Since this is a new application, this section is Preliminary Studies rather than a Progress Report.

PRELIMINARY STUDIES

Brings in collaborators to round out expertise. Addresses review criteria: PI. Shows PI and collaborators can do it since they already did the basic preliminary work. Gives lots of detail. Stays within page limit (in original version, prior to annotation).

This proposal is a collaboration between the PI and Dr. Barry Hurlburt in the Department of Biochemistry and Molecular Biology. The collaboration takes advantage of the expertise of the PI in the molecular genetics of *S. aureus* and the biochemical expertise of Dr. Hurlburt in transcription factor structure and function (14,15). The overall goals are 1) correlation of the expression of the sarA, sarB and sarC transcripts with the production and activity of SarA, 2) characterization of the mechanism by which *sar* regulates expression of the *S. aureus* collagen adhesin gene (*cna*) and 3) identification and characterization of additional *S. aureus* genes under the direct regulatory control of SarA. We have assembled all of the experimental tools required to accomplish these objectives. Specifically, we have (i) purified SarA in a form capable of binding an appropriate DNA target, (ii) generated an affinity-purified antibody against purified SarA, (iii) constructed a *xylE* reporter plasmid that can be used to assess the functional activity of SarA (Specific Aim #1) and define the sequence characteristics required for the regulation of *cna* transcription (Specific Aim #2), (iv) cloned the regions encoding the sarA, sarB and sarC transcripts for use in complementation experiments, (v) demonstrated that SarA binds a DNA target upstream of *cna* and begun the process of localizing the SarA binding site and (vi) obtained or generated *sar* and *agr* mutants in both *cna*-positive and *cna*-negative *S. aureus* strains. The experiments done to accomplish each of these tasks are described in detail below.

Cloning and expression of sarA. The polymerase chain reaction (PCR) was used to amplify the *sarA* coding region from *S. aureus* strain RN6390. Utilizing *Nde*I and *Bam*HI restriction sites incorporated into the oligonucleotide primers, the fragment containing the *sarA* coding region was cloned into the *E. coli* expression vector pET9A. Because the *Nde*I site (CATATG) in the vector overlaps an ATG start codon, cloning of the *sarA* coding region into the *Nde*I site places the *sarA* structural gene in perfect register with the vector-derived ribosome binding site. Recombinant proteins are therefore expressed as full-length, wild type proteins without fusions to exogenous peptide or protein tags. After cloning the *sarA* PCR fragment into pET9A and confirming the identity of the cloned fragment by DNA sequencing (data not shown), the recombinant plasmid (pETSarA) was used to transform *E. coli* strain BL21(DE3)pLysS. Transformants were grown to mid-log phase before

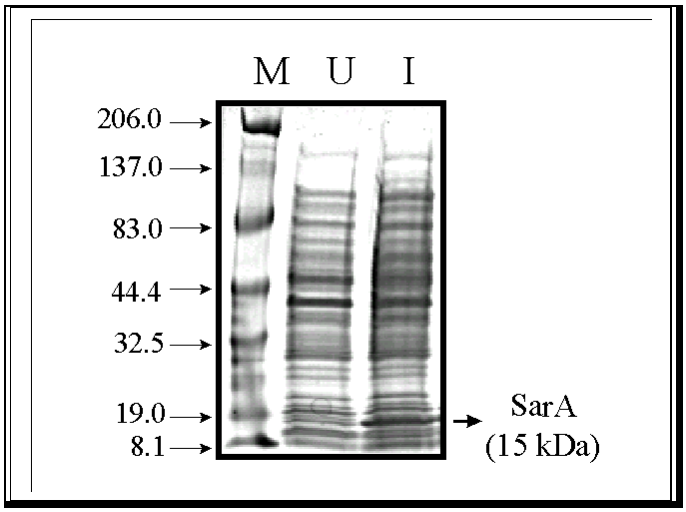

Fig. 7. Expression of SarA in *E. coli*. Whole cell lysates were resolved using 10-20% SDS-PAGE gradient gels. Lane designations: M, molecular weight markers (sizes indicated in kDa); U, uninduced; I, induced.

Figure illustrates results of previous work.

inducing SarA expression by adding IPTG to a final concentration of 0.4 mM. After two hours, cells were harvested and lysed by sonication. The presence of SarA in the crude lysate was confirmed by SDS-PAGE followed by Coomassie Brilliant Blue staining (Fig. 7).

Purification of SarA. A 500 ml culture of the BL21(DE3)pLysS *E. coli* strain containing pETSarA was induced and lysed as described above. After removing the insoluble material in the crude lysate by centrifugation, the soluble fraction was subjected to a series of ammonium sulfate precipitations culminating at 70% saturation. The pellet from each precipitation was resuspended in SDS-PAGE buffer and examined along with an aliquot of the supernatant (Fig. 8, left). The supernatant remaining after the final precipitation was found to contain ~70% SarA.

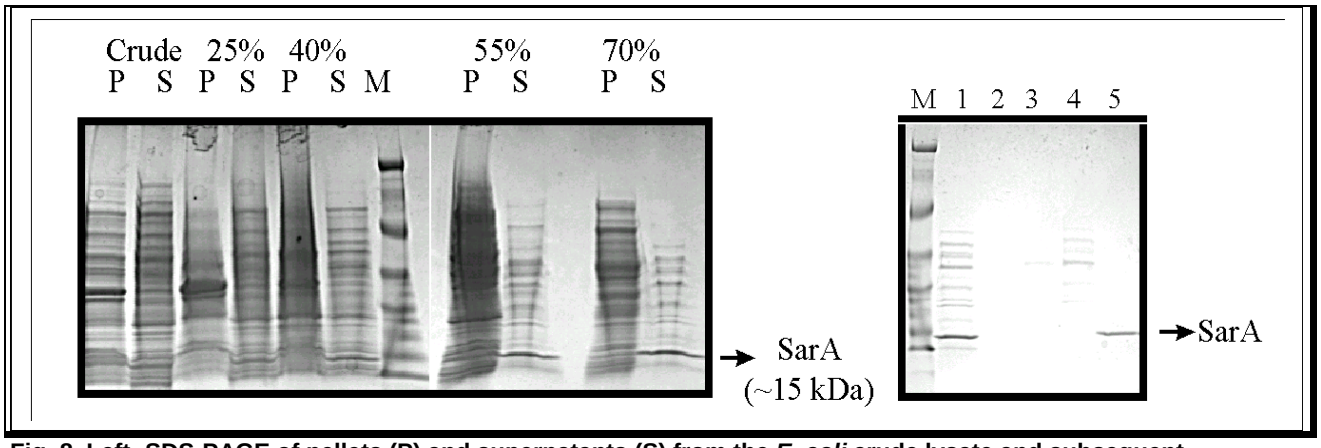

Fig. 8. Left: SDS-PAGE of pellets (P) and supernatants (S) from the *E. coli* crude lysate and subsequent ammonium sulfate precipitations. The lane marked "M" contains the same molecular weight markers shown in Fig.7. Right: Purification of SarA. Lane 1 contains an aliquot of the supernatant from the 70% ammonium sulfate precipitation. Lanes 2 and 3 contain fractions of the "flow-through" obtained after applying the 70% supernatant to the heparin sepharose column. Lanes 4 and 5 contain elution fractions obtained by increasing the salt concentration of the elution buffer. Fraction 5 contained SarA in the absence of detectable levels of other proteins.

SarA was subsequently purified to apparent homogeneity by ion exchange chromatography using Heparin-Sepharose (Fig. 8, right).

Details of preliminary data provide great credibility for investigator.

Confirmation that the purified protein is SarA. To ensure that the protein purified from our *E. coli* lysates is SarA, we performed electrophoretic mobility shift assays (EMSA) with a 45 bp DNA fragment containing the heptad repeats *cis* to the *agr* P_3 promoter (see Fig. 5, Background and Significance). The 45 bp fragment was generated by

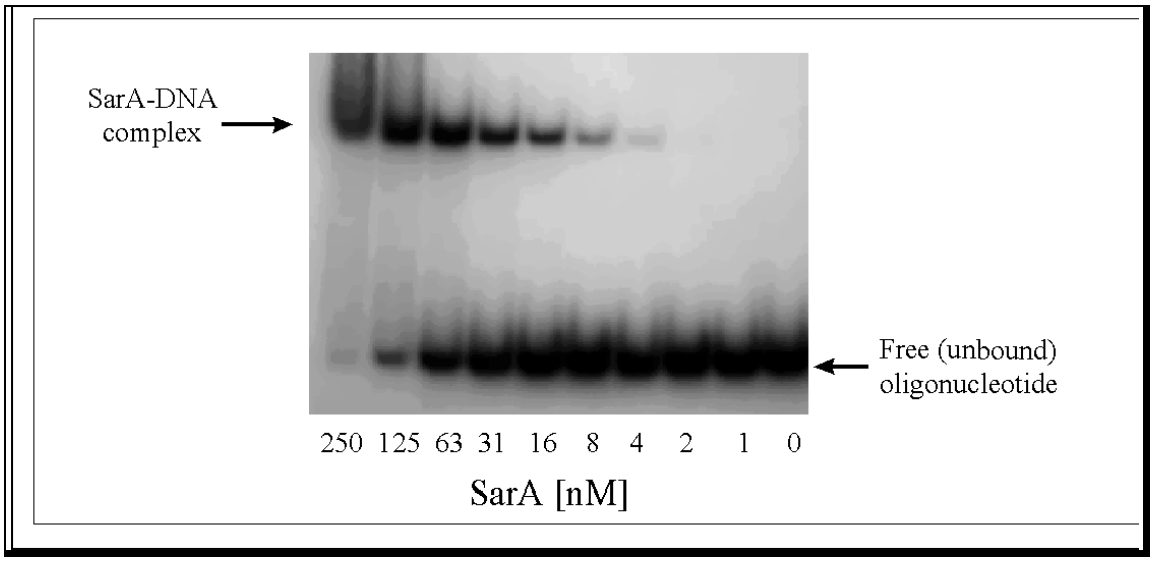

Fig. 9. EMSA experiment done with a 45 bp DNA fragment containing the heptad repeats upstream of the *agr* P3 promoter and purified SarA. Lane designations indicate the concentration of SarA in each binding reaction. All lanes contain an equal amount (~10 pmole) of ^{32}P-labeled DNA.

synthesizing and annealing complementary oligonucleotides. After labeling the target DNA with ^{32}P, the fragment was allowed to equilibrate in solution with varying amounts of the purified protein. The mixture was then resolved by native gel electrophoresis. The fact that a mobility shift was observed with the *agr*-derived target DNA (Fig. 9) confirms that the protein we purified from the *E. coli* lysates is SarA. **Moreover, these results, together with the results of our EMSA experiments employing DNA fragments derived from the region upstream of *cna* (see below), demonstrate that our *E. coli*-derived SarA preparation is appropriate for the experiments aimed at the identification of additional SarA targets within the *S. aureus* genome (Specific Aim #3).**

Shows PI can interpret results critically.

Generation of SarA-specific antiserum. To generate affinity-purified anti-SarA antibodies, two rabbits were immunized by sequential subcutaneous injections of 1) 75 µg SarA suspended in Freund's Complete Adjuvant (day -35), 2) 75 µg SarA suspended in Freund's Incomplete Adjuvant (day -20) and 3) 150 µg SarA without adjuvant (day -5). To test for seroconversion, rabbits were bled from the ear vein on day -5 and the sera tested by enzyme-linked immunosorbent assay (ELISA) using microtiter plates coated with purified SarA and goat anti-rabbit IgG conjugated to horseradish peroxidase. One of two rabbits was found to have antibody reactive against SarA (data not shown). The seropositive rabbit was subsequently bled by cardiac puncture on day 0. SarA antibody present in the antiserum was purified by affinity chromotography using CNBr-activated sepharose coupled to purified SarA. The specificity of the antibody will be determined by blotting cell lysates prepared from wild-type strains with lysates prepared from *sar* mutants (Table 1) and by EMSA supershift assays using purified SarA and an appropriate DNA target. The affinity-purified antibody preparation will be used in Western blot assays (Specific Aim #1) and for the immunoprecipitation of SarA-DNA complexes (Specific Aim #3).

Correlates preliminary findings to specific aims.

Construction of xylE reporter fusion vector. A plasmid containing a promoter-less *xylE* reporter gene was obtained from Dr. Ken Bayles at the University of Idaho. The plasmid was constructed by amplifying the *xylE* structural gene using oligonucleotide primers containing engineered *Hind*III and *Bam*HI restriction sites in the 5' and 3' primers respectively. These restriction sites were used to clone the amplification product into pUC19. We used the same restriction sites to subclone the promoter-less *xylE* gene from pUC19 into the *E. coli-S. aureus* shuttle vector pLI50 (data not shown). By introducing DNA fragments containing the *cis* elements associated with the *agr* P_2 and P_3 promoters upstream of the promoter-less *xylE* gene, this construct can be used to assess the activity of SarA as a transcriptional activator (Specific Aim #1). Additionally, by introducing DNA fragments corresponding to the DNA upstream of *cna* (P_{cna}), the *xylE* reporter fusion can be used to define the sequence requirements necessary for the SarA-mediated regulation of *cna* transcription (Specific Aim #2).

Studies with the S. aureus collagen adhesin gene (cna). Recent studies in our laboratory have focused on the role of collagen binding in the pathogenesis of staphylococcal osteomyelitis. These studies have led to the following conclusions. **First**, we have established that the collagen-binding capacity (CBC) of *S. aureus* is a direct function of the presence and expression of *cna* (21) and that *cna* is encoded within a discrete chromosomal

element that is not present in most *S. aureus* isolates (19,39). The *cna* element does not encode any additional genes and its presence does not disrupt a gene present in *cna*-negative strains (19). These results are relevant to this proposal because they established that 1) a fragment containing ~500 bp of the DNA upstream of *cna* is sufficient for the regulated expression of *cna* from a pLI50 construct (20) and 2) the regulatory loci that control *cna* transcription are present in both *cna*-positive and *cna*-negative strains (20). **Second**, as discussed in the Background and Significance section of this proposal, we have established that the expression of *cna* is regulated by *sar* in an *agr*-independent manner (20, 21). Specifically, when we examined the collagen binding capacity (CBC) of *cna*-positive strains that carry mutations in the *sar* and/or *agr* loci, we found that mutation of *agr* results in only a slight increase in CBC that is not apparent in all strains (20) while mutation of *sar* results in a dramatic increase in the ability to bind collagen (Fig. 10) and a corresponding increase in *cna* transcription (Fig. 11). Moreover, mutation of both *sar* and *agr* had an additive effect on CBC and *cna* transcription. **Because these results would not be expected if the regulatory effects of *sar* were a function of the SarA-dependent activation of *agr* transcription, we conclude that *sar* is the primary regulatory element controlling expression of *cna* and that the regulatory effects of *sar* on *cna* transcription are independent of the regulatory impact of *sar* on *agr*.**

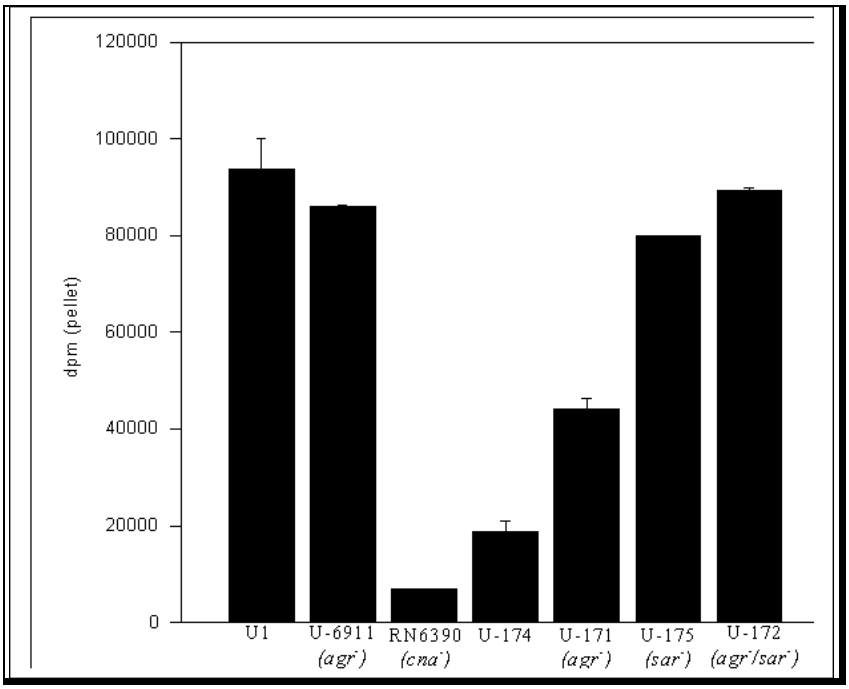

Fig. 10. Collagen binding capacity (CBC) of RN6390 *sar* and *agr* mutants. The prefix "U" designates a UAMS strain (Table 1). UAMS-1 is a clinical isolate that encodes and expresses *cna*. U-6911 is an *agr*-null mutant of U-1. Collagen binding assays were done using ^{125}I-labeled collagen as previously described (20).

Shows ability to interpret data critically.

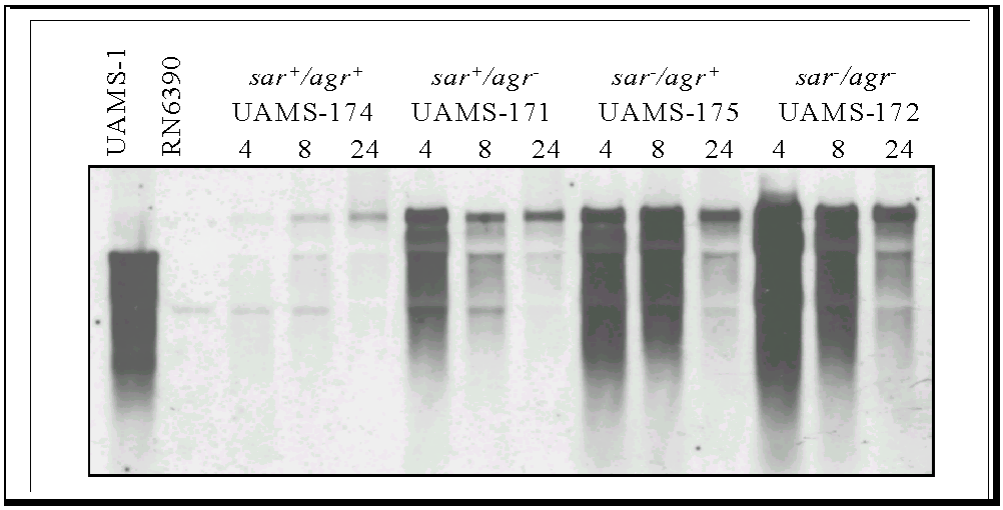

Fig. 11. Northern blot analysis of *cna*-positive *sar* and *agr* mutants. RNA samples taken at various time points (top) were blotted using a *cna* probe. Growth curves indicate that the transition between exponential and post-exponential growth occurred between the 4 and 8 hr time points (data not shown). UAMS-1 encodes a *cna* gene with a single B domain and was included as a positive control. RN6390 does not encode *cna* and was included as a negative control. The *cna* gene introduced into RN6390 and the RN6390 *sar* and *agr* mutants is derived from FDA574 and includes three B domains; as a result, the *cna* mRNA in these strains is larger than the *cna* mRNA in UAMS-1.

Studies on the mechanism of sar-mediated regulation of cna transcription. To determine whether the regulation of *cna* transcription is a direct function of SarA or is due to an unidentified regulatory factor under the regulatory control of SarA, we generated a series of short (~140 bp) overlapping fragments that collectively span the region extending 360 bp region upstream of the *cna* coding region.

EMSA experiments done with purified SarA and each of these fragments suggests that at least one SarA-binding site exists within ~200 bp of the *cna* start codon (Fig. 12). Although sequence analysis of the DNA region upstream of *cna* reveals three sites that match the heptanucleotide *agr* target at 6 of 7 bp (19), it is important to note that, by comparison to the repeats upstream of the *agr* promoters, none of these sites are appropriately placed with respect to the putative *cna* promoter or with respect to each other (19, 30). These results suggest that the SarA-binding sites upstream of the *agr* promoters may be relatively unique by comparison to other SarA targets. Whether the distinction between the SarA-binding sites upstream of *agr* and the SarA-binding sites upstream of *cna* is related to the fact that SarA-binding results in the activation of *agr* transcription and the repression of *cna* transcription is one of the issues that will be addressed during the course of these studies. **Most importantly, we believe our preliminary results regarding the *sar*-mediated regulation of *cna* transcription provide us with a unique system by which we can investigate the *agr*-independent branch of the *sar* regulatory pathway.**

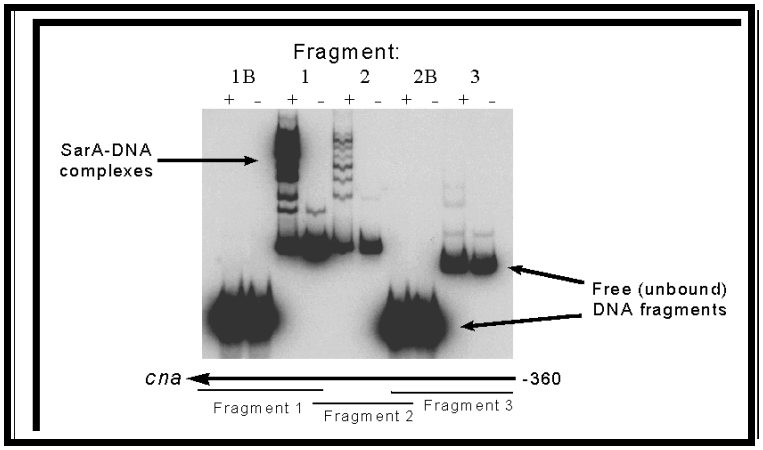

Fig. 12. EMSA done with DNA fragments derived from the region immediately upstream of *cna* (19) and purified SarA. The presence (+) or absence (-) of SarA is indicated above each lane. All binding assays were done with a constant amount of ^{32}P-labeled DNA (10 pmole) and 55 nM SarA. Fragments 1, 2 and 3 collectively span the 360 bp region upstream of the *cna* start codon as indicated below the figure. Fragments 1B and 2B are extraneous PCR products obtained during the amplification of fragments 1 and 2 respectively. To ensure binding specificity, both fragments were purified and included in the EMSA experiments.

Describes why investigators believe they have unique capability to conduct the proposed research.

Cloning of the DNA regions encoding the sarA, sarB and sarC transcripts. We used PCR to amplify individual DNA fragments corresponding to the regions encoding the sarB, sarC and sarA transcripts. All three fragments were amplified using an antisense primer that corresponds to the region downstream of a *rho*-independent terminator ownstream of *sarA* (2). The 5' primers used for each amplification were designed to result in PCR products that include all potentially relevant DNA sites associated with the *sar* P_1, P_2 and

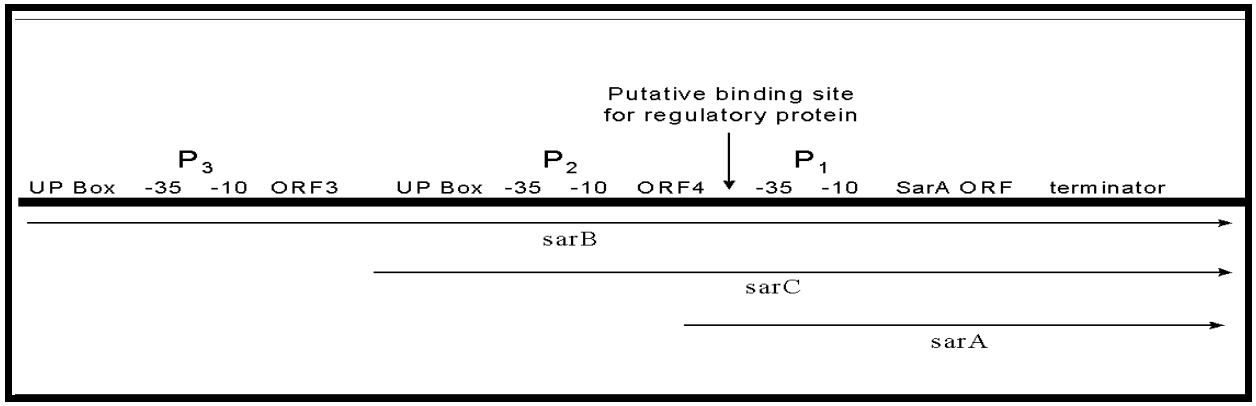

Fig. 13. Molecular architecture of the *sar* locus as defined by Bayer et al. (2). DNA sites with the potential to impact transcription are indicated above the bold line. Thin arrows below the bold line indicate the relative location of the DNA fragments amplified and cloned for use in our complementation studies.

P_3 promoters (Fig. 13).

Specifically, Bayer et al. (2) identified a potential “UP box” (an A/T-rich DNA site that may facilitate binding of the RNA polymerase subunit) upstream of the *sar* P_2 and P_3 promoters. The region upstream of the *sar* P_1 promoter also contains what Bayer et al. (2) identified as a putative binding site for regulatory DNA-binding proteins (Fig. 13). Although no function for these DNA sites has been demonstrated, it nevertheless seemed prudent to include each site within the respective PCR product. Therefore, the 5' end of the largest fragment (1288 bp) is located 20 bp upstream of the UP box adjacent to the *sar* P_2 promoter while the 5' end of the second fragment (978 bp) is located 16 bp upstream of the UP box adjacent to the *sar* P_3 promoter. These fragments span the regions encoding sarB and sarC respectively (2). The 5' end of the smallest fragment (708 bp) is located 226 bp upstream of the *sarA* start codon and includes the region encoding both the putative binding site for regulatory DNA-binding proteins (Fig. 13). The relatively long 5’ extension was chosen because the region extending 189 bp upstream of the sarA transcript appears to be required for efficient transcription of the genes encoded within *agr* (6). Each amplification product was cloned using the TOPO-TA vector (Invitrogen, Carlsbad, CA) and subcloned into the pLI50 shuttle vector. By introducing each of these plasmids into a *cna*-positive *S. aureus sar* mutant (see below), we can confirm the results of our experiments correlating the production of each *sar* transcript with the production of SarA and the temporal pattern of *cna* transcription (Specific Aim #2).

Construction of S. aureus sar mutants. We have generated all combinations of *sar* and *agr* mutants in both *cna*-positive and *cna*-negative *S. aureus* strains. Specifically, we obtained two *sar* mutants (Sar R and 11D2) and their corresponding wild-type parent strains (RN6390 and DB respectively) from Dr. Ambrose Cheung at Rockefeller University. The *cna* gene was introduced into Sar R and RN6390 by 11-mediated transduction from CYL574 (21). This transduction results in the integration of *cna* into the chromosomally-encoded lipase gene (*geh*). The RN6390 (*geh::cna*) derivative has been designated UAMS-174 while the Sar R (*geh::cna*) derivative has been designated UAMS-175. The successful introduction of *cna* was confirmed by Southern blot (data not shown) and by demonstrating that, unlike the RN6390 and Sar R parent strains, UAMS-174 and UAMS-175 bind collagen (Fig. 10). To generate *sar/agr* double mutants, we used 11-mediated transduction to introduce the *agr*-null mutation from RN6911 into Sar R, UAMS-174 and UAMS-175. RN6911 is an RN6390 strain in which the entire *agr* locus including the region encoding RNAIII has been replaced with *tetM*. Importantly, the *geh::cna* insertion in UAMS-174 and UAMS-175 confers tetracycline resistance by virtue of the *tetK* gene. That is relevant because, while both *tetK* and *tetM* confer tetracycline-resistance, only *tetM* confers resistance to minocycline.

Table summarizes data.

Table 1. *S. aureus agr* and *sar* mutants.

Strain	Relevant genotype	Source/Comments
RN6390	*agr*$^+$, *sar*$^+$	*S. aureus* 8325-4. Dr. Ambrose Cheung, Rockefeller University.
Sar R	*agr+, sar-*	RN6390 *sarA* mutant. Dr. Ambrose Cheung.
RN6911	*agr-, sar+*	RN6390 *agr*-null mutant. Dr. Richard Novick, New York University.
UAMS-173	*agr-, sar-*	Transduction from RN6911 into Sar R.
UAMS-174	*agr+, sar+, cna+*	Transduction from CYL574 into RN6390.
UAMS-175	*agr+, sar-, cna+*	Transduction from CYL574 into Sar R.
UAMS-171	*agr-, sar+, cna+*	Transduction from RN6911 into UAMS-174.
UAMS-172	*agr-, sar-, cna+*	Transduction from RN6911 into UAMS-175.
DB	*agr+, sar+*	*S. aureus* wild-type strain. Provided by Dr. Ambrose Cheung.
11D2	*agr+, sar-*	DB *sarA* mutant. Dr. Ambrose Cheung.

It was therefore possible to select for the *agr*-null mutation by plating transductants on medium containing 2 µg/ml minocycline. The same selection was used to identify the *agr*-null derivative of Sar R. The *agr*-null derivatives of UAMS-174, UAMS-175 and Sar R were designated UAMS-171, UAMS-172 and UAMS-173 respectively (Table 1).

The mutations in each of the strains described in Table 1 were confirmed at the genotypic level by Southern blot (data not shown) and at the phenotypic level by collagen binding assays (Fig. 10) and Northern blot (Fig. 11). The availability of these strains will be important in the experiments aimed at 1) correlating the production of each of the *sar* transcripts with the production of functional SarA (Specific Aim #1), 2) complementation analysis aimed at defining the *sar* transcripts required to restore control of *cna* transcription (Specific Aim #2) and 3) confirmation of the *sar*-mediated regulatory control of additional *S. aureus* target genes (Specific Aim #3).