Characterization of the *Streptococcus gordonii* chromosomal region immediately downstream of the glucosyltransferase gene

M. M. Vickerman,1 P. E. Minick2 and N. M. Mather2

Author for correspondence: M. M. Vickerman. Tel: +1 317 278 3250. Fax: +1 317 278 6244. e-mail: mvickerm@iupui.edu

Department of Oral Surgery and Hospital Dentistry, School of Dentistry1,2 and Department of Microbiology and Immunology,1 School of Medicine, Indiana University, Indianapolis, IN 46202, USA

The *Streptococcus gordonii* glucosyltransferase gene, *gtfG*, is positively regulated by the upstream determinant *rgg*. In the present study, two ORFs, transcribed on the opposite DNA strand, were identified immediately downstream of *gtfG*. The first, designated *dsg*, shares a convergent putative transcriptional terminator with *gtfG*, and encodes a predicted 46 kDa transmembrane protein similar to the *Yersinia enterocolitica* TrsA involved in polysaccharide biosynthesis. Insertional inactivation of *dsg* resulted in only ~60% of the parental level of glucosyltransferase activity. The 870 bp gene *gtfG* to *dsg* is similar to the *gtfG* regulatory determinant. Designated *rggD*, this *rgg*-like determinant downstream of *gtfG* encodes a putative 33-kDa cytoplasmic protein. Despite their sequence similarity, the functions of *rgg* and *rggD* appear specific. Strains in which *rggD* was insertionally inactivated and strains containing plasmid-borne *rggD* had parental levels of glucosyltransferase activity. Northern blot hybridization analyses showed ~1.3 kb *dsg*-specific and ~10 kb *rggD*-specific mRNA transcripts associated with this region; no polycistronic transcript was observed. Although *rgg*-like gene products have been demonstrated to function as positive transcriptional regulators of adjacent genes in several streptococcal species, Northern blot analysis suggested that *rggD* did not influence the transcription of *dsg* or the divergent downstream *yblN*-like determinant under the conditions in the present study. Comparison of this *S. gordonii* chromosome region to other streptococcal genomes, which do not contain the *rgg*/*rggD*-flanked region involved in glucan synthesis, raised intriguing possibilities about the origins of this chromosomal region, and also suggested that *rggD* might regulate a distally located gene.

Keywords: oral streptococci, glucans, *rgg*

INTRODUCTION

Glucan polymers synthesized by the extracellular glucosyltransferase (GTF) enzymes of oral streptococci are associated with the ability of bacteria to accumulate on tooth surfaces in dental plaque. Mutans streptococci and *Streptococcus salivarius* have multiple enzymes in each strain that synthesize primarily water-soluble or water-insoluble glucans from sucrose. Although environmental factors have been shown to influence expression of *gtf* genes in these species, no genetic regulatory determinants for their *gtf* genes have been identified. In contrast, the commensal species *Streptococcus gordonii* has a single GTF enzyme that synthesizes both α1,3- and α1,6-linked glucans (Grahame & Mayer, 1984; Haisman & Jenkinson, 1991). Moreover, *S. gordonii* has a positive regulatory determinant, *rgg*, located immediately upstream from the GTF structural gene, *gtfG*, which encodes a trans-acting product that increases *gtfG* transcription (Sulavik et al., 1992). Both *rgg* and *gtfG* have their own putative promoter and ribosome-binding site. Three mRNA transcripts have been associated with the *rgg*/*gtfG* locus: a 1 kb *rgg*-specific transcript, a 5.5 kb *gtfG*-
specific transcript and a 6.5 kb rgg/ftG polycistronic transcript (Sulavik & Clewell, 1996).

At the time rgg was first identified, there were no similar genes in the genetic databases (Sulavik et al., 1992). Since then, determinants with significant similarity to rgg have been identified in a number of related bacterial species. Nucleotide sequence and Southern blot hybridization analyses have identified rgg-like determinants in Streptococcus oralis (Fujisawa et al., 2000) and Streptococcus sanguis (Vickerman et al., 1995), early colonizers of dental plaque (Frandsen et al., 1991) which, like S. gordonii, appear to have only one GTF enzyme; these rgg-like determinants appear to be located near their gtf determinants, suggesting that GTF expression in these species may be regulated in a manner similar to that of S. gordonii (Vickerman et al., 1995). rgg-like determinants have also been characterized in other streptococcal (Lyon et al., 1998; Chaussee et al., 1999; Qi et al., 1999) and lactococcal (Sanders et al., 1998) species, and have been found to regulate a variety of proteins with different functions. These data suggest that rgg-like genes are members of a family of important streptococcal regulatory determinants.

Genetic data suggest that the regulation of S. gordonii GTF activity is complex and may involve genes in addition to rgg and gftG. Both rgg and gftG are preceded by DNA inverted repeats, suggestive of possible regulatory factor binding sites (Sulavik et al., 1992). Two distinct genetic loci that influence GTF activity have been identified by chemical mutagenesis (Haisman & Jenkinson, 1991). Nucleotide sequence analysis has shown that S. gordonii strains with only 20–30% of the parental level of GTF activity have no differences in the 5′-95 kb genome region encoding rgg, gftG and their immediate flanking regions (Vickerman et al., 1997a). Because functionally linked genes are often located in close proximity on bacterial chromosomes, it was hypothesized that genes flanking the rgg/gftG locus could be involved in GTF regulation. However, insertional inactivation of the lemA and htpX genes immediately upstream of rgg did not affect the level of GTF expression (Vickerman et al., 2001). Therefore, the present study was undertaken to examine the S. gordonii chromosomal region downstream of gftG and to determine potential influences of any genes identified in this region on GTF activity.

METHODS

Bacteria, medium, and culture conditions. Bacterial strains and plasmids used in this study are shown in Table 1. All strains were stored at −70 °C in 50% (v/v) glycerol. S. gordonii strains were grown in Todd–Hewitt broth (TH; Difco) or in a defined FMC (Terlecky et al., 1975) medium and incubated at 36 °C in an anaerobic chamber (Coy Industries) with a gas mixture of 85% N2, 10% H2, 5% CO2. The sucrose-promoted colony phenotype (Spe® for parental strain Challis CH1) was determined on 3% (v/v) sucrose TH agar plates incubated for 48 h in 5% CO2. This is a hard colony phenotype which has been associated with glucan synthesis (Tardif et al., 1989). S. gordonii strains carrying replicative or chromosomally integrated plasmids were selected and grown with 5 µg erythromycin ml−1. Escherichia coli DH5α (Invitrogen Life Technologies) strains used to construct pVA891 derivatives for streptococcal integration were grown in Luria–Bertani medium containing 300 µg erythromycin ml−1 and incubated aerobically at 37 °C. Growth phase was determined by measuring OD600 or OD900 for FMC and TH broth cultures, respectively, in a Pharmacia Ultraspec 2000 spectrophotometer.

DNA isolation and manipulations. DNA preparation and manipulations were done according to standard molecular biology procedures (Ausubel et al., 1987). Modifications for S. gordonii included growing strains in 0.5% (w/v) glycine and incubating cells with mutanolysin and lysozyme to facilitate lysis, as described by Sulavik et al. (1992). Plasmid DNA was prepared using Qiagen purification columns according to the manufacturer’s directions. Double-stranded PCR products were obtained using Elongase enzyme (Invitrogen Life Technologies). In some cases, oligonucleotide primers were designed with engineered restriction sites (indicated by underlining: see Results) and random flanking 5′ nucleotides (shown in lower-case letters) to ensure efficient digestion. DNA fragments were eluted from agarose gels with a QiaexII bead kit (Qiagen). For cloning, DNA fragments were digested with appropriate restriction enzymes, ligated with T4 DNA ligase into convenient restriction sites of the vector and transformed into either CaCl2–competent E. coli DH5α or S. gordonii cells made competent (Lawson & Gooder, 1970) with horse serum.

Southern blot hybridization analyses. S. gordonii chromosomal DNA was digested with appropriate restriction enzymes, electrophoresed on 0.7% agarose gels and transferred to a Hybond-N membrane via capillary action under neutral conditions (Ausubel et al., 1987). Probe DNA was labelled with digoxigenin-dUTP, hybridized to membranes and washed under stringent conditions. Hybridized probe was detected by chemiluminescence with the Genius System (Roche Molecular Biochemicals), according to the manufacturer’s directions.

Recovery of S. gordonii chromosomal DNA downstream of gftG. Nucleotide sequencing of the downstream region was originally done using the E. coli plasmid pAM21, which carries the downstream 3·4 kb HindIII fragment of pAM40 subcloned in pBluescript (Vickerman et al., 1997b; Fig. 1a). However, difficulty identifying ORFs raised the possibility that there were difficulties cloning this region. Therefore, using primers derived from the pAM21 sequence, three independent identical PCR products from this region were sequenced directly. The results indicated that deletions had occurred in the ORF immediately downstream of gftG in pAM21. Subsequent attempts to clone these PCR products confirmed that deletions occurred in E. coli cloning vectors. Consequently, the nucleotide sequence of the region downstream of gftG was determined directly from PCR products. At least three overlapping independent products were sequenced for each region.

Inverse PCR (IPCR) was used to recover additional downstream DNA. A Southern blot of the S. gordonii chromosome showed that the ~600 bp Sfi–HindIII fragment, located at the 3′ end of pAM40 (Fig. 1a), hybridized to a ~3 kb BspHI chromosomal fragment. Therefore, chromosomal DNA was digested with BspHI, electrophoresed on 0.7% agarose and fragments ranging from 3 to 4 kb were eluted, re-circularized and used as a template for IPCR. Amplification with primers
**Table 1. Bacterial strains and plasmids**

<table>
<thead>
<tr>
<th>Strain/plasmid</th>
<th>Relevant characteristics</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Bacteria</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CH1</td>
<td><em>S. gordonii</em> parental strain Challis CH1</td>
<td>Tardif et al. (1989)</td>
</tr>
<tr>
<td>CH8942</td>
<td>CH1 with insertional inactivation of <em>dag</em> via chromosomally integrated pMI8942; Em&lt;sup&gt;R&lt;/sup&gt;</td>
<td>This study</td>
</tr>
<tr>
<td>CH8961</td>
<td>CH1 with insertional inactivation of <em>rggD</em> via chromosomally integrated pMI8961; Em&lt;sup&gt;R&lt;/sup&gt;</td>
<td>This study</td>
</tr>
<tr>
<td>DS512</td>
<td>CH1 with premature translational stop in <em>rgg</em>, with ~ 3% of the parental level of GTF activity</td>
<td>Sulavik et al. (1992)</td>
</tr>
<tr>
<td><strong>Plasmids</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pVA749</td>
<td>Streptococcal cloning vector; Em&lt;sup&gt;R&lt;/sup&gt;</td>
<td>Macrina et al. (1981)</td>
</tr>
<tr>
<td>pAMS57</td>
<td><em>rgg</em> cloned into <em>HaeIII</em> site of pVA749</td>
<td>Sulavik et al. (1992)</td>
</tr>
<tr>
<td>pMI226</td>
<td><em>rggD</em> cloned into <em>HaeIII</em> site of pVA749</td>
<td>This study</td>
</tr>
<tr>
<td>pVA891</td>
<td><em>E. coli</em> vector, no streptococcal origin of replication, Em&lt;sup&gt;R&lt;/sup&gt;</td>
<td>Macrina et al. (1983)</td>
</tr>
<tr>
<td>pMI8942</td>
<td>pVA891 with 315 bp internal fragment of <em>dsf</em></td>
<td>This study</td>
</tr>
<tr>
<td>pMI8961</td>
<td>pVA891 with 322 bp internal fragment of <em>rggD</em></td>
<td>This study</td>
</tr>
<tr>
<td>pAMS21</td>
<td>3.4 kb <em>HindIII</em> fragment of pAMS40 cloned into pBluescript SK&lt;sup&gt;+&lt;/sup&gt;; Ap&lt;sup&gt;R&lt;/sup&gt;</td>
<td>Vickerman et al. (1997b)</td>
</tr>
</tbody>
</table>

5'–CTCTCAAGATCTTCAAACG–3' and 5'–GAAAAGCG-ACTTGAGAGC–3' generated a ~ 2.8 kb product (Fig. 1a) which was sequenced directly.

**Nucleotide sequence determination and analysis.** Both strands of the DNA template were sequenced using a PRISM-Ready Reaction Dye Deoxy Terminator Sequencing kit (Applied Biosystems) and an automated DNA sequencer (model 373, Applied Biosystems), using customized oligonucleotide primers. Searches for homologues to *S. gordonii* nucleotide and amino acid sequences were done using the BLAST algorithm (Altschul et al., 1990) on the National Center for Biotechnology Information database (http://www.ncbi.nlm.nih.gov), streptococcal genome databases at the University of Oklahoma (http://genome.ou.edu) and The Institute for Genomic Research database (http://www.tigr.org). Comparisons of entire genes and their encoded proteins were done using MacVector (version 7.0, Oxford Molecular Group) and the GCG package (Wisconsin Package version 10) for nucleotide and protein similarity. Encoded signal peptide and transmembrane helices were predicted with the Signal P (Nielsen et al., 1997) and TMpred (Hofmann & Stoffel, 1993) programs. Potential protein structural motifs were identified via the pfam database (Bateman et al., 2000).

**Insertional inactivation of ORFs.** Oligonucleotide primers with engineered flanking restriction sites and translational stops were used with pAMS40 template to produce internal gene fragments by PCR. The resulting fragments were digested and directionally cloned into the *BamHI* and *HindIII* sites of pVA891 (Macrina et al., 1983). After construction and verification in *E. coli* DH5α, purified plasmid DNA was transformed into strain CH1. Putative *S. gordonii* transformants with the expected plasmid insertion were verified by two methods. Southern blots of digested chromosomal DNA were probed with both pVA891 and the cloned internal gene fragment. Transformant strains that had the expected hybridization patterns were confirmed by direct sequence analysis of PCR products. Primers designed to anneal adjacent to the *BamHI* and *HindIII* sites of the vector (5’–ACGATGCGTCCGGCCTAGAG–3’ and 5’–AGGTCGTGACTTCTAGACTGC–3’, respectively) were used with primers designed to anneal to *S. gordonii* chromosomal sequences flanking the region of each expected plasmid insertion. Resulting PCR products with each transformant template were sequenced to confirm the correct integration of each plasmid for gene disruption.

**Determination of GTF activity.** Relative amounts of GTF activity for each strain were measured via glucan production in acrylamide gels, as described by Tardif et al. (1989). Briefly, strains to be compared were grown to the same mid- to late-exponential stage (*OD<sub>600</sub> ~ 1.6). Cell pellets were extracted with 1% (w/v) SDS (Vickerman & Clewell, 1997) and equal volumes of cell extracts and cell-free culture supernatants were run on an 875% acrylamide SDS-PAGE. After electrophoresis, gels were incubated overnight at 37 °C in a solution of 3% sucrase and 0.5% Triton X-100 in 10 mM sodium phosphate (pH 6.8). The resulting glucan bands were stained with pararosaniline, as described by Tardif et al. (1989). Band intensities reflect the relative amount of GTF protein and activity (Vickerman et al., 1996). Relative GTF activity for each strain was determined via laser densitometry (LKB Ultrascan XL) in at least four independent gels. GTF activity was determined as a percentage of the parental strain activity (set at 100% for the parental strain on each gel).

**Northern hybridization analyses.** *S. gordonii* cells were grown in 0.5% (w/v) glycine medium to the same density. Total RNA was prepared using a Purescript RNA isolation kit (Gentera Systems) essentially to the manufacturer’s directions, but with the following modifications: Cell pellets were resuspended in 161 μl lysozyme ml<sup>–1</sup> and 100 units of mutanolysin in GET buffer (50 mM glucose, 10 mM EDTA, 25 mM Tris/HCl, pH 8.0) and incubated at 37 °C for 10 min. After addition of the SDS solution, cells were incubated on ice for 5 min, followed by 30 s at 65 °C for lysis. Preparations were further purified with two phenol/chloroform/isoamyl alcohol extractions and an ethanol precipitation. Equal amounts of RNA were electrophoresed on a 2% formaldehyde/1% agarose gel, transferred to Hybond-N membranes, probed with digoxigenin-labelled DNA fragments and washed under stringent conditions, according to the instructions for the Genius system (Roche Biochemicals).
RESULTS
Identification and preliminary characterization of dsg

Immediately downstream of gtfG is a 1236 bp ORF on the opposite reading strand that shares the gtfG putative transcriptional terminator (Fig. 1a). This determinant, designated dsg, encodes a putative 46 kDa transmembrane protein. Although dsg does not have highly significant similarity to genes of known function in genomic databases, dsg is 51% similar to a Lactococcus lactis plasmid-borne gene encoding a hypothetical protein of unknown function. Although no other high-probability homologues were found in GenBank, the encoded Dsg protein is 44% similar and shares structural similarities with the protein encoded by the trsA gene of Yersinia enterocolitica (Fig. 2; Skurnik et al., 1995). TrsA is a 47 kDa lipopolysaccharide O-side chain transferase involved in LPS biosynthesis and transport in the Gram-negative cell wall. The trsA gene is a member of the wzx (formerly rfbX; Reeves et al., 1996) family of polysaccharide biosynthetic genes. Difficulties in cloning members of this gene family due to deletions, as were seen with dsg, have been noted (Schnaitman & Klena, 1993). The sequence similarity suggests that dsg may serve a related function involving extracellular polysaccharides and transport in the Gram-positive S. gordonii.

To examine potential functions of dsg, S. gordonii strains were constructed in which dsg was insertionally

![Fig. 1.](image-url)
**Fig. 2.** Comparison of the amino acids of the encoded *S. gordonii* Dsg protein with the *Y. enterocolitica* 47 kDa TrsA protein, using the GCG gap program with the Blosum62 scoring matrix. The proteins are 44.25% similar. Identical amino acids are designated with a vertical line. Similar amino acids with a scoring matrix greater than 2 are represented by a double dot; those with a value of 1 are represented by a single dot. Computer analysis programs predict that both proteins have similar signal sequences and 11 transmembrane regions with a NH2-terminus orientation. Predicted pI values are 9.77 and 9.52 for Dsg and TrsA, respectively.

**Fig. 3.** GTF activities of *S. gordonii* strains. Bacteria were grown in FMC medium to the same cell density and equal volumes of culture supernatants were run on an 8.75% acrylamide SDS-PAGE. Gels were incubated with sucrose and Triton X-100 detergent at 37°C. Synthesized glucan bands were treated with periodic acid and stained with pararosaniline. Extracellular GTF from culture supernatants is shown in each gel. Results for cell-associated GTF activity, which made up less than 10% of the total GTF activity for each strain, were proportional to the extracellular activity for each strain (data not shown). The position of the 174 kDa native protein band is indicated. The lower molecular mass forms of GTF, with various levels of GTF activity, are thought to be due to the degradation of the native enzyme by endogenous proteases and occur even in the presence of protease inhibitors, as described by Grahame & Mayer (1984) and Vickerman & Clewell (1997). Strain names are shown above each lane. Gels shown are representative of a minimum of four independent experiments. (a) Effects of insertional inactivation of *dsg* and *rggD*. Comparison of strains CH1 and CH8961 indicates that insertional inactivation of *rggD* did not affect the level of GTF activity, whereas inactivation of *dsg* in strain CH8942 resulted in a decrease to about 60% of the parental level. (b) Effects of plasmid-borne *rgg* and *rggD*. Although *rgg* in pAMS57 increased the GTF activity of strain CH1 and the *rgg*-mutant strain DS512, *rggD* in pMI226 did not affect the level of GTF activity in either strain.

Disruption of *dsg* influences GTF activity

Colonies of strain CH8942 had the hard, cohesive Spp+ colony phenotype on sucrose agar plates, indicating that they synthesized glucan polymers. However, the colonies were macroscopically more glassy and translucent than parental strain CH1 Spp+ colonies, suggesting some difference in the GTF activities of these two strains. Scanning of GTF activity gels indicated that strain CH8942 only had 57.4 ± 8.2% of the extracellular GTF activity of strain CH1 (Fig. 3a). Similar results were seen for cell-associated GTF activity: strain CH8942 had only 53 ± 8.4% of the parental level of cell-associated GTF activity. However, Northern blots

inactivated. Oligonucleotide primers 5′-taGGATCC-TAGAGTTCGCTTAGGCA-3′ and 5′-tagAAAGCTTGGGTGCTGCTTAGC-3′ were used to amplify an internal region of *dsg* by PCR. For each primer, underlined bases indicate engineered restriction sites; lower-case letters indicate random flanking 5′ nucleotides; the bold T in the second primer was added to create a translational stop site in the integrated plasmid to avoid possible read-through from vector sequences. The 315 bp product was cloned into pVA891 and the resulting plasmid, pMIS942, was transformed into strain CH1. The representative confirmed transformant strain CH8942, in which *dsg* was insertionally inactivated, was further characterized.
Fig. 4. Identification of S. gordonii transcripts by Northern blot analysis. Equal amounts of total RNA from S. gordonii strains (listed above each lane) were electrophoresed on 2% formaldehyde-1% agarose gels, transferred to nylon membranes and hybridized with digoxigenin-labelled probes for internal regions of gtfG, rggD, dsg and ylbN-like genes (represented by hatched bars in Fig. 1a). (a) gtfG-probed RNA of strains CH1, CH8942 and CH8961 indicates that the levels of the 5.4 kb gtfG-specific and the 6.4 kb rggD-gtfG polycistronic transcripts are similar in all three strains. The lower levels of polycistronic transcript compared with gtfG-specific transcript are consistent with previously published results (Sulavik & Clewell, 1996). (b) rggD-probed RNA of strains CH1, CH1(pMI226), CH1(pVA749) demonstrates the ~1 kb rggD transcript. There is an increase in the quantity and size of transcripts in the strain carrying plasmid-borne rggD, which may be due to readthrough in the multicopy pVA749 vector. This has been noted for pAMS57, by Sulavik & Clewell (1996). (c) dsg-probed RNA of strains CH1, CH1(pMI226) and CH8961 shows similar amounts of the ~1.3 kb dsg-specific transcript. (d) ylbN-like-gene-probed RNA of strains CH1, CH1(pMI226) and CH8961 shows similar amounts of a ~2.4 kb transcript. Blots shown are from mid-exponential-phase cultures grown in FMC medium, and are representative of a minimum of three independent experiments. Similar results were seen when bacteria were grown in Todd–Hewitt broth and when cultures were grown to late-exponential to stationary phase (data not shown). RNA standards were run as indicated, with sizes in kb shown on the right.

Identification of an rgg-like determinant downstream of gtfG

On the same reading strand, dsg is immediately preceded by a DNA inverted repeat and an 867 bp ORF which is 42% identical to the gtfG positive regulatory determinant, rgg (Fig. 1a). This rgg-like determinant downstream of gtfG, designated rggD, encodes a putative 33.6 kDa cytoplasmic protein with a pl of 5.43. Like Rgg, the encoded RggD has a helix–turn–helix motif (indicated by italics for amino acids 11–64 for Rgg and amino acids 8–61 for RggD in Fig. 5), suggesting a potential DNA-binding function (Wintjens & Rooman, 1996). In addition to its similarity to S. gordonii Rgg, the encoded RggD shares conserved amino acids with characterized Rgg-like proteins from other species (Fig. 5), which may provide functional insights into specific domains of these regulatory proteins.

To examine potential functional roles of rggD, strain CH8961, in which rggD was insertionally inactivated, was constructed. Primer 5'-taGGATCCTATCCACACAGAAAACGAAG-3', with an engineered BamHI site, and primer 5'-AACCGGAAGTGCAGATAGA-3', designed to anneal 22 bp upstream of the HindIII site at the 3' end of pAMS40 (Fig. 1a), were used to amplify an internal fragment of rggD. The PCR product was digested with restriction enzymes and the 322 bp fragment was cloned into the BamHI and HindIII sites of pVA891 to create pMI8961. Strain CH8961 resulted from the integration of pMI8961 into the strain CH1 chromosome. Despite the similarity of rgg and rggD, there was no cross-reactivity of the internal gene fragments used as probes of S. gordonii chromosomal DNA under stringent conditions in Southern blot hybridization analyses (data not shown).

To compare trans effects of rggD to those of rgg in S. gordonii, the rggD, its putative promoter and its transcriptional termination regions were amplified by PCR using primers 5'-tataGGATCCTATCCACACAGAAAACGAAG-3' and 5'-tataGGATCCTATCCACACAGAAAACGAAG-3'. The resulting 1397 bp EcoRI-flanked fragment was cloned into the compatible HaeIII site of the streptococcal plasmid pVA749. Nucleotide sequence analysis confirmed that the orientation of the rggD fragment in the resulting plasmid, pMI226, was the same as that of rgg in the plasmid pAMS57 previously characterized by Sulavik et al. (1992).

rggd does not affect GTF activity

Although their nucleotide and deduced amino acid sequences were similar, the specificities of rgg and rggD appeared to differ. Northern blots indicated that disruption of rggD in strain CH891 did not affect levels of rgg/gtfG or gtfG transcript (Fig. 4a). The GTF activities of strains CH1 and CH8961 were similar (Fig. 3a), indicating that disruption of rggD did not affect potential post-transcriptional modifications that would result in altered levels of GTF activity. Effects of plasmid-borne rgg and rggD were also specific. Plasmid-borne rgg in pAMS57 has been shown to increase GTF activity approximately sixfold in the parental strain CH1; pAMS57 has also been shown to increase GTF activity to the same level in strain DS512 (Sulavik et al., 1992). Due to a frameshift mutation resulting in a premature translation stop in rgg, strain DS512 has only

Downloaded from www.microbiologyresearch.org by IP: 54.70.40.11
Genes downstream of \( S. gordonii \) gtfG

Fig. 5. Comparison of deduced amino acid sequences of \( S. gordonii \) Rgg and RggD to other characterized Rgg-like proteins encoded by \( S. mutans \) mutR (Qi et al., 1999), \( S. pyogenes \) ropB (Lyon et al., 1998) and \( L. lactis \) gadR (Sanders et al., 1998) using CLUSTAL analysis (Thompson et al., 1994). Deduced amino acids of \( S. gordonii \) Rgg and RggD in the putative helix–turn–helix motif (Bateman et al., 2000), which may function in DNA binding (Wintjens & Rooman, 1996), are shown in italics. Conserved regions of the similar proteins are shown in shaded blocks.

Effects of \( rggD \) on transcription of flanking genes

To determine if \( rggD \) and \( dsg \) were co-transcribed, strain CH1 RNA was probed with internal fragments of \( rggD \) and \( dsg \) in Northern blots (Fig. 4b, c). A \( \sim 1.0 \) kb \( rggD \)-specific transcript and a \( \sim 1.3 \) kb \( dsg \)-specific transcript were detected. No polyclstronic transcript was noted.

Previously characterized \( rgg \)-like genes positively regulate transcription of adjacent genes (Sulavik & Clewell, 1996; Lyon et al., 1998; Sanders et al., 1998; Qi et al., 1999; Chaussee et al., 1999). Furthermore, \( rgg \)-like genes are often separated from the gene they regulate by DNA secondary structures, such as inverted repeats. Accordingly, the inverted repeats between \( rggD \) and \( dsg \) suggested the possibility that \( rggD \) might affect \( dsg \) transcription. However, in \( Streptococcus pyogenes \), the \( rgg \)-like gene that regulates the gene for streptococcal erythrogenic toxin B, speB, is transcribed in the divergent reading direction (Lyon et al., 1998). Therefore, the possibility that \( rggD \) regulates the gene further downstream was also investigated. Nucleotide sequencing identified a 531 bp ORF downstream of \( rggD \), in the divergent reading direction (Fig. 1a). This gene encodes a predicted 19–9 kDa cytoplasmic protein with a pI of 3–89. Examination of genome databases indicated that this gene is most similar to the \( ylbN \)-like gene of unknown function of \( Lactococcus lactis \) subsp. cremoris. The \( S. gordonii ylbN \)-like gene is associated with a \( \sim 2.4 \) kb transcript (Fig. 4d), suggesting that it is co-transcribed with determinants further downstream. Indeed, partial sequence of an ORF, designated \( orf8 \) in Fig. 1(a), showed a putative ribosome-binding site for translation of the encoded protein, but no apparent transcriptional termination sequences were noted in the region between the \( ylbN \)-like gene and \( orf8 \).

Northern blot analyses were used to examine the transcription of \( dsg \) and \( ylbN \)-like genes in strain CH8961 (in which \( rggD \) was insertionally inactivated) and in a parental strain carrying additional copies of \( rggD \) in pMI226 (Fig. 4b, c). Unexpectedly, the results did not indicate any differences in transcription of either \( dsg \) or \( ylbN \) in these strains compared with the parental strain CH1. Recent studies in \( S. pyogenes \) have indicated

\( \sim 3\% \) of the parental level of GTF activity (Sulavik & Clewell, 1996). However, plasmid-borne \( rggD \) did not increase GTF activity in strains CH1 or DS512 to a level above that of the plasmid-free strain (Fig. 3b). Thus, \( rggD \) in trans did not affect GTF activity, even in the absence of a functional chromosomal \( rgg \). These results suggest that \( rggD \) does not influence \( gtfG \) expression and may provide important insights into the specificity of \( rgg \)-like genes.
that an rgg-like determinant influences transcription in stationary-phase cells (Chaussée et al., 2001); hence the S. gordonii strains were also examined in late-exponential to stationary growth phase. However, as seen in the mid-exponential-phase cells, no differences were evident (data not shown). Similar results were also seen when the S. gordonii strains were grown in defined FMC medium and in complex Todd–Hewitt broth, indicating that components of these different media did not influence potential regulation of either gene by rggD. These results do not preclude the possibility that rggD affects transcription of the dsg or ylbN-like genes under conditions other than those in the present study.

Comparison of streptococcal chromosomal regions

The finding that rggD did not influence transcription of the adjacent genes raised the possibility that rggD regulates distally located genes on the S. gordonii chromosome. This hypothesis is supported by the comparison of similar chromosomal regions using available genome data for other streptococcal species (Fig. 1b). S. pyogenes has htxX- and ylbN-like genes adjacent to each other. However, S. gordonii has four genes between the htxX and ylbN-like genes, i.e. two convergent genes, gtfG and dsg, which are each preceded by similar rgg-like genes. Both Streptococcus mutans and Streptococcus pneumoniae have hypothetical ATP-binding proteins with no rgg-like determinants on the opposite reading strand between their htxX- and ylbN-like genes. These hypothetical genes, encoding ATP-binding proteins, do not have significant similarities to the S. gordonii rgg-flanked gtfG and dsg genes. However, there is a dsg-like gene (37\% identity at the nucleotide level; 57\% similarity of the encoded putative proteins) distally located on the S. pneumoniae chromosome between putative histidine RNA ligase and dihydroxy-acid dehydratase genes. These findings suggest that the S. gordonii determinants between the htxX and ylbN-like genes may have resulted from recombinations between similar DNA regions within the S. gordonii chromosome, or via horizontal gene transfer. It is possible that such a recombination or rearrangement resulted in distancing rggD from the S. gordonii gene which it regulates.

DISCUSSION

Examination of the region downstream of gtfG led to the identification of a gene, designated dsg, which affects the level of GTF activity. The similar amounts of gtfG seen in Northern blots, together with the sequence and structural similarities of the encoded Dsg to a protein involved in polysaccharide biosynthesis in Gram-negative organisms (Skurnik et al., 1995), support the possibility that dsg may be involved in some post-transcriptional and/or post-translational modifications that affect the level of GTF. Insertional inactivation of dsg resulted in a decrease of approximately 40\% in GTF activity. Although colonies of these dsg mutants were hard and cohesive Spp\(^+\) on sucrose agar plates, they also had a glassy, more translucent appearance than parental colonies. It is not known if this phenotypic change is due to the approximately 60\% level of GTF activity, or due to changes in the glucan polymers themselves. Such changes in synthesized glucans could potentially result from conformational changes of the active enzyme when it is near or attached to the cell surface (Kopec et al., 1997), as might occur if there were differences in the extracellular transport of GTF. Mutations in S. gordonii gtfG can lead to changes in the proportions of \(\alpha_1,3-\) and \(\alpha_1,6\)-glucosidic linkages in the synthesized glucans, and result in changes in the detectable levels of GTF activity: as measured by both [\(\text{glucose}^{14}\text{C}\)]sucrose incorporation into glucan and GTF activity gel assays (Vickerman et al., 1996). Analysis of the glucans produced by colonies of strain CH8942 grown on sucrose agar plates was not done in the present study. Spp\(^+\) strains, which produce soft colonies on sucrose agar plates, have been reported to have less than 30\% the parental level of GTF activity (Tardif et al., 1989). Changes in Spp have also been associated with changes in the amino acid composition of the GTF enzyme (Vickerman et al., 1996). Development of the Spp\(^-\) and changes in GTF activity have been associated with the ability of S. gordonii cells to be released from biofilm surfaces (Vickerman et al., 1991), which may provide important ecological advantages for cells growing in dental plaque. Thus, genes such as rgg, gtfG and dsg, involved in varying the quantity or quality of glucans produced by the GTF enzyme of the commensal S. gordonii, may play important roles in oral microbial ecology.

Directly 5\’ to dsg, an rgg-like determinant, rggD, was identified: this is the third rgg-like determinant reported in S. gordonii. In addition to the originally identified positive gtfG regulator, rgg (Sulavik et al., 1992), an rgg-like determinant, ivb, was recovered from a rabbit endocarditis model using in vivo expression technology (Kilic et al., 1999); this suggests that ivb, or possibly a gene(s) that it regulates, is required for survival in vivo. Multiple rgg-like genes are present in S. pyogenes (Ferretti et al., 2001), S. mutans (http://www.genome.ou.edu) and S. pneumoniae (http://www.tigr.org), suggesting that rgg-like genes make up a relatively widely occurring family of streptococcal regulatory genes. Under the conditions in the present study, the S. gordonii gene that rggD regulates was not identified. Based upon the findings that most rgg-like determinants regulate adjacent genes (Sulavik & Clewell, 1996; Lyon et al., 1998; Sanders et al., 1998; Qi et al., 1999; Chaussee et al., 1999) it was hypothesized that rggD would regulate the S. gordonii dsg or ylbN-like genes. Northern blots did not confirm this hypothesis. However, it is possible that changes in the levels of the dsg and ylbN transcript in strains with disruption or additional copies of rggD were too subtle to be detected by Northern blot hybridization analyses. Nevertheless, the dramatic sixfold differences in gtfG transcription due to rgg (Sulavik & Clewell, 1996), along with the effects of rgg (ropB) on transcription of speB (Lyon et al., 1998), were readily apparent on Northern blots, suggesting that the magnitude of posi-
tive regulation by rgg-like genes should be clearly evident by this assay. It is possible that rggD regulates either of the adjacent genes under conditions other than those in these studies. Although transcripts were examined from mid-exponential through to stationary growth phase in these studies. Although transcripts were examined from mid-exponential through to stationary growth phase in these studies. Although transcripts were examined from mid-exponential through to stationary growth phase in these studies. Although transcripts were examined from mid-exponential through to stationary growth phase in these studies. Although transcripts were examined from mid-exponential through to stationary growth phase in these studies. Although transcripts were examined from mid-exponential through to stationary growth phase in these studies. Although transcripts were examined from mid-exponential through to stationary growth phase in these studies. Although transcripts were examined from mid-exponential through to stationary growth phase in these studies. Although transcripts were examined from mid-exponential through to stationary growth phase in these studies. Although transcripts were examined from mid-exponential through to stationary growth phase in these studies. Although transcripts were examined from mid-exponential through to stationary growth phase in these studies. Although transcripts were examined from mid-exponential through to stationary growth phase in these studies. Although transcripts were examined from mid-exponential through to stationary growth phase in these studies. Although transcripts were examined from mid-exponential through to stationary growth phase in these studies. Although transcripts were examined from mid-exponential through to stationary growth phase in these studies. Although transcripts were examined from mid-exponential through to stationary growth phase in these studies. Although transcripts were examined from mid-exponential through to stationary growth phase in these studies. Although transcripts were examined from mid-exponential through to stationary growth phase in these studies. Although transcripts were examined from mid-exponential through to stationary growth phase in these studies. Although transcripts were examined from mid-exponential through to stationary growth phase in these studies. Although transcripts were examined from mid-exponential through to stationary growth phase in these studies. Although transcripts were examined from mid-exponential through to stationary growth phase in these studies. Although transcripts were examined from mid-exponential through to stationary growth phase in these studies. Although transcripts were examined from mid-exponential through to stationary growth phase in these studies. Although transcripts were examined from mid-exponential through to stationary growth phase in these studies. Although transcripts were examined from mid-exponential through to stationary growth phase in these studies. Although transcripts were examined from mid-exponential through to stationary growth phase in these studies. Although transcripts were examined from mid-exponential through to stationary growth phase in these studies. Although transcripts were examined from mid-exponential through to stationary growth phase in these studies. Although transcripts were examined from mid-exponential through to stationary growth phase in these studies. Although transcripts were examined from mid-exponential through to stationary growth phase in these studies. Although transcripts were examined from mid-exponential through to stationary growth phase in these studies. Although transcripts were examined from mid-exponential through to stationary growth phase in these studies. Although transcripts were examined from mid-exponential through to stationary growth phase in these studies. Although transcripts were examined from mid-exponential through to stationary growth phase in these studies. Although transcripts were examined from mid-exponential through to stationary growth phase in these studies. Although transcripts were examined from mid-exponential through to stationary growth phase in these studies. Although transcripts were examined from mid-exponential through to stationary growth phase in these studies. Although transcripts were examined from mid-exponential through to stationary growth phase in these studies. Although transcripts were examined from mid-exponential through to stationary growth phase in these studies. Although transcripts were examined from mid-exponential through to stationary growth phase in these studies. Although transcripts were examined from mid-exponential through to stationary growth phase in these studies. Although transcripts were examined from mid-exponential through to stationary growth phase in these studies. Although transcripts were examined from mid-exponential through to stationary growth phase in these studies. Although transcripts were examined from mid-exponential through to stationary growth phase in these studies. Although transcripts were examined from mid-exponential through to stationary growth phase in these studies. Although transcripts were examined from mid-exponential through to stationary growth phase in these studies. Although transcripts were examined from mid-exponential through to stationary growth phase in these studies. Although transcript

ACKNOWLEDGEMENTS

This work was supported by Public Health Service grant DE11090 from the National Institutes of Health.

REFERENCES


Received 6 June 2001; revised 6 July 2001; accepted 24 July 2001.