Regulation of the α-Galactosidase Activity in *Streptococcus pneumoniae*: Characterization of the Raffinose Utilization System

Carsten Rosenow,[1] Mita Maniar, and Joaquim Trias

Versicor, Inc., Fremont, California 94555 USA

A 10.2-kb gene region was identified in the *Streptococcus pneumoniae* genome sequence that contains eight genes involved in regulation and metabolism of raffinose. The genes *rafR* and *rafS* are transcribed as one operon, and their gene products regulate the raffinose-dependent stimulation of a divergently transcribed second promoter (P_A) directing the expression of *aga*, the structural gene for α-galactosidase. Raffinose-mediated transcription from P_A results in a 500-fold increase in α-galactosidase activity in the cell. A third promoter within the cluster is responsible for the transcription of the remaining five genes (*rafE*, *rafF*, *rafG*, *gtfA*, and *rafX*), whose gene products might be involved in transport and metabolism of raffinose. The presence of additional internal promoters cannot be excluded. The *aga* promoter P_A is negatively regulated by the presence of sucrose in the growth medium. Consistent with catabolite repression (CR), a DNA sequence with high homology to the CRE (*cis*-active element) was identified upstream of the *aga* promoter. Sucrose-mediated CR depends on the phosphoenolpyruvate: sucrose phosphotransferase system (PTS) but is unaffected by a mutation in a gene encoding a homolog of the CRE regulatory protein CcpA.

Streptococcus pneumoniae is limited in its sugar fermentation ability, being able to use fructose, galactose, sucrose, glucose, raffinose, lactose, inulin, threhalose, and maltose as sources of energy (Holt et al. 1994). The only characterized sugar transport system in *S. pneumoniae* is the maltodextrin (Mal) utilization system (Stassi et al. 1982). The *mal* operon has two negatively regulated promoters, both of which are only activated in the presence of maltose, maltotriose, or maltotetraose (Nieto et al. 1997). Little is known about the mechanism for the utilization of other sugars by *S. pneumoniae*. In contrast, carbohydrate utilization by *Streptococcus mutans* is well studied (Vadeboncoeur and Pelletier 1997). In *S. mutans*, sugar metabolism contributes to the initiation and progression of dental caries and is an important virulence determinant (Cvitkovitch et al. 1995). A well-characterized sugar utilization system in *S. mutans* is the multiple-sugar metabolism operon (*msm*) (Russell et al. 1992). This gene cluster contains eight genes, products of which are involved in the uptake of melibiose, raffinose, and isomaltotriose and the metabolism of melibiose, sucrose, and isomaltosaccharides. The cluster also contains the gene for a regulatory protein that acts as a positive effector.

The phosphoenolpyruvate: sugar phosphotransferase system (PTS) is one of the preferred carbohydrate uptake systems in bacteria. The PTS is a complex enzyme system that is responsible for the detection, transmembrane transport, and phosphorylation of numerous sugar substrates in both Gram-negative and Gram-positive prokaryotes (Postma and Lengeler 1985). The PTS in bacteria is involved in the regulation of other, non-PTS carbohydrate uptake systems. This regulation is triggered by the presence of a preferred carbon source in the growth medium that inhibits the expression of proteins involved in the utilization of alternative carbon sources (catabolite repression) (Postma et al. 1993; Saier 1996). Catabolite repression in Gram-positive organisms can be controlled by the binding of the catabolite control protein (CcpA) to CREs (*cis*-acting element), a 14-bp region of dyad symmetry that has been identified upstream of genes under CcpA regulation (Hueck et al. 1994; Saier et al. 1996). Such regulatory mechanisms provide bacteria with the ability to select their preferred carbon and energy source from the environment by switching off genes that code for transport and metabolizing enzymes of other PTS and non-PTS systems. In many Gram-positive bacteria, glucose has been shown to repress synthesis of both PTS and non-PTS carbohydrate catabolic enzymes. This function of glucose in catabolite repression is linked to its role as the preferred carbon source in many bacteria (Saier et al. 1996).

We have identified in the *S. pneumoniae* genome a gene cluster that is involved in raffinose metabolism. Some of the genes show homology to genes of the *S. mutans msm* gene cluster. The *S. pneumoniae* gene cluster includes genes encoding α-galactosidase (*aga*), an activator (*rafR*) and repressor (*rafS*) for *aga* expression, and genes whose products are homologous to sugar transport systems in other prokaryotes. The expression

[1]Corresponding author.
E-MAIL carsten_rosenow@affymetrix.com; FAX (408) 481-0422.

of *aga* is induced in the presence of raffinose and repressed in the presence of sucrose in the growth medium. The newly identified gene cluster enables *S. pneumoniae* to use raffinose as a carbon source. We demonstrate by insertional gene inactivation that in *S. pneumoniae* the sucrose-specific PTS, but not a CcpA homolog, is required for sucrose repression of *aga*.

RESULTS

Gene Identification and Analysis

Figure 1 shows the gene cluster retrieved from the partial genomic sequence database of *S. pneumoniae* (www.tigr.com). Some members of this gene cluster show high homology to members of the multiple-sugar metabolism cluster (*msm*) from *S. mutans*. The region containing the *rafS*, *rafR*, *aga*, and *rafE* genes was PCR-amplified, using the high-fidelity DNA polymerase *Pfu*, and sequenced. There was a difference of 12 nucleotide substitutions between the PCR-amplified sequence and The Institute for Genomic Research (TIGR) database that resulted in changes in protein sequence. In all but one case, the amino acid replacements were conservative. The sole exception was the change of Gly to Asp at residue 430 of the *aga* gene product.

The sequence similarities between the genes of the identified cluster to those of the *msm* system from *S. mutans* suggest that the pneumococcal gene products could be involved in transport and metabolism of α-galactosides and/or other carbon sources. The *dexB* gene (dextran glucosidase) and the *msmK* gene (ATP-binding protein), both present in the *msm* cluster, are absent in the pneumococcal gene cluster. Two genes, *rafS* and *rafX*, are only found in the pneumococcal gene cluster.

Comparison of the RafS amino acid sequence with proteins in the public database revealed 34% identity and 52% similarity to BirA from *Bacillus subtilis*, a bifunctional protein involved in biotin-operon repression and biotin-protein ligation (Bower et al. 1995). The amino acid sequence of RafR shows 40% identity and 60% similarity to the MsmR protein from *S. mutans* and contains a sequence signature (RMHRARQL-LENTQESIKVIAYSVGFSDPLHFSKAYKQYFNQTP) of the AraC/XylS family of transcriptional regulators (Russell et al. 1992; Gallegos et al. 1997). *Aga* is transcribed divergently from *rafS* and *rafR* and encodes a protein with 64% identity and 79% similarity to α-galactosidase from *S. mutans*.

There is a noncoding region of 91 bp between the termination codon of *aga* and the initiation codon of the *rafE* gene. The putative translational start codon (ATG) of *rafE* is preceded by a sequence with homology to a ribosome binding site and the promoter consensus sequence from *S. pneumoniae* (Sabelnikov et al. 1995) (Fig. 2B). The *rafE* gene encodes a protein with high homology to the *S. mutans* MsmE protein and other sugar binding proteins. A PROSITE search (GCG; Wisconsin Package) using the RafE sequence revealed the peptide sequence Arg-Gly-Asp (263-RGD-265) within RafE. This sequence has been shown to play a role in cell adhesion in various systems (Ruoslahti and Pierschbacher 1986; d'Souza et al. 1991) but its relevance in RafE is not known. In addition, the RafE sequence contains the ATP/GTP-binding site motif A (P-loop) (22-ACSNYGKS-29). This sequence is known to interact with one of the phosphate groups of the nucleotide and is present in ABC transporters (Higgins et al. 1990). The third identified motif (138-PFTANAYGIYYNKDK-FEE-155) is present in the family of bacterial extracellular solute-binding proteins (Tam and Saier 1993). The first 20 amino acids of the RafE protein specify a potential signal peptide, with a cleavage site (19-Gly-Leu-Gly-Ala-Cys-Ser-25) similar to the bacterial lipoprotein consensus sequence (Leu-Ala-Gly/Ala-Cys) (von Heijne 1998).

The amino acid sequences of RafF and RafG are homologous of the corresponding Msm proteins, which are involved in sugar transport in *S. mutans* (Russell et al. 1992). Based on homology studies, GtfA could be a sucrose phosphorylase, which cleaves sucrose into fructose and glucose and phosphorylates the glucose for further metabolization (Russell et al. 1988). The last ORF in the cluster, *rafX*, encodes a protein with no significant homology to any known protein in the SWISS-PROT database and no motifs were identified using the PROSITE search.

Induction of the *aga* Promoter by Raffinose

The α-galactosidase encoded by the *aga* gene was used

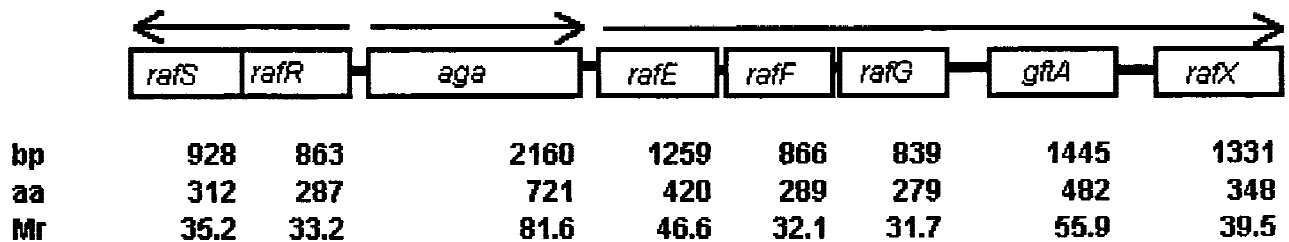|	rafS	rafR	aga	rafE	rafF	rafG	gtfA	rafX
bp	928	863	2160	1259	866	839	1445	1331
aa	312	287	721	420	289	279	482	348
Mr	35.2	33.2	81.6	46.6	32.1	31.7	55.9	39.5

Figure 1 Map of the *raf* genes in the *S. pneumoniae* contig. Arrows indicate operons and the orientation of transcription. The number of base pairs (bp), amino acids (aa), and the molecular mass (M_r) of the corresponding proteins are indicated.
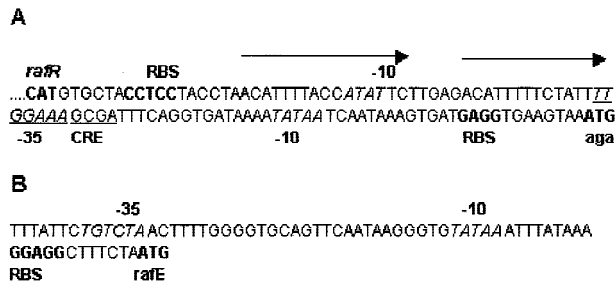
A

```
rafR        RBS                          -10
....CATGTGCTACCTCCTACCTAACATTTTACCATATTCTTGAGACATTTTTCTATTTT
GGAAAGCGATTTCAGGTGATAAAATATAATCAATAAAGTGATGAGGTGAAGTAAATG
 -35     CRE                  -10                RBS        aga
```

B

```
          -35                              -10
TTTATTCTGTCTAACTTTTGGGGTGCAGTTCAATAAGGGTGTATAAATTTATAAA
GGAGGCTTTCTAATG
RBS           rafE
```

Figure 2 Sequence analysis of the intergenic region between *rafR* and *aga* (*A*) and *aga* and *rafE* (*B*). Potential ribosome binding sites (RBS), −10 and −35 promoter elements (italics), and the direct repeats (arrows) are indicated. A 15-bp sequence with homology to the CRE consensus sequence is underlined.

as a reporter protein to analyze the regulation of its promoter (P_A). Aga activity is measured in the cell lysate using a colorimetic assay with the substrate *p*-nitrophenyl-α-D-galactopyranoside.

Very low α-galactosidase activity (0.4 U/mg of protein) could be detected in cell lysates of *S. pneumoniae* grown in semidefined medium (C+Y) in which glucose and sucrose are the only carbon sources. A 500-fold increase in activity was observed when sucrose and glucose in the growth medium were replaced by the α-galactoside sugar raffinose [α-galactosyl (1-6) α-glucosyl (1-2) β-fructose] at a concentration of 0.2% (wt/vol) (Table 1). Another α-galactoside sugar, melibiose [α-galactosyl (1-6) α-glucosyl], does not support growth of pneumococcus when provided as the sole carbon source. In the presence of glucose, melibiose did not induce α-galactosidase activity (data not shown). None of the other sugars tested (glucose, fructose, sucrose, galactose, lactose, maltose, inulin, and trehalose) induced α-galactosidase activity, suggesting that raffinose is the only sugar capable of inducing *aga* expression (Table 1).

To determine the correlation between α-galactosidase expression and raffinose concentration in the medium, pneumococcus strain R6x was grown in C+Y medium supplemented with 0.2% glucose and different concentrations of raffinose. After 3 hr of incubation, the α-galactosidase activity in the cell lysate was measured. As shown in Figure 3, α-galactosidase activity is dependent on the raffinose concentration in the medium and increases with higher raffinose concentrations up to saturation at ~200 ± 23 U/mg of protein at a raffinose concentration of 0.03% (wt/vol) in the medium.

Regulation of the *aga* Promoter by RafR and RafS

The homology of RafR to MsmR and other regulatory proteins within the AraC/XylS family of transcriptional regulators suggested that *aga* might be positively regulated by *rafR*. A *rafR* mutant showed only an 80-fold increase in α-galactosidase activity after induction

with raffinose, compared with a 500-fold increase in the wild type strain (Table 1).

Similarly, homology of the RafS protein to the biotin repressor protein suggested a regulatory function for RafS. Insertional inactivation of the gene resulted in a mutant with 2.2-fold higher α-galactosidase activity than the wild-type strain when both were grown in the presence of the inducer raffinose (Table 1).

To define the *aga* promoter by heterogeneous expression in *Escherichia coli*, a PCR fragment extending from the translational stop codon of *rafR* through the transcriptional start codon of *rafE* was amplified and cloned into pR326 (Claverys et al. 1995) resulting in plasmid pR326aga. The 2-kb amplified fragment was expected to include the *aga* gene as well as the extended *aga* promoter region necessary for *aga* regulation. *E. coli* harboring the plasmid pR326aga showed high levels of α-galactosidase activity (Table 1). Introduction of a compatible plasmid carrying the *rafR* gene (including the putative promoter) into the same *E. coli* strain resulted in a fivefold increase in α-galactosidase activity (Table 1). The endogenous levels of α-galactosidase activity in *E. coli* were negligible.

Specificity of the *raf* Cluster

Mutants in each of the remaining genes (*aga*, *rafE*, *rafF*, *rafG*, *gtfA*, and *rafX*) were constructed by insertion–duplication mutagenesis (see Methods). These strains and the parent strain were compared in their ability to ferment raffinose, trehalose, fructose, lactose, inulin,

Table 1. α-Galactosidase Actvities in Cell Lysates of Different Strains Grown on C + Y Medium Supplemented with the Indicated Carbon Source

Strain[a]	Carbon source	α-Galactosidase activity (U/mg protein)[b]
sp. R6x (wild type)	glucose	0.7 ± 0.1
	glucose/sucrose	0.4 ± 0.1
	glucose/raffinose	200 ± 40
	sucrose/raffinose	0.7 ± 0.2
sp. R6x (Δ*ptsII*)	glucose	0.8 ± 0.1
	glucose/raffinose	220 ± 38
	sucrose/raffinose	203 ± 45
sp. R6x (Δ*ccpA*)	glucose	0.7 ± 0.1
	glucose/raffinose	238 ± 44
	sucrose/raffinose	1.0 ± 0.4
sp. R6x (Δ*rafR*)	glucose	0.4 ± 0.1
	glucose/raffinose	34 ± 8
sp. R6x (Δ*rafS*)	glucose	0.8 ± 0.1
	glucose/raffinose	444 ± 92
E. coli	L broth	0
E. coli pR326aga	L broth	800 ± 120
E. coli rafR/aga	L broth	4600 ± 410

[a](sp.) *S. pneumoniae*.
[b]The activities are expressed in units (nmoles of *p*-nitrophenol produced per min of incubation at room temperature) per mg protein with their S.D. Shown is the mean of three experiments.
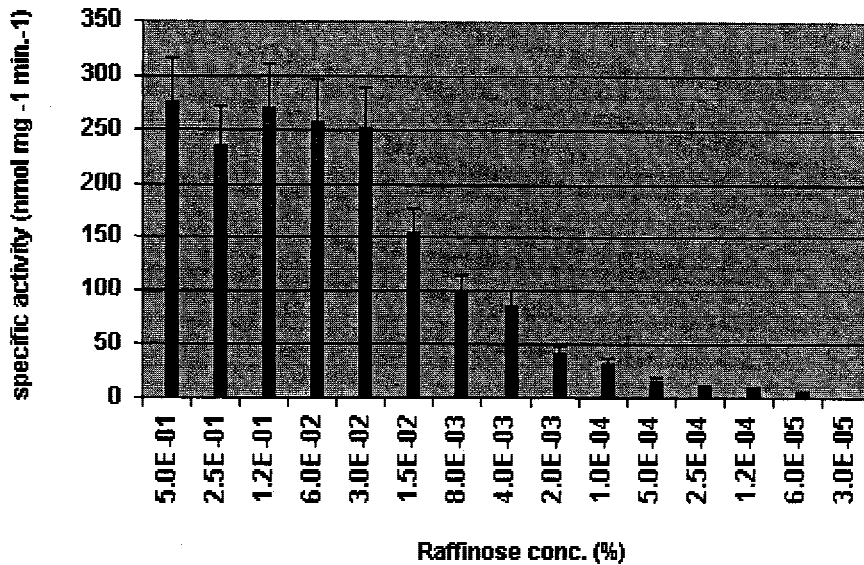
Figure 3 α-Galactosidase activities in the cell lysates of *S. pneumoniae* strain R6x grown in the presence of 0.2% glucose and different raffinose concentrations in the growth medium. The activities are expressed in units (nmoles of *p*-nitrophenol produced per min of incubation at room temperature) per mg protein with their S.D..

maltose, sucrose, and glucose and to induce α-galactosidase in the presence of raffinose. Of the six mutant strains, only the *gtfA* mutant could still ferment raffinose and induce *aga* in the presence of raffinose. All the mutants were able to ferment the other sugars, demonstrating that the *raf* cluster proteins are required for the metabolism of raffinose but not for that of the other sugars tested (Table 2).

Unlike *S. mutans*, *S. pneumoniae* does not ferment melibiose, isomaltose, or isomaltotriose (data not shown). Consistent with this property, none of these sugars serve as inducers of *aga* expression in *S. pneumoniae* (data not shown).

Catabolite Repression by Sucrose

To investigate whether the *aga* promoter is regulated by other mechanisms, pneumococcus strain R6x was grown in C+Y + raffinose medium supplemented with various other sugars at a final concentration of 0.2% (see Methods). The resulting α-galactosidase activities in the cell lysate were measured (data for sucrose and glucose are shown in Table 1). Only sucrose was able to catabolite-repress *aga* expression (Table 1). Because sucrose can be transported by the PTS, the *ptsII* gene of the sucrose-specific PTS was identified in the *S. pneumoniae* genome sequence (www.tigr.com) by homology to the sucrose-specific PTS in *B. subtilis*. Upon inactivation of the *ptsII* gene in *S. pneumoniae* by insertion-duplication mutagenesis, the mutant strain was unable to repress *aga* induction in the presence of sucrose and raffinose (Table 1).

In other Gram-positive bacteria, catabolite repression is mediated through a defined *cis*-active element (CRE) in the nucleotide sequence of the regulated promoter (Hueck et al. 1994). A 15-bp DNA sequence located 50 bp upstream of the putative *aga* translation initiation codon (ATG) (Fig. 2A) shows high homology to the CRE consensus sequence (Hueck et al. 1994) and to the CRE sequence identified in the *S. mutans aga* promoter (Simpson and Russell 1998) (Table 3).

In *B. subtilis*, catabolite repression is mediated by CcpA, a homolog of which was identified within the *S. pneumoniae* genome (www.tigr.com). Insertional inactivation of the *S. pneumoniae ccpA* homolog had no effect on raffinose induction nor sucrose repression of *aga* expression (Table 1).

Promoter Sequence Analysis of *aga*

Translation of the *aga* gene is most likely initiated from the putative ATG start codon (Fig. 2) that is preceded by a ribosome binding site (TGAGG). Preceding the ribosome binding site was a sequence (TATAA-10 and TTTGGAAA-35) that shows similarity to the previously defined pneumococcal promoter consensus sequence (Sabelnikov et al. 1995). Most striking, however, is the pres-

Table 2. Ability of *S. pneumoniae* Mutants Inactivated in Genes of the *raf* Cluster to Ferment (Acid Production) Raffinose, other Sugars and to Induce α-Galactosidase

Gene inactivated	Raffinose fermentation	Sugar fermentation[a]	α-Galactosidase activity (U/mg protein)[b]
none	+	all	228 ± 38
aga	−	all	0.7 ± 0.1
rafE	−	all	0.9 ± 0.1
rafF	−	all	0.8 ± 0.1
rafG	−	all	0.7 ± 0.1
gtfA	+	all	219 ± 25
rafX	−	all	0.7 ± 0.1

[a]Glucose, fructose, sucrose, galactose, lactose, maltose, inulin, and trehalose.
[b]The α-galactosidase activities are expressed in units (nmoles of *p*-nitrophenol produced per min of incubation at room temperature) per mg protein with their S.D.

Table 3. Homology of the CRE Consensus Sequence with the *S. pneumoniae* CRE Sequence Within the *aga* Promoter and the *S. mutans* CRE Sequence Within the *aga* Promoter

Consensus sequence	T G A A A A C G G T T T C A
S. pneumoniae	T G G A A A G C G A T T T C A
S. mutans	T T A A A A C G G T T A C A

ence of a direct repeat (ACATTTTACCATATT) at nucleotide positions 21 and 42 (Fig. 2). This repeat motif has high homology to the consensus binding region for AraC (AGCN₇TCCATA) (Gallegos et al. 1997). This potential RafR binding site overlaps the −35 promoter region by 3 bp. The *cis*-active element (CRE) is located

following the last repeat and overlaps with the rest of the −35 promoter.

Analysis of mRNA Transcripts

Putative promoters were assigned upstream of the *aga* gene, the *rafR* gene, and the *rafE* gene (Fig. 2). All promoter sequences show homology to the consensus sequence for *S. pneumoniae* promoters (Sabelnikov et al. 1995). To characterize the mRNA transcripts spanning the *raf* cluster, reverse transcriptase–PCR (RT–PCR) analysis was performed on total RNA obtained from R6x strain grown in C+Y medium supplemented with glucose and raffinose. Figure 4 shows the DNA regions covered by the primers and the results from the RT–PCR experiment. A 1.0-kb fragment was amplified using sense and antisense primers within *rafR*. A 1.5-kb fragment was amplified using the sense primer within *rafR* and the antisense primer within *rafS*, confirming that *rafR* and *rafS* are cotranscribed, presumably from a promoter upstream of *rafR*. Primers within the *aga* gene amplified the expected 0.3-kb DNA fragment. However, primers spanning the *aga*–*rafE* genes resulted in no DNA amplification after RT–PCR, indicating that the *aga* transcript does not extend into *rafE* and suggesting the presence of an additional promoter between *aga* and *rafE*. To investigate how many genes are cotranscribed from the *rafE* promoter, combinations of primers, each one from a different gene, were generated and tested using the RT–PCR approach. All different primer combinations amplified a DNA fragment of the predicted size (Fig. 4). The additional cDNA bands in the RT–PCR reactions presumably reflect nonspecific binding of the oligonucleotides to other gene sequences or additional internal promoters.
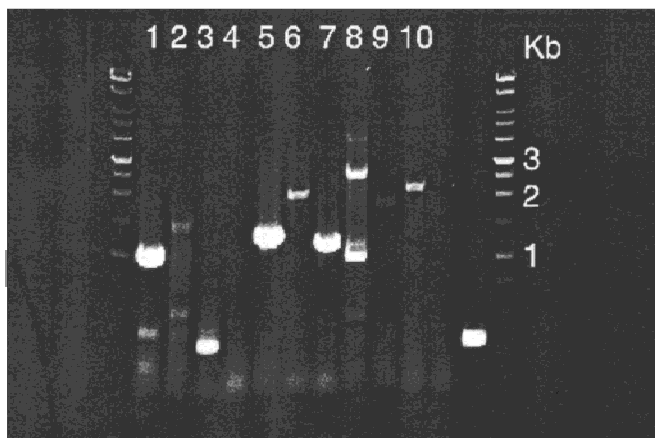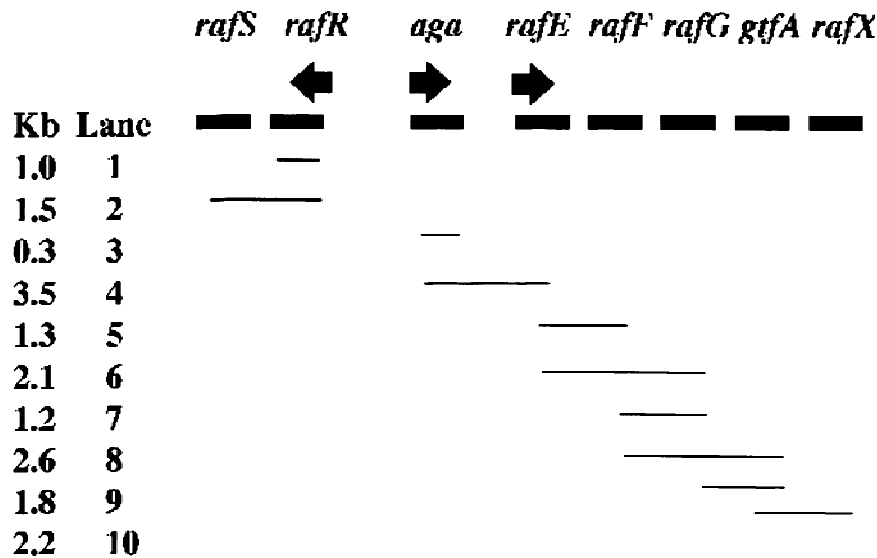
DISCUSSION

A gene cluster responsible for utilization of raffinose was identified in *S. pneumoniae* and named the *raf* cluster. The sequence of ~10,200 bp contains eight open reading frames, which were named *rafS*, *rafR*, *aga*, *rafE*, *rafF*, *rafG*, *gtfA*, and *rafX* to indicate

Figure 4 Amplified cDNA from total RNA from the *S. pneumoniae* strain R6x grown in C+Y supplemented with 0.2% raffinose and 0.2% glucose. In each reaction (lanes *1–10*) one primer pair was used, as indicated by the lines. The expected size for each fragment is given in kb. Lane *11* is a negative control (−) without RT and the primer pair from lane *3* in the reaction mixture; Lane *12* is a positive control (+) with total RNA and primers provided with the Access RT–PCR introductory System. The arrows indicate the identified potential promoters.

their role in raffinose utilization. Because most of the proteins are homolog to the proteins of the *S. mutans* multiple sugar metabolism (*msm*) system, we have named the pneumococcal components according to the *S. mutans* nomenclature (Russell et al. 1992). Raffinose was identified as the inducer for *aga* when added to the growth medium. The extent of induction in the presence of raffinose is 500-fold based on increased α-galactosidase activity. The mechanism of regulation by raffinose and the identity of the actual trigger for induction remain unknown. The level of *aga* induction depends on the raffinose concentration in the medium, which results in a raffinose titratable promoter system with different levels of gene expression, which could be a useful tool for regulated gene expression in *S. pneumoniae*.

Based on the homology, the *raf* cluster contains all the genes needed for the early steps of raffinose metabolism. Aga is the α-galactosidase that cleaves raffinose into galactose and sucrose. Sucrose might be subsequently hydrolyzed and phosphorylated by GtfA (sucrose phosphorylase), which is a bifunctional enzyme to yield fructose and glucose-1-phosphate for further metabolism (Russell et al. 1988). Inactivation of *gtfA* in the *raf* cluster was unique in that it had no effect on the phenotype in regards to raffinose metabolism or *aga* induction. Apparently, the mutation in *gtfA* was not polar on *rafX*, and GtfA is not required for raffinose metabolism. Structural and functional analysis indicates that RafS is a repressor of the *aga* promoter. In contrast, the decreased α-galactosidase activity in the *rafR* mutant suggests that RafR is an activator of *aga* expression. This is in agreement with the homology of RafR to members of the AraC/XylS family of transcriptional regulators. The remaining α-galactosidase activity (34 U/mg protein in the mutant vs. 200 U/mg protein in wild type) in the *rafR* mutant could reflect basal level promoter activity.

Based on homology studies and insertional inactivation of the remaining genes (*rafE*, *rafF*, *rafG*, and *rafX*) in the *raf* cluster, the corresponding proteins are involved in utilization of raffinose. The results from the gene inactivation experiments could indicate a polar effect from *rafE* on *rafF* or *rafG*; it will be necessary to construct nonpolar mutations to draw conclusions on the role of RafF and RafG in raffinose utilization. However, the strong homology of RafE, RafF, and RafG to other sugar transport proteins suggests that these proteins are involved in raffinose transport. Despite the similarities between the *raf* and *msm* gene clusters, the products of *raf* genes are involved in the metabolism of raffinose alone. In contrast, the products of the *msm* genes are involved in uptake of raffinose, melibiose, and isomaltotriose and metabolism of melibiose, sucrose, and isomaltosaccharides (Russell et al. 1992). This difference could be due to the absence of ho-

mologs of *msmK* or *dexB* in the *raf* cluster or differences in substrate specificity of the metabolic enzymes or transport systems.

Three promoters were ascribed in the gene cluster using sequence analysis and RT–PCR on total RNA. One transcript was found for the *rafR* and *rafS* genes, suggesting a mutual promoter for these genes. The gene products from *rafR* and *rafS* regulate the *aga* promoter, which is transcribed in the opposite direction. The inverted repeat identified within the *aga* promoter shows high homology to the AraC binding region (Gallegos et al. 1997) and could function as the binding region for the activator RafR. This repeat was shown to be essential for activation of the promoter P_{araBAD} in *E. coli*, which also overlaps the -35 promoter region by 4 bp.

Consistent with a putative third promoter sequence upstream of *rafE*, RT–PCR experiments showed that the remaining genes (*rafE*, *rafF*, *rafG*, *gtfA*, and *rafX*) are likely to be organized in an operon transcribed from a proximal promoter. However, the presence of internal promoters cannot be excluded.

Protein sequence analysis of the RafE protein sequence revealed a putative signal sequence with homology to a lipid anchor sequence at the cleavage region. The identified ATP/GTP and solute binding motifs are close to the amino-terminal end of the protein and might be exposed to the extracellular site of the bacteria, because an identified lipid-anchor motif at the amino-terminal end of RafE suggests that this protein is anchored in the membrane. Members of the solute binding motif's family are known to be bound to the membrane via an amino-terminal lipid anchor and probably serve as receptors to trigger or initiate translocation of the solute (Tam and Saier 1993), in this case the transport of raffinose.

The RGD motif was first identified in extracellular matrix proteins, where it is crucial for their interaction with integrin, a cell surface receptor (Ruoslahti and Pierschbacher 1986). The RGD sequence was subsequently characterized as a putative cell-binding sequence and has been shown to promote eukaryotic cell attachment (d'Souza et al. 1991). Whether the tripeptide Arg-Gly-Asp (RGD) found within the RafE sequence is involved in cell attachment of *S. pneumoniae* is unclear, and further experiments are necessary. However, the connection of sugar metabolism and the expression of virulence factors has been established in several pathogenic bacteria (Martinez-Cadena et al. 1981; Smith et al. 1986; Busque et al. 1995; Milenbachs et al. 1997), and it may provide an advantage to colonize a niche in the host.

Because the PTS is the preferred sugar transport system in bacteria, other non-PTS transport systems can be regulated by PTS-mediated catabolite repression. In the presence of sucrose in the medium, no *aga*

induction was detected, whereas none of the other sugars tested, including glucose, showed repression of *aga* induction. This is highly unusual, because most of the catabolite repression in prokaryotes is mediated by glucose and not sucrose in the medium (Saier et al. 1996). To further investigate the catabolite repression mechanism, the enzyme II (*ptsII*) of the sucrose-specific PTS was inactivated. The *ptsII* mutant lost its ability to repress the *aga* promoter in the presence of sucrose. This observation supports the theory that the sucrose-specific PTS is either directly or indirectly involved in the regulation of raffinose utilization. This suggests that *S. pneumoniae* uses the PTS as the preferred sugar transport system, with sucrose being most likely the preferred carbon source among the sugars tested.

Upstream of the translation initiation codon for *aga*, a highly homologous sequence to the *cis*-active element (CRE) consensus sequence was identified. Because catabolite repression in *B. subtilis* (Miwa et al. 1994), *B. Megaterium* (Deutscher et al. 1995), *Staphylococcus xylosus* (Egeter and Brückner 1996), and *Lactococcus lactis* (Luesink et al. 1998) is mediated by the protein CcpA, it has been hypothesized that this protein is a common mediator between the PTS and the regulated gene in Gram-positive bacteria (Hueck et al. 1994; Saier et al. 1996; Luesink et al. 1998). A homolog of the *B. subtilis* CcpA protein was identified in the *S. pneumoniae* genome, but inactivation of the gene had no effect on the regulation of *aga* expression. Because a mutation in a CcpA homolog in *S. mutans* (Simpson and Russell 1998) showed the same phenotype, the regulation of catabolite repression in *Streptococcus* seems to be either mediated through PTS directly or a, so far, unidentified regulatory protein.

METHODS

Strains and Media

The nonvirulent *S. pneumoniae* strain R6x (Tiraby and Fox 1973) was used for the experiments. *S. pneumoniae* strains were routinely plated on tryptic soy agar supplemented with sheep blood (TSAB) to a final concentration of 3% (vol/vol). Cultures were grown in a liquid semisynthetic casein hydrolysate medium supplemented with yeast extract but without any carbon source (C+Y medium) (Lacks and Hotchkiss 1960). The carbon source was added to a final concentration of 0.2% (wt/vol) if not otherwise indicated. Mutants harboring vector pR326 were grown in the presence of 2.5 µg/ml chloramphenicol. *E. coli* strain JM109 (Promega, Madison, WI) was used for cloning experiments and grown in L broth-based medium with aeration at 37°C. Recombinant *E. coli* strains carrying vector pR326 were grown in the presence of 25 µg/ml chloramphenicol.

DNA Techniques and Sequence Analysis

DNA techniques, including plasmid preparations, restriction endonuclease digestion, ligations, PCR, and genetic manipulation of *E. coli*, were performed according to standard protocols (Maniatis 1989) and the specifications of the manufac-

turers. Chromosomal DNA of *S. pneumoniae* was isolated as described previously (Pearce et al. 1993). The *raf* cluster was identified by searching the *S. pneumoniae* genome database (at www.tigr.com), accessible from the National Center for Biotechnology Information (NCBI) Web page (http://www.ncbi.nlm.nih.gov/BLAST/), with AraC from *E. coli* as the query sequence. The gene region starting at *rafS* and including *rafR*, *aga*, and *rafE* was PCR-amplified using *Pfu* turbo (Stratagene, La Jolla, CA) and sequenced. DNA sequencing was done at Sequatech (Mountain View, CA). Protein and DNA sequences were analyzed with the GCG Wisconsin package (GCG, Madison, WI).

Insertion–Duplication Mutagenesis in *S. pneumoniae*

An internal DNA fragment of the target gene was PCR-amplified and ligated into the T-tailed (Marchuk et al. 1991) vector pR326 (Claverys et al. 1995), which replicates in *E. coli* but not *S. pneumoniae*. After transformation into *E. coli* JM109, a single transformant that contained the vector with the DNA fragment was identified. Plasmid DNA from this clone was transformed into *S. pneumoniae* strain R6x as described previously (Pearce et al. 1993). Transformants were selected on TSAB plates containing chloramphenicol (2.5 µg/ml). PCR analysis of the mutant's chromosomal DNA was used to confirm homologous recombination of the cloned internal fragment into the target gene.

Expression of *aga* and *rafR* in *E. coli*

The *aga* gene including the promoter region was amplified by PCR on chromosomal DNA from pneumococcus R6x using the following primers: sense, 5'-GCGCCGGAATTCCATGTGCTACCTCCTACCTAACATTTTACC-3' and antisense, 5'-CGCGGATCCTCATAGTTTTCTAAAAATATAC-3'. The amplified 2-kb fragment includes the ribosomal binding site, the -10 and -35 promoter elements, and an additional 70-bp upstream sequence. The PCR product was cloned in pR326 (Claverys et al. 1995) using the *Bam*HI and *Eco*RI restriction sites. After transformation in *E. coli* JM109, colonies were tested for α-galactosidase activity and showed high α-galactosidase expression. The *rafR* gene was PCR-amplified using the sense primer 5'-GGCCCTTCCATATGCATTTACTTCACCTCATCACTTTATTG-3' and the antisense primer 5'-CCCGGAATTCAGCTTGGTAGGATTTCATAATGTTGCC-3'. The PCR product was cloned in the T-tailed vector pGEMT-Easy (Promega, Madison, WI) and transformed into the *E. coli* strain expressing pneumococcal α-galactosidase. From the strain *E. coli* aga/RafR, the α-galactosidase activity was measured.

RNA Analysis

Total RNA was isolated from *S. pneumoniae* strain R6x grown in C+Y medium containing 0.2% glucose and 0.2% raffinose using RNAeasy mini kit (Qiagen, Santa Clarita, CA) according to the recommended procedure, except that cell lysis was performed using 0.25% Triton X-100 in TE buffer for 10 min at 37°C. After resuspension in DEPC-treated water, RNA was treated with RQ1 RNase free DNase (Promega, Madison, WI). The isolated RNA was used in RT–PCR reactions using Access RT–PCR Introductory System (Promega, Madison, WI). The reaction was performed using 1 µg of total RNA, 1 µM sense primer, 1 µM antisense primer, 1× AMV/*Tfl* reaction buffer, 0.2 mM dNTP mix, 1 mM MgSO₄, 0.1 U/µl AMV reverse transcriptase, 0.1 U/µl *Tfl* DNA polymerase, and nuclease-free water to a final volume of 50 µl. The RT reaction was performed

at 48°C for 45 min; an initial denaturation was performed at 94°C for 2 min, followed by 40 cycles at 94°C for 30 sec, 52°C for 1 min, and at 68°C for 2 min. An additional extension was performed at 68°C for 7 min. The RT–PCR products obtained were loaded on an 0.7% agarose gel for analysis.

Sugar Utilization Assay

S. pneumoniae strain R6x and mutant strains obtained by insertion–duplication mutagenesis in individual genes in the *raf* cluster were grown in 1.5 ml of C+Y medium supplemented with glucose and chloramphenicol for the mutants, to an optical density (OD_{600}) of 0.5. Cells were harvested by centrifugation, washed twice with C+Y medium containing no sugars, and resuspended in 1 ml of the same medium. This inoculum was diluted 1:300 in 5 ml of C+Y medium containing chloramphenicol (2.5 µg/ml) when needed, and 0.2% (wt/vol) of the following sugars was tested: glucose, sucrose, raffinose, inulin, galactose, maltose, lactose, fructose, and trehalose. Phenol red solution was added to each tube to a final concentration of 0.02% (vol/vol) and incubated at 37°C for 4 hr. Sugar utilization by the bacteria results in acid production changing the color of the medium from red to yellow. A control culture without sugar added was included in each experiment. All assays were made in duplicates.

α-Galactosidase Enzyme Assay

S. pneumoniae strains were grown to an OD_{600} of 0.5 in 3 ml of C+Y medium containing the test sugar. A 200-µl aliquot was pelleted by centrifugation, and cells were resuspended in 100 µl of sodium phosphate buffer (pH 7.5) containing 0.25% Triton X-100 and incubated at 37°C for 10 min. Ten microliters of the cell lysate was used for protein quantification using the BCA Protein Assay Kit (Pierce, Rockford, Il). To measure α-galactosidase activity, 10 µl of the cell lysate was incubated with 90 µl of the substrate solution [1 mM $MgCl_2$, 4.5 mM β-mercaptoethanol, 90 µg *p*-nitrophenyl-α-D-galactopyranoside (Sigma, St. Louis, MO) in 0.1 M sodium phosphate buffer at pH 7.5] for 5 min at room temperature. When using *E. coli*, cells were disintegrated by ultrasonication, and cell debris were removed by centrifugation. α-Galactosidase activity was followed by measuring absorption (405 nm) of the product *p*-nitrophenol. A standard curve was obtained using p-nitrophenol at 800, 600, 400, 200, 100, and 0 µM. Units of α-galactosidase were defined as nmoles of *p*-nitrophenol per minute.

ACKNOWLEDGMENTS

REFERENCES

Bower, S., J. Perkins, R.R. Yocum, P. Serror, A. Sorokin, P. Rahaim, C.L. Howitt, N. Prasad, S.D. Ehrlich, and J. Pero. 1995. Cloning and characterization of the *Bacillus subtilis* birA gene encoding a repressor of the biotin operon. *J. Bacteriol.* **177:** 2572–2575.

Busque, P., A. Letellier, J. Harel, and J.D. Dubreuil. 1995. Production of *Escherichia coli* STb enterotoxin is subject to catabolite repression. *Microbiology* **141:** 1621–1627.

Claverys, J.P., A. Dintilhac, E.V. Pestova, B. Martin, and D.A. Morrison. 1995. Construction and evaluation of new drug-resistance cassettes for gene disruption mutagenesis in *Streptococcus pneumoniae*, using an *ami* test platform. *Gene* **164:** 123–128.

Cvitkovitch, D.C., D.A. Boyd, and I.R. Hamilton. 1995. Regulation of sugar transport via the multiple sugar metabolism operon of *Streptococcus mutans* by the phosphoenolpyruvate phosphotransferase system. *J. Bacteriol.* **177:** 5704–5706.

Deutscher, J., E. Küster, U. Bergstedt, V. Charrier, and W. Hillen. 1995. Protein kinase-dependent HPr/CcpA interaction links glycolitic activity to carbon catabolite repression in Gram-positive bacteria. *Mol. Microbiol.* **15:** 1049–1053.

d'Souza, S.E., M.H. Ginsberg, and E.F. Plow. 1991. Arginyl-glycyl-aspartic acid (RGD): A cell adhesion motif. *Trends Biochem. Sci.* **16:** 246–250.

Egeter, O. and R. Brückner. 1996. Catabolite repression mediated by the catabolite control protein CcpA in *Staphylococcus xylosus*. *Mol. Microbiol.* **21:** 739–749.

Gallegos, M.-T., R. Schleif, A. Bairoch, K. Hofmann, and J.L. Ramos. 1997. AraC/XylS family of transcriptional regulators. *Microbiol. Mol. Biol. Rev.* **61:** 393–410.

Higgins, C.F., S.C. Hyde, M.M. Mimmack, U. Gileadi, D.R. Gill, and M.P. Gallagher. 1990. Binding protein-dependent transport systems. *J. Bioenerg. Biomembr.* **22:** 571–592.

Holt, J., N.R. Krieg, P.H.A. Sneath, J.T.A. Staley, and S.T. Williams. 1994. *Bergey's manual of determinative bacteriology*. Williams & Wilkins, Baltimore, MD.

Hueck, C.J., W. Hillen, and J.M.H. Saier. 1994. Analysis of a *cis*-active sequence mediating catabolite repression in Gram-positive bacteria. *Res. Microbiol.* **145:** 503–518.

Lacks, S. and R.D. Hotchkiss. 1960. A study of the genetic material determining an enzyme activity in pneumococcus. *Biochim. Biophys. Acta* **39:** 508–517.

Luesink, E.J., R.E. van Herpen, O.P. Kuipers, and W.M. de Vos. 1998. Transcriptional activation of the glycolytic *las* operon and catabolite repression of the *gal* operon in *Lactococcus lactis* are mediated by the catabolite control protein CcpA. *Mol. Microbiol.* **30:** 789–798.

Maniatis, T., E.F. Fritsch, and J. Sambrook. 1989. *Molecular cloning: A laboratory manual.* Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

Marchuk, D., M. Drumm, A. Saulino, and F.S. Collins. 1991. Construction of T-vectors, a rapid and general system for direct cloning of unmodified PCR products. *Nucleic Acids Res.* **19:** 1154.

Martinez-Cadena, M.G., L.M. Guzman-Verduzco, H. Stieglitz, and Y.M. Kupersztoch-Portnoy. 1981. Catabolite repression of *Escherichia coli* heat stable enterotoxin activity. *J. Bacteriol.* **145:** 722–728.

Milenbachs, A.A., D.P. Brown, M. Moors, and P. Youngman. 1997. Carbon source regulation of virulence gene expression in *Listeria monocytogenes*. *Mol. Microbiol.* **23:** 1075–1085.

Miwa, Y., M. Saikawa, and Y. Fujita. 1994. Possible function and some properties of the CcpA protein in *Bacillus subtilis*. *Microbiology* **140:** 2567–2575.

Nieto, C., M. Espinosa, and A. Puyet. 1997. The maltose/maltodextrin regulon of *Streptococcus pneumoniae*. *J. Biol. Chem.* **272:** 30860–30865.

Pearce, B.J., Y.B. Yin, and H.R. Masure. 1993. Genetic identification of exported proteins in *Streptococcus pneumoniae*. *Mol. Microbiol.* **9:** 1037–1050.

Postma, P.W. and J.W. Lengeler. 1985. Phosphoenolpyruvate:carbohydrate phosphotransferase system of bacteria. *Microbiol. Rev.* **49:** 232–269.

Postma, P.W., J.W.A. Lengeler, and G.R. Jacobson. 1993. Phosphoenolpyruvate:carbohydrate phosphotransferase system of bacteria. *Microbiol. Rev.* **57:** 543–594.

Ruoslahti, E. and M.D. Pierschbacher. 1986. Arg-Gly-Asp: A versatile cell recognition signal. *Cell* **44:** 517–518.

———. 1987. New perspective in cell adhesion: RGD and integrins. *Science* **238:** 491–497.

Russell, R.R., H. Mukasa, A. Shimamura, and J.J. Ferretti. 1988.

Streptococcus mutans gtfA gene specifies sucrose phosphorylase. *Infect. Immun.* **56:** 2763–2765.

Russell, R.R., J. Aduse-Opoku, I.C. Sutcliffe, L. Tao, and J.J. Ferretti. 1992. A binding protein-dependent transport system in *Streptococcus mutans* responsible for multiple sugar metabolism. *J. Biol. Chem.* **267:** 4631–4637.

Sabelnikov, A.G., B. Greenberg, and S.A. Lacks. 1995. An extended −10 promoter alone directs transcription of the *Dpn*II operon of *Streptococcus pneumoniae*. *J. Mol. Biol.* **250:** 144–155.

Saier, M.H. 1996. Cyclic AMP-independent catabolite repression in bacteria. *FEMS Microbiol. Lett.* **138:** 97–103.

Saier, M.H.J., S. Chauvaux, G.M. Cook, J. Deutscher, I.T. Paulsen, J. Reizer, and J.-J. Ye. 1996. Catabolite repression and inducer control in Gram-positive bacteria. *Microbiology* **142:** 217–230.

Simpson, C.L. and R.R.B. Russell. 1998. Identification of a homolog of CcpA catabolite repressor protein in *Streptococcus mutans*. *Infect. Immun.* **66:** 2085–2092.

Smith, J.L., M.M. Bencivengo, and C.A. Kunsch. 1986. Enterotoxin a synthesis in *Staphylococcus aureus*: Inhibition by glycerol and maltose. *J. Gen. Microbiol.* **132:** 3375–3380.

Stassi, D.L., J.J. Dunn, and S.A. Lacks. 1982. Nucleotide sequence of DNA controlling expression of genes for maltosaccharide utilization in *Streptococcus pneumoniae. Gene* **20:** 359–366.

Tam, R. and M.H.J. Saier. 1993. Structural, functional, and evolutionary relationships among extracellular solute-binding receptors of bacteria. *Microbiol. Rev.* **57:** 320–346.

Tiraby, G. and M.S. Fox. 1973. Marker discrimination in transformation and mutation of pneumococcus. *Proc. Natl. Acad. Sci.* **70:** 3541–3545.

Vadeboncoeur, C. and M. Pelletier. 1997. The phosphoenolpyruvate: sugar phosphotransferase system of oral streptococci and its role in the control of sugar metabolism. *FEMS Microbiol. Rev.* **19:** 187–207.

von Heijne, G. 1998. The structure of signal peptides from bacterial lipoproteins. *Protein Eng.* **2:** 531–534.