JOURNAL OF BACTERIOLOGY, Jan. 2003, p. 413–421 Vol. 185, No. 2
0021-9193/03/$08.00+0 DOI: 10.1128/JB.185.2.413–421.2003

Transcriptional Regulation in the *Streptococcus pneumoniae rlrA* Pathogenicity Islet by RlrA

David L. Hava, Carolyn J. Hemsley, and Andrew Camilli*

Department of Molecular Biology and Microbiology, Tufts University School of Medicine, Boston, Massachusetts 02111

Received 29 August 2002/Accepted 28 October 2002

The proper temporal expression of virulence genes during infection is crucial to the infectious life cycle of microbial pathogens, particularly in pathogens that encounter a multitude of environments in eukaryotic hosts. *Streptococcus pneumoniae* normally colonizes the nasopharynges of healthy adults but can cause a range of diseases at a variety of host sites. Transcriptional regulators that are essential for full virulence of *S. pneumoniae* in different animal models have been identified. One such regulator, *rlrA*, is required for colonization of the nasopharynx and lung infection but is dispensable for systemic infection. Previous work has shown that *rlrA* lies in a 12-kb pathogenicity islet, divergently opposed to three putative sortase-anchored surface proteins and three sortase enzymes. In addition to *rlrA*, one of the putative surface proteins and one of the sortases have also been shown to be essential for lung infection. In this work, we demonstrate that RlrA is a positive regulator of all seven genes in the *rlrA* pathogenicity islet, with transcriptional activation occurring at four different promoters in the islet with AT-rich sequences. These promoters direct the expression of *rlrA* itself, the three sortases, *rrgA*, and *rrgBC*. These data are consistent with the model whereby the *rlrA* pathogenicity islet acts in an autonomous manner to alter the bacterial surface components that interact with the pulmonary and nasopharyngeal environments.

Streptococcus pneumoniae remains a major cause of morbidity and mortality in undeveloped and developed parts of the world, and resistance to common antibiotics is widespread (2, 4, 27). *S. pneumoniae* is a component of the normal flora of the nasopharynges of approximately 50% of all adults, where it coexists with other microflora in a nonpathogenic state. In immunocompromised people, the elderly, and young children, *S. pneumoniae* bacteria that initially colonize the nasopharynx may spread to distal sites, such as the inner ear, lower respiratory tract, or bloodstream, and cause diseases ranging from otitis media to pneumonia to meningitis (7, 18). Factors that lead to its spread from the nasopharynx to other sites of infection are not understood. Several studies have identified *S. pneumoniae* virulence factors that are essential to the survival of the bacterium in different host environments by signature-tagged mutagenesis (STM) with murine models of infection (10, 13, 23). A subset of these factors has been shown to be specific to certain host environments (10), and therefore, these genes code for proteins that have tissue-specific roles during infection and colonization. Among these are a number of putative transcriptional regulators that may regulate tissue-specific virulence factors in response to different host environments.

One putative transcriptional regulator identified by STM is RlrA (10), a homologue of RofA and Nra from *S. pyogenes* (6, 22). Through sequence analysis, *rlrA* has been shown to be one of seven genes in a pathogenicity islet of approximately 12 kb (Fig. 1) (10) that is not highly conserved in other *S. pneumoniae* strains (26). Of the six genes that are divergently transcribed from *rlrA*, three have homology to the LPXTG family of cell wall-anchored surface proteins (*rrgA*, *rrgB*, and *rrgC*). RrgA, RrgB, and RrgC have C-terminal sorting signals that are characteristic of LPXTG-containing proteins, except that the leucine of LPXTG is deviant in each protein. RrgB and RrgC have conservative changes to isoleucine and valine, respectively, whereas RrgA has a change to tyrosine. The C-terminal sorting signals predict that these proteins are covalently anchored to the cell wall by sortases, which are transpeptidases found in most gram-positive bacteria (15, 20). Interestingly, three of the four sortase homologues (*srtB*, *srtC*, and *srtD*) encoded in the TIGR4 genome lie within the *rlrA* pathogenicity islet; however, no proteins are known to be sorted by these sortases (Fig. 1) (10, 20, 26).

In addition to *rlrA*, *srtD* was also identified as an essential virulence gene through STM and each was confirmed to be essential to the survival of *S. pneumoniae* during lung infection by testing strains with transposon insertions in each gene in competition assays against the wild-type parental strain (10). The *rlrA* gene was also found to be essential for colonization of the nasopharynx, but not bacteremia, whereas *srtD* was dispensable in both of these models (10). The generation of transposon insertion mutations in each of the remaining genes in the locus and subsequent analysis of each mutant strain in murine models of infection demonstrated that *rrgA* is also essential for colonization of the nasopharynx and for lung infection, whereas *srtB* is essential only for colonization of the nasopharynx (10).

On the basis of the homology of RlrA to other gram-positive transcriptional regulators, the organization of the islet, and the phenotypes of certain mutant strains in animal assays, we previously proposed a model of regulation in the *rlrA* pathogenicity islet in which RlrA positively regulates the transcription of each *rlrA* pathogenicity islet gene. In this report, we confirm

* Corresponding author. Mailing address: Department of Molecular Biology and Microbiology, Tufts University School of Medicine, 136 Harrison Ave., Boston, MA 02111. Phone: (617) 636-2144. Fax: (617) 636-0337. E-mail: andrew.camillli@tufts.edu.

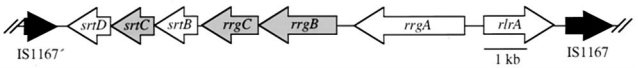

FIG. 1. The *rlrA* pathogenicity islet. The 12-kb locus includes a positive regulator, three surface proteins, and three sortase homologues. The four genes that are required for virulence in one or more animal models are in white (10).

this model by demonstrating that transcription of each gene in the islet is dependent upon RlrA. Furthermore, RlrA is shown to act at four different promoters within the islet at a consensus sequence that is found elsewhere in the *S. pneumoniae* chromosome, suggesting that although the *rlrA* pathogenicity islet may function autonomously at both the level of transcription and the protein secretion level, there may be additional targets of regulation in the TIGR4 chromosome.

MATERIALS AND METHODS

Bacterial strains, plasmids, and primers. The bacterial strains and plasmids used in this study are listed in Table 1. The parental strain for all *S. pneumoniae* genetic manipulations was AC353, a streptomycin-resistant (Sm^r) derivative of TIGR4 (10). *S. pneumoniae* strains were grown in Todd-Hewitt broth plus 5% yeast extract (THY) and supplemented with 0.8% maltose when indicated. Unless otherwise stated, the antibiotic concentrations used in this study were as follows: streptomycin at 100 μg/ml, chloramphenicol (CM) at 4 μg/ml, and spectinomycin at 200 μg/ml for *S. pneumoniae* and ampicillin at 100 μg/ml, CM at 10 μg/ml, and spectinomycin at 100 μg/ml for *Escherichia coli*. The primers used in this study are listed in Table 2. Unless otherwise noted, all PCRs were performed in reaction buffer containing 1× *Taq* reaction buffer (Promega), 250 μM deoxynucleoside triphosphates, each primer at 1 μM, and a 10:1 mixture of *Taq* and *Pfu* DNA polymerases. Reaction conditions consisted of 25 cycles of 95°C for 30 s, 50 to 52°C for 30 s, and 72°C for 30 s/kb of DNA, followed by a 5-min postdwell period at 72°C.

Construction of an *rlrA*-overexpressing strain. To construct a strain that expresses *rlrA* from an inducible promoter, the coding sequence of *rlrA* was introduced into the *S. pneumoniae* maltose locus downstream of *malM* (24). To this end, DNA fragments containing the 3′ end of the *malM* gene and the 5′ end of *malP* were PCR amplified from AC353 with the primer pairs MALFX/MALRP and MALPF2/MALPRP, respectively. Similarly, the *cat* gene, conferring CM resistance (Cm^r) on both *E. coli* and *S. pneumoniae*, was PCR amplified from pAC1000 with the primer set PCATF1/PCATR1 and the coding sequence of *rlrA* was PCR amplified from AC353 with the primer set RLRAFR/RLRARX. In the latter case, the Shine-Dalgarno sequence of the *S. pneumoniae rpoB* gene was engineered into the RLRAFR sequence to allow optimal translation efficiency of *rlrA* at the maltose locus. Each of these fragments was subcloned separately into pCR-Script Amp SK(+) (Stratagene) and subsequently inserted into pAC1000 to generate pCH84. pAC1000 is a derivative of pEVP3 (3) that was created by PCR amplifying the pEVP3 vector backbone with the primer set PEVPF1/PEVPR1 to delete the promoterless *lacZ* gene in pEVP3. The resulting product was digested with *Bam*HI, gel purified, and ligated overnight at 4°C. The final construct contains the 3′ *malM* sequence and the 5′ *malP* sequence flanking the *rlrA* coding sequence and the *cat* gene. To generate AC1278, the *S. pneumoniae* strain overexpressing *rlrA*, pCH84, was linearized by digestion with *Xho*I and the gel-purified fragment was transformed into naturally competent AC353 as pre-

TABLE 1. Strains and plasmids used in this study

Strain or plasmid	Relevant genotype	Reference(s)
E. coli		
DH5αλpir	F^- Δ(*lacZYA-argF*)*U169 recA1 endA1 hsdR1 supE44 thi-1 gyrA96 relA1* λ::*pir*	9, 12
XL-1 Blue	*recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac*[F′ *proAB lacI*^{q}*ZDM15* Tn*10* (Tet^r)]	Stratagene
AC1287	DH5αλpir; contains pAC1287	This work
AC1288	DH5αλpir; contains pAC1288	This work
AC1289	DH5αλpir; contains pAC1289	This work
AC1290	DH5αλpir; contains pAC1290	This work
AC1291	DH5αλpir; contains pAC1291	This work
AC1292	XL-1 Blue; contains pAC1292	This work
S. pneumoniae		
AC353	TIGR4 Sm^r derivative	10
AC1213	*rlrA′*::*magellan2* Sm^r Cm^r	10
AC1278	*malM*::*rlrA*::*cat*::*malP* Sm^r Cm^r	This work
Plasmids		
pGEM-T	Cloning vector; Ap^r	Promega
pCR-Script Amp SK(+)	Cloning vector; Ap^r	Stratagene
pQE60	His_6 expression vector; Ap^r	Qiagen
pAC1000	*S. pneumoniae* suicide vector	This work
pCH84	pAC1000 *′malM*::*rlrA*::*cat*::*malP′*; Sm^r Cm^r	This work
pAC1279	pGEM-T *rlrA* RPA probe; Ap^r	This work
pAC1280	pGEM-T *rrgA* RPA probe; Ap^r	This work
pAC1281	pGEM-T *rrgB* RPA probe; Ap^r	This work
pAC1282	pGEM-T *rrgC* RPA probe; Ap^r	This work
pAC1283	pGEM-T *srtB* RPA probe; Ap^r	This work
pAC1284	pGEM-T *srtC* RPA probe; Ap^r	This work
pAC1285	pGEM-T *srtD* RPA probe; Ap^r	This work
pAC1286	pGEM-T *rpoB* RPA probe; Ap^r	This work
pAC1287	pGEM-T *rlrA-rrgA* promoter fragment	This work
pAC1288	pGEM-T *rrgB* promoter fragment	This work
pAC1289	pGEM-T *rrgC* promoter fragment	This work
pAC1290	pGEM-T *srtB* promoter fragment	This work
pAC1291	pGEM-T *srtC-srtD* promoter fragment	This work
pAC1292	pQE60 *rlrA*-His_6; Ap^r	This work
pAC1293	pGEM-T *srtA* RPA probe	This work

TABLE 2. Sequences of the primers used in this study

Primer	Sequence (5′ to 3′)
MALFX	CCCTCGAGTGAAAGCTATCGTGAGCAATT
MALRP	CCGAGCTCAAGATCTGGATCCTTATTTCTTTAAATCTACC
MALPF2	CCCTCTAGAGAGCATGCGACAATAATCAGGAGACAAC
MALPRP	CCGCGGCTCGAGTTCAAGAGGCCATTTTTCAAG
PCATF1	CCCGGTCTAGAGTCGACGGTATCGATAAGCT
PCATR1	CCGGCGCATGCTTATAAAAGCCAGTCATTAG
RLRAFR	CGCGGATCCAAAGGAGAATCATCATGCTAAACAAATACATTGA
RLRARX	CCCTCTAGATTATAACAAATAGTGAGCCTT
PEVPF1	GAGGATCCTATACCGCGGCCATGTCTGCCCGTATT
PEVPR1	TTCACCACCTTTTCCCTAT
RLRAF2	TTACATGCTGTTTTATCAATAA
RLRAR7	AGTAGAAAGAAGCGGAGTATT
RRGAF3	CACTTTTATACGCTTTTGCTA
RRGAR3	TAATACGACTCACTATAGGTGCCATCCGTATTGTTTTTC
RRGBF2	AAACTATCATTGAAAGGGGAG
RRGBR1	TAATACGACTCACTATAGGGGCATTGCCCTGAGAGTTTA
RRGCF2	GGCTGCGATTATGGGTATT
RRGCR2	TAATACGACTCACTATAGGGGTCATCTCAAACGAAGTCT
SRTBF2	AGGACTGGGATTCTGATTTA
SRTBR1	TAATACGACTCACTATAGGATCGCCACTCACTACATTATT
SRTCF2	GATTCTTTTATGGATTATTCG
SRTCR2	TAATACGACTCACTATAGGGACGCCTTTCTTTTTCTCTTG
SRTDF2	GCGGTCATCCTTCTCTTGCT
SRTDR2	TAATACGACTCACTATAGGGTCGTCAGACACTTGGTAAT
SRTAF1	AAAAGAAAAACAAGCGAAAAA
SRTAR1	TCCTTCTCCCATTACTTGCTC
RPOBF3	TGCTTATGACTTGGCAGCAG
RPOBR3	GGCTTTCAATGCTTTCAATC
RLRAPE2	AGTTAAAGTAGACAGTTCATC
RRGAP2	ACGGATTACTTATGTTCTGAT
RRGBPE	GCTGAAAACAGGCTACTCGCT
RRGCPE	CCATAACAAAGAAGATACGACTAAT
SRTBPB	TTTTAAATCAGAATCCCAGTC
SRTCPE	GCGAATCCTACTAAGAAAATC
SRTDPE	TATCCCAATAAGGCTCGTAG
RLRA2	TGTGTGACCCAATCCATACTT
RRGA2	CCCTGTTTGTGGATACTGGTC
RRGB2	GGGTTACGAGTTTACGAATGA
RRGC2	CAATTGACTAACCACCTCCTG
SRTBP1	TCAGCAGTACCAGCATAAACC
SRTBP2	TTAAAAATAACAAGCGACCAC
SRTCD1	CCAAAACAATAAATAGGAATC
SRTCD2	CAAGTGGATCAAGTAAAGGTG
RLRAC1	CCATGGTTCTAAACAAATACATTGAAAAAA
RLRAC2	AGATCTTAACAAATAGTGAGCCTTTTTA
REGF1	TCTAGACATGTGTGTCTCCCTGTT
IIR1	TCTAGACATAGTTACCGAATCTTAGTT
AP2	AACAACTTCCATCACAATAGA
AP3	AGGATAGTTAATAGTAATACTATAC
AP4	TAACTATCCTAGTATAAATTAAAAC
AP5	TAAAACTCCACCAATACTCAT
AP6	ATGAGTATTGGTGGAGTTTTA

viously described (10). The double-recombination event was selected by plating on CM and confirmed by PCR and DNA sequencing.

RPAs. Total RNA was isolated from 10 ml of exponential-phase *S. pneumoniae* with the Qiagen RNeasy kit in accordance with the manufacturer's (Qiagen) recommendations. Template DNA for the generation of riboprobes was PCR amplified with the primer sets RLRAF2/RLRAR7, RRGAF3/RRGAR3, RRGBF2/RRGBR1, RRGCF2/RRGCR2, SRTBF2/SRTBR1, SRTCF2/SRTCR2, SRTDF2/SRTDR2, SRTAF1/SRTAR1, and RPOBF3/RPOBR3. The resulting products were purified with a QIAquick PCR purification kit, subsequently cloned into pGEM-T (Promega), and confirmed by PCR with both an SP6 or T7 primer and a primer specific to the cloned insert. These plasmids (AC1279 to AC1286 and AC1293; Table 1) were used as templates for the generation of riboprobes as previously described (19). Synthesized probes were purified on a 4% denaturing polyacrylamide gel containing 7 M urea. RNase protection assays (RPAs) were carried out as described by the manufacturer with an RPAII kit (Ambion). The protected fragments were visualized by exposing each gel to a phosphorimaging screen (Kodak) and analyzed with a Storm 860 scanner and IQMac V1.2 imaging software. The relative amount of each protected fragment in each assay was normalized to the amount of *rpoB* protected RNA in each lane.

Northern blotting. Northern blots were performed with the NorthernMax analysis kit (Ambion) exactly as described by the manufacturer with 5 μg of total RNA. Riboprobes were synthesized as described above. Total RNA was separated on a 1% formaldehyde agarose gel by electrophoresis and subsequently transferred to Hybond-N+ nitrocellulose membranes. Membranes were then probed with 10^6 cpm of gel-purified riboprobe per ml of hybridization buffer and washed as recommended by the manufacturer (Ambion). Processed blots were exposed to a phosphorimaging screen (Kodak) and analyzed as described above.

Primer extension and DNA sequencing. Primer extension reactions were carried out with the avian myeloblastosis virus reverse transcriptase primer exten-

sion system (Promega). RNA was isolated from AC1278 as described above. A primer corresponding to the 5′ end of each coding sequence was end labeled with [γ-^{32}P]ATP with T4 polynucleotide kinase for 10 min at 37°C. The primers used were RLRAPE2, RRGAP2, RRGBPE, RRGCPE, SRTBPB, SRTCPE, and SRTDPE (Table 1). End-labeled primers were annealed to total RNA extracted from AC1278 by incubation at 58°C for 20 min, followed by cooling to room temperature for 10 min. Avian myeloblastosis virus extension mixture was added to each annealed primer, and cDNA synthesis was carried out at 42°C for 30 min.

Preliminary experiments were carried out with primers that were designed to be within 50 to 100 bp of the translation start site for each respective gene. In the case of *rrgA* and *srtB*, the resulting cDNA products were too large to be resolved by electrophoresis and only a single product was observed, indicating that no additional transcriptional start sites were present between the larger product and the translation start site (data not shown). Thus, two new primers (RRGAP2 and SRTBPB) were designed in the 5′ untranslated region of the RNA, resulting in products of approximately 150 bp.

DNA fragments predicted to contain promoter regions in the islet were PCR amplified from AC353 with the primer sets RLRA2/RRGA2, RRGB2/RRGBR1, RRGC2/RRGBF2, SRTBP1/SRTBP2, and SRTCD1/SRTCD2. PCR products were purified with the Stratagene PCR purification kit in accordance with the protocol provided (Stratagene), and purified products were subsequently cloned into pGEM-T (Promega) to generate plasmids AC1287, AC1288, AC1289, AC1290, and AC1291, respectively. DNA sequencing of *rlrA* pathogenicity islet promoter regions was performed with the Sequenase 2.0 DNA sequencing kit in accordance with the manufacturer's (USB) specifications. Briefly, strains AC1287, AC1288, AC1289, AC1290, and AC1291 were grown in 4 ml of Luria-Bertani broth and plasmid DNA was purified with a QIAGEN Plasmid Mini Kit (Qiagen). Plasmid DNA was resuspended in 100 μl of TE (10 mM Tris [pH 8.0], 1 mM EDTA) and subsequently denatured by the addition of 25 μl of 1 N NaOH–10 mM EDTA and incubation at 37°C for 30 min. Single-stranded DNA was ethanol precipitated by the addition of 1/10 volume of 3 M sodium acetate (pH 5.2) and 2 volumes of 100% ethanol. Precipitated DNA was resuspended in 1× Sequenase reaction buffer, and 60 pmol of the appropriate primer was annealed by incubation at 37°C for 30 min. Sequencing reactions were performed by the addition of a Sequenase 2.0 reaction mixture containing [α-^{35}S]dATP and incubation at room temperature for 5 min. Next, 3.5 μl of each reaction mixture was added to 2.5 μl of each dideoxynucleotide at 37°C and the termination reaction mixture was incubated for 5 min, at which time the reaction was stopped by the addition of stop solution.

Primer extension products and sequencing reactions were denatured for 10 min at 80°C prior to electrophoresis on a 5% polyacrylamide–7 M urea sequencing gel (National Diagnostics). Gels were run at 45 W, dried with a Bio-Rad 853 gel drying apparatus, and analyzed as described above.

RlrA-His$_6$ purification. The predicted coding sequence of RlrA was PCR amplified from AC353 with primers RLRAC1 and RLRAC2, subcloned into pGEM-T, and liberated by digestion with *Nco*I and *Bgl*II. The liberated fragment was ligated into similarly digested pQE60 to create AC1292. The resulting strain containing the coding sequence for RlrA with a C-terminal His$_6$ tag was grown in 2 ml of Luria-Bertani broth containing ampicillin to an optical density at 600 nm of 0.5, and expression of RlrA was induced by the addition of isopropyl-β-D-thiogalactopyranoside (IPTG) to 1 mM for 2 h. Proper expression of RlrA-His$_6$ was assessed by separation of induced and uninduced cell culture extracts by sodium dodecyl sulfate (SDS)-polyacrylamide gel electrophoresis and subsequently by Western blotting with anti-His$_6$ antibody (Roche) in accordance with the ECL Western blotting protocol (Amersham Pharmacia Biotech).

For the purification of RlrA-His$_6$, 2 liters of AC1292 was grown as described above and induced with IPTG for 2 h. RlrA-His$_6$ was subsequently purified on an Ni^{2+}-nitrilotriacetic acid agarose column in accordance with the manufacturer's (Qiagen) protocols under native conditions. Briefly, induced cell cultures were lysed by sonication for five cycles of 30 s with 15-s rests between cycles. The cellular debris was removed by centrifugation, and protein in the supernatant was bound to Ni^{2+}-nitrilotriacetic acid agarose beads for 1 h at 4°C. Bound protein was washed two times with wash buffer (50 mM NaH_2PO_4, 300 mM NaCl, 20 mM imidazole, pH 8.0) and subsequently eluted with 2 ml of elution buffer (50 mM NaH_2PO_4, 300 mM NaCl, 250 mM imidazole, pH 8.0) in 0.5-ml fractions. Fractions containing RlrA-His$_6$ were identified by Coomassie-stained gels and Western blotting with anti-His$_6$ antibody (Roche). RlrA-His$_6$-containing fractions were combined and concentrated with Centricon centrifugation filters (Amicon) to a final concentration of 800 nM.

Gel mobility shift assays of the *rrgA-rlrA* promoter region. Overlapping DNA fragments of the *rrgA-rlrA* intergenic region were amplified by PCR with the primer set REGF1-AP3, IIR1-AP5, AP4-AP6, or IIR1-AP4 (AP7) and used in gel mobility shift assays with RlrA-His$_6$. In each experiment, 60 pmol of a selected primer was end labeled with T4 polynucleotide kinase (New England Biolabs) and [γ-^{32}P]ATP (6,000 Ci/mmol, 150 mCi/ml) for 30 min at 37°C. Labeled primers were ethanol precipitated twice with ammonium acetate, resuspended in 10 μl of distilled H_2O, and used in PCRs with pAC1287 as the template. Amplified products were separated on a 4% polyacrylamide gel, gel purified, and eluted overnight in gel mobility shift assay elution buffer (0.5 mM ammonium acetate, 10 mM magnesium acetate, 1 mM EDTA, 0.1% SDS) at 37°C. Gel mobility shift assay binding reactions were carried out with 5,000 cpm of each probe with increasing concentrations of RlrA-His$_6$ at 30°C for 15 min in gel mobility shift assay binding buffer (20 mM Tris [pH 8.0], 50 mM KCl, 2 mM $MgCl_2$, 1 mM EDTA, 1 mM dithiothreitol, 0.05% Nonidet P-40, 5% glycerol) supplemented with 1 μg of poly(dI) · poly(dC)-poly(dI) · poly(dC) and bovine serum albumin as nonspecific inhibitors. For the supershift experiments, binding reaction mixtures were prepared as described above, chilled on ice, and incubated with 0.5 μg of anti-His$_6$ antibody (Roche) for 30 min on ice. Reaction mixtures were subsequently separated on 5% nondenaturing polyacrylamide gel at approximately 10 V cm^{-1}(Protogel; National Diagnostics) and visualized as described above.

DNase I footprinting. DNase I footprinting experiments were carried out with the gel mobility shift protocol with 2×10^4 cpm of each probe. Following protein binding, the concentrations of $MgCl_2$ and $CaCl_2$ were adjusted to 5 and 10 mM, respectively, and each reaction mixture was incubated with DNase I (0.5 to 2 U) for 1 min at room temperature. Reactions were stopped by the addition of stop solution (200 mM NaCl, 30 mM EDTA, 1% SDS), and the digested products were extracted with an equal volume of phenol and chloroform and subsequently ethanol precipitated. Precipitated DNA was resuspended in loading buffer (98% formamide, 10 mM EDTA, 0.1% bromophenol blue, 0.1% xylene cyanol) and separated on a 5% polyacrylamide–7 M urea sequencing gel (National Diagnostics). Sequencing reactions of the footprinted region were performed as described above with primers specific to the region.

Determination of RlrA consensus binding sites. A consensus RlrA binding site was determined by PRETTY (Genetics Computer Group software package) with the four RlrA binding sites determined by DNase I footprinting. The resulting consensus sequence was used to query the complete TIGR4 genomic sequence with FINDPATTERNS (Genetics Computer Group software package). The resulting sequences were analyzed to determine if the sequences were present in regions likely to contain *S. pneumoniae* promoters. Intergenic regions were examined for potential promoter sequences with the Neural Network Promoter Prediction program at the Berkeley Drosophila Genome Project (http://www.fruitfly.org/seq_tools/promoter.html), and the resulting output was visually examined for *S. pneumoniae* σ^{70} consensus −10 binding sites. Consensus sequences located within 300 bp of a putative translational start site and within 70 bp of putative +1 transcriptional start sites were considered as possible sites of RlrA binding and regulation.

RESULTS

RlrA is required for wild-type levels of expression of each gene in the islet. To assess the effect of an *rlrA* mutation on the steady-state levels of mRNA for each gene in the islet, RPAs were performed with RNA isolated from wild-type AC353 or AC1213, a strain that harbors a transposon insertion in *rlrA*. Riboprobes specific to each of the islet genes, as well as to *rpoB*, were synthesized. The *rpoB* gene, which codes for the β subunit of RNA polymerase, was used to probe the same RNA preparations as the *rlrA* islet probes to serve as a loading control in each experiment. In each case, the steady-state levels of the mRNAs of the genes were decreased in the *rlrA* strain compared with those in the wild-type strain, albeit to differing degrees (Fig. 2). The greatest decreases were observed for the *rrgB* and *rrgC* messages, which were reduced by 10- and 11-fold, respectively. The *rrgA* message was only decreased by 2.5-fold in AC1213, suggesting that *rrgA* is transcribed from a promoter distinct from *rrgB* or *rrgC*. Lastly, the *srtB*, *srtC*, and *srtD* mRNAs were also dependent upon RlrA, with observed decreases in mRNA levels of six-, seven-, and eightfold, respectively. It is worth noting that the *srtB* probe protected three different-size messages, suggesting the possibility that there are

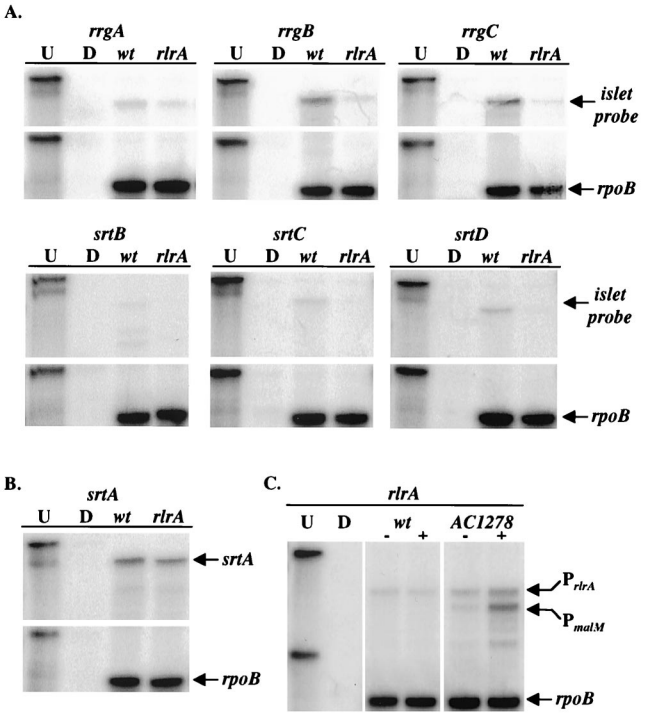

FIG. 2. RPAs were performed to analyze the steady-state mRNA levels of each gene in the *rlrA* pathogenicity islet in both the wild-type (*wt*; AC353) and *rlrA* mutant (AC1213) strain backgrounds. (A) Riboprobes to each gene in the islet, as well as to *rpoB*, were generated and hybridized to 10 μg of total *S. pneumoniae* RNA from either the wild-type or the mutant strain. Lanes U and D contained undigested riboprobes or riboprobes digested by RNase in the absence of *S. pneumoniae* RNA. In each case, the experimental probe for the given gene in the upper part of the panel and the control *rpoB* probe in the lower part of the panel are from the same gel. (B) Riboprobes to *srtA* and *rpoB* were hybridized to the same samples in panel A and presented in the same manner. (C) A riboprobe that differentially recognizes the two *rlrA* transcripts in AC1278 was used to determine if RlrA is autoregulatory. The larger fragment in each lane represents the mRNA from the native *rlrA* pathogenicity islet promoter. Lanes marked with a plus sign are RNA samples that were harvested from cells grown in the presence of maltose.

multiple transcriptional start sites within the sequence of the riboprobe.

To test for a possible role of RlrA in the regulation of *srtA*, the fourth sortase homologue in *S. pneumoniae* that is unlinked from the *rlrA* islet, a riboprobe specific to the *srtA* coding sequence was generated. As described above, an RPA was performed with total RNA harvested from either AC353 or AC1213. As shown in Fig. 2B, there was no difference in the amount of protected *srtA* message in either strain, indicating that *srtA* transcription occurs independently of RlrA.

RlrA is autoregulatory. In *S. pyogenes*, RofA positively regulates its own expression (5). To investigate the possibility that RlrA functions in a similar manner, a merodiploid strain that overexpressed *rlrA* from an inducible promoter was constructed (AC1278). This strain contained two copies of *rlrA*; one present in the *rlrA* pathogenicity islet and a second copy integrated into the maltose utilization operon downstream of *malM* (24). In the latter case, expression of *rlrA* was under the control of the *malM* promoter (P_{malM}), and thus, its expression was inducible by the addition of maltose to the growth medium (1). In addition, the Shine-Dalgarno site of *rpoB* was engineered into the *rlrA* construct upstream of the *rlrA* initiation codon to ensure maximal efficiency of RlrA translation from the maltose utilization locus.

To determine if overexpression of *rlrA* (from P_{malM}) activated transcription from the native *rlrA* promoter (P_{rlrA}), RPAs were performed with a single riboprobe specific for *rlrA* that differentiates between the two transcripts. The riboprobe was completely complementary to the P_{rlrA} transcript, as it overlapped the coding sequence of *rlrA* and the 5′ untranslated mRNA, resulting in a 409-bp protected band. Alternatively, the *rlrA* riboprobe was only partially complementary to the P_{malM} transcript and resulted in a smaller protected fragment since the sequence upstream of the *rlrA* coding sequence in this locus is different from that in the *rlrA* pathogenicity islet. Because of these differences, the two different-size protected messages detected with the same riboprobe were used to assess the quantity of steady-state mRNA from either of these promoters.

As shown in Fig. 2C, an increase in the amount of *rlrA* mRNA initiated from P_{rlrA} was observed in strain AC1278 compared to AC353 when each strain was grown in the absence of maltose. The increase observed in the absence of inducer compared to AC353 was due to the fact that AC1278 contained two copies of *rlrA* and transcription from P_{malM} is not completely repressed during growth in THY. A sixfold increase in expression from P_{rlrA} was observed in strain AC1278 compared to AC353 when each strain was grown in the presence of maltose. No increase in *rlrA* expression was observed when AC353 was grown in maltose compared to the same strain grown in THY, confirming that the increase in *rlrA* expression in AC1278 is not due simply to growth in the presence of maltose. Together, these data show that RlrA is autoregulatory in addition to activating the expression of the other six genes in the *rlrA* pathogenicity islet.

Transcription in the *rlrA* pathogenicity islet initiates at four different promoters. The finding that AC1213 (*rlrA*::*magellan2*) exhibited decreased levels of steady-state mRNA of different genes in the islet by differing levels led to the hypothesis that RlrA acts at numerous sites within the locus. To identify sites of transcription initiation and thus sites of potential RlrA activity, a primer specific to each gene in the locus was synthesized and used for primer extension analysis. By this method, transcription initiation sites upstream of *rlrA*, *rrgA*, *rrgB*, and *srtB* were identified (Fig. 3). By analysis of the sequences upstream of the predicted transcriptional start sites, σ^{70} consensus −10 and −35 sequences were identified for *rlrA* and an extended −10 sequence (25) was found for *rrgA* and *rrgB*; however, no such sequences were found for the *srtB* promoter (Fig. 3A and B). These results support the model in which there are multiple promoters within the islet and thus multiple sites at which RlrA acts.

Efforts to identify transcriptional start sites upstream of the predicted open reading frames of *rrgC*, *srtC*, and *srtD* proved unsuccessful. This suggested that each of these genes is transcribed from a distal promoter and that each is cotranscribed with an upstream gene(s). To test this, we carried out Northern blot assays with total RNA extracted from either AC1278 or AC1213 grown in THY-maltose. By using the same riboprobes

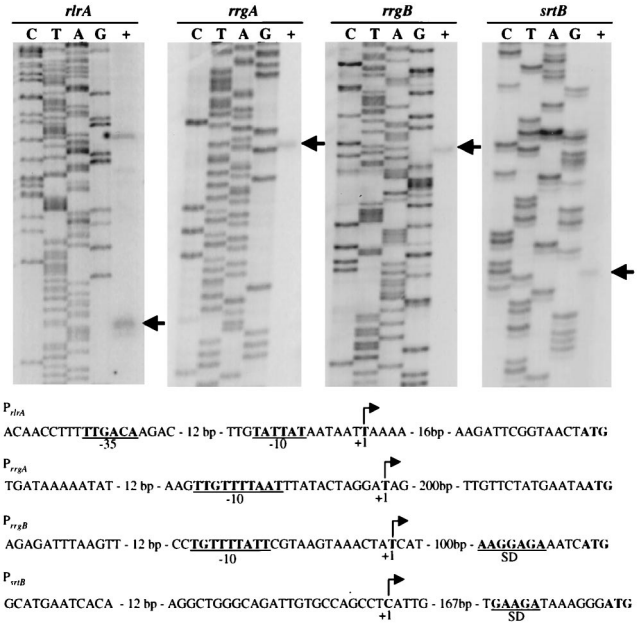

FIG. 3. (A) Transcriptional start sites of promoters upstream of *rlrA*, *rrgA*, *rrgB*, and *srtB* were mapped by primer extension analysis. The arrows indicate primer extension products. (B) Graphical depiction of the four *rlrA* pathogenicity islet promoters. A rightward arrow indicates the +1 start site. When present, −10 and −35 σ^{70} consensus sequences and predicted Shine-Dalgarno (SD) sequences are underlined and in bold.

that we used for RPAs, we found that the *rrgC* probe hybridized to an mRNA of approximately 3.8 kb, the predicted size of a mRNA that would include both *rrgC* and *rrgB*. In support of this, a Northern blot probed with *rrgB* indicated a message of the same size (Fig. 4A, lanes 1). No message corresponding to *rrgC* or *rrgB* was detected in the *rlrA* mutant strain, consistent with the RPA-based finding that transcription of both genes is dependent upon RlrA (Fig. 4A, lanes 2).

When the same RNA samples were probed with riboprobes complementary to *srtB*, *srtC*, and *srtD*, an mRNA of approximately 2.7 kb was detected with all three probes in the AC1278 background (Fig. 4B). A similar-size message was detected in

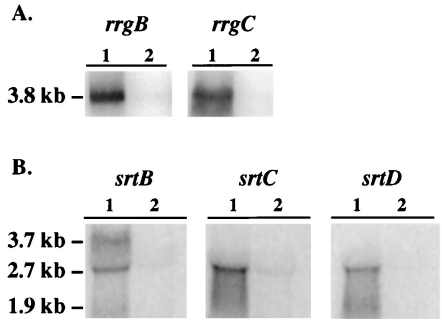

FIG. 4. Northern blot of *rlrA* pathogenicity islet mRNAs. Riboprobes to selected genes were synthesized and used to hybridize to total RNA recovered from AC1278 (lane 1) or AC1213 (lane 2) grown under maltose-inducing conditions. (A) Northern blots probed with *rrgB* and *rrgC* riboprobes. (B) Northern blots probed with *srtB*, *srtC*, and *srtD* riboprobes.

AC1213 at sharply decreased levels, although the same amount of RNA was loaded in each lane as determined by the quantity of rRNA on ethidium bromide-stained agarose gels (data not shown). An additional mRNA species of approximately 3.7 kb was also detected with the *srtB* probe that was not found with the *srtC* or *srtD* probe. Given the size of the message, its dependence on RlrA, and the position of the *srtB* riboprobe (which is predicted to overlap the terminal 3′ region of the *rrgBC* message), this mRNA is predicted to be the *rrgBC* message that terminates immediately upstream of the *srtB* coding sequence.

RlrA-His$_6$ acts at the *rrgA* and *rlrA* promoters. To determine if RlrA directly acts at one or more of the promoters in the *rlrA* pathogenicity islet, a C-terminally His$_6$-tagged version of RlrA was purified from *E. coli*. To test if RlrA-His$_6$ was able to bind to *rlrA* pathogenicity islet promoter sequences, the noncoding sequence between *rrgA* and *rlrA* was amplified by PCR with the primer set REGF1/IIR1 and an end-labeled REGF1 primer (Fig. 5A). The resulting fragment was incubated with purified RlrA-His$_6$ and separated on a nondenaturing polyacrylamide gel. In this gel mobility shift assay, RlrA-His$_6$ retarded the mobility of the probe evinced by the presence of multiple species that migrated more slowly on the gel than the probe alone (data not shown). This shows that RlrA-His$_6$ retains DNA binding activity and indicates that it binds to multiple sequences between *rrgA* and *rlrA*.

To more finely map the regions to which the purified protein bound, smaller overlapping fragments of the same region of DNA were generated by PCR and used in gel mobility shift experiments (Fig. 5B). When incubated with the AP4 fragment, RlrA-His$_6$ retarded the mobility of the probe, resulting in a single band that increased in intensity as the concentration of protein was increased (35% mobility shift at 4 nM and 70% mobility shift at 16 nM). A similar result was observed when RlrA-His$_6$ was incubated with the AP5 fragment; however, as the concentrations of protein were increased, two retarded species were observed (50% mobility shift at 4 nM RlrA-His$_6$). With both the AP4 and AP5 probes, as well as with the AP3 probe that spans the intergenic region downstream of the *rrgA* transcriptional start site, a retarded species running at the top of the gel was observed at high protein concentrations (RlrA-His$_6$, >130 nM; lanes 7, 13, and 14). We believe that this band is the result of nonspecific binding of RlrA-His$_6$ at high concentrations, an idea supported by the binding of RlrA-His$_6$ to nonpromoter regions of *rrgA* (AP3) and to two other *S. pneumoniae* promoters that are not regulated by RlrA (data not shown).

To confirm that the retarded mobility of the probe was due to the binding of RlrA-His$_6$ and not a contaminating species in the purified protein preparation, anti-His$_6$ antibody was added at the conclusion of the binding reaction to supershift RlrA-His$_6$-specific species. Figure 5C shows that incubation of RlrA-His$_6$-bound complexes with anti-His$_6$ antibody resulted in the appearance of a third complex migrating higher on the gel. This tertiary complex demonstrates that it is indeed RlrA-His$_6$ that is bound to the AP4 and AP5 probes. Together, these data suggest that RlrA-His$_6$ specifically binds to three distinct sites between the *rrgA* and *rlrA* transcription initiation sites, resulting in activation of transcription from both the *rrgA* and *rlrA* promoters.

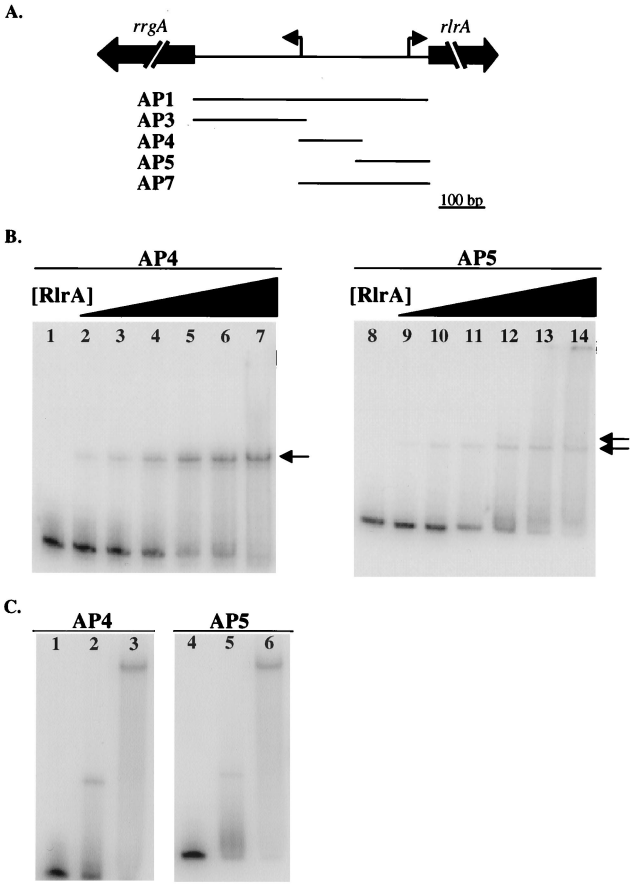

FIG. 5. Gel mobility shift analysis with RlrA-His$_6$. (A) The four ^{32}P-labeled probes that span the *rrgA-rlrA* intergenic region and were used in gel mobility shift analyses are depicted. The sizes of the PCR fragments were as follows: AP1, 522 bp; AP3, 250 bp; AP4, 139 bp; AP5, 163 bp; AP7, 290 bp. (B) Gel mobility shift analysis of AP4 and AP5. ^{32}P-labeled probes were incubated with increasing concentrations of RlrA-His$_6$. The protein concentrations used were as follows: lanes 1 and 8, 0 nM; lanes 2 and 9, 0.25 nM; lanes 3 and 10, 1 nM; lanes 4 and 11, 4 nM; lanes 5 and 12, 16.4 nM; lanes 6 and 13, 33 nM; lanes 7 and 14, 66 nM. The arrows indicate shifted species. (C) Supershift of RlrA-His$_6$ complexes by the addition of anti-His$_6$ antibody to the binding reaction mixture. The protein concentrations used were as follows: lanes 1 and 4, no protein; lanes 2 and 5, 16.4 nM RlrA-His$_6$; lanes 3 and 6, 16.4 nM RlrA-His$_6$ and 0.5 μg of anti-His$_6$ antibody.

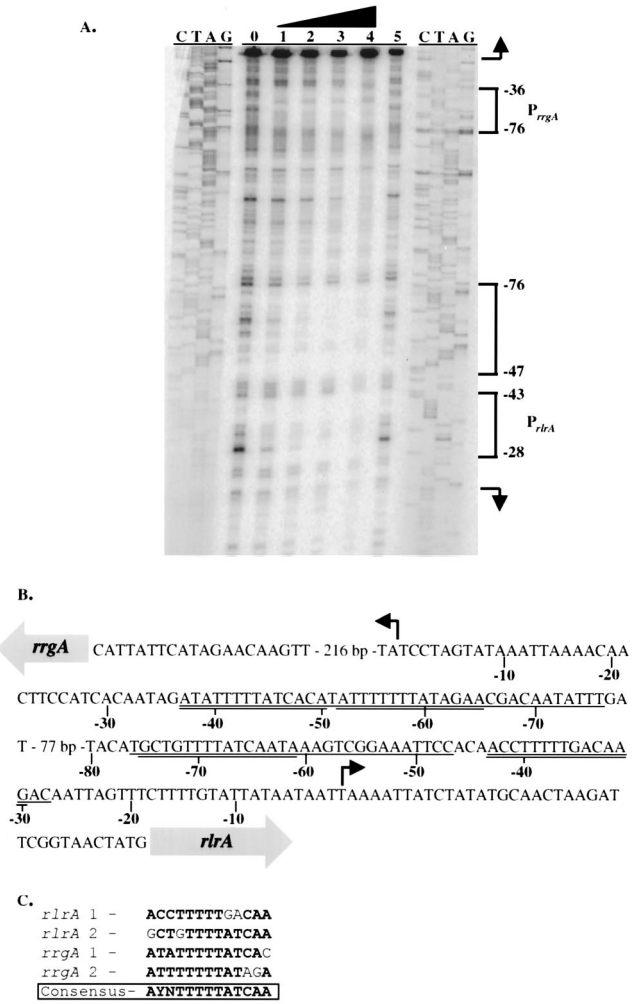

FIG. 6. (A) DNase I footprinting analysis of the *rrgA-rlrA* promoter regions. The ^{32}P-labeled AP7 probe was incubated with increasing amounts of RlrA-His$_6$ and subsequently treated with DNase I. The protein concentration used were as follows: lane 1, 0 nM; lane 2, 0.5 nM; lane 3, 2.05 nM; lane 4, 8.2 nM; lane 5, 32.8 nM. The amounts of DNase I used were as follows: lanes 1 and 2, 0.5 U; lane 3, 1 U; lanes 4 and 5, 2 U. Brackets indicate areas protected by RlrA-His$_6$. (B) The *rlrA* and *rrgA* promoter regions. The nucleotide sequence of each promoter is shown, and a hooked arrow denotes the +1 transcription initiation site. RlrA binding sites are singly underlined, and consensus binding sites are doubly underlined. (C) Alignment of the four consensus RlrA binding sites within the *rlrA* and *rrgA* promoter regions. Nucleotides that are identical to those in the consensus sequence are in bold, and the consensus sequence is in bold and boxed. In both cases, the sequence labeled number 1 is the consensus site closest to the transcriptional start site for that gene.

Determination of RlrA-His$_6$ binding sites. DNase I footprinting was used to precisely map the sites of RlrA binding in the *rrgA* and *rlrA* promoter regions. RlrA-His$_6$ was incubated with the AP7 fragment as described above, and the resulting bound complexes were subjected to DNase I digestion. Consistent with the findings of the gel mobility shift experiments, RlrA-His$_6$ protected three discrete regions of DNA (Fig. 6A). Two of these regions were present within 80 bp of the *rlrA* transcriptional start site (−28 to −43 and −47 to −76), and a third, larger region was present within 80 bp of the *rrgA* transcriptional start site (−36 to −76; Fig. 6A and B), which we believe constitutes two binding sites that are similar in arrangement to the *rlrA* binding sites. When the complementary strand of DNA was end labeled and used in the same assay, the same binding patterns were identified for both the *rlrA* and *rrgA* promoters (data not shown). Alignment of the protected regions revealed that RlrA binds to AT-rich regions close to the RNA polymerase binding site and transcriptional start site of each gene (Fig. 6B). The four binding sites identified were aligned, and a 13-bp RlrA consensus binding site was determined, as shown in Fig. 6C. The consensus sequence was used to query the TIGR4 genome sequence for additional possible RlrA binding sites. This search resulted in the identification of 153 sequences, 29 of which are present in putative promoter regions and 14 of which lie within 70 bp of the putative +1

transcriptional start site. These data suggest that RlrA may regulate additional genes outside of the *rlrA* pathogenicity islet.

DISCUSSION

The *rlrA* gene was initially identified as an essential gene for the colonization of *S. pneumoniae* in the murine nasopharynx and for its ability to infect the murine lung (10). In addition, several genes that are divergently transcribed from *rlrA* and lie within a 12-kb stretch of DNA that is flanked by two insertion elements have also been shown to be essential for either or both of these two models of infection (Fig. 1). The *rrgA* gene codes for a predicted cell wall-anchored protein of the LPXTG family of gram-positive surface proteins (10, 20). The LPXTG motif is part of a larger C-terminal sorting signal that targets the protein to a specific pathway that ultimately covalently anchors the protein to the cell wall (16). The enzymes that anchor proteins to the cell wall in this manner are called sortases. Sortases are transpeptidases that cleave between the threonine and glycine of the LPXTG motif, resulting in anchoring of the N-terminal half of the protein by a peptide bond between the threonine and the cell wall. Interestingly, also divergently transcribed from *rlrA* are three sortase homologues, *srtBCD* (Fig. 1). Two of these three genes have been shown to have a role during in vivo survival; *srtD* is essential for lung infection, and *srtB* is essential for colonization of the nasopharynx (10).

RlrA exhibits amino acid sequence similarity to a number of *S. pyogenes* transcriptional regulators, including RofA and Nra, a positive and a negative regulator, respectively. Both RofA and Nra regulate their own expression, as well as a number of different surface proteins that interact with eukaryotic extracellular matrices and thus are important to the pathogenesis of *S. pyogenes* (6, 8, 22). In each case, the gene divergently transcribed from the regulator is one target of regulation.

To determine if RlrA is a regulator of neighboring genes and of its own transcription, we used RPAs to measure the steady-state levels of transcription of each gene in the *rlrA* pathogenicity islet. We found that RlrA positively regulates the transcription of each gene (Fig. 2). The decrease in each message in the *rlrA* mutant strain was compared to that in the wild-type strain. On the basis of this analysis, *rlrA*-dependent expression fell into three categories: expression of the *rrgA* gene was only slightly affected, *srtBCD* expression was decreased to an intermediate level, and *rrgBC* expression was drastically reduced. The role of RlrA in its own expression was analyzed with a merodiploid strain that expressed *rlrA* from the *malM* promoter, allowing inducible expression in the presence of maltose. Analysis of this strain revealed that RlrA positively regulates its own transcription.

The different levels of expression of each gene in the islet suggest that RlrA regulates multiple promoters within the islet. Indeed, by using primer extension analysis, we mapped transcripts initiating upstream of *rlrA*, *rrgA*, *rrgB*, and *srtB*. A consensus σ^{70} −35 and −10 binding site was found upstream of *rlrA*, indicating that *rlrA* is expressed constitutively but may also be subject to positive autoregulation by increased RlrA levels under unknown conditions. In contrast, in three cases, *rrgA*, *rrgB*, and *srtB*, σ^{70} consensus −35 boxes could not be identified upstream of the transcriptional start sites; however, extended −10 sequences for *rrgA* and *rrgB* were identified. Previous studies of other *S. pneumoniae* promoters have shown that −35 consensus sequences cannot always be found within DNA fragments with known promoter activity (25). It is conceivable that genes such as those in the *rlrA* pathogenicity islet are transcribed by alternative σ factors, such as ComX, which regulates a subset of competence-induced genes (14). Comparison of the sequences upstream of the *srtB* promoter to the consensus *comX* box did not reveal an obvious binding site (21), indicating that this promoter is ComX independent and may be transcribed with an unknown sigma factor that is aided by RlrA binding. Alternatively, RlrA may enhance transcription by stimulating binding of σ^{70} RNA polymerase holoenzyme to the poor −35 elements in the *rrgA*, *rrgB*, and *srtB* promoters.

The identification of multiple promoters that are regulated by RlrA indicated that RlrA must bind multiple sites within the islet to regulate gene expression. This was indeed shown to be the case by gel mobility shift analyses and DNase I footprinting. In these experiments, RlrA was demonstrated to directly bind to four sites within the *rlrA-rrgA* intergenic region: two sites upstream of *rlrA* and two sites upstream of *rrgA*. In each case, there is a smaller RlrA binding site near the transcriptional start site and a larger binding site at a more distal location. A 13-bp consensus sequence in present in all four sites, and we propose that it is this sequence that is bound directly by RlrA. It is curious that the smaller site in each promoter overlaps the −35 sequence in each promoter, which is expected to be bound by RNA polymerase. As mentioned above, a consensus σ^{70} −35 promoter sequence could be identified in the *rlrA* promoter but not in the *rrgA*, *rrgB*, or *srtB* promoter. Thus, these data suggest that RlrA may compete with σ^{70} for the smaller binding site in the *rlrA* promoter, possibly when RlrA is expressed at high levels, resulting in repression of RlrA expression.

An interesting aspect of the biology of the *rlrA* pathogenicity islet is that it is not conserved among all pneumococcal serotypes (11, 26). Therefore, this islet may require a means of autonomous regulation, as we demonstrate here. On the other hand, it may seem unlikely that RlrA would regulate chromosomal genes outside the islet. However, we identified a number of putative targets of *rlrA* regulation outside the islet and scattered throughout the genome. It will be interesting to analyze these loci to see if they are indeed regulated by RlrA.

The *srtBCD* genes represent three of the four sortase homologues in the TIGR4 *S. pneumoniae* genome. The presence of multiple sortase homologues is a common occurrence in gram-positive bacteria genomes. The role of sortases in the anchoring of surface proteins important for the pathogenicity of various organisms is well documented. To our knowledge, however, prior to the finding that *srtBCD* are regulated by RlrA, only one other sortase has been shown to be regulated at the transcriptional level (17). The finding reported here that three of the four pneumococcal sortases are coordinately regulated by a single regulator suggests that RlrA may indirectly regulate the expression of numerous cell wall-anchored proteins by controlling sortase expression from a single promoter. It remains formally possible that the multiple sortase homologues in the *rlrA* pathogenicity islet do not have substrates that

lie outside of the islet. In this case, the role of SrtB, SrtC, and SrtD may be to specifically anchor one or more of the Rrg proteins to the cell wall. This would add a second, posttranslational level of autogenous regulation to the *rlrA* pathogenicity islet.

ACKNOWLEDGMENTS

We are indebted to Boris Belitsky and Linc Sonenshein for technical assistance and enlightening discussions concerning the promoter mapping and footprinting studies. We are grateful to Michael Angelichio and Julianna LeMieux for invaluable discussions and suggestions.

This work was supported by Pew Scholars Award P0168SC to A.C. and D.L.H. and by Welcome Trust Research Training Fellowship 061036 to C.J.H.

REFERENCES

1. **Acebo, P., C. Nieto, M. A. Corrales, M. Espinosa, and P. Lopez.** 2000. Quantitative detection of *Streptococcus pneumoniae* cells harbouring single or multiple copies of the gene encoding the green fluorescent protein. Microbiology **146:**1267–1273.
2. **Charpentier, E., and E. Tuomanen.** 2000. Mechanisms of antibiotic resistance and tolerance in *Streptococcus pneumoniae*. Microbes Infect. **2:**1855–1864.
3. **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.
4. **Crook, D. W., and B. G. Spratt.** 1998. Multiple antibiotic resistance in *Streptococcus pneumoniae*. Br. Med. Bull. **54:**595–610.
5. **Fogg, G. C., and M. G. Caparon.** 1997. Constitutive expression of fibronectin binding in *Streptococcus pyogenes* as a result of anaerobic activation of *rofA*. J. Bacteriol. **179:**6172–6180.
6. **Fogg, G. C., C. M. Gibson, and M. G. Caparon.** 1994. The identification of *rofA*, a positive-acting regulatory component of *prtF* expression: use of an mγδ-based shuttle mutagenesis strategy in *Streptococcus pyogenes*. Mol. Microbiol. **11:**671–684.
7. **Gillespie, S. H., and I. Balakrishnan.** 2000. Pathogenesis of pneumococcal infection. J. Med. Microbiol. **49:**1057–1067.
8. **Granok, A. B., D. Parsonage, R. P. Ross, and M. G. Caparon.** 2000. The RofA binding site in *Streptococcus pyogenes* is utilized in multiple transcriptional pathways. J. Bacteriol. **182:**1529–1540.
9. **Hanahan, D.** 1983. Studies on transformation of Escherichia coli with plasmids. J. Mol. Biol. **166:**557–580.
10. **Hava, D. L., and A. Camilli.** 2002. Identification of tissue-specific pneumococcal virulence factors. Mol. Microbiol. **45:**1389–1405.
11. **Hoskins, J., W. E. Alborn, Jr., J. Arnold, L. C. Blaszczak, S. Burgett, B. S. DeHoff, S. T. Estrem, L. Fritz, D. J. Fu, W. Fuller, C. Geringer, R. Gilmour, J. S. Glass, H. Khoja, A. R. Kraft, R. E. Lagace, D. J. LeBlanc, L. N. Lee, E. J. Lefkowitz, J. Lu, P. Matsushima, S. M. McAhren, M. McHenney, K. McLeaster, C. W. Mundy, T. I. Nicas, F. H. Norris, M. O'Gara, R. B. Peery, G. T. Robertson, P. Rockey, P. M. Sun, M. E. Winkler, Y. Yang, M. Young-Bellido, G. Zhao, C. A. Zook, R. H. Baltz, S. R. Jaskunas, P. R. Rosteck, Jr., P. L. Skatrud, and J. I. Glass.** 2001. Genome of the bacterium *Streptococcus pneumoniae* strain R6. J. Bacteriol. **183:**5709–5717.
12. **Kolter, R., M. Inuzuka, and D. R. Helinski.** 1978. Trans-complementation-dependent replication of a low molecular weight origin fragment from plasmid R6K. Cell **15:**1199–1208.
13. **Lau, G. W., S. Haataja, M. Lonetto, S. E. Kensit, A. Marra, A. P. Bryant, D. McDevitt, D. A. Morrison, and D. W. Holden.** 2001. A functional genomic analysis of type 3 *Streptococcus pneumoniae* virulence. Mol. Microbiol. **40:**555–571.
14. **Lee, M. S., and D. A. Morrison.** 1999. Identification of a new regulator in *Streptococcus pneumoniae* linking quorum sensing to competence for genetic transformation. J. Bacteriol. **181:**5004–5016.
15. **Mazmanian, S. K., G. Liu, H. Ton-That, and O. Schneewind.** 1999. *Staphylococcus aureus* sortase, an enzyme that anchors surface proteins to the cell wall. Science **285:**760–763.
16. **Mazmanian, S. K., H. Ton-That, and O. Schneewind.** 2001. Sortase-catalysed anchoring of surface proteins to the cell wall of *Staphylococcus aureus*. Mol. Microbiol. **40:**1049–1057.
17. **Mazmanian, S. K., H. Ton-That, K. Su, and O. Schneewind.** 2002. An iron-regulated sortase anchors a class of surface protein during *Staphylococcus aureus* pathogenesis. Proc. Natl. Acad. Sci. USA **99:**2293–2298.
18. **McCullers, J. A., and E. I. Tuomanen.** 2001. Molecular pathogenesis of pneumococcal pneumonia. Front. Biosci. **6:**D877–D889.
19. **Merrell, D. S., and A. Camilli.** 2000. Regulation of *Vibrio cholerae* genes required for acid tolerance by a member of the "ToxR-like" family of transcriptional regulators. J. Bacteriol. **182:**5342–5350.
20. **Pallen, M. J., A. C. Lam, M. Antonio, and K. Dunbar.** 2001. An embarrassment of sortases—a richness of substrates? Trends Microbiol. **9:**97–101.
21. **Pestova, E. V., and D. A. Morrison.** 1998. Isolation and characterization of three *Streptococcus pneumoniae* transformation-specific loci by use of a *lacZ* reporter insertion vector. J. Bacteriol. **180:**2701–2710.
22. **Podbielski, A., M. Woischnik, B. A. Leonard, and K. H. Schmidt.** 1999. Characterization of *nra*, a global negative regulator gene in group A streptococci. Mol. Microbiol. **31:**1051–1064.
23. **Polissi, A., A. Pontiggia, G. Feger, M. Altieri, H. Mottl, L. Ferrari, and D. Simon.** 1998. Large-scale identification of virulence genes from *Streptococcus pneumoniae*. Infect. Immun. **66:**5620–5629.
24. **Puyet, A., and M. Espinosa.** 1993. Structure of the maltodextrin-uptake locus of *Streptococcus pneumoniae*: correlation to the Escherichia coli maltose regulon. J. Mol. Biol. **230:**800–811.
25. **Sabelnikov, A. G., B. Greenberg, and S. A. Lacks.** 1995. An extended −10 promoter alone directs transcription of the DpnII operon of *Streptococcus pneumoniae*. J. Mol. Biol. **250:**144–155.
26. **Tettelin, H., K. E. Nelson, I. T. Paulsen, J. A. Eisen, T. D. Read, S. Peterson, J. Heidelberg, R. T. DeBoy, D. H. Haft, R. J. Dodson, A. S. Durkin, M. Gwinn, J. F. Kolonay, W. C. Nelson, J. D. Peterson, L. A. Umayam, O. White, S. L. Salzberg, M. R. Lewis, D. Radune, E. Holtzapple, H. Khouri, A. M. Wolf, T. R. Utterback, C. L. Hansen, L. A. McDonald, T. V. Feldblyum, S. Angiuoli, T. Dickinson, E. K. Hickey, I. E. Holt, B. J. Loftus, F. Yang, H. O. Smith, J. C. Venter, B. A. Dougherty, D. A. Morrison, S. K. Hollingshead, and C. M. Fraser.** 2001. Complete genome sequence of a virulent isolate of *Streptococcus pneumoniae*. Science **293:**498–506.
27. **Tomasz, A.** 1997. Antibiotic resistance in *Streptococcus pneumoniae*. Clin. Infect. Dis. **24:**S85–S88.