FUNCTIONAL DOMAINS IN THE MULTIGENE REGULATOR OF THE GROUP A STREPTOCOCCUS

APPROVED BY SUPERVISORY COMMITTEE

Kevin S. McIver, Ph.D., Assistant Professor
Department of Microbiology
U.T. Southwestern Medical Center at Dallas, Dallas, Texas

Eric J. Hansen, Ph.D., Professor
Department of Microbiology
U.T. Southwestern Medical Center at Dallas, Dallas, Texas

David Hendrixson, Ph.D., Assistant Professor
Department of Microbiology
U.T. Southwestern Medical Center at Dallas, Dallas, Texas

Margaret A. Phillips, Ph.D., Professor
Department of Pharmacology
U.T. Southwestern Medical Center at Dallas, Dallas, Texas

DEDICATION

To my mother, who has always been my shining light on the sea of life and has never steered me wrong. You will always be my role model.

I would like to thank my mentor, Dr. Kevin McIver, for his patience and persistence in answering all of my questions, the members of my graduate committee: Dr. Simon Daefler, Dr. Eric Hansen, and Dr. Margaret Phillips for their guidance and helpful insight with all of my projects and the newest member Dr. David Hendrixson, who was not afraid to jump aboard at the last moment. In addition, I would like to thank my family and friends that have been there to support me every step of the way.

FUNCTIONAL DOMAINS IN THE MULTIGENE REGULATOR OF THE GROUP A STREPTOCOCCUS

By

CHERYL M. VAHLING

DISSERTATION

Presented to the Faculty of the Graduate School of Biomedical Sciences

The University of Texas Southwestern Medical Center at Dallas

In Partial Fulfillment of the Requirements

For the Degree of

DOCTOR OF PHILOSOPHY

The University of Texas Southwestern Medical Center at Dallas

Dallas, Texas

April, 2006

FUNCTIONAL DOMAINS IN THE MULTIGENE REGULATOR OF THE GROUP A STREPTOCOCCUS

Publication No. _____

Cheryl M. Vahling, Ph.D.

The University of Texas Southwestern Medical Center at Dallas, 2006

Supervising Professor: Kevin S. McIver, Ph.D.

The group A streptococcus (GAS) is a human pathogen capable of causing a broad range of symptoms from mild erythema to severe tissue destruction. The multigene regulator of the GAS, or Mga, has been shown to bind DNA and activate transcription of virulence genes implicated in colonization and immune evasion in response to environmental conditions. Several different avenues of investigation were undertaken in the present study with an overall goal of identifying and characterizing functional domains within the

protein. The first section of this study uses a naturally occurring mutant, previously identified to be deficient in transcriptional activation, as a guide to map functional residues within Mga. Electrophoretic mobility shift assays using purified proteins determined that the defect was not a result of the inability to bind DNA, so gain-of-function mutants were used to restore transcriptional activity to the mutant and pinpoint the residues necessary for full activation. In the second section, a genetic screen was undertaken to identify novel domains within the protein. From this screen, a domain was found within the extreme N-terminus. Sequence homology revealed several proteins in other pathogenic streptococci that shared this motif. As a result, it was named CMD-1 for conserved Mga domain 1, while the group of proteins was subsequently designated as the Mga family of putative virulence gene regulators. Alanine scanning mutagenesis demonstrated the importance of CMD-1 for transcriptional activation in several members of this family. This established that the results yielded from investigations of Mga could be broadened beyond the GAS to provide global insight into possible mechanisms of virulence regulation in other pathogenic streptococci. Finally, an *in silico* analysis of Mga was performed, revealing several areas with differing degrees of structural similarity to known domains of other bacteria including a receiver domain and a PTS-regulatory domain. Each region was explored to determine if it was functionally active in Mga. Since both domains have been implicated in forming higher order structures, the oligomeric state of Mga was also determined. Overall, characterization of Mga has helped unravel the mechanisms of virulence regulation in pathogenic streptococci.

TABLE OF CONTENTS

PRIOR PUBLICATIONS

Vahling, C. M., and K. S. McIver. Colony Immunoblot for the Detection of Intracellular Proteins in Gram-Positive Bacteria. *Submitted.*

\Vahling, C. M., and K. S. McIver. 2006. Domains required for transcriptional activation show conservation in the Mga family of virulence gene regulators. Journal of Bacteriology **188:**863-73.

Vahling, C. M., and K. S. McIver. 2005. Identification of residues responsible for the defective virulence gene regulator Mga produced by a natural mutant of *Streptococcus pyogenes*. Journal of Bacteriology **187:**5955-66.

THE UNIVERSITY OF TEXAS
SOUTHWESTERN MEDICAL CENTER
AT DALLAS

Department of Microbiology

Feb. 28, 2006

Journals Department
American Society for Microbiology
1752 N Street, N.W.
Washington, DC 20036-2904

To whom it may concern:

I am writing to request permission to reuse the information in the two papers listed below for my dissertation. This work, entitled "Functional Domains in the Multigene Regulator of the Group A Streptococcus," will be used to fulfill the requirements set forth by the University of Texas Southwestern Medical Center for graduation from the Molecular Microbiology Program.

1) **Vahling, C. M., and K. S. McIver.** 2005. Identification of residues responsible for the defective virulence gene regulator Mga produced by a natural mutant of Streptococcus pyogenes. Journal of Bacteriology **187:**5955-66.

2) **Vahling, C. M., and K. S. McIver.** 2006. Domains required for transcriptional activation show conservation in the mga family of virulence gene regulators. Journal of Bacteriology **188:**863-73.

Thank you for attention to this matter,

Cheryl Vahling

Cheryl Vahling

PERMISSION GRANTED
CONTINGENT ON AUTHOR PERMISSION (which you must obtain)
AND APPROPRIATE CREDIT
American Society for Microbiology
Journals Department
Date 2-28-06

UT Southwestern Medical Center 5323 Harry Hines Blvd., NA6.300 / Dallas, TX 75390-9048
Phone: (214) 648-1255 Fax: (214) 648-5907 Email: kevin.mciver@utsouthwestern.edu

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

A_{420} – absorbance at 420 nanometers

aa – amino acid

Ala – alanine

AmrA – activation of the Mga regulon

ARF – acute rheumatic fever

Arg – arginine

Asn – asparagine

Asp – aspartic acid

bp– base pair

BHI – brain heart infusion

B. subtilis – Bacillus subtilis

C4BP – C4b-binding protein

CcpA – catabolite control protein A

CheY – chemotactic protein Y

CMD-1 – conserved Mga domain 1

COMPASS – comparative mapping by annotation and sequence similarity

c.p.m. – counts per minute

cre– catabolite response element

Da – dalton

DNA – deoxyribonucleic acid

DTT – dithiothreitol

dH_2O – deionized water

EI – enzyme I

EII – enzyme II

E. coli – *Escherichia coli*

EDTA – ethylenediamine tetracetic acid disodium salt

EMSA – electrophoretic mobility shift assay

EtOH – ethanol

FHL-1 – factor H-like protein

GAS – group A Streptococcus

GCS – group C Streptococcus

GusA – β-glucuronidase

Glu – glutamic acid

HEPES – N-2-hydroxyethylpiperazine-N′-2-ethanesulfonic acid

His – histidine

HPr – heat stable protein

HTH – helix-turn helix motif

HVR – N-terminal hypervariable region

IgA – immunoglobulin A

IgG – immunoglobulin G

IPTG – isopropyl-β-D-galactopyranoside

Iso – isoleucine

kb – kilobase pair

LB – Luria-Bertani broth

Lys – lysine

MBP – maltose binding proteins

Mga – multigene regulator of the group A Streptococcus

Mrp – M-related protein

mry – M protein RNA yield

NEB – New England Biolabs

Nra – negative regulator of GAS

NtrC – nitrogen regulator I

OD – optical density

P(gene name) – promoter of the named gene

PAGE – polyacrylamide gel electrophoresis

PAM – plasminogen-binding group A streptococcal M-like protein

PBS – phosphate buffer solution

PCR – polymerase chain reaction

PEP – phosphoenolpyruvate

PMNs – polymorphonuclear leucocytes

PRD – PTS-regulatory domain

Pro – proline

PTS – phosphoenolpyruvate:sugar phosphotransferase system

RcsB – regulator of capsule synthesis

RD-1 – receiver domain 1

RD-2 – receiver domain 2

RNA – ribonucleic acid

RNAP – RNA polymerase

RT – room temperature

rpm – revolutions per minute

SDS – sodium dodecyl sulfate

Slo – streptolysin O

Sls – streptolysin S

Spe – streptococcal pyrogenic exotoxins

S. pyogenes – *Streptococcus pyogenes*

STSS – streptococcal toxic shock-like syndrome

S. typhimurium – *Salmonella typhimurium*

THY – Todd-Hewitt medium supplemented with 0.2% yeast extract

v – volume

virR – virulence regulator

VIT – vectors for integration

w – weight

WT – wild type

X-glu – pNitrophenyl β-D glucuronide

CHAPTER ONE:

INTRODUCTION

The group A streptococcus (GAS) or *Streptococcus pyogenes* is a Gram-positive pathogen that can elicit a wide variety of diseases at numerous locations throughout its human host. The severity of a GAS infection can range from an asymptomatic carrier state to life threatening invasive disease. With over a million cases of GAS pharyngitis and 10,000-15,000 cases of invasive GAS disease occurring per year in the United States alone (http://www.cdc.gov/ncidod/biotech/strep/doc.htm), understanding how GAS is able to infect and overcome the defenses of a healthy immune system is important for the development of novel vaccines and drugs to combat this pathogen.

Interpretation of studies involving a GAS infection is complicated by the large number of virulence genes that it possesses and their regulation at the different stages of disease progression. Thus, elucidation of how the GAS regulates its arsenal of virulence genes is essential not only for a complete understanding of how the bacteria cause disease but also to shed light on mechanisms that may prevent an infection from becoming severe, eliminate associated sequelae and identify specific targets for drug therapy. One common mechanism of gene regulation is through global transcriptional regulators, which act as an efficient way to simultaneously activate or inactivate transcription of groups of genes in response to a particular stimulus. The present study looks at one such

global regulator in the GAS known as the multigene regulator of the group A streptococcus or Mga. Previously, it has been shown that in different GAS serotypes, Mga regulates as many as nine different virulence genes involved in colonization and immune evasion of the host in a growth phase-dependent manner (126). Exactly how Mga is able to sense its environment and interact with only the specific subset of virulence genes that it regulates is of particular interest. Here, mutational analysis of this protein was undertaken to help unravel its mechanism of control.

This study begins with a further characterization of the *mga-2* allele found in B514, an M50 serotype strain naturally deficient in activation of genes regulated by Mga. Using gain-of-function mutants and a heterologous GAS background, the particular amino acids of Mga involved in the transcriptional deficiency are defined and the conclusion that Mga alone is responsible for the deficiency is supported. Differences between the two divergent alleles of *mga* are also illuminated by demonstrating that an amino acid substitution equivalent to the one found to make M50 Mga dysfunctional does not lead to a loss of function in an M6 Mga, which is encoded by the *mga-1* allele.

This report also describes a novel Mga domain involved in the transcriptional activation of Mga-regulated promoters identified using a random PCR mutagenesis screen. A sequence alignment reveals that this domain exists not only in all GAS Mga proteins but is conserved amongst orthologues of Mga found in several other pathogenic streptococci. Mutational analysis provides evidence that this domain is functionally

equivalent in orthologues implying that mechanisms of activation common to all members of the Mga family of virulence gene regulators may also exist.

In addition, an initial investigation into two putative domains of the Mga protein is also provided. The first motif shows homology to a receiver domain of the response regulator of a two-component system while the second is homologous to a phosphoenolpyruvate phosphotransferase system regulatory domain. Taken as a whole, the data lends insight into the regulatory processes by which Mga governs the expression of important virulence genes in the progression of a GAS infection.

CHAPTER TWO:

REVIEW OF THE LITERATURE

I. Identification and Classification of Streptococci

A. Classical features of the *Streptococcus*

The genus *Streptococcus* is a large collection of Gram-positive non-motile cocci, which are often found growing in pairs or chains. These facultative anaerobes grow by fermentation regardless of the presence of oxygen and are differentiated from most *Staphylococci* because they are catalase-negative. Each species within the genus differs in its requirements for optimum growth and demonstrates a defining hemolytic reaction on blood agar plates. Although not all members are pathogenic, one unifying theme of those that are is their ability to colonize a broad range of tissue sites in humans and/or animals (63), causing a wide variety of infections throughout the host (Table 1).

B. Identification based on hemolytic activity

Over time, multiple methods have arisen in an attempt to categorize the different species of streptococci. One of the earliest classification schemes was based on the ability or inability of the bacteria to lyse red blood cells (19). This method separates the

Table 1: Species classification of commonly isolated pathogenic streptococci

Species	Lancefield Group	Pattern of Hemolysis	Origin	Commonly Caused Infections
S. pyogenes	A	β	Human	Variety of respiratory, skin and other infections. Also causes sequelae.
S. agalactiae	B	β, γ	Human, Bovine	Sepsis and meningitis in neonates. Invasive infections in susceptible adults.
S. pneumoniae	Not applicable	α	Human	Respiratory infections, otitis media and meningitis
S. anginosus (group)	A, C, F, G or None	α, β, γ	Human	Purulent infections
S. bovis (group)	D	α, γ	Human	Endocarditis and bacteremia
S. mutans (group)	Not applicable	α, γ, β	Human	Dental caries and endocarditis
S. salivarius (group)	Not applicable	α,	Human	Endocarditis and meningitis in compromised hosts
S. mitis (group)	Not applicable	α	Human	Endocarditis and bacteremia in neutropenic patients
S. dysgalactiae				
subsp. *equisimilis*	A, C, G, L	β	Human, Animal	Cellulitis, septicemia, respiratory and deep tissue infections
subsp. *dysgalactiae*	C, L	α, γ	Animal	Mastitis in cows and suppurative polyarthritis in lambs
S. equi				
subsp. *equi*	C	β	Animal	Strangles in horses
subsp. *zooepidemicus*	C	β	Animal, Human	Bovine mastitis, equine respiratory infections and infertility. Nephritis in humans
S. canis	G	β	Dog, Human	Necrotizing fasciitis and STSS in dogs. Sepsis in humans.
S. porcinus	E, P, U, V or None	β	Swine, Human	Cervical lymph node infections in swine. Stillbirths in humans.
S. iniae	None	β	Dolphin, Fish Human	Skin lesions in dolphins and meningitis in fish. Bacteremia and cutaneous infections in humans.
S. phocae	C, F	β	Seal	Pneumonia and sinusitis in seals
S. didelphis	None	β	Opossum	Suppurative dermatitis, septicemia and hepatic fibrosis in opossums
S. suis	R, S, T	α	Swine, Human	Arthritis and pneumonia in pigs. Meningitis in humans and pigs.

bacteria by the three different phenotypes seen when plated on blood-containing agar: 1) a zone of clearing caused by complete lysis of the cells (β-hemolytic); 2) a greenish colored area caused by a partial destruction of red blood cells (α-hemolytic); and 3) an inability to lyse the red blood cells (γ- or nonhemolytic).

Even though this system is still used by both clinicians and taxonomists, several caveats do exist. For example, ambiguous results may be seen with certain strains such as *S. suis*, which show α-hemolysis on sheep blood agar but demonstrate β-hemolysis on horse blood agar (97). Differentiating between the three types of hemolytic reactions can be important because in humans β-hemolysis is the dominant pattern seen with infectious strains (Table 1).

C. Sherman and Lancefield's classification methodologies

Sherman used not only hemolysis but also other biochemical properties such as tolerance to salt, high pH and temperature to subdivide the streptococci into four groups (205). His four classes were named pyogenic, viridans, enterococcal and lactic. The pyogenic division remains relatively the same today although members of many of the other divisions have been reclassified.

Rebecca Lancefield devised a different classification system that used immunological properties as a way to subdivide the most prevalent group pathogenic to humans or the β-hemolytic streptococci (Table 1) (132). This system is based on the different carbohydrates that exist within the bacterial cell wall. Again, this system is not

without its flaws including the inability to identify non-β-hemolytic streptococci because these isolate may display a variety of antigens or none at all (Table 1). An additional flaw to the system is that the groupings do not always correspond with the genetic relatedness as predicted by 16S rRNA (20). Regardless, most human pathogenic strains contain the group A carbohydrate and thus the name the GAS.

D. Classification based on M protein

Due to their clinical importance, Lancefield also developed a system for classifying members of the group A streptococci based on the surface protein known as the M protein. To type the different species, the M protein was first solublized by treating the bacteria with a solution of boiling hydrochloric acid and then used in an antibody precipitin test (130). Structurally, the N-terminal portion of the M protein is an alpha-helical coiled-coil dimer (175) that extends outwards from the surface of the cell, while the C-terminus acts as the membrane anchor (71). Within the N-terminus is a hypervariable region (HRV), which can differ from strain to strain. Antibodies produced against this region are type-specific and form the basis of the antibody test (70, 131). Currently, 124 different M-genotypes have been identified (64), which have been subdivided into two classes (class I and class II) to reflect the ability to react with antibodies against a more conserved domain within the surface-exposed region of the M protein (15). M classes tightly correlate with virulence (133) and serotypes containing class I *emm* genes (e.g., M1, M6 and M24) are generally associated with throat infections whereas those possessing class II *emm* genes (M2, M4 and M49) are associated with both

skin and throat infections, although these lines cannot be considered absolute (17, 113). Recently, molecular methods have been used to sequence the *emm* gene, which encodes for the M protein. The nucleotide differences within the sequences reflect the specific genotypes to which each bacterium belongs. This method has replaced the need for every laboratory to possess antisera to every M type and has eliminated the difficulties involved in sample solubility.

E. T-protein slide agglutination method

In 1934 Griffith adopted a separate approach to differentiate amongst the streptococci. He used a different surface protein, known as T protein, for the basis of his slide agglutination test to type the organisms. This method is used on whole cells, thus eliminating the need for soluble antigen and it is also not limited to typing only the group A streptococci (111, 133). Unfortunately, the M and T typing systems do not always correlate. In general, there appears to be a greater diversity among *tee* alleles compared to *emm* alleles (111) and unlike the M antigen, the T antigen does not seem to correlate with virulence (58). Regardless of the system used, a particular subtype has yet to be shown to correlate 100% with a given disease.

II. Clinical Presentation, Diseases, Diagnosis and Treatment

Although several different species of streptococci cause human disease (Table 1), this study shall focus mainly on the most prominent cause or the GAS. Interestingly, the

GAS is able to cause various diseases throughout the human host ranging from a mild self-limiting infection such as pharyngitis, impetigo and erysipelas to life threatening illnesses such as necrotizing fasciitis and streptococcal toxic shock-like syndrome. After the initial infection has subsided, a variety of immune mediated sequelae may arise including acute rheumatic fever and poststreptococcal glomerulonephritis. Specific disease manifestations of several of the more common diseases along with the clinical diagnosis of GAS infections and treatments are discussed in detail below. For a more thorough review Mandell, Douglas and Bennett's Principles and Practice of Infectious Diseases should be consulted (21).

A. Superficial diseases

Tonsillopharyngitis is a self-limiting infection, lasting between three to five days, which localizes to the tonsillopharynx. Infected patients often present with the following symptoms: a rapid onset of fever, pharyngeal erythema and exudates, cervical adenopathy, and an elevated white blood count. Presence of a cough or nasal discharge can rule out infection by the GAS over viral causes of pharyngitis. From mid-winter to early spring infection rates are at their highest point with the most susceptible host being children ranging from toddlers to adolescents (176). Asymptomatic carriage is also common in these individuals, with rates estimated to be as high as 15 to 20% for school-aged children (21). This population is targeted because children younger than two are protected from an acquisition of maternal immunity and a decreased ability of the GAS to

attach to their epithelial cells. By adolescents, a protective immunity has been built from contact throughout the previous years of life.

Pyoderma or impetigo is characterized by lesions that are usually located on the face or lower extremities. The lesions begin as papules that swiftly progress into vesicles and then to pustules. The pustules enlarge and then break down over the course of four to six days, leaving a characteristic thick crust and depigmented area on the surface of the skin. Infectious strains can also be transferred to and carried asymptomatically in the upper respiratory tract of infected individuals. Initially, the bacteria colonize the skin and only infect once this barrier is breached. For this reason, athletes and those whose occupations entail recurrent lacerations or skin abrasions have an increased risk of infection (1, 66). The most common population infected are impoverished children between two and five that live in tropical or subtropical regions because their poor hygiene promotes colonization of the skin.

Erysipelas is another superficial cutaneous GAS infection occurring most often on the face and/or lower extremities. It affects only layers of the epidermis and is distinguish by a well-defined raised red rash, which demarcates infected from uninfected areas. Erysipelas may cause localized pain, swelling and erythema in conjunction with more systemic symptoms like fever and chills. With facial erysipelas, a patient's eyes are often swollen shut.

B. Invasive infections

Not all GAS infections are superficial. Several are considered invasive, which are defined by the presence of bacteria at a normally sterile site. The subcutaneous invasive disease known as cellulitis is noted by symptoms similar to that of erysipelas, except the infected area is not raised or demarcated. In addition, cellulitis often results from a more substantial injury such as a burn or surgical incision than from a mild skin trauma.

Necrotizing fasciitis, also known as the "flesh-eating disease", is an infection of the deep subcutaneous tissues. Approximately 24 hours after a minor trauma, the initial infection will rapidly progress, completely destroying the superficial and deep fascia while the only signs on the skin's surface is a slight purple hue and possibly a fluid filled bullae. Massive inflammation and bacteremia may also be present at this time. Four to five days later, myositis (swelling of the muscle) and gangrene become evident and their effects are far reaching. This life-threatening infection often strikes otherwise apparently healthy young adults (215).

C. Toxin-mediated diseases

Scarlet fever is a complication of a GAS pharyngeal infection, resulting only when the infectious strain carries a phage with the ability to produces pyrogenic toxins (199). The toxins cause a diffuse erythematous rash, which starts on the upper chest and spreads to the rest of the body excluding the area around the mouth (circumoral pallor), palms, and soles of the feet (85). The tongue may be covered by a yellowish white coat, which will shed to make the tongue appear beefy red in color (strawberry tongue). The

rash lasts for about a week and is followed by desquamation. This disease was prevalent in the late 19th century but is not as common today (189).

Streptococcal toxic shock-like syndrome or STSS is thought to result from the production of pyrogenic exotoxins by the GAS acting as superantigens (84) and is characterized by three phases of disease progression. During the first phase, flu-like symptoms including fever, chills, nausea, vomiting and diarrhea may be present with patients often complaining of pain, which will increase in severity throughout the remaining phases. In the second phase, patients develop a high-grade fever, tachycardia (rapid heart beat) and tachypnea (rapid breathing). By the third phase, patients are hypotensive and there is a rapid onset of shock and multiorgan failure. Patients that have progressed to this stage will die within 24 to 48 hours (216). This infection is also seen in apparently healthy individuals.

D. Non-suppurative sequelae

Acute rheumatic fever (ARF) is a non-suppurative sequelae, which appears between one to five weeks after the clearing of an initial GAS pharyngeal infection. ARF manifests itself as an acute systemic febrile illness and is known to affect several areas including the heart, joints, subcutaneous tissues and the central nervous system. Although for the most part ARF is self-limiting, the damage that it causes can be chronic and, years later, lead to cardiac failure and possibly an early death. Since this disease is associated with prior pharyngeal infections that are most commonly found in childhood, ARF is also found primarily in children between the ages of five to fifteen. Recurrent

attacks later in life are only common if a patient has previously suffered from ARF. Current theories predict that ARF is an autoimmune disorder, which results from antistreptococcal antibodies that cross-react with host tissues (115).

Another sequelae known as poststreptococcal glomerulonephritis is the leading cause of acute nephritic syndrome (85). It occurs primarily in young children, with a bias towards males over females (49) and can take anywhere between seven to forty-two days after a streptococcal infection to develop. Latency periods at the lower end of this spectrum follow throat infections, while those at the upper end tend to follow infections of the skin. Edema, hypertension, hematuria and red or smoky urine are all clinical signs of infection. Like ARF, this disease also appears to be the result of an autoimmune reaction. Interestingly, although both rheumatic fever and poststreptococcal glomerulonephritis can follow a throat infection, they rarely coexist in patients (206).

E. Spread

GAS infections are spread through direct contact with an infected individual and via inhalation of contaminated droplets of saliva or nasal secretions. Crowded living conditions such as barracks or schools often facilitate spread. Documented epidemics cause by a contaminated food source can also be found although they are not especially common (74, 141). Contaminated clothing, bedding, or other fomites, however, have not been reported to play a significant role in transmission (21). Transmission rates between individuals are drastically reduced after only 24 hours of treatment with penicillin (see below) (176).

F. Diagnosis

Early diagnosis and treatment of GAS infections can be critical for decreasing both the associated morbidity and mortality, especially with the severe invasive diseases and GAS sequelae. A positive culture for β-hemolytic streptococci on sheep blood agar that shows sensitivity to bacitracin is considered the gold standard (19), while an elevated white blood count exceeding 12,000/mm^3 in conjunction with an increase in polymorphonuclear leukocytes can also be diagnostic of a GAS infection (21). Because the results from a throat culture can take to two to three days to obtain, a rapid strep test has been developed. This test is based on antibody-antigen interactions and can detect a GAS pharyngitis infection within minutes. For infections other than pharyngitis where a positive culture is not practical or obtainable (as in the case of sequelae), clinical manifestations alone are used to make the diagnosis. This is common for infections such as impetigo, cellulitis, scarlet fever and glomerulonephritis. In addition, clinically relevant algorithms have been developed, such as the one by Centor and co-workers or Jones, that use the presence or absence of a multitude of symptoms to diagnose GAS infections like pharyngitis or ARF, respectively (36, 211). Certain severe invasive streptococcal diseases such as necrotizing fasciitis require surgical investigation for diagnosis and determination of the extent of destruction before an appropriate treatment can be given.

G. Treatment

Currently, penicillin is the antibiotic recommended by the American Academy of Pediatrics Red Book Committee, the American Heart Association, and the Infectious Diseases Society of America for treatment of most GAS infections (21, 176). A ten-day course of oral Penicillin V is frequently prescribed, although a single intermuscular dose of penicillin G benzathine can be given instead (19). Treatment of skin infections often require a penicillinase-resistant penicillin because the area is commonly co-inhabitated by beta-lactamase producing *Staphylococcus* (176). Erythromycin is the suggested alternative for any individual with an allergy to penicillin (21, 176). Although the immune system is capable of clearing most infections, antimicrobials are still used because they not only ease the clinical manifestations and stop spread, but can help prevent the occurrence of rheumatic fever if given within the first nine days of the initial throat infection (34). Because of this, asymptomatic carriers usually do not undergo treatment since they are not at risk for developing sequelae (21).

When antibiotics are used however, they are not always effective. Studies over the past 30 years have reported a failure rate between 5 to 35 % for penicillin treatment (176). Several hypotheses concerning the reasons for treatment failure have evolved, including the co-colonization with beta-lactamase producing bacteria as discussed earlier (176), an inability of the antibiotic to affect the bacteria once it has reached stationary phase (also know as the “Eagle” effect), and an inaccessibility due to internalization of the streptococcus into the epithelial cells (165, 171). A tonsillectomy is sometimes

performed on those with reoccurring episodes of pharyngitis but is considered a radical measure and should be reserved for only the most severe cases.

Treatment of the more serious deep infections like necrotizing fasciitis and STSS should include surgical debridement of the infected area. With necrotizing fasciitis, amputation of a limb is frequently necessary to prevent spread to uninfected areas. Mortality rates are high for these types of invasive infections without early detection and treatment. Other modes of treatment besides antimicrobials are also under investigation. For example, the use of intravenous immunoglobulin from recovered patients to treat STSS may become a useful tool in the future (167).

H. Vaccines

Although both human and animal studies have shown that antibodies to the highly antigenic amino-terminus region of the M protein are protective, they are also type specific (70, 131). Because it is neither economical nor feasible to make a vaccine for each M type, research into a multivalent vaccine is currently underway (122). Unfortunately, antibodies to the M protein can be cross-reactive with human epitopes (52, 148) and thus a multitude of candidates other than the M protein are also being investigated as possible vaccine targets (reviewed in (157)).

III. Model of GAS Infections

The GAS is considered a pathogen because it not only colonizes the human host, but also causes harm to its host in the form of the diseases discussed above. The infection process begins when the GAS is able to bind and adhere to host tissues. This adherence keeps the bacteria from being repelled by the many mechanical barriers that are meant to protect the host. Adhesion is thought to occur as a two-step process. The first step brings the organism into close contact with host cells and allows for a second tighter binding reaction to occur through a variety of different molecules along the bacterial surface (89). The organism then establishes itself by acquiring nutrients and employs several mechanisms to avoid elimination by the immune system. Depending upon the initial site of colonization and the infecting strain, the GAS may disseminate to other locations resulting in both local and/or systemic damage. The ability of the GAS to exist as a pathogen has evolved from the assortment of virulence genes that the organism has at its disposal.

IV. Virulence Factors

Encoded within its genome are factors that allow the bacteria to adhere, circumvent the natural defenses of the host, internalize, spread to distant sites, and damage the host. Even though these factors are known to play a role in virulence, it is possible that they have additional cellular functions that have yet to be determined. Regardless, only the role(s) that each virulence factor plays in pathogenesis is presented

below. To avoid redundancy because so many products have multiple functions, the virulence factors have been divided according to their placement within the context of the cell. For an overview of their contribution to the infection process see

Figure 1.

A. Cell-associated virulence factors

Lipoteichoic acid: The cell wall of the GAS is composed of lipoteichoic acid (LTA), which aids in its ability to bind to epithelial cells (9). LTA is thought to act at the initial step in the binding process where the organism is brought into close proximity with host cells, allowing subsequent higher affinity adhesins to bind (89). This initial binding appears critical because treating mice with LTA or pretreating bacteria with anti-LTA blocks colonization of the animals (51).

Fibronectin-binding proteins: The GAS contains many fibronectin-binding proteins. Collectively, these proteins help the organism to bind to host tissues and include PrtFI/SfbI, PrtF2 (106), SfbII (127), SOF (44), Fbp54 (45), Pfbp (194), FbaA (220), FbaB (221), and M-protein (200); (which is discussed in greater detail below). The GAS also produces a protease known as SpeB, which is discussed with the other secreted proteins, that actively removes these proteins from the cell surface to abrogate the binding (38).

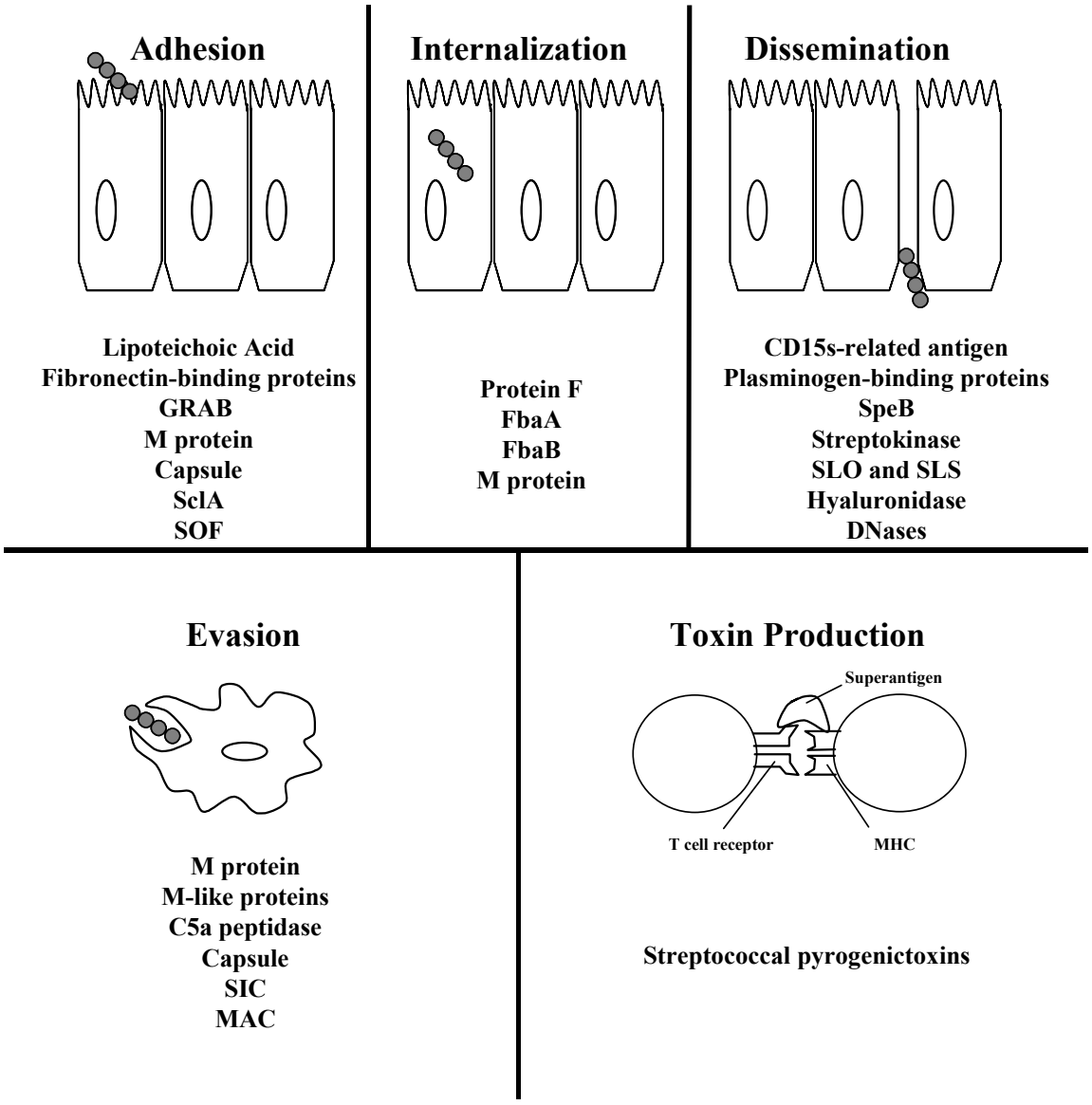


Figure 1: Interaction between the GAS and the host.

The GAS uses many different virulence factors throughout the course of infection. The key factors involved are listed beneath the corresponding stage in the infection process.

Protein F1 (*prtF*), also known as streptococcal fibronectin-binding protein SfbI (*sfb*), facilitates adherence to respiratory epithelial (88) and Langerhans cells (169). It has also been shown to bind to IgG through the Fc portion of the molecule (158). Because expression is enhanced under oxygen rich conditions, PrtF1 has been suggested to be important for binding to the surface of the skin (77). Besides being an adhesin, PrtF1 is also thought to be a critical component for internalization of the GAS into nonphagocytic cells (105, 160). Fbp54 is similar to PrtF in that it facilitates binding to epithelial cells (42), but unlike PrtF, Fbp54 is highly conserved and present in all strains tested (119, 227).

Strains that lack PrtF1 but still bind fibronectin contain PrtF2 or protein F2. PrtF2, like PrtF1, is regulated by changes in the oxygen concentration (106). Both Pfbp and FbaB have been found to be homologous to PrtF 2 (186). FbaB along with FbaA, another fibronectin-binding protein identified in a genome search for proteins containing the surface-anchoring motif LPXTG, are involved not only with adherence but also invasion (220, 221).

The protein SfbII is antigentically distinct from PrtF/SfbI, yet serves the same purpose (127). It has been shown to have an apoproteinase activity in addition to its fibronectin-binding function and thus is classified as a serum opacity factor (125). The serum opacity factor (*sof*) is similar to SfbII, in that it is bifunctional with a domain involved in binding fibronectin and another involved in opacification of serum (43, 44, 185). It exists as both a membrane-bound and free form (86, 238) and is usually

associated with strains that cause skin infections (234). The *sof* gene is transcribed in an operon with another fibronectin-binding protein encoded by *sfbX* (107).

Streptococcal collagen-like proteins: Even though the role of the streptococcal collagen-like proteins SclA/Scl1 and Scl2 in virulence is not entirely understood, it has been shown that mutants affect adherence to human fibroblasts and are attenuated for virulence in mice (144). Both proteins can be produced simultaneously, although the mode of regulation for each protein is different, with SclA being controlled at the level of transcription by Mga (188) and Scl2 being controlled at the level of translation (144).

CD15s-related antigen and plasminogen-binding proteins: The exact benefit of the CD15s-related antigen and plasminogen binding proteins is unclear, although it is known that both mimic host proteins involved in binding. For example, neutrophils and monocytes express CD15s to allow them to bind to vessel walls and pass from the vessel into the adjacent tissues (163). It has been suggested that the GAS uses its CD15s-related antigen in a similar fashion, allowing the bacteria to spread (93). Because the host expresses a similar antigen, this mimicry may also act to camouflage the bacteria from immune detection.

Three streptococcal proteins, the plasmin receptor protein or glyceraldhyde-3-phosphate dehydrogenase (239), α-enolase (172), and plasminogen-binding group A streptococcal M-like protein (PAM) (12), also mimic the functions of their eukaryotic receptor counterparts in binding to plasminogen or the proteolytic active form, plasmin. Although a definitive reason for this has not been demonstrated, it is thought that either

the binding of plasmin aids in dissemination by breaking down fibrin barriers or that binding of the hosts plasminogen helps decrease the amount available to the host (55, 155)

HtrA: HtrA, a homologue of the DegP protease of Gram-negative bacteria involved in degrading proteins damaged by thermal or oxidative stress, was found during an *in silico* analysis of the GAS. HtrA seems to serve the same function in the streptococcus, which may also play a part in virulence (110). Insertional inactivation mutants overexpress streptolysin S and do not process the SpeB zymogen, both of which are discussed below (145).

α_2-macroglobulin and GRAB: One regulator of proteolytic activity along the bacterial surface is α_2-macroglobulin. Although this molecule is a host protease inhibitor, it is thought to bind to the bacterial cell surface and provide protection for important virulence factors from cleavage by host or GAS proteases (187). It is bound to the bacterial surface via GRAB or the protein G-related alpha$_2$-M-binding protein. Since GRAB can bind host alpha$_2$-macroglobulin, current theory predicts that it may play a role as an adhesion.

M protein: The M protein is a multifaceted protein, which has a role in adhesion (46), immune evasion, and internalization (47). It has been shown to bind to plasma, matrix proteins (200), collagen (57), and kininogen (11). Protein production is upregulated in CO_2-rich conditions, implying that it may be more important for binding to deeper layers of the skin or the dermis (18). M protein has also been shown to inhibit

phagocytosis by human neutrophils, allowing the GAS to grow rapidly in whole blood (70). Exactly how M protein inhibits phagocytosis is unclear although several models have been proposed. The first two models are based upon limiting complement deposition via factor H, with one hypothesizing that the M protein binds directly to factor H (99) and the second hypothesizing that the M protein binds fibrinogen, which in turn binds factor H (100). A third possibility involves the binding of the hypervariable region (HVR) of the M-protein to the human complement control components known as C4b-binding protein (C4BP) and factor H-like protein (FHL-1) to inhibit phagocytosis (13, 109, 162). However, not all strains can bind C4BP, implying that another mechanism must also exist (222).

Mrp, Enn and Arp: The M-protein is part of a superfamily of proteins, which are share structural similarity with the M protein. Included in this family are the protein Mrp/Fcr and Enn. Like M protein, they also affect phagocytosis resistance (182) and are capable of binding immunoglobulins. Different family members show different affinities for a particular substance. For example, Mrp binds to both IgG and fibrinogen (213), while Arp or the IgA receptor protein, binds tightly to IgA, weakly to IgG and does not bind fibrinogen at all (139, 213). Although the role of binding immunoglobulins in virulence remains unclear, it is thought to be involved with the clinical manifestations of streptococcal sequelae (27, 28).

C5a peptidase: Another way that the GAS avoids being destroyed by the immune system is through C5a peptidase (ScpA). This endopeptidase specifically cleaves the

complement component C5a in order to inactivate the chemotactic signal (41). Inactivation of the signal prevents immune cells from infiltrating the area and destroying the bacteria (108, 236).

Capsule: The capsule of the GAS is composed solely of hyaluronic acid, which is structurally identical to that found in humans. Its production can vary from strain to strain (3), giving the highly encapsulated strains a mucoid phenotype. Like many other GAS virulence factors, encapsulation protects the bacteria in multiple ways. The first way is by allowing the bacteria to resist phagocytic killing (235). This is thought to work by providing a physical barrier, which prevents the phagocytes from interacting with the opsonic complement proteins bound to the surface of the bacteria and thus may take precedence over M protein in heavily encapsulated strains (53). Capsule may also play a role in adhesion, by binding to collagen (57), and invasion, by inducing cytoskeleton rearrangement of host cells by binding to the host hyaluronic acid receptor, CD44, on epithelial cells in the pharynx (50, 201).

Streptolysin S: Most strains of the GAS produce at least two known hemolysins, streptolysin S and streptolysin O (explained below); (6). Streptolysin S (Sls), is oxygen-stable and exists primarily in a surface-bound form (78). This hemolysin actively forms pores in erythrocytes, neutrophils, lymphocytes and a variety of other cells (79, 102) by insertion of a lysin complex into the cell membrane (33). Sls is encoded for on the streptolysin S-associated gene (*sag*) locus, an operon encoding for nine genes and a regulatory RNA known as *pel* within the promoter region of the first gene in the operon.

Although *pel* affects the transcription of other virulence genes (137), single-gene complementation analysis has demonstrated that Sls is involved in both resistance to phagocytic killing and host cell injury (54).

B. Secreted virulence factors

Streptolysin O: Streptolysin O (Slo), is an oxygen-labile hemolysin secreted by the streptococcus (6). Slo produces its toxic affects by binding cholesterol and forming aggregates in the membrane of cells including erythrocytes, macrophages, leukocytes and platelets (204). These aggregates have been shown to form pores in the cell through which effector molecules can be introduced into the host cell cytosol in a similar fashion to the type III secretion apparatus (146).

SpeB: The pyrogenic exotoxin type B or SpeB is one of the most widely studied virulence factors in the GAS, however, these studies often give conflicting results and thus its role is still under intense debate. It is known that this enzyme is produced as a zymogen (140), which in its mature form acts as a cysteine protease, cleaving matrix metalloproteases (26), fibronectin, vitronectin (118) and the IL-1β precursor (117). This cleavage contributes to the inflammation and can possibly lead to shock (117). In addition, it can solublize both M and M-like proteins (60). Since it shows a high degree of conservation and is found in all GAS strains (118, 240), it has been investigated as a possible vaccine target (116).

Streptokinase: Streptokinase (Ska) was originally called fibrinolysin and was found to be involved in the spread of the organism by specifically lysing human fibrin

clots (224). It is now known that this activity stems from its ability to convert plasminogen into the proteolytically active plasmin form (142), and that this activity is not inhibited by the mechanisms that normally inhibit plasmin (35). Deposits of Ska has been suggested to play a role in the development of acute poststreptococcal glomerulonephritis by activating the complement cascade locally in the kidneys (166).

Secreted inhibitor of complement: The secreted inhibitor of complement (Sic) serves many functions in the streptococcus. Its name originates from its ability to inhibit complement-mediated killing by binding to the C5-7 components of the membrane attack complex (68). It has also been shown to inactivate two different antibacterial peptides, human neutrophil alpha-defensin and LL37, produced by the host innate immune response (72). In addition, Sic helps avoid internalization by PMNs by binding to ezrin, a protein involved in bridging the actin cytoskeleton with the plasma membrane (94). Other components of the innate immune response such as lysozyme were also found to be inhibited by Sic (67)

Mac or IdeS: Mac, also known as immunoglobulin G-degrading enzyme of *S. pyogenes* (IdeS), protects the streptococcus in multiple ways. This proteinase is able to cleave the hinge region of IgG, which prevents phagocytosis (231). In addition, it serves to protect cells from killing by inhibiting the production of reactive oxygen species by PMNs (135).

Hyaluronidases: Several different hyaluronidases are produced by the GAS. Not only is a hyaluronate lyase encoded for on the chromosome but one is also encoded for

on a bacteriophage (discussed in the subsequent section). Hyaluronidases work by breaking down the hyaluronic acid in host connective tissues and allowing the organism or its toxins to spread (104). Exactly why the GAS produces an enzyme that can degrade its own protective capsule is not known although is has been speculated that regulation of capsule production is necessary only during certain stages of disease progression with the capsule being produced early and then being degraded in later stages after other protective factors have been produced.

C. Phage-encoded virulence factors

Phages are defined as movable genetic elements that can infect bacteria. There are approximately four to six phages in each of the three streptococcal genomes sequenced (18). Phages are thought to play a critical role in the passage of virulence factors via horizontal gene transfer amongst the streptococci (233).

Hyaluronidase: In addition to the hyaluronate lyase encoded on the chromosome of the streptococcus discussed above, one is also encoded on the phages that infect the organism. This enzyme is thought to degrade the capsule to allow entry of the phage into the cell (104). Whether or not it plays an additional role in the disease process is not known although antibodies to it are detectable post infection (87).

DNases: The GAS produces four different DNases (A-D) all of which degrade DNA. They appear to play a role in immune evasion through degradation of innate immune structures that allow neutrophils to trap bacteria and kill them in an extracellular fashion (218). DNase B, also known as streptodornase, has been shown to be identical to

SpeF. (62). This molecule contains two separate epitopes, one responsible for the DNase activity and the other involved in mitogentic activity. Why almost all of the DNases are phage-encoded (25) is not known, but antibodies to them can be found in patient serum (101, 114).

Streptococcal pyrogenic exotoxins: Superantigens are antigens that are potent stimulators of the T cells. They act by binding to the T-cell receptor and the major histocompatibility complex protein outside of the normal peptide-binding site, allowing for stimulation of a large number of cells. Stimulation in this fashion does not lead to a pathogen-specific immune response but instead to a massive release of cytokines, resulting in systemic toxicity and suppression of the adaptive immune response. The streptococcal pyrogenic exotoxins (Spe) are toxins that act as superantigens, with many of them including SpeA, SpeC (A and C are scarlet fever toxins), SpeF, SpeH, SpeI and SSA (29) being encoded for by prophages. A direct link between the toxin production and its role in the infection process has yet to be established, outside of their implication as the causative agent of the clinical symptoms such as hypotension and multiorgan failure seen with scarlet fever and streptococcal toxic shock syndrome (198, 199). It has been suggested that the effects may depend more on susceptibility of the host to the toxin than on the properties of the toxin itself (37).

V. Immune Response/Evasion

Opsonization, phagocytosis and subsequent killing by polymorphonuclear leucocytes (PMNs) is the major host defense against GAS infections. In order to cause disease, the bacteria must be able to evade the host immune response. Thus, the GAS has developed ways to interfere with the different steps of the classical and alternative complement pathways. Because each virulence factor was discussed in detail above, a summary of the factors and their effects is presented (Table 2), which describe the interplay between the bacteria and the immune response.

Table 2: Mechanisms of immune evasion

Ab Interaction	Opsonization	Phagocytosis	Recruitment	Reactive Oxygen Species
M-like proteins MAC SpeB	M protein Fibrinogen-binding proteins	Capsule Fibrinogen-binding proteins Collagen-binding proteins SIC MAC	C5A peptidase	HtrA protease

Products produced by the GAS are listed below the corresponding step of the immune response that they inhibit.

VI. Modes of Transcriptional Regulation

Bacterial survival depends not only on the repertoire of virulence genes that the pathogen possesses, but also on its ability to regulate their expression, creating an appropriate response to the specific environmental conditions. Differential gene expression in bacteria is controlled at the level of transcription by a variety of mechanisms, with the two most common being alternative sigma factors and

transcriptional regulators. The only known alternative sigma factor in the group A streptococcus is ComX (170), which shares 40% identity to the known regulator in *Streptococcus pneumoniae* involved in competence (134). Since ComX has not been reported to be directly involved with virulence regulation, it is considered beyond the scope of this report and will not be discussed here.

The other method that bacteria have evolved to regulate large networks of genes is through transcriptional regulators. Regulators respond to a variety of different cues ranging from environmental conditions occurring outside the cell to the metabolic components that exist within the cell. Overall, the GAS has been estimated to encode for more than 100 transcriptional regulators, with more than thirty-six transcription factors shared amongst the sequenced stains (69, 209).

A. Metabolism regulators

Of particular interest to this study are transcription factors that use metabolic intermediates as a signal for regulation in the cell. Many regulators that respond to this type of signal contain a phosphoenolpyruvate phosphotransferase regulatory domain or PRD. These regulators exert their affects by acting as either antiterminators or transcriptional activators, which bind to RNA or DNA, respectively. The subset that act as transcriptional activators often contain a DNA-binding motif and two PRDs. The PRDs undergo phosphorylation on conserved His residues (each domain may contain one or two conserved His residues), which causes the protein to change its oligomeric state. Whether this state is active or inactive depends upon the regulator because some, such as

LicR, are activated by phosphorylation (225) and others, such as LevR, are inactivated (147).

Regardless of the activated state, phosphorylation of a PRD is part of a cascade in which the phosphate is initially transferred from phosphoenol pyruvate (PEP) to the soluble enzyme I (EI) component, and then to the heat stable protein (HPr) (Figure 2). Depending upon which residue of HPr is phosphorylated, either the Ser at position 46 or a His at position 15, two different outcomes may arise. HPr is phosphorylated on Ser-46 when glycolytic intermediates are at high levels, which allows it to form a complex with a regulator known as the catabolite control protein (CcpA) (56). This complex binds to catabolite-responsive elements (*cre*) within promoters and either inhibits or activates transcription of genes involved in sugar metabolism (73). However, when HPr is phosphorylated on His-15, it will either transfer the phosphate to the membrane-bound sugar specific enzyme II (EII) components, to import the sugar if external levels of substrate are high, or it will phosphorylate a regulatory protein that contains a PRD, if substrate levels are low. In addition to phosphorylation by HPr, EII can also donate its phosphate directly to a PRD if a sugar is not available for uptake/ phosphorylation by the protein (147).

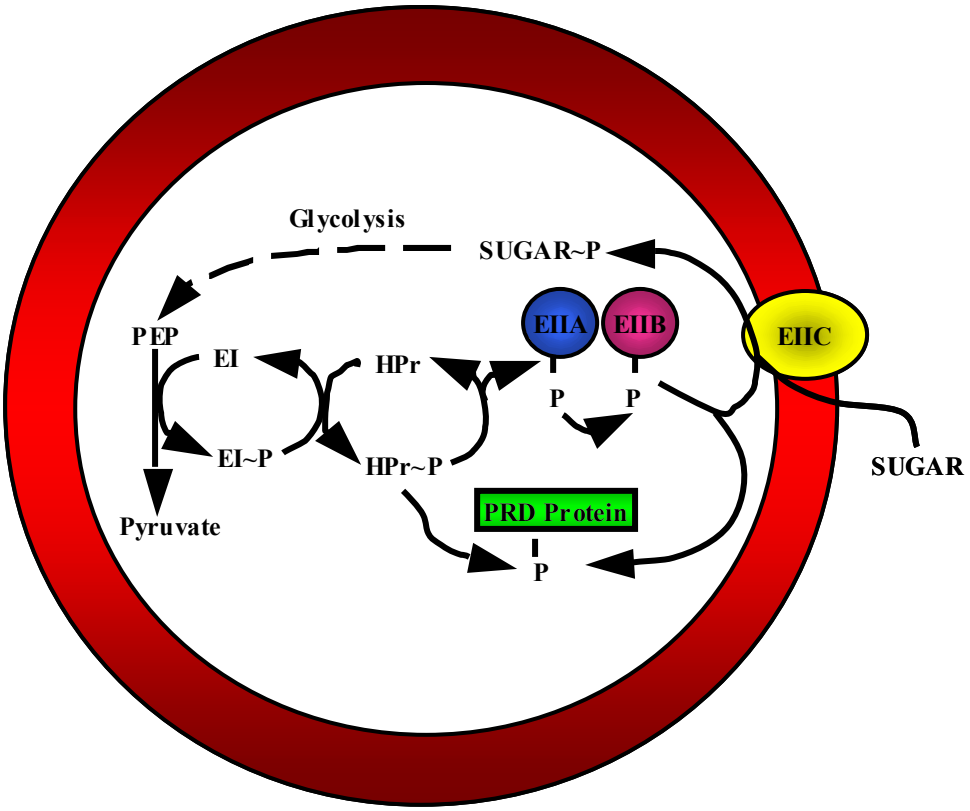


Figure 2: PTS-regulatory system in the GAS.

The PTS system is responsible not only for the uptake of a number of different sugars but also for the regulation of genes encoding enzymes used in the metabolic pathways. This system is regulated by a phosphorylation cascade that begins with the transfer of a phosphate from phosphoenolpyruvate (PEP) to enzyme I (EI), which in turn phosphorylates HPr. The phosphate on HPr can then be transferred to either the enzyme II components (EIIA and EIIB) ending with the membrane bound sugar specific enzyme (EIIC), which allows for phosphorylation of the incoming sugar, or to regulatory proteins that contain phosphoenol-pyruvate regulatory domains (PRDs). This phosphorylation occurs on a His residue in the PRD and acts to activates or inactivates the PRD containing protein. HPr is not the only source of phosphorylation for PRDs. These domains may also be phosphorylated by IIB if the sugar concentration on the exterior of the cell is low and thus there is not an incoming sugar to phosphorylate.

B. Two-component regulatory systems

Bacterial two-component signal transduction systems are another type of transcriptional regulator that play a critical role in converting an environmental stimulus into a cellular response. The system is made up of a sensor protein, which often spans the membrane of the cell acting as a receptor for the environmental stimulus, and a response regulator, which transmits the signal from the receptor to either activate or inhibit transcription from specific gene promoters. Upon receipt of an activating signal, the sensor kinase undergoes an autophosphorylation reaction on a histidine within the protein. The phosphate is then transferred from the sensor kinase to an aspartate residue that lies within an acidic pocket of a response regulator. This causes the response regulator to undergo a conformational change (involving a critical lysine residue), which activates/inactivates the protein, resulting in an alteration of transcriptional activity. Mutational analysis using both the key aspartate and lysine residues has been undertaken in several different response regulators including the model system CheY.

Mutational analysis of model response regulators: CheY mutants have provided several insights into the way a response regulator functions and its requirements for activation *in vivo*. One key feature is that when the site of phosphorylation (D57) was mutated to an Ala (D57A) a loss of activity was seen, demonstrating the necessity of this Asp (23). However, when this position was changed to an Asn (D57N) two different phenotypes were noted. This mutant could be phosphorylated at an alternative site (the serine at position 56) and was inactive in the presence of the dephosphorylating molecule

CheZ (24, 143, 197) but active in its absence (8). In contrast, when the same residue was changed to a Glu (D57E), the mutant could still be phosphorylated at position 57, yet it no longer retained its *in vivo* activity. It was concluded that activation of the response regulator was not merely determined by the phosphorylation state of the protein, but also by the ability of the phosphorylation to "reposition key residues".

A Lys residue also appears to be critical for the proper repositioning of the protein, although it is not directly involved in the phosphorylation event. Mutational analysis of CheY at this Lys revealed that a K109R mutant was not active even though the protein was phosphorylated (143). It was suggested that the inactivation resulted from an inability of the mutated side chain to interact with the phosphorylated D57 residue and thus it was not able to change into an active conformation.

After the initial studies with CheY, corresponding mutations were produced in other response regulators such as NtrC (nitrogen regulator I in both *E. coli* and *S. typhimurium*) and RcsB (a regulator of capsule synthesis in *E. coli*). Analysis of these regulators demonstrated that a change from an Asp to an Asn at the active site resulted in an inactive phenotype (83, 121, 161). Mutants exchanging the Asp for a Glu in these systems were shown to be constitutively active, which sharply contrasted to that seen with CheY. Even proteins mutated at the conserved Lys residue were active, once again a phenotype opposite to that for CheY. Overall, it appears as though mutations that act universally on all response regulators may not exist and thus each regulator must be considered individually within its system (208).

Two-component regulatory systems in the GAS: An understanding of two-component systems is imperative when studying pathogenesis since virulence factors are often controlled by such systems. This type of system is advantageous for a bacterium because it allows a rapid adaptation to varying conditions by changing the repertoire of genes produced. To date, thirteen two-component systems have been identified in the GAS (69), of those characterized, three have been found to be involved in virulence regulation.

CovR/CovS: The first two-component regulatory system involved in virulence was initially called CsrR/CsrS for capsule synthesis regulator because it was originally identified during a transposon screen as a mutant with an increased production of capsule (136). It has since been renamed CovR/CovS for control of virulence genes and has been found to affect approximately 15% of the GAS genome including genes such as *sagA, speB* (91), *ska*, and *mac/ides* in addition to its own expression (65, 135). CovR binding to target DNA is phosphorylation-dependent (14) and appears to influence the transcription of genes expressed during log or stationary phase (65). So far the only environmental signal identified to activate CovS is Mg^{2+}, which is thought to initiate the phosphorylation of CovR by CovS (82).

FasBCAX: The second two-component system, called FASBCAX, regulates fibronectin binding, hemolysis, and transcription of streptokinase (124). This system is unique because it contains two potential histidine kinases (FasB and FasC), one response regulator (FasA) and one effector RNA molecule (FasX). A homologue has been found

in the GCS, although the homologue does not contain the second histidine kinase (212). Like CovR, it influences the transcription of late stationary phase genes, including the global regulator known as *ropB/rgg*.

Ihk/Irr: The Ihk/Irr two-component regulatory system or *isp*-adajacent histidine kinase and *isp*-adajacent response regulator was named after its placement in the genome upstream of *isp*. Its proximity to the virulence gene regulator known as Mga, discussed below, implicates that it too may play a role in virulence. Not much is known about how this system works except that it plays an essential role in survival following phagocytosis by PMNs (232).

C. Stand-alone regulators

Not all transcriptional regulators that respond to environmental stimuli are part of an established two-component system. For example, the GAS contains several transcriptional regulators, which respond to environmental stimuli but contain no known sensory components (126). The term "stand-alone response regulator" has been coined to define these regulators. The three that are known to be involved in virulence regulation are Mga, RALPs, and RopB/Rgg. Since Mga is the main focus of this work, it will be discussed in great detail in a subsequent section.

RALPs: Regulators within the RALP family are all functionally similar to *rofA* and are appropriately named the RofA-like proteins. This family contains four members: RofA, Nra, and other forms which are similar to these (81). Most genes under the control of these regulators, such as *prtF* and *prtF2,* are positively regulated and encode factors

involved in adhesion (10). However, RALPs have been shown to repress expression of some genes such as *mga* and *speA*. Even though both RofA and Nra regulate expression of genes during stationary phase, their activities appear to be regulated by different mechanisms in the cell.

RopB/Rgg: The regulator of protease or RopB is homologous to an activator of the glucosyl transferase gene known as Rgg. *ropB* is located immediately upstream of *speB*, one of the genes that RopB positively regulates (39) primarily during stationary phase (40). RopB/Rgg can be considered a regulator of regulators because a mutant lacking this factor affects expression of *mga, covR/covS, fasBCAX*, and *ihk/irr*. Therefore, RopB/Rgg appears to act both directly and indirectly on the genes that it regulates (164).

Regulation of virulence gene expression by sensing changing host environments is a major reason for the success of the GAS as a pathogen. Because of this, deciphering these mechanisms will be an important step towards controlling GAS infections. The present study focuses on the stand-alone regulator Mga and how it is able to function within the context of the cell.

VII. The Multigene Regulator of the Group A Streptococcus

A. Discovery of Mga

The multigene regulator of the group A streptococcus (Mga) was identified independently by two different research laboratories. In 1987, the Cleary lab began investigations of a deletion strain originally described by Spanier *et al.* to be reduced in M protein production even though the deletion was located 400 bp upstream from the *emm* coding sequence (192, 210). They called this upstream region the *virR* (virulence regulator) locus (207). In the same year, Caparon and Scott identified a gene through a transposon screen that was deficient in *emm* transcription (32). They named this gene *mry* for M protein RNA yield.

It was not until 1995 that the protein came to be known as Mga (202). Mga is a *trans*-acting factor (174) that is composed of approximately 530 amino acids (a slight variation is seen amongst the different GAS serotypes), has a molecular weight of 62 kDa, and a predicted pI of 6.3. Since its initial discovery, much has been learned about its function in the cell; the highlights of which are discussed in the following sections.

B. Regulation by Mga

Mga has been classified as a "stand-alone" response regulator. Like other regulators in this class, Mga has been shown to positively activate the expression of genes, including itself (154, 168), in response to signals such as increased CO_2 levels, body temperature, and exponential phase growth (30, 150, 152, 181). Although the

mechanisms by which Mga is able to sense this stimuli remain a mystery, many of the genes that it regulates have been identified.

The *mga* locus on the GAS chromosome always includes at least one *emm*-family gene (*emm*, *arp, emmL*) and the downstream gene *scpA* (178). Certain serotypes encode additional genes such as the M-like proteins (*mrp*, *enn*), which reside between *mga* and *scpA* (49, 95, 237), or the fibronectin-binding protein (*fba/orfX*) (184, 220), which is located immediately 3' to *scpA*. These genes make up the core *mga* regulon (156, 179) and are regulated in *trans* by Mga (174) (Table 3). Unlike the genes located downstream, the gene immediately upstream of *mga,* known as *isp* or the immunogenic secreted protein, does not appear to be regulated by Mga (153). Not all genes regulated by Mga reside within the *mga* regulon. Several genes such as *speB* and *sof* have been shown to protein, does not appear to be regulated by Mga (153). Not all genes regulated by Mga reside within the *mga* regulon. Several genes such as *speB* and *sof* have been shown to exist outside the core regulon. Regardless of gene placement, all Mga-regulated genes appear to be involved in adhesion or immune evasion. In addition, they all are positively regulated, with the exception of streptolysin S, whose expression appears to be negatively regulated by Mga (120).

Table 3: Mga-regulated genes

Gene Name	In Core *mga* Regulon	Function	Reference
Cell associated Factors			
scnA	No	Bacteriocin (Streptococcin A)	(184)
speB	No	Cysteine protease	(184)
emm	Yes	Adhesion, anti-phagocytic surface molecule	(156, 179, 181)
sph	Yes	IgG Fc-binding protein (protein H)	(120)
arp	Yes	IgA binding protein	(139)
mrp/fcrA	Yes	IgG binding protein	(156, 179)
enn	Yes	IgG binding protein	(179)
scpA	Yes	C5a peptidase	(156, 179)
fbaA/ orfX	Yes	Fibronectin-binding protein	(184, 220)
sclA	No	Collagen-like adhesion	(4, 181, 188)
sof	No	Fibronectin binding lipoprotease	(156, 185)
sfbX	No	Fibronectin-binding protein	(107)
dpp operon	No	Dipeptide permease	(180)
opp operon	No	Oligopeptide permease	(184)
Secreted Factors			
sic	Yes	Inhibitor of complement	(2, 120)
sagA	No	Hemolysin (streptolysin S)	(120)
Autoregulation			
mga	Yes	Regulator of virulence genes	(154)

Using a maltose-binding fusion protein, it has been shown that Mga binds directly to the promoter regions of the genes that it regulates (149, 154). Further analysis of the promoter regions using DNase-protection assays revealed that Mga binds to a single 45 bp region overlapping the –35 region in the promoter of both *emm* and *scpA* (149). The consensus sequence derived from the study is unusual in that it contains a string of 13 bases that are highly conserved followed by a variable region of 11 bp and terminates with another string of 16 that are highly conserved (149). Since that initial study, other Mga-regulated promoters have been investigated such as P*mga* and P*sclA* (4, 154). From these studies differences in both the number of binding sites and the area of protection has been noted; thus, it was determined that the binding of Mga is variable at the different promoters.

C. Regulators of *mga*

One unique promoter that Mga regulates is that of *mga* itself. The *mga* gene is transcribed from two different promoters, one distal to the start of transcription (P1) and another located proximal to the transcriptional start site (P2) (75, 168). Studies have shown that the *mga* promoter is stimulated by signals such as growth phase (152), increased levels of CO_2 and decreased levels of O_2 (168). Although it has not been determined how these environmental signals are sensed, it is possible that Mga itself is directly involved. DNase protection assays of the *mga* promoter revealed that Mga binds to two separate 59 base-pair regions located between P1 and P2 in the *mga* promoter (154). However, Mga may not be the only factor responsible for the promoter

stimulation, since both binding sites reside between the two promoters, yet 473 base pairs of DNA upstream of the coding sequence is necessary for autoregulation (75, 168). These 473 base pairs include sequence upstream of even the P1 promoter.

In contrast to the positive regulation by Mga, *mga* expression has been shown to be negatively regulated by the RALPs (10, 183). In addition, a mutation in the gene for the lipoprotein of *S. pyogenes* (Lsp) was found to have a decreased level of transcription of *mga*, *nra* and *csrRS* (61). Whether the effect on *mga* is direct or indirect (due to decreased transcription of another gene that regulates *mga*) remains to be seen.

Another Mga-regulating factor besides those involved in global regulation was identified during a transposon screen (191). This factor is known as AmrA (activation of the Mga regulon) and is a putative integral membrane protein involved in cell wall polysaccharide synthesis. How this protein regulates *mga* expression is unknown, but several hypotheses exist including either a direct interaction or the possibility that Mga senses a signal produced by the transporter, like cell wall turnover, which leads to activation of the *mga* regulon.

D. Two alleles of *mga*

While all strains of the GAS contain *mga* (16), the architecture of the *mga* regulon differs from strain to strain in both gene composition and arrangement (95, 237). Furthermore, two divergent *mga* alleles have been described within the GAS (*mga-1* and *mga-2*) based upon the ability to hybridize to an oligonucleotide probe (96). Each allele has been associated with different gene patterns at the *mga* locus and tissue tropism of the

serotype (95). The Mga proteins produced from the two alleles are most divergent within the C-terminal end of the protein (7), showing a maximal amino acid divergence of 20.7% (16). Despite the divergence, a *mga-2* allele has been shown to functionally complement a *mga-1* deletion strain (7).

E. Orthologues

S. pyogenes is not the only species of streptococci that appears to contain a gene analogous to *mga*. Geyer and Schmidt identified an orthologue in *S. dysgalactiae* subsp. *equisimilis* named *mgc* (76) for the multigene regulator of the group C streptococcus (GCS). Mgc is 51% identical and 64% similar to the Mga (*mga-1*) from the serotype M6 GAS strain D471 (76). Another orthologue of the GCS designated as DmgB, exhibiting 45% identity and 61% similarity to Mga, has been identified in *S. dysgalactiae* subsp. *dysgalactiae* (230). The *mgc* and *dmgB* loci closely resemble that of the GAS *mga* in that a gene encoding an M protein homologue (*emm* and *demB,* respectively) lies directly downstream of the aforementioned GCS genes. Recently, a virulence regulator showing 51% similarity and 25% identity to Mga from the GAS was found in *S. pneumoniae* called MgrA for the Mga-like repressor A (90, 92). In contrast to Mga and DmgB/Mgc, MgrA appears to repress transcription of genes in the unlinked *rlrA* pathogenicity islet in the *S. pneumoniae* TIGR4 genome (92). Chapter five of this study describes a family of virulence gene regulators showing similarity to Mga that includes additional predicted regulators found in the genome sequences of other streptococcal species, including *S. equi, S. gordonii,* and *S. mitis*. Currently, only Mga from the GAS has been

characterized, therefore, it serves as the model system for this family of Mga-like virulence regulators.

F. Domains of Mga

Despite its established role in the pathogenesis of GAS infections, we currently know very little about how Mga functions. Previous studies using the serotype M6 Mga (*mga-1*) identified a major (HTH-4) and a minor (HTH-3) helix-turn-helix DNA-binding domain within the amino-terminus of Mga (151). HTH-4 was determined to be absolutely essential for binding to all targets while HTH-3 serves an accessory role primarily involved in autoregulation from P*mga* (151). Two putative response regulator receiver domains have also been suggested to exist in Mga based on homology to other two-component systems (174); however, no molecular evidence for these domains has been demonstrated. Aside from its DNA-binding abilities, little is known about how Mga is able to sense environmental stimuli and activate virulence gene transcription. Hence, the purpose of the following study is to identify novel domains within Mga that are necessary for its function in the GAS.

CHAPTER THREE:

MATERIALS AND METHODS

I. Bacterial Strains and Plasmids

A. *E. coli* strains

E. coli strains were grown in Luria-Bertani broth (LB) and growth was measured by absorbance at 600 nanometers on an Ultrospec 2000 spectrophotometer. Broth cultures were grown at 37°C with agitation under normal atmospheric conditions. All stock cultures were stored at -80°C in a 20% (v/v) glycerol solution. When required, the following concentrations of antibiotics were used: ampicillin at 100 μg/ml, erythromycin at 500 μg/ml, kanamycin at 50 μg/ml, and spectinomycin at 100 μg/ml. *E. coli* DH5α (*hsdR17 recA1 gyrA endA1 relA1*) (New England Biolabs) was used as a host for plasmid constructions, while *E. coli* SA2817 (149), BL21[DE3] containing the T7 RNA polymerase (217), and JM109 were used for protein purification (Pharmacia).

B. GAS strains

GAS strains were grown in Todd-Hewitt medium supplemented with 0.2% yeast extract (THY). Growth of the GAS was measured by absorbance on a Klett-Summerson photoelectric colorimeter using the A filter. When required, the following concentrations of antibiotics were used: erythromycin at 1 μg/ml, kanamycin at 300 μg/ml,

spectinomycin at 100 μg/ml, and streptomycin at 1000 μg/ml. All strains of the GAS that were used in this study are listed in Table 4 along with a brief description of their corresponding phenotype. Cultures were grown statically at 37°C under atmospheric conditions. All stock cultures were stored at -80°C in a 20% (v/v) glycerol solution.

Table 4: GAS strains

Strain	Relevant Characteristics	Ref. or Source
AL168-*mga*	Heterologous Δ*mga* M22 strain	(223)
AP4	Serotype M4 GAS clinical strain	(213)
B514-Sm	Streptomycin-resistant derivative of the M50 GAS strain B514	(138)
JRS4	Streptomycin-resistant derivative of the M6 serotype M6 strain	(203)
JRS519	Δ*mga* kanamycin-resistant derivative of the JRS4	(152)
KSM148	RTG229 derivative containing a single copy P*emm-gusA* reporter at an ectopic location within the chromosome	(191)
KSM148.150	KSM148 derivative containing an insertional inactivation of *mga* and the single copy P*emm-gusA* reporter in the chromosome	J. Lawson (Unpublished)
KSM148.174	Δ*mga* M6 KSM148 derivative containing a single copy P*emm-gusA* reporter at an ectopic location within the chromosome	(229)
KSM149	Δ*mga* M6 VIT231 derivative containing a single copy P*mrp-gusA* reporter at an ectopic location within the chromosome	(229)
KSM231.310	Δ*mga* M6 VIT231 derivative containing a single copy P*mga6-gusA* reporter at an ectopic location within the chromosome	(228)
RTG229	JRS4 derivative containing an erythromycin-resistant transposon inserted at an ectopic location within the chromosome	(75)
VIT230	RTG229 derivative containing a spectinomycin-resistant transposon inserted at an ectopic location within the chromosome	(229)
VIT231	Δ*mga* derivative of the VIT230 containing a spectinomycin-resistant transposon at an ectopic chromosomal location	(229)

C. Strain and plasmid construction

Construction of the Δmga allele in the M6 Pemm-gusA reporter strain KSM148.174. A deletion encompassing the 2.1 kb *mga* coding sequence and its 493 bp promoter region (P*mga*) was generated by inverse PCR from the plasmid pJRS515 (149), which contains the wild-type serotype M6 *mga* locus, using the diverging primers Mgadel-L2 and Mgadel-R2 (Table 5). The resulting 5.4 kilobase (kb) product was re-ligated to form the *mga-16* deletion allele marked with a new StuI restriction site in the plasmid pJRS547. A 2.5 kb BamHI/HindIII fragment containing the *mga-16* deletion allele from pJRS547 was subsequently cloned into the counter-selectable gene-replacement vector pJRS9160 (151) to generate pKSM174 (Table 6). An *mga*-deleted derivative of the serotype M6 VIT strain KSM148 (191) was constructed through an allelic replacement of the wild-type *mga* at the native locus with the unmarked *mga-16* deletion allele using pKSM174 as previously described (151) to produce KSM148.174 (Table 4).

Table 5: PCR primers and their relevant targets

Non-complementary sequences are in lower case, mutagenic nucleotides are in bold and introduced restriction sites are underlined.

Target/Primer	Sequence	Reference
23S rRNA		
rRNA-23SL	GGAAGGTAAGCCAAAGAGAG	(191)
rRNA-23SR	TCCTAGTTGTCTGTGCAACC	(191)
mga6		
Mga6-150	gcgtcaaagcttctaATCTCCTGATACTTGTACGG	(228)
Mgadel-L2	CGAGGCCTTAGCTTTTTGATGGCATCATGG	(229)
Mgadel-R2	CCCTTGGACTTTCATCGC	(229)
MgaHis-BglII	gccgagatctGCGAGAAAGG	(229)
Mga-Pet1_Nde	ggggcatATGTATGTAAGTAAGTTGTTT	(4)
Mga-Pet2_Xho	aactcgagAGTTGTGGAGGGGG	(4)
OYR-29	AAACCAACGCCTATTTGACGCATAC	(229)
RMut1-R2	cccgcctcgagAAAGAAGGGTATACAAGG	(228)
RMut2-L	TCGACCTGCAGgcatgcaaa	(228)
Mutagenic *mga6* oligos		
DR1A-a	GGGATACTCTGCCGTCTACG**c**CAACAAAAAAACCAGTCACCGT	This study
DR1A-b	ACGGTGACTGGTTTTTTTGTTG**g**CGTAGACGGCAGAGTATCCC	This study
DR1E-a	CAAGGGATACTCTGCCGTCTACGA**g**AACAAAAAAACCAGTCACCG	This study
DR1E-b	CGGTGACTGGTTTTTTTGTT**c**TCGTAGACGGCAGAGTATCCCTTG	This study
DR1N-a	CAAGGGATACTCTGCCGTCTAC**a**ACAACAAAAAAACCAGTCACCG	This study
DR1N-b	CGGTGACTGGTTTTTTTGTTGT**t**GTAGACGGCAGAGTATCCCTTG	This study
KR1R-a	GCGTTTGATAGCATCA**g**GCAAGACTCGCCAACGGGC	This study
KR1R-b	GCCCGTTGGCGAGTCTTGC**c**TGATGCTATCAAACGC	This study
DR2A-a	ACAGTATGATGTGATCGTGACAG**c**TGTTATGGTAGGAAAAAGCGA	This study
DR2A-b	TCGCTTTTTCCTACCATAACA**g**CTGTCACGATCACATCATACTGT	This study
DR2E-a	AGTATGATGTGATCGTGACAGA**g**GTTATGGTAGGAAAAAGCGATG	This study
DR2E-b	CATCGCTTTTTCCTACCATAAC**c**TCTGTCACGATCACATCATACT	This study
DR2N-a	GTATGATGTGATCGTGACA**a**ATGTTATGGTAGGAAAAAGCGATGA	This study
DR2N-b	TCATCGCTTTTTCCTACCATAACAT**t**TGTCACGATCACATCATAC	This study
KR2R-a	CCCCTTGGACTTTCATCGCA**g**AGAGCTTACCTTACCCACTCCC	This study

Table 5: continued

Target/Primer	Sequence	Reference
KR2R-b	GGGAGTGGGTAAGGTAAGCTCT**c**TGCGATGAAAGTCCAAG GGG	This study
M6 M461R-a	GATGTGATCGTGACAGATGTTAgGGTAGGAAAAAGCGATG AG	(229)
M6 M461R-b	CTCATCGCTTTTTCCTACCcTAACATCTGTCACGATCACATC	(229)
M6 mga Q10R-a	GTTGTTTACAAGTC**g**ACAGTGGAGAGAACTAAAATTAATC TCATACGTAACGG	(228)
M6 mga Q10R-b	CCGTTACGTATGAGATTAATTTTAGTTCTCTCCACTGT**c**GA CTTGTAAACAAC	(228)
M6 mga Q11A-a	GTTGTTTACAAGTCAA**gc**GTGGAGAGAACTAAAATTAATCT CATACGTAACGG	(228)
M6 mga Q11A-b	CCGTTACGTATGAGATTAATTTTAGTTCTCTCCAC**gc**TTGAC TTGTAAACAAC	(228)
M6 mga Q11R-a	GTTGTTTACAAGTCAAC**g**GTGGAGAGAACTAAAATTAATC TCATACGTAACGG	(228)
M6 mga Q11R-b	CCGTTACGTATGAGATTAATTTTAGTTCTCTCCAC**c**GTTGA CTTGTAAACAAC	(228)
M6 mga W12A-a	GTTGTTTACAAGTCAACAG**gc**GAGAGAACTAAAATTAATCT CATACGTAACGG	(228)
M6 mga W12A-b	CCGTTACGTATGAGATTAATTTTAGTTCTCTC**gc**CTGTTGAC TTGTAAACAAC	(228)
M6 mga W12R-a	GTTGTTTACAAGTCAACAG**c**GGAGAGAACTAAAATTAATC TCATACGTAACGG	(228)
M6 mga W12R-b	CCGTTACGTATGAGATTAATTTTAGTTCTCTCC**g**CTGTTGAC TTGTAAACAAC	(228)
M6 mga R13A-a	GTTTACAAGTCAACAGTGG**gc**AGAACTAAAATTAATCTCAT ACGTAACGG	(228)
M6 mga R13A-b	CCGTTACGTATGAGATTAATTTTAGTTCT**gc**CCACTGTTGAC TTGTAAAC	(228)
M6 mga H204A-a	CGATGTGAGGGTTAATTTTACGTTATTTCAG**gc**CCTTAAAA TACTAAGCTCAG	This study
M6 mga H204A-b	CTGAGCTTAGTATTTTAAGG**gc**CTGAAATAACGTAAAATTA ACCCTCACATCG	This study
M6 mga H204D-a	CGATGTGAGGGTTAATTTTACGTTATTTCAG**g**ACCTTAAAA TACTAAGCTCAG	This study
M6 mga H204D-b	CTGAGCTTAGTATTTTAAGGT**c**CTGAAATAACGTAAAATTA ACCCTCACATCG	This study
M6 mga H324A-a	CCAACAAATATGAAGTAGCTGTCATCCTT**gc**TAACACTACC GTTTTGAAAG	This study
M6 mga H324A-b	CTTTCAAAACGGTAGTGTTA**gc**AAGGATGACAGCTACTTCA TATTTGTTGG	This study
M6 mga H324D-a	CCAACAAATATGAAGTAGCTGTCATCCTT**g**ATAACACTACC GTTTTGAAAG	This study
M6 mga H324D-b	CTTTCAAAACGGTAGTGTTAT**c**AAGGATGACAGCTACTTCA TATTTGTTGG	This study

Table 5: continued

Target/Primer	Sequence	Reference
M6 mga E14A-b	CCGTTACGTATGAGATTAATTTTAGT**g**CTCTCCACTGTTGACTTGTAAACAAC	(228)
M6 mga L15A-a	GTTTACAAGTCAACAGTGGAGAGAA**gc**AAAATTAATCTCATACGTAACGG	(228)
M6 mga L15A-b	CCGTTACGTATGAGATTAATTTT**gc**TTCTCTCCACTGTTGACTTGTAAAC	(228)
mga4, mga50		
DivMga-5'blunt	CATGTAAGTAAATTGTTTACTAGCCAACAATGGAG	(229)
DivMga-3'HIII	gcgtcaaagcttCTGAAATCCTATGATGATGTTGC	(229)
DivMga-Pet1	gggcatatgCATGTAAGTAAATTG	(229)
DivMga-Pet2	gggctcgagTGATGATGTTGCTTGTTT	(229)
DivMga-R5	TGGATCCATCTATTAGATGAG	(229)
DivMga-R9	ccccaagcttGATAAGGACATGAAGTTAAT	(229)
DivMga-Sph	ccgcatgcctatgatgatgttgcttg	(229)
MgaHis-Sph	cgcatgcCTATGATGATGTTGCTTG	(229)
Mutagenic *mga4* oligos		
M4 mga Q11R-a	GCATGTAAGTAAATTGTTTACTAGCCAAC**g**ATGGAGAGAATTGAAACTG	(228)
M4 mga Q11R-b	CAGTTTCAATTCTCTCCAT**c**GTTGGCTAGTAAACAATTTACTTACATGC	(228)
M4 mga R13A-a	GTTTACTAGCCAACAATGG**gc**AGAATTGAAACTGATTTCATATTTAACAG	(228)
M4 mga R13A-b	CTGTTAAATATGAAATCAGTTTCAATTCT**gc**CCATTGTTGGCTAGTAAAC	(228)
M4 mga HTH4-a	GCTGAAGAGCTGTTTGTCAGC**gc**A**g**CTACCCTCAAACGCC	(228)
M4 mga HTH4-b	GGCGTTTGAGGGTAG**c**T**gc**GCTGACAAACAGCTCTTCAGC	(228)
Mutagenic *mga50* oligos		
Mga50 P361A-a	CACCCTCGCATTTATGAG**g**CCTTTGTGACAAGTGTCGAGAAGC	(229)
Mga50 P361A-b	TCTTCTCGACACTTGTCACAAAGG**c**CTCATAAATGCGAGGGTG	(229)
Mga50 R461M-a	CAGTACGATGTTATCGTGACAGATGTTA**t**GGTGGGTAAAAGCGAAG	(229)
Mga50 R461M-b	CTTCGCTTTTACCCACC**a**TAACATCTGTCACGATAACATCGTACTG	(229)
mrp49		
mrp49-L1	CAAGCTAAGCTAGATACAGCAACT	(229)
mrp49-R1	CGCCTGTTGACGGTAATT	(229)
Mutagenic *dmgB* oligos		
dmgB Q11R-a	CTCTTTACAACAAAAC**g**GTGGAGAGAATTGGAGCTAATTGCGC	(228)
dmgB Q11R-b	GCGCAATTAGCTCCAATTCTCTCCAC**c**GTTTTGTTGTAAAGAG	(228)
dmgB R13A-a	CTCTTTACAACAAAACAGTGG**gc**AGAATTGGAGCTAATTGCGC	(228)
dmgB R13A-b	GCGCAATTAGCTCCAATTCT**gc**CCACTGTTTTGTTGTAAAGAG	(228)

Table 5: continued

Target/Primer	Sequence	Reference
dmgB L42A-a	GTGTGAGAGATTAAACTGCTCACTC**gc**AACTTTACAATCATGTG	(228)
dmgB L42A-b	CACATGATTGTAAAGTT**gc**GAGTGAGCAGTTTAATCTCTCACAC	(228)
dmgB HTH4-a	GAGCTGTTTGTCAGC**gc**G**g**CAACACTCAAGCGTTTGATTG	(228)
dmgB HTH4-b	CAATCAAACGCTTGAGTGTTG**c**C**gc**GCTGACAAACAGCTC	(228)
P*arp*		
Parp4-L	CATTGACAGTGATCGCATCT	(229)
Parp4-R	AAGCGAATACTGTTTATTCG	(229)
P*emm*		
Pemm-L1	GCATGGATCCCATCGCAAAGAGCTTA	(151)
Pemm-R1	GCGGCTCGAGTAGTGTCTATTCGTGTTATT	(151)
P*mga*		
OYL-25	TACCATAAAATACCTTTC	(151)
OYR-25	GGTTGTACCATAACAGTC	(151)
Pmga-B	gcgggatccTAAGTTAACCAGTTCACAAA	(154)
Pmga-X	ggctcgagACCTTGTATACCCTTCTTTT	(154)
P*mrp*		
Pmrp4-L	AGCCAGACAATTCAGTTAAA	(229)
Pmrp4-R	CTCAGTGAATAGAGTTTGTTTG	(229)
Pmrp4-EcoRI	gcgggaattcTCAATTTCTAAGAATTGTG	(229)
Pmrp4-BglII	gcggagatctGGATTTCAGACGTCAT	(229)
P*rpsL*		
GASrpsL-EcoRI	gggcgaattcTGTCTAAAATCACATCTTCG	(229)
GASrpsL-Hind	ggggaagcttGGTTGATATAGCACTTGGTGAC	(229)
Other		
SpacI	CTACATCCAGAACAACCTCTGC	(229)
T7-term	CTAGTTATTGCTCAGCGG	Novagen

Table 6: Plasmids

Plasmid	Relevant Characteristics	Ref. or Source
pBlue-gus#2	*gusA* under P*mga6* in Bluescript vector	
pBluescriptII KS-	Blue/white cloning vector	Stratagene
pE14A	M6 *mga* mutant E14A under constitutive P*spac*	This study
pET21a	6X-His vector	Novagen
pH204A	M6 *mga* mutant (H204A) under the native P*mga6*	This study
pH324A	M6 *mga* mutant (H324A) under the native P*mga6*	This study
pJRS312	Spectinomycin cassette	(196)
pJRS515	WT *mga6* in pBluescript II KS-	(149)
pJRS547	Deletion of 2.1 kb of *mga* and P*mga* (*mga-16*)	(229)
pJRS2050	WT *mga6* under the native P*mga6*	(7)
pJRS9160	Counter-selectable gene-replacement vector	(151)
pKSM140	promoterless *gusA* in pCIV2 suicide vector	(229)
pKSM152	*S. dysgalactiae* locus containing both WT *dmgB* and *demB*	(228)
pKSM155	N-terminal MBP fusion to M4 Mga	(229)
pKSM156	N-terminal MBP fusion to M50 Mga	(229)
pKSM162	WT *mga6* under constitutive P*spac*	(151)
pKSM164	WT *mga6-his* under native P*mga6*	This study
pKSM170	WT M6 *mga-his* in *E. coli* expression vector	(4)
pKSM174	*mga-16* deletion allele in pJRS9160	(229)
pKSM300.A	M6 *mga* mutant (D225A) under the native P*mga6*	This study
pKSM300.N	M6 *mga* mutant (D225N) under the native P*mga6*	This study
pKSM301.A	M6 *mga* mutant (D459A) under the native P*mga6*	This study
pKSM301.N	M6 *mga* mutant (D459N) under the native P*mga6*	This study
pKSM302.A	M6 *mga* mutant (D225A) under constitutive P*spac*	This study
pKSM302.N	M6 *mga* mutant (D225N) under constitutive P*spac*	This study
pKSM303.A	M6 *mga* mutant (D459A) under constitutive P*spac*	This study
pKSM303.N	M6 *mga* mutant (D459N) under constitutive P*spac*	This study
pKSM315	WT *mga50* under constitutive P*spac*	(229)
pKSM317-H	*mga50-his* (S26N) under constitutive P*spac*	(229)
pKSM318	WT *mga6-his* under the constitutive P*spac*	(229)
pKSM318.1H	M6 *mga* mutant in HTH-1 under constitutive P*spac*	(228)
pKSM318.4H	M6 *mga* mutant in HTH-4 under constitutive P*spac*	(228)
pKSM320	WT *mga4-his* under constitutive P*spac*	(229)
pKSM321	WT *mga50-his* under constitutive P*spac*	(229)
pKSM322	WT *mga4-his* under native P*mga4*	(229)
pSKM323	*mga50-his* (S26N) under native P*mga4*	(229)
pKSM324	WT *mga4-his* under constitutive P*rpsL*	(229)
pKSM325	WT *mga50-his* under constitutive P*rpsL*	(229)
pKSM326	*mga50-his* (S26N) under constitutive P*rpsL*	(229)
pKSM327	*mga50-his* (P361A) under constitutive P*rpsL*	(229)
pKSM328	*mga50-his* (R461M) under constitutive P*rpsL*	(229)
pKSM329	WT *mga50-his* under native P*mga4*	(229)
pKSM330	*mga50-his* (R461M) under native P*mga4*	(229)

Table 6: continued

Plasmid	Relevant Characteristics	Ref. or Source
pKSM331	*mga50-his* (S26N/R461M) under native P*mga4*	(229)
pKSM332	*mga50-his* (S26N/R461M) under constitutive P*rpsL*	(229)
pKSM333	WT *mga6-his* under constitutive P*rpsL*	(229)
pKSM336	*mga50*-his (P361A) under native P*mga4*	(229)
pKSM337	*mga6*-his (M461R) under the native P*mga6*	(229)
pKSM338	*mga6-his* (M461R) under constitutive P*rpsL*	(229)
pKSM339	*dmgB* mutant (Q11R) under the native P*dmgB* promoter	(228)
pKSM340	*dmgB* mutant (R13A) under the native P*dmgB* promoter	(228)
pKSM341	*dmgB* mutant in HTH-4 under the native P*dmgB* promoter	(228)
pKSM342	*dmgB* mutant (L42A) in HTH-1 under the native promoter	(228)
pKSM344	M6 *mga-his* mutant (Q11R) in *E. coli* expression vector	(228)
pKSM345	M6 *mga-his* mutant (R13A) in *E. coli* expression vector	(228)
pKSM346	M4 *mga-his* mutant (Q11R) under native P*mga4*	(228)
pKSM347	M4 *mga-his* mutant (R13A) under native P*mga4*	(228)
pKSM348	M4 *mga-his* mutant in HTH-4 under native P*mga4*	(228)
pKSM349	M6 *mga* mutant (H204A) under constitutive P*rpsL*	This study
pKSM350	M6 *mga* mutant (H324A) under constitutive P*rpsL*	This study
pKSM352	M6 *mga* mutant (H204D) under constitutive P*rpsL*	This study
pKSM353	M6 *mga* mutant (H324D) under constitutive P*rpsL*	This study
pKSM354	M6 *mga-his* mutant (W12R) in *E. coli* expression vector	(228)
pKSM355	M6 *mga-his* mutant (E14A) in *E. coli* expression vector	(228)
pKSM358.A	M6 *mga* mutant (Q11A) under native P*mga6*	(228)
pKSM358.R	M6 *mga* mutant (Q11R) under native P*mga6*	(228)
pKSM359	M6 *mga* mutant (R13A) under native P*mga6*	(228)
pL15A	M6 *mga* mutant (L15A) under constitutive P*spac*	(228)
pMal-c2	*malE* MBP fusion vector	NEB
pMga4-4	WT *mga4* in pLZ12-Spec	(7)
pMga50-His	WT *mga50* gene with C-terminal 6X-His	(229)
pP361A-H	*mga50-his* (P361A) under constitutive P*spac*	(229)
pP*mga*-blue	P*mga* in pBluescript II KS-	(229)
pP*mga-gusA*	*gusA* under P*mga6* in the VIT vector	(229)
pP*mrp-gusA*	*gusA* under P*mrp* in the VIT vector	(229)
pQ10R	M6 *mga* mutant (Q10R) under constitutive P*spac*	(228)
pQ11A	M6 *mga* mutant (Q11A) under constitutive P*spac*	(228)
pQ11R	M6 *mga* mutant (Q11R) under constitutive P*spac*	(228)
pR13A	M6 *mga* mutant (R13A) under constitutive P*spac*	(228)
pR461M-H	*mga50*-his (R461M) under constitutive P*spac*	(229)
pVIT164	The GAS vector for integration	(31)
pW12A	M6 *mga* mutant (W12A) under constitutive P*spac*	(228)
pW12R	M6 *mga* mutant (W12R) under constitutive P*spac*	(228)

Construction of the Δmga allele in the M6 Pmrp-gusA reporter strain KSM149. An *mga*-deleted derivative of the serotype M6 VIT strain RTG229 (75) was constructed as follows: to allow the use of the erythromycin-resistant pKSM174, the erythromycin cassette found in the VIT locus of RTG229 was replaced with the spectinomycin resistance cassette from pJRS312 (196) to generate VIT230 (Table 4). An *mga*-deleted VIT230 strain was created by allelic replacement of its wild-type *mga* at the native locus with the *mga-16* deletion allele from pKSM174 as previously described (151) to produce VIT231 (Table 4).

An *mga*-deleted serotype M6 VIT strain containing a single-copy transcriptional fusion of the serotype M4 Mga-regulated *mrp* promoter to *gusA* (P*mrp-gusA*) was constructed as follows: a P*mga-gusA* fusion was first produced by amplification of a 493 bp P*mga* fragment from serotype M6 JRS4 gDNA using the primers Pmga-B and Pmga-X (Table 5). The resulting fragment was digested with BamHI and cloned into BamHI/SmaI-digested pBluescript II KS- (Table 6) to produce pP*mga*-blue. A 1.9 kb EcoRI/HindIII fragment containing the promoterless *gusA* gene from pKSM140 (191) was inserted into EcoRI/HindIII-digested pP*mga*-blue to generate pBlue-Gus#2. The resulting 2.4 kb HpaI/SalI P*mga-gusA* fragment was excised from pBlue-Gus#2 and cloned into SmaI/SalI digested pVIT164 (31) to produce pP*mga-gusA*. P*mrp* was amplified from serotype M4 AP4 gDNA using primers Pmrp4-EcoRI and Pmrp4-BglII (Table 5). P*mga* was excised from pP*mga-gusA* following digestion with EcoRI/HpaI and replaced with EcoRI-digested P*mrp* to produce pP*mrp-gusA*. This construct was

linearized with XmnI and transformed into the *mga*-deleted strain VIT231 to generate KSM149 (Table 4).

Construction of an M6 Pmga-gusA reporter strain KSM231.310 containing a Δ*Pmga-mga allele. KSM231.310 (Table 4) was generated by transforming the *mga*-deleted M6 VIT strain VIT231 with pP*mga-gusA* (229) linearized by XmnI, which created a Δ*mga* reporter strain containing a single-copy transcriptional fusion of the native M6 *mga* promoter to *gusA* (P*mga-gusA*).

MBP-Mga fusion proteins for over expression of Mga in E. coli. Plasmids expressing amino-terminal fusions of maltose binding protein (MBP) to the wild-type (WT) Mga from serotypes M4 and M50 (MBP-Mga) were constructed as follows: a 1.9-kb DNA fragment of only the *mga* coding sequence was amplified from genomic DNA (gDNA) from both AP4 (M4) or B514-Sm (M50) using the primers DivMga-5'blunt and DivMga-3'HIII (Table 5). The resulting PCR fragments were digested with HindIII and inserted into XmnI/HindIII-digested vector pMal-c2 (New England Biolabs) to generate the *malE-mga* fusion alleles for purification of M4 (pKSM155) and M50 (pKSM156) MBP-Mga, respectively (Table 6).

Construction of the Pspac-mga-his plasmids. A wild-type *mga50* gene containing a 3' 6X-*his* tag under the constitutive promoter P*spac* (P*spac-mga50-his*) was

generated as follows: the *mga50* gene was amplified from serotype M50 B514-Sm gDNA using primers DivMga-R9 and DivMga-Sph (Table 5), digested with HindIII/SphI and cloned downstream of P*spac* in HindIII/SphI digested pKSM162 (152), to create pKSM315. To produce an *mga50-his* fusion, the *mga50* allele was amplified from serotype M50 B514-Sm gDNA using primers DivMga-Pet1 and DivMga-R9 (Table 5), digested with NdeI/XhoI and cloned into the NdeI/XhoI-digested pET21a (Novagen), creating pMga50-His. The C-terminal 996 bp fragment of the *mga50-his* gene was amplified from pMga50-His using primers DivMga-R5 and MgaHis-Sph (Table 5), digested with BamHI/SphI and cloned into the BamHI/SphI-digested pKSM315, producing pKSM321 (Table 6). The P*spac-mga4-his* allele was generated as follows: *mga4* was amplified from pMga4-4 (7) using primers DivMga-R9 and DivMga-Pet2 (Table 5), digested with HindIII/XhoI and cloned into the HindIII/XhoI-digested pKSM321 to produce pKSM320 (Table 6). The S26N mutation in M50 *mga* was constructed as follows: the 917 bp HindIII/BamHI fragment from pKSM320 was cloned into HindIII/BamHI-digested pKSM321, producing pKSM317-H. Site-specific mutants of Mga50 P361A and R431M were generated in pKSM321 as described above using the mutagenic primers Mga50 P361A-a, Mga50 P361A-b, Mga50 R461M-a, and Mga50 R461M-b (Table 5) resulting in pP361A-H and pR431M-H, respectively (Table 6).

Construction of the PrpsL-mga-his plasmids. The wild-type and mutant *mga* alleles were placed under the constitutive promoter P*rpsL* as follows: a 385 bp region of

the *rpsL* promoter was amplified from serotype M6 JRS4 gDNA using the primers GASrpsL-EcoRI and GASrpsL-Hind (Table 5). P*spac* was excised from pKSM320 following digestion with XmnI/HindIII and replaced with the HindIII-digested P*rpsL* fragment to produce pKSM324 (Table 6). Plasmids pKSM325, pKSM326 and pKSM327 were all constructed in a similar fashion using pKSM321, pKSM317-H, and pP361A-H, respectively, as template vectors instead of pKSM320. The P*rpsL-mga50-his* allele containing the R461M mutation was constructed by inserting a 1.9 kb HindIII/SphI fragment containing the R461M mutation from pR431M-H into HindIII/SphI-digested pKSM324 to produce pKSM328 (Table 6). A double mutant of Mga50 (S26N and R461M) was generated in pKSM326, which already contains the S26N mutation, using the mutagenic primers Mga50 R461M-a and Mga50 R461M-b (Table 5) resulting in pKSM332 (Table 6).

Construction of the Pmga4-mga-his plasmids. An *mga4-his* fusion under the native *mga4* promoter (P*mga4*) was generated as follows: the 988 bp BamHI/SphI fragment containing *mga4-his* from pKSM320 was cloned into BamHI/SphI-digested pMga4-4 (7), which contains *Pmga4-mga4*, producing pKSM322. An S26N allele was constructed using the 988 bp BamHI/SphI fragment from pKSM321 cloned into the BamHI/SphI-digested pKSM322 to produce pKSM323 (Table 6). P*mga4-mga50-his* was generated as follows: the 1.9 kb NsiI/SphI fragment containing the *mga50-his* fusion from pKSM321 was cloned into NsiI/SphI-digested pMga4-4 producing pKSM329.

Plasmid pKSM330 (P*mga4-mga50*[R461M]-*his*) was constructed in a similar fashion extracting the NsiI/SphI fragment from pR431M-H instead of pKSM321. The P*mga4-mga50*[S26N, R461M]-*his* double mutant was produced by excising a 707 bp BamHI/XhoI fragment from pKSM332 and inserting it into BamHI/XhoI-digested pKSM322 vector producing pKSM331 (Table 6). A P*mga4-mga50*[P361A]-*his* was generated as described above using the mutagenic primers Mga50 P361A-a, and Mga50 P361A-b (Table 5) in pKSM329 resulting in pKSM336.

Construction of M6 mga Mutant Alleles (M461R) Under Pmga6 and PrpsL. A P*mga6-mga6-his* plasmid containing the M461R mutation was generated in pKSM164 (4) via site-specific mutagenesis using the mutagenic primers M6 M461R-a, and M6 M461R-b (Table 5) to produce pKSM337 (Table 6). The wild-type and mutant *mga6* genes were placed under the constitutive expression of P*rpsL* as follows*:* a 3' 1.5 kb region of *mga6-his* was amplified from pKSM164 (4) using the primers OYR-29 and MgaHis-BglII (Table 5) and digested with SpeI/BglII. The 3' end of P*spac-mga6* was excised from pKSM162 (151) following digestion with SpeI/BglII and replaced with the digested PCR fragment to produce pKSM318. The *mga6-his* gene was then amplified from pKSM318 using the primers SpacI and T7-term (Table 5), digested with XhoI/HindIII and into XhoI/HindIII-digested pKSM328, producing pKSM333 (Table 6). An M6 Mga M461R was generated using the mutagenic primers M6 M461R-a, and M6 M461R-b (Table 5) in pKSM333 resulting in pKSM338 (Table 6).

Construction of the site-specific Pspac-mga6-his plasmids pQ10R, pQ11A, pW12A, pW12R, pR13A, pE14A, and pL15A. Site-specific mutants of M6 Mga produced from the constitutive promoter P*spac* and containing a carboxy-terminal 6X-His tag were generated as described above using the QuickChange Site-Directed Mutagenesis Kit (Stratagene), pKSM318 (Table 6) template DNA, and the corresponding *mga6* mutagenic oligonucleotides (Table 5), resulting in the respective plasmids.

Construction of the site-specific Pmga6-mga6-his mutant plasmids pKSM358.A, pKSM358.R and pKSM359. Site-specific mutants of M6 Mga produced from the native promoter P*mga6* and containing a carboxy-terminal 6X-His tag were generated using the QuickChange Site-Directed Mutagenesis Kit (Stratagene), pKSM164 (Table 6) template DNA, and the mutagenic oligonucleotides pairs M6 mga Q11A-a and M6 mga Q11A-b; M6 mga Q11R-a and M6 mga Q11R-b; or M6 mga R13A-a and M6 mga R13A-b (Table 5) to create plasmids pKSM358.A, pKSM358.R and pKSM359, respectively.

Construction of the control Pspac-mga6-his plasmids pKSM318.1H and pKSM318.4H. Plasmids that contained mutations to either disrupt the essential DNA-binding HTH-4 motif or two non-essential residues of M6 *mga* under the constitutive P*spac* were constructed as follows: the 464 bp HindIII/SpeI fragment containing either the HTH-4 motif or K31A/D32A from pKSM164.4c or pKSM164.1c (151), respectively,

was cloned into HindIII/SpeI digested pKSM318 (229) to produce pKSM318.4H and pKSM318.1H.

Construction of the Pmga4-mga4-his mutant plasmids pKSM346, pKSM347, and pKSM348. Site-specific mutants of the class II serotype M4 Mga Q11R and R13A were generated as described above using the mutagenic oligos M4 *mga* Q11R-a, M4 *mga* Q11R-b, M4 *mga* R13A-a, and M4 *mga* R13A-b (Table 5) in pKSM322 (Table 6), resulting in pKSM346 and pKSM347, respectively. A mutant disrupting the HTH-4 domain, named pKSM348, was also constructed by site-specific mutagenesis of pKSM322 using the mutagenic oligos M4 *mga* HTH4-a and M4 *mga* HTH4-b (Table 5).

Construction of the PdmgB-dmgB mutant plasmids pKSM339, pKSM340, pKSM341 and pKSM342. Site-specific mutations at positions 11 and 13 were constructed in the Mga orthologue from *S. dysgalactiae,* DmgB, using the mutagenic oligos *dmgB* Q11R-a, *dmgB* Q11R-b, *dmgB* R13A-a, and *dmgB* R13A-b (Table 5) in pKSM152 (Table 6), resulting in pKSM339 and pKSM340, respectively. Control plasmids containing either mutations in the HTH-4 domain, pKSM341, or an arbitrary mutation at position 42, pKSM342, were also constructed by site-specific mutagenesis of pKSM152 using the mutagenic oligos *dmgB* HTH4-a, *dmgB* HTH4-b, *dmgB* L42A-a and *dmgB* L42A-b, respectively (Table 5).

Mga-His fusion proteins for overproduction of Mga mutants in E. coli. Plasmids producing carboxy-terminal 6X-His fusions to M6 Mga mutants Q11R, W12R, R13A and E14A were constructed as follows: a 1.6-kb DNA fragment of the *mga* coding sequence was amplified from plasmid DNA encoding the corresponding mutant pQ11R, pW12R, R13A and pE14A using the primers Mga-Pet1_Nde and Mga-Pet2_Xho (Table 5). The resulting PCR fragments were purified, digested with NdeI/XhoI, and inserted into NdeI/XhoI digested vector pKSM170 (4) to generate the *mga-his* fusion alleles for purification of Q11R-His (pKSM344), W12R-His (pKSM354), R13A-His (pKSM345) and E14A-His (pKSM355), respectively (Table 6).

Construction of the M6 mga receiver domain mutations under the constitutive Pspac promoter. Plasmids that contained a mutation changing the Asp residues to either an Ala, Asn or an Glu in both receiver domain 1 and 2 under the constitutive promoter P*spac* were constructed as follows: the 1.3 kb SpeI/SphI-digested fragment from plasmids pKSM300.A, pKSM300.N, pKSM300.E pKSM301.A, pKSM301.N and pKSM301.E (Table 6), which contain the receiver domain mutations, were cloned into the SpeI/SphI-digested vector containing P*spa*c (pKSM162), resulting in plasmids pKSM302.A, pKSM302.N, pKSM303.A and pKSM303.N. The precursor plasmids pKSM300.A, pKSM300.N, pKSM300.E, pKSM301.A, pKSM301.N, and pKSM301.E containing the site specific receiver domain mutants of the M6 Mga (D225A, D225N, D225E, D459A, D459N, and D459E, respectively) were under the control of the native

promoter P*mga6* and were generated using the QuickChange Site-Directed Mutagenesis Kit (Stratagene) with the *mga6* mutagenic oligonucleotides DR1A-a, DR1A-b, DR1E-a, DR1E-b, DR1N-a, DR1N-b, DR2A-a, DR2A-b, DR2E-a, DR2E-b, DR2N-a and DR2N-b (Table 5) and pJRS2050 (Table 6) as template DNA.

Plasmids that contained a mutation changing the Lys residues to an Arg in both receiver domain 1 and 2 under the constitutive promoter P*spac* were constructed as above except the 1.3 kb SpeI/SphI digested fragments were from plasmids pKSM304.R or pKSM305.R (Table 6), which contain the receiver domain mutations. They were cloned into the SpeI/SphI digest vector containing P*spac* (pKSM162) resulting in plasmids pKSM306.R and pKSM307.R. The precursor plasmids pKSM304.R and pKSM305.R containing the site-specific receiver domain mutants of the M6 Mga (K285R and K512R respectively) were under the control of the native promoter P*mga6* and were produced using the QuickChange Site-Directed Mutagenesis Kit (Stratagene), pJRS2050 (Table 6) template DNA, and the *mga6* mutagenic oligonucleotides KR1R-a, KR1R-b, KR2R-a, and KR2R-b, (Table 5).

Construction of the M6 mga PRD domain mutations under the constitutive PrpsL promoter. Site specific-mutants of the class I serotype M6 Mga H204A and H324A were first generated under the native *mga* promoter using the mutagenic oligos M6 mga H204A-a, M6 mga H204A-b, M6 mga H324A-a, and M6 mga H324A-b (Table 5) in pKSM164 (Table 6), resulting in plasmids pH204A and pH324A, respectively.

Then, a 793 bp SpeI/XmnI fragment containing the mutated *mga* from pH204A and pH324A was ligated into the SpeI-XmnI-digested vector pKSM333 creating pKSM349 and pKSM350, respectively. Transformants were sequenced for accuracy.

Site-specific mutants H204D and H324D under P*rpsL* were generated using the mutagenic oligos M6 mga H204D-a, M6 mga H204D-b, M6 mga H324D-a, and M6 *mga* H324D-b (Table 5) in pKSM333 (Table 6) according to the Quickchange Mutagenesis kit (Stratagene) resulting in pKSM352 and pKSM353, respectively.

II. DNA Manipulation

A. Isolation and manipulation of plasmid DNA from *E. coli*

Plasmid DNA was isolated from *E. coli* via alkaline lysis using the Wizard Miniprep kit (Promega), the Plasmid Midi Kit (Qiagen), or the HiSpeed Plasmid Maxi Kit (Qiagen) according to the manufacture's protocol. DNA fragments were purified from agarose gels with the QIAquick gel extraction kit (Qiagen) using the manufacturer's microcentrifugation purification instructions. Plasmid DNA was manipulated with the appropriate enzymes from New England Biolabs and buffer exchanges were performed using the QIAquick PCR purification system (Qiagen) as necessary.

B. Isolation of plasmid DNA from the GAS

Cells from a 50 ml overnight culture were resuspended in P1 solution (Qiagen) (Table 7) containing 20 mg/ml lysozyme, mixed and incubated for 30 min at 37°C. Samples were lysed with the FastPrep Cell Disruptor (Bio101), allowed to settle and the supernatant was mixed with 300 µl of a 1% SDS, 0.2 M NaOH solution. After 5 min at room temperature, 300 µl of 2.5 M potassium acetate pH 4.5 was added, mixed by inversion, and centrifuged (5 min, 13,000 rpm). The supernatant was applied to a mini prep column (Qiagen) and plasmid DNA eluted using the manufacturer's protocol.

C. Fast DNA GAS chromosomal prep

Genomic DNA was isolated from a 50 ml overnight culture using the FastDNA prep and a FastPrep cell disruptor (Bio101). Briefly, cells were homogenized using two ¼ inch ceramic spheres in conjunction with the garnet matrix and the cell lysis solution TC. Sample was centrifuged at 14, 0000 X g to pellet debris and lysing matrix. The supernatant was transferred and a binding matrix added. The matrix was washed with an EtOH solution and then the DNA eluted from the matrix using DNase-free water.

Table 7: Solutions and buffers

The solution or buffer name is underlined with its components written directly beneath.

5X Band Shift Buffer
10X Transcription buffer
50 ng/µl poly[d(I-C)]
500 µg/ml BSA
(Stored at 4ºC)

Coomassie Blue Stain
0.25% (w/v) Coomassie brilliant blue
45.4% (v/v) methanol
9.2 % (v/v) glacial acetic acid

Coomassie Blue Destain
5% (v/v) methanol
7.5% (v/v) glacial acetic acid

5X Cracking Buffer
0.3 M Tris pH 6.8
25% (v/v) 2-mercaptoethanol
51% (v/v) glycerol
10% (w/v) SDS
0.01% (w/v) bromophenol blue
(Heated at 55ºC until
SDS goes into solution)

CsCl Cushion
5.7 M CsCl in 0.1 M EDTA

Gel Drying Solution
3% (v/v) glycerol
20% (v/v) EtOH

Lower Gel Stock
1.5 M Tris HCl pH 8.8
0.4% (w/v) SDS

Lysis Solution
0.5 M Tris pH 7.0
20 mM EDTA
10% (w/v) SDS

Lysozyme Solution
100 mM Tris pH 7.0
2 mM EDTA
2mg/ml lysozyme

MBP Column buffer
2 M Tris base
2 M glycine
1% (w/v) SDS

P1 Buffer
0.05M Tris pH 7.5
10 mM EDTA
200 µg/ml RNase A
(Stored at 4ºC)

8X PBS for chromatography
79 mM $Na_2HPO_4 \cdot 2H_2O$
80 mM $NaH_2PO_4 \cdot H_2O$
4 M NaCl
(adjust pH to 7.4)

PBS-Tween
0.01 M PBS pH 7.4
0.5% Tween-20

10X SDS Running Buffer
3% (w/v) Tris base
14.4% (w/v) glycine
1% (w/v) SDS

10% SDS Resolving Gel
26% (v/v) Lower gel stock
25% (v/v) 40% Acrylamide/bis solution
0.5% (v/v) 10% ammonium persulfate
0.1% (v/v) tetramethylethylene diamine

12% SDS Resolving Gel
26% (v/v) Lower gel stock
32% (v/v) 40% Acrylamide/bis solution
0.5% (v/v) 10% ammonium persulfate
0.1% (v/v) tetramethylethylene diamine

6% SDS Stacking Gel
25% (v/v) Upper gel stock
15% (v/v) 40% Acrylamide/bis solution
0.5% (v/v) 10% ammonium persulfate
0.1% (v/v) tetramethylethylene diamine

5X TBE 5.4% (w/v) Tris base 2.75% (v/v) boric acid 0.5 M EDTA pH 8.0 **TE Buffer** 10 mM Tris pH 7.4 1 mM EDTA pH 7.5 **10X Transcription Buffer** 120 mM HEPES pH 7.5 10 mM EDTA 6 mM DTT 600 mM KCl 50 mM $MgCl_2$ (Store at –20 ºC) **10X Transfer Buffer** 250 mM Tris base 2 M glycine (dilute 10X to 1X in 20% methanol and store at 4ºC)	**Tris Solution** 100 mM Tris pH 7.0 2 mM EDTA 0.06% Sodium azide **Upper Gel Stock** 0.5 M Tris HCl pH 6.8 0.4% (w/v) SDS **Western Stripping Solution** 2% (w/v) SDS 50 mM DTT 50 mM Tris HCl pH 7.0 **X-Gluc** 40 mg/ml in N-N-dimethyl formamide

D. Polymerase chain reaction

Polymerase chain reaction (PCR) was performed using either Taq (NEB), Platinum Taq (Invitrogen) or *Pfu* Turbo DNA polymerase (Stratagene) according to the manufacturer's protocol. Amplification was performed on an Eppendorf Mastercycler gradient using the following parameters: an initial denaturation step of 95ºC for 3 min followed by thirty cycles of denaturation at 95ºC for 30 seconds, annealing for 1 min at a temperature ranging from 48-65ºC depending upon the specific primer set (Table 5), and extension at 72ºC for 1.5 min/kb of DNA to be amplified. The final cycle consisted of an extra extension step at 72ºC for 3 min to ensure the production of full-length products.

Resulting amplification products were purified with the QIAquick PCR purification system (Qiagen).

E. PAGE purification of oligonucleotides

Oligonucleotides were purified by resuspending the lyophilized oligonucleotide pellet in 100 µl dH_2O plus 20 µl of 5x Formamide Stop Solution (Epicentre SequiTherm Excel II Kit). The entire solution was loaded onto a 10% sequencing gel (Amresco) and electrophoresed at 400 volts for 2 hrs. The top plate was removed, the gel was covered by saran wrap, and exposed to short wave UV light to locate the oligonucleotides. The major product was excised and crushed via pestle. The slice was soaked in 1 ml of 10 mM Tris pH 8.0 and incubated at 37ºC with agitation overnight. The next day the gel debris was separated by centrifugation and the soluble fraction transferred to clean tube. The supernatant was separated using a phenol:chloroform extraction before undergoing ethanol precipitation. The precipitated pellet was resuspended in 100 µl 10 mM Tris pH 8.0 and the concentration determined from the OD_{260}.

F. Phenol:chloroform extraction

The following phenol:chloroform extraction was performed as a step in various protocols involving both DNA and oligonucleotide purification. One volume of phenol:chloroform (1:1) was added to the sample. The sample was mixed via inversion and spun at 13,000 rpm for 3 to 6 min at RT. The aqueous phase (top) was transferred to a clean tube. The phenol was subsequently removed from the sample by the addition of

one volume of chloroform:Iso-Amyl alcohol (24:1). The sample was mixed, spun and transferred to a clean tube as above. The nucleic acids were then pelleted via ethanol precipitation.

G. Ethanol precipitation

Both DNA and RNA were precipitated in the following manner. Cold 100% ethanol in an amount equivalent to 2.5 times the total reaction volume and 3 M sodium acetate pH 5.2 equivalent to 1/10 the reaction volume was added to the DNA or RNA containing sample. The solution was mixed by inversion and incubated at -20ºC for 30 min. The solution was then spun at 13,000 rpm for 15 min at 4ºC to pellet the sample. The sample was washed with 500 µl 75% ethanol. The ethanol was removed and the pellet was dried for 10 min at RT. The pellet was then resuspended in either 10 mM Tris pH 8.0 or H_2O.

H. Site-specific mutagenesis and random mutagenesis screen

All site-specific mutations were generated with the QuickChange Site-Directed Mutagenesis Kit (Stratagene) using mutagenic oligonucleotides synthesized by Integrated DNA Technologies. DNA sequencing was done by the McDermott Center sequencing core facility at UT Southwestern Medical Center.

Random mutagenic PCR was performed across the 5' region of *mga* using the GeneMorph PCR Mutagenesis Kit (Stratagene) on plasmid pKSM162 (151) using primers RMut1-R2 and Mga6-150 (Table 5) with an initial amount of DNA of 0.110 µg.

The resulting PCR product was digested with AvaI/SpeI, ligated with AvaI/SpeI-digested pKSM318 (229), transformed into *E. coli* DH5α and grown overnight in 10 mL LB broth containing spectinomycin. A similar procedure was performed for the 3' portion of *mga* except primers OYR-29 and RMut2-L were used on various concentrations of pKSM162 ranging from 0.055 μg to 0.838 μg. Also both the PCR product and the vector pKSM318 were digested with SpeI/SphI for this fragment. Plasmid DNA was extracted from the entire overnight culture, and 1 μg of plasmid DNA was transformed into the reporter strain KSM148.174 (229). Following selection on THY agar containing spectinomycin and 175 μg/ml X-glu (Gold Biotechnology), GAS transformants that were either light blue or white in color on plates were subjected to western analysis to determine protein levels as described. Plasmid DNA was extracted from the selected clones and amplified via subsequent transformation into *E. coli* DH5α for DNA sequencing. Sequence alignments were performed with the Vector NTI software (Invitrogen) using the ClustalW algorithm, and genes containing more than two mutations were discarded from further analysis.

III. Transformation of Bacterial Strains

A. Drop dialysis

To remove excess salts prior to electroporation samples were dialyzed against dH_2O using a 25 mm, type-VS Millipore membrane with a pore size of 0.025 μm. Samples were incubated for at least 30 min at room temperature.

B. Preparation of electrocompetent *E. coli* cells

Two 500 ml cultures were grown to an OD_{600} of 0.5 to 0.8 before being chilled on ice for 30 min. Cultures were pelleted by centrifugation at 5,500 rpm at 4°C for 15 min. The pellet was washed once with 300 ml cold dH_2O and once with 20 ml of a cold 10% glycerol (v/v) solution before the final resuspension in 800 μl cold 10% glycerol. Cells were stored in 50 μl aliquots at -80°C.

C. Preparation of electrocompetent GAS cells

Cultures were grown in 150 ml THY containing 20 mM glycine to an OD_{600} of 0.2-0.4 except the GAS strain B514-Sm, which was grown in 2X BHI supplemented with 2% isovitalex, 85 μg/ml hyaluronidase and 20 mM glycine to an OD_{600} of 0.6. Cells were pelleted by centrifugation at 8,000 rpm for 20 min at 4°C. Pellets were washed twice with a cold 10% glycerol solution. Cells were resuspended in 1ml of a cold 10% glycerol solution before being stored at -80°C in 200 μl aliquots.

D. Electroporation of *E. coli* and the GAS strains

Dialyzed plasmid samples were placed in a prechilled cuvette containing 50 μl of the appropriate *E. coli* strain. Electroporation was carried out using a Bio-Rad gene pulser set at a resistance of 200 ohms, a capacitance of 25 μF and a voltage of 2.5 kV. After electroporation, samples were grown for 1 hour at 37°C in 1 ml of LB before being plated on LB agar containing the appropriate antibiotics. The GAS strains were electroporated as above with the following modifications: 200 μl of the appropriate GAS strain was used, the Bio-rad gene pulser was set at a resistance of 400 ohms and a voltage of 1.75 kV and samples were outgrown for 2 hour at 37°C in 10 ml of THY before being plated on THY agar with the appropriate antibiotics.

E. Allelic replacement using a temperature-sensitive plasmid

To inactivate *mga* through double homologous recombination in the GAS, a fragment internal to the coding region was cloned into the vector pJRS9160, which is temperature sensitive for replication. Approximately 30 μg of the resulting plasmid was electroporated into the parent GAS strain and outgrown at the permissive temperature of 30°C for 3 hours. The cells are then pelleted at 6,000 rpm for 20 min at 4°C, resuspended in 500 μl saline, and plated onto THY agar with erythromycin. The plates were incubated at 30°C for 3 days at which time single colonies were chosen for overnight growth in 10 ml THY containing erythromycin at 30°C. Fresh THY broth with erythromycin was inoculated with 10 μl of culture from the previous day and incubated at

the non-permissive temperature of 37°C overnight. 10 µl of this culture was then passaged at 37°C overnight in THY broth that does not contain any antibiotics. From this culture, serial dilutions were plated on THY agar plates containing streptomycin and grown overnight at 37°C. Transformants were screened using PCR to ensure the allelic exchange event occurred.

IV. Protein Production, Purification and Analysis

A. Protein expression and purification

MBP-Mga proteins were purified from *E. coli* as previously described (149). Briefly, cultures of *E. coli* SA2817 containing either pKSM155 or pKSM156 were grown in LB broth containing ampicillin at 30°C and protein producation induced by the addition of 6 mM IPTG (isopropyl-β-D-thiogalactopyranoside) for 2 hours. Whole-cell lysates were obtained by passing cells twice through a pre-chilled French pressure cell. MBP-Mga proteins were purified from soluble lysates on an amylose resin column (New England Biolabs) and proteins were eluted from the affinity resin with 10 mM maltose. Protein concentrations were determined using a protein assay kit (Bio-Rad). Protein purity was assessed by SDS-PAGE stained with Sypro Ruby (Sigma) and by western blot analysis probed with αMBP antisera (New England Biolabs) as described below.

Mga-His proteins were purified from *E. coli* as previously described (4). Briefly, *E. coli* strains containing the appropriate fusion plasmid were grown in LB broth

containing ampicillin at 30°C and protein producation induced by the addition of 6 mM IPTG (isopropyl-β-D-thiogalactopyranoside) for 2 hours. Whole-cell lysates were obtained by passing cells twice through a pre-chilled French pressure cell. His tagged proteins were purified from soluble lysates on an NiNTA resin column (Amersham) and proteins were eluted from the affinity resin with 250-500 mM imidazole. Protein concentrations were determined using a protein assay kit (Bio-Rad). Protein purity was assessed by SDS-PAGE stained with Coomassie and by western blot analysis probed with αHis antisera (Novagen) as described below.

B. Whole-cell protein extraction from *E. coli*

Whole-cell lysates were extracted from *E. coli* by resuspending pelleted samples in cracking buffer (Table 7). Samples are then boiled for 5 min and centrifuged to remove debris. Soluble lysates are separated by SDS-PAGE and visualized by Coomassie stain.

C. Whole-cell protein extraction from the GAS

Whole-cell GAS proteins were extracted as previously described (151). Briefly, mid-logarithmic phase GAS cultures were harvested by centrifugation and resuspended in saline containing 1X Complete protease inhibitor cocktail (Roche). Cells were lysed using a FastPrep cell disruptor (Bio101, Inc.) and soluble lysates recovered by centrifugation. Total protein concentrations were determined using the Protein Assay kit (Bio-Rad).

D. Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE)

Samples for SDS-PAGE were mixed with 1X cracking buffer (Table 7) and boiled for 5 min. Samples were separated on either 10 or 12% polyacrylamide gels with a 6% stacking gel (Table 7) using the Mini-Protean 3 cell (Bio-Rad). Gels were electrophoresed at 150 volts at 4ºC in 1X SDS running buffer (Table 7).

E. Western blot analysis

Proteins were analyzed by western blot as previously described (151). Briefly, proteins were separated by either 10% or 12% SDS-PAGE, transferred to nitrocellulose membrane, and incubated with a 1:2,000 dilution of either αHis-tag (Novagen), αMBP antiserum (New England Biolabs), or a 1:1,000 dilution αMga-pep2 antiserum (151). Blots were then incubated with a 1:25,000 dilution of either anti-mouse (Chemicon), anti-rabbit (Sigma), or α-rat (Santa Cruz Biotechnologies) horseradish peroxidase-conjugated secondary antibody, respectively, and visualized using the Western lighting chemiluminescense system (Perkin Elmer) followed by exposure to film. As a loading control, blots were stripped at 70°C for 30 min in a solution of 2% SDS, 50 mM DTT and 50 mM Tris-HCl pH 7.0 and then probed with a 1:50,000 dilution of mouse αHsp60 antiserum (StressGen Biotechnologies Corp.) and visualized as described above.

F. Sequence Alignments

A sequence alignment of different Mga and Mga-like protein sequences was performed using the ClustalW algorithm in the AlignX module of VectorNTI (Invitrogen). The alignment included Mga protein sequences representing *mga-1* alleles from the GAS serotypes M1 SF370 (GI-15675800) and M6 JRS4 (GI-153733) as well as *mga-2* alleles from the GAS serotypes M4 AP4 (GI-1246852) and M49 531 (GI-56808536). Protein sequences for *S. pneumoniae* MgrA (GI-17368568), *S. dysgalactiae* subsp. *equisimilis* Mgc (GI-6782393), and *S. dysgalactiae* subsp. *dysgalactiae* DmgB (GI-6689248) were also utilized. Additional Mga-like proteins were obtained by BLAST interrogation of the unfinished genomes for *S. equi* (Sanger Institute), *S. mitis* (TIGR), and *S. gordonii* (TIGR). Conserved regions were defined as amino acids with $\geq$ 70% identity amongst homologues and spanned more than two adjacent residues.

G. Size exclusion chromatography

Size exclusion chromatography was performed at 4°C using a phosphate buffer solution (PBS) or PBS with 500 mM imidazole added as the eluent (Table 7). Elution profiles were obtained using an Mga-His protein (purified from *E. coli* lysates as described above) on either a Superdex 200 preparative grade or a Superdex 200 10/300 GL analytical grade column using an AKTA FPLC system (Amersham). The flow rate was set at 0.5 ml/min with fractions collected at 0.5 or 0.3 ml intervals.

V. Electrophoretic Mobility Shift Assay (EMSA)

EMSA was performed as previously described (149). Briefly, promoter probes were generated by PCR amplification of the appropriate genomic DNA using the relevant primer pairs (Table 5). PCR fragments were end-labeled with [γ-^{32}P]-ATP by T4 polynucleotide kinase (NEB) and the resulting labeled probes were purified from a 5% polyacrylamide gel by crush-and-soak elution. Constant amounts of probe end-labeled with [γ-^{32}P]-ATP were incubated with increasing concentrations of purified MBP-Mga or Mga-His protein for 15 min at 16°C before being separated on a 5% polyacrylamide gel. Gels were dried under vacuum at 80°C and exposed to film or a phosphoimager cassette.

VI. RNA Analysis

A. FastRNA RNA isolation

Total RNA was isolated from a 10 ml culture of the GAS using the FastRNA Pro Blue isolation kit (Bio101). Cultures were spun at 7,000 rpm to pellet the cells, pellet was resuspended in 1 ml RNApro Solution and transferred to a 1 ml tube containing lysing matrix B. Cells were lysed using the FastPrep cell disruptor at a speed of 6 for 40 sec. The tubes were chilled on ice for 5 min and then spun at 13,000 rpm for 5 min at 4ºC. The liquid phase was removed and incubated for 5 min at RT before 300 µl of chloroform was added. The solution was vortexed for 10 sec and respun in order to separate the phases. The soluble phase was extracted and precipitated with 500 µl cold

100% ethanol. The precipitated pellet was resuspended in 100 µl DEPC-dH_2O. Concentrations were determined using the NanoDrop 3.0.0 (NanoDrop Technologies) and samples were stored at –20ºC. When necessary, samples were treated with DNase to remove DNA contaminants.

B. Triton X-100 RNA isolation

RNA was isolated by using Triton X-100 according the protocol by Kidon Sung (219). Briefly, cell pellet was resuspended in 1 ml TE buffer containing 0.2% Triton X-100. The sample was boiled for 10 min, cooled and the RNA was isolated from the lysate using a chloroform extraction followed by ethanol precipitation. RNA concentrations were determined using the NanoDrop 3.0.0 (NanoDrop Technologies) and samples were stored at -20ºC.

C. CsCl RNA isolation

RNA was prepared in large quantities using the CsCl isolation method of Glisin (80). Briefly, a 50 ml culture (grown to mid-log phase) was mixed with 50 ml frozen Tris solution and gently swirled until thawed. Cells were pelleted by centrifugation at 10,000 rpm for 15 min at 4ºC. Pellets were brought up in 2.7 ml cold lysozyme solution and incubated for 20 min on ice. Lysis buffer (0.3 ml) was added and samples were boiled for 5 min. Samples were then mixed with 60 µl of 1M KCl, incubated on ice for 30 min and centrifuged at 7,000 rpm for 20 min at 4ºC. A 0.5 ml mixture of 0.4 M EDTA, 50 µl 2-mercaptoehtanol and CsCl (1 g/2.5 ml) was added to the supernatants and

incubated at 37ºC. The RNA-CsCl solution was layered onto a CsCl cushion and spun overnight at 35,000 rpm at 20ºC. RNA pellets were washed and resuspended in DEPC-dH_20. The NanoDrop 3.0.0 (NanoDrop Technologies) was used to determine RNA concentrations and samples were stored at -20ºC.

D. Random labeling of probes

DNA Probes were radiolabeled using the RadPrime DNA Labeling System (Invitrogen) according to the manufactures protocol. In short, 25 ng of a DNA probe was incubated with random hexanucleotide primers, dNTPs, the Klenow large fragment of DNA polymerase, and α^{32}P-dATP at 37ºC for 30 min. Unincorporated label was removed from the reaction using a TE select-D, G-25 Sephadex column (Roche). The activity of 1µl of probe was determined on a Beckman LS 6500 scintillation counter.

E. Northern and slot blot analysis

Northern blot analysis was performed on total RNA using the NorthernMax system (Ambion) as previously described (191). Briefly, a specified amount of total RNA was separated on a formaldehyde-agarose gel before being transferred to a nylon membrane. Blots were prehybridized for 30 min at 42ºC in ULTRAhyb solution followed by an overnight hybridization with 5 X 10^6 c.p.m. of a randomly labeled DNA probe. Blots were exposed to a phosphor screen, and scanned using a Storm 820 phosphor imager (Amersham) for visualization. As a loading control, blots were stripped with 2X SSC/50% formamide at 68°C for 2 hrs. and probed again with a 23S rRNA

probe amplified from serotype M6 JRS4 genomic DNA (Table 5). Slot blot analysis was performed as previously described (152). In brief, heat denatured samples of total RNA were transferred to a nylon membrane using the Bio-dot SF apparatus (BioRad). Blots were prehybridized, hybridized with a randomly labeled DNA probe and visualized as described above.

VII. GusA-Reporter Assay

GusA activity was determined for each lysate as described by Ribardo *et al.* (191). Concisely, whole-cell lysates of a GAS culture were harvested from 2-3 ml of mid-log cells using the FastPrep cell disruptor (Bio101) as described previously. GusA activity was determined by a colorimetric assay using 100 μl aliquots of sample incubated at 37ºC with pNitrophenyl β-D glucuronide until a yellow color develops. Total protein concentrations were determined using a protein assay kit (Bio-Rad). GusA units are defined as the OD_{420} / total protein concentration of the lysate.

CHAPTER FOUR:

RESIDUES RESPONSIBLE FOR THE DEFECTIVE MGA PRODUCED BY THE NATURAL OCCURING M50 MUTANT

I. Introduction

Although the GAS is considered a strict human pathogen, one serotype, classified as M50, has been found to primarily infect mice. Historically, this serotype was uncovered as the causative agent of several outbreaks of fatal cervical lymphadentis in mouse colonies at various institutions dating back to the 1930's (98). Since that time, serotype M50 B514 has only been isolated from the throats of individuals if they were working directly with the organism, demonstrating that it retains the capacity to colonize humans although this is not its natural reservoir (128). Initial investigations of the serotype M50 B514 strain using the spontaneous streptomycin resistant derivative B514-Sm revealed a Mga protein that was defective in activation of transcription of the downstream genes *mrp50,* encoding an M-like protein, and *emm50,* encoding a class II M protein (241, 242). Furthermore, the defective M50 *mga-2* allele exhibited 98% nucleotide identity to *mga* encoding a fully functional Mga found in other class II serotype M4 and M49 strains, which express *mga-2* alleles (242). In that study, the four amino acids at positions 26, 361, 461, and 521 (denoted as a star in Figure 3) were reported to be different between the M50 Mga protein and M4 and M49 Mga. However,

in this study only the first three amino acid differences were found to exist and an amino acid alignment revealed that none of the changes resided within the established HTH DNA-binding motifs (151) or at other sites predicted to be important for Mga activity (242) (Figure 3). To date, the specific residues responsible for the reduction in activation have not been characterized; therefore, mutational analysis of the differing residues was used to establish which amino acids are important for Mga-specific transcription of the downstream gene *mrp* and autoactivation of *mga-2* in these strains.

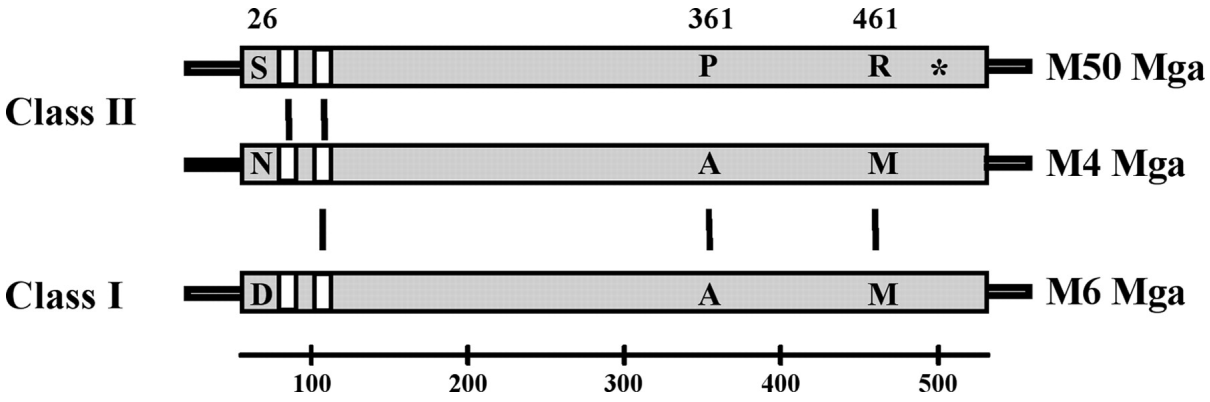


Figure 3: Schematic comparison of class I and class II Mga proteins

Mga homologues are found in all serotypes of the GAS and can range in size from 530 amino acids (M6) to 533 amino acids (M4 and M50). Based upon amino acid alignment, both class I and class II Mga proteins contain two N-terminal HTH domains (white boxes). The three non-conservative amino acids differences detected between Mga4 and Mga50: positions 26, 361, and 461 are indicated with their respective amino acids shown in each homologue. A fourth non-conservative change at position 521 identified in a previous study (242), but not in the present study, is also indicated (*). Solid black lines represent residues or domains relevant to this study that are 100% conserved among the proteins.

II. Results

A. M50 MBP-Mga fusion protein exhibits DNA-binding activity.

Previous studies have shown that the *mga-2* alleles found in serotypes M50 and M4 are almost identical at the amino acid level. Yet, Mga from the M50 strain is defective in activating transcription of downstream Mga-regulated genes compared to its fully functional M4 counterpart (241, 242). M50 Mga was originally thought to possess four amino acid changes that differed from Mga proteins produced by serotypes M4 and M49 (242). We now report that after multiple independent rounds of sequencing from the original serotype M50 strain B514-Sm gDNA only three of these non-conservative amino acid changes were observed (Figure 3; S26, P361, and R461). Although none of the changes were located within the two recently established helix-turn-helix DNA-binding domains of Mga (151), it was important to rule out the possibility that the three mutations in M50 Mga resulted in a structural alteration affecting its DNA-binding activity, leading to the observed reduction in gene activation.

Electrophoretic mobility shift assays (EMSAs) were performed to investigate the ability of purified M4 and M50 Mga proteins to bind to DNA targets *in vitro*. Plasmids containing *malE-mga* fusion alleles encoding MBP-Mga4 (pKSM155) and MBP-Mga50 (pKSM156) were generated and Mga was purified from *E. coli* lysates containing these plasmids using an amylose affinity resin column. Comparable amounts of each 103 kDa protein were assessed for purity by SDS-PAGE followed by staining for total protein using Sypro Ruby and specific detection of the fusion protein using western analysis with

αMBP antisera (Figure 4A). Increasing amounts of the purified MBP-Mga proteins were incubated with a constant amount of radiolabeled promoter probes corresponding to two Mga-regulated promoters from the serotype M4 strain AP4. Because the *mga* operon of the AP4 strain does not contain an *emm* gene, the two native promoters P*arp* (Figure 4B) and P*mrp* (Figure 4C) were chosen for use as probes. EMSA reactions containing 5 μg of either M4 MBP-Mga or M50 MBP-Mga showed reduced mobility of each probe, indicative of a protein-DNA interaction when compared to 5 μg of the purified MBP control protein alone (Figure 4B and C). Thus, the activation-defective M50 Mga exhibited the ability to bind to DNA from the promoter regions of the Mga-regulated genes *mrp* and *arp,* although the bands resulting from M50 MBP-Mga bound to each promoter probe were more diffuse and required more protein for a definitive shift than observed for M4 MBP-Mga bound to the same targets (Figure 4B and C).

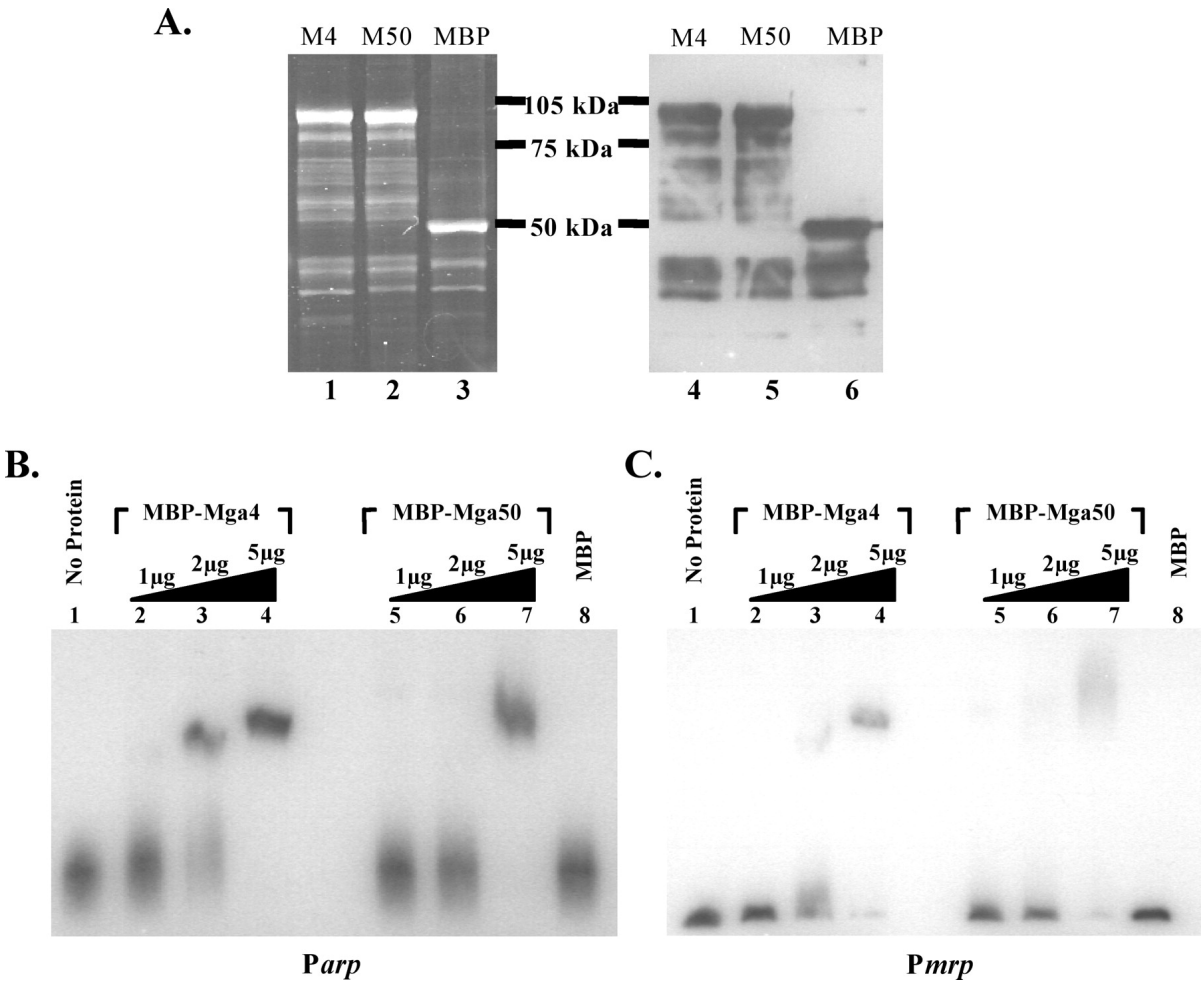


Figure 4: Electrophoretic mobility shift analysis of M4 and M50 MBP-Mga binding to class II Mga-regulated promoters

A) Purification of M4 and M50 MBP-Mga fusion proteins from *E. coli* lysates using an amylose affinity resin. Purified protein was assessed by SDS-PAGE stained with Sypro Ruby (left) and western blot probed with αMBP antisera (right). Purified M4 MBP-Mga (lanes 1 and 4), M50 MBP-Mga (lanes 2 and 5), and MBP control (lanes 3 and 6) are shown. Electrophoretic mobility shift assays (EMSA) were performed on two M4 Mga-regulated promoters, **B)** P*arp* and **C)** P*mrp*. Identical amounts of each radiolabeled promoter probe were incubated for 15 min at 16°C with an increasing amount (1, 2, and 5 µg) of either the purified M4 MBP-Mga (lanes 2-4) or M50 MBP-Mga (lanes 5-7) fusion protein before being separated on a 5% polyacrylamide gel. Binding was also assessed in the presence of no protein (lane 1) or 5 µg MBP alone (lane 8) for each.

B. Establishment of P*rpsL* as a constitutive promoter in the GAS.

In order to remove any contribution that the promoter may have on overall transcript levels as a result of autoregulation by Mga, each allele was placed under the control of a constitutive promoter. Although the *spac* promoter (P*spac*) has been used in other studies for low-level constitutive production of M6 Mga in the GAS (59, 151, 152), expression of M4 and M50 Mga proteins from P*spac* did not produce a detectable level of protein, nor did it show activation of a Mga-regulated GusA reporter to a level significantly above background (Figure 5). Therefore, the promoter for the ribosomal protein S12 gene *rpsL* (P*rpsL*) from the GAS was tested as a constitutive promoter. Expression of *mga4* from P*rpsL* resulted in Mga-regulated GusA expression levels significantly above background and produced a detectable amount of Mga-His, albeit at a lower level than *mga4* expressed from its native promoter P*mga* (Figure 5). Thus, different *mga* alleles were cloned under the control of P*rpsL* for low-level constitutive expression in the GAS (Table 6).

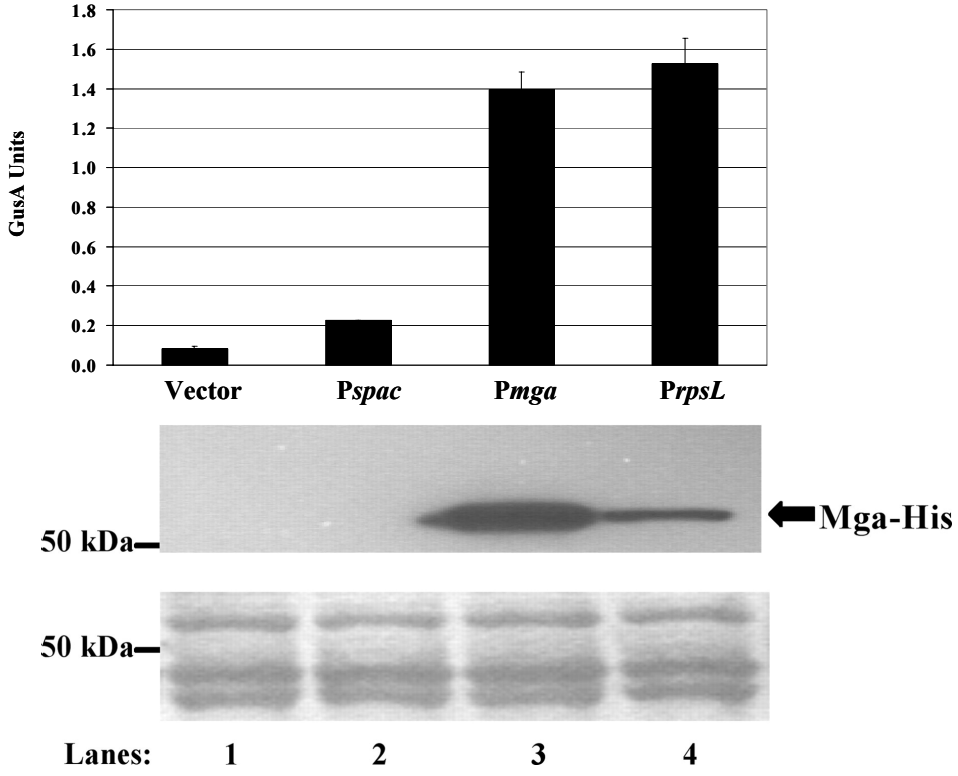


Figure 5: Steady-state protein levels of Mga produced from both the native and various constitutive promoters

GusA activity of whole-cell lysates (top). Production of β-glucoronidase activity was determined for lysates from an *mga*-deleted P*emm-gusA* reporter strain KSM148.174 containing plasmids expressing either vector only (lane 1) or the *mga4* allele from the following promoters: P*spac* (lane 2), P*mga4* (lane 3), and P*rpsL* (lane 4). GusA units are a measure of absorbance (A_{420})/protein concentration (μg/mL). Western analysis (middle) was performed on whole-cell lysates using a αHis antibody for the detection of Mga-His protein in samples. Amido black stain (bottom) of total protein on membranes was used as a loading control. Blots are a representation of data from two independent experiments.

C. Low-level constitutive expression demonstrates a functional importance of residue 461.

To determine the contribution of the three amino acid differences found in the M50 Mga (Figure 3; S26, P361, and R461) on its reduced transcriptional activation *in vivo*, each residue was mutated to the corresponding residue found in M4 Mga to test for a 'gain-of-function' phenotype. Plasmids were constructed which contained *mga4, mga50*, or *mga50* in which either one or two of the targeted amino acids in M50 Mga was changed to resemble M4 Mga (Table 6). Since antibodies recognizing M4 and M50 Mga are not available, each construct also contains a carboxy-terminal fusion to 6X-His to allow for the detection of protein by western blots using αHis monoclonal antibodies (Novagen). Transformation of the plasmids into a class I *mga*-deleted GAS reporter strain (KSM149) containing *gusA* fused to the promoter of a native Mga-regulated gene *mrp* from the M4 strain AP4 (P*mrp-gusA*) in single copy (Table 4) allowed for direct quantitation of Mga-regulated activity by measuring the level of GusA activity in cell lysates (see Materials and Methods).

Mga-regulated GusA activity and steady-state levels of Mga-His were assessed in the P*mrp-gusA* reporter strain KSM149 expressing the different P*rpsL-mga-his* plasmids (Figure 6A). Expression of M50 Mga from P*rpsL* in the reporter strain produced reduced activity (ca. 67%) compared to the M4 Mga, but well above the background seen for vector alone (Figure 6A). This finding agrees with previous studies comparing M50 Mga activity to that for M4 Mga (242). Two of the mutated *mga50* alleles producing amino

acid changes S26N (pKSM326) and P361A (pKSM327) did not substantially increase the level of GusA activity above that seen for the native M50 Mga (pKSM325), indicating that these mutations did not improve the ability of M50 Mga to activate P*mrp-gusA*. In fact, the P361A mutation actually resulted in a more defective protein (Figure 6A). In contrast, GusA levels similar to those observed for the M4 Mga were seen for the strain containing both R461M and the S26N/R461M double mutation (Figure 6A). The levels of Mga-His protein detected in each strain were not equivalent compared to the loading control Hsp60 (Figure 6A). These data suggest that the amino acid change at residue 461 is important for the reduced activity of M50 Mga compared to M4 Mga and may affect its steady-state levels in the cell when expressed from the constitutive promoter P*rpsL*.

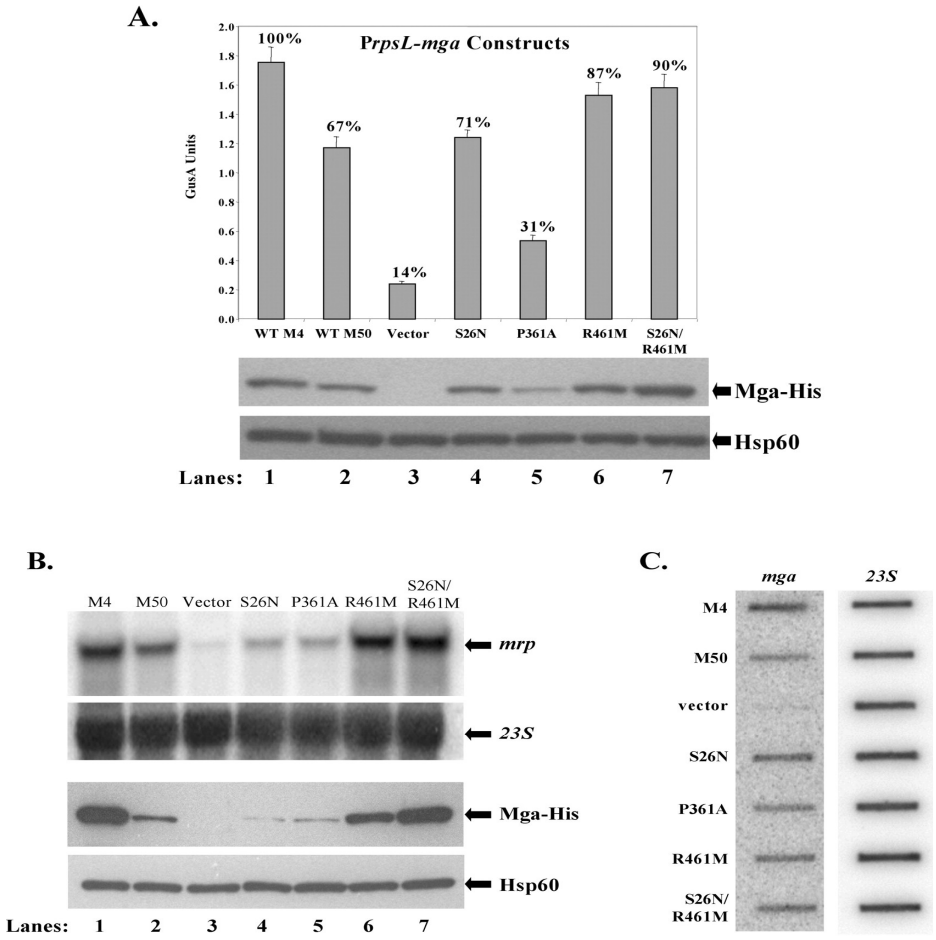


Figure 6: ***In vivo*** **transcriptional activity of mutant M50** ***mga*** **alleles expressed from a constitutive P*rpsL* promoter.**

A) GusA activity of whole-cell lysates (top). Production of β-glucoronidase activity was determined for lysates from an *mga*-deleted P*mrp-gusA* reporter strain KSM149 containing plasmids expressing the following *mga* alleles from the P*rpsL* promoter: M4 Mga (lane 1), M50 Mga (lane 2), vector only (lane 3) and M50 Mga mutants S26N (lane 4), P361A (lane 5), R461M (lane 6), and S26N/R461M (lane 7). GusA units are a measure of absorbance (A_{420})/protein concentration (μg/mL). Percent activity compared to M4 Mga is indicated above each bar. Western analysis was performed on whole-cell lysates using both a αHis antibody for Mga-His protein levels (middle) and with αHsp60 antibodies as a control for loading (bottom). **B)** Northern analysis of Mga-specific transcriptional activation. Transcript levels for the Mga-regulated gene *mrp* was determined using total RNA (5 μg) isolated from an *mga*-inactivated M22 strain AL168-*mga* containing plasmids expressing the following Mga alleles from the P*rpsL* promoter: M4 Mga (lane 1), M50 Mga (lane 2), vector only (lane 3) and M50 Mga mutants S26N (lane 4), P361A (lane 5), R461M (lane 6), and S26N/R461M (lane 7). Blots were stripped and reprobed with 23S rRNA to serve as a loading control (directly below). Western analysis was performed on whole-cell lysates using both a αHis antibody (third panel) and αHsp60 antibody (bottom). **C)** Slot blot analysis of *mga* transcripts produced from the constitutive P*rpsL* promoter. Total RNA (1μg) was isolated from the *mga*-inactivated M22 strains above and probed for M50 *mga* (column A) and *23S rRNA* as a loading control (column B).

To investigate the effect of the different M50 Mga mutants on transcriptional activation in a class II background, the P*rpsL-mga-his* plasmids (Table 6) were introduced into the *mga*-inactivated M22 strain AL168-*mga* (223) (Table 4). Northern blot analysis was performed on total RNA isolated from each strain using a radiolabeled probe to the Mga-regulated gene *mrp22*. As observed above, the strain expressing M50 Mga was reduced in activation of *mrp22* compared to wild-type M4 Mga. Furthermore, *mrp22* transcript levels in only those strains containing the R461M or the S26N/R461M mutations were comparable to the levels found with M4 Mga (Figure 6B). In this system, both the S26N and P361A mutations appeared to further reduce activation of *mrp22* compared to wild-type M50 Mga (Figure 6B). Again, the steady-state levels of Mga-His detected in each strain directly correlated to the observed transcript levels for *mrp22*; although this result was more apparent in the M22 strain AL168-*mga* (Figure 6B). This variation in levels of Mga-His protein was not reflected at the level of transcription, since slot blot analysis of *mga*-specific transcripts in the samples exhibited equivalent expression in contrast to the variable protein levels detected (Figure 6C). Taken together, changing only amino acid 461 from arginine to methionine in the defective M50 Mga resulted in a restoration of activity comparable to M4 Mga. Furthermore, the varying level of Mga-His protein produced amongst the mutant alleles, all of which were expressed from the low level constitutive P*rpsL*, may reflect a destabilized conformation when certain residues are present.

D. Expression from the native promoter demonstrates that both residues 26 and 461 are important for transcriptional activation.

To determine whether the varying protein levels observed for the mutants expressed from P*rpsL* reflected *in vivo* expression levels, each allele was placed under the control of the native P*mga4* found upstream of the fully functional M4 *mga*. The resulting plasmids (Table 6) were transformed into the *mga*-deleted P*mrp-gusA* reporter strain KSM149, and the level of Mga-regulated GusA activity was measured. In contrast to the P*rpsL* results (Figure 6A), levels of Mga-His protein resulting from P*mga4* were comparable across strains, excluding the vector alone negative control (Figure 7A and B). Similar to previous results, P*mrp-gusA* expression was reduced in the strain expressing M50 Mga to approximately 52% of the levels seen for M4 Mga strain (Figure 7A). However, only the amino acid change P361A (pKSM336) failed to restore Mga-regulated GusA levels to those seen in M4 Mga (pKSM322). As before, the P361A mutation appeared to further reduce M50 Mga activity. Under P*mga4*, mutations in both R461M (pKSM330) and S26N (pKSM323) resulted in GusA levels comparable to M4 Mga (Figure 7A), indicating that both positions are involved in transcriptional activation of Mga-regulated genes when expressed from P*mga4*. In addition, the S26N/R461M double mutant had a slightly increased activation of P*mrp-gusA* over either S26N or R461M mutants alone.

The results of the P*mga4* studies in the GusA reporter strain were confirmed via Northern analysis of transcript levels of *mrp22* in the *mga*-inactivated M22 strain AL168-

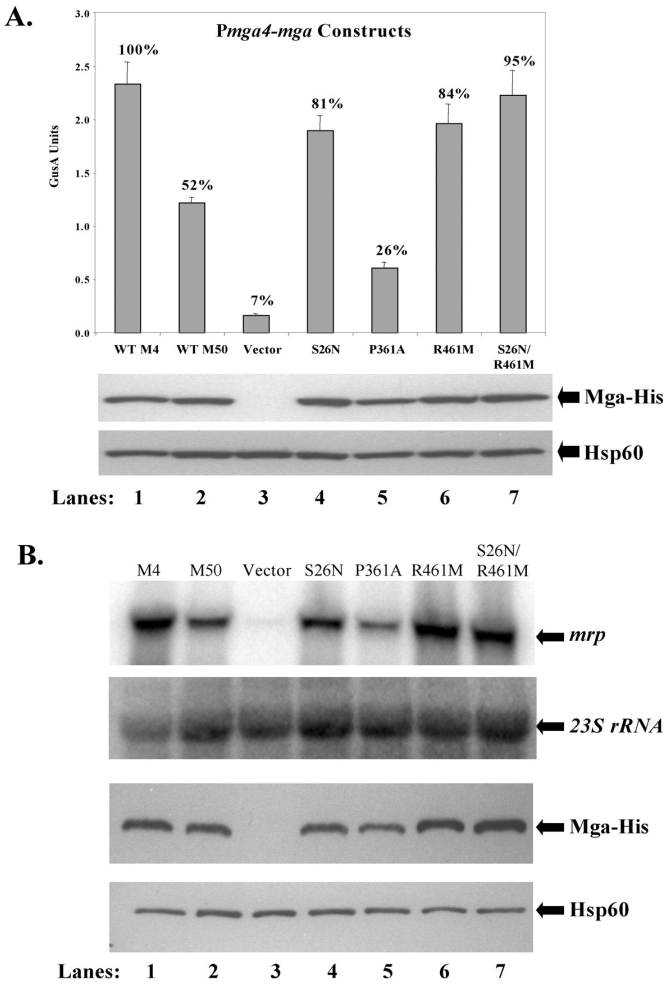


Figure 7: ***In vivo*** **transcriptional activity of mutant M50** ***mga*** **alleles expressed from a native** ***mga4*** **promoter.**

A) GusA activity of whole-cell lysates (top). Production of β-glucoronidase activity was determined for lysates from an *mga*-deleted P*mrp-gusA* reporter strain KSM149 containing plasmids expressing the following Mga alleles from the P*mga4* promoter: M4 Mga (lane 1), M50 Mga (lane 2), vector only (lane 3) and M50 Mga mutants S26N (lane 4), P361A (lane 5), R461M (lane 6), and S26N/R461M (lane 7). GusA units represent a measure of absorbance (A_{420})/protein concentration (μg/mL) and is an average of three independent experiments. Percent activity compared to M4 Mga is indicated above each bar. Western analysis was performed on whole-cell lysates using both an αHis antibody for Mga-His protein levels (middle) and with antibodies to Hsp60 as a control for loading (bottom). **B)** Northern analysis of Mga-specific transcriptional activation. Transcript levels for the Mga-regulated *mrp* was determined using total RNA (5 μg) isolated from an *mga*-inactivated M22 strain AL168-*mga* containing plasmids expressing the following *mga* alleles from the P*mga4* promoter: M4 Mga (lane 1), M50 Mga (lane 2), vector only (lane 3) and M50 Mga mutants S26N (lane 4), P361A (lane 5), R461M (lane 6), and S26N/R461M (lane 7). Blots were stripped and reprobed with 23S rRNA to serve as a loading control (directly below). Western analysis was performed on whole-cell lysates using both a αHis (third panel) and αHsp60 antibodies (bottom). All blots shown are representative of data from thee independent experiments.

mga. Transcript levels of *mrp22* in each strain exhibited the same profile seen in the GusA reporter (Figure 7B), with the R461M mutant and the S26N/R461M double mutant producing *mrp* transcript levels similar to those seen for strains expressing M4 Mga (Figure 7B). The S26N mutant did show a positive effect on M50 Mga activity but to a somewhat lesser degree than R461M. Once again, protein levels produced from P*mga4* were consistent from strain to strain. Overall, these data indicate that both residues 461 and 26 are responsible for the reduced transcriptional activation of M50 Mga when expressed from the native P*mga4*. Furthermore, they suggest that the high expression level from P*mga4* may be able to mask the instability of several M50 mutant proteins. However, this is not the case when expression is from the lower level *rpsL* promoter.

E. Residue 461 is not important for activity in a divergent serotype M6 Mga.

Upon alignment of primary amino acid sequences for Mga from multiple GAS strains, it was noted that the methionine residue at position 461, but not the serine at residue 26, was conserved amongst all Mga proteins, with the exception of M50 Mga (Figure 3). To test whether the conserved M461 was also important for full activation and stability in Mga expressed from an *mga-1* allele, a site-specific mutation leading to a M461R change was made in the *mga-1* allele from the serotype M6 strain JRS4 under its native promoter P*mga6* (pKSM337). Since *mrp* is not a native target in the M6 strain, the resulting M461R mutant plasmid (Table 6), a wild-type P*mga6-mga6* (pKSM164) and a vector control (pLZ12-Spc), were transformed into the *mga*-deleted P*emm6-gusA* reporter strain KSM148.174 (Table 4) to assess activation of a single copy Mga-regulated

promoter (P*emm6*) in its native M6 background. No significant differences in Mga-regulated GusA levels were seen between the wild-type M6 Mga protein and the M461R Mga (Figure 8A). To ensure that the activation defect observed in M50 Mga was not restricted to certain Mga-regulated genes, the plasmids were also transformed into the P*mrp-gusA* reporter strain KSM149 and analyzed for GusA activity (Figure 8B). Alteration of M461R in M6 Mga again did not significantly change its ability to activate expression of Mga-regulated P*mrp* (Figure 8B). Western analysis of all samples showed steady-state levels of Mga were comparable for each strain, indicating the M461R mutation does not appear to play a significant role in the stability of M6 Mga. Thus, the conserved M461 residue does not appear to be important in a M6 Mga, which is produced from an *mga-1* allele, regardless of the downstream target gene.

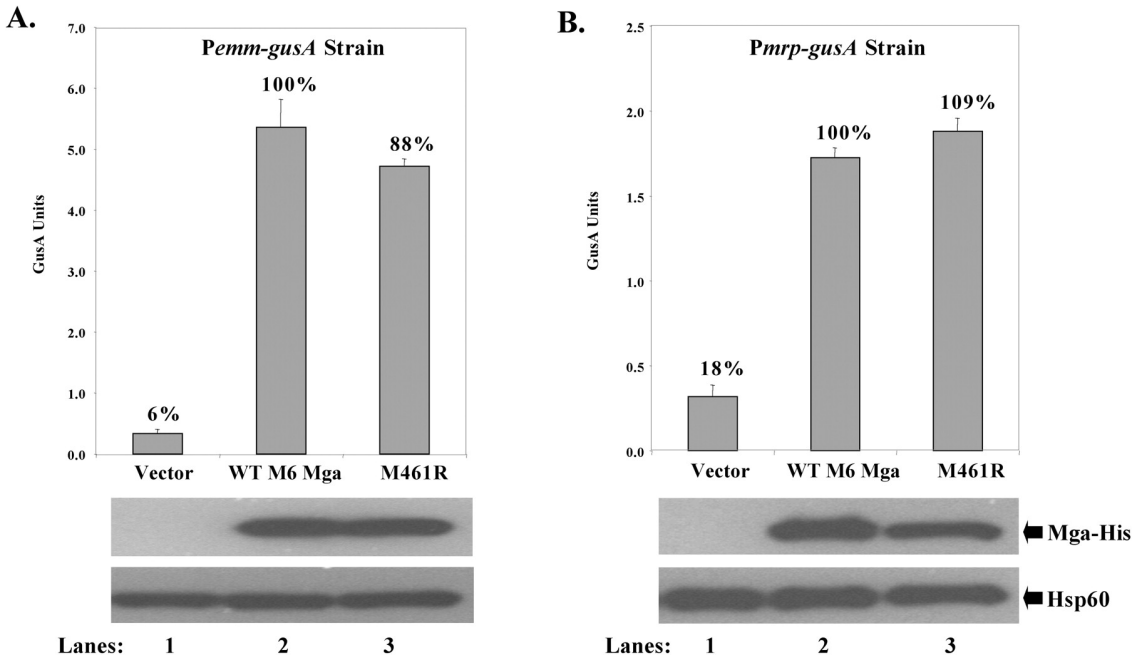


Figure 8: *In vivo* transcriptional activity of a mutant M6 *mga* allele at different promoters.

GusA activity of whole-cell lysates are shown (top). Production of β-glucoronidase activity was determined for lysates from either **A)** an *mga*-deleted P*emm-gusA* reporter strain KSM148.174 or **B)** an *mga*-deleted P*mrp-gusA* reporter strain KSM149. Both strains contained plasmids expressing wild-type M6 Mga (lane 1), mutant M461R M6 Mga (lane 2) and vector alone (lane 3) from the native P*mga6* promoter. GusA units represents a measure of absorbance (A_{420})/protein concentration (μg/mL) and is the average of at least three independent experiments. The percent activity compared to M6 Mga is indicated above each bar. Western analysis was performed on whole-cell lysates using both an αHis antibody for Mga-His protein levels (middle) and with antibodies to Hsp60 as a control for loading (bottom). Blots shown are representative of data from three independent experiments.

III. Discussion

A. Significance of the naturally occurring serotype M50 Mga.

Initial characterization of M50 B514-Sm revealed an Mga regulatory protein defective in activation of its Mga-regulated virulence genes (*mrp* and *arp*) even though it was 98% identical to other fully functional Mga proteins expressed in other similar class II serotypes such as M4 and M49 possessing an *mga-2* allele (242). Yung *et al.* previously proposed that the activation defect of M50 Mga might result from an inability to bind properly to target promoter DNA based on the observation that one of the three amino acids differences in M50 Mga (S26N) was located within a putative HTH DNA-binding motif (242). However, the DNA-binding domains have since been shown to be located elsewhere (151) and do not overlap any of the amino acids in question. Furthermore, our results indicate that M50 Mga does retain its ability to bind to DNA (Figure 4). Thus, determination of why the naturally occurring M50 Mga mutant is defective could provide insight into the relationship between the primary amino acid sequence of Mga and novel aspects of its ability to function as a transcriptional activator.

B. Amino acids important for full activity of Mga proteins.

To determine which of the three amino acid changes alone or in combination affected the ability of Mga to activate target gene transcription, 'gain-of-function' mutants were introduced in M50 Mga by changing amino acids to their corresponding residues in the functional M4 Mga. All *in vivo* analyses of the proteins, from the transcriptional

reporter fusion in the serotype M6 to the Northern analysis in the M22 serotype, was performed in heterologous systems to rule out the possibility that a defect other than the M50 Mga protein alone was the cause of the loss of transcriptional activity seen in the B514-Sm M50 strain.

Transcription levels resulting from a change in Mga at position 26 from the serine to an asparagine varied based upon the promoter used to express the mutant *mga*. When the *rpsL* promoter was used, transcription remained at a level equal to that of M50 Mga and protein levels were variable. In contrast, when the same *mga* allele was transcribed from its native promoter, transcriptional activation was elevated and protein levels did not vary. This fluctuation of steady-state protein levels under the lower level constitutive *rpsL* promoter could be an end result of a less stable protein conformation leading to a more rapid degradation of the protein in the cell. Since a change at position 26 restored full activation under the native *mga4* promoter, where protein levels are higher, one can assume that in the native bacteria residue 26 plays a more significant role in producing a functional Mga. This also stresses the importance of P*mga* for proper Mga production in a virulent GAS cell.

It was suggested by Yung *et al.* that the second amino acid difference, which changed residue 361 from a flexible alanine to a constrained proline, could alter the protein conformation of M50 Mga and possibly disrupt its activity (242). However, altering residue 361 to an alanine never exhibited a gain-of-function phenotype. Instead, the outcome of a change at residue 361 was independent of the specific amino acid at this

position. Interestingly, the most detrimental effect was seen with an alanine, not a proline at position 361 in the context of the M50 Mga (i.e., P361A M50 mutant). This mutation resulted in a decrease of transcriptional activation below that of the already deficient native M50 Mga to only slightly above the level seen for background vector alone. These data suggest that this area may be acting in concert with the other parts of the protein because the level of activation can fall depending upon the context in which it is placed and not solely on the amino acid residue at this exact location.

This is not the case for the third mutation at position 461. Changing residue 461 from an arginine to a methionine in M50 Mga always restored the activity to wild-type levels despite the promoter or context in which the mutant was placed. Therefore, this position must play an essential role in the activation of *mrp* and possibly other Mga-regulated genes.

C. R461 plays an important role in the conformation of Mga.

Several lines of evidence point to a methionine at position 461 as an amino acid important for correct conformation and thus overall stability of Mga. The first piece of data comes from a comparison of the R461M mutation under the constitutive P*rpsL* promoter. When the level of transcript is constitutive although at a lower level than seen for the native P*mga* (Figure 5) and M50 Mga is mutated to contain a methionine at position 461, an increase in overall Mga-His protein levels is observed compared to all other constructs containing an arginine at this position. However, the level of *mga* transcription did not increase amongst these constructs (Figure 6). This finding would

imply that a methionine at position 461 is essential for correct conformation and that the substitution of an arginine leads to a protein more sensitive to degradation. The *mga* promoter is able to mask the effects of degradation (Figure 7) possibly because P*mga* produces more protein than P*rpsL* (Figure 5). This production could possibly conceal the low level degradation apparent only with P*rpsL*.

The second line of evidence stems from the EMSA assays performed in this study comparing wild-type M50 and M4 MBP-Mga. These data clearly demonstrate that both proteins bind to DNA from the promoter regions of relevant Mga-regulated genes (Figure 4). Thus, it was inferred that the reduction in M50 Mga activity *in vivo* is not due simply to a loss of overall DNA-binding ability. Although both M50 and M4 MBP-Mga each bound to the promoter probes when 5 μg of protein was used, differences in the overall binding pattern were observed. For instance, at lower levels of M50 MBP-Mga protein (2 μg), binding to P*arp* was not detected compared to M4 MBP-Mga (Figure 4B). Also, the bound complex formed by M50 MBP-Mga and promoter DNA appeared more diffuse with both targets than those observed for M4 MBP-Mga, suggesting that differences may exist in how the two Mga proteins interact at regulated promoters. One possibility for this difference could be a slight structural change that does not disrupt the ability to bind but does alter the overall shape of the bound complex and therefore its binding pattern.

Because no higher order structural data exists for the Mga protein, the exact location of amino acid 461 in relation to the rest of the protein cannot be determined. The polarity of the position is altered when the non-polar methionine is changed to a

polar arginine; however, the possible consequences of this change does not appear to be overly dramatic in a region predicted to have less than 25% but greater than 5% solvent accessibility according to JNET (48). In addition, the consensus secondary structure prediction generated from the Jpred server (http://www.compbio.dundee.ac.uk/~www-jpred) found amino acid 461 located within a beta sheet in which case the presence of a methionine over an arginine would not appear to have a significant effect.

D. Differences between divergent Mga proteins.

Because secondary structural predictions did not reveal a compelling need for a methionine at position 461, a primary sequence alignment of all sequenced Mga proteins was performed to determine the conservation of methionine at this position. Surprisingly, even though the Mga proteins diverge the most from one another in the extreme carboxy-terminus, the primary amino acid alignment of all Mga proteins revealed that M461 was 100% conserved with the exception of M50 Mga. Since this was the only amino acid of the three investigated in this study that demonstrated 100% conservation and was shown to be essential for activation in M50 Mga (from an *mga-2* allele), this residue was mutated to an arginine in a divergent M6 strain containing an *mga-1* allele to determine if a loss of activation would be seen. However, this mutation did not show a transcriptional defect at either a native M6 Mga-regulated gene *emm* or the M4/M50 Mga-regulated gene *mrp*. This finding illustrates that the importance of M461 is not universal, contrary to what may be predicted based on its conservation and the fact that the Mga proteins have been shown to be functionally equivalent *in vivo* (7). These data also reinforce the

divergence between the two classes of Mga; implying fundamental differences may also exist between the functional residues used by each class to activate transcription of Mga-regulated virulence genes.

CHAPTER FIVE:

DOMAINS REQUIRED FOR TRANSCRIPTIONAL ACTIVATION SHOW CONSERVATION IN THE MGA FAMILY OF VIRULENCE GENE REGULATORS

I. Introduction

Not only do all serotypes of the GAS possess one of two divergent *mga* alleles (*mga-1* or *mga-2*), but Mga orthologues have also been identified in other pathogenic streptococci besides the GAS (76, 90, 92, 230). Of these orthologues, only Mga from a GAS *mga-1* allele has been extensively characterized and thus serves as the model system for this family of Mga-like virulence regulators. It has been demonstrated previously that a serotype M6 Mga (*mga-1*) activates transcription by binding to sites located within regulated (P*emm,* P*scpA,* P*sclA*) and autoregulated (P*mga*) promoters (4, 149, 154). Two amino-terminal helix-turn-helix (HTH) domains were deemed necessary for this DNA-binding and transcriptional activation (151). Aside from the HTH motifs, little is known about Mga so a genetic screen was devised to uncover functional domains of this protein within a serotype M6 Mga (*mga-1*) important for transcriptional activation of the Mga-regulated *emm* gene. In addition to mutations in the established DNA-binding domains, the screen revealed mutations within a region that was conserved

amongst several Mga orthologues. Alanine scanning was performed across this conserved Mga domain (CMD-1), which spanned from residue 10 to 15 of the protein, to help define its function. Western blotting analysis revealed that the two flanking amino acid residues were likely involved in protein stability while transcriptional reporter analyses demonstrated a role of CMD-1 for activation of P*emm* and autoactivation of P*mga* in the serotype M6 Mga. Mutational analyses showed that both CMD-1 and HTH-4 were also necessary for activation of the promoter target P*mrp* in a divergent serotype M4 Mga (*mga-2*), signifying a conserved functionality for these domains. However, in contrast to M6, the M4 Mga mutants did not show a defect in autoregulation suggesting that differences exist between divergent Mga proteins and their contribution to autoregulation. Mutagenesis of conserved residues in the Mga-like regulator DmgB from *S. dysgalactiae* subsp. *dysgalactiae* showed that CMD-1 and HTH-4 are also critical for transcriptional activation in this orthologue, implying that some common mechanisms of virulence gene activation may also exist among members of the Mga family of regulators.

II. Results

A. Random mutagenesis screen for transcriptionally defective Mga mutants.

A non-biased strategy using PCR-generated mutations was devised to identify functional residues within Mga involved in activation of Mga-dependent genes (Figure 9). PCR amplification using a defective DNA polymerase was utilized to generate

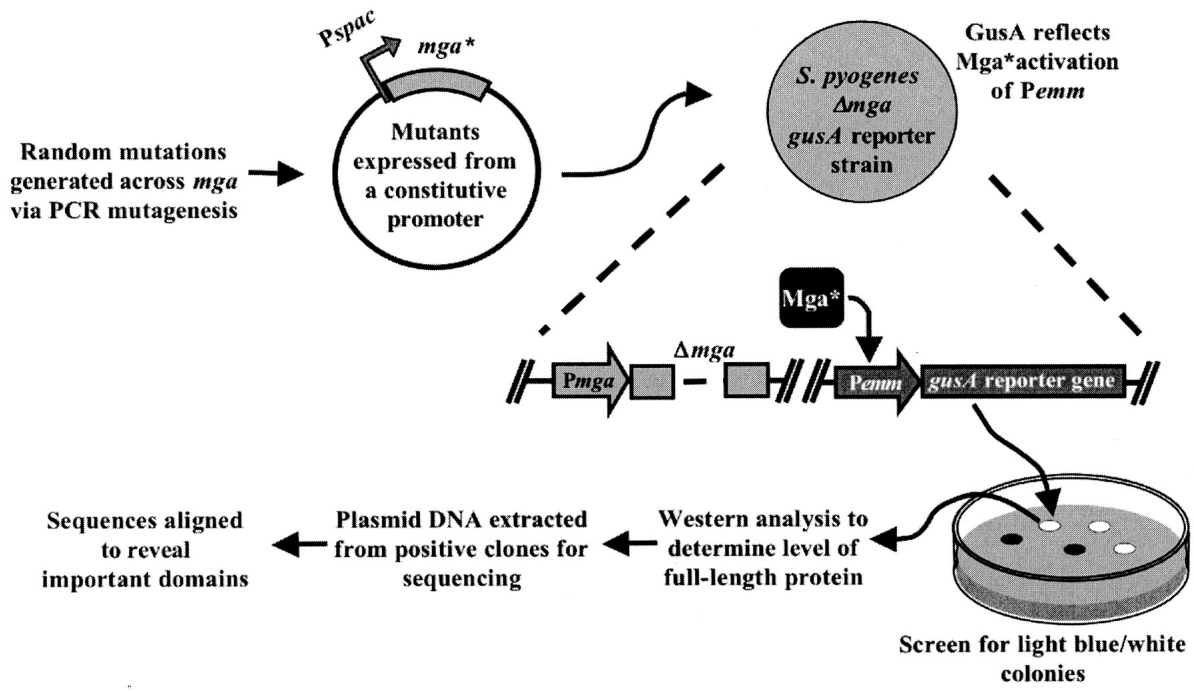


Figure 9: Random PCR mutagenesis screen used to identify domains involved in Mga-dependent transcriptional activation.

Random mutations were generated in *mga* using the GeneMorph PCR Mutagenesis kit (Stratagene). Mutated M6 *mga* alleles were placed on a plasmid under the constitutive promoter, P*spac,* in an *mga*-deleted strain containing a promoter fusion of the Mga-regulated P*emm* to a promoterless *gusA* reporter gene in the chromosome of the M6 GAS. Resulting strains were plated onto THY plates containing X-glu to reveal clones deficient in transcriptional activation (light blue or white). Whole-cell lysates from clones showing a defect in activation were extracted, and Mga protein levels were determined using western blot analysis. Plasmid DNA from clones producing WT levels of protein was isolated, sequenced, and aligned to the wild-type M6 *mga* gene to identify mutations.

random mutations across the 1590 base pair *mga* gene from the serotype M6 strain JRS4 (see Materials and Methods), which was arbitrarily divided into two fragments (N- and C-terminal) at a unique SpeI site for ease of handling. Mutated *mga* fragments were then cloned under the constitutive P*spac* promoter in the pKSM318 (229) such that transcription levels in the screen would not reflect autoactivation from the native P*mga* promoter. The resulting plasmids were transformed into a *mga*-deleted GAS reporter strain (KSM148.174) containing a single-copy transcriptional fusion of *gusA* to the Mga-regulated P*emm*, allowing for direct quantitation of Mga-regulated activity based on a colorimetric assay (see Materials and Methods). Transformed GAS strains were plated onto media containing X-glu, which allowed for visual determination of a defective Mga based on colony color (white or light blue) compared to blue seen with the WT control. Whole-cell lysates were prepared from clones chosen for further analysis and the steady-state level of Mga was determined using western blot analysis. Finally, plasmid DNA from defective clones producing WT levels of protein was sequenced, and mutation sites were recorded. Although identical *mga* mutations were found, possibly as a result of the amplification steps in *E. coli*, they were not included as separate entries in the overall results.

In total, 12 independent mutants were identified within the N-terminal Mga fragment that resulted in either one or two amino acid changes per molecule and produced WT levels of protein (Table 8 gray highlights). Eight mutants were identified that contained mutations within the known DNA-binding domains HTH-3 and HTH-4

(151), resulting in colonies that were either light blue (HTH-3 mutations) or white (HTH-4 mutations) in color. These activities correlated to the effects previously observed for mutations in the two Mga HTH domains (151) and served as a strong validation of the screen. Four additional mutations located outside of the known DNA-binding regions were also identified that exhibited a light blue colony phenotype in the screen. Two pairs of mutants (Q11R, W12R/A38T and K33T, V30I/T139I) involved neighboring residues, suggesting that each might indicate potential functional domains in the N-terminus of Mga.

Table 8: Mutants found from the random PCR mutagenesis screen

Domain	Mutation producing WT protein levels	Phenotype*
DNA-binding domain (HTH-3)	F55C	LB
	I62N	LB
	I71F	LB
DNA-binding domain (HTH-4)	S117T	W
	S117N	W
	S119L	W
	T120K	W
	R123H	W
Other	Q11R	LB
	W12R/A38T	LB
	K33T	LB
	V30I/T139I	LB
C-terminal fragment	None	LB / W

*LB: light blue, W: white

Interestingly, no light blue or white colonies producing WT levels of protein were obtained from repeated screens of the C-terminal fragment of Mga. In fact, 272 C-terminal mutants exhibiting a defect in Mga activity were analyzed via western analysis and found to produce little to no detectable Mga (data not shown). Further analysis of a subset of these mutants found they each contained multiple amino acid changes, regardless of attempts to decrease the mutation frequency. Therefore, the C-terminus of Mga appears to be quite sensitive to mutagenesis and, as a result, was not amenable to the screen used in this study.

B. Two mutations reside within a region conserved amongst Mga homologues.

Conserved regions in protein families often represent important functional domains that are retained during evolution. Therefore, a protein sequence alignment was performed on Mga homologues from various streptococcal species, including Mga proteins representing divergent *mga-1* and *mga-2* alleles in the GAS, Mgc from *S. dysgalactiae* subsp. *equisimilis,* DmgB from *S. dysgalactiae* subsp. *dysgalactiae*, MgrA from *S. pneumoniae* as well as putative Mga-like transcriptional regulators found in the unfinished genomes of *S. equi* (Mge)*, S. gordonii* (Mgg)*,* and *S. mitis* (Mgm) via BLAST analysis (Table 9 and Figure 10). This group of homologous and orthologous proteins is termed the Mga family of virulence gene regulators.

Table 9: Mga family of virulence gene regulators

Gene Name	Organism	GI #	Similarity to M6 Mga[+]	Length (aa)
M6 Mga	*S. pyogenes*	153733	100%	531
M1 Mga	*S. pyogenes*	1675800	97%	529
M4 Mga	*S. pyogenes*	1246852	79%	533
M49 Mga	*S. pyogenes*	56808536	79%	533
Mgc	*S. dysgalactiae* subsp. *equisimilis*	6782393	49%	513
DmgB	*S. dysgalactiae* subsp. *dysgalactiae*	6689248	41%	496
Mge	*S. equi*	N/A *	42%	505
Mgg	*S. gordonii*	N/A**	20%	493
Mgm	*S. mitis*	N/A**	19%	494
MgrA	*S. pneumoniae*	17368568	16%	357

[+]: Sequence has been reanalyzed for similarity using the ClustalW algorithm and thus may not correspond exactly with previously published results. N/A: Not applicable. Sequence is part of the unfinished genome from the * Sanger Institute or **TIGR.

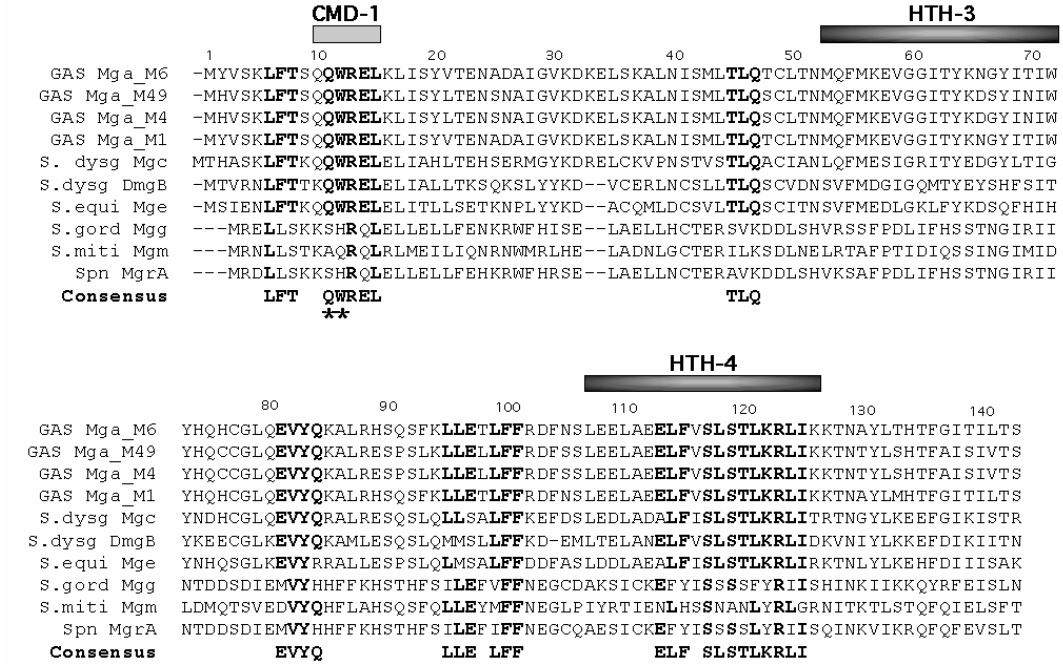


Figure 10: Sequence alignment of Mga orthologues reveals conserved domains.

A sequence alignment of the Mga orthologues from various streptococcal species, including the GAS (M6, M49, M4, and M1), *S. dysgalactiae* (subsp. *equisimilis* and *dysgalactiae,* respectively), *S. equi, S. gordonii, S. mitis* and *S. pneumoniae*, was used to derive a consensus sequence for conserved domains located within the first 143 amino acids of the proteins (see Materials and Methods). A conserved domain was defined as an area containing more than two consecutive residues exhibiting $\geq$ 70% identity amongst homologues. Amino acids identical to the consensus are in bold type. Black bars depict the locations of the two known DNA-binding domains in the GAS M6 Mga while the gray bar denotes CMD-1. The two mutations found outside of the DNA-binding domains during the random mutagenesis screen are indicated (*).

A large region of conservation encompassed the essential HTH-4 domain of Mga, suggesting that this DNA-binding domain is likely serving a similar function in many of the orthologues (Figure 10). Several other groups of conserved residues were also found by the alignment outside of the HTH domains (Figure 10). Two of the four mutants identified in our activity screen (Table 8; K33T, V30I/T139I) did not fall within one of these conserved domains and were not investigated further in this study. However, the remaining two Mga mutants (Table 8; Q11R, W12R/A38T) contained mutations within a region of amino acid conservation (noted as asterisks in Figure 10) spanning from the glutamine at position 11 to the leucine at position 15. The glutamine at position 10 also showed a high degree of conservation, and thus the region encompassing residues 10 to 15 was named conserved Mga domain 1 (CMD-1).

C. CMD-1 of an M6 Mga (*mga-1*) is involved in transcriptional activation.

To establish whether the entire CMD-1 was important for Mga activity, we utilized site-specific mutagenesis to construct mutant *mga* alleles that encoded for a single amino acid change in Mga at residues 10-15 (Q10R, Q11A, W12A, R13A, E14A and L15A). In addition, the original Q11R mutant and a site-specific W12R mutant, which recapitulated the original W12R/A38T double mutant, were used. As in the initial screen, the resulting CMD-1 site-directed mutant alleles were cloned under the constitutive P*spac* in pKSM318 to produce pQ10R, pQ11A, pQ11R, pW12A, pW12R, pR13A, pE14A and pL15A (Table 6). A double Mga mutant K31A/D32A (pKSM318.1H), which has been shown not to affect Mga-dependent activation, and an

established HTH-4 mutation (pKSM318.4H), which leads to a DNA-binding deficiency and loss of Mga-dependent activation, were constructed as controls (151). All of the plasmids were introduced into two *mga*-deleted serotype M6 GAS GusA reporter strains. The first, the Δ*mga* P*emm-gusA* strain KSM148.174 (229), determined Mga-regulated activity at a downstream promoter; while the second, the Δ*mga* P*mga-gusA* strain KSM231.310, assessed the ability to autoregulate.

Mga mutants Q10R, Q11A, Q11R, W12R, R13A, E14A, L15A and the HTH-4 control mutant showed a reduction in activity at the *mga*-regulated promoter P*emm* when compared to WT Mga (Figure 11A). These same mutants also showed a reduction in autoactivation at the native P*mga* with the exception of Q11A, which appeared to have WT activity only in the P*mga* strain (Figure 11B). In contrast, mutations W12A and the negative control mutant K31A/D32A demonstrated WT promoter activity at both P*emm* and P*mga* (Figure 11B). The steady-state levels of Mga detected in the strains expressing CMD-1 mutants in residues 11 to 14, as well as the control strains, were equivalent to WT when compared to the loading control Hsp60 (Figure 11C). However, flanking mutations at either end of CMD-1 (residues 10 and 15) show a reduction in protein level. Taken together, mutations in residues 11 through 14 in CMD-1 appear to affect activation without changing the overall steady-state levels of Mga; whereas, amino acids 10 and 15 likely play a role in protein stabilization.

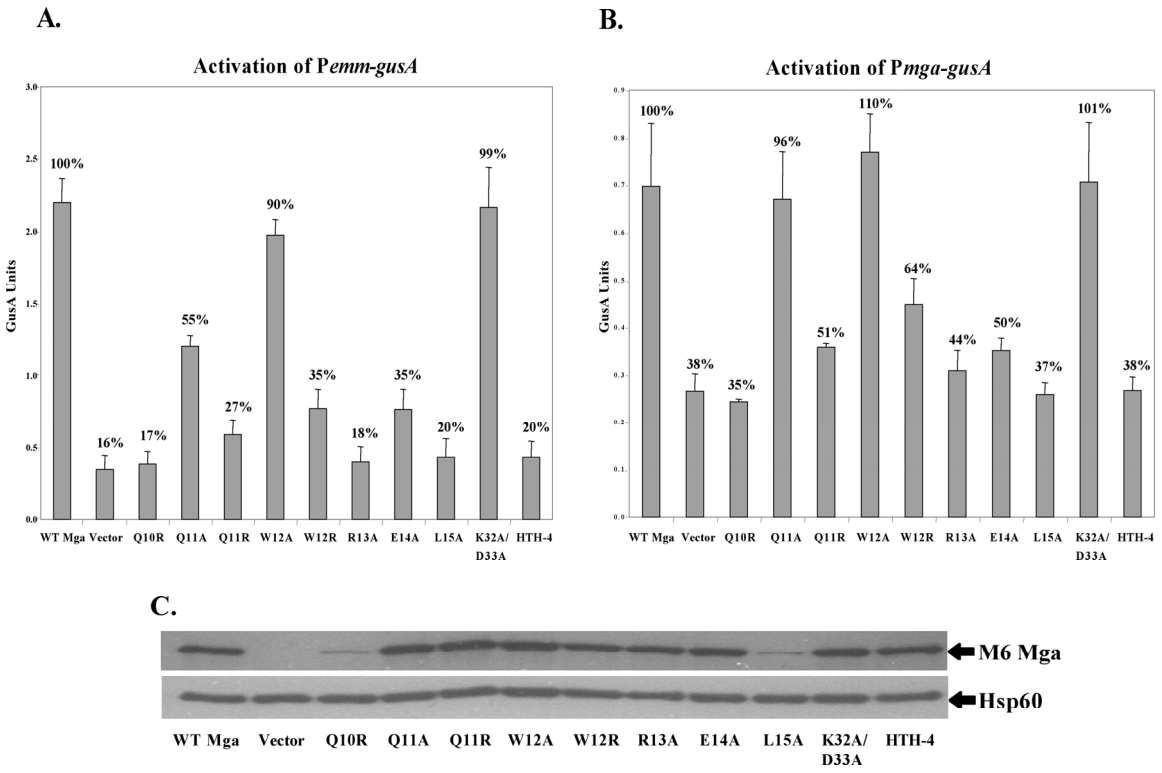


Figure 11: ***In vivo*** **transcriptional activity of mutant M6** ***mga*** **alleles under a constitutive promoter.**

A) GusA activity of whole-cell lysates. Production of β-glucuronidase activity was determined for lysates from a *mga*-deleted P*emm-gusA* reporter strain KSM148.174 or **B)** a *mga*-deleted P*mga-gusA* reporter strain KSM231.310 containing plasmids expressing the following M6 Mga alleles from a constitutive P*spac* promoter: WT M6 Mga, vector only, the M6 Mga mutants Q10R, Q11A, Q11R, W12A, W12R, R13A, E14A, and L15A. An arbitrary mutant K32A/D33A (no defect) and a known DNA-binding mutant in HTH-4 (defective) are included as controls. GusA units represent a measure of absorbance (A_{420})/protein concentration (μg/mL) and is the average of three independent experiments. **C)** Western analysis was performed on whole-cell lysates from the above samples using both an αMga antibody for detection of protein levels (top) and an αHsp60 antibody as a control for loading (bottom).

As an independent verification of the role of CMD-1 on autoregulation at P*mga*, the Q11A, Q11R, and R13A mutants in M6 Mga were produced from the native P*mga* in the Δ*mga* P*emm-gusA* reporter strain KSM148.174 (Figure 12). As predicted by the P*mga* reporter studies (Figure 11B), those mutants showing a defect in P*mga-gusA* activation (Q11R and R13A) did not produce wild-type levels of mutant protein from P*mga*; whereas, Q11A was normal for P*mga* activation and showed levels of protein comparable to the wild-type control (Figure 12). Overall, most mutants in CMD-1 in M6 Mga are defective for both activation and autoregulation; however Q11A is only defective at downstream promoters.

D. CMD-1 and HTH-4 are important for transcriptional activation in a divergent Mga.

Since an *mga-2* allele has been shown to functionally complement an *mga-1* allele *in vivo* (7), we predicted that CMD-1 would play a similar role in a divergent Mga. To investigate the effects of CMD-1 mutations in a divergent Mga, we introduced mutations into the *mga-2* allele from the serotype M4 strain AP4. Given that an antibody against the M4 Mga is not available, all of the alleles were modified to produce a carboxy-terminal 6x-His fusion to allow detection with α-His monoclonal antibodies. Plasmids containing either the WT M4 *mga-his* allele or *mga-his* possessing the Q11R, R13A and HTH-4 mutations under their native P*mga4* promoter were transformed into the P*mrp-gusA* GAS reporter strain KSM149. This strain has been used previously to study transcriptional activation by divergent Mga proteins (229).

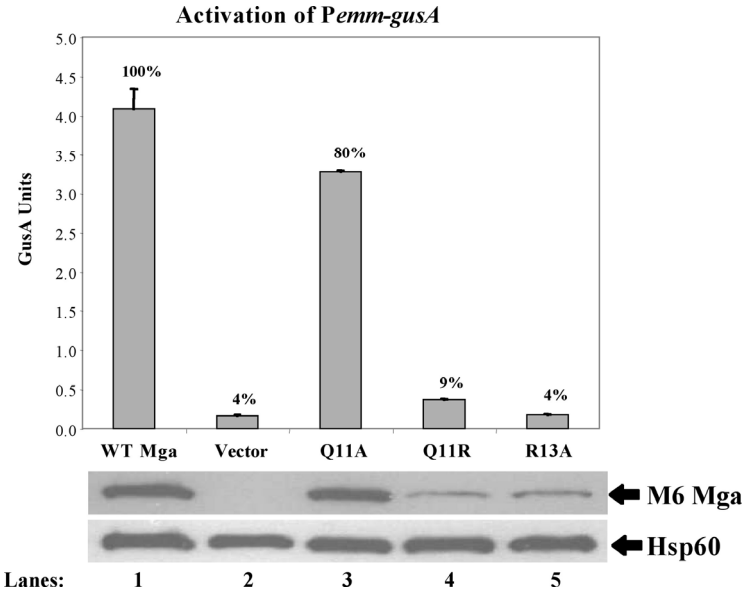


Figure 12: ***In vivo*** **transcriptional activity of mutant M6** ***mga*** **alleles expressed from the native promoter.**

Activity of mutants expressed from the native promoter. Activity levels for both WT and the M6 Mga mutants Q11A, Q11R and R13A were examined for each allele when expressed from the native *mga* promoter using the aforementioned Δ*mga* P*emm-gusA* reporter strain (top). Protein production was also determined via western analysis (bottom) as described above. Percentages indicate level of GusA activity compared to wild-type control.

GusA analysis demonstrated that the HTH-4 DNA-binding domain, which shows 100% identity within all Mga proteins from the GAS, is necessary for full transcriptional activation of P*mrp* by a divergent Mga. Furthermore, a decrease in activation at P*mrp* was also observed when the M4 Mga was mutated at either amino acids 11 or 13, implying that CMD-1 is important for Mga-specific activation in a divergent Mga as well (Figure 13). Interestingly, all of the mutant alleles produced steady-state levels of Mga-His from P*mga4* equivalent to WT protein levels (Figure 13), suggesting that the R13A

and HTH-4 mutations are not defective for autoregulation as observed for the divergent M6 Mga (Figure 11B).

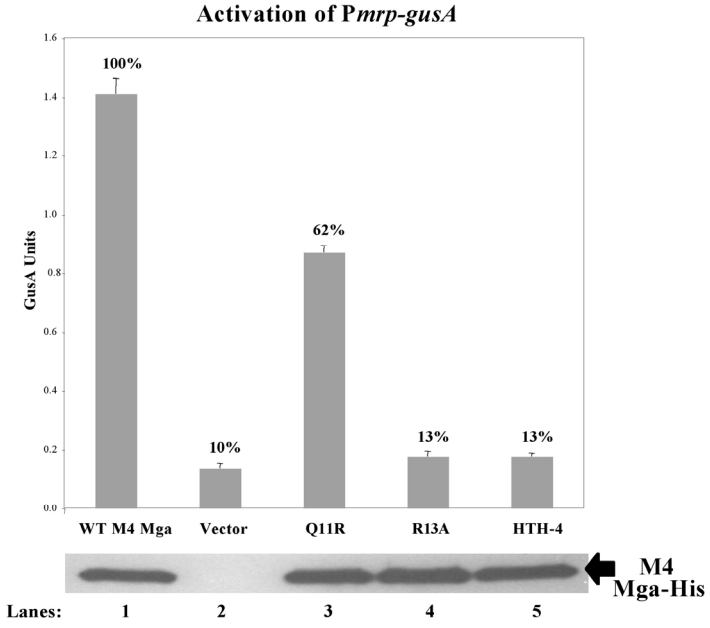


Figure 13: *In vivo* transcriptional activity of mutant M4 *mga* alleles expressed from the native P*mrp* promoter.

GusA activity of whole-cell lysates (top). Production of β-glucoronidase activity was determined for lysates from a *mga*-deleted P*mrp-gusA* reporter strain KSM149 containing plasmids expressing the following different M4 *mga* alleles from the native P*mga* promoter: WT M4 Mga (lane 1), vector only (lane 2), the M4 Mga mutants Q11R (lane 3) and R13A (lane 4), and a DNA-binding mutant in HTH-4 (lane 5). GusA units are a measure of the absorbance (A_{420})/protein concentration (μg/mL) and are the average of three independent experiments. Western analysis was performed on whole-cell lysates of each sample using a αHis antibody for detection of Mga-His protein levels (bottom). Percentages indicate level of GusA activity compared to wild-type control.

E. CMD-1 and HTH-4 are important for transcriptional activation in the Mga orthologue DmgB.

Since the CMD-1 and HTH-4 domains are important in both *mga* alleles from the GAS, we investigated whether these domains were also necessary for transcriptional activation in an Mga orthologue from another pathogenic streptococcus. A plasmid containing both the Mga-like regulator gene *dmgB* and the linked DmgB-regulated gene *demB* from *S. dysgalactiae* subsp. *dysgalactiae* (230) under the native P*dmgB* promoter were introduced into the Δ*mga* M6 GAS strain JRS519 (152). Northern analysis was then used to monitor DmgB-mediated activation of *demB* in the Δ*mga* GAS background (Figure 14). Expression of WT *dmgB* leads to high levels of *demB* transcripts compared to vector alone (Figure 14: lanes 1 and 2). Introduction of the HTH-4 mutation or the CMD-1 mutations Q11R and R13A into DmgB resulted in a dramatic reduction in *demB* transcript levels compared to the WT allele (Figure 14). As a control, mutation of an arbitrary amino-terminal residue (L42A) of DmgB had little effect on its ability to regulate *demB* expression (Figure 14: lane 5). Overall, both CMD-1 and HTH-4 are necessary for activation of corresponding virulence genes, not only in Mga, but also in other members of the Mga family of virulence regulators.

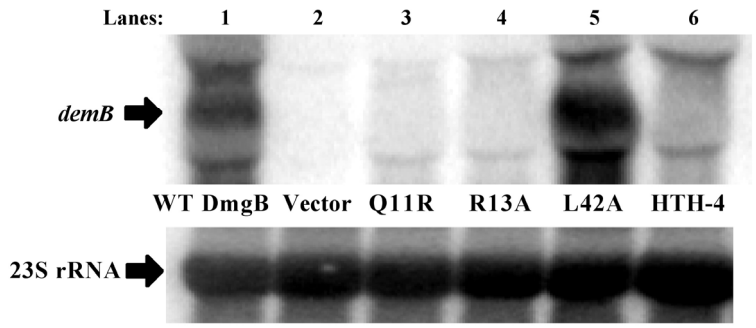


Figure 14: Transcriptional activation of *demB* by mutant Mga orthologue DmgB from *S. dysgalactiae* subsp. *dysgalactiae*.

Northern analysis of transcript levels of the DmgB-regulated gene *demB* was determined using total RNA (5 µg) isolated from the GAS *mga*-inactivated M6 strain JRS519 harboring plasmids containing *demB* and the following *dmgB* alleles under the native *dmgB* promoter: WT *dmgB* (lane 1), vector only (lane 2) and the *dmgB* mutants Q11R (lane 3), R13A (lane 4), an arbitrary mutant L42A (lane 5), and an HTH-4 DNA-binding domain mutant (lane 6). Blots were stripped and reprobed with 23S rRNA to serve as a loading control (directly below). Blot shown is representative of data from three independent experiments.

F. DNA-binding activity of CMD-1 Mga mutants.

Binding of Mga to promoter targets is essential for transcriptional activation of Mga-regulated genes in the GAS (151). To elucidate the contribution of CMD-1 to DNA-binding, electrophoretic mobility shift assays (EMSAs) were performed to determine the ability of Mga mutants to bind to Mga-regulated DNA targets *in vitro*. Plasmids containing either WT M6 *mga-his* (pKSM162) or the mutant M6 *mga-his* alleles: Q11R (pKSM344), W12R (pKSM354), R13A (pKSM345), and E14A (pKSM355) were transformed into *E. coli* and mutant Mga proteins purified from each respective lysate. Increasing amounts of the purified Mga-His proteins were incubated with a constant amount of radiolabeled promoter probe corresponding to the M6 Mga-regulated promoters P*emm* (Figure 15A) and P*mga* (Figure 15B). EMSA reactions

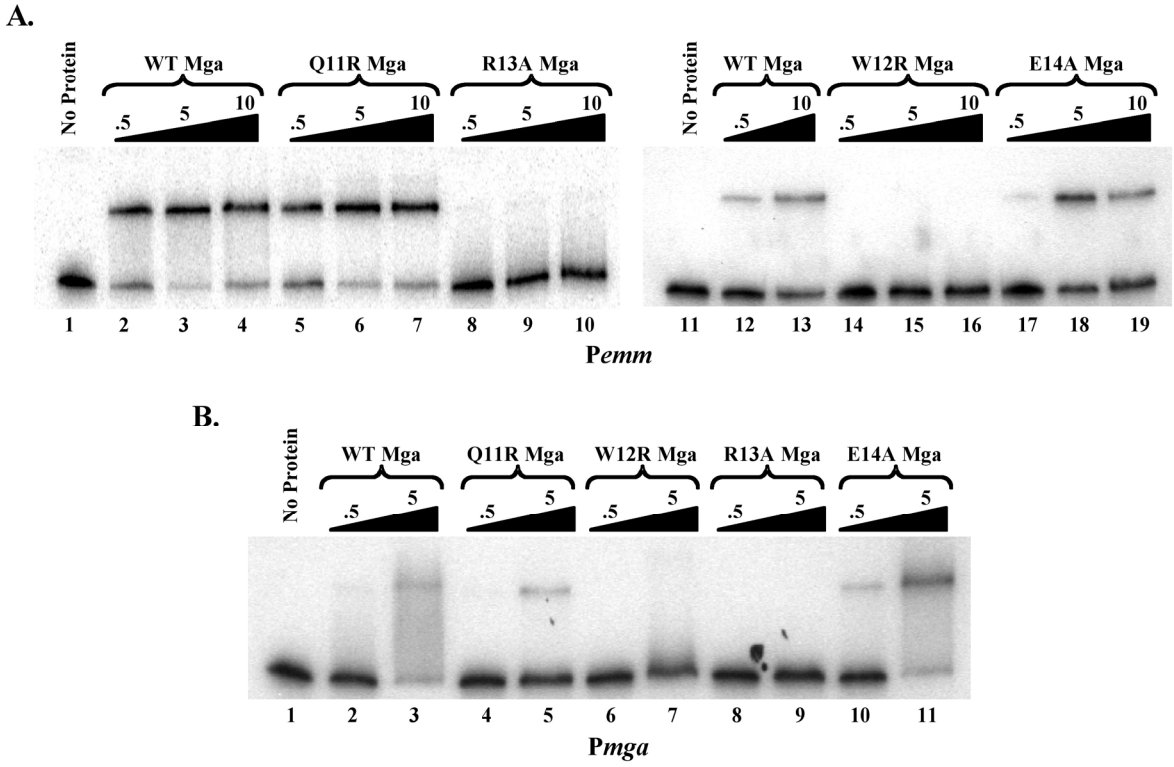


Figure 15: Electrophoretic mobility shift assays of CMD-1 Mga mutants binding to Mga-regulated promoters.

A) Electrophoretic mobility shift assays (EMSAs) of Mga-regulated promoter P*emm*. C-terminal 6X-His fusion proteins (Mga-His) were purified from *E. coli* lysates. Identical amounts of the radiolabeled promoter probe P*emm* were incubated for 15 min at 16°C with an increasing amount (0.5µg, 5 µg and 10 µg) of either the WT (lanes 2-4 and 12-13), Q11R (lanes 5-7), W12R (lanes 8-10), R13A (lanes 14-16), E14A (lanes 17-19), or no M6 Mga-His (lanes 1, 11) prior to separation on a 5% polyacrylamide gel. **B)** EMSAs of the native P*mga* promoter were performed as described above using 0.5µg and 5 µg of the WT (lanes 2-3), Q11R (lanes 4-5), W12R (lanes 6-7), R13A (lanes 8-9), E14A (lanes 10-11) or no Mga-His protein (lane 1).

containing the WT Mga (Figure 15A: lanes 2-4, 12, 13 and Figure 15B: lanes 2-3) and mutant Mga proteins Q11R (Figure 15A: lanes 5-7 and Figure 15B: lanes 4-5) and E14A (Figure 15A: lanes 17-19 and Figure 15B: lanes 10-11) showed reduced mobility of both promoter probes compared to wild type. However, neither W12R (Figure 15A: lanes 14-16 and Figure 15B: lanes: 6-7) nor R13A (Figure 15A: lanes 8-10 and Figure 15B: lanes 8-9) were able to bind at WT levels to either of the two probes. Thus, two of the CMD-1 mutants (W12R, R13A) were defective for binding to DNA targets, whereas two other flanking CMD-1 mutants (Q11R, E14) exhibited normal DNA-binding activity. Therefore, the loss of its DNA-binding ability alone does not explain the defect in transcriptional activation observed with some CMD-1 mutations.

III. Discussion

A. HTH-4 and CMD-1 are functional domains of Mga.

Despite its established role in the pathogenesis of GAS infections, we currently know very little about how Mga functions. Previous studies using a serotype M6 Mga (*mga-1*) identified a minor (HTH-3) and a major (HTH-4) helix-turn-helix DNA-binding domain within the amino-terminus of Mga (151). Importantly, mutations in HTH-4 led to a defect in M6 Mga-dependent activation *in vivo* and established DNA-binding as an essential function of Mga. Finding additional functional domains in Mga would provide insights into its mechanism of action.

To address this issue, we utilized a genetic screen to look for random mutations in the serotype M6 Mga (*mga-1*) that lead to a defect in its ability to activate Mga-regulated gene transcription *in vivo*. The screen identified both the HTH-3 and HTH-4 DNA-binding domains of M6 Mga (Table 8), providing further evidence of their importance and acting as a strong validation of the screen. The analysis also revealed two amino acids (Q11 and W12) involved in M6 Mga activity that were conserved amongst divergent Mga proteins in the GAS and orthologues found in other pathogenic streptococci (Figure 10). Site-directed mutagenesis within the conserved Mga domain 1 (CMD-1; amino acids 10-15) verified the importance of CMD-1 for transcriptional activation of M6 Mga-regulated genes (*emm*) as well as its own autoregulation (*mga*) in this background (Figure 11). Therefore, CMD-1 represents a new Mga functional domain involved in its ability to regulate virulence gene expression in the GAS.

To date, most of the functional studies performed on Mga have been done using the serotype M6 (*mga-1*) allele (151, 154). The previous chapter found that some amino acids important for M4 Mga (*mga-2*) activity did not serve the same function in a divergent M6 Mga (*mga-1*). Thus, it was important to see if functional domains identified in the M6 Mga were also required in a divergent M4 Mga (*mga-2*) allele. Site-directed mutations in both the HTH-4 and CMD-1 domains of the M4 Mga (*mga-2*) resulted in a protein that was unable to activate transcription of the M4 Mga-regulated gene *mrp* (Figure 13). This result demonstrates that both motifs appear to serve a similar

role in divergent Mga alleles and represent a conserved functional domain in all Mga proteins.

B. Functional role for CMD-1 in Mga activity.

Given that CMD-1 is important for Mga-dependent transcriptional activation, exactly how the domain contributes to Mga function is of particular interest. The mutational analysis provided some clues as to the role of specific residues. It was clear from the mutagenesis of CMD-1 that mutations in either of the two amino acids (position 10 and 15) flanking the conserved region showed diminished activation due to low steady-state protein levels (Figure 11C), suggesting that these two residues are likely important in overall Mga stability. Most alanine substitutions within the rest of CMD-1 in M6 Mga resulted in a significant decrease in activity by as much as 83% for P*emm* and 65% for P*mga* without affecting the levels of protein (Figure 11A and B). Similar trends were seen for those CMD-1 mutations tested in the divergent M4 Mga, although differences in the degree of reduction were observed for the Mga Q11R mutation (Figure 13). One M6 Mga mutant (W12A) did not show a decrease in transcription at either promoter; however, when W12 was mutated to an arginine instead of an alanine, a loss of activity was observed. This implies that, at least for this residue in CMD-1, the ability of Mga to activate transcription is contingent upon the particular amino acid at that position, possibly reflecting the charge or polarity of this site.

Because DNA binding is essential for Mga activity (151), CMD-1 M6 Mga mutants were tested for their ability to bind probes corresponding to the Mga-regulated

P*emm* and P*mga* promoters. It was expected that all CMD-1 mutations would either be wild type or defective in their ability to bind to regulated promoter targets. Surprisingly, two of the four transcriptionally defective mutants tested (Q11R and E14A) retained the ability to bind DNA, while the other two mutants (W12R and R13A) did not (Figure 15A and B). The inability of the W12R mutant to bind normally may be dependent upon the amino acid change at that position as discussed above. Since a W12A mutation at this same position shows wild-type Mga activity (Figure 11A and B), it is predicted to retain normal binding activity as well. The other binding mutant R13A occurs at an arginine residue that is 100% conserved amongst Mga orthologues (Figure 10), and we can only speculate what effects different mutations at R13 would have on the ability to bind or activate Mga-regulated promoters. Therefore, at least two residues in CMD-1 are important for DNA-binding in addition to the established HTH-3 and HTH-4 domains.

The ability of two mutants (Q11R and E14A) to bind DNA normally, while still being defective in transcriptional activation *in vivo*, clearly suggests that portions of CMD-1 contribute to Mga activity independent of DNA-binding. Secondary structure predictions (Jpred; http://www.compbio.dundee.ac.uk/~www-jpred) suggest that CMD-1 is part of an alpha helix. Mutants at positions 11 and 14 not only demonstrate a similar phenotype, but they also would reside on the same face of the helix. This aspect of CMD-1 potentially represents a novel function for this regulator, and further study may provide us with new insights into the mechanism of Mga regulation.

C. Autoregulation and divergent Mga proteins.

The M6 Mga (*mga-1*) has been shown to bind directly to its own promoter, resulting in activation of *mga* expression and amplification of the Mga response (75, 151, 154). In this chapter, constitutive expression of M6 Mga CMD-1 and HTH-4 mutants unable to activate P*emm* also demonstrated a corresponding defect in autoactivating P*mga* (Figure 11A and B). Furthermore, expression of several CMD-1 M6 Mga mutants from their native M6 P*mga* did not produce wild-type levels of protein (Figure 12), further supporting a direct role of an active Mga in its own regulation. However, because of the fact that the M6 Mga mutant Q11A when produced from a constitutive promoter was transcriptionally defective at P*emm* and not P*mga* (Figure 11A and B), it appears possible to unlink the ability of Mga to activate itself from its ability to activate downstream promoters. This hypothesis is further supported by both the differences in the number, size, and location of Mga binding sites at each promoter when compared to the start of transcription (149, 151, 154).

Previous studies have suggested a similar ability of an M49 Mga produced from a divergent *mga-2* allele to undergo autoregulation based on Northern analysis (179). Interestingly, when the same mutations in CMD-1 as well as HTH-4 that affected autoactivation in M6 Mga were created in the divergent M4 Mga (*mga-2*) no effect was seen on their expression from the native M4 P*mga* (Figure 13). From this, it appears that normal expression of M4 Mga from its own promoter is not dependent on an active Mga. Since the M4 and M49 strains share 99.5% sequence identity across P*mga,* it would be

interesting to determine if all *mga-2* alleles share a common mechanism of *mga* regulation.

D. Conserved functional domains define an Mga family of virulence regulators.

Regions of amino acid conservation found between related proteins will often highlight those areas that are indispensable for function in the cell. In this chapter, we propose a new family of transcriptional regulators found within various pathogenic streptococcal species, including established virulence regulators from *S. dysgalactiae* and *S. pneumoniae*, based on their sequence homology to Mga proteins from the GAS. Even though individual members can vary considerably from one to another, regions of 70-100% identity were observed in the different regulatory proteins. Two such conserved domains, CMD-1 and HTH-4, were subsequently found in our genetic screen as being essential for transcriptional activation in the two divergent Mga proteins tested. Based on these results, one would predict that these regions might likely serve similar roles in other members of this family. Using the Mga-like transcriptional regulator DmgB from *S. dysgalactiae* subsp. *dysgalactiae* to test this hypothesis, we found that mutations in both the conserved HTH-4 and CMD-1 domains resulted in an inactive DmgB as predicted from our results with Mga (Figure 14). Due to technical reasons, our screen only investigated the amino-terminal 150 residues of Mga. Thus, there may be a number of domains in the full-length molecules that will provide interesting targets for further analysis. Although functional differences are bound to exist amongst the proteins to

correspond with the variations in regulated genes, vast amounts of knowledge applicable to the entire Mga family can be gained through exploration of the conserved regions within family members.

CHAPTER SIX:

PUTATIVE DOMAINS INVOLVED IN POST-TRANSLATIONAL MODIFICATIONS AND OLIGOMERIZATION

I. Introduction

Expression of the Mga regulon has been shown to be activated in response to different signals such as increased CO_2 levels, body temperature, and the exponential phase of growth (30, 150, 152, 181). However, the mechanisms by which Mga regulates the expression of its genes in response to such cues remain undefined. Several lines of investigation will be discussed in this chapter addressing the question of how Mga responds to its environment.

The first possible mode of regulation is via a two-component regulatory system. This system is made up of both a sensor protein that spans the membrane of the cell acting as a receptor for an environmental stimulus and a response regulator, which transmits the signal from the receptor to the appropriate gene promoters. The domain within the response regulator that accepts the signal is known as the receiver domain. Two putative response regulator receiver domains have been suggested to exist in Mga based on their homology to other known two-component systems (174). Evidence for the existence of these response regulator domains would define a way for Mga to convert an environmental stimulus into a transcriptional response. In order to address this, site-

directed mutational analysis of the key aspartate and lysine residues found in these receiver domains was undertaken to determine if these residues were essential for proper function in the cell.

In addition to two-component regulatory systems, the GAS also contains a phosphoenolpyruvate:sugar phosphotransferase system (PTS), which allows the cell to respond to a change carbohydrate availability within a particular environment. This response involves a phosphorylation cascade where a phosphate from phosphoenol pyruvate (PEP) is ultimately transferred to a regulatory protein containing a PTS-regulatory domain (PRD). This protein, upon changing its oligomeric state, is then able to positively regulate expression of other genes. Existence of these PRD domains has been well characterized in many different systems including both Gram-negative and Gram-positive organisms. An *in silico* analysis of Mga using a consensus protein revealed an area of structural homology to the PRD domain in LicT, an antiterminator protein in *Bacillus subtilis*. Site-directed mutational analysis experiments were conducted to determine if Mga contained a functional PRD domain while gel filtration chromatography was used to establish the oligomeric state of the Mga protein.

II. Results

A. The Mga protein exists in a transcriptionally inactive state in stationary phase.

Previous studies have demonstrated that expression of *mga* and genes in the Mga regulon are growth phase-dependent, reaching a maximal level during exponential phase and drastically decreasing in early stationary phase (152). Protein levels were determined at various time points across the growth curve (corresponding with the different phases of growth) in order to determine whether the loss of expression during stationary phase was the direct result of a drop in protein level or a post-translational modification leading to an inactive protein state within the cell. Samples were grown statically in THY broth at 37ºC and taken during both exponential (4, 4.5 and 5 hrs post-inoculation) and stationary phases (9.5, 29.5, 53.5 and 77.5 hrs post-inoculation) of growth with each sample being adjusted for total cell number (Figure 16A). Both RNA and protein were analyzed at each time point to determine not only the relative amount of *mga* transcript versus Mga protein level, but also for the ability of Mga to activate transcription. Identical samples of total RNA (0.5 and 5 μg) from each time point were probed for *mga* to establish transcript levels, the Mga-regulated gene *emm* to establish the ability of Mga to activate transcription and *23S RNA* to control for loading amongst time points (Figure 16A and B). Both the level of *mga* and *emm* transcripts drastically decreased between 5 hrs and 9.5 hrs of growth post inoculation, which corresponds to late log and stationary phase (Figure 16B: lane 3 and 4), respectively. However, the level of *23S RNA* transcript did

not decrease between these two time points demonstrating that the cells were not transcriptionally defective during stationary phase. The level of Mga protein was also determined for these time points by western analysis of whole-cell lysates with αMga antibodies (Figure 16A and C). In contrast to the level of *mga* transcripts, the steady-state level of Mga protein did not drastically decrease upon entrance into stationary growth phase. Surprisingly, protein was detected in late stationary phase, up to 3 days (4650 min) post inoculation (Figure 16C: lane 7). Thus, during late stationary phase, when transcription of *mga* has ceased, the Mga protein exists but does not activate transcription of regulated genes. This finding suggests that Mga may be modified post-translationally.

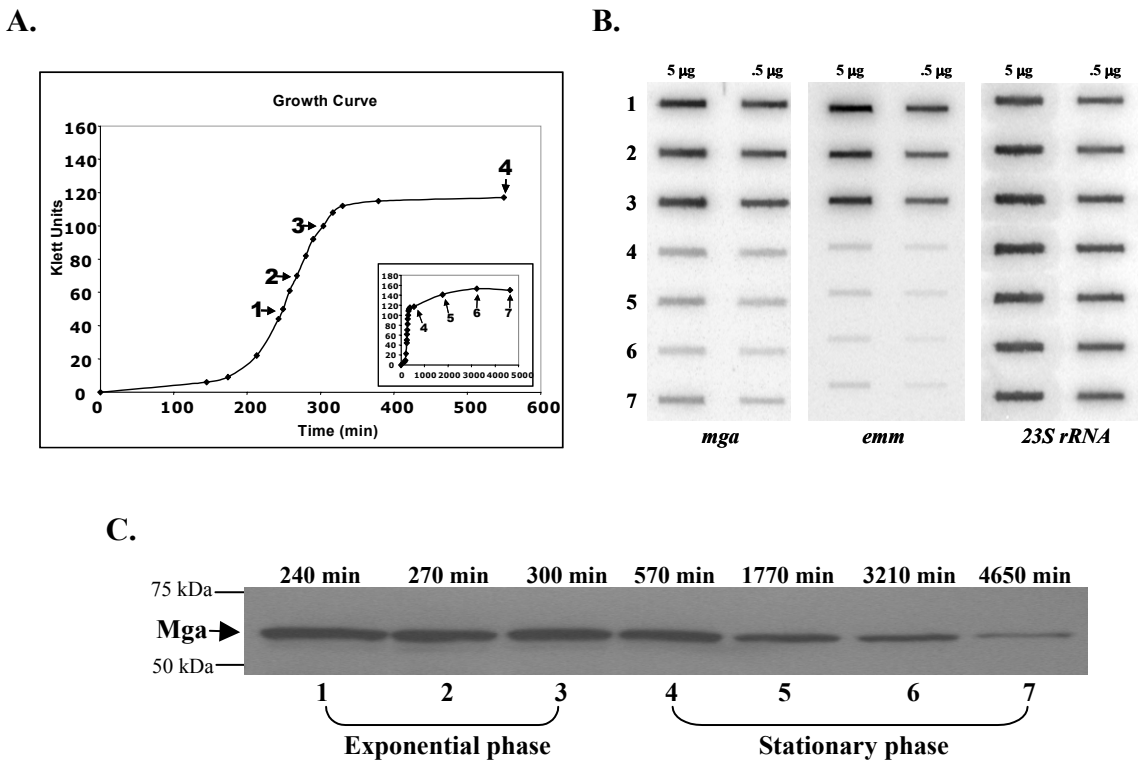


Figure 16: Expression of Mga and the Mga regulon during exponential and stationary phases of growth.

A) Growth curve of JRS4. Cells were grown statically in THY broth at 37ºC for 4 days. Culture growth was measures by turbidity on a Klett-Summerson colorimeter. Arrows indicate time points at which samples were removed for RNA and protein isolation. **B)** Slot blot analysis of transcripts from the Mga regulon in JRS4 across the growth curve. Total RNA (5 or .5µg) was isolated from samples taken at the corresponding time points and identical blots were probed for either *mga* (left), the Mga-regulated gene *emm* (middle) or *23S rRNA* as a loading control (right). **C)** Western analysis of Mga across the growth curve. Whole-cell lysates were extracted from samples taken at the corresponding time points and the level of Mga was detected using αMga antibodies.

B. Receiver domain mutations have various phenotypes *in vivo.*

Perez-Casal *et al.* suggested that Mga could be a response regulator in a two-component regulatory system based on two areas, receiver domain 1 (RD-1) and receiver domain 2 (RD-2), which show limited homology to the receiver domains in a family of response regulator proteins (174). To establish the existence of the putative RD-1 and RD-2 domains, site-directed mutagenesis of Mga was undertaken at both the Asp residue corresponding to the putative site of phosphorylation and the Lys residue corresponding to the site deemed necessary for the structural change involved in activation of a response regulator (Figure 17). Based on published studies, the Asp residues 225 and 459 of M6 Mga, which are located in RD-1 and RD-2, respectively, were changed to Ala (to abolish activation and phosphorylation) (121, 161), Asn (to allow phosphorylation of alternative sites) (8, 22), and Glu (to mimic the phosphorylated state and constitutively activate) (83, 121, 129, 161). The Lys residues 285 and 519 were changed to Arg so that the protein would no longer be able to undergo the proper structural change and thus remain in an inactive state (143).

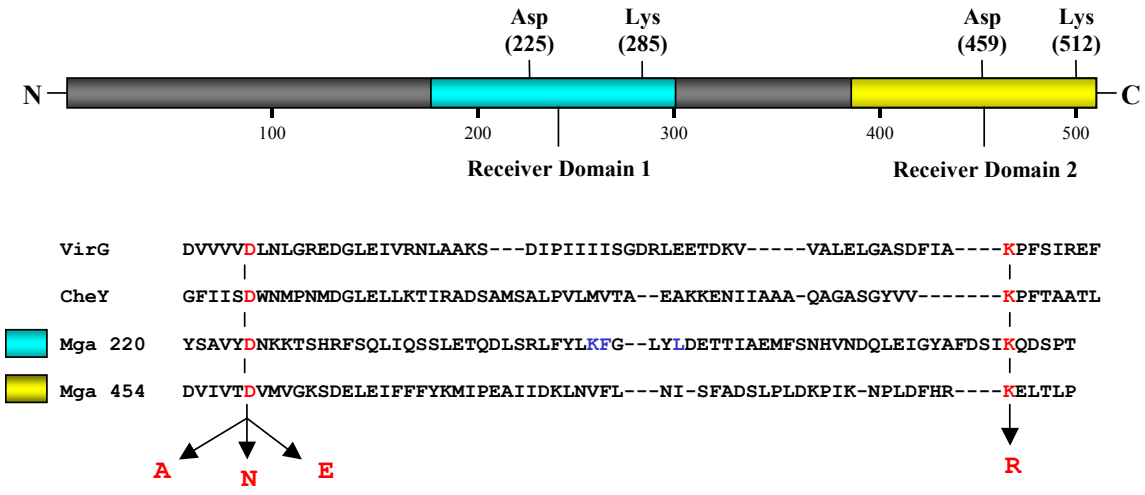


Figure 17: Schematic of the putative receiver domains and the corresponding mutation locations.

The location of the putative receiver domain 1 (teal box) and 2 (yellow box) within the M6 Mga protein (top). Positions of the key Asp and Lys residues of each domain are noted above the schematic. An amino acid alignment of the two receiver domains of Mga as predicted by Perez-Casal *et al*. (174) with two other proteins (VirG and CheY) known to contain receiver domains (bottom). The key Asp and Lys resides along with the corresponding mutations are listed in red while resides listed in blue have been changed from or added to the original sequence published by Perez-Casal *et al*. (174). Dashed lines indicate gaps introduced to improve alignment amongst sequences.

Plasmids were constructed that contained the above mutant *mga* alleles under the control of the constitutive P*spac* promoter to remove any contribution of autoactivation that is seen with the native promoter. The plasmids were transformed into the KSM148.150 strain, which contains the *emm* promoter fused to a promoterless *gusA* and an insertional inactivation of *mga* at its native locus. The level of β-glucuronidase from each P*emm-gusA* reporter strain was measured to examine the effects that each point mutation had on the activity level of Mga *in vivo* (Figure 18). Western analysis was also performed on whole-cell lysates to determine if the mutations altered steady-state protein levels within the cell.

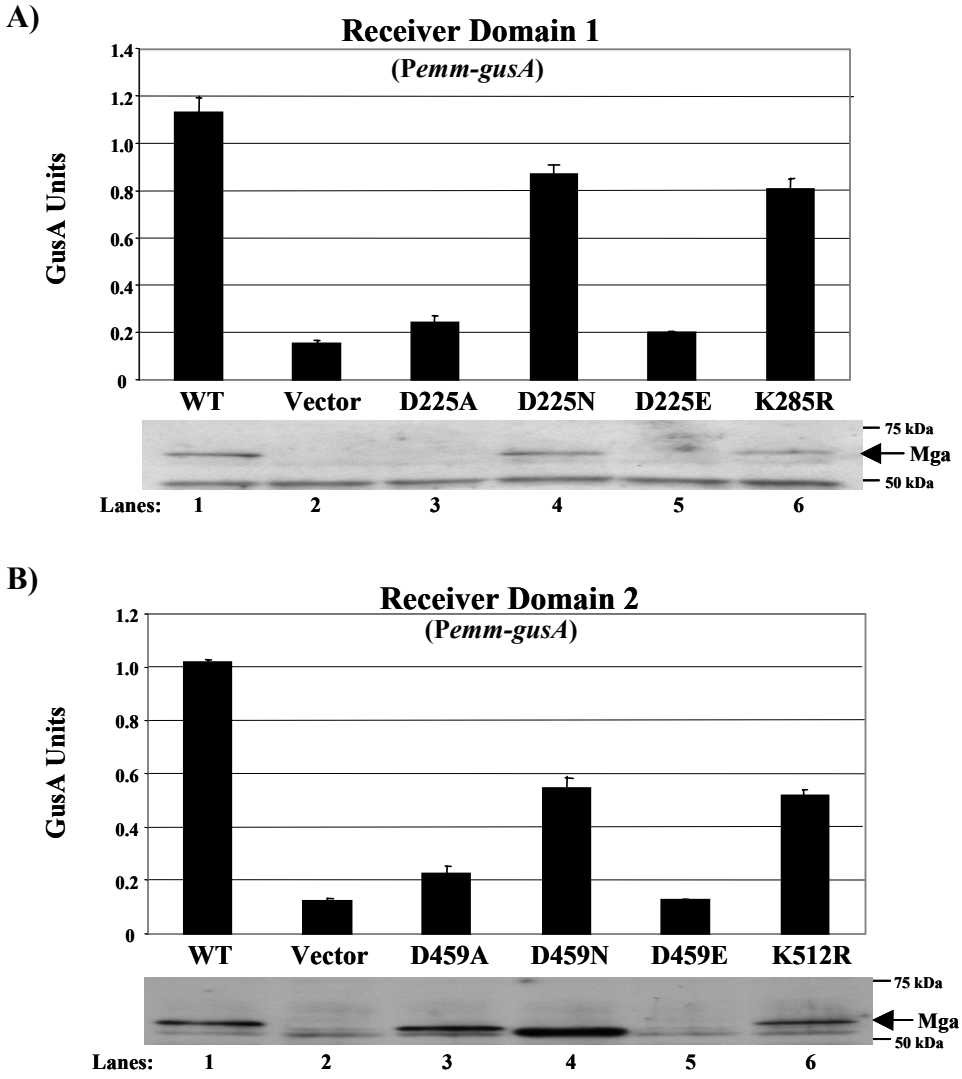


Figure 18: ***In vivo*** **transcriptional activity of receiver domain mutants under a constitutive promoter.**

A) GusA activity of whole-cell lysates (top). The β-glucuronidase activity was determined for *mga*-deleted P*emm-gusA* reporter strains KSM148.150 containing plasmids that constitutively express mutations at either the Asp (D225A, D225N and D225E) or Lys (K285R) residue within the first putative receiver domain of *mga*. Plasmids expressing both WT *mga* and a vector control were also included for comparison. GusA units are a measure of absorbance (A_{420})/protein concentration (μg/mL). Western analysis on whole-cell lysates was also performed using an αMga antibody for detection of protein levels (bottom). **B)** Strains containing plasmids with mutations at either the Asp (D459A, D459N and D459E) or Lys (K512R) residue within the second putative receiver domain were also analyzed as above for both β-glucuronidase activity (top) and the level of Mga produced within the cell (bottom).

If Mga was indeed a response regulator, both mutants D225E in RD-1 and D459E in RD-2 would be predicted to constitutively activate Mga based on prior mutational analysis involving receiver domains in RcsB from *E. coli*, NtrC from *S. typhimurium* and *E. coli*, and OmpR from *E. coli* (83, 121, 129, 161). However, neither mutant produced a level of β-glucuronidase activity above background or a detectable quantity of protein via western analysis (Figure 18A and B: lane 5). Because the GusA assay directly correlates with the amount of Mga within the cell, the results of these mutations were inconclusive in this study due to the lack of protein. Mutant D225A in RD-1 was similar to the above, once again showing a lack of both protein and β-glucuronidase activity (Figure 18A: lane 3).

The Mga mutants D225N in RD-1 along with D459A and D459N in RD-2 were expected to abolish Mga-regulated activity *in vivo* based on previous studies (83, 121, 161). While mutations in RD-2 drastically reduced the level of activity (Figure 18B: lanes 3 and 4) even though protein levels were similar to WT, the RD-1 mutant (Figure 18A: lane 4) showed only a slight reduction in activity. In order to provide more substantial evidence for the existence of a receiver domain, an additional mutation at the Lys residue in both RD-1 and RD-2 was also analyzed. This residue is thought to play a key role in the conformational change involved in the activation of CheY (143). Once again, the RD-1 mutant K285R (Figure 18A: lane 6) was only slightly below WT activity levels whereas the RD-2 mutant K512R (Figure 18B: lane 6) was decreased to approximately half the WT level. Both mutations produced protein at a comparable level

to WT. Taken together, RD-2 contains residues important for activity, and appears to be the only predicted domain to demonstrate characteristics similar to those seen with the previously characterized response regulators.

C. *In silico* analysis of the Mga protein predicts new putative domains.

Since all of the known domains of Mga lie within the N-terminal one-fourth of the protein and the random mutagenesis screen was not applicable to the C-terminus of the protein, an *in silico* analysis of Mga was performed (in collaboration with Sara Cheek from the laboratory of Dr. Grichin) to search for putative domains that demonstrate three-dimensional homology to other known domains. First a profile for the protein was constructed by running a PSI-BLAST search for five iterations against the non-redundant protein database using the M6 Mga as the query sequence. Either the original query sequence or the profile was then used as the input sequence for the following prediction methods: RPS-BLAST (available at http://www.ncbi.nlm.nih.gov/), COMPASS (195), SUPERFAMILY and 3-D Jury (available at http://BioInfo.PL/Meta/). In addition to the known DNA-binding domains and the putative CheY-like fold discussed previously, two new domains known as PTS-regulatory domain or PRDs were also predicted based on structural similarity (Table 10).

Table 10: Structural prediction methods and values

	Residues	RPS-BLAST e-value	PSI-BLAST e-value *	COMPASS e-value *	SUPERFAMILY e-value	3D-JURY Score *
DNA-binding domain 1	1-70	N/A	8.2	1.60e-22	1.1	43.75
DNA-binding domain 2	85-150	N/A	1.1	1.63e-07	0.012	44.38
CheY like fold domain	395-530	N/A	N/A	N/A	N/A	41.75
PRD domain	170-390	N/A	2e-18	4.18e-46	N/A	136.12

***** The highest e-value/ score obtained from all hits that covered all but 10 residues at either end of the database domain is reported here.

D. Mutational analysis of the two putative PTS-regulatory domains in Mga.

Mutational analysis of the His residues within each of the two PRD domains was undertaken to determine whether the three-dimensional predictions represented functional domains in Mga. Previous structural/functional mutational analysis has been performed on several well-characterized PRD containing proteins. From these prior studies it was demonstrated that by changing the His to an Ala, phosphorylation of the protein was abolished; changing the same His to an Asp mimicked the phosphorylated state resulting in a constitutively active protein (147, 226).

Based upon this knowledge, a similar analysis was undertaken on alleles of the M6 serotype *mga* in which the His residue of each of the PRDs (Figure 19) was mutated to either an Ala or an Asp and placed on a plasmid under the control of the constitutive *rpsL* promoter. The ability of the mutants to activate downstream genes was determined

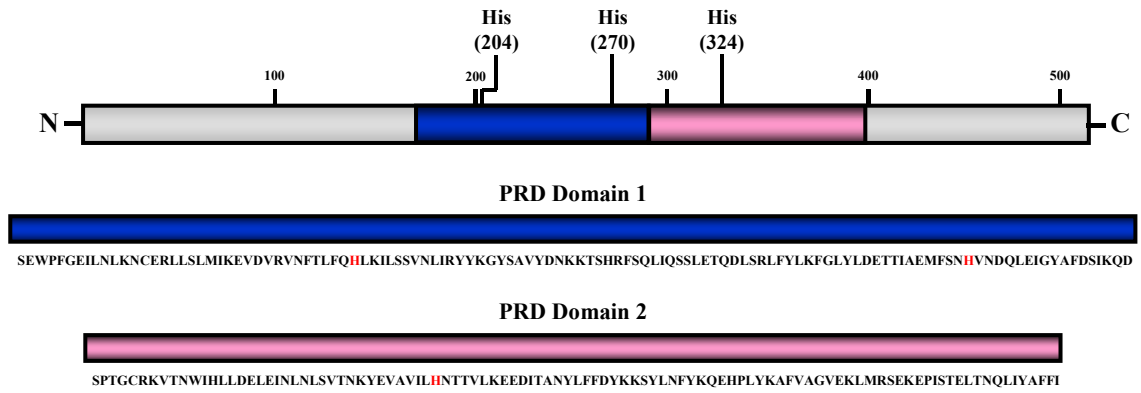


Figure 19: Diagram of the two putative PTS-regulatory domains within the Mga protein.

Based on structural homology, two putative PTS-regulatory domains were predicted to exist within Mga. The first domain, shown in blue, or PRD-1 spans from reside 170 to 287. The second domain, shown in pink, or PRD-2 spans from residue 288 to 390. The composition of each domain in the M6 Mga protein is listed below with the specific His resides predicted to be phosphorylated highlighted in red.

by transforming the plasmids into an Δ*mga* M6 strain, JRS519, and detecting the level of transcription of the Mga-regulated gene *emm* in each strain via Northern analysis (Figure 20). Western blotting analysis was also performed on whole-cell lysates to ensure that the mutation did not alter the overall protein level. Samples were taken at three different time points corresponding to mid-log (70 Klett), late-log (100 Klett), and stationary phase (3 hours post 100 Klett) in case the mutational affects were growth phase-dependent.

Interestingly, the overall pattern of Mga regulation for WT Mga (Figure 20A and B: lanes 1, 5, and 8) was similar to that seen for Mga-regulated transcription *in vivo* when *mga* is produced from its native promoter, with *emm* transcripts seen in both mid and late-log phase but absent in stationary phase (152) even though in this study *mga* was under a constitutive promoter. Protein production, albeit at lower levels, is also

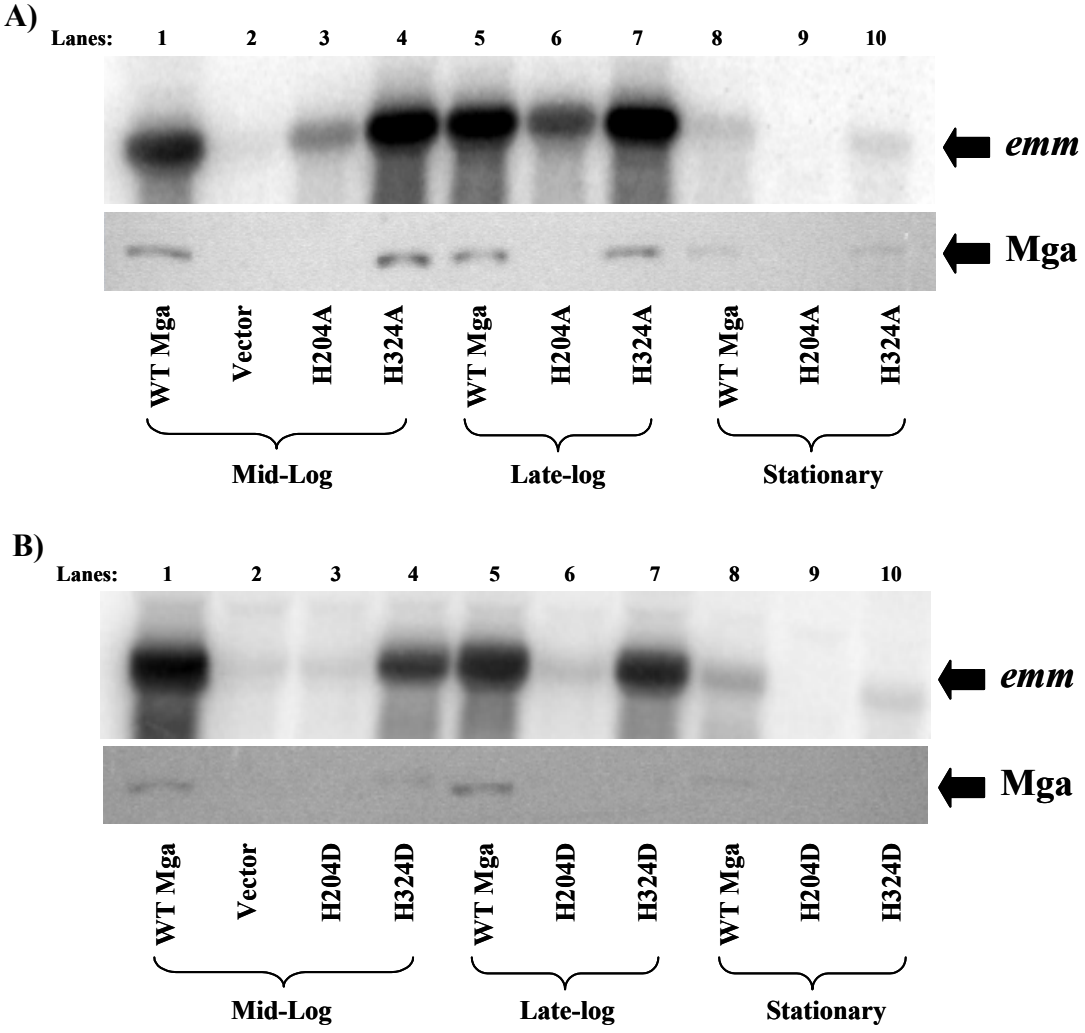


Figure 20: Transcriptional activation ability of an Mga protein containing mutations in the PRD.

A) Northern analysis of Mga-regulated transcriptional activation produced from a Mga containing an His to Ala substitution at different phases of growth (top). Transcript levels were determined for the Mga-regulated gene *emm* in 5 μg of total RNA isolated from the Δ*mga* M6 strain containing plasmids expressing either WT *mga* (lanes 1, 5 and 8), vector only (lane 2), or an *mga* gene mutated in PRD-1 (H204A [lanes 3, 6, and 9]) and PRD-2 (H324A [lanes 4, 7 and 10]) from the constitutive *rpsL* promoter. Western analysis using an αMga antibody was also performed on 7 μg of whole-cell lysates to determine the amount of Mga in each strain (bottom). **B)** Analysis of Mga-regulated transcriptional activation produced from an Mga containing a His to Asp substitution at different phases of growth. Northern (top) and western analysis (bottom) was performed as above for strains containing WT *mga* (lanes 1, 5 and 8), vector only (lane 2), or an *mga* gene mutated in the PRD-1 (H204D [lanes 3, 6, and 9]) and PRD-2 (H324D [lanes 4, 7 and 10]) from the constitutive *rpsL* promoter.

consistent with that seen for P*mga-mga* (Figure 16). Similarly, both PRD-2 mutants H324A, which was expected to eliminate phosphorylation, and H324D, which was expected to mimic the phosphorylated state, demonstrated the same overall transcriptional and translational profiles as the WT Mga (Figure 20A and B: lanes 4, 7, and 10). Because both mutants demonstrated the same phenotype instead of different phenotypes as would be expected considering the opposing nature that the two mutations have on the phosphorylation state, it is implied that PRD-2 is not an active domain in the M6 Mga protein.

In contrast to the PRD-2 mutants, *emm* transcription was greatly diminished and Mga was not detected at any of the three phases of growth tested when the His at position 204 was changed to an Ala to eliminate phosphorylation at the site in PRD-1 (Figure 20A: lanes 3, 6, and 9). However, when the same His in PRD-1 is changed to an Asp, which was expected to mimic the phosphorylated protein, the mutant H204D once again did not produce WT levels of *emm* transcription or a detectable level of Mga (Figure 20B: lanes 3, 6, and 9). The fact that both mutants were unable to produce protein indicates that position 204 is sensitive to mutation and may be involved in stability; however, a definite conclusion concerning the ability of the putative PRD-1 to act as a function domain cannot be deduced from these results.

E. Mga can bind to DNA as a monomer

Since many proteins that contain PRDs often function as dimers, the oligomeric state of Mga was also investigated. Mga-His was first purified from *E. coli* using a

NiNTA affinity column before its apparent molecular mass was determined on a Superdex 200 size exclusion column. The initial eluent (Figure 21A) was a phosphate buffer solution that contained 20 mM phosphate and 0.5 M NaCl. Using this solution, three different peaks were seen: the first peak corresponded to the void volume, the second peak had an apparent molecular mass much greater than the expected size of a dimer (based on a theoretical molecular mass of 62 kDa for Mga), and the final peak was equivalent to the calculated size of a monomer. From this, it was determined that in phosphate buffer alone most of the protein was forming precipitants too large for the column to separate based on size and thus eluting in the void volume.

Since the Mga protein is known to precipitate out of solution when the concentration of imidazole in the solution reaches a minimum threshold (A. Almengor, unpublished data), the experiment was repeated using a second eluent that contained 20 mM phosphate, 0.5 M NaCl, and 500 mM imidazole (Figure 21B). Under these conditions, only one peak was seen with an apparent molecular weight comparable to that of a monomer Mga. These results indicate that in the presence of 500 mM imidazole the Mga protein exists only in a momeric form.

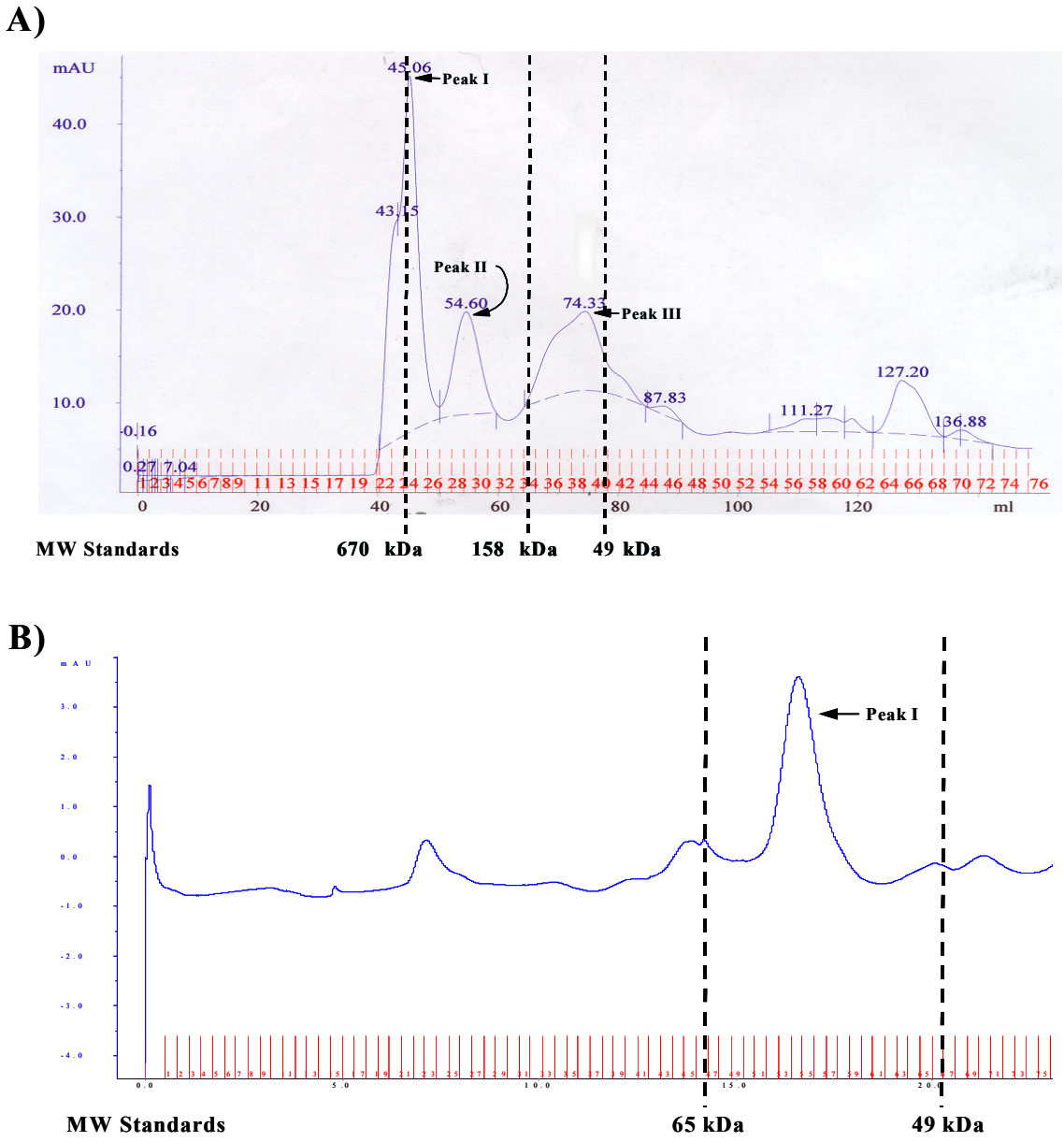


Figure 21: Oligomerization state of Mga.

Size exclusion chromatographs of Mga-His purified from *E. coli* on a Superdex 200 column. **A)** Elution profile of Mga in a solution of 20 mM phosphate and 0.5 M NaCl. Both the elution volume of each peak along with the closest corresponding molecular weight (MW) standard is shown at the bottom of the graph. **B)** Elution profile of Mga in a solution of 20 mM phosphate, 0.5 M NaCl and 500 mM imidazole. MW standards and elution volumes are listed along the bottom.

Because the binding of Mga to its DNA promoter targets is an essential step in activation of the Mga-regulated genes (151), electrophoretic mobility shift assays were performed to determine if the purified protein retained its ability to bind DNA under the same conditions that were used for its elution from the Superdex 200 column (Figure 22). WT M6 Mga-His was purified from *E. coli* lysates using a NiNTA affinity resin column as described previously. 5 μg of the purified Mga-His protein was incubated with a constant amount of radiolabeled promoter probe corresponding to the M6 Mga-regulated P*emm* in a solution containing either no additional imidazole or 500 mM imidazole (consistent with the conditions used for the elution of Mga from the size exclusion column (Figure 21)). Both EMSA reactions showed a reduced mobility of the probe indicating that Mga was able to perform this initial binding step even in the presence of 500 mM imidazole when only a monomer is present in solution.

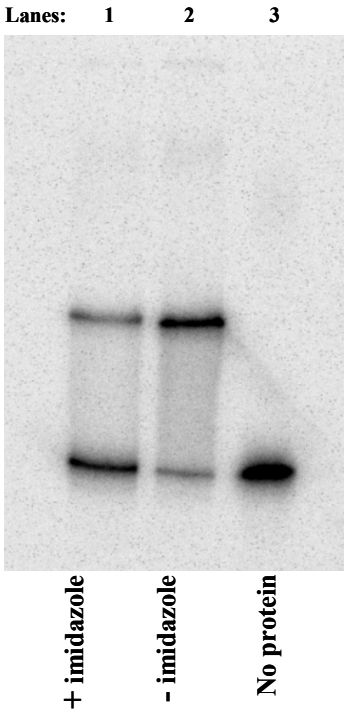


Figure 22: Electrophoretic mobility shift analysis of Mga binding to P*emm* in the presence of imidazole.

Binding analysis of Mga to the Mga-regulated promoter P*emm*. C-terminal 6X-His fusion proteins (Mga-His) were purified from *E. coli* lysates. Identical amounts of both protein (5 µg) and radiolabeled probe were incubated for 15 min at 16°C in a solution with (+) or without (-) 500 mM imidazole prior to separation on a 5% polyacrylamide gel (lanes 1 and 2). A lane that did not contain protein (lane 3) was added as a control.

III. Discussion

A. Loss of Mga activity in stationary phase is not due to a loss of protein.

A previous study by McIver and Scott investigated transcription of *mga* and the Mga-regulated genes *emm* and *scpA* throughout the growth cycle (152), determining that transcription peaked during exponential phase and that the growth phase dependence

required the native promoter of *mga*. However, the mechanism behind this growth phase dependence was not established. It was shown here that in late stationary phase, hours after transcription of *mga* was abrogated, the Mga protein could be detected although it is no longer able to activate transcription at downstream targets. This finding implies that a mechanism other than proteolytic degradation must influence Mga-regulated activity within the cell. Many different explanations could account for this phenomenon such as the existence of an inhibitor during stationary phase or a post-translationally modified Mga protein. These possibilities are discussed further in chapter seven of this study, although further experimentation is necessary to decipher the exact mechanism of control.

In addition to demonstrating that the protein existed in stationary phase, it was also shown that the level remained relatively constant over time (Figure 16C: lanes 5 and 6). Although a slight drop was seen in late stationary phase (Figure 16C: lane 7), this could be the result of an overall drop in viable cell number at this time point, which was not accounted for in this experiment. Because the level of Mga remains relatively constant after transcription has ceased, it can be inferred that Mga is relatively stable in the context of the cell. Nevertheless, until the half-life of the protein is determined this cannot be stated explicitly.

B. Comparison of Mga with other response regulators.

Previously, several arguments have been made suggesting Mga acts as a response regulator in a two-component system. The first argument is based on the fact that Mga

and the genes that it regulates have been shown to respond to environmental stimuli (30, 150, 152, 181). The second is derived from the homology between Mga and other response regulators such as CheY from *E. coli* and VirG from *Agrobacterium tumefaciens* (174). In addition, the present study has demonstrated that transcription of the Mga-regulated genes is not always activated in the presence of the protein (Figure 16), thus the protein may exist in both an active and inactive state within the cell. Since previous attempts at direct phosphorylation were unsuccessful (K. S. McIver, unpublished data), site-directed mutagenesis was used to determine if either receiver domain 1 or receiver domain 2 was functional in Mga.

Upon mutation of the key Asp residue in RD-2, a loss of Mga activation at P*emm* was detected. The drop in activity of the RD-2 mutant D459N was not as drastic as that seen for D459A, possibly because D459N may act like the equivalent D57N mutation in CheY being phosphorylated at another residue and thus retaining some activity (8). In addition, the activity level of the K512R mutant was also reduced, indicating that RD-2 may act as a functional receiver domain. Unfortunately, the Glu mutation at position 459 did not yield protein and thus the results cannot be confirmed by a mutation leading to a constitutively active phenotype as seen with other response regulators.

In contrast to RD-2, the experimental results of RD-1 conflicted with those obtained from other response regulators. Of the two mutants that made protein, neither D225N nor K285R showed a loss-of-function phenotype. The very slight drop in RD-1 activity seen for both mutants was too small to be indicative of a loss-of-function

mutation indicating that RD-1 probably does not act as a receiver domain in Mga. However, the possibility cannot be completely ruled out for at least two reasons. The first reason is that a particular mutation may not have the same effect in all response regulators. In a study looking for activating mutations, it was suggested that substitutions that could act universally on all response regulators "are rare in receiver domains, if they exist at all" (208). This is echoed by the finding that when the key Asp is changed to a Glu it makes CheY inactive, an effect opposite from the constitutive activity seen with NtrC (161). The second line of reasoning is based on the experimental evidence that when a proline is placed four amino acids downstream from the conserved lysine (resulting in an Ala to Pro exchange at position 113 in CheY and an Iso to Pro exchange at position 106 in VirG) a constitutively active protein is produced for both CheY and VirG (23, 173). The corresponding amino acid in Mga at 289 is normally a proline so it is possible that this domain is already constitutively active and thus the subsequent mutations show no effect.

Structural analysis of the known regulatory domains reveals that all response regulators demonstrate a high degree of conservation consisting of five strands with an alternating pattern of β-strands and α-helices (214). Based on the predicted structure of Mga, neither domain has the conserved alternating β/α pattern. Jpred predicts RD-1 to consist mainly of α-helices while RD-2 consists mainly of β-strands suggesting that if either domain is functional it may not act in the traditional sense.

C. PRD domain mutations are inconclusive.

In addition to the receiver domains, Mga was also predicted to contain two PTS-regulatory domains. Support for the existence of these domains includes homology to other known PRD domains (Table 10), the influence of different sugars in the environment on the transcription of the Mga-regulated gene *emm* (177), and the possibility of an active and inactive protein state. Several additional observations also support the role of sugar metabolism in the regulation of Mga, including evidence that the catabolite control protein, CcpA, binds to the Mga promoter at a *cre* site (A. Almengor, unpublished data) and the regulation of Mga by other proteins involved in sugar regulation (191). Again, mutagenesis of the putative sites of phosphorylation for each PRD domain was conducted in order to provide some initial evidence that these residues were functionally important in the protein. In total two of the three His residues, one in PRD-1 and the other in PRD-2, was changed to produce the following four mutants: H204A and H324A, which should no longer accept a phosphate, and H204D and H324D, which should mimic the phosphorylated state.

Both the WT and all four mutants lacked *emm* transcript during late stationary phase even though Mga was under a constitutive promoter (Figure 20A and B lanes 8, 9, and 10). Since no other transcripts were probed for at this time point, the possibility that there is an overall degradation of mRNA cannot be ruled out. Alternatively, this could result from specific inhibition of *emm* by either loss of an activator that works in conjunction with Mga, an inactivation of Mga itself, or an active repressor of this gene.

Aside from the loss of transcription during stationary phase, H204A and H204D were transcriptionally defective across all phases of growth and showed no detectable level of protein production. When a loss of protein is seen although the gene is constitutively transcribed, the general conclusion drawn is that the mutation affected the stability of the protein. Because this position is located within the C-terminal mutagenesis fragment from chapter five, this finding once again demonstrates the sensitivity of the protein to mutation within the region. It is also interesting that this phenotype is observed upon mutation of amino acid 204 because this His residue is only conserved amongst the GAS Mga protein compared to the His at position H324, which is conserved amongst other orthologues of Mga besides Mgc from *Streptococcus dysgalactiae*. Overall, further investigations are necessary to deduce the ability of this domain to function in Mga.

D. Oligomeric state of Mga in the cell

Through the use of size exclusion chromatography, it was demonstrated that Mga is in a monomeric state in the presence of 500 mM imidazole and that it is able to bind to the DNA target P*emm* in this state. In this study only binding at the *emm* promoter was tested, so future investigations should be aimed at determining whether the monomeric form can also bind to different categories of Mga-regulated promoters such as P*mga*. Unfortunately, these experiments were not able to determine the activity of the monomeric state. Since a high concentration of imidazole inhibits the *in vitro*

transcription system at all promoters (A. Almengor, unpublished data), the ability of the monomeric state to activate P*emm* could not be tested directly.

CHAPTER SEVEN:

CONCLUSIONS AND RECOMMENDATIONS

I. Summary

Overall, this study has provided valuable insight into how the Mga protein may function within the context of the streptococcal cell. Before this study began, most investigations involving Mga revolved around identification of the genes activated by Mga and the binding patterns at the individual promoters. Very little was known about the protein itself except that it acted as a positive transcriptional regulator and bound directly to DNA using residues within the two HTH domains spanning from amino acid 53 to 72 and 107 to 126, respectively (151). Now this list of functional residues can be expanded to include the necessity of amino acid 461, located within the C-terminus, for activation of the divergent M4 Mga (Chapter Four). Despite the variation that is seen in the C-terminus compared to the N-terminus of different Mga proteins, residue 461 is conserved in both GAS *mga* alleles (*mga-1 and mga-2*). Nevertheless, it was shown to be essential only for the function of Mga proteins from the *mga-2* alleles. This demonstrates that these two divergent alleles may use different mechanisms to activate gene transcription in their respective GAS serotypes.

A novel functional domain was also identified in the extreme N-terminus (amino acids 10 to 15) of the M6 Mga protein, encoded by an *mga-1* allele (Chapter Five).

Mutagenesis revealed that the flanking residues of this domain were important for stability, while all other residues tested within this domain were important for activation at the downstream target promoter of *emm*. They were also necessary for activation at the native promoter of *mga* with the exception of one mutant, Q11A. This mutant was transcriptionally defective at P*emm* but not at P*mga,* demonstrating that the ability to autoactivate can be separated from the ability to activate downstream genes. In contrast to that seen with residue 461, the importance of this domain is very broad and affects transcriptional activation not only from the two different GAS Mga proteins but also the Mga-like regulator from the group C streptococcus, known as DemB (Chapter Five). Interestingly, this domain is conserved amongst the members of an entire family of proteins identified as putative virulence gene regulators. Thus, it is possible that this domain is functionally important in all members of this family and indicates that a common mechanism of activation may also be present.

This study has also divulged new information about how the Mga protein itself is regulated. It was shown that Mga persists in a transcriptionally dormant state during stationary growth phase (Chapter Six). With the aid of bioinformatics, different regulatory domains that may be responsible for this inactivation of the protein were revealed. Although these investigations were preliminary, they provided several new areas of study that may help to further explain the control of Mga activity. Several of these topics related to the Mga protein and its regulation are discussed further in the remainder of this chapter.

II. Putative protein-protein interaction domains of Mga

Many new and interesting questions were generated throughout the course of this work. A case in point was the identification of two Mga mutants found to retain full DNA-binding capabilities but yet were transcriptionally silent (Figure 11 and Figure 15). From this data, it can be concluded that the Mga protein potentially interacts with other cellular constituents, in addition to DNA, in order to activate transcription. All protein-protein interactions require a domain or specific contact residue(s) where the associations occur. Mga could contain a variety of possible domains but only the three most likely are discussed here.

It is highly probable that Mga may possess a domain allowing for a direct interaction with RNA polymerase (RNAP). Transcriptional activators often work by binding directly to RNAP at the promoter. The most common interactions occur between the activators and either the α or σ subunits of RNAP, although some have been shown to bind to the β or β' subunits (190). Interactions between Mga and the transcriptional machinery have yet to be described, but our recent establishment of an *in vitro* system for Mga regulation (5) and the use of RNAP mutants will allow us to investigate this possibility in detail.

A second possible domain would be one that allowed the protein to bind to itself and form a higher order multimeric structure. Almost all proteins that contain HTH domains for binding function as homodimers, binding to palindromic or pseudopalindromic sites on the DNA (103). It is also common for proteins that contain

PRDs to use multimerization as a way to confer activity to the protein. Dimerization is also necessary for some response regulators to be able to bind DNA (193). Because bioinformatics has predicted homologous areas within Mga to all of these motifs, the possibility exists that Mga may act in the same fashion. Whether the protein forms higher order structures *in vivo* is unknown; however, initial *in vitro* results suggest that Mga is capable of binding to DNA as a monomer (Figure 22). Nevertheless, this study did not demonstrate that Mga was transcriptionally active leaving open the possibility that Mga retains the ability to bind to DNA as a monomer but is not transcriptionally active in this form. Conversely, Mga could be active in a monomeric state and upon dimerization become inactive.

Another potential domain would be one involved in a protein-protein interaction of Mga with either a repressor or an activator of the protein. Given that Mga purified from *E. coli* is able to activate transcription *in vitro* (5) argues against this possibility, because it implies that the interacting protein must also exist in *E. coli* and that it must be co-purifying with the Mga protein. Regardless, this possibility cannot be ruled out until pure reagents are available for use in the system. Experimental methods such as bacterial two-hybrid systems or pull-downs may aid in the resolution of any possible interacting partners. Studies aimed at identifying domains involved in potentially novel functions for this regulator should be continued because they will provide new insights into the mechanism of Mga regulation.

III. Growth-phase regulation of Mga

The existence of Mga in stationary phase, long after not only its own transcription but also transcription of the genes that it regulates has ceased (Figure 16), signifies that there must be a mechanism within the cell to turn off the transcriptional activity of Mga. One explanation outside of the aforementioned protein-protein interactions is that the Mga protein itself remains active but another protein or regulatory molecule found in stationary phase inhibits transcription at each of the Mga-regulated promoters. This explanation seems unlikely given that the point of peak transcriptional activation of *emm* was shifted into stationary phase when *mga* was placed under the control of the heterologous promoter Pϕ2 (152). However, it cannot be disregarded because the level of activation seen in stationary phase was much lower than the levels seen during exponential phase. This implies that there may be a partial inhibition or incomplete activation of Mga-regulated genes during stationary phase. Another explanation for this effect would be that Pϕ2 is a weaker promoter than P*mga* and thus, if this were the case, the overall level of active protein would also be lower in this strain. The ability to produce a constitutively active Mga could rule out this possibility since, in theory, a repressor that acts independently of Mga should not be masked by a mutation in the Mga protein.

Another possible mechanism that Mga may use to turn off its activity would be to undergo a post-translational modification during a particular stage in the growth cycle, which serves to activate or inactivate the protein at the appropriate times. There are

many advantages to controlling activation at the level of the protein including a quick response time after stimulation and the ability to reverse the reaction allowing for the components to be reused without being resynthesized. Given that Mga possesses homology to both a two-component receiver domain and a PRD, phosphorylation would be the most probable post-translational modification for Mga to undergo. *In vitro* systems exist for the phosphorylation of either a receiver domain (using small phosphodonors) or a PRD containing protein (using PEP, EI, and HPr) and could be used to address this issue. Previous efforts at *in vitro* phosphorylation using the system designed for receiver domains were unsuccessful (K. S. McIver, unpublished results), thus future attempts will be directed towards phosphorylation of the PRD.

IV. Autoregulation of Mga

The M6 Mga (*mga-1*) has been shown previously to bind directly to its own promoter, resulting in activation of *mga* expression and amplification of the Mga response (75, 151, 154). Supporting this model is the fact that DNA-binding mutants that affected transcriptional activation at downstream promoters also affected transcription from the native promoter in the same M6 strain (151). In contrast, mutations within the divergent M4 Mga (*mga-2*) that were shown to have an effect on downstream Mga-regulated genes did not affect the overall level of protein expressed from the native M4 *mga* promoter as would be expected from an autoregulated protein (Chapters Four and Five). Therefore, it appears that unlike what was demonstrated for M6 Mga, normal

expression of M4 Mga from its own promoter is not dependent on the existence of an active Mga within the cell.

Previous analysis of the *mga* promoter region from the M49 strain, which also contains a divergent *mga-2* allele, found significant differences at the nucleotide level when compared to the same region from an M6 (*mga-1*) strain (179) (Figure 23). An S1 nuclease assay of the M49 promoter region deduced different transcriptional start sites (Figure 23 gray vs. yellow highlights). Different Mga-binding sites were also predicted compared to those identified in the M6 strain (154). Even though the binding sites and thus the possible mechanism of action was different, Northern analysis suggested that *mga* in the M49 strain was still autoregulated (179). Reevaluation of these studies in the M49 strain using the sequence of the Mga-binding site found in the M6 strain (154) suggests that only one of the two Mga-binding sites may be functional due to a nine nucleotide insertion at the start of the second Mga-binding site (Figure 23 blue box). Given that the M4 and M49 strain share 99.5% sequence identity across P*mga*, it is likely that these *mga-2* alleles share a common mechanism of *mga* regulation.

To verify that purified Mga can bind to each of the two binding sites within the promoter region of the divergent *mga,* electrophoretic mobility shift assays could be used. Also, autoactivation of a divergent P*mga* could be analyzed using a transcriptional reporter fusion to determine if activity levels are directly affected by mutations shown to prevent transcription of Mga such as the HTH-4 DNA-binding deficient mutants. These experiments should be extended to include both P*mga4* and P*mga49* to help determine

whether all divergent alleles are regulated in the same manner or if the lack of autoregulation is a phenomenon restricted to the M4 strain.

Autoregulation serves as a mechanism to amplify the signal within the cell. One would assume that in a strain that does not possess this ability, the signal would be dampened. The response time after stimulation may also be slower in these strains because there would be less overall protein to activate downstream targets. However, this may not be the case if Mga is not the limiting substance for activation. For example, if the expression level of Mga from P*mga4* was equivalent to that seen with P*mga6* under autoactivating conditions, then the argument over signal strength would not be relevant. Until antibodies are produced to the divergent M4 Mga, the steady-state level of protein in this strain under native conditions cannot be deduced.

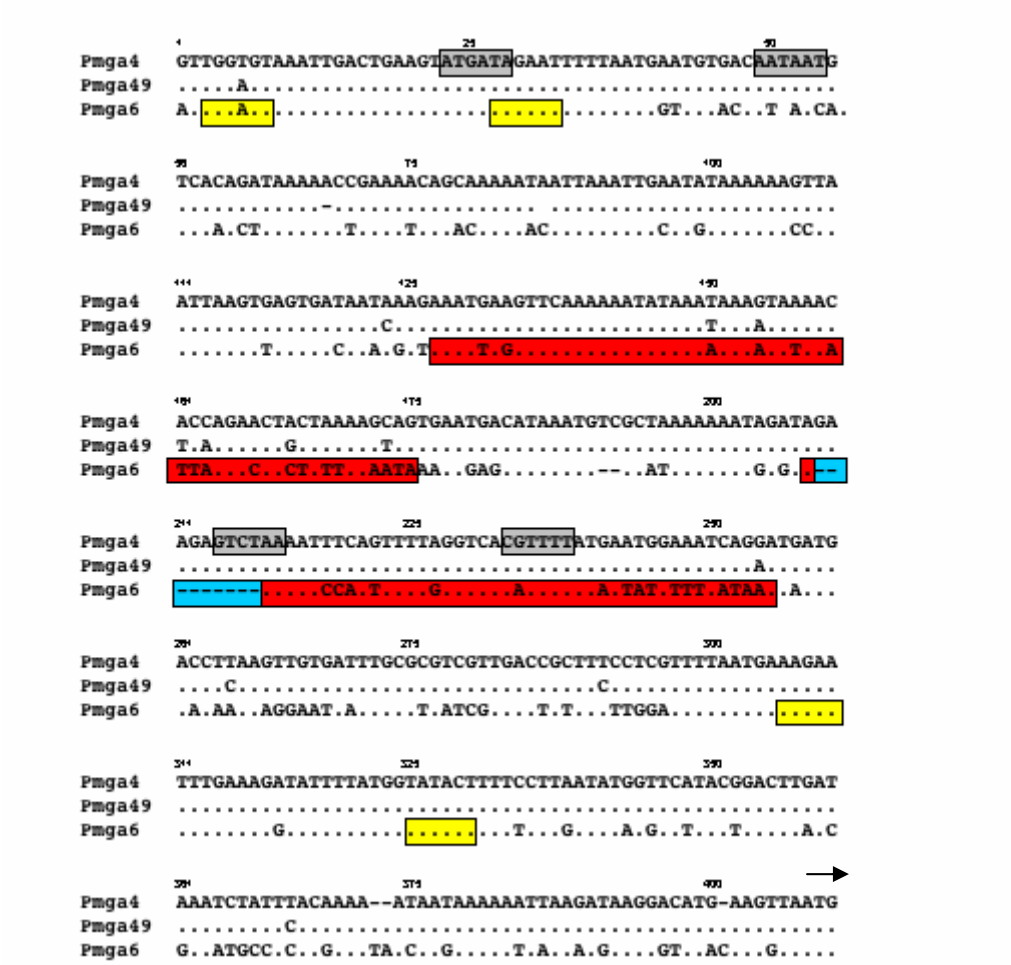


Figure 23: Nucleotide alignment of the promoter regions from M4, M49 and M6 *mga*.

The promoter sequences of M4, M49 and M6 *mga* are aligned with its corresponding identification designated on the left. Both gray (M49 *mga)* and yellow (M6 *mga*) highlighted areas indicate the positions of the –35 and –10 regions of the promoters while the Mga-binding sites for the M6 Mga are highlighted in red. The nine nucleotide insertion site is marked by a blue box in P*mga6*. Identical amino acids are denoted by dots (**.**), while gaps (**-**) have been introduced for maximal alignment. The translational start site is marked with an arrow (⟶).

V. C-terminal mutants fail to produce WT levels of protein

Another interesting point about the Mga protein raised by this study is that the C-terminus appears to be highly sensitive to mutations. This sensitivity is underscored by the fact that during a random mutagenesis screen not a single mutant found in the C-terminal region produced a WT level of protein. Although many of the mutants from the screen contained two amino acid changes, single site-specific mutations in the area such as D225E and D459E also resulted in a lack of a detectable level of protein. Because of this, it appears as though the protein may be easily misfolded or otherwise affected by issues of stability. This is interesting because in nature the two divergent groups of Mga proteins show the largest percentage of diversity in this C-terminal region. Since the original mutagenesis was performed on an *mga-1* allele, it is not currently known if mutations in a divergent *mga-2* allele would have this same phenotype. A similar strategy to the screen used with the *mga-1* allele could be employed to answer this question.

From an evolutionary standpoint, this would imply that most natural mutations in Mga would be deleterious to the proteins ability to activate its target genes. Considering that humans are the only known reservoir for the GAS and that Mga regulates multiple factors necessary for establishment within the human host, it would be reasonable to assume that most of the strains containing a mutant protein would be less virulent. This idea offers an alternative explanation for the cyclic nature of GAS infections that has been seen throughout time (123). The cycle would begin when a hypervirulent strain

quickly spread to susceptible hosts causing the start of an epidemic. Over time, point mutations would reduce the virulence of the strains, which would be considered endemic at this point, until they are merely carried and passed asymptomatically from person to person. Eventually, after the bacteria has undergone many rounds of mutations causing it to become less virulent, a single fully virulent mutant would be produced, possibly through additional mutations or a horizontal gene transfer event. This single strain could then start the cycle over by rapidly spreading and causing a new epidemic. Whether this affect is a direct result of mutations in Mga remains to be seen.

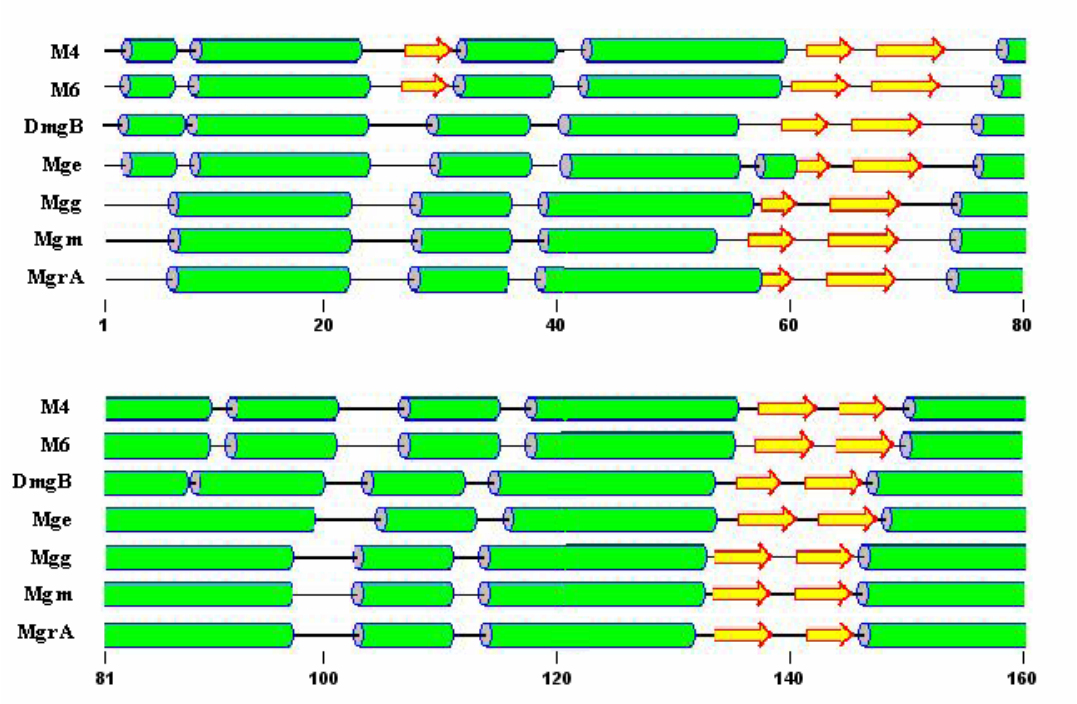
Despite what appears to be a low tolerance for mutation, Mga proteins that are highly divergent within the C-terminus to the M6 Mga are found in nature. These homologous proteins are not only stable but can complement the activity of the protein (7). Since the C-terminus does not appear to be conducive to mutagenesis, other tools must be employed in order to functionally analyze the protein. A high-resolution crystal structure of Mga would serve as an excellent means to overcome this challenge.

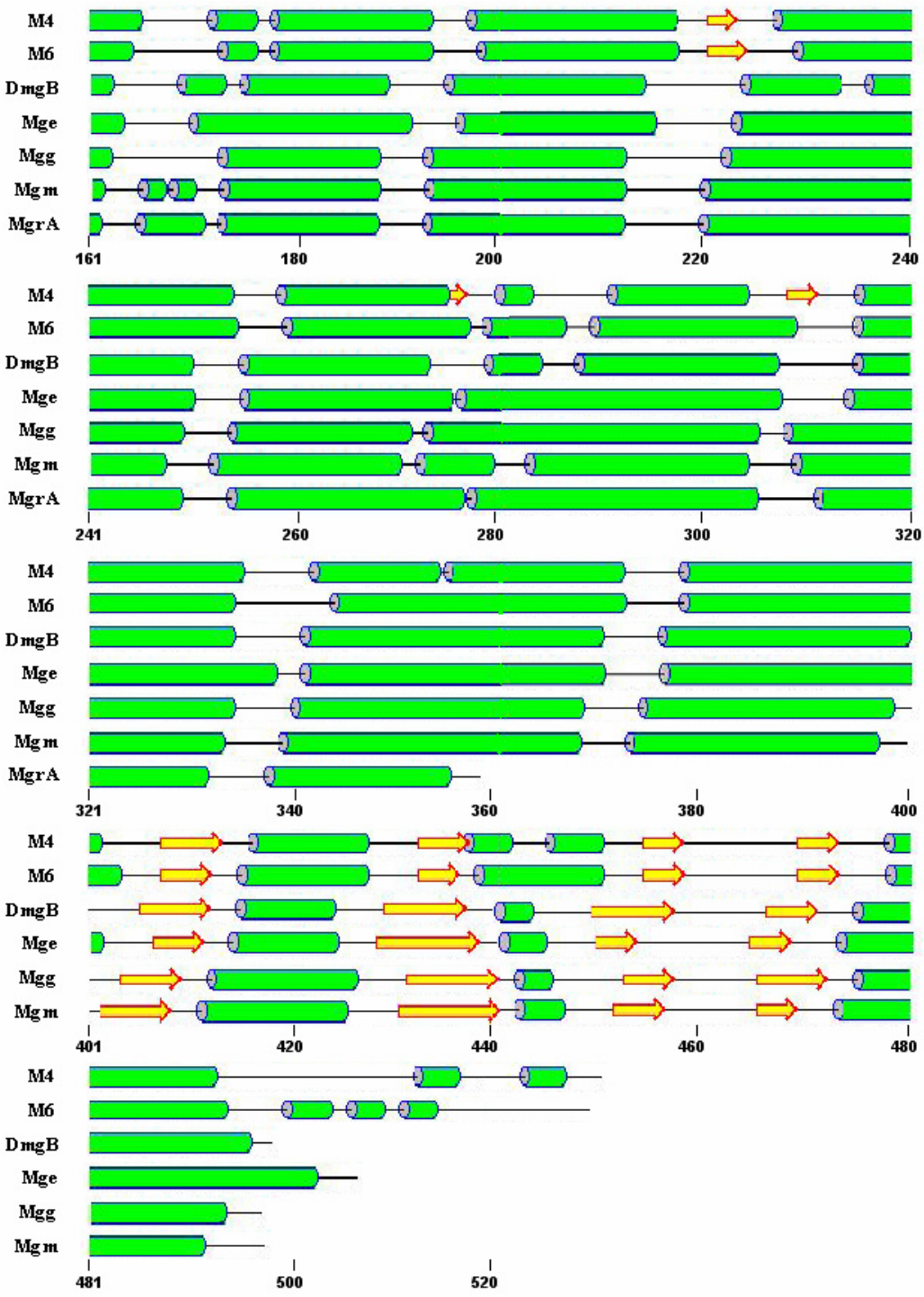
VI. Mga family of virulence gene regulators

Not only are there homologues of Mga but orthologues of this protein can also be found. These orthologues exist in a variety of different pathogenic streptococci, ranging from those that infect primarily humans such as *S. pneumoniae,* to species that infect primarily animals such as *S. dysgalactiae* subsp. *dysgalactiae.* The similarities in secondary structure (Figure 24) amongst the different orthologues are astonishing

Figure 24: Secondary structure alignment of Mga orthologues.

Secondary structures for Mga and its orthologues as predicted by Jpred. Each α-helix is represented by a cylinder, β-sheets by and arrow and random coils by a line. The relative positions compared to the to the primary amino acid sequence is indicated by the numbers below. The first 160 amino acids, otherwise called the N-terminal fragment in this study is presented on the current page while the final amino acids, otherwise called the C-terminal fragment is presented on the following page due to space limitations of the page. The proteins of the different species are identified as follows: M4 (M4 Mga from the GAS), M6 (M6 Mga from the GAS), DmgB (from *S. dysgalactiae* subsp. *dysgalactiae*), Mge (*S. equi*), Mgg (*S. gordonii*), Mgm (*S. mitis*) and MgrA (*S. pneumoniae*).




M4
M6
DmgB
Mge
Mgg
Mgm
MgrA
161
180
200
220
240
241
260
280
300
320
321
340
360
380
400
401
420
440
460
480
481
500
520

considering the different niches that each occupies and the primary amino acid sequence homologies (Table 9). In general, all of the orthologues contain either a α-helix, β-sheet or random coil in approximately the same order and at similar positions within the primary sequence. Because of the strong similarity seen with the secondary structure, it is reasonable to assume that all of these proteins may also share some common tertiary structure folds or domains. Until these proteins are crystallized or another method is used to resolve their structures, these questions will remain largely unanswered.

Not only are these orthologues similar in structure, but they are also similar in the types of genes that they regulate. Of those studied, all but one appears to be a positive transcriptional regulator for virulence genes that appear on the surface of the cell and are important for colonization. The one negative regulator is MgrA from *S. pneumoniae* (see Orthologues section in the Introduction for a more detailed discussion). It does not appear to regulate any genes that are similar to those regulated in the GAS (92). This orthologue is also the smallest member in the family, with a primary amino acid sequence that is more than 100 amino acid residues shorter than any of the others identified (Figure 24). Whether or not the loss of primary sequence has resulted in its unusual function has yet to be determined. Future experimentation utilizing deletion mutants or protein chimeras may be useful in answering these questions.

It is interesting that both Mga and, in most cases, a Mga-regulated virulence gene such as *emm* have been identified in different species considering that in a study of the genome-wide distribution of genes between four different streptococci species, it was

concluded that only a limited number of virulence-associated genes were conserved (159). It was found that most of the conserved genes were housekeeping genes involved with protein synthesis or metabolism. This same study stated that gene positioning does not appear to be highly conserved among the different species, yet *mga* and its orthologues often reside by the genes in which they regulate. Based on this information, it can be hypothesized that *mga* and its corresponding genes were transferred horizontally as a unit. Although it is currently unclear as to which species the genes originated in, there has been a report suggesting that more genes are transferred from the GAS to *S. dysgalactiae* subsp. *dysgalactiae* and *S. equi* subsp. *equisimilis* (112).

VII. Concluding remarks

In conclusion, the results of this study have provided insight into domains within the Mga protein functionally important for regulation. Further exploration of a defective Mga protein pinpointed the exact residues responsible for the loss of transcription that was noted previously in the strain. In addition, a novel domain was uncovered that remains to be biochemically characterized, while a bioinformatic analysis of the protein has provided new avenues for exploration. Many of the tools developed throughout the course of this study will be useful for future investigations of this regulator. Most importantly, this study has established that the results yielded from investigations of the Mga protein can be broadened beyond one specific regulator in the GAS to provide

global insight into possible mechanisms of virulence regulation in other pathogenic streptococci.

Bibliography

1. **Adams, B. B.** 2002. Dermatologic disorders of the athlete. Sports Medicine **32:**309-21.
2. **Akesson, P., A. G. Sjoholm, and L. Bjorck.** 1996. Protein SIC, a novel extracellular protein of *Streptococcus pyogenes* interfering with complement function. Journal of Biological Chemistry **271:**1081-1088.
3. **Alberti, S., C. D. Ashbaugh, and M. R. Wessels.** 1998. Structure of the *has* operon promoter and regulation of hyaluronic acid capsule expression in group A streptococcus. Molecular Microbiology **28:**343-53.
4. **Almengor, A. C., and K. S. McIver.** 2004. Transcriptional activation of *sclA* by Mga requires a distal binding site in *Streptococcus pyogenes*. Journal of Bacteriology **186:**7847-7857.
5. **Almengor, A. C., M. S. Walters, and K. S. McIver.** 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:**2038-47.
6. **Alouf, J. E.** 1980. Streptococcal toxins (streptolysin O, streptolysin S, erythrogenic toxin). Pharmacology and Therapeutics **11:**661-717.
7. **Andersson, G., K. McIver, L. O. Heden, and J. R. Scott.** 1996. Complementation of divergent *mga* genes in group A streptococcus. Gene **175:**77-81.
8. **Appleby, J. L., and R. B. Bourret.** 1999. Activation of CheY mutant D57N by phosphorylation at an alternative site, Ser-56. Molecular Microbiology **34:**915-25.
9. **Beachey, E. H., and I. Ofek.** 1976. Epithelial cell binding of group A streptococci by lipoteichoic acid on fimbriae denuded of M protein. Journal of Experimental Medicine **143:**759-71.
10. **Beckert, S., B. Kreikemeyer, and A. Podbielski.** 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.
11. **Ben Nasr, A. B., H. Herwald, W. Muller-Esterl, and L. Bjorck.** 1995. Human kininogens interact with M protein, a bacterial surface protein and virulence determinant. Biochemical Journal **305 (Pt 1):**173-80.
12. **Berge, A., and U. Sjobring.** 1993. PAM, a novel plasminogen-binding protein from *Streptococcus pyogenes*. Journal of Biological Chemistry **268:**25417-24.
13. **Berggard, K., E. Johnsson, E. Morfeldt, J. Persson, M. Stalhammar-Carlemalm, and G. Lindahl.** 2001. Binding of human C4BP to the hypervariable region of M protein: a molecular mechanism of phagocytosis resistance in *Streptococcus pyogenes*. Molecular Microbiology **42:**539-51.

14. **Bernish, B., and I. van de Rijn.** 1999. Characterization of a two-component system in *Streptococcus pyogenes* which is involved in regulation of hyaluronic acid production. Journal of Biological Chemistry **274:**4786-93.
15. **Bessen, D., K. F. Jones, and V. A. Fischetti.** 1989. Evidence for two distinct classes of streptococcal M protein and their relationship to rheumatic fever. Journal of Experimental Medicine **169:**269-283.
16. **Bessen, D. E., A. Manoharan, F. Luo, J. E. Wertz, and D. A. Robinson.** 2005. Evolution of transcription regulatory genes is linked to niche specialization in the bacterial pathogen *Streptococcus pyogenes*. Journal of Bacteriology **187:**4163-72.
17. **Bessen, D. E., C. M. Sotir, T. L. Readdy, and S. K. Hollingshead.** 1996. Genetic correlates of throat and skin isolates of group A streptococci. Journal of Infectious Diseases **173:**896-900.
18. **Bisno, A. L., M. O. Brito, and C. M. Collins.** 2003. Molecular basis of group A streptococcal virulence. Lancet Infectious Diseases **3:**191-200.
19. **Bisno, A. L., M. A. Gerber, J. M. Gwaltney, Jr., E. L. Kaplan, and R. H. Schwartz.** 2002. Practice guidelines for the diagnosis and management of group A streptococcal pharyngitis. Infectious Diseases Society of America. Clinical Infectious Diseases **35:**113-25.
20. **Bisno, A. L., and K. L. Ruoff.** 2005. Classification of Streptococci, p. 2360-2362. *In* G. L. Mandell, J. E. Bennett, and R. Dolin (ed.), Principles and Practice of Infectious Diseases, Sixth ed, vol. 2. Elsevier Churchill Livingstone, Inc., Philadelphia.
21. **Bisno, A. L., and D. L. Stevens.** 2005. *Streptococcus pyogenes*, p. 2362-2392. *In* G. L. Mandell, J. E. Bennett, and R. Dolin (ed.), Principles and Practice of Infectious Diseases, Sixth ed, vol. 2. Elsevier Churchill Livingstone, Inc., Philadelphia.
22. **Bourret, R. B., K. A. Borkovich, and M. I. Simon.** 1991. Signal transduction pathways involving protein phosphorylation in prokaryotes. Annual Reviews in Biochemistry **60:**401-441.
23. **Bourret, R. B., S. K. Drake, S. A. Chervitz, M. I. Simon, and J. J. Falke.** 1993. Activation of the phosphosignaling protein CheY. II. Analysis of activated mutants by 19F NMR and protein engineering. Journal of Biological Chemistry **268:**13089-96.
24. **Bourret, R. B., J. F. Hess, and M. I. Simon.** 1990. Conserved aspartate residues and phosphorylation in signal transduction by the chemotaxis protein CheY. Proceedings of the National Academy of Sciences of the United States of America **87:**41-5.
25. **Broudy, T. B., V. Pancholi, and V. A. Fischetti.** 2002. The *in vitro* interaction of *Streptococcus pyogenes* with human pharyngeal cells induces a phage-encoded extracellular DNase. Infection and Immunity **70:**2805-11.

26. **Burns, E. H., Jr., A. M. Marciel, and J. M. Musser.** 1996. Activation of a 66-kilodalton human endothelial cell matrix metalloprotease by *Streptococcus pyogenes* extracellular cysteine protease. Infection and Immunity **64:**4744-50.
27. **Burova, L. A., I. V. Koroleva, R. P. Ogurtzov, S. V. Murashov, M. L. Svensson, and C. Schalen.** 1992. Role of streptococcal IgG Fc receptor in tissue deposition of IgG in rabbits immunized with *Streptococcus pyogenes*. APMIS : Acta Pathologica, Microbiologica et Immunologica Scandinavica **100:**567-74.
28. **Burova, L. A., V. A. Nagornev, P. V. Pigarevsky, M. M. Gladilina, V. G. Seliverstova, C. Schalen, and A. A. Totolian.** 1998. Triggering of renal tissue damage in the rabbit by IgG Fc-receptor-positive group A streptococci. APMIS : Acta Pathologica, Microbiologica et Immunologica Scandinavica **106:**277-87.
29. **Canchaya, C., C. Proux, G. Fournous, A. Bruttin, and H. Brussow.** 2003. Prophage genomics. Microbiology Molecular Biology Reviews **67:**238-76.
30. **Caparon, M. G., R. T. Geist, J. Perez-Casal, and J. R. Scott.** 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.
31. **Caparon, M. G., and J. R. Scott.** 1991. Genetic manipulation of pathogenic streptococci. Methods in Enzymology **204:**556-586.
32. **Caparon, M. G., and J. R. Scott.** 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.
33. **Carr, A., D. D. Sledjeski, A. Podbielski, M. D. Boyle, and B. Kreikemeyer.** 2001. Similarities between complement-mediated and streptolysin S-mediated hemolysis. Journal of Biological Chemistry **276:**41790-6.
34. **Catanzaro, F. J., C. A. Stetson, A. J. Morris, R. Chamovitz, C. H. Rammelkamp, Jr., B. L. Stolzer, and W. D. Perry.** 1954. The role of the streptococcus in the pathogenesis of rheumatic fever. American Journal of Medicine **17:**749-56.
35. **Cederholm-Williams, S. A., F. De Cock, H. R. Lijnen, and D. Collen.** 1979. Kinetics of the reactions between streptokinase, plasmin and alpha 2-antiplasmin. European Journal of Biochemistry **100:**125-32.
36. **Centor, R. M., J. M. Witherspoon, H. P. Dalton, C. E. Brody, and K. Link.** 1981. The diagnosis of strep throat in adults in the emergency room. Medical Decision Making: An International Journal of the Society for Medical Decision Making **1:**239-46.
37. **Chatellier, S., N. Ihendyane, R. G. Kansal, F. Khambaty, H. Basma, A. Norrby-Teglund, D. E. Low, A. McGeer, and M. Kotb.** 2000. Genetic relatedness and superantigen expression in group A streptococcus serotype M1 isolates from patients with severe and nonsevere invasive diseases. Infection and Immunity **68:**3523-34.

38. **Chaussee, M. S., R. L. Cole, and J. P. van Putten.** 2000. Streptococcal erythrogenic toxin B abrogates fibronectin-dependent internalization of Streptococcus pyogenes by cultured mammalian cells. Infection and Immunity **68:**3226-32.
39. **Chaussee, M. S., G. L. Sylva, D. E. Sturdevant, L. M. Smoot, M. R. Graham, R. O. Watson, and J. M. Musser.** 2002. Rgg influences the expression of multiple regulatory loci to coregulate virulence factor expression in *Streptococcus pyogenes*. Infection and Immunity **70:**762-70.
40. **Chaussee, M. S., R. O. Watson, J. C. Smoot, and J. M. Musser.** 2001. Identification of Rgg-regulated exoproteins of *Streptococcus pyogenes*. Infection and Immunity **69:**822-31.
41. **Cleary, P. P., U. Prahbu, J. B. Dale, D. E. Wexler, and J. Handley.** 1992. Streptococcal C5a peptidase is a highly specific endopeptidase. Infection and Immunity **60:**5219-23.
42. **Courtney, H. S., J. B. Dale, and D. I. Hasty.** 1996. Differential effects of the streptococcal fibronectin-binding protein, FBP54, on adhesion of group A streptococci to human buccal cells and HEp-2 tissue culture cells. Infection and Immunity **64:**2415-9.
43. **Courtney, H. S., J. B. Dale, and D. L. Hasty.** 2002. Mapping the fibrinogen-binding domain of serum opacity factor of group A streptococci. Current Microbiology **44:**236-40.
44. **Courtney, H. S., D. L. Hasty, Y. Li, H. C. Chiang, J. L. Thacker, and J. B. Dale.** 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.
45. **Courtney, H. S., Y. Li, J. B. Dale, and D. L. Hasty.** 1994. Cloning, sequencing, and expression of a fibronectin/fibrinogen-binding protein from group A streptococci. Infection and Immunity **62:**3937-46.
46. **Courtney, H. S., C. von Hunolstein, J. B. Dale, M. S. Bronze, E. H. Beachey, and D. L. Hasty.** 1992. Lipoteichoic acid and M protein: dual adhesins of group A streptococci. Microbial Pathogenesis **12:**199-208.
47. **Cue, D., P. E. Dombek, H. Lam, and P. P. Cleary.** 1998. *Streptococcus pyogenes* serotype M1 encodes multiple pathways for entry into human epithelial cells. Infection and Immunity **66:**4593-601.
48. **Cuff, J. A., and G. J. Barton.** 2000. Application of multiple sequence alignment profiles to improve protein secondary structure prediction. Proteins **40:**502-11.
49. **Cunningham, M. W.** 2000. Pathogenesis of group A streptococcal infections. Clinical Microbiology Reviews **13:**470-511.
50. **Cywes, C., and M. R. Wessels.** 2001. Group A Streptococcus tissue invasion by CD44-mediated cell signalling. Nature **414:**648-652.

51. **Dale, J. B., R. W. Baird, H. S. Courtney, D. L. Hasty, and M. S. Bronze.** 1994. Passive protection of mice against group A streptococcal pharyngeal infection by lipoteichoic acid. Journal of Infectious Diseases **169:**319-23.
52. **Dale, J. B., and E. H. Beachey.** 1985. Epitopes of streptococcal M proteins shared with cardiac myosin. Journal of Experimental Medicine **162:**583-91.
53. **Dale, J. B., R. G. Washburn, M. B. Marques, and M. R. Wessels.** 1996. Hyaluronate capsule and surface M protein in resistance to opsonization of group A streptococci. Infection and Immunity **64:**1495-501.
54. **Datta, V., S. M. Myskowski, L. A. Kwinn, D. N. Chiem, N. Varki, R. G. Kansal, M. Kotb, and V. Nizet.** 2005. Mutational analysis of the group A streptococcal operon encoding streptolysin S and its virulence role in invasive infection. Molecular Microbiology **56:**681-95.
55. **Derbise, A., Y. P. Song, S. Parikh, V. A. Fischetti, and V. Pancholi.** 2004. Role of the C-terminal lysine residues of streptococcal surface enolase in Glu- and Lys-plasminogen-binding activities of group A streptococci. Infection and Immunity **72:**94-105.
56. **Deutscher, J., E. Kuster, U. Bergstedt, V. Charrier, and W. Hillen.** 1995. Protein kinase-dependent HPr/CcpA interaction links glycolytic activity to carbon catabolite repression in Gram-positive bacteria. Molecular Microbiology **15:**1049-53.
57. **Dinkla, K., M. Rohde, W. T. Jansen, E. L. Kaplan, G. S. Chhatwal, and S. R. Talay.** 2003. Rheumatic fever-associated *Streptococcus pyogenes* isolates aggregate collagen. The Journal of Clinical Investigation **111:**1905-12.
58. **Dole, V. P.** 1946. The properties of T antigens extracted from Group A hemolytic streptococci. Journal of Experimental Medicine **84:**449-471.
59. **Eichenbaum, Z., M. J. Federle, D. Marra, W. M. de Vos, O. P. Kuipers, M. Kleerebezem, and J. R. Scott.** 1998. Use of the lactococcal *nisA* promoter to regulate gene expression in gram-positive bacteria: comparison of induction level and promoter strength. Applied and Environmental Microbiology **64:**2763-2769.
60. **Elliott, S. D.** 1945. A proteolytic enzyme produced by group A streptococci with special reference to its effect on the type-specific M antigen. Journal of Experimental Medicine **81:**573-92.
61. **Elsner, A., B. Kreikemeyer, A. Braun-Kiewnick, B. Spellerberg, B. A. Buttaro, and A. Podbielski.** 2002. Involvement of Lsp, a member of the LraI-lipoprotein family in *Streptococcus pyogenes*, in eukaryotic cell adhesion and internalization. Infection and Immunity **70:**4859-69.
62. **Eriksson, A., B. Eriksson, S. E. Holm, and M. Norgren.** 1999. Streptococcal DNase B is immunologically identical to superantigen SpeF but involves separate domains. Clinical and Diagnostic Laboratory Immunology **6:**133-6.
63. **Facklam, R.** 2002. What happened to the streptococci: overview of taxonomic and nomenclature changes. Clinical Microbiology Reviews **15:**613-30.

64. **Facklam, R. F., D. R. Martin, M. Lovgren, D. R. Johnson, A. Efstratiou, T. A. Thompson, S. Gowan, P. Kriz, G. J. Tyrrell, E. Kaplan, and B. Beall.** 2002. Extension of the Lancefield classification for group A streptococci by addition of 22 new M protein gene sequence types from clinical isolates: *emm103* to *emm124*. Clinical Infectious Diseases **34:**28-38.

65. **Federle, M. J., K. S. McIver, and J. R. Scott.** 1999. A response regulator that represses transcription of several virulence operons in the group A streptococcus. Journal of Bacteriology **181:**3649-3657.

66. **Fehrs, L. J., K. Flanagan, S. Kline, R. R. Facklam, K. Quackenbush, and L. R. Foster.** 1987. Group A beta-hemolytic streptococcal skin infections in a US meat-packing plant. JAMA **258:**3131-4.

67. **Fernie-King, B. A., D. J. Seilly, A. Davies, and P. J. Lachmann.** 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:**4908-16.

68. **Fernie-King, B. A., D. J. Seilly, C. Willers, R. Wurzner, A. Davies, and P. J. Lachmann.** 2001. Streptococcal inhibitor of complement (SIC) inhibits the membrane attack complex by preventing uptake of C567 onto cell membranes. Immunology **103:**390-8.

69. **Ferretti, J. J., W. M. McShan, D. Ajdic, D. J. Savic, G. Savic, K. Lyon, C. Primeaux, S. Sezate, A. N. Suvorov, S. Kenton, H. S. Lai, S. P. Lin, Y. Qian, H. G. Jia, F. Z. Najar, Q. Ren, H. Zhu, L. Song, J. White, X. Yuan, S. W. Clifton, B. A. Roe, and R. McLaughlin.** 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:**4658-63.

70. **Fischetti, V. A.** 1989. Streptococcal M protein: molecular design and biological behavior. Clinical Microbiology Reviews **2:**285-314.

71. **Fischetti, V. A., V. Pancholi, and O. Schneewind.** 1990. Conservation of a hexapeptide sequence in the anchor region of surface proteins from gram-positive cocci. Molecular Microbiology **4:**1603-5.

72. **Frick, I. M., P. Akesson, M. Rasmussen, A. Schmidtchen, and L. Bjorck.** 2003. SIC, a secreted protein of *Streptococcus pyogenes* that inactivates antibacterial peptides. Journal of Biological Chemistry **278:**16561-6.

73. **Fujita, Y., Y. Miwa, A. Galinier, and J. Deutscher.** 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:**953-60.

74. **Gallo, G., R. Berzero, N. Cattai, S. Recchia, and G. Orefici.** 1992. An outbreak of group A food-borne streptococcal pharyngitis. European Journal of Epidemiology **8:**292-7.

75. **Geist, R. T., N. Okada, and M. G. Caparon.** 1993. Analysis of *Streptococcus pyogenes* promoters by using novel Tn916-based shuttle vectors for the

construction of transcriptional fusions to chloramphenicol acetyltransferase. Journal of Bacteriology **175:**7561-7570.

76. **Geyer, A., and K. H. Schmidt.** 2000. Genetic organisation of the M protein region in human isolates of group C and G streptococci: two types of multigene regulator-like (*mgrC*) regions. Molecular & General Genetics **262:**965-76.
77. **Gibson, C., G. Fogg, N. Okada, R. T. Geist, E. Hanski, and M. Caparon.** 1995. Regulation of host cell recognition in *Streptococcus pyogenes*. Developments in Biological Standardization **85:**137-44.
78. **Ginsburg, I.** 1999. Is streptolysin S of group A streptococci a virulence factor? Apmis **107:**1051-9.
79. **Ginsburg, I.** 1972. Mechanisms of cell and tissue injury induced by group A streptococci: relation to poststreptococcal sequelae. The Journal of Infectious Diseases **126:**294-340.
80. **Glisin, V., R. Crkvenjakov, and C. Byus.** 1974. Ribonucleic acid isolated by cesium chloride centrifugation. Biochemistry **13:**2633-7.
81. **Granok, A. B., D. Parsonage, R. P. Ross, and M. G. Caparon.** 2000. The RofA binding site in *Streptococcus pyogenes* is utilized in multiple transcriptional pathways. Journal of Bacteriology **182:**1529-1540.
82. **Gryllos, I., J. C. Levin, and M. R. Wessels.** 2003. The CsrR/CsrS two-component system of group A streptococcus responds to environmental Mg^{2+}. Proceedings of the National Academy of Sciences of the United States of America **100:**4227-32.
83. **Gupte, G., C. Woodward, and V. Stout.** 1997. Isolation and characterization of *rcsB* mutations that affect colanic acid capsule synthesis in *Escherichia coli* K-12. Journal of Bacteriology **179:**4328-35.
84. **Hackett, S. P., and D. L. n. Stevens.** 1993. Superantigens associated with staphylococcal and streptococcal toxic shock syndrome are potent inducers of tumor necrosis factor-beta synthesis. Journal of Infectious Diseases **168:**232-5.
85. **Hahn, R. G., L. M. Knox, and T. A. Forman.** 2005. Evaluation of poststreptococcal illness. American Family Physician **71:**1949-54.
86. **Hallas, G., and J. P. Widdowson.** 1982. The opacity factor of group-A streptococci. Journal of Medical Microbiology **15:**451-64.
87. **Halperin, S. A., P. Ferrieri, E. D. Gray, E. L. Kaplan, and L. W. Wannamaker.** 1987. Antibody response to bacteriophage hyaluronidase in acute glomerulonephritis after group A streptococcal infection. Journal of Infectious Diseases **155:**253-61.
88. **Hanski, E., P. A. Horwitz, and M. G. Caparon.** 1992. Expression of protein F, the fibronectin-binding protein of *Streptococcus pyogenes* JRS4, in heterologous streptococcal and enterococcal strains promotes their adherence to respiratory epithelial cells. Infection and Immunity **60:**5119-25.
89. **Hasty, D. L., I. Ofek, H. S. Courtney, and R. J. Doyle.** 1992. Multiple adhesins of streptococci. Infection and Immunity **60:**2147-52.

90. **Hava, D. L., and A. Camilli.** 2002. Large-scale identification of serotype 4 *Streptococcus pneumoniae* virulence factors. Molecular Microbiology **45:**1389-406.
91. **Heath, A., V. J. DiRita, N. L. Barg, and N. C. Engleberg.** 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.
92. **Hemsley, C., E. Joyce, D. L. Hava, A. Kawale, and A. Camilli.** 2003. MgrA, an orthologue of Mga, Acts as a transcriptional repressor of the genes within the rlrA pathogenicity islet in *Streptococcus pneumoniae*. Journal of Bacteriology **185:**6640-7.
93. **Hirota, K., H. Kanitani, K. Nemoto, T. Ono, and Y. Miyake.** 1995. Cross-reactivity between human sialyl Lewis(x) oligosaccharide and common causative oral bacteria of infective endocarditis. FEMS Immunology and Medical Microbiology **12:**159-64.
94. **Hoe, N. P., R. M. Ireland, F. R. DeLeo, B. B. Gowen, D. W. Dorward, J. M. Voyich, M. Liu, E. H. Burns, Jr., D. M. Culnan, A. Bretscher, and J. M. Musser.** 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:**7646-51.
95. **Hollingshead, S. K., J. Arnold, T. L. Readdy, and D. E. Bessen.** 1994. Molecular evolution of a multigene family in group A streptococci. Molecuar & Biological Evolution **11:**208-219.
96. **Hollingshead, S. K., T. L. Readdy, D. L. Yung, and D. E. Bessen.** 1993. Structural heterogeneity of the *emm* gene cluster in group A streptococci. Molecular Microbiology **8:**707-17.
97. **Hommez, J., L. A. Devriese, J. Henrichsen, and F. Castryck.** 1986. Identification and characterization of *Streptococcus suis*. Veterinary Microbiology **11:**349-55.
98. **Hook, E. W., R. R. Wagner, and R. C. Lancefield.** 1960. An epizootic in Swiss mice caused by a group A Streptococcus, newly designated type 50. American Journal of Hygiene **72:**111-9.
99. **Horstmann, R. D., H. J. Sievertsen, J. Knobloch, and V. A. Fischetti.** 1988. Antiphagocytic activity of streptococcal M protein: selective binding of complement control protein factor H. Proceedings of the National Academy of Sciences of the United States of America **85:**1657-61.
100. **Horstmann, R. D., H. J. Sievertsen, M. Leippe, and V. A. Fischetti.** 1992. Role of fibrinogen in complement inhibition by streptococcal M protein. Infection and Immunity **60:**5036-41.

101. **Hostetler, C. L., K. P. Sawyer, and I. Nachamkin.** 1988. Comparison of three rapid methods for detection of antibodies to streptolysin O and DNase B. Journal of Clinical Microbiology **26:**1406-8.
102. **Hryniewicz, W., and J. Pryjma.** 1977. Effect of streptolysin S on human and mouse T and B lymphocytes. Infection and Immunity **16:**730-3.
103. **Huffman, J. L., and R. G. Brennan.** 2002. Prokaryotic transcription regulators: more than just the helix-turn-helix motif. Current Opinion in Structural Biology **12:**98-106.
104. **Hynes, W. L., and S. L. Walton.** 2000. Hyaluronidases of Gram-positive bacteria. FEMS Microbiology Letters **183:**201-7.
105. **Jadoun, J., V. Ozeri, E. Burstein, E. Skutelsky, E. Hanski, and S. Sela.** 1998. Protein F1 is required for efficient entry of *Streptococcus pyogenes* into epithelial cells. Journal of Infectious Diseases **178:**147-58.
106. **Jaffe, J., S. Natanson-Yaron, M. G. Caparon, and E. Hanski.** 1996. Protein F2, a novel fibronectin-binding protein from *Streptococcus pyogenes*, possesses two binding domains. Molecular Microbiology **21:**373-84.
107. **Jeng, A., V. Sakota, Z. Li, V. Datta, B. Beall, and V. Nizet.** 2003. Molecular genetic analysis of a group A streptococcus operon encoding serum opacity factor and a novel fibronectin-binding protein, SfbX. Journal of Bacteriology **185:**1208-17.
108. **Ji, Y., L. McLandsborough, A. Kondagunta, and P. P. Cleary.** 1996. C5a peptidase alters clearance and trafficking of group A streptococci by infected mice. Infection and Immunity **64:**503-10.
109. **Johnsson, E., K. Berggard, H. Kotarsky, J. Hellwage, P. F. Zipfel, U. Sjobring, and G. Lindahl.** 1998. Role of the hypervariable region in streptococcal M proteins: binding of a human complement inhibitor. Journal of Immunology **161:**4894-901.
110. **Jones, C. H., T. C. Bolken, K. F. Jones, G. O. Zeller, and D. E. Hruby.** 2001. Conserved DegP protease in gram-positive bacteria is essential for thermal and oxidative tolerance and full virulence in *Streptococcus pyogenes*. Infection and Immunity **69:**5538-45.
111. **Jones, K. F., O. Schneewind, J. M. Koomey, and V. A. Fischetti.** 1991. Genetic diversity among the T-protein genes of group A streptococci. Molecular Microbiology **5:**2947-52.
112. **Kalia, A., M. C. Enright, B. G. Spratt, and D. E. Bessen.** 2001. Directional gene movement from human-pathogenic to commensal-like streptococci. Infection and Immunity **69:**4858-69.
113. **Kalia, A., B. G. Spratt, M. C. Enright, and D. E. Bessen.** 2002. Influence of recombination and niche separation on the population genetic structure of the pathogen *Streptococcus pyogenes*. Infection and Immunity **70:**1971-83.
114. **Kaplan, E. L., and B. B. Huew.** 1980. The sensitivity and specificity of an agglutination test for antibodies to streptococcal extracellular antigens: a

quantitative analysis and comparison of the Streptozyme test with the anti-streptolysin O and anti-deoxyribonuclease B tests. The Journal of Pediatrics **96:**367-73.

115. **Kaplan, M. H., and M. Meyeserian.** 1962. An immunological cross-reaction between group-A streptococcal cells and human heart tissue. Lancet **1:**706-10.
116. **Kapur, V., J. T. Maffei, R. S. Greer, L. L. Li, G. J. Adams, and J. M. Musser.** 1994. Vaccination with streptococcal extracellular cysteine protease (interleukin-1 beta convertase) protects mice against challenge with heterologous group A streptococci. Microbial Pathogenesis **16:**443-50.
117. **Kapur, V., M. W. Majesky, L. L. Li, R. A. Black, and J. M. Musser.** 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:**7676-80.
118. **Kapur, V., S. Topouzis, M. W. Majesky, L. L. Li, M. R. Hamrick, R. J. Hamill, J. M. Patti, and J. M. Musser.** 1993. A conserved *Streptococcus pyogenes* extracellular cysteine protease cleaves human fibronectin and degrades vitronectin. Microbial Pathogenesis **15:**327-46.
119. **Kawabata, S., E. Kunitomo, Y. Terao, I. Nakagawa, K. Kikuchi, K. Totsuka, and S. Hamada.** 2001. Systemic and mucosal immunizations with fibronectin-binding protein FBP54 induce protective immune responses against *Streptococcus pyogenes* challenge in mice. Infection and Immunity **69:**924-30.
120. **Kihlberg, B. M., J. Cooney, M. G. Caparon, A. Olsen, and L. Bjorck.** 1995. Biological properties of a *Streptococcus pyogenes* mutant generated by Tn916 insertion in *mga*. Microbial Pathogenesis **19:**299-315.
121. **Klose, K. E., D. S. Weiss, and S. Kustu.** 1993. Glutamate at the site of phosphorylation of nitrogen-regulatory protein NTRC mimics aspartyl-phosphate and activates the protein. Journal of Molecular Biology **232:**67-78.
122. **Kotloff, K. L., M. Corretti, K. Palmer, J. D. Campbell, M. A. Reddish, M. C. Hu, S. S. Wasserman, and J. B. Dale.** 2004. Safety and immunogenicity of a recombinant multivalent group a streptococcal vaccine in healthy adults: phase 1 trial. JAMA **292:**709-15.
123. **Krause, R. M.** 2002. Evolving microbes and re-emerging streptococcal disease. Clinics in Laboratory Medicine **22:**835-48.
124. **Kreikemeyer, B., M. D. Boyle, B. A. Buttaro, M. Heinemann, and A. Podbielski.** 2001. Group A streptococcal growth phase-associated virulence factor regulation by a novel operon (Fas) with homologies to two-component-type regulators requires a small RNA molecule. Molecular Microbiology **39:**392-406.
125. **Kreikemeyer, B., D. R. Martin, and G. S. Chhatwal.** 1999. SfbII protein, a fibronectin binding surface protein of group A streptococci, is a serum opacity factor with high serotype-specific apolipoproteinase activity. FEMS Microbiology Letters **178:**305-11.

126. **Kreikemeyer, B., K. S. McIver, and A. Podbielski.** 2003. Virulence factor regulation and regulatory networks in *Streptococcus pyogenes* and their impact on pathogen-host interactions. Trends in Microbiology **11:**224-32.
127. **Kreikemeyer, B., S. R. Talay, and G. S. Chhatwal.** 1995. Characterization of a novel fibronectin-binding surface protein in group A streptococci. Molecular Microbiology **17:**137-45.
128. **Kurl, D. N.** 1981. Laboratory-acquired human infection with group A type 50 streptococci. Lancet **2:**752.
129. **Lan, C. Y., and M. M. Igo.** 1998. Differential expression of the OmpF and OmpC porin proteins in *Escherichia coli* K-12 depends upon the level of active OmpR. Journal of Bacteriology **180:**171-4.
130. **Lancefield, R. C.** 1928. The antigenic complex of the *Streptococcus haemolyticus*: II. Chemical and immunological properties of the protein fractions. Journal of Experimental Medicine **47:**469-80.
131. **Lancefield, R. C.** 1959. Persistence of type-specific antibodies in man following infection with group A streptococci. Journal of Experimental Medicine **110:**271-92.
132. **Lancefield, R. C.** 1933. A serological differentiation of human and other groups of hemolytic streptococci. Journal of Experimental Medicine **57:**571-95.
133. **Lancefield, R. C.** 1941. Specific relationship of cell composition to biological activity of haemolytic streptococci., vol. 36. Rockefeller Institute for Medical Research.
134. **Lee, M. S., and D. A. Morrison.** 1999. Identification of a new regulator in Streptococcus pneumoniae linking quorum sensing to competence for genetic transformation. Journal of Bacteriology **181:**5004-16.
135. **Lei, B., F. R. DeLeo, N. P. Hoe, M. R. Graham, S. M. Mackie, R. L. Cole, M. Liu, H. R. Hill, D. E. Low, M. J. Federle, J. R. Scott, and J. M. Musser.** 2001. Evasion of human innate and acquired immunity by a bacterial homolog of CD11b that inhibits opsonophagocytosis. Nature Medicine **7:**1298-305.
136. **Levin, J. C., and M. R. Wessels.** 1998. Identification of *csrR*/*csrS*, a genetic locus that regulates hyaluronic acid capsule synthesis in group A streptococcus. Molecular Microbiology **30:**209-219.
137. **Li, Z., D. D. Sledjeski, B. Kreikemeyer, A. Podbielski, and M. D. Boyle.** 1999. Identification of *pel*, a *Streptococcus pyogenes* locus that affects both surface and secreted proteins. Journal of Bacteriology **181:**6019-6027.
138. **Limbago, B., K. S. McIver, V. Penumalli, B. Weinrick, and J. R. Scott.** 2001. Restoration of Mga Function to a *Streptococcus pyogenes* strain (M type 50) that is virulent in mice. Infection and Immunity **69:**1215-1220.
139. **Lindahl, G., and B. Akerstrom.** 1989. Receptor for IgA in group A streptococci: cloning of the gene and characterization of the protein expressed in *Escherichia coli*. Molecular Microbiology **3:**239-47.

140. **Liu, T., and S. Elliott.** 1965. Streptococcal proteainase: the zymogen to enzyme transformation. The Journal of Biological Chemistry **240:**1138-1142.
141. **Lossos, I. S., I. Felsenstein, R. Breuer, and D. Engelhard.** 1992. Food-borne outbreak of group A beta-hemolytic streptococcal pharyngitis. Archives of Internal Medicine **152:**853-5.
142. **Lottenberg, R., D. Minning-Wenz, and M. D. Boyle.** 1994. Capturing host plasmin(ogen): a common mechanism for invasive pathogens? Trends in Microbiology **2:**20-4.
143. **Lukat, G. S., B. H. Lee, J. M. Mottonen, A. M. Stock, and J. B. Stock.** 1991. Roles of the highly conserved aspartate and lysine residues in the response regulator of bacterial chemotaxis. Journal of Biological Chemistry **266:**8348-54.
144. **Lukomski, S., K. Nakashima, I. Abdi, V. J. Cipriano, B. J. Shelvin, E. A. Graviss, and J. M. Musser.** 2001. Identification and characterization of a second extracellular collagen-like protein made by group A streptococcus: control of production at the level of translation. Infection and Immunity **69:**1729-38.
145. **Lyon, W. R., and M. G. Caparon.** 2004. Role for serine protease HtrA (DegP) of *Streptococcus pyogenes* in the biogenesis of virulence factors SpeB and the hemolysin streptolysin S. Infection and Immunity **72:**1618-25.
146. **Madden, J. C., N. Ruiz, and M. Caparon.** 2001. Cytolysin-mediated translocation (CMT): a functional equivalent of type III secretion in Gram-positive bacteria. Cell **104:**143-52.
147. **Martin-Verstraete, I., V. Charrier, J. Stulke, A. Galinier, B. Erni, G. Rapoport, and J. Deutscher.** 1998. Antagonistic effects of dual PTS-catalysed phosphorylation on the *Bacillus subtilis* transcriptional activator LevR. Molecular Microbiology **28:**293-303.
148. **Massell, B. F., L. H. Honikman, and J. Amezcua.** 1969. Rheumatic fever following streptococcal vaccination. Report of three cases. JAMA **207:**1115-9.
149. **McIver, K. S., A. S. Heath, B. D. Green, and J. R. Scott.** 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.
150. **McIver, K. S., A. S. Heath, and J. R. Scott.** 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.
151. **McIver, K. S., and R. L. Myles.** 2002. Two DNA-binding domains of Mga are required for virulence gene activation in the group A streptococcus. Molecular Microbiology **43:**1591-1602.
152. **McIver, K. S., and J. R. Scott.** 1997. Role of *mga* in growth phase regulation of virulence genes of the group A streptococcus. Journal of Bacteriology **179:**5178-5187.

153. **McIver, K. S., S. Subbarao, E. M. Kellner, A. S. Heath, and J. R. Scott.** 1996. Identification of *isp*, a locus encoding an immunogenic secreted protein conserved among group A streptococci. Infection and Immunity **64:**2548-2555.
154. **McIver, K. S., A. S. Thurman, and J. R. Scott.** 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.
155. **McKay, F. C., J. D. McArthur, M. L. Sanderson-Smith, S. Gardam, B. J. Currie, K. S. Sriprakash, P. K. Fagan, R. J. Towers, M. R. Batzloff, G. S. Chhatwal, M. Ranson, and M. J. Walker.** 2004. Plasminogen binding by group A streptococcal isolates from a region of hyperendemicity for streptococcal skin infection and a high incidence of invasive infection. Infection and Immunity **72:**364-70.
156. **McLandsborough, L. A., and P. P. Cleary.** 1995. Insertional inactivation of *virR* in *Streptococcus pyogenes* M49 demonstrates that VirR functions as a positive regulator of ScpA, FcRA, OF, and M protein. FEMS Microbiology Letters **128:**45-51.
157. **McMillan, D. J., and G. S. Chhatwal.** 2005. Prospects for a group A streptococcal vaccine. Current Opin in Molecular Therapeutics **7:**11-6.
158. **Medina, E., G. Molinari, M. Rohde, B. Haase, G. S. Chhatwal, and C. A. Guzman.** 1999. Fc-mediated nonspecific binding between fibronectin-binding protein I of *Streptococcus pyogenes* and human immunoglobulins. The Journal of Immunology **163:**3396-402.
159. **Mitchell, T. J.** 2003. The pathogenesis of streptococcal infections: from tooth decay to meningitis. Nature Reviews. Microbiology **1:**219-30.
160. **Molinari, G., S. R. Talay, P. Valentin-Weigand, M. Rohde, and G. S. Chhatwal.** 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:**1357-63.
161. **Moore, J. B., S. P. Shiau, and L. J. Reitzer.** 1993. Alterations of highly conserved residues in the regulatory domain of nitrogen regulator I (NtrC) of *Escherichia coli*. Journal of Bacteriology **175:**2692-701.
162. **Morfeldt, E., K. Berggard, J. Persson, T. Drakenberg, E. Johnsson, E. Lindahl, S. Linse, and G. Lindahl.** 2001. Isolated hypervariable regions derived from streptococcal M proteins specifically bind human C4b-binding protein: implications for antigenic variation. The Journal of Immunology **167:**3870-7.
163. **Munro, J. M., S. K. Lo, C. Corless, M. J. Robertson, N. C. Lee, R. L. Barnhill, D. S. Weinberg, and M. P. Bevilacqua.** 1992. Expression of sialyl-Lewis X, an E-selectin ligand, in inflammation, immune processes, and lymphoid tissues. The American Journal of Pathology **141:**1397-408.

164. **Neely, M. N., W. R. Lyon, D. L. Runft, and M. Caparon.** 2003. Role of RopB in growth phase expression of the SpeB cysteine protease of *Streptococcus pyogenes*. Journal of Bacteriology **185:**5166-74.
165. **Neeman, R., N. Keller, A. Barzilai, Z. Korenman, and S. Sela.** 1998. Prevalence of internalisation-associated gene, prtF1, among persisting group-A streptococcus strains isolated from asymptomatic carriers. Lancet **352:**1974-7.
166. **Nordstrand, A., M. Norgren, J. J. Ferretti, and S. E. Holm.** 1998. Streptokinase as a mediator of acute post-streptococcal glomerulonephritis in an experimental mouse model. Infection and Immunity **66:**315-21.
167. **Norrby-Teglund, A., R. Kaul, D. E. Low, A. McGeer, D. W. Newton, J. Andersson, U. Andersson, and M. Kotb.** 1996. Plasma from patients with severe invasive group A streptococcal infections treated with normal polyspecific IgG inhibits streptococcal superantigen-induced T cell proliferation and cytokine production. Journal of Immunology **156:**3057-64.
168. **Okada, N., R. T. Geist, and M. G. Caparon.** 1993. Positive transcriptional control of *mry* regulates virulence in the group A streptococcus. Molecular Microbiology **7:**893-903.
169. **Okada, N., A. P. Pentland, P. Falk, and M. G. Caparon.** 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.
170. **Opdyke, J. A., J. R. Scott, and C. P. Moran.** 2001. A secondary RNA polymerase sigma factor from *Streptococcus pyogenes*. Molecular Microbiology **42:**495-502.
171. **Osterlund, A., R. Popa, T. Nikkila, A. Scheynius, and L. Engstrand.** 1997. Intracellular reservoir of *Streptococcus pyogenes in vivo*: a possible explanation for recurrent pharyngotonsillitis. Laryngoscope **107:**640-7.
172. **Pancholi, V., and V. A. Fischetti.** 1998. Alpha-enolase, a novel strong plasmin(ogen) binding protein on the surface of pathogenic streptococci. Journal of Biological Chemistry **273:**14503-15.
173. **Pazour, G. J., C. N. Ta, and A. Das.** 1992. Constitutive mutations of Agrobacterium tumefaciens transcriptional activator *virG*. Journal of Bacteriology **174:**4169-74.
174. **Perez-Casal, J., M. G. Caparon, and J. R. Scott.** 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.
175. **Phillips, G. N., Jr., P. F. Flicker, C. Cohen, B. N. Manjula, and V. A. Fischetti.** 1981. Streptococcal M protein: alpha-helical coiled-coil structure and arrangement on the cell surface. Proceedings of the National Academy of Sciences of the United States of America **78:**4689-93.
176. **Pichichero, M. E.** 1998. Group A beta-hemolytic streptococcal infections. Pediatr Rev **19:**291-302.

177. **Pine, L., and M. W. Reeves.** 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.
178. **Podbielski, A.** 1992. Ubiquitous occurrence of *virR* and *scpA* genes in group A streptococci. Medical Microbiology and Immunology **181:**227-40.
179. **Podbielski, A., A. Flosdorff, and J. Weber-Heynemann.** 1995. The group A streptococcal *virR49* gene controls expression of four structural *vir* regulon genes. Infection and Immunity **63:**9-20.
180. **Podbielski, A., and B. A. Leonard.** 1998. The group A streptococcal dipeptide permease (Dpp) is involved in the uptake of essential amino acids and affects the expression of cysteine protease. Molecular Microbiology **28:**1323-34.
181. **Podbielski, A., J. A. Peterson, and P. Cleary.** 1992. Surface protein-CAT reporter fusions demonstrate differential gene expression in the vir regulon of *Streptococcus pyogenes*. Molecular Microbiology **6:**2253-2265.
182. **Podbielski, A., N. Schnitzler, P. Beyhs, and M. D. Boyle.** 1996. M-related protein (Mrp) contributes to group A streptococcal resistance to phagocytosis by human granulocytes. Molecular Microbiology **19:**429-441.
183. **Podbielski, A., M. Woischnik, B. A. Leonard, and K. H. Schmidt.** 1999. Characterization of *nra*, a global negative regulator gene in group A streptococci. Molecular Microbiology **31:**1051-64.
184. **Podbielski, A., M. Woischnik, B. Pohl, and K. H. Schmidt.** 1996. What is the size of the group A streptococcal *vir* regulon? The Mga regulator affects expression of secreted and surface virulence factors. Medical Microbiology and Immunology (Berl) **185:**171-181.
185. **Rakonjac, J. V., J. C. Robbins, and V. A. Fischetti.** 1995. DNA sequence of the serum opacity factor of group A streptococci: identification of a fibronectin-binding repeat domain. Infection and Immunity **63:**622-31.
186. **Ramachandran, V., J. D. McArthur, C. E. Behm, C. Gutzeit, M. Dowton, P. K. Fagan, R. Towers, B. Currie, K. S. Sriprakash, and M. J. Walker.** 2004. Two distinct genotypes of *prtF2*, encoding a fibronectin binding protein, and evolution of the gene family in *Streptococcus pyogenes*. Journal of Bacteriology **186:**7601-9.
187. **Rasmussen, M., H. P. Muller, and L. Bjorck.** 1999. Protein GRAB of *Streptococcus pyogenes* regulates proteolysis at the bacterial surface by binding α2-macroglobulin. Journal of Biological Chemistry **274:**15336-44.
188. **Rasmussen, R., A. Eden, and L. Bjorck.** 2000. SclA, a novel collagen-like surface protein of *Streptococcus pyogenes*. Infection and Immunity **68:**6370-6377.
189. **Reichardt, W.** 2001. *Streptococcus pyogenes*. Contributions to Microbiology **8:**90-101.
190. **Rhodius, V. A., and S. J. Busby.** 1998. Positive activation of gene expression. Current Opinion in Microbiology **1:**152-9.

191. **Ribardo, D. A., and K. S. McIver.** 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.
192. **Robbins, J. C., J. G. Spanier, S. J. Jones, W. J. Simpson, and P. P. Cleary.** 1987. *Streptococcus pyogenes* type 12 M protein gene regulation by upstream sequences. Journal of Bacteriology **169:**5633-40.
193. **Robinson, V. L., D. R. Buckler, and A. M. Stock.** 2000. A tale of two components: a novel kinase and a regulatory switch. Nature Structural & Molecular Biology **7:**626-33.
194. **Rocha, C. L., and V. A. Fischetti.** 1999. Identification and characterization of a novel fibronectin-binding protein on the surface of group A streptococci. Infection and Immunity **67:**2720-8.
195. **Sadreyev, R., and N. Grishin.** 2003. COMPASS: a tool for comparison of multiple protein alignments with assessment of statistical significance. Journal of Molecular Biology **326:**317-36.
196. **Saile, E., and T. M. Koehler.** 2002. Control of anthrax toxin gene expression by the transition state regulator *abrB*. Journal of Bacteriology **184:**370-80.
197. **Sanders, D. A., B. L. Gillece-Castro, A. M. Stock, A. L. Burlingame, and D. E. Koshland, Jr.** 1989. Identification of the site of phosphorylation of the chemotaxis response regulator protein, CheY. Journal of Biological Chemistry **264:**21770-8.
198. **Schlievert, P. M.** 1993. Role of superantigens in human disease. Journal of Infectious Diseases **167:**997-1002.
199. **Schlievert, P. M.** 1981. Staphylococcal scarlet fever: role of pyrogenic exotoxins. Infection and Immunity **31:**732-6.
200. **Schmidt, K. H., K. Mann, J. Cooney, and W. Kohler.** 1993. Multiple binding of type 3 streptococcal M protein to human fibrinogen, albumin and fibronectin. FEMS Immunology and Medical Microbiology **7:**135-43.
201. **Schrager, H. M., S. Alberti, C. Cywes, G. J. Dougherty, and M. R. Wessels.** 1998. Hyaluronic acid capsule modulates M protein-mediated adherence and acts as a ligand for attachment of group A streptococcus to CD44 on human keratinocytes. Journal of Clinical Investigation **101:**1708-16.
202. **Scott, J. R., P. Cleary, M. G. Caparon, M. Kehoe, L. Heden, J. M. Musser, S. Hollingshead, and A. Podbielski.** 1995. New name for the positive regulator of the M protein of group A streptococcus. Molecular Microbiology **17:**799.
203. **Scott, J. R., P. C. Guenthner, L. M. Malone, and V. A. Fischetti.** 1986. Conversion of an M- group A streptococcus to M+ by transfer of a plasmid containing an M6 gene. Journal of Experimental Medicine **164:**1641-51.
204. **Shatursky, O., A. P. Heuck, L. A. Shepard, J. Rossjohn, M. W. Parker, A. E. Johnson, and R. K. Tweten.** 1999. The mechanism of membrane insertion for a cholesterol-dependent cytolysin: a novel paradigm for pore-forming toxins. Cell **99:**293-9.

205. **Sherman, J. M.** 1937. The Streptococci. Bacteriological Reviews **1:**3-97.
206. **Sieck, J. O., M. Awad, J. Saour, H. Ali, W. Qunibi, and E. Mercer.** 1992. Concurrent post-streptococcal carditis and glomerulonephritis: serial echocardiographic diagnosis and follow-up. European Heart Journal **13:**1720-3.
207. **Simpson, W. J., D. LaPenta, C. Chen, and P. P. Cleary.** 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.
208. **Smith, J. G., J. A. Latiolais, G. P. Guanga, J. D. Pennington, R. E. Silversmith, and R. B. Bourret.** 2004. A search for amino acid substitutions that universally activate response regulators. Molecular Microbiology **51:**887-901.
209. **Smoot, J. C., K. D. Barbian, J. J. Van Gompel, L. M. Smoot, M. S. Chaussee, G. L. Sylva, D. E. Sturdevant, S. M. Ricklefs, S. F. Porcella, L. D. Parkins, S. B. Beres, D. S. Campbell, T. M. Smith, Q. Zhang, V. Kapur, J. A. Daly, L. G. Veasy, and J. M. Musser.** 2002. Genome sequence and comparative microarray analysis of serotype M18 group A streptococcus strains associated with acute rheumatic fever outbreaks. Proceedings of the National Academy of Sciences of the United States of America **99:**4668-73.
210. **Spanier, J. G., S. J. Jones, and P. Cleary.** 1984. Small DNA deletions creating avirulence in *Streptococcus pyogenes*. Science **225:**935-8.
211. **Special Writing Group of the Committee on Rheumatic Fever Endocarditis and Kawasaki Disease of the Council on Cardiovascular Disease in the Young of the American Heart Association.** 1992. Guidelines for the diagnosis of rheumatic fever: Jones Criteria, 1992 update. JAMA **268:**2069-73.
212. **Steiner, K., and H. Malke.** 2002. Dual control of streptokinase and streptolysin S production by the *covRS* and *fasCAX* two-component regulators in *Streptococcus dysgalactiae* subsp. *equisimilis*. Infection and Immunity **70:**3627-36.
213. **Stenberg, L., P. O'Toole, and G. Lindahl.** 1992. Many group A streptococcal strains express two different immunoglobulin-binding proteins, encoded by closely linked genes: characterization of the proteins expressed by four strains of different M-type. Molecular Microbiology **6:**1185-94.
214. **Stephenson, K., and J. A. Hoch.** 2004. Developing inhibitors to selectively target two-component and phosphorelay signal transduction systems of pathogenic microorganisms. Current Medicinal Chemistry **11:**765-73.
215. **Stevens, D. L.** 2000. Group A beta-hemolytic streptococci: virulence factors, pathogenesis, and spectrum of clinical infections, p. 37-56. *In* D. L. Stevens and E. L. Kaplan (ed.), Streptococcal Infections: Clinical Aspects, Microbiology, and Molecular Pathogenesis. Oxford University Press, New York.
216. **Stevens, D. L., M. H. Tanner, J. Winship, R. Swarts, K. M. Ries, P. M. Schlievert, and E. Kaplan.** 1989. Severe group A streptococcal infections associated with a toxic shock-like syndrome and scarlet fever toxin A. The New England Journal of Medicine **321:**1-7.

217. **Studier, F. W., and B. A. Moffatt.** 1986. Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. Journal of Molecular Biology **189:**113-30.

218. **Sumby, P., K. D. Barbian, D. J. Gardner, A. R. Whitney, D. M. Welty, R. D. Long, J. R. Bailey, M. J. Parnell, N. P. Hoe, G. G. Adams, F. R. Deleo, and J. M. Musser.** 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:**1679-84.

219. **Sung, K., S. A. Khan, M. S. Nawaz, and A. A. Khan.** 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.

220. **Terao, Y., S. Kawabata, E. Kunitomo, J. Murakami, I. Nakagawa, and S. Hamada.** 2001. Fba, a novel fibronectin-binding protein from *Streptococcus pyogenes*, promotes bacterial entry into epithelial cells, and the *fba* gene is positively transcribed under the Mga regulator. Molecular Microbiology **42:**75-86.

221. **Terao, Y., S. Kawabata, M. Nakata, I. Nakagawa, and S. Hamada.** 2002. Molecular characterization of a novel fibronectin-binding protein of *Streptococcus pyogenes* strains isolated from toxic shock-like syndrome patients. Journal of Biological Chemistry **277:**47428-35.

222. **Thern, A., L. Stenberg, B. Dahlback, and G. Lindahl.** 1995. Ig-binding surface proteins of Streptococcus pyogenes also bind human C4b-binding protein (C4BP), a regulatory component of the complement system. The Journal of Immunology **154:**375-86.

223. **Thern, A., M. Wastfelt, and G. Lindahl.** 1998. Expression of two different antiphagocytic M proteins by *Streptococcus pyogenes* of the OF+ lineage. Journal of Immunology **160:**860-9.

224. **Tillett, W. S., and R. L. Garner.** 1933. The fibrinolytic activity of hemolytic streptococci. Journal of Experimental Medicine **58:**485-502.

225. **Tobisch, S., J. Stulke, and M. Hecker.** 1999. Regulation of the *lic* operon of *Bacillus subtilis* and characterization of potential phosphorylation sites of the LicR regulator protein by site-directed mutagenesis. Journal of Bacteriology **181:**4995-5003.

226. **Tortosa, P., N. Declerck, H. Dutartre, C. Lindner, J. Deutscher, and D. Le Coq.** 2001. Sites of positive and negative regulation in the *Bacillus subtilis* antiterminators LicT and SacY. Molecular Microbiology **41:**1381-93.

227. **Towers, R. J., P. K. Fagan, S. R. Talay, B. J. Currie, K. S. Sriprakash, M. J. Walker, and G. S. Chhatwal.** 2003. Evolution of *sfbI* encoding streptococcal fibronectin-binding protein I: horizontal genetic transfer and gene mosaic structure. Journal of Clinical Microbiology **41:**5398-406.

228. **Vahling, C. M., and K. S. McIver.** 2006. Domains required for transcriptional activation show conservation in the Mga family of virulence gene regulators. Journal of Bacteriology **188:**863-73.
229. **Vahling, C. M., and K. S. McIver.** 2005. Identification of residues responsible for the defective virulence gene regulator Mga produced by a natural mutant of *Streptococcus pyogenes*. Journal of Bacteriology **187:**5955-66.
230. **Vasi, J., L. Frykberg, L. E. Carlsson, M. Lindberg, and B. Guss.** 2000. M-like proteins of S*treptococcus dysgalactiae*. Infection and Immunity **68:**294-302.
231. **von Pawel-Rammingen, U., B. P. Johansson, and L. Bjorck.** 2002. IdeS, a novel streptococcal cysteine proteinase with unique specificity for immunoglobulin G. The EMBO Journal **21:**1607-15.
232. **Voyich, J. M., D. E. Sturdevant, K. R. Braughton, S. D. Kobayashi, B. Lei, K. Virtaneva, D. W. Dorward, J. M. Musser, and F. R. DeLeo.** 2003. Genome-wide protective response used by group A streptococcus to evade destruction by human polymorphonuclear leukocytes. Proceedings of the National Academy of Sciences of the United States of America **100:**1996-2001.
233. **Wagner, P. L., and M. K. Waldor.** 2002. Bacteriophage control of bacterial virulence. Infection and Immunity **70:**3985-93.
234. **Wannamaker, L. W.** 1970. Differences between streptococcal infections of the throat and of the skin. I. The New England Journal of Medicine **282:**23-31.
235. **Wessels, M. R., A. E. Moses, J. B. Goldberg, and T. J. DiCesare.** 1991. Hyaluronic acid capsule is a virulence factor for mucoid group A streptococci. Proceedings of the National Academy of Sciences of the United States of America **88:**8317-8321.
236. **Wexler, D. E., D. E. Chenoweth, and P. P. Cleary.** 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.
237. **Whatmore, A. M., V. Kapur, D. J. Sullivan, J. M. Musser, and M. A. Kehoe.** 1994. Non-congruent relationships between variation in *emm* gene sequences and the population genetic structure of group A streptococci. Molecular Microbiology **14:**619-31.
238. **Widdowson, J. P., W. R. Maxted, and D. L. Grant.** 1970. The production of opacity in serum by group A streptococci and its relationship with the presence of M antigen. The Journal of General Microbiology **61:**343-53.
239. **Winram, S. B., and R. Lottenberg.** 1996. The plasmin-binding protein Plr of group A streptococci is identified as glyceraldehyde-3-phosphate dehydrogenase. Microbiology **142 (Pt 8):**2311-20.
240. **Yu, C. E., and J. J. Ferretti.** 1991. Frequency of the erythrogenic toxin B and C genes (*speB* and *speC*) among clinical isolates of group A streptococci. Infection and Immunity **59:**211-5.

241. **Yung, D. L., and S. K. Hollingshead.** 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.
242. **Yung, D. L., K. S. McIver, J. R. Scott, and S. K. Hollingshead.** 1999. Attenuated expression of the *mga* virulence regulon in an M serotype 50 mouse-virulent group A streptococcal strain. Infection and Immunity **67:**6691-4.

Vitae

Cheryl M. Vahling, the daughter of Connie and John Vahling was born in Pekin, IL, on May 31, 1977. In 1995, after completing her work at Richwoods High School in Peoria, IL, she attended Illinois Central Community College until 1997. Upon earning her associates degree in the arts and sciences, she continued her education at the University of Illinois in Urbana/Champaign, Illinois. She received the degree of Bachelor of Sciences with a major in biochemistry from the University of Illinois in 2000. She entered the Graduate School of Biomedical Sciences at the University of Texas Health Science Center at Dallas in the fall of 2000 and upon graduation plans to continue working in the field of microbial pathogenesis.

Permanent Address: 2566 E. Meadows Blvd.

Mesquite, TX 75150