Transcriptional Activation of *sclA* by Mga Requires a Distal Binding Site in *Streptococcus pyogenes*

Audry C. Almengor and Kevin S. McIver*

*Department of Microbiology, University of Texas Southwestern Medical Center, Dallas, Texas*

Received 16 July 2004/Accepted 30 August 2004

*A hallmark of the group A streptococcus (GAS; *Streptococcus pyogenes*) is its ability to elicit a wide array of human diseases at varied locations throughout the human body. These syndromes may include relatively minor purulent infections that target either the upper respiratory tract or the skin as well as life-threatening invasive infections of deeper tissues, bloodstream, and organs (6, 30). GAS pathogenesis involves a number of different factors that allow the organism to survive within the human host. The initial stages of infection involve cell-associated factors that contribute to colonization, cell tropism, and immune evasion, including M proteins, capsule, and extracellular matrix binding proteins. Later stages often utilize secreted bacterial factors such as proteases, toxins, and other enzymes that allow the organism to spread and invade deeper tissues (6). In order to coordinate the expression of these different determinants with the changing environments that GAS encounters during an infection, the pathogen has developed multigene networks to regulate virulence.

One such regulator is the multiple virulence gene regulator of GAS known as Mga, which activates the transcription of genes encoding primarily surface-associated proteins involved in GAS pathogenesis. These factors include the M family of proteins (*emm*, *mtp*, *arp*, and *emm*), C5a peptidase (*scpA*), fibronectin-binding proteins (*sof* and *fbx*), and the secreted inhibitor of complement (*sic*) (4, 5, 12, 20). Mga is also required for its own positive regulation or autoactivation (19, 23). Environmental growth conditions such as increased CO$_2$ and temperature have been shown to activate transcription of genes in the Mga regulon (3, 16, 26). Furthermore, expression of the Mga regulon is maximal during exponential phase and becomes transcriptionally silent as the bacteria enter stationary phase (18). This temporal regulation correlates with the requirement for Mga-regulated surface proteins during early stages of infection when GAS is actively growing but down regulates its expression during times of nutrient deprivation when the proteins may actually hinder the bacterium from spreading to new tissues. Thus, Mga controls its set of virulence factors in response to both environmental and temporal signals through an unknown mechanism.

Mga functions by binding directly to specific sequences found in the promoters of target genes, as demonstrated for *emm* (*Pemm*) and *scpA* (*PscpA*) in the serotype M6 GAS strain JRS4 (15). An alignment of these two sites was used as the basis for a published consensus Mga-binding site (15). Binding of Mga in these two promoters occurs at similar locations consisting of a single site overlapping the −35 hexamer and just upstream of the start of transcription (15). Interestingly, autoregulation of *mga* transcription also involves binding of Mga within its own promoter (*Pmga*); however, Mga interacts with two separate binding sites located more than 100 bp upstream of the major P2 start site (19). Thus, based upon the variable location of binding sites in these three promoters, the mechanisms for Mga activation (*Pemm* and *PscpA*) versus autoreactivation (*Pmga*) may be quite different. Although all four of the identified Mga-binding sites are less than 50% identical at the DNA sequence level, recent studies have shown that the same two helix-turn-helix DNA-binding domains of Mga are responsible for recognizing each of these divergent sites (17). Furthermore, the ability of Mga to bind to these promoters...
in vivo was shown to be essential for virulence gene activation (17). Exactly which nucleotides within a given binding site are required for Mga recognition and binding at these various promoters has not been established.

Recently, a gene encoding a cell-wall-anchored streptococcal collagen-like protein (sclA or sclI) was identified in all serotypes of GAS tested (13, 27). A mutant of the newly described surface protein was found to be defective for adherence to lung epithelial cells, but not pharyngeal cells, and was attenuated for virulence in a mouse model of soft-tissue infection (13, 28). The role of Mga in transcriptional regulation of the sclA promoter remains unestablished.

**MATERIALS AND METHODS**

**Bacterial strains and media.** S. pyogenes (GAS) SF370 (31) represents an available serotype M1 GAS genome sequence (8). JRS4 is a streptomycin-resistant derivative of serotype M6 strain D471 (29). JRS519 (mga-10) is an Mga derivative of JRS4 (18). MGAS166 is a streptomycin-resistant M1 strain (22). *Escherichia coli* strain DH5α (Invitrogen) was used as the host for plasmid constructions. The T7 polymerase E. coli strain BL21 (DE3) [F’ompT hsdSB (r5’-m5’) gal dcm met] was used for protein expression and purification.

All E. coli strains were grown in Luria–Bertani broth. GAS was cultured in Todd-Hewitt medium supplemented with 0.2% yeast extract (THY), and growth was assayed by optical density with a Klett-Summmerson photoelectric colorimeter with the A filter. Antibiotics were used at the concentrations indicated: ampicillin at 100 μg/ml for *E. coli*, spectinomycin at 100 μg/ml for both *E. coli* and GAS, kanamycin at 50 μg/ml for *E. coli* and 300 μg/ml for GAS, and erythromycin at 500 μg/ml for *E. coli* and 1.0 μg/ml for GAS.

**DNA manipulations.** Plasmid DNA was isolated from *E. coli* by using either the Wizard Miniprep system (Promega) or Maxi/Mini Prep purification systems (Qiagen). DNA fragments were isolated from agarose gels by using the Wizard Miniprep system or Maxi/Midi Prep purification systems (Qiagen). DNA manipulations were performed using the *E. coli* Genetic Manipulation guide (30). GAS chromosomal DNA was isolated using the FastDNA kit and a FastPrep cell disruptor (Bio 101, Inc.). PCR for cloning and promoter probes was performed using Pfu high-fidelity DNA polymerase (Promega), and reaction mixtures were purified using the QIAquick PCR purification system. PCR for diagnostic assays was performed using Vent DNA polymerase (New England Biolabs). DNA sequencing was performed using the dye-terminator cycle sequencing kit (Applied Biosystems) or the automated sequencing core facility in the Mc Dermott Center at University of Texas Southwestern Medical Center.

**Construction of the Mga** - *M1* strains KS165-L and MGAS166,165-L. To inactivate mga in SF370 and MGAS166, a suicide plasmid was constructed as follows: a 983-bp PCR fragment containing an internal fragment of mga amplified from MGAS166 genomic DNA by using primers OYR-4 and DivMgaL6 (Table 1). The resulting fragment was cloned blunt into a SmallAlgI site containing a stuffer gene, pC2V (23), to produce pKS165-L. Serotype M1 SF370 and MGAS166 were transformed with 200 to 300 μg of pKS165-L, and chromosomal integrations were selected for by growth on THY plates containing kanamycin at 37°C. Mutants were verified by PCR analysis of the mga locus and slot blot analysis of mga transcripts (data not shown).

**Northern blot analysis.** Total RNA was isolated from samples in late exponential phase (65 Klett units) by using the FastRNA kit and a FastPrep cell disruptor as previously described (17). Northern blot assays of total RNA were performed using the NorthernMax kit (Ambion). Briefly, either 1 (emm) or 5 (sclA) μg of total RNA was separated on a 1% agarose denaturing gel (formaldehyde) and transferred to a positively charged nylon membrane. The blots were then UV cross-linked and prehybridized for 30 min at 50°C. Blots were hybridized overnight at 50°C with 5 × 10⁶ cpm of [α-32P]dATP-labeled *emm* or *sclA* probes (RadPrime labeling system; Invitrogen) followed by two low-stringency washes at room temperature (RT) for 5 min and two high-stringency washes at 50°C for 10 min.
5°C for 15 min. Blots were visualized by exposure to a phosphorimaging cassette for 2 h. Probes were PCR amplified from a serotype M1 strain by using the primers listed in Table 1 for *emm* and *sclA*.

**Primer extension analysis.** Total RNA was extracted from samples in late exponential phase as described previously (17) by using the FastRNA kit and a FastPrep cell disruptor. Primer extensions were performed on 25 μg of total RNA as described previously (19) with the primers *P* *emm*-L-R and *sclA*-L-R (Table 1). The primer extension products were run on a 6% denaturing polyacrylamide gel (Amresco), and gels were processed as described above. Sequence was determined using a FLUOSTar Optima (BMG Lab Technology) fluorometer by using an excitation filter at 485 nm and an emission filter at 510 nm. Relative fluorescence units for each sample were expressed as fold above the level of the promoterless control strain containing pKSM410.

**Construction and use of a chromosomal Gusa-based transcriptional reporter plasmid.** A 257-bp fragment of *PscA* containing *Mga*-binding site I was PCR amplified from a serotype M1 strain SF370 by using the *P* *sclA*-L-R fragment pair (Table 1) and ligated into the EcoRV-digested *pBluescript II KS*(−) (Stratagene) to form pKSM417. A 619-bp fragment of *PscA* containing sequence upstream of Mga-binding site II was PCR amplified from SF370 by using the *P* *sclA*-L5-R4 primer pair (Table 1) and cloned into the Smal-digested *pBluescript II KS*(−) to form pKSM418. To form the plasmid pKSM424, an 890-bp fragment of pKSM418 was PCR amplified using the *P* *sclA*-L5-R primer pair (Table 1) and ligated into the Smal-digested *pBluescript II KS*(−) (Table 1). The presence of the site I mutation was confirmed by digestion of this PCR product with *MluI*.

**RESULTS**

Identification of two potential Mga-binding sites upstream of *sclA*. To identify new Mga-regulated virulence genes in GAS, a homology search of the serotype M1 SF370 genome database (AE004092) was performed using the published Mga consensus binding sequence (15). As expected, binding sites were found within promoters of known Mga-regulated genes, including *emm, scpA*, and *scl* (data not shown). A putative Mga-binding site was also identified centered 51 bp upstream of *sclA*.
of the Mga-regulated \( \text{sclA} \) (Fig. 1A; \( \text{P}_{\text{sclA}}\text{-II} \)) and corresponds to a site described previously (13, 27). The site shares 66% identity to the Mga consensus sequence at conserved sites and was located in \( \text{P}_{\text{sclA}}\text{-II} \) overlapping the predicted \(-35\) hexamer of this promoter as described previously (13, 27). \( \text{P}_{\text{sclA}}\text{-II} \) represents a novel site and is located farther upstream. Bold, capital letters and lines represent nucleotide identity to the Mga consensus binding sequence, while lowercase letters represent nonidentity. Asterisks indicate nucleotides in \( \text{P}_{\text{sclA}}\text{-I} \) that are identical to the Mga-binding consensus but are not conserved in \( \text{P}_{\text{sclA}}\text{-II} \). Northern blot analysis of total RNA isolated from M1 GAS strains SF370 (Mga\(^{+}\), lane 1), KSM165-L (Mga\(^{-}\), lane 2), MGAS166 (Mga\(^{+}\), lane 3), and MGAS166.165-L (Mga\(^{-}\), lane 4). Blots were hybridized with probes to \( \text{emm} \) and \( \text{sclA} \) amplified from the PCR primers listed in Table 1.

FIG. 1. Mga regulation of \( \text{sclA} \) in M1 strains. (A) A search of the serotype M1 GAS SF370 genomic sequence upstream of the known Mga-regulated \( \text{sclA} \) gene with the published Mga consensus binding sequence (15) revealed two potential Mga-binding sites. An alignment of \( \text{P}_{\text{sclA}}\text{-I} \) and site II (\( \text{P}_{\text{sclA}}\text{-II} \)) with published consensus Mga-binding sequence is shown. \( \text{P}_{\text{sclA}}\text{-II} \) overlaps the predicted \(-35\) hexamer of this promoter as described previously (13, 27). \( \text{P}_{\text{sclA}}\text{-I} \) represents a novel site that is centered 153 bp upstream from \( \text{P}_{\text{sclA}}\text{-II} \). Scanning of the \( 3' \) sequence adjacent to \( \text{P}_{\text{sclA}}\text{-I} \) did not reveal an obvious second start of transcription, suggesting that Mga bound at this potential element may target the downstream transcriptional start site next to \( \text{P}_{\text{sclA}}\text{-II} \) (Fig. 2B). Furthermore, sequences that are 100% identical to the conserved consensus nucleotides for both sites (Fig. 1A) are found upstream of \( \text{sclA} \) in strains representing a wide array of GAS serotypes, including M2, M3, M6, M12, M18, M28, M49, M52, M56, and M77 (data not shown). Thus, \( \text{P}_{\text{sclA}}\) appears to possess two potential Mga-binding elements arranged in a pattern that is unique to this Mga-regulated promoter and is conserved across GAS serotypes.

Previous studies found that expression of \( \text{sclA} \) is Mga regulated in the M1 strains AP1 (27) and JRS301 (14). In order to confirm that \( \text{sclA} \) is also Mga regulated in the M1 SF370 and MGAS166 strains used in this study, a Northern analysis was performed on wild-type and \( \text{mga} \)-inactivated versions of each strain on RNA isolated at the exponential phase of growth (Fig. 1B). As a control, Northern analysis was also performed on the established Mga-regulated gene \( \text{emm} \) (Fig. 1B). As expected, transcript levels for both \( \text{emm} \) and \( \text{sclA} \) are dramatically lower in the \( \text{mga} \)-inactivated versions of SF370 and MGAS166 than in their wild-type counterparts, confirming Mga regulation of \( \text{sclA} \) in these M1 strains.

Identification of the Mga-regulated start of transcription for \( \text{sclA} \). Since the predicted location of the Mga-regulated start site for \( \text{sclA} \) expression has not been experimentally verified, primer extension analysis was performed on total RNA isolated from both wild-type (SF370) and \( \text{mga} \)-inactivated (KSM165-L) serotype M1 GAS strains by using an antisense primer, \( \text{P}_{\text{sclA}}\text{-R} \) (Table 1), located in the \( 5' \) end of \( \text{sclA} \). An extension product was clearly observed in a background producing Mga (Fig. 2A) that corresponded exactly to the start of transcription predicted in the literature (13, 27). However, the product was clearly absent in the \( \text{mga} \)-inactivated KSM165-L (Fig. 2A). As a control for the primer extension analysis, the transcriptional start for the divergently transcribed \( \text{nrdI} \) gene found directly upstream of \( \text{sclA} \) was mapped in both strains.
tested (Fig. 2A). Interestingly, although the nrdI start site is located within the putative distal PsclA-I Mga-binding element, there was no alteration in nrdI transcription either in the presence or in the absence of Mga. These data indicate that sclA contains a single Mga-regulated start of transcription that is located directly downstream of the potential PsclA-II Mga-binding site (Fig. 2B).

Mga-His purified from E. coli binds to the distal PsclA-I site but not the promoter-proximal PsclA-II site. To determine whether Mga directly interacts with the two potential binding elements within PsclA, EMSAs were performed on overlapping PCR-generated promoter probes encompassing the PsclA region (Fig. 3A). Previous EMSA experiments have successfully utilized an MBP-Mga protein generated from a serotype M6 GAS strain (15, 17, 19); however, this chimera is much larger in size than Mga alone (105 versus 62 kDa) and exhibits only 60% of wild-type Mga activity when expressed in vivo (18). In order to bypass these limitations, a carboxy-terminal fusion of six histidines to the full-length Mga from the M6 GAS strain JRS4 (Mga-His; 63 kDa) was constructed and purified.

FIG. 2. Identification of transcriptional start sites for sclA and nrdI. (A) Primer extension analysis was performed on total RNA isolated from serotype M1 SF370 (Mga+), KSM165-L (Mga−) by using the radiolabeled antisense primers PsclA-R for sclA and nrdI-PE for nrdI (Table 1) as described in Materials and Methods. The start of transcription for sclA and nrdI (asterisks) as well as their corresponding −10 hexamers is shown. (B) Sequence of sclA and nrdI intergenic region from serotype M1 SF370 indicating the identified start sites (asterisk and dark arrow) along with −10 and −35 regions (solid bars). Potential Mga-binding sites (black boxes) and starts of translation (gray arrows) are designated. The sequence is numbered using the PsclA transcriptional start site as +1.
from *E. coli* lysates as described in Materials and Methods. Purified Mga-His showed DNA-binding activity to known binding sites in vitro comparable to that of MBP-Mga and was capable of wild-type transcriptional activation of Mga-regulated genes in vivo (data not shown).

All three promoter probes that contained the distal PsclA-I binding site demonstrated a slower migration consistent with binding when incubated in increasing amounts of the purified Mga-His protein (Fig. 3B; L2-R, L2-R2, and L4-R). In contrast, no detectable binding of Mga-His was observed to a probe that lacked the distal site I but possessed the proximal PsclA-II binding element and the sclA start of transcription (see above) at any of the concentrations tested (Fig. 3B; L-R). Furthermore, no additional binding sites were found in the region located upstream of PsclA-I (Fig. 3B; L2-R4). Thus, Mga appears to show in vitro DNA binding only to the distal PsclA-I binding site situated more than 100 bp upstream of the Mga-regulated start of transcription.

**Specific binding of Mga-His to PsclA-I in vitro.** The specificity of Mga-His binding to the PsclA-I element was investi-
Mga may bind to the proximal P\textsuperscript{sclA}\textsubscript{2} site but with markedly lower affinity than that observed for the distal P\textsuperscript{sclA}\textsubscript{4} site.

Mga-dependent regulation of P\textsuperscript{sclA} is not observed if located on a multicopy plasmid. To investigate the role of Mga binding to the upstream P\textsuperscript{sclA}\textsubscript{4} site in regulation of P\textsuperscript{sclA}, a transcriptional reporter system based on GFP was developed. Use of GFP as a reporter allows direct quantification of promoter activity via levels of fluorescence without the need to lyse the bacteria or add reagents. Unfortunately, attempts to generate single-copy promoter fusions of the strong P\textsuperscript{emm} promoter to gfp in the chromosome of GAS did not allow detection over background levels (data not shown). Therefore, a multicopy GFP transcriptional reporter plasmid called pKSM410 was constructed that is capable of replicating in GAS as well as E. coli (see Materials and Methods).

Promoter fragments corresponding to P\textsuperscript{emm} (pKSM425), the entire P\textsuperscript{sclA}-nrnl intergenic region (pKSM412), and P\textsuperscript{sclA} lacking the distal P\textsuperscript{sclA}\textsubscript{1} Mga-binding site (pKSM414) were cloned in front of the promoterless gfp in pKSM410 (Fig. 5A) and introduced into both wild-type MGAS166 and mga-inactivated MGAS166.165-L serotype M1 GAS. Strains were grown to late logarithmic phase, and the levels of fluorescence were quantified using a fluorometer as described in Materials and Methods. As expected for an Mga-regulated promoter, the control P\textsuperscript{emm} construct pKSM425 showed extremely high expression in MGAS166 but was reduced 7.5-fold in MGAS166.165-L (Fig. 5B; pKSM425). However, the full-length P\textsuperscript{sclA}\textsubscript{4} construct pKSM412, which has both the P\textsuperscript{sclA}\textsubscript{4} binding site and the Mga-regulated start of transcription, demonstrated very little promoter activity over background and was not affected by the presence or absence of Mga (Fig. 5B; pKSM412). Likewise, the transcriptional start site without the distal P\textsuperscript{sclA}\textsubscript{1} binding site showed little change in activity in the presence or absence of Mga (Fig. 5B; pKSM414). Therefore, the presence of the upstream Mga-binding site P\textsuperscript{sclA}\textsubscript{1} in the P\textsuperscript{sclA} promoter was not sufficient to show Mga-regulated gfp expression in vivo and suggests that additional cis sequences not included in this construct are required.

P\textsuperscript{sclA}\textsubscript{1} Mga-binding site is essential for Mga regulation of P\textsuperscript{sclA} at its native chromosomal location. Due to the lack of Mga-dependent regulation of P\textsuperscript{sclA} in a plasmid system, analysis of the promoter at its native chromosomal location was undertaken. Initial attempts to introduce a deletion of the P\textsuperscript{sclA}\textsubscript{1} binding site into the MGAS166 chromosome via allelic replacement of the divergently transcribed nrnl promoter (Fig. 2B; data not shown). Therefore, a merodiploid approach was used to allow the temperature-sensitive integration of single-copy P\textsuperscript{sclA}-gusA reporter fusions into the native chromosomal location while retaining a second wild-type promoter expressing both nrnl and sclA (Fig. 6A). Merodiploid insertion-duplication mutants containing either wild-type P\textsuperscript{sclA} (KSM421) or the P\textsuperscript{sclA}\textsubscript{1} deletion allele (KSM423) fused to a promoterless gusA reporter were constructed in the MGAS166 chromosome via allelic exchange with plasmids. The wild-type KSM421 exhibited strong Mga-regulated GusA activity, whereas the P\textsuperscript{sclA}\textsubscript{1} deletion strain KSM423 showed GusA activity at background levels regardless of the presence or absence of Mga (Fig. 6B). These
results support in vivo the results of the in vitro EMSAs and indicate an essential role for the distal PsclA-I binding site in the Mga-dependent expression of sclA.

**DISCUSSION**

What constitutes an Mga-binding site? Since every Mga-regulated gene described to date encodes a protein involved in GAS pathogenesis, the ability to identify potential Mga-regulated genes through an Mga-binding site in their promoter would be highly advantageous. The published consensus Mga-binding element was constructed using known binding sites (Pemn and PscpA from serotype M6) as well as predicted sites from sequenced promoters corresponding to emn, mpr, arp, enn, and scpA from various other GAS serotypes (15). The resulting 45-bp consensus element contains three regions of conserved nucleotides totaling 29 nucleotides separated by several stretches of irrelevant nucleotides (Fig. 1A). However, this consensus was not able to predict the location of the two Mga-binding sites found in Pmga (19), indicating that Mga can interact with diverse sequences and that any true consensus must reflect these various sites. Since Mga appears to use the same two helix-turn-helix domains of the protein for binding to all of the known promoter sites (17), there likely exists a core consensus binding element that is shared by all sites.

Although the two sites identified within PsclA shared significant identity (72 and 66%, respectively) to the nucleotides in the published consensus, we have shown here that Mga was able to bind only the distal PsclA-I site in vitro (Fig. 2 to 4). In fact, PsclA-I contains only four nucleotides identical to conserved consensus sequences that are lacking in PsclA-II (Fig. 1A, asterisks). Thus, it is likely that one or more of these nucleotides are critical for the in vitro interaction of Mga and provide excellent targets for mutational analysis of Mga-binding sites. Attempts to introduce both Pmga and the PsclA-I binding site into the Mga-binding consensus resulted in a very nonspecific sequence consisting of two short stretches of primarily adenine nucleotides (data not shown). Therefore, future experiments will focus on changing the nonbinding PsclA-II sequence to resemble PsclA-I to allow for a gain-of-function phenotype and identification of those nucleotides that are essential for binding. These types of directed biochemical and mutational analyses of Mga-binding sites should provide a
more reliable way than sequence alignment to establish the essential core Mga-binding sequence.

Interestingly, when Mga is purified from GAS compared to being purified from E. coli, binding to PscI-A is noticeably enhanced (Fig. 4). In addition, Mga purified from E. coli is still sufficient to activate transcription of Mga-regulated promoters in an in vitro transcription assay (A. C. Almengor and K. S. McIver, unpublished results), but not as effectively as Mga purified from GAS. This difference may indicate several possibilities: that an additional factor present only in GAS copurifies with Mga and augments its activity, that a specific modification of Mga occurs in GAS and not E. coli, or simply that the recombinant expression of Mga in a heterologous host is not optimal for its activity (e.g., improper folding). Regardless of the mechanisms involved, these results support the importance of using Mga purified from GAS in future experiments dealing with its activity.

FIG. 6. Analysis of PscI-A activity in vivo with a chromosomal GusA reporter. (A) Construction of an insertion-duplication allele that produces a chromosomal PscI-A fused to a gusA transcriptional reporter while allowing the preservation of the nrdl promoter. Briefly, the PscI-A-gusA transcriptional fusion was cloned into a temperature-sensitive plasmid (pKSM423). Integrants into the chromosome of MGAS166 were isolated after growth at the nonpermissive temperature, which created the strain KSM423. Genes (thick black arrows), gusA (thick gray arrows), starts of transcription (circles with arrows), and wild-type (open boxes) and mutated (gray boxes) Mga-binding sites are indicated. (B) GusA reporter assay on M1 GAS strains MGAS166 (Mga⁺), MGAS166.165-L (Mga⁻), KSM421 (Mga⁺, PscI-A-gusA), KSM421.165-L (Mga⁻, PscI-A-gusA), KSM423 (Mga⁺, PscI-A-Δ site I-gusA), and KSM423.165-L (Mga⁻, PscI-A-Δ site I-gusA). Data are reported in GusA units (OD 420/concentration of total protein [micrograms per microliter]) and represent an average of the results of three independent experiments. The error bars express the standard deviation for each strain measured.
location of Mga binding at these promoters falls into two types. In the case of Pemm and PscpA, Mga binds to a single site centered at −51 bp from the start of transcription and juxtaposed with the −35 hexamer. This places Mga in a position similar to class I transcriptional activators that stabilize RNA polymerase through direct contact with the carboxy-terminal domain of the α subunit and increase initiation (11, 15). In contrast, Prnga contains two Mga-binding sites situated further upstream (−100 bp and −181 bp) from the P2 start of transcription, and both of these sites are required for the autoactivation of the downstream start site (19). Currently, it is not clear by what mechanism Mga is able to activate transcription initiation at such a distance.

In this study, we have shown that Mga binds specifically within the promoter of the Mga-regulated sclA, which encodes a streptococcal collagen-like protein involved in GAS virulence. Although the sclA promoter (PscI) contains two potential binding sites based upon homology to the published consensus sequence, Mga interacts only with a single binding site (PscI-I) located upstream (−168 bp) of the Mga-regulated start of transcription (Fig. 2 to 4). Additionally, the distal PscI-I was found to be necessary for activation of sclA transcription in vivo (Fig. 6). The location of PscI-I is somewhat similar to the most distal Pnga Mga-binding site (−181 bp) and may suggest some commonality of activation between these two promoters. Activators that bind at a distance will often use DNA looping or bending to bring the regulator into physical contact with the transcriptional machinery (2, 10). On the other hand, the fact that PscI uses only a single bound Mga, while autoactivation at Prnga requires two binding sites, suggests that PscI represents a unique type of Mga-regulated promoter. Unpublished work from our laboratory indicates that an Mga-regulated promoter (Psof) from a class II strain of GAS also possesses a single distal Mga-binding site of this type.

Based on the differences observed in the location of Mga-binding sites with respect to the start of transcription in Mga-regulated promoters, we propose that the promoters be grouped into three categories: a single proximal binding site (category A, Pemm and PscpA), a single distal binding site (category B, PscI and Psof), and two distal binding sites (category C, Prnga). At some time in the past, the second putative PscI binding site (PscI-II) may have been fully functional but is no longer utilized to regulate sclA expression. Alternatively, PscI-II may be functional for binding only under very specific temporal or environmental conditions in vivo. Evidence of the possible functionality of PscI-II comes from our observation of barely detectable Mga binding to this site upon significant overexpression of the EMSA, several orders of magnitude below that seen at PscI-I (data not shown). Therefore, PscI may have once represented an Mga-regulated promoter containing both a proximal (category A) and a distal (category B) binding site. However, PscI has evolved such that it no longer requires the proximal site for Mga-specific activation.

Mga regulation of PscI occurs at its native locus but is not observed on a multicopy plasmid. Expression of sclA has been shown previously to be strongly activated at the level of transcription in wild-type M1 GAS strains but not in isogenic strains lacking Mga (Fig. 1B) (14, 27). Our primer extension data confirm these results, showing a strong start of transcription in total RNA isolated from wild-type M1 strain SF370 but not in the mga-inactivated KSM165-L (Fig. 2A). Thus, Mga appears to control sclA by activating transcriptional initiation at PscI. The ability of Mga to bind to the promoter-distal PscI-I suggests strongly that interaction at this site is directly involved in the observed activation of the downstream transcriptional start site. Further, the presence of PscI-I at its native position in the GAS chromosome is essential to the activation of sclA transcription in vivo with a GusA reporter system (Fig. 6). However, when a large (500-bp) region of PscI, including the PscI-I binding site and the downstream start site, was inserted in a transcriptional reporter plasmid based on a promoterless gfp, it did not exhibit detectable Mga-specific activation, nor did it demonstrate significant promoter activity over background levels of the system (Fig. 5). In contrast, a control plasmid containing the category A Pemm-GFP produced very strong Mga-specific activation in the GFP reporter system. There are several possible explanations for why Mga regulation of sclA occurs only in a single copy at its native locus and not on a multicopy plasmid. First, one or more cis-acting elements located either further upstream or within sclA itself may be required for Mga regulation. Alternatively, there may be a gene dosage effect observed in the plasmid system, resulting in the titration of a factor necessary for Mga regulation to occur. Finally, DNA looping may be required to bring Mga in contact with the RNA polymerase, and this event may not be occurring properly in our plasmid system.

Upstream of sclA in the SF370 M1 genome is the divergently transcribed nrdI gene, which is predicted to encode a putative ribonucleotide reductase (Fig. 2B). As an internal control for primer extension of PscI, we mapped the start of transcription for nrdI and found that it was located in the center of the PscI-I Mga-binding site on the antisense strand. Although this might be expected to result in the Mga-specific repression of nrdI transcription, a primer extension product was clearly observed in SF370, an Mga′ strain (Fig. 2A). In fact, the presence of the nrdI promoter in the middle of the distal Mga-binding site may have hindered our attempts to delete PscI-I at its native site in the chromosome (data not shown), suggesting the importance of nrdI in growth under the conditions used in this study. This suggests an even more complex interplay among Mga, PscI, and the overlapping PndII regions that must be further explored.

In order to assess the importance of the distal Mga-binding site of PscI while avoiding the deletion of PndII, single-copy PscI-GusA reporter fusions were integrated into the native chromosomal location such that a second wild-type promoter expressing both nrdI and sclA was preserved in each case. These integrations clearly showed that Mga-regulated activity requires the distal Mga-binding site (Fig. 6). As stated above, several reasons for Mga regulation of sclA only at its native locus can be proposed, including DNA topology and gene dosage from a multicopy plasmid. These possibilities might explain why a category A Mga-regulated promoter (Pemm), which presumably does not require DNA looping, shows significant Mga-specific activity in the plasmid-based GFP reporter whereas a category B PscI does not function in the same way. However, it is clear that the ability of Mga to regulate PscI from a distal binding site is more complex than activation from a proximal site and may require additional factors.

In conclusion, we have demonstrated that Mga binds to a
single distal binding site within the sclA promoter but not to a second putative site located proximal to the Mga-regulated start of transcription. Further, PsclA-I is required for in vivo transcriptional activation of sclA. Notably, an alignment of sequences for PsclA-I found in multiple serotypes of GAS showed no differences in key conserved nucleotides. Therefore, PsclA-I and the role of Mga in its regulation appear to be conserved among most serotypes of GAS. Analysis of this novel Mga-regulated promoter should help build our overall knowledge of the molecular mechanisms by which Mga controls virulence in GAS.

ACKNOWLEDGMENTS

We thank Deborah Ribardo, Cheryl Vahling, and Temekka Leday for critical reading of the manuscript; Steve Lindow for providing the pGreenTIR plasmid; Slawomir Lukomski for unpublished PsclA promoter sequences; and Rhonda Myles for technical assistance.

This work was supported by grants from the American Heart Association, Texas Affiliate (beginning-grant-in-aid no. 0060073Y to K.S.M.) and the National Institutes of Health (NIH AI-47928 to K.S.M.).

REFERENCES