JOURNAL OF BACTERIOLOGY, Aug. 1996, p. 4563–4570 Vol. 178, No. 15
0021-9193/96/$04.00+0

The Molecular Architecture of the *sar* Locus in *Staphylococcus aureus*

MANFRED G. BAYER, JON H. HEINRICHS, AND AMBROSE L. CHEUNG*

Laboratory of Bacterial Pathogenesis and Immunology, The Rockefeller University, New York, New York 10021

Received 22 January 1996/Accepted 16 May 1996

The global regulator *sar* in *Staphylococcus aureus* controls the synthesis of a variety of cell wall and extracellular proteins, many of which are putative virulence factors. The *sar* locus in strain RN6390 contains a 339-bp open reading frame (*sarA*) and an 860-bp upstream region. Transcriptional analyses of this locus revealed three different transcripts of 0.58, 0.84, and 1.15 kb (designated *sarA*, *sarC*, and *sarB*, respectively). All three transcripts seemed to be under temporal, growth cycle-dependent regulation, with *sarA* and *sarB* being most abundant in early log phase and the *sarC* concentration being highest toward the late stationary phase. Mapping of the 5′ ends of the *sar* transcripts by primer extension and modified S1 nuclease protection assays demonstrated that transcription is initiated from three separate, widely spaced promoters. The 3′ ends of all three *sar* transcripts are identical, and transcriptional termination occurs upstream of a typical prokaryotic poly(T) termination signal. Northern (RNA) analysis of *sar* mutant clones containing plasmids that comprised various promoters and the termination signal revealed that individual transcripts can be generated from each of the three promoters, thus suggesting possible activation as independent promoters. The multipromoter system, from which transcription is initiated, bears conserved features for recognition by homologous σ^{70} transcription factors and also by those expressed in the general stress response. Downstream of the two distal promoters (P3 and P2) are two regions potentially encoding short peptides. It is conceivable that posttranslational cooperation between these short peptides and the *sarA* gene product occurs to modulate *sar*-related functions. Complementation studies of a *sar* mutant with a clone expressing all three *sar* transcripts showed that this clone was able to restore the *sar* wild-type phenotype to the *sar* mutant.

Staphylococcus aureus is a major human pathogen (30). Infections caused by this organism range from abscesses, endocarditis, pneumonia, meningitis, and sepsis to severe toxemia including toxic shock syndrome. The recent increase in antibiotic resistance and a lack of potential vaccine candidates have highlighted the importance in finding new approaches to control this important pathogen.

Infections caused by *S. aureus* are probably related to the organism's striking capability to react to changing environments during the infection process. In any stage of surface colonization, entry, and invasion, this highly coordinated response seems to be modulated by the expression of appropriate genes via signal transduction pathways. These modulators may interact with other regulatory elements which, in turn, control the transcription of a wide variety of unlinked genes, many of which are involved in pathogenesis (12, 30). Accordingly, an understanding of the genetic basis in the expression of virulence factors in *S. aureus* is a prerequisite to the identification of target sites for the development of novel antimicrobial agents based on the global regulatory networks.

Temporal expression of many of the proteins involved in virulence has been shown to be under the control of global regulatory systems in which a common regulator directs the expression of multiple genes. A global regulator, called *agr*, has been shown to play an important role in regulating the synthesis of extracellular and cell wall proteins in *S. aureus* (19). The *agr* locus is composed of two divergent transcripts, RNAII and RNAIII, with sizes of 3 and 0.8 kb, respectively. Complementation analysis indicates that RNAIII is the *agr* regulatory molecule and is responsible for the restoration of parental phenotypes to *agr* mutants (19).

During the growth cycle, many of the exoproteins (e.g., α-hemolysin and toxic shock syndrome toxin) are secreted during the postexponential phase (12), while most surface proteins (e.g., protein A and coagulase- and fibronectin-binding proteins) are synthesized primarily during the exponential phase and are repressed postexponentially (12, 19). In *agr* mutants, however, the synthesis of many extracellular proteins is downregulated while the expression of surface proteins is upregulated. Recently, our laboratory has discovered a new regulatory locus, designated *sar*, which is also involved in the expression of extracellular and cell wall proteins (9). A gene within the *sar* locus of strains DB and RN450, called *sarA*, has been cloned and sequenced, revealing an open reading frame (ORF) of 372 bp. Transcriptional studies indicated that the *agr* locus itself is in part under the control of *sar* (15).

In a previous study, we showed that the complementation of *sar* mutants with a 732-bp DNA fragment encoding the entire *sarA* gene resulted in only partial restoration of the parental phenotype. In contrast, a derivative clone carrying a 1.7-kb DNA fragment encompassing *sarA* and a 1.2-kb upstream fragment was able to restore the production of exoproteins to parental levels in a *sar* mutant. More importantly, cell extracts of this complemented mutant were found to bind to the P2 promoter region of the *agr* locus, which controls RNAII transcription. As the P2 gene products have been postulated to promote the transcription of the *agr* regulatory molecule RNAIII (19), the *sar* locus may regulate exoprotein synthesis by controlling the *agr*-mediated pathway.

In this study, we analyzed the nucleotide sequence of *sarA* of strain RN6390, the upstream sequence of *sarA*, and the termination signal downstream of *sar*. In analyzing the transcriptional organization of the *sar* locus, we present data that three

* Corresponding author. Mailing address: The Laboratory of Bacterial Pathogenesis and Immunology, The Rockefeller University, 1230 York Ave., New York, NY 10021. Phone: (212) 327-8163. Fax: (212) 327-7385. Electronic mail address: cheunga@rockvax.rockefeller.edu.

TABLE 1. Bacterial strains and plasmids used

Strain or plasmid	Reference or source	Comments
S. aureus		
RN4220	24	A mutant of strain 8325-4 that accepts foreign DNA
RN6390	24	Laboratory strain that maintains its hemolytic pattern when propagated on sheep erythrocytes
ALC136	11	Isogenic mutant of RN6390 carrying a *sar*::Tn*917*LTV1 mutation
ALC103	15	ALC136 with pALC103
ALC556	This work	ALC136 with pALC556
ALC557	This work	ALC136 with pALC557
ALC561	This work	ALC136 with pALC561
ALC529	This work	ALC136 with pALC529
E. coli		
XL1-Blue	22	Highly transformable strain
InvαF′	Invitrogen	Host strain for the TA cloning vector (pCRII)
DU5384	11	Strain carrying a pBR322 plasmid with a 3-kb *Eco*RI-*Hin*dIII fragment of the α-hemolysin gene
RN6929	11	JM109 containing a pBluescript phagemid with a 2.2-kb *Cla*I fragment of the β-hemolysin gene
Plasmids		
pCRII	Invitrogen	*E. coli* cloning vector for direct cloning of PCR fragments
pSPT181	17	Shuttle vector
pALC103	15	pSPT181 with *sar* fragment nt[a] 620–1349
pALC556	This work	pSPT181 with *sar* fragment nt 531–1349
pALC557	This work	pSPT181 with *sar* fragment nt 364–1349
pALC561	This work	pSPT181 with *sar* fragment nt 1–1349
pALC529	This work	pSPT181 with *sar* fragment nt 1–1231 and additional 300 bp upstream sequence

[a] nt, nucleotides.

overlapping transcriptional units originate from the *sar* locus and are controlled in a temporal manner by a parallel multiple-promoter system. In addition, we demonstrated that the expression of the overlapping transcriptional *sar* units was required to restore parental phenotypes to the *sar* mutant.

MATERIALS AND METHODS

Bacterial strains, plasmids, and growth media. The bacterial strains and plasmids used in this study are listed in Table 1. Phage ϕ11 was used as the transducing phage for *S. aureus* strains. CYGP and 0.3GL media (25) were used for the growth of *S. aureus*, while Luria-Bertani broth (LB) was used for growing *Escherichia coli*. Antibiotics were used at the following concentrations: erythromycin at 10 μg/ml, tetracycline at 5 μg/ml, and ampicillin at 50 μg/ml.

Sequence analysis. A pBluescript phagemid with a 6-kb insert encompassing the *sar* locus was used. Plasmids were purified by the Magic Midiprep procedure (Promega, Madison, Wis.). By using both T3 and T7 primers, bidirectional plasmid sequencing was performed and completed by primer walking with a Sequenase kit (version 2.0; U.S. Biochemicals, Cleveland, Ohio) according to the dideoxy-chain termination method.

Construction of 5′ and 3′ deletion clones. DNA fragments of the *sar* locus at several positions (Fig. 1), 1 to 1349, 364 to 1349, 531 to 1349, and 1 to 1231 with an additional 300 bp of upstream sequence, were amplified by PCR using genomic DNA of *S. aureus* RN6390 as the template and cloned into the TA cloning vector pCRII (Invitrogen, San Diego, Calif.). The authenticity of the amplified fragments was confirmed by DNA sequencing. The PCR fragments were cleaved from pCRII, ligated to the shuttle vector pSPT181, and transformed into *E. coli* XL1-Blue. The presence of the correct inserts in the plasmids was confirmed by restriction mapping and sequencing.

Protoplast transformation of *S. aureus* RN4220 with shuttle plasmids containing various *sar* fragments was performed as previously described (5). For transduction, phage ϕ11 was used to produce a phage lysate of strain RN4220 containing the modified pSPT181 shuttle vector with various *sar* DNA fragments. The phage lysate was then used to infect the *sar* transposon mutant (ALC136; previously designated mutant R) as described previously (9), producing ALC561, ALC557, ALC556, and ALC529 (Table 1).

Phenotypic characterization. A *sar* mutant clone carrying a shuttle plasmid with a *sar* fragment was tested in duplicate for the production of hemolysins as previously described (9). To determine levels of protein A production, cell wall-associated proteins were extracted from overnight cultures of *S. aureus* with lysostaphin in a hypertonic medium (30% raffinose) as previously described (8). Equivalent volumes (1 μl each) of cell wall protein extracts were separated on sodium dodecyl sulfate–10% polyacrylamide gels, electroblotted onto nitrocellulose, and probed with chicken anti-staphylococcal protein A antibody (Accurate Chemicals, Westbury, N.Y.) at a 1:3,000 dilution. Bound antibody was detected with rabbit anti-chicken immunoglobulin G conjugated to alkaline phosphatase (Jackson Immunoresearch, West Grove, Pa.) (1:5,000 dilution) and visualized as described by Blake et al. (1).

The ability of whole cells to bind to ^{125}I-labeled fibronectin was assayed as described previously (7). The Student *t* test was used to compare the binding of radiolabeled fibronectin to parental and *sar* mutant strains. *P* values of ≤0.05 were considered significant.

Isolation of RNA and Northern (RNA) analysis. Overnight cultures of *S. aureus* were diluted 1:100 in CYPG and grown to early log, mid-log, late log, and postexponential phases. The cells were pelleted and processed with a FastRNA isolation kit (Bio 101, Vista, Calif.) in combination with 0.1-mm-diameter zirconia-silica beads and a FastPrep reciprocating shaker (Bio 101) as described previously (6). Ten micrograms of each sample was electrophoresed through a 1.5% agarose–0.66 M formaldehyde gel in morpholinepropanesulfonic acid (MOPS) running buffer (20 mM MOPS, 10 mM sodium acetate, 2 mM EDTA [pH 7.0]). Blotting of RNA onto a Hybond N membrane (Amersham, Arlington Height, Ill.) was performed with the Turbo Blotter alkaline transfer system (Schleicher & Schuell, Keene, N.H.). For detection of specific transcripts, DNA probes were radiolabeled with [α-^{32}P]dCTP (Amersham) by the random prime method (Ready to Go labeling kit; Pharmacia, Piscataway, N.J.) and hybridized under high-stringency conditions. The blots were subsequently autoradiographed. In some blots, band intensities were quantitated by densitometric scanning with SigmaGel software (Jandel Scientific, San Rafael, Calif.); these values are presented as integrated area units.

Primer extension analysis. Mapping of the 5′ ends of the *sarA*, *-B*, and *-C* transcripts by primer extension was performed with a series of synthesized oligonucleotides complementary to both DNA strands. The 23-mer primers used in Fig. 3A to C correspond to the complementary strand of the sequence given in Fig. 1 as follows: 5′-TAAGTGACCATTGATAACAACTC-3′ (positions 918 to 896), 5′-TGTTTCGATATTCAATGTGTGTC-3′ (positions 525 to 503), and 5′-TGTCTATACGTTATTCCGATTGA-3′ (positions 309 to 287), respectively. The primers were end labeled with [γ-^{32}P]ATP and purified by Sephadex G-25 spin columns (Boehringer Mannheim, Indianapolis, Ind.).

For primer extension, 30 μg of RNA was coprecipitated with the appropriate end-labeled primer (5×10^4 cpm) and then annealed at 35°C overnight. The reaction mixture was ethanol precipitated, washed with 70% ethanol, and dried. Reverse transcription was carried out by using SuperScriptII (Gibco-BRL, Gaithersburg, Md.) at 42°C for 90 min, and the reaction mixture heated at 65°C for 10 min to inactivate the enzyme. The reaction product was incubated with RNase H (2 U) for 15 min at 55°C, ethanol precipitated, resuspended in 10 μl of Sequenase stop solution, denatured, and applied onto a 4% sequencing gel. Sequencing reaction mixtures primed by an oligonucleotide identical to that used for primer extension were applied in parallel lanes on the gel.

Mapping of transcriptional start and stop sites. Mapping of the 5′ and 3′ ends of different *sar* mRNAs was performed with several single-stranded riboprobes complementary to the appropriate sequence of the transcripts. All probes were tested on sequencing gels for their integrity prior to S1 analysis.

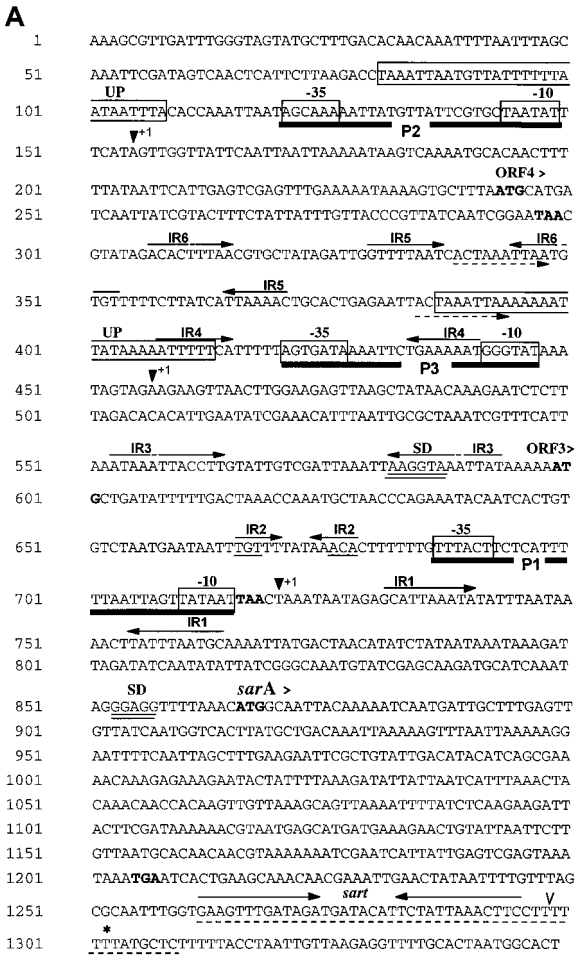

FIG. 1. (A) Nucleotide sequence of the *sar* regulatory region. Shown are the region encoding *sarA*, an 890-bp upstream sequence, and a sequence 200 bp downstream of the *sarA* stop codon. Putative ribosome-binding sites are emphasized with a double line and labeled SD. Start and stop codons of *sarA*, ORF3, and ORF4 are in boldface. The mapped 5′ ends of the mRNAs identified by primer extension and S1 nuclease protection are indicated by solid arrowheads and labeled +1. The mapped 3′ end is marked by an open arrowhead. The termination site predicted by a computer program is identified by an asterisk. The termination signal (*sart*) forming a stem-loop is labeled by a broken line. Putative promoter regions are highlighted by boxed sequences and labeled −10 and −35 (core promoter) and UP (extended promoter). The promoters are designated P1, P3, and P2. Broken arrows below the sequence indicate direct repeats; solid arrows above the sequence mark a structure of dyad symmetry and are numerically designated with the prefix IR (inverted repeat). The potential core target site known for DNA-binding proteins is underlined. (B) Deduced amino acid sequence of SarA protein of strain RN6390. The difference in composition in comparison with strain DB (10) is indicated by the boldface underlined letter. In addition, the 11 C-terminal amino acids of the SarA protein of strain DB are absent in strain RN6390 because of a nonsense mutation.

To generate the 3′ riboprobe, a 358-bp fragment (positions 992 to 1349) covering ~200 bp of the 3′ end of the *sar*-encoded sequence plus a 160-bp downstream fragment including the putative transcription termination signal was amplified by PCR using the primer 5′-ATCAGCGAAAACAAAGAGAAAGAAT-3′ (positions 992 to 1016) and 5′-AGTGCCATTAGTGCAAAAC-3′ (positions 1349 to 1331). The PCR product was cloned into pCRII, transformed to XL1-Blue, and sequenced to determine the orientation of the insert. Using this modified pCRII vector as a template, a second PCR was performed with the vector-specific M13 reverse primer and a *sar*-specific primer identical to positions 992 to 1016, thus generating a PCR product that carries a SP6 phage promoter in front of the *sar*-specific probe encompassing the 3′ end. SP6-driven in vitro transcription of a single-stranded probe complementary to the 3′ end of the mRNA was performed by using a standard transcription assay (MAXIscript SP6/T7; Ambion, Austin, Tex.). To ensure the synthesis of a full-length probe, the radiolabeled [α-^{32}P]UTP was supplemented with unlabeled UTP to a final concentration of 7.5 μM. The reaction mixture was incubated at 16°C for 2 h and then subjected to RNase-free DNase I treatment at 37°C for 15 min.

In an additional approach to map the 5′ ends of different *sar* mRNAs, we used a similar procedure to generate single-stranded 5′ riboprobes. Specific DNA fragments of approximately 350 bp in length encompassing the putative 5′ end of various transcripts were used as templates for PCR amplification. After cloning of the PCR fragments into pCRII, the inserts were subjected to a second PCR with a T7 phage promoter-specific primer, thus generating a T7 phage promoter in front of the *sar* probe. In vitro transcription for the synthesis of full-length antisense RNA probes was allowed to proceed at 10°C for 3 h.

The specific activity of probes generated this way usually ranged between 2 and 8×10^7 cpm/μg of DNA template.

Modified S1 nuclease protection assay. Ten micrograms of RNA was mixed with 10^5 cpm of riboprobe, heated to 85°C for 3 min, and then incubated in hybridization buffer [40 mM piperazine-*N*,*N*′-bis(2-ethanesulfonic acid (PIPES; pH 6.4), 400 mM NaCl, 1 mM EDTA, 80% (wt/vol) formamide] at 42°C overnight. Digestion of unprotected fragments was performed in 30 mM sodium acetate (pH 4.8)–10 mM $ZnSO_4$–5% glycerol–100 mM NaCl–60 U of S1 nuclease at 37°C for 30 min. After ethanol precipitation, the protected RNA hybrids were resuspended in water and gel loading buffer, heat denatured, and run on a 4% sequencing gel. For the exact determination of the 3′ and 5′ ends of mRNA, a concomitant sequencing reaction using a primer identical to the ones used for the generation of riboprobes was applied onto parallel lanes and subsequently autoradiographed.

Computer analysis. Sequence analysis was performed with the Wisconsin Genetics Computer Group, Inc., package. For the prediction of prokaryotic termination signals, we used the program TERMINATOR (3, 4) from the HUSAR package of the DFKZ, Heidelberg, Germany.

Nucleotide sequence accession number. The nucleotide sequence reported here is available from GenBank under accession number U46541.

RESULTS

In previous studies, we identified the *sar* locus as a global regulator of several virulence determinants in *S. aureus* (7, 9, 11). In particular, the *sar* locus may regulate hemolysin production by controlling RNAII and RNAIII transcription in the *agr* locus (10, 15). Recently, we found that the *sarA* gene together with a 1.2-kb upstream fragment was required to complement *agr*-related transcription in a *sar* mutant (15). To determine the genetic requirements for *sar* function and transcription, we sequenced and analyzed the *sarA* gene in strain RN6390 and the region upstream.

Sequence analysis of *sarA* and the upstream region. The nucleotide sequence of the 1.35-kb fragment encompassing *sarA* is given in Fig. 1. The ORF encoding the *sarA* gene begins at position 866 with a start codon and extends to position 1204 followed by a stop codon. A highly conserved Shine-Dalgarno sequence is centered 11 bp upstream of the initiation codon. Within the *sarA* gene, two 10-bp direct repeats are located at positions 927 and 940 (Fig. 1A). The mature protein has a predicted molecular size of 13,469 Da and a deduced basic pI of 9.37. With the exception of a slightly truncated C terminus and a conservative substitution at residue 53 in which a leucine is exchanged for phenylalanine, this SarA protein is highly homologous to that of strains RN450 and DB (Fig. 1B) (13). As with the SarA protein for strain RN450, this protein also has a preponderance of charged amino acids (34.3%).

Examination of the 865-bp DNA sequence upstream of *sarA* indicates that this region contains no additional large ORF. To exclude possible sequencing errors, both strands have been sequenced several times by both the dideoxy-chain termination protocol and ABI dye termination fluorescence-based cycle sequencing. A computer analysis using the programs TESTCODE and CODONPREFERENCE revealed neither a pattern of alternating ORFs located on different frames nor an

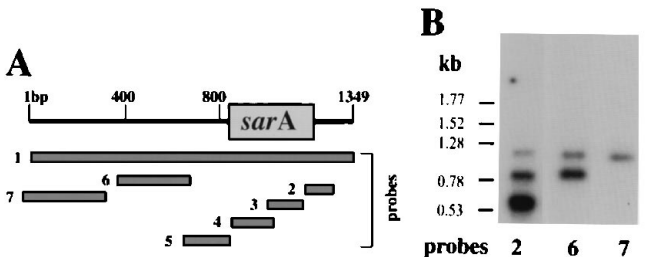

FIG. 2. Transcriptional analysis of *sar*. (A) Graphic representation of the location of the probes within the sequences given in Fig. 1. (B) Northern blot of *sar* transcripts by probe walking along the *sar* locus with PCR-generated DNA fragments. The transcriptional pattern obtained with probe 2 is identical to that from probes 1, 3, 4, and 5.

extended *sarA* coding region. However, two additional small ORFs potentially encoding peptides of 39 and 18 amino acids (Fig. 1) located at positions 599 to 715 and 243 to 296, respectively, could be detected. The region immediately upstream of ORF3 is defined by a putative but weakly conserved ribosome-binding site 11 bp upstream of a possible initiation codon together with features that are in compliance with the rules obtained from translational initiation sites in *E. coli*. Interestingly, this ribosome-binding site is part of an inverted repeat whose axis of symmetry lies 26 bp upstream of the initiation codon of ORF3. In contrast, no obvious Shine-Dalgarno sequence could be identified upstream of ORF4, nor does this region encompass any of the known required features for translation.

Transcriptional analysis of the *sar* locus. Northern blot analysis of the *sar* transcripts with several DNA probes representing various segments of *sarA* (probes 2 to 5 in Fig. 2) as well as the entire *sar* locus (probe 1) revealed three transcriptional units sizing at 0.58, 0.8, and 1.15 kb, respectively. In a previous study (19), we hypothesized the 0.8-kb transcript to be a degradation product on the basis of its weak hybridization signal in a Northern blot analysis using log-phase cultures. However, subsequent probe walking using only selected upstream probes (probes 6 and 7) resulted in abolition of first the 0.58-kb transcript and then the 0.8-kb transcript, suggesting three different sites of transcriptional initiation. Probes complementary to sequences either further upstream of probe 7 or downstream of probe 2 were not able to detect any of these transcripts (not shown).

To verify this model of transcriptional initiation, we mapped the 5′ ends of all three *sar* transcripts by primer extension, walking in 200- to 300-bp increments along both strands of the entire 1.3-kb region. Only three primer extension products derived from the *sarA* coding strand could be detected (Fig. 1 and 3), with the transcription start sites positioned at 720, 457, and 155 (designated P1 at −146, P3 at −409, and P2 at −711 relative to the *sarA* start codon in Fig. 4A). This finding also excludes the possibility of overlapping transcripts derived from the DNA strand complementary to the one listed in Fig. 1A. To confirm the primer extension data, a second mapping approach with a modified S1 nuclease protection assay involving single-stranded riboprobes was performed. The S1 mapping data were in complete concordance with those of the primer extension (data not shown). The transcription initiation sites of the three *sar* transcripts together with the putative −10 and −35 promoter boxes as well as other regulatory regions are shown in Fig. 1A; analysis presented in Discussion.

Mapping of the termination signal. To demonstrate that all three *sar* transcripts possess the very same termination signal recognized by the RNA polymerase, an S1 nuclease mapping of the 3′ ends of the *sar* transcripts was performed. We initially carried out a computer analysis of the downstream region of *sarA* to facilitate the construction of an appropriate mapping probe. The program TERMINATOR (3, 4), which searched for prokaryotic RNA polymerase terminators independent of *rho* factors, predicted a constitutive termination site at position

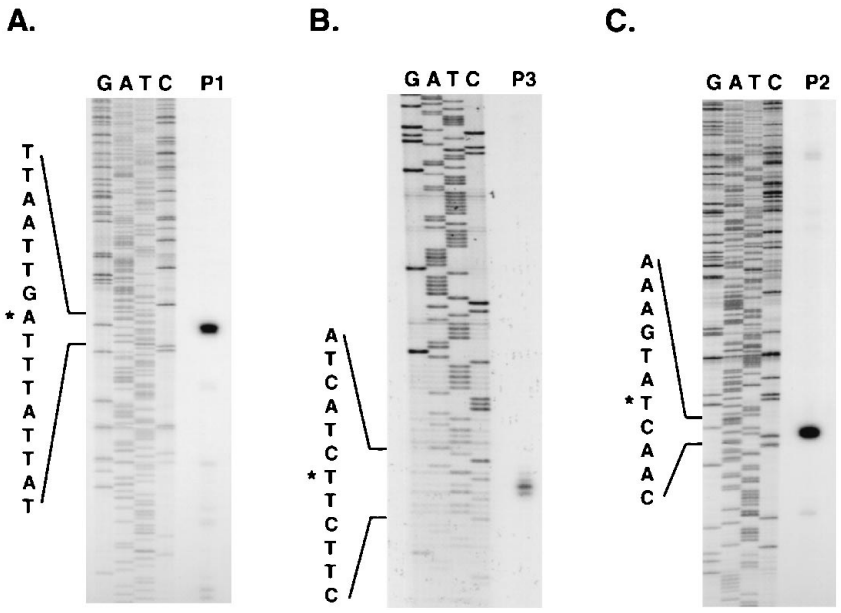

FIG. 3. Mapping of the 5′ ends of three *sar* transcripts by primer extension analysis. Total RNA was hybridized with several oligonucleotides complementary to the mRNA in different parts of the *sar* locus and extended by reverse transcriptase. The precise base mapping was done by comparing the migration of the extended product with a parallel sequencing reaction primed by an identical oligonucleotide. The sequence encompassing the initiation start (marked by asterisks) is enlarged on the left. The transcription start sites of the 0.58-kb *sarA* (A), the 0.8-kb *sarC* (B), and the 1.2-kb *sarB* (C) mRNAs are labeled P1, P3, and P2, respectively.

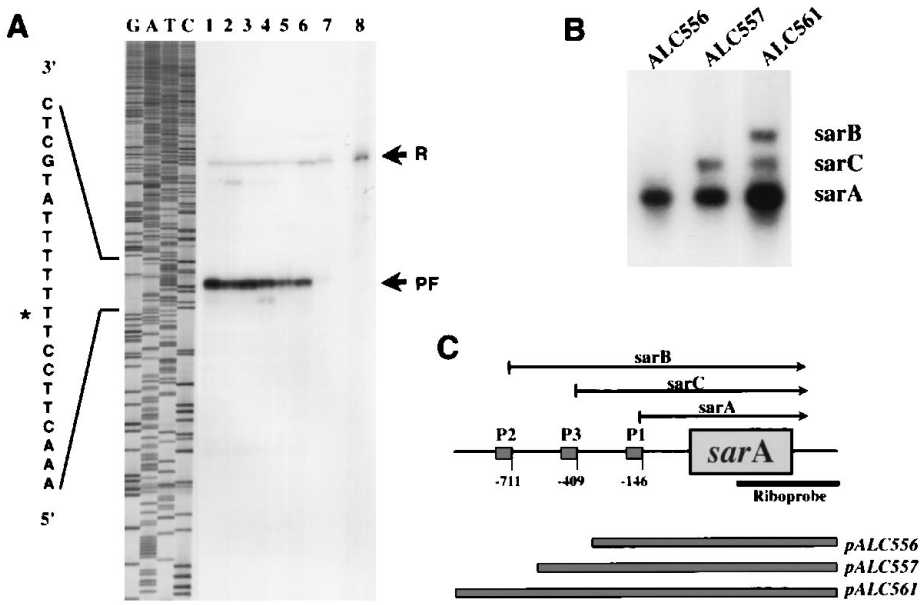

FIG. 4. 3′ mapping of the *sar* transcripts. (A) S1 nuclease protection assay. A concomitant sequencing reaction with a primer identical to the 5′ end of the riboprobe was run in parallel lanes. The termination region is outlined on the left, with the exact 3′ end indicated by an asterisk. Protected fragments (PF) of the wild type RN6390 from early-, mid-, and late-log-phase cells (lane 1, 5, and 6, respectively), the 5′ deletion clones ALC556 (lane 2), ALC557 (lane 3), and ALC561 (lane 4), a reaction of riboprobe (R) and S1 nuclease without RNA (lane 7), and the riboprobe alone (lane 8) were applied onto a sequencing gel. (B) Transcriptional analysis of 5′ deletion clones. RNA from cells grown to late log phase was prepared as described in Materials and Methods. Northern analysis of clones encompassing the promoters P1 (ALC556), P1 and P3 (ALC557), or P1, P3, and P2 (ALC561), as predicted by the primer extension data, was performed with a DNA probe encompassing the entire *sarA* coding region. (C) Physical map of the *sar* region. The riboprobe used in the S1 nuclease mapping is indicated by a solid line. Various 5′ deletion *sar* clones are represented by dark bars. Promoter positions are labeled P1, P2, and P3. The arrows depict the *sar* transcripts. The distances between the 5′ ends of the transcripts and the start codon of *sarA* are expressed by negative position numbers.

1301 ($P = 3.88$) by applying a normalized dinucleotide distribution matrix ($P > 3.5$) alone. Using a combined search based on an additional complementarity weight matrix (degree of dyad symmetry), no higher score could be obtained. This is not surprising because an inspection of the sequence revealed a strong 26-bp potential hairpin loop formed at positions 1262 to 1295 (Fig. 1A), with the center of the hairpin lying between positions −23 and −24 (with the predicted termination site as position −1) and a loop size of 8 bp, thus being 1 nucleotides too long to meet the criteria of Brendel et al. (3, 4), even though all other parameters fit quite well.

The S1 nuclease assays were performed with single-stranded RNA probes generated by in vitro transcription. These probes are derived from clones carrying the appropriate DNA fragment downstream of an SP6 promoter. For this experiment, we used the wild-type strain showing the regular transcription pattern and *sar* transposon mutants carrying shuttle vectors with various 5′-deleted regions of the *sar* locus (Fig. 4C) to initiate transcription at either P1 alone, P1 and P3, or P1, P3, and P2 (Fig. 4B). As indicated, all *sar* transcripts terminated at nucleotide 1298 (Fig. 4A), thus including the complete C terminus of *sarA* in strain RN6390. This is close to the putative termination site as predicted by computer analysis of *E. coli* sequences. Furthermore, Northern analysis of the derivative clone ALC529 lacking this signal demonstrated readthrough transcription into the vector (data not shown).

Phenotypic characterization. Given our primer extension data, we wanted to determine if the deficiency in *sar*-related functions in a *sar* mutant could be complemented by a plasmid carrying the 1,349-bp fragment that encompassed all three *sar* transcripts. This fragment was successfully introduced into *sar* mutant ALC136 via a shuttle vector to yield strain ALC561. Phenotypic analysis of this strain (ALC561) demonstrated that the cloned *sar* fragment was able to restore parental functions to the *sar* mutant (Table 2). For instance, levels of α- and δ-hemolysin production, which were reduced in the *sar* mutant ALC136, were restored close to parental levels in this clone, as determined by the zone of hemolysis on blood agar plates. Quantitative densitometry on Northern blots probed with α- and δ-hemolysin probes also appeared to verify this result (Table 2). Immunoblot analysis revealed that the level of protein A production was upregulated in the *sar* mutant ALC136 compared with the parental strain RN6390. As anticipated, protein A production returned to parental levels in the *sar* mutant clone containing the plasmid with the 1,349-bp *sar* fragment. The binding of whole cells to ^{125}I-labeled fibronectin was slightly less in *sar* mutant ALC136 than in parental strain

TABLE 2. Phenotypic characterization of an *sar* mutant clone encompassing the *sar* locus

Strain	Production		
	α-Hemolysin[a] (U)	δ-Hemolysin[a] (U)	Fibronectin-binding protein (cpm of ^{125}I-fibronectin bound to 10^9 CFU ± SEM)
RN6390	+ (7,473)	± (5,630)	3,320 ± 503
ALC136	−	−	2,393 ± 124
ALC561	+ (4,881)	± (7,963)	19,288 ± 2,295

[a] Assayed by measuring the zones of hemolysis on cross-streaked rabbit erythrocyte agar plates. ++, strong producer; +, moderate producer; ±, very weak producer; −, nonproducer. Levels of production were also assayed on Northern blots (11); band intensities were quantitated densitometrically as integrated area units.

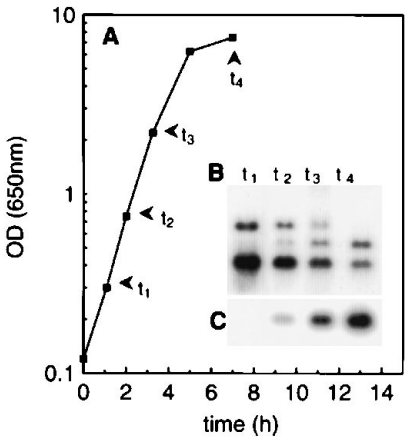

FIG. 5. Growth cycle-dependent regulation of the *sar* transcripts. (A) Growth of *S. aureus* RN6390. Time points at which RNA has been prepared for Northern analysis are indicated by arrows. Also shown is Northern analysis of the *sar* transcripts (B) and the RNAIII transcript of the *agr* locus (C) with PCR-generated probes derived from positions 620 to 1349 (Fig. 1) and 555 to 1773 (19), respectively. OD, optical density.

RN6390. However, introduction of the *sar* mutant with the *sar* fragment encompassing all three *sar*-related transcripts (ALC561) produced a dramatic increase in fibronectin binding (Table 2). In contrast, the transposon mutant complemented with *sarA* alone (ALC103) did not exhibit a similar increase in fibronectin-binding capacity (data not shown).

Growth cycle dependency of *sar*. *S. aureus* RNA harvested at various time points during the growth cycle at early log, mid-log, late log, and stationary phases revealed a temporal dependency of all three *sar* transcripts (Fig. 5). Whereas the concentrations of *sarB* and the more abundant *sarA* transcripts tapered during the late growth cycle, the *sarC* transcript began to appear during the log phase and increased in the late stationary phase (Fig. 5B). In contrast, the expression of RNAIII displayed the expected increase in transcription toward the stationary phase as shown in Fig. 5C, clearly demonstrating a delay in expression compared with *sar*.

DISCUSSION

In an earlier study, we found that the global regulator *sar* encodes a *trans*-acting regulatory molecule (*sarA*) which positively mediates the expression of RNAII from the P2 promoter of the *agr* operon (15), thereby indirectly modulating the *agr* regulatory molecule RNAIII. However, this effect on *agr* transcription has been found to depend not only on the *sarA* gene alone but also extensive additional upstream sequences (15). More specifically, restoration of the parental phenotype could be achieved only by the inclusion of *sarA* and an additional 1.2-kb upstream region. To analyze this genetic requirement in detail, we sequenced *sarA* of strain RN6390 and its upstream region and tested for additional divergent or convergent transcriptional units by performing a careful mapping study.

The data presented in this paper demonstrated that there are three distinct transcripts with the *sar* locus that require the participation of extensive *cis*-acting sequences located upstream of the 339-bp *sarA* gene. This finding was supported by data derived from different mapping approaches as well as Northern analysis showing that three RNA transcripts are initiated from three distinct promoters in a tandem array. The 0.58-kb transcript is initiated most proximal to *sarA* at position −146 (P1 at nucleotide 720 in Fig. 1A), with the putative −10 (TATAAT) and −35 (TTTACT) core promoter boxes spaced by 17 bp. The putative sequences correspond with the canonical hexamers known for σ^{70}-dependent promoter boxes derived from *E. coli*. Consequently, recognition by *E. coli* σ^{70} of some of these *S. aureus* promoters has been observed. However, not all σ^{70}-dependent promoters from *E. coli* are recognized by the RNA polymerase of *S. aureus* (27). The same promoter elements are also recognized by σ^{A}-dependent RNA polymerase in *Bacillus subtilis* and probably resemble a conserved signal for eubacterial promoters (16). A dyad symmetry formed by two 10-bp inverted repeats (IR1) is located downstream of P1. An additional inverted repeat (IR2) of 5 bp is identified upstream of the putative −35 sequence. The nucleotide sequence at the core of IR2 (GTT-N_5-AAC) closely resembles the consensus sequence (GTG-N_5-10-CAC) of regulatory DNA-binding proteins involved in DNA bending (20). Sequences further upstream comprise a 12-bp inverted repeat (IR3) with its dyad symmetry center lying 9 bp upstream of a putative, but suboptimally spaced, Shine-Dalgarno binding site in front of a potential translational start codon for ORF3. Interestingly, a 5′ deletion clone (ALC103) of the *sar* locus, lacking this particular inverted repeat (IR3), resulted in a significant transcriptional upregulation of the *sarA* transcript, whereas a clone (ALC556) with an 80-bp extended upstream sequence (but still lacking the P3 promoter) exhibited a parental level of *sarA* transcription (data not shown). This effect might well be due to the modulation of DNA curvature of the promoter region, since curved DNA is recognized to be the major component in the control of bacterial gene expression (26). The function of the putative peptide coding region of ORF3 remains to be defined and is currently under investigation. Nonetheless, we speculate that the 260-bp region between P1 and P3 comprising this inverted repeat participates in *cis*-acting negative regulation of *sarA* transcription.

The central (P3) and distal (P2) promoters correspond less well with the σ^{70}-dependent −10 and −35 consensus sequences. The extremely high A+T content (75%) of this upstream regulatory region limits the predictive value for the identification of putative promoter boxes. However, the P3 promoter possesses a striking homology (−37 AGTGAT-N_{14}-GGGTAT) to σ^{B}-dependent promoters, which share the consensus −37 RGGXTT-N_{14}-GGGTAT; the activation of these promoters is involved in the general stress response of gram-positive organisms such as *B. subtilis* (2). In *E. coli*, the recognition of stationary-phase-inducible promoters is mediated by a set of σ factors (σ^{32}, σ^{38}, and σ^{54}), the expression of which seems to be regulated by various triggers, usually in response to the environmentally enforced transition of logarithmic growth (23). In both gram-negative and gram-positive bacteria, the corresponding consensus sequences for promoters activated in response to heat shock, stress, starvation, or postexponential growth are generally G/C rich and distinct from housekeeping promoters (14, 21, 29). Given the G/C-rich nature of the putative P3 promoter boxes, the possibility that the *sar* P3 promoter is activated in response to growth phase transition is supported by the data derived from the transcriptional analysis of *sar* during the growth cycle (Fig. 5). More specifically, *sarC* remained almost undetectable in the log phase and increased toward the late stationary phase. Upstream of the P3 promoter is an 8-bp sequence centered 46 bp upstream of the transcription initiation site, and it is perfectly repeated in an inverse orientation (IR4) at the interspace of the putative core promoter elements. Upstream from this element is a 100-bp stretch with high secondary structure potential including two

10-bp direct repeats and four alternating inverted repeats. The region upstream of the putative −35 promoter box encompasses an A/T-rich DNA sequence (UP box) between −43 and −60 of the P3 transcription start, previously described as a third recognition element in bacterial promoters. This UP-box element may be recognized by the alpha subunit of RNA polymerase (18, 28). An UP element is also featured upstream of the P2 promoter. However, the P2 core promoter boxes did not exhibit unambiguous consensus sequence. Further inspection of this sequence stretch revealed no known pattern involved in transcriptional activation.

As an additional confirmation of the transcription start sites derived from primer extension and S1 analysis, we verified by Northern analysis that a *sar* mutant clone containing a shuttle plasmid with the 1,349-bp fragment (ALC561) was able to reconstitute all three *sar* transcripts to the mutant (Fig. 4B). Likewise, a shuttle plasmid (pALC557) carrying a *sar* fragment that encodes P3 and P1 promoters was capable of restoring *sarC* and *sarA* transcription. Phenotypic analysis also demonstrated that the 1,349-bp DNA fragment encoding all three *sar* transcripts was able to complement the phenotypic deficiency in the *sar* mutant to near parental levels. Thus, it is highly likely that the function of the *sar* locus is contained within this fragment encoding all three *sar* transcripts.

In comparing the SarA proteins from strains DB, RN450, and RN6390, it is evident that this protein is highly conserved among *S. aureus* isolates. Although strains RN450 and RN6390 are supposed to derive from a similar genetic background, two base pair changes have occurred within the *sarA* gene coding region in strain RN6390, thus leading to a conservative substitution at residue 53 as well as a truncation of the 11 C-terminal amino acids as a result of a nonsense mutation (TGA). Nonetheless, the SarA protein in strain RN6390 can restore hemolysin production to a *sar* mutant (15). On the basis of this finding, we reasoned that the 11 C-terminal amino acids of the SarA protein in strains RN450 and DB are probably not required for *sar* function. Nevertheless, it is important to note that the *sarA* gene alone, together with the P1 promoter but excluding the distal P3 and P2 promoters and their intervening sequences, is not able to fully complement the phenotypic deficiency of the *sar* mutant, even when the *sarA* gene is being expressed. As the complete restoration of *sar* functions to the mutant requires the expression of all three *sar* transcripts (pALC561), our current hypothesis is that the extensive sequence upstream of the P1 promoter not only may modulate *sarA* expression but also have a function divergent from that of *sarA*.

Both the *sarA* and *sarB* transcripts are expressed throughout the exponential growth phase and are dramatically decreased at the transition to stationary phase. The activation of these two *sar* transcripts thus precedes the expression of RNAIII, which increases in exponential phase and is most abundant in the stationary phase. This observation is in agreement with earlier results showing that the *sar* gene product(s) likely binds to the P2 promoter region of *agr*, thus regulating RNAII and subsequently RNAIII transcription.

Current experiments using different *sar* promoters fused to a reporter gene are directed toward the analysis of environmental triggers of *sar* promoter activation. In addition, we are focusing on the functional role of ORF3 and ORF4 by physically separating *sarA* and these coding regions in divergent transcriptional units on the same plasmid and trying to complement the wild-type phenotype posttranslationally.

ACKNOWLEDGMENTS

This study was made possible by a Norman and Rosita Winston Fellowship Award to M.G.B. and was supported in part by grants-in-aid from the American Heart Association and the New York Heart Association and by NIH grants AI30061 and AI37142. A. L. Cheung is a recipient of the Irma T. Hirshl Career Scientist Award as well as the AHA-Genentech Established Investigator Award from the American Heart Association.

REFERENCES

1. **Blake, M. S., K. H. Johnston, G. J. Russell-Jones, and E. C. Gotschlich.** 1984. A rapid sensitive method for detection of alkaline phosphatase-conjugated anti-antibody on Western blots. Anal. Biochem. **136:**175–179.
2. **Boylan, S. A., M. D. Thomas, and C. W. Price.** 1991. Genetic method to identify regulons controlled by nonessential elements: isolation of a gene dependent on alternate transcription factor σ^B of *Bacillus subtilis*. J. Bacteriol. **173:**7856–7866.
3. **Brendel, V., G. H. Hamm, and E. N. Trifonov.** 1986. Terminators of transcription with RNA polymerase from *E. coli*. J. Biomol. Struct. Dyn. **3:**705–723.
4. **Brendel, V., and E. N. Trifonov.** 1984. A computer algorithm for testing potential prokaryotic terminators. Nucleic Acids Res. **12:**4411–4427.
5. **Chang, S., and S. N. Cohen.** 1979. High frequency transformation of *Bacillus subtilis* protoplasts by plasmid DNA. Mol. Gen. Genet. **168:**111–115.
6. **Cheung, A. L., K. Eberhardt, and V. A. Fischetti.** 1994. A method to isolate RNA from gram-positive bacteria and mycobacteria. Anal. Biochem. **222:** 511–514.
7. **Cheung, A. L., K. J. Eberhardt, E. Chung, M. R. Yeaman, P. M. Sullam, M. Ramos, and A. S. Bayer.** 1994. Diminished virulence of a *sar*$^-$ *agr*$^-$ mutant of *Staphylococcus aureus* in the rabbit model of endocarditis. J. Clin. Invest. **94:**1815–1822.
8. **Cheung, A. L., and V. A. Fischetti.** 1988. Variation in the expression of cell wall proteins of *Staphylococcus aureus* grown on solid and liquid media. Infect. Immun. **56:**1061–1065.
9. **Cheung, A. L., J. M. Koomey, C. A. Butler, S. J. Projan, and V. A. Fischetti.** 1992. Regulation of exoprotein expression in *Staphylococcus aureus* by a locus (*sar*) distinct from *agr*. Proc. Natl. Acad. Sci. USA **89:**6462–6466.
10. **Cheung, A. L., and S. J. Projan.** 1994. Cloning and sequencing of *sarA*: a gene required for the expression of *agr*. J. Bacteriol. **176:**4168–4172.
11. **Cheung, A. L., and P. Ying.** 1994. Regulation of α and β hemolysins by the *sar* locus of *Staphylococcus aureus*. J. Bacteriol. **176:**580–585.
12. **Easmon, C. S. F., and C. Adlam.** 1983. Staphylococci and staphylococcal infections. Academic Press, New York.
13. **Greene, C., D. McDevitt, P. Francois, P. E. Vaudaux, D. P. Lew, and T. J. Foster.** 1995. Adhesion properties of mutants of *Staphylococcus aureus* defective in fibronectin-binding proteins and studies on the expression of *fnb* genes. Mol. Microbiol. **17:**1143–1152.
14. **Haldenwang, W. G.** 1995. The sigma factors of *Bacillus subtilis*. Microbiol. Rev. **59:**506–531.
15. **Heinrichs, J. H., M. G. Bayer, and A. L. Cheung.** 1996. Characterization of the *sar* locus and its interaction with *agr* in *Staphylococcus aureus*. J. Bacteriol. **178:**418–423.
16. **Helmann, J. D.** 1995. Compilation and analysis of *Bacillus subtilis* sigmaA-dependent promoter sequences: evidence for extended contact between RNA polymerase and upstream promoter DNA. Nucleic Acids Res. **23:** 2351–2360.
17. **Janzon, L., and S. Arvidson.** 1990. The role of the δ-hemolysin gene (*hld*) in the regulation of virulence genes by the accessory gene regulator (*agr*) in *Staphylococcus aureus*. EMBO J. **9:**1391–1399.
18. **Jeon, Y. H., T. Negishi, M. Shirakawa, T. Yamazaki, N. Fujita, A. Ishihama, and Y. Kyogoku.** 1995. Solution structure of the activator contact domain of the RNA polymerase α subunit. Science **270:**1495–1497.
19. **Kornblum, J., B. Kreiswirth, S. J. Projan, H. Ross, and R. P. Novick.** 1990. *Agr*: a polycistronic locus regulating exoprotein synthesis in *Staphylococcus aureus*, p. 373–402. *In* R. P. Novick (ed.), Molecular biology of the staphylococci. VCH Publishers, New York.
20. **Liu-Johnson, H. N., M. R. Gartenberg, and D. M. Crothers.** 1986. The DNA binding domains and bending angle of the *E. coli* CAP promoter. Cell **47:**995–1005.
21. **Mager, W. H., and A. J. J. De Kruiff.** 1995. Stress-induced transcriptional activation. Microbiol. Rev. **59:**506–531.
22. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
23. **Miksch, G., and P. Dobrowolski.** 1995. Growth phase-dependent induction of stationary-phase promoters of *Escherichia coli* in different gram-negative bacteria. J. Bacteriol. **177:**5374–5378.
24. **Novick, R. P.** 1990. The staphylococcus as a molecular genetic system, p.

1–40. *In* R. P. Novick (ed.), Molecular biology of the staphylococci. VCH Publishers, New York.
25. **Novick, R. P., G. K. Adler, S. J. Projan, S. Carleton, S. Highlander, A. Gruss, S. A. Khan, and S. Iordanescu.** 1984. Control of pT181 replication. I. The pT181 copy control functions acts by inhibiting the synthesis of a replication protein. EMBO J. **3:**2399–2405.
26. **Perez-Martin, J., F. Rojo, and V. de Lorenzo.** 1994. Promoters responsive to DNA bending: a common theme in prokaryotic gene expression. Microbiol. Rev. **58:**268–290.
27. **Rao, L., R. K. Karls, and M. J. Betley.** 1995. In vitro transcription of pathogenesis-related genes by purified RNA polymerase from *Staphylococcus aureus*. J. Bacteriol. **177:**2609–2614.
28. **Ross, W., K. K. Gosink, J. Salomon, K. Igarashi, C. Zou, A. Ishihama, K. Severinov, and R. L. Gourse.** 1993. A third recognition element in bacterial promoters: DNA binding by the α subunit of RNA polymerase. Science **262:**1407–1413.
29. **Vicente, M., S. R. Kushner, T. Garrido, and M. Aldea.** 1991. The role of the gearbox in the transcription of essential genes. Mol. Microbiol. **5:**2085–2091.
30. **Waldvogel, F. A.** 1985. *Staphylococcus aureus*, p. 1097–1116. *In* G. L. Mandell, R. G. Douglas, Jr., and J. E. Bennett (ed.), Principles and practice of infectious diseases. John Wiley & Sons, New York.