiple gene regulator of the group A streptococcus
Mrp	M-related protein
MSCRAMM	microbial surface component recognizing adhesive matrix molecules
NADH	nicotinamide adenine dinucleotide
NEB	New England Biolabs
NET	neutrophil extracellular trap
NF	necrotizing fasciitis
Ni-NTA	nickel-nitriloacetic acid
n	nano
Nra	negative regulator of the group A streptococcus
OF	opacity factor
ORF	open reading frame
ov	omnilog values
PAMP	pathogen associated molecular pattern
PANDAS	pediatric autoimmune neuropsychiatric disorders associated with streptococcal infections

PCR	polymerase chain reaction
PCVR	PRD-containing virulence regulators
PEP	phosphoenolpyruvate
Phe, F	phenylalanine
PMN	polymorphonuclear leuokocyte
PNK	polynucleotide kinase
PRD	phosphoenolpyruvate phosphotransferase regulatory domain
PTS	phosphoenolpyruvate phosphotransferase system
q	quantitative
RALPs	RofA-like proteins
RD	receiver domain
RNA	ribonucleic acid
RNA-Seq	RNA-sequencing
RR	response regulator
RivR	RofA-like protein IV
RofA	regulator of protein F
RT	reverse transcriptase
S. aureus	*Staphylococcus aureus*
SC	Sydenham's chorea
SDS	sodium dodecyl sulfate
Ser, S	serine
S. gordonii	*Streptococcus gordonii*
S. iniae	*Streptococcus iniae*
S. mutans	*Streptococcus mutans*
S. pneumoniae	*Streptococcus pneumoniae*
S. pyogenes	*Streptococcus pyogenes*
S. suis	*Streptococcus suis*
SclA	streptococcal collagen-like protein
ScpA	streptococal C5a peptidase
Sic	streptococcal inhibitor of complement
Ska	streptokinase

Slo	streptolysin O
SLS	streptolysin S
SOF	serum opacity factor
Spe	streptococcal pyogenic exotoxin
Srv	streptococcal regulator of virulence
STSS	streptococcal toxic shock syndrome
TBE	tris/borate/EDTA
TCA	tricarboxylic acid cycle
TCS	two-component system
TE	tris/EDTA
THY	Todd-Hewitt yeast extract medium
TraSH	transposon-site hybridization
TrxSR	two-component regulatory system X
Tyr, Y	tyrosine
UV	ultraviolet
Val, V	valine
WT	wild type
wHTH	winged helix-turn-helix
ZYP	N-Z-amine/yeast extract/phosphate

Chapter 1: Introduction

The Group A Streptococcus (GAS) or *Streptococcus pyogenes* is a Gram-positive pathogen that can colonize a variety of tissues in the human host, resulting in a wide range of invasive or non-invasive diseases. Worldwide, GAS is estimated to cause over 700 million cases of non-invasive infections such as pharyngitis and impetigo yearly (1). In contrast, invasive forms of disease, including necrotizing fasciitis (NF), puerperal sepsis, and streptococcal toxic shock syndrome (STSS) (2, 3) are much more life threatening. GAS infections can also lead to nonsuppurative sequelae such as acute rheumatic fever (ARF) and acute poststreptococcal glomerulonephritis (3) that combined lead to over half a million deaths worldwide each year (1). While a licensed vaccine is not readily available for prevention of GAS disease, self-limiting GAS infections are still treatable with penicillin. However, treatment of invasive GAS infections requires intensive care, and treatment with antibiotics is not always successful.

GAS can successfully colonize and adapt to variety of niches in the human host by coordinating environmental cues and its nutritional status with global transcriptional networks. One of the major factors in the host environment that GAS can sense is the availability of nutrients such as carbohydrates. Given that GAS is a fastidious fermentative organism, it is dependent on the phosphoenolpyruvate-phosphotransferase system (PTS) as the primary source for carbon uptake and utilization for the cell. Interestingly, metabolic operons and genes under carbon catabolite repression (CCR) are induced *in vivo* and are often required for full

virulence in GAS (4-9), which directly links GAS carbohydrate metabolism and virulence factor production. In addition, since GAS does not encode for alternate sigma factors to coordinately regulate expression of genes, GAS utilizes global transcriptional regulators to coordinate transcriptome changes, including 13 two-component systems (TCS) and several "stand-alone" regulators (e.g., RofA-like proteins [RALPs], Rgg, and Mga) that control virulence gene expression in response to environmental stimuli (10).

This study begins with the characterization of a putative PTS-fructose operon encoded by the *fruRBA* and its role in virulence. The expression of the *fruRBA* operon was highly induced in the presence of fructose as determined by RNA-Seq. FruR acted as a repressor of the operon likely in conjunction with carbon catabolite protein A (CcpA), and FruA was the main transporter of fructose for GAS. *fruR* and *fruB* were also identified as being important for GAS survival in whole human blood and for evasion of human neutrophils. However, neither showed a defect during growth in whole mouse blood nor intraperitoneal (*i.p.*) infection in mice. The results of this study showed that FruR and FruB, but not FruA, were important for GAS survival in a human-specific environment.

The multiple gene regulator of GAS, or Mga, is ubiquitous in all sequenced strains of GAS (11). Mga has been shown to be critical for multiple *in vivo* phenotypes and is critical for genetic fitness in *ex vivo* screens (11-16). Several studies have shown that Mga and its regulon are intimately linked to the PTS and sugar utilization. For example, CcpA was found to regulate *mga* expression via an upstream *cre* site (17). M protein, a major virulence factor that is directly regulated by

Mga, exhibited increased expression in medium with glucose as the sole carbon source in comparison to alternative sugar sources (18). The most salient piece of evidence demonstrating the link between sugar utilization and Mga lies within the recent studies that show the PTS of GAS to directly phosphorylate and affect the activity of Mga both *in vitro* (19) and *in vivo* (20).

The second focus of my research assessed how glucose availability affects Mga activity and its regulon. We demonstrated the mechanism by which Mga alters its activity and regulon expression was a direct response to glucose through phosphorylation on the phosphohistidines, presumably via the PTS. We used RNA-Seq to identify transcriptomic differences between the Mga regulon under either a high or low glucose environment. Interestingly only 19 genes were found to overlap between the high and low glucose regulons, most of which were part of the core Mga regulon. Mga was found to be necessary for expression of all three phage associated nucleases under low glucose conditions, but this could not be correlated to an *in vivo* phenotype. Moreover, the majority phenotypes that were associated with a low glucose environment were complemented upon supplementation of glucose in the growth medium. Thus, the results from this study show that Mga activity is directly linked to glucose availability and this is correlated to the phosphorylation of Mga through its PTS-regulatory domain (PRD) histidines.

PTS phosphorylation of Mga is hypothesized to impact the dimerization state and the ability to activate virulence gene transcription (19). The final study of this dissertation was designed to identify individual amino acid residues within the C-terminus of Mga in order to characterize their importance in Mga dimerization and

activity. Data generated from a bacterial two-hybrid screen was used to select a representative group of mutations that showed a putative defect in dimerization. Mga was also modeled to the crystal structure of AtxA, a functional homolog from *B. anthracis,* to identify additional residues important for dimerization. Interestingly, only two mutations that were identified resulted in expression of a full-length protein and a functional defect in transcriptional activation. Overall, the results of these experiments demonstrated that mutations in the C-terminal tail of Mga are highly unstable because they most likely contribute to proper protein conformation, making it difficult to assess individual amino acid contributions to dimerization and activity.

The results presented within this dissertation emphasize that the PTS and carbohydrate metabolism, through the ability to stimulate signal transduction and regulation of global transcriptional networks, play an important role in the pathogenesis for GAS. Furthermore, the C-terminal of Mga is crucial for not only dimerization but for overall protein stability, confirming the critical role for this domain.

Chapter 2: Literature Review

Historical Perspectives

The Group A Streptococcus (GAS, *Streptococcus pyogenes*) is a strict human pathogen that has led to several disease outbreaks throughout human history. Skulls dating back to 8000-6500 BC presented the first evidence of GAS infection (21). The first detailed epidemic of GAS infection was documented to have occurred during the Age of Pericles (430 BC) in ancient Greece; those afflicted were documented as having a red rash over their body, which today correlates with scarlet fever (22).

With the advent of modern microbiology and the microscope, GAS was first described as a spherical chain-growing microorganism found at the site of infection, allowing for its identification as the etiologic agent of disease. However, it was not until 1874 that Billroth named these organisms *Streptococcus*; the word streptococcus itself contains two roots: *strepto*, meaning twisted or chain and *kokhos*, meaning berry or seed (21). Concurrently, streptococci began to be classified based on the type of disease they caused. In 1903, Hugo Schottmüler began categorizing streptococci based on their ability to hemolyze red blood cells (21). In 1933, Rebecca Lancefield developed a classification system for β-hemolytic streptococci that is based on the type of carbohydrate present on the cell surface (21, 23). All β-hemolytic streptococci that belonged to the serotype A group in Lancefield's classification included the majority of the pathogenic strains that infected humans, and thus these were collectively renamed as *Streptococcus pyogenes*.

Classification of GAS

Standard microbiology and growth requirements

GAS is non-motile, Gram-positive cocci that can form chains of varying length. GAS is catalase and oxidase negative, and is non-spore forming. Typically, colonies are opaque in color, and have a variety of appearances, ranging from flat, dry, glossy, matte or mucoid. The colonies that are mucoid appear large, wet, and glistening, which is attributed to an increased production of hyaluronic acid capsule (24). GAS can be distinguished from other streptococci by the type of hemolysis present on a 5% blood agar plate, where it exhibits β-hemolysis (total lysis of red blood cells). GAS is a fastidious organism that relies on the fermentation of sugars for growth, as it does not encode for the necessary enzymes of a functional tricarboxylic acid cycle (TCA). In the laboratory, GAS can be grown in nutrient-rich media such as Todd-Hewitt yeast extract (THY) either statically at 37^{o}C in ambient air or 5-10% CO_2 (25, 26). GAS is considered auxotrophic for multiple amino acids, and therefore requires all amino acids to be present in its growth media with the exception of alanine, asparagine, aspartic acid, and glutamic acid (27). GAS can also grow in a chemically defined media (CDM) supplemented with all of the necessary amino acids (28).

Lancefield grouping

Rebecca Lancefield established the Lancefield serological differentiation of streptococci in 1933 by splitting hemolytic streptococci into five different groups (A, B, C, D, E) using an acid-precipitin reaction. In brief, serum was collected from

rabbits immunized with 106 different strains of hemolytic streptococci. The 106 strains were washed with hot acid in order to extract their C-substance, which is now known to be a carbohydrate. The acid extracts were incubated with the anti-sera to all other streptococcal strains and the formation of precipitate was observed when the anti-sera of one strain reacted with cell wall extract of a strain within the same group (23). This agglutination assay was later adapted and employed during the creation of the rapid strep test for pharyngitis that is used in the clinical laboratory (29).

M, T, and OF typing

The Group A Streptococcus can be broken down further into serotypes based on the presence of M and T proteins. The M protein is present on the cell surface of all GAS strains, and historically is the primary way to categorize different GAS strains. Much like the detection of the "C-substance", the M protein is detected by a precipitation of extracted M protein using M type-specific antisera prepared against reference type strains (30, 31). Currently, there are over 200 different M types, based on sequence typing of the *emm* gene (32). Recently, a new adjunct method of classification, known as the "*emm*-cluster system" has been developed, which organizes the 223 *emm*-types into 48 distinct *emm*-clusters based on the phylogenetic analysis of the whole M protein and its structural and binding properties (33). The T typing assay, which detects the presence of pili, is also performed as a slide agglutination test (34). T typing can be used in conjunction or as a replacement for M typing as certain T types correlate to M types (35).

The presence of a serum opacity factor (OF) can also be utilized to subdivide GAS into two classes. Approximately half of GAS produces OF, which is a

lipoproteinase that causes various types of serum to increase in opacity. Antibodies against OF are used in the OF-inhibition test (36), and these results can be used to divide GAS into two different categories, Class I and Class II. The OF type is closely correlated to the M type (37), where Class I typically contains an extended conserved surface region of M protein and cannot opacify serum. Class II, however, lacks this conserved region of the M protein but contains the ability to opacify serum (3).

Disease Manifestations

The Group A Streptococcus colonizes many different tissues in the human host, eliciting a wide range of disease. Worldwide, GAS is estimated to cause over 700 million cases of non-invasive infections each year. In addition, invasive GAS diseases and nonsuppurative sequelae lead to over half a million deaths worldwide each year making GAS the ninth leading cause of mortality (1). This section will focus on the hallmarks of different GAS diseases along with treatments currently used to treat each type of disease.

Non-invasive Infections

Pharyngitis

Pharyngitis is one of the most common forms of GAS infections. In 2005, the World Health Organization (WHO) estimated that there are over 616 million cases of GAS pharyngitis each year (1). Symptoms include an abrupt sudden onset of severe throat pain accompanied by tonsilar exudates, a fever greater than 38^{o}C that usually lasts 3-5 days, malaise, and headache (38, 39). Pharyngitis usually affects school-

aged children, infecting approximately 15% of children in developed countries, and incidence is estimated to be five to ten times greater in under developed countries (1).

Clinical procedures to identify GAS attributed pharyngitis are as follows: swabbing the patients' throat, culturing the swab on blood agar, and observing the presence of β-hemolytic colonies within 24-48 hours. Since streptococcal pharyngitis symptoms can be similar to viral pharyngitis, a Centor score can be used to evaluate a patient presenting symptoms. This test uses a point system that determines the presence or absence of symptoms including presence of cough, swollen or tender anterior cervical nodes, temperature greater than 38^{o}C, tonsilar exudates, and the age of the patient. Patients with a score of 0 to 1 are at very low risk for streptococcal pharyngitis and do not require testing. Patients with a score of 2 or 3 should be tested using a rapid strep test or throat culture; positive results warrant antibiotic therapy. Patients with a score of 4 or higher are at high risk of having contracted streptococcal-associated pharyngitis, and empiric treatment may be considered (39).

Although pharyngitis infections are self-limiting and usually resolve themselves within a few days, treatment with antibiotics is often recommended, not only to decrease the risk of spreading the disease and decrease the amount of time presenting symptoms, but also to avoid the risk of suppurative complications such as scarlet fever and acute rheumatic fever (ARF) (both discussed below). The antibiotics that are prescribed for treatment are usually penicillin and penicillin derivatives, but macrolides and cephalosporins can be prescribed for those who have allergic reactions to penicillin (38, 39).

Scarlet Fever

Scarlet fever, or scarlatina, is usually associated with pharyngitis. Today, scarlet fever is a mild disease with a significantly decreased morbidity and mortality rate, which is a stark contrast to the devastating infections seen in the 19th century (40). During the 1800s and prior, scarlet fever was more malignant due to the lack of antibiotic treatment and manifested in three types of disease: atactic, hemorrhagic, and anginose (40). Today, symptoms of scarlet fever often present themselves as a deep red, finely papular, erythematous rash, otherwise called "strawberry tongue" in conjunction with pharyngitis (32). GAS strains that produce scarlet fever often carry one or more of the pyrogenic exotoxins A, B, and C that are either phage- or chromosomally-encoded (3).

Impetigo

Impetigo or pyoderma is estimated to affect 111 million children under the age of 15 in under developed countries at any one time (1). Infections are usually seasonal, occurring mostly in the summer months and in a temperate climate, and are often linked to the development of acute glomerulonephritis (discussed below) (3). Symptoms include pustules that gradually enlarge and rupture, forming thick, honey-colored scabs (32). Lesions often occur on exposed skin such as the face and extremities (41). Local symptoms are minimal and include itching and an erythematous halo around the lesion (42). The disease is spread through direct skin contact and most commonly affects children with poor hygiene that are living in crowded living conditions. Treatment for impetigo includes administration of penicillin and penicillin derivatives.

Cellulitis and Erysipelas

Cellulitis and Erysipelas are caused when GAS enters the skin through abrasions and other types of lesions and penetrate the epidermis. Erysipelas, or St Anthony's fire, is a distinct form of cellulitis where the superficial layer of the skin and cutaneous lymphatics are affected, as opposed to cellulitis that affects subcutaneous tissues (3, 41). Erysipelas is characterized by an abrupt onset of a fiery-red, salmon, or scarlet rash and well-defined margins particularly on the nasolabial fold. There is rapid progression and intense pain also associated with this disease. Erysipelas mostly affects elderly people, and recurring infections are likely (43).

Symptoms associated with cellulitis include local swelling, erythema and pain. The skin becomes indurated and unlike the brilliant redness of erysipelas, turns a pinkish color. In addition, the clear demarcation that is present between infected and uninfected skin is not present in cellulitis as it is in erysipelas. It is important to diagnose GAS cellulitis correctly, as many other bacteria (e.g., *S. aureus*) can cause similar symptoms and the correct course of antibiotics must be given to treat the disease (41).

Invasive Infections

Puerperal fever

During the 18th century, puerperal fever or childbed fever had a mortality rate of 50%, which made it one of the most deadly diseases of its time. Historically, puerperal fever was contracted during pregnancy or abortion when an attending physician infected women post childbirth. During infection, GAS invades and colonizes the endometrium and surrounding tissue as well as the lymphatics and

bloodstream and produces severe abdominal pain (43). The Hungarian physician, Semmelweiss, introduced the concept of septic technique and in conjunction with the use of antibiotics, the number of cases of childbed fever has been greatly reduced and is now rarely fatal (21).

Streptococcal Toxic Shock Syndrome

Streptococcal toxic shock syndrome (STSS) is associated with bacteremia (discussed below) (44). It can affect persons of all ages and most do not have a predisposing disease, although most cases occur sporadically in individuals who are between the ages of 20 and 50 years old (43). The abrupt onset of severe pain, often in an extremity, is a common initial symptom. This is followed by fever that can lead to a coma, although some patients have hypothermia secondary to shock. Other symptoms may include deep soft-tissue infection, shock, multiple organ failure, and confusion. Eighty percent of patients have clinical signs of soft tissue infection, such as localized swelling and erythema, that progresses to necrotizing fasciitis or myositis in 70% of patients and requires surgical debridement, fasciotomy or amputation (44). Penicillin can be used to treat STSS, however there is a decrease in efficacy at clearing infection due to the invasiveness of this disease.

Bacteremia

Streptococcal-associated bacteremia, or the presence of bacteria in the blood stream, is historically most common in the very young and elderly. Although rare, scarlet fever can be a predisposing factor, occurring previously in 0.3% of bacteria presenting patients. In children, predisposing factors including burns, infection with Varicella zoster, malignant neoplasm, and immunosuppression increase the

probability of contracting bacteremia. For the elderly, susceptibility to bacteremia increases after skin infections, diabetes, peripheral vascular disease, malignancy, and corticosteroid use (40). Mortality rates between 35-80% have been reported. Since the mid-1980's, the 14 to 40 year old demographic has seen a surge in cases of bacteremia due to drug use (40). STSS presents itself in conjunction to bacteremia in 30-70% of cases (44).

Necrotizing Fasciitis

Hippocrates first reported necrotizing fasciitis, also known as the "flesh eating disease", in the fifth century BCE (45). The first modern report that gives a detailed clinical description of the disease is from a U.S. Army surgeon who practiced during the American Civil War (46). The hallmark of necrotizing fasciitis is a rapidly progressing infection of the subcutaneous skin, deep soft tissues, and muscle. The infection can spread from being a small lesion to a highly lethal disease within 24 hours (40, 46). In addition to surgical intervention to remove large portions of affected tissues, aggressive use of intravenous antibiotics with volume resuscitation is crucial to successful therapy (46). However, even with aggressive surgical amputation of the affected area, mortality rates are often greater than 50%. This disease can affect patients of any age, but mortality rates increases in those who are 50 or older (40).

Sequelae

Acute Rheumatic Fever

Acute rheuamatic fever (ARF) was first described by Guillaume and Sydenham in the 17th century. However, Cheadle was the first to provide a clear

definition of the symptoms associated with ARF in 1898. It wasn't until the 1930s, that Coburn first showed convincingly that the development of rheumatic fever is correlated with previous GAS infection (21). ARF typically affects school-aged children, with approximately 2.4 million 5-14 year olds being affected worldwide; there are 282,000 new cases annually. Taken together, ARF is the most common cause of pediatric heart disease (1).

ARF is an autoimmune collagen-vascular disease that manifests about 3 weeks after GAS infection of the upper respiratory tract, specifically pharyngitis (3). Symptoms of ARF include arthritis, carditis, Sydenham's chorea (discussed below), and subcutaneous nodules (3, 47). It is believed that the underlying mechanism for this autoimmune disorder is attributed to either the formation of antibodies directed against streptococcal antigens that cross-react with host tissues, or a cell-mediated response that is toxic to tissues (48). In 1944, T. Duckett Jones developed a series of guidelines, which are now commonly referred to as the Jones criteria, for identifying clinical manifestations of ARF (49). The patient must exhibit two major (arthritis, carditis, chorea) or one major and two minor (fever, increased erythrocyte sedimentation, arthralgia, prolonged P-R intervals on electrocardiograms, etc.) criteria for a positive diagnosis (3, 47). Patients diagnosed with ARF receive antibiotics, anti-inflammatories, and supportive care (50).

Acute Post-Streptococcal Glomerulonephritis

Acute post-streptococcal glomerulonephritis (APSGN) is an immune complex-mediated disorder of the kidneys, resulting in symptoms such as edema, hypertension, hematuria, urinary sediment abnormalities, and decreased levels of

complement components in serum (3, 32). It is estimated that there are over 470,000 cases of APSGN annually, which result in approximately 5,000 deaths (1). APSGN arises after skin associated GAS infections, specifically impetigo (21). APSGN presents itself in children and young adults, with males being affected twice as often as females. Those living in southern and temperate climates are more susceptible to APSGN, especially in the summer months, correlating with the disease cycle of impetigo (3). Symptoms usually arise 3 to 6 weeks after GAS infection, lasting for about a week (51). Treatment for APSGN includes blood pressure control with vasodilators, dialysis (used to treat uremia), and prophylactic antibiotic treatment (51).

Neurologic Disorders

Sydenham's chorea (SC) is a major neurological manifestation of acute rheumatic fever and symptoms can appear in 10 to 50% of patients presenting with ARF (52). Sir Thomas Sydenham is credited for first description of the disease in 1686, although the symptoms of the disease were known for several decades before as "St. Vitus Dance" (52). Children and adolescents, specifically females, are the predominant population who are affected by SC. However, cases in adults have also been reported. Symptoms of SC include hyper kinetic moves, typically of the face and limbs, motor impersistence, and motor and vocal tics (47, 52). Behavioral alterations are also common in SC, including emotional labiality and anxiety. SC is a self-limiting, non-progressing disorder that resolves itself after 3-18 months, however relapses can occur (52).

Pediatric autoimmune neuropsychiatric disorders associated with streptococcal infections, or PANDAS, were first described in 1998 with five clinical diagnostic criteria: presence of 1) obsessive-compulsive disorder (OCD) and/or tic disorder, 2) pre-pubescent symptom onset, 3) abrupt onset or exacerbation of symptoms with relapses, 4) temporal association between symptoms and GAS infections, and 5) associated neurological abnormalities including choreiform movements (52, 53). Symptoms can arise at any point after GAS infection, although a six-week interval has been suggested as the predominant timeframe (53).

Although there is evidence of a correlation with GAS infection and the development of the neuropsychiatric disorders described above, several question the postulated etiologic role GAS plays in the development of these diseases. In fact, recent studies have failed to provide a conclusive link to the exacerbation of diseases symptoms (54-56). The exact pathogenesis of PANDAS remains unknown, but it is hypothesized that it is a result of anti-neuronal antibodies developed posted infection (32).

Vaccination and Prevention

Vaccine Development

The need for a GAS vaccine is evident with 18.1 million cases of severe GAS disease and over half a million deaths each year worldwide (57). Currently, a licensed vaccine is not readily available, due in part to high rate of cross-reactivity of GAS antigens with host tissues and the lack of standardization in immunoassays (58). An

ideal vaccine would provide protection against all GAS strains, and with over 200 different serotypes (32), the abundance of diversity among GAS strains makes an all-encompassing vaccine nearly impossible. However, to date, several vaccine candidates have reached various stages of preclinical development, and a limited number have also reached early phase clinical trials (59, 60).

Historically, the M protein – a major virulence factor – has been the main target for vaccine development, using either the variable N-terminal or conserved C-terminal region as an epitope (3, 61). However, due to the hypervariablity of the N-terminal and the highly immunogenic C-terminal (cross reacts with human cardiac myosin) (62), these strategies have not been successful. Two vaccine strategies have been developed that utilize the N-terminus of the M protein: a hexavalent- and a 26-valent preparation (59, 60). The newest 30-valent vaccine was developed by Dale in 2011 and contains M protein peptides from the most prevalent strains in North America and Europe. This vaccine strategy was immunogenic in rabbits and had bactericidal effects; however, this formulation has not been tested in human to date (63).

Vaccines have also been developed that target non-M protein virulence factors such as C5a peptidase, lipoteichoic acid, protein F, the group A carbohydrate, streptococcal protective antigen, serum opacity factor, streptococcal pili, streptcococcal cell envelope protease, and streptococcal fibronectin-binding protein. All have been shown in animal studies to raise protective antibodies that contribute to host immunity, particularly in preventing colonization (3, 64-66). The streptococcal pyrogenic exotoxins SpeA, SpeB, and SpeC have also been proposed to play a role in

immunity, and have been proposed as potential vaccine candidates (3).

Antibiotic Treatment

Since the advent of antibiotic therapy with penicillin and penicillin derivatives, GAS has remained expressly sensitive to treatment with this family of antibiotics. In some cases however, GAS infections may fail to be cleared by penicillin treatment and may lead to asymptomatic carriage. There are two different hypotheses about this phenomenon: 1) GAS can enter epithelial cells and thereby evades penicillin, or 2) there is a co-infection with another β-lactamase-producing bacteria, making GAS indirectly resistant to the antibiotic treatment (32).

For patients who are allergic to penicillin, macrolides are often used. However, a recent longitudinal study saw a dramatic increase in macrolide resistance genes present in different GAS isolates; with resistance genes increasing from 12.1% in 2003 to 18.8% in 2006 (67). In addition, there has been evidence for both tetracycline and fluoroquinolone resistance in GAS (32).

Virulence Factors

In order to successfully colonize its host, a bacterial pathogen must possess features that allow for adhesion, internalization, and immune evasion. *Streptococcus pyogenes* produces a wide array of virulence factors (**Fig. 2.1**) that are utilized at different points of the infection process. Some of these are surface associated, while others are secreted into the extracellular milieu. In this section more comprehensive

details will be reserved for those virulence factors that are regulated by the Multiple Gene Activator (Mga), which is the focus of this dissertation.

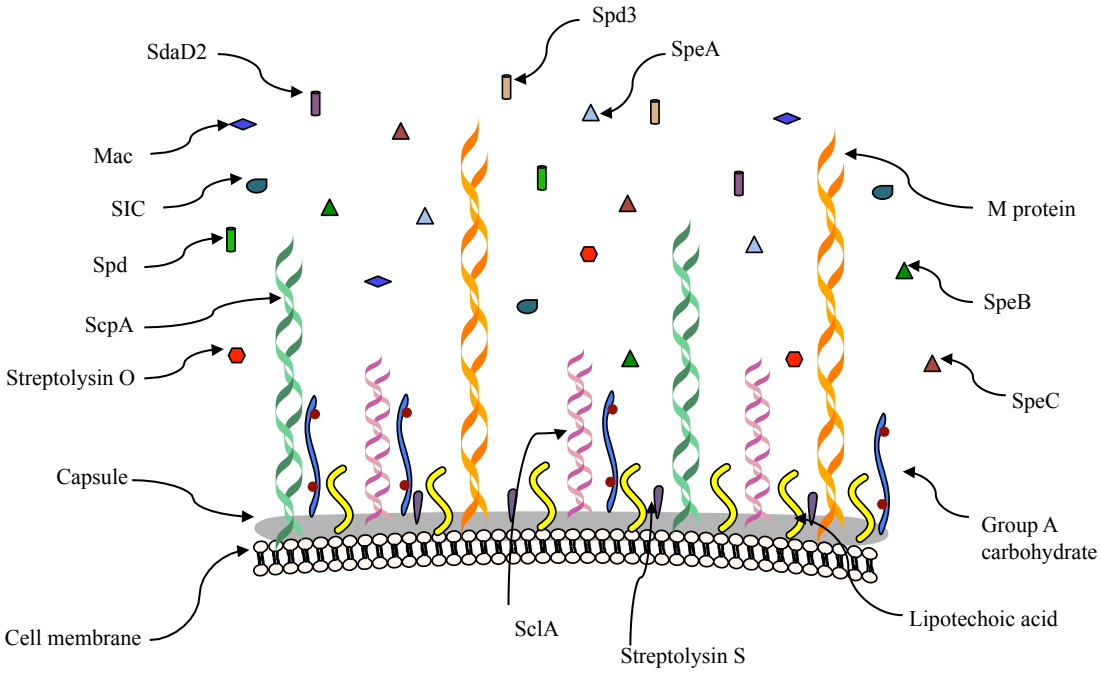


Figure 2.1. Virulence factors produced *by S. pyogenes*.
Graphical depiction of select virulence factors that are surface-associated, secreted, and phage-associated encoded by GAS. Virulence factors with similar function are shown as the same shape but with different colors.

Secreted proteins

Streptococcal Inhibitor of Complement

The streptococcal inhibitor of complement (SIC) is a signal sequence-containing secreted protein possessing a central region with three tandem repeats that is present in only four serotypes (M1, M12, M55 and M57) of GAS. The gene is located within the *mga* locus and is part of the core Mga regulon. It was discovered in 1996 as a putative inhibitor of the membrane attack complex (MAC), which is a part of the complement pathway (68, 69). Fernie-King and colleagues later confirmed that the direct mechanism of inhibition was attributed to disruption of the formation of the C5b67 MAC complex, thereby preventing lysis (69). SIC was also shown to bind secretory leukocyte proteinase inhibitor (SLPI), lysozyme, cystatin C, LL-37, β-defensin, ezrin, and moesin, which could aid the modification of host defense systems and in the initial colonization of GAS on mucosal surfaces (70-73) (74).

In addition to the interaction with human plasma proteins, a SIC-deficient M1 strain showed a significant decrease in the ability to colonize and persist in the upper respiratory tract of mice after intranasal infection (75). It is also necessary for GAS resistance in a murine model of systemic infections, for growth in both whole human blood and human serum, as well as for intracellular macrophage survival (76, 77). Therefore, SIC contributes to GAS immune resistance through binding of complement-associated proteins and is a major contributor to GAS virulence phenotypes.

Streptolysin O

Streptolysin O, or SLO, is one of two GAS hemolysins. SLO is a thiol-

activated, cholesterol-dependent cytolysin that creates 30 nm pores within host cells (78). SLO can form pores in a number of cell types, including erythrocytes, leukocytes, macrophages, and platelets (79, 80). This cytolysin helps deliver the enzymatically active co-toxin NAD^+-glycohydrolase (NADase) into the cytoplasm of eukaryotic cells, and is co-transcribed with *slo*. Together, these two virulence factors significantly contribute to GAS pathogenicity by inducing premature apoptosis of keratinocytes and affecting host cell signaling pathways (81). Most recently, it has been shown that these two effectors work together to mediate GAS intracellular survival and cytotoxicity for macrophages by preventing phagolysosome acidification (82).

Streptococcal pyrogenic exotoxin B (SpeB)

The cysteine protease, also known as the streptococcal pyrogenic exotoxin B or SpeB (encoded by *speB*), is a secreted zymogen that is auto-catalytically cleaved into its active form (83). Although SpeB is ubiquitous in GAS, its expression varies greatly among strains (84), and is highly correlated with the switch between localized to invasive GAS infections (85). SpeB is expressed at stationary phase of growth, and is associated with nutrient depletion, acid pH, and NaCl (4, 86). In addition, SpeB is responsible for cleavage of a variety of GAS surface-exposed and secreted antigens, including the M protein (discussed below) and SIC (32), that aid in dissemination. Furthermore, SpeB also degrades many host surface components such as fibronectin and IL-1β precursor (87) to inhibit internalization of GAS into host cells. SpeB is highly expressed in C Media, a low-glucose media, thereby linking *speB* expression to catabolite repression (88).

Mac/IdeS

Immunoglobulin G-degrading enzyme of *S. pyogenes*, Mac or IdeS, is a conserved secreted GAS protein with homology to a human β_2-integrin (CD11b; Mac-1), which binds to CD16 on the surface of human PMNs and inhibits phagocytosis and bacterial killing (89). It has two alleles; where complex I is the most conserved to CD11b, and is present in serotypes M1 and M3. Mac was later found to also cleave immunoglobulin G (IgG) through endopeptidase activity (90, 91). Crystal structure of Mac confirmed its specificity for IgG (92). The cleavage of IgG by Mac results in the formation of high amounts of circulating ½Fc fragments that are specific for GAS and can still bind to the surface antigens, but no longer possess the ability to signal and activate complement. These antibody fragments exhibit a proactive effect and interfere with the recognition of surface-associated antigens (93). Therefore, it was found that antibodies against Mac were found to decrease bacterial survival in whole human blood (94).

Surface-associated proteins

M and M-like proteins

The M protein is a major virulence factor that plays a role in immune evasion and cellular adhesion based on its complex structure. The initial interaction of GAS with the human host occurs at the epithelial surface of the pharynx or the external skin. M protein mediates adherence to skin keratinocytes via the C repeat region of the keratinocyte membrane cofactor CD46 (95, 96). The antiphagocytic properties of the M protein interferes with opsonization via the alternate complement pathway (97). This can occur by one of two mechanisms: 1) direct binding of complement proteins

via the C repeat region of the M protein (98), or 2) by binding of fibrinogen to the M protein via the N-terminus (99).

The M protein is a surface anchored, α-helical coiled-coil protein that is comprised of four distinct repeat regions (A, B, C, D) (3) and is anchored to the cell wall from the C-terminus by a LPXTG motif (100). The hypervariable A repeat, located on the N-terminus, is the basis for serological/sequence differentiation of GAS strains (3, 32). As mentioned in the GAS classification section, M Proteins can be divided into two subclasses: Class I with an extended conserved surface region (no serum opacification, OF^-) and Class II that lacks this conserved region (serum opacification, OF^+) (101).

The M protein was first described by Lancefield (30), and has been implicated in various GAS diseases and sequelae, specifically rheumatic fever (101). It is involved in mediating adhesion to host epithelial cells (96), and immune evasion through resistance to opsonophagocytosis by the binding of fibrinogen, complement inhibitory factor H, C4b-binding protein, and immunoglobulin (102, 103). During the later phases of infection, the M protein is released into the host by the cysteine proteinase SpeB (104, 105), which induces neutrophil activation and consequently vascular leakage. It also has been shown that soluble M protein also induced T cell proliferation and the Th1 response (106).

Additional surface proteins related to M protein have now been identified. Although the structure of these so-called "M-like" proteins is overall quite similar to that of M protein, they differ in the types of repeats and in their ability to interact with different human proteins. GAS encodes for other genes that are part of the M protein

superfamily due to their structural similarity in the cell wall associated region. There have been more than 20 genes (*mrp*, *enn*, *arp*, *sir*, *fcrA*, *etc*.) that have been identified in this superfamily, and typically GAS may encode for one to three of these genes (107). The salient characteristics of the M protein family is its ability to bind to a wide range of host proteins, including but not limited to, albumin, fibrinogen, and plasminogen (108, 109).

Streptococcal Collagen-like Surface Protein

Streptococcal collagen-like protein, SclA or Scl-1, is a surface associated protein that contains 50 contiguous Gly-X-X triplet amino acid motifs that mimic human collagen (110, 111). It is a cell wall-anchored protein containing the canonical LPXTG motif on its C-terminus, and is identified in all serotypes of GAS (110, 111). SclA is important for adherence to alveolar epithelial cells, fibronectin, and laminin, and is important for soft tissue infection (110, 112). SclA has also been shown to mediate bacterial survival in neutrophil extracellular traps (NETs) by interfering with myeloperoxidase (MPO) release, which is a prerequisite for NET production. It was also found to provide protection for GAS from antimicrobial peptides found within NETs (113). SclA participates in biofilm formation through its interaction with cellular fibronectin in the extracellular matrix produced by dermal fibroblasts, directly correlating its role in host tissue colonization (114).

C5a Peptidase

The streptococcal C5a peptidase, or ScpA (encoded by *scpA*), is a surface-associated serine protease that is present in all GAS isolates (115, 116). It can cleave the human serum chemotaxin C5a at a polymorphonuclear leukocyte (PMNL)

binding site on the carboxyl terminus (117) allowing GAS to evade opsonophagocytosis. The crystal structure of active ScpA confirmed its interaction with C5a (118) and it has been implicated in the clearance and dissemination of GAS during infection (119). Due to the conservation of this gene across GAS serotypes, it has also been proposed as a vaccine candidate. Two studies have successfully used recombinant ScpA as an immunogen in order to confer protective immunity in mice when infected intranasally with GAS (119, 120).

MSCRAMMs

In order to adhere to host extracellular matrix components, specifically fibronectin and collagen, GAS encodes for several microbial surface component recognizing adhesive matrix molecules, or MSCRAMMs (121). Currently, there are 15 known MSCRAMMS encoded by GAS, 11 of which are specific for fibronectin binding, two that uniquely bind collagen, and one that binds both (M protein) (122). The presence of so many hints at redundancy of these surface adhesins, but most of these are not ubiquitous in the GAS genome. In addition, they are expressed at different times during infection, and are under the tight control of different virulence regulators (122).

Protein F1, or Sfb1 (encoded by *prtF*), is one of the most well characterized MSCRAMMs present in the GAS genome. Initially, attachment occurs when the N-terminus of fibronectin is bound by dual binding sites on Protein F1; consequently, conformational changes in Protein F1 and alternative interaction with fibronectin lead to the internalization of the GAS cell (123, 124).

Serum Opacity Factor

The serum opacity factor, or SOF (encoded by *sof*), of GAS is present in Class II strains (see M protein). This α-lipoproteinase is known for its ability to opacify mammalian serum and bind fibronectin (125). The molecule is composed of three domains: an N-terminal domain responsible for the opacity reaction, a repeat domain with fibronectin-binding activity, and a C-terminal (LPXTG motif) cell attachment domain (126). It is an important virulence factor, as a strain lacking SOF saw significant attenuation during a murine intraperitoneal model of infection (127).

Capsule

The hyaluronic acid (HA) capsule (encoded by *hasABC*) is a significant virulence determinant present on the surface of GAS. Composed of a polymer of glucuronic acid and N-acetylglucosamine, it exhibits molecular mimicry with connective tissue, playing a role in adhesion and immune evasion. Capsule is under tight transcriptional control by the CovRS (CsrRS) two-component system, RocA, and the more recently discovered sRNA HasS (128-130). Additionally, degrees of encapsulation have been implicated in severity of disease, where those strains that are highly encapsulated are often associated with rheumatic fever and invasive disease (131), and acapsular strains are significantly attenuated in a variety of infection models (132, 133). Interestingly, capsule was found to impede GAS adherence and internalization to Hep-2 cells and human keratinocytes. However, this paradox was clarified when capsule was identified to bind CD44 on human epithelial cells, which induces cytoskeletal rearrangements that disrupt cellular junctions and allow GAS to remain extracellular as they penetrate the epithelium (134).

Streptolysin S

Streptolysin S, or SLS, is an oxygen stable hemolysin that is responsible for the characteristic zone of β-hemolysis of GAS. It is encoded by a nine-gene operon (*sagA-I*), where *sagA* is the actual structural toxin and the remaining genes are necessary for production of an active toxin (135). SLS is primarily found in a to the cell surface of GAS in a cell-bound form (136) and is delivered to target cells by direct contact (137). SLS has a broad range of cytotoxic effects, leading to lysis of lymphocytes, neutrophils, platelets, cancer cells, and even cellular organelles (138-141). It is also found to be critical for full GAS pathogenesis, since a strain defective in toxin production has a significant reduction in virulence in a murine skin model of infection through enhanced phagocytic clearing (135).

Phage-associated proteins

DNases

GAS M1T1 serotype possesses three different DNases, two of which are encoded by prophages (*spd3* and *sdaD2/sda1*) and one that is chromosomally encoded (*spd*). Streptodornase 1 (Sda1) has been shown to degrade the DNA in NETs (142, 143), which allows GAS to escape from NET-mediated killing. It is believed that upon transitioning to an invasive form of infection, Sda1 is beneficial for GAS survival in whole human blood (144). A M1T1 strain lacking all three DNases was attenuated in two mouse models of invasive infection and showed a decrease in its ability to colonize cynomolgus macaques (145). Interestingly, a lysogenic M1 strain expressing Sda1 showed no capacity to survive human neutrophil killing or cause invasive lethal infection in a humanized plasminogen mouse model, suggesting that

the hypervirulence of the M1T1 clone is due to a unique synergic effect of Sda1 acting in concert with the M1T1 clone-specific genetic scaffold (146).

Streptococcal pyrogenic exotoxins

The streptococcal pyrogenic exotoxins (Spe) are a family of bacterial superantigens believed to be associated with streptococcal toxic shock syndrome (STSS). This family includes the bacteriophage encoded SpeA and SpeC, historically known as the scarlatina toxins due to their association with scarlet fever (44). Other toxins include: SpeG, SpeH, SpeJ, SpeK, SpeL, streptococcal superantigen SSA, streptococcal mitogenic exotoxins SmeZ and SmeZ-2, SpeG, SpeJ, SpeH, SpeL, and SpeM. SmeZ interacts with host MHC class II and the T cell receptor, leading to activation of both cells. *In vitro*, these interactions lead to proliferation and cytokine production by these cells. Studies using humanized HLA class II transgenic mice and isogenic streptococcal strains have characterized the *in vivo* responses to superantigens produced in the context of live infection. However, due to the deleterious role of superantigens in toxic shock of the host, the evolutionary advantage conferred by these toxins remains a subject of speculation (147).

Regulation of Virulence

Due to its vast repertoire of virulence factors, GAS must have tight control over their expression in order to successfully adhere, colonize, and invade the host. Unlike other prokaryotes that use alternative sigma factors as a means of regulation for gene expression, GAS utilizes different methods of regulation, including its 13 two-component systems and "stand-alone" regulators (32).

Two Component Systems

Prokaryotes' main method of signal transduction is through two-component systems (TCS). TCS consist of a membrane bound histidine sensor kinase (HK) and a cytosolic response regulator (RR). An external environmental signal causes auto-phosphorylation of the histidine residue on the sensor kinase. The phosphate is then transferred to the response regulator on an aspartate residue, thus allowing for activation of the protein and transcriptional regulation (148).

CovRS

The control of virulence regulatory system (Cov; *covRS*), or capsule synthesis regulon (CsrRS), is the most well characterized TCS in the GAS genome. It was first identified as a regulator for the capsule synthesis operon (*hasABC*), where inactivation of *covRS* leads to hypermucoid colonies (129). It was subsequently found to regulate the transcription of other virulence factors, including streptokinase (*ska*), streptolysin S (*sagA*), and the cysteine protease (*speB*) (149-151). CovRS is functionally connected to the general stress response in GAS and is activated by changes in heat, low pH, salts, low iron, antimicrobial peptides, and Mg^{2+}(152-155). A genome-wide regulatory study identified that 15% of the GAS genome is regulated by CovRS; genes regulated include other TCSs, as well as surface and secreted proteins mediating host–pathogen interaction (156). Interestingly, mutations in *covRS* have been linked to the switch between non-invasive and invasive disease. In addition to altering the expression of important virulence factors, these mutations lead to increased resistance to phagocytosis and killing by neutrophils (157) and cause hypervirulence in a mouse model of infection (158). The mechanism by which

CovRS affects the transcription of these genes is not completely understood. Although a consensus binding site (ATTARA) has been identified that maps at or near the putative promoter sequence of many genes that are directly regulated by CovRS, other promoters lacking this sequence exist (159).

TrxSR

The two-component regulatory system X or Trx (encoded by *trxSR*) is repressed by CovRS (156). It has been shown to play a role in virulence regulation whereby a *trxR* mutant was significantly attenuated in a subcutaneous murine module of infection (160). TrxR can also directly regulate Mga and its regulon via its ability to bind directly to P*mga in vitro* (161).

Stand-Alone Regulators

A "stand-alone regulator" is a term used to describe a single regulator that does not have an identified sensor component and responds to environmental stimuli directly to alter gene expression. This subsection will focus on two of the main families of stand-alone regulators, and the third such regulator, Mga, which is the focus of this dissertation, will be discussed in greater detail in its own section below.

Rgg/RopB

The regulator of protease, or Rgg/RopB (encoded by *ropB*), is known to affect the transcription of several genes in stationary phase of growth such as the cysteine protease, SpeB (162). In a global transcriptional study, Rgg was shown to affect the expression of several regulatory networks including CovRS, and Mga in order to co-regulate virulence factor networks (163). Rgg also regulates several amino acid catabolism operons (164), contributes to utilization of alternative sugars associated

with secondary metabolism and thermal stress responses (165), and participates in GAS quorum sensing through small-peptide signaling (166). Rgg has been shown to bind directly to about 65 targets through a ChIP-chip experiment (167), and it was later found that DNA binding of Rgg is dependent on environmental changes in a LacD.1-dependent manner (4).

RALPs

The RofA-like protein (RALPs) family contains four regulators that act at the transition between logarithmic and stationary phase of growth to control virulence gene expression. Each protein in this family (RALP1-4) will be further described in the subsection below.

RofA

The regulator of F, or RofA/RALP1 (encoded by *rofA*), was first found to be a positive transcriptional activator of the fibronectin-binding adhesion protein F under anaerobic growth conditions (168). It is also shown to negatively regulate Streptolysin S, the cysteine protease SpeB, and Mga (169). It binds directly to DNA in its target promoters using a 17 base pair binding site (170), and autoregulates its own expression (168). Additionally, RofA is important for internalization and survival of GAS (169).

Nra

The negative regulator of GAS, Nra/RALP2, was first found to negatively regulate its own expression alongside the MSCRAMM fibronectin-binding protein F2, the collagen-binding protein Cpa, Mga, SLS, the superantigen SpeA, and SpeB (171-173). It is predominantly found in opacity factor negative (OF^-) serotypes (174).

Nra-deficient strains also showed no defect on bacterial adhesion to epithelial cells, but showed a reduction in the internalization rates of the bacteria, and these mutant bacteria could also escape phagocytic vacuoles more efficiently (172). In addition, Nra has been shown to have varying effects on pilus gene expression, where in some serotypes it is negatively regulated and others it is an activator of pilus expression (175).

RALP3

The third member of the RALPs, RALP3, is thought to have been acquired from *Streptococcus suis* through a horizontal gene transfer event, and as such is only found in certain serotypes (M1, M4, M12, M28, and M49). *ralp3* is located in a discrete genetic region of the genome designated as the ERES (for the genes *eno, ralp3, epf,* and *sagA*) pathogenicity region (173). RALP3 is known to negatively regulate numerous virulence factor genes, including those involved in hyaluronic acid capsule and cysteine protease production. It also contributes to GAS epithelial cell invasion and antimicrobial peptide resistance (176), and is critical for GAS survival in whole human blood (15).

RivR

The ralp IV protein (RivR) is cotranscribed with a sRNA *rivX* (177). Both of these genes are repressed by CovRS (178) and appear to enhance the transcription of Mga and its regulon in an M1T1 background (MGAS5005; cov^-)(177). *In vivo* studies have shown that a *rivR* mutant is attenuated for virulence in an invasive subcutaneous murine model of infection (177). However a recent study in a different M1T1 strain (MGAS2221; cov^+) demonstrated that RivR was a negative regulator of the

hyaluronic acid capsule biosynthesis proteins and the G-related-α_2-macroglobulin binding (GRAB) protein, and showed no effect on the Mga regulon. Moreover, RivR was shown to be important for adherence to human keratinocytes and in a bacteremia murine model of infection (179).

The phosphoenolpyruvate-phosphotransferase system (PTS) and its role in virulence

Overview

The phosphoenolpyruvate phosphotransferase system (PTS) is a conserved multi-protein pathway that couples transport of sugar with their phosphorylation and further metabolism (180). In addition, the PTS is also a signal transduction system where the flow of phosphorylated sugars is monitored. The phosphate from PEP generated though glycolysis flows to the EI protein, then to Hpr, which can be phosphorylated on the His-15 residue, and continues to EII and then to the incoming sugar where it cycles back into glycolysis (**Fig 2.2**). Under favorable nutrient conditions (high glucose) Hpr can also be phosphorylated on Ser-46 residue, allowing for interaction with CcpA to mediate carbon catabolite repression (CCR) in the presence of glucose (180, 181).

In Gram-positive bacteria, the HPr-CcpA complex mediates CCR by binding to catabolite response elements (*cre)* present within the promoters or coding regions of regulated genes. A 14 base-pair consensus *cre* site (TGWAARCGYTWNCW) was determined in *B. subtilis* (182). Interaction between CcpA and the *cre* site can either

represses alternative sugar utilization genes or activate the transcription of genes that may function in glucose metabolism (183) (184).

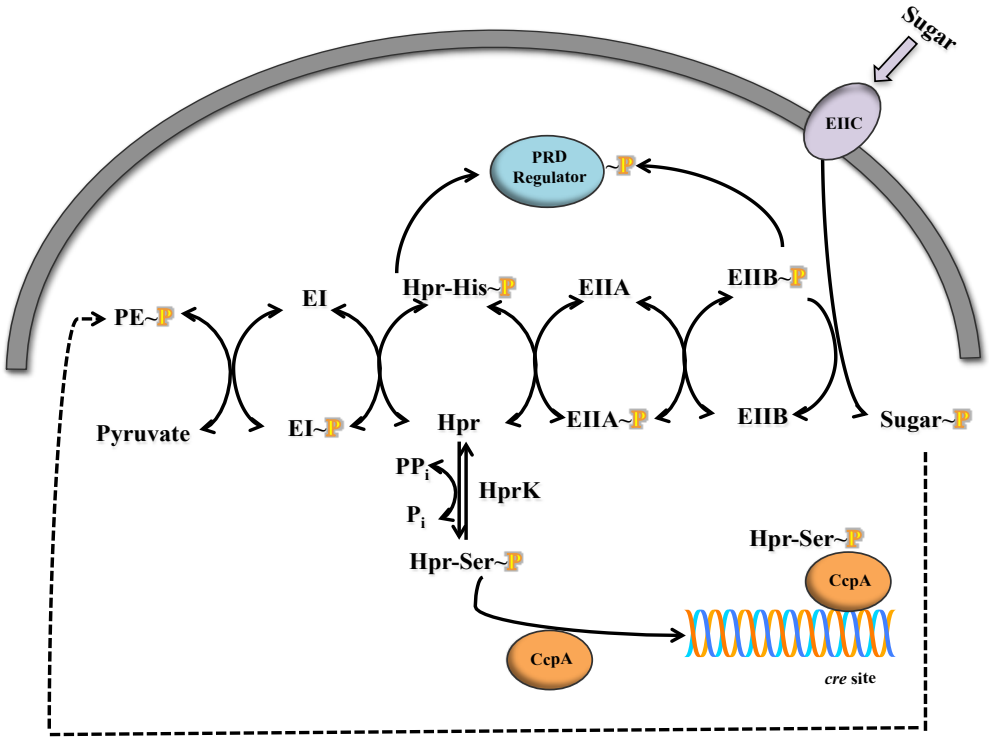


Figure 2.2. Phosphoenol-pyruvate phosphotransferase phosphorelay.
PTS-specific sugars are brought into the cell via the PTS EIIABC(D) permease and enters glycolysis. The resulting PEP transfers the phosphate to EI, which in turn transfers the phosphate to the His^{15} residue of HPr. P~His^{15}-HPr and P~EIIB can also phosphorylate PTS regulatory domains (PRDs) of sugar regulators, thereby modulating their activity. Alternatively, Hpr kinase recognizes intermediates of glycolysis and phosphorylates HPr on the Ser^{46} residue. P~Ser^{46}-Hpr then complexes with CcpA and mediates CCR. (Adapted from (19)).

Carbon catabolite repression (CCR) is a global regulatory mechanism of carbon source utilization that both Gram-negative and Gram-positive bacteria use to promote efficient utilization of carbon sources when the preferred substrate (glucose) is present (185). In a nutrient rich environment, such as glucose, bacteria use CCR to inhibit the expression of enzymes and transporters necessary for consumption of alternative carbon sources. In addition, CCR and other PTS related genes have been shown to play an important role in virulence gene regulation for Gram-positive bacteria, specifically pathogens in the phylum *Firmicutes*. To mention a few, low glucose availability conveys methicillin resistance in *Staphylococcus aureus*, promotes successful colonization in *Streptococcus pneumoniae*, stimulates host cell invasion in *Listeria monocytogenes*, and increases toxin production in *Clostridium perfringens* (186). These pathogens, like all bacteria, rely on global regulatory networks and dedicated sugar transporters in order to detect the presence of preferred carbon sources as a reflection of the nutrient status of the host environment. This allows for the appropriate coordination of virulence gene expression and disease manifestations in response to surrounding conditions.

GAS is a fastidious organism that lacks a functional tricarboxylic acid cycle (TCA) cycle, and relies on the transport of carbon sources through the PTS in order to generate energy. Components of the PTS and CcpA have been shown to be very important for the pathogenicity of GAS. The rest of this section will further outline and PTS and its importance in pathogenesis, as well as discuss the link between carbon metabolism and virulence in GAS.

Virulence Regulation and Sugar Metabolism in Gram-Positive Pathogens

Listeria monocytogenes

The major activator of virulence in *L. monocytogenes,* PrfA, is under control of CCR (187). PrfA controls the expression of genes required for invading host cells and for release from phagolysosomes. However, the mechanism for CCR-mediated repression of PrfA was not a result of direct regulation by CcpA (188), but was a result of P~Ser46-HPr (189). PrfA was found to be repressed in the presence of glucose and fructose. Interestingly, several other virulence genes such as a hemolysin (*hly*) and a phospholipase (*plcA*) were also repressed in *L. monocytogenes* when it was grown on cellobiose, glucose, or other readily metabolized carbon source (181).

Clostridium perfringens

In the anaerobic, spore-forming pathogen *C. perfringens,* expression of a variety of virulence genes are shown to mediated by CCR. For example, CcpA regulates expression of enterotoxin (*cpe*), sporulation, collagenase, polysaccharide capsule, and motility through Type IV pilus (190-192).

Staphylococcus aureus

Studies on CcpA and CCR in the pathogen *S. aureus* found that CcpA-mediates expression of genes involved in antibiotic resistance and other virulence determinants. In fact, a CcpA mutant showed reduced resistance to the antibiotic oxacillin (193). In addition, CcpA was also found to negatively regulate the expression of the staphylococcal α-hemolysin, capsule, and protein A. Furthermore, this study showed CcpA-mediated activation of the virulence regulator RNAIII, an effector molecule for the *agr* locus involved in global regulation (193).

Streptococcus pneumoniae

In the pathogen *S. pneumoniae,* the CcpA ortholog RegM represses α-glucosidase and β-galactosidase activities and the capsular polysaccharide synthesis (*cps*) locus, which is important for cell adhesion (180). A mutation in RegM is attenuated for both nasopharygeal colonization and virulence in the mouse pneumonia model of infection (194).

Streptococcus mutans

The pathogenic oral streptococci *S. mutans* relies on carbohydrates for energy production in order to successfully colonize and the oral cavity and cause dental caries. A Δ*ccpA* mutant exhibited diminished fructosyltransferase and glucosyltransferase gene expression, which are both important for biofilm formation (195). Interestingly, it has been demonstrated that CcpA has very little influence on CCR in *S. mutans;* however, HPr can exert effects on CCR in *S. mutans*. CCR in this organism is dominantly controlled by the glucose/mannose EIIAB permease (180).

Virulence Regulation and Sugar Metabolism in GAS

Enzyme I (EI)

The first enzyme in the PTS, EI (*ptsI*), initiates the phosphorelay of the PTS after autophosphorylation by PEP (180). Recently, Gera *et al*. demonstrated that a Δ*ptsI* mutant exhibited a hyper virulence phenotype in a subcutaneous murine model of infection. In addition, an increase in transcript levels of Streptolysin S, in conjunction with the early onset of its activity, correlated with the increased lesion size of the Δ*ptsI* mutant. Taken together, this data suggested that a functional PTS is

not required for subcutaneous skin infection in mice; however, it does influence correct timing of virulence factor expression and limits disease progression (8).

CcpA

The global catabolite repressor protein, a LacI/GalR family member, is a key regulator of carbon catabolite repression (CCR) in GAS. Numerous metabolic operons are under the control of CcpA through its ability to bind to the catabolite response element (*cre*) sequence when complexed with P~Ser46-Hpr (196). In GAS, CcpA has been shown to have an indirect regulatory effect on streptolysin S, however the directionality of this effect is not exactly clear. Shelburne *et al.* found SLS to be down regulated and a Δ*ccpA* mutant was attenuated in a mouse invasive model of infection, whereas Kinkel *et al.* found SLS to be up regulated and a Δ*ccpA* mutant was hypervirulent in a systemic infection model (5, 6). CcpA was also found to negatively regulate *cfa* (CAMP factor), but positively regulate the cysteine protease gene *speB* (7). In addition, CcpA is proposed to bind directly to the promoter region of *mga* (P*mga*) via an identified alternative *cre* site, thereby affecting the expression of this major virulence regulator (17).

Maltose Operon

The maltose PTS transporter, MalE, in conjunction with another transporter MalT (*ptsG*), is responsible for the transport of maltose into the cell. Mutations in these genes, in addition to abrogating maltose uptake by GAS, were shown to influence the ability of GAS to grow in human saliva and impair colonization of the mouse oropharynx (197). In addition, MalR, a LacI/GalR-family repressor protein has also been shown to affect the ability of GAS to colonize the mouse oropharynx but

not affect invasive infection. It also regulates the virulence gene *pulA,* which encodes a cell wall-anchored carbohydrate binding and degrading enzyme (198).

CodY

The regulator CodY links carbon metabolism with the nitrogen cycle by sensing amino-acid starvation in several low G+C Gram-positive bacteria. This regulator has been shown to affect virulence gene regulation by subsequent activation of CovRS in GAS (199). Moreover, CodY affects the transcription of *mga*, *ropB*/*rgg*, *fasBCAX*, and *pel* (200) Therefore, it is not surprising that CodY contributes to GAS survival and virulence gene regulation in whole human blood (201).

M protein

In 1970, Pine and Reeves investigated the role of metabolism and sugar source on the production of M protein (18, 202), and their studies showed that expression of M protein was highest during exponential growth phase. However, they also observed a surge of M protein production in stationary phase as all growth substrates were being depleted. Additionally, they found that the highest levels of M protein were produced when GAS was grown in medium with glucose as the sole carbon source as compared to sucrose and trehalose (18). Thus, this suggested a link between sugar metabolism and production of this important virulence factor of the GAS.

PRD Containing Regulators

PTS regulatory domains (PRD) are found in a large number of bacterial regulator proteins, otherwise known as PRD regulators, and are known for their

ability to respond to the availability of carbohydrates (**Fig. 2.2**). These metabolic regulatory proteins can be RNA-binding transcriptional anti-terminators or DNA-binding transcription factors. Their activity is dependent on the concomitant phosphorylation of conserved histidine residues on their reiterated PRD domains (PRD1/2), which in turn influences homodimerization (180). Inactivity occurs in the absence of an inducing sugar, where the phosphorylated EIIB component of the sugar permease transfers its phosphate to a PRD, often inactivating the regulator. In the presence of the preferred sugar, the regulator is dephosphoryalated and the phosphate is transferred to the incoming sugar for further metabolism.

PRD-containing regulators have been demonstrated to aid in CCR, whereby P~His15-Hpr can transfer its phosphate to the regulator and stimulate its activity (203). Importantly, newly recognized members of this regulator family have also been shown to regulate virulence pathways, which directly links sugar metabolism and growth to virulence and the disease process. This section will describe two paradigm PRD containing regulators as well as two homologs of Mga that are PRD-containing virulence regulators (PCVRs).

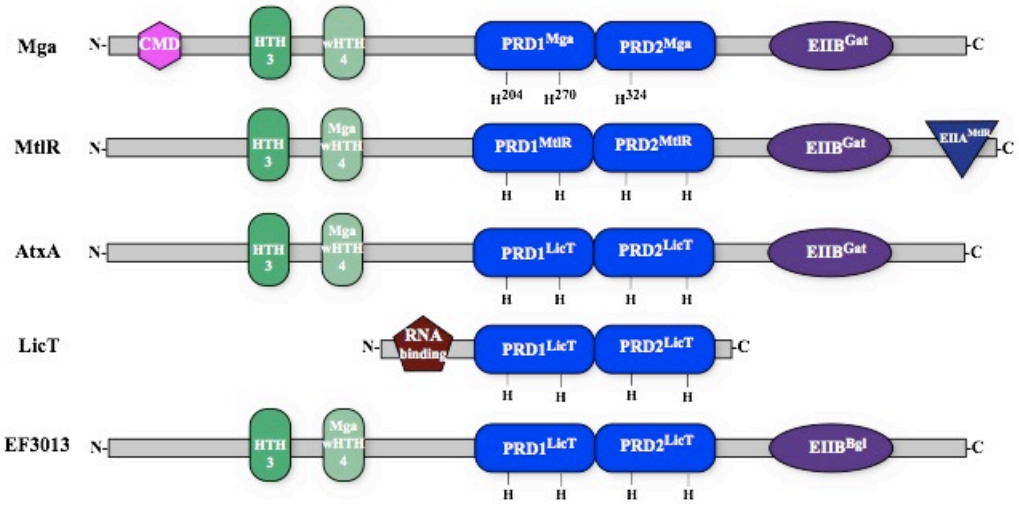


Figure 2.3. Domain alignment of GAS Mga with other PRD containing regulators.
Shown are *B. subtilis* MtlR and LicT, *B. anthracis* AtxA, and *E faecalis* EF3013. EII (purple), PRD (blue), DNA-binding (green) and conserved (pink) domains are indicated with conserved histidines (H). PRD^{Mga} and PRD^{LicT} refer to Pfam 08270 and Pfam 00874 respectively. (Adapted from (19)).

Archetype PRD Regulators

LicT

The *Bacillus subtilis* archetype PRD-containing regulator, LicT, is a BglG-family anti-terminator that controls expression of the metabolic genes for lichenan and other β-glucosides utilization and import (204). Antiterminators are activated in the presence of a sugar inducer, and function by binding to an antiterminator RNA (RAT) site. This binding occludes the terminator on the nascent mRNA and prevents the formation of the terminator stem-loop structure allowing for the transcription of

the operon. The crystal structure of the active form of LicT revealed a N-terminal RNA-binding domain followed by two PRD domains (**Fig. 2.3**). The two PRD domains each contain two conserved histidines (205) that are phosphorylated by P~His15-HPr (206) for full activity. Activity is reduced in the presence of P~Ser46-Hpr and P~EIIB^{Bgl} (207). LicT froms a homodimer through interactions of its two PRD domains; PRD1 of one monomer interacts with PRD1 of the other, and the same holds true for PRD2. *In silico* models predicted that phosphorylation of the histidines on PRD lead to strong electrostatic repulsions, whereas phosphorylation of PRD2 leaders to increased electrostatic complementarity between the homodimer (208).

MtlR

The DeoR-family mannose operon activator MtlR of *Geobacillus stearothermophilus* was the first PTS-regulated transcriptional activator identified (209). This sugar regulator is composed of a DNA-binding domain at the amino terminus, followed by two PRD domains and an EIIA^{Mtl} domain at its C-terminus (210). MtlR can be phosphorylated in a PEP-dependent manner through EI/Hpr (increase affinity for DNA) or through EIICB^{Mtl} (reduces affinity for DNA). Similar to LicT, MtlR encodes for two histidines on each PRD domain, and activity of the protein is dependent on the phosphorylation of both PRD2 histidines by Hpr. The EIIA domain of the protein can be phosphorylated by EIICB^{Mtl}, and can then commence a phosphorelay to PRD1 back to EIICB^{Mtl} and the incoming mannitol (210). MtlR is predicted to form a homodimer using the PRD1 and PRD2 of each subunit as the dimer interface (180). MtlR also is found in *Bacillus subtilis*, and is found to have similar activity with its PRD phosphohistidines (**Fig. 2.3**). However, a

secondary mechanism of regulation is also present in *B. subtilis* that involves the sequestration of MtlR to the cell membrane by $EIIB^{Mtl}$ mediated by the $EIIB^{Gat}$ domain on MtlR, which is necessary for full activity of the protein (211).

PRD-Containing Virulence Regulators (PCVRs)

AtxA

The *Bacillus anthracis* transcription factor AtxA (**Fig 2.3**) is a master regulator that controls the expression of virulence genes encoded on both pathogenicity plasmids pXO1 and pXO2. It directly regulates toxin gene expression (pXO1) and indirectly regulates, capsule synthesis (pXO2) (212). Its activity is elevated during growth in media containing glucose and CO_2/bicarbonate (213). This virulence regulator shares homology with Mga, and also shares similar predicted domain structures.

Tsvetanova *et al*. identified two PRD domains in AtxA *in silico*, and determined that protein activity was dependent on phosphorylation by a functional PTS *in vivo*. Specifically, phosphorylation of H199 on PRD1 appeared to be necessary for activation, and phosphorylation of the H379 residue on PRD2 was shown to have inhibitory effects on toxin gene transcription. They also addressed the requirement of the histidine residues by performing phosphomimetic/phosphoablative experiments that mutagenized the H199 to H199D or A (214). Taken together, these findings suggest that AtxA is a PRD containing regulator, and virulence factor production is linked to sugar metabolism in *B. anthracis*.

AtxA also contains an EIIB-like domain at its C-terminus that has been implicated in dimerization. Hammerstrom *et al*. showed that AtxA forms dimers, as

well as higher order structures under native conditions. From this study, it was also determined that the C402 residue that is located in the predicted EIIB-like domain is crucial for protein cross-linking. However, this residue was not found to be important for protein-protein interaction, as a C402S mutant exhibited multimerization similar to the native protein. Thereby, this data suggests that the EIIB-like region of AtxA is important for oligomerization. Additionally, under elevated CO_2/bicarbonate conditions, AtxA exhibited higher rates of dimer formation, suggesting that host-associated signals also enhance AtxA function by favoring the dimer configuration of the protein (215).

Recently, Hammerstrom *et al*. resolved the crystal structure of AtxA, and it was confirmed that the EIIB region of the protein is important for multimerization. The dimer interface formed contact between helices in PRD2 on one monomer and the EIIB-like domain of the other. The key residues for these interactions are L375, T382, L386 and N389 from PRD2, and I403, Y407, E413 and K414 from the EIIB-like domain. Additionally, they demonstrated the functional importance of the phosphohistidines located on PRD1/2 domains (216). As mentioned above, when AtxA is phosphorylated on H379 or a phosphomimetic H379D mutant was tested, AtxA was found to be inactive (214). Interestingly, the H379 residue is found on PRD2 and phosphorylation of this residue was found to inhibit dimer formation of AtxA. This indicates that phosphorylation of H379 destabilizes homodimer interaction, therefore affecting activity. Additionally, the H199 phospohistidine, which was also implicated in AtxA's activity (214), was not found to affect dimerization when mutated to a phosphomimetic. However, this residue sits in PRD1

adjacent to the linker between this region and the HTH2 domain. Therefore, it was hypothesized that phosphorylation of H199 may not affect dimerization, but affect DNA binding affinity (216).

Mga

The multiple gene regulator of the Group A Streptococcus (Mga) is another PCVR and is the main focus of this dissertation, and its regulation by the PTS of GAS will be discussed below.

EF3013/MafR

The *Enterococcus faecalis* protein is a PRD-containing regulator with a resolved crystal structure (PDB 3SQN) (217). It has overall structural homology to Mga, and up until very recently has had no biological significance associated with it. A recent study by Ruiz-Cruz characterized the role of this protein in *E. faecalis* pathogenesis rnamed it as *mafR*, for Mga/AtxA-like *faecalis* regulator. *mafR* was found to activate genes important for various steps in carbohydrate metabolism and affects the utilization of glycerol, maltose, and mannitol. In addition a mutation in *mafR* was less virulent in a mouse peritonitis model, indicating that it may be important for *E. faecalis* survival in host niches (218). Thus, it appears that MafR exhibits high degrees of similarities with other archetype PRD regulators, by its ability to regulate sugar metabolism, however it is similar to AtxA/Mga in that it also has a broader effect on gene regulation and plays a role pathogenesis.

Multiple Gene Regulator of the Group A Streptococcus (Mga)

Overview

The multiple gene regulator of the Group A Streptococcus (Mga) is ubiquitous in all strains of GAS (219, 220). Its discovery is attributed to two independent studies that aimed at identifying regulators of the M protein (**Fig. 2.1**), a major virulence determinant that contributes to GAS adhesion and immune evasion. The first study aimed at identifying the genetic factor for strains lacking M protein. Strains lacking M protein expression exhibited random, 50 bp deletions that were mapped to a distinct locus upstream of the *emm* gene (221). This locus was subsequently found to affect the expression of the streptococcal C5a peptidase (*scpA*), and was consequently named *virR* (virulence regulator)(115). Concurrently, Caparon *et al*. identified a mutant in a Tn916 mutagenesis screen that caused a 50-fold reduction of *emm* transcript. The locus was then named *mry* for M protein RNA yield (222). Further characterization of Mry found it to act as a positive regulator of *emm in trans* (223). In 1995, VirR and Mry were found to be homologous and were renamed as Mga (Multiple gene regulator of the group A Streptococcus) (224).

Mga has since been shown to be critical for multiple pathogenic phenotypes, including biofilm formation, growth in whole human blood, and resistance to phagocytosis, (11-14). Mga is also critical for survival in a mouse model of systemic infection and GAS attachment to keratinocytes (13). Two genetic fitness screens showed that *mga* is critical for GAS survival in both an *ex vivo* (15) and *in vivo* (16) environments. The *mga* gene was found to be maximally expressed 30 minutes post exposure to either whole human blood or human saliva (12, 225), correlating this

finding to the *in vitro* expression of the Mga regulon during exponential growth (226, 227). Additionally, using a cynomolgus macaque pharyngeal model, the Mga regulon was shown to be maximally expressed during the acute phase of pharyngitis infection (228). Taken together, this data demonstrates that Mga is critical for GAS pathogenicity and its activity greatly depends on cellular metabolism.

Mga-regulated genes

A microarray study found Mga to affect the transcription of approximately 10% of the GAS genome during exponential phase of growth under nutrient rich conditions (THY) (229). However, the 'core' Mga regulon consists mostly of cell wall-associated surface molecules important for adherence to host tissues, internalization into non-phagocytic cells and avoidance of host immune responses. Mga transcriptionally activates genes that are in its immediate chromosomal region (Mga locus), which includes the *emm*-superfamily of genes (*emm*, *mrp*, *enn*, *arp*, *etc*.), *scpA*, *fba*, and *sic* (13). It also acts as an autoregulator (230) by directly binding to its promoter (231). Additional virulence factors outside of the "Mga locus" have also been found to be regulated, and these include *sof* and *sfbX*, *sclA* and *speB* (232, 233). In addition, microarray results found Mga to negatively regulate genes, albeit indirectly, but this was the first evidence that Mga can act as both an activator and a repressor. This study also revealed that Mga influenced several sugar transport and utilization operons as well as other metabolic processes such as amino acids and fatty acid biosynthesis (229).

Structural Characteristics

General specificities

Mga is a 530 amino acid protein with a molecular weight of 62 kDa, and is comprised of six domains (**Fig. 2.3**). There are two divergent alleles of *mga*, *mga-1* and *mga-2*, which possess 24.5% sequence identity at the nucleotide level. The *mga-1* allele is strongly associated with the class I/SOF$^-$ strains, and the *mga-2* allele is class II/SOF$^+$ strains (234). The amino-terminus contains a conserved Mga domain (CMD), followed by two helix-turn-helix DNA binding domains all of which are critical for binding to targets. Mga also contains two PRD-domains, PRD-1 and PRD-2, which are involved in detecting nutrient status and play an important role in its activity, and an EIIB^{Gat}-like domain at the C-terminus that appears to be important for dimerization. This section will provide more in-depth detail of each of these domains in the Mga protein.

DNA-binding

The N-terminal portion of the Mga protein encodes for two helix-turn-helix (HTH) DNA binding domains, HTH-3 and wHTH-4, which are critical for DNA binding. Mutants in HTH-3 were found to have a slight effect on binding to *emm* and *scpA*; however, those that lack wHTH-4 completely abrogate the ability of Mga to bind to its promoter targets (235). In addition, both HTH-3 and wHTH-4 are necessary for the transcriptional activation of the same virulence genes *in vivo* (236). These domains, in conjunction with the consensus Mga binding site (discussed further below), enable direct binding to target promoters and activation of the core regulon genes.

PRD domains

Both Mga alleles possess three histidine residues (PRD1: H204 and H270; PRD2: H324) that are conserved and ubiquitous among all sequenced GAS genomes. In addition, these histidine residues correlate with those present on AtxA, (discussed above). Hondorp *et al*. created an *in vivo* phosphorelay in order to verify that the PTS system of GAS can phosphorylate and modify the activity of Mga in an M1 and M4 serotype (representatives of Class I and Class II, respectively). Indeed, all three histidines can be phosphorylated by the PTS *in vivo*. The histidines on PRD1 are phosphorylated by P~His15-Hpr in the absence of glucose, and render the protein inactive; this is the dominant form of the protein. When no phosphorylation was present on any histidine, which occurs in the presence of the preferred sugar, glucose the phosphate is shunted to EIIB transporter, and Mga is active. Interestingly, when the histidine in PRD2 is phosphorylated, Mga showed an even greater increase in activity. However, overall inactivation of phosphorylation on PRD1 caused significant attenuation in a mouse model of disseminating subcutaneous skin infection, signifying the importance of proper expression of Mga and its regulon during infection (19).

Other studies have shown that altering the phosphorylation status of the M59 Mga using phosphomimetics on PRD1 (H207 and H273) and PRD2 (H327) were inhibitory to Mga-dependent gene expression. However, replacement with phosophoablative residues on PRD1/2 (H273A and H327A) was shown to relieve inhibition and restore the wild type phenotype. In virulence studies, the phosphomimetic (H273D) Mga was attenuated in mouse model of necrotizing

fasciitis whereas the phosphoablative (H273A) Mga was comparable to wild type in virulence (20). Furthermore, Mga59 showed an altered expression profile using phosphomimetics, and the PRD1 residue H273 was critical for proper regulon expression.

While Hondorp *et al*. showed that the ability for Mga to be phosphorylated/dephosphorylated on PRD1 is essential for full virulence (19), Sanson *et al*. showed that only phosphorylation on PRD1 caused an inhibitory effect of regulon expression (20). Regardless, the results from both studies confirms that Mga is PRD-containing virulence regulator (PCVR) that interacts with the PTS to sense carbohydrate availability and utilization within the GAS cell, and phosphorylation status of the protein is important for proper expression of the Mga regulon and thus full virulence for GAS.

C-terminal domain

The C-terminal region of Mga was once predicted to be a CheY-like receiver domain (RD), indicating that Mga could be part of a TCS (223). However, Ribardo *et al*. found that when 12 of the 13 response regulators from GAS TCS were mutagenized, there was no reduction in the levels of *emm* transcripts or Mga protein levels in comparison to wild type (233). More recently, protein alignment with another PRD-containing regulator, MtlR, revealed that the C-terminus of Mga contains structural homology to an $EIIB^{Gat}$-like domain. A study conducted by Hondorp *et al*. revealed that the C-terminal tail also acts a dimer interface for Mga, and increasing NaCl concentrations caused higher-order structures to form (dimers, tetramers, and hexamers). A truncation of the last 139 amino acids (Δ139-Mga),

which eliminates the putative $EIIB^{Gat}$-like domain, led to the loss of dimerization. While this truncation did not affect DNA binding, Mga failed to activate the transcription of its target genes (237). Taken together, the data suggests that DNA binding is important for Mga activity but not sufficient to induce full activity, and oligomerization is necessary for full activation of the Mga regulon.

Regulation of Mga

Optimal expression of the Mga regulon *in vitro* requires several growth conditions including elevated CO_2 levels, 37°C, increased iron levels, and metabolizable sugars (18, 238, 239). Mga is also directly phosphorylated by the PTS of GAS within conserved PRD residues, and altering virulence gene expression leading to attenuation *in vivo* (19). Other up stream regulators have been identified and include AmrA (hypothetical flippase involved in O-antigen/cell wall synthesis) (240), RivR (RALP IV; stand alone regulator) (177), TrxR (TCS response regulator) (160), CcpA (carbon catabolite repressor protein) (17), and DltA (D-alanylated lipoteichoic acid) (241); however, the exact mechanism by which they regulate Mga is not clear, and are often serotype specific.

Consensus Mga-binding site

Mga can directly bind to three different categories of promoters (Category A-C), which are separated based on size, number, and location of the Mga binding sites (MBS). In an M6 GAS background, Mga was found to bind to a single 45 base pair site 52 nucleotides upstream of the transcriptional start site of *emm* and *scpA* (235) (**Fig. 2.4**). Analysis of these two loci generated a Category A consensus, which is

associated with genes within the Mga locus. Further studies with P*emm1*, a Category A promoter identified critical nucleotides that were necessary Mga binding, Mga-dependent transcriptional activation, or both. This resulted in a minimal consensus of 35 base pairs (GGTCAAAAAGGTGGCAAAAGCTAAAAAAGCTGGTC) for P*emm1* that was just as effective as the original 45 base pair site (242).

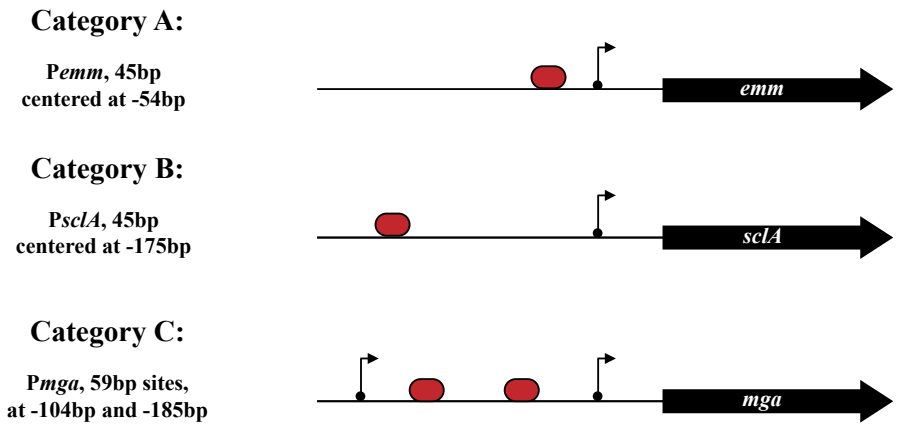


Figure 2.4. Promoter architecture of Mga regulated genes.
Three types of promoters regulated by Mga (shown in red) are categorized according to size, number and location from TSS (243).

Category B promoters are roughly 45 base pairs in length, and the MBS is centered at either 278 base pairs (P*sof*/s*fbX*) (244) or 153 bp (P*sclA*) (243) upstream of the regulated start of transcription (**Fig. 2.4**). Category B promoters are also associated with Mga regulated genes that are outside of the "Mga locus". Mga is also autoregulated in some serotypes and therefore binds to its own promoter (P*mga*) (11). However, the mechanisms for its binding are different than the other categories mentioned above. This type of binding, designated as Category C, contains two binding sites that are 59 bp in length and are centered around 108 bp and 180 bp upstream of the major P2 transcriptional start site (231). Although all three categories of MBS are considerably different at the DNA sequence level, they can compete with each other for binding of Mga *in vitro* (231, 235). The mechanism by which Mga is able to recognize and bind to these divergent binding sites is currently unknown.

Chapter 3: Materials and Methods

Bacterial Strains and growth

E. coli strain, media, and growth conditions

E. coli strains used in this study are listed in **Table 3.1**. Strains DH5α or C43 [DE3] were used as the host strain for plasmid construction and were cultured in Luria-Bertani (LB) medium (EMD Biosciences). C41 [DE3] was used for protein expression and cultured in ZYP-5052 auto-induction media (245). *E. coli* was grown at 37^{o}C with shaking under normal aerobic conditions. Growth was measured by a spectrophotometer (Ultraspec 2100 pro, Amersham Biosciences) at OD_{600}. Antibiotics for *E. coli* were used at the following concentrations: ampicillin 100 μg/ml, spectinomycin 100 μg/ml, kanamycin 50 μg/ml, chloramphenicol 30 μg/ml.

GAS strains, media, and growth conditions

GAS strains used and constructed in the study are listed in **Table 3.1**. GAS was cultured in Todd-Hewitt medium (Alpha Biosciences) supplemented with 0.2% yeast extract (Fisher Scientific) (THY). Growth was assayed by optical density with a Klett-Summerson colorimeter (A filter) and expressed in Klett Units. Alternatively, GAS were grown on blood agar plates overnight at 37^{o}C with 5% ambient CO_2. Colonies were scraped off the plate and resuspended in 10 ml of sterile saline. C Media was prepared using a previously described method (246). Prior to use, magnesium sulfate (0.4 mM), potassium phosphate (10 mM), and sodium chloride

(150 mM) were added and the pH of the media was buffered to 7.5 with concentrated HCl. Otherwise, a 2x Chemically Defined Media (CDM) (Alpha Biosciences) was prepared, as previously described (28). Prior to use, freshly prepared sodium bicarbonate (59.51 μM) and L-cysteine (11.68 μM) were added along with a carbohydrate and source at a final concentration of either 0.5% (glucose) or 1% (all other sugar sources).

Table 3.1: Bacterial strains and plasmids

Strain	Description	Reference
***E. coli* Strains**		
DH5α	*hsdR17 recA1 hyrA endA1 relA1*	(247)
C41 [DE3]	*F⁻ ompT gal dcm hsdS$_B$* ($r_B^-m_B^-$)(DE3)	(248)
C43 [DE3]	*F⁻ ompT gal dcm hsdS$_B$* ($r_B^-m_B^-$)(DE3)	(248)
BTH101	*F⁻ cya-99 araD139 galE15 galK16 rpsL1* (Str^R) *hsdR2 mcaA1 mcrB1*	Gift, D. Kearns
***S. pyogenes* Strains**		
5448	M1T1 GAS, clinical isolate, SpeB(+) SpeA(-) Sda1(low)	(85)
5448.930	Δ*mga* mutant in 5448 (insertional inactivation)	This study
5448.930$_R$	*mga*+ 5448, cured strain	This study
KM16.1	Δ*mga* mutant in 5448 (transposon)	This study
5448.Δ*fruR*	Δ*fruR* mutant in 5448 (allelic exchange)	This study
5448.Δ*fruB*	Δ*fruB* mutant in 5448 (allelic exchange)	This study
5448.Δ*fruA*	Δ*fruA* mutant in 5448 (allelic exchange)	This study
5448.Δ*fruR*$_R$	Δ*fruR* rescue strain in 5448	This study
5448.Δ*fruB*$_R$	Δ*fruB* rescue strain in 5448	This study
5448.Δ*fruA*$_R$	Δ*fruA* rescue strain in 5448	This study
GA40634	M4 GAS, clinical isolate	Georgia Emerging Infections Program (GaEIP);(229)
KSM547	Δ*mga* mutant in GA40634	(243)
SF370	M1 GAS	(249)
KSM165L	Δ*mga* mutant in SF370	(229)
Plasmids		
pACYC	Vector containing Cm^R gene	(250)
pCRS	Temperature-sensitive conditional vector; Sp^R	(15)
pJRS525	Replicating vector for GAS with Sp^R	(226)
p5448.ΔfruR	Δ*fruR* mutagenic plasmid; nonpolar *cat*	This study
p5448.ΔfruB	Δ*fruB* mutagenic plasmid; nonpolar *cat*	This study
p5448.ΔfruA	Δ*fruA* mutagenic plasmid; nonpolar *cat*	This study
pKSM201	Broad host range cloning vector, Km^R	(5)
pKSM720	GAS replicating plasmid with promoterless firefly luciferase (*luc*) and ribosomal binding site	(5)
pKSM807	WT *mga1* under native P*mga1*	(251)
pKSM808	WT *mga4-his6* (GA40634) under native P*mga4*	(251)
pKSM809	WT *mga1-his6* under native P*mga1*	(251)
pKSM813	Δ139 *mga1-his6* under native P*mga1*	(251)
pKSM871	A/A/A M4 Mga-His6	(251)
pKSM930	pCRS derivative for insertional inactivation of *mga*	This study
pKSM934	pUC19 containing P*mga1-mga*1 for SDM constructs	This study
pKSM935	GAS replicating plasmid P*mga1-mga1* F430Y	This study
pKSM936	GAS replicating plasmid P*mga1-mga1*S432G	This study
pKSM937	GAS replicating plasmid P*mga1-mga1*L409F	This study
pKSM938	GAS replicating plasmid P*mga1-mga1*D445V	This study
pKSM939	GAS replicating plasmid P*mga1-mga1*N453S	This study
pKSM940	GAS replicating plasmid P*mga1-mga1*M461I	This study
pKSM941	GAS replicating plasmid P*mga1-mga1*K425A	This study
pKSM942	GAS replicating plasmid P*mga1-mga1*E394A	This study

pKSM943	GAS replicating plasmid P*mga1-mga1*Y387A	This study
pKSM944	GAS replicating plasmid with full length P*fru* driving expression of *luc*	This study
pKSM945	GAS replicating plasmid with P*fru* no FruR consensus driving expression of *luc*	This study
pKSM948	GAS replicating plasmid with P*sloR* driving expression of *luc*	This study
pLZ12-Spec	Broad host range cloning vector, Sp^R	(252)
pMga1-His	WT *mga1* (SF370) with C-terminal His6	(251)
pT18N-link	N-terminal T18 expressing two-hybrid vector with linker region Ap^R	Gift, D. Kearns
pT25N-link	N-terminal T25 expressing two-hybrid vector with linker region Km^R	Gift, D. Kearns
pUC19	ColE1 ori Ap^R *lacZ*α	(253)
pT25N-F430Y	T25 fragment of *cya*- fused to Mga1-F430Y	This study
pT25N-S432G	T25 fragment of *cya*- fused to Mga1-S432G	This study
pT25N-N453S	T25 fragment of *cya*- fused to Mga1-N435S	This study
pT25N-M461I	T25 fragment of *cya*- fused to Mga1-M461I	This study
pT25N-L409F	T25 fragment of *cya*- fused to Mga1-L409F	This study
pT25N-D445V	T25 fragment of *cya*- fused to Mga1-D455V	This study

Modified Diauxic Growth

GAS strains were cultured on blood agar plates (TSA + 5% sheep blood) (Beckton Dickinson) overnight at 37°C with 5% ambient CO_2. Colonies were scraped off and resuspended in sterile saline. Klett flasks containing 150 ml of CDM + 0.5% glucose were inoculated to an OD_{600} = 0.05. Strains were grown to mid-log (T_o) then pelleted and washed in 10 ml of warm saline. The pellet was then resuspended in 1 ml of saline and split into two 500 µl aliquots for inoculation into fresh CDM + 0.5% glucose or CDM + 1% fructose (75 ml). A 10 ml sample of culture was harvested after T_{30min}, T_{1h}, and T_{2h} of growth for RNA extraction, and growth continued until samples reached stationary phase.

DNA Manipulations

Plasmid Isolation

Plasmid DNA was isolated from *E. coli* by alkaline lysis using the Wizard Plus SV Miniprep kit (Promega) or the Midiprep purification system (Qiagen) following the manufacture's instructions. Alternatively, DNA fragments were isolated from agarose gels using the Wizard SV Gel and PCR Clean-up kit (Promega).

Genomic DNA extraction

GAS genomic DNA (gDNA) was isolated as previously described (254). Briefly, 10 ml of an overnight culture grown at 37°C were pelleted by centrifugation for 10 min at 8,000 x *g*, 4°C, and the supernatant was discarded. Cells were then

resuspended in 10 mM Tris pH 8.0 supplemented with fresh lysozyme (100 mg/ml) and incubated at 37^{o}C for 2 h with vigorous agitation in a shaking heat block (Eppendorf). The proceeding extraction was then carried out following the manufacturer's instructions for gDNA isolation of Gram-positive bacteria using the MasterPure Complete DNA and RNA Purification kit (Epicentre).

Polymerase Chain Reaction (PCR)

PCR was performed using Phusion (New England Biolabs; NEB) or Accuprime *pfx* (Life Technologies) for cloning and Taq DNA polymerase (NEB) was used for diagnostic assays. To use Phusion DNA polymerase, annealing temperatures for primers (**Table 3.2**) were determined using the T_m calculator (http://tmcalculator.neb.com) selecting the HF buffer conditions. Amplification was carried out by an initial denaturation step at 98^{o}C for 4 min, followed by 30 cycles of denaturation at 98^{o}C for 10 s, a 30 s annealing step at the pre-determined temperature, and extension step of 72^{o}C for approximately 1 min/kb of DNA. A final extension step of 4 min was added for the completion of the reaction. For *pfx*, the parameters were as follows: An initial denaturation at 98^{o}C for 2 min followed by 30 cycles of denaturation at 98^{o}C for 30 s, a 30 s annealing step at 55^{o}C, and extension step of 68^{o}C for approximately 1 min/kb of DNA. PCR reactions were purified by the Wizard SV Gel and PCR Clean-up kit (Promega) according to the manufacturer's instructions. All DNA sequencing was performed by Genewiz, Inc.

Table 3.2: Primers used in this study

Target	Primer Name	Sequence (5'-3')	Reference
fruR	FruR RevTrans F	TATGTCAAACAAACGACCGA	This study
	FruR SP1	TCAGCACCTCCATGAACACG	This study
	FruR SP2	ATGTAGCCGTCCTTCTTGCT	This study
	FruR SP3	CCAACTCTCCTAGATCTCTA	This study
	oAXSpy0660.5	ccc**GGATCC**CAGAGCTAGTTCACGCTAAAG	This study
	oAXSpy0660.2	TATCCAGTGATTTTTTTCTCCATGATTAATTGTTTTCGTTTTGATTTT	This study
	oAXSpy0660.3	CAGGGCGGGGCGTAATTTATACCGTGACCTTAAACCC	This study
	oAXSpy0660.6	ccc**GGATCC**AATTCGGTCTCTTCTTGTGAC	This study
	FruR luc_1	ggg**GGATCC**TTGTGAGCCTAGATTATGAG	This study
	FruR luc_4	ggg**CTCGAG**GATTAATTGTTTTCGTTTTG	This study
fruB	FruB RevTrans F	TATGAAGAGTCAAAATCCAA	This study
	FruB RevTrans R	GGGAAAGCTTAGTTTTCAAA	This study
	oAXSpy0661.1	ccc**GGATCC**ATTACTCAACTCCTCTGAGTC	This study
	oAXSpy0661.2	TATCCAGTGATTTTTTTCTCCATAATTCAATCACCTTTGCCTTTTC	This study
	oAXSpy0661.3	GATGAGTGGCAGGGCGGGGCGGCTTTCTCTGATGACTTGGC	This study
	oAXSpy0661.4	ccc**GGATCC**TAGCTACTCCCATTTCTGCAG	This study
fruA	FruA RevTrans R	ATCATAAAGAAGAGATCCGT	This study
	oAXSpy0662.1	ccc**GGATCC**AAAGGTGATCAAGATACTCG	This study
	oAXSpy0662.2	TATCCAGTGATTTTTTTCTCCATCGTTTTTCTACCTCAACTTTATG	This study
	oAXSpy0662.3	GGCAGGGCGGGGCGTAAAAGGTAAAGAACTTTTTTCTTAC	This study
	oAXSpy0662.4	ccc**GGATCC**AAAATTAGTGAGTTGATACC	This study
mga	Mga InIn F	ccc**GGATCC**CAGTGGAGAGAACTAAAATT	This study
	Mga InIn R	ccc**GGATCC**TAATTCTTCCAAAGAGTTGA	This study
sloR	sloR InIn F	ggg**GGATCC**ATTCTTGTTTTAGCTATTTT	This study
	sloR InIn R	ggg**GGATCC**CCCAGCCACAAAACCTGGGA	This study
	sloR luc F	ggg**GGATCC**GTCAAACTTCCTATATCTATCTT	This study
	sloR luc R	ggg**CTCGAG**ATACGAACCTCCTCATTGATAATAT	This study
cat	oCatFwd	ATGGAGAAAAAAATCACTGGATA	This study
	oCatRev	TTACGCCCCGCCCTGCC	This study
	oCatSeq1	TTCCATGAGCAAACTGAAACG	This study
	oCatSeq2	CAGGTTTTCACCGTAACACG	This study
M13	1201	AACAGCTATGACCATGATTACG	Genewiz
	1211	GTTGTAAAACGACGGCCAGT	Genewiz
aad3	SPR1	ccc**CTCGAG**AATAAAGCCCCGCTTCGGCG	(15)
	SPR1	ccc**ACTAGT**AAGGACTTAAAAGAGCGTATTG	(15)

†Bold type denotes restriction site
* Underline segments denote overlap with *cat* cassette for PCR-SOEing
^ Lowercase letters denote nucleotides not complementary to target DNA

Enzymatic DNA Modifications

Enzymatic DNA modifications were performed using enzymes with conditions suggested by the manufacturer (NEB). Restriction digests were performed for 2 h at their optimal temperatures. Ligation reactions using T4 DNA ligase were set up using a 1:16 vector to insert ratio with overnight incubations at 16^{o}C. Antarctic phosphatase was used to dephosphorylate the vector ends for cloning with incubations at 37^{o}C for 1 h followed by 10 min inactivation at 80^{o}C. T4 polynucleotide kinase (PNK) was used to radioactively end-label probes at 25^{o}C for 10 min.

Bacterial Transformation

E. coli competent cells

To prepare DH5α or C43 [DE3] competent cells, 500 ml of LB broth was inoculated with a 5 ml overnight culture and grown to an OD_{600} of 0.5. Cells were placed on ice for 30 min prior to centrifugation at 7,000 x *g* for 30 min at 4^{o}C. After pelleting, cells were washed and resuspended three times in ice-cold sterile 10% glycerol. After the final wash, cells were resuspended in 1 ml of 10% glycerol and split into 75 µl aliquots. Aliquots were stored at -80^{o}C for no more than 1 year. GAS competent cells

To prepare GAS competent cells, 150 ml of THY broth containing 20 mM glycine and 85 µg/ml of hyaluronidase was inoculated with 7.5 ml of an overnight culture and incubated static at 37^{o}C until OD_{600} of 0.4. Cells were placed on ice prior

to centrifugation at 7,000 x *g* for 30 min. Cells were washed three times in 20 ml of 10% glycerol and resuspended in a final volume of 1 ml. Cells were split into aliquots of 175 µl and stored at -80^{o}C for no more than 6 months.

Electroporation

To remove excess salts prior to electroporation of DNA into either *E. coli* or GAS, DNA was drop dialyzed with dH_2O on a 0.025 µm membrane filter (Millipore) for 30 min for *E. coli* and 1 h for GAS. Electroporation of both *E. coli* and GAS was performed using a GenePulser Xcell (Bio-Rad). The 75 µl aliquot of *E. coli* along with the drop dialyzed DNA was added to a pre-chilled 2 mm cuvette and electroporated using the following settings: 2.5 kV, 200 Ω and 25 µF. Cells were immediately transferred to 1 ml of LB broth and outgrown for 1 h at 37^{o}C shaking, prior to centrifugation at 6,000 x *g*. Pelleted cells were resuspended in 100 µl of saline and plated on the appropriate antibiotic selection. For electroporation of GAS, the parameters were as follows: 1.75 kV, 400 Ω and 25 µF. Cells were transferred to 10 ml of THY and outgrown statically for 2 h at 37^{o}C. Cells were pelleted at 8,000 x *g* and resuspended in 100 µl of THY broth prior to plating on the appropriate antibiotic for selection.

Genetic Constructions

Plasmid constructions

Construction of pKSM930

An internal fragment of *mga* that began 30 nucleotides downstream of the ATG was amplified from 5448 gDNA using primers Mga_InIn_F and Mga_InIn_R (**Table 3.2**). The 300 bp fragment was digested with *BamH*I and ligated into *BamH*I-digested pCRS to create pKSM930 (**Table 3.1**).

Construction of Pfru-luc and PsloR-luc vectors, pKSM944, pKSM945, and pKSM948

The wild type P*fru* promoter was amplified from 5448 gDNA using primers FruR_luc_1 and FruR_luc_4 (**Table 3.2**). The 163 bp PCR fragment was digested with *Bgl*II/*Xho*I and ligated into a *BamH*I/*Xho*I-digested pKSM720 (**Table 3.1**) to form pKSM944.

A truncated P*fruR* was amplified from 5448 gDNA using primers FruR_luc_2 and FruR_luc_4 (**Table 3.2**), deleting the first 60 bp from the promoter region and eliminating the putative FruR binding site. The 103 bp promoter PCR fragment was digested with *Bgl*II/*Xho*I and ligated into a *BamH*I/*Xho*I-digested pKSM720 (**Table 3.1**) to form pKSM945.

Wild type P*sloR* was amplified from 5448 gDNA using primers sloR_luc_F and sloR_luc_R (**Table 3.2**). The 175 bp promoter PCR fragment was digested with *Bgl*II/*Xho*I and ligated into a *BamH*I/*Xho*I digested pKSM720 (**Table 3.1**) to form pKSM948.

Construction of pKSM934

A wild type P*mga1-mga1* was digested from pKSM807 (**Table 3.1**) using *BamH*I and *Xho*I, and ligated into a *BamH*I/*Xho*I-digested pUC19 to create pKSM934.

Construction of EIIB mutated complementation vectors, pKSM935-pKSM943

Using pKSM934 (**Table 3.1**) as a template, point mutations were made in the EIIB domain of Mga1 using the Quickchange II Site-Directed Mutagenesis kit (Stratagene). Briefly, oligonucleotides containing the mutated residue(s), each complementary to the opposite strands of the vector, were extended during 16 cycles by PfuUltra HF DNA polymerase (Stratagene) with an extension time of 5 min. To create point mutations, primer pairs Mga1-xxxx-L and Mga1-xxxx-R (**Table 3.2**) were used to generate the following amino acid replacements: K425A, E394A, Y387A, F430Y, S432G, L409F, D445V, N453S, and M461I. Following temperature cycling, the PCR product was DpnI treated for 1.5 h, followed by enzyme heat inactivation at 80^{o}C for 15 min. Then, 1.5 µl of the reaction was transformed into DH5α electrocompetent cells. Up to 14 clones were selected for sequence analysis for mutation verification. The confirmed mutants were then digested with *BamH*I/*Xho*I, gel purified, and the Pmga1-mga1 fragment was moved into pJRS525 digested with *BamH*I/*Xho*I in a DH5α background. The resulting plasmids were named pKSM935, pKSM936, pKSM937, pKSM938, pKSM939, pKSM940, pKSM941, pKSM942, and pKSM943 (**Table 3.1**).

Strain constructions

Inactivation of mga in 5448

The *mga* gene in the M1T1 strain 5448 (**Table 3.1**) was inactivated using the temperature-sensitive integration method as previously described (223). Briefly, plasmid pKSM930 (**Table 3.1**) was electroporated into 5448, followed by growth at 30^{o}C with spectinomycin selection. To allow for integration of the plasmid, cells were passed twice at 37^{o}C under spectinomycin selection. Integrants were screened for plasmid/genomic junctions by PCR using the primers Mga_InIn_F and M13R (**Table 3.2**).

Rescuing of 5448.930

The strain 5448.930 (**Table 3.1**) was cured of the plasmid inactivating *mga* by passage in liquid culture four times at 30^{o}C with no drug selection. Cultures were then plated, and patched on THY supplemented with spectinomycin and THY alone. Patches that showed spectinomycin sensitivity were then screened for loss of the spectinomycin gene by PCR using the primer SPR1 and SPR 2 (**Table 3.2**).

Transposon mutagenesis of mga in 5448

A transposon insertion in *mga* was identified from a pool of Krmit (for kanamycin-resistant transposon for massive identification of transposants) transposon mutants generated in 5448. Library #16 was screened using AP-PCR (254) followed by sequencing and alignment of sequence to the 5448 genome. Transposant #1 was mapped to the 5'-end of *mga* in the sense direction.

RNA Analysis

RNA extraction and DNase treatment

RNA was isolated using a Triton X-100 method as previously described (255). Briefly, 10 ml of THY broth was inoculated 1:20 from an overnight culture and grown to the appropriate Klett unit. Cells were pelleted by centrifugation at 8,000 x *g* for 10 min at 4^{o}C. Cells were then resuspended in TE buffer containing 0.2% Triton X-100 (v/v) and boiled for 10 min. A chloroform:isoamyl alcohol (24:1) extraction was performed twice, and ethanol was used to precipitate the RNA at -80^{o}C for 30 min. The reaction was then pelleted at 10,000 *x g* for 10 min and washed with 1 ml of 70% ethanol. The RNA was resuspended in DEPC-treated water and quantified using a spectrophotometer. The RNA was assessed for quality on formaldehyde gel: 18% (v/v) formaldehyde, 1% (w/v) agarose, 72% DEPC-treated water, and 10% (v/v) 10x MOPS.

RNA was then treated with Turbo DNA-*free* kit (Life Technologies) to remove gDNA from the sample. Approximately 20 µg of RNA was incubated with 1 µl (20 U) of Turbo DNase along with 1 µl of RNase Inhibitor (NEB) at 37^{o}C for 1 h as per manufacturer's recommendations. A 1x volume of the DNase Inactivation Reagent was added to the reaction and incubated at room temperature for 2 min with occasional mixing. The sample was then pelleted at 10,000 *x g* for 1.5 min at room temperature, and the supernatant containing the DNAsed RNA was transferred to a new 1.5 ml tube and stored at -80^{o}C for no more than 1 year. A diagnostic PCR reaction was performed on 1.5 µl of the DNased RNA to verify complete degradation of gDNA alongside a denaturing RNA gel to measure RNA integrity.

RNA-Seq

For RNA sequencing (RNA-Seq), total RNA was extracted using the Direct-zol RNA MiniPrep kit (Zymo Research) using a modified procedure to improve GAS cell disruption. Cells from frozen pellets were resuspended in 700 µl of Trizol supplemented with approximately 300 mg of acid-washed glass beads (Sigma Life Science) and disrupted by vortexing for 5 min. Beads were collected by brief centrifugation and cell lysate was used for RNA purification as recommended by the manufacturer. RNA samples were treated with the Turbo DNase-free kit (Life Technologies) to avoid gDNA contamination. A total of 5 µg of DNased RNA was treated for Ribosomal RNA (rRNA) removal using the Ribo-Zero Magnetic kit (Epicentre) for Gram-positive bacteria and rRNA-depleted RNA was then purified with the RNAClean XP kit (Agencourt). Sample quality was assessed using a 2100 Bioanalyzer (Agilent) and sample quantity was determined using a NanoDrop 8000 spectrophotometer (Thermo Scientific).

RNA-Seq directional libraries were generated using the ScriptSeq v2 RNA-Seq Library Preparation kit (Illumina) according to the manufacturer's recommendations. Briefly, 45 ng rRNA-depleted RNA was fragmented and used for reverse-transcription with random primers containing a 5'-tagging sequence. The 5'-tagged cDNA was then modified at its 3' end by a terminal-tagging reaction to generate di-tagged, single-stranded cDNA that was then purified using the AMPure System (Agencourt). The purified di-tagged cDNA was used as a template to generate second strand cDNA containing Illumina adaptor sequences, to incorporate index barcodes and amplify the library by limited-cycle PCR. The resulting RNA-Seq

libraries were purified using the AMPure System (Agencourt) and RNA-Seq library quality was verified as described above. A rapid-run 100 bp single-read DNA sequencing was then performed at the Institute for Bioscience and Biotechnology Research (IBBR) Sequencing Facility at the University of Maryland, College Park using the Illumina HiSeq 1500 platform. Data were generated in the standard Sanger FastQ format and raw reads were deposited with the Sequence Read Archive (SRA) at the National Center for Biotechnology Institute.

Read quality was measured using FastQC (256), filtered and trimmed using trimmomatic (257), and mapped against the MGAS5005 genome (accession CP000017) using bowtie (258), bowtie2 (259), and tophat (260) with options to allow 0,1,2 mismatches and randomly map multihit reads. The resulting alignments were converted (261) and counted (262) by coding and intergenic region. Initial visualizations of the sequencing mapping were performed using the Integrative Genomics Viewer (IGV) (263). Differential expression analyses were performed after library quantification, normalization, and batch effect estimation using limma (264) and/or DESeq (265). The resulting metrics of expression were visualized using KEGG (266) and circos (267) as well as tested for ontology enrichment using goseq (268, clusterProfiler {Yu, 2012 #1997), GOstat (269), and topGO (270).

Quantitative Real-time PCR

The quantitative real-time PCR (qRT-PCR) experiments were performed as follows: 25 ng of DNase-treated RNA was added to a SYBR green master mix (Life Technologies) with 0.25 µl of of each gene-specific real-time primer from a 20 nM stock. cDNA was generated using 5.75 U of Multiscribe reverse transcriptase (Life

Technologies) using the one-step protocol on a Light Cycler 480 (Roche). The reaction conditions were as follows: 48°C for 30 min, 95°C for 10 min, and 40 cycles of 95°C for 15 s followed by 60°C for 1 min.

Alternatively, for any RNA-Seq validation, cDNA was generated using a two-step protocol in order to mimic the strand-specificity of the RNA-Seq results. Briefly, 100 ng of DNase-treated total RNA was incubated with a gene-specific primer using the manufacturers recommendations (Quanta Biosciences). 25 ng of cDNA was then added to a SYBR green master mix (Quanta Biosciences) with 0.25 µl of gene-specific primers from a 20 nM stock. The reaction conditions were as follows: an initial denaturation at 95°C for 3 min, 40 cycles of 95°C for 15 s followed by 60°C for 45 s. Correlation coefficients for RNA-Seq were determined by plotting the log value of the array on the X-axis to the log value of the quantitative real-time RT-PCR (qPCR) on the Y-axis. Linear regression was used to determine the line of best fit, and the resulting R^2 value was calculated, which represented the fitness of the data.

For all experiments, standard error was calculated from three biological replicates, and differences over 2-fold in expression were considered significant. Transcript levels presented represent ratios of wild type/experimental relative to the level of *gyrA* transcript as an internal control. All real-time primers were designed using the interactive tool Primer3 (http://biotools.umassmed.edu/bioapps/primer3_www.cgi) (**Table 3.3**).

Table 3.3: Real-time qPCR primers

Target	Primer	Sequence (5'-3')	Reference
adh.2	adh.2 M1T1 RT L	ATGGACTCGGTCACTCAGC	This study
	adh.2 M1T1 RT R	ATTCCAGATGACGCGGATA	This study
amyA	amyA M1T1 RT R	TGGGTGATGTTTTTGAGATGG	(5)
	amyA M1T1 RT L	GTTTGGGTACTTGGCAATGG	(5)
emm	emm1 RT-L	ACTCCAGCTGTTGCCATAACAG	(233)
	emm1 RT-R	GAGACAGTTACCATCAACAGGTGAA	(233)
fruA	fruA M1 RT L	ACCGGGCTTAGTAGCTGGT	This study
	fruA M1 RT R	ACTTCTCCTCCTGCTGCAA	This study
fruB	fruB M1 RT L	GTTCAGCGCCAGCTAATCT	This study
	fruB M1 RT R	TCACAAACCACCTGAGCAC	This study
fruR	fruR M1 RT L	GTCAAACAAACGACCGATC	This study
	fruR M1 RT R	TCCAAGAAATGCCTTGTCA	This study
glnH	glnH M1T1 RT L	TGAGGGCATTTCTACTGGTG	This study
	glnH M1T1 RT R	GCAGCTAAGCCATTGTTGAA	This study
gyrA	gyrA M1 RT-L	CGACTTGTCTGAACGCCAAAGTC	(233)
	gyrA M1 RT-R	ATCACGTTCCAAACCAGTCAAAC	(233)
lacD.2	lacD.2 M1 RT-L	GCTGCTTGTGGTCCTTCTTC	(160)
	lacD.2 M1 RT-R	TTGGTGTTTGCAGCAGAATC	(160)
mac	mac M1T1 RT L	ACGGCTACCGCCTTAGTCT	This study
	mac M1T1 RT R	ACGGCGTCAAAATACCAC	This study
malX	malX M1T1 RT R	TTTGCTTTTGCCTCTGAACC	(5)
	malX M1T1 RT L	CCATAACCGGCAATTAAACC	(5)
mga	mga M1 RT-L	CGCTGAGTTGAGCCTGATTTC	(160)
	mga M1 RT-R	AGACTCACCAACGGGCTGTC	(160)
nagB	nagB M1T1 RT L	TGTTTTGGCACATCCTCAAT	This study
	nagB M1T1 RT R	AAGAAGAAACGCATGTGGTG	This study
opuAA	opuAA M1 RT-L	TGATTTGCAAGACAGCATGA	(160)
	opuAA M1 RT-R	CATCAAAGCAATCCGATCAC	(160)
ptsA	ptsA M1T1 RT R	TTCATTACAGCGGTTACCATTTTC	(5)
	ptsA M1T1 RT L	TGAGCAGGGCCTTATTTTTACTG	(5)
sagC	sagC M1T1 RT L	TGCTGATGCTTGAGGATGAC	(5)
	sagC M1T1 RT R	GCGCCATAAAACGGAAATTA	(5)
salB	salB M1T1 RT L	TGGTGTTTTAGCAGCATTTT	This study
	salB M1T1 RT R	AATCTCCGCCTGAAACAAT	This study
sclA	sclA M1 RT-L	TGCTGACAAAGAAGCTAACCAAAC	(233)
	sclA M1 RT-R	CGTCTGTGGTTGTTGGCTACAG	(233)
sdaD2	sdaD2 M1T1 RT L	AGCAAACCAGAGCCAAGTG	This study
	sdaD2 M1T1 RT R	CACATTTTTGGGGCTATGC	This study
sdhA	sdhA M1T1 RT L	TGCAGGAGCTTTTGGTCTT	This study
	sdhA M1T1 RT R	GCAGCTGCAGAACCAACTT	This study
sic	sic M1 RT-L	GGTAAAGTAGGCGGACATGCC	(233)
	sic M1 RT-R	CCTCGTGTGCCAGAAAAACC	(233)
ska	ska M1T1 RT L	TGCAAACGGCAAGGTCTAC	This study
	ska M1T1 RT R	GCGCACATGTCCCTTTAAC	This study
slo	slo M1 RT-L	AAAACAAACCAGACGCGGTAGT	(233)
	slo M1 RT-R	TTGTCTCCCATACCTGGTAAATCA	(233)
sloR	sloR M1T1 RT L	GTGAAGTAGGCGGAGCAAA	This study
	sloR M1T1 RT R	CGGCATAACCTGGAGACAC	This study
speB	speB M1T1 RT L	GGTAAAGTAGGCGGACATGCC	(160)
	speB M1T1 RT R	CACCCCAACCCCAGTTAACA	(160)
spxA.2	spxA.2 M1T1 RT L	ATGCCAAAGCTCTCGATTG	This study
	spxA.2 M1T1 RT R	TAAGGATGCGCGGGTTAT	This study

Reverse Transcriptase PCR (RT-PCR)

Reverse transcriptase PCR (RT-PCR) was carried out on 500 ng of total RNA with M-MuLV reverse transcriptase (NEB) according to the manufacturer's protocol, using primers complementary to the RNA. Reverse transcription was performed for 1 h at 42°C using 1 µl of cDNA as a template in reaction, followed by enzyme inactivation at 90^{o}C for 10 min. The subsequent PCR was done with the following parameters: 5 min at 95°C (initial activation); 30 cycles of 30 s at 95°C, 30 s at 55°C, and 1.5 min at 72^{o}C (PCR); and a final extension step of 5 min at 72^{o}C. PCR reactions were performed on the RT+ and the RT- cDNA samples, as well as on gDNA, using Taq polymerase. The following primer pairs were used to amplify intergenic regions with the *fruRBA* operon: *fruR-fruA* (FruR_RevTrans_F and FruA_RevTrans_R); *fruB-fruA* (FruB_RevTrans_F and FruA_RevTrans_R); *fruR-fruB* (FruR_RevTrans_F and FruB_RevTrans_R) (**Table 3.1**).

5' RACE

The 5'/3' Rapid Amplification of cDNA Ends (RACE) kit (Roche) was used according to the supplier's instructions. A 500 ng aliquot of total DNase-treated RNA was used to obtain the cDNA by primer extension with primer complementary to the sequence (fruR_SP1; **Table 3.1**). Following the 3' tailing reaction with dATPs, the cDNA was amplified by PCR using the reverse primer (fruR_SP2; **Table 3.1**) and oligo(dT)-anchor (forward primer; supplied with the kit). The 5' end of the transcript was then determined by sequencing the PCR product using the primer (fruR_SP3; **Table 3.1**).

Protein Analysis

Mga-1 purification and expression

GAS Mga1-His_6 was purified as previously described (251). Briefly, *E. coli* containing the plasmid pMga1-His was grown in ZYP auto-induction media for ~50 h at 37°C. Cells were harvested by centrifugation at 4°C. Pellets were then resuspended in Lysis Buffer A (20 mM NaP_i, 500 mM NaCl, 20 mM imidazole, pH 7.4) and 1x EDTA-free cOmplete protease inhibitor (Roche), and incubated on ice with lysozyme for 30 min followed by sonication using a Branson Sonifier 450 with a tapered microtip (setting 6, 50% duty cycle) pulsing 3 x 1 min with 3 min breaks on ice between cycles. The lysate was spun for clarification at 20,000 x *g* for 30 min at 4°C 3-4 times. The lysate was loaded on a 5 ml NiNTA agarose resin (Qiagen) and rotated at 4°C for 1 h. The column was then washed twice with NiNTA Wash Buffer (50 mM NaH_2PO_4, 300 mM NaCl, 20mM imidazole, pH 8.0). Protein was eluted with 10 ml of NiNTA elution buffer (50 mM NaH_2PO_4, 300 mM NaCl, 250 mM imidazole, pH 8.0) and chosen fractions were stored at -20°C. Protein concentration was analyzed using the Bio-Rad protein assay kit, reading the absorbance at 595 nm on a spectrophotometer.

GAS whole cell lysates

Soluble GAS protein extractions were performed using the bacteriophage lysine, PlyC (kindly provided by D. Nelson) as previously described (251). Briefly, GAS cells were inoculated 1:20 into 12 ml of THY or C media (246), grown to late-log, and pelleted by centrifugation for 15 min at 6,000 x *g*. Pellets were resuspended

in 200 µl of Buffer D (30 mM Tris pH 7.5, 0.1 mM d*ithiothreitol [DTT],* 40% (v/v) glycerol), 1x cOmplete EDTA-free protease inhibitor (Roche), 10 µl (20 U) of Turbo DNase (Life Technologies), and 5 µl (250 U) of PlyC. The cells were mixed by flicking and incubated on ice for 20 min, followed by centrifugation at 13,000 x *g* for 10 min at 4°C. Clarified supernatants containing soluble proteins were extracted and boiled with 3x cracking buffer (0.15% [w/v] Bromophenol Blue, 0.6% [w/v] SDS, 3 ml glycerol, 3.9 ml Tris-HCl [500 mM, pH 6.8], 1.5 ml β-mercaptoethanol, and 4 ml dH_2O) at 95°C for 5 min, and stored at -20°C. Protein concentration was assayed using the Bio-Rad protein assay kit, reading the absorbance at 595 nm on a spectrophotometer.

SDS-PAGE Analysis

SDS-polyacrylamide gels were made as follows: a 10% resolving gel was mixed and poured followed by a 6% stacking gel using the reagents found in **Table 3.4**. Protein samples were mixed with 1x cracking buffer, boiled at 95°C, and loaded into the wells. SDS-PAGE was run at 180V for 1 h.

Western Blots

Selected protein samples were run on a 10% SDS-PAGE, with 6% stacking gel, for 1 h at 180V. The gels were then transferred to nitrocellulose membranes using the Mini-Protean apparatus (Bio-Rad) in 1x transfer buffer (25 mM Tris base, 0.2 M glycine, 20% methanol). Membranes were blocked overnight at 4°C in blocking solution (5% (w/v) dried milk in PBS-tween). Primary antibodies were incubated at a 1:1,000 dilution, unless otherwise stated. The following primary antibodies were used

at a 1:1,000 dilution: α-His antibody (Roche), α-Mga4, α-Mga1-HRP, and α-Sda1. Blots were then washed three times for 5 min in PBS-tween. The primary antibody Hsp60 (Enzo Life Sciences) was used at a 1:2,500 dilution. Blots were then incubated with secondary antibodies: α-rabbit-HRP (1:5,000) or α-mouse-HRP (1:20,000) for 1 h. The blots were then washed with PBS-tween three times for 5 min, and visualized using SuperSignal West Femto Substrate (Thermo Scientific) and a LAS-3000 CCD camera (FUJIFILM).

Table 3.4: SDS-PAGE buffer

10- 15% Resolving gel	**6% Stacking gel**
26% (v/v) Lower gel stock 25-37.5% (v/v) 40% Acrylamide/ bis sol. 0.5% (v/v) 10% ammonium persulfate 0.1% (v/v) tetramehylethylene diamine Brought up with H_2O	25% (v/v) Upper gel stock 15% (v/v) 40% Acrylamide/ bis sol. 0.5% (v/v) 10% ammonium persulfate 0.1% (v/v) tetramehylethylene diamine Brought up with H_2O

Lower gel stock	**Upper gel stock**
1.5M Tris HCl pH 8.8 0.4% (w/v) SDS Brought up with H_2O	0.5M Tris-HCl pH 8.8 0.4% (w/v) SDS Brought up with H_2O

SDS running buffer	**Transfer buffer**
25mM Tris 190mM Glycine 0.1% (w/v) SDS Brought up with H_2O	25mM Tris base 0.2M glycine Brought up with 20% methanol

5x Cracking buffer	**PBS-Tween**
0.3M Tris pH 6.8 25% (v/v) 2-mercaptoethanol 51% (v/v) glycerol 10% (w/v) SDS 0.01% (w/v) bromophenol blue Brought up with H_2O Heated at 55ºC Heated at 55ºC	0.01M PBS pH 7.4 0.5% Tween-20 Brought up with H_2O

Phos-tag phosphate affinity gel

The Phos-tag gel system was used to visualize the phosphorylation state of Mga. Briefly, GAS whole cell lysates were run on a Mn^{2+}-Phos-tag 10% SDS-PAGE gel (WAKO), with 6% stacking gel. The resolving gel was made as follows: 1.75 ml of 1.5 M Tris-HCl, 1.75 ml of 40% bis-acrylamide, 140 µl of 10 mM MnCl, 140 µl of phos-tag solution, 3.1 ml of dH_2O, 100 µl of 10% ammonium persulfate, and 20 µl of TEMED (Bio-Rad). Protein samples were run for 2 h at room temperature and 1 h on ice, for a total of 3 h at 180 V.

Transcriptional Reporter Assay

Luciferase Assay

For strains containing the P*fru-luc* or the P*sloR-luc* constructs, luciferase assays were performed as follows: 5448 (WT), 5448.Δ*fruR* (*fruR* mutant) were transformed with pKSM944 and pKSM948 and grown static in 12.5 ml CDM supplemented with the appropriate carbon source under spectinomycin selection at 37ºC. The construct pKSM944 is a P*fru-luc* reporter plasmids, and pKSM948 is a P*sloR-luc* (**Table 3.1**) reporter plasmid that are both capable of replicating in GAS. 500 µl samples were taken at mid-logarithmic (Klett 50) and late-logarithmic (Klett 80) phase, centrifuged, and pellets were placed at -20ºC overnight. The luciferase assay was performed using the Luciferase Assay system (Promega). Pellets were resuspended in various amounts of 1x Lysis buffer to normalize for cell unit according to the equation 4.5 = (x ml)((Klett 65)/2). The luciferase assay was read

using a Centro XS^3 LB 960 luminometer (Berthold Technologies) after 50 µl of Luciferin-D reagent was directly injected.

Virulence Assays

Lancefield bactericidal assay

The ability of GAS strains to survive in whole human blood was tested as previously described (271). Briefly, GAS were grown to early mid-exponential phase (OD_{600} ~ 0.15) and serially diluted in saline. Blood donation is approved by the University of Maryland Institutional Review Board (IRB) (protocol 10-0735) with approved consent from donors; records have been archived. A 50 µl volume of a 10^{-4} dilution (ca. 50 to 200 CFU) was added to 500 µl of fresh heparinized whole human blood and rotated for 3 h at 37°C.

A murine blood bactericidal assay was performed as follows: Whole blood was obtained from 8- to 10-week old female CD1 mice (weight > 25g; Charles River Laboratories). The mice were anesthetized with ketamine, then a blood volume ≥ 500µl was withdrawn by terminal cardiac puncture and the blood was immediately transferred to a heparinized vacutainer (Beckton Dickinson). Cervical dislocation was performed as a secondary method of euthanasia to ensure animals were deceased following cardiac puncture. A 30-µl volume of a 10^{-4} dilution (ca. 50 to 200 CFU) was added to 300 µl of fresh whole mouse blood and rotated for 3 h at 37°C.

The multiplication factor (MF) was calculated for both assays by dividing the CFU obtained after blood challenge by the initial CFU inoculated. Data are presented as percent growth in blood corresponding to the MF of the mutant divided by the MF

of the wild type (WT) × 100. The human bactericidal assay MF of each strain was calculated from 6 biological replicates, whereas the murine bactericidal assay obtained blood from multiple mice (n= 4-6) in duplicate. Significance (*p*-value) was determined using a unpaired Student *t* test.

Opsonophagocytic killing assays

Isolated PMNs or differentiated HL60 cells were seeded at a density of 10^6 cells/ml in 24-well plates with RPMI 1640 cell culture medium supplemented with 2.05 mM L-Glutamine. GAS was grown overnight, and cultures were diluted into fresh THY and grown to mid-log phase (OD_{600} ~ 0.4). Bacteria were opsonized prior to neutrophil challenge by re-suspension in donor plasma (PMN assays) or FBS (HL60 assays) for 30 min at 37˚C. GAS were then added to seeded neutrophils to desired multiplicity of infection (MOI = 0.1 unless otherwise indicated) in a final volume of 1 ml RPMI 1640 cell culture medium (HyClone) + 2.05 mM L-Glutamine, 20% plasma or FBS. Neutrophil-challenged bacteria were incubated at 37˚C in 5% CO_2 for 2 h. GAS were also incubated in RPMI 1640 + 2.05 mM L-Glutamine, 20% plasma or FBS in the absence of PMNs, monocytes or differentiated HL60 cells for the purpose of survival comparison.

Following incubation, surviving GAS were harvested by collecting supernatants and neutrophils. Neutrophils were immediately lysed by re-suspension in sterile H_2O and the intracellular contents pelleted and recombined with corresponding well supernatants for plating on THY agar to obtain total viable bacterial counts after overnight incubation at 37˚C; 5% CO_2. Resistance of GAS to opsonophagocytic killing following neutrophil challenge was assessed by comparing CFUs obtained

from plating of viable bacteria isolated from killing assays to CFUs obtained from GAS incubation in cell culture media in the absence of neutrophils $(\frac{CFU\ obtained\ in\ PMNs}{CFU\ obtained\ in\ media} \times 100)$. All survival indices are normalized to survival of wild type 5448 (100%) in the condition tested. Data presented is the result of at least 3 biological replicates, each performed in triplicate with *p* values determined by unpaired *t* test.

Murine intraperitoneal model of infection

All animal work was performed in AAALAC-accredited ABSL-2 facilities at the University of Maryland following IACUC-approved protocols (R12-112) for humane treatment of animal subjects in accordance with guidelines set up by the Office of Laboratory Animal Welfare (OLAW) at NIH, Public Health Service, and the Guide for the Care and Use of Laboratory Animals. Every effort to limit distress and pain to animals were taken.

An overnight culture (10 ml) was used to inoculate 80 ml of THY and incubated static at 37°C until late-logarithmic phase (Klett 100). Briefly, 6- to 7-week-old female CD-1 mice (Charles River Laboratories) were injected in the peritoneal cavity with 100 µl of 4 x 10^9 CFU/ml dilution (4 x 10^8 CFU). Mice were monitored three times daily for a period not exceeding 72 h and were euthanized using CO_2 asphyxiation consistent with the recommendations of the Panel on Euthanasia of the American Veterinary Medical Association upon signs of systemic morbidity (hunching, lethargy, hind leg paralysis). Survival data were assessed by Kaplan-Meier survival analysis and tested for significance by log rank test. Data shown represent 10 mice for each strain.

Chapter 4: The *fruRBA* operon is necessary for Group A Streptococcal growth in fructose and for resistance to neutrophil killing during growth in whole human blood

DOI: 10.1128/IAI.01296-15.

Introduction

Bacterial pathogenesis is intimately linked to the availability of nutrients that a pathogen encounters in the host, such as preferred carbohydrate sources for carbon and energy. This paradigm holds true for Gram-positive pathogens in the phylum *Firmicutes*, and is discussed extensively in Chapter 2. The main system responsible for the appropriate coordination of virulence gene expression in response to available sugars is the phosphoenolpyruvate (PEP)-phosphotransferase system (PTS) (detailed in Chapter 2). The Group A Streptococcus (GAS) utilizes the PTS as an environmental sensor in order to coordinate the expression of virulence factors. Examples of such synchronization include: the stand-alone regulator Mga (19), LacD.1 (4), and CcpA and PtsI (5-8). Taken together, these findings point to the direct link between GAS carbohydrate metabolism and virulence factor production.

Several PTS EIIs have been implicated in the pathogenesis of low-G+C Gram-positive bacteria, including the fructose-specific PTS system (Fru) in

pathogenic species of *Streptococci*. In the oral pathogen *S. gordonii,* the fructose EIIABC (*fruA*) has been linked to biofilm formation (272). In the zoonotic pathogen *S. iniae*, *fruA* was found in the genomes of five virulent strains, but was absent in non-virulent strains (273). In GAS, a transposon-site hybridization (TraSH) screen found *fruA* as a gene critical for the survival of GAS in whole human blood (15). Transcription of *fruA* was also induced in an analysis of global heme stress in GAS (274), a condition likely encountered during bloodstream infection.

Despite clear links to pathogenesis, the exact role of FruA and fructose utilization in the physiology and virulence of GAS has yet to be examined. *fruA* encodes one of the 14 EII PTS sugar-transporters found in the GAS M1T1 genome (197), and is predicted to be located within a three-gene *fruRBA* operon (*fruRBA*) conserved across low G+C Gram-positive bacteria. In the model organism *Lactococcus lactis,* the *fruRKA* operon is composed of *fruR*, encoding a DeoR family transcriptional repressor; *fruK* (*fruB*), a 1-phosphofructokinase; and *fruA*, an EIIABC PTS sugar transporter (275).

In this study, we have characterized the *fruRBA* operon in the M1T1 strain 5448, a representative of one of the most prevalent serotypes of GAS isolated from invasive forms of infections worldwide. We found that expression of the *fruRBA* genes are highly induced as an operon in the presence of fructose, and FruR acted as a repressor of the operon likely in conjunction with CcpA (CCR). Growth and metabolism studies demonstrated that FruA was the likely primary transporter of fructose for GAS and that all three *fru* genes were important for growth in fructose. Interestingly, *fruR* and *fruB* are important for GAS survival in whole human blood

and for evasion of human neutrophils whereas the FruA EII transporter was not. Surprisingly, neither growth in whole mouse blood nor intraperitoneal (*i.p.*) infection in mice was attenuated by mutations in the *fruRBA* operon. These data strongly suggest FruB, and its regulation by FruR, affect a human-specific mechanism for evasion of neutrophil killing independently of FruA fructose transport.

Results

Transcriptome of M1T1 5448 growing with fructose as the sole carbon source

Our previous genome-wide screen to identify genes required for GAS 5448 fitness in whole human blood identified the putative PTS fructose EII gene *fruA* (15), suggesting a role for fructose utilization in host survival. To determine the effect of fructose on GAS gene expression, we conducted RNA-sequencing (RNA-Seq) in order to assess how fructose affects global transcription. Total RNA was isolated from the wild type 5448 grown in either chemically defined media (CDM) supplemented with 0.5% glucose or 1% fructose to late logarithmic phase in two biological replicates and processed for RNA-Seq. Data obtained from fructose-grown cells was compared to glucose-grown cells (WT-glu/WT-fru) with changes in gene expression over two-fold ($\log_2 \leq -1.0$ or ≥ 1.0) with a p value ≤ 0.05 were considered significant (**Table 4.1**).

Table 4.1: Genes in 5448 differentially expressed in CDM-glu compared to CDM-fru

Spy #	Annotation	Gene name	Log_2 FC[a]
Induced by Fructose			
M5005_Spy_0039	alcohol dehydrogenase/acetaldehyde dehydrogenase	*adh.2*	-2.10
M5005_Spy_0123	translation initiation inhibitor		-1.82
M5005_Spy_0124	**serine catabolism regulator; FruA-like EIIC domain protein**	***sloR***	**-2.32**
M5005_Spy_0125	hypothetical protein		-1.09
M5005_Spy_0126	V-type sodium ATP synthase subunit I	*ntpI*	-2.01
M5005_Spy_0127	V-type sodium ATP synthase subunit K	*ntpK*	-1.97
M5005_Spy_0128	V-type sodium ATP synthase subunit E	*ntpE*	-2.22
M5005_Spy_0129	V-type sodium ATP synthase subunit C	*ntpC*	-1.79
M5005_Spy_0130	V-type sodium ATP synthase subunit F	*ntpF*	-1.22
M5005_Spy_0131	V-type sodium ATP synthase subunit A	*ntpA*	-1.24
M5005_Spy_0132	V-type sodium ATP synthase subunit B	*ntpB*	-0.96
M5005_Spy_0133	V-type sodium ATP synthase subunit D	*ntpD*	-1.21
M5005_Spy_0150	PTS system, 3-keto-L-gulonate specific IIA	*ptxA*	-0.99
M5005_Spy_0474	transcription antiterminator, BglG family	*licT*	-1.36
M5005_Spy_0476	6-phospho-beta-glucosidase	*bglA*	-1.22
M5005_Spy_0521	PTS system, N-acetylgalactosamine-specific IIB	*agaV*	-1.75
M5005_Spy_0660	**fructose repressor**	***fruR***	**-1.95**
M5005_Spy_0661	**1-phosphofructokinase**	***fruB***	**-2.52**
M5005_Spy_0662	**PTS system, Fructose-specific IIABC**	***fruA***	**-3.33**
M5005_Spy_0798	IFN-response binding factor 1		-1.01
M5005_Spy_1062	maltodextrose utilization protein	*malA*	-1.12
M5005_Spy_1063	ABC Cyclomaltodextrin permease protein	*malD*	-1.17
M5005_Spy_1081	PTS system, Cellobiose-specific IIA component		-0.98
M5005_Spy_1307	hypothetical membrane spanning protein		-1.02
M5005_Spy_1308	ABC unknown sugar-binding protein		-1.16
M5005_Spy_1309	ABC unknown sugar permease protein		-1.82
M5005_Spy_1310	ABC unknown sugar permease protein		-1.40
M5005_Spy_1395	tagatose-bisphosphate aldolase; SpeB regulator	*lacD.1*	-1.01
M5005_Spy_1396	tagatose-6-phosphate kinase; regulatory	*lacC.1*	-1.35
M5005_Spy_1397	galactose-6-phosphate isomerase lacB subunit; regulatory	*lacB.1*	-1.54
M5005_Spy_1400	PTS system, Galactose-specific IIB component		-1.45
M5005_Spy_1401	PTS system, Galactose-specific IIA component		-1.14
M5005_Spy_1509	pyruvate, phosphate dikinase		-1.36
M5005_Spy_1510	pyruvate, phosphate dikinase		-1.26
M5005_Spy_1629	lantibiotic transport ATP-binding protein	*salX*	-1.89
M5005_Spy_1631	lantibiotic salivaricin A	*salA*	-1.67
M5005_Spy_1632	6-phospho-beta-galactosidase	*lacG*	-1.66
M5005_Spy_1633	PTS system, Lactose-specific IIBC component	*lacE*	-2.73
M5005_Spy_1634	PTS system, Lactose-specific IIA component	*lacF*	-2.22
M5005_Spy_1635	tagatose-bisphosphate aldolase	*lacD.2*	-2.18
M5005_Spy_1636	tagatose-6-phosphate kinase	*lacC.2*	-1.49
M5005_Spy_1637	galactose-6-phosphate isomerase lacB subunit	*lacB.2*	-1.67
M5005_Spy_1638	galactose-6-phosphate isomerase lacA subunit	*lacA.2*	-1.50
M5005_Spy_1663	PTS system, Mannitol-specific IIB component		-1.69
M5005_Spy_1744	PTS system, Cellobiose-specific IIC component	*celC*	-1.34
M5005_Spy_1745	PTS system, Cellobiose-specific IIB component	*celB*	-0.95
M5005_Spy_1841	L-serine dehydratase	*sdhB*	-1.06
M5005_Spy_1842	L-serine dehydratase	*sdhA*	-1.27
Repressed by Fructose			
M5005_Spy_0667	exotoxin type C precursor		2.07
M5005_Spy_1139	glucosamine-6-phosphate isomerase	*nagB*	0.98
M5005_Spy_1388	N-acetylglucosamine-6-phosphate deacetylase	*nagA*	0.97
M5005_Spy_1574	universal stress protein family		2.05
M5005_Spy_1575	quinolone resistance protein	*norA*	2.33
M5005_Spy_1746	PTS system, **Cellobiose**-specific IIA component	*celA*	1.02

[a] Fold change determined by comparing CDM-glu/CDM-fru

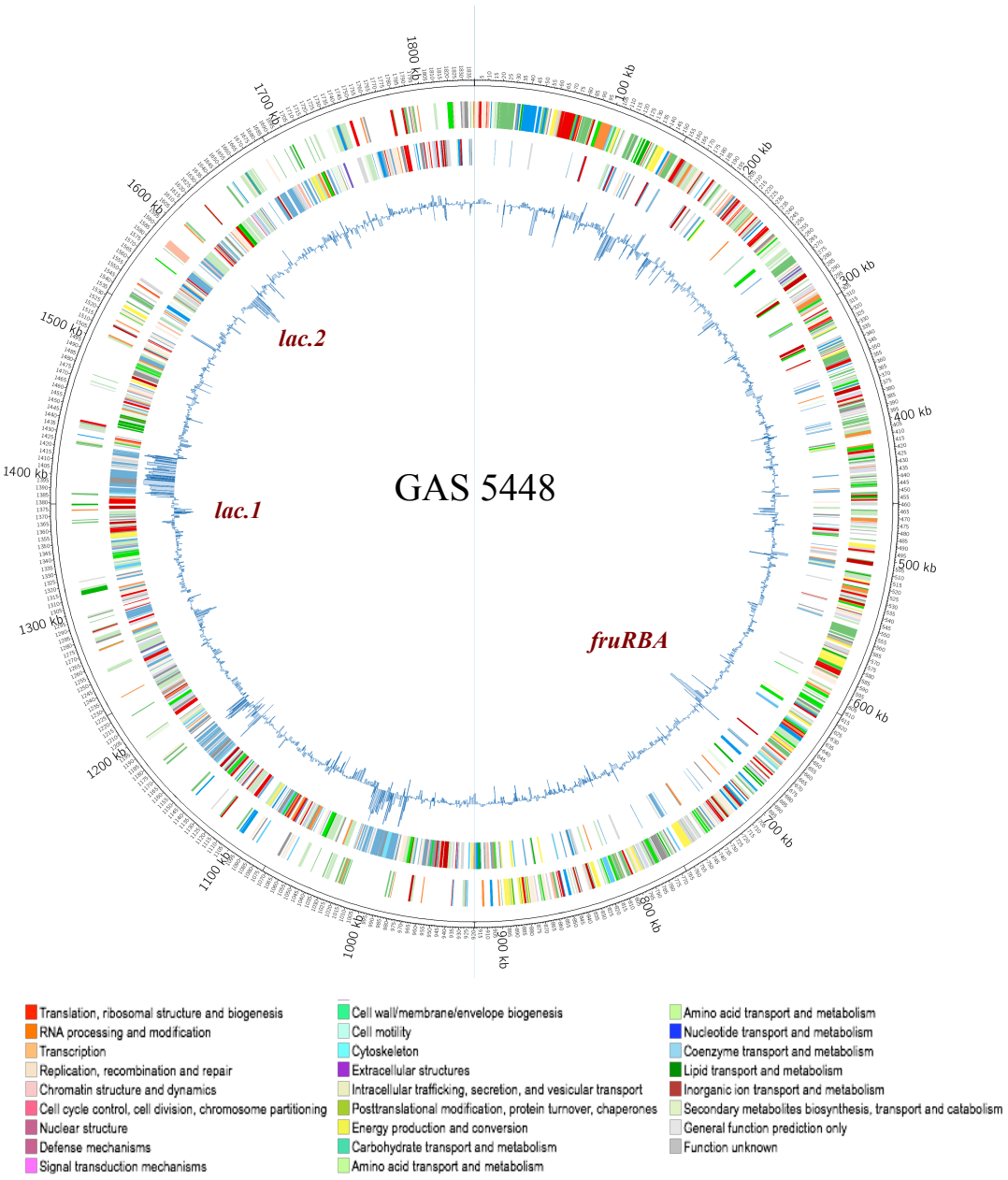


Figure 4.1. Transcriptomic landscape of GAS 5448 during growth in fructose. Circos plot of wild type GAS 5448 gene global expression altered under growth in CDM + 0.5% glucose versus CDM + 1% fructose. The outer most ring represents a size ruler. The next two rings represent the GAS open reading frames on the (+) and (-) strand of the genome, respectively. The fourth ring shows transcript levels of genes, where the height of the bars represent the $\log_2$ fold change in differential expression in glucose vs. fructose (inside is induced, outside is repressed). The *fruRBA* and *lacD* loci induced by fructose are indicated. *Performed by K.M.V. and A.T.B.

A total of 195 genes (103 non-phage genes; 5.5% of genome) were altered by the presence of fructose in comparison to glucose as the sole carbon source (**Table 4.1 and Fig. 4.1**). Several predicted and established sugar metabolism operons were highly induced during growth in fructose (**Table 4.1, shaded gray**), including the *lacD.1* and *lacD.2* β-glucoside-specific PTS operons, the maltose-specific ABC transport system (*malADC*), and the predicted fructose operon (*fruRBA*). Of these, the putative *fruRBA* operon was the most highly induced by fructose with over a 5-fold increase in expression (**Table 4.1**). In addition, several members of the SerR regulon involved in serine metabolism were induced by fructose (276), including genes encoding the putative regulator SloR, the V-type sodium ATP synthase (*ntp*) operon and the serine dehydratase (*sdhAB)* operon. Interestingly, *sloR* is predicted to encode a transmembrane protein that exhibits homology to the EIIC domain of the putative FruA PTS transporter. Notably, only six genes, including *nagB* and *nagA* involved with N-acetylglucosamine metabolism, appeared to be repressed in the presence of fructose. Overall, the two highest gene ontology categories found to be induced by fructose were carbohydrate transport and metabolism (33 genes; 32%) and energy production conversion (9 genes; 9%). Therefore, the presence of fructose influences the expression of a subset of sugar metabolism genes, particularly the predicted *fruRBA* fructose utilization operon.

In order to validate the RNA-Seq results, seven genes were selected for analysis by qPCR (**Fig. 4.2A**). Overall, the qPCR results confirmed the RNA-Seq data with a correlation coefficient (R^2) calculated to be 0.83 (n=7) (**Fig. 4.2B**).

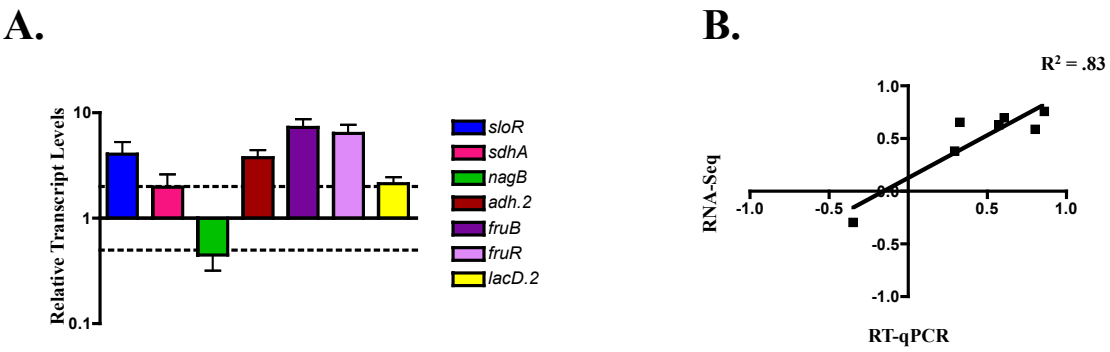


Figure 4.2. Validation of RNA-Seq results using qPCR.
(A) Transcript levels were determined using qPCR on RNA taken from GAS grown in fructose compared to glucose as the sole carbon source. Error bars represent the standard error from three biological replicates. Differences greater than 2-fold in expression for fructose grown cells compared to glucose grown cells (denoted by a dashed line) are considered significant. (B) Coefficient of correlation between the RNA-Seq and RT-qPCR results. Generated a confidence in the data with an R^2 =83%. *Performed by K.M.V.

fruRBA represents an operon in M1T1 GAS

In the 5448 genome, the *fru* locus is comprised of three contiguous genes (**Fig. 4.3A and Table 4.1**): *fruR*, encoding a predicted transcriptional repressor; *fruB*, encoding the enzyme 1-phosphofructokinase; and *fruA*, encoding a predicted fructose-specific EIIIABC component of the sugar phosphotransferase system (PTS). Each of the three genes slightly overlaps, where the start codon of the gene downstream is found in the upstream gene, suggesting they might form an operon. RT-PCR of transcripts from 5448 grown in CDM + 1% fructose showed transcriptional linkage between *fruB*/*fruA* and *fruR*/*fruB*, as well as a 1.5 kb full-length transcript detected between *fruR*/*fruA* (**Fig. 4.3AB**). These data demonstrate that *fruRBA* are expressed as an operon during inducing growth in fructose as the sole carbon source.

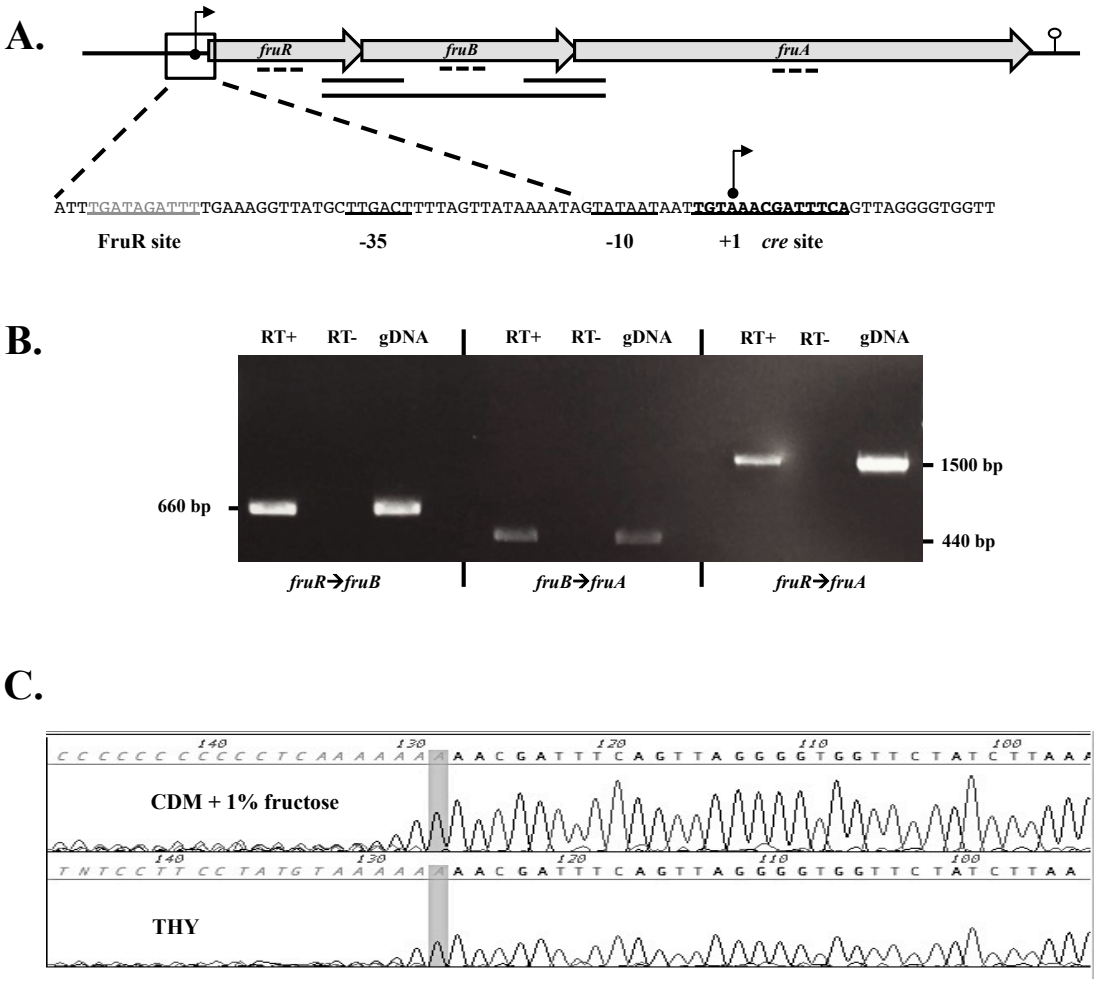


Figure 4.3. *fruRBA* is an operon in GAS M1T1 5448.
(A) Genetic organization of *fruR* (encoding a transcriptional regulator), *fruB* (encoding 1-phosphofructokinase), and *fruA* (encoding PTS fructose-specific EIIABC) is shown. Magnified region indicates promoter sequences, including putative FruR and CcpA (*cre*) binding sites (underlined). Solid lines indicate transcription linkage found by RT-PCR. Dashed lines indicate location of qPCR amplification target. The transcriptional start site is indicated by an arrow along with the -10 and -35 shown in (A). (B) RT-PCR on RNA isolated from 5448 grown in CDM + 1% fructose demonstrating transcriptional linkage between *fruR/fruB, fruB/fruA,* and *fruR/fruA*. Reactions were run using specific primers (**Table 3.2**) either with reverse transcriptase (+RT), without reverse transcriptase (-RT) as a negative control, and using genomic DNA (gDNA) as a positive control. Expected sizes of bands are indicated. (C) The transcriptional start site (+1) was determined by 5'-RACE and subsequent DNA sequencing (gray bar) from 5448 grown in THY and CDM + 1% fructose. *Performed by K.M.V.

To determine the transcriptional start site (TSS) of the *fruRBA* operon, 5'-rapid amplification of cDNA ends (5'-RACE) was used in conjunction with a primer complementary to *fruR* (fruR_SP1; **Table 3.2**). A TSS was identified 56 bp upstream of the start codon for *fruR* (**Fig. 4.3AC**) that possessed classical -10 (TATAAT) and -35 (TTGACT) promoter hexamers upstream. A predicted CcpA- (*cre* site, (7)) and a putative FruR-binding site (275) were also identified overlapping and upstream of the promoter, respectively (**Fig. 4.3A**). Finally, a putative *rho*-independent terminator was detected downstream of *fruA* consistent with the generation of the full-length *fruRBA* transcript (**Fig. 4.3A**). Thus, the promoter of the *fruRBA* operon appears to be under CcpA-mediated catabolite repression by glucose and FruR-mediated induction by fructose.

Non-polar deletion mutants of genes in the *fruRBA* operon

To interrogate each gene in the *fruRBA* operon for its role in fructose metabolism and virulence, we generated non-polar deletions for *fruR*, *fruB*, and *fruA* by allelic exchange in the 5448 genome with an in-frame chloramphenicol (*cat*) cassette (250), generating the mutant strains 5448.Δ*fruR*, 5448.Δ*fruB*, and 5448.Δ*fruA*, respectively (**Table 3.1**). These mutations were subsequently verified by PCR (data not shown). Both 5448.Δ*fruR* and 5448.Δ*fruA* showed growth kinetics comparable to the parental 5448 in THY (**Fig. 4.4B**). 5448.Δ*fruB* had a slightly slower growth rate than the WT, but reached the same yield after overnight growth (**Fig. 4.4B**). When the mutants were streaked on TSA supplemented with 5% sheep blood, 5448.Δ*fruB* exhibited a small colony phenotype compared to WT 5448 (**Fig. 4.4C**).

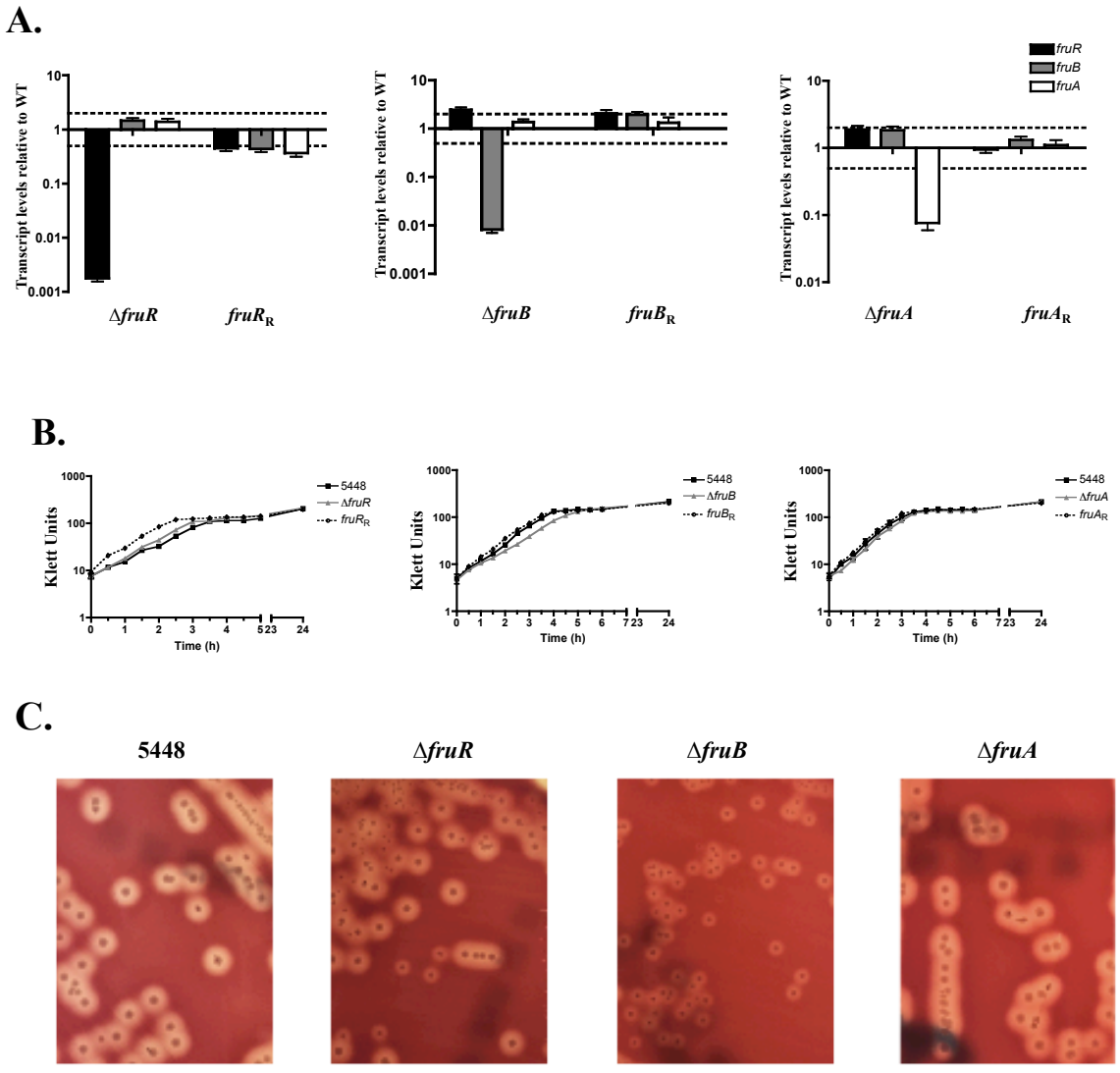


Figure 4.4. Growth and expression of *fruRBA* genes in individual *fru* mutants and their rescue strains.

(**A**) Shown are real-time qPCR analyses of relative transcript levels for each gene in the *fruRBA* operon in individual mutants and their respective rescue strains in *fruR*, *fruB*, and *fruA* compared to wild type 5448 grown in THY to late log. Error bars represent the standard error from three biological replicates. Dashed lines indicate 2-fold significance. Significance was determined using relative transcript level comparisons to the gene *gyrA*. (**B**) Each *fru* mutant was grown in THY with their respective revertant strain and the parental 5448. (**C**) Colony morphology of each *fru* mutant was streaked on blood agar (TSA + 5% sheep blood) plates, cultured overnight at 37°C + 5% CO_2 and imaged. Shown are the parental 5448, Δ*fruR*, Δ*fruB*, and Δ*fruA* as indicated. *Performed by K.M.V.

Due to the translational overlap between the genes in the *fruRBA* operon, we wanted to further validate that the mutations generated for each individual gene did not influence expression of the gene downstream. qPCR for each individual mutant was utilized to assess this using probes centered on each gene as shown in **Fig. 4.3A** with RNA isolated during late logarithmic phase in THY. The data for Δ*fruR,* Δ*fruB,* and Δ*fruA* show that the transcript levels for each mutant, respectively, was significantly reduced (**Fig. 4.4A**). Revertant strains for each mutant restored the transcript levels of that gene. Importantly, the transcript levels of the other genes in the operon were unaffected by other individual mutations as expected due to the operon being under CCR in THY. Thus, the mutations made in the *fruRBA* operon genes were non-polar and their transcript levels could be restored via reversion.

*fru*A is required for optimal growth and utilization of fructose in GAS 5448

FruA is predicted to be a fusion of the EIIA, EIIB, and EIIC enzyme subunits of a fructose-specific PTS Enzyme II (EII) involved in import and concomitant phosphorylation of fructose. To confirm the role of FruA in the utilization of fructose, 5448.Δ*fruA* was grown on a panel of PTS sugars and compared to the parental GAS 5448. In order to establish what PTS sugars are amenable for 5448 growth, growth curves in CDM supplemented with 1% (0.5% for glucose) of 11 readily available PTS sugars were determined (**Fig. 4.5A**). The wild type 5448 was able to grow in 9 of the 11 PTS sugars tested; however, mannitol and cellobiose were not conducive for significant growth of 5448 and were eliminated from further analysis. When 5448.Δ*fruA* was grown in the remaining 9 PTS sugars, only fructose resulted in

reduced growth (**Fig. 4.5BC and 4.6C**). Importantly, reversion of the Δ*fruA* mutation (5448.$fruA_R$) restored growth in fructose back to wild type levels (**Fig. 4.5BC and 4.6C**). Both 5448.Δ*fruB* and 5448.Δ*fruR* also displayed significant growth defects in CDM supplemented with 1% fructose that was restored in the revertant (**Fig. 4.6AB**). These data suggest that FruA is the primary EII transporter of fructose in GAS 5448 and that FruB and FruR are also required for growth on fructose as the sole carbon source.

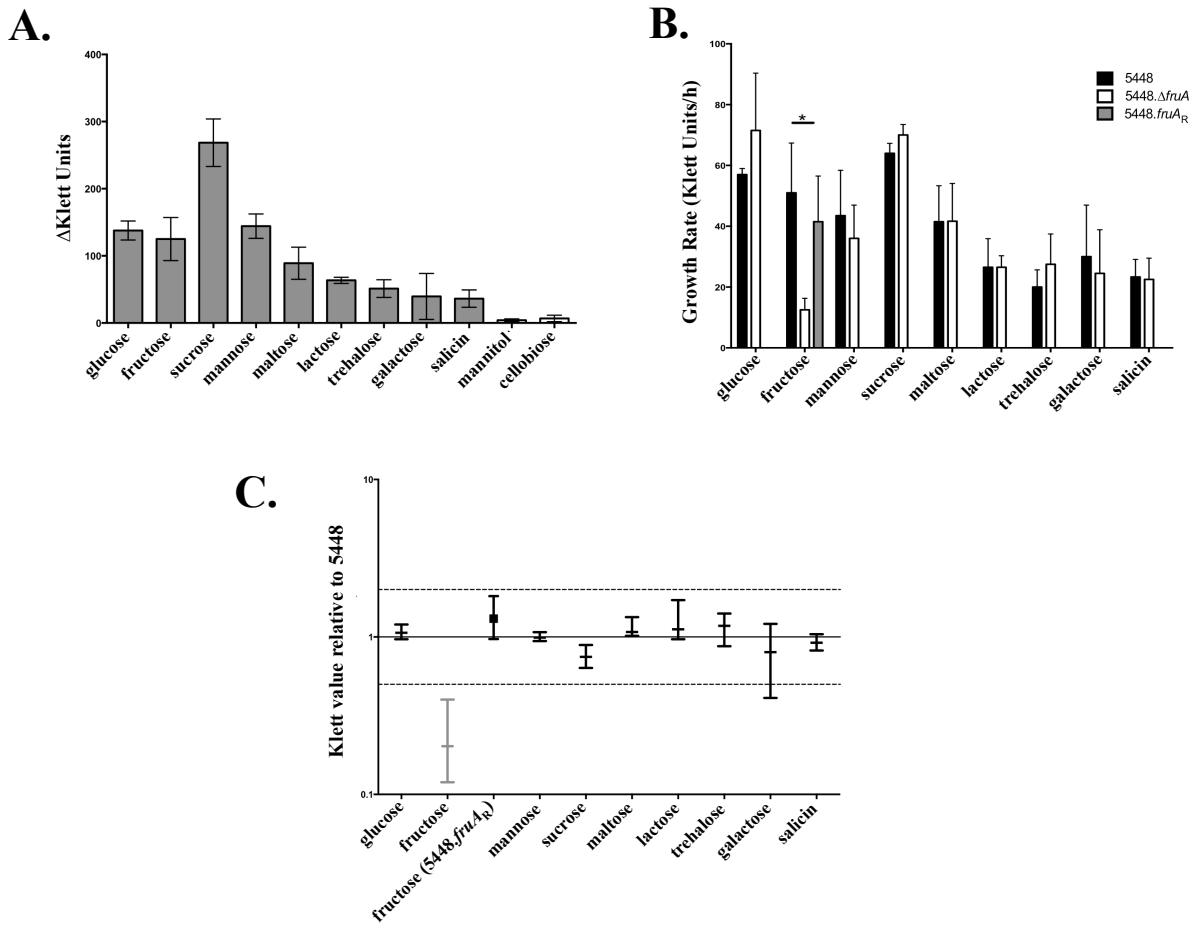


Figure 4.5. Growth analysis of 5448 and *fruRBA* operon mutants in PTS sugars. (**A**) Total yield was calculated for GAS 5448 grown for 26 h at 37°C, 5% CO_2 in various PTS sugars as the sole carbon source. (**B**) Growth rates were determined for 5448 (black bars) and 5448.Δ*fruA* (white bars) in 9 PTS sugars (p value < 0.05). Growth of the $fruA_R$ revertant (gray bar) was assessed for fructose only. Error bars represent standard error of the mean from three biological replicates. Statistical analysis was performed using an unpaired t test. Data is representative of three independent experiments. (**C**) Klett unit relative to 5448 at each time point was calculated by taking the percent WT of each individual time point, and plotting the median percent WT along with the interquartile range (25-75% of the data set). Black bars represent sugars in which 5448.Δ*fruA* grows similarly to the parental 5448. Gray bars indicate sugars in which 5448.Δ*fruA* grows poorly compared to 5448. Black square represents growth comparison of 5448.$fruA_R$ to wildtype. Dashed lines represent 2-fold growth above and below 5448. *Performed by G.S.S.

Due to the growth defect that was exhibited by a Δ*fruA* mutant, we also wanted to ascertain if *fruA* affects the utilization of other carbon sources. We used Biolog PM1 and PM2A Phenotype Microarrays (PMs) to monitor the metabolism of 190 different carbon sources by the Δ*fruA* strain. The parental GAS 5448 and 5448.Δ*fruA* were different in their utilization of 72 carbon sources (data not shown), 3 of which are designated as PTS sugars (N-acetyl-D-galactosamine, cellobiose, and mannitol) (**Table 4.2**). Surprisingly, fructose was not identified in the Biolog assay as having a defect in metabolism (**Table 4.2**). This suggests that there is another PTS transporter that is able to transport fructose, although at levels that do not sustain wild type growth (**Fig. 4.5B and 4.6C**). To confirm this observation, wild type 5448 and 5448.Δ*fruA* were also tested using a Biomérieux API50® strip that tests utilization of a subset of 50 carbohydrates (**Fig. 4.7**). As seen with BIOLOG, 5448.Δ*fruA* was able to utilize fructose similarly to 5448 in this assay; however, the fructose utilization is insufficient to support growth. We also tested 5448.Δ*fruB* and 5448.Δ*fruR* using the API50 strips. Like 5448.Δ*fruA*, both the Δ*fruR* and Δ*fruB* mutants were able to utilize fructose, even though neither was able to grow on fructose as the sole carbon source (**Fig. 4.7**). Taken together, these data suggest that *fruA* is required for optimal growth on fructose, and that FruA is the main fructose transporter for GAS. Since a PTS-defective mutant (Δ*ptsI*) of M1T1 GAS is defective for utilization of fructose in the BIOLOG assay (8), there appears to be another PTS EII that is able to transport fructose to be metabolized, albeit not at the levels necessary to support GAS growth on fructose. Finally, loss of *fruA* affects the utilization of other carbon sources,

however this is unsurprising as a Δ*ptsI* also affects utilization of non-PTS carbon sources through currently unknown mechanisms.

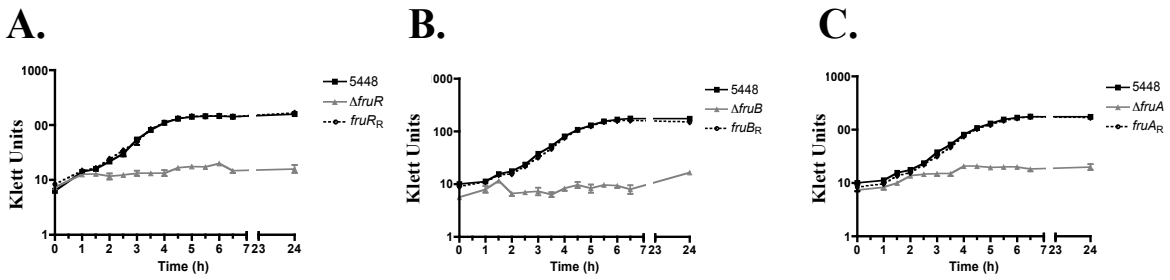


Figure 4.6. Growth curves of individual *fruRBA* operon mutants in minimal media.
Growth curves for (**A**) 5448.Δ*fruR* mutant and its $fruR_R$ revertant, (**B**) 5448.Δ*fruB* mutant and its $fruB_R$ revertant, and (**C**) 5448.Δ*fruA* and 5448.$fruA_R$ in chemically-defined media (CDM) supplemented with 1% fructose was measured by absorbance using a Klett-Summerson colorimeter (Klett units) and compared to wild type 5448. Data is representative of three independent experiments. *Performed by K.M.V.

Table 4.2: Carbon sources with altered utilization

Carbon Source	5448	5448.Δ*fruA*
N-Acetyl-D-Galactosamine	+	+/−
D-Cellobiose	+	+/−
D-Mannitol	+	+/−
D-Galactose	+	+
D-Fructose	+	+
α-D-Glucose	+	+
D-Mannose	+	+
Sucrose	+	+
α-D-Lactose	+	+
Maltose	+	+
Salicin	+	+
D-Trehalose	+	+
β-Methyl-D-Glucoside	+	+
Maltotriose	+	+

*Listed are only PTS carbon sources.
** +, Omnilog Value (OV) >200; +/-, 125<OV<200; -, OV<125.
*** Carbon sources with differences in utilization are bolded.

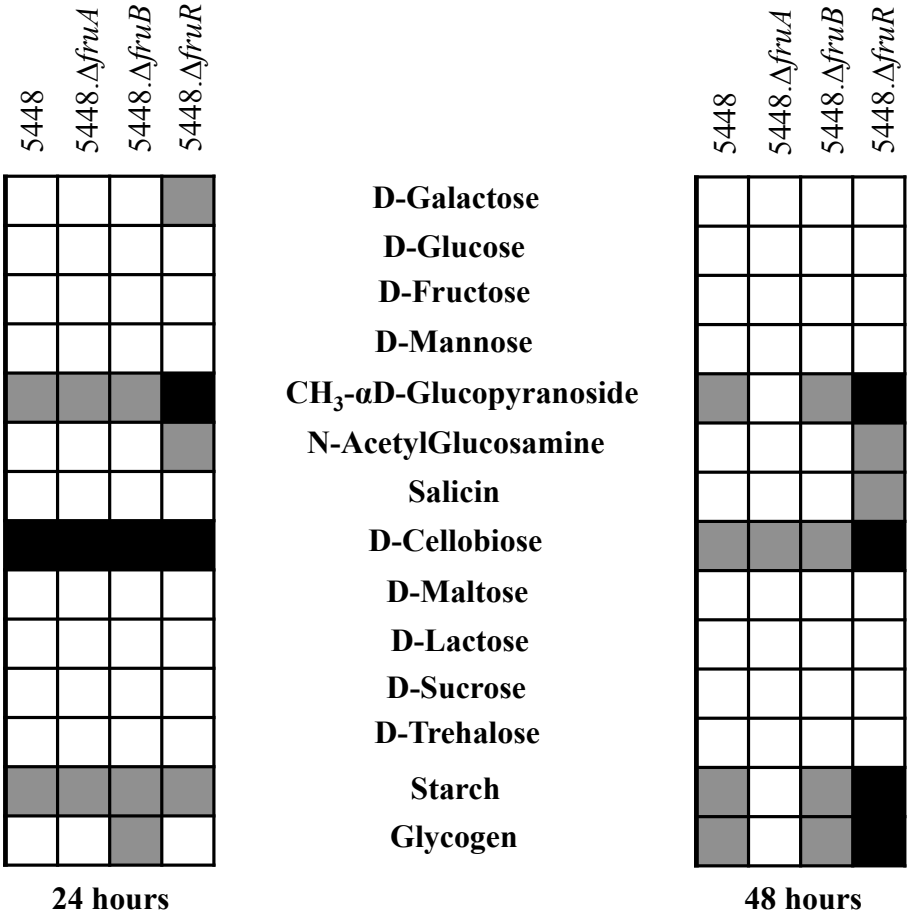


Figure 4.7. Limited carbohydrate metabolism profile of 5448 and the *fruRBA* operon mutants.
Wild type 5448 and the 5448.Δ*fruA*, 5448.Δ*fruB*, and 5448.Δ*fruR* mutants were assayed using an API® 50 CH results after incubation at 37°C for 24 h and 48 h. Listed are select carbon sources from a panel of 50 carbohydrate sources: complete utilization (white boxes); partial utilization (gray boxes), and no utilization (black boxes) based on the colorimetric indicator dye shown for each. *Performed by G.S.S.

The effect of fructose on *sloR*

The RNA-Seq data (**Table 4.1**) showed that several members of the SerR regulon (276) were up regulated in fructose, including *sloR*, the putative Streptolysin O regulator (277). *In silico* analysis of *sloR* showed nine transmembrane domains that exhibited a high degree of homology to a fructose-specific EIIC. Given that a Δ*fruA* mutant could still utilize fructose (**Table 4.2, Fig. 4.5B, Fig. 4.6C, Fig. 4.7**), we wanted to explore the role of SloR in fructose utilization and transport in addition to confirming its regulation in fructose.

An insertional inactivation mutant in *sloR* (5448.Δ*sloR*; **Table 3.1**) was generated using a temperature sensitive replicating plasmid. Growth of 5448.Δ*sloR* in CDM + 1% fructose was found to be comparable to the wild type 5448 (**Fig. 4.8B**), suggesting that SloR likely does not play a role in fructose uptake. We also tested the effect SloR had on the metabolism of other sugar sources using the API50 assay, and saw no observable difference between a Δ*sloR* strain and wild type 5448 (data not shown). However, the effect of fructose on the induction of *sloR* was confirmed using qPCR (**Fig. 4.2**, blue bar). Additionally, an expression reporter plasmid was constructed by cloning the promoter of *sloR* (P*sloR*) upstream of the firefly luciferase gene (*luc*) and activity was determined through quantification of relative luciferase units (LU). A significant increase in P*sloR*-luciferase activity was observed when cells were grown in fructose in comparison to sucrose (**Fig. 4.8A**). To rule out that FruR regulates *sloR*, we also tested *sloR* transcript levels in a Δf*ruR* strain grown in fructose and glucose using qPCR and saw no observable difference (data not shown). Taken together, we show that *sloR* induction is a fructose specific phenotype,

however this induction is independent of FruR. We also show that that *sloR* is most likely not an alternative transporter of fructose, but this cannot be ruled out until a double Δs*loR*Δ*fruA* strain is tested.

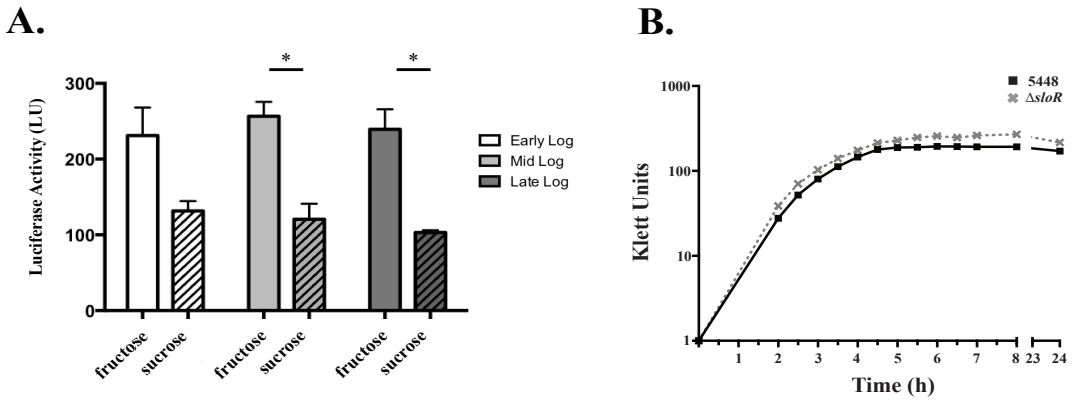
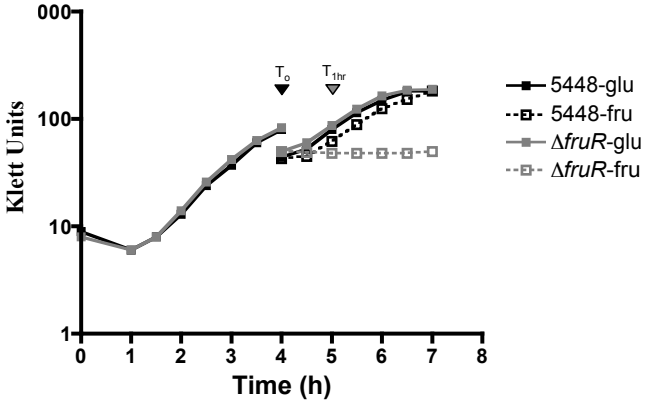


Figure 4.8. The effect of fructose on *sloR*.
(**A**) P*sloR* promoter activity in WT 5448 containing the P*sloR-luc* luciferase reporter plasmid was grown in CDM containing either 1% fructose (solid bars) or sucrose (diagonal bars). Samples were taken at three time points in logarithmic phase of growth (early: white bars; mid: light gray bars; late: dark gray bars) and assayed for luciferase production, expressed in relative luciferase units. Error bars for represent mean ± SD of the results from three biological replicates performed in triplicate. Significance for (**A**) was determined using a Student *t* test ($p \leq 0.01$). (**B**) Growth of 5448.Δ*sloR* in CDM + 1% fructose compared to 5448. *Performed by K.M.V. and G.S.S.

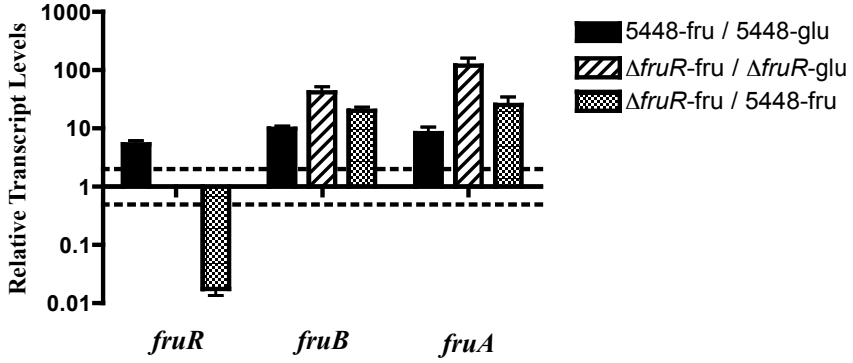
FruR is a repressor of the *fruRBA* operon

In other streptococci, *fruR* encodes a putative DeoR-like repressor in the family of carbohydrate regulatory proteins (272, 275, 278). To confirm that FruR is a repressor of the *fruRBA* operon, we grew wild type 5448 and the 5448.Δ*fruR* mutant in both catabolite repressing (glucose) and inducing (fructose) conditions in order to isolate RNA for qPCR. Growth in CDM + 0.5% glucose (**Fig. 4.9A**) was successful, however, when fructose was the sole carbon source available, the Δ*fruR* strain failed to grow (**Fig. 4.6A and 4.9A**). In order to generate appropriate cell density to extract RNA, we applied a "diauxic growth" condition. Both the WT and the Δ*fruR* strains were grown in CDM + 0.5% glucose to late logarithmic phase, washed twice, and transferred for growth in either CDM + 1% fructose or CDM + 0.5% glucose for one hour. qPCR was used to determine the effect of the 5448.Δ*fruR* mutant on the transcription of *fruB* and *fruA* using RNA from cells grown in either fructose or glucose. When comparing WT 5448 grown in glucose then fructose, we saw an increase in transcript levels across the entire operon (**Fig. 4.9B**, black bars), mirroring what we saw in the RNA-Seq results (**Table 4.1**). Comparing WT 5448 grown in fructose to 5448.Δ*fruR* grown in fructose, an increase in *fruB* and *fruA* transcript levels is observed, indicating that FruR is acting as a repressor of this operon (**Fig. 4.9B**, checkered bars). Furthermore, an even larger increase in transcript levels for *fruB* and *fruA* in the 5448.Δ*fruR* mutant was observed when grown in glucose compared to fructose, indicating that the operon is also under carbon catabolite repression (CCR) (**Fig. 4.9B**, diagonal bars).

A.


000
100
10
1
Klett Units
0 1 2 3 4 5 6 7 8
Time (h)
T0
T1hr
5448-glu
5448-fru
ΔfruR-glu
ΔfruR-fru

B.


1000
100
10
1
0.1
0.01
Relative Transcript Levels
fruR
fruB
fruA
5448-fru / 5448-glu
ΔfruR-fru / ΔfruR-glu
ΔfruR-fru / 5448-fru

C.

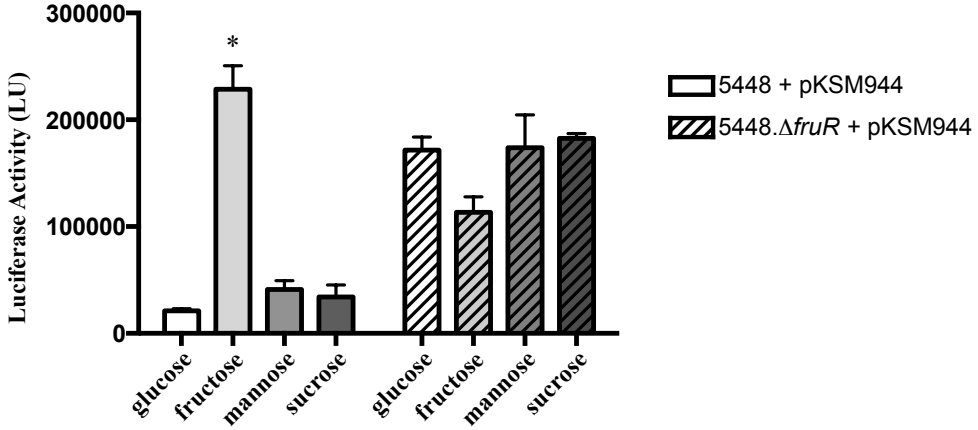

300000
200000
100000
0
Luciferase Activity (LU)
*
glucose
fructose
mannose
sucrose
glucose
fructose
mannose
sucrose
5448 + pKSM944
5448.ΔfruR + pKSM944

Figure 4.9. FruR represses expression of the *fruRBA* operon.
(**A**) Growth of 5448 (black) or the 5448.Δ*fruR* mutant (gray) strain under modified diauxic conditions. Cells were grown in CDM + 0.5% glucose (solid line) to late log (T_o), washed and transferred into CDM with either 0.5% glucose (solid line) or 1% fructose (dashed line). Cells were outgrown for 1hr (T_{1hr}) in either sugar for RNA extraction. **(B)** Transcript levels of *fruRBA* operon genes were measured by RT-qPCR with respect to growth condition. 5448 grown in fructose compared to glucose (black bars), the 5448.Δ*fruR* mutant grown in fructose compared to glucose (diagonal bars), and the 5448.Δ*fruR* mutant grown in fructose compared to the WT 5448 grown in fructose (checkered bars) are shown. Error bars represent the standard error from two biological replicates. Differences greater than 2-fold in expression for the mutant compared with wild type (denoted by a dashed line) are considered significant. Significance was determined using relative transcript level comparisons to the gene *gyrA*. (**C**) P*fruR* promoter activity in WT 5448 containing the P*fruR-luc* luciferase reporter plasmid was grown in CDM containing 0.5% glucose or 1% fructose, mannose, or sucrose. Samples were taken at mid logarithmic phase of growth and assayed for luciferase production, expressed in relative luciferase units. Error bars for represent mean ± SD of the results from three biological replicates performed in triplicate. Significance for (**C**) was determined in comparison to glucose using a unpaired *t* test ($p \leq .001$). *Performed by K.M.V.

To confirm the role of FruR on repression of the *fruRBA* operon, a luciferase reporter assay was performed using the mapped promoter region of the operon (P*fruR*). A significant increase in luciferase activity was observed only when the WT 5448 containing the P*fruR-luc* construct was grown in CDM + 1% fructose as opposed to 0.5% glucose, or 1% sucrose or mannose (**Fig. 4.9C**) indicating induction of the operon is fructose-dependent. When the same assay was performed in the Δ*fruR* background, there was an overall increase in luciferase activity for all sugars

tested (**Fig. 4.9C**). Therefore, both FruR and CcpA negatively control the expression of the *fruRBA* operon, and fructose (after conversion to fructose-1-phosphate) likely acts as an inducer of the operon through inactivation of FruR.

fruB and *fruR*, but not *fruA*, are important for 5448 survival in whole human blood

The *fruA* gene was identified as important for GAS survival in whole human blood during a genome-wide competitive TraSH (15). To determine the step in fructose utilization that might be involved in GAS survival in blood, we performed Lancefield bactericidal assays on each of the non-polar *fruRBA* operon mutants and their respective revertant strains. Interestingly, both 5448.Δ*fruR* and 5448.Δ*fruB* exhibited decreased survival in whole human blood relative to the parental GAS 5448 (**Fig. 4.10A**). Surprisingly, the 5448.Δ*fruA* did not show a significant decrease in survival in human blood, although it trended downward and was variable. The revertant strains for all three mutants showed a restoration of growth comparable to wild type 5448 (**Fig. 4.10A**). Thus, although our TraSH screen had indicated that *fruA* was needed for GAS survival in human blood, (15), our data using defined non-polar mutants in the *fruRBA* operon indicate that *fruR* and *fruB* are mostly responsible for this phenotype.

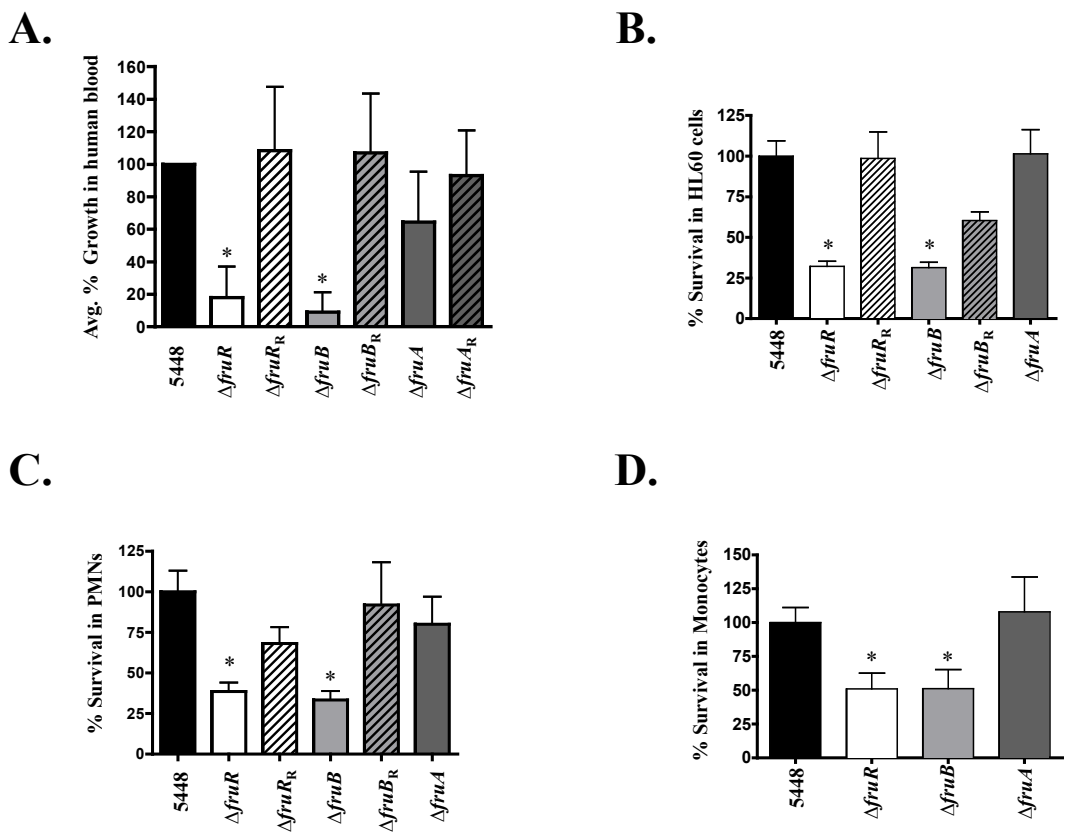


Figure 4.10. *fruR* and *fruB* are important for survival in human blood and phagocytic cells.

In all assays, wild type 5448 (black bars) was compared to 5448.Δ*fruR* (white bars), 5448.Δ*fruB* (light gray bars) and 5448.Δ*fruA* (dark gray bars) for growth and survival. Respective rescued mutant strains were indicated with diagonal lines. (**A**) Lancefield bactericidal assay was performed by monitoring growth in whole human blood with rotation at 37°C for 3 h. Data is presented as percent growth in blood corresponding to the multiplication factor (MF) of the mutant divided by the MF of the wild type × 100. Error bars represent mean ± SD of the results of six independent experiments. **(B)** Survival in neutrophil-like HL60 cells. GAS (1×10^5 cfu/well) and HL60 cells (1×10^6 cfu/well) were incubated for 2 h at 37˚C. Cells were lysed and percentage of surviving GAS was determined relative to survival of 5448. (**C**) Survival in human neutrophils (PMNs). GAS (1×10^5 cfu/well) and neutrophils (1×10^6 cfu/well) were incubated for 2h at 37˚C. Neutrophils were then lysed and the percentage of surviving GAS was determined and shown as relative survival to the WT (5448). **(D)** Survival in human monocytes. GAS (1×10^5 cfu/well) and human monocytic cells (1×10^6 cfu/well) were incubated for 2 h at 37˚C. Cells were lysed and percentage of surviving GAS was determined relative to survival of 5448.Error bars for (**B-D**) represent mean ± SD of the results from three biological replicates performed in triplicate. Significance for (**A**) was determined using a Student *t* test ($p \leq 0.0015$), while significance for (**B-D**) was determined by unpaired *t* test ($p < 0.05$). *Performed by Y.L.B. and L.A.V.

fruR and *fruB* are important for GAS 5448 survival in neutrophils and monocytes

The phenotype exhibited by 5448.Δ*fruR* and 5448.Δ*fruB* in whole human blood can be attributed to defects in metabolic processes or in immune evasion. To clarify this question, we subjected 5448.Δ*fruR* and 5448.Δ*fruB* to opsonophagocytic killing assays using the human-derived promyelocytic leukemia cell line HL60 differentiated into neutrophil-like cells (**Fig. 4.10B**) and in freshly isolated human polymorphonuclear leukocytes (PMNs) (**Fig. 4.10C**). Neutrophils are the predominant cell type that act as a first line of defense at the site of infection, and can kill pathogens through multiple extra- or intra-cellular means such as antimicrobial peptides, phagocytosis, and DNA neutrophil extracellular traps (NETs) (279). Both neutrophil-like HL60 cells and human-isolated PMNs were infected using an MOI of 0.1, and strains were plated for survival and compared to the parental 5448. When either 5448.Δ*fruR* or 5448.Δ*fruB* was incubated with HL60 cells, each showed a significant decrease in survival compared to 5448 (**Fig. 4.10B**). However, 5448.Δ*fruA* survived comparable to the wild type GAS. Importantly, revertant strains 5448.*fruR*$_R$ and 5448.*fruB*$_R$ either partially or completely restored the phenotype to wild type levels (**Fig. 4.10B**). These phenotypes were recapitulated with freshly isolated human PMNs (**Fig. 4.10C**), supporting a requirement for FruB and FruR, but not FruA, for survival in human neutrophils. There was also a survival defect observed for 5448.Δ*fruR* or 5448.Δ*fruB*, but not 5448.Δ*fruA*, in human monocytes, suggesting a role for these *fruRBA* operon genes in evading killing by different phagocytic cell

types (**Fig. 4.10D**). To eliminate the possibility that any of the *fruRBA* operon mutants had a reduced ability to grow on cell culture media and artificially cause the phenotypes we observed, we also grew each strain in RPMI + 20% plasma, and saw no growth defect in comparison to the parental 5448 (**Fig. 4.11**). Thus, *fruR* and *fruB*, but not *fruA,* are important for GAS resistance to killing by human neutrophils and primary monocytes.

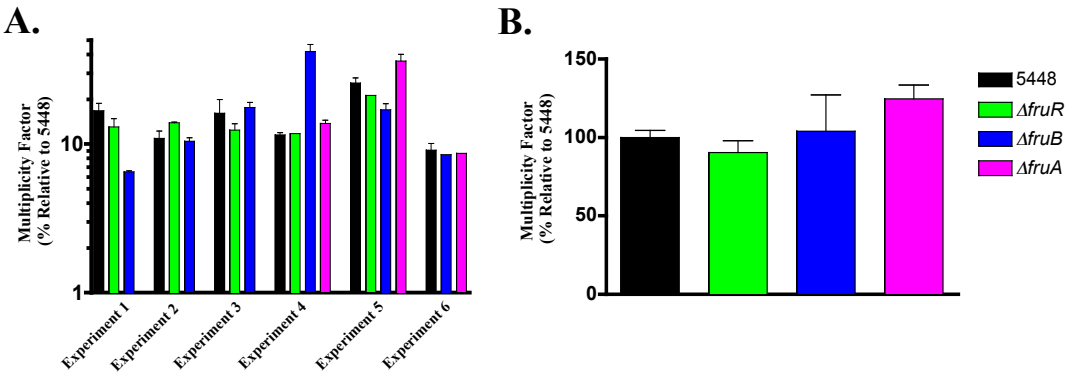


Figure 4.11. 5448 and Δ*fruRBA* Growth in RPMI+ 20% Plasma.
Each *fru* mutant was grown along with the parental strain 5448 in RPMI + 20% plasma. Data from 6 independent experiments, with at least three replicates for each *fruRBA* operon mutant in comparison to WT 5448 are shown in (A). The cumulative results of (A) are shown as an average of the replicates in (B). *Performed by L.A.V.

fruR and *fruB* phenotypes are not observed in murine models of infection

Murine models of infection have long been used as the gold standard for simulating disease of human pathogens, with GAS being no exception. The significant reduction in survival of both 5448.Δ*fruR* and 5448.Δ*fruB* in human blood and phagocytic cells led us to hypothesize that both mutants would show attenuated virulence in mouse models. We first subjected 5448.Δ*fruR* and 5448.Δ*fruB* to a modified Lancefield bactericidal assay using whole murine blood. Unexpectedly, we found that both 5448.Δ*fruR* and 5448.Δ*fruB* mutants did not show a defect for growth in mouse blood as seen in human blood (**Fig. 4.12A**). To test the impact of the *fruRBA* operon mutants *in vivo*, an intraperitoneal (*i.p.*) murine model of systemic GAS infection was used. Mice were infected *i.p.* with 4 x 10^8 CFU/ml of 5448, 5448.Δ*fruR* and 5448.Δ*fruB*, and monitored three times daily during a 72 h period. Similar to the murine Lancefield bactericidal assay, we did not see a decrease in survival for the 5448.Δ*fruR* or the 5448.Δ*fruB* mutants when compared to the parental 5448 (**Fig. 4.12B**). Thus, while *fruR* and *fruB* appear to be important for survival against neutrophil and monocyte killing in human blood, they do not affect GAS pathogenesis during murine models of infection. This strongly suggests a species-specific mechanism may be involved.

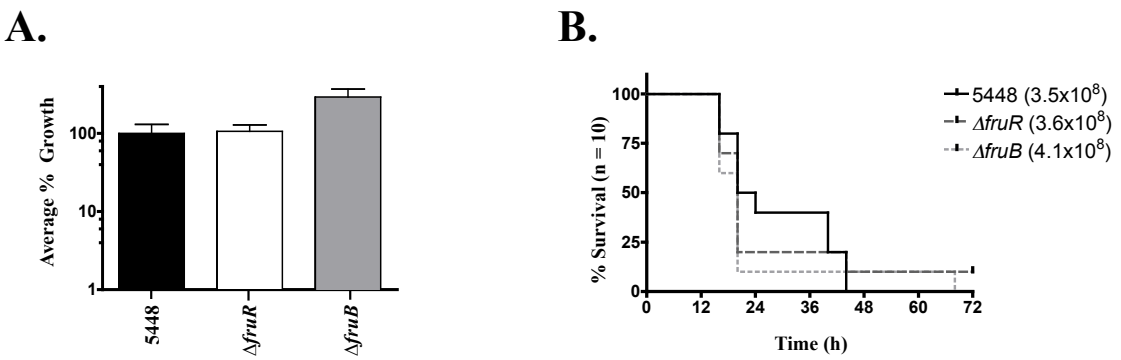


Figure 4.12. Loss of *fruR* and *fruB* does not affect survival in murine models of infection.

The 5448.Δ*fruR* and 5448.Δ*fruB* mutants were compared to the wild type 5448 in the following two assays; (**A**) Survival in a modified Lancefield bactericidal assay using whole mouse blood. The data are presented as percent growth in blood corresponding to the multiplication factor (MF) of the mutant divided by the MF of the wild type × 100. Data represents GAS survival in blood obtained from multiple mice (n=4-6). Error bars represent mean ± SD of the results of two independent experiments. Significance was determined using a Student *t* test ($p \leq 0.0015$) (**B**) Survival in a murine intraperitoneal (*i.p.*) model of infection. GAS (~4×10^8 Cfu/mouse) were injected *i.p* into mice (n=10/strain) and monitored for morbidity over the course of 72 h. Significance was determined using Kaplan–Meier survival analysis and log rank test. *Performed by L.A.V.

Discussion

The current study was undertaken to assess the importance of fructose utilization in GAS pathogenesis. Our results reveal a fructose-induced operon (*fruRBA*) that was found to be required for growth of GAS in fructose and implicate the PTS-transporter, FruA, as the primary transporter of fructose in GAS. FruR acts as a repressor at the *fruRBA* operon promoter, likely along with CcpA, to allow for fructose induction of the system. Mutants lacking *fruR* and *fruB*, but not *fruA*, were found to be sensitive to survival in whole human blood and phagocytic killing by neutrophils and monocytes. Interestingly, neither mutant was found to be required for survival in murine blood or for virulence in a murine model of systemic infection. Thus, the *fruRBA* operon plays a key role in both GAS metabolism and pathogenesis, revealing a mechanism by which this human pathogen can sense its nutritional environment during infection and directly linking this operon to innate immune evasion in the host.

Fructose-induced regulon of GAS

Metabolic genes for utilization of alternative (non-glucose) sugars are typically induced only during times of low glucose (carbon catabolite repression, CCR) and in the presence of the inducing sugar itself. For pathogenic streptococci and related low G+C Gram-positive bacteria, a genome-wide fructose-induced regulon has previously been determined only for *Streptococcus mutans* and the probiotic *Lactobacillus acidophilus* (280, 281). In *S. mutans*, microarray analysis revealed 68 genes (3.5% of genome) that were differentially expressed at least 2-fold

in fructose compared to glucose, whereas a similar analysis in *L. acidophilus* found that 110 genes (6% of genome) were regulated. Our RNA-Seq analysis discovered that growth of GAS in CDM with fructose affected transcription of roughly 5.5% of the non-prophage genome (102 genes) compared to glucose (**Tables 4.1**). This is in line with what was observed in both *L. acidophilus* and *S. mutans*, highlighting a fairly focused fructose-specific response in these cells. The most highly fructose-induced genes in GAS were *fruRBA*, followed by several other sugar-specific PTS operons such as *lacD.1* and *lacD.2*. The comparable *L. acidophilus fruRBA* genes were also highly induced by fructose (281). In contrast, the *S. mutans* genes for *fruRBA* (Smu.0870-72) were constitutively expressed and a separate mannose/fructose family PTS system was found to be inducible (280).

We also observed an increase in expression of genes in the serine catabolism regulon including *sloR*, *sdhAB*, and *ntpA-K* (276). We further validated the fructose induction of *sloR* expression via qPCR (**Fig. 4.2A**) and through promoter-luciferase fusions (**Fig. 4.8A**). SloR was initially implicated in streptolysin O expression (277), but more recently was shown to be involved in serine acquisition (276). Interestingly, *in silico* analysis (282) of SloR shows significant homology to a fructose EIIC transporter, suggesting that it might play a role in the transport of fructose, and potentially aid in signal transduction to coordinate gene expression. However, a *sloR* mutant did not show any significant effect on utilization of fructose or any other PTS sugar by GAS (**Fig. 4.8B)**.

Role of the *fruRBA* operon in fructose metabolism

(i)FruA

The GAS M1T1 genome is predicted to encode 14 sugar-specific Enzyme II (EII) translocation components of the phosphoenolpyruvate-phosphotransferase system (PTS). This is the first study in GAS characterizing fructose utilization and the contribution of the *fruRBA* operon, confirming its annotated role in fructose metabolism. We found that an in-frame mutant in *fruA* (5448.Δ*fruA*), encoding the putative PTS EIIABC, exhibited a significant growth defect in comparison to WT (5448), but did show some residual growth on fructose (**Fig. 4.5BC and 4.6C**). Non-polar mutants in *fruB* (5448.Δ*fruB*), encoding 1-phosphofructokinase, and *fruR* (5448.Δ*fruR*), a putative repressor, also showed a comparable defect in growth on fructose, confirming that the entire *fruRBA* operon is necessary for fructose utilization in GAS. As mentioned above, the residual growth and utilization of fructose may involve other PTS EII pathways as other PTS-transporters were up regulated in a fructose-specific manner (**Fig. 4.7 and Table 4.2**).

(ii) FruR

The DeoR-family transcriptional regulator, FruR, is directly involved in repression of the *fruRBA* operon in *L. lactis* (275) and other Gram-positive bacteria (272, 278, 283). Through transcriptional analysis of the 5448.Δ*fruR* mutant and its rescue strain (5448.Δ$fruR_R$), we demonstrated that FruR represses transcription of the *fruRBA* operon through its promoter (P*fru*) and that this promoter is required for fructose-induction (**Fig.4.9**). Furthermore, expression of the operon exhibits evidence of CCR likely through a CcpA-mediated CCR mechanism. *In silico* analysis of the

fruRBA promoter region identified the presence of a putative FruR binding site (275) upstream of the -35 hexamer and a catabolite response element (*cre*) site for CcpA binding (5-7) overlapping with the start of transcription (**Fig. 4.3A**). The physiological inducer of FruR de-repression is fructose-1-phosphate (F1P) and only 1 mM of F1P is sufficient to disrupt FruR binding to its operator in *P. putida* (284, 285). Fructose-1-phosphate is also the inducer of the *fruRBA* operon in other Gram-positive species such as *L. lactis* (275). This correlates with FruR being a member of the DeoR-family of repressors, which are typically induced by sugar-phosphates produced by the regulatory pathway they control (180, 272, 275).

The *fruR* mutant also exhibited a significant growth defect when grown in fructose alone, similar to loss of *fruA* and *fruB* (**Fig. 2 and Fig S5**). Given that the *fruRBA* operon is up regulated in this mutant, it would be expected to metabolize fructose at wild type levels. One explanation might be that, in a Δ*fruR* strain increased *fruA* and *fruB* transcripts leads to an increase in the intracellular levels of F1P that is converted to fructose-1,6-bisphosphate (FBP) by FruB. This mechanistic conversion of fructose into FBP directly feeds into gluconeogenesis and glycolysis. High levels of FBP have been shown to be important for signal transduction in Gram-positive bacteria as it stimulates activity of HprK, a serine/threonine kinase involved in catabolite repression through phosphorylation of Hpr(Ser~P), which can then interact with CcpA and trigger CCR (181, 286). Therefore, high levels of FBP can mimic high levels of glucose and elicit CCR conditions (glucose), which then shuts down transcription of CcpA regulated operons such as *fruRBA*, and this then would lead to a growth arrest on fructose. However, this is most likely not the case as there were

high levels of the *fruB and fruA* transcript in the qPCR results (**Fig 4.9B**) observed in the Δ*fruR* mutant, and there were high levels of activity exhibited in the P*fru-luc* reporter assays (**Fig. 4.9C**).

(iii) FruB

1-phosphofructokinase (*fruB*) converts imported fructose-1-phosphate to fructose-1,6-bisphosphate and our data shows that it is also necessary for GAS growth on fructose (**Fig. 4.4B**). The *fruB* mutant (5448.Δ*fruB*) also exhibited a small colony phenotype when plated on blood agar (**Fig. 4.4C**) and showed a slightly longer doubling time when compared to the parental 5448 and its revertant (5448.Δ$fruB_R$) in rich media implicating this enzyme in playing role outside of the conversion of F1P to FBP (discussed below). Although other alternate PTS EII pathways exist in GAS that allow for the transport of fructose (enters the cell as fructose-6-phosphate; F6P), as demonstrated by both the RNA-Seq and Biolog/API CH50® metabolic data (**Tables 4.1, and 4.2; Fig. 4.7**), the efficiency of transport is not able to support growth. In addition, when fructose is transported by alternative PTS systems, a 6-phosphofructokinase (*pfkA*) enzyme to convert F6P to FBP would need to be present (287). However, *pfkA* was not up regulated when GAS was grown in fructose as the sole carbon source (**Table 4.1**). Therefore, while the data presented in this study demonstrates that the *fruRBA* operon is the main operon responsible for the uptake and growth in fructose for the GAS, there are other PTS transporters that still enable the cell to utilize this sugar, but not to levels that support growth.

Role of FruR and FruB in innate immune evasion

A phenotype for *fruA* in either whole human blood or in phagocytic killing was not observed, which is surprising as *fruA* was originally observed to be important for survival of GAS in whole human blood using a genome-wide TraSH screen (15). However, the TraSH screen is a competitive assay, therefore it is possible that *fruA* is important in competition. Regardless, since a *fruB* mutant (5448.Δ*fruB*) is defective for survival in whole blood and human-derived phagocytic cells (**Fig. 4.10A**), it would appear that 1-phosphofructokinase enzymatic activity (F1P conversion to FBP) is crucial for GAS to evade the innate immune response through an undefined mechanism. It could simply be that the buildup of F1P triggers a downstream regulatory cascade that influences genes important for resistance to phagocytic killing. However, FruB may be playing a more direct role, such as the LacD tagatose-1,6-bisphosphate aldolase, which controls expression of the SpeB exotoxin through interacting with the transcriptional regulator RopB (4).

FruR, which also exhibits decreased survival in whole human blood and phagocytic killing, is likely contributing to this phenotype through its regulation of *fruB*. Although there is an increase in *fruB* transcript levels (**Fig. 4.9B**), and presumably FruB protein levels, in a Δ*fruR* mutant, it is still sensitive to phagocytic killing (**Fig. 4.10**). These data suggest that there is a stoichiometry effect of the levels of 1-phosphofructokinase in the cell leads to a similar defect in immune evasion. *In silico* analysis of the MGAS5005 genome shows that a consensus FruR binding site (275) only occurs once in the genome, upstream of FruR. Therefore, it is unlikely that FruR is directly affecting transcription of genes beyond the *fruRBA* operon. However,

this hypothesis cannot be ruled out, and a transcriptional analysis of a *fruR* mutant could be performed to answer this question.

fruRB-mediated immune evasion is specific to the human host environment

Although *fruR* and *fruB* mutants were defective for survival in human blood and phagocytes (**Fig. 4.10**), we saw no growth defects in whole mouse blood or any attenuation in a mouse model of systemic GAS infection (**Fig. 4.12**). This strongly suggests that the *fruRB*-mediated mechanism of innate immune evasion is highly adapted to its human host. This is most likely attributed to the composition of the neutrophils present in human blood versus mouse blood, which is further addressed in Chapter 7. While mice have been used extensively for studying GAS pathogenesis, there are aspects related to species specificity that have been problematic for disease modeling. GAS is able to produce a wide range of disease severities in different mouse strains, causing different strains of mice to be either better or worse models for simulating GAS infections (288). It would appear that coevolution between the human immune system and GAS has resulted in adaptation to its natural host environment. Whether 1-phosphofructokinase (FruB) is directly sensing a human-specific environmental stimulus leading to downstream virulence gene regulation, or whether it merely results in the signal remains to be determined. However we hypothesize that it is not the ability for GAS to transport fructose into the cell, but the metabolic byproducts and activity of FruB itself that contributes to the virulence phenotype observed in this study. Regardless, it is clear that GAS possesses a link

between fructose metabolism and regulation of immune evasion molecules that are important during human infections.

**Author contributions: K.M.V. performed the research on the following sections: RNA-Seq and transcriptome analysis, the fructose operon structure and start of transcription, characterization of *sloR*, and the role of *fruR*. K.M.V. in conjunction with G.S.S. performed growth analysis on fructose on the individual fructose operon genes. G.S.S. performed all other growth experiments. Lancefield studies were performed by YLB, and mouse work was performed by LAV. ATB aided in RNA-Seq data analysis.

Chapter 5: Glucose differentially phosphorylates Mga and alters its activation and regulon dynamics

Introduction

Coordination of the timely expression of virulence factors is key for bacterial pathogens in order to successfully colonize the host and elicit an infection. For Gram-positive bacteria, expression of virulence genes is often linked to the availability of essential nutrients, specifically carbohydrates (186). Therefore, it is not surprising that these bacteria have tied their main system for carbohydrate uptake, the phosphoenolpyruvate (PEP)-phosphotransferase system (PTS), (180, 186) to coordinate the expression of carbohydrate metabolism genes (14, 19, 289-293) (9).

The Group A Streptococcus (GAS) or *Streptococcus pyogenes* is a Gram-positive pathogen that can colonize a variety of tissues in the human host, resulting in a wide range of invasive or non-invasive diseases. GAS can successfully colonize and adapt to a variety of niches in the human host through coordination of environmental cues and nutritional status with global transcriptional networks, as has been shown in several studies (12, 14, 157, 228, 294). Since GAS does not encode for alternate sigma factors to coordinately regulate expression of genes, GAS utilizes global transcriptional regulators to synchronize transcriptomic changes. GAS encodes for 13 two-component systems (TCS) and several "stand-alone" regulators (e.g., RofA-like proteins [RALPs], Rgg, and Mga) that control virulence gene expression in response to environmental stimuli (10).

The multiple gene regulator of GAS, or Mga, is ubiquitous in all sequenced strains thus far (11). A microarray study found Mga to regulate approximately 10% of the GAS genome during exponential phase of growth under nutrient rich conditions (THY) (229). The core Mga regulon consists of genes important for colonization and immune evasion. Recently, the PTS of GAS has been shown to directly phosphorylate Mga on conserved histidines located on the two PTS regulatory domains (PRDs) both *in vitro* (19) and *in vivo* (20). The PTS acts antagonistically on the phosphorylation of the PRDs, which has been shown to affect the activity of Mga *in vivo* (19). Taken together, Mga is a PRD-containing virulence regulator (PCVR) that is critical for GAS pathogenicity, and its activity is intimately linked to carbohydrate utilization.

Although it is known that Mga activity is modulated by the PTS, how specific carbohydrates affect Mga-dependent gene regulation has not been investigated. In this study, we characterized the effects of limited glucose availability on GAS physiology. Particularly, we were interested in the effects this may have on Mga and the expression of its regulon. We observed that Mga was differentially phosphorylated in THY in comparison to C media, which led to a high degree of plasticity of the regulon that is directly correlated to glucose availability. Thus, we were able to correlate activation of the Mga regulon, through glucose, to differential phosphorylation of Mga through its PRD histidines.

Results

Transcriptome analysis of M1T1 5448 growing in a low-glucose environment.

When GAS colonizes the host, it must adapt to both the hosts' immune response and to altered availability of nutrients. Often, this signal for global transcriptional change within the cell is critical for pathogen survival. When GAS invades deep tissue niches in the host, it encounters a limited carbohydrate (*ie.* carbon) source supply, where specifically, the preferred carbohydrate, glucose, is not readily available. Previous work by Loughman *et al*. has shown that C media acts as an accurate representation of the environment encountered by GAS during a deep-tissue infection due to its low glucose, high peptide composition (88), and because it stimulates the expression of *speB*, the cysteine protease.

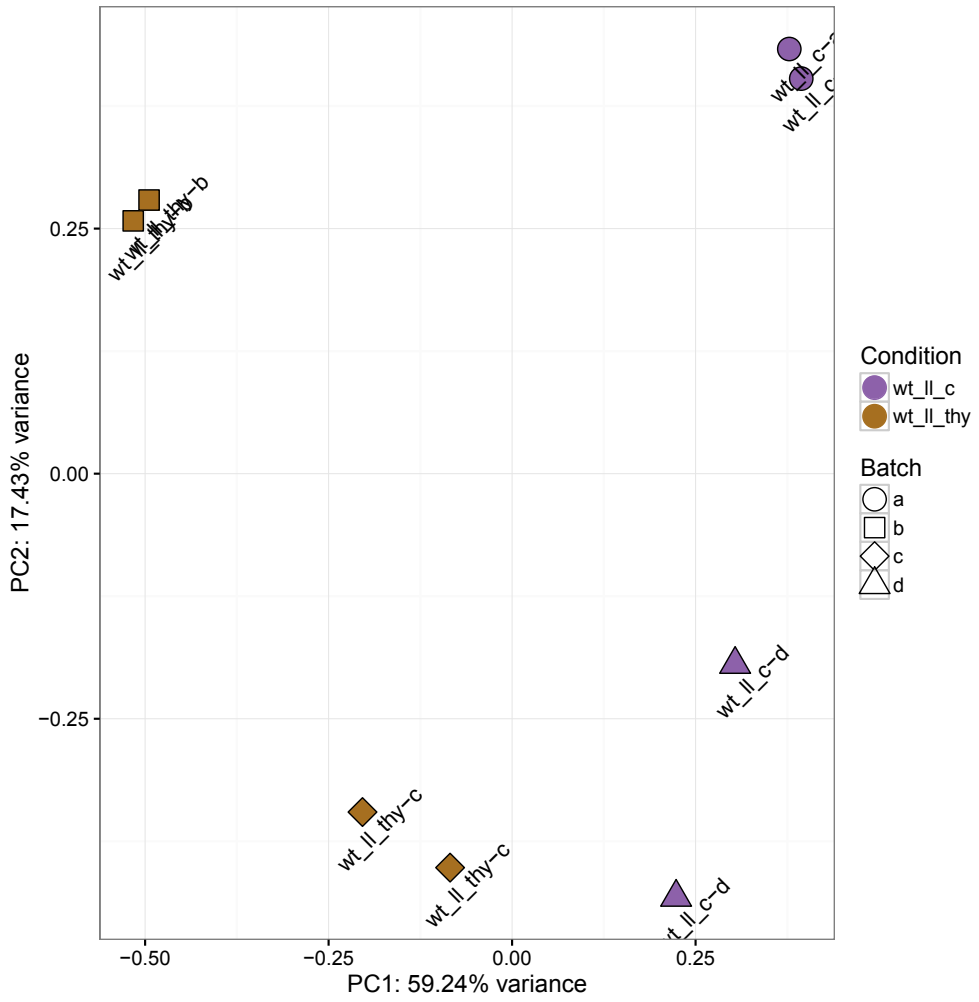


Figure 5.1. Global gene expression profiles of the WT 5448 at late logarithmic growth in different medias.
A principal component analysis (PCA) plot is shown for the transcriptome of the WT 5448 grown in either THY (high glucose) or C media (low glucose) at late logarithmic growth phase prior to normalization for batch effects. The PCA plots show the first two principal components on the X and Y axes, respectively, and the percentage of total variance specific to that principle component is indicated. Each biological replicate is represented as a single point with color corresponding to the media condition (THY, brown; C media, purple) and shapes represent different experimental batch. *Performed by A.T.B.

This study focuses on identifying whether a low-glucose versus a high-glucose environment has any effect on the global transcriptome of the M1T1 strain 5448. Thus, as representatives of these two conditions, we used the rich media THY as a representative of a high-glucose environment, and C media as a representative of a low-glucose environment. Total RNA was isolated from the WT 5448 grown in either THY (0.5% glucose (w/v)) or C media (0.05% glucose (w/v)) to late logarithmic phase in four biological replicates and processed for RNA-Seq. Data obtained from THY-grown cells was compared to C media-grown cells (WT-THY/WT-C media) with changes in gene expression over 1.9-fold ($\log_2 \leq -0.95$ or ≥ 0.95) and a p value of ≤ 0.05 was considered significant (**Appendix 2**).

We used a principal component analysis (PCA) plot in order to visualize the correlation between the two media conditions (**Fig. 5.1**). The plot showed a strong simialrities between the replicates generated from the same media, which demonstrates that reproducibility of the datasets (**Fig 5.1**; y axis). In addition, a stray batch effect was also observed. It also demonstrated that the transcriptional profile between WT 5448 grown in THY, is different from 5448 grown in C media, as demonstrated by PC1 (**Fig 5.1**; x axis).

CcpA controls the expression of genes for carbon catabolite repression (CCR) through directly binding to catabolite response elements (*cre* sites). Thus, during favorable growth conditions (high glucose), CCR occurs. Since C media contains low amounts of glucose (0.05% w/v), we hypothesize that repression of genes that are normally under CCR to be relieved in C media. A total of 510 genes (476 non-phage or tRNA genes) were differentially expressed (249 repressed, 261 activated) in C

media compared to THY (**Fig. 5.2**; **Appendix 2**). 75 of those genes (~15%), were also found to regulated by CcpA using microarrays, and are likely under direct CCR, (5, 6). Additionally, 28 *cre* sites were observed upstream of genes that were activated and repressed in C media, further supporting that genes that are activated in C media are most likely directly regulated through CcpA. Thus, the data presented here shows that growth of the WT 5448 in C media mimics a CcpA regulon (5, 6), although there are media specific phenotypes as well, and confirms that C media can act a good representative for a low glucose environment.

Alternative methods of sugar uptake are usually necessary for carbon acquisition during times of low glucose, thus metabolic genes involved in utilization and transport of alternative (non-glucose) sugars were induced during growth in C media. We observed that most of the alternative PTS carbohydrate EII transporters such as 3-keto-L-gulonate (*ptxABC*), mannitol (*spy_1664*), galactose (*spy_1399*), both cellobiose operons (*spy_1079-1083* and *celB*), and trehalose (*treCB*) were all up regulated in C media. The ABC transporters for sialic acid *(spy_0213-0215*), cyclomaltodextrin (*malADC*), and an unknown sugar permease (*spy_1308-1310*) were also up regulated. Interestingly, the β-glucoside-specific EIIABC and both mannose operons *ptsABCD* and *manLMNO* were repressed in C media (**Appendix 2**). This indicates, that these two operons are potentially not important for sugar uptake under this condition.

We also found several virulence-related genes to be activated in C media, including the *has* operon (capsule synthesis) and the *sag* operon (Streptolysin S), both of which are also under CCR. Although there is a large overlap between the C media

regulon and a CcpA regulon, there were also media-specific regulatory phenotypes. Interestingly, we found that an additional 25 genes found in this study have been previously shown to be regulated by CcpA (5, 6), but were regulated in a different direction. Interestingly, we observed a significant increase in the expression of both the *lacD.1* operon and *rgg/ropB*, both of which are regulators of the secreted cysteine protease and virulence factor gene *speB*. As mentioned above, C media is an inducer of *speB* expression (88); however, we did not see *speB* being differentially expressed at this time point in C media under the conditions tested.

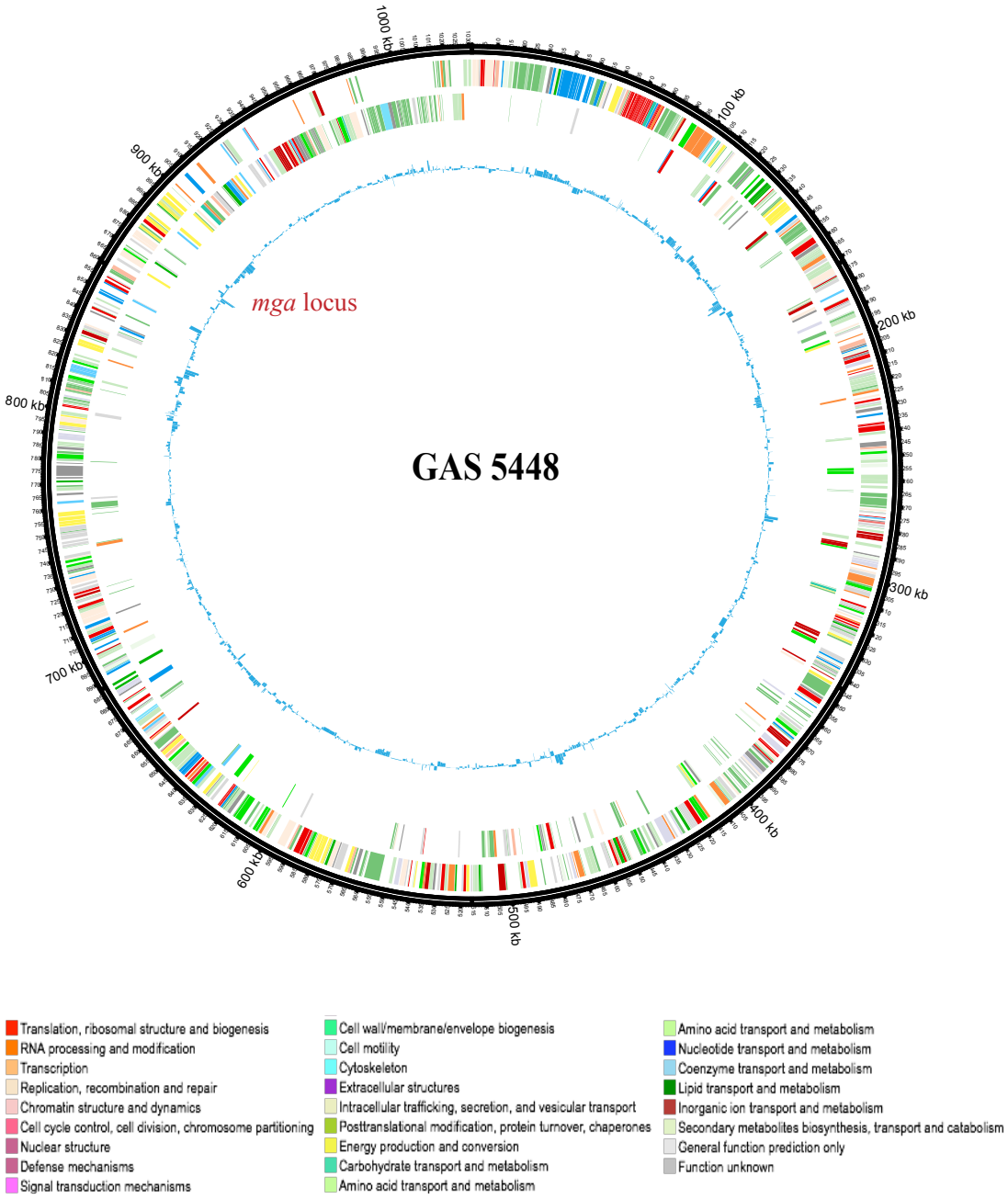


Figure 5.2. Transcriptomic landscape of 5448 during growth in C Media.

Circos plot of wild-type GAS 5448 gene global expression altered under growth in C media versus THY. The outer most ring represents a size ruler. The next two rings represent the GAS open reading frames on the (+) and (-) strand of the genome, respectively. The fourth ring shows transcript levels of genes, where the height of the bars in the fourth ring represent the $\log_2$ fold-change in deferential expression in C media vs. THY (inside is repressed, outside is induced). The *mga* locus, which is repressed by C media, is indicated. *Performed by K.M.V. and A.T.B.

We also found the majority of the core Mga regulon to be down regulated in C media (**Table 5.1**) in comparison to THY. Interestingly, we did not find the transcript levels of *emm* and *mga* to be affected. We further confirmed this observation by real-time qPCR using the same RNA from the original RNA-Seq experiments, where *mga* and *emm* transcript levels continued to be unaffected by growth in C media and the genes encoding the fibronectin binding protein (*fba*) and the C5A peptidase precursor (*scpA*) were down regulated. Interestingly, we were unable to observe a phenotype for the secreted inhibitor of complement gene *sic* using qPCR (**Table 5.1**).

Table 5.1: Repression of the Mga locus in C media

Spy #	Annotation	Gene	RNA-Seq Log_2 FC	qPCR Log_2 FC
M5005_Spy1714	cell surface protein; Mga-regulated	*fba*	2.51	1.32 ± 0.17
M5005_Spy1715	C5A peptidase precursor; Mga-regulated	*scpA*	2.20	1.26 ± 0.33
M5005_Spy1716	transposase		-1.66	--
M5005_Spy1717	transposase		-2.47	--
M5005_Spy1718	streptococcal inhibitor of complement; Mga-regulated	*sic1.0*	2.32	0.49 ± 0.14
M5005_Spy1719	M protein; Mga-regulated	*emm1*	0.74	0.84 ± 0.21
M5005_Spy1720	Multi-virulence gene regulator Mga	*mga*	0.67	0.32 ± 0.11
M5005_Spy1721	hypothetical protein		-2.86	--

Analysis of gene ontology categories found to be the most affected by growth in C media revealed that amino acid transport and metabolism (51 genes; 10.7% of total genes differentially expressed), carbohydrate transport and metabolism (59 genes; 12.4% of total genes differentially expressed), and translation, ribosomal structure and biogenesis (61 genes; 12.8% total genes differentially expressed) were the top three COG categories. These three categories correlate with the composition of C media, which is high in peptides and low in glucose, thus it is likely that *de novo*

synthesis of metabolites is necessary for GAS growth on this media. Overall, 25.5% of the genome is differentially regulated in C media in comparison THY. Taken together, the transcriptome of 5448 in C media in comparison to THY exhibits similarity to a CcpA regulon likely due to the low glucose conditions present in C media. Additionally, expression of some of the Mga regulon is activated in THY in comparison to C media during late logarithmic phase of growth.

Mga is differentially phosphorylated in C Media.

As mentioned in the previous section, we observed some of the core Mga regulon genes to be activated in THY (high glucose), but Mga itself was not differentially expressed (**Table 5.1**). Because it is know that Mga activity is dependent on phosphorylation by the PTS of GAS, and there was a global transcriptional change in the WT 5448 response to limited glucose availability, we wanted to see if the phosphorylation of Mga was altered between these two mediums. To test this, we used a Phos-tag SDS-PAGE gel matrix. The Phos-tag compound specifically forms a complex with phosphorylated amino acid residues present on the protein of interest and retards its movement through the gel matrix, as opposed to their unphosphorylated counterparts which run quicker. We hypothesized that under favorable growth conditions, such as high glucose (THY), we would observe an abundance of unphosphorylated Mga than phosphorylated Mga. Under low glucose conditions, the converse would be true, where there would be more phosphorylated than unphosphorylated Mga. To test whether Mga was differentially phosphorylated *in vivo* under these conditions, we attempted to grow a M1T1 strain containing a mutant in *mga* that was complemented *in trans* with a multi-copy plasmid, expressing

P*mga1-mga-his*$_6$ (pKSM809) (**Table 3.1**) in either THY or C media. However, we never saw sufficient expression of Mga1 by western blot and SDS-PAGE analysis (data not shown) in the conditions tested.

As an alternative, we used an existing plasmid expressing an M4 *mga* construct, P*mga4-mga4-his*$_6$, (pKSM808, WT M4) in the strain KSM547 (GA40634.Δ*mga4*) (**Table 3.1**) that has been successfully expressed in GAS for PTS analysis previously (19). The phosphoablative construct, pKSM871, which substitutes all three PTS-phosphorylatable histidines with alanines (A/A/A M4 Mga construct) was also used. The two strains were grown in either THY or C media to late logarithmic growth and whole cell lysates were isolated. Phos-tag gel analysis of the lysates followed by a western blot with anti-Mga4 antibodies revealed two distinct bands corresponding to WT Mga, a slow-migrating band arising from phosphorylated Mga and a fast-migrating band corresponding to non-phosphorylated Mga for the THY grown strain. When the strain expressing the WT Mga4 was grown in C media, only a single band corresponding to the unphosphorylated A/A/A protein was present (**Fig 5.3A**). As expected, the A/A/A M4 Mga control mutant showed the unphosphorylated species regardless of the media (**Fig 5.3A**).

As a control for equal loading, cell lysates were separated by SDS-PAGE and subjected to a western blot revealing only one Mga-specific band (**Fig. 5.3B**). Thus, the retarded migration of Mga occurs only on the Phos-tag gel and likely results from alternative phosphorylation status of Mga. Interestingly, while an equal amount of protein was detected for all strains in the Western blot, there was consistently less protein detected for the A/A/A M4 Mga mutant when grown in C media and assessed

by Phos-tag gel. Together, these data provide the first evidence that Mga can be alternatively phosphorylated between high and low glucose conditions, and that phosphorylation is most likely occurring on the PRD1/2 phosphohistidines. Overall, these data provide the first evidence that Mga can be alternatively phosphorylated under a low glucose condition.

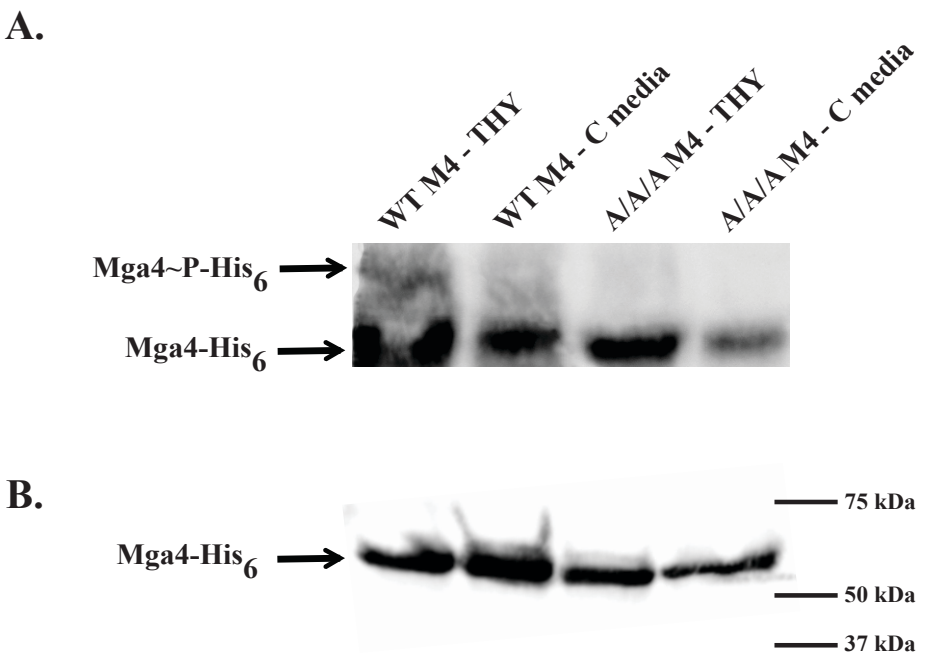


Figure 5.3. Mga is differentially phosphorylated based on glucose availability. The isogenic M4 *mga* mutant strain KSM547 complemented *in trans* with the His_6-tagged WT strain (pKSM808) or phosphoablative (A/A/A) *mga* mutant (pKSM871) was grown in THY or C medium to the late exponential growth phase, pictured is a representative image of multiple replicates. Cells were lysed using PlyC. (**A**) Equal concentrations of cell lysates of the WT and A/A/A strain were resolved on a Phos-tag-$MnSO_4$ gel and probed using a polycolonal α-Mga4 antibody. Bands corresponding to non-phosphorylated Mga and phosphorylated Mga are labeled and indicated by arrows. (**B**) Protein levels in cell lysates of the *mga* strains were determined by immunoblotting on a 10% SDS-PAGE gel using the polycolonal α-Mga4 antibody.

The M1T1 Mga regulon in THY versus C media.

Since we observed differential phosphorylation patterns for a WT Mga strain under different glucose availability and media conditions, we hypothesized that the Mga regulon in these two medias would vary. A previous study (229) characterized the Mga regulon in THY at late logarithmic growth across three serotypes (M4, M6, and a lower virulence M1) using a microarray. Due to the in-depth resolution of RNA-Seq over microarrays, the clinical relevance of the M1T1 clinical isolate 5448, and the need for comparison to C media, determination of the THY Mga regulon again was warranted. We constructed an insertional inactivation of *mga* in 5448 using a temperature sensitive replicating plasmid to generate the *mga* mutant, 5448.930 (**Table 3.1)**. Growth kinetics of 5448.930 were found to be comparable to parental 5448 in both THY and C media (data not shown). Total RNA was isolated from the WT 5448 and the Δ*mga* strain grown in either THY or C media to late logarithmic phase in two biological replicates and processed for RNA-Seq. Data obtained from the parental 5448 was compared to the Δ*mga* grown cells in either media (WT/Δ*mga*) with changes in gene expression over 1.8-fold ($\log_2 \leq -0.85$ or ≥ 0.85) and a p value of ≤ 0.05 were considered significant (**Appendix 3 and 4**).

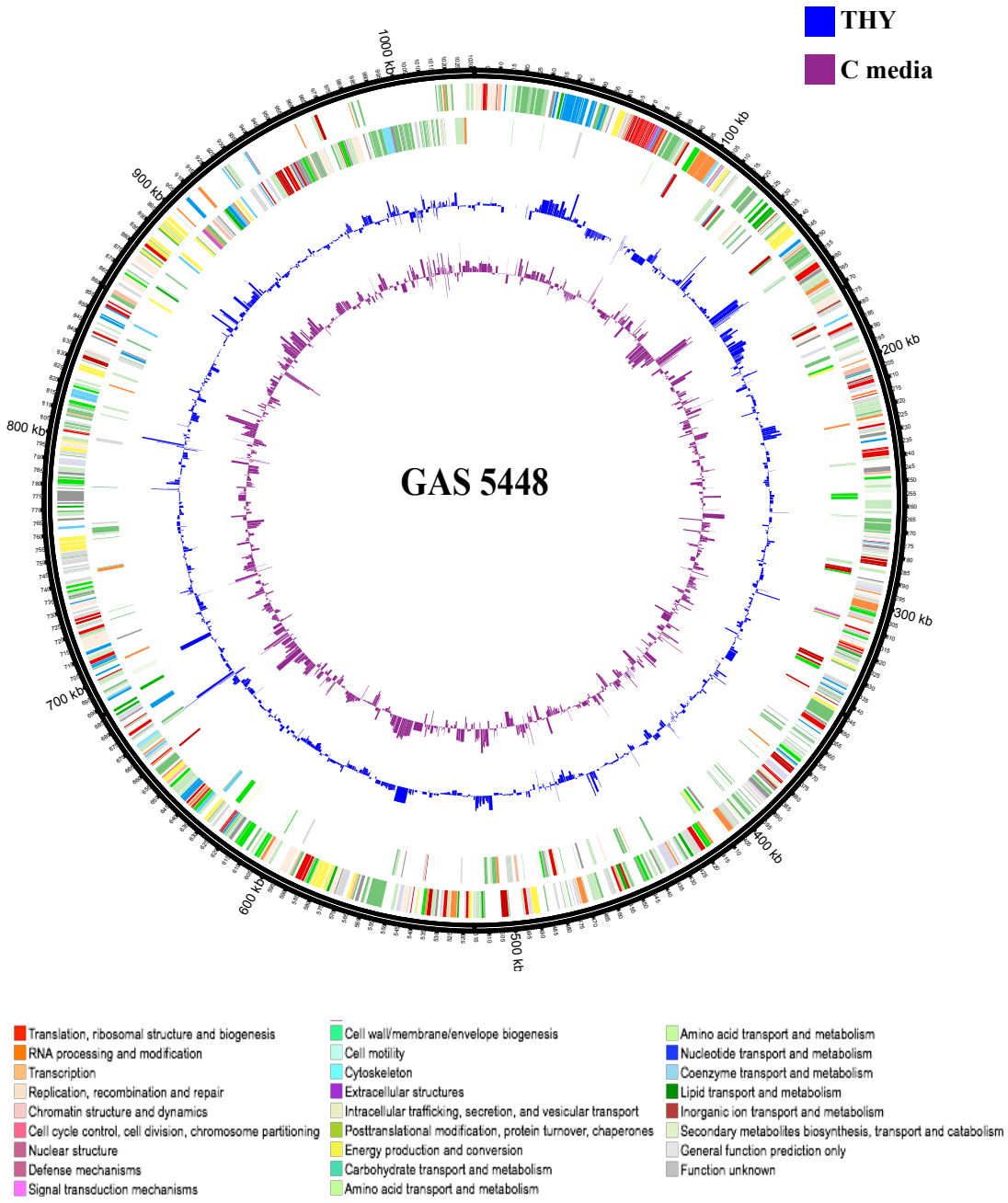


Figure 5.4. Transcriptomic landscape of Δ*mga* mutant during growth under varying glucose conditions.

Circos plot of the differentially expressed (DE) genes between wild-type GAS 5448 versus a Δ*mga* mutant under growth in either C media or THY at late logarithmic growth. The outer most ring represents a size ruler. The next two rings represent the GAS open reading frames on the (+) and (-) strand of the genome, respectively. The fourth ring shows transcript levels of genes, where the height of the bars in the fourth ring represent the $\log_2$ fold-change in deferential expression in THY (blue) vs. C media (purple) (inside is repressed, outside is induced). *Performed by K.M.V. and A.T.B.

As expected for both medias, the most activated genes were those of the core regulon (*emm, sic, scpA,* etc.), which all showed between an 18- and 7-fold reduction in transcript levels (**Table 5.2**; **Appendix 3 and 4)**. Interestingly, we did not observe a Mga-specific phenotype for *sclA* in either RNA-Seq datasets. *sclA* is a gene consider to be apart of the core Mga regulon that contains a category B promoter (see Chapter 2, Mga DNA binding section) and is directly regulated by Mga (20, 111, 229, 243). Lack of finding *slcA* in our data could indicate that *sclA* is under the control of a different regulatory mechanism in the M1T1 5448, is not expressed in this background, or it could signify that the computational analysis had difficulty mapping this gene. Thus, we performed real-time qPCR on both the WT 5448 and 5448.930 Δ*mga* strain using the same RNA that was used in the RNA-Seq for both conditions. We observed hat *sclA* was still regulated by Mga via real-time qPCR, and could be complemented (**Fig 5.5AB and Fig. 5.6AB**), therefore it appears that there was some issue with read alignment or calculations of differential expression for this gene. There are two *scl* genes (*sclA* and *scl2*) encoded in M1 GAS strains, the second of which is not regulated by Mga. However, because of their homology and the restrictions of mapping reads with Bowtie (258, 259) this second gene may artificially add reads to the Mga-regulated *sclA* gene that would otherwise correspond to *scl2*. This could help explain the abnormal phenotype of *sclA* regulation in the RNA-Seq data, which was not recapitulated in the qPCR results.

(i) M1T1 Mga regulon in THY

A total of 155 genes (including 7 tRNA genes) were found to be significantly regulated by Mga with approximately equal numbers of repressed (75) and activated

(80) genes (**Appendix 3; Fig. 5.4**, blue ring). The chromosomally encoded DNase *spd*, was also shown to be activated by Mga; transcript levels were reduced by 4-fold in comparison to WT 5448. As opposed to what was seen from the microarray study (229), *speB* is repressed by Mga in a M1T1 background (**Appendix 3**; **Fig 5.5A**). Mga also repressed *spi*, an inhibitor of *speB*, to similar levels (**Appendix 3**), indicating that proper expression of SpeB may be critical for GAS under this condition. The microarray study also showed Mga to indirectly regulate operons involved in sugar metabolism, this phenotype was recapitulated in our RNA-Seq experiment as well; however, here, the majority of sugar-related operons were activated by Mga and not repressed as was previously observed (229).

In order to validate the RNA-Seq results generated from THY, a total of 8 genes were chosen for analysis via qPCR; 6 genes which are activated, 2 which are repressed (**Fig. 5.5A**). In addition, 5448.930 was passaged at the permissive temperature in order to generate the rescue strain, 5448.930_R, which was used to complement the phenotypes seen in both the RNA-Seq and qPCR data (**Fig. 5.5B**). The real time qPCR results confirmed the RNA-Seq data with a correlation coefficient of 87% (**Fig. 5.5C**). Overall, a total of 8.3% (7.9% without tRNA genes) of the genome was regulated by Mga in THY at late logarithmic phase in an M1T1 background. While there appears to be similarities between the broad categories of genes regulated by Mga in our study with those observed by Ribardo *et al*. in the microarray experiment, there were several differences on both the individual gene level as well as directionality of gene regulation by Mga. Taken together, these data

show a Mga regulon that is specific to the M1T1 GAS in THY (high glucose) at late logarithmic growth, indicating that the Mga regulon can vary across serotypes.

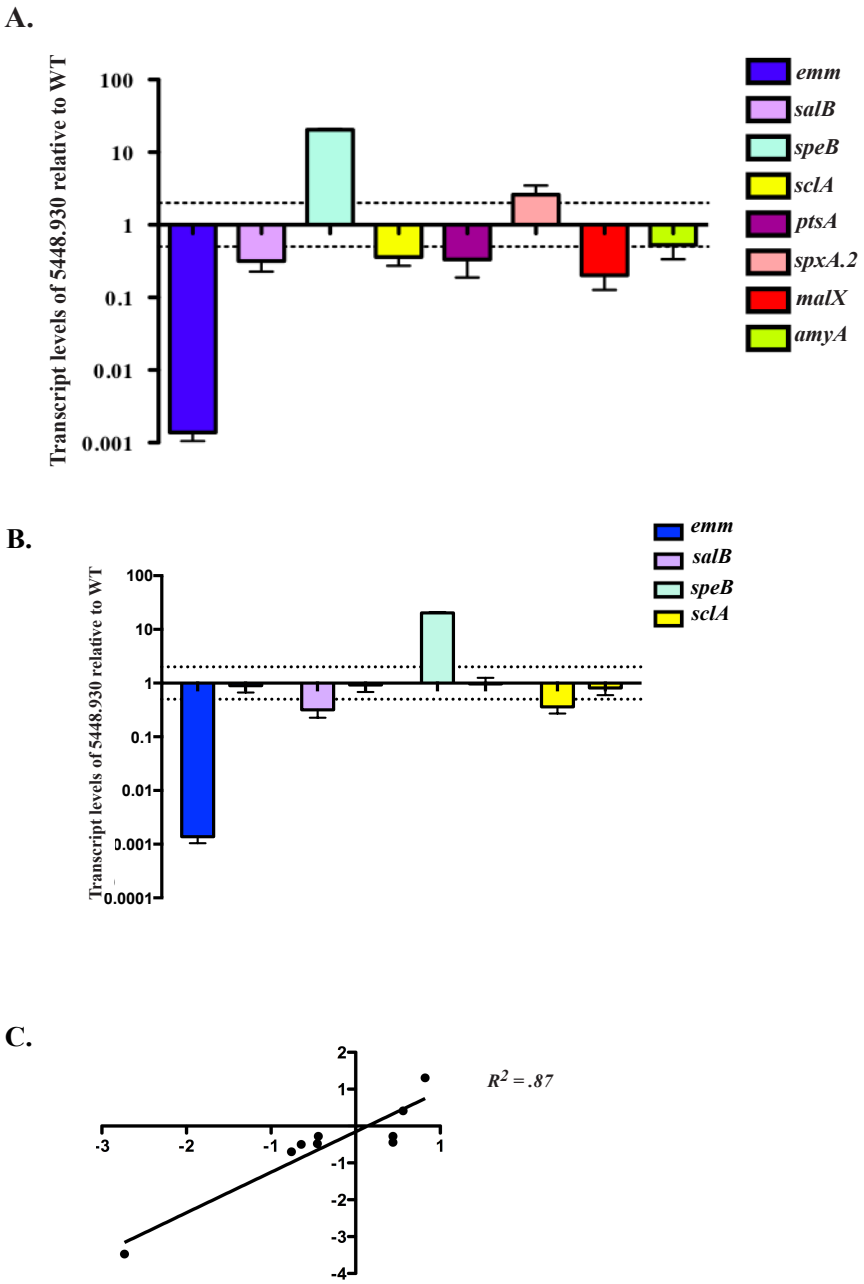


Figure 5.5. Validation of THY RNA-Seq results using qPCR.
(**A**) Transcript levels were determined using qPCR on RNA taken from WT GAS 5448 grown in THY at late-log compared to 5448.930 grown under the same conditions. Error bars represent the standard error from three biological replicates. Differences greater than 2-fold in expression for fructose grown cells compared to glucose grown cells (denoted by a dashed line) are considered significant. (**B**) Complementation of the Δ*mga* mutant using the rescue strain 5448.930_R restores the transcriptional phenotypes observed in THY. (**C**) Coefficient of correlation between the RNA-Seq and RT-qPCR results generated a confidence in the data with an R^2 =87%.

(ii) M1T1 Mga regulon in C media

For C media, a total of 148 genes (including 13 tRNA genes) were regulated by Mga (57 repressed, 91 activated) in C media, representing 7.9% (7.2% without tRNA genes) of the genome (**Fig. 5.4**; purple ring). We also observed novel virulence factors and metabolic genes to be regulated in C media that were not observed in THY. For example, both hemolysin operons Streptolysin S (*sagA-H*) and Streptolysin O (*slo*), Streptokinase (*ska*), Streptococcal secreted esterase (*sse*), the heme utilization operon (*hupYZ*), and the osmotic stress operon (*opuAA/opuABC*) were regulated by Mga in C media. In addition, Mga did not regulate *speB* in this condition (**Fig 5.6**, light blue bar), indicating that *speB* is Mga regulated in potentially a glucose-specific manner. We also observed *grm* to be Mga regulated in a C media-specific manner. *grm* encodes for a 78 amino acid hypothetical cytosolic protein that was directly regulated by Mga at its Category B promoter (P*grm*), and was suggested to be a part of the core Mga regulon (229), however its role in GAS physiology has never been identified.

A total of 9 genes were chosen for validation via qPCR. We also used the rescue strain, 5448.930_R, to complement the phenotypes seen in both the RNA-Seq and qPCR data (**Fig. 5.6B**). The real time qPCR results confirmed the RNA-Seq data with a correlation coefficient of 95% (**Fig. 5.6C**). Overall, we observed a broader induction of known virulence factors in a Δ*mga* background in C media. Taken together, these data indicates that there is a glucose-specific effect on the Mga regulon in the M1T1 GAS that occurs most likely due to the altered phosphorylation of Mga at PRD1/PRD2.

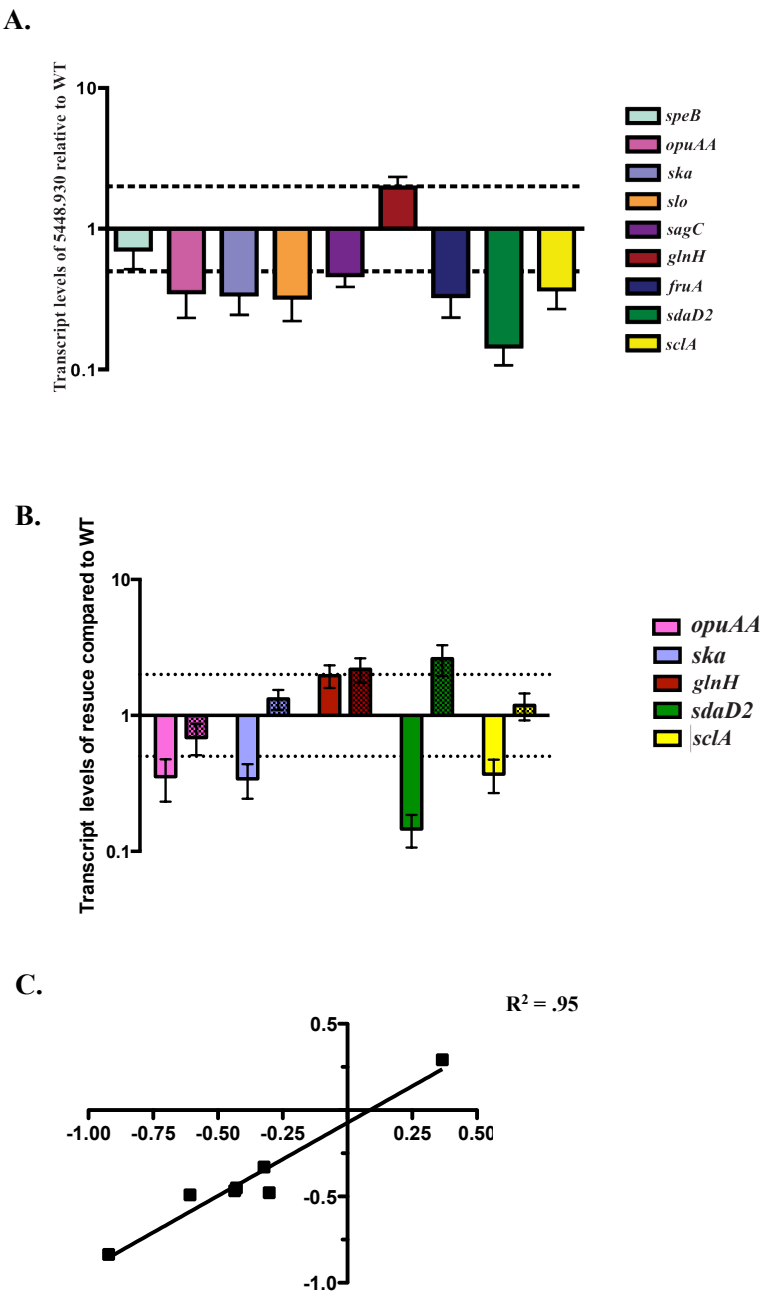


Figure 5.6. Validation of C media RNA-Seq results using qPCR.
(**A**) Transcript levels were determined using qPCR on RNA taken from WT GAS 5448 grown in C media at late-log compared to 5448.930 grown under the same conditions. Error bars represent the standard error from three biological replicates. Differences greater than 2-fold in expression for fructose grown cells compared to glucose grown cells (denoted by a dashed line) are considered significant. (**B**) Complementation of the Δ*mga* mutant using the rescue strain 5448.930_R (hashed bars) restores the transcriptional phenotypes observed in C media. (**C**) Coefficient of correlation between the RNA-Seq and RT-qPCR results. Generated a confidence in the data with an $R^2 = 95\%$.

Comparison of the Mga regulon under varying glucose availability.

Analysis of the Mga regulon in either THY or C media revealed a significant difference between the two data sets. Not only were the genes that are regulated by Mga in either condition different, but also the number of genes that are activated/repressed in these conditions were also different; THY had equal numbers whereas C media had more activated genes. Furthermore, while we expected a difference in gene regulation between the high and low glucose conditions, we did not expect to see the so little in common between the two data sets. In total, only 19 genes, 3 of which encode for tRNAs, were shown to be in common between the two Mga regulons (**Fig.5.7; Table 5.2**). The majority of genes that were consistently regulated between both medias were the core regulon genes (**Table 5.2;** bold, shaded grey). In addition, *glnH* and *glnQ.2* (bicarbonate transporter), *mac* (IgG-degrading protease), *spd* (chromosomally encoded streptodornase), *M5005_spy0522* (sugar hydrolase), *M5005_ spy0143* (hypothetical protein), and *spxA.2* (a transcriptional regulator) were also regulated in both medias.

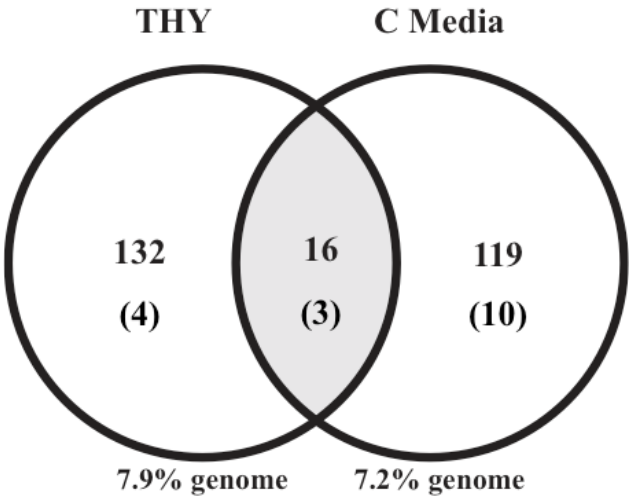


Figure 5.7. Genes in common between the Mga regulons in medias.
Venn diagram showing the numbers of Mga-regulated genes by either THY (left) or C media (right), and their overlap with each other. A total of 16 genes were in common between the two data sets (shaded grey). In parenthesis is the number of tRNA genes found in each dataset. Percentage of regulation was calculated by excluding tRNA genes.

Albeit the stark difference in genes that are regulated by Mga in a high glucose versus a low glucose condition, the COG categories that were most affected were not different. In THY, the top three COGs regulated by Mga were "unknown function" (23 total; 8/15) "carbohydrate utilization" (32 total; 31/2) "defense mechanism" (18 total; 12/6) (activated/repressed) (**Fig. 5.8**). In C media, the "unknown function" COG still had the highest representation (32 total; 13/19), this was followed by "defense mechanism" (19 total; 18/1), and "energy production and conservation" (17 total; 3/14) (activated/repressed) (**Fig. 5.8**). Taken together, while the genes within the COG categories vary drastically between the two different media

conditions, the broad categories that these genes belong to do not vary. This confirms that Mga is a master regulator for genes involved in virulence factor production and also energy production/carbohydrate metabolism.

Table 5.2: Genes regulated by Mga in both THY and C Media (WT 5448/Δmga)

Spy Number	Annotation	Gene	THY Log_2 FC	C Media Log_2 FC
M5005_Spy0113	transposase		1.47	1.32
M5005_Spy0143	hypothetical protein		2.54	3.35
M5005_Spy0522	unsaturated glucuronyl hydrolase		1.45	0.98
M5005_Spy0668	IgG-degrading protease	*mac*	1.62	1.37
M5005_Spy0824	2-amino-4-hydroxy-6-hydroxymethyldihydropteridine pyrophosphokinase	*folK*	-1.00	-0.87
M5005_Spy1076	Putative Bicarbonate transporter	*glnH*	-1.31	-0.85
M5005_Spy1077	Putative Bicarbonate transporter	*glnQ.2*	-1.04	-1.04
M5005_Spy1714	cell surface protein; Mga-regulated	***fba***	3.20	4.79
M5005_Spy1715	C5A peptidase precursor; Mga-regulated	***scpA***	3.43	5.04
M5005_Spy1718	streptococcal inhibitor of complement; Mga-regulated	***sic1.0***	4.63	5.53
M5005_Spy1719	M protein; Mga-regulated	***emm1***	9.08	10.06
M5005_Spy1720	Multi-virulence gene regulator Mga	***mga***	4.60	5.84
M5005_Spy1721	hypothetical protein		-2.78	-5.00
M5005_Spy1733	hypothetical protein		-2.54	1.45
M5005_Spy1738	phage-associated deoxyribonuclease	*spd*	0.98	1.03
M5005_Spy1798	suppressor of clpP/X, SpxA homolog, allele 2	*spxA2*	-1.86	-1.07
M5005_SpyT0041	tRNA-Leu		1.47	2.38
M5005_SpyT0046	tRNA-Tyr		-2.63	1.48
M5005_SpyT0065	tRNA-Asn		-1.43	1.76

*Bold, shaded grey denotes the core Mga regulon

Importantly, in addition to complementation, we were also able to control for Mga regulation of the genes identified in these two datasets by using an alternative, independent Δ*mga* mutant. The stable transposon mutant, KM16.1 (**Table 3.1**), was used for real time qPCR on selected targets in either C media or THY. KSM16.1 contains a KRMIT (295) insertion at the 5'-end of *mga*. Four genes originally chosen for validation in THY (**Fig. 5.9A**) and nine genes that were identified in C media (**Fig. 5.9B**) were tested using this mutant. We observed an identical phenotype in gene regulation for KM16.1 in both medias in comparison to WT 5448 (**Fig. 5.5A and 5.6A**). Therefore, the regulon observed in either THY or C media is consistent in M1T1 GAS.

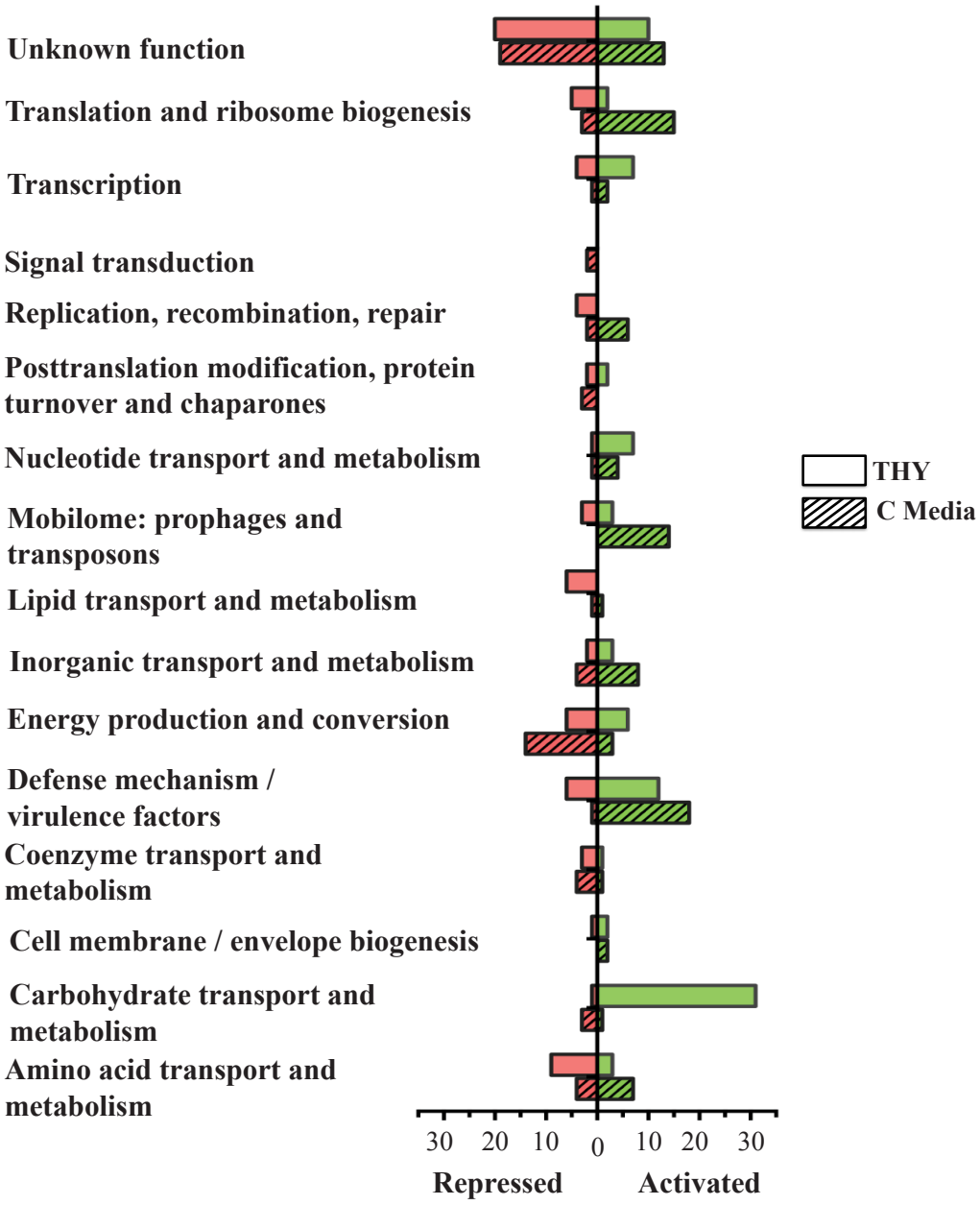


Figure 5.8. Comparison of the COG categories observed in THY and C media for a Δ*mga* strain.

A comparison of the differentially expressed (DE) genes in various cluster of orthologous genes (COG) categories from THY versus C media of a Δ*mga* mutant at late logarithmic growth when compared to the WT 5448. Bar length depicts the number of DE genes either down regulated (red) or up regulated (green) at *p* value of ≤0.05 with the total number of down and up regulated genes shown. Plain (THY) or hashed (C media) boxes indicate media condition.

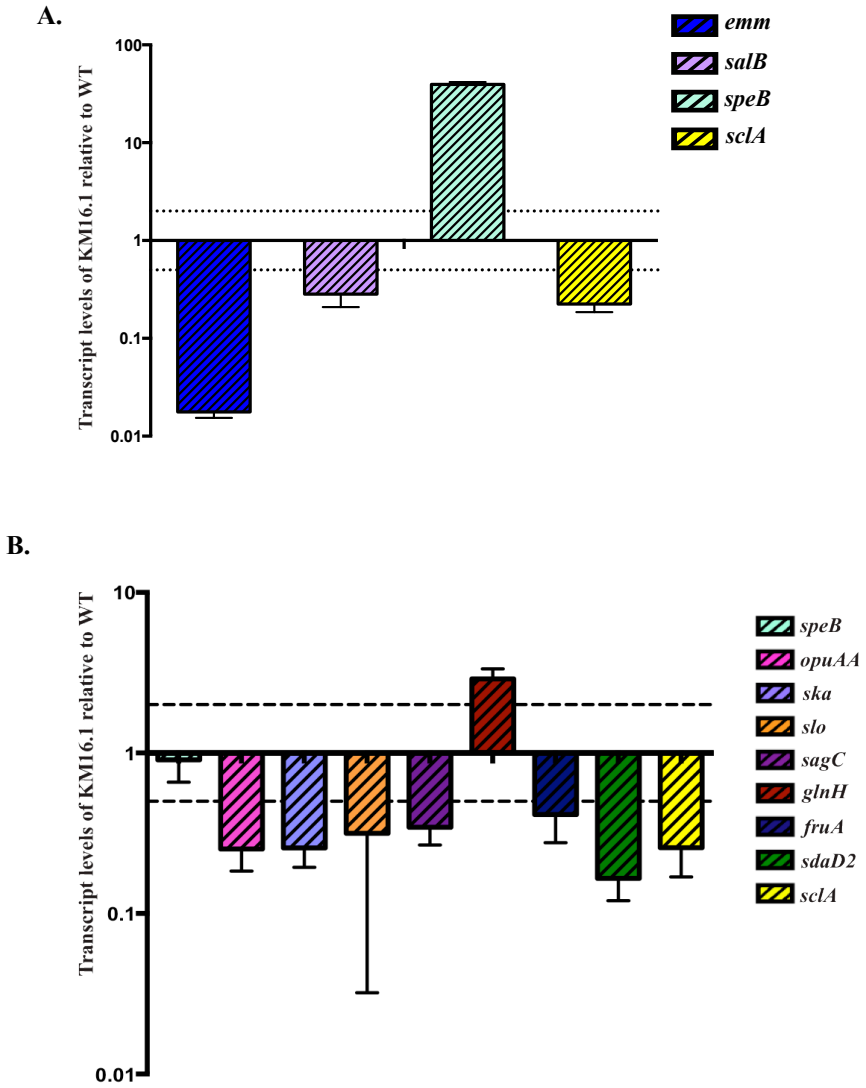


Figure 5.9. Transcriptional regulation of genes by Mga is recapitulated in an alternative Δ*mga* mutant.
Transcript levels were determined using qPCR on RNA taken from WT GAS 5448 grown in THY (**A**) or C media (**B**) at late-log compared to KM16.1 grown under the same conditions. Error bars represent the standard error from three biological replicates. Differences greater than 2-fold in expression for fructose grown cells compared to glucose grown cells (denoted by a dashed line) are considered significant.

Supplementation of glucose in C media restores some phenotypes.

Mga has an altered regulon in C media (low glucose) in comparison to THY (high glucose) that is likely to due to the phosphorylation status of the protein (**Fig. 5.3A**). We were interested in testing whether supplementing C media, which has 0.05% (w/v) of glucose to a final concentration of 0.5% (w/v) (the amount that is present in THY), would mimic the THY regulon. RNA was isolated for real-time qPCR from 5448.930 and the parental 5448 grown in C media supplemented to 0.5% glucose. We chose a total of 5 of differentially-regulated genes that were exclusive to the C media RNA-Seq dataset (**Appendix 4; Fig. 5.6A and 5.9B**) including *speB,* which was only found to be regulated in THY (**Fig. 5.5A, 5.6A, 5.9; Appendix 3 and 4**) and *emm* which was observed in both datasets. We saw that 4 out of the 5 differentially regulated genes specific to C Media were no longer regulated by Mga when glucose was supplemented back to levels present in THY, which includes the phage-encoded DNase *sdaD2* (**Fig. 5.10;** green bar). Interestingly, *opuAA* (**Fig.5.10**; pink bar) was the only gene that was not complemented by glucose, which potentially indicates that growth in C media causes osmotic stress and is likely regulated, indirectly by Mga.

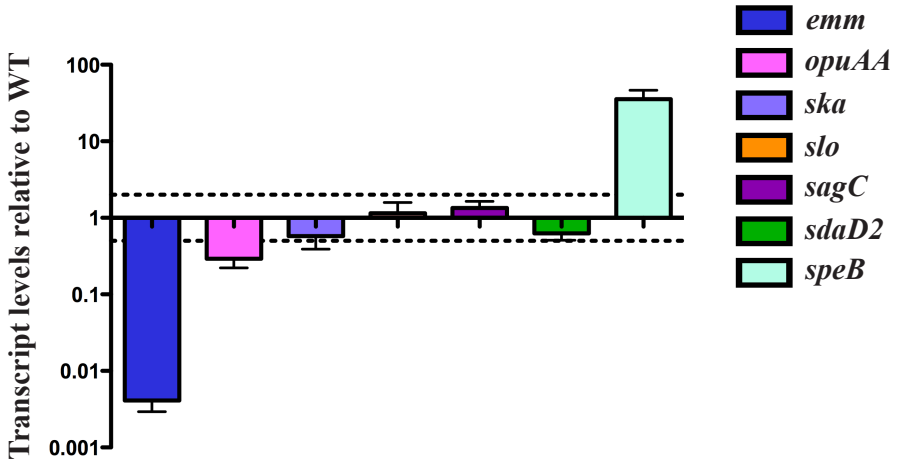


Figure 5.10. Supplementation of C media with glucose mimics some phenotypes observed for a Δ*mga* strain in THY.
Transcript levels were determined using qPCR on RNA taken from WT GAS 5448 grown in C media supplemented to 0.5% glucose at late-log compared to 5448.930 grown under the same conditions. Error bars represent the standard error from three biological replicates. Differences greater than 2-fold in expression for fructose grown cells compared to glucose grown cells (denoted by a dashed line) are considered significant.

Interestingly, Mga was shown to activate the transcription of *speB* in THY at late log phase (**Fig. 5.5A and Fig. 5.9A**; light blue bar) and this regulation phenotype was abolished in C media (**Fig. 5.6A and Fig 5.9B**; light blue bar). Therefore, we wanted to test whether *speB* regulation could be restored when glucose was supplemented back into C media. Indeed, upon the addition of glucose to 0.5% (w/v), *speB* was regulated by Mga once again (**Fig 5.10**; light blue bar). Therefore, while the mechanism by which Mga regulates *speB* is most likely indirect, there is a direct correlation between glucose availability Mga regulation of *speB*. Overall, there

appears to be two causes to the changes observed between THY and C media: 1) the majority of the changes are specific to glucose availability, and therefore are probably directly related to the phosphorylation status of Mga. 2) A minor subset of changes in the regulon dynamics are due to C media-specific differences.

Phage-encoded DNase, Sda1, is subject to C media-specific Mga regulation.

In C media, Mga transcriptionally activates all three phage-associated DNases or streptodornases, including the M1T1 distinguishing phage, *sdaD2* (encoding Sda1). The RNA-Seq data showed that there was over a 5-fold reduction in transcript of *sdaD2* in a Δ*mga* mutant (~8-fold qPCR; **Fig. 5.6A and Fig 5.9B;** green bar) in a low glucose environment. Interestingly, regulation of *sdaD2* by Mga (**Fig. 5.10**; green bar) was complemented upon glucose supplementation, confirming that regulation of this streptodornase occurs in a glucose-dependent manner. Sda1 is encoded on Φ5448.3, one of the mobile genetic elements that is responsible for the emergence of the M1T1 serotype (145, 296). Due to the reduction of transcript level observed in a low glucose environment (C media), we wanted to test if this would be reflected at the protein level. We grew WT 5448 and the two independently constructed Δ*mga* mutants (5448.930 and KM16.1) in C Media in the presence of E-64 protease inhibitor, and observed expression of Sda1 in culture supernatants. There was significant decrease in protein production of Sda1 in both Δ*mga* mutants in comparison to WT 5448 in several independent replicates (**Fig. 5.11**). We also attempted correlate this phenotype to a functional defect in a Δ*mga* mutant to degrade DNA. Unfortunately, after several attempts using a variety of DNase assay methods,

we were unable to see a significant decrease in the ability for a Δ*mga* mutant to degrade DNA. However, while Mga does activate all three streptodornases in 5448 in C media, the amount of DNase produced from a Δ*mga* background may be sufficient enough to not observe a difference between the two strains. Regardless, Mga clearly regulates this serotype-specific virulence determinant in a glucose-dependent manner in the M1T1 strain 5448.

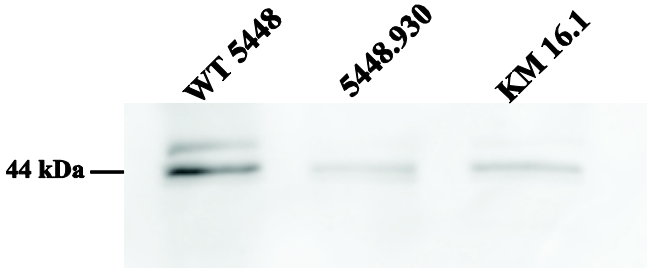


Figure 5.11. Sda1 expression is reduced in a Δ*mga* background in M1T1 GAS. Western blot analysis of the expression Sda1 (encoded by *sdaD2*) (44 kDa), the Φ5448.3 phage-encoded streptodornase in culture supernatants grown to late-log phase (Klett ~70) in the WT 5448 compared to two independent Δ*mga* mutants (5448.930 and KM16.1). Culture supernatants were concentrated using methanol-chloroform precipitation, and equal amounts were run on a 10% SDS-page gel and probed using a polyclonal α-Sda1.

Discussion

The current study was undertaken to assess how glucose availability affects global transcriptional changes in the cell; in particular how it affects Mga activity. Our results reveal that some of the Mga regulon is up regulated in THY, a high glucose environment. In addition, there was a strong correlation between the genes activated in the C media dataset and those previously shown to be repressed by CcpA, indicating that C media can relieve CCR. We show that the difference in Mga activity in THY is likely due to alternative phosphorylation patterns of Mga in comparison to C media through its PRD domains. We then tested the differences of the Mga regulon under these two glucose levels and found that while a similar percentage of the genome was regulated under both THY and C media, there was very little overlap. Importantly, we show that a high percentage of the differences in the Mga regulons are directly attributed to glucose. Thus, this is the first time that Mga and its regulon are influenced by glucose, revealing an *in vivo* link between glucose availability and activation of a global transcriptional network that plays a key role in virulence gene regulation.

The C media-mediated regulon of M1T1 GAS.

For low G+C Gram-positive bacteria, carbon catabolite repression (CCR) is controlled by the master regulator, CcpA through direct binding of catabolite response elements (*cre*) sites. Metabolic genes for utilization of alternative (non-glucose) sugars are typically induced during lack of the preferred carbohydrate source, glucose. For GAS, two genome-wide CcpA-induced regulons in M1T1 GAS

have been determined (5, 6). Both studies found that ~8% of the GAS genome is regulated by CcpA in THY. We found that 15.8% of the total number of genes differentially expressed in C media overlapped with the CcpA regulon (5, 6), indicating that the low glucose environment encountered by GAS in C media is capable of relieving CCR and stimulating the transcription of alternative carbohydrate uptake systems. Although there is clear overlap between these two dataset, our data also showed that GAS undergoes a much more broader global transcriptomic change during growth in C media (low glucose) as 25.5% of the entire non-prophage GAS genome (476 genes) were differentially expressed when compared to THY (high glucose) (**Appendix 2**). Thus, there is a media specific regulon that encompasses some of the CcpA regulon, which indicates that growth in C media can mimic a limited glucose environment, but it also creates other changes in the cell that are most likely attributed to its high levels of peptides.

The C media-induced regulon has been established for the M5 strain, HSC5, using altered buffered and osmotic conditions at early stationary growth phase (88). However, this is the first study to investigate the global regulatory changes that occur when GAS is grown in THY compared to C media using RNA-Seq in an M1T1 strain. As mentioned previously, C media is a low in glucose but is also peptide rich. This media has been shown to be a representative of conditions encountered by GAS during deep tissue infection due to up regulation of the virulence factor and cysteine protease, *speB*, to levels comparable to those observed during *in vivo* infection (88). Thus, C media has been used in several studies as an inducer of *speB* expression. Surprisingly, we did not observe a significant induction of *speB* transcript levels in C

media as expected. However, this study was conducted in late logarithmic phase, and not stationary phase, where *speB* is known to be most highly expressed. However, we did observe an increase in expression both the *lacD.1* and *rgg/ropB*, which have been shown to interact and together stimulate *speB* activity (297). Therefore, due to the growth phase constraint of our data set, the conditions tested here may not be optimal for the induction of *speB* expression. To answer this questions, we could conduct qPCR on RNA isolated a various stages in growth, including stationary phase to assess *speB* expression. Another possibility of the lack of observable *speB* expression is that since C media acts as a growth media that mimics deep tissues infection (88), it is possible that *speB* is not being expressed in the WT 5448. A study found that *in vivo*/environmental signals suppress the expression of *speB*, and induce the expression of *speA* (298). While C media can induce the expression of *speB* in other serotypes, it is possible that it cannot function in the same for the M1T1 serotype. However, there was no observable change in transcript level for *speA* in our RNA-Seq results of the WT 5448 grown in either THY or C media. Therefore, it is likely that the first scenario is what is occurring.

Plasticity of the M1T1 Mga virulence regulon under varying glucose availability.

The multiple gene regulator of GAS, or Mga, has a well-characterized regulon encompassing both virulence and carbohydrate metabolism genes (229). However, questions remain with regards to how Mga, through activation of its regulon, may differ due to availability of a preferred carbohydrate source (glucose). To date, a direct global comparison of the M1T1 5448 Mga regulated genes in a limited glucose

(C media) to a rich glucose (THY) environment has not been undertaken. In this study, we used RNA-Seq in order to identify differences in the Mga regulon under different nutrient conditions.

(i) The M1 versus the M1T1 Mga regulon in THY

It was previously reported that 10% of the genome was regulated by Mga at late-logarithmic growth in THY; however, this study was conducted across 3 serotypes (M1, M6, and M4) using a microarray (229). Our RNA-Seq experiments revealed that the Mga regulon in an M1T1 GAS strain under similar growth conditions regulated 7.9% of the genome. While RNA-Seq is a much more sensitive technique than a microarray, we performed our analysis on only a single serotype (M1T1) as, opposed to three serotypes (M1, M6 and M4). Therefore, it is possible that the regulon is smaller when only one genome is under consideration. Although the M1 and M1T1 serotypes only have a 7% difference between their core genomes, and the majority of these differences lie in their prophage content (299), comparison of the Mga regulon datasets revealed very little overlap. Only 8 genes were regulated by Mga in the same direction, with the majority (5 genes) being the core regulon (**Table 5.3**). The other three genes were *lacG* (6-phospho-beta-galactosidase), *lctO* (lactate oxidase), and *M5005_spy0143* (hypothetical protein). Interestingly, *M5005_spy0143* was also found to be regulated in the C media dataset (**Fig. 5.7** and **Table 5.2**), therefore further experimentation should be done to characterize this small hypothetical protein as it is the only gene in common across the three different experiments besides the core regulon. Both *lacG* and *lctO* are involved in lactose metabolism. Interestingly, *lctO* converts lactate into pyruvate, but is also involved in

production of hydrogen peroxide (H_2O_2) in a growth phase- and glucose-dependent manner through direct repression by CcpA (7). In *S. pneumoniae*, increase of H_2O_2 production acts as a virulence trait that aids in colonization by slowing the clearance of bacteria through the inhibition of ciliary beating in the upper respiratory tract and by promoting bacteriocidal activity to kill other competing with other naturally occurring microflora (300-302). Therefore it's possible that Mga activates the transcription of *lctO* in order to up regulate hydrogen peroxide production just enough to aid in colonization, which is the stage of infection that correlates with the highest level of Mga activation (228).

Table 5.3. Comparison between the M1 and M1T1 Mga regulon in THY at late-log

MGAS5005	Gene name	Annotation	Log_2 FC - RNA-Seq	SF370	Log_2 FC - Microarray
M5005_spy0143		hypothetical protein	2.54	SPy0169	1.95
M5005_spy0340	*lctO*	L-lactate oxidase	2.47	SPy0414	1.30
M5005_spy1632	*lacG*	6-phospho-beta-galactosidase	2.06	SPy1916	1.59
M5005_spy1714	*fba*	cell surface protein	3.20	SPy2009	5.59
M5005_spy1715	*scpA*	C5A peptidase precursor	3.43	SPy2010	5.67
M5005_spy1718	*sic1.0*	streptococcal inhibitor of complement	4.63	SPy2016	6.13
M5005_spy1719	*emm1*	M protein; Mga-regulated	9.08	SPy2018	6.13
M5005_spy1720	*mga*	Multi-virulence gene regulator Mga	4.60	SPy2019	1.96

The THY Mga regulon showed several sugar-specific operons to be activated by Mga, which was in contrast to what was observed previously, where all metabolic operons were repressed (229). However, for this invasive clinical isolate M1T1 strain, it appears that Mga regulates carbon metabolism genes in a manner that mimics what was observed for MafR, a recently characterized Mga-homolog and PRD-containing virulence regulator (PCVR) that has also been implicated in the pathogenesis of *E. faecalis* (218). Thus, this could present a much more accurate representation of the Mga regulon in invasive GAS strains. Regardless of the difference that were observed between the microarray and our RNA-Seq results, it is apparent that Mga still regulates a variety of genes that are important for adhesion and initial colonization and plays as a role, albeit indirect, in regulation of metabolic operons under high glucose conditions. As expected, under both media conditions, the genes that showed the strongest levels of regulation corresponded to the "core" Mga regulon genes (**Table 5.2**). These genes are defined as those that are directly regulated by Mga (235, 243) and are located in the Mga locus (*emm*, *sic*, *scpA*, and *fba*).

Of note, the cysteine protease, *speB,* and the SpeB inhibitor (*speI*) were both observed as being repressed by Mga (**Appendix 3; Fig 5.4A,** light blue bar). It is curious, that both the protease and its inhibitor are simultaneously repressed, and may point to a feedback loop between these two genes in M1T1 GAS that keeps expression levels this virulence factor in under tight control. *speB* was previously observed as being activated by Mga in a serotype specific manner (229), however here we observe *speB* to be repressed. Interestingly, no change in the transcript of *ropB*/*rgg* was observed in a Δ*mga* mutant, which is known to regulate *speB*

expression (4). In addition, Mga only activates one gene in the regulatory *lacD.1* operon, *lacB.1*, in THY at late log. Interestingly, regulation of *speB* by Mga was restored when glucose was added back to C media (**Fig 5.10**). Taken together, this demonstrates that there is a Mga-dependent, glucose-dependent regulation of *speB* within the M1T1 strain 5448. Furthermore, there appears to be several serotype specificities for a Mga regulon even amongst closely related strains.

(ii) Comparison between THY and C media Mga regulons

In both THY and C media at logarithmic growth, ~8% (155 genes, THY; 148 genes, C media) of the non-prophage genome was regulated by Mga (**Fig. 5.4 and 5.7**). Interestingly Mga also affects the same COG categories in both of these medias as well (**Fig. 5.8**); however, there was a stark contrast in the genes represented between the two datasets. A total of 16 genes overlapped between THY and C media, half of which are a part of the core Mga regulon (**Table 5.2, Fig. 5.4 and 5.7**). The remaining genes, were not previously identified as a part of the Mga regulon (229), and therefore could be unique to the M1T1 core Mga regulon. Interestingly, Mga repressed both genes (*glnH* and *glnQ*) that encode for the putative bicarbonate transporter in both medias. Mga has been shown to be activated under elevated CO_2 conditions (238), and this environmental signal may act independent of other activation signals, as is seen in other Mga homologs, such as the *B. anthracis* PCVR regulator AtxA and the group G Streptococcus (GBS) DmgB (11). Therefore, a potential link between bicarbonate transport and Mga gene regulation may exist. Experiments that elucidate bicarbonate transport in a Δ*mga* mutant may shed more light on this connection. *M5005_spy0522,* which encodes for a sugar hydrolase, was

also found in both data sets. Since Mga is intimately tied to sugar status of the cell via the PTS is would not be surprising that if it directly regulates a sugar hydrolase. Since the number of genes that overlap between the THY and C media regulon datasets are so few, we propose that this set of these novel 9 genes may also be a part of the core M1T1 Mga regulon. Further DNA binding assays need to be performed in order to determine if Mga can directly regulate these genes via binding to their promoter regions.

(iii) Regulation of the M1T1 Streptodornases by Mga.

Of note, all three phage encoded DNases (*spd, sda1/sdaD.2,* and *spd3*) were activated by Mga in C media (**Appendix 4**); however, the chromosomally encoded *spd* was activated in both conditions (**Table 5.2**). This is not the first instance demonstrating that a stand-alone regulator can play a role in regulation of a phage-encoded streptodornase. In the M49 strain, NZ131, Rgg was found to bind directly to the upstream region of *spd3*, and directly represses the transcription of this phage-encoded nuclease (303) (167). Further studies need to be conducted in order to determine if Mga can directly bind to this gene and activate its transcription.

In addition, we also demonstrated that Mga regulates expression of *sdaD2* in a glucose-dependent manner (**Fig. 5.6A, 5.9B, and 5.10**). *sdaD2* is one of two horizontally transferred genes that distinguishes the invasive M1T1 serotype (304). This horizontally acquired streptodornase has been attributed to provide selective pressure for the switch to the hypervirulent *covS*$^{-}$ mutant in the M1T1 strain 5448 (*covS*$^{+}$) *in vivo* (85, 144). However, introduction of the *sdaD2*-encoding prophage into the closely related M1 strain, SF370, does not lead to increased virulence *in vivo*

(146). There has also been evidence in an alternative *covS*$^{+}$ M1T1 train, MGAS2221, that the Mga regulon, not Sda1, plays a much more important role in the *covS* switch *in vivo* (305). Taken together, we hypothesize that during invasive infection with *covS*$^{+}$ M1T1 strains, where a low glucose environment is encountered, there is an alternative activation of the Mga regulon, where there is an increase in *sdaD2* expression (along with other streptodornases) that plays a role in stimulating and favoring the switch to a *covS*$^{-}$ mutant.

Glucose availability affects Mga phosphorylation and subsequent activation of its regulon.

We observed a 4- to 5-fold repression of some core Mga regulon genes (*fba, scpA, sic,* etc.) in C Media (**Table 5.1**); however, a significant change in transcription of *emm* was not observed. Since *emm* is one of the most highly expressed genes in the GAS genome, the potential modest levels of reduction that occur in C media are above the limit of detection for differential gene expression calculations. There are two different possibilities that can contribute to the reduction in expression of the Mga regulon in C media when compared to THY: (1) CcpA has been shown to bind directly to a *cre* site located upstream of the P1 promoter region of *mga* (P*mga* P1) and stimulate transcription (17). The low-glucose environment in C media relieves CCR, as seen by the large overlap between the data generated from our RNA-Seq experiments and microarray generated from a Δ*ccpA* mutant. Therefore, under these conditions it is possible that CcpA is released from the *cre* site upstream of the P1 promoter region of *mga*, causing a decrease in *mga* transcript, and therefore causing a decrease in expression of its regulon under a low glucose condition. An alternate

possibility for this phenotype is that (2) Mga activity is modulated by the PTS of GAS through phosphorylation events on the PRD domains, which directly affects its activity and ability to regulate core regulon genes (19, 20). Furthermore, phosphorylation of Mga is hypothesized to be connected to availability of a preferred carbohydrate source. We believe that the latter scenario is the most probable, as the transcript levels of *mga* were not significantly different between THY and C media, indicating that the changes in the regulon exhibited in C media is most likely due to a posttranslational modification.

To test this hypothesis we looked for differential phosphorylation patterns of Mga using Phos-tag gel. Phosphorylation is hypothesized to affect the transcription of the regulon genes in one of three ways: 1) no phosphorylation produces an active Mga, 2) phosphorylation on only PRD2 can cause an increase in activity, and 3) phosphorylation on PRD1 inhibits Mga activity, and is dominant over PRD2 phosphorylation (19) (**Fig 5.12**). Therefore, we hypothesized that the low glucose environment encountered in C media would likely cause phosphorylation on PRD1, causing a reduction in Mga activity and regulon expression. In contrast, growth in THY (high glucose) would result in an unphosphorylated Mga and remain active. Hence this would help explain the activation of the regulon in THY over C media (**Table 5.1**). Furthermore, if this phosphorylation were a direct result of the PTS, we would predict that the phosphohistidines would be the target of modification(s), and using a phosphoablative construct would eliminate the phosphorylation phenotype regardless of media condition.

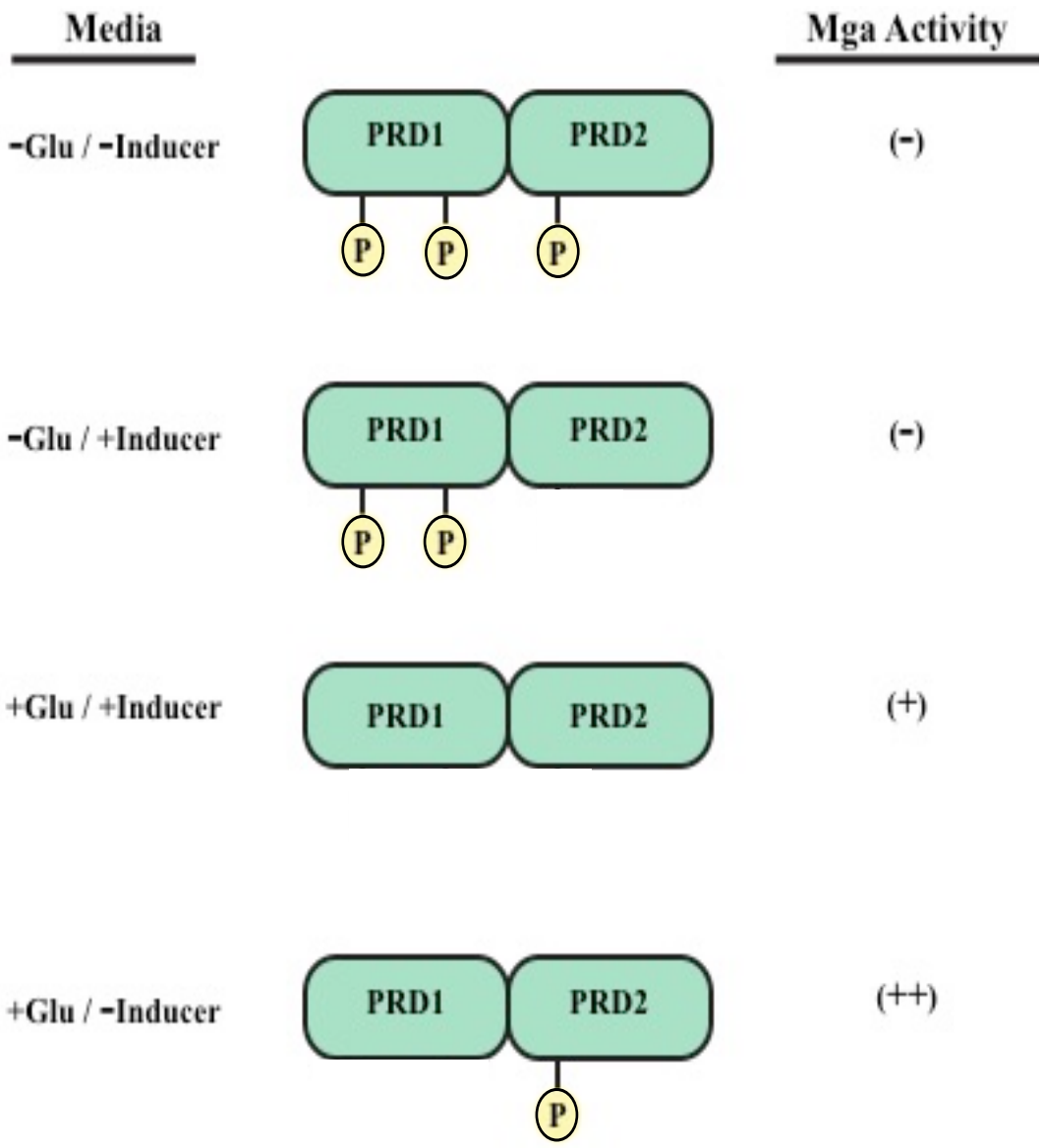


Figure 5.12. Graphical representation of the hypothesized role of PTS-mediated Mga phosphorylation on activity based on the presence of glucose.
In the absence of glucose, Mga would be inactivated through phosphorylation of PRD1, regardless of the presence of an inducer. However, when glucose and inducer are present, Mga is not phosphorylated at either PRD and is active. In the presence of glucose only phosphorylation of Mga PRD2 leads a more dramatic increase in activity. Adapted from (19).

Surprisingly, under the conditions tested in this study, when the WT M4 Mga was grown in THY, two bands were observed on a Phos-tag gel, corresponding to a phosphorylated and unphosphorylated species. In C media, only an unphosphorylated Mga was observed; the same pattern was seen for the phosphoablative (A/A/A) construct under both media conditions (**Fig 5.3**). Therefore, we now hypothesize that in C media, Mga still may be in its "active" form (condition 1), and THY represents a "hyper active" form of Mga (condition 2). This would be in line with the model that was previously proposed by Hondorp *et al*. (19), since glucose is present in both THY and C media. Therefore, it may not only be the level of glucose that is important for Mga activity, but the mere presence of it. Furthermore, since we observed only a nonphosphorylated band for the A/A/A construct in both C media and THY, this indicates that the phosphorylation is occurring at the phosphohistidines. However, this experiment does not verify which phosphohistidine(s) this is occurring at, although we hypothesize that it is located on PRD2 (H324). In order to confirm this, a Phos-tag gel would need to be performed on a Mga mutant that changed both histidines on PRD1 to alanines (A/A construct). If the banding pattern is identical for this strain when grown in THY as the WT grown in C media, it would implicate that the phosphate on PRD2 is responsible for the banding pattern observed in this study. Taken together, this is the first study that identifies that glucose is the sugar that dictates the activity of Mga, and the phosphorylation by the PTS occurs at the phosphohistidines on the PRD domains of Mga.

As discussed above, Mga has very different regulons in high versus low glucose growth environments. If the mechanism by which Mga is differentially

regulating genes is dependent on the level of glucose, we hypothesized that C media supplemented to 0.5% glucose should show a reversion to the phenotypes observed in THY. Indeed, out of the 6 genes that were specific to the C media regulon, 5 that were tested via qPCR showed a glucose specific phenotype (**Fig. 5.10**). However, we also saw differences in the regulon (**Fig. 5.10**; *opuAA*) that are most likely due to broader attributes of C media itself. In order to further confirm that glucose is the inducer of the effect observed here, it would be important to conduct a Phos-tag gel comparing the WT Mga in THY, C media, and C media-0.5% glucose. Based on our data, we would expect the lack of phosphorylation on the WT that was observed in C media would now have a phosphate(s)-specific band, and mirror what was seen in THY.

Taken together, this work shows that glucose directly affects Mga activity and the composition of the Mga regulon. Furthermore, the differences between the regulons at varying glucose levels is most likely attributed to a phosphorylation event that occurs on the phosphohistidines located on the PRD domains of Mga. While it is unclear which phosphohistdine(s) are affected and the mechanism by which Mga is able to regulate genes alternatively, it is clear that in the presence of glucose, the PTS plays an important role in activating this global transcriptional regulator and virulence genes downstream.

**Author contributions: K.M.V. performed the research for all the sections. A.T.B. aided in RNA-Seq data analysis.

Chapter 6: Identification of critical amino acids necessary for Mga dimerization and transcriptional activation

Introduction

The group A streptococcus (GAS) relies on the timely expression of a wide range of virulence factors in order to successfully evade the host immune system and cause infection (3). Expression of virulence determinants are under tight control by multiple regulatory networks, which include two-component signal transduction systems (TCS) and "stand-alone" regulators, in order to produce adaptive responses to the host environment (10). The multiple gene activator of the group A streptococcus, or Mga, is a highly conserved stand-alone regulator that is ubiquitous in GAS (219, 220). It activates the transcription of a variety of virulence factors important for immune evasion and adhesion, including the M and M-like proteins, during exponential phase growth. Expression of the Mga regulon is also dependent on different environmental signals such as increased CO_2 levels, body temperature, and exponential-phase growth (238, 239). Mga activity is also dictated by the sugar utilization status of the cell via its PEP-phosphotransferase system (PTS) regulation domains (PRDs), demonstrating that virulence factor production is intimately linked to carbohydrate utilization in GAS. Thus, Mga has been categorized as a PRD-containing virulence regulator, PCVR, which is part of a broad family of Gram-positive virulence regulators whose activity is dictated by carbohydrates (19).

GAS strains have been divided into two classes; Class I strains contain the *mga-1* allele (M1 serotype), and Class II strains posses the *mga-2* allele (M4

serotype) (306). The two different *mga* alleles share approximately 88% sequence identity and can complement each other across serotypes, indicating functional similarity (307). Mga is a 530 amino acid protein with a molecular weight of 62 kDa comprised of six domains. The amino-terminus contains a three domains (CMD, HTH-3, wHTH-4) that are all important for DNA binding (235, 308). Mga also contains PRD-1 and PRD-2, which directly phosphorylated by the PTS to modulate functional activity (19). The carboxy-terminus was originally suggested to have structural similarity to a CheY-like receiver domain (223). However, *in silico* (282) analysis later determined that this region shared close homology to an $EIIB^{Gat}$-like domain (**Fig 6.1**). (251). Interestingly, while Mga is highly homologous across all GAS serotypes, the very extreme C-terminus is not conserved between GAS strains (309).

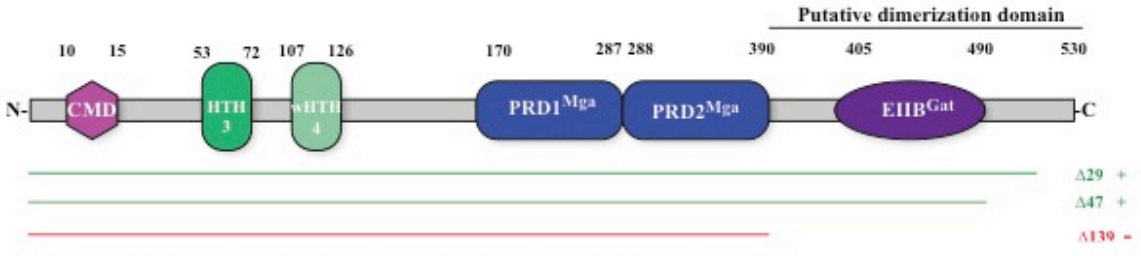


Figure 6.1. Mga EIIB domain is necessary for full activation.
Schematic representation of Mga, shown are the N-terminal regions implicated in DNA binding (CMD-1, HTH-3 and HTH-4) and the central two domains have structural homology to PTS regulatory domains (PRD-1 and PRD-2). The C-terminal region with structural homology to an EIIB domain is depicted. Lines representing the different C-terminal truncation mutants and their effect on Mga activity; only the Δ139 construct showed a defect in Mga activation. Adapted from (310).

Hondorp *et al*. utilized different truncations in the C-terminus of a M4 Mga to establish its effect on Mga activity. Two mutants that deleted either the last 29 or 47 amino acid residues, removing the non-conserved region of the protein, showed no functional difference in activation of *arp* and *sof*, two genes that are a part of the Class II Mga regulon. However, a mutation that deleted the last 139 amino acids (Δ139-Mga4), which eliminates the putative $EIIB^{Gat}$-like domain, led to a significant decrease in Mga activity and also loss of oligomerization. This truncation did not affect the ability for Mga to bind DNA. Importantly, this phenotype was recapitulated in an M1 serotype as well (237). Therefore, while Mga needs its helix-turn-helix DNA binding domains in order to bind to its core regulon genes, this activity is not sufficient for full Mga activity, and oligomerization is critical for full activation of the Mga regulon *in vivo* across all serotypes.

The *E. faecalis* protein MafR, a Mga homolog with unknown functional activity, has a readily available crystal structure deposited in the Protein Databank (PDB 3SQN) (217); however, a crystal structure for Mga is not available. Therefore, using the information concurrently available from both Hondorp *et al*. and the MafR protein structure, a modified bacterial two-hybrid screen (311) was undertaken using an M1 Mga in conjunction with low-frequency random mutagenesis of the carboxy-terminal region, which aimed at identifying amino acid residues that were critical for Mga oligomerization and activity (310). Subsequent to the two-hybrid study, a crystal structure for the established Mga homolog and PCVR AtxA from *B. anthracis* was published (PDB 4R6I) (216). In this study, we used the results generated from the two-hybrid screen and the availability of the AtxA crystal structure to further

characterized the amino acid residues that were shown to abrogate Mga dimerization and functional activity *in vitro*.

Results

Expression of the Mga1 EIIB-mutants shown to abrogate Mga oligomerization in *E. coli*.

A bacterial two-hybrid screen revealed 20 different mutations that showed a loss of interaction with the wild type Mga. Out of 20, 17 contained single amino acid substitutions, and 3 contained a frameshift insertion. Given they would likely not produce a full-length Mga, the 3 frame-shift mutants were disregarded in our analysis. We also identified two amino acid residues that were found more than once in the screen (T458 and M461). Interestingly, the mutations that were identified appeared to cluster within two distinct regions in the EIIB domain (**Fig 6.2; red triangles**). We also identified single amino acid substitutions that generated a positive interaction phenotype (**Fig 6.2; green triangles**), which were subsequently used as a positive controls for dimerization.

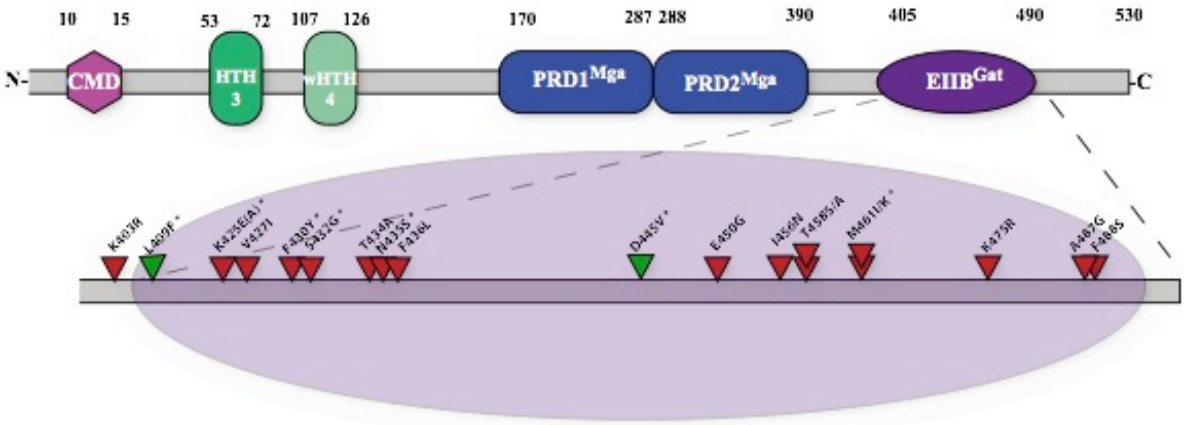


Figure 6.2. Localization of mutants identified in a bacterial two-hybrid screen. A random mutagenesis approach was used to generate a low-frequency mutation rate in the EIIB region of Mga. These mutants were then screened in a two-hybrid system, and mutants showing lack of interaction (white colonies) were mapped. A total of 17 individual amino acid substitutions (red triangles) were identified, mutations that generated positive interactions were also found (green triangles). Residues that were targeted for site-directed mutagenesis are indicated with a *. (*Performed by A.K. and K.M.V.)

One explanation for a negative interaction phenotype would be the inability to produce Mga; therefore, mutants were verified for production of a functional protein. Attempts were made to express the proteins in their original strains used in the 2-hybrid screen (contains both bait and prey plasmids). However, due to the similarity in size between the two adenylate cyclase fragments (T25N, ~88kDa; T18N; ~84kDa), we could not resolve the two bands by SDS-PAGE for comparison by western blot. We next attempted to resolve the two plasmids from the original background by using a limited dilution approach. Plasmids (pT25N-xxxx) were extracted from each 2-hybrid strain (**Table 3.1**), serially diluted, and transformed into

the *E. coli cya*$^{-}$ strain BTH101 (D. Kearns, gift) or C43 [DE3] (248) with appropriate antibiotic selection. Strains were grown, induced and whole cell lysates were then analyzed for Mga expression by western blot. Despite several attempts, we were unable to observe expression of full-length Mga for any of the original mutants generated in the EIIB screen or positive controls in either background (data not shown).

Expression of select Mga mutants in GAS.

Due to the expression issues encountered with the 2-hybrid constructs in *E. coli*, we generated a representative group of mutants that were found in the screen via site directed mutagenesis (SDM) in a native M1 *mga* background. A total of 7 different EIIB mutants were constructed (K425A, F430Y, S432G, N435S, and M461I), two of which were dimerization controls (L409F and D445V), generating plasmids pKSM935-941 (**Table 3.1, and Fig. 6.2**). The EIIB mutant plasmids were expressed from their native promoter (P*mga1*), and transformed in a M1 Δ*mga* background strain, KSM165L. Using this approach, we were able to express 5 out of the 7 EIIB mutants, including the two dimerization controls (K425A, F430Y, S432G, L409F and D445V). Interestingly, one of our dimerization controls, D445V, runs at a slightly smaller protein, however a similar phenotype was observed before, when other mutations in Mga were generated (19). Importantly, each of these mutants expressed at the same level as both the wild type Mga1 and the Δ139-Mga1 (**Fig 6.3**).

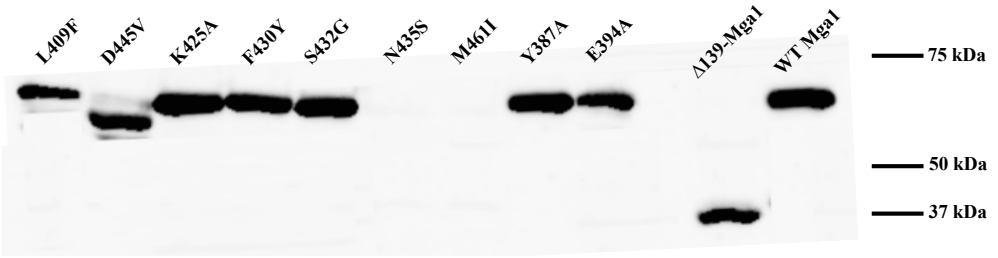


Figure 6.3. Expression of site-directed EIIB mutants of M1 Mga in GAS.
Western blot of whole cell lysates of the EIIB and PRD mutants compared to WT Mga1 and Δ139-Mga1 expressing using α-Mga1. Predicted size for Δ139-Mga1 is 47kDa, but runs aberrant at 35kDa (251).

Mapping of Mga4 onto the crystal structure of AtxA.

The *B. anthracis* master regulator, AtxA, controls the expression of a variety of virulence genes including toxin and capsule production (213). AtxA is a functional homolog of Mga and a member of the PCVR family of transcriptional regulators. The crystal structure of AtxA was recently resolved by Hammerstrom et al. to 2.4-3.1 angstrom (Å) resolution (PDB accession number 4R6I) (216). Contrary to what was observed with *E. faecalis* MafR, where both EIIB domains on each monomer act as the dimer interface, the AtxA dimer utilizes the EIIB domain of one monomer and the PRD2 domain of the second monomer for dimerization.

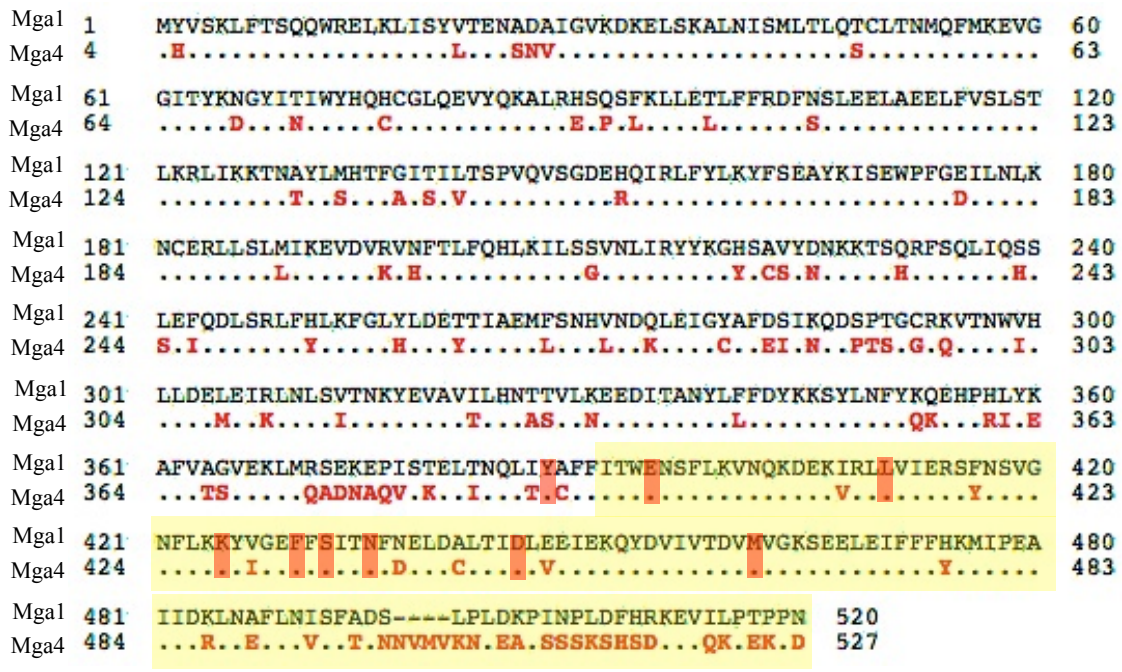


Figure 6.4. Alignment of Mga1 and Mga4.
Mga homologues are found in all serotypes of GAS and are highly conserved, however the C-terminus of the protein seems to exhibit a high degree or variability. A pBLAST alignment of the Mga1 and Mg4 protein demonstrates differences between these two strains (red letters), with conserved residues indicated as a dot. Highlighted in yellow is the EIIB domain starting at aa 405. Mutants identified are highlighted in red.

In collaboration with Dr. Osnat Herzberg (IBBR), we were able to thread an M4 Mga amino acid sequence onto this newly available crystal structure. The programs, Raptor (312), I-tasser (313), and Robetta (314) were each used to determine the best structure of Mga4. As stated earlier, Mga1 and Mga4 posses 88% sequence identity (307) (**Fig. 6.4**), and M4 Mga was initially chosen for comparison since we possess a large amount of functional data for M4 (19, 251). Both Raptor and

I-tasser computed the model with a high degree of probability, and both shared a significant similarities between the outputs, and were therefore chosen for further analysis (**Figure 6.5**).

Interestingly, while the N-terminus of the protein mapped perfectly to the AtxA crystal structure, the segment of the protein immediately after PRD2 (aa 391-530) was very hard to map and remains relatively unstructured. This is attributed to a large insertion that Mga possesses between the last helix of PRD2 and the EIIB domain. This indicates that Mga has a unique tertiary structure in the C-terminal of the protein compared to AtxA (**Fig. 6.4;** labeled EIIB).

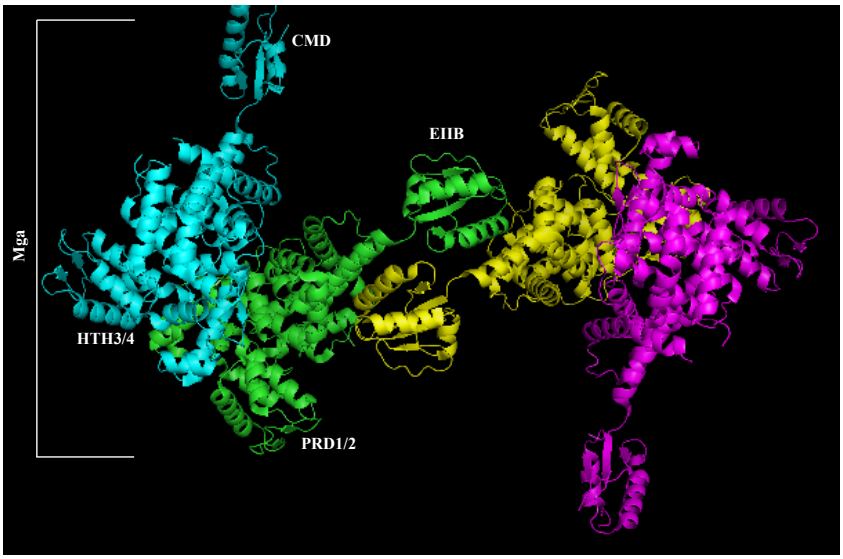


Figure 6.5. AtxA-based structure of C-terminal Mga interactions.
Shown is a modular 3-dimensional structure of the Mga dimer as deduced from the crystal structure of AtxA. Mga was threaded onto the AtxA crystal structure using the Raptor and I-tasser protein modeling programs. Two Mga monomers are shown in dimer conformation. The N-terminals are depicted in cyan (dimer 1) or pink (dimer 2) and the PRD and EIIB domains are shown in green (dimer 1) or yellow (dimer 2). *Performed by O.H.

Identification of PRD2 interaction residues.

Since the original region targeted in the bacterial two-hybrid screen was based on the MafR crystal structure, which shows the EIIB domain as the sole interaction region for dimerization, PRD2 was never interrogated. However, given the new information provided by the AtxA crystal structure, it was possible that both the PRD2 and the EIIB domains of Mga are critical for dimerization. Therefore, the predicted Mga4 crystal structure (using both Raptor and I-Tasser) was used to identify several residues in PRD2 that were putative interaction partners with the EIIB domain. Raptor identified residues I381, Y387, or E394, whereas I-tasser identified: Y387, I391, F397, E394 as putative residues critical for dimerization. Since they were found in both models, we chose to modify both Y387 and E394 via alanine substitutions (**Table 3.1;** pKSM942 and pKSM943). In addition, both residues putatively interact with residue K425 on the EIIB domain (**Figure 6.6**), a residue that was identified in the two-hybrid screen, and expressed in western blot assays (**Fig 6.3**). After creation of the site-directed mutants, Y387A and E394A, we tested their expression in the M1 Δ*mga* GAS background KSM165L using western blot with α-Mga1 antibodies; both showed expression comparable to WT Mga1 levels (**Fig 6.3**).

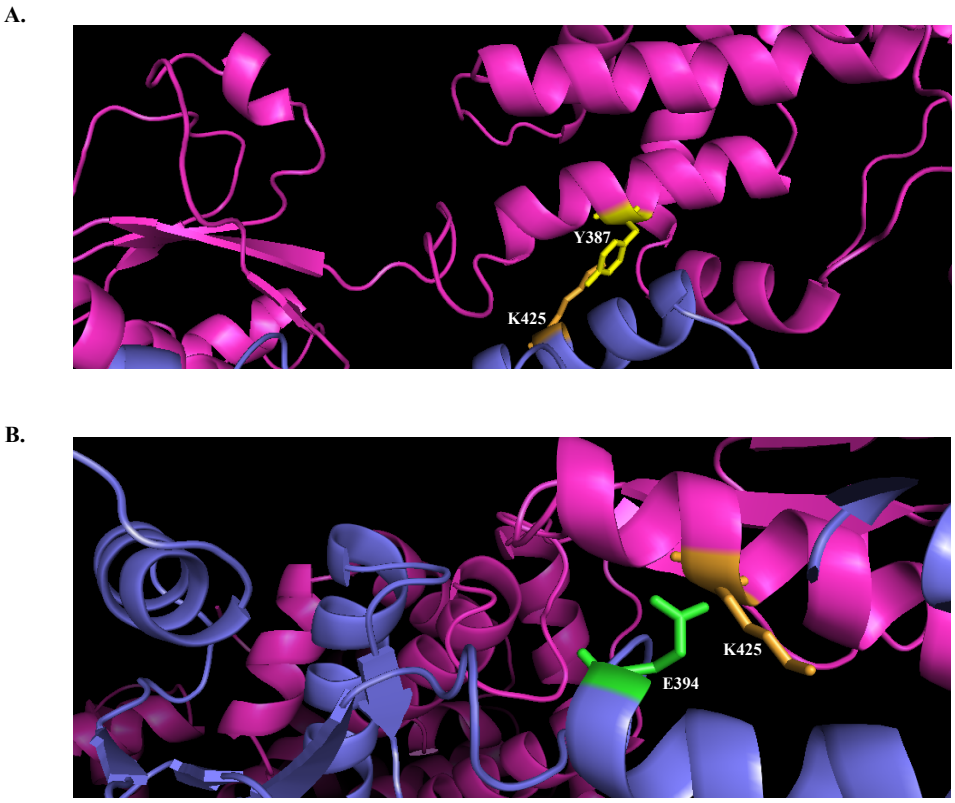


Figure 6.6. Interactions of Mga-PRD2 with residue K425 during dimerization. Two PRD2 residues, Y387 (yellow; **A**) and E394 (green; **B**) show putative interactions with residue K425 (orange) located on the EIIB domain. *Performed by O.H. and K.M.V.

Functional activity of Mga EIIB mutants.

To determine the *in vivo* effects of the EIIB mutants on Mga activity, RNA was isolated from strains grown in THY to late exponential phase. Mga activity was then assayed by real-time qPCR using the Mga-regulated virulence gene encoding for the M protein (*emm*). Changes in transcript levels for the EIIB mutant strains that were greater than twofold compared to the strain expressing full-length wild type

Mga1 were considered significant (**Figure 6.7**, dotted line). The F430Y, N435S, and M461I showed a significant decrease in transcript levels for *emm* compared to the wild type strain. Interestingly, one of our dimer interface positive controls, D445V, showed a significant decrease in transcript levels of *emm* as well.

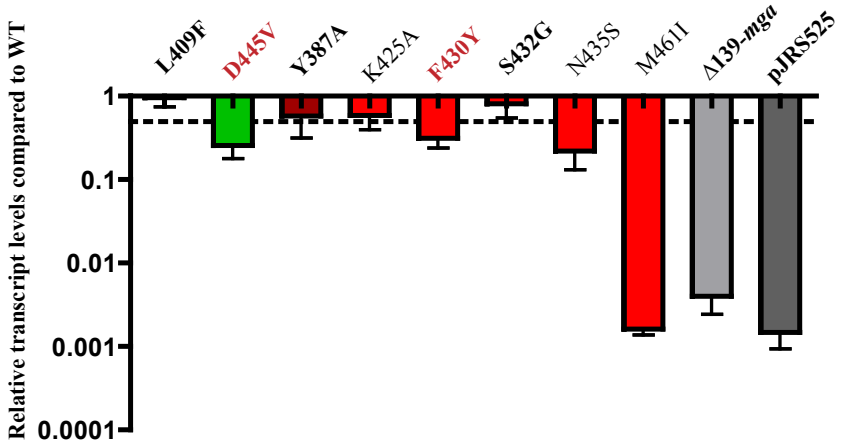


Figure 6.7. *In vivo* activity of putative residues critical for dimerization.
Mga activity was assayed *in vivo* by real-time qPCR analysis of the Mga-regulated gene *emm*. Positive dimerization controls are shown in green, the PRD2 mutant is shown in dark red, and the EIIB mutants are shown in red. The Δ139-*mga1* (light grey) and empty vector Δ*mga* (dark grey; pJRS525) results are shown alongside for comparison. Relative transcript levels for each strain compared to the wild type are presented. Error bars represent the standard error from at least three biological replicates. Differences greater than two-fold in expression were considered significant (dashed line). Mutants bolded and red are those that were expressed and showed a functional defect in *emm* activation, and only bolded indicates those mutants that are expressed in a western blot using α-Mga1.

Although M461I and N435S showed a decrease in transcript levels of *emm*, with M461I showing a decrease comparable to both the Δ139 and the empty vector strain, this is most likely due to the lack of expression of a full-sized protein (**Fig. 6.3**). Unfortunately, the mutation in residue K425, which was predicted based on the mapped crystal structure to be critical for interaction with PRD2, showed no significant effect on Mga activity. Therefore, out of the 9 site-directed mutants that were constructed, 6 showed expression in GAS, and only two showed a significant decrease in Mga transcriptional activation, albeit not to the levels of either the Δ139-*mga1* or a Δ*mga* (pJRS525) background.

Discussion

While it has recently been established by our laboratory that the C-terminus plays an important role in both Class I and Class II Mga dimerization and activation (251), the exact residues that are important for this interaction have yet to be established. This study is the first to look at individual residues within the C-terminus of M1 Mga in order to characterize their functional importance. Furthermore, while Mga has been modeled to the AtxA crystal structure previously using I-Tasser (20), this is the first time that Mga was modeled using multiple structure prediction programs in order to help predict residues involved in the interactions between the EIIB domain that are important for dimerization.

General mutability and protein stability of Mga's C-terminus.

Interestingly, while there were 17 mutations that were found to abrogate Mga1 dimerization using the two-hybrid screen, we experienced difficulty in obtaining full-length expression of these Mga mutants in *E. coli*. We attribute this to the potential stability of the adenylate cyclase fragment (T25N) when fused to Mga, since we observed degradation products but never a full-length (~84 kDa) protein (data not shown). The two adenylate cyclase fragments, T25 and T18, may aid in stabilizing protein expression in order to have a functional two-hybrid system. However, when one fragment is removed, induction and expression of a single fragment fused to the protein of interest may result in a protein that is susceptible to degradation.

To overcome this issue, we generated 7 representative site-directed mutants for expression in GAS. Interestingly, while this eliminated the degradation issue we

were experiencing in *E. coli*, it did not resolve the issue of protein expression for some of the mutants. We were only able to express a full-length Mga in 5 out of the 7 constructed mutants; N435S and M461I were never observed by western blot of GAS whole cell lysates. In addition, when expressing PRD mutations predicted as interaction partners with K425, both Y387A and E394A were expressed. In 2006, Vahling *et al*. conducted a screen using a P*emm-gusA* reporter that aimed to identify domains critical for Mga activity. While the study identified 272 C-terminal mutants that showed little to no functional Mga activity, western blot analysis revealed that none produced wild type levels of protein (308). Taken together with our results, this suggests that the C-terminus of Mga is extremely sensitive to mutagenesis, and the majority of mutations result in the inability to produce a full-length protein in GAS. These results support the importance of the C-terminus and EIIB domain in Mga structure and function, but further study of this region will be very difficult and another method of identifying amino acid residues critical for Mga dimerization may be required. This could be potentially accomplished using chemical cross-linking assays, as was done originally for the Mga homolog, AtxA, prior to the availably of its crystal structure (215).

Amino acids important for full activity of Mga.

(i) F430 and D445 play a role in Mga1 dimerization and activation

Out of the original 7 residues that were interrogated by mutagenesis, only two (F430 and D445) were able to express full-length Mga at levels comparable to wild type (**Fig. 6.3**) and also resulted in a significant decrease in transcriptional activation

of the Mga-regulated gene *emm* (**Fig. 6.7,** bold and red). Based on the modeled crystal structure of the M4 Mga, it appears that both F430 and D445 face away from the dimer interface (**Fig. 6.8AB**). However, the exact positioning and interaction that these key residues in Mga may be altered in the AtxA-modeled Mga. As mentioned above, the threading of the C-terminus of Mga was extremely difficult due to a large insertion between the last helix of PRD2 and the EIIB domain in Mga that is not present in AtxA. This insertion may alter the exact positions of both F430 and D445 in Mga, and these residues may actually be closer to the dimer interface than predicted by the model.

Surprisingly, one of the original amino acid substitutions that was chosen as a control that would not effect dimerization based on its 2-hybrid phenotype, actually showed a decrease in transcriptional activation of *emm*. We attribute this finding to two possibilities: (1) it was a false positive in the original bacterial two-hybrid screen. Since none of the residues found were screened in the opposite direction, where EIIB mutants were cloned into the T18 containing vector and screened against the wild type T25-Mga1, it is possible that this mutation could have been missed as an actual negative interaction. (2) It has no affect on Mga dimerization, but does affect Mga activity. We believe that the latter is most likely outcome since it showed a positive interaction in on iteration of the two-hybrid screen. Regardless, it is clear that a mutation in either F430 or D445 results in a decrease in Mga activity, and potentially disrupt dimerization. Of interest, neither F430 nor D445 showed a decrease in activity to the levels of either the Δ139-*mga* or Δ*mga* strains. Therefore, a single mutation may not be sufficient to disrupt dimer formation, as several amino acids are usually

implicated in these types of interactions. In order to further confirm these two mutants for defects in dimerization, either a native gel or gel-filtration assays need to be preformed.

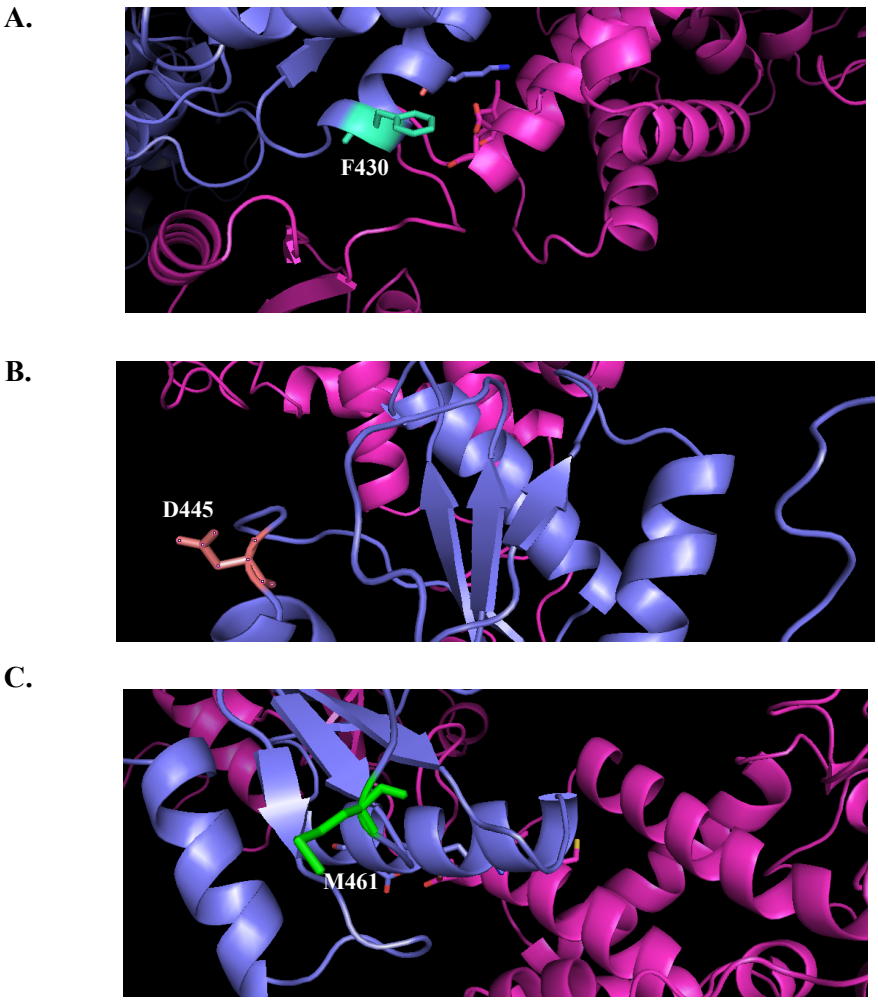


Figure 6.8. Localization of residues important for Mga dimerization and activity. Shown are F430 (cyan, A), D445 (orange, B) and M461 (green, C) on one monomer of Mga (blue) in reference to the second monomer (pink). *Performed by O.H. and K.M.V.

(ii) M461 plays an important role in the conformation of Mga.

A mutation in residue M461 was identified twice (M461I and M461K) in the two-hybrid screen as potentially abrogating Mga-Mga interaction (**Fig. 6.2**). Further analysis of this residue showed that it exhibited over a 100-fold decrease in transcriptional activity (**Fig. 6.7**). However, when this residue was tested for expression of Mga using Western blot analysis (**Fig. 6.3**), it was not shown to express a full-length protein. Upon alignment of primary amino acid sequences for Mga from multiple GAS strains, it was noted that the methionine residue at position 461 was conserved among all Mga proteins.

A previous study compared two Class II *mga* alleles found in M50 and M4 GAS that are nearly identical at the amino acid level, yet the M50 strain is defective in activating downstream Mga-regulated genes (309, 315, 316). Vahling *et al*. found only three non-conserved amino acid between these two *mga* alleles. Interestingly, one of the residues containing an amino acid substitution was residue 461, where the methionine was changed to an arginine in the M50 strain (316). This study also showed that the levels of Mga produced by the wild type M50 strain was not equivalent to either the M4 Mga or a M50 Mga R461M mutant. In addition, a reporter assay using the promoter of the Mga regulated gene *mrp* (P*mrp-gusA)* showed the M50 strain had a 52% decrease in transcriptional activation in comparison to M4 and to an M50 strain that was restored the amino acid residue to a methionine (R461M) (316). Of particular interest, this study also tested the importance of this residue in the Class I M6 strain JRS4, and no significant difference was seen either protein expression levels or transcriptional activation (316). Interestingly, it is now known

that JRS4 is a lab-adapted strain that exhibits an altered Mga regulon from other GAS strains (229). This could therefore attribute the differences seen from this study.

Taken together, residue M461 is of great importance to the stability of Mga as demonstrated both here and in other studies (315, 316). This is can be attributed to position 461 being essential for correct protein conformation, and any amino acid substitution could lead to a protein that is sensitive to degradation. This also points to the importance of the correct conformation of the EIIB domain for overall protein stability, since very few mutations generated a fully functional protein in several studies. Therefore, it is possible that Mga may need to exist in a dimer within the cell in order to be fully functional and not be susceptible to degradation. In order to test this hypothesis, further work is needed. In vivo dimerization assays, such as a native gel can be used to answer this question.

While, it is yet unclear the physiological implications these amino acid mutations studies here cause on Mga dimerization, it is clear that several affect Mga activity. Overall, this study has elucidated two potential residues, F430 and D445, located in the EIIB domain that have implication on Mga activation, and thus could have effects in dimer formation. In addition, a mutation in residue M461 has been proven in multiple studies to cause a general destabilization of Mga, and thus generates a complete loss of protein expression and down-stream gene activation.

**Author contributions: A.M.K. performed the two-hybrid assay. A.M.K. and K.M.V. performed the research on expression of the EIIB mutants in *E. coli*. O.H. performed threading of Mga crystal structure. K.M.V performed the rest of the experiments described here.

Chapter 7: Conclusions and Recommendations

The Group A Streptococcus (GAS) is a strict human pathogen that is capable of causing a wide range of diseases in diverse niches in the host. In order to adapt to these locations, GAS utilizes environmental cues, such as carbohydrates, to act as a signal to coordinate the expression of virulence factors. Research efforts to date have focused on identifying how either components of the phosphoenolpyruvate-phosphotransferase system (PTS) or global transcriptional networks affect the regulation of virulence factors, but not the synergistic relationship between the two. In this dissertation, the role of putative PTS-fructose sugar transport operon in GAS pathogenesis and sugar utilization was investigated. In addition, while it is known that the PTS can affect activity of the Mga virulence regulator, further characterization of the mechanism by which sugars and its protein domains affect its activity have not been studied. Here, two aspects of Mga activity were investigated: (1) how glucose availability affects Mga activity and changes the dynamics of its regulon, and (2) how amino acid interactions aid in the formation of Mga homodimers, a target of PTS regulation, and activate transcription of regulon genes. Overall, this dissertation sought to shed light on how external environmental cues, specifically carbohydrates, control complex regulatory networks used by GAS to adapt to various nutrient conditions encountered and contribute to pathogenesis.

Fructose metabolism by the M1T1 GAS

The PTS is the main system that is utilized by bacteria for the uptake of sugar and sugar derivatives as well as for signal transduction (180). The GAS genome encodes for 14 putative sugar-specific EII proteins (197), whereas EI and Hpr are common for metabolism of all the carbohydrates imported in the cell. In this dissertation, the *fruRBA* operon, its role in fructose metabolism and utilization in the physiology of the M1T1 GAS was characterized. The clinical isolate, 5448 was utilized as a representative of the M1T1 serotype, which is the most common serotype associated with invasive infection worldwide.

The global transcriptional changes of wild type 5448 grown in chemically defined media supplemented with either glucose or fructose showed that the *fruRBA* operon was the most highly induced followed by several other sugar-specific PTS operons. There are three potential mechanisms by which fructose could enter GAS: (i) the predominant pathway in Gram-positive bacteria involves the canonical PTS^{fru} EIIABC (*fruA/fruI*), a 1-phosphofructokinase (*fruB/fruK*), and a fructose-responsive regulator (*fruR*); (ii) an alternate PTS that recognizes and translocates fructose as fructose-6-phosphate (F6P) with lower affinities than (i), and; (iii) facilitated diffusion transport using a membrane spanning permease that would esterify free fructose at position-6 using ATP (287). The induction of *fruRBA* by fructose in GAS suggests (i) is the main pathway; however, induction of other PTS systems (e.g., *lacD.1, lacD.2*) suggests that mechanism (ii) may also play a role in fructose utilization. Thus this raises the question of the role other PTS operons play in the utilization of fructose. Currently, studies are being undertaken in the lab to

characterize of all 14 EII sugar-specific transporters and their specificity for transporting sugars. Interestingly, current data indicates that several PTS EII's are redundant in their ability to transport a single sugar. However, the reason behind this redundancy is not understood and will need to be explored further. Furthermore, questions remain in regards to the EII transporters, such as, 1) how the does the chemical structure of the sugar affect its ability to be transported by certain EIIs? 2) Does the specificity for the sugar(s) the EII transports come from the structure of the protein itself or another mechanism?

sloR and its regulon were also activated in fructose. SloR was initially implicated in streptolysin O expression (277), but more recently was shown to be involved in serine acquisition (276). This gene shows significant homology to a fructose EIIC transporter based on *in silico* analysis (282), thus suggesting that it might play a role in the transport of fructose (ii, intake as F6P) and potentially aid in signal transduction to coordinate gene expression. However, a *sloR* mutant did not show any significant effect on utilization of fructose or any other PTS sugar by GAS. In other *Streptococcus* and *Listeria* species, orphan EIICs have been suggested to act as environmental sensors that interact with other regulatory proteins to influence gene expression (317, 318). It is possible that SloR serves a similar sensory role in GAS, and this hypothesis still needs to be investigated.

Prior to this study, functional characterization of only one of the putative PTS EIIs encoded in the GAS genome has been carried out so far. This resulted in a discrepancy between the predicted function of the EII and the experimentally validated role of the transporter (197). Therefore, homology and subsequent

annotation does not necessarily follow functionality. *fruA* (encoded in the *fruRBA* operon) encodes one of the 13 other uncharacterized EII PTS sugar-transporters found in the M1T1 GAS genome. Interestingly, deletion of *fruA* showed a defect in the utilization of 3 other PTS sugars (N-Acetyl-D-Galactosamine, Cellobiose, and Mannitol) using two carbon utilization assays. Although a *fruA* mutant could utilize fructose, it exhibited a significant growth defect when this sugar was the sole carbon source available. Residual growth was observed for a *fruA* mutant, and this can be potentially attributed to the activation of other EII PTS transporters that were observed in the RNA-Seq data, however the exact role of these alternative fructose transporters need to be confirmed. In addition, questions still remain on why FruA affects the utilization of the three alternative PTS sugars, and how the metabolism of those sugars feed into gluconeogenesis through FruA. In addition, a definitive experiment confirming the transport of fructose into the cell is necessary to conclusively state that FruA is the main transporter of fructose.

FruR was demonstrated to be the fructose repressor through transcriptional assays, and is necessary for growth in fructose. However, demonstration that FruR can directly bind to its recognized sequence has not been shown and will need to be explored further. The physiological inducer of FruR de-repression is fructose-1-phosphate (F1P) and only 1 mM of F1P is sufficient to disrupt FruR binding to its operator in *P. putida* (284, 285). Fructose-1-phosphate is also the inducer of the *fruRBA* operon in other Gram-positive species, such as *L. lactis* (275). This correlates with FruR being a member of the DeoR-family of repressors, which are typically induced by sugar-phosphates produced by the regulatory pathway they control (180,

272, 275). However, the role of F1P in derepression of the fructose operon in GAS has yet to be explored. The question still remains if FruR is able to bind F1P and subsequently relieve repression of the *fruRBA* operon.

FruB encodes for 1-phosphofructokinase, the enzyme that converts F1P to FBP. Surprisingly, a *fruB* mutant exhibited a significant growth defect in fructose. This is likely due to the fact that the *fruB* mutant (5448.Δ*fruB*) would be expected to build up F1P due to an inability to convert it to FBP following transport by FruA. Since the majority of intracellular fructose cannot be further metabolized, gluconeogenesis of FBP for glycolysis would be blocked thus resulting in growth arrest.

Results presented in this dissertation demonstrates that not only FruA, but also FruB (1-phosphofructokinase) and FruR (DeoR-family repressor) are necessary for metabolism and growth on fructose, and the *fruRBA* operon is the primary transporter of fructose in M1T1 GAS. This is further supported given that *fruRBA* is the only operon that encodes a 1-phosphofructokinase (M5005_Spy_0061; *fruB*) in the 5448 genome (i, above) (319). Additionally, this operon has been confirmed as the primary transporter of fructose in several other Gram-positive bacterial species (272, 275, 278, 283). However, several questions are left unanswered about the complex metabolism and transport of sugars in GAS. In particular, studies will be necessary to determine the role that other EIIs play in fructose metabolism and GAS physiology.

fruR and fruB contribute to GAS pathogenesis in a human-specific manner.

Several PTS EIIs have been implicated in the pathogenesis of low-G+C Gram-positive bacteria, including the fructose-specific PTS system in other pathogenic species of Streptococci (272, 273). The putative fructose-specific EII (*fruA*) was originally identified as important for GAS fitness during growth in whole human blood in a TraSH screen performed in our lab (15), however this study showed that a *fruA* mutant did not exhibit a virulence defect. However, in the current study, non-polar defined GAS mutants in *fruR* and *fruB*, but not *fruA*, exhibited a significant survival defect in whole human blood as well as within neutrophil-like HL60 cells, human neutrophils, and peripheral blood monocytic cells or peripheral blood monocytic cells (PBMCs). One possible explanation for the lack of a *fruA* phenotype here is that TraSH screens are based on large libraries of mutants that must compete to survive. Therefore, *fruA* may only show a defect in blood survival and killing by phagocytes in a competitive environment, and a competition assay in whole human blood or PBMCs may elucidate this further. Furthermore, *fruR* and *fruB* may have not been identified in the original TraSH screen because it relied on a GAS oligonucleotide microarray that was found to miss almost 40% of the genome due to technical limitations of hybridization and array probes (15). Regardless, because *fruA* was not found to be important for GAS survival in human blood or neutrophils, this suggests that the blood phenotype is not linked to fructose uptake and subsequent phosphorylation.

The virulence defect exhibited in either whole human blood or neutrophils by a *fruB* mutant is hypothesized to be attributed to its enzymatic activity (F1P conversion to FBP). This is not the first study in GAS demonstrating PTS-related enzymes having an effect on virulence outcomes. The GAS tagatose-1,6-bisphosphate aldolase (LacD.1) enzyme has been adapted to play a regulatory role in controlling expression of the SpeB exotoxin through interacting with the transcriptional regulator RopB (4). LacD.1 still functions as an aldolase, although it is not the preferred aldolase used in galactose metabolism in GAS (297). Therefore, a metabolically functional enzyme can play a regulatory role via indirect interaction with regulatory proteins. However, the question remains as to the mechanism by which FruB is impacting immune evasion. Thus the following questions remain: Does FruB directly interact with a transcriptional regulator? If the virulence defect is attributed to a build up of F1P, and therefore a direct result of a lack of enzymatic activity through FruB, how does this act as a signal for the cell? In turn, this raises the question if FBP acts as more than an intermediate for gluconeogenesis and can actually be a regulatory signal for GAS and affect pathogenesis? Making a mutation in the catalytic domain of FruB, which presumably will cause a build up of F1P, and observing the effect of this mutation on GAS pathogenesis can help answer these questions. Additionally, if it is not the build up of F1P but an interaction of FruB with another regulator, this catalytically inactive mutant may be able to complement the virulence defect observed in a Δ*fruB* strain.

FruR also exhibits decreased survival in either whole human blood or neutrophils. This phenotype is most likely indirect and occurring through regulation

of *fruB*, which suggests that there is a stoichiometry effect during the buildup of 1-phosphofructokinase in the cell that leads to a defect in immune evasion. It is unlikely that FruR is regulating other virulence factors as a canonical FruR binding site (275) only occurs once in the genome, upstream of FruR. However, this cannot currently be ruled out and needs to be investigated further as GAS may have an alternative consensus for FruR binding. In addition, although RNA-Seq would provide information with regards to whether FruR is capable of regulating virulence genes, it would be difficult to perform this experiment due to the growth phenotype FruR exhibits in fructose. As an alternative, RNA-seq could be performed in glucose as an alternative, however if fructose is the primary signal for this gene and the *fru* operon, it is possible that a phenotype would not be observed under these conditions.

Although the *fruB* and *fruR* phenotypes observed in this study are most likely not due to a fructose-specific signal, the RNA-Seq results of the WT 5448 in fructose did reveal three virulence related genes to be up regulated during growth on fructose. These genes and the proteins that they encode for are as follows: *M5005_spy_0798* (IFN-response binding factor 1), *M5005_spy_1629* (Lantibiotic transport ATP-binding protein *salX*), and *M5005_spy_1631* (Lantibiotic salivaricin A). However, it is important to consider for further studies the environment to conduct these experiments in. The neutrophil/whole blood environment is different from CDM + fructose/glucose, therefore, the transcriptome results generated in CDM are not necessarily representative of the *ex vivo* environments. While the regulation of these genes is fructose-dependent, they are not necessarily *fruR/B*-dependent, and this needs to be investigated further. Regardless, the RNA-Seq data supports the stated

hypothesis that the virulence defect is attributed to a signaling effect due to feedback of *fruB/R*.

Surprisingly, we were unable to observe a virulence phenotype for either *fruR* or *fruB* during an intraperitoneal (*i.p.*) model of infection or during growth in whole mouse blood. Recapitulation of human disease in a susceptible laboratory animal has been the gold standard for pathogenesis since the inception of Koch's postulates. Host-pathogen research has favored mouse models due to reduced cost, reproducibility, and availability of reagents, amongst others. However, this poses a challenge with testing human-restricted pathogens in mice, where striking differences in the immune system between these two species can lead to varied disease progression. Immune system differences can cause misleading results as has been seen with using mouse models to identify protective antigens for a *S. aureus* vaccine, where identified targets in mice have not been a good predictor of outcomes in human clinical trails (320). Of particular note, the composition of blood between the two species is quite different; human blood is neutrophil rich (50-70% of blood), whereas mouse blood has more circulating lymphocytes and fewer neutrophils (10-25% of blood) (321). Therefore it is necessary to investigate what component(s) of whole human blood, which are not present in murine blood, are necessary for clearly GAS infections. This leads to the question whether or not the mere difference in neutrophil composition between these two species blood is enough to cause a difference in virulence? Or if it is a difference in recognition of the pathogen itself, and a host-specific signal generated during infection which causes an altered outcome in pathogenesis?

Furthermore, to elicit disease in a mouse, the inoculums of GAS required are often several times higher than what would be needed in humans (5, 8, 160, 322, 323). It can also be difficult to reiterate certain GAS diseases in mice. Mice do not elicit symptoms of GAS pharyngitis (strep throat) when infected intranasally or in the pharynx directly. An *in vivo* model of streptococcal impetigo, one of the most prevalent forms of non-invasive disease (1), required the grafting of human foreskin onto huSCID mice (324). Known GAS virulence factors, such as streptokinase (325, 326) and protein H (327) (328), were not testable until the development of humanized mice expressing either human plasminogen (329) or human C4BP/Factor H (330), respectively. Additionally, murine T cells appear not to proliferate and respond to GAS superantigens such as SMEZ, SpeG, or SpeC, while human T cells do (331). Taken together, it would appear that coevolution between the human immune system and GAS has resulted in adaptation to its natural host environment.

It cannot be ruled out, nevertheless, that the *i.p* model of infection was not the appropriate environment to exhibit a virulence phenotype for either *fruB* or *fruR*. Virulence genes and their transcriptional regulators are often implicated in specific stages of infection and growth. It would be important to attempt virulence studies using several different models of infection, for example the subcutaneous model of infection since it was shown that a mutant in EI of the PTS (*ptsI*) causes a hypervirulent phenotype in this model of infection through its altered temporal regulation of streptolysin S (SLS) (8). In addition, perhaps using huSCID mice may be more appropriate for these studies since the results presented in this dissertation suggest that *fruB* and *fruR* have a human-specific phenotype.

Overall, this dissertation has shown that FruB and FruR, through its regulation of *fruB*, are necessary for full virulence of GAS in a host-specific manner. It is hypothesized that FruB is directly stimulating downstream virulence gene regulation through a currently unknown signaling mechanism. Nonetheless, there is a clear link between fructose metabolism and regulation of immune evasion-related genes that are important for GAS during human infection.

The effect of glucose on Mga phosphorylation and activity.

The stand-alone regulator, Mga, possesses homology to PRD domains found in sugar-specific regulators, suggesting that Mga may be linked to the PTS (11). Thus, global virulence gene regulation can be directly influenced by carbohydrate metabolism. Work conducted by Hondorp *et al*. demonstrated that the PTS, through EI/Hpr, can modulate Mga *in vitro* by phosphorylating histidines located on the PRD domains of the protein; thus affecting activity *in vivo* through expression of its regulon (19). PRD regulators are also often phosphorylated by an $EIIB^{sugar}$ and can affect protein activity; however, this phosphodonor(s) has yet to be identified for Mga. Furthermore, while it is hypothesized that glucose is the main inducer of Mga regulon expression, it has never been conclusively demonstrated. This dissertation characterizes the effect of glucose on Mga phosphorylation and the subsequent effect it has on the expression of the Mga regulon in M1T1 GAS.

Transcription of several of the core Mga-regulated genes were repressed during growth in C media (low glucose) as opposed to THY (high glucose). Phos-tag gel analysis revealed that in the presence of high glucose (THY), Mga exhibited a phosphorylated and unphosphorylated species whereas in low glucose (C media) it

exhibited only an unphosphorylated species. For PRD-containing regulators, phosphorylation often creates antagonistic effects on activity depending on which domain is targeted (332). Hondorp *et al*. was able to demonstrate, using real-time qPCR and a phosphomimetic construct of the two phosphohistidine located on PRD1, that a there was a complete loss of activity regardless of the phosphorylation status of PRD2, indicating that PRD1-inactivation of Mga by phosphorylation is dominant (19). In addition, this condition is meant to mimic growth conditions in the absence of glucose. Importantly, the aforementioned study also suggested that phosphorylation of residue H324 (PRD2) enhances Mga activity, as a 2- to 3-fold loss of activity was observed for a complete phosphoablative mutant (PRD1/PRD2; A/A/A) compared to a construct that prevents phosphorylation on just PRD1 (A/A); the former occurring in the presence of both glucose and an inducing sugar, and the latter occurring solely in the presence of glucose. Interestingly, qPCR results from a H324A mutant (PRD2) alone showed no difference in regulon activity in comparison to the WT, which was attributed to the high level of protein produced due to the use of the multi-copy plasmid-based system used, causing a compensation for loss of regulon activity. Thus, based on the results of this study in combination with the model presented by Hondorp *et al*., we propose that Mga is phosphorylated on PRD2 during high levels of glucose (growth on THY), causing a hyperactive phenotype. This is validated by what was seen previously, where loss of phosphorylation on all three PRDs (A/A/A) shows a more dramatic loss in activation of regulon genes when compared to the wild type because in theory the WT is in the "hyperactive state" due to high glucose in THY and phosphorylation on PRD2. During growth in C media, since glucose is

present at low levels, Mga is still in an active state albeit to a lesser extent than observed in THY. Based on this model, C media must also have the hypothetical $EIIB^{sugar}$ active to import and inducing sugar, and shunt the flow of phosphates to the phosphohistidine on PRD2 of Mga.

The PTS is not the primary system for GAS glucose uptake based on the finding that a Δ*ptsI* mutant was defective on growth and utilization of the majority of PTS sugars, but not growth and utilization of glucose (8), indicating an alternative uptake system for glucose. Other species of low G+C Gram-positive bacteria encode for either a PTS-independent active transport system called the multiple sugar metabolism (MSM) system, or a non-PTS glucose uptake system called GlcU (333-335). However, genes encoding for a homologous MSM system are not found in the genome of GAS, but *glcU* is encoded for by GAS. Thus, since glucose plays a key role in not only the metabolism of GAS, but also the regulation of Mga, the search for the alternative glucose transport system(s) is currently under investigation in our lab. Interestingly, preliminary studies show that GlcU is not the only alternative non-PTS transport system in GAS because mutagenesis of *glcU* in a Δ*ptsI* background revealed that it can still grow on glucose. Therefore, it appears that GlcU may not be the sole glucose transport system encoded in GAS, and is not the preferred route of glucose uptake, which points to yet a third unknown method for glucose to be transported into the cell.

This currently unknown primary glucose transport system would be the predominant pathway for glucose uptake, and it would most likely be sufficient for growth in THY due to the high levels of glucose present in this media. During growth

on low glucose, such as in C media, sole utilization of this primary transport system may not be efficient for growth, and activation of the EII PTS transporters that are specific for glucose may occur. This can be seen in the RNA-Seq data presented here, where the PTS transporters for mannitol, cellobiose (both), trehalose, galactose, sialic acid, and the 3-keto-L-gulonate were all activated in C media. Therefore, this may provide the first evidence of a PTS EII(s) that is able to transport glucose into the cell. However, this raises the question as to which one(s) are actually transporting glucose? Additionally, do any of these EIIs act as a phosphodonor for Mga activation? In order to determine which PTS EII is directly interacting with Mga, a bacterial two-hybrid screen between Mga and all 14 EIIs could be used. In addition, if glucose is being transported in to the cell via an alternate PTS transport system it would mean that glucose itself could act as an inducing sugar. Thus the interplay between glucose, the PTS, and Mga activations would be more nuanced than previously thought.

It is also possible that the model proposed by Hondorp *et al*. is incorrect, and an alternative mechanism is used to regulate Mga activity in a high versus a low glucose environment. Phosphorylation is often a transient property of proteins, and so the current bimodal view of Mga phosphorylation and activity may not be an accurate representation of what is happening within the cell. It was originally hypothesized that C media, with its low glucose composition, would be a condition where Mga would be phosphorylated (on PRD1) and less active than in THY; however, only a band corresponding to an unphosphorylated protein was observed in C media. A recent study demonstrated that a putative protein tyrosine phosphatase (SP-PTP)

affects the Mga regulon expression, presumably through posttranslational modification of Mga (336). Thus, although no phosphorylation of Mga in C media was observed, it is possible that this phosphatase may be playing a role in the observed phosphorylation status of Mga in C media and masking alternative phosphorylation in these conditions. However, it is important to note that this phosphatase was not found to be differentially regulated in the wild type 5448 RNA-Seq in THY versus C media. Still, questions still remain as to whether this phosphatase affects Mga activity directly? It is possible that other regulatory proteins may be interacting with Mga and causing a change in activity. Questions regarding these interactions can be addressed using Tn-Seq in both THY and C media and to identify genetic interaction of *mga*. Tn-Seq allows for epistatic relationships between genes to be elucidated and thus, this could provide a clearer picture to not only what Mga is interacting with downstream but also upstream.

Regardless of the physical mechanism by which glucose is affecting the phosphorylation of Mga, it is clear that glucose is crucial for activation of the protein and the types of genes it regulates. The Mga virulence regulon was highly variable under different glucose levels. While 7.9% of the genome was regulated in both THY and C media, the genes within these two datasets were different. Only 16 genes were found in common, most of which were the core regulon genes (*emm, sic, scpA,* and *fba*). From the data generated here, several questions remain: How does phosphorylation of Mga enable the protein to differentially regulate genes based on glucose availability? Does the phosphorylation (hypothesized to occur on PRD2) cause a conformational change in the protein that allows it to bind to target DNA

differentially than when it is completely unphosphorylated? Or, does the variation in the regulon demonstrate a change in the cell on a much more global level; does removal of Mga in C media signal for a sort of stress response for GAS, which causes a Mga-independent transcriptional change that allows for compensation for loss of this stand-alone regulator? In order to answer this question, ChIP-Seq will need to be performed on Mga when grown in either THY or C media in order to detect direct DNA binding differences on a genome level.

Interestingly, when comparing the microarray study with the RNA-Seq data generated in THY, only 8 genes overlapped and 5 of these were part of the core regulon. Importantly, analysis of this dataset revealed that the only other gene that was in common in the microarray, THY RNA-Seq and the C media RNA-Seq was the gene *spy_0143*. This gene encodes for a 78 amino acid hypothetical protein, with unknown function. This small protein is reminiscent of the gene *grm*, which was previously shown to be directly regulated by Mga, but the function of this gene has never been identified (229), and no role in virulence was ever determined (Gold, unpublished data). Intriguingly, *grm* was found to be regulated in C media by Mga but not in THY in an M1T1 background, potentially indicating a low glucose-specific role for this gene. Therefore, there are questions that remain concerning why Mga regulates these two small peptides. For example, why does Mga regulate several of these smaller proteins? What are the roles of these peptides in GAS physiology and could they act as a signaling mechanism for peptide-mediated quorum sensing? In addition, why are some of these small peptides, such as *grm*, regulated in a glucose-specific manner?

Finally, the regulon in C media contained a number of novel genes that were involved in adhesion and immune evasion, such as all three phage-associated nucleases. Interestingly, it was demonstrated that much of the genes that were differentially-regulated in C media were due to the low glucose availability, and supplementation of C media with glucose restored the regulon to what was observed in THY. This further supports the conclusion that glucose is the inducer of Mga activity. Importantly, C media also acts a representative media condition for a deep tissue infection, including cellulitis and necrotizing fasciitis (88). Therefore, it would be interesting to analyze the Mga regulon in a subcutaneous model of skin infection to confirm that the profile of gene regulation that was observed *in vitro* (C media) mimics what is occurring *in vivo*.

Overall, this dissertation has shown that Mga is differentially phosphorylated due to glucose availability, which correlates with a change in the genes within the Mga regulon. It is hypothesized that THY acts as a condition where Mga is in a "hyperactive" condition, and phosphorylated on PRD2. In contrast, C media would act as a more physiologically-relevant condition that limits the source of glucose. However, since some glucose is still present in the media, it maintains Mga in an unphosphorylated and active form. While the location of the phosphorylation event on Mga can be directly attributed to the phosphohistdines, the residue(s) that are affected were not determined. How this affects the conformation of the protein *in vivo* has yet to be studied. Nonetheless, it is clear that glucose is a main inducer of Mga activity, and can alter the expression of the regulon.

Importance of the C-terminus of Mga for activity.

Through previous work, it is known that the key components for Mga activity are DNA-binding, dimerization, and PTS phosphorylation (19, 236, 251). Specifically, the N-terminus of Mga is important for binding to the promoters of target virulence genes, but is not sufficient for Mga activity *in vivo*. The C-terminus, specifically the EIIB domain, is necessary for dimerization and full activation of Mga (251). It is hypothesized that Mga forms homodimers using multiple amino acid residues on the EIIB domain based on the crystal structures of Mga homologs. A bacterial two-hybrid screen was used to identify residues that could play a role in dimerization of Mga. In this dissertation, we characterized nine different single amino acid mutants in the EIIB domain of Mga for their effect on transcriptional activation of known regulon genes. In addition, we modeled Mga onto its homolog, AtxA, from *B. anthracis* (PDB 4R6I) (216).

The two-hybrid screen was originally conducted based on the crystal structure of the Mga homolog, MafR (formally known as EF3013) from *E. faecalis* (PDB 3SQN) (217, 218). MafR was demonstrated to form homodimers based on interactions between its EIIB domains. Therefore, using targeted random mutagenesis, the EIIB domain of Mga was probed for putative interaction residues. A total of 20 putative interaction residues were identified in the screen as abolishing interaction of Mga1-Mga1. In each case, the mutation identified led to a change in amino acid polarity or size. Therefore, it is possible that such chemical changes may have altered local hydrogen bonding or intermolecular interactions of the protein fold enough to prevent homodimerization. From those mutants identified, nine were

validated by confirming protein expression levels and transcriptional activation of the Mga-regulated gene, *emm,* as an indirect method of measuring dimerization. Out of the nine residues, two (F430Y and D445V) were found to be expressed at levels comparable to wild type M1 Mga, and also affect the expression of *emm*. Thus, the data presented in this dissertation suggests that C-terminal Mga1 residues F430 and D445 normally function to mediate Mga homodimerization. However, based on the AtxA crystal structure, both of these residues exist outside of the putative dimer interface. Therefore, the question remains whether these residues are affecting dimerization *in vivo*, or if they are just affecting protein activity? This question can only be addressed by determining the dimerization state of the resulting Mga proteins, using techniques such as gel filtration or native gel analysis. However, if loss of *emm* activation is solely due to protein activity, how do these residues play a regulatory role in Mga activation?

In addition, since the two-hybrid study was conducted, a crystal structure of *B. anthracis* AtxA has revealed that this PRD-containing regulator forms dimers using both its EIIB and PRD2 domains (216). Based on this new information, and the modeling of Mga to the AtxA crystal structure, two putative residues were identified on PRD2 (E349 and Y387) as potential interaction partners; however, no defect in activation of *emm* was observed for either of the residues. A number of questions regarding Mga's ability to dimerize still remain and need to be explored. For example, does Mga form dimers similar to MafR (EIIB-EIIB) or like AtxA (PRD2-EIIB)? Further comprehensive mutational analysis of PRD2 may help answer this question. Most importantly, until a crystal structure of Mga is obtained, it is difficult

to make accurate predictions pertaining to the quaternary structure of Mga and the tertiary effects of individual amino acids in the EIIB region.

This study used the Mga homolog and transcriptional regulator, AtxA from *B. anthracis* (PDB 4R6I) (216) as a model for Mga dimerization. Interestingly, the C-terminal tail was extremely unstructured in all three predicted protein structures, while the rest of the protein was mapped with a high degree of precision. In addition, it was extremely difficult to produce mutations in Mga in this study and in others (308). Taken together, Mga possesses unique properties in its C-terminus, and therefore it crucial for future studies involving interrogation of this region to have a precise Mga crystal structure. However, due to difficulties in the quantity of Mga generated from protein purification, it has been impossible to produce a crystal structure thus far.

Interestingly, while mutational analysis revealed that complete removal of the EIIB domain abolished dimer formation in solution, and consequently loss of activation of regulon genes, it was still able to bind cognate DNA *in vitro*. This points to a much more complex mechanism by which Mga is able to activate transcription of its regulon genes, and multimerization plays a key role in this activation. Thus, it would be important to test the mutants identified in this study for their effect on DNA binding. Based on previous studies it is hypothesized that they would maintain their ability to bind to target DNA.

Historically, dimerization is not the only important attribute of EIIB domains on PRD-containing regulators. Many possess a conserved regulatory amino acid on their EIIB domain that can be phosphorylated to modulate protein activity. For

example, the *B. subtilis* PRD regulator MtlR contains a cysteine in its $EIIB^{Gat}$-like domain that can modulate its activity (337) through phosphorylation by a PTS EIIA (211). Mga does not have the conserved catalytic cysteine residue, and instead a glutamate (E413) exists in its place. Therefore, it was hypothesized that this glutamate may serve as a phosphomimetic and further play a role in Mga activity (251). However, in a study conducted by Kruszewski *et al*., mutational analysis of the glutamate to an alanine (E413A) revealed that this residue was not important for Mga activity (310). However, this study was conducted using a multi-copy plasmid system, and therefore it may alter the phenotypes that were observed. Therefore, this residue may still play a role in Mga activity, and further investigation of E413 should be conducted in single-copy using ectopic integration on the chromosome. Identification of a putative "activation switch amino acid residue(s)" would have potentially not been identified by the two-hybrid screen, as they may not affect dimerization. Therefore, if this residue or another is phosphorylated, it would be important to 1) determine if activity is affected by loss of dimerization, and 2) identify what EIIA is responsible for the phosphorylation of this putative residue. In order to answer this question alanine scanning on the EIIB domain and subsequent analysis of activation of Mga may answer this question. However, this may prove to be difficult due to the general issues with generating mutation in this region stated above.

Since Mga also relies on phosphorylation events on its PRD domains for activation, phosphorylation on PRD1 may play antagonistic effect on Mga activity not only by affecting general activation but also can abrogate dimer formation by

causing a conformational change that occludes interaction residues on the EIIB domain. Conversely, dimerization may hinder phosphorylation of PRD1 by occluding the phosphohistidine and prolong activation of the Mga regulon. A similar scenario was demonstrated in the archetype PRD-containing anti-terminator, LicT. The phosphohistidines of LicT are buried deep within the dimer interface, and therefore are inaccessible to phosphorylating enzymes (205). *In silico* models predicted that phosphorylation of the histidines on PRD1 lead to strong electrostatic repulsions, whereas phosphorylation of PRD2 leads to increased electrostatic complementarity between the homodimer (208). Thus, it would be interesting to investigate the dynamics of Mga phosphorylation, not only in the context of activation of its regulon, but also in the context of dimer formation.

Overall, this study demonstrated that two amino acid (F430 and D445) substitutions within the EIIB domain affect Mga activity, and potentially mediate Mga dimerization. Furthermore, while Mga homologs have played a useful role in characterizing potential interaction residues for Mga dimerization, future work would greatly benefit from a direct Mga crystal structure.

Summary

In conclusion, the results of this work have characterized the role of a PTS-fructose operon in both fructose metabolism and its role in GAS pathogenesis. FruA was identified as the main transporter of fructose in the M1T1 GAS, and all three operon genes (*fruRBA*) were necessary for growth in fructose. In addition, FruB and FruR, through its regulation of *fruB*, were found to be are necessary for full virulence of GAS in a host-specific manner. Furthermore, we characterize the regulatory inputs

that are necessary for full Mga activation. Glucose was found to be the inducing sugar that promotes Mga activity and causes differential phosphorylation patterns of the protein on its PTS-regulatory domains (PRD), which subsequently alter activation of the Mga regulon. Mga activity was also affected by individual amino acid residues located in the EIIB domain, and implicate these residues in dimer formation. Overall, the results provided within this dissertation have established the functional characterization of the role of sugars as major contributors to virulence factor production through either directly eliciting genome-wide transcriptional changes or through their interaction with global transcriptional regulators such as Mga.

Appendices

Appendix 1: Genes in 5448 differentially expressed in CDM-glucose compared to CDM-fructose

Spy #	Annotation	Gene Name	Log_2 FC [a]
M5005_Spy_0039	alcohol dehydrogenase/acetaldehyde dehydrogenase (acetylating)	*adh2*	-2.10
M5005_Spy_0086	comG operon protein 1	*comYA*	-0.99
M5005_Spy_0089	comG operon protein 4		1.72
M5005_Spy_0120	acetate CoA-transferase alpha subunit	*atoD.2*	1.76
M5005_Spy_0123	translation initiation inhibitor		-1.82
M5005_Spy_0124	serine catabolism regulator	*sloR*	-2.32
M5005_Spy_0125	hypothetical protein		-1.09
M5005_Spy_0126	V-type sodium ATP synthase subunit I	*ntpI*	-2.01
M5005_Spy_0127	V-type sodium ATP synthase subunit K	*ntpK*	-1.97
M5005_Spy_0128	V-type sodium ATP synthase subunit E	*ntpE*	-2.22
M5005_Spy_0129	V-type ATP synthase subunit C	*ntpC*	-1.79
M5005_Spy_0130	V-type sodium ATP synthase subunit F	*ntpF*	-1.22
M5005_Spy_0131	V-type sodium ATP synthase subunit A	*ntpA*	-1.24
M5005_Spy_0132	V-type sodium ATP synthase subunit B	*ntpB*	-0.96
M5005_Spy_0133	V-type sodium ATP synthase subunit D	*ntpD*	-1.21
M5005_Spy_0145	hypothetical protein		-1.31
M5005_Spy_0150	PTS system, **3-keto-L-gulonate** specific IIA	*ptxA*	-0.99
M5005_Spy_0152	L-xylulose 5-phosphate 3-epimerase		1.37
M5005_Spy_0154	hypothetical protein		1.57
M5005_Spy_0166	transposase		-1.32
M5005_Spy_0167	transposase		-1.15
M5005_Spy_0168	transposase		-1.70
M5005_Spy_0170	nicotinate-nucleotide pyrophosphorylase (carboxylating)	*nadC*	-1.95
M5005_Spy_0184	hypothetical cytosolic protein		-1.28
M5005_Spy_0189	hypothetical protein		-1.66
M5005_Spy_0256	competence-specific sigma factor		-1.68
M5005_Spy_0297	transposase		-1.12
M5005_Spy_0395	transposase		-1.39
M5005_Spy_0451	transcriptional regulator		-1.07
M5005_Spy_0460	hypothetical protein		1.13
M5005_Spy_0474	transcription antiterminator, BglG family	*licT*	-1.36
M5005_Spy_0476	6-phospho-beta-glucosidase	*bglA*	-1.22
M5005_Spy_0479	hypothetical membrane spanning protein		-1.80
M5005_Spy_0518	oligohyaluronate lyase		1.35
M5005_Spy_0521	PTS system, **N-acetylgalactosamine**-specific IIB	*agaV*	-1.75
M5005_Spy_0616	ferredoxin		-1.02
M5005_Spy_0650	hypothetical protein		-1.95
M5005_Spy_0659	2-dehydropantoate 2-reductase	*apbA*	-0.95
M5005_Spy_0660	fructose repressor	*fruR*	-1.95
M5005_Spy_0661	1-phosphofructokinase	*fruB*	-2.52
M5005_Spy_0662	PTS system, **Fructose**-specific IIABC	*fruA*	-3.33
M5005_Spy_0666	hypothetical protein		-1.70
M5005_Spy_0667	exotoxin type C precursor		2.07
M5005_Spy_0773	hypothetical protein		0.99
M5005_Spy_0798	IFN-response binding factor 1		-1.01

M5005_Spy_0800	DNA-cytosine methyltransferase		-1.38
M5005_Spy_0853	short chain dehydrogenase		-1.25
M5005_Spy_0912	hypothetical protein		-1.41
M5005_Spy_0997	phage protein		1.20
M5005_Spy_0998	phage protein		-1.40
M5005_Spy_0999	phage protein		1.92
M5005_Spy_1000	phage protein		2.52
M5005_Spy_1010	phage protein		1.56
M5005_Spy_1013	antigen B		2.32
M5005_Spy_1017	phage protein		2.65
M5005_Spy_1023	terminase large subunit		1.52
M5005_Spy_1025	phage encoded transcriptional regulator, ArpU family		0.96
M5005_Spy_1026	phage protein		2.04
M5005_Spy_1027	phage protein		1.00
M5005_Spy_1028	phage protein		1.12
M5005_Spy_1029	phage protein		1.92
M5005_Spy_1030	phage protein		1.38
M5005_Spy_1032	phage protein		1.84
M5005_Spy_1033	phage protein		1.69
M5005_Spy_1034	phage protein		1.48
M5005_Spy_1035	phage protein		1.41
M5005_Spy_1036	phage single-strand DNA binding protein	*ssb2*	1.48
M5005_Spy_1037	phage single-strand DNA binding protein	*ssb1*	2.07
M5005_Spy_1038	phage protein		1.87
M5005_Spy_1039	phage protein		2.09
M5005_Spy_1040	phage protein		1.71
M5005_Spy_1041	phage protein		2.76
M5005_Spy_1042	phage replication protein		2.03
M5005_Spy_1043	phage protein		1.88
M5005_Spy_1045	transcriptional regulator		1.80
M5005_Spy_1062	maltodextrose utilization protein	*malA*	-1.12
M5005_Spy_1063	ABC **Cyclomaltodextrin** permease protein	*malD*	-1.17
M5005_Spy_1078	hypothetical protein		-1.53
M5005_Spy_1081	PTS system, **Cellobiose**-specific IIA component		-0.98
M5005_Spy_1089	hypothetical protein		-1.19
M5005_Spy_1090	transposase		-1.58
M5005_Spy_1139	glucosamine-6-phosphate isomerase	*nagB*	0.98
M5005_Spy_1172	holin		0.99
M5005_Spy_1173	phage protein		-1.33
M5005_Spy_1174	phage protein		2.37
M5005_Spy_1175	phage protein		1.47
M5005_Spy_1176	phage infection protein		1.34
M5005_Spy_1187	phage structural protein		1.05
M5005_Spy_1190	phage protein		-1.19
M5005_Spy_1200	phage protein		1.14
M5005_Spy_1201	phage protein		1.94
M5005_Spy_1203	phage protein		2.16
M5005_Spy_1204	recT protein		1.59
M5005_Spy_1205	phage protein		1.75
M5005_Spy_1206	phage protein		1.91
M5005_Spy_1208	phage protein		2.23
M5005_Spy_1209	DNA replication protein		1.82
M5005_Spy_1210	phage replication protein		1.45
M5005_Spy_1211	phage protein		1.42
M5005_Spy_1213	phage protein		1.72
M5005_Spy_1216	phage protein		1.57

M5005_Spy_1217	phage antirepressor protein		1.20
M5005_Spy_1218	phage protein		1.79
M5005_Spy_1307	hypothetical membrane spanning protein		-1.02
M5005_Spy_1308	ABC unknown sugar-binding protein		-1.16
M5005_Spy_1309	ABC unknown sugar permease protein		-1.82
M5005_Spy_1310	ABC unknown sugar permease protein		-1.40
M5005_Spy_1388	N-acetylglucosamine-6-phosphate deacetylase	*nagA*	0.97
M5005_Spy_1395	tagatose-bisphosphate aldolase; SpeB regulator	*lacD.1*	-1.01
M5005_Spy_1396	tagatose-6-phosphate kinase	*lacC.1*	-1.35
M5005_Spy_1397	galactose-6-phosphate isomerase lacB subunit	*lacB.1*	-1.54
M5005_Spy_1400	PTS system, **Galactose**-specific IIB component		-1.45
M5005_Spy_1401	PTS system, **Galactose**-specific IIA component		-1.14
M5005_Spy_1414	phage protein		1.42
M5005_Spy_1416	phage-associated cell wall hydrolase		1.90
M5005_Spy_1417	phage protein		1.44
M5005_Spy_1418	phage protein		2.34
M5005_Spy_1419	phage protein		2.03
M5005_Spy_1420	phage protein		1.67
M5005_Spy_1421	phage infection protein		1.81
M5005_Spy_1422	phage protein		1.48
M5005_Spy_1423	hyaluronoglucosaminidase		1.95
M5005_Spy_1424	phage endopeptidase		1.89
M5005_Spy_1425	phage protein		2.07
M5005_Spy_1426	phage protein		2.49
M5005_Spy_1427	phage protein		2.18
M5005_Spy_1428	phage protein		2.41
M5005_Spy_1429	phage protein		2.27
M5005_Spy_1430	phage protein		1.99
M5005_Spy_1431	phage protein		1.95
M5005_Spy_1432	phage protein		2.30
M5005_Spy_1433	phage protein		2.00
M5005_Spy_1434	phage protein		2.38
M5005_Spy_1435	phage scaffold protein		2.40
M5005_Spy_1436	phage protein		2.18
M5005_Spy_1437	hypothetical protein		2.26
M5005_Spy_1438	phage protein		1.81
M5005_Spy_1439	portal protein		2.14
M5005_Spy_1440	terminase large subunit		1.79
M5005_Spy_1441	phage terminase small subunit		1.83
M5005_Spy_1442	phage transcriptional activator		2.50
M5005_Spy_1444	adenine-specific methyltransferase		1.89
M5005_Spy_1446	phage protein		1.20
M5005_Spy_1447	phage-related DNA helicase		1.49
M5005_Spy_1448	hypothetical protein		1.22
M5005_Spy_1449	DNA primase		2.37
M5005_Spy_1450	phage-encoded DNA polymerase		1.80
M5005_Spy_1451	phage protein		1.69
M5005_Spy_1452	phage protein		1.74
M5005_Spy_1453	phage protein		1.60
M5005_Spy_1454	phage protein		2.02
M5005_Spy_1455	phage protein		1.82
M5005_Spy_1457	phage protein		2.66
M5005_Spy_1458	phage protein		1.70
M5005_Spy_1459	phage protein		1.57
M5005_Spy_1460	phage protein		1.35
M5005_Spy_1461	phage protein		2.43

M5005_Spy_1463	phage protein		1.79
M5005_Spy_1509	pyruvate, phosphate dikinase		-1.36
M5005_Spy_1510	pyruvate, phosphate dikinase		-1.26
M5005_Spy_1536	transposase		-1.17
M5005_Spy_1541	hypothetical protein		-1.10
M5005_Spy_1574	universal stress protein family		2.05
M5005_Spy_1575	quinolone resistance protein	*norA*	2.33
M5005_Spy_1629	lantibiotic transport ATP-binding protein	*salX*	-1.89
M5005_Spy_1631	lantibiotic salivaricin A	*salA*	-1.67
M5005_Spy_1632	6-phospho-beta-galactosidase	*lacG*	-1.66
M5005_Spy_1633	PTS system, **Lactose**-specific IIBC component	*lacE*	-2.73
M5005_Spy_1634	PTS system, **Lactose**-specific IIA component	*lacF*	-2.22
M5005_Spy_1635	tagatose-bisphosphate aldolase	*lacD.2*	-2.18
M5005_Spy_1636	tagatose-6-phosphate kinase	*lacC.2*	-1.49
M5005_Spy_1637	galactose-6-phosphate isomerase lacB subunit	*lacB.2*	-1.67
M5005_Spy_1638	galactose-6-phosphate isomerase lacA subunit	*lacA.2*	-1.50
M5005_Spy_1643	hypothetical protein		0.99
M5005_Spy_1663	PTS system, **Mannitol**-specific IIB component		-1.69
M5005_Spy_1689	hypothetical protein		1.42
M5005_Spy_1744	PTS system, **Cellobiose**-specific IIC component	*celC*	-1.34
M5005_Spy_1745	PTS system, **Cellobiose**-specific IIB component	*celB*	-0.95
M5005_Spy_1746	PTS system, **Cellobiose**-specific IIA component	*celA*	1.02
M5005_Spy_1786	transcriptional regulator, MarR family		1.04
M5005_Spy_1832	hypothetical protein		-1.15
M5005_Spy_1841	L-serine dehydratase	*sdhB*	-1.06
M5005_Spy_1842	L-serine dehydratase	*sdhA*	-1.27
M5005_SpyR0011	large subunit 23S ribosomal RNA		-1.05
M5005_SpyT0020	tRNA-Pro		-1.84
M5005_SpyT0030	tRNA-Val		1.00
M5005_SpyT0031	tRNA-Gly		-1.80
M5005_SpyT0032	tRNA-Ile		-1.52
M5005_SpyT0033	tRNA-Glu		-1.35
M5005_SpyT0036	tRNA-Phe		-1.13
M5005_SpyT0040	tRNA-Gln		-2.59
M5005_SpyT0041	tRNA-Leu		-1.13
M5005_SpyT0046	tRNA-Tyr		-1.02
M5005_SpyT0052	tRNA-Thr		-1.00
M5005_SpyT0065	tRNA-Asn		-0.97

[a] Limma analysis exhibited Log_2 fold-change of $-.85 \leq$ or $\geq .85$

Appendix 2: Genes in 5448 differentially expressed in THY compared to C Media

Spy #	Annotation	Gene	Log_2 FC [a]	*p* value [b]
M5005_Spy0011	tRNA(Ile)-lysidine synthetase	*tilS*	-1.45	3E-04
M5005_Spy0017	CHAP-domain-containing cell wall hydrolase; pcsB/sibA homolog	*cdhA*	2.34	7E-07
M5005_Spy0019	DNA repair protein	*recO*	-1.05	1E-03
M5005_Spy0022	phosphoribosylaminoimidazole-succinocarboxamide synthase		2.45	1E-04
M5005_Spy0023	phosphoribosylformylglycinamidine synthase		1.97	3E-04
M5005_Spy0024	amidophosphoribosyltransferase	*purF*	2.76	2E-06
M5005_ Spy0025	phosphoribosylformylglycinamidine cyclo-ligase	*purM*	2.38	1E-04
M5005_Spy0026	phosphoribosylglycinamide formyltransferase	*purN*	2.05	1E-04
M5005_Spy0027	phosphoribosylaminoimidazolecarboxamide formyltransferase/IMP cyclohydrolase		2.00	3E-05
M5005_Spy0029	phosphoribosylamine--glycine ligase	*purD*	2.21	7E-04
M5005_Spy0030	phosphoribosylaminoimidazole carboxylase catalytic subunit	*purE*	1.63	8E-03
M5005_Spy0031	phosphoribosylaminoimidazole carboxylase ATPase subunit	*purK*	1.27	5E-03
M5005_Spy0032	hypothetical protein		1.06	9E-03
M5005_Spy0034	transcriptional regulator		1.24	3E-03
M5005_Spy0035	holliday junction DNA helicase	*ruvB*	1.46	2E-04
M5005_Spy0039	alcohol dehydrogenase/acetaldehyde dehydrogenase (acetylating)	*adh2*	-2.07	7E-03
M5005_Spy0040	alcohol dehydrogenase	*adhA*	-1.94	2E-02
M5005_Spy0041	Na+ driven multidrug efflux pump		-1.03	9E-03
M5005_Spy0043	SSU ribosomal protein S10P	*rpsJ*	1.15	2E-02
M5005_Spy0044	LSU ribosomal protein L3P	*rplC*	1.04	1E-02
M5005_Spy0052	LSU ribosomal protein L29P	*rpmC*	1.54	1E-02
M5005_Spy0064	protein translocase subunit	*secY*	1.22	8E-03
M5005_Spy0065	adenylate kinase	*adk*	1.97	4E-04
M5005_Spy0066	bacterial protein translation initiation factor 1	*infA*	1.54	4E-04
M5005_Spy0068	SSU ribosomal protein S13P	*rpsM*	1.25	2E-03
M5005_Spy0070	DNA-directed RNA polymerase alpha chain	*rpoA*	1.49	9E-04
M5005_Spy0071	LSU ribosomal protein L17P	*rplQ*	1.70	3E-04
M5005_Spy0073	hypothetical protein		2.57	5E-03
M5005_Spy0076	4-diphosphocytidyl-2-C-methyl-D-erythritol kinase		1.16	2E-02
M5005_Spy0077	transcriptional regulator, MarR family	*adcR*	1.09	9E-05
M5005_Spy0078	high-affinity zinc uptake system ATP-binding protein	*adcC*	1.10	1E-03
M5005_Spy0085	putative DNA binding protein		2.62	9E-06
M5005_Spy0086	comG operon protein 1	*comYA*	2.00	5E-03
M5005_Spy0088	comG operon protein 3	*comYC*	1.76	2E-03
M5005_Spy0094	acetate kinase	*ackA*	1.21	2E-03
M5005_Spy0095	hypothetical protein		2.59	2E-04
M5005_Spy0106	transcriptional regulator; Regulator of protein F	*rofA*	-1.39	8E-04
M5005_Spy0107	fibronectin-binding protein; FCT Pilus Region	*cpa*	2.50	4E-07
M5005_Spy0108	signal peptidase I; FCT Pilus Region	*sipA1*	2.52	3E-06
M5005_Spy0109	fibronectin-binding protein; FCT Pilus Region		2.37	1E-04
M5005_Spy0110	hypothetical protein	*srtC2*	2.61	8E-06
M5005_Spy0111	hypothetical protein		1.63	2E-03
M5005_Spy0113	transposase		2.14	4E-03

M5005_Spy0114	sortase; FCT Pilus Region	*srtB*	2.59	3E-04
M5005_Spy0115	hypothetical protein		-1.26	4E-04
M5005_Spy0134	tellurite resistance protein		-1.33	1E-05
M5005_Spy0137	nucleoside-binding protein		1.39	2E-05
M5005_Spy0142	hypothetical protein		2.34	2E-03
M5005_Spy0143	hypothetical protein		3.97	2E-03
M5005_Spy0148	PTS system, **3-keto-L-gulonate** specific IIC	*ptxC*	-2.37	6E-03
M5005_Spy0149	PTS system, **3-keto-L-gulonate** specific IIB	*ptxB*	-4.53	2E-05
M5005_Spy0150	PTS system, **3-keto-L-gulonate** specific IIA	*ptxA*	-4.61	3E-04
M5005_Spy0151	3-keto-L-gulonate-6-phosphate decarboxylase		-3.52	4E-03
M5005_Spy0152	L-xylulose 5-phosphate 3-epimerase		-3.57	7E-04
M5005_Spy0153	L-ribulose-5-phosphate 4-epimerase	*araD*	-5.77	3E-06
M5005_Spy0154	hypothetical protein		-3.95	6E-05
M5005_Spy0156	metal-dependent hydrolase		-2.14	4E-04
M5005_Spy0157	glycine betaine transport ATP-binding protein	*opuAA*	2.11	2E-07
M5005_Spy0158	glycine betaine transport system permease protein	*opuABC*	2.01	2E-06
M5005_Spy0174	hypothetical protein		1.96	1E-02
M5005_Spy0175	queuine tRNA-ribosyltransferase	*tgt*	1.06	2E-03
M5005_Spy0177	bioY protein		1.21	2E-02
M5005_Spy0185	glucose-6-phosphate isomerase	*pgi*	1.09	4E-02
M5005_Spy0192	UTP--glucose-1-phosphate uridylyltransferase	*hasC.2*	-1.19	1E-03
M5005_Spy0194	glycerol-3-phosphate dehydrogenase; GAPDH	*gapA*	-1.27	2E-04
M5005_Spy0195	Heme-sensitive transcriptional repressor, MarR-family	*pefR*	1.17	1E-02
M5005_Spy0204	TCS-1 histidine kinase FasB	*fasB*	1.10	1E-03
M5005_Spy0206	TCS-1 response regulator FasA	*fasA*	1.13	1E-02
M5005_Spy0211	LSU ribosomal protein L34P	*rpmH*	1.64	2E-03
M5005_Spy0213	ABC **Sialic Acid**; N-acetylneuraminate-binding protein		-1.98	2E-03
M5005_Spy0214	ABC **Sialic Acid**; N-acetylneuraminate permease protein		-2.11	6E-04
M5005_Spy0215	ABC **Sialic Acid**; N-acetylneuraminate permease protein		-1.83	3E-03
M5005_Spy0216	hypothetical membrane spanning protein		-2.87	1E-04
M5005_Spy0217	N-acetylneuraminate lyase	*nanH*	-1.94	3E-03
M5005_Spy0218	N-acetylmannosamine kinase		-1.95	3E-03
M5005_Spy0221	ribonuclease M5		1.31	2E-03
M5005_Spy0224	ribulose-phosphate 3-epimerase	*rpe*	-1.36	1E-04
M5005_Spy0230	SSU ribosomal protein S12P	*rpsL*	1.71	4E-04
M5005_Spy0231	SSU ribosomal protein S7P	*rpsG*	1.32	4E-03
M5005_Spy0238	putative undecaprenol kinase	*bacA*	-1.20	2E-03
M5005_Spy0241	hypothetical cytosolic protein	*rgpG*	1.62	3E-02
M5005_Spy0247	D-alanyl-D-alanine carboxypeptidase		1.79	1E-05
M5005_Spy0273	ABC transporter permease protein		1.32	4E-02
M5005_Spy0274	branched-chain amino acid transport system carrier protein	*braB*	4.77	6E-07
M5005_Spy0275	serine/threonine sodium symporter		3.26	2E-07
M5005_Spy0295	60 kDa inner membrane protein YIDC		-1.20	2E-03
M5005_Spy0296	acylphosphatase		-1.23	2E-03
M5005_Spy0307	DNA integration/recombination/invertion protein		-1.22	2E-04
M5005_Spy0318	pyruvate formate-lyase activating enzyme	*pflC*	1.35	4E-04
M5005_Spy0321	ferrichrome transport system permease protein	*fhuG*	1.34	5E-04
M5005_Spy0323	ferrichrome-binding protein	*fhuD*	1.13	8E-03
M5005_Spy0324	ferrichrome transport ATP-binding protein	*fhuA*	1.85	1E-04

M5005_Spy0329	hypothetical cytosolic protein		-1.11	5E-05
M5005_Spy0333	signal peptidase-like protein		-1.23	5E-04
M5005_Spy0335	corrin/porphyrin methyltransferase		-0.97	2E-04
M5005_Spy0336	hypothetical membrane associated protein		-1.13	1E-02
M5005_Spy0338	arsenate reductase family protein		1.24	1E-03
M5005_Spy0341	lactocepin; SpyCEP; IL-8 protease	*spyCEP*	-1.78	1E-04
M5005_Spy0346	hypothetical protein		-1.41	3E-02
M5005_Spy0347	ribonucleoside-diphosphate reductase beta chain	*nrdF.1*	-1.25	2E-03
M5005_Spy0348	hypothetical protein	*nrdI*	-1.49	5E-03
M5005_Spy0349	ribonucleoside-diphosphate reductase alpha chain	*nrdE.1*	-1.56	4E-03
M5005_Spy0361	phosphoglycerate transporter protein		-1.05	3E-02
M5005_Spy0374	LSU ribosomal protein L11P	*rplK*	1.84	5E-03
M5005_Spy0375	LSU ribosomal protein L1P	*rplA*	1.86	3E-03
M5005_Spy0385	67 kDa Myosin-crossreactive antigen		-1.23	5E-03
M5005_Spy0387	uracil DNA glycosylase superfamily protein		-0.98	1E-02
M5005_Spy0398	bacteriocin		1.53	4E-02
M5005_Spy0421	lactoylglutathione lyase	*gloA*	-1.70	1E-05
M5005_Spy0422	NAD(P)H-dependent quinone reductase		-1.94	4E-06
M5005_Spy0423	Xaa-Pro dipeptidase	*pepQ*	-1.61	2E-04
M5005_Spy0425	glycosyltransferase		-1.29	2E-04
M5005_Spy0426	1,2-diacylglycerol 3-glucosyltransferase		-1.52	3E-04
M5005_Spy0457	plasmid stabilization system protein		-1.08	2E-02
M5005_Spy0461	hypothetical cytosolic protein		-1.67	2E-04
M5005_Spy0462	hypothetical protein		-1.50	7E-03
M5005_Spy0463	hypothetical cytosolic protein		-1.13	5E-03
M5005_Spy0464	microcin C7 self-immunity protein	*mccF*	-1.14	1E-03
M5005_Spy0465	hypothetical protein		-0.95	2E-02
M5005_Spy0466	hypothetical protein		-1.43	3E-05
M5005_Spy0467	transposase		-1.09	3E-02
M5005_Spy0469	hypothetical protein		-1.74	9E-04
M5005_Spy0473	multidrug resistance protein B		1.46	1E-03
M5005_Spy0474	transcription antiterminator, BglG family	*licT*	2.43	8E-05
M5005_Spy0475	PTS system, **b-Glucoside**-specific IIABC		3.78	3E-08
M5005_Spy0476	6-phospho-beta-glucosidase	*bglA*	2.94	4E-06
M5005_Spy0477	hypothetical membrane spanning protein		-1.67	4E-05
M5005_Spy0478	hypothetical membrane spanning protein		-1.65	4E-04
M5005_Spy0479	hypothetical membrane spanning protein		-3.11	5E-02
M5005_Spy0483	stress-responsive transcriptional regulator		-2.31	4E-08
M5005_Spy0485	prolipoprotein diacylglyceryl transferase	*lgt*	-0.96	6E-04
M5005_Spy0486	hypothetical protein		-1.21	3E-03
M5005_Spy0496	hydrolase, HAD superfamily		1.90	5E-06
M5005_Spy0503	glutathione peroxidase		-1.45	1E-04
M5005_Spy0504	oligoendopeptidase F	*pepF*	-1.70	1E-05
M5005_Spy0512	hydrolase, HAD superfamily		-1.02	9E-05
M5005_Spy0523	hypothetical protein		-1.63	2E-02
M5005_Spy0528	beta-phosphoglucomutase/glucose-1-phosphate phosphodismutase		-1.13	1E-03
M5005_Spy0539	ATP-binding protein		-1.30	2E-05
M5005_Spy0540	transporter		-1.27	4E-04
M5005_Spy0541	hypothetical cytosolic protein		-1.50	6E-05
M5005_Spy0542	dipeptidase	*pepD*	-1.77	9E-05
M5005_Spy0543	high-affinity zinc uptake system protein znuA precursor		1.49	3E-03
M5005_Spy0545	galactosamine-6-phosphate deaminase (isomerizing)	*agaS*	-0.93	1E-02

M5005_Spy0546	LSU ribosomal protein L31P	*rpmE*	1.77	5E-06
M5005_Spy0549	chorismate mutase		1.13	1E-02
M5005_Spy0551	LSU ribosomal protein L19P	*rplS*	2.14	5E-05
M5005_Spy0555	hypothetical cytosolic protein		-1.40	8E-04
M5005_Spy0558	transposase		-1.01	5E-02
M5005_Spy0562	streptolysin S precursor	*sagA*	-2.37	2E-04
M5005_Spy0563	streptolysin S biosynthesis protein	*sagB*	-1.86	2E-04
M5005_Spy0564	streptolysin S biosynthesis protein	*sagC*	-1.87	1E-04
M5005_Spy0565	streptolysin S biosynthesis protein	*sagD*	-1.25	2E-03
M5005_Spy0566	streptolysin S putative self-immunity protein	*sagE*	-2.32	6E-05
M5005_Spy0567	streptolysin S biosynthesis protein	*sagF*	-1.82	1E-03
M5005_Spy0568	streptolysin S export ATP-binding protein	*sagG*	-1.88	8E-05
M5005_Spy0569	streptolysin S export transmembrane protein	*sagH*	-1.76	2E-04
M5005_Spy0570	streptolysin S export transmembrane protein	*sagI*	-1.14	3E-02
M5005_Spy0575	ATP synthase C chain	*atpE*	1.60	1E-04
M5005_Spy0576	ATP synthase A chain	*atpB*	1.26	2E-03
M5005_Spy0579	ATP synthase alpha chain	*atpA*	1.43	2E-02
M5005_Spy0593	neutral zinc metallopeptidase family		-1.35	8E-05
M5005_Spy0594	ATP-dependent nuclease subunit B	*rexB*	-1.18	7E-04
M5005_Spy0595	ATP-dependent nuclease subunit A	*rexA*	-1.26	4E-04
M5005_Spy0597	SSU ribosomal protein S21P		1.39	1E-03
M5005_Spy0611	Group A Carbohydrate; Membrane protein	*gacJ*	-1.02	9E-03
M5005_Spy0615	pore forming protein	*ebsA*	-1.66	4E-04
M5005_Spy0616	ferredoxin		-1.78	5E-02
M5005_Spy0618	cytidylate kinase	*cmk*	-1.08	4E-04
M5005_Spy0619	bacterial protein translation initiation factor 3	*infC*	2.03	4E-04
M5005_Spy0620	LSU ribosomal protein L35P	*rpl36*	1.93	2E-03
M5005_Spy0621	LSU ribosomal protein L20P	*rplT*	2.34	1E-05
M5005_Spy0624	3-dehydroquinate dehydratase	*aroD*	1.25	4E-04
M5005_Spy0625	chorismate synthase	*aroF*	1.20	5E-03
M5005_Spy0628	folylpolyglutamate synthase/dihydrofolate synthase	*folC.2*	-1.18	2E-04
M5005_Spy0633	LSU ribosomal protein L21P	*rplU*	1.84	4E-05
M5005_Spy0634	hypothetical ribosome-associated protein		1.05	1E-03
M5005_Spy0635	LSU ribosomal protein L27P	*rpmA*	2.05	3E-05
M5005_Spy0636	transcriptional regulator, LysR family		-1.36	2E-05
M5005_Spy0637	lipoprotein signal peptidase	*lsp*	-1.45	5E-05
M5005_Spy0638	ribosomal large subunit pseudouridine synthase D		-1.21	1E-05
M5005_Spy0639	pyrR bifunctional protein	*pyrR*	1.57	2E-03
M5005_Spy0640	uracil permease	*pyrP*	1.30	5E-03
M5005_Spy0641	aspartate carbamoyltransferase	*pyrB*	2.18	2E-05
M5005_Spy0642	carbamoyl-phosphate synthase small chain	*carA*	1.73	4E-04
M5005_Spy0643	carbamoyl-phosphate synthase large chain	*carB*	1.62	2E-03
M5005_Spy0648	SSU ribosomal protein S16P	*rpsP*	1.74	4E-04
M5005_Spy0649	RNA binding protein		1.59	7E-05
M5005_Spy0651	Heme-utilization cell surface protein	*hupY*	1.11	1E-02
M5005_Spy0652	Heme-oxygenase cytosolic protein	*hupZ*	1.76	5E-04
M5005_Spy0653	cobalt-zinc-cadmium resistance protein	*czcD*	-3.16	3E-05
M5005_Spy0655	16S rRNA processing protein	*rimM*	1.25	6E-04
M5005_Spy0668	IgG-degrading protease	*mac*	1.82	4E-03
M5005_Spy0669	phage protein		1.84	3E-02
M5005_Spy0676	hypothetical protein		1.25	2E-02
M5005_Spy0686	3-hydroxy-3-methylglutaryl-coenzyme A reductase		-1.29	5E-04
M5005_Spy0687	hydroxymethylglutaryl-CoA synthase	*mvaS.1*	-0.95	1E-02

M5005_Spy0696	phosphopentomutase	*deoB*	-1.00	7E-03
M5005_Spy0698	purine nucleoside phosphorylase	*punA*	-1.01	2E-02
M5005_Spy0703	orotidine 5'-phosphate decarboxylase	*pyrF*	1.62	6E-03
M5005_Spy0704	orotate phosphoribosyltransferase	*pyrE*	2.61	1E-03
M5005_Spy0706	cystine-binding protein		1.67	5E-05
M5005_Spy0707	cystine transport system permease protein		1.07	8E-03
M5005_Spy0717	rhodanese-related sulfurtransferases		1.23	7E-04
M5005_Spy0718	hypothetical protein		1.11	7E-03
M5005_Spy0719	glutathione S-transferase		2.40	2E-06
M5005_Spy0721	hypothetical protein		-0.99	3E-02
M5005_Spy0732	NIF3-related protein		1.24	2E-02
M5005_Spy0735	dTDP-4-dehydrorhamnose 3,5-epimerase	*cpsFP*	-1.04	8E-04
M5005_Spy0737	mutator mutT protein/7,8-dihydro-8-oxoguanine-triphosphatase		-1.48	6E-05
M5005_Spy0738	hypothetical membrane spanning protein		-1.64	2E-06
M5005_Spy0739	tetratricopeptide repeat family protein		-1.19	1E-03
M5005_Spy0751	pyruvate dehydrogenase E1 component alpha subunit	*acoA*	1.95	5E-03
M5005_Spy0752	pyruvate dehydrogenase E1 component beta subunit	*acoB*	1.85	9E-04
M5005_Spy0753	dihydrolipoamide acetyltransferase component of pyruvate dehydrogenase complex	*acoC*	1.02	5E-03
M5005_Spy0774	nucleoside diphosphate kinase		2.02	5E-03
M5005_Spy0775	nucleoside diphosphate kinase		1.61	6E-04
M5005_Spy0780	PTS system, **Mannose/Fructose** family IIA component	*ptsA*	1.51	8E-03
M5005_Spy0781	PTS system, **Mannose/Fructose** family IIB component	*ptsB*	1.33	1E-03
M5005_Spy0783	PTS system, **Mannose/Fructose** family IID component	*ptsD*	1.12	1E-02
M5005_Spy0790	succinate-semialdehyde dehydrogenase	*gabD*	-1.23	1E-02
M5005_Spy0791	excinuclease ABC subunit C	*uvrC*	-1.98	5E-04
M5005_Spy0792	NAD(P)H-dependent quinone reductase		-2.64	1E-04
M5005_Spy0793	Xaa-His dipeptidase		-1.74	2E-02
M5005_Spy0794	tRNA (5-carboxymethylaminomethyl-2-thiouridylate) synthase	*thdF*	2.16	5E-06
M5005_Spy0795	LSU ribosomal protein L10P	*rplJ*	1.52	9E-03
M5005_Spy0796	LSU ribosomal protein L12P	*rplL*	2.69	1E-03
M5005_Spy0804	TCS-6 response regulator; Streptin biosynthesis	*srtR*	-0.94	4E-04
M5005_Spy0806	lantibiotic precursor	*srtA*	-1.24	2E-03
M5005_Spy0807	ABC transporter (ATP binding)-lantibiotic associated	*srtT*	-1.63	6E-06
M5005_Spy0808	lantibiotic transport ATP-binding protein	*srtF*	-0.95	3E-03
M5005_Spy0809	lantibiotic transport permease protein	*srtE*	-1.98	4E-05
M5005_Spy0810	lantibiotic transport permease protein	*srtG*	-1.62	1E-04
M5005_Spy0811	transcriptional regulator, Cro/CI family		-1.29	5E-02
M5005_Spy0820	folylpolyglutamate synthase/dihydrofolate synthase	*folC.1*	-2.04	6E-08
M5005_Spy0821	GTP cyclohydrolase I	*folE*	-1.20	4E-04
M5005_Spy0822	dihydropteroate synthase	*folP*	-1.11	2E-02
M5005_Spy0823	dihydroneopterin aldolase	*folQ*	-1.34	3E-04
M5005_Spy0824	2-amino-4-hydroxy-6-hydroxymethyldihydropteridine pyrophosphokinase	*folK*	-1.51	1E-04
M5005_Spy0825	UDP-N-acetylenolpyruvoylglucosamine reductase	*murB*	1.29	2E-04

M5005_Spy0826	spermidine/putrescine transport system ATP-binding protein	*potA*	2.58	4E-05
M5005_Spy0827	spermidine/putrescine transport system permease protein	*potB*	1.36	8E-04
M5005_Spy0828	spermidine/putrescine transport system permease protein	*potC*	1.52	4E-03
M5005_Spy0829	spermidine/putrescine-binding protein		2.29	3E-04
M5005_Spy0835	class B acid phosphatase		-1.26	2E-03
M5005_Spy0841	glutamine amidotransferase, class I		-1.06	3E-03
M5005_Spy0843	hypothetical cytosolic protein		-1.46	8E-05
M5005_Spy0845	ribose-phosphate pyrophosphokinase		-1.00	2E-03
M5005_Spy0852	short chain dehydrogenase		2.02	2E-06
M5005_Spy0853	short chain dehydrogenase		2.86	4E-07
M5005_Spy0858	xanthine phosphoribosyltransferase	*xpt*	4.52	2E-06
M5005_Spy0859	xanthine permease		4.41	2E-09
M5005_Spy0873	L-lactate dehydrogenase	*ldh*	1.60	1E-03
M5005_Spy0879	hypothetical protein		-2.03	3E-04
M5005_Spy0880	conserved membrane protein		-5.13	1E-06
M5005_Spy0881	hypothetical cytosolic protein		-5.63	7E-09
M5005_Spy0883	ribonuclease HII		1.06	4E-05
M5005_Spy0884	hypothetical protein	*smf*	2.51	1E-08
M5005_Spy0890	D-lactate dehydrogenase	*ddh*	1.57	7E-03
M5005_Spy0911	hypothetical protein		-1.01	8E-03
M5005_Spy0914	phage transcriptional repressor		1.30	5E-03
M5005_Spy0943	cytidine deaminase	*cdd*	1.42	5E-04
M5005_Spy0945	pantothenate kinase	*coaA*	-1.61	4E-04
M5005_Spy0965	ABC transporter permease protein		-1.09	6E-04
M5005_Spy0982	histidine-binding protein		1.14	1E-03
M5005_Spy1017	phage protein		3.15	6E-03
M5005_Spy1023	terminase large subunit		1.45	4E-02
M5005_Spy1028	phage protein		3.67	2E-03
M5005_Spy1046	phage protein		-1.28	5E-03
M5005_Spy1062	maltodextrose utilization protein	*malA*	-1.96	4E-02
M5005_Spy1063	ABC **Cyclomaltodextrin** permease protein	*malD*	-1.85	4E-02
M5005_Spy1064	ABC **Cyclomaltodextrin** permease protein	*malC*	-1.57	5E-02
M5005_Spy1065	alpha-amylase	*amyA*	-1.70	3E-02
M5005_Spy1066	neopullulanase/cyclomaltodextrinase/maltogenic alpha-amylase	*amyB*	-1.65	5E-02
M5005_Spy1069	esterase		-0.99	3E-03
M5005_Spy1070	protein precursor	*dltD*	-0.92	1E-02
M5005_Spy1072	protein DltB	*dltB*	-1.47	7E-05
M5005_Spy1073	D-alanine-activating enzyme	*dltA*	-1.43	1E-06
M5005_Spy1074	hypothetical protein		-1.79	3E-04
M5005_Spy1075	excinuclease ABC subunit B	*uvrB*	-1.57	6E-07
M5005_Spy1079	PTS system, **Cellobiose**-specific IIC component		-2.09	3E-03
M5005_Spy1080	hypothetical protein		-1.43	7E-03
M5005_Spy1081	PTS system, **Cellobiose**-specific IIA component		-1.76	8E-03
M5005_Spy1082	PTS system, **Cellobiose**-specific IIB component		-1.99	4E-03
M5005_Spy1083	PTS system, **Cellobiose**-specific IIA component		-1.83	6E-04
M5005_Spy1084	outer surface protein		-1.25	2E-02
M5005_Spy1085	beta-glucosidase	*bglA.2*	-1.61	7E-03
M5005_Spy1093	hypothetical protein		-3.31	2E-06
M5005_Spy1096	thioesterase superfamily protein		1.85	8E-03
M5005_Spy1097	phosphorylase, Pnp/Udp family		-0.97	2E-02
M5005_Spy1100	shikimate kinase	*aroK*	-1.23	5E-03
M5005_Spy1102	ribonuclease BN		-1.46	2E-04

M5005_Spy1106	protein G-related alpha 2M-binding protein	*grab*	1.90	7E-05
M5005_Spy1107	UDP-N-acetylglucosamine 1-carboxyvinyl transferase	*murZ*	1.34	7E-03
M5005_Spy1109	internalin protein	*inlA*	1.27	1E-03
M5005_Spy1114	hypothetical membrane spanning protein		-1.93	9E-04
M5005_Spy1117	ATP-dependent RNA helicase	*deaD2*	-1.00	4E-03
M5005_Spy1120	phosphoenolpyruvate-protein phosphotransferase	*pstI*	1.55	6E-04
M5005_Spy1122	glutaredoxin	*nrdH*	-1.87	5E-05
M5005_Spy1123	ribonucleoside-diphosphate reductase alpha chain	*nrdE.2*	-1.62	1E-03
M5005_Spy1124	ribonucleoside-diphosphate reductase beta chain	*nrdF.2*	-1.07	5E-03
M5005_Spy1125	chloride channel protein		-1.59	3E-06
M5005_Spy1131	transcriptional regulator, Cro/CI family		2.11	4E-05
M5005_Spy1132	alanyl-tRNA synthetase	*alaS*	2.46	9E-06
M5005_Spy1137	putative competence protein/transcription factor		1.62	2E-03
M5005_Spy1139	glucosamine-6-phosphate isomerase	*nagB*	1.24	3E-03
M5005_Spy1142	hypothetical protein		-2.17	9E-03
M5005_Spy1143	hypothetical protein		-2.16	2E-02
M5005_Spy1144	hypothetical protein		-2.26	1E-03
M5005_Spy1149	1-acyl-sn-glycerol-3-phosphate acyltransferase		1.23	3E-02
M5005_Spy1152	kup system potassium uptake protein		1.56	3E-04
M5005_Spy1167	lead, cadmium, zinc and mercury transporting ATPase; MtsR, PerR regulated	*pmtA*	1.43	2E-03
M5005_Spy1199	phage protein		2.11	3E-02
M5005_Spy1205	phage protein		1.51	1E-02
M5005_Spy1227	hypothetical protein		1.08	3E-03
M5005_Spy1229	arginine repressor	*argR1*	1.72	7E-03
M5005_Spy1230	hemolysin		1.18	5E-02
M5005_Spy1231	dimethylallyltransferase/geranyltranstransferase	*fps*	1.55	2E-02
M5005_Spy1232	exodeoxyribonuclease VII small subunit	*xseB*	2.43	1E-03
M5005_Spy1237	arginine transport ATP-binding protein	*artP*	2.19	8E-05
M5005_Spy1238	arginine transport system permease protein	*artQ*	1.52	1E-03
M5005_Spy1241	mutator protein (7,8-dihydro-8-oxoguanine-triphosphatase)	*mutT*	-0.95	2E-03
M5005_Spy1256	rhodanese-related sulfurtransferases		1.12	8E-03
M5005_Spy1257	glucokinase/Xylose repressor	*glcK*	1.98	3E-04
M5005_Spy1258	hypothetical cytosolic protein		2.50	3E-06
M5005_Spy1263	hypothetical protein		-1.65	3E-03
M5005_Spy1270	carbamate kinase	*arcC*	-3.88	3E-05
M5005_Spy1271	Xaa-His dipeptidase		-4.11	6E-09
M5005_Spy1272	arginine/ornithine antiporter		-4.09	2E-07
M5005_Spy1273	ornithine carbamoyltransferase	*arcB*	-4.04	4E-06
M5005_Spy1274	acetyltransferase		-3.79	2E-07
M5005_Spy1275	arginine deiminase	*arcA*	-3.22	2E-04
M5005_Spy1281	TCS-9 response regulator yesN-like	*tcs9R*	-0.99	2E-02
M5005_Spy1297	phospho-2-dehydro-3-deoxyheptonate aldolase		1.91	7E-04
M5005_Spy1298	3-dehydroquinate synthase	*aroB*	2.02	8E-03
M5005_Spy1299	hypothetical protein		-1.91	4E-02
M5005_Spy1305	TCS-10 response regulator; TCS-X; YesN-like	*trxR*	-1.20	1E-02
M5005_Spy1308	ABC unknown sugar-binding protein		-1.79	9E-03
M5005_Spy1309	ABC unknown sugar permease protein		-2.78	7E-03
M5005_Spy1310	ABC unknown sugar permease protein		-2.82	6E-03
M5005_Spy1311	glucokinase		-1.95	2E-03
M5005_Spy1312	hypothetical protein		-3.23	3E-03
M5005_Spy1319	tRNA (uracil-5-)-methyltransferase		1.24	2E-02

M5005_Spy1323	transposase		-1.12	2E-02
M5005_Spy1329	cysteine synthase	*cysM*	3.54	1E-05
M5005_Spy1334	transporter	*yvqF*	-1.13	1E-02
M5005_Spy1340	DNA-directed RNA polymerase omega chain		2.04	2E-05
M5005_Spy1341	guanylate kinase	*gmk*	1.23	4E-03
M5005_Spy1342	hydrolase, HAD superfamily		1.30	4E-04
M5005_Spy1344	acetyl-CoA acetyltransferase	*atoB*	-1.16	5E-04
M5005_Spy1346	acetate CoA-transferase beta subunit	*atoA*	-1.44	7E-04
M5005_Spy1347	D-beta-hydroxybutyrate dehydrogenase		-1.76	1E-03
M5005_Spy1348	D-beta-hydroxybutyrate permease		-2.18	6E-08
M5005_Spy1354	recombination protein	*recU*	-1.14	3E-03
M5005_Spy1356	aminopeptidase C	*pepC*	-1.25	2E-02
M5005_Spy1359	amino acid permease		1.36	2E-04
M5005_Spy1362	transporter		1.78	4E-06
M5005_Spy1363	amino acid ABC transporter permease protein		1.17	2E-03
M5005_Spy1374	hypothetical protein		1.17	1E-02
M5005_Spy1376	transaldolase		-2.82	1E-04
M5005_Spy1377	trans-acting positive regulator		-2.26	3E-03
M5005_Spy1378	NADH peroxidase		-1.90	2E-02
M5005_Spy1379	glycerol uptake facilitator protein	*glpF*	-4.84	1E-05
M5005_Spy1380	alpha-glycerophosphate oxidase	*glpO*	-4.44	2E-07
M5005_Spy1381	glycerol kinase	*glpK*	-4.13	5E-08
M5005_Spy1391	degV family protein		-1.76	8E-06
M5005_Spy1393	hydrolase, HAD superfamily		-1.17	5E-05
M5005_Spy1395	tagatose-bisphosphate aldolase; SpeB regulator	*lacD.1*	-2.93	3E-04
M5005_Spy1396	tagatose-6-phosphate kinase	*lacC.1*	-2.29	6E-03
M5005_Spy1397	galactose-6-phosphate isomerase lacB subunit	*lacB.1*	-2.54	8E-04
M5005_Spy1398	galactose-6-phosphate isomerase lacA subunit	*lacA.1*	-2.63	4E-03
M5005_Spy1399	PTS system, **Galactose**-specific IIC component		-2.16	2E-02
M5005_Spy1408	ribosome-binding factor A	*rbfA*	1.44	2E-03
M5005_Spy1409	bacterial protein translation initiation factor 2	*infB*	1.69	2E-04
M5005_Spy1410	LSU ribosomal protein L7AE		1.47	3E-04
M5005_Spy1412	N utilization substance protein A	*nusA*	1.09	4E-03
M5005_Spy1420	phage protein		1.52	3E-03
M5005_Spy1428	phage protein		1.79	3E-03
M5005_Spy1432	phage protein		1.08	3E-02
M5005_Spy1441	phage terminase small subunit		1.52	1E-02
M5005_Spy1442	phage transcriptional activator		2.16	3E-04
M5005_Spy1444	adenine-specific methyltransferase		2.09	6E-03
M5005_Spy1445	phage protein		2.05	5E-03
M5005_Spy1447	phage-related DNA helicase		1.20	2E-02
M5005_Spy1448	hypothetical protein		1.52	4E-02
M5005_Spy1449	DNA primase		1.45	2E-02
M5005_Spy1450	phage-encoded DNA polymerase		1.45	8E-03
M5005_Spy1454	phage protein		1.20	1E-02
M5005_Spy1455	phage protein		1.60	2E-02
M5005_Spy1459	phage protein		2.58	1E-03
M5005_Spy1466	phage protein		1.29	3E-02
M5005_Spy1472	bis(5'-nucleosyl)-tetraphosphatase (asymmetrical)	*hit*	-1.32	2E-03
M5005_Spy1476	ATP/GTP hydrolase		1.16	2E-04
M5005_Spy1477	guanine-hypoxanthine permease		2.27	2E-05
M5005_Spy1479	PTS system, **Mannose**-specific IIAB; Primary Glucose PTS	*manL*	2.99	7E-05
M5005_Spy1480	PTS system, **Mannose**-specific IIC; Primary Glucose PTS	*manM*	3.36	1E-05

M5005_Spy1481	PTS system, **Mannose**-specific IID; Primary Glucose PTS	*manN*	2.78	4E-05
M5005_Spy1482	hypothetical cytosolic protein	*manO*	-0.97	3E-02
M5005_Spy1493	acyl carrier protein	*acpP*	1.39	2E-02
M5005_Spy1497	chaperone protein	*dnaJ*	1.27	3E-03
M5005_Spy1504	hypothetical membrane spanning protein		1.14	2E-02
M5005_Spy1508	glutamyl-tRNA(Gln) amidotransferase subunit C	*gatC*	1.04	4E-04
M5005_Spy1511	pyrazinamidase/nicotinamidase		1.01	3E-02
M5005_Spy1518	transporter		1.52	4E-03
M5005_Spy1519	ATP-dependent DNA helicase	*recG*	-1.14	5E-04
M5005_Spy1520	hypothetical protein		-4.09	7E-04
M5005_Spy1523	ABC transporter, ATP-binding protein	*siaF*	-1.09	7E-04
M5005_Spy1533	holo-[acyl-carrier protein] synthase	*alr*	1.13	8E-03
M5005_Spy1538	mannose-6-phosphate isomerase	*pmi*	-1.39	8E-04
M5005_Spy1539	fructokinase	*scrK*	1.59	5E-04
M5005_Spy1553	SSU ribosomal protein S18P	*rpsR*	1.93	1E-03
M5005_Spy1559	thioredoxin	*trx*	-1.26	3E-03
M5005_Spy1569	formate acetyltransferase	*pfl*	2.57	4E-07
M5005_Spy1574	universal stress protein family		8.91	2E-10
M5005_Spy1575	quinolone resistance protein	*norA*	9.81	1E-12
M5005_Spy1581	transcriptional regulator, MerR family		1.32	8E-03
M5005_Spy1582	DNA polymerase III, epsilon chain	*dnaQ*	1.26	7E-03
M5005_Spy1598	hypothetical protein		-1.38	2E-04
M5005_Spy1606	LSU ribosomal protein L28P	*rpmB*	2.04	2E-05
M5005_Spy1610	CTP synthase	*pyrG*	2.63	8E-06
M5005_Spy1611	DNA-directed RNA polymerase delta chain	*rpoE*	1.33	1E-04
M5005_Spy1627	ABC transporter permease protein	*salY*	-1.16	5E-04
M5005_Spy1644	DNA integration/recombination/invertion protein		2.93	3E-02
M5005_Spy1646	SSU ribosomal protein S9P	*rpsI*	1.92	4E-04
M5005_Spy1647	LSU ribosomal protein L13P	*rplM*	1.51	1E-03
M5005_Spy1661	transaldolase		-1.72	2E-03
M5005_Spy1662	putative transport protein		-1.70	1E-03
M5005_Spy1664	PTS system, **Mannitol**-specific IIA component		-1.17	3E-02
M5005_Spy1666	SSU ribosomal protein S15P		1.83	4E-04
M5005_Spy1667	hypothetical protein		1.58	1E-03
M5005_Spy1671	transcriptional regulator, MarR family		-1.00	8E-03
M5005_Spy1675	phosphatidate cytidylyltransferase	*cdsA*	-1.21	5E-04
M5005_Spy1691	exodeoxyribonuclease III		-1.63	2E-04
M5005_Spy1697	para-aminobenzoate synthetase component I/4-amino-4-deoxychorismate lyase		-1.09	3E-04
M5005_Spy1698	anthranilate synthase component II/para-aminobenzoate synthase glutamine amidotransferase component II	*trpG*	-1.45	1E-04
M5005_Spy1699	ATPase, AAA family		-1.70	9E-06
M5005_Spy1710	streptococcal histidine triad protein		3.01	4E-08
M5005_Spy1711	laminin binding lipoprotein; virulence assoc; AdcR- repressed	*lsp*	2.25	3E-05
M5005_Spy1714	cell surface protein; Mga-regulated	*fba*	2.51	2E-03
M5005_Spy1715	C5A peptidase precursor; Mga-regulated	*scpA*	2.20	3E-03
M5005_Spy1716	transposase		-1.66	9E-05
M5005_Spy1717	transposase		-2.47	1E-03
M5005_Spy1718	streptococcal inhibitor of complement; Mga-regulated	*sic1.0*	2.32	2E-03
M5005_Spy1721	hypothetical protein		-2.86	8E-03

M5005_Spy1730	hypothetical protein		-1.45	1E-04
M5005_Spy1736	hypothetical protein		-2.04	2E-02
M5005_Spy1737	Stand-alone transcriptional regulator Rgg, RopB	*rgg, ropB*	-1.43	1E-02
M5005_Spy1738	phage-associated deoxyribonuclease	*spd*	1.45	8E-04
M5005_Spy1741	glycerol dehydrogenase	*gldA*	-1.50	1E-04
M5005_Spy1742	transaldolase	*mipB*	-1.35	9E-04
M5005_Spy1743	formate acetyltransferase	*pflD*	-1.35	2E-05
M5005_Spy1745	PTS system, **Cellobiose**-specific IIB component	*celB*	-1.64	3E-02
M5005_Spy1747	sorbitol operon regulator		1.07	3E-02
M5005_Spy1765	cold shock protein	*csp*	1.67	8E-03
M5005_Spy1768	peroxiredoxin reductase (NAD(P)H)	*ahpC*	-1.07	6E-03
M5005_Spy1769	peroxiredoxin reductase (NAD(P)H)	*ahpF*	-1.24	6E-03
M5005_Spy1770	imidazolonepropionase	*hutI*	-5.24	1E-03
M5005_Spy1771	urocanate hydratase	*hutU*	-5.28	1E-02
M5005_Spy1773	formiminotetrahydrofolate cyclodeaminase		-7.25	3E-03
M5005_Spy1774	formate--tetrahydrofolate ligase	*fhs.2*	-7.37	2E-03
M5005_Spy1775	hypothetical cytosolic protein		-6.67	9E-03
M5005_Spy1776	amino acid permease		-7.13	4E-03
M5005_Spy1777	histidine ammonia-lyase	*hutH*	-6.67	2E-03
M5005_Spy1778	formiminoglutamase	*hutG*	-1.82	6E-03
M5005_Spy1780	SSU ribosomal protein S2P	*rpsB*	1.79	3E-04
M5005_Spy1781	protein translation elongation factor Ts		1.72	6E-04
M5005_Spy1782	neutral endopeptidase	*pepO*	-1.38	2E-04
M5005_Spy1783	trehalose-6-phosphate hydrolase	*treC*	-1.55	3E-04
M5005_Spy1784	PTS system, **Trehalose**-specific IIBC component	*treB*	-0.95	4E-02
M5005_Spy1786	transcriptional regulator, MarR family		2.52	3E-04
M5005_Spy1787	glyoxalase family protein		-1.05	8E-03
M5005_Spy1789	anaerobic ribonucleoside-triphosphate reductase activating protein	*nrdG*	1.44	2E-03
M5005_Spy1790	acetyltransferase		2.23	1E-02
M5005_Spy1792	hypothetical protein		1.02	2E-03
M5005_Spy1797	hypothetical cytosolic protein		-1.38	1E-04
M5005_Spy1798	suppressor of clpP & X SpxA homolog, allele 2	*spxA2*	-2.61	2E-04
M5005_Spy1802	holliday junction DNA helicase	*ruvA*	-1.19	6E-04
M5005_Spy1803	multidrug resistance protein B	*lmrP*	-1.12	2E-03
M5005_Spy1806	hypothetical cytosolic protein		1.45	4E-03
M5005_Spy1810	hypothetical membrane spanning protein		1.47	6E-03
M5005_Spy1811	hypothetical membrane spanning protein		2.32	1E-02
M5005_Spy1812	hypothetical membrane spanning protein		1.18	1E-02
M5005_Spy1815	LSU ribosomal protein L32P	*rpmF*	2.49	1E-03
M5005_Spy1821	hypothetical protein		1.20	3E-03
M5005_Spy1823	integral membrane protein		-1.65	1E-05
M5005_Spy1830	transcriptional regulator, TetR family		1.12	7E-03
M5005_Spy1831	SSU ribosomal protein S4P	*rpsD*	1.43	2E-03
M5005_Spy1834	hypothetical protein		1.49	1E-03
M5005_Spy1840	tRNA (5-methylaminomethyl-2-thiouridylate)-methyltransferase	*trmU*	1.32	2E-02
M5005_Spy1841	L-serine dehydratase	*sdhB*	3.26	1E-03
M5005_Spy1842	L-serine dehydratase	*sdhA*	3.50	6E-04
M5005_Spy1843	transglycosylase SLT domain family protein		2.76	1E-04
M5005_Spy1847	CDP-diacylglycerol--glycerol-3-phosphate 3-phosphatidyltransferase		-0.96	1E-02
M5005_Spy1849	zinc protease		-0.93	9E-03
M5005_Spy1851	hyaluronan synthase	*hasA*	1.30	8E-04

M5005_Spy1853	UTP--glucose-1-phosphate uridylyltransferase	*hasC*	1.24	2E-03
M5005_Spy1862	ABC transporter, permease protein		-1.42	5E-05
M5005_SpyT0010	tRNA-Pro		2.48	3E-04
M5005_SpyT0021	tRNA-Met		1.91	3E-02
M5005_SpyT0027	tRNA-Ile		2.55	4E-02
M5005_SpyT0028	tRNA-Ser		2.46	4E-04
M5005_SpyT0030	tRNA-Val		3.10	1E-02
M5005_SpyT0038	tRNA-Tyr		2.01	1E-02
M5005_SpyT0041	tRNA-Leu		1.71	5E-02
M5005_SpyT0048	tRNA-Arg		-1.43	2E-02
M5005_SpyT0051	tRNA-Ile		4.60	4E-05
M5005_SpyT0052	tRNA-Thr		2.23	2E-03
M5005_SpyT0059	tRNA-Leu		2.56	5E-03
M5005_SpyT0065	tRNA-Asp		1.45	4E-02
M5005_SpyT0067	tRNA-Arg		2.30	3E-03

[a] Both Limma & DESeq analyses exhibited Log_2 fold-change of $-0.95 \leq$ or ≥ 095. Limma data only shown.
[b] p value ≤ 0.05 in both analyses

Appendix 3: RNA-Seq results of a *mga* mutant in THY (WT/Δ*mga*)

Spy #	Annotation	Gene	Log_2 FC[a]	*p* value[b]
M5005_Spy0012	hypoxanthine-guanine phosphoribosyltransferase		-1.73	4.25E-05
M5005_Spy0013	cell division protein	*ftsH*	-1.02	1.01E-03
M5005_Spy0022	phosphoribosylaminoimidazole-succinocarboxamide synthase		-1.89	1.35E-05
M5005_Spy0040	alcohol dehydrogenase	*adhA*	1.21	5.05E-02
M5005_Spy0041	Na+ driven multidrug efflux pump		-1.16	1.49E-04
M5005_Spy0080	bis(5'-nucleosyl)-tetraphosphatase (asymmetrical)		-1.09	4.93E-03
M5005_Spy0113	transposase		1.47	6.56E-03
M5005_Spy0115	hypothetical protein		1.21	2.16E-05
M5005_Spy0117	transcriptional regulator, LysR family		2.06	1.59E-04
M5005_Spy0118	transcriptional regulator, LysR family		1.32	2.67E-02
M5005_Spy0143	hypothetical protein		2.54	5.24E-03
M5005_Spy0156	metal-dependent hydrolase		1.16	9.06E-03
M5005_Spy0189	hypothetical protein		1.17	4.56E-02
M5005_Spy0212	N-acetylmannosamine-6-phosphate 2-epimerase		1.39	1.97E-03
M5005_Spy0213	ABC **Sialic Acid**; N-acetylneuraminate-binding protein		1.44	2.84E-03
M5005_Spy0260	putative lipase		-1.04	1.51E-03
M5005_Spy0261	GTP-binding protein		-0.98	1.75E-02
M5005_Spy0297	transposase		2.55	1.50E-02
M5005_Spy0313	riboflavin transporter		-1.20	1.40E-04
M5005_Spy0318	pyruvate formate-lyase activating enzyme	*pflC*	1.52	8.23E-07
M5005_Spy0319	manganese-dependent inorganic pyrophosphatase	*ppaC*	0.90	4.79E-05
M5005_Spy0340	L-lactate oxidase	*lctO*	2.47	2.56E-04
M5005_Spy0341	lactocepin; SpyCEP; IL-8 protease	*spyCEP*	1.69	4.90E-06
M5005_Spy0344	permease		-1.10	5.79E-04
M5005_Spy0361	phosphoglycerate transporter protein		1.01	7.77E-03
M5005_Spy0383	D-alanyl-D-alanine carboxypeptidase		-1.04	2.70E-03
M5005_Spy0432	acetyl-CoA acetyltransferase		-1.22	4.35E-04
M5005_Spy0456	plasmid stabilization system antitoxin protein		-1.13	1.13E-02
M5005_Spy0469	hypothetical protein		-1.04	3.96E-03
M5005_Spy0471	hydrolase, HAD superfamily		-1.09	3.40E-03
M5005_Spy0494	hypothetical protein		-1.19	4.26E-03
M5005_Spy0518	oligohyaluronate lyase		1.58	2.76E-03
M5005_Spy0519	PTS system, **N-acetylgalactosamine**-specific IID	*agaD*	1.02	3.58E-02
M5005_Spy0520	PTS system, **N-acetylgalactosamine**-specific IIC	*agaS*	1.89	4.88E-05
M5005_Spy0522	unsaturated glucuronyl hydrolase		1.45	8.61E-04
M5005_Spy0549	chorismate mutase		-1.15	3.80E-04
M5005_Spy0585	epuA protein	*epuA*	-1.20	6.36E-03
M5005_Spy0660	fructose repressor	*fruR*	1.06	1.84E-02
M5005_Spy0667	exotoxin type C precursor		2.07	7.07E-03
M5005_Spy0668	IgG-degrading protease	*mac*	1.62	6.08E-04
M5005_Spy0714	hypothetical cytosolic protein		-1.24	2.46E-04
M5005_Spy0742	hypothetical protein		-1.97	1.97E-05

M5005_Spy0743	ABC transporter substrate-binding protein		-1.59	1.83E-05
M5005_Spy0744	hypothetical protein		-1.55	2.63E-04
M5005_Spy0745	ABC transporter permease protein		-1.26	3.78E-05
M5005_Spy0746	ABC transporter ATP-binding protein		-1.47	5.94E-07
M5005_Spy0780	PTS system, **Mannose/Fructose** family IIA component	*ptsA*	1.50	6.80E-04
M5005_Spy0781	PTS system, **Mannose/Fructose** family IIB component	*ptsB*	1.30	7.22E-05
M5005_Spy0782	PTS system, **Mannose/Fructose** family IIC component	*ptsC*	1.03	1.35E-03
M5005_Spy0786	iron(III)-binding protein		0.89	1.97E-03
M5005_Spy0793	Xaa-His dipeptidase		-1.04	4.77E-02
M5005_Spy0797	hypothetical protein		-2.85	1.38E-02
M5005_Spy0798	IFN-response binding factor 1		-3.53	2.18E-04
M5005_Spy0824	2-amino-4-hydroxy-6-hydroxymethyldihydropteridine pyrophosphokinase	*folK*	-1.00	3.83E-04
M5005_Spy0834	Zn-dependent alcohol dehydrogenases and related dehydrogenases		1.20	2.59E-03
M5005_Spy0853	short chain dehydrogenase		0.96	7.07E-03
M5005_Spy0858	xanthine phosphoribosyltransferase	*xpt*	-2.16	7.40E-05
M5005_Spy0859	xanthine permease		-1.81	6.01E-06
M5005_Spy0914	phage transcriptional repressor		-1.32	1.43E-04
M5005_Spy0920	UDP-N-acetylmuramoylpentapeptide-lysine N(6)-alanyltransferase		-1.70	8.64E-05
M5005_Spy0921	ABC transporter ATP-binding protein		-1.54	8.39E-05
M5005_Spy0932	luciferase-like monooxygenase		0.94	1.63E-03
M5005_Spy0933	probable NADH-dependent flavin oxidoreductase		1.00	6.80E-04
M5005_Spy0934	lipoate-protein ligase A		1.07	5.14E-04
M5005_Spy0978	hypothetical protein		-1.17	6.29E-04
M5005_Spy0985	phnA protein		-1.39	6.11E-07
M5005_Spy0986	glucosamine--fructose-6-phosphate aminotransferase (isomerizing)	*glmS*	-1.02	4.97E-04
M5005_Spy1059	ABC **Maltodextrin** permease protein	*malF*	0.94	2.81E-04
M5005_Spy1062	maltodextrose utilization protein	*malA*	1.49	4.11E-02
M5005_Spy1063	ABC **Cyclomaltodextrin** permease protein	*malD*	1.54	2.66E-02
M5005_Spy1064	ABC **Cyclomaltodextrin** permease protein	*malC*	1.38	2.28E-02
M5005_Spy1065	alpha-amylase	*amyA*	1.46	1.05E-02
M5005_Spy1066	neopullulanase/cyclomaltodextrinase/maltogenic alpha-amylase	*amyB*	1.74	7.27E-03
M5005_Spy1067	ABC **Cyclomaltodextrin** -binding protein	*malX*	2.52	1.94E-05
M5005_Spy1076	Putative Bicarbonate transporter	*glnH*	-1.31	2.34E-03
M5005_Spy1077	Putative Bicarbonate transporter	*glnQ.2*	-1.04	1.32E-02
M5005_Spy1082	PTS system, **Cellobiose**-specific IIB component		1.16	2.98E-02
M5005_Spy1083	PTS system, **Cellobiose**-specific IIA component		1.00	1.12E-02
M5005_Spy1084	outer surface protein		1.02	1.76E-02
M5005_Spy1085	beta-glucosidase	*bglA.2*	1.46	2.32E-03
M5005_Spy1093	hypothetical protein		1.15	1.48E-02
M5005_Spy1135	oxalate/formate antiporter		-1.16	3.94E-06
M5005_Spy1148	COME operon protein 1		-1.32	2.04E-02
M5005_Spy1165	dihydroorotate dehydrogenase	*pyrD*	-1.19	4.50E-03

M5005_Spy1194	phage protein		-1.11	3.50E-03
M5005_Spy1235	phosphoglucomutase		1.14	9.24E-05
M5005_Spy1261	radical SAM family enzyme		-1.25	2.85E-03
M5005_Spy1274	acetyltransferase		0.90	3.67E-02
M5005_Spy1275	arginine deiminase	*arcA*	1.75	4.22E-03
M5005_Spy1284	cytochrome c-type biogenesis protein	*ccdA*	1.31	2.87E-02
M5005_Spy1299	hypothetical protein		-2.03	3.07E-03
M5005_Spy1314	hyaluronoglucosaminidase	*hyl*	0.85	1.02E-02
M5005_Spy1315	transcriptional regulator, GntR family		1.18	3.01E-04
M5005_Spy1316	hypothetical protein		1.35	1.94E-03
M5005_Spy1317	alpha-mannosidase		1.16	5.48E-03
M5005_Spy1340	DNA-directed RNA polymerase omega chain		-1.50	1.71E-05
M5005_Spy1347	D-beta-hydroxybutyrate dehydrogenase		-0.97	9.06E-03
M5005_Spy1376	transaldolase		1.00	5.20E-02
M5005_Spy1377	trans-acting positive regulator		1.09	5.66E-02
M5005_Spy1387	aldo/keto reductase family		1.35	2.37E-05
M5005_Spy1397	galactose-6-phosphate isomerase lacB subunit	*lacB.1*	1.10	5.00E-02
M5005_Spy1477	guanine-hypoxanthine permease		-1.29	2.08E-04
M5005_Spy1488	biotin carboxyl carrier protein of acetyl-CoA	*accB*	-1.20	1.61E-02
M5005_Spy1489	3-oxoacyl-[acyl-carrier-protein] synthase	*fabF*	-1.11	2.96E-02
M5005_Spy1491	malonyl-CoA-[acyl-carrier-protein] transacylase	*fabD*	-1.21	4.90E-03
M5005_Spy1492	enoyl-[acyl-carrier protein] reductase (NADH)	*fabK*	-1.13	1.78E-02
M5005_Spy1569	formate acetyltransferase	*pfl*	1.06	8.55E-04
M5005_Spy1572	hypothetical membrane spanning protein		-1.36	1.35E-04
M5005_Spy1587	uridine phosphorylase	*udp*	1.12	3.29E-05
M5005_Spy1589	transcriptional regulator, Cathelicidin resistance	*crgR*	-1.13	7.00E-03
M5005_Spy1618	competence-specific sigma factor	*comX.2*	-2.41	2.07E-02
M5005_Spy1628	ABC transporter ATP-binding protein		1.40	3.76E-03
M5005_Spy1629	lantibiotic transport ATP-binding protein	*salX*	1.53	4.27E-03
M5005_Spy1630	serine (threonine) dehydratase	*salB*	2.13	6.52E-06
M5005_Spy1631	lantibiotic salivaricin A	*salA*	1.81	3.35E-02
M5005_Spy1632	6-phospho-beta-galactosidase	*lacG*	2.06	2.70E-03
M5005_Spy1633	PTS system, **Lactose**-specific IIBC component	*lacE*	2.12	4.34E-03
M5005_Spy1634	PTS system, **Lactose**-specific IIA component	*lacF*	2.06	1.43E-02
M5005_Spy1635	tagatose-bisphosphate aldolase	*lacD.2*	2.11	5.38E-03
M5005_Spy1636	tagatose-6-phosphate kinase	*lacC.2*	2.52	1.48E-03
M5005_Spy1637	galactose-6-phosphate isomerase lacB subunit	*lacB.2*	2.24	4.25E-03
M5005_Spy1638	galactose-6-phosphate isomerase lacA subunit	*lacA.2*	1.99	2.07E-02
M5005_Spy1644	DNA integration/recombination/invertion protein		-3.39	6.67E-04
M5005_Spy1649	hypothetical membrane spanning protein		-1.00	2.85E-03
M5005_Spy1682	multiple sugar transport ATP-binding protein	*msmK*	1.01	1.18E-03
M5005_Spy1702	mitogenic exotoxin Z	*smeZ*	-1.15	1.79E-04
M5005_Spy1705	dipeptide transport system permease protein	*dppB*	-1.09	6.94E-03
M5005_Spy1709	hypothetical protein		-1.47	7.88E-04
M5005_Spy1714	cell surface protein; Mga-regulated	*fba*	3.20	4.52E-06
M5005_Spy1715	C5A peptidase precursor; Mga-regulated	*scpA*	3.43	2.07E-07
M5005_Spy1717	transposase		-2.57	8.10E-06
M5005_Spy1718	streptococcal inhibitor of complement; Mga-	*sic1.0*	4.63	6.22E-09

	regulated			
M5005_Spy1719	M protein; Mga-regulated	*emm1*	9.08	8.56E-16
M5005_Spy1720	Multi-virulence gene regulator Mga	*mga*	4.60	2.76E-15
M5005_Spy1721	hypothetical protein		-2.78	6.42E-04
M5005_Spy1733	hypothetical protein		-2.54	4.21E-03
M5005_Spy1734	streptopain SpeB inhibitor	*speI*	-2.94	3.09E-03
M5005_Spy1735	streptococcal pyrogenic exotoxin B	*speB*	-2.73	2.05E-04
M5005_Spy1738	phage-associated deoxyribonuclease	*spd*	0.98	2.19E-03
M5005_Spy1750	hypothetical protein		-1.14	2.46E-04
M5005_Spy1758	probable dipeptidase B		1.09	5.27E-04
M5005_Spy1759	transcriptional regulator, MutR family		-1.10	6.96E-03
M5005_Spy1762	10 kDa chaperonin	*groES*	-1.04	7.08E-03
M5005_Spy1779	transcriptional regulator, LuxR family		1.55	1.46E-03
M5005_Spy1798	suppressor of clpP & X SpxA homolog, allele 2	*spxA2*	-1.86	2.98E-04
M5005_Spy1809	bacteriocin	*uviB*	-1.13	1.46E-03
M5005_Spy1812	hypothetical membrane spanning protein		-1.01	2.96E-03
M5005_Spy1828	phage infection protein		-1.71	6.65E-07
M5005_SpyT0022	tRNA-Met		0.91	2.97E-02
M5005_SpyT0041	tRNA-Leu		1.47	2.28E-02
M5005_SpyT0045	tRNA-Arg		-1.34	1.79E-03
M5005_SpyT0046	tRNA-Tyr		-2.63	1.33E-05
M5005_SpyT0049	tRNA-Thr		-2.53	3.53E-03
M5005_SpyT0059	tRNA-Leu		-1.59	1.03E-02
M5005_SpyT0065	tRNA-Asp		-1.43	2.75E-03

[a] Both Limma & DESeq analyses exhibited Log_2 fold-change of $-0.85 \leq$ or ≥ 0.85. Limma data only shown.
[b] p value ≤ 0.05 in both analyses

Appendix 4: RNA-Seq results of the *mga* mutant in C Media (WT/Δ*mga*)

Spy #	Annotation	Gene	Log_2 FC[a]	*p* value[b]
M5005_Spy0009	hypothetical protein		1.61	4.86E-02
M5005_Spy0030	phosphoribosylaminoimidazole carboxylase catalytic subunit	*purE*	1.09	2.88E-02
M5005_Spy0036	protein tyrosine phosphatase		-0.93	3.99E-03
M5005_Spy0091	comG operon protein 6	*comYD*	1.45	1.39E-02
M5005_Spy0106	transcriptional regulator; Regulator of protein F	*rofA*	-1.09	1.79E-04
M5005_Spy0113	transposase		1.32	1.76E-02
M5005_Spy0129	V-type ATP synthase subunit C	*ntpC*	-1.78	6.66E-03
M5005_Spy0130	V-type sodium ATP synthase subunit F	*ntpF*	-1.44	8.76E-03
M5005_Spy0131	V-type sodium ATP synthase subunit A	*ntpA*	-1.30	1.38E-02
M5005_Spy0132	V-type sodium ATP synthase subunit B	*ntpB*	-1.50	1.15E-02
M5005_Spy0133	V-type sodium ATP synthase subunit D	*ntpD*	-1.57	8.51E-03
M5005_Spy0136	adenylosuccinate synthetase	*purA*	-1.11	2.90E-05
M5005_Spy0139	NAD glycohydrolase	*nga/ spn*	1.04	4.63E-02
M5005_Spy0140	NAD glycohydrolase inhibitor	*ifs*	1.14	4.89E-02
M5005_Spy0141	streptolysin O	*slo*	1.12	2.94E-02
M5005_Spy0143	hypothetical protein		3.35	2.11E-03
M5005_Spy0157	glycine betaine transport ATP-binding protein	*opuAA*	1.23	4.82E-06
M5005_Spy0167	transposase		1.58	2.94E-02
M5005_Spy0183	hypothetical membrane associated protein		0.91	5.42E-03
M5005_Spy0202	hypothetical protein		-1.34	6.89E-05
M5005_Spy0247	D-alanyl-D-alanine carboxypeptidase		0.99	1.27E-04
M5005_Spy0257	transposase		-1.55	2.24E-02
M5005_Spy0279	hypothetical protein	*lemA*	-0.91	6.34E-04
M5005_Spy0296	acylphosphatase		-1.04	2.92E-04
M5005_Spy0385	67 kDa Myosin-crossreactive antigen		-1.32	6.12E-05
M5005_Spy0403	hypothetical protein		0.92	4.61E-02
M5005_Spy0422	NAD(P)H-dependent quinone reductase		-1.11	4.80E-05
M5005_Spy0423	Xaa-Pro dipeptidase	*pepQ*	-1.23	4.72E-05
M5005_Spy0477	hypothetical membrane spanning protein		-0.97	2.83E-04
M5005_Spy0478	hypothetical membrane spanning protein		-1.08	5.98E-04
M5005_Spy0486	hypothetical protein		-1.08	2.60E-04
M5005_Spy0522	unsaturated glucuronyl hydrolase		0.98	5.02E-03
M5005_Spy0527	4-hydroxy-2-oxoglutarate aldolase/2-dehydro-3-deoxyphosphogluconate aldolase	*kgdA*	1.03	5.35E-04
M5005_Spy0555	hypothetical cytosolic protein		-1.09	2.67E-04
M5005_Spy0560	transcriptional regulator		-0.94	1.36E-02
M5005_Spy0562	streptolysin S precursor	*sagA*	1.33	7.39E-04
M5005_Spy0564	streptolysin S biosynthesis protein	*sagC*	0.99	6.77E-04
M5005_Spy0565	streptolysin S biosynthesis protein	*sagD*	1.09	1.51E-04
M5005_Spy0567	streptolysin S biosynthesis protein	*sagF*	1.46	1.07E-04
M5005_Spy0568	streptolysin S export ATP-binding protein	*sagG*	1.30	2.50E-05
M5005_Spy0569	streptolysin S export transmembrane protein	*sagH*	1.09	3.36E-04
M5005_Spy0570	streptolysin S export transmembrane protein	*sagI*	1.42	1.25E-04
M5005_Spy0571	endonuclease/exonuclease/phosphatase family protein		1.21	1.05E-06
M5005_Spy0598	large-conductance mechanosensitive channel	*mscL*	-0.88	2.64E-02
M5005_Spy0615	pore forming protein	*ebsA*	-1.03	9.57E-04
M5005_Spy0616	ferredoxin		-2.15	1.10E-03

M5005_Spy0641	aspartate carbamoyltransferase	*pyrB*	0.90	5.65E-03
M5005_Spy0651	Heme-utilization cell surface protein	*hupY*	2.28	4.55E-09
M5005_Spy0652	Heme-oxygenase cytosolic protein	*hupZ*	2.16	5.20E-08
M5005_Spy0668	IgG-degrading protease	*mac*	1.37	1.24E-02
M5005_Spy0679	GTP pyrophosphokinase		-1.01	3.03E-04
M5005_Spy0703	orotidine 5'-phosphate decarboxylase	*pyrF*	1.15	9.19E-03
M5005_Spy0708	uracil-DNA glycosylase	*ung*	-0.86	4.09E-02
M5005_Spy0715	SSU ribosomal protein S1P		1.28	9.43E-04
M5005_Spy0720	putative exfoliative toxin		0.86	1.14E-02
M5005_Spy0753	dihydrolipoamide acetyltransferase component of pyruvate dehydrogenase complex	*acoC*	-0.85	8.15E-04
M5005_Spy0754	hypothetical protein		-1.50	5.80E-02
M5005_Spy0755	dihydrolipoamide dehydrogenase	*acoL*	-1.16	2.91E-05
M5005_Spy0794	tRNA (5-carboxymethylaminomethyl-2-thiouridylate) synthase	*thdF*	1.10	2.70E-04
M5005_Spy0820	folylpolyglutamate synthase/dihydrofolate synthase	*folC.1*	-0.98	1.95E-05
M5005_Spy0821	GTP cyclohydrolase I	*folE*	-0.95	6.65E-05
M5005_Spy0822	dihydropteroate synthase	*folP*	-0.92	5.66E-03
M5005_Spy0824	2-amino-4-hydroxy-6-hydroxymethyldihydropteridine pyrophosphokinase	*folK*	-0.87	1.42E-3
M5005_Spy0879	hypothetical protein		-1.52	7.92E-05
M5005_Spy0880	conserved membrane protein		-2.45	2.47E-05
M5005_Spy0881	hypothetical cytosolic protein		-1.93	5.24E-05
M5005_Spy0890	D-lactate dehydrogenase	*ddh*	1.38	5.23E-03
M5005_Spy0898	2-(5''-triphosphoribosyl)-3'-dephosphocoenzyme-A synthase		1.05	2.95E-02
M5005_Spy1005	phage protein		-1.47	2.11E-02
M5005_Spy1035	phage protein		1.44	6.01E-02
M5005_Spy1076	Putative Bicarbonate transporter	*glnH*	-0.85	3.52E-02
M5005_Spy1077	Putative Bicarbonate transporter	*glnQ.2*	-1.04	1.02E-02
M5005_Spy1093	hypothetical protein		-1.16	1.95E-04
M5005_Spy1145	superoxide dismutase	*sodA*	-0.93	5.73E-03
M5005_Spy1169	streptodornase	*spd3*	2.35	1.01E-06
M5005_Spy1170	hypothetical membrane associated protein		2.39	6.14E-07
M5005_Spy1199	phage protein		-1.87	9.59E-03
M5005_Spy1211	phage protein		1.19	4.99E-02
M5005_Spy1213	phage protein		1.28	7.48E-03
M5005_Spy1227	hypothetical protein		1.06	6.16E-05
M5005_Spy1228	DNA repair protein	*recN*	0.99	6.69E-03
M5005_Spy1229	arginine repressor	*argR1*	0.99	3.20E-02
M5005_Spy1230	hemolysin		0.92	3.15E-02
M5005_Spy1237	arginine transport ATP-binding protein	*artP*	1.36	7.76E-04
M5005_Spy1238	arginine transport system permease protein	*artQ*	1.40	2.08E-04
M5005_Spy1239	hypothetical cytosolic protein		1.14	1.48E-02
M5005_Spy1285	CRISPR2-associated protein Cas2	*cas2*	1.11	4.91E-04
M5005_Spy1287	CRISPR2-associated protein Cas4	*cas4*	1.08	2.63E-04
M5005_Spy1289	CRISPR2-associated protein Cds1	*cds1*	1.06	4.01E-07
M5005_Spy1297	phospho-2-dehydro-3-deoxyheptonate aldolase		1.18	3.04E-03
M5005_Spy1320	regulatory protein	*recX*	-1.55	1.37E-05
M5005_Spy1323	transposase		0.96	9.62E-03
M5005_Spy1325	ribosome-associated factor Y		-1.16	5.79E-04
M5005_Spy1369	hypothetical protein		1.18	2.44E-02

M5005_Spy1372	ABC transporter permease protein	*proB*	1.11	2.21E-04
M5005_Spy1382	hypothetical membrane associated protein		-1.40	1.37E-03
M5005_Spy1400	PTS system, Galactose-specific IIB component		-2.47	2.38E-03
M5005_Spy1401	PTS system, Galactose-specific IIA component		-2.66	2.89E-04
M5005_Spy1407	streptococcal secreted esterase Sse	*sse*	1.81	6.49E-04
M5005_Spy1415	phage-encoded streptodornase	*sdaD2*	2.22	8.65E-06
M5005_Spy1440	terminase large subunit		1.23	1.17E-02
M5005_Spy1441	phage terminase small subunit		1.12	3.41E-02
M5005_Spy1514	universal stress protein family		-1.22	1.16E-03
M5005_Spy1525	conserved hypothetical protein	*siaD*	1.09	2.23E-03
M5005_Spy1526	ABC transporter, ATP-binding protein, putative	*siaC*	1.33	7.51E-07
M5005_Spy1527	ABC transporter, ATP-binding protein, putative	*siaB*	1.39	1.21E-03
M5005_Spy1528	iron compound ABC transporter, ATP-binding protein	*siaA*	1.44	9.46E-04
M5005_Spy1529	iron compound ABC transporter, permease protein	*shp*	1.22	6.56E-05
M5005_Spy1530	iron compound ABC transporter, substrate-binding protein	*shr*	1.27	4.35E-07
M5005_Spy1559	thioredoxin	*trx*	-0.96	9.74E-04
M5005_Spy1590	SSU ribosomal protein S14P	*rpsN2*	-1.22	7.50E-03
M5005_Spy1622	type I restriction-modification system specificity subunit	*hsdS*	1.20	9.40E-07
M5005_Spy1714	cell surface protein; Mga-regulated	*fba*	4.79	8.72E-08
M5005_Spy1715	C5A peptidase precursor; Mga-regulated	*scpA*	5.04	2.90E-10
M5005_Spy1716	transposase		-1.07	1.26E-04
M5005_Spy1718	streptococcal inhibitor of complement; Mga-regulated	*sic1.0*	5.53	2.06E-09
M5005_Spy1719	M protein; Mga-regulated	*emm1*	10.06	8.06E-17
M5005_Spy1720	Multi-virulence gene regulator Mga	*mga*	5.84	4.18E-17
M5005_Spy1721	hypothetical protein		-5.00	2.44E-08
M5005_Spy1727	ABC transporter ATP-binding protein		0.99	1.00E-02
M5005_Spy1731	gene regulated by Mga;; Mga-regulated	*grm*	3.85	7.69E-05
M5005_Spy1733	hypothetical protein		1.45	5.22E-02
M5005_Spy1736	hypothetical protein		-1.31	2.98E-02
M5005_Spy1738	phage-associated deoxyribonuclease	*spd*	1.03	3.17E-04
M5005_Spy1739	hypothetical protein		-3.02	3.24E-05
M5005_Spy1754	translation initiation inhibitor		-0.98	4.57E-02
M5005_Spy1768	peroxiredoxin reductase (NAD(P)H)	*ahpC*	-0.92	1.06E-03
M5005_Spy1774	formate--tetrahydrofolate ligase	*fhs.2*	1.84	2.49E-02
M5005_Spy1775	hypothetical cytosolic protein		1.59	5.02E-02
M5005_Spy1776	amino acid permease		1.83	2.75E-02
M5005_Spy1777	histidine ammonia-lyase	*hutH*	2.16	9.13E-03
M5005_Spy1778	formiminoglutamase	*hutG*	1.01	9.55E-03
M5005_Spy1798	suppressor of clpP & X SpxA homolog, allele 2	*spxA2*	-1.07	1.68E-02
M5005_Spy1823	integral membrane protein		-0.88	3.46E-04
M5005_Spy1843	transglycosylase SLT domain family protein		1.33	4.54E-03
M5005_SpyT0010	tRNA-Pro		1.33	4.54E-03
M5005_SpyT0021	tRNA-Met		1.07	1.87E-02
M5005_SpyT0028	tRNA-Ser		1.68	2.07E-02
M5005_SpyT0038	tRNA-Tyr		1.43	4.17E-03
M5005_SpyT0039	tRNA-His		2.21	7.82E-04
M5005_SpyT0041	tRNA-Leu		1.86	3.70E-03

M5005_SpyT0046	tRNA-Tyr	2.38	2.22E-03
M5005_SpyT0047	tRNA-Gln	1.48	1.47E-02
M5005_SpyT0051	tRNA-Ile	2.04	4.62E-04
M5005_SpyT0052	tRNA-Thr	1.78	3.51E-02
M5005_SpyT0058	tRNA-Ser	1.29	3.34E-02
M5005_SpyT0064	tRNA-Cys	1.37	8.30E-03
M5005_SpyT0065	tRNA-Asn	1.27	3.24E-02

[a] Both Limma & DESeq analyses exhibited Log_2 fold-change of $-0.85 \leq$ or ≥ 0.85. Limma data only shown.
[b] p value ≤ 0.05 in both analyses

Appendix 5: Differential Expression of intergenic region in WT 5448 in CDM-glucose compared to CDM-fructose at late log (CDM-glu/CDM-fru)

Spy #[a]	Gene Name	Log_2 FC[b]	*p* value[c]
M5005_Spy0040	*adhA*	-1.89	1.04E-02
M5005_Spy0154		2.43	2.10E-02
M5005_Spy0170	*nadC*	-1.34	7.19E-04
M5005_Spy0207	*fasX*	-1.37	1.52E-05
M5005_Spy0216		1.40	7.63E-03
M5005_Spy0660	*fruR*	-1.12	8.86E-03
M5005_Spy0661	*fruB*	-2.47	9.49E-03
M5005_Spy0662	*fruA*	-2.60	2.21E-03
M5005_Spy0663	*mur1.1*	-2.37	1.55E-07
M5005_Spy0968		1.37	1.32E-02
M5005_Spy1005		2.69	1.57E-03
M5005_Spy1010		3.45	9.78E-04
M5005_Spy1011		2.14	2.83E-02
M5005_Spy1012		2.91	1.03E-03
M5005_Spy1017		2.16	1.30E-02
M5005_Spy1021		2.76	2.08E-03
M5005_Spy1022		2.89	2.43E-03
M5005_Spy1026		1.69	2.89E-03
M5005_Spy1027		2.06	9.45E-04
M5005_Spy1029		2.32	1.66E-03
M5005_Spy1030		3.04	5.72E-04
M5005_Spy1031		2.04	7.79E-03
M5005_Spy1032		3.56	1.47E-05
M5005_Spy1034		2.26	2.07E-05
M5005_Spy1035		2.73	4.20E-05
M5005_Spy1036	*ssb2*	2.39	1.01E-03
M5005_Spy1038		3.00	1.46E-06
M5005_Spy1039		2.98	2.39E-06
M5005_Spy1040		2.73	1.18E-04
M5005_Spy1041		2.16	3.14E-03
M5005_Spy1042		2.30	9.72E-05
M5005_Spy1043		2.59	1.26E-05
M5005_Spy1044		2.27	2.76E-06
M5005_Spy1049		1.46	4.41E-04
M5005_Spy1139	*nagB*	1.33	1.22E-03
M5005_Spy1154	*deaD*	-1.18	1.77E-02
M5005_Spy1185		2.59	6.44E-03
M5005_Spy1186		2.20	1.15E-02
M5005_Spy1188		3.15	1.48E-03
M5005_Spy1198		1.80	8.37E-05
M5005_Spy1200		2.66	2.42E-04
M5005_Spy1201		3.02	2.66E-04
M5005_Spy1202		2.03	4.92E-04
M5005_Spy1203		2.11	6.27E-03
M5005_Spy1204		2.48	1.23E-04
M5005_Spy1205		2.38	3.53E-03
M5005_Spy1206		3.04	5.57E-05
M5005_Spy1207		2.54	3.67E-04
M5005_Spy1209		2.67	1.57E-03
M5005_Spy1210		2.75	1.52E-04
M5005_Spy1211		3.56	3.53E-05
M5005_Spy1213		1.55	1.85E-02

M5005_Spy1217		1.76	2.81E-04
M5005_Spy1218		1.72	4.43E-04
M5005_Spy1307		-1.87	2.77E-02
M5005_Spy1415	*sdaD2*	1.93	3.73E-04
M5005_Spy1416		1.61	8.70E-03
M5005_Spy1417		2.53	6.97E-07
M5005_Spy1418		2.79	9.09E-06
M5005_Spy1419		3.41	9.83E-07
M5005_Spy1420		3.06	1.37E-06
M5005_Spy1421		2.50	2.19E-06
M5005_Spy1422		2.02	4.90E-05
M5005_Spy1423		2.45	4.87E-07
M5005_Spy1424		2.27	5.27E-04
M5005_Spy1425		2.80	7.23E-07
M5005_Spy1426		2.84	5.46E-06
M5005_Spy1427		2.68	2.51E-07
M5005_Spy1428		3.11	4.14E-07
M5005_Spy1429		2.66	1.47E-07
M5005_Spy1430		2.77	1.05E-07
M5005_Spy1431		2.53	8.65E-09
M5005_Spy1432		2.59	2.36E-08
M5005_Spy1433		2.88	1.96E-06
M5005_Spy1434		2.77	1.63E-08
M5005_Spy1435		2.68	2.04E-08
M5005_Spy1436		2.89	1.23E-08
M5005_Spy1437		2.66	3.40E-09
M5005_Spy1438		2.35	8.25E-04
M5005_Spy1439		2.61	2.39E-07
M5005_Spy1440		2.13	4.07E-05
M5005_Spy1441		2.34	5.74E-05
M5005_Spy1442		2.44	1.90E-03
M5005_Spy1444		1.79	3.88E-03
M5005_Spy1446		1.98	1.05E-02
M5005_Spy1447		2.19	8.63E-04
M5005_Spy1448		3.41	7.60E-08
M5005_Spy1449		2.17	2.44E-04
M5005_Spy1451		2.25	2.46E-07
M5005_Spy1452		3.15	9.52E-09
M5005_Spy1453		1.92	1.57E-03
M5005_Spy1454		2.52	9.63E-04
M5005_Spy1455		2.44	4.07E-05
M5005_Spy1456		2.63	9.01E-05
M5005_Spy1457		2.62	3.06E-05
M5005_Spy1458		1.90	2.04E-04
M5005_Spy1459		2.19	2.38E-05
M5005_Spy1460		1.71	1.21E-03
M5005_Spy1461		2.49	1.48E-06
M5005_Spy1462		1.72	7.27E-05
M5005_Spy1575	*norA*	2.69	2.53E-06
M5005_Spy1746		2.77	1.08E-02
M5005_Spy1779		2.50	1.33E-02
M5005_Spy1838	*gidA*	1.63	9.30E-04

[a] Analysis conducted on intergenic region 5' to spy number listed
[b] Both Limma & DESeq analyses exhibited Log_2 fold-change of $-1 \leq$ or ≥ 1. Limma data only shown.
[c] p value ≤ 0.05 in both analyses

Appendix 6: Differential Expression WT 5448 in C media compared to THY at late log (WT C media/ WT THY)

Spy #[a]	Gene Name	Log_2 FC[b]	*p* value[c]
M5005_Spy0012		1.60	9.79E-04
M5005_Spy0013	*ftsH*	1.16	7.43E-03
M5005_Spy0014		1.93	3.97E-04
M5005_Spy0022		2.37	3.43E-03
M5005_Spy0023		-2.32	1.57E-02
M5005_Spy0024	*purF*	-1.38	1.69E-02
M5005_Spy0025	*purM*	-2.15	1.78E-02
M5005_Spy0026	*purN*	-2.23	1.63E-03
M5005_Spy0027		-2.94	2.05E-06
M5005_Spy0028		-2.35	4.11E-04
M5005_Spy0029	*purD*	-2.99	2.13E-05
M5005_Spy0030	*purE*	-2.31	2.13E-03
M5005_Spy0031	*purK*	-1.92	9.53E-03
M5005_Spy0032		-1.55	1.02E-03
M5005_Spy0033	*purB*	-1.90	8.23E-04
M5005_Spy0034		-1.67	9.58E-05
M5005_Spy0039	*adh2*	1.44	4.03E-02
M5005_Spy0040	*adhA*	2.34	3.06E-03
M5005_Spy0041		2.56	3.01E-03
M5005_Spy0043	*rpsJ*	-1.77	9.07E-04
M5005_Spy0069	*rpsK*	-2.10	2.20E-04
M5005_Spy0070	*rpoA*	-1.60	1.33E-03
M5005_Spy0071	*rplQ*	-2.04	6.24E-05
M5005_Spy0073		-1.92	9.29E-05
M5005_Spy0074		-1.19	1.39E-02
M5005_Spy0076		-1.79	2.25E-03
M5005_Spy0085		-2.66	4.03E-06
M5005_Spy0089		-2.03	1.48E-02
M5005_Spy0095		-1.86	5.30E-03
M5005_Spy0103		-1.45	9.43E-03
M5005_Spy0104		-1.32	4.15E-03
M5005_Spy0108		-1.85	4.13E-03
M5005_Spy0109		-2.82	7.35E-06
M5005_Spy0110	*eftLSL.B*	-2.14	1.38E-03
M5005_Spy0111		-3.26	7.33E-07
M5005_Spy0112		-2.22	2.49E-05
M5005_Spy0114		-1.44	3.84E-03
M5005_Spy0116	*atoE*	3.12	1.62E-03
M5005_Spy0144		-3.16	2.16E-03
M5005_Spy0149		2.94	3.44E-04
M5005_Spy0150		2.80	3.39E-03
M5005_Spy0151		3.37	5.66E-03
M5005_Spy0152		3.08	5.45E-03
M5005_Spy0153	*araD*	3.55	8.84E-04
M5005_Spy0154		3.88	4.43E-04
M5005_Spy0157	*opuAA*	-1.18	6.43E-03
M5005_Spy0158	*opuABC*	-2.72	3.10E-07
M5005_Spy0159	*polA*	-3.04	2.17E-07
M5005_Spy0165		-1.88	9.89E-03
M5005_Spy0166		-2.21	5.74E-03
M5005_Spy0175	*tgt*	-1.86	1.47E-04

M5005_Spy0185	*pgi*	-1.51	6.33E-03
M5005_Spy0195		1.61	1.91E-03
M5005_Spy0211	*rpmH*	-2.19	4.38E-05
M5005_Spy0214		1.86	6.86E-03
M5005_Spy0215		2.04	1.13E-02
M5005_Spy0217	*nanH*	1.77	1.34E-02
M5005_Spy0222	*ksgA*	-1.69	3.43E-04
M5005_Spy0226		1.37	1.36E-02
M5005_Spy0230	*rpsL*	-1.68	3.92E-04
M5005_Spy0231	*rpsG*	-2.29	5.41E-05
M5005_Spy0232	*fus*	-1.53	1.19E-03
M5005_Spy0236		-1.74	4.40E-05
M5005_Spy0273		-1.70	2.76E-02
M5005_Spy0274	*braB*	-2.76	7.25E-03
M5005_Spy0275		-1.61	4.18E-02
M5005_Spy0285	*dnaB*	1.55	3.23E-04
M5005_Spy0296		1.43	1.39E-03
M5005_Spy0306		1.06	8.70E-03
M5005_Spy0307		1.73	3.80E-04
M5005_Spy0308		1.74	1.96E-03
M5005_Spy0324	*fhuA*	-2.38	2.53E-04
M5005_Spy0334		1.03	1.26E-02
M5005_Spy0337	*cutC*	1.81	4.46E-02
M5005_Spy0341		1.50	3.55E-02
M5005_Spy0348	*nrdI*	3.18	8.13E-04
M5005_Spy0350		2.49	1.14E-03
M5005_Spy0353		2.96	1.57E-02
M5005_Spy0362	*gcaD*	1.25	1.85E-02
M5005_Spy0375	*rplA*	-1.96	1.54E-03
M5005_Spy0386	*phoH*	1.68	7.76E-04
M5005_Spy0399		-2.46	1.38E-03
M5005_Spy0423	*pepQ*	1.72	4.77E-05
M5005_Spy0426		1.10	3.14E-02
M5005_Spy0436	*vicK*	1.61	1.14E-03
M5005_Spy0450	*mefE*	2.00	1.91E-03
M5005_Spy0466		1.39	4.43E-05
M5005_Spy0467		1.81	3.04E-04
M5005_Spy0468		1.32	3.05E-03
M5005_Spy0469		1.74	4.98E-04
M5005_Spy0471		1.38	5.57E-03
M5005_Spy0476	*bglA*	-2.17	1.66E-05
M5005_Spy0477		1.66	7.95E-04
M5005_Spy0478		1.50	8.82E-03
M5005_Spy0479		1.55	2.92E-04
M5005_Spy0483		2.60	2.78E-06
M5005_Spy0484	*ptsK*	2.01	5.40E-04
M5005_Spy0486		1.16	1.43E-02
M5005_Spy0492		1.21	3.94E-02
M5005_Spy0495	*lysS*	1.23	3.82E-04
M5005_Spy0497		-1.50	4.14E-03
M5005_Spy0504	*pepF*	2.64	2.76E-05
M5005_Spy0511	*murM*	1.11	1.00E-02
M5005_Spy0529		1.03	2.63E-03
M5005_Spy0540		1.38	1.23E-02
M5005_Spy0542	*pepD*	1.34	1.95E-02
M5005_Spy0546	*rpmE*	-1.67	6.33E-04

M5005_Spy0551	*rplS*	-1.78	1.80E-02
M5005_Spy0562	*sagA*	3.54	1.60E-06
M5005_Spy0563	*sagB*	2.55	4.95E-05
M5005_Spy0564	*sagC*	1.27	3.80E-02
M5005_Spy0565	*sagD*	1.69	3.87E-03
M5005_Spy0566	*sagE*	3.00	1.33E-07
M5005_Spy0567	*sagF*	3.02	4.68E-04
M5005_Spy0569	*sagH*	2.03	4.59E-04
M5005_Spy0570	*sagI*	1.50	2.01E-02
M5005_Spy0576	*atpB*	-1.50	1.16E-03
M5005_Spy0577	*atpF*	-1.72	1.72E-03
M5005_Spy0596		1.38	4.80E-04
M5005_Spy0599	*dnaG*	-1.73	3.79E-03
M5005_Spy0600	*rpoD*	-1.61	4.16E-05
M5005_Spy0615	*ebsA*	1.90	1.76E-05
M5005_Spy0616		1.28	1.02E-02
M5005_Spy0619	*infC*	-2.02	5.05E-05
M5005_Spy0620	*rpl36*	-2.05	1.01E-03
M5005_Spy0621	*rplT*	-2.61	6.60E-07
M5005_Spy0622		-2.12	1.04E-03
M5005_Spy0629		1.21	4.33E-02
M5005_Spy0633	*rplU*	-1.71	4.84E-04
M5005_Spy0635	*rpmA*	-1.63	2.38E-04
M5005_Spy0637	*lsp*	2.19	4.52E-05
M5005_Spy0640	*pyrP*	-1.68	3.28E-03
M5005_Spy0641	*pyrB*	-1.48	2.08E-03
M5005_Spy0642	*carA*	-2.79	2.14E-06
M5005_Spy0643	*carB*	-1.95	1.30E-03
M5005_Spy0645		-2.20	1.26E-04
M5005_Spy0648	*rpsP*	-1.47	2.92E-02
M5005_Spy0649		-1.57	1.99E-03
M5005_Spy0653	*czcD*	1.19	3.06E-02
M5005_Spy0664	*mur1.2*	1.29	4.83E-03
M5005_Spy0677	*fms*	-1.20	2.26E-02
M5005_Spy0693		1.61	1.64E-04
M5005_Spy0694	*clpL*	1.21	1.75E-02
M5005_Spy0696	*deoB*	1.14	2.01E-03
M5005_Spy0699	*deoD2*	1.06	3.13E-02
M5005_Spy0700	*cpsX*	1.13	2.93E-02
M5005_Spy0702		1.50	1.24E-03
M5005_Spy0704	*pyrE*	-2.32	3.00E-03
M5005_Spy0705	*amiC*	-1.72	2.59E-02
M5005_Spy0717		-1.16	1.78E-02
M5005_Spy0719		-1.86	2.64E-05
M5005_Spy0722	*miaA*	1.29	1.14E-02
M5005_Spy0731		-1.34	2.70E-02
M5005_Spy0735	*cpsFP*	1.29	1.91E-02
M5005_Spy0738		1.68	2.83E-05
M5005_Spy0740	*fbp*	1.62	4.48E-02
M5005_Spy0751	*acoA*	-2.37	9.72E-03
M5005_Spy0752	*acoB*	-2.16	4.65E-03
M5005_Spy0753	*acoC*	-1.62	1.02E-03
M5005_Spy0759		1.82	1.30E-03
M5005_Spy0774		-2.54	2.06E-04
M5005_Spy0776	*lepA*	-1.45	1.55E-03
M5005_Spy0779		-1.32	4.05E-03

M5005_Spy0780		-1.80	4.63E-03
M5005_Spy0781	*ptsB*	-1.81	1.72E-03
M5005_Spy0790	*gabD*	1.39	6.85E-03
M5005_Spy0791	*uvrC*	2.22	4.62E-05
M5005_Spy0792		2.35	1.70E-04
M5005_Spy0793		1.70	4.05E-02
M5005_Spy0797		-1.87	1.46E-03
M5005_Spy0805	*srtK*	1.52	1.02E-03
M5005_Spy0807	*srtT*	1.59	2.56E-03
M5005_Spy0808	*srtF*	2.10	4.34E-06
M5005_Spy0810	*srtG*	1.92	4.03E-03
M5005_Spy0811		2.73	1.15E-03
M5005_Spy0820	*folC.1*	1.85	3.38E-05
M5005_Spy0823	*folQ*	1.70	6.80E-05
M5005_Spy0824	*folK*	1.34	1.24E-03
M5005_Spy0826	*potA*	-2.22	7.37E-05
M5005_Spy0828	*potC*	-1.65	1.39E-03
M5005_Spy0829		-1.51	2.77E-02
M5005_Spy0835		3.90	1.80E-03
M5005_Spy0841		1.94	2.36E-04
M5005_Spy0845		1.22	3.25E-03
M5005_Spy0853		-2.33	1.66E-04
M5005_Spy0854		-2.98	4.45E-09
M5005_Spy0858	*xpt*	-2.06	1.65E-03
M5005_Spy0859		-3.79	5.06E-05
M5005_Spy0869		1.19	7.93E-03
M5005_Spy0879		-1.55	9.12E-03
M5005_Spy0880		2.30	1.40E-04
M5005_Spy0881		3.93	4.58E-06
M5005_Spy0884	*smf*	-2.83	7.27E-05
M5005_Spy0893	*gid*	1.19	1.28E-02
M5005_Spy0899	*citG*	1.69	2.19E-02
M5005_Spy0900		2.42	3.99E-02
M5005_Spy0913	*xerD*	1.68	3.30E-02
M5005_Spy0917		1.50	2.14E-03
M5005_Spy0918		1.66	5.90E-03
M5005_Spy0931		1.84	4.41E-04
M5005_Spy0940		1.39	1.27E-03
M5005_Spy0948	*ciaR*	1.81	3.94E-03
M5005_Spy0950	*phoU*	1.44	9.02E-04
M5005_Spy0956		1.12	2.21E-02
M5005_Spy0972		-1.65	7.94E-03
M5005_Spy0978		-1.49	2.28E-02
M5005_Spy0983		-1.47	1.76E-03
M5005_Spy0984		-1.83	2.71E-05
M5005_Spy0994		2.06	1.18E-03
M5005_Spy1001		-2.27	7.41E-03
M5005_Spy1035		-1.91	8.86E-03
M5005_Spy1038		-1.95	6.46E-03
M5005_Spy1057	*malR*	1.76	3.10E-02
M5005_Spy1060	*malG*	2.63	3.73E-03
M5005_Spy1064	*malC*	2.95	1.33E-03
M5005_Spy1067	*malX*	1.52	2.53E-02
M5005_Spy1071		1.86	1.26E-02
M5005_Spy1072		1.59	2.05E-04
M5005_Spy1073	*dltA*	1.83	5.61E-04

M5005_Spy1074		1.06	7.55E-03
M5005_Spy1075	*uvrB*	2.33	6.30E-06
M5005_Spy1076	*glnH*	1.31	1.98E-02
M5005_Spy1080		1.28	2.50E-02
M5005_Spy1081		2.00	1.09E-02
M5005_Spy1082		1.61	3.87E-02
M5005_Spy1083		3.29	1.81E-04
M5005_Spy1084		1.62	7.59E-03
M5005_Spy1085	*bglA.2*	2.78	1.30E-02
M5005_Spy1103	*map*	1.07	1.94E-02
M5005_Spy1107	*murZ*	-1.97	6.31E-04
M5005_Spy1112		2.56	8.14E-07
M5005_Spy1121	*ptsH*	-1.50	1.30E-03
M5005_Spy1124	*nrdF.2*	1.13	4.71E-02
M5005_Spy1131		-2.22	8.31E-07
M5005_Spy1132	*alaS*	-1.54	2.15E-03
M5005_Spy1137		-1.77	7.12E-04
M5005_Spy1143		2.73	2.39E-03
M5005_Spy1144		2.18	1.10E-02
M5005_Spy1145	*sodA*	1.23	1.61E-02
M5005_Spy1197		-1.79	1.66E-02
M5005_Spy1206		-2.25	4.92E-03
M5005_Spy1217		-1.59	6.89E-03
M5005_Spy1218		-2.27	3.87E-04
M5005_Spy1226		1.17	1.89E-03
M5005_Spy1230		-1.41	1.69E-02
M5005_Spy1232	*xseB*	-2.40	3.70E-03
M5005_Spy1233		-1.49	9.87E-03
M5005_Spy1238	*artQ*	-1.79	1.44E-03
M5005_Spy1240	*clpE*	1.70	1.38E-03
M5005_Spy1250	*ftsA*	-1.56	1.59E-02
M5005_Spy1257	*glcK*	-1.55	6.87E-04
M5005_Spy1258		-3.04	5.56E-06
M5005_Spy1270	*arcC*	3.06	4.08E-04
M5005_Spy1271		4.19	2.69E-05
M5005_Spy1272		5.05	1.69E-06
M5005_Spy1273	*arcB*	4.51	2.45E-06
M5005_Spy1274		2.98	1.16E-04
M5005_Spy1275	*arcA*	3.71	4.03E-06
M5005_Spy1276		2.61	1.82E-02
M5005_Spy1282	*msrA*	1.83	1.73E-03
M5005_Spy1291		-1.35	1.18E-02
M5005_Spy1298	*aroB*	-1.59	4.77E-03
M5005_Spy1299		-2.05	8.42E-03
M5005_Spy1310		2.75	3.73E-02
M5005_Spy1311		2.31	1.46E-02
M5005_Spy1320	*recX*	-2.78	1.57E-03
M5005_Spy1323		1.31	1.71E-02
M5005_Spy1324		-1.65	2.74E-04
M5005_Spy1330		-1.66	9.04E-03
M5005_Spy1331		1.64	5.53E-05
M5005_Spy1342		-1.87	7.39E-04
M5005_Spy1357		1.72	2.37E-02
M5005_Spy1363		-2.21	2.21E-05
M5005_Spy1376		1.59	8.22E-03
M5005_Spy1377		3.03	1.36E-05

M5005_Spy1379	*glpF*	3.95	4.72E-04
M5005_Spy1380	*glpO*	4.93	1.37E-06
M5005_Spy1381	*glpK*	3.37	6.85E-04
M5005_Spy1382		2.16	7.97E-05
M5005_Spy1390		1.14	1.60E-02
M5005_Spy1392		2.06	2.41E-05
M5005_Spy1395	*lacD.1*	2.39	3.14E-05
M5005_Spy1396		2.86	1.90E-04
M5005_Spy1397	*lacB.1*	2.92	3.80E-03
M5005_Spy1398	*lacA.1*	3.08	5.23E-04
M5005_Spy1399		2.20	3.92E-03
M5005_Spy1408	*rbfA*	-2.21	8.50E-04
M5005_Spy1409	*infB*	-1.23	1.47E-02
M5005_Spy1410		-2.02	3.43E-04
M5005_Spy1412	*nusA*	-1.52	1.77E-03
M5005_Spy1418		-2.40	6.42E-04
M5005_Spy1420		-3.10	4.71E-05
M5005_Spy1422		-1.34	2.82E-02
M5005_Spy1424		-2.08	6.83E-03
M5005_Spy1440		-2.07	1.78E-03
M5005_Spy1444		-1.48	3.30E-02
M5005_Spy1445		-2.16	3.82E-02
M5005_Spy1447		-1.73	2.35E-02
M5005_Spy1451		-1.74	3.32E-04
M5005_Spy1452		-1.52	3.78E-03
M5005_Spy1455		-1.45	2.47E-02
M5005_Spy1456		-1.69	2.40E-02
M5005_Spy1462		-1.35	9.47E-03
M5005_Spy1469		-1.58	6.34E-03
M5005_Spy1477		-1.28	1.54E-02
M5005_Spy1479	*manL*	-1.55	2.46E-03
M5005_Spy1480	*manM*	-2.50	1.53E-03
M5005_Spy1481	*manN*	-4.09	6.81E-07
M5005_Spy1482	*manO*	1.03	2.78E-02
M5005_Spy1494	*fabH*	-1.31	3.17E-02
M5005_Spy1495		-1.70	2.27E-03
M5005_Spy1505		-1.89	2.54E-04
M5005_Spy1521		1.26	1.67E-03
M5005_Spy1523		2.14	1.79E-04
M5005_Spy1524		1.28	1.15E-03
M5005_Spy1537		1.63	2.02E-02
M5005_Spy1538	*pmi*	1.49	1.75E-02
M5005_Spy1553	*rpsR*	-1.31	8.46E-03
M5005_Spy1555	*rpsF*	-1.55	3.44E-03
M5005_Spy1560		1.27	1.02E-02
M5005_Spy1566	*recD*	1.06	2.45E-02
M5005_Spy1569	*pfl*	-2.43	1.24E-03
M5005_Spy1574		-1.73	4.53E-03
M5005_Spy1575	*norA*	-9.04	5.80E-13
M5005_Spy1582	*dnaQ*	-2.20	1.48E-04
M5005_Spy1592		-1.48	1.00E-03
M5005_Spy1607	*fba*	-1.84	3.34E-05
M5005_Spy1610	*pyrG*	-2.54	5.62E-03
M5005_Spy1611	*rpoE*	-1.55	2.51E-04
M5005_Spy1613		1.09	2.04E-02
M5005_Spy1641		1.27	3.04E-02

M5005_Spy1645		-4.02	1.06E-04
M5005_Spy1647	*rplM*	-2.45	4.26E-05
M5005_Spy1648		-1.68	4.89E-03
M5005_Spy1667		-1.86	5.24E-05
M5005_Spy1668		-1.52	2.27E-03
M5005_Spy1674		1.16	9.97E-03
M5005_Spy1692		1.82	2.90E-03
M5005_Spy1698	*trpG*	2.11	4.70E-05
M5005_Spy1699		1.54	2.94E-04
M5005_Spy1708	*dppE*	1.28	6.20E-03
M5005_Spy1711	*lmb*	-2.35	1.01E-03
M5005_Spy1718	*sic1.01*	-2.01	2.63E-03
M5005_Spy1722		1.70	4.74E-03
M5005_Spy1730		1.73	1.65E-04
M5005_Spy1743	*pflD*	1.10	1.33E-02
M5005_Spy1744		1.06	3.54E-02
M5005_Spy1745		2.83	7.05E-03
M5005_Spy1754		1.51	4.05E-03
M5005_Spy1757		-1.42	1.57E-03
M5005_Spy1769	*ahpF*	1.19	2.95E-02
M5005_Spy1771	*hutU*	4.32	1.36E-03
M5005_Spy1772		5.39	1.07E-02
M5005_Spy1773		8.04	7.67E-07
M5005_Spy1774	*fhs.2*	5.03	4.50E-02
M5005_Spy1775		7.58	2.57E-03
M5005_Spy1776		6.35	2.81E-03
M5005_Spy1777	*hutH*	5.76	1.92E-02
M5005_Spy1778	*hutG*	4.69	5.64E-03
M5005_Spy1780	*rpsB*	-2.03	7.56E-04
M5005_Spy1781		-2.09	5.40E-05
M5005_Spy1783	*dexS*	1.76	3.00E-04
M5005_Spy1786		-1.68	2.82E-03
M5005_Spy1787		-2.56	1.48E-04
M5005_Spy1789	*nrdG*	-1.34	1.31E-03
M5005_Spy1790		-2.37	1.05E-02
M5005_Spy1797		1.93	6.97E-06
M5005_Spy1799	*recA*	3.32	1.38E-05
M5005_Spy1802	*ruvA*	1.09	2.08E-02
M5005_Spy1811		-2.16	7.67E-04
M5005_Spy1812		-1.69	4.74E-03
M5005_Spy1815	*rpmF*	-2.76	1.72E-03
M5005_Spy1816	*rpmG*	-2.69	3.41E-05
M5005_Spy1830		-2.89	1.23E-02
M5005_Spy1835	*holB*	-2.57	2.20E-06
M5005_Spy1842	*sdhA*	-3.69	9.02E-04
M5005_Spy1843		-2.79	7.55E-04
M5005_Spy1848		1.15	3.44E-02
M5005_Spy1850		1.53	2.74E-02
M5005_Spy1852	*hasB*	-1.45	3.26E-03
M5005_Spy1860		-1.69	1.13E-02
M5005_Spy1863		1.10	6.10E-03
M5005_Spy1866	*parB*	1.32	5.42E-03

[a] Analysis conducted on intergenic region 5' to spy number listed
[b] Both Limma and DESeq analyses exhibited Log_2 fold-change of $-1 \leq$ or ≥ 1. Limma data only shown.
[c] p value ≤ 0.05 in both analyses

Appendix 7: Differential Expression of Mga regulated intergenic region in THY at late log (Δ*mga*/wt 5448)

Spy #[a]	Gene Name	Log_2 FC[b]	*p* value[c]
M5005_Spy0028		-2.48	5.46E-06
M5005_Spy0029	*purD*	-2.56	1.93E-06
M5005_Spy0034		-2.33	3.09E-09
M5005_Spy0065	*adk*	1.21	5.57E-03
M5005_Spy0115		-2.89	7.49E-07
M5005_Spy0140		-2.88	3.39E-04
M5005_Spy0144		-3.85	1.48E-05
M5005_Spy0157	*opuAA*	-1.24	2.30E-04
M5005_Spy0158	*opuABC*	-1.60	1.30E-05
M5005_Spy0159	*polA*	-3.15	4.48E-10
M5005_Spy0160		-1.07	5.31E-04
M5005_Spy0213		-1.37	1.32E-03
M5005_Spy0623		-1.43	1.52E-03
M5005_Spy0719		-1.19	1.80E-04
M5005_Spy0776	*lepA*	-1.53	2.47E-05
M5005_Spy0778	*msrB*	-1.82	2.38E-04
M5005_Spy0882		-2.13	8.29E-08
M5005_Spy0884	*smf*	-2.38	1.09E-05
M5005_Spy0912		-1.51	1.54E-04
M5005_Spy1139	*nagB*	-1.80	4.30E-06
M5005_Spy1348		-1.17	1.33E-03
M5005_Spy1539	*scrK*	-1.19	9.08E-05
M5005_Spy1540	*endoS*	-1.98	2.45E-05
M5005_Spy1543	*scrB*	-2.16	1.63E-05
M5005_Spy1631	*salA*	-2.00	3.43E-04
M5005_Spy1661		1.19	2.95E-03
M5005_Spy1668		-2.77	1.03E-08
M5005_Spy1714		-3.91	1.65E-05
M5005_Spy1715	*scpA*	-6.27	3.29E-09
M5005_Spy1716		-4.79	2.32E-09
M5005_Spy1718	*sic1.01*	-5.62	7.13E-13
M5005_Spy1719	*emm1.0*	-5.59	1.06E-09
M5005_Spy1720	*mga*	-3.61	1.17E-09
M5005_Spy1721		-1.65	6.48E-06
M5005_Spy1729		-1.25	7.03E-04
M5005_Spy1730		-1.14	6.40E-04
M5005_Spy1731		-1.46	2.37E-04
M5005_Spy1733		2.31	2.49E-05
M5005_Spy1734	*spi*	4.68	2.91E-06
M5005_Spy1735	*speB*	4.94	1.24E-05
M5005_Spy1736		4.57	7.03E-09
M5005_Spy1737	*rgg*	3.00	1.99E-07
M5005_Spy1787		-3.03	2.60E-07
M5005_Spy1852	*hasB*	-1.54	7.45E-05
M5005_Spy1860		-1.52	2.65E-03
M5005_Spy1864		-1.47	5.86E-06

[a] Analyses conducted on intergenic region 5' to spy number listed
[b] Both Limma & DESeq analyses exhibited Log_2 fold-change of $-1 \leq$ or ≥ 1. Limma data only shown.
[c] *p* value ≤ 0.05 in both analyses

Appendix 8: Differential Expression of Mga regulated intergenic region in C media at late log (Δ*mga*/wt 5448)

Spy #[a]	Gene Name	Log_2 FC[b]	*p* value[c]
M5005_Spy0006	*trcF*	-1.20	7.56E-03
M5005_Spy0010		-1.27	4.31E-03
M5005_Spy0018	*prsA.2*	-1.57	3.63E-04
M5005_Spy0023		-2.16	4.84E-03
M5005_Spy0031	*purK*	-1.45	7.10E-03
M5005_Spy0043	*rpsJ*	-1.47	7.16E-05
M5005_Spy0073		-2.31	5.70E-09
M5005_Spy0107		1.70	5.99E-03
M5005_Spy0115		-1.17	1.36E-02
M5005_Spy0116	*atoE*	-1.18	2.51E-02
M5005_Spy0128	*ntpE*	1.96	4.20E-03
M5005_Spy0130	*ntpF*	1.77	4.09E-03
M5005_Spy0131	*ntpA*	1.83	3.32E-03
M5005_Spy0132	*ntpB*	1.50	1.28E-02
M5005_Spy0133	*ntpD*	2.09	2.16E-03
M5005_Spy0136	*purA*	1.46	9.68E-05
M5005_Spy0137		1.34	4.60E-04
M5005_Spy0140		-3.74	1.35E-05
M5005_Spy0141	*slo*	-2.93	9.45E-05
M5005_Spy0142		-3.33	1.04E-04
M5005_Spy0144		-3.57	7.19E-04
M5005_Spy0157	*opuAA*	-1.25	8.81E-05
M5005_Spy0158	*opuABC*	-1.62	3.33E-06
M5005_Spy0159	*polA*	-1.22	2.76E-03
M5005_Spy0167		-2.07	2.19E-02
M5005_Spy0177		-1.34	4.96E-03
M5005_Spy0191		-1.18	2.64E-03
M5005_Spy0197		1.66	2.53E-03
M5005_Spy0202		1.31	4.86E-04
M5005_Spy0205	*fasC*	-1.30	1.41E-03
M5005_Spy0207	*fasX*	-1.47	3.38E-06
M5005_Spy0224	*rpe*	1.41	7.49E-03
M5005_Spy0230	*rpsL*	-1.31	7.53E-05
M5005_Spy0231	*rpsG*	-1.47	1.16E-04
M5005_Spy0271		1.39	2.19E-03
M5005_Spy0274	*braB*	-2.97	9.41E-04
M5005_Spy0276		-1.28	9.29E-03
M5005_Spy0280		1.03	4.42E-04
M5005_Spy0296		1.01	9.03E-04
M5005_Spy0297		1.81	4.26E-03
M5005_Spy0318	*pflC*	-1.39	3.64E-04
M5005_Spy0319	*ppaC*	1.09	6.59E-04
M5005_Spy0340	*lctO*	1.94	4.17E-04
M5005_Spy0352		-1.92	2.81E-03
M5005_Spy0375	*rplA*	-1.40	9.02E-04
M5005_Spy0376		-1.60	2.27E-03
M5005_Spy0385		1.88	5.65E-03
M5005_Spy0386	*phoH*	1.05	1.31E-03
M5005_Spy0404		-2.60	5.79E-05
M5005_Spy0423	*pepQ*	1.23	2.84E-05
M5005_Spy0424	*ccpA*	1.59	1.85E-03

M5005_Spy0432		-1.68	8.33E-05
M5005_Spy0453		1.90	1.34E-02
M5005_Spy0455		1.74	1.26E-02
M5005_Spy0477		1.47	1.14E-05
M5005_Spy0478		1.20	3.04E-03
M5005_Spy0479		1.03	1.59E-04
M5005_Spy0504	*pepF*	1.52	1.14E-04
M5005_Spy0518		1.38	5.03E-03
M5005_Spy0520		-1.68	2.21E-05
M5005_Spy0521	*agaV*	-1.35	2.94E-03
M5005_Spy0522		-2.14	6.38E-08
M5005_Spy0523		-1.50	4.37E-04
M5005_Spy0535		1.04	1.17E-03
M5005_Spy0563	*sagB*	-1.65	1.65E-04
M5005_Spy0564	*sagC*	-1.13	1.43E-03
M5005_Spy0565	*sagD*	-1.28	1.01E-03
M5005_Spy0566	*sagE*	-1.35	8.19E-06
M5005_Spy0567	*sagF*	-1.21	9.23E-03
M5005_Spy0568	*sagG*	-2.23	4.08E-05
M5005_Spy0569	*sagH*	-1.62	1.31E-05
M5005_Spy0570	*sagI*	-1.27	3.04E-03
M5005_Spy0571		-1.51	1.22E-05
M5005_Spy0572		-1.26	1.71E-05
M5005_Spy0573	*lig*	-1.87	6.75E-04
M5005_Spy0592		1.11	3.04E-03
M5005_Spy0599	*dnaG*	1.18	4.56E-03
M5005_Spy0620	*rpl36*	-1.29	2.05E-03
M5005_Spy0621	*rplT*	-1.23	1.34E-04
M5005_Spy0640	*pyrP*	-1.49	2.40E-04
M5005_Spy0642	*carA*	-1.33	2.74E-04
M5005_Spy0650		-1.48	4.33E-03
M5005_Spy0651		-1.94	9.40E-06
M5005_Spy0652		-2.10	5.52E-07
M5005_Spy0653	*czcD*	-2.39	1.12E-07
M5005_Spy0665		-1.81	3.41E-04
M5005_Spy0666		-1.64	4.56E-03
M5005_Spy0668	*mac*	-1.39	5.37E-03
M5005_Spy0694	*clpL*	-1.34	7.20E-04
M5005_Spy0710		-1.09	4.42E-02
M5005_Spy0714		-1.43	2.47E-04
M5005_Spy0715		-2.51	2.39E-07
M5005_Spy0722	*miaA*	1.28	4.32E-04
M5005_Spy0731		-1.41	4.84E-03
M5005_Spy0734	*cpsFO*	1.40	2.37E-02
M5005_Spy0752	*acoB*	1.04	3.75E-02
M5005_Spy0753	*acoC*	1.19	4.50E-04
M5005_Spy0754		1.31	4.14E-04
M5005_Spy0755	*acoL*	1.72	5.13E-04
M5005_Spy0756		1.22	1.28E-03
M5005_Spy0791	*uvrC*	1.47	5.66E-05
M5005_Spy0803	*srtI*	-2.53	4.69E-03
M5005_Spy0811		1.19	2.29E-02
M5005_Spy0820	*folC.1*	1.31	2.24E-05
M5005_Spy0826	*potA*	-1.27	7.11E-04
M5005_Spy0832	*malP*	-3.00	9.19E-03
M5005_Spy0856		1.23	2.48E-02

M5005_Spy0858	*xpt*	1.31	5.12E-03
M5005_Spy0877		1.36	1.57E-02
M5005_Spy0878		1.66	2.58E-02
M5005_Spy0880		2.87	1.08E-08
M5005_Spy0881		2.83	1.18E-07
M5005_Spy0882		1.20	1.54E-04
M5005_Spy0898		-1.53	1.22E-02
M5005_Spy0900		-1.98	1.47E-02
M5005_Spy0921		-1.74	1.19E-04
M5005_Spy0922	*pdxK*	-1.66	3.08E-02
M5005_Spy0924		-1.18	6.13E-04
M5005_Spy0947	*ciaH*	-1.42	9.70E-04
M5005_Spy0978		-1.31	5.94E-03
M5005_Spy0980		-1.53	1.26E-04
M5005_Spy0998		-1.27	6.19E-03
M5005_Spy1000		-2.77	2.82E-03
M5005_Spy1022		-2.49	1.10E-02
M5005_Spy1027		-2.24	1.67E-03
M5005_Spy1028		-1.23	3.07E-02
M5005_Spy1034		-1.16	1.45E-02
M5005_Spy1040		-1.51	2.03E-02
M5005_Spy1041		-1.57	3.29E-02
M5005_Spy1076	*glnH*	1.50	2.69E-04
M5005_Spy1077	*glnQ.2*	1.26	4.16E-04
M5005_Spy1090		-1.90	1.01E-02
M5005_Spy1093		1.23	2.80E-03
M5005_Spy1108	*metK2*	-1.33	1.82E-02
M5005_Spy1110	*birA*	-1.38	2.12E-02
M5005_Spy1114		-1.50	1.30E-04
M5005_Spy1116	*udk*	-1.49	2.39E-04
M5005_Spy1124	*nrdF.2*	1.06	7.41E-03
M5005_Spy1134		-1.34	2.23E-04
M5005_Spy1139	*nagB*	-1.37	6.93E-05
M5005_Spy1146		1.72	1.07E-03
M5005_Spy1170		-2.38	8.22E-06
M5005_Spy1171		-2.76	2.26E-06
M5005_Spy1172		-1.82	1.36E-02
M5005_Spy1211		-1.40	3.90E-02
M5005_Spy1227		-1.17	1.01E-03
M5005_Spy1228	*recN*	-1.27	3.08E-04
M5005_Spy1229	*argR1*	-1.51	2.26E-03
M5005_Spy1231	*fps*	-1.27	1.45E-02
M5005_Spy1232	*xseB*	-1.24	2.42E-02
M5005_Spy1238	*artQ*	-2.37	1.07E-06
M5005_Spy1239		-1.65	1.05E-02
M5005_Spy1240	*clpE*	1.05	3.06E-03
M5005_Spy1241	*mutT*	1.45	7.61E-05
M5005_Spy1260		1.43	1.80E-02
M5005_Spy1261		-1.48	4.54E-03
M5005_Spy1290		-1.23	6.94E-03
M5005_Spy1297		-1.33	3.01E-03
M5005_Spy1298	*aroB*	-1.52	2.50E-04
M5005_Spy1324		-1.96	9.19E-08
M5005_Spy1325		1.03	2.70E-03
M5005_Spy1326	*comFC*	1.57	4.69E-04
M5005_Spy1344	*atoB*	1.10	6.05E-04

M5005_Spy1357		1.56	3.30E-03
M5005_Spy1369		-1.22	1.49E-02
M5005_Spy1374		-1.28	3.67E-04
M5005_Spy1379	*glpF*	1.37	5.62E-03
M5005_Spy1387		1.11	7.91E-04
M5005_Spy1389		-1.15	5.17E-04
M5005_Spy1415	*sdaD2*	-1.92	3.04E-04
M5005_Spy1416		-2.50	1.86E-05
M5005_Spy1422		-1.37	6.97E-03
M5005_Spy1427		-1.09	3.23E-02
M5005_Spy1430		-1.01	2.75E-02
M5005_Spy1432		-1.08	8.77E-03
M5005_Spy1435		-1.18	5.01E-03
M5005_Spy1436		-1.06	1.28E-02
M5005_Spy1439		-1.08	3.75E-02
M5005_Spy1440		-2.33	2.81E-04
M5005_Spy1451		-1.63	1.00E-04
M5005_Spy1453		-2.24	1.74E-03
M5005_Spy1479	*manL*	-1.20	3.89E-03
M5005_Spy1483	*serS*	1.13	2.57E-03
M5005_Spy1486	*accC*	1.43	2.45E-02
M5005_Spy1487	*fabZ*	1.08	4.24E-02
M5005_Spy1488	*accB*	1.12	1.83E-02
M5005_Spy1489	*fabF*	1.38	9.68E-03
M5005_Spy1490		1.05	1.85E-02
M5005_Spy1514		1.61	1.55E-02
M5005_Spy1525		-2.40	6.16E-03
M5005_Spy1526	*fhuC*	-1.99	4.69E-04
M5005_Spy1528		-1.85	1.33E-03
M5005_Spy1529	*shp*	-1.96	2.24E-04
M5005_Spy1530		-1.30	1.46E-03
M5005_Spy1537		1.25	2.16E-03
M5005_Spy1562		-1.27	7.43E-05
M5005_Spy1570		-1.26	1.83E-02
M5005_Spy1575	*norA*	1.23	4.85E-02
M5005_Spy1582	*dnaQ*	-1.79	9.53E-05
M5005_Spy1606	*rpmB*	-1.21	9.35E-04
M5005_Spy1607	*fba*	-1.24	3.36E-05
M5005_Spy1612	*ropA*	-1.95	1.04E-04
M5005_Spy1628		-2.60	3.86E-03
M5005_Spy1630	*salB*	-2.93	6.65E-05
M5005_Spy1631	*salA*	-2.58	4.65E-05
M5005_Spy1632	*lacG*	-1.61	6.44E-03
M5005_Spy1633	*lacE*	-1.94	4.61E-03
M5005_Spy1635	*lacD.2*	-1.83	1.06E-02
M5005_Spy1637	*lacB.2*	-1.62	1.92E-02
M5005_Spy1642		-3.44	1.81E-03
M5005_Spy1647	*rplM*	-1.28	9.15E-04
M5005_Spy1648		-1.40	6.81E-04
M5005_Spy1649		-1.63	3.53E-05
M5005_Spy1671		1.07	1.11E-02
M5005_Spy1685		-1.43	1.81E-03
M5005_Spy1702	*smeZ*	-1.15	7.51E-03
M5005_Spy1714		-6.96	4.20E-09
M5005_Spy1715	*scpA*	-5.79	3.10E-08
M5005_Spy1716		-3.82	3.49E-08

M5005_Spy1718	*sic1.01*	-5.29	9.87E-13
M5005_Spy1719	*emm1.0*	-6.93	9.47E-13
M5005_Spy1720	*mga*	-4.64	5.18E-13
M5005_Spy1721		-1.50	1.53E-05
M5005_Spy1732		-1.35	1.98E-02
M5005_Spy1733		-1.12	1.37E-02
M5005_Spy1738	*spd*	-1.59	4.25E-03
M5005_Spy1754		1.54	2.49E-05
M5005_Spy1768	*ahpC*	1.55	1.69E-03
M5005_Spy1776		-1.69	4.56E-02
M5005_Spy1780	*rpsB*	-2.10	2.02E-06
M5005_Spy1781		-1.35	8.98E-05
M5005_Spy1800	*cinA*	1.05	5.23E-03
M5005_Spy1832		-1.23	4.36E-04
M5005_Spy1865	*htrA*	1.10	3.07E-02

[a] Analysis conducted on intergenic region 5' to the spy number listed.
[b] Both Limma and DESeq analyses exhibited Log_2 fold-change of $-1 \leq$ or ≥ 1. Limma data only shown.
[c] p value ≤ 0.05 in both analyses.

Bibliography

1. **Carapetis JR, Steer AC, Mulholland EK, Weber M.** 2005. The global burden of group A streptococcal diseases. Lancet Infectious Disease **5:**685-694.
2. **Bisno AL, Brito MO, Collins CM.** 2003. Molecular basis of group A streptococcal virulence. Lancet Infectious Disease **3:**191-200.
3. **Cunningham MW.** 2000. Pathogenesis of group A streptococcal infections. Clin Microbiol Rev **13:**470-511.
4. **Loughman JA, Caparon MG.** 2006. A novel adaptation of aldolase regulates virulence in *Streptococcus pyogenes*. EMBO Journal **25:**5414-5422.
5. **Kinkel TL, McIver KS.** 2008. CcpA-mediated repression of streptolysin S expression and virulence in the group A streptococcus. Infection and Immunity **76:**3451-3463.
6. **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:**1698-1703.
7. **Kietzman CC, Caparon MG.** 2010. CcpA and LacD.1 affect temporal regulation of Streptococcus pyogenes virulence genes. Infection and Immunity **78:**241-252.
8. **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:**1192-1204.
9. **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 doi:10.1128/IAI.01296-15.
10. **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:**224-232.
11. **Hondorp ER, McIver KS.** 2007. The Mga virulence regulon: infection where the grass is greener. Molecular Microbiology **66:**1056-1065.
12. **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:**455-465.
13. **Kihlberg BM, Cooney J, Caparon MG, Olsén A, Björck L.** 1995. Biological properties of a *Streptococcus pyogenes* mutant generated by Tn916 insertion in *mga*. Microbial Pathogensis **19:**299-315.

14. **Cho KH, Caparon MG.** 2005. Patterns of virulence gene expression differ between biofilm and tissue communities of *Streptococcus pyogenes*. Molecular Microbiology **57:**1545-1556.
15. **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.
16. **Kizy AE, Neely MN.** 2009. First Streptococcus pyogenes signature-tagged mutagenesis screen identifies novel virulence determinants. Infection and Immunity **77:**1854-1865.
17. **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.
18. **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.
19. **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:**1176-1193.
20. **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:**2382-2395.
21. **Denny FW, Jr.** 2000. History of hemolytic streptococci and associated diseases, p 1-18. *In* Stevens DL, Kaplan EL (ed), Streptococcal Infections: Clinical Aspects, Microbiology, and Molecular Pathogenesis. Oxford University Press, New York.
22. **Magner LN.** 1992. A History of Medicine. New York Marcel Dekker, Inc.
23. **Lancefield RC.** 1933. A serological differentiation of human and other groups of hemolytic streptococci. Journal of Experimental Medicine **57:**571-595.
24. **Martin DR.** 2000. Laboratory evaluation of streptococci, p 266-279. *In* Stevens DL, Kaplan EL (ed), Streptococcal Infections: Clinical Aspects, Microbiology, and Molecular Pathogenesis. Oxford University Press, New York.
25. **Todd EW, Hewitt LF.** 1932. A new culture medium for the production of antigenic streptococcal haemolysin. Journal of Pathology and Bacteriology **35:**973.
26. **Vincent WF, Lisiewski KJ.** 1969. Improved growth medium for group A streptococci. Applied Microbiology **18:**954-955.
27. **Slade HD, Knox GA, Slamp WC.** 1951. The amino acid nutrition of group A hemolytic streptococci, with reference to the effect of glutathione on the cysteine requirement. Journal of Biochemistry **62:**669-675.
28. **van de Rijn I, Kessler RE.** 1980. Growth characteristics of group A streptococci in a new chemically defined medium. Infection and Immunity **27:**444-448.

29. **Gerber MA, Spadaccini LJ, Wright LL, Deutsch L.** 1984. Latex agglutination tests for rapid identification of group A streptococci directly from throat swabs. Journal of Pediatrics **105:**702-705.
30. **Lancefield RC.** 1928. The antigenic complex of *Streptococcus haemolyticus*: I. Demonstration of a type-specific substance in extracts of *Streptococcus haemolyticus*. Journal of Experimental Medicine **47:**91-103.
31. **Lancefield RC.** 1928. The antigenic complex of the *Streptococcus haemolyticus*: II. Chemical and immunological properties of the protein fractions. Journal of Experimental Medicine **47:**469-480.
32. **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. Clin Microbiol Rev **27:**264-301.
33. **McMillan DJ, Dreze PA, Vu T, Bessen DE, Guglielmini J, Steer AC, Carapetis JR, Van Melderen L, Sriprakash KS, Smeesters PR.** 2013. Updated model of group A Streptococcus M proteins based on a comprehensive worldwide study. Clinical Microbiology and Infection **19:**E222-229.
34. **Griffith F.** 1934. The Serological Classification of Streptococcus pyogenes. Journal of Hygiene **34:**542-584.
35. **Beall B, Facklam RR, Elliott JA, Franklin AR, Hoenes T, Jackson D, Laclaire L, Thompson T, Viswanathan R.** 1998. Streptococcal emm types associated with T-agglutination types and the use of conserved emm gene restriction fragment patterns for subtyping group A streptococci. J Med Microbiol **47:**893-898.
36. **Maxted WR, Widdowson JP, Fraser CA, Ball LC, Bassett DC.** 1973. The use of the serum opacity reaction in the typing of group A streptococci. Journal of Medical Microbiology **6:**83-90.
37. **Johnson DR, Kaplan EL.** 1993. A review of the correlation of T-agglutination patterns and M-protein typing and opacity factor production in the identification of group A streptococci. Journal of Medical Microbiology **38:**311-315.
38. **Wessels MR.** 2011. Clinical practice. Streptococcal pharyngitis. New England Journal of Medicine **364:**648-655.
39. **Choby BA.** 2009. Diagnosis and treatment of streptococcal pharyngitis. American Family Physician **79:**383-390.
40. **Stevens DL.** 2000. Life-threatening streptococcal infections: scarlet fever, necrotizing fasciitis, myositis, bacteremia, and streptococcal toxic shock syndrome, p 163-179. *In* Stevens DL, Kaplan EL (ed), Streptococcal Infections: Clinical Aspects, Microbiology, and Molecular Pathogenesis. Oxford University Press, New York.
41. **Bisno AL, Stevens DL.** 1996. Streptococcal infections of skin and soft tissues. The New England Journal of Medicine **334:**240-245.
42. **Anthony BF.** 2000. Streptococcal pyoderma, p 144-151. *In* Stevens DL, Kaplan EL (ed), Streptococcal Infections: Clinical Aspects, Microbiology, and Molecular Pathogenesis. Oxford University Press, New York.

43. **Stevens DL.** 2000. Group A beta-hemolytic streptococci: virulence factors, pathogenesis, and spectrum of clinical infections, p 37-56. *In* Stevens DL, Kaplan EL (ed), Streptococcal Infections: Clinical Aspects, Microbiology, and Molecular Pathogenesis. Oxford University Press, New York.
44. **Stevens DL.** 1995. Streptococcal toxic-shock syndrome: spectrum of disease, pathogenesis, and new concepts in treatment. Emerging Infectious Diseases **1:**69-78.
45. **Descamps V, Aitken J, Lee MG.** 1994. Hippocrates on necrotising fasciitis. Lancet **344:**556.
46. **Olsen RJ, Musser JM.** 2010. Molecular pathogenesis of necrotizing fasciitis. Annual Review of Pathology **5:**1-31.
47. **Markowitz M, Kaplan EL.** 2000. Rheumatic Fever, p 133-143. *In* Stevens DL, Kaplan EL (ed), Streptococcal Infections: Clinical Aspects, Microbiology, and Molecular Pathogenesis. Oxford University Press, New York.
48. **Ayoub EM, Kotb M, Cunningham MW.** 2000. Rheumatic fever pathogenesis, p 102-132. *In* Stevens DL, Kaplan EL (ed), Streptococcal Infections: Clinical Aspects, Microbiology, and Molecular Pathogenesis. Oxford University Press, New York.
49. **Jones TD.** 1944. The diagnosis of rheumatic fever. Journal of the American Medical Association **126:**481-484.
50. **Tompkins DG, Boxerbaum B, Liebman J.** 1972. Long-term prognosis of rheumatic fever patients receiving regular intramuscular benzathine penicillin. Circulation **45:**543-551.
51. **Holm SE, Nordstrand A, Stevens DL, Norgreen M.** 2000. Acute poststreptococcal glomerulonephritis, p 152-162. *In* Stevens DL, Kaplan EL (ed), Streptococcal Infections: Clinical Aspects, Microbiology, and Molecular Pathogenesis. Oxford University Press, New York.
52. **Williams KA, Swedo SE.** 2014. Post-infectious autoimmune disorders: Sydenham's chorea, PANDAS and beyond. Brain Research doi:10.1016/j.brainres.2014.09.071.
53. **Swedo SE, Leonard HL, Garvey M, Mittleman B, Allen AJ, Perlmutter S, Lougee L, Dow S, Zamkoff J, Dubbert BK.** 1998. Pediatric autoimmune neuropsychiatric disorders associated with streptococcal infections: clinical description of the first 50 cases. American Journal of Psychiatry **155:**264-271.
54. **Leckman JF, King RA, Gilbert DL, Coffey BJ, Singer HS, Dure LSt, Grantz H, Katsovich L, Lin H, Lombroso PJ, Kawikova I, Johnson DR, Kurlan RM, Kaplan EL.** 2011. Streptococcal upper respiratory tract infections and exacerbations of tic and obsessive-compulsive symptoms: a prospective longitudinal study. Journal of the American Academy of Child and Adolescent Psychiatry **50:**108-118 e103.
55. **Kurlan R, Johnson D, Kaplan EL.** 2008. Streptococcal infection and exacerbations of childhood tics and obsessive-compulsive symptoms: a prospective blinded cohort study. Pediatrics **121:**1188-1197.
56. **Esposito S, Bianchini S, Baggi E, Fattizzo M, Rigante D.** 2014. Pediatric autoimmune neuropsychiatric disorders associated with streptococcal

infections: an overview. European Journal of Clinical Microbiology and Infectious Diseases **33:**2105-2109.
57. **Carapetis JR, Currie BJ, Kaplan EL.** 1999. Epidemiology and prevention of group A streptococcal infections: acute respiratory tract infections, skin infections, and their sequelae at the close of the twentieth century. Clinical Infectious Disease **28:**205-210.
58. **Tsoi SK, Smeesters PR, Frost HR, Licciardi P, Steer AC.** 2015. Correlates of Protection for M Protein-Based Vaccines against Group A Streptococcus. Journal of Immunological Research **2015:**167089.
59. **Kotloff KL, Corretti M, Palmer K, Campbell JD, Reddish MA, Hu MC, Wasserman SS, Dale JB.** 2004. Safety and immunogenicity of a recombinant multivalent group a streptococcal vaccine in healthy adults: phase 1 trial. Journal of the American Medical Association **292:**709-715.
60. **Hu MC, Walls MA, Stroop SD, Reddish MA, Beall B, Dale JB.** 2002. Immunogenicity of a 26-valent group A streptococcal vaccine. Infection and Immunity **70:**2171-2177.
61. **Hirst GK, Lancefield RC.** 1939. Antigenic properties of the type-specific substance derived from group A hemolytic streptococci. Journal of Experimental Medicine **69:**425-445.
62. **Dale JB, Beachey EH.** 1985. Epitopes of streptococcal M proteins shared with cardiac myosin. Journal of Experimental Medicine **162:**583-591.
63. **Dale JB, Penfound TA, Chiang EY, Walton WJ.** 2011. New 30-valent M protein-based vaccine evokes cross-opsonic antibodies against non-vaccine serotypes of group A streptococci. Vaccine **29:**8175-8178.
64. **Cole JN, Henningham A, Gillen CM, Ramachandran V, Walker MJ.** 2008. Human pathogenic streptococcal proteomics and vaccine development. Proteomics Clinical Applicaation **2:**387-410.
65. **Henningham A, Gillen CM, Walker MJ.** 2013. Group a streptococcal vaccine candidates: potential for the development of a human vaccine. Current Topics in Microbiology and Immunology **368:**207-242.
66. **Steer AC, Dale JB, Carapetis JR.** 2013. Progress toward a global group a streptococcal vaccine. The Pediatric Infectious Disease Journal **32:**180-182.
67. **Michos AG, Bakoula CG, Braoudaki M, Koutouzi FI, Roma ES, Pangalis A, Nikolopoulou G, Kirikou E, Syriopoulou VP.** 2009. Macrolide resistance in Streptococcus pyogenes: prevalence, resistance determinants, and emm types. Diagnosis of Microbiology and Infectious Disease **64:**295-299.
68. **Åkesson P, Sjöholm AG, Björck L.** 1996. Protein SIC, a novel extracellular protein of *Streptococcus pyogenes* interfering with complement function. Journal of Biological Chemistry **271:**1081-1088.
69. **Fernie-King BA, Seilly DJ, Willers C, Wurzner R, Davies A, Lachmann PJ.** 2001. Streptococcal inhibitor of complement (SIC) inhibits the membrane attack complex by preventing uptake of C567 onto cell membranes. Immunology **103:**390-398.
70. **Fernie-King BA, Seilly DJ, Davies A, Lachmann PJ.** 2002. Streptococcal inhibitor of complement inhibits two additional components of the mucosal

innate immune system: secretory leukocyte proteinase inhibitor and lysozyme. Infection and Immunity **70:**4908-4916.

71. **Akesson P, Herwald H, Rasmussen M, Hakansson K, Abrahamson M, Hasan AA, Schmaier AH, Muller-Esterl W, Bjorck L.** 2010. Streptococcal inhibitor of complement-mediated lysis (SIC): an anti-inflammatory virulence determinant. Microbiology **156:**3660-3668.
72. **Hoe NP, Ireland RM, DeLeo FR, Gowen BB, Dorward DW, Voyich JM, Liu M, Burns EH, Jr., Culnan DM, Bretscher A, Musser JM.** 2002. Insight into the molecular basis of pathogen abundance: group A streptococcus inhibitor of complement inhibits bacterial adherence and internalization into human cells. Proceedings of the National Academy of Sciences of the United States of America **99:**7646-7651.
73. **Fernie-King BA, Seilly DJ, Lachmann PJ.** 2004. The interaction of streptococcal inhibitor of complement (SIC) and its proteolytic fragments with the human beta defensins. Immunology **111:**444-452.
74. **Frick IM, Åkesson P, Rasmussen M, Schmidtchen A, Björck L.** 2003. SIC, a secreted protein of *Streptococcus pyogenes* that inactivates antibacterial peptides. Journal of Biological Chemistry **278:**16561-16566.
75. **Lukomski S, Hoe NP, Abdi I, Rurangirwa J, Kordari P, Liu M, Dou SJ, Adams GG, Musser JM.** 2000. Nonpolar inactivation of the hypervariable streptococcal inhibitor of complement gene (*sic*) in serotype M1 *Streptococcus pyogenes* significantly decreases mouse mucosal colonization. Infection and Immunity **68:**535-542.
76. **Pence MA, Rooijakkers SH, Cogen AL, Cole JN, Hollands A, Gallo RL, Nizet V.** 2010. Streptococcal inhibitor of complement promotes innate immune resistance phenotypes of invasive M1T1 group A Streptococcus. Journal of Innate Immunity **2:**587-595.
77. **Frick IM, Shannon O, Akesson P, Morgelin M, Collin M, Schmidtchen A, Bjorck L.** 2011. Antibacterial activity of the contact and complement systems is blocked by SIC, a protein secreted by Streptococcus pyogenes. Journal of Biological Chemistry **286:**1331-1340.
78. **Bhakdi S, Tranum-Jensen J, Sziegoleit A.** 1985. Mechanism of membrane damage by streptolysin-O. Infection and Immunity **47:**52-60.
79. **Hirsch JG, Bernheimer AW, Weissmann G.** 1963. Motion Picture Study of the Toxic Action of Streptolysins on Leucocytes. Journal of Experimental Medicine **118:**223-228.
80. **Launay JM, Alouf JE.** 1979. Biochemical and ultrastructural study of the disruption of blood platelets by streptolysin O. Biochim Biophys Acta **556:**278-291.
81. **Bricker AL, Cywes C, Ashbaugh CD, Wessels MR.** 2002. NAD+-glycohydrolase acts as an intracellular toxin to enhance the extracellular survival of group A streptococci. Molecular Microbiology **44:**257-269.
82. **Bastiat-Sempe B, Love JF, Lomayesva N, Wessels MR.** 2014. Streptolysin O and NAD-glycohydrolase prevent phagolysosome acidification and promote group a streptococcus survival in macrophages. MBio **5:**e01690-01614.

83. **Musser JM, Stockbauer K, Kapur V, Rudgers GW.** 1996. Substitution of cysteine 192 in a highly conserved Streptococcus pyogenes extracellular cysteine protease (interleukin 1beta convertase) alters proteolytic activity and ablates zymogen processing. Infection and Immunity **64:**1913-1917.
84. **Chaussee MS, Liu J, Stevens DL, Ferretti JJ.** 1996. Genetic and phenotypic diversity among isolates of *Streptococcus pyogenes* from invasive infections. Journal of Infectious Diseases **173:**901-908.
85. **Aziz RK, Pabst MJ, Jeng A, Kansal R, Low DE, Nizet V, Kotb M.** 2004. Invasive M1T1 group A Streptococcus undergoes a phase-shift in vivo to prevent proteolytic degradation of multiple virulence factors by SpeB. Molecular Microbiology **51:**123-134.
86. **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:**1956-1959.
87. **Kapur V, Majesky MW, Li LL, Black RA, Musser JM.** 1993. Cleavage of interleukin 1 beta (IL-1 beta) precursor to produce active IL-1 beta by a conserved extracellular cysteine protease from *Streptococcus pyogenes*. Proceedings of the National Academy of Sciences of the United States of America **90:**7676-7680.
88. **Loughman JA, Caparon M.** 2006. Regulation of SpeB in Streptococcus pyogenes by pH and NaCl: a model for in vivo gene expression. Journal of Bacteriology **188:**399-408.
89. **Lei B, DeLeo FR, Hoe NP, Graham MR, Mackie SM, Cole RL, Liu M, Hill HR, Low DE, Federle MJ, Scott JR, Musser JM.** 2001. Evasion of human innate and acquired immunity by a bacterial homolog of CD11b that inhibits opsonophagocytosis. Nature Medicine **7:**1298-1305.
90. **Lei B, DeLeo FR, Reid SD, Voyich JM, Magoun L, Liu M, Braughton KR, Ricklefs S, Hoe NP, Cole RL, Leong JM, Musser JM.** 2002. Opsonophagocytosis-inhibiting mac protein of group a streptococcus: identification and characteristics of two genetic complexes. Infection and Immunity **70:**6880-6890.
91. **von Pawel-Rammingen U, Johansson BP, Bjorck L.** 2002. IdeS, a novel streptococcal cysteine proteinase with unique specificity for immunoglobulin G. EMBO Journal **21:**1607-1615.
92. **Agniswamy J, Nagiec MJ, Liu M, Schuck P, Musser JM, Sun PD.** 2006. Crystal structure of group A streptococcus Mac-1: insight into dimer-mediated specificity for recognition of human IgG. Structure **14:**225-235.
93. **von Pawel-Rammingen U.** 2012. Streptococcal IdeS and its impact on immune response and inflammation. Journal of Innate Immunity **4:**132-140.
94. **Akesson P, Moritz L, Truedsson M, Christensson B, von Pawel-Rammingen U.** 2006. IdeS, a highly specific immunoglobulin G (IgG)-cleaving enzyme from Streptococcus pyogenes, is inhibited by specific IgG antibodies generated during infection. Infection and Immunity **74:**497-503.
95. **Okada N, Pentland AP, Falk P, Caparon MG.** 1994. M protein and protein F act as important determinants of cell-specific tropism of *Streptococcus pyogenes* in skin tissue. Journal of Clinical Investigation **94:**965-977.

96. **Okada N, Liszewski MK, Atkinson JP, Caparon M.** 1995. Membrane cofactor protein (CD46) is a keratinocyte receptor for the M protein of the group A streptococcus. Proceedings of the National Academy of Sciences of the United States of America **92:**2489-2493.
97. **Bisno AL.** 1979. Alternate complement pathway activation by group A streptococci: role of M-protein. Infection and Immunity **26:**1172-1176.
98. **Johnsson E, Berggard K, Kotarsky H, Hellwage J, Zipfel PF, Sjobring U, Lindahl G.** 1998. Role of the hypervariable region in streptococcal M proteins: binding of a human complement inhibitor. Journal of Immunology **161:**4894-4901.
99. **Whitnack E, Beachey EH.** 1985. Inhibition of complement-mediated opsonization and phagocytosis of *Streptococcus pyogenes* by D fragments of fibrinogen and fibrin bound to cell surface M protein. Journal of Experimental Medicine **162:**1983-1997.
100. **Ghosh P.** 2011. The Nonideal Coiled Coil of M Protein and Its Multifarious Functions in Pathogenesis, p 197-211. *In* Linke D, Goldman A (ed), Bacterial Adhesion vol 715. Springer Dordrecht Heidelberg, New York, London.
101. **Bessen D, Jones KF, Fischetti VA.** 1989. Evidence for two distinct classes of streptococcal M protein and their relationship to rheumatic fever. Journal of Experimental Medicine **169:**269-283.
102. **Horstmann RD, Sievertsen HJ, Leippe M, Fischetti VA.** 1992. Role of fibrinogen in complement inhibition by streptococcal M protein. Infection and Immunity **60:**5036-5041.
103. **Carlsson F, Berggard K, Stalhammar-Carlemalm M, Lindahl G.** 2003. Evasion of phagocytosis through cooperation between two ligand-binding regions in Streptococcus pyogenes M protein. Journal of Experimental Medicine **198:**1057-1068.
104. **Pahlman LI, Morgelin M, Eckert J, Johansson L, Russell W, Riesbeck K, Soehnlein O, Lindbom L, Norrby-Teglund A, Schumann RR, Bjorck L, Herwald H.** 2006. Streptococcal M protein: a multipotent and powerful inducer of inflammation. Journal of Immunology **177:**1221-1228.
105. **Smeesters PR, McMillan DJ, Sriprakash KS.** 2010. The streptococcal M protein: a highly versatile molecule. Trends in Microbiology **18:**275-282.
106. **Pahlman LI, Olin AI, Darenberg J, Morgelin M, Kotb M, Herwald H, Norrby-Teglund A.** 2008. Soluble M1 protein of Streptococcus pyogenes triggers potent T cell activation. Cell Microbiology **10:**404-414.
107. **Whatmore AM, Kehoe MA.** 1994. Horizontal gene transfer in the evolution of group A streptococcal *emm*-like genes: gene mosaics and variation in Vir regulons. Molecular Microbiology **11:**363-374.
108. **Ji Y, Schnitzler N, DeMaster E, Cleary P.** 1998. Impact of M49, Mrp, Enn, and C5a peptidase proteins on colonization of the mouse oral mucosa by *Streptococcus pyogenes*. Infection and Immunity **66:**5399-5405.
109. **Podbielski A, Schnitzler N, Beyhs P, Boyle MD.** 1996. M-related protein (Mrp) contributes to group A streptococcal resistance to phagocytosis by human granulocytes. Molecular Microbiology **19:**429-441.

110. **Lukomski S, Nakashima K, Abdi I, Cipriano VJ, Ireland RM, Reid SD, Adams GG, Musser JM.** 2000. Identification and characterization of the *scl* gene encoding a group A streptococcus extracellular protein virulence factor with similarity to human collagen. Infection and Immunity **68:**6542-6553.
111. **Rasmussen R, Eden A, Björck L.** 2000. SclA, a novel collagen-like surface protein of *Streptococcus pyogenes*. Infection and Immunity **68:**6370-6377.
112. **Caswell CC, Lukomska E, Seo NS, Höök M, Lukomski S.** 2007. Scl1-dependent internalization of group A Streptococcus via direct interactions with the α2β(1) integrin enhances pathogen survival and re-emergence. Molecular Microbiology **64:**1319-1331.
113. **Dohrmann S, Anik S, Olson J, Anderson EL, Etesami N, No H, Snipper J, Nizet V, Okumura CY.** 2014. Role for streptococcal collagen-like protein 1 in M1T1 group A Streptococcus resistance to neutrophil extracellular traps. Infection and Immunity **82:**4011-4020.
114. **Oliver-Kozup H, Martin KH, Schwegler-Berry D, Green BJ, Betts C, Shinde AV, Van De Water L, Lukomski S.** 2013. The group A streptococcal collagen-like protein-1, Scl1, mediates biofilm formation by targeting the extra domain A-containing variant of cellular fibronectin expressed in wounded tissue. Molecular Microbiology **87:**672-689.
115. **Simpson WJ, LaPenta D, Chen C, Cleary PP.** 1990. Coregulation of type 12 M protein and streptococcal C5a peptidase genes in group A streptococci: evidence for a virulence regulon controlled by the *virR* locus. Journal of Bacteriology **172:**696-700.
116. **Wexler DE, Chenoweth DE, Cleary PP.** 1985. Mechanism of action of the group A streptococcal C5a inactivator. Proceedings of the National Academy of Sciences of the United States of America **82:**8144-8148.
117. **Cleary PP, Prahbu U, Dale JB, Wexler DE, Handley J.** 1992. Streptococcal C5a peptidase is a highly specific endopeptidase. Infection and Immunity **60:**5219-5223.
118. **Kagawa TF, O'Connell MR, Mouat P, Paoli M, O'Toole PW, Cooney JC.** 2009. Model for substrate interactions in C5a peptidase from Streptococcus pyogenes: A 1.9 A crystal structure of the active form of ScpA. Journal of Molecular Biology **386:**754-772.
119. **Ji Y, McLandsborough L, Kondagunta A, Cleary PP.** 1996. C5a peptidase alters clearance and trafficking of group A streptococci by infected mice. Infection and Immunity **64:**503-510.
120. **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:**7878-7886.
121. **Patti JM, Höök M.** 1994. Microbial adhesins recognizing extracellular matrix macromolecules. Curr Opin Cell Biol **6:**752-758.
122. **Kreikemeyer B, Oehmcke S, Nakata M, Hoffrogge R, Podbielski A.** 2004. *Streptococcus pyogenes* fibronectin-binding protein F2: expression profile, binding characteristics, and impact on eukaryotic cell interactions. Journal of Biological Chemistry **279:**15850-15859.

123. **LaPenta D, Rubens C, Chi E, Cleary PP.** 1994. Group A streptococci efficiently invade human respiratory epithelial cells. Proceedings of the National Academy of Sciences of the United States of America **91:**12115-12119.
124. **Molinari G, Talay SR, Valentin-Weigand P, Rohde M, Chhatwal GS.** 1997. The fibronectin-binding protein of *Streptococcus pyogenes*, SfbI, is involved in the internalization of group A streptococci by epithelial cells. Infection and Immunity **65:**1357-1363.
125. **Oehmcke S, Podbielski A, Kreikemeyer B.** 2004. Function of the fibronectin-binding serum opacity factor of *Streptococcus pyogenes* in adherence to epithelial cells. Infection and Immunity **72:**4302-4308.
126. **Rakonjac JV, Robbins JC, Fischetti VA.** 1995. DNA sequence of the serum opacity factor of group A streptococci: identification of a fibronectin-binding repeat domain. Infection and Immunity **63:**622-631.
127. **Courtney HS, Hasty DL, Li Y, Chiang HC, Thacker JL, Dale JB.** 1999. Serum opacity factor is a major fibronectin-binding protein and a virulence determinant of M type 2 *Streptococcus pyogenes*. Molecular Microbiology **32:**89-98.
128. **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:**5293-5307.
129. **Levin JC, Wessels MR.** 1998. Identification of *csrR*/*csrS*, a genetic locus that regulates hyaluronic acid capsule synthesis in group A streptococcus. Molecular Microbiology **30:**209-219.
130. **Lynskey NN, Goulding D, Gierula M, Turner CE, Dougan G, Edwards RJ, Sriskandan S.** 2013. RocA truncation underpins hyper-encapsulation, carriage longevity and transmissibility of serotype M18 group A streptococci. PLoS Pathog **9:**e1003842.
131. **Johnson DR, Stevens DL, Kaplan EL.** 1992. Epidemiologic analysis of group A streptococcal serotypes associated with severe systemic infections, rheumatic fever, or uncomplicated pharyngitis. Journal of Infectious Diseases **166:**374-382.
132. **Wessels MR, Bronze MS.** 1994. Critical role of the group A streptococcal capsule in pharyngeal colonization and infection in mice. Proceedings of the National Academy of Sciences of the United States of America **91:**12238-12242.
133. **Schrager HM, Rheinwald JG, Wessels MR.** 1996. Hyaluronic acid capsule and the role of streptococcal entry into keratinocytes in invasive skin infection. Journal of Clinical Investigation **98:**1954-1958.
134. **Stollerman GH, Dale JB.** 2008. The importance of the group a streptococcus capsule in the pathogenesis of human infections: a historical perspective. Clin Infectious Disease **46:**1038-1045.
135. **Datta V, Myskowski SM, Kwinn LA, Chiem DN, Varki N, Kansal RG, Kotb M, Nizet V.** 2005. Mutational analysis of the group A streptococcal

operon encoding streptolysin S and its virulence role in invasive infection. Molecular Microbiology **56:**681-695.
136. **Ginsburg I.** 1999. Is streptolysin S of group A streptococci a virulence factor? APMIS : Acta Pathologica, Microbiologica et Immunologica Scandinavica **107:**1051-1059.
137. **Ofek I, Zafriri D, Goldhar J, Eisenstein BI.** 1990. Inability of toxin inhibitors to neutralize enhanced toxicity caused by bacteria adherent to tissue culture cells. Infection and Immunity **58:**3737-3742.
138. **Keiser H, Weissmann G, Bernheimer AW.** 1964. Studies on Lysosomes. Iv. Solubilization of Enzymes During Mitochondrial Swelling and Disruption of Lysosomes by Streptolysin S and Other Hemolytic Agents. Journal of Cell Biology **22:**101-113.
139. **Ginsburg I.** 1972. Mechanisms of cell and tissue injury induced by group A streptococci: relation to poststreptococcal sequelae. Journal of Infectious Diseases **126:**294-340.
140. **Taketo Y, Taketo A.** 1966. Cytolytic effect of streptolysin S complex on Ehrlich ascites tumor cells. Journal of Biochemistry **60:**357-362.
141. **Hryniewicz W, Pryjma J.** 1977. Effect of streptolysin S on human and mouse T and B lymphocytes. Infection and Immunity **16:**730-733.
142. **Buchanan JT, Simpson AJ, Aziz RK, Liu GY, Kristian SA, Kotb M, Feramisco J, Nizet V.** 2006. DNase expression allows the pathogen group A Streptococcus to escape killing in neutrophil extracellular traps. Current Biology **16:**396-400.
143. **Aziz RK, Ismail SA, Park HW, Kotb M.** 2004. Post-proteomic identification of a novel phage-encoded streptodornase, Sda1, in invasive M1T1 Streptococcus pyogenes. Molecular Microbiology **54:**184-197.
144. **Walker MJ, Hollands A, Sanderson-Smith ML, Cole JN, Kirk JK, Henningham A, McArthur JD, Dinkla K, Aziz RK, Kansal RG, Simpson AJ, Buchanan JT, Chhatwal GS, Kotb M, Nizet V.** 2007. DNase Sda1 provides selection pressure for a switch to invasive group A streptococcal infection. Nat Med **13:**981-985.
145. **Sumby P, Barbian KD, Gardner DJ, Whitney AR, Welty DM, Long RD, Bailey JR, Parnell MJ, Hoe NP, Adams GG, Deleo FR, Musser JM.** 2005. Extracellular deoxyribonuclease made by group A streptococcus assists pathogenesis by enhancing evasion of the innate immune response. Proceedings of the National Academy of Sciences of the United States of America **102:**1679-1684.
146. **Venturini C, Ong CIY, Gillen CM, Ben-Zakour NL, Maamary PG, Nizet V, Beatson SA, Walker MJ.** 2013. Acquisition of the Sda1-Encoding Bacteriophage Does Not Enhance Virulence of the Serotype M1 Streptococcus pyogenes Strain SF370. Infection and Immunity **81:**2062-2069.
147. **Sriskandan S, Faulkner L, Hopkins P.** 2007. Streptococcus pyogenes: Insight into the function of the streptococcal superantigens. International Journal of Biochemistry and Cellular Biology **39:**12-19.
148. **Stock AM, Robinson VL, Goudreau PN.** 2000. Two-component signal transduction. Annual Reviews in Biochemistry **69:**183-215.

149. **Federle MJ, McIver KS, Scott JR.** 1999. A response regulator that represses transcription of several virulence operons in the group A streptococcus. Journal of Bacteriology **181:**3649-3657.
150. **Heath A, DiRita VJ, Barg NL, Engleberg NC.** 1999. A two-component regulatory system, CsrR-CsrS, represses expression of three *Streptococcus pyogenes* virulence factors, hyaluronic acid capsule, streptolysin S, and pyrogenic exotoxin B. Infection and Immunity **67:**5298-5305.
151. **Bernish B, van de Rijn I.** 1999. Characterization of a two-component system in *Streptococcus pyogenes* which is involved in regulation of hyaluronic acid production. J Biol Chem **274:**4786-4793.
152. **Dalton TL, Scott JR.** 2004. CovS inactivates CovR and is required for growth under conditions of general stress in *Streptococcus pyogenes*. Journal of Bacteriology **186:**3928-3937.
153. **Velarde JJ, Ashbaugh M, Wessels MR.** 2014. The Human Antimicrobial Peptide LL-37 Binds Directly to CsrS, a Sensor Histidine Kinase of Group A Streptococcus, to Activate Expression of Virulence Factors. Journal of Biological Chemistry **289:**36315-36324.
154. **Churchward G.** 2007. The two faces of Janus: virulence gene regulation by CovR/S in group A streptococci. Molecular Microbiology **64:**34-41.
155. **Gryllos I, Grifantini R, Colaprico A, Jiang S, Deforce E, Hakansson A, Telford JL, Grandi G, Wessels MR.** 2007. Mg(2+) signalling defines the group A streptococcal CsrRS (CovRS) regulon. Molecular Microbiology **65:**671-683.
156. **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:**13855-13860.
157. **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.
158. **Engleberg NC, Heath A, Miller A, Rivera C, DiRita VJ.** 2001. Spontaneous mutations in the CsrRS two-component regulatory system of *Streptococcus pyogenes* result in enhanced virulence in a murine model of skin and soft tissue infection. Journal of Infectious Diseases **183:**1043-1054.
159. **Federle MJ, Scott JR.** 2002. Identification of binding sites for the group A streptococcal global regulator CovR. Molecular Microbiology **43:**1161-1172.
160. **Leday TV, Gold KM, Kinkel TL, Roberts SA, Scott JR, McIver KS.** 2008. TrxR, a new CovR-repressed response regulator that activates the Mga virulence regulon in group A Streptococcus. Infection and Immunity **76:**4659-4668.
161. **Gold KM.** 2011. Characterization of the TrxSR two-component signal transduction system of *Streptococcus pyogenes* and its role in virulence regulation. Ph.D. University of Maryland, College Park.

162. **Chaussee MS, Ajdic D, Ferretti JJ.** 1999. The *rgg* gene of *Streptococcus pyogenes* NZ131 positively influences extracellular SPE B production. Infection and Immunity **67:**1715-1722.
163. **Chaussee MS, Sylva GL, Sturdevant DE, Smoot LM, Graham MR, Watson RO, Musser JM.** 2002. Rgg influences the expression of multiple regulatory loci to coregulate virulence factor expression in *Streptococcus pyogenes*. Infection and Immunity **70:**762-770.
164. **Chaussee MS, Somerville GA, Reitzer L, Musser JM.** 2003. Rgg coordinates virulence factor synthesis and metabolism in *Streptococcus pyogenes*. Journal of Bacteriology **185:**6016-6024.
165. **Chaussee MA, Callegari EA, Chaussee MS.** 2004. Rgg regulates growth phase-dependent expression of proteins associated with secondary metabolism and stress in *Streptococcus pyogenes*. Journal of Bacteriology **186:**7091-7099.
166. **Mashburn-Warren L, Morrison DA, Federle MJ.** 2010. A novel double-tryptophan peptide pheromone controls competence in Streptococcus spp. via an Rgg regulator. Molecular Microbiology **78:**589-606.
167. **Anbalagan S, McShan WM, Dunman PM, Chaussee MS.** 2011. Identification of Rgg binding sites in the Streptococcus pyogenes chromosome. Journal of bacteriology **193:**4933-4942.
168. **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.
169. **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.
170. **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.
171. **Podbielski A, Woischnik M, Leonard BA, Schmidt KH.** 1999. Characterization of *nra*, a global negative regulator gene in group A streptococci. Molecular Microbiology **31:**1051-1064.
172. **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.
173. **Kreikemeyer B, Nakata M, Koller 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.
174. **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:**1501-1511.
175. **Luo F, Lizano S, Bessen DE.** 2008. Heterogeneity in the Polarity of Nra Regulatory Effects on Streptococcal Pilus Gene Transcription and Virulence. Infection and Immunity.

176. **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:**1322-1329.
177. **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:**1506-1522.
178. **Roberts SA, Churchward GG, Scott JR.** 2007. Unraveling the regulatory network in *Streptococcus pyogenes*: the global response regulator CovR represses *rivR* directly. Journal of Bacteriology **189:**1459-1463.
179. **Trevino J, Liu Z, Cao TN, Ramirez-Pena E, Sumby P.** 2013. RivR is a negative regulator of virulence factor expression in group A Streptococcus. Infection and Immunity **81:**364-372.
180. **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.
181. **Deutscher J, Herro R, Bourand A, Mijakovic I, Poncet S.** 2005. P-Ser-HPr--a link between carbon metabolism and the virulence of some pathogenic bacteria. Biochimica Biophysica Acta **1754:**118-125.
182. **Stulke J, Hillen W.** 2000. Regulation of carbon catabolism in *Bacillus* species. Annu Rev Microbiol **54:**849-880.
183. **Titgemeyer F, Hillen W.** 2002. Global control of sugar metabolism: a gram-positive solution. Antonie Van Leeuwenhoek **82:**59-71.
184. **Fujita Y.** 2009. Carbon catabolite control of the metabolic network in Bacillus subtilis. Bioscience, Biotechnology, and Biochemistry **73:**245-259.
185. **Deutscher J.** 2008. The mechanisms of carbon catabolite repression in bacteria. Current Opinion in Microbiology **11:**87-93.
186. **Gorke B, Stulke J.** 2008. Carbon catabolite repression in bacteria: many ways to make the most out of nutrients. Nature Reviews Microbiology **6:**613-624.
187. **Milenbachs AA, Brown DP, Moors M, Youngman P.** 1997. Carbon-source regulation of virulence gene expression in *Listeria monocytogenes*. Molecular Microbiology **23:**1075-1085.
188. **Behari J, Youngman P.** 1998. A homolog of CcpA mediates catabolite control in Listeria monocytogenes but not carbon source regulation of virulence genes. Journal of Bacteriology **180:**6316-6324.
189. **Herro R, Poncet S, Cossart P, Buchrieser C, Gouin E, Glaser P, Deutscher J.** 2005. How seryl-phosphorylated HPr inhibits PrfA, a transcription activator of Listeria monocytogenes virulence genes. Journal of Molecular Microbiology and Biotechnology **9:**224-234.
190. **Varga J, Stirewalt VL, Melville SB.** 2004. The CcpA protein is necessary for efficient sporulation and enterotoxin gene (*cpe*) regulation in *Clostridium perfringens*. Journal of Bacteriology **186:**5221-5229.
191. **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:**5454-5461.

192. **Mendez M, Huang IH, Ohtani K, Grau R, Shimizu T, Sarker MR.** 2008. Carbon catabolite repression of type IV pilus-dependent gliding motility in the anaerobic pathogen *Clostridium perfringens*. Journal of Bacteriology **190:**48-60.
193. **Seidl K, Stucki M, Ruegg M, Goerke C, Wolz C, Harris L, Berger-Bachi B, Bischoff M.** 2006. Staphylococcus aureus CcpA affects virulence determinant production and antibiotic resistance. Antimicrobial Agents and Chemotherapy **50:**1183-1194.
194. **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:**8340-8349.
195. **Krzysciak W, Jurczak A, Koscielniak D, Bystrowska B, Skalniak A.** 2014. The virulence of Streptococcus mutans and the ability to form biofilms. European Journal of Clinical Microbiology and Infectious Disease **33:**499-515.
196. **Fujita Y, Miwa Y, Galinier A, Deutscher J.** 1995. Specific recognition of the *Bacillus subtilis gnt* cis-acting catabolite-responsive element by a protein complex formed between CcpA and seryl-phosphorylated HPr. Molecular Microbiology **17:**953-960.
197. **Shelburne SA, 3rd, Keith DB, Davenport MT, Horstmann N, Brennan RG, Musser JM.** 2008. Molecular characterization of group A Streptococcus maltodextrin catabolism and its role in pharyngitis. Molecular Microbiology **69:**436-452.
198. **Shelburne SA, 3rd, Sahasrobhajane P, Suber B, Keith DB, Davenport MT, Horstmann N, Kumaraswami M, Olsen RJ, Brennan RG, Musser JM.** 2011. Niche-specific contribution to streptococcal virulence of a MalR-regulated carbohydrate binding protein. Molecular Microbiology **81:**500-514.
199. **Steiner K, Malke H.** 2000. Life in protein-rich environments: the *relA*-independent response of *Streptococcus pyogenes* to amino acid starvation. Molecular Microbiology **38:**1004-1016.
200. **Malke H, Steiner K, McShan WM, Ferretti JJ.** 2006. Linking the nutritional status of *Streptococcus pyogenes* to alteration of transcriptional gene expression: the action of CodY and RelA. International Journal of Medicinal Microbiology **296:**259-275.
201. **Malke H, Ferretti JJ.** 2007. CodY-affected transcriptional gene expression of Streptococcus pyogenes during growth in human blood. Journal of Medical Microbiology **56:**707-714.
202. **Pine L, Reeves MW.** 1972. Correlation of M protein production with those factors found to influence growth and substrate utilization of *Streptococcus pyogenes*. Infection and Immunity **5:**668-680.
203. **Stulke J, Arnaud M, Rapoport G, Martin-Verstraete I.** 1998. PRD--a protein domain involved in PTS-dependent induction and carbon catabolite repression of catabolic operons in bacteria. Molecular Microbiology **28:**865-874.

204. **Schnetz K, Stulke J, Gertz S, Kruger S, Krieg M, Hecker M, Rak B.** 1996. LicT, a Bacillus subtilis transcriptional antiterminator protein of the BglG family. Journal of Bacteriology **178:**1971-1979.
205. **van Tilbeurgh H, Le Coq D, Declerck N.** 2001. Crystal structure of an activated form of the PTS regulation domain from the LicT transcriptional antiterminator. EMBO Journal **20:**3789-3799.
206. **Lindner C, Galinier A, Hecker M, Deutscher J.** 1999. Regulation of the activity of the Bacillus subtilis antiterminator LicT by multiple PEP-dependent, enzyme I- and HPr-catalysed phosphorylation. Molecular Microbiology **31:**995-1006.
207. **Tortosa P, Declerck N, Dutartre H, Lindner C, Deutscher J, Le Coq D.** 2001. Sites of positive and negative regulation in the *Bacillus subtilis* antiterminators LicT and SacY. Molecular Microbiology **41:**1381-1393.
208. **Ben-Zeev E, Fux L, Amster-Choder O, Eisenstein M.** 2005. Experimental and computational characterization of the dimerization of the PTS-regulation domains of BglG from Escherichia coli. Journal of molecular biology **347:**693-706.
209. **Henstra SA, Tuinhof M, Duurkens RH, Robillard GT.** 1999. The Bacillus stearothermophilus mannitol regulator, MtlR, of the phosphotransferase system. A DNA-binding protein, regulated by HPr and iicbmtl-dependent phosphorylation. Journal Biology Chemistry **274:**4754-4763.
210. **Henstra SA, Duurkens RH, Robillard GT.** 2000. Multiple phosphorylation events regulate the activity of the mannitol transcriptional regulator MtlR of the *Bacillus stearothermophilus* phosphoenolpyruvate-dependent mannitol phosphotransferase system. Journal of Biological Chemistry **275:**7037-7044.
211. **Bouraoui H, Ventroux M, Noirot-Gros MF, Deutscher J, Joyet P.** 2013. Membrane sequestration by the EIIB domain of the mannitol permease MtlA activates the Bacillus subtilis mtl operon regulator MtlR. Molecular Microbiology **87:**789-801.
212. **Mock M, Fouet A.** 2001. Anthrax. Annu Rev Microbiol **55:**647-671.
213. **Fouet A.** 2010. AtxA, a Bacillus anthracis global virulence regulator. Research Microbiology **161:**735-742.
214. **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:**644-655.
215. **Hammerstrom TG, Roh JH, Nikonowicz EP, Koehler TM.** 2011. *Bacillus anthracis* virulence regulator AtxA: oligomeric state, function and CO(2) - signalling. Molecular Microbiology **82:**634-647.
216. **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:**426-441.
217. **Osipiuk J, Wu R., Jedrzejczak, R., Moy, S., and Joachimiak, A.** 2011. Structure for the Putative Mga family transcriptional regulator from *Enterococcus faecalis*, Protein Data Bank (PDB).

218. **Ruiz-Cruz S, Espinosa M, Goldmann O, Bravo A.** 2015. Global Regulation of Gene Expression by the MafR Protein of Enterococcus faecalis. Frontiers in Microbiology **6:**1521.
219. **Podbielski A.** 1992. Ubiquitous occurrence of *virR* and *scpA* genes in group A streptococci. Medical Microbiol ogy and Immunology **181:**227-240.
220. **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:**4163-4172.
221. **Spanier JG, Jones SJ, Cleary P.** 1984. Small DNA deletions creating avirulence in *Streptococcus pyogenes*. Science **225:**935-938.
222. **Caparon MG, Scott JR.** 1987. Identification of a gene that regulates expression of M protein, the major virulence determinant of group A streptococci. Proceedings of the National Academy of Sciences of the United States of America **84:**8677-8681.
223. **Perez-Casal J, Caparon MG, Scott JR.** 1991. Mry, a trans-acting positive regulator of the M protein gene of *Streptococcus pyogenes* with similarity to the receptor proteins of two-component regulatory systems. Journal of Bacteriology **173:**2617-2624.
224. **Scott JR, Cleary P, Caparon MG, Kehoe M, Heden L, Musser JM, Hollingshead S, Podbielski A.** 1995. New name for the positive regulator of the M protein of group A streptococcus. Molecular Microbiology **17:**799.
225. **Shelburne SA, 3rd, Granville C, Tokuyama M, Sitkiewicz I, Patel P, Musser JM.** 2005. Growth characteristics of and virulence factor production by group A Streptococcus during cultivation in human saliva. Infection and Immunity **73:**4723-4731.
226. **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.
227. **Unnikrishnan M, Cohen J, Sriskandan S.** 1999. Growth-phase-dependent expression of virulence factors in an M1T1 clinical isolate of *Streptococcus pyogenes*. Infection and Immunity **67:**5495-5499.
228. **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:**9014-9019.
229. **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.
230. **Okada N, Geist RT, Caparon MG.** 1993. Positive transcriptional control of *mry* regulates virulence in the group A streptococcus. Molecular Microbiology **7:**893-903.

231. **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.
232. **Podbielski A, Flosdorff A, Weber-Heynemann J.** 1995. The group A streptococcal *virR49* gene controls expression of four structural *vir* regulon genes. Infection and Immunity **63:**9-20.
233. **Ribardo DA, Lambert TJ, McIver KS.** 2004. Role of *Streptococcus pyogenes* two-component response regulators in the temporal control of Mga and the Mga-regulated virulence gene *emm*. Infection and Immunity **72:**3668-3673.
234. **Haanes EJ, Heath DG, Cleary PP.** 1992. Architecture of the *vir* regulons of group A streptococci parallels opacity factor phenotype and M protein class. Journal of Bacteriology **174:**4967-4976.
235. **McIver KS, Heath AS, Green BD, Scott JR.** 1995. Specific binding of the activator Mga to promoter sequences of the *emm* and *scpA* genes in the group A streptococcus. Journal of Bacteriology **177:**6619-6624.
236. **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.
237. **Hondorp ER, S.C. H, 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:**953-967.
238. **McIver KS, Heath AS, Scott JR.** 1995. Regulation of virulence by environmental signals in group A streptococci: influence of osmolarity, temperature, gas exchange, and iron limitation on *emm* transcription. Infection and Immunity **63:**4540-4542.
239. **Caparon MG, Geist RT, Perez-Casal J, Scott JR.** 1992. Environmental regulation of virulence in group A streptococci: transcription of the gene encoding M protein is stimulated by carbon dioxide. Journal of Bacteriology **174:**5693-5701.
240. **Ribardo DA, McIver KS.** 2003. *amrA* encodes a putative membrane protein necessary for maximal exponential phase expression of the Mga virulence regulon in *Streptococcus pyogenes*. Molecular Microbiology **50:**673-685.
241. **Cox KH, Ruiz-Bustos E, Courtney HS, Dale JB, Pence MA, Nizet V, Aziz RK, Gerling I, Price SM, Hasty DL.** 2009. Inactivation of DltA modulates virulence factor expression in Streptococcus pyogenes. PLoS One **4:**e5366.
242. **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:**4904-4919.
243. **Almengor AC, McIver KS.** 2004. Transcriptional activation of *sclA* by Mga requires a distal binding site in *Streptococcus pyogenes*. Journal of Bacteriology **186:**7847-7857.
244. **Almengor AC, Walters MS, McIver KS.** 2006. Mga is sufficient to activate transcription *in vitro* of *sof/sfbX* and other Mga-regulated virulence genes in the group A streptococcus. Journal of Bacteriology **188:**2038-2047.

245. **Studier FW.** 2005. Protein production by auto-induction in high density shaking cultures. Protein Expression Purification **41:**207-234.
246. **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.
247. **Hanahan D, Meselson M.** 1983. Plasmid screening at high colony density. Methods in Enzymology **100:**333-342.
248. **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:**289-298.
249. **Ferretti JJ, McShan WM, Ajdic D, Savic DJ, Savic G, Lyon K, Primeaux C, Sezate S, Suvorov AN, Kenton S, Lai HS, Lin SP, Qian Y, Jia HG, Najar FZ, Ren Q, Zhu H, Song L, White J, Yuan X, Clifton SW, Roe BA, McLaughlin R.** 2001. Complete genome sequence of an M1 strain of *Streptococcus pyogenes*. Proceedings of the National Academy of Sciences of the United States of America **98:**4658-4663.
250. **Rose RE.** 1988. The nucleotide sequence of pACYC184. Nucleic acids research **16:**355.
251. **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:**953-967.
252. **Husmann LK, Scott JR, Lindahl G, Stenberg L.** 1995. Expression of the Arp protein, a member of the M protein family, is not sufficient to inhibit phagocytosis of *Streptococcus pyogenes*. Infection and Immunity **63:**345-348.
253. **Yanisch-Perron C, Vieira J, Messing J.** 1985. Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. Gene **33:**103-119.
254. **Le Breton Y, McIver KS.** 2013. Genetic Manipulation of *Streptococcus pyogenes* (The Group A Streptococcus, GAS). Current Protocols in Microbiology.
255. **Sung K, Khan SA, Nawaz MS, Khan AA.** 2003. A simple and efficient Triton X-100 boiling and chloroform extraction method of RNA isolation from Gram-positive and Gram-negative bacteria. FEMS Microbiology Letters **229:**97-101.
256. **Andrews S.** 2010. FASTQC. A quality control tool for high throughput sequence data, *on* Babraham Institute. http://www.bioinformatics.babraham.ac.uk/projects/fastqc/ Accessed
257. **Bolger AM, Lohse M, Usadel B.** 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics **30:**2114-2120.
258. **Langmead B.** 2010. Aligning short sequencing reads with Bowtie. Current Protocols in Bioinformatics **11**.
259. **Langmead B, Salzberg SL.** 2012. Fast gapped-read alignment with Bowtie 2. Nature methods **9:**357-359.
260. **Trapnell C, Roberts A, Goff L, Pertea G, Kim D, Kelley DR, Pimentel H, Salzberg SL, Rinn JL, Pachter L.** 2012. Differential gene and transcript

expression analysis of RNA-seq experiments with TopHat and Cufflinks. Nature protocols **7:**562-578.
261. **Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, Subgroup GPDP.** 2009. The Sequence Alignment/Map format and SAMtools. Bioinformatics **25:**2078-2079.
262. **Anders S, Pyl PT, Huber W.** 2015. HTSeq--a Python framework to work with high-throughput sequencing data. Bioinformatics **31:**166-169.
263. **Thorvaldsdóttir H, Robinson JT, Mesirov JP.** 2013. Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. Briefings in Bioinformatics **14:**178-192.
264. **Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, Smyth GK.** 2015. limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic acids research **43:**e47.
265. **Love MI, Huber W, Anders S.** 2014. Moderated estimation of fold change and dispersion for RNA-Seq data with DESeq2.
. bioRxiv doi:doi:10.1101/002832.
266. **Ogata H, Goto S, Sato K, Fujibuchi W, Bono H, Kanehisa M.** 1999. KEGG: Kyoto Encyclopedia of Genes and Genomes. Nucleic acids research **27:**29-34.
267. **Krzywinski M, Schein J, Birol I, Connors J, Gascoyne R, Horsman D, Jones SJ, Marra MA.** 2009. Circos: an information aesthetic for comparative genomics. Genome Research **19:**1639-1645.
268. **Young MD, Wakefield MJ, Smyth GK, Oshlack A.** 2010. Gene ontology analysis for RNA-seq: accounting for selection bias. Genome Biology **11:**R14.
269. **Beissbarth T, Speed TP.** 2004. GOstat: find statistically overrepresented Gene Ontologies within a group of genes. Bioinformatics **20:**1464-1465.
270. **Alexa A, Rahnenfuhrer J, Lengauer T.** 2006. Improved scoring of functional groups from gene expression data by decorrelating GO graph structure. Bioinformatics **22:**1600-1607.
271. **Lancefield RC.** 1957. Differentiation of group A streptococci with a common R antigen into three serological types, with special reference to the bactericidal test. Journal of Experimental Medicine **106:**525-544.
272. **Loo CY, Mitrakul K, Voss IB, Hughes CV, Ganeshkumar N.** 2003. Involvement of an inducible fructose phosphotransferase operon in Streptococcus gordonii biofilm formation. Journal of Bacteriology **185:**6241-6254.
273. **Pridgeon JW, Li Y, Yildirim-Aksoy M, Song L, Klesius PH, Srivastava KK, Reddy PG.** 2013. Fitness cost, gyrB mutation, and absence of phosphotransferase system fructose specific IIABC component in novobiocin-resistant Streptococcus iniae vaccine strain ISNO. Veterinary Microbiology **165:**384-391.
274. **Sachla AJ, Le Breton Y, Akhter F, McIver KS, Eichenbaum Z.** 2014. The crimson conundrum: heme toxicity and tolerance in GAS. Frontiers in Cellular Infection and Microbiology **4:**159.

275. **Barriere C, Veiga-da-Cunha M, Pons N, Guedon E, van Hijum SA, Kok J, Kuipers OP, Ehrlich DS, Renault P.** 2005. Fructose utilization in *Lactococcus lactis* as a model for low-GC gram-positive bacteria: its regulator, signal, and DNA-binding site. Journal of Bacteriology **187:**3752-3761.
276. **LaSarre B, Federle MJ.** 2011. Regulation and consequence of serine catabolism in Streptococcus pyogenes. Journal of Bacteriology **193:**2002-2012.
277. **Savic DJ, McShan WM, Ferretti JJ.** 2002. Autonomous expression of the slo gene of the bicistronic nga-slo operon of Streptococcus pyogenes. Infection and Immunity **70:**2730-2733.
278. **Wen ZT, Browngardt C, Burne RA.** 2001. Characterization of two operons that encode components of fructose-specific enzyme II of the sugar:phosphotransferase system of Streptococcus mutans. FEMS Microbiology Letters **205:**337-342.
279. **Kolaczkowska E, Kubes P.** 2013. Neutrophil recruitment and function in health and inflammation. Nature reviews Immunology **13:**159-175.
280. **Ajdic D, Pham VT.** 2007. Global transcriptional analysis of Streptococcus mutans sugar transporters using microarrays. Journal of Bacteriology **189:**5049-5059.
281. **Barrangou R, Azcarate-Peril MA, Duong T, Conners SB, Kelly RM, Klaenhammer TR.** 2006. Global analysis of carbohydrate utilization by *Lactobacillus acidophilus* using cDNA microarrays. Proceedings of the National Academy of Sciences of the United States of America **103:**3816-3821.
282. **Kelley LA, Sternberg MJ.** 2009. Protein structure prediction on the Web: a case study using the Phyre server. Nature Protocols **4:**363-371.
283. **Voigt C, Bahl H, Fischer RJ.** 2014. Identification of PTS(Fru) as the major fructose uptake system of Clostridium acetobutylicum. Applied Microbiology and Biotechnology **98:**7161-7172.
284. **Chavarria M, Durante-Rodriguez G, Krell T, Santiago C, Brezovsky J, Damborsky J, de Lorenzo V.** 2014. Fructose 1-phosphate is the one and only physiological effector of the Cra (FruR) regulator of Pseudomonas putida. FEBS Open Bio **4:**377-386.
285. **Chavarria M, Santiago C, Platero R, Krell T, Casasnovas JM, de Lorenzo V.** 2011. Fructose 1-phosphate is the preferred effector of the metabolic regulator Cra of Pseudomonas putida. Journal of Biological Chemistry **286:**9351-9359.
286. **Reizer J, Hoischen C, Titgemeyer F, Rivolta C, Rabus R, Stulke J, Karamata D, Saier MH, Jr., Hillen W.** 1998. A novel protein kinase that controls carbon catabolite repression in bacteria. Molecular Microbiology **27:**1157-1169.
287. **Kornberg HL.** 2001. Routes for fructose utilization by Escherichia coli. Journal of Molecular Microbiology and Biotechnology **3:**355-359.

288. **Medina E, Lengeling A.** 2005. Genetic regulation of host responses to group A streptococcus in mice. Briefings in Functional Genomics and Proteomics **4:**248-257.
289. **Munoz-Elias EJ, McKinney JD.** 2005. *Mycobacterium tuberculosis* isocitrate lyases 1 and 2 are jointly required for in vivo growth and virulence. Nature Medicine **11:**638-644.
290. **Iyer R, Camilli A.** 2007. Sucrose metabolism contributes to in vivo fitness of *Streptococcus pneumoniae*. Molecular Microbiology **66:**1-13.
291. **Son MS, Matthews WJ, Jr., Kang Y, Nguyen DT, Hoang TT.** 2007. In vivo evidence of *Pseudomonas aeruginosa* nutrient acquisition and pathogenesis in the lungs of cystic fibrosis patients. Infection and Immunity **75:**5313-5324.
292. **Rollenhagen C, Bumann D.** 2006. *Salmonella enterica* highly expressed genes are disease specific. Infection and Immunity **74:**1649-1660.
293. **Moyrand F, Fontaine T, Janbon G.** 2007. Systematic capsule gene disruption reveals the central role of galactose metabolism on *Cryptococcus neoformans* virulence. Molecular Microbiology **64:**771-781.
294. **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:**927-942.
295. **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.
296. **Aziz RK, Edwards RA, Taylor WW, Low DE, McGeer A, Kotb M.** 2005. Mosaic prophages with horizontally acquired genes account for the emergence and diversification of the globally disseminated M1T1 clone of *Streptococcus pyogenes*. Journal of Bacteriology **187:**3311-3318.
297. **Loughman JA, Caparon MG.** 2007. Comparative functional analysis of the *lac* operons in *Streptococcus pyogenes*. Molecular Microbiology **64:**269-280.
298. **Kazmi SU, Kansal R, Aziz RK, Hooshdaran M, Norrby-Teglund A, Low DE, Halim AB, Kotb M.** 2001. Reciprocal, temporal expression of SpeA and SpeB by invasive M1T1 group a streptococcal isolates in vivo. Infection and Immunity **69:**4988-4995.
299. **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:**771-782.
300. **Hirst RA, Sikand KS, Rutman A, Mitchell TJ, Andrew PW, O'Callaghan C.** 2000. Relative roles of pneumolysin and hydrogen peroxide from Streptococcus pneumoniae in inhibition of ependymal ciliary beat frequency. Infection and Immunity **68:**1557-1562.
301. **Pericone CD, Overweg K, Hermans PW, Weiser JN.** 2000. Inhibitory and bactericidal effects of hydrogen peroxide production by Streptococcus

pneumoniae on other inhabitants of the upper respiratory tract. Infection and Immunity **68:**3990-3997.
302. **Regev-Yochay G, Trzcinski K, Thompson CM, Malley R, Lipsitch M.** 2006. Interference between Streptococcus pneumoniae and Staphylococcus aureus: In vitro hydrogen peroxide-mediated killing by Streptococcus pneumoniae. Journal of Bacteriology **188:**4996-5001.
303. **Anbalagan S, Chaussee MS.** 2013. Transcriptional regulation of a bacteriophage encoded extracellular DNase (Spd-3) by Rgg in Streptococcus pyogenes. PLoS One **8:**e61312.
304. **Maamary PG, Ben Zakour NL, Cole JN, Hollands A, Aziz RK, Barnett TC, Cork AJ, Henningham A, Sanderson-Smith M, McArthur JD, Venturini C, Gillen CM, Kirk JK, Johnson DR, Taylor WL, Kaplan EL, Kotb M, Nizet V, Beatson SA, Walker MJ.** 2012. Tracing the evolutionary history of the pandemic group A streptococcal M1T1 clone. FASEB Journal **26:**4675-4684.
305. **Horstmann N, Sahasrabhojane P, Saldana M, Ajami NJ, Flores AR, Sumby P, Liu CG, Yao H, Su X, Thompson E, Shelburne SA.** 2015. Characterization of the effect of the histidine kinase CovS on response regulator phosphorylation in group A Streptococcus. Infection and Immunity **83:**1068-1077.
306. **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.
307. **Andersson G, McIver K, Heden LO, Scott JR.** 1996. Complementation of divergent *mga* genes in group A streptococcus. Gene **175:**77-81.
308. **Vahling CM, McIver KS.** 2006. Domains required for transcriptional activation show conservation in the Mga family of virulence gene regulators. Journal of Bacteriology **188:**863-873.
309. **Yung DL, McIver KS, Scott JR, Hollingshead SK.** 1999. Attenuated expression of the *mga* virulence regulon in an M serotype 50 mouse-virulent group A streptococcal strain. Infection and Immunity **67:**6691-6694.
310. **Kruszewski AM, Hausse LL, Hondorp EH.** 2013. The Group A Streptococcus Mga virulence regulator: determination of C-terminal amino acids respnsible for oligomeriation and transcriptional activation. Bachelors of Science. University of Maryland, College Park.
311. **Karimova G, Ullmann A, Ladant D.** 2000. A bacterial two-hybrid system that exploits a cAMP signaling cascade in Escherichia coli. Methods in Enzymology **328:**59-73.
312. **Xu J, Li M, Kim D, Xu Y.** 2003. RAPTOR: optimal protein threading by linear programming. Jorunal of Bioinformatics and Computational Biology **1:**95-117.
313. **Zhang Y.** 2008. I-TASSER server for protein 3D structure prediction. BMC Bioinformatics **9:**40.
314. **Chivian D, Kim DE, Malmstrom L, Bradley P, Robertson T, Murphy P, Strauss CE, Bonneau R, Rohl CA, Baker D.** 2003. Automated prediction of CASP-5 structures using the Robetta server. Proteins **53 Suppl 6:**524-533.

315. **Yung DL, Hollingshead SK.** 1996. DNA sequencing and gene expression of the *emm* gene cluster in an M50 group A streptococcus strain virulent for mice. Infection and Immunity **64:**2193-2200.
316. **Vahling CM, McIver KS.** 2005. Identification of residues responsible for the defective virulence gene regulator Mga produced by a natural mutant of *Streptococcus pyogenes*. Journal of Bacteriology **187**.
317. **Van den Bogert B, Boekhorst J, Herrmann R, Smid EJ, Zoetendal EG, Kleerebezem M.** 2013. Comparative genomics analysis of Streptococcus isolates from the human small intestine reveals their adaptation to a highly dynamic ecosystem. PloS One **8:**e83418.
318. **Kreft J, Vazquez-Boland JA.** 2001. Regulation of virulence genes in Listeria. International Journal of Medical Microbiology **291:**145-157.
319. **Ferenci T, Kornberg HL.** 1971. Role of fructose-1,6-diphosphatase in fructose utilization by Escherichia coli. FEBS Letters **14:**360-363.
320. **Kim HK, Missiakas D, Schneewind O.** 2014. Mouse models for infectious diseases caused by Staphylococcus aureus. Journal of Immunology Methods **410:**88-99.
321. **Mestas J, Hughes CC.** 2004. Of mice and not men: differences between mouse and human immunology. Journal of Immunology **172:**2731-2738.
322. **Kugelberg E, Norstrom T, Petersen TK, Duvold T, Andersson DI, Hughes D.** 2005. Establishment of a superficial skin infection model in mice by using Staphylococcus aureus and Streptococcus pyogenes. Antimicrobial Agents and Chemotherapy **49:**3435-3441.
323. **Roberts S, Scott JR, Husmann LK, Zurawski CA.** 2006. Murine models of Streptococcus pyogenes infection. Curr Protoc Microbiol **Chapter 9:**Unit 9D 5.
324. **Scaramuzzino DA, McNiff JM, Bessen DE.** 2000. Humanized in vivo model for streptococcal impetigo. Infection and Immunity **68:**2880-2887.
325. **Gladysheva IP, Turner RB, Sazonova IY, Liu L, Reed GL.** 2003. Coevolutionary patterns in plasminogen activation. Proceedings of the National Academy of Sciences of the United States of America **100:**9168-9172.
326. **Khil J, Im M, Heath A, Ringdahl U, Mundada L, Cary Engleberg N, Fay WP.** 2003. Plasminogen enhances virulence of group A streptococci by streptokinase-dependent and streptokinase-independent mechanisms. Journal of Infectious Diseases **188:**497-505.
327. **Accardo P, Sanchez-Corral P, Criado O, Garcia E, Rodriguez de Cordoba S.** 1996. Binding of human complement component C4b-binding protein (C4BP) to *Streptococcus pyogenes* involves the C4b-binding site. Journal of Immunology **157:**4935-4939.
328. **McArthur JD, Walker MJ.** 2006. Domains of group A streptococcal M protein that confer resistance to phagocytosis, opsonization and protection: implications for vaccine development. Molecular microbiology **59:**1-4.
329. **Sun H, Ringdahl U, Homeister JW, Fay WP, Engleberg NC, Yang AY, Rozek LS, Wang X, Sjobring U, Ginsburg D.** 2004. Plasminogen is a

critical host pathogenicity factor for group A streptococcal infection. Science **305:**1283-1286.
330. **Ermert D, Shaughnessy J, Joeris T, Kaplan J, Pang CJ, Kurt-Jones EA, Rice PA, Ram S, Blom AM.** 2015. Virulence of Group A Streptococci Is Enhanced by Human Complement Inhibitors. PLoS Pathogens **11:**e1005043.
331. **Proft T, Moffatt SL, Berkahn CJ, Fraser JD.** 1999. Identification and characterization of novel superantigens from Streptococcus pyogenes. Journal of Experimental Medicine **189:**89-102.
332. **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.
333. **Vadeboncoeur C, Pelletier M.** 1997. The phosphoenolpyruvate:sugar phosphotransferase system of oral streptococci and its role in the control of sugar metabolism. FEMS Microbiology Reviews **19:**187-207.
334. **Cvitkovitch DG, Boyd DA, Thevenot T, Hamilton IR.** 1995. Glucose transport by a mutant of Streptococcus mutans unable to accumulate sugars via the phosphoenolpyruvate phosphotransferase system. Journal of Bacteriology **177:**2251-2258.
335. **Bidossi A, Mulas L, Decorosi F, Colomba L, Ricci S, Pozzi G, Deutscher J, Viti C, Oggioni MR.** 2012. A functional genomics approach to establish the complement of carbohydrate transporters in Streptococcus pneumoniae. PLoS One **7:**e33320.
336. **Kant S, Agarwal S, Pancholi P, Pancholi V.** 2015. The Streptococcus pyogenes orphan protein tyrosine phosphatase, SP-PTP, possesses dual specificity and essential virulence regulatory functions. Molecular Microbiology **97:**515-540.
337. **Greenberg DB, Stulke J, Saier MH, Jr.** 2002. Domain analysis of transcriptional regulators bearing PTS regulatory domains. Research in Microbiology **153:**519-526.