Regulation of virulence gene expression in Staphylococcus aureus

Karin Tegmark Wisell

STOCKHOLM 2000
ABSTRACT

The pathogenic bacterium *Staphylococcus aureus* has the ability to cause a wide variety of human diseases, ranging from superficial abscesses and wound infections to deep and systemic infections such as osteomyelitis, endocarditis and septicaemia. The ability to cause disease has been attributed to a large number of toxins and digesting enzymes as well as to proteins at the bacterial surface that bind various host molecules. These so-called virulence factors are accessory, and are supposed to be synthesised in response to the specific needs during the course of the infectious process. The work described in this thesis aims at a better understanding of the mechanisms that regulate the expression of virulence factors.

Two interacting regulatory systems, *agr* (accessory gene regulator) and *sar* (staphylococcal accessory regulator), are involved in this regulation. The *agr* locus, which encodes a two-component signal transduction system responding to cell density, controls the expression of at least 25 different virulence factors. The effector molecule of the *agr* system is a regulatory RNA molecule, named RNAIII. The *sar* locus has been shown to regulate several staphylococcal virulence genes by modulating the activity of *agr*, but also via *agr*-independent mechanisms. The effector of the *sar* locus is a 14.7 kDa DNA binding protein, SarA. In animal models of infection both *agr* and *sarA* have been shown to affect virulence. How RNAIII and *sarA* function at the molecular level is, however, poorly understood.

The structure and function of the RNAIII promoter have been studied in detail showing that SarA, which regulates the synthesis of RNAIII under certain growth conditions, binds to multiple sites within the RNAIII promoter region. It has also been shown that a region of 93 bp upstream of the transcription start point is sufficient for *agr*-dependent regulation of RNAIII synthesis.

The RNAIII genes of several coagulase negative staphylococcal species – *S. epidermidis*, *S. simulans* and *S. warneri*– have been identified and analysed. The RNAIII molecules from the coagulase negative staphylococci were able to partially complement an RNAIII deficient *S. aureus* mutant. By the construction of hybrid RNAIII molecules it has also been demonstrated that highly conserved primary and secondary structures in both the 5´- and the 3´-half of the RNAIII molecule are required for regulation of virulence genes, and that separate parts of the molecule were involved in regulation of different target genes.

Several genes known to be regulated by RNAIII have been demonstrated to be regulated by the *sarA* locus, independent of its effect on expression of RNAIII. By electrophoresis mobility shift experiments and DNase footprinting, SarA has been found to bind in a very similar way to the promoter regions of genes that are either activated or repressed by *sarA*. SarA does not appear to recognise a conserved DNA sequence motif but rather binds to AT-rich sequences.

New potential regulators of *agr* (RNAIII), *hla* (alpha-hemolysin), *ssp* (serine protease) and *spa* (protein A) have been searched for using specific promoter DNA linked to magnetic beads. Of several new candidate regulators, one protein with a high degree of similarity to SarA, named SarH1 (Sar Homologue 1) has been characterised and found to be part of the *agr-sarA* regulatory network controlling virulence gene expression. By computer searches in the unfinished *S. aureus* genome databases four additional Sar homologues have been found, some of which may also be involved in this regulatory network.
MAIN REFERENCES

This thesis is based on the following papers, which will be referred to in the text by their Roman numerals:

I. Eva Morfeldt, Karin Tegmark and Staffan Arvidson. (1996)

II. Karin Tegmark, Eva Morfeld and Staffan Arvidson. (1998)

III. Karin Tegmark, Anna Karlsson and Staffan Arvidson. (2000)

IV. Karin Tegmark, Eva Morfeld and Staffan Arvidson. (2000)
   The virulence gene regulator, SarA, in *Staphylococcus aureus*, appears to be a non-specific DNA binding protein. Manuscript.

Reprints were made with permission from the publishers.
To my sister Anna

If only science had reached further
# CONTENTS

INTRODUCTION ......................................................................................................................2

Regulation of Bacterial Virulence Factors .................................................................2
Signal transduction ...............................................................................................................2
Gene regulation.......................................................................................................................4
  Transcriptional regulation...............................................................................................5
  Posttranscriptional regulation ..........................................................................................8
Regulatory RNAs .................................................................................................................9

Staphylococcus aureus ..................................................................................................10
Virulence factors...............................................................................................................11
Expression and regulation of virulence factors ..........................................................14
  The agr locus ................................................................................................................14
  The sarA locus ..............................................................................................................18
  1E3 ................................................................................................................................20
  sae ................................................................................................................................20
  rot ..................................................................................................................................20
  Sigma B ........................................................................................................................21
  Environmental factors affecting virulence gene expression ........................................21

THE PRESENT STUDY ..........................................................................................................24

Aim of the study ...................................................................................................................24
Results and discussion .................................................................................................24
  Transcriptonal control of RNAIII ..................................................................................24
  The role of RNAIII in regulation of transcription ........................................................28
  Several virulence genes regulated by RNAIII are also regulated by sarA ....................29
  SarA binds DNA without apparent sequence specificity ..............................................32
  Effects of deletions in the hla and ssp promoter regions ...............................................33
  Analysis of proteins binding to agr-regulated promoters .............................................35
  Characterisation of sarH1 ..............................................................................................36
  Additional Sar homologues ...........................................................................................38
  RNAIII is present in several coagulase-negative species .............................................40
  Identification of functional domains of RNAIII ............................................................41

GENERAL CONCLUSIONS AND A HYPOTHETICAL REGULATORY MODEL ..........43

ACKNOWLEDGEMENT ........................................................................................................47

REFERENCES .........................................................................................................................48
INTRODUCTION
To guide the reader into the work described in this thesis a short introduction with some relevant points of bacterial gene regulation as well as aspects of the bacterium *Staphylococcus aureus* is given.

**Regulation of Bacterial Virulence Factors**
Most bacterial pathogens that enter a human host experience several different environmental conditions during the course of the infectious process. In addition to the ability to survive in different organ systems of a host, many pathogenic bacteria can also live as free organisms outside the host. It can easily be envisioned that the demands on the bacterial physiology vary greatly between the different environments it might experience. Virulence factors can be defined as components of a pathogen that when deleted, specifically impairs virulence but not viability (Wood and Davis, 1980), or as microbial products that permits a pathogen to cause disease (Smith, 1977). Bacteria express virulence factors as a mean of survival and not with the prime purpose to cause disease. From the bacterial point of view virulence factors must be considered as accessory survival factors that only should be expressed when absolutely needed. Expression of virulence factors at the wrong time or place can not only be a disadvantage for the bacteria but also means waste of valuable nutrients and energy.

To be able to cope with the different environmental conditions and demands that may prevail in different locations, the bacteria have to be able to sense the environment and accordingly adapt to new milieus. Factors that make the bacteria optimally fit in specific contexts have to be expressed and conversely, factors that renders the bacteria less fit, repressed. For this, elaborate sensing systems which convey information via signal transduction systems to intricate regulatory networks have evolved.

**Signal transduction**
Different systems have been developed to enable the bacteria to sense the environment and subsequently couple a stimulus to a specific response. Many of these systems are two-component signal transduction systems, which typically consist of a sensor that is a histidine protein kinase (HPK) and a response regulator (RR). Upon a specific stimulus the membrane-located HPK autophosphorylates at a conserved histidine residue, and this phosphate is subsequently relayed to a conserved aspartate residue in the RR. Upon phosphorylation the
RR undergoes a conformational change. In many cases the RR is a DNA binding transcriptional regulator but may also be an enzyme that regulates the activity of specific target molecules through covalent modifications such as demethylation (Amsler and Matsumura, 1995).

Many cellular responses have been shown to be controlled by two-component signal transduction. Starvation for phosphate or nitrogen, responses to oxygen limitation, and adaptation to new carbon and nitrogen sources are only a few environmental changes to which bacteria adapt by the means of two-component systems (Stock et al., 1989). The stimuli sensed by different HPKs are hence diverse, and are in many systems unknown. The diversity of the stimulus is reflected in that the sensing domains of the HPKs are highly variable while the transmitter domains are conserved.

Virulence factors have also been shown to be regulated by typical two-component systems e.g. production of capsule in Pseudomonas aeruginosa (algR-algD), Echerichia coli (rcsB-rcsC) and Klebsiella pneumoniae (rmpA2); resistance to antibiotics in Bacteroides fragilis (rprX-rprY) and Enterococcus faecalis (vanR-vanS); global virulence in Salmonella (phoQ-phoR and spv) and Staphylococcus aureus (agrA-agrC) (Arthur et al., 1992; Fields et al., 1989; Goldberg and Dahnke, 1992; Grob et al., 1997; Jayaratne et al., 1993; Morfeldt et al., 1988; Rasmussen and Kovacs, 1993; Wacharotayankun et al., 1993).

More complex pathways such as sporulation or cell cycle control seem to involve more complicated signal transduction pathways with additional regulator and phosphotransfer domains, called phosphorelay signal transduction. In some instances these additional domains are individual proteins but in many systems the domains are associated with polydomain proteins. This organisation offers more regulatory targets, especially for phosphatases, and can thus provide the regulatory pathway with additional levels of control. Examples of such systems are arcB-arcA involved in respiratory regulation in E. coli, bvgA-bvgAS in the regulation of virulence in Bordetella pertussis and during sporulation in Bacillus subtilis (kinA-spoOF-spoOB-spoOA) (Beier et al., 1995; Fabret et al., 1999; Georgellis et al., 1998).

A special kind of signal transduction is quorum sensing, through which bacteria, by cell to cell signalling, can monitor the population density. In Gram positive bacteria quorum sensing seems to be mediated by small secreted peptides (“pheromones”) that typically interact with the sensor element of a HPK in a two-component system. At low population densities the concentration of pheromone is insufficient to trigger a response. When the bacterial population increases above a specific density, the threshold level of pheromone is reached and the cognate HPK sense the signal and undergoes auto-phosphorylation.
Activation usually involves autoinduction of the transcriptional unit that is responsible for pheromone production, as well as the specific target gene(s), meaning that an exponential response is achieved. In *Staphylococcus aureus* production of virulence factors is controlled by quorum sensing through the *agr* system (Ji et al., 1995) (detailed below “The *agr* locus”). Other traits known to be regulated by quorum sensing in Gram positive bacteria are development of competence in *Bacillus subtilis* (through Com-X and CSF) and *Streptococcus pneumoniae* (through Com-C) and conjugation in *Enterococcus feacialis* (e.g. through cCF10) are (Cheng et al., 1997; Leonard et al., 1996; Magnuson et al., 1994; Pestova et al., 1996; Solomon et al., 1995).

Not all signalling in bacteria involve sensing of the stimuli at the surface of the cell. In Gram negative bacteria quorum sensing utilises N-acyl homoserine lactones (AHL) as signalling molecules. AHL are small diffusible molecules that function intracellularly where they bind and modulate the function of a transcriptional regulator. AHL has been shown to regulate light production in *Vibrio ficheri* and *Vibrio harveyi* (by *luxR* and *luxI*), proteases in *Pseudomonas aeruginosa* (by *lasR* and *lasI*), plant tissue degrading enzymes in *Erwinia carotovara* (by *expR* and *expI*) and conjugation in *Agrobacterium tumefaciens* (by *traA* and *traI*) (Bassler et al., 1994; Passador et al., 1993; Pirhonen et al., 1993; Zhang et al., 1993). In addition to AHLs, alternative Gram negative diffusible quorum sensing molecules have recently been identified, e.g. 3-hydroxy palmitic acid methyl ester (*Xanthomonas campestris*), 2-heptyl-3-hydroxy-4-quinolone (*Pseudomonas aeruginosa*), butyrolactones (*Pseudomonas aureofaciens*) and diketopiperazines (several pseudomonas species, *Enterobacter agglomerans* and *Citrobacter freundii*) (de Kievit and Iglewski, 2000).

**Gene regulation**

As discussed above, bacteria have different ways to couple a specific signal to a specific response. A quick and transient response might be accomplished by activation of already existing proteins while for a permanent and maybe not so quick response, increased transcription or translation might be the best alternative. The amount of virulence factor produced under different conditions can be controlled at a number of levels. In this text the regulatory control mechanisms have been separated into transcriptional and posttranscriptional regulation.
Transcriptional regulation

Productive transcription involves binding of the RNA polymerase (RNAP; holoenzyme including α, α, β, β′ and the σ-factor) to promoter DNA, formation of a closed binary complex, followed by a series of steps leading to reversible opening (isomerisation) of a 10 to 15 bp region at the start site of transcription (open complex formation), incorporation of the first nucleoside triphosphates (initiation) and finally transition into the progressive elongation complex (clearance/escape) with concomitant release of the σ subunit. Each of these steps is critical for successful transcription and constitute possible targets for transcriptional regulation (Record et al., 1996).

A large group of proteins have been shown to activate transcription by facilitating binding of RNAP to the promoter, a mechanism called recruitment (Ptashne and Gann, 1997). Recruitment is primarily seen at promoters with sub optimal –10 and –35 sigma recognition element. Activators that function through recruitment bind to operator sequences located upstream of the RNAP binding site. A direct contact between activator and RNAP is indicated by the demonstration that mutations in the flexible αCTD (C-terminal domain of the alpha subunit) of the RNAP that abrogate activation can be compensated by mutations in the activator (Zhou et al., 1993). It has also been shown that activators functioning through recruitment can be made indispensable by optimising the sigma recognition elements of the promoter and/or adding an UP (upstream recognition element) (Ross et al., 1993). The cyclic AMP receptor protein (CRP) (also referred to as catabolite gene activator protein [CAP]) is one of the best known examples of an activator that functions through recruitment (Dove et al., 1997; Igarashi and Ishihama, 1991).

Similarly, direct contact between the λcI activator and CRP, respectively, to the RNAP σ unit has been proven (Kuldell and Hochschild, 1994). In contrast to activators that interact with the αCTD, activators interacting with the σ unit of the RNAP bind to DNA sequences overlapping the –35 region. For a subset of CRP dependent promoters (Class II dependent promoters) interaction of CRP with the RNAP σ unit has been shown to facilitate open complex formation in addition to recruitment (Niu et al., 1996).

Activators can also function by altering the promoter conformation. MerR, involved in resistance to mercury in E. coli, binds to a site located between the –35 and –10 elements at the Tn501 merP promoter. In the absence of mercury, MerR interacts with RNAP allowing an inactive closed complex to form. Upon interaction with Hg(II), MerR is converted to an activator by a mechanism that involves a DNA-bend modulation, which mediates the
formation of an active open complex. It is shown that MerR induced bending of the suboptimal –35 to –10 spacer region (19 bp) raises the energy required by the polymerase to form an open complex, whereas Hg-MerR untwists the spacer which lowers the energy required and favours formation of an open complex and thereby initiation of RNA synthesis (Ansari et al., 1995). Transmission of duplex destabilisation is another mechanism by which changes in DNA structure leads to activation of transcription by increasing the rate of open complex formation. The ilvP \textsubscript{G} promoter of \textit{Echerichia coli} contains an A+T-rich (88\%) region located between +1 and –160, relative to the transcription start point. At physiological superhelical densities the DNA duplex of this region is destabilised (super-coiling-induced DNA duplex destabilisation [SIDD]). A bend in the DNA, introduced by IHF (Integration Host Factor) binding to this region, activates the promoter through stabilisation of the upstream SIDD region, which results in destabilisation of the –10 region and decreases the energy required for open complex formation.

Gene expression can also be controlled by the interaction of several regulators. In eucaryotic promoters the binding of some activator requires binding of a first activator. In \textit{E. coli} binding of CRP to the \textit{malK} promoter triggers repositioning of the MalK protein to a location where it is able to interact with the RNAP and thereby function as an activator (Richet et al., 1991). The function of an activator can be suppressed by a repressor, that in turn can be neutralised by the binding of a second activator (“anti-repressor”) as shown for the \textit{nir} promoter (Tyson et al., 1997).

Gralla and Collado-Vides (1996) compared the location of DNA binding activators and repressors to 132 \(\sigma\)\textsuperscript{70} dependent \textit{E. coli} promoters. A major conclusion from their analysis was that the location of a regulator is critically related to its function. Repressor binding sites were found to be located between –900 to +400, with the highest score of interaction overlapping the +1 to +10 region, followed by an even distribution between location –50 to +20. In contrast, essentially no activator binding sites were found downstream of –30, this zone was therefore considered to be an “exclusive zone of repression” (indeed, repositioning of an activator from an upstream location to downstream of –30 converted the activator to a repressor (Kredich, 1992)). The majority of activators were found to bind between –30 to –80, however activators were shown to interact with DNA elements as far up as to –252. From the same study it was also concluded that 75\% of the promoters were controlled by a single regulator (49\% by a single repressor and 25\% by a single activator). Only two of the promoters were regulated by a highly specific protein, whereas the remaining 130 promoters were controlled by regulators that also regulated other promoters.
Transcriptional repression is predominately seen in genes having promoters with strong intrinsic activity, that bind the RNA polymerase well. Negative control can be thought of as originating from a system of constitutive expression and the development of repressors a way of increasing economy and efficiency (Jacob and Monod, 1961). Repressors can exert their activity by simply binding to the same site as RNA polymerase and thereby hinder the RNA polymerase from binding to promoter DNA. This mechanism of “competition” is common and can be exemplified by LacI binding to the lacZ promoter (Lewis et al., 1996; Schlax et al., 1995). In agreement with the competition model, DNA mutations that unable LacI binding to the operator site, makes the promoter insensitive to repression by LacI. In addition to the primary operator site many repressors also interact with auxiliary operators, whereby the local concentration of repressor increases and repressor efficiency is increased (Lewis, et al., 1996). Binding of LacI to operator sites, like many other repressors involved in catabolite repression (e.g. GalR, MalT, TrpR) is dependent on the presence or absence of the specific inducer.

Transcriptional silencing, first described in eukaryotes, is a process whereby a large region of DNA is made inaccessible to the RNA-polymerase as well as to other DNA-binding and modifying proteins. Proteins involved in silencing have been shown to spread from sites of nucleation to multiple sites in a process involving co-operative binding and protein-protein interactions (Henikoff, 1996; Wolffe and Matzke, 1999). In E. coli and Salmonella the abundant nucleoid-associated protein, H-NS (histone-like nucleoid –structuring protein) appear to mediate repression through a mechanism resembling that of silencing (Caramel and Schnetz, 1998; Goransson et al., 1990; Schnetz, 1995). H-NS binds with high affinity to a 100 bp AT-rich region, designated SE (silencer element) upstream of the bgl operon and has the capacity to polymerise along DNA from this region. Several features of H-NS function differ from that of classical repressors: i.) The orientation of the SE is of no importance for the silencing function; ii.) Silencing can be alleviated by large deletions or insertions, whilst point mutations or small insertion have no effect on H-NS function; iii.) Substitution of the natural downstream gene for a non H-NS repressed gene alleviates silencing; iv.) Binding of CRP to its natural site within SE or establishment of an artificial tight DNA-protein complexes to this region alleviates silencing. The alleviated repression achieved by tight binding of a protein to the H-NS silencer region strongly suggest that HN-S function can be disrupted by blocking H-NS polymerisation/spreading by “road-blocks”.

Regulation of gene expression at the level of transcriptional elongation by premature termination is an important mechanism for control of gene expression in both in Gram positive and Gram negative bacteria. Premature transcription termination can either be
achieved by intrinsic terminators, composed of a G+C-rich stem-loop followed by a series of U residues (Yarnell and Roberts, 1999) or by terminators dependent on protein factors binding to the nascent transcript (Rho-dependent termination) (Henkin, 1996). In E. coli, Nus factors that bind to the growing RNA and inhibit Rho-dependent termination play a key role in transcription of ribosomal RNA (rrn) operons (Vogel and Jensen, 1997). In Bacillus, Klebsiella and Pseudomonas the SacY, NasR and AmiR proteins, respectively, act as RNA binding anti-terminators (Chai and Stewart, 1999; Manival et al., 1997; O’Hara et al., 1999). A common mechanism for anti-terminating proteins is that they stabilise the DNA-RNAP elongating complex and thereby counteract the destabilising effects exerted by terminator structures (Henkin, 1996). A well studied regulatory mechanism involving termination is transcriptional attenuation. In the trp operon of Bacillus the rate of translation of a leader peptide determines whether a terminating hairpin is formed or not and consequently if transcription is allowed to proceed (Shimotsu et al., 1986).

Posttranscriptional regulation

The steady state level of an mRNA is determined by the rate of transcription initiation as well as by the stability of the RNA. Thus, mRNA stability is an important way of regulating gene expression and degradation of mRNA molecules is therfore strictly controlled. The rate-limiting step in RNA decay seems to be an initial endonucleolytic cleavage of the RNA, which is subsequently followed by degradation to nucleotides by 3’→5’ exonucleases (Melefors et al., 1993). The stability of a specific mRNA molecule is determined by its nucleotide sequence, and secondary structure. Proteins, or antisense RNAs, interacting with specific mRNA molecules may promote stability while others may facilitate degradation (Cleveland, 1989; Wagner et al., 1997) (see below; Regulatory RNAs).

The effectiveness of translation can be regulated by secondary structure of the specific mRNA (see below; “Regulatory RNAs”). In E. coli the post-transcriptional regulator Hfq affect translation and/or stability of several mRNAs (hence a global translational regulator), e.g. the rpoS, mutS, ompA and its own mRNA (Muffler et al., 1996; Tsui et al., 1997; Vytvytska et al., 1998). Attenuation regulates translation of the ermC (adenine methylase) gene of S. aureus (Horinouchi and Weisblum, 1980) and the cat (chloramphenicol adenylate transferase) gene in Staphylococcus and Bacillus (Lovett, 1990). The presence of erythromycin and chloramphenicol, respectively, leads to stalling of the ribosome during translation of a leader peptide. Ribosomal stalling results in resolvment of an intramolecular
structure and freeing of the ribosomal binding site allowing high level translational initiation of the *ermC* and *cat* mRNA, respectively.

Negative feedback autoregulation by the interaction of a protein with its own mRNA is an efficient way to inhibit superfluous amounts of protein being produced and assures balanced levels in the cell. The threonyl-tRNA synthetase in *E. coli* inhibits its own translation through interactions with the leader sequence of its own mRNA (Moine *et al.*, 1990). Also the adenine methylase protein of *S. aureus* inhibits its own synthesis by interaction with the *ermC* mRNA (Denoya *et al.*, 1986).

**Regulatory RNAs**

In addition to the role as a messenger of genetic information, RNA molecules have been shown to exhibit structural, catalytic and regulatory functions. Antisense RNAs, have the ability to base pair with the “sense” or coding RNA, thereby regulating translation of the target mRNA. Most antisense RNAs are small (60–150 nt), untranslated and encoded by the opposite strand to that of its target. One such antisense RNA, CopA, regulates bacterial plasmid copy number through base pairing with the RepA (replication rate limiting Rep protein) mRNA and thereby blocks translation. In *E. coli* the 93-nt MicF RNA inhibits translation and destabilises the OmpF mRNA (encoding an outer membrane porin) by direct RNA-RNA base pairing. Unlike most antisense RNAs, MicF is transcribed at a locus distinct from *ompF*.

The newly described regulatory RNAs involved in stress response in *E. coli*, OxyS, DsrA, and CsrB act as promiscuous regulators that control or modulate the functions of multiple RNAs. OxyS, a 109-nt untranslated RNA, which is induced in response to oxidative stress, activates or represses as many as 40 genes including the *fhlA*-encoded transcriptional activator and the *rpoS*-encoded σ^54_ subunit of RNA polymerase. By base-pairing with a short sequence (7 bp) overlapping the ribosome binding site, *fhlA* mRNA translation is inhibited by OxyS (Altuvia *et al.*, 1998). In the case of *rpoS* mRNA, OxyS inhibits translation by binding the Hfq protein which is required for *rpoS* mRNA translation (Zhang *et al.*, 1998a). The 87-nt, DsrA RNA inhibits translation of *hns* mRNA by base pairing to a 13 nt region located immediately downstream of the AUG start codon of H-NS. There are also examples of antisense RNAs that activate translation as shown for the regulatory RNAs, RNAIII and DsrA (Lease *et al.*, 1998; Majdalani *et al.*, 1998; Morfeldt *et al.*, 1995). RNAIII which is transcribed from the *agr* locus in *S. aureus* (see below, “The *agr* locus”), activates translation.
of alpha-hemolysin mRNA by resolving an intramolecular basepairing that unables the ribosome to access the mRNA. The same mechanism has been demonstrated for DsrA in the activation of RpoS translation. The CsrB (carbon storage regulator) RNA is a ~360 nt non-coding molecule which antagonise the activity of CsrA, by forming a globular complex with 18 CsrA polypeptides. The function of CsrA is to bind and facilitate decay of specific mRNA molecules of exponential-phase metabolic pathway genes (Romeo, 1998).

Interestingly, naturally occurring antisense RNAs, basepair with very short regions of their target RNA. A recent study by Franch and co-workers (1999) revealed a specific U-turn motif in the recognition loops of naturally occurring antisense RNAs and their targets. The U-loop structure was found to act as a rate-of-binding enhancer in the hok/sok antisense system of plasmid R1, and seems to be a mechanism that is conserved to the majority of naturally occurring antisense RNA-regulated gene systems (Franch and Gerdes, 2000; Franch et al., 1999).

**Staphylococcus aureus**

Staphylococci constitute a considerable part of the normal skin-flora of humans and several different mammals. The genus which consists of over 40 defined species is characterised as catalase-positive, highly lipid and salt tolerant Gram positive cocci with a low G + C content. Staphylococci have the ability to divide in more than one plane and therefore forms characteristic irregular clusters when grown in liquid as well as solid media (the word staphyle is Greek for cluster).

The genus contains several species that can cause disease in humans. *Staphylococcus aureus* is the most frequent and versatile pathogen and can distinguished from the other members of the staphylococcal genus by its ability to produce the extracellular enzyme coagulase. Although *S. aureus* is a member of the normal flora in 25-30% of the population it is also one of the most common pathogens causing minor lesions such as skin abscesses and wound infections as well as more serious infections, e.g. septicaemia, endocarditis, septic arthritis, and toxic shock (Tabel. 1). The primary site of infection is often the skin from where the infecting organism spreads and infects deeper tissues where it causes a variety of symptoms depending on the organ colonised. Osteomyelitis, septic arthritis and endocarditis are the most common deep infections but abscesses in brain, lungs and kidneys may occasionally occur. Patients with breaches in the skin due to surgery, central venous lines or
burns are particularly prone to *S. aureus* infections. Due to its ubiquity and ability to survive outside the body, infections caused by *S. aureus* are highly prevalent both in the community and in hospital settings.

*Staphylococcus saprophyticus* is a common cause of urinary tract infections in fertile women, and *Staphylococcus epidermidis* one of the most common sources of foreign body associated infections and septicaemia in newborn and immunocomprised individuals.

*S. aureus* appears to differ considerably from the coagulase-negative staphylococci (CoNS) with respect to its pathogenicity, in that it produces a large number of soluble extracellular virulence factors (Table 3.) while the CoNS generally produce few of these factors.

**Table 1. Diseases caused by *Staphylococcus aureus***

<table>
<thead>
<tr>
<th>Disease</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abscesses in all tissues</td>
</tr>
<tr>
<td>Endocarditis</td>
</tr>
<tr>
<td>Folliculitis</td>
</tr>
<tr>
<td>Food poisoning</td>
</tr>
<tr>
<td>Osteomyelitis</td>
</tr>
<tr>
<td>Pneumonia</td>
</tr>
<tr>
<td>Scalded skin syndrome</td>
</tr>
<tr>
<td>Septic arthritis</td>
</tr>
<tr>
<td>Septicaemia</td>
</tr>
<tr>
<td>Skin and wound infections</td>
</tr>
<tr>
<td>Toxic Shock Syndrome</td>
</tr>
</tbody>
</table>

**Virulence factors**

With the exception of the toxinsoses, toxic shock syndrome (caused by toxic shock syndrome toxin-1 (Bohach *et al.*, 1990)), staphylococcal scalded skin syndrome (caused by the exfoliative toxins A and B (Iandolo, 1989)) and staphylococcal food-poisoning (caused by the staphylococcal enterotoxins (Iandolo, 1989)) the virulence of *S. aureus* is considered multifactorial. Different sets of virulence factors are thought to be important for pathogenicity in different stages during the course of infection. During early course of infection, factors involved in attachment of the bacteria to cells or extracellular matrixes are believed to be of key importance while factors involved in invasion and evasion of host defence mechanisms come into play in later stages. All together more than 40 different extracellular and cell-surface associated proteins have been identified (Table 3.). In addition several genes have
been found in the *S. aureus* genome database, which have been proposed to encode virulence factors, based on their homology to known virulence factors. Based on their biological activities virulence factors may be divided into three functional categories, *i.e.* those that mediate adhesion of bacteria to cells or tissues, those that promote invasion and those that protect the bacteria from the host immune system. Some virulence factors may have more than one function *e.g.* alpha-hemolysin and the fibrinogen binding proteins. Alpha-hemolysin, which in addition to its ability to promote invasion by lysing tissue cells, has a pronounced activity against human monocytes and neutrophils and may therefore also be regarded as an immune escape factor (Bhakdi and Tranum-Jensen, 1991; Gemmell *et al.*, 1982). The fibrinogen binding proteins can both make the bacteria adhere to surfaces coated with fibrinogen, and function as an immuno protection factor by masking the bacteria with soluble fibrinogen (Wann *et al.*, 2000).

The pathology of *S. aureus* disease is highly variable, and most likely, different sets of virulence factors are important in different types of disease. Several animal models have been developed to study pathogenic mechanisms of *S. aureus* infection. Pathogenicity of isogenic strains with mutations in specific virulence genes, as well as the protective effect of immunisation with specific virulence factors, has been evaluated. In a rat model of endocarditis, fibronectin binding protein (Kuypers and Proctor, 1989; Schennings *et al.*, 1993), capsular polysaccharide (Baddour *et al.*, 1992; Lee *et al.*, 1997), clumping factor (Moreillon *et al.*, 1995) and collagen adhesin (Hienz *et al.*, 1996) have been shown to be important for pathogenesis. In a rat wound infection model the extracellular fibrinogen binding protein, Efb, was demonstrated to be important (Palma *et al.*, 1996). A pivotal role for collagen adhesin (Patti *et al.*, 1994), bone sialoprotein (Bremell *et al.*, 1991) and polysaccharide capsule (Nilsson *et al.*, 1997) have been proven in a murine septic arthritis model. Alpha-hemolysin has been shown to be important in several animal models of infection such as the endocarditis, keratitis, arthritis and septicemia models, respectively (Bayer *et al.*, 1997; Callegan *et al.*, 1994; Gemmell *et al.*, 1997; Kernodle *et al.*, 1997). It has also been shown that passive immunisation with anti-bodies against alpha-hemolysin confer protection against lethal challenge of alpha hemolysin using a rabbit model (Menzies and Kernodle, 1996).

Genes important for survival *in vivo* have been identified by the use of signature-tagged mutagenesis (STM) and *in vivo* expression technology (IVET). In the two STM screens performed in *S. aureus* (Coulter *et al.*, 1998; Mei *et al.*, 1997), the majority of the identified genes encoded proteins involved in housekeeping functions such as transport, energy
metabolism and amino acid biosynthesis. However more classical virulence genes such as the global regulator SarA (sarA), V8 serine protease (ssp) and the type 5 capsular polysaccharide (cap5) were also found, as well as a new lipase and a protein similar to streptococcal M proteins. In both screens a homologue to the clpX stress response gene was found, suggesting a link between stress response and in vivo survival. Also several genes encoding proteins of unknown function were identified of which some showed homology to members of two-component signal transduction systems. As many as 45 in vivo transcriptionally induced genes were identified in the IVET screen performed by Lowe et al (Lowe et al., 1998). Of these, three were known virulence factors, i.e. the genes coding for the accessory gene regulator AgrA, glycerol ester hydrolase (geh) and type 8 capsular polysaccharide (cap8), respectively, while the majority encoded proteins involved in transport and other biological functions. Many of the genes identified by IVET had an unknown function. Interestingly, seven of these seemed to affect virulence when re-tested in a murine renal abscess model.

Though the importance of some adhesins and toxins has been demonstrated in animal models of infection, the suggested role for most virulence factors is based on their biological effects in vitro. The ability of protein A to bind to the Fc portion of IgG anti-bodies is well documented, whether this could interfere with opsonisation as suggested is however not demonstrated in vivo (Kronvall and Gewurz, 1970; Verhoef et al., 1979). Mutants deficient in protein A do not show significant attenuation in animal models of infection (Callegan, et al., 1994; Patel et al., 1987). Also the ability of coagulase to convert soluble fibrinogen to fibrin is well documented in vitro. However, no difference in virulence has been demonstrated between coagulase expressing and non-expressing strains when tested in several animal models of infection (e.g. experimental endocarditis, subcutaneous wound infection and mastitis (Baddour et al., 1994)).

Thus it can be concluded that the pathogenesis of S. aureus is complex and in most cases multifactorial. This is demonstrated by the fact that most mutant strains lacking single toxins or enzymes show little attenuation of virulence and are still highly pathogenic in most animal models of infection. On the other hand strains defective in the global regulatory systems, affecting the expression of several virulence factors, typically show a significantly decreased virulence in animal models of infection (Abdelnour et al., 1993; Cheung et al., 1994; Giraudo et al., 1996; Nilsson et al., 1999) (Table 2.).
Table 2. Effect on virulence of agr and sarA

<table>
<thead>
<tr>
<th>Animal model</th>
<th>wild type</th>
<th>agr⁻</th>
<th>sarA⁻</th>
<th>agr⁻ sarA⁻</th>
<th>ref.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mureine arthritis</td>
<td>70% (severe)</td>
<td>10%</td>
<td>21%</td>
<td>n.d.</td>
<td>2, 129</td>
</tr>
<tr>
<td>Rabbit endocarditis</td>
<td>90%</td>
<td>40%</td>
<td>0%</td>
<td>0%</td>
<td>34</td>
</tr>
<tr>
<td>-low infectious dose</td>
<td>100%</td>
<td>80%</td>
<td>91%</td>
<td>29%</td>
<td></td>
</tr>
<tr>
<td>Rabbit endophthalmitis</td>
<td>80%</td>
<td>37%</td>
<td>72%</td>
<td>4%</td>
<td>22</td>
</tr>
<tr>
<td>Rabbit osteomyelitis</td>
<td>92% (severe)</td>
<td>46%</td>
<td>n.d.</td>
<td>n.d.</td>
<td>63</td>
</tr>
<tr>
<td>-low infectious dose</td>
<td>83% (severe)</td>
<td>17%</td>
<td>n.d.</td>
<td>n.d.</td>
<td></td>
</tr>
</tbody>
</table>

The numbers represent percentage of animals with pathological findings. n.d.=not determined

Expression and regulation of virulence factors

Abbis-Ali and Coleman (1977) found that most soluble extracellular proteins were produced mainly during post-exponential growth (5 to 10 times higher than during exponential growth). In the 70-ties and 80-ties, several authors reported on pleiotropic mutants with simultaneous changes in the production of several exoproteins indicating the existence of global virulence regulators (Björklind and Arvidson, 1980; Duval-Iflah et al., 1977; Forsgren et al., 1971; Kondo and Katsuno, 1973; Omenn and Friedman, 1970; Yoshikawa et al., 1974). Until now at least six global regulatory loci (agr, sarA, sarH1, sae, rot and 1E3) have been identified. In addition, some of the genes identified by STM (see above) seem to represent regulators of virulence genes (Coulter, et al., 1998). The staphylococcal alternative sigma factor, σB, has also been shown to be involved in regulation of virulence factors, both directly (coa and clfA (Nicholas et al., 1999)) and through its regulation of sarA and sarH1.

The agr locus

The agr (accessory gene regulator) locus was originally identified as a chromosomal Tn551 insertion resulting in a pleiotropic phenotype with increased production of secreted toxins and enzymes, e.g. alpha-hemolysin and proteases, and a decreased production of coagulase, protein A and other cell wall associated proteins (Morfeldt, et al., 1988; Peng, et al., 1988; Recsei, et al., 1986). The locus consists of two divergent transcriptional units, one coding for the agrB, D, C, and A genes and the other for the regulatory molecule RNAIII (Janzon, 1989; Janzon, et al., 1989; Novick et al., 1995; Novick, et al., 1993) (Fig. 1.). AgrA and AgrC are homologous to protein belonging to the family of response regulators and histidine protein kinase sensors, respectively, of classical two-component signal transduction systems. The membrane located AgrC is activated by an octapeptide pheromone encoded within the agrD
gene, designated AIP, for auto inducing peptide (Ji, et al., 1995; Lina et al., 1998). AgrD is modified and secreted through the involvement of the membrane located AgrB protein (Ji et al., 1997; Mayville et al., 1999; Novick, 2000; Otto et al., 1998). The system is double autocatalytic in the sense that it produces its own transcriptional activator (AgrA) and its own inducing pheromone (AIP). Induction of the agr system leads to activation of the divergent transcribed RNAlII gene that appears to be the actual effector molecule of the agr dependent virulence gene regulation (Janzon, et al., 1989; Novick, et al., 1993). RNAlII is also an mRNA coding for delta-hemolysin (hld) (Janzon, et al., 1989). By different deletion analysis and non-sense mutations in the hld gene it was demonstrated that it is in fact the 514 nt RNA molecule that is responsible for the regulation and not the delta-hemolysin or any other gene product encoded within RNAlII. As delta-hemolysin is only 26 amino acids long most of the RNAlII molecule is composed of non-coding sequence.

**Figure 1.** The agr system (adapted from E. Morfeldt, (1996))
RNAIII upregulates at least 16 genes encoding toxins and enzymes and repress at least seven genes encoding surface associated proteins (Table 3.). Some virulence factors have been shown to be unaffected by \textit{agr}, e.g. fibronectin binding protein B (\textit{fnbB}), clumping factor A (\textit{clfA}) and enterotoxin A (\textit{entA}) (Gillaspy, \textit{et al}, 1997; Tremaine, \textit{et al}, 1993; Wolz, \textit{et al}, 1996). Generally RNAIII regulates the level of transcribed mRNA. A direct transcriptional effect of RNAIII is indicated by fusion experiments using the promoters of alpha-hemolysin (\textit{hla}), beta-hemolysin (\textit{hlb}) protein A (\textit{spa}) and exfoliative toxin A (\textit{etaA}), toxic shock syndrome toxin-1 (\textit{tst}) genes, respectively (Chan and Foster, 1998b; Novick, \textit{et al}, 1993; Patel, \textit{et al}, 1992; Sheehan, \textit{et al}, 1992). The case of alpha-hemolysin RNAIII activates translation in addition to transcription. By \textit{in vitro} experiments an interaction between the untranslated leader sequence of the \textit{hla} mRNA and the 5´-part of RNAIII was demonstrated and suggested to mediate the activation of alpha-hemolysin translation (Morfeldt, \textit{et al}, 1995). Though alpha-hemolysin is the only gene which has been demonstrated to be activated on the level of translation, several observations suggests that additional genes might be regulated on this level. Indications that RNAIII can interact with other mRNAs and thereby might affect translation or mRNA stability comes from data in our lab e.g. migration of whole cell RNA in a native polyacrylamide gel shows that the RNAIII molecule migrates at several different positions indicating retardation through complex formation with other RNA molecules, also has complementarity between the untranslated leader sequence of the enterotoxin D gene (\textit{sed}) and RNAIII has been demonstrated (Morfeldt, 1996). In addition Novick \textit{et al} (1993) found that deletions in the 5´-end of RNAIII affected translation of several exoproteins.

have diverged in concert in order to maintain specificity in processing and activation of the system. How this has occurred is not known. The fact that AIP molecules repress the *agr* response of non-cognate subgroups is an interesting observation of bacterial interference. A therapeutic potential of non-cognate AIP molecules as inhibitors of the *agr* response has been demonstrated by Mayville *et al* (1999). By simultaneous inoculation of wild type bacteria with inhibiting AIP molecules, *S. aureus* induced lesions in mice was efficiently suppressed.

In addition to the octapeptide, a 38 kDa protein RAP (RNAIII activating protein), has been proposed to induce RNAIII by an *agr* independent pathway (Balaban *et al*., 1998). However, the significance of RAP has been questioned (Novick *et al*., 2000).

<table>
<thead>
<tr>
<th>Low bacterial density:</th>
<th>High bacterial density:</th>
<th>As bacteria escape from the microcolony</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>agr</em> off, no production of toxins and enzymes.</td>
<td><em>agr</em> on, downregulation of adhesins.</td>
<td>Adhesins are up-regulated again</td>
</tr>
<tr>
<td>Adhesins are expressed</td>
<td>Toxins and enzymes are produced</td>
<td></td>
</tr>
</tbody>
</table>

*Figure 2*. Possible role of *agr* in the infectious process.
Thus the *agr* system is a quorum sensing system that senses cell density through the octapeptide pheromone (AIP), produced from within the *agr* operon. Considering the pathogenesis of *S. aureus* infections with formations of focal infections such as soft tissue abscesses and endocardial vegetations that commonly are associated with episodes of bacterimia or septicaemia, a role for the *agr* system can be hypothesised. When bacteria are few (e.g. the blood stream) the concentration of pheromone is low and the *agr* system is therefore not activated and production of coagulase and cell surface adhesion molecules is allowed. With the production of fibronectin binding protein, collagen binding protein, fibrinogen binding protein, etc., adhesion to various tissues is promoted. Successful colonisation leads to an increased bacterial population and consequently activation of the *agr* system. Activation leads to production of the regulatory molecule, RNAIII, which repress the synthesis of cell surface adhesion molecules and reciprocally activate the expression of extracellular toxins and enzymes facilitating tissue degradation and anew spread and dissemination (Fig. 2.).

**The sarA locus**

In 1992 a second regulatory locus with pleiotropic effects on the production of virulence genes was identified by Cheung *et al* (1992) (Cheung *et al.*, 1992). The *sarA* locus was found in a transposon mutagenesis screen, searching for a fibrinogen binding negative *S. aureus* mutant. It has later been shown that the most pronounced phenotype of a *sarA* mutant is the increased production of proteases, which might explain the reduced expression of cell surface proteins (Chan and Foster, 1998b; McGavin *et al.*, 1997). Under microaerobic growth conditions, *sarA* also stimulates the expression of RNAIII, this means that *sarA* mutants show increased production of cell surface molecules and a decreased expression of secreted toxins and enzymes when grown under oxygen limiting conditions (Cheung, *et al.*, 1992; Cheung and Projan, 1994; Lindsay and Foster, 1999). The product of the *sarA* locus is a small (14.7 kDa), dimeric and basic protein (pI 9.2) with predominantly alpha helical structure (Cheung and Projan, 1994; Rechtin *et al.*, 1999). SarA shows no significant similarity to known regulators. SarA is transcribed from three distinct promoters (P1, P2 and P3) and terminates at a common 3´end, thus generating three transcripts of different size (0.56, 0.8 and 1.2 kb) (Bayer *et al.*, 1996) (Fig. 3.). The P1 and P2 promoters are recognised by the vegetative sigma factor, σA, and are mainly expressed during early exponential phase of growth, while P3 which is recognised by the alternative sigma factor, σSB, is induced as the cells enter postexponential phase of growth (Bayer, *et al.*, 1996; Deora *et al.*, 1997; Manna *et al.*, 1998).
Upstream of P1 and P3 are two small open reading frames (orf). Deletions in these orfs affects the regulatory ability of the \textit{sarA} locus, however promoter fusion experiments indicated that these effects are due to altered production of SarA rather than to the disruption of the orfs (Manna, \textit{et al}., 1998). Though the presence of a triple promoter system (also conserved in \textit{S. epidermidis}) indicates that the amount of SarA produced is strictly controlled, quantitative analysis of SarA in the bacteria under different growth stages \textit{in vitro} show surprisingly little variation (Blevins \textit{et al}., 1999). To investigate if the individual \textit{sar} promoters are differentially expressed \textit{in vitro} and \textit{in vivo}, promoter constructs with each of the \textit{sar} promoters (P1, P2 and P3, respectively) was made with the \textit{gfp} (green fluorescent protein) reporter gene (Cheung \textit{et al}., 1998). \textit{In vitro} the \textit{sar} P1 promoter was expressed at very high levels as compared to P2 and P3 which were almost silent. Further, the P1 an P2 promoters expressed the highest levels in early logarithmic phase with a slight decrease in expression as the bacteria enter stationary phase of growth while P3 showed the opposite kinetics with low levels during logarithmic phase and increasing expression as the bacteria enter stationary phase of growth. Interestingly \textit{in vivo} (rabbit endocarditis model), both the P1 and P2 promoters were highly active while the P3 promoter was silent. In particular the P2 promoter was activated on the surface of the endocardial vegetations (Cheung, \textit{et al}., 1998).

Inactivation of the \textit{sarA} locus led to attenuation of virulence in several animal models of infection (Booth \textit{et al}., 1997; Cheung, \textit{et al}., 1994; Gillaspy, \textit{et al}., 1995; Nilsson, \textit{et al}., 1997) (Table 2.). In the endocarditis model, where bacterial adhesins are believed to be of key importance, the \textit{sarA} mutant was completely avirulent (Cheung, \textit{et al}., 1994).

\textbf{Figure 3.} The \textit{sarA} locus
1E3

1E3 was defined by a Tn551 transposon insertion into a unique chromosomal locus of *S. aureus* (Cheung *et al.*, 1995). The insertion, which has not been mapped or sequenced, gives a pleiotropic effect on expression of both extracellular and cell wall associated proteins. In particular transcription of protein A, clumping activity with fibrinogen and fibrinogen binding was significantly decreased, while mRNA levels of alpha-hemolysin and RNAIII was only modestly increased. The effect of the 1E3 mutation on virulence has been assessed by a rabbit endocarditis model of infection. At low infectious doses ($10^3$ colony forming units (cfu)) the mutant was avirulent as compared to the wild type which caused endocarditis in 66% of the rabbits using the same inoculum. At a higher infectious dose ($10^5$ cfu), 87% of the rabbits developed endocarditis when infected with the wild type bacteria as compared to 66% when inoculated with the mutant strain.

sae

The two-component signal transduction system *sae* (*S. aureus* exoprotein expression), was identified by transposon mutagenesis, screening for altered exoprotein production (Giraudo *et al.*, 1999; Giraudo, *et al.*, 1994). *sae* has been shown to stimulate the transcription of alpha-hemolysin, beta-hemolysin and coagulase by a pathway that does not involve *agr* (Giraudo, *et al.*, 1997). The actual stimuli activating *sae* is unknown. A *sae* mutant shows decreased virulence in an intraperitoneal mouse model of infection. The lethal dose of the *sae* mutant was 32 times higher than that of the parental strain (Giraudo, *et al.*, 1994).

rot

Based on the observation that protein synthesis inhibitors suppress production of several virulence factors that are positively regulated by *agr*, Balaban and Novick (1995) suggested that protein factors must be required for RNAIII to exert its regulatory function. To identify such factors, mutations that suppressed the *agr* phenotype were screened for using Tn917 transposon mutagenesis (McNamara, *et al.*, 2000). Three independent clones with a mutation that suppressed both the protease- and alpha-hemolysin negative phenotype of the *agr* mutant were found. Southern blot analysis revealed that all three mutants had the Tn917 insertion in the same gene. Nucleotide sequence analysis showed that the transposon interrupted a 498 bp open reading frame predicted to encode a 161 amino acid protein. The gene was named *rot* for repressor of toxins. BLAST searches with *rot* revealed a high degree of similarity to the SarA
protein of both \textit{S. aureus} and \textit{S. epidermidis} over the entire length of the protein. Northern blot analysis demonstrated that the \textit{rot} mutation affected \textit{hla} expression on the level of mRNA. Comparison of alpha-hemolysin activity in a wild type, \textit{agr} mutant and \textit{agr-rot} mutant strain showed that while the \textit{rot} mutation restored alpha-hemolysin activity in the \textit{agr} mutant, the mutation had no effect on alpha-hemolysin activity in the wild type strain. Since an effect of the \textit{rot}-mutation was seen only in the absence of \textit{agr} it was hypothesised that \textit{rot}-associated activity is altered by an \textit{agr} product, presumably RNAIII.

**Sigma B**
Under less favourable conditions, such as lack of nutrients and exposure to high salt or low pH many bacteria induce stress response programs that renders the organism more apt to endure and survive various environmental stresses. In several Gram negative bacteria it has been shown that stress-responses in part are orchestrated by the alternative sigma factor, $\sigma^S$, encoded by \textit{rpoS} (Hengge-Aronis, 1996; Kolter \textit{et al}., 1993; Spector, 1998). In \textit{S. aureus} an alternative sigma factor, $\sigma^{SB}$, is induced upon entry into stationary phase and has been shown to be important in the recovery from heat shock and in acid and hydrogen peroxide resistance (Chan \textit{et al}., 1998; Kullik and Giachino, 1997). Recognition sequences for $\sigma^{SB}$ has been found upstream of the genes coding for coagulase (\textit{coa}), SarA (\textit{sarA}, P3), SarH1,(\textit{sarH1}, P2) alkaline shock protein (\textit{asp23}), thermonuclease (\textit{nuc}), clumping factor A (\textit{clfA}) and the staphyloxantin biosynthesis operon (Deora, \textit{et al}., 1997; Kullik \textit{et al}., 1998; Nicholas, \textit{et al}., 1999) (Paper III). Of note is that one of the most commonly used laboratory strains, 8325-4, is defect in the regulation of $\sigma^{SB}$, due to a mutation in the anti sigma factor, \textit{RsbU}.

Although $\sigma^{SB}$ seems to be involved in the regulation of many virulence genes, a role of $\sigma^{SB}$ in pathogenicity could not be demonstrated in various animal models, such as the murine subcutaneous abscess-, wound infection- and hematogenous pyelonefritis model as well as the rat osteomyelitis model (Chan, \textit{et al}., 1998; Nicholas, \textit{et al}., 1999).

**Environmental factors affecting virulence gene expression**
Several environmental factors affect the expression of exoproteins. Glucose has been shown to have a general repressive effect on exoprotein production independent of pH (Coleman, 1983; Coleman \textit{et al}., 1989). Specifically, glucose repressed the transcription of enterotoxins A and C (Hallis \textit{et al}., 1991; Regassa, \textit{et al}., 1991; Regassa \textit{et al}., 1992), alpha-hemolysin
(Chan and Foster, 1998; Ohlsen et al., 1997), TSST-1 (Chan and Foster, 1998a), and protein A (Chan and Foster, 1998a; Taylor and Arvidson, 2000). Repression of agr was seen with glucose at a pH below 5.5 (glucose or pH 5.5 alone did not repress agr) further it was shown that expression of agr was maximal at pH 7 while a pH above 8 inhibited agr expression (Regassa, et al., 1992).

The effect of adding NaCl to the growth medium has also been studied. Interestingly several toxins (enterotoxins B and C (Iandolo and Shafer, 1977; Regassa and Betley, 1993), exfoliative toxin A (etaA) (Sheehan, et al., 1992), alpha-hemolysin (Chan and Foster, 1998a; Lindsay and Foster, 1999), TSST-1 (Chan and Foster, 1998a), epidermolytic toxin (Gillaspy, et al., 1998; Gillaspy, et al., 1997; Sheehan, et al., 1992) and protein A (Chan and Foster, 1998a) were repressed by NaCl while transcription of serine protease was stimulated (Lindsay and Foster, 1999). Changes in osmolarity have been shown to affect supercoiling and thereby gene expression, (Dorman, 1991). Sheehan et al., (1992) showed that addition of 0.7 M NaCl to the growth medium resulted in an increased degree of negative supercoiling. It was further hypothesised that DNA topology was involved in regulation of etaA since addition of novobiocin (a DNA gyrase inhibitor) reversed the etaA repression induced by NaCl. This is however not the case for all genes shown to be affected by salt. In the study performed by Chan and Foster (1998a) addition of novobiocin did not counteract the NaCl induced repression of tst, hla and spa.

Adding metal chelators to the medium, such as EDTA and EGTA, decreased the expression of protein A but increased the expression of TSST-1 (Chan and Foster, 1998a). The effect of divalent metal ions (Mg 2+) on the expression on TSST-1 is however controversial (Kass et al., 1988; Mills et al., 1986; Sarafian and Morse, 1987; Taylor and Holland, 1988).
### Table 3. Virulence factors in S. aureus; production and effect of global regulators

<table>
<thead>
<tr>
<th>Virulence factor</th>
<th>Gene</th>
<th>Effect of global regulators</th>
<th>Refs</th>
<th>1=refers to regulatory reference, + activation, - repression a=mRNA level, b=protein level, c=protein activity level, *= highly aerobic conditions, **=microaerophilic conditions, n.d.=not determined</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Toxins</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>alpha hemolysin</td>
<td>hla</td>
<td>+abc</td>
<td></td>
<td></td>
</tr>
<tr>
<td>beta hemolysin</td>
<td>hlb</td>
<td>+abc</td>
<td></td>
<td></td>
</tr>
<tr>
<td>delta hemolysin</td>
<td>hld</td>
<td>+abc</td>
<td></td>
<td></td>
</tr>
<tr>
<td>enterotoxin A and J</td>
<td>ent A, ent J</td>
<td>0ab</td>
<td></td>
<td></td>
</tr>
<tr>
<td>enterotoxin B</td>
<td>entB</td>
<td>+ab</td>
<td></td>
<td></td>
</tr>
<tr>
<td>enterotoxin C and D</td>
<td>entC and D</td>
<td>+ab</td>
<td></td>
<td></td>
</tr>
<tr>
<td>enterotoxin E-I</td>
<td>entE-J</td>
<td>n.d.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>exfoliative/epidermolytic toxin A</td>
<td>etaA</td>
<td>+ab</td>
<td></td>
<td></td>
</tr>
<tr>
<td>exfoliative/epidermolytic toxin B</td>
<td>etaB</td>
<td>+b</td>
<td></td>
<td></td>
</tr>
<tr>
<td>gamma hemolysin/leukocidin R</td>
<td>hlgA-C/lukR</td>
<td>+c</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Panton-Valentini toxin</td>
<td>lukS/F-PV</td>
<td>n.d.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>toxic shock syndrome toxin-1</td>
<td>tst</td>
<td>+a</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Enzymes</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>alkaline/acid phosphatase</td>
<td>blaZ</td>
<td>0b</td>
<td></td>
<td></td>
</tr>
<tr>
<td>beta lactamase</td>
<td>kat</td>
<td>–a</td>
<td></td>
<td></td>
</tr>
<tr>
<td>catalase</td>
<td>coa</td>
<td>–bc</td>
<td></td>
<td></td>
</tr>
<tr>
<td>coagulase</td>
<td>sspB</td>
<td>n.d.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>fatty acid modifying enzyme (FAME)</td>
<td>geh</td>
<td>+a</td>
<td></td>
<td></td>
</tr>
<tr>
<td>glycerol ester hydrolase</td>
<td>hysA</td>
<td>+c</td>
<td></td>
<td></td>
</tr>
<tr>
<td>hyaluronate lyase</td>
<td>plc</td>
<td>+c</td>
<td></td>
<td></td>
</tr>
<tr>
<td>lipase/esterase</td>
<td>lip</td>
<td>+c</td>
<td></td>
<td></td>
</tr>
<tr>
<td>metalloprotease/proteaseIII/aureolysin</td>
<td>aura</td>
<td>+ab</td>
<td></td>
<td></td>
</tr>
<tr>
<td>nuclease/thermonuclease</td>
<td>nuc</td>
<td>+ab</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PI-phospholipase C</td>
<td>plc</td>
<td>+c</td>
<td></td>
<td></td>
</tr>
<tr>
<td>staphopain/proteaseII</td>
<td>sscp</td>
<td>+ab</td>
<td></td>
<td></td>
</tr>
<tr>
<td>staphylokinase</td>
<td>sak</td>
<td>+c</td>
<td></td>
<td></td>
</tr>
<tr>
<td>V8 serine protease</td>
<td>sspA, sasp</td>
<td>+abc</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Surface proteins</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>bone sialoprotein binding protein</td>
<td>cap5</td>
<td>+b</td>
<td></td>
<td></td>
</tr>
<tr>
<td>capsular polysaccharide type 5</td>
<td>cap8</td>
<td>n.d.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>capsular polysaccharide type 8</td>
<td>elf A</td>
<td>0abc</td>
<td></td>
<td></td>
</tr>
<tr>
<td>clumping factor A</td>
<td>elfA</td>
<td>0abc</td>
<td></td>
<td></td>
</tr>
<tr>
<td>clumping factor B</td>
<td>elfB</td>
<td>n.d.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>collagen binding protein</td>
<td>cna</td>
<td>0abc</td>
<td></td>
<td></td>
</tr>
<tr>
<td>extracellular fibrinogen binding protein</td>
<td>efb/fib</td>
<td>n.d.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>fibronectin binding protein A</td>
<td>fnb A</td>
<td>–ab</td>
<td></td>
<td></td>
</tr>
<tr>
<td>fibronectin binding protein B</td>
<td>fnb B</td>
<td>0a</td>
<td></td>
<td></td>
</tr>
<tr>
<td>lactoferrin-binding protein</td>
<td>n.d.</td>
<td>n.d.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>laminin binding protein</td>
<td>n.d.</td>
<td>n.d.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>lecithin-like protein</td>
<td>n.d.</td>
<td>n.d.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>major histocompatibility complex class</td>
<td>map</td>
<td>n.d.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>II analogous protein</td>
<td>n.d.</td>
<td>n.d.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>plasminogen binding protein</td>
<td>spa</td>
<td>–ab</td>
<td></td>
<td></td>
</tr>
<tr>
<td>protein A</td>
<td>spa</td>
<td>–ab</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sdr A-D</td>
<td>n.d.</td>
<td>n.d.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>thrombospondin binding protein</td>
<td>n.d.</td>
<td>n.d.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>vitronectin binding protein</td>
<td>n.d.</td>
<td>n.d.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

1=refers to regulatory reference, + activation, - repression a=mRNA level, b=protein level, c=protein activity level, *= highly aerobic conditions, **=microaerophilic conditions, n.d.=not determined
THE PRESENT STUDY

Aim of the study
The aim of this work was to study the molecular mechanisms of action of the global regulatory systems \textit{agr} and \textit{sarA} in the pathogenic bacteria \textit{Staphylococcus aureus}. The importance of these systems for \textit{Staphylococcus aureus} virulence has been demonstrated in several animal models of infection. Understanding how these regulatory networks govern the expression of virulence factors is of vital importance for a full comprehension of \textit{S. aureus} pathogenesis.

Results and discussion

Transcriptonal control of RNAIII
Though transcription of RNAIII seems to depend primarily on the products of the \textit{agrBDCA} operon, it is also affected by other factors such as \textit{sarA}, IE3, glucose and pH. The promoter P2, driving \textit{agrBDCA} transcription is separated from the divergent RNAIII promoter, P3, by only 120 bp. The intergenic region contains several elements of repetition and dyad symmetry. An inverted repeat of 17 bp, that is located exactly between the two promoters, has been suggested to constitute the binding site of a bi-directional common regulator, presumably AgrA (Janzon, 1989; Novick, \textit{et al}., 1995). By a more thorough analysis of the intergenic region, four heptanucleotide repeats centred at -45 and -66 upstream of the transcription start point of both P2 and P3, respectively, were found (\textit{Paper I}). The two middle repeats overlapped the 17 bp inverted repeat reported previously. Sequence analysis of RNAIII homologues in \textit{S. epidermidis}, \textit{S. simulans} and \textit{S. warneri} showed that the heptanucleotide repeats were highly conserved in all species (\textit{Paper II}). As the CoNS RNAIII promoters were regulated by the \textit{S. aureus agrBDCA}, as well as by their cognate \textit{agr} operons, it was suggested that the heptanucleotide repeats constitute the \textit{agr} responsive element (\textit{Paper II}). This was further supported by the observation that almost identical heptanucleotide repeats in the promoter region of the \textit{ssp} and \textit{pln} operons bind the AgrA homologues SppR, PlnC and PlnD, respectively (Risoen \textit{et al}., 2000). SppR, PlnC and PlnD, are the response regulators of two-component systems regulating bacteriocin production in \textit{L. sake} and \textit{L. plantarum}, respectively. Like \textit{agr}, the \textit{ssp} and \textit{pln} systems respond to a peptide pheromone. Attempts to demonstrate a direct interaction of AgrA with the P2-P3 region were however unsuccessful (Morfeldt \textit{et al}., 1996) (\textit{Paper I}) (see below). One reason for this could be that AgrA needs to be in a phosphorylated form and that the experimental conditions used by
Morfeldt et al., (1996) did not provide this. We have later tried to phosphorylate AgrA *in vitro* by the use of acetyl phosphate, preliminary experiments have however failed to demonstrate binding of AgrA also in the presence of acetyl phosphate. Another reason could be that a cofactor is needed for the interaction that was not present, or not active in the extract used.

According to the model for *agr* autoregulation (Novick, *et al.*, 1995), the components of the *agr* signal transduction pathway need to be present at low levels to allow autocatalytic induction. It therefore seems likely that *agrBDCA* promoter, P2, has a low constitutive activity, either by intrinsic promoter activity or through regulatory factors ensuring low level of expression. The RNAlIII promoter, P3, on the other hand is completely silent when the *agr* system is not induced (Janzon, *et al.*, 1989).

In addition to the autoregulation of *agr*, sarA and the less well characterised 1E3 locus have been reported to affect transcription of *agr* and RNAlIII. In microaerophilic cultures, sarA seems to be an activator of RNAlIII (Lindsay and Foster, 1999), while in highly aerated cultures a *sarA* mutant and a wild type express the same levels of RNAlIII (*Paper III*). In contrast, inactivation of 1E3 resulted in an increased transcription of RNAlIII, indicating that the product(s) of the 1E3 locus has a repressive effect on P3.

To find out which sequences upstream of P3 were important for transcription, different parts of the promoter region were deleted on a plasmid RNAlIII construct (*Paper I*). The expression of RNAlIII from these constructs was analysed in the RNAlIII deletion mutant WA400 that contains an intact *agrBDCA* operon. A plasmid containing only 93 bp (two copies of the heptanucleotide repeat) upstream of the transcription start point expressed the same level of RNAlIII as a control containing the entire P2-P3 intergenic region. In contrast, no RNAlIII was expressed from a plasmid containing only 52 bp (one copy of the heptanucleotide repeat) of the upstream sequence. This showed that only the P3 proximal half, including two copies of the heptanucleotide repeat, was sufficient for full *agr* dependent transcription of RNAlIII. In addition, it could also be concluded that transcription from P3 can take place in the absence of simultaneous transcription from P2. As a consequence of these experiments, it was also assumed that the corresponding sequences in front of P2 are required for activation of *agrBDCA*. Analysis of the RNAlIII promoter revealed an unusually long spacer of 19 to 20 bp between the putative –10 and –35 promoter elements (Janzon, *et al.*, 1989), as compared to 16 to 17 bp in most of the known *S. aureus* promoters. By deleting 3 bp of the spacer from the RNAlIII plasmid construct containing 52 bp upstream of the transcription start point, the promoter was made active in both wild type and *agr* mutant background (*Paper I*). Thus by making the P3 promoter more according to the consensus,
upstream regulatory sequences were no longer required for expression. From this experiment it can be concluded that the wild type P3 promoter belongs to the class of promoters that are dependent of activators due to poor intrinsic activity.

In an attempt to demonstrate the presumed binding of AgrA to the P2 and P3 intergenic region, electrophoretic mobility shift experiments (EMSA) were performed with DNA fragments containing various parts of the P2-P3 regulatory regions (Paper I). Surprisingly, the mobility of the DNA fragments were specifically retarded by proteins extracted from both agr positive and agr negative cells, indicating that AgrA was not involved. Depending on the length of the DNA fragments used, ladderlike patterns with different number of discrete bands were seen. Further analysis revealed that the promoter regions of P2 and P3 competed for the same protein(s) in gel mobility shift experiments, suggesting that the two promoters independently could bind the protein(s). The shortest fragment to which binding was demonstrated was only 27 bp long.

By using magnetic beads to which a DNA fragment containing the P2 and P3 regulatory region had been coupled, we were able to absorb the DNA binding entity from crude extracts. After thorough washings with salt and unspecific DNA a single protein with an apparent molecular mass of 15 kDa was eluted from P2-P3 DNA coupled to magnetic beads. By N-terminal amino acid sequence analysis the protein was identified as SarA. Purified SarA protein generated an EMSA pattern identical to that obtained with the crude extracts, suggesting that the multiple bands seen in EMSA represents the interaction of SarA protein with the DNA fragments. This was further confirmed by the finding that the P2-P3 DNA fragment was not shifted in EMSA using protein extracts from a sarA mutant strain. The different bands formed in EMSA must therefore either represent DNA fragments with different number of SarA molecules attached or different conformations of the DNA-SarA complex.

As inactivation of sarA resulted in diminished production of agrBDCA and RNAIII (Heinrichs, et al., 1996) it was assumed that the demonstrated binding of the SarA protein to the P2-P3 promoter region activates transcription from these promoters. However, we have never been able to see this effect using the same wild type strain (8325-4) as Heinrich et al., (1996). On the other hand a difference in RNAIII expression between wild type and a corresponding sarA mutant was observed by Lindsay et al., (1999) when growing the bacteria under micro-aerophilic conditions. Apparently, in highly aerated cultures, SarA binding is dispensable for expression of RNAIII.
Taken together our results show that 93 bp but not 52 bp upstream of the transcription start point is sufficient for \textit{agr} regulated RNAIII expression. Within this region two heptanucleotide repeats, conserved within the staphylococcal genus, have been identified. Exactly the same repeats, at the same relative location, are found upstream of the P2 promoter. Though we have not been able to demonstrate binding of AgrA to these repeats, recent results of homologous systems in \textit{Lactobacillus} suggest that these repeats constitute the binding site for AgrA. We have also demonstrated that the SarA protein binds to several sites within the P2-P3 regulatory region. The reported stimulatory effect of \textit{sarA} on transcription of \textit{agrBDCA} and RNAIII can thus be explained by a direct interaction of the SarA protein with promoter DNA.

The binding of SarA to the \textit{agr} P2-P3 region was analysed in two separate reports (Chien and Cheung, 1998; Rechtin, \textit{et al}., 1999). By the use of a recombinant glutathione S-transferase (GST) -SarA fusion protein, Chien and Cheung (1998) determined the SarA binding site to a 29 bp region located between the two sets of heptanucleotide repeats described above. By deletion analysis of the SarA protein they also showed that with the exception of 15 N-terminal amino acids, the entire length of the protein seemed to be needed for DNA binding. Rechtin \textit{et al}., (1999) analysed the interaction of native SarA protein to the \textit{agr} P2-P3 region. In their report SarA was found to protect six defined areas within the regulatory region, each being 16 to 18 bp in length. It was further demonstrated that SarA binds DNA in the form of a dimer, and that one SarA dimer probably interacts with two protected areas (each protected area being a half-site) and that three SarA dimers would bind to the P2-P3 fragment used in the experiment. One binding site overlaps the –35 recognition sequence of P3 and another the –10 recognition sequence of P2. The third protected area, that was located roughly in the middle of the intergenic region, overlap partially the SarA binding site demonstrated by Chien and Cheung. The reason for the discrepancy is not clear. One explanation might be that Chien and Cheung used a fusion protein while Rechtin \textit{et al}., unmodified native SarA protein. As the GST-SarA fusion only generated one distinct DNA-protein complex in EMSA while a number of discrete complexes were observed when the same DNA fragment was shifted with purified native SarA (Rechtin, \textit{et al}., 1999) (\textit{Paper I}) it seems likely that the fusion with the ~15 kDa GST protein affects the DNA binding properties of SarA.
The role of RNAIII in regulation of transcription

RNAIII is a unique molecule in many aspects. In addition to being an mRNA, encoding the delta-hemolysin, it regulates mRNA levels of a vast number of potential virulence factors both in a positive and negative way. As discussed above, RNAIII also regulates translation of alpha-hemolysin. The expression of RNAIII is initiated during exponential phase of growth and reaches maximal levels during the postexponential phase. The amount of RNAIII in postexponential phase of growth is so high that it can be visualised by ethidium bromide staining of an agarose gel, which is uncommon for an mRNA. Usually only the ribosomal RNA molecules that constitute about 90% of the cells total RNA can be seen as discrete bands. The extremely long half-life of RNAIII (12 to 15 minutes, (Janzon, et al., 1989)) most likely explains the extreme abundance of RNAIII.

Expression of RNAIII results in repression of cell surface associated proteins while production of extracellular soluble toxins and enzymes are activated. How RNAIII exerts this reciprocal mode of regulation is not known. Considering the large amounts of RNAIII being produced it seems likely that several RNAIII molecules are needed for each of the regulated genes. One possibility would be that RNAIII interacts with the nascent target mRNA and thereby affecting productive transcription through termination or antitermination. RNAIII could also affect the stability of the target mRNAs by a direct interaction and thereby regulate the levels of mRNA. Another possibility would be that RNAIII regulates the translation of a transcriptional regulator that in turn either activates or repress the target genes. A forth possibility is that RNAIII affects the ability of the RNAP to initiate transcription through interactions with the promoter region, either directly or by affecting binding of regulatory proteins. Promoter-fusion studies have indicated that the agr responsive element is located in the promoter region. Novick et al., (1993) showed that a fusion of the hla promoter (-1371 to +1) to blaZ (beta lactamase) resulted in agr-dependent regulation of beta lactamase. With this fusion the same degree of agr regulation was seen as with the wild type allele. Patel et al., (1992) fused the eta (epidermolytic toxin A) promoter to a promoterless protein A gene (normally repressed by agr), resulting in agr dependent activation of protein A production. The increased amount of protein A was however only two-fold as compared to the amount of epidermolytic toxin A, which was increased 35-fold. The difference in level of regulation might be explained by degradation of protein A by proteases, which are upregulated in the wild type (Patel, et al., 1992). Another explanation would be that sequences downstream of the transcription start are needed for full agr dependent stimulation of eta. Similarly Sheehan et al., (1992) showed that a 220 bp DNA fragment containing the eta promoter (nt –215 to
+5) fused to the luciferase gene operon (luxAB) resulted in agr regulated luciferase activity. The increase of luciferase activity was 5-fold as compared to the 35-fold increase of epidermolytic toxin A in a wild type strain as compared to an agr mutant strain. Fusions of the spa, tst and hla promoter regions including several hundred basepairs downstream of the transcription start, to a reporter gene resulted in agr dependent regulation of the reporter to the same degree as the respective wild type alleles (Chan and Foster, 1998a). A recent publication by Zhang and Stewart (2000) demonstrated that 18 bp downstream of the transcription start point were needed for agr dependent regulation of the enterotoxin D gene (sed).

Several regulatory antisense RNAs regulate their target genes by affecting steady state concentration of a specific mRNA, either by facilitating degradation or by promoting stability. Björklind and Arvidson (1980) measured the half life of two agr regulated transcripts, encoding alpha-hemolysin (hla) and serine protease (ssp) in a wild type and in the corresponding isogenic agr mutant without noticing any difference in stability.

Thus, the promoter–fusion experiments together with the results obtained by Björklind and Arvidson suggest that RNAIII regulates mRNA levels by affecting productive transcription rather than affecting mRNA stability.

Several virulence genes regulated by RNAIII are also regulated by sarA. It seems likely that RNAIII regulates target gene transcription through interactions with protein factors rather than by direct interaction with the DNA helix. Some observations have indicated the presence of such factors. For example, activation of hla transcription was not observed until the onset of postexponential phase, regardless of what time point RNAIII was induced, suggesting the existence of a temporal signal, distinct from RNAIII (Vandenesch et al., 1991). In addition, Novick et al., (1993) found that incubation of RNAIII with crude S. aureus lysates decreased the sedimentation rate of RNAIII in a neutral sucrose gradient significantly as compared to the control, indicating interaction of RNAIII with proteins.

As protein factors might be involved in the RNAIII dependent regulation of target genes, proteins interacting with target gene promoters were searched for. Promoter DNA of two activated genes (hla and ssp) and one repressed gene (spa), was mixed with crude protein extracts of wild type S. aureus bacteria and analysed by electrophoretic mobility shift assay (EMSA). Interestingly all three promoter DNA fragments gave rise to almost identical EMSA patterns (Fig. 4.). The patterns formed were ladder like and looked very similar to those observed with purified SarA and the agr P2-P3 fragment (Paper I). The hla, ssp and spa
fragments were consequently tested for their ability to bind purified SarA. Indeed SarA was shown to bind all three fragments giving rise to identical patterns as those observed with the crude extracts (data not shown).

Figure 4. Electrophoretic Mobility Shift Assay of 90 bp DNA fragments containing promoter upstream regions (-25 to –115, relative the transcription start) of the alpha-hemolysin (hla), serine protease (ssp) and protein A (spa), gene, respectively. 5nM of 32P-labeled DNA was incubated with the indicated amounts of total protein extracts from the wild type strain 8325-4. The large arrow indicates migration of unbound DNA and small arrows indicate migration of protein-DNA complexes.

To investigate whether in vivo expression of hla, ssp and spa was affected by sarA and if such an effect would be influenced by agr, Northern blot analysis was performed with RNA from a wild type (8325-4), agr-mutant (RN6911), sarA-mutant (PC1839) and agr-sarA double mutant (WA513) strain. Using highly aerated cultures conditions total RNA was analysed in samples harvested hourly, at 1 to 7 hours of growth. As can be seen in Fig. 5. mRNA levels of both ssp (serine protease) and spa (protein A) were upregulated in the sarA mutant while mRNA levels of hla (alpha-hemolysin) were down regulated. Except for a one-hour delay in the appearance of RNAIII, sarA did not affect the final levels of RNAIII with the conditions employed. As the final levels of RNAIII were unaffected by sarA and since concordant results were obtained when sarA was inactivated in the wild type as well as in the
agr mutant, it was concluded that sarA activates hla and repressed ssp and spa by mechanisms that do not involve activation of agr.

![Northern blot analysis](image)

**Figure 5.** Northern blot analysis of RNAIII, alpha-hemolysin (hla), serine protease (ssp) and protein A (spa) transcripts in strains 8325-4 (wt), RN6911 (agr\(^-\)), PC1839 (sarA\(^-\)) and WA513 (agr\(^-\)sarA\(^-\)) at different time points during growth. 16S ribosomal RNA was used as an internal control of the amount of total RNA loaded. The same filter was hybridised with each of the specific probes.

During the course of this study several reports were published supporting our results that sarA has an agr-independent effect on agr regulated genes. Using promoter fusions of hla and tst (toxic shock syndrome toxin-1) to a reporter gene, Chan and Foster (1998b) demonstrated that both of these genes were transcriptionally up regulated by sarA, independently of agr. Further, fnbA (fibrinogen binding protein A) was shown to be activated by sarA, while ssp (serine protease), spa (protein A) and cna (collagen adhesin) were repressed (Cheung, *et al*., 1997; Gillaspy, *et al*., 1998; Wolz, *et al*., 2000).

Apparently many virulence genes are regulated by both RNAIII and SarA. Whether this involves direct physical interaction between the two regulators remain to be determined.
SarA binds DNA without apparent sequence specificity
Since the *sarA* locus has the ability to either stimulate or suppress transcription of several genes known to be regulated by *agr*, the molecular mechanism of SarA regulation was investigated. We analysed the interaction of the SarA protein with the promoter region of *hla*, a gene that appears to be stimulated by both *agr* and *sarA*, *ssp* which is stimulated by *agr* and repressed by *sarA* and *zntR* (regulator of zinc resistance operon) reported to be unaffected by SarA in an *in vitro* transcription assay ([Paper IV](#)).

DNA fragments encompassing nucleotides –25 to –115 relative the transcription start point of the *hla*, *ssp*, *spa* and *znt* genes, respectively, were shown to bind purified SarA protein in a very similar way. All four fragments generated similar ladder-like patterns as those obtained with the *agr* P2-P3 DNA probe when tested in EMSA. Specifically, each of the four 90 bp fragments generated four major SarA-DNA complexes with different electrophoretic mobility. No difference in the EMSA-pattern was observed between fragments of genes activated by *sarA* as compared to those repressed by *sarA*. The apparent dissociation constant (K_D) for the interaction of SarA with the *hla*, *znt* and *ssp* fragment was determined to be very similar (3 nM, 5 nM and 12 nM respectively). No conserved nucleotide sequence element that could define a SarA binding motif was found. However, all DNA fragments that bound SarA were highly AT-rich.

To further characterise the SarA binding, DNaseI footprinting analysis was performed with DNA encompassing the promoter regions of *hla* (activated) and *ssp* (repressed) ([Paper IV](#)). For both genes very similar footprints were observed. At low concentrations of SarA, a 40 to 50 bp region, overlapping the −35 consensus sequence of the *hla* promoter was protected. In the case of *ssp*, two regions (position −20 to −80 and −95 to −125) were protected at the same concentration of SarA. A two-fold increase of the SarA concentration (from 50 to 100 nM in the case of *hla* and 25 to 50 nM in the case of *ssp*) resulted in protection of the entire fragments used. For both *hla* and *ssp* identical protection patterns were seen with both the coding and non-coding strand. The regions protected in the footprinting analysis were part of the fragments shown to bind SarA in the EMSA experiments described above.

In a more detailed analysis of the *hla* promoter region using EMSA with promoter fragments of different lengths, it was confirmed that SarA has the ability to interact with an extended region of the *hla* locus and that a region of higher affinity overlaps the −35 consensus element. It was further demonstrated that the number of DNA-SarA complexes formed in EMSA could be directly related to the length of the DNA probe, indicating that the
individual bands represent DNA fragments with increasing numbers of SarA molecules bound. On average 15 to 20 nucleotides were required for binding of one SarA molecule. However, more than 30 bp seemed to be required for high affinity binding of SarA.

Chien et al., (1999) demonstrated a direct interaction of GST-SarA to DNA fragments encompassing promoter sequences of spa, hla, fnbA, fnbB and sec, respectively. Through DNaseI footprinting analysis, GST-SarA was shown to protect a stretch of 235 and 302 bp of the hla and spa promoter upstream regions, respectively. Within the SarA binding region of the hla locus, six smaller footprints were seen, each ranging from 9 to 28 bp in length. Based on DNaseI footprinting analysis and DNA sequence comparisons Rechtin et al., (1999) and Chien et al., (1999) have suggested two different SarA binding consensus sequences. Both consensus sequences are extremely AT-rich (92% and 100% respectively) and seem to allow a high degree of variation in most positions.

Our results together with the recent report by Chien et al., (1999) show that purified SarA binds similarly to the promoters of genes that are activated or repressed by sarA. Most of the reported SarA binding sites seems to overlap the −35 promoter element. A comparison of all sequences reported to bind SarA revealed no common consensus binding sequence. However, a high AT:GC ratio seemed to correlate with high affinity binding of SarA. Thus in some respect SarA appears to belong to the group of regulatory proteins that bind DNA with limited sequence specificity (e.g. H-NS and HU).

Effects of deletions in the hla and ssp promoter regions
In an attempt to define the regions of the promoter responsible for agr and sarA dependent regulation of hla and ssp, different promoter DNA fragments were fused to a promoter-less lacZ gene and the expression of the fusion analysed by Northern blotting probing for lacZ (Paper IV). To avoid possible side effects of increased promoter dosage all constructs were inserted as single copies in the chromosome. From analysis of the different fusions in wild type (8325-4), agr-mutant (RN6911), sarA-mutant (PC1839) and agr-sarA mutant background, respectively, it was concluded that 100 bp upstream of the transcription start point, but not 45 bp, were sufficient for agr-dependent regulation of ssp and hla. Similar results were reported by Mahmood and Kahn (1990) who found that 93 bp but not 59 bp upstream of the transcription start was sufficient for agr dependent regulation of seb (enterotoxin B). However in the case of hla, full agr responsiveness was only seen when sequences downstream of the transcription start point were included. Interestingly, the fusions
with downstream sequences were less expressed in an *agr* mutant than the fusion without
downstream sequences, suggesting that downstream sequences might bind a repressor. In an
*agr* positive background the fusion containing downstream sequences was twice as active as
the fusion without these sequences. Possibly RNAIII stimulate transcription of *hla* by
counteracting this putative repressor.

DNA sequences downstream of the *ssp* promoter also stimulated expression of the *ssp-
fusion transcript. However, this was only seen in the absence of the repressor SarA,
suggesting that sequences downstream of the transcription start point influence both the
intrinsic activity of the *ssp* promoter and its sensitivity to SarA repression. For both *ssp* and
*hla* it was only the longest promoter fusions (containing upstream and downstream sequences)
that responded to *sarA* to the same degree as the corresponding wild type alleles. Analysis of
a 48 bp promoter construct (-45 to +3) showed that the activities of the *hla* and *ssp* “core
promoters” were different. In the case of *ssp*, the 48 bp construct was highly expressed in wild
type background, while the corresponding *hla* construct showed almost no activity. A
requirement of DNA sequences upstream of –45 for expression of *hla* was also demonstrated
by Chien *et al.*, (1999) who found that deletion of 33 bp immediately upstream of the –35
promoter element (-36 to –69 relative the transcription start) abolished expression of *hla.*

Footprinting and EMSA experiments indicate that SarA has the ability to interact with
extended regions of the *hla* locus, both upstream and downstream of the transcription start
point. In DNaseI footprinting experiments a two-fold increase of the SarA concentration
determined whether SarA protected a 40 to 50 bp region or whether the entire fragment was
protected. To investigate whether increasing concentrations of SarA in growing bacteria had a
specific effect on transcription of *hla*, *sarA* was overexpressed by the use of the xylose
inducible promoter, p*xyLA*. When SarA was moderately overexpressed, transcription of both
RNAIII and *hla* appeared to be activated already during early exponential phase of growth as
compared to mid- as late-exponential phase, in the wild type. However, at higher
concentrations of SarA expression of both RNAIII and *hla* was repressed. Interestingly
bacterial growth was completely inhibited when SarA was expressed at even higher levels. It
thus seems like low concentrations of SarA could either stimulate, (*e.g. agr*, RNAIII and *hla*)
or repress (*e.g. spa* and *ssp*), while at higher concentrations SarA acts as a more general
repressor of gene expression. The repressive effect of SarA was recently demonstrated *in vitro*
with the *agr* P2 and P3 promoters, which are activated by *sarA in vivo* (Chakrabarti and
Misra, 2000).
Analysis of proteins binding to agr-regulated promoters
As described above SarA appears to bind in a very similar way to the promoter regions of both positively and negatively regulated genes. To search for additional regulators that could explain the differential effects of SarA, proteins with affinity for the promoter regions of the RNAIII gene, hla and ssp, respectively, were isolated (Paper III). By the use of magnetic beads, to which promoter DNA was attached, several proteins were recovered from total protein extracts of both wild type and sarA mutant S. aureus bacteria. The pattern of proteins recovered by this method showed both similarities and differences between the different promoters and between wild type, sarA and agr mutant extracts, indicating some degree of specificity in the system and the presence of proteins with both general and specific DNA binding characteristics. As expected, SarA was recovered from wild type and agr mutant extracts, but not from sarA mutant extracts with all three promoter fragments tested.

One protein with an apparent molecular mass of 29 kDa bound to all three promoters when incubated with extracts from sarA mutant and agr mutant but not from the wild type. The N-terminal amino acid sequence of the protein was identical to the first ten amino acids of a 250 residues open reading frame present in both the TIGR S. aureus COL and University of Oklahoma S. aureus 8325-4 unfinished genome databases. A BLAST search with the predicted amino acid sequence revealed that the 29.9 kDa protein consisted of two homologous halves, each with a high degree of identity to SarA (32 and 34% respectively). Like SarA, the 29.9 kDa protein was basic (calculated pI = 9.6), typical of DNA binding proteins. Because of its similarity to SarA and its DNA-binding ability, the protein was named SarH1 for Sar Homologue 1. Due to the location of sarH1 immediately upstream of the protein A gene (spa), (see below), which could implicate a functional coupling, the binding of SarH1 to spa promoter DNA was also analysed by the magnetic beads technique. Binding of SarH1 to the spa promoter fragment was indeed observed with both agr and sarA mutant extracts.

In addition to SarA and SarH1, a protein of almost identical electrophoretic mobility as SarA was found to bind the hla promoter fragment when wild type extracts were used. The N-terminal amino acid sequence matched the N-terminal of a 115 amino acid orf in the TIGR S. aureus 8325-4 unfinished genome database. BLAST searches with the predicted protein showed a high degree of identity to SarA (30%) throughout the protein. Like SarA and SarH1 the 115 amino acid protein was highly basic (calculated pI = 10.1). The 115 amino acid protein was named SarH5 for Sar Homologue 5. The SarH5 protein was not recovered in extracts from neither the sarA not the agr mutant strain. One possible explanation for this
might be that the increased levels of SarH1 in these mutants repress expression of SarH5 (Tegmark et al., unpublished). Interestingly SarH5, which is identical to SarR (Sar Repressor) described by A.L. Cheung (Personal communication), has been shown to repress transcription of sarA from the P1 and P2 promoters, respectively. Moreover SarH5/SarR has been demonstrated to bind promoter DNA of sarA P1 and P2. Like for SarA and SarH1 (see below), binding of SarH5/SarR to promoter DNA generated a ladder-like pattern (A.L., Personal communication).

The binding of several additional proteins to the promoter regions of RNAIII gene, hla, ssp and spa suggests that the activity of each promoter might be influenced by a number of regulatory proteins. For example, a protein with an apparent molecular size of 28 kDa present in both wild type and sarA mutant extracts bound the ssp, but not the agr and hla promoter. A different protein, with an apparent molecular size of 26 kDa, only present in wild type extracts was recovered from the hla and ssp but not the agr promoter. A unique protein of 27 kDa was recovered from the agr promoter using wild type but not sarA and agr mutant extracts, possibly AgrA. The identity and the role of the 28, 26 and 27 kDa proteins have however not yet been determined.

Characterisation of sarH1

The sarH1 locus was found to be located between the staphylococcal iron regulated genes A, B and C (sirABC) and the protein A gene (spa) (Paper III). By Northern blot analysis three sarH1 specific transcripts were identified. Primer extension analysis determined the transcription start point for two of these transcripts. Like sarA, sarH1 was transcribed from both SigA- and SigB-dependent promoters suggesting that expression of sarH1 might be affected by various stresses that influence SigB activity. The SigA-dependent P1 promoter was located 150 nt upstream, and the SigB-dependent, P2 promoter, 800 nt upstream of the SarH1 translation start point.

Northern blot analysis of sarH1 specific mRNA in agr and sarA mutant background revealed that the sarH1 P1 promoter was under strong repression of both agr and sarA. The increased transcription of sarH1 in agr and sarA mutants most likely explain why the SarH1 protein was recovered from agr and sarA mutant, but not wild type extracts in the promoter affinity experiments described above.

To analyse the possible role of sarH1 in the regulation of virulence genes, a sarH1-knockout mutant was created. The effect of the sarH1 mutation was analysed in wild type,
agr-, sarA-, and agr-sarA- double mutant backgrounds, respectively. Different patterns of extracellular proteins were found to be produced in the different backgrounds. When sarH1 was inactivated in the wild type or in the sarA mutant an increasing number of extracellular proteins were found while sarH1 inactivation in the agr mutant backgrounds resulted in a decreased production of extracellular proteins. Apparently SarH1 stimulate or suppress target gene expression depending of the activity of agr and/or sarA. The regulatory role of sarH1 was further studied by Northern blot analysis. Although SarH1 bound in vitro to the promoter regions of RNAIII, hla, ssp and spa, a significant effect of the sarH1 mutation on mRNA levels was only observed with hla and spa. Transcription of spa was found to be stimulated by sarH1 and it was shown that the increased transcription of spa in the agr and sarA mutants was predominantly the result of increased production of SarH1, contrary to the previous suggestion that RNAIII and SarA directly repress spa transcription (Cheung, et al., 1997; Recsei, et al., 1986). However, both RNAIII and SarA also seemed to have a slight suppressive effect on spa transcription independent of sarH1. Transcription of hla was shown to be repressed by sarH1. Interestingly, the suppression of hla was only seen in the absence of SarA and in the presence of RNAIII. It has recently been suggested that a product of rot represses hla transcription in the absence of RNAIII (McNamara, et al., 2000). The reason why sarH1 had no effect on hla expression in an agr mutant background is probably because hla is already completely repressed by Rot.

In vitro experiments have shown that purified SarA binds to DNA predominately as a homodimer (Blevins, et al., 1999). Considering that SarH1 consists of two halves, each with a high degree of identity to SarA, it is reasonable to assume that SarH1 can fold into a structure similar to that of a SarA dimer. It therefore seems likely that SarA and SarH1 could interact with the same DNA element in a competitive way. Indeed, purified SarH1 protein bound to the same hla, ssp and znt promoter DNA fragments as SarA (Fig. 6.) (data not shown). The ladder-like pattern indicates the binding of several SarH1 molecules to each of the DNA fragments. The EMSA pattern with SarH1 was similar to that observed with SarA. However the 90 bp promoter DNA fragments of hla, ssp and znt gave four discrete bands with SarA, while the same fragments incubated with SarH1 generated eight discrete bands. The Kd value for SarH1 binding to the hla, ssp and znt fragments was roughly 100 times higher than those observed for SarA (Tegmark et al., unpublished). Overexpression of sarH1, by use of the xylose inducible promoter, Pxy, abolished both hemolysin and protease production as determined by blood and casein agar plates, respectively, (data not shown). Like SarA,
overexpression of SarH1 also had a negative effect on bacterial growth, though it was not completely inhibited.

Figure 6. Electrophoretic Mobility Shift Assay of a 90 bp DNA fragment containing promoter upstream regions (-25 to –115, relative the transcription start) of the zinc resistance operon (znt). 5nM of 32P-labeled DNA was incubated with the indicated amounts of purified SarH1 protein (nM). The large arrow indicates migration of unbound DNA and small arrows indicate migration of protein-DNA complexes. Identical protein-DNA complexes were obtained using DNA fragments containing the corresponding region of the alpha-hemolysin (hla) and serine protease (ssp) gene.

Additional Sar homologues
In addition to SarH1 and SarH5, identified by their ability to bind promoter DNA, three more Sar homologues were found in the TIGR and University of Oklahoma S. aureus genome databases, (Paper III) (Table 4.). Two of the homologues, SarH2 and SarH4, had predicted molecular masses of 14 kDa and showed 35% and 26% identity to SarA, respectively. The third homologue, SarH3, had the same predicted molecular mass as SarH1 and a similar primary structure, with two homologous halves each with a high degree of identity to SarA (33% and 30%, respectively). SarH2 and SarH3 are located close to fnbB (fibronectin binding protein B) and a gene coding for a putative cell wall-associated protein, similar to pls (S. aureus surface protein Pls), suggesting that they may also be involved in the regulation of virulence genes. If and when sarH2, sarH3 and sarH4 are expressed remains to be
determined. Preliminary results from our lab indicate that all homologues are present in total protein extracts prepared from bacteria harvested in post exponential phase of growth. Two dimensional gel electrophoresis of proteins recovered from the *hla* promoter region by the magnetic bead experiments, show that four very basic proteins in the 13 to 14 kDa range and two in the 30 kD range interact with the *hla* promoter. Of the 13 to 14 kDa proteins, the most acidic protein has been identified to be SarA. Of the remaining three, two have the same apparent molecular mass as SarA, one is however slightly more basic than the other (most likely SarH2 and Sar H4), the third which has an apparent molecular mass of 13 kDa and a pI of 9.5, has been identified to be SarH5. Of the two proteins in the 30 kDa size rage, one has been identified as SarH1. The other which is as basic as SarH1 has a slightly lower apparent molecular mass and is most likely SarH3.

**Table 4.** Sar homologues

<table>
<thead>
<tr>
<th></th>
<th>Length</th>
<th>Molecular mass</th>
<th>Isoelectric point</th>
<th>Percent identity</th>
</tr>
</thead>
<tbody>
<tr>
<td>SarA</td>
<td>124 a.a.</td>
<td>14.7 kDa</td>
<td>8.5</td>
<td>100%</td>
</tr>
<tr>
<td>SarH1</td>
<td>250 a.a.</td>
<td>29.9 kDa</td>
<td>9.6</td>
<td>32%/34%e</td>
</tr>
<tr>
<td>SarH2</td>
<td>118 a.a.</td>
<td>14.0 kDa</td>
<td>10.3</td>
<td>35%</td>
</tr>
<tr>
<td>SarH3</td>
<td>247 a.a.</td>
<td>29.7 kDa</td>
<td>10.7</td>
<td>30%/33%e</td>
</tr>
<tr>
<td>SarH4</td>
<td>116 a.a.</td>
<td>14.0 kDa</td>
<td>9.9</td>
<td>26%</td>
</tr>
<tr>
<td>SarH5</td>
<td>115 a.a.</td>
<td>13.7 kDa</td>
<td>10.1</td>
<td>30%</td>
</tr>
<tr>
<td>Rot</td>
<td>133 a.a.</td>
<td>15.7 kDa</td>
<td>5.0</td>
<td>26%</td>
</tr>
</tbody>
</table>

|a| calculated molecular mass and isoelectric point, based on the amino acid sequence, respectively;  
|b| Identity to SarA;  
|c| N-terminal and C-terminal half of the protein, respectively. |

Comparison of SarA and the Sar homologues reveal a highly conserved motif located in the C-terminal half of the proteins. The motif, which is also found in Rot and partly in the Mar family of regulators, (a transcriptional regulator in several Gram negative species), consists of a series of equally spaced charges amino acid residues in the C-terminal half of the respective proteins (Fig. 7.). The presence of a conserved motif within the Sar family of proteins as well as in Rot, suggests a common function.
RNAIII is present in several coagulase-negative species

Since production of delta-like haemolytic activity has been demonstrated in several coagulase-negative staphylococcal species we asked the question whether delta-hemolysins from *S. epidermidis* and other CoNS are encoded by RNAIII-like mRNAs which also have a regulatory function (Paper II). In three out of seven tested coagulase-negative staphylococcal species (*S. cohnii, S. epidermidis, S. haemolyticus, S. saprophyticus, S. simulans, S. warneri, and S. xylosus*) RNAIII homologues were found. In *S. epidermidis*, *S. simulans* and *S. warneri* RNAIII molecules of 560, 573 and 684 nucleotides, respectively, were identified. All three RNAIII genes contained open reading frames with high degrees of sequence identity to the *S. aureus* delta-hemolysin. Like in *S. aureus* RNAIII the predicted delta-hemolysin genes were located in the 5’ half of the molecule but at different distances from the transcription start point. Surprisingly, *S. warneri* RNAIII was predicted to encode two nonidentical copies of delta-like proteins, both 25 amino acids in length. Compared to *S. aureus* RNAIII, the gene in *S. simulans* had a 107 bp insertion upstream of the delta-hemolysin open reading frame and a deletion in the 3’ part of the gene. The RNAIII gene in *S. epidermidis* and *S. warneri* both had small insertions of 25 bp upstream of their respective delta-hemolysin open reading frames. In addition, the gene in *S. warneri* had a 130 bp insertion, containing the second delta-hemolysin determinant. Except for these differences the RNAIII molecules from the CoNS species turned out to be very similar to that of *S. aureus*, especially the 3’ halves of the molecules were highly conserved. Moreover, predicted computer analysis of the secondary structure revealed several conserved stem-loop structures of comparable energy throughout the molecule. The inability to detect an RNAIII gene in *S. cohnii, S. xylosus, S. haemolyticus,* and

<table>
<thead>
<tr>
<th>Protein</th>
<th>Sequence</th>
<th>Amino Acid Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>SarA</td>
<td>VKAVKILSQEDYFDKKRNEHDERTVLILVNAQQR</td>
<td>a.a. 68 to 102</td>
</tr>
<tr>
<td>SarH1-N</td>
<td>VQHIKLVKHYSISKVRSDKERNTYISISEEQQR</td>
<td>a.a. 67 to 101</td>
</tr>
<tr>
<td>SarH1-C</td>
<td>VRALNLLKKQGGLKHERSTEDKILIHMDAQKQ</td>
<td>a.a. 191 to 225</td>
</tr>
<tr>
<td>SarH2</td>
<td>VKSVKDLSSKGFLNKRNEADERRIFVSVTIPQR</td>
<td>a.a. 70 to 104</td>
</tr>
<tr>
<td>SarH3-N</td>
<td>VKNIKSLSKGMLIKERSLADERIVLIKINKIQY</td>
<td>a.a. 68 to 102</td>
</tr>
<tr>
<td>SarH3-C</td>
<td>VRSVNRLSSKGYLNKRDPHDSSRNVIIVSVKQH</td>
<td>a.a. 193 to 227</td>
</tr>
<tr>
<td>SarH4</td>
<td>DVIRIKIYKKELSKLRS重大ERQVFYSTSTQK</td>
<td>a.a. 66 to 100</td>
</tr>
<tr>
<td>SarH5</td>
<td>TKAQLQKLDKLKLSSKRS重大雀LQDERIVVYYVCTIQK</td>
<td>a.a. 69 to 103</td>
</tr>
<tr>
<td>Rot</td>
<td>TRTYNNLVELEWIYKERPVDDERTVIHFNENKLQ</td>
<td>a.a. 66 to 100</td>
</tr>
</tbody>
</table>

**Figure 7.** Alignment of parts of SarA, SarH1-N terminal half, SarH1-C terminal half, SarH2, SarH3-N terminal half, sarH3-C terminal half, SarH4, SarH5 and Rot, demonstrating the presence of a highly conserved motif. Conserved amino acid residues in bold. Co-ordinates of the protein shown to the right.
S. saprophyticus by the PCR based method used in paper II does not unambiguously exclude the possibility that these species lack the gene. However, similar results were obtained by Donvito et al., (1997) in a screen for RNAIII and delta-hemolysin in several CoNS species by PCR and southern blotting. In addition Vandenesch et al., (1993) have identified RNAIII in S. lugdunensis. In contrast to the RNAIII molecules analysed by us, the S. lugdunensis RNAIII did not contain an open reading frame encoding delta-hemolysin. However, the remaining parts of S. lugdunensis RNAIII are very similar to the RNAIII molecules identified in paper II, suggesting a common function.

Identification of functional domains of RNAIII
To test the ability of the CoNS RNAIII molecules to regulate target gene expression in S. aureus plasmids containing the S. epidermidis, S. simulans and S. warneri RNAIII genes were introduced into an RNAIII deficient S. aureus strain (Paper II). The levels of alpha-hemolysin (hla), serin protease (ssp) and protein A (spa) mRNA were analysed by Northern blot analysis. All three CoNS RNAIII molecules stimulated the expression of hla and ssp. However, the levels of transcripts were generally reduced compared to those obtained with the S. aureus RNAIII molecule. Interestingly, transcription of spa was completely repressed with all RNAIII molecules. The finding that the CoNS RNAIII repressed transcription of spa as efficient as S. aureus RNAIII, while stimulation of hla and ssp transcription was impaired, suggest that these regulatory functions are independent. To test which part of RNAIII that is most important for the regulatory function, hybrid molecules of RNAIII from S. aureus and S. epidermidis were created and tested for their ability to regulate hla, ssp and spa. The results from the fusion experiments showed that when either the 5´ or 3´ part of RNAIII of S. epidermidis was exchanges for that of S. aureus an increased ability to stimulate hla and ssp mRNA levels were obtained. As expected both fusions suppress spa mRNA levels completely. These experiments indicated that both the 5´ and 3´ halves of RNAIII are important in the regulation of hla and ssp.

Similar results were obtained by Benito et al., (1998) who found that RNAIII from S. lugdunensis only partially restored hla, hlb, geh, nuc and ssp mRNA levels in a RNAIII deficient S. aureus strain. S. lugdunensis RNAIII did however not repress the levels of spa mRNA. A fusion of the S. lugdunensis RNAIII 5´ half to that of the 3´ half of S. aureus resulted in a molecule that had an improved ability to stimulate hla, hlb, geh and ssp mRNA levels as compared to the S. lugdunensis RNAIII molecule, but not to the same levels as those
obtained with RNAIII from *S. aureus*. Interestingly, the *S. lugdunensis- S. aureus* RNAIII hybrid completely repressed the expression of *spa* mRNA suggesting that the 3’-half of RNAIII is responsible for repression of *spa* mRNA levels. In a recent report from the same laboratory it was indeed showed that the 3’ end of the RNAIII molecule alone, was sufficient to suppress *spa* mRNA levels (Benito *et al*., 2000).

When comparing the sequences of the CoNS RNAIII molecule to the proposed structure of RNAIII (Benito, *et al*., 2000) certain stemloops were found to be conserved both in primary and secondary structure while others were only conserved in the secondary structure and yet others not conserved at all (Fig. 8.). Most likely these conserved hairpin structures represent functional domains of the RNAIII molecule.

*Figure 8.* Secondary structure model of RNAIII showing regions with conserved primary structure (shaded areas). Hairpin structures are numbered 1 to 14; long distance interactions are designated A, B and C; the Shine-Dalgarno sequence and the start codon of the delta-hemolysin gene are in bold face. The picture is adapted from Benito *et al*., 2000.
GENERAL CONCLUSIONS AND A HYPOTHETICAL REGULATORY MODEL

Clearly, regulation of virulence genes in *S. aureus* is very complex and involves several regulatory factors that appear to interact to determine the expression of specific target genes. In addition to the global regulators identified so far, additional uncharacterised regulators are most likely involved in the regulation virulence genes. With the limited knowledge we have today about the function of the known regulators, (*e.g.* agr, sarA, sarH1, rot and sae), together with the assumption that additional, unknown regulators exists, any model for how *S. aureus* virulence genes are regulated must be highly speculative. Based on the following general conclusions and assumptions a hypothetical model for the function of the Sar proteins and RNAIII is however proposed.

SarA and SarH1 belong to a family of six highly basic DNA binding proteins with a conserved primary sequence motif, suggesting a common function. Inactivation of either sarA or sarH1 results in pleiotropic phenotypes with activation of certain genes and repression of others. Studies of the interaction of SarA with DNA *in vitro*, showed that SarA interacts with the promoters of both positively and negatively regulated genes. Most of the reported binding sites for SarA overlap the –35 promoter element, typical for transcriptional repressors. This is consistent with the observation that SarA functions solely as a repressor in an *in vitro* transcription assay (Chakrabarti and Misra, 2000). The same has also been shown for SarH1 (Tegmark, *et al.*, unpublished). It thus seems that SarA and SarH1 are basically repressors. There is also strong evidence that SarH5 (SarR) is a repressor (A. L. Cheung, Personal communication). In addition to being strong repressors of certain virulence genes (*e.g.* repression of ssp and cna by SarA) some of the Sar homologues also repress other Sar homologues (*e.g.* repression of SarH1 by SarA, and SarA by SarH5 (SarR)). This means that inactivation of one sar homologue can lead to decreased expression of a specific virulence gene due to derepression of another sar homologue that is the actual repressor of the target virulence gene. For example, the decreased transcription of *hla* in a sarA mutant was due to increased expression of SarH1, which repressed *hla* (*Paper III*). Similar to SarA, that appears to directly repress ssp, cna and sarH1, it seems likely that at least some of the other Sar homologues also have several target genes. Considering that there are six Sar homologues which might interact in a similar way, very complex regulatory networks can be envisioned (Fig. 9.). Depending on the number of Sar homologues involved in a regulatory circuit a specific target virulence gene will either be activated or repressed upon inactivation of a
certain *sar* gene. This means that it is the relative concentrations of the different Sar homologues that determine the level of expression of virulence genes.

![Figure 9](image)

**Figure 9.** Hypothetical model of Sar proteins functioning as repressors

Interestingly, most of the genes regulated by SarA and its homologues are also regulated by RNAIII. The RNAIII responsive element of target genes seems to be located to the promoter, suggesting a close interaction between RNAIII and the Sar homologues. One possible function for RNAIII could be that of an anti-repressor. Preliminary experiments in our laboratory have shown that RNAIII binds SarA in a mobility shift assay. By binding Sar proteins, RNAIII would modulate the free pool of repressors, thereby either activating or repressing target genes. RNAIII might have different affinities for different Sar proteins. It has been shown that different parts of the RNAIII molecule are involved in the regulation of different target genes (Benito, *et al.*, 2000; Benito, *et al.*, 1998) (*PaperII*), possibly because specific stem-loop structures interact with different Sar proteins.

According to this model the low expression of *ssp* during early exponential phase of growth would be due to high expression of SarA and to the lack of anti-repressor, RNAIII. The increased *ssp* transcription, following the increased RNAIII levels at late exponential phase of growth, would be the result of out-titration of SarA by RNAIII. Consequently the decreased *ssp* transcription in an RNAIII mutant would be due to the lack of anti-repressor rather than to lack of activator. This is consistent with the fact that inactivation of *sarA* in an RNAIII mutant resulted in the same level of *ssp* transcript as in the wild type.

In the case of *hla*, both SarH1 and Rot appears to be repressors (see above). In the wild type the amounts of SarH1 and Rot are relatively low while the concentration of RNAIII is high, therefore allowing high level expression of *hla*. In a *sarA* mutant where *sarH1*
expression is derepressed, the relative concentration of SarH1 is high enough to repress \textit{hla}, as indicated by the high level \textit{hla} expression in a \textit{sarA sarH1} double mutant as compared to low levels in the \textit{sarA} mutant (\textbf{Paper III}). The higher level of \textit{hla} expression in the \textit{sarA sarH1} double mutant as compared to the \textit{sarA, sarH1, RNAIII} triple mutant could be explained by RNAIII also being an anti-repressor of Rot, as suggested (McNamara, \textit{et al.}, 2000). (Fig. 10.)

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{schematic.png}
\caption{Schematic figure of the hypothetical model of transcriptional regulation of the alpha-hemolysin gene (\textit{hla}).
A. Wild type background with the presence of SaeR, high levels of RNAIII, relative low levels of SarH1 and Rot allowing high level of transcription.
B. \textit{agr}-mutant background with the presence of SaeR, absence of RNAIII, relative high levels of SarH1 and Rot rendering the \textit{hla} promoter inaccessible for the RNA-polymerase and consequently low levels of transcription.
C. \textit{agr sarA} double mutant background with the presence of SaeR, absence of RNAIII and SarH1 and relative high levels of Rot rendering the \textit{hla} promoter inaccessible for the RNA-polymerase and consequently low levels of transcription.}
\end{figure}
This is a hypothetical model that needs to be tested by further experiments demonstrating the specific binding between RNAIII and each of the Sar homologues as well as the interaction between the Sar homologues and their target virulence genes. In addition to further <i>in vitro</i> experiments studies of the expression of these regulators <i>in vivo</i> will be of critical importance for the understanding of how these regulatory networks function and for their role in <i>S. aureus</i> pathogenesis.
ACKNOWLEDGEMENT

I thank all those who have supported and helped me during the work of this thesis, especially I want to thank:

Staffan Arvidson, my supervisor, for his never ending enthusiasm for the work on the regulatory networks of \textit{Staphylococcus aureus}. For his generosity and patience in teaching me bacteriology and bacterial genetics. For believing in me, especially during the early years when I was an unexperienced medical student. For having confidence in me and allowing me to persevere any idea I might have had, but still guiding me in doing the right experiments. For freedom and understanding during the hard time of my sisters disease and death.

Eva Morfeldt, my co-worker and mentor, for enormous support and positivism. For key contributions to this thesis. For being so generous in sharing time and expertise. For teaching me all those practical and theoretical things. For being one of my best friends during both sorrow and joy.

Agnete Wahlquist for help with everything from the first DNA-prep to the last gel. For saving me all those times when I had tons of experiments running, no buffers and only two hands. For keeping order and a nice atmosphere in the lab and for all those biscuits that I have stolen late at night.

Caterina Guzman-Verri, my dear friend on the other side of the Atlantic, for being such a good role model with your strong character, ambition and solid knowledge in the field of bacterial pathogenesis. For being supportive though being so far away.

Patricia Sarravia-Otten, my friend and former group mate, for all the laughs, your glowing temperament and for a shared interest in \textit{sar} and \textit{agr}.

Anna Karlsson, my co-worker and friend. For contributions to this work and for fun and enjoyable times in and out of the lab. Lena Norenius, for all the practical help in the lab and for keeping order of my strains. For bringing coffeepaeks, delicious biscuits and art to the group.

Fredrik Åslund, my recent group mate, for shared interest in bacterial genetics, new ideas, good advice and critical reading of manuscripts.

Mark Clements for withstanding my never ending staph-talk though having converted to salmonella, for critical reading of manuscripts, correction of my english and for being a nice friend.

Roland Möllby and the members of his group, my former room mates, Patricia Colque-Navarro, Lena Guldevall, Beatrice Vecsey-Semjen, Aina Iversen, Inger Kühn, Jenny Gabrielson, Mokhlasur Rahmann, Anette Frost, Maj Ringman and Mats Söderhäll for always being helpful and good friends.

Staffan Normark and the present and former members of his group, Anders Folkesson, Anna-Karin Pernestig, Agneta Richter Dahlflors, Christina Nilsson, Christine Jacobs, Fredrik Bäckhed, Hannes Loeflerer, Jeanette Çelik, Katrin Pütsep, Margareta Krabbe, Margret Wahlström, Mathias Hornef, Mårten Hammar, Tomas Söderblom, Vjole Dzikaite, Ute Roemling, Zhao Bian and Öjjar Melefors for scientific discussions and allowing me to participate in your groupmeetings as well as for creating a nice scientific and social atmosphere. A special gratitude to: Anders for interest in my work, slaving with me at Cold Spring Harbor and for friendship. Anna-Karin for the energy in everything you do, for badminton and two-component systems and for sharing fun and sad moments. Katrin for being a nice friend and for good advice in everything from gardening and home decorating to thesis writing and how to overcome the obstacles of a thesis application.

Monica Thelestam and the members of her group, my present room mates, Alberto Alape, Esteban Chaves-Olarte, Juan-Carlos Higueta, Marietta Flores-Diaz, Teresa Frisan and Ximena Cortes-Bratti for fun times and nice chats.

Mikael Rehn and the members of his group, Anna Syk, Dilek Tezcan and Sofia Eriksson. To Reini Hurme for friendship and encouragement and for your "PhD-student-integrity-survival-advice" without which I would have been lost so many more times.

The SMI-ers, Ann-Beth Jonsson, Barbara Albigier, Birgitta Henriques Normark, Dan Andersson, Farzat Olfat, Helena Källström, Johanna Björkman, Paola Hansson-Palo, Sture Karlsson and Tomas Åkerlund for scientific exchange and many more things.

To Mats Wahlgren and the members of his group, Andreas Heddini, Antonio Barragan, Carin Scholander, Carl-Johan Treutiger, Chen Qijun, Gabriel Guzman, Johan Lind, Karima Mashoo, Martha Schlichtherle, Silvia Botero and Victor Fernandez for your friendliness and positivism.

Linda Engberg and Carolina Tapper for help with economical and administrative tasks.

Gregor Blomkvist and Alex Feldetoe for help with computer problems.

Birgitta Karlsson for all the plates, buffers and media that I have used.

My brothers Sven and Mats and their families, my sister Maria and my brother-in-law, Kalle, for all support and for being the best brothers and sisters there are.

My parents, Barbro and Erik, for always believing in me, for interest and encouragement in my studies.

My beloved Mats, for your loyalty and invaluable support. For taking care of me after failed experiments and making me realise that there are other things in life than work. Thank you for always being there for me.
REFERENCES


