2008

Identification of Virulence Determinants for Streptococcus sanguinis Infective Endocarditis

Lauren Turner
Virginia Commonwealth University

IDENTIFICATION OF VIRULENCE DETERMINANTS FOR *STREPTOCOCCUS SANGUINIS* INFECTIVE ENDOCARDITIS

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy at Virginia Commonwealth University.

by

LAUREN SENTY TURNER
B.S., Virginia Tech, 2002

Director: TODD KITTEN, PH.D.
ASSOCIATE PROFESSOR, PHILIPS INSTITUTE, SCHOOL OF DENTISTRY

Virginia Commonwealth University
Richmond, Virginia
July, 2008

Acknowledgement

Dr. Todd Kitten has been an excellent mentor to me. I especially appreciate his commitment to his students in reviewing and revising their written documents, discussion of data, and identifying opportunities for their individual professional growth. I am also grateful to the members of my graduate committee, Dr. Cornelissen, Dr. Munro, Dr. Peterson, and Dr. Tew for their advice and direction regarding my research and graduate education. Dr. Munro has also been an integral contributor to many of the experiments discussed in this dissertation. I am appreciative of her assistance, as well as Dr. Takeshi Unoki's and Dr. Taisei Kanamoto's collaboration in animal experiments. I am also thankful to past and present members of Dr. Todd Kitten's lab, Jill Callahan, Dr. Sankar Das, Dr. Taisei Kanamoto, Brian Mahoney, Da' Shekia Mitchell, Dr. Shemi Paik Lee, Lt. Col. Jody Noe, Alejandro Rodriguez, and Nicai Zollar for their friendship over the years. I am particularly grateful to Nicai Zollar for her timely acquisition of reagents and materials required for experiments, and her assistance with animal experiments. Additional faculty, staff and students of the Philips Institute deserve recognition for the success of our research group. I am especially thankful to Cecilia Anaya, Kevin Jones and Dr. Hiro Miyasaki for their scientific input and thought-provoking conversations. Also, the assistance and friendship of Kim Holloway, Margaret Polland and Jennifer Kunc is much appreciated.

The support of my family and friends has been unequivocal in the past few years, as in every other part of my life. Foremost I must acknowledge my parents, Laura and David Senty and maternal and paternal grandparents, Peggy Baird, W. Calvin Poston, Herb Baird, and Alice and Robert Senty for their steadfast encouragement and support in my pursuit of learning. Both my parents and their parents have provided inspiring examples of dedication in their service to our country and its citizens, and in nurturing of their children. I also want to thank members of my extended family, as well as Betty Lou Turner, and Linda and Mark Turner for their encouragement and loving support in the pursuit of my doctorate. Luke, my husband and I have made many friends while in Richmond that will keep this place dear to our hearts for years to come. I am especially grateful to Katelyn Sevin, Chantal Ayres, Geoffrey Waugh, Chris Sander, Mike Smith, Lawton Grinter and Emily Tran for the balance they have set in my recent years. I also want to thank my dearest friend of over twenty years, Elaine Vidal for her ceaseless enthusiasm in our friendship. My husband, Luke Turner has been a catalyst for all my accomplishments at VCU. I especially admire in Luke his infectious ambition in pursuit of his goals, his encouragement of mine, and will forever be thankful for his dedication to our marriage.

Table of Contents

List of Tables

Page

List of Figures

Page

List of Abbreviations

%	percentage
°C	degree Celsius
µg	microgram
µl	microliter
µM	micromolar
x *g*	times gravity
aa	amino acid
ABC	ATP-binding cassette
ADP	adenosine diphosphate
ANOVA	analysis of variance
AP	alkaline phosphatase
ARNR	anaerobic ribonucleotide reductase
ATP	adenosine triphosphate
BHI	brain heart infusion broth
BHIT	brain heart infusion broth with threonine
bp	base pair
BM	biofilm medium
BSA	bovine serum albumin
Ca^{2+}	calcium
CFU	colony-forming unit
$CHCl_3$	chloroform
CI	competitive index
cm	centimeter
Cm	chloramphenicol
CSP	competence stimulating peptide
C-terminus	carboxyl terminus
CV	crystal violet
dH_2O	deionized water
DIG	digoxigenin
DNA	deoxyribonucleic acid
DNase	deoxyribonuclease
dNTP	deoxynucleotide triphosphate
E. coli	*Escherichia coli*
EDTA	ethylenediaminetetraacetic acid
Em	erythromycin
EtOH	ethanol

fmol	femtomole
g	gram
gDNA	genomic DNA (chromosomal DNA)
glu	glucose
HGT	horizontal gene transfer
HS	horse serum
IE	infective endocarditis
IgA	immunoglobulin A
IgG	immunoglobulin G
in.	inch
kb	kilobase
kDa	kiloDalton
kg	kilogram
Kn	kanamycin
LA	Luria-Bertani agar
LB	Luria-Bertani broth
LC	liquid chromatography
LMP	low melting point
M	molar
mAb	monoclonal antibody
Mb	mega base pairs
Mg^{2+}	magnesium
min	minute(s)
ml	milliliter
mm	millimeter
mM	millimolar
Mn^{2+}	manganese
mol	mole
MS	mass spectrometry
MudPIT	multidimensional protein identification technology
NaOAc	sodium acetate
NaOH	sodium hydroxide
ng	nanogram
nm	nanometer
nM	nanomolar
N-terminus	amino terminus
OD	optical density
ORF	open reading frame
PAGE	polyacrylamide gel electrophoresis
PBS	phosphate buffered saline

PCR	polymerase chain reaction
PG	peptidoglycan
Pi	protease inhibitor
p-NPP	p-nitrophenyl phosphate
pmol	picomole
PS	polystyrene
PVDF	polyvinylidene fluoride
qRT-PCR	quantitative (real time) reverse transcriptase PCR
RE	restriction endonuclease
RNA	ribonucleic acid
RNR	ribonucleotide reductase
rpm	revolutions per minute
RT	room temperature
RT-PCR	reverse transcriptase PCR
SAM, S-AdoMet	s-adenosylmethionine
SAP	shrimp alkaline phosphatase
Sc	spectinomycin
SEB	protoplasting buffer
SDS	sodium dodecyl sulfate
sec	second(s)
spp	species (plural)
SSC	saline-sodium citrate
ssp	sub-species
SOEing	splicing by overlap-extension
STM	signature-tagged mutagenesis
suc	sucrose
TB	Todd Hewitt broth
TBS	Tris buffered saline
TCA	trichloroacetic acid
TE	Tris-EDTA
Tet	tetracycline
Tnp	transposase
Tris-HCl	Tris hydrochloride
TSB	tryptic soy broth
TSA	tryptic soy agar
U	unit
UV	ultraviolet
X-gal	5-bromo-4-chloro-3-indolyl-β-$_D$-galactopyranoside
WT	wild type
yrs	years

Abstract

IDENTIFICATION OF VIRULENCE DETERMINANTS FOR *STREPTOCOCCUS SANGUINIS* INFECTIVE ENDOCARDITIS

By Lauren Senty Turner, Ph.D.

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy at Virginia Commonwealth University.

Virginia Commonwealth University, 2008

Major Director: Todd O. Kitten, Ph.D.
Associate Professor, Philips Institute, School of Dentistry

Streptococcus sanguinis is the second most common causative agent of bacterial infective endocarditis (IE). Risk of *S. sanguinis* IE is dependent on pre-disposing damage to the heart valve endothelium, which results in deposition of clotting factors for formation of a sterile thrombus (referred to as vegetation). Despite medical advances, high mortality and morbidity rates persist. Molecular characterization of *S. sanguinis* virulence determinants may enable development of prevention methods. In a previous screen for *S. sanguinis* virulence determinants by signature-tagged mutagenesis (STM) an attenuated mutant was identified with a transposon insertion in the *nrdD* gene, encoding an anaerobic

ribonucleotide reductase. Evaluation of this mutant, as well as an *nrdD* in-frame deletion mutant, JFP27, by a soft-agar growth assay confirmed the anaerobic growth sensitivity of these strains. These studies suggest that an oxygen gradient occurs at the site of infection which selects for expression of anaerobic-specific genes at the nexus of the vegetation. The random STM screen failed to identify any favorable streptococcal surface-exposed prophylactic candidates. It was also apparent that additional genetic tools were required to facilitate the *in vivo* analyses of mutant strains. As it was desirable to insert antibiotic resistance markers into the chromosome, we identified a chromosomal site for ectopic expression of foreign genes. *In vitro* and *in vivo* analyses verified that insertion into this site did not affect important cellular phenotypes. The genetic tools developed facilitated further *in vivo* screening of *S. sanguinis* cell wall-associated (Cwa) protein mutants. A directed application of STM was employed for a comprehensive analysis of this surface protein class in the rabbit model of IE. Putative sortases, upon which Cwa proteins are dependent for cell surface localization, were also evaluated. No single *S. sanguinis* Cwa protein was determined essential for IE by STM screening; however competitiveness for colonization of the infection site was reduced for the mutant lacking expression of sortase A. The studies described here present a progressive picture of *S. sanguinis* IE, beginning with surface protein-dependent colonization of the vegetation in early IE, that later shifts to a bacterial persistence *in situ* dependent on condition-specific housekeeping genes, including *nrdD*.

INTRODUCTION

The human oral flora

More than 700 different bacterial species and phylotypes have been detected in the human oral cavity; roughly 50 % of detected species (spp.) are uncultivable *ex vivo* by current laboratory methods. Bacterial spp. native to the oral cavity typically exhibit a predilection for specific anatomical sites within the mouth, so that certain sites have a subset of commonly associated isolates. The largest biomasses of oral bacteria in the mouth exist on the teeth and at the back of the tongue (referred to as the dorsum of the tongue). Bacteria on the teeth, which comprise dental plaque, can accumulate to ~10^{11} organisms per gram wet weight. The actual content and density of dental plaque is dependent on host factors including oral hygiene, and food intake. Epithelial cells sloughed from the surface of the tongue often contain greater than 100 organisms per cell (81). Resident oral bacteria preferentially colonize different oral surfaces through the specific recognition of oral surface receptors by bacterial cell surface adhesins.

Species identification by 16S rRNA gene sequencing has facilitated the detection of previously unrecognized oral species, which do not thrive once isolated from the complex oral environment. Aas and colleagues recently described a technique for direct extraction of DNA from plaque samples from multiple oral sites, including the dorsum and lateral sides of the tongue, buccal fold, hard and soft palates of the mouth,

labial gingiva and soft tissue surfaces, and supra- and sub-gingival plaques from the tooth surfaces. In their comprehensive analysis the researchers identified 141 taxa from six different bacterial phyla, including the *Firmicutes* (low G+C gram-positive species), the *Actinobacteria* (high G+C gram-positive species), the *Proteobacteria*, the *Bacteroidetes* (including pathogenic *Porphyromonas* and *Prevotella* spp.), the *Fusobacteria*, and a phylum of uncultivatable species, TM7 (1). DNA probe detection has also identified species that may be clustered as "bacterial complexes", as general indicators of oral health (237, 238). For example, species associated with clinical aspects of periodontal disease, including *Porphyromonas gingivalis*, *Treponema denticola*, and *Prevotella intermedia* typically cluster, and are exclusive from bacterial species commonly associated with healthy oral sites, including *Streptococcus gordonii*, *Streptococcus intermedius*, *Streptococcus oralis*, and *Streptococcus sanguinis* (237). Indeed, a healthy oral cavity contains distinctively different flora from one afflicted with oral disease. In healthy oral sites sampled by Aas and colleagues, neither spp. associated with periodontal disease, nor dental caries (dental cavities) were detected (1). Periodontitis is demarcated by the shift in dental plaque flora from gram-positive facultative anaerobes to gram-negative obligate anaerobes (237). Dental caries is associated with an increase in focal acidogenicity attributed to breakdown of dietary sucrose by *Streptococcus mutans* and *Lactobacillus* spp. (150). Fluctuations in the oral environment are in part defined by the oral microflora; however host behaviors and physiology also influence the composition of the flora.

Oral streptococci

The genus *Streptococcus* consists of more than 50 spp., the majority of which are divisible into one of six phylogenetic clusters on the basis of 16S rRNA gene sequence similarity. *Streptococcus* spp. are gram-positive and spherical or ovoid in shape, with a diameter less than 2 μm. The pattern of cellular division in streptococci distinguishes this genus from other gram-positive coccus shaped bacteria. In streptococci division occurs along a single axis, resulting in chains or pairs of cocci (233). The name *Streptococcus* derives from *sreptos*, meaning easily bent or twisted, like a chain, and *kokkos*, meaning berry, as a descriptive of the spherical shape of the cell. Streptococci are catalase-negative, facultatively anaerobic, non-motile and do not produce spores or gas. Unlike enterococci, streptococci do not grow in 6.5 % sodium chloride. Streptococci are also distinct from lactococci, which are Lancefield (cell surface carbohydrate) group N organisms.

Streptococci constitute ~20 % of the normal human oral flora, and are abundant at most oral sites, although species within the *Streptococcus* genera demonstrate preferential association with different oral locales. For example, *S. salivarius*, *S. parasanguinis* and *S. australis* are typically found on the dorsum of the tongue, whereas *S. sanguinis* and *S. gordonii* exhibit a predilection for the tooth surface (1). *S. mutans* associates with teeth along the gingival margin of the tooth and gum tissue, and at pits and fissures at the surface of molars (81). In contrast, *S. mitis* is nearly ubiquitous in the oral cavity (1). Among the oral microbes streptococci have been heavily investigated as

these bacteria are readily cultivatable, are a major component of the microflora on most oral surfaces, and in the case of *S. mutans*, may be relevant for oral disease.

S. sanguinis is of particular interest to the research described here. *S. sanguinis* is a member of the mitis group of streptococci, which also includes *S. gordonii*, *S. parasanguinis*, *S. oralis*, *S. mitis*, *S. pneumoniae*, *S. pseudopneumoniae*, *S. infantis*, *S. cristatus*, *S. peroris*, *S. australis* and *S. sinensis* (106, 120). With the exception of *S. pneumoniae* and *S. pseudopneumoniae* the mitis group streptococci are oral commensals, and pioneer colonizers of the tooth surface. *S. sanguinis* has also been classified as a member of its own group, the sanguinis group, in recent literature, which was expanded to include *S. gordonii* and *S. parasanguinis* as well. In this case, the separation between sanguinis and mitis groups is based on the predominance of sanguinis group strains that are positive for two catalytic tests, the deamination of arginine and the hydrolysis of esculin (71). The majority of mitis group strains are negative for these assays.

Despite these differences in grouping, both mitis and sanguinis group streptococci, and additional oral streptococcal groups are subgroups of a more general classification, the viridans group of streptococci. (This term does not encompass *S. pneumoniae*, however; 71.) The term viridans derives from the Latin word *viridis*, or green, which refers to the coloration on blood agar resulting from partial hemolysis of erythrocytes by α-hemolytic streptococci, due to the release of streptococcal H_2O_2 (10). While *S. sanguinis* is positive for α-hemolysis, the designation of viridans for this group is confusing given that members of the group have also been documented as γ-

hemolytic (no lytic zones observed surrounding colonies on blood agar), and β-hemolytic (zones of clearing surrounding colonies on blood agar due to complete hemolysis of erythrocytes) (233). The viridans group has been described as characteristically lacking in features which distinguish other major streptococcal pathogens, including β-hemolysis, Lancefield typing (cell wall carbohydrate) of β-hemolytic streptococci, CAMP test (for Lancefield group B streptococci), and optochin sensitivity (*S. pneumoniae*) (233).

The complete genome sequences of many streptococcal species, including a few oral streptococcal species, *S. mutans* UA159, *S. gordonii* CH1 Challis, and *S. sanguinis* SK36 are currently available (3, 270, 285). The SK36 strain of *S. sanguinis* was chosen for genome sequencing because this strain exhibits definitive characteristics of this species, including robust glucan production, superior platelet adhesion and aggregation phenotypes (relative to other sequencing candidates), genetic competence, and infectivity in the rat model of IE comparable to virulent strains (www.sanguinis.mic.vcu.edu/strain_selection.htm). The streptococcal chromosome is uniformly a single circular DNA molecule. Among published streptococcal genomes, *S. sanguinis* SK36 has the largest; at 2,388,435 bp, it is roughly 7 to 24 % larger than other published genomes (285). The difference in genome size between these species is also reflected in the number of predicted encoded proteins. *S. sanguinis* SK36 is predicted to code for 2,274 proteins, whereas *S. mutans* UA159, and *S. gordonii* CH1 Challis are predicted to encode 1,960 and 2,051 proteins, respectively (3, 270, 285). *S. sanguinis*

SK36 also has the highest G+C genome content of the low G+C streptococci, at 43.4 % (285).

S. sanguinis has a general lack of sequence elements suggestive of genetic transfer by transposons or prophage. This may be attributed to the presence of a putative DpnII-like system that reduces the efficiency of horizontal gene transfer (HGT) by phage infection, conjugative plasmid transfer and plasmid transformation in *S. pneumoniae* (134). However, at least 270 genes of *S. sanguinis* were predicted as candidates for HGT based on phyletic patterns of gene distribution (285). It is likely that transformation is the major method accounting for HGT in *S. sanguinis* SK36. (Transformation by chromosomal DNA would be unaffected by the putative DnpII restriction system.) *S. sanguinis* SK36, like several other oral streptococci encodes a genetic system for the uptake and integration of foreign DNA in response to a strain-specific competence inducer, referred to as competence stimulating peptide (CSP) (79). In *S. pneumoniae* 22 proteins are required for chromosomal transformation by this quorum sensing system. Orthologs of 20 of these proteins were identified for *S. sanguinis* (285). These reports suggest that the genetic complexity of *S. sanguinis* may be attributed to the acquisition of novel genes by HGT through transformation *in situ*.

Streptococcal colonization of the oral cavity

Shortly after birth, bacteria appear in the mouths of infants as their oral microflora begins to take shape. Early on, *S. salivarius* is detectible (primarily on the dorsum of the tongue), and by one month of age all infants are colonized with at least

one viridans streptococcal species (201). It is thought that most endogenous flora microorganisms are acquired from parents or attendants of the infant. The emergence of teeth results in more drastic changes in an infant's oral flora. The hydroxyapatite surface of the tooth, once covered in salivary proteins (the salivary pellicle) presents an abundance of binding sites utilized by streptococcal colonizers for retention in the oral cavity. *S. sanguinis* has been shown to bind specific salivary proteins including salivary IgA and α-amylase, among the heterogeneous mix of salivary components and glycoproteins on the tooth surface (40, 85). The window of colonization of the tooth surface by *S. sanguinis* has been narrowed to ~9 months of age, with the first appearance of teeth. In a study of tooth colonization by *S. sanguinis*, Caufield and colleagues discovered that infants with higher salivary proportions of *S. sanguinis* experienced delayed colonization by *S. mutans*, suggesting competition between these streptococci for potential binding sites, nutrient resources, or by release of inter-species inhibitors. The latter was supported by the observation that eventual emergence of *S. mutans* resulted in decreased levels of *S. sanguinis* (40). *S. sanguinis*, as a major component of dental plaque and primary colonizer of the tooth surface, is thought to act as a scaffold to bring other bacterial species closer to the surface of the tooth for formation of complex multi-species dental plaque. While *S. mutans* is also a dental plaque species, mounting evidence supports the theory that these two species do not complacently coexist. Comparisons of sequential colonization of *S. sanguinis* and *S. mutans* in an *in vitro* mixed biofilm assay suggest that once one species gains a foothold, the other is at a loss and can not overcome the initial disadvantage. However,

when co-administered the two strains are capable of forming a mixed species biofilm, and thus interact in a way that is much more complex than general antagonism. It has been demonstrated that the effectors aiding in competitive colonization *in vitro* of these two species include, for *S. sanguinis*, H_2O_2 production, and for *S. mutans*, production of the bacteriocin, mutacin (132). However whether a similar interplay occurs *in vivo*, in which bacteriocins may be inactivated by proteolytic enzymes of dental plaque has not been determined. Despite the apparent benefit of *S. sanguinis* in the oral cavity—as an indicator of oral health and putative competitor of pathogenic oral bacteria—*S. sanguinis* is a medically relevant opportunistic agent of infective endocarditis.

Infective endocarditis

Infective endocarditis (IE) is a life threatening, non-contagious chronic infection of the cardiac valves, or lining of the heart mainly caused by bacterial species, although fungi can also be associated with this infection (20). Viral infections are also related to development of IE. Cardiac complications, including IE are of great concern for HIV patients. A risk factor for bacterial IE in this population is intravenous drug use (45, 263). Cardiac complications leading to bacterial IE have also been described following varicella zoster virus infection (2).

A comprehensive review of 26 publications, encompassing more than 3,500 cases of IE revealed important epidemiological trends for this disease. The mean incidence for disease was 3.6 cases per 100,000, per year. The incidence of disease increased with age, with fewer than 5 cases per 100,000 in individuals less than 50 yrs

to more than 15 cases per 100,000 in individuals over 65 yrs (176). The mean age of patients ranged from 36 to 69 yrs. In the preantibiotic era this disease commonly affected persons between 30 and 40 yrs. The emergence of antibiotic therapy and increased longevity in developed countries has resulted in a shift to more aged IE patients (181). The male to female ratio was determined to be roughly two-to-one.

IE may be classified as sub-acute or acute related to the severity of the clinical presentation and progression of the disease. Additional classifications include the nature of the valve involved (native *vs.* prosthetic), and the source of the causative agent, such as community acquired infection, nosocomial IE, or IE attributed to intravenous (IV) drug use. Risk factors for IE may be classified as cardiac risk factors or risk factors for bacteremia. A prosthetic valve is an example of a cardiac risk factor, whereas IV drug use and some medical procedures are considered risk factors for bacteremia.

Native-valve IE (NVE) is classically associated with congenital heat disease and chronic rheumatic heart disease. In addition valvular lesions due to valve dysfunction—attributed to mitral and aortic valve regurgitation, as well as aortic valve stenosis and mitral valve prolapse—are major predisposing factors in elderly populations (186). NVE most commonly presents as “left sided” IE, with infection of the mitral and aortic valves of the heart most common. Prosthetic valve IE (PVE) is divided into late or early infection depending on whether infection presents within 60 days of surgery (early) or later. This condition is most common within the first 2 months of valve implantation, and is often caused by *S. aureus* and *S. epidermidis*. Late PVE is more commonly attributed to non-staphylococcal organisms, including streptococci, and gram-negative

bacteria of the HACEK group (*Haemophilus* spp., *Actinobacillus actinomycetemcomitans*, *Cardiobacterium hominus*, *Eikenella corrodens*, and *Kingella kingae*) (181). IE resulting from IV drug use is more common among younger cohorts (30 to 40 yrs.). IV drug users more often present with the tricuspid valve of the heart affected. Organisms resulting in IE in this group are typically skin isolates, with *S. aureus* as the most common causative agent. In nosocomial IE, many patients lack predisposing cardiac damage associated with NVE. Predominant causative agents of nosocomial IE include enterococci and staphylococci, which may be introduced into the bloodstream on contaminated catheters or other indwelling devices (176).

Staphylococci are the most common causative agents of IE, accounting for 44 % of NVE, 69 % of intravenous drug user IE, 67 % of early PVE, and 46 % of late PVE. The second most common causative agents of NVE and late PVE are streptococci, which account for ~31 % and 35 % of infections, respectively (176). The oral streptococci are the most common agents of streptococcal IE, and among viridans group streptococci, *S. sanguinis* is most often implicated as the cause of NVE (57, 67, 216, 272).

Progression of oral streptococcal NVE

The development of NVE proceeds in a series of general steps. As described above, the first step is predisposing damage to heart valve endothelium. During transient bacteremia, streptococci adhere to, and colonize the site of endothelial damage. The characteristic lesion of IE, referred to as a vegetation, includes a platelet

and fibrin thrombus, embedded with infective microorganisms and inflammatory immune cells. Once attached, the bacteria become engulfed in vegetation components as the vegetation enlarges in size, and persist at the site of infection. Dissemination of septic emboli may then result in peripheral manifestations of this disease (reviewed in 176, 177).

Following damage to endothelia, sub-endothelial components including endothelial matrix proteins triggering blood coagulation are exposed or secreted. During thrombosis, induction of pro-inflammatory cytokines results in the up-regulation of proteins mediating coagulation, such as tissue factor, a glycoprotein that in surface-bound or soluble form interacts with soluble factors to form proteolytically active thrombin. Thrombin converts plasma fibrinogen to fibrin, which envelops and stabilizes the developing blood clot (232).

The bacteremic event preceding IE by oral streptococci is asymptomatic and may result from dental manipulation or less invasive daily tasks. Dental procedures resulting in the greatest levels of subsequent bacteremic events include dental extraction, periodontal probing in patients with periodontitis, tartar removal from the tooth surface, endodontic procedures, and tooth scaling to remove plaque from below the gum-line (207). Bacteremic events may also result from mundane daily activities like eating, chewing gum, tooth brushing, flossing, or use of toothpicks (75, 86).

Development of *S. sanguinis* IE putatively requires adhesin-dependent bacterial colonization of the vegetation within minutes of transient bacteremia onset. Surface adhesins of streptococci are considered indispensable in the primary interaction with

host proteins. Several *S. sanguinis* surface-exposed cell wall-anchored proteins have been determined to be mediators of platelet interaction and activation *in vitro* (98, 122, 204). This phenotype has been correlated with virulence in the rabbit model of IE (97). The capacity for adherence and aggregation of platelets could be an important factor in colonization of the platelet and fibrin matrix during transient bacteremia, and may also encourage further deposition of platelets and fibrin for engulfment of streptococci at the site of infection. In spite of the course of disease progression described here, and the observation that duplication of these conditions incites IE in the rabbit model of the disease, there is evidence to suggest that streptococcal colonization of damaged endothelium can occur in the absence of the pre-formed aseptic thrombus. In the study resulting in this observation, rabbits with induced valve damage, treated with warfarin (an inhibitor of fibrin-platelet matrix formation) still had bacterial colonization on the surface of the valve. As expected, vegetation formation was altered in these rabbits; however they presented with characteristic morbidities of IE, including high fever and constant bacteremia. However, the antibiotic therapies administered to these rabbits were more effective than in control rabbits that did not receive warfarin, therefore suggesting that vegetation formation has an additional benefit for the infective bacteria (104). Oral streptococci encode many surface proteins vital for maintenance in the oral cavity; yet several proteins which bind to human salivary proteins have also been shown to bind extracellular matrix proteins (ECM) *in vitro* (14, 52). Therefore these proteins could potentially facilitate the association of oral streptococci with sub-endothelial ECM proteins exposed at the site of tissue damage.

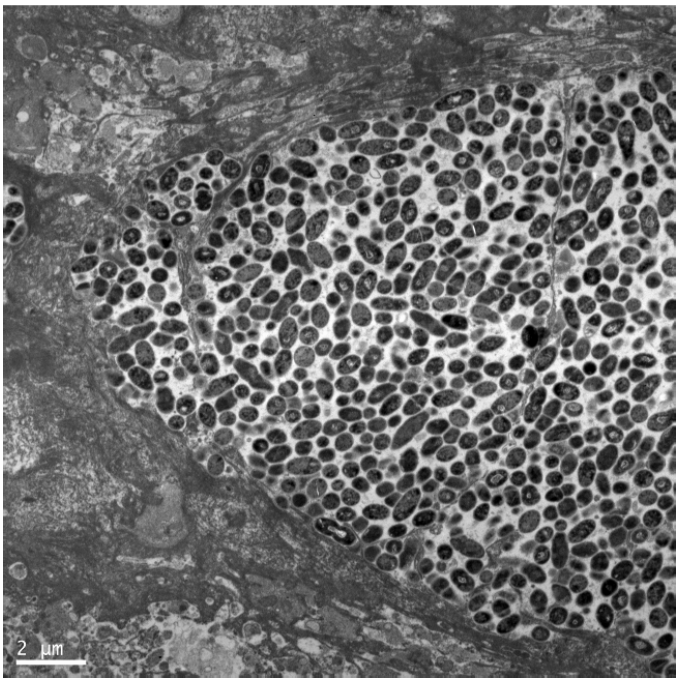
Following streptococcal colonization of the vegetation, the progression of the disease is, at least in part, due to the pro-inflammatory response initiated by monocytes attracted to the site of infection. At the site of infection peripheral blood monocytes are proposed to undergo activation resulting in production of pro-inflammatory cytokines and tissue factor activity, leading to further coagulation of blood components for engulfment of adherent streptococci (176). In the rabbit model of IE, bacteria are completely engulfed within three hours of administration (63). *In vitro* assays of *S. sanguinis* infected endothelial cell monolayers have indicated that pro-coagulant activity from tissue factor production is enhanced by soluble IL-1 and TNF-α during monocyte co-culture (268, 269). Additionally, ELISA of patient sera has revealed increased presence of the pro-inflammatory cytokine Interleukin-6 (IL-6), and the cytokine receptor IL-2R, which diminish during treatment with antibiotic alone, or antibiotic and cell wall-active antibacterial agents (4, 211). In this scenario, the progressive vegetation enlargement in response to cytokine production and tissue factor activity would cause multiple pathologies associated with IE (181).

Transmission electron microscopy of vegetations isolated from infected rabbits shows streptococci embedded in layers of fibrin meshwork (Fig. 1). Oftentimes streptococci are found in multiple zones of the vegetation. Multiple rounds of streptococcal deposition on the vegetation may occur by re-seeding of vegetation components dislodged in the bloodstream. As streptococci are known to not only persist, but also grow within the vegetation, it is also possible that the layers observed are due to outgrowth from the initial colonization site (63). Within 24 hours of

Figure 1. *S. sanguinis* SK36 vegetation

A transmission electron micrograph of an SK36 vegetation cross-section, resulting from an experimental case of IE, is shown. Bacterial cells are electron dense and are spherical or ovoid in shape. The bacterial zone is surrounded by fibrin layers.

Figure 1. *S. sanguinis* SK36 vegetation



administration, streptococci begin to exhibit biofilm-like sessile community characteristics, including the development of antibiotic resistant microcolonies (51, 63, 64).

If left untreated, IE is uniformly fatal, with death resulting from embolization of vegetation components to major organs, or congestive heart failure. Infection-induced valvular damage is typically the cause of congestive heart failure in patients with IE. In rare cases myocardial infarction results from embolization of vegetation fragments. Up to 60 % of embolic IE events involve the central nervous system, and neurological complications, including stroke and mycotic aneurism, occur in ~20 to 40 % of patients. The spleen, kidney and liver are also commonly affected by systemic embolization and seeding of septic vegetation fragments (181).

Symptoms, treatment and prevention of bacterial IE

Symptoms of IE are general and include fever, anorexia, weight loss, malaise and night sweats. Heart arrhythmias like heart murmur are also common indicators of IE, in conjunction with other described symptoms. However, in most cases the heart murmur is preexisting and due to valve dysfunction (181, 186). Peripheral manifestations include petechiae on the skin, conjunctiva, or oral mucosa owing to small hemorrhages following embolization of vegetation pieces. In addition, Osler's and Janeway's lesions, attributed to the deposition of immune complexes subcutaneously and in small blood vessels, respectively, are pathognomonic of IE (181).

In 1994 the “Duke Criteria” for diagnosis of IE was proposed by a group at Duke University. The criteria were validated worldwide and then revised in 2000 for creation of diagnostic criteria that would identify culture-negative IE with greater accuracy (66, 147). The criteria incorporate clinical, laboratory and echocardiographic information for informed diagnosis of IE. The proposed criteria were subdivided into “major criteria”, including blood culture and echocardiography findings, and “minor criteria”.

The major blood culture criteria include two positive blood cultures of a typical IE organism drawn more than 30 minutes apart, as well as blood cultures described as “persistently positive”, meaning two culture sets drawn at least 12 hours apart, or three to four sets drawn over more than a 1-hour window. The culture criteria are not applicable to culture-negative agents, which has prompted the expansion of detection methods to include assays for DNA detection, and immunodetection.

Echocardiography may provide strong indication of the presence of a vegetation. Transthoracic echocardiography (TTE) is a quick and non-invasive detection tool. However some patients are not good candidates for TTE due to issues with visualization within the chest cavity. (Obesity, chronic obstructive pulmonary disease, or chest-wall deformities complicate the interpretation of TTE results.) Transesophageal echocardiography, in which an ultrasound transducer probe is inserted into a patient’s esophagus, is an alternative to TTE for enhanced visualization of the vegetation mass (181).

Minor criteria for IE diagnosis include predisposing cardiac conditions and intravenous drug use, as well as the more general symptoms and peripheral indicators of IE described above. The combination of positive factors between the major and minor criteria sub-groups results (theoretically) in diagnoses that range from definite IE, to possible IE, to rejection of IE being the affliction.

The treatment regimen for resolution of IE requires prolonged administration of antibiotics. In the case of viridans streptococci the antibiotics suggested differ according to the penicillin-susceptibility of the infective agent. The aggressive treatment strategy implemented typically requires IV or intramuscular administration of antibiotics daily, for a minimum of two weeks. For penicillin-sensitive streptococci ceftriaxone alone, or in combination with netilmicin may be recommended. Alternatives include procaine benzylpenicillin alone or with gentamycin, or vancomycin for patients allergic to β-lactam antibiotics. For elimination of intermediate penicillin-resistant viridans streptococci procaine benzylpenicillin with gentamycin, or vancomycin may be used (176). In extreme cases surgery is required for treatment of IE. Surgical intervention encompasses radical valve replacement and removal of the infective vegetation, or repair of the damaged heart valve (176, 181).

As no vaccine for prevention of IE exists, the most effective prophylactic measures include identification of at-risk patients and their education of behaviors that might cause bacteremia. American Heart Association (AHA) recommendations for patients at high risk for IE emphasize the maintenance of proper oral health and antibiotic prophylaxis during dental procedures. Antibiotic prophylaxis is generally

recommended for procedures associated with significant bleeding from hard or soft tissues, including periodontal surgery, scaling and tooth cleaning. For efficacy antibiotics should be administered, at the latest, 2 hours post-procedure (282). Despite improvements in endocarditis outcome through advances in anti-microbial therapy and better diagnostic and treatment tools, substantial morbidity and mortality still result from IE. Therefore development of alternative preventative methods is called for.

Streptococcal virulence factors for IE

The investigation of streptococcal virulence factors for IE has included surface-exposed proteins that would mediate the initial stages of disease. These include the major surface-exposed protein class of gram-positive bacteria, the cell wall anchored (Cwa) proteins. Cwa proteins are considered particularly important in pathogenesis as many have been grouped as MSCRAMMs, or microbial surface components recognizing adhesive matrix molecules of host tissue and host cells. The staphylococcal MSCRAMMs, clumping factor A (ClfA) and fibronectin binding proteins A and B (FnbA, FnbB) have been demonstrated by experimental IE to be important pathogenicity factors (177). As their name implies, FnbA and FnbB are capable of binding to fibronectin *in vitro*, and ClfA has fibrinogen binding capacity *in vitro*; both of these proteins are soluble in blood, and fibronectin exists in the extracellular matrix of most host tissues (114, 167, 206).

Streptococcal MSCRAMMs identified as mediators of IE include Hsa of *S. gordonii*, and CbpA of *S. sanguinis*. Hsa is a serine rich protein with fibronectin binding

capacity *in vitro*, and was shown to contribute to infectivity under competitive conditions in the rabbit model of IE (248). CbpA of *S. sanguinis* was an identified mediator of platelet aggregation *in vitro* (98). A mutant defective for expression of CbpA also exhibited reduced infectivity in the rabbit model of IE, suggesting that the platelet interaction phenotype contributes to development of IE (97).

Extracellular production of glucan polymers has also been linked to viridans streptococcal adherence phenotypes *in vitro* and infectivity *in vivo*. The extracellular glucans are synthesized by bacterially encoded glucosyltransferase (Gtf) and fructosyltransferase (Ftf) enzymes from a sucrose substrate. In a previous study the production of water-soluble glucan by *S. sanguinis* grown in 5 % sucrose enhanced adherence to aortic, mitral and tricuspid valves *in vitro*, thus suggesting that exopolysaccharide production may enhance colonization in the development of IE (210). *In vivo* analysis of an *S. mutans* mutant deficient for glucan and fructan synthesis confirmed that exopolysaccharide production was positively correlated with the development of IE in the rat model of infection. *In vitro* phenotypic analysis suggests that glucan and fructan production both enhance streptococcal survival post-phagocytosis, and putatively mediate adherence to vegetation-like matrices (180). However, an *S. gordonii gtf* mutant maintained infectivity in the rat model of IE; therefore the role of sugar-polymer production in development of IE may differ between *Streptococcus* species (277).

A family of related lipoproteins, including FimA of *S. parasanguinis*, has also been implicated in streptococcal adherence and virulence phenotypes. A *fimA* mutant

displayed decreased binding to fibrin monolayers *in vitro*, and was decreased for infectivity in the rat model of IE (34). The *fimA* ortholog in *S. mutans*, SloC is also important for development of IE (126, 197). FimA and SloC are members of ABC transport systems for acquisition of extracellular manganese and iron; this specific function has also been evidenced as important for *S. mutans* IE (197).

The random signature-tagged mutagenesis (STM) technique (technical details of which are further described in Chapter V) was recently applied to *S. sanguinis* SK36 for identification of virulence determinants in the rabbit model of IE. Of 800 transposon insertion mutants screened *in vivo*, five were identified in which insertion mutagenesis resulted in a drastic decrease in recovery. In the five strains, genes encoding housekeeping functions, including amino acid and nucleic acid synthesis, and the ability to survive in anaerobic conditions (further detail is provided in Chapter III) were mutated (198). Genes considered virulence factors for development of IE, including MSCRAMMs and lipoproteins were not identified in this screen.

MATERIALS AND METHODS

Bacterial strains, growth media and chemicals

Strains and plasmids used in the described studies are listed in Table 1. *S. sanguinis* SK36, a human oral isolate was obtained from Mogens Killian (University of Aarhus, Denmark) (124). *S. sanguinis* strains were routinely cultured under a reduced O_2 condition created by gas exchange with the final gas mixture unknown, a 6 % O_2 condition (10 % H_2, 10 % CO_2, and 80 % N_2), or in an anaerobe chamber (Forma Scientific, Marietta, Ohio) with a palladium catalyst to maintain O_2 levels less than 1 ppm. All culture media were purchased from Difco. *S. sanguinis* strains were routinely cultured in brain heart infusion (BHI) (Difco) media for liquid culture, or on Tryptic soy broth (TSB) (Difco) containing 1.5 % agar (TSA). Soft-agar studies required growth of *S. sanguinis* strains in TSB supplemented with 1 % low melting point (LMP) agarose. For transformation experiments *S. sanguinis* strains were cultured in Todd-Hewitt broth containing 2.5 % heat-inactivated horse serum (TH-HS). For this purpose TH broth was adjusted to a pH of 7.6, filter sterilized, and stored at -20 ºC. TH broth was thawed and supplemented with 2.5 % HS immediately before use. For chromosomal DNA preparation *S. sanguinis* was grown in BHI supplemented with 20 mM DL-threonine (BHIT). When required for selective growth of *S. sanguinis*, erythromycin (Em), chloramphenicol (Cm), spectinomycin (Sc),

Table 1. Bacterial strains and plasmids used

S. sanguinis	**Genotype or description**	**Source**
SK36	Human plaque isolate	(124)
6-26	Cm^r; *nrdD*::*magellan2*, derived from SK36	(198)
3-24	Cm^r; *SSA_0707-SSA_0708* intergenic region::*magellan2*, derived from SK36	(198)
VT1614	Kn^r; Δ*srpA*::*aphA-3*, derived from SK36	(204)
JFP27	Δ*nrdD*, derived from SK36	This study
JFP36	Em^r; Δ*0169*::*pSerm*, derived from SK36	This study
JFP56	Sp^r; Δ*0169*::*aad9*, derived from SK36	This study
JFP76	Tet^r; Δ*0169*::*tetM*, derived from SK36	This study
CWA STM	Cm^r; *SSA_#*::*Himar1* derived *magellan2*	This study
SRT18	Cm^r; *srtC*::*magellan2*	This study
SRT21	Cm^r; *srtA*::*magellan2*	This study
SRT24	Cm^r; *srtB*::*magellan2*	This study
JFP42	Sc^r; Δ*srtA*::*aad9*	This study
JFP44	Kn^r; Δ*srtB*::*kan*	This study
JFP47	Sc^r, Kn^r; Δ*srtA*::*aad9*, 0169::*srtAkan*	This study
JFP42 (pJFP50)	Sc^r, Em^r; Δ*srtA*::*aad9* (pJFP50)	This study
JFP52	Sc^r, Em^r; Δ*srtA*::*aad9* (pJFP52)	This study
JFP62	Sp^r, Em^r; Δ*srtA*::*aad9* (pJFP62)	This study
Plasmid	**Genotype or description**	**Source**
pVA2606	Kn^r; suicide vector derived from pUC19 by replacement of *bla* with *aphA-3*	(198)
pJFP27	Kn^r; pVA2606 containing Δ*nrdD*	This study
pVA838	Em^r, Cm^r; *E. coli-Streptococcus* shuttle plasmid	(154)
pJFP29	Em^r; pVA838 containing the 2.7-kb *nrdD* locus	This study
pJFP30	Em^r; pVA838 containing the 2.4-kb nrdD locus	This study
pJFP34	Kn^r; derivative of pVA2606 with *0169*	This study
pJFP36	Kn^r, Em^r; derivative of pJFP34 with *pSerm*	This study
pJFP46	Kn^r; derivative of pJFP34 with *Asc*I, *Not*I in 0*169*	This study
pJFP56	Sc^r; derivative of pJFP46 with *aad9*	This study
pJFP76	Tet^r; derivative of pJFP46 with *tetM*	This study
pJFP16	Kn^r; derivative of pVA2606 containing 3.3 kb *nrdD*::*cat*	(196)
pCM18	Em^r; pTRKL2-PCP25-RBSII-*gfp*mut3-T0-T1	(90)
pJFP50	Em^r, Cm^r; pVA838 containing the 1.112 kb *srtA* locus	This study
pJFP52	Em^r; pCM18 containing the 0.8 kb *srtA* locus	This study
pJFP62	Em^r, Cm^r; pVA838-PCP25-*srtA*	This study

kanamycin (Kn) and tetracycline (Tet) were used at 10, 5, 200, 500, and 5 μg ml^{-1}, respectively.

Electromax E. *coli* DH10B (Invitrogen) was used as a bacterial host for plasmid construction. All *E. coli* strains were cultured in Luria-Bertani (LB) medium, or on LB containing 1.5 % agarose (LA). When required for *E. coli* X-gal was added at 40 or 60 μg ml^{-1}. For selective culture of *E. coli*, Kn, Em, Cm, Tet, and Sc were used at 50, 300, 5, 2.5, and 100 μg ml^{-1}, respectively.

DNA Manipulations

Chromosomal DNA was isolated from *S. sanguinis* SK36 as described previously (126). Labeling and detection in Southern dot-blotting was performed by the Genius digoxigenin system (Roche Molecular Biochemicals). PCR was routinely performed in a GeneAmp 960 thermal cycler (PE Biosystems), MyCycler (Bio-Rad), or an MJ Research PCT-200 thermal cycler (Bio-Rad). Real-time quantitative PCR was preformed with a 7500 Fast Real Time PCR System (Applied Biosystems). PCR amplification was performed using Platinum® PCR Supermix (Supermix) or High Fidelity Platinum® PCR Supermix (HIFI Supermix) (Invitrogen). PCR reaction mixtures typically included 0.2 pmol μl^{-1} primers, and *S. sanguinis* SK36 chromosomal DNA or PCR amplicon as template at 0.2 ng μl^{-1}. Plasmid DNA as PCR template was used at 0.01 ng μl^{-1}. Oligonucleotide primers were synthesized by the Nucleic Acids Research Facility of Virginia Commonwealth University or Integrated DNA Technologies. Restriction enzymes were purchased from New England Biolabs Inc. and used according to the manufacturer's

instructions. DNA purification from agarose was routinely performed using Quantum Prep Freeze 'N Squeeze columns (Bio-Rad). For plasmid DNA extraction form *E. coli* the Quantum Prep Plasmid Mini-Prep Kit (Bio-Rad) or the QiaPrep Spin Miniprep Kit (Qiagen) was used. Purification of PCR products was performed using the MinElute PCR Purification kit (Qiagen). T4 DNA ligase (Invitrogen) was routinely used for cloning purposes, as directed by the manufacturer.

Development and screening of *S. sanguinis* SK36 mutant strains

The sequences of most primers described in this section are given in Appendix A.

***Construction of JFP27 (ΔnrdD)*:** Gene Splicing by Overlap Extension (Gene SOEing (105) was used to construct an internal, in-frame deletion of the *nrdD* gene with the fusion of the first 12 codons and last 14 codons. Oligonucleotide pairs ARTR-SalI-24495-up and ARTR-23327-dn, and ARTR-21212-up and ARTR-SalI-20031-dn were used in the first PCR with S. *sanguinis* SK36 chromosomal DNA (gDNA) template and Supermix. The cycling conditions were as follows: One cycle at 94 ºC for five min; followed by 30 cycles at 94 ºC for five min, 55 ºC for 30 sec, and 72 ºC for one min; and a final extension step at 72 ºC for 7 min. The resulting PCR products were gel extracted, purified (Bio-Rad), and concentrated by precipitation with 3 M NaOAc and 95 % EtOH. The purified products were adjusted to 4 ng μl^{-1} and combined at a 1:1 ratio for amplification in a second PCR reaction with primers ARTR-SalI-24495-up and ARTR-SalI-20031-dn. The cycling conditions used were as follows: one cycle at 94 ºC for 40 sec; followed by 30 cycles of

94 ºC for 40 sec, 58 ºC for 30 sec, and 72 ºC for 6 min; and a final extension step at 72 ºC for 7 min.

The band of the anticipated size was gel extracted, purified and concentrated as described above. The ARTR-24495/20031 product was then cloned into suicide plasmid pVA2606 (Fig. 2A) via a *Sal*I restriction site for creation of the pJFP27 (Fig. 2B). *E. coli* DH10B was electrotransformed with pJFP27 and plated on LA containing Kn and X-gal. Blue-white screening of resulting colonies identified pJFP27 transformants. Plasmid DNA was prepared (Bio-Rad) and screened by PCR with ARTR-SalI-24495-up and ARTR-SalI-20031-dn primers. DNA sequence analysis confirmed proper insertion and in-frame deletion of *nrdD*.

S. sanguinis SK36 was transformed with 500 ng *Ssp*I-linearized pJFP27 and plated on TSA. Colonies were inoculated into overnight broth cultures for rapid PCR analysis by the boiled cell extraction technique. Briefly, 1 ml of each overnight culture was centrifuged (18,000 x *g* for 2 min) and the cell pellet was washed in 1 ml 10 mM Tris-HCl (pH 8.0), and suspended in 100 µl 10 mM Tris-HCl. Tubes were then boiled for 5 min, cooled and centrifuged again, and 100 µl of supernatant was transferred to a new tube. Two µl of the supernatant served as template for PCR screening reactions with primers ARTR-SalI-24495-up and ARTR-SalI-20031-dn. A clone that produced a PCR amplicon consistent with *nrdD* deletion was selected for further study and named JFP27.

***Construction of plasmids for nrdD complementation*:** Two *nrdD* gene region constructs were designed for complementation studies. Primers specific for the two constructs, ARTR-EcoI-up1, ARTR-EcoRI-up2 and ARTR-EcoRI-dn were used in PCR reactions

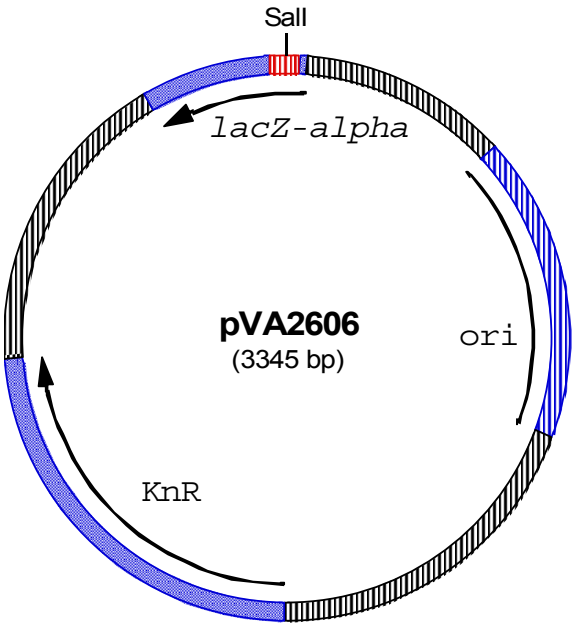
Figure 2. Plasmid maps of pVA2606 and pJFP27

A) The pVA2606 vector was derived from the *E. coli* plasmid pUC19 by replacement of the *bla* ampicillin resistance determinant with the *aphA-3* kanamycin resistance (KnR) cassette (198). The plasmid contains the pMB1 origin of replication (ori) from pBR322, but includes additional modifications for temperature-determined copy number, roughly 75 plasmid copies per cell at 37 ºC (and >200 copies per cell at 42ºC). The multiple cloning site (MCS), in red, is in frame with the *lacZα* gene (*lacZ-alpha*) for screening of MCS insertion by α-complementation (289).

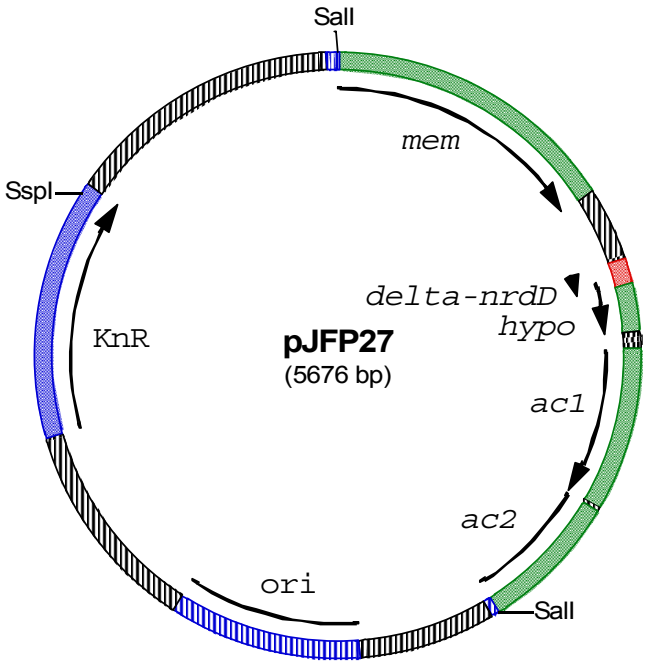
B) The Gene splicing by overlap extension (Gene SOEing) PCR technique was used to create an in-frame deletion of the *nrdD* gene. Flanking primers were designed for amplification and cloning of the mutation site, and neighboring ORFs, including a putative membrane protein gene (*mem*) upstream of *nrdD*, a hypothetical protein gene overlapping the 3' end of *nrdD* (*hypo*), as well as two putative downstream acetyltransferase genes (*ac1* and *ac2*). The *Sal*I restriction sites incorporated into flanking primers permitted cloning of the construct into pVA2606 to create pJFP27. The plasmid was made linear by *Ssp*I digestion prior to transformation of SK36.

Figure 2.

A.


SalI
lacZ-alpha
pVA2606
(3345 bp)
ori
KnR

B.


SalI
mem
SspI
delta-nrdD
hypo
KnR
pJFP27
(5676 bp)
ac1
ac2
ori
SalI

with *S. sanguinis* SK36 gDNA template, and Supermix (Invitrogen). Cycling conditions for amplification with the up1/dn primer pair were: one cycle at 94 ºC for five min; five cycles of 94 ºC for 40 sec, 41 ºC for 30 sec, and 68 ºC for 3 min; then 25 cycles of 94 ºC for 40 sec, 52 ºC for 30 sec, and 68 ºC for 3 min; followed by a final extension at 68 ºC for 7 min. Cycling conditions for amplification with the up2/dn primer pair were: one cycle at 94 ºC for five min; five cycles of 94 ºC for 40 sec, 41 ºC for 30 sec, and 68 ºC for 3 min; then 25 cycles of 90 ºC for 40 sec, 54 ºC for 30 sec, and 68 ºC for 3 min; followed by a final extension at 68 ºC for 7 min.

PCR products were gel extracted and purified (BioRad) as described above and cloned into the pVA838 shuttle vector (Fig. 3A) via *Eco*RI restriction sites, creating plasmids pJFP29 and pJFP30 (shown in Fig. 3B and 3C**,** respectively). *E. coli* DH10B was electrotransformed with pJFP29 and pJFP30, and transformants were selected on LA containing Em. Plasmid DNA was extracted (BioRad), and then screened for the insert by *Eco*RI digestion. Sequencing with pVA838 specific primers, pVA838-2939 and pVA838-5331, confirmed correct construct insertion at splicing junctions. The plasmids were introduced into *S. sanguinis* strains by transformation with 250 ng of DNA. Transformants were selected by plating on TSA containing Em. PCR with pVA838-specific primers was used to confirm the presence of appropriate plasmids.

***PCR analysis of 6-26 (pJFP29) and 6-26 (pJFP30)*:** Lenticular colonies of 6-26 were isolated from TSB-LMP agarose tubes by a sterile pipette tip and cultured in 5 ml BHI with Em overnight, under a reduced O_2 condition. An aliquot of the overnight culture was used to prepare a -70 ºC, 30 % glycerol freezer stock, for later analysis by the soft-agar

Figure 3. Plasmid maps of pVA838, pJFP29 and pJFP30

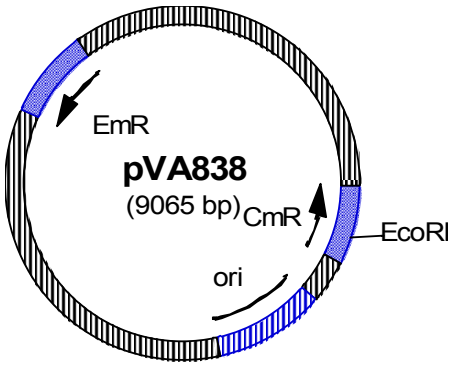
A) The pVA838 cloning vector is capable of replication in *S. sanguinis* and *E. coli*. It confers Em^r to both *E. coli* and streptococci. The Cm^r *cat* marker is expressed only in *E. coli*, and may be inactivated by insertion at the internal *Eco*RI site. This is a low copy number plasmid in *E. coli* with ~8 plasmid copies/chromosomal equivalent in *E. coli* V854 (154). In the described studies this vector was employed for cloning *S. sanguinis* SK36 genes for complementation studies.

B) The *nrdD* gene of *S. sanguinis* SK36 was PCR amplified and cloned into the pVA838 *Eco*RI restriction site to interrupt expression of *cat*, the chloramphenicol resistance determinant gene to create pJFP29. In the insert *nrdD* was preceded by its predicted native upstream terminator, promoter, and ribosome binding site.

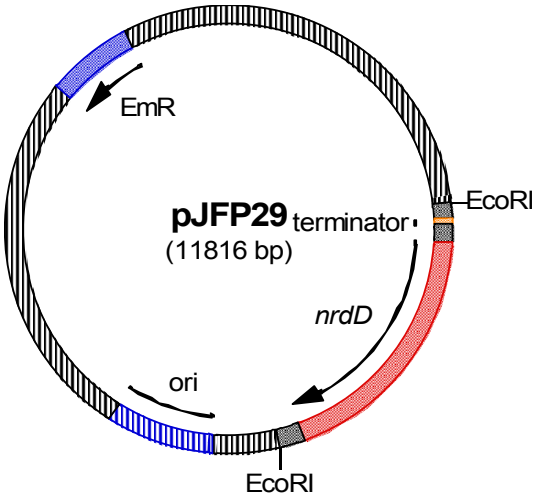
C) The *nrdD* gene of *S. sanguinis* SK36 was PCR amplified and cloned into the pVA838 *Eco*RI restriction site to interrupt expression of *cat*, the chloramphenicol resistance determinant gene, to create pJFP30. In the insert *nrdD* was preceded by its predicted ribosome binding site only.

Figure 3.

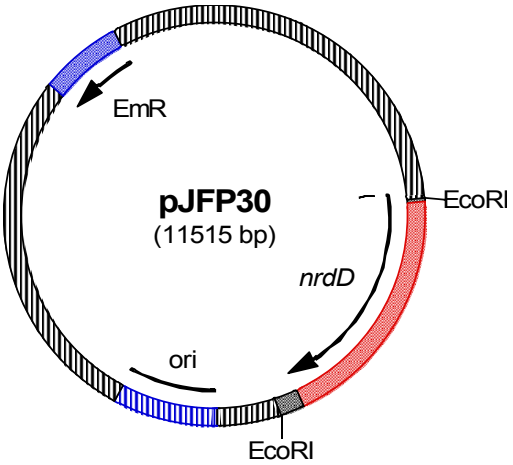
A.


EmR
pVA838
(9065 bp)
CmR
EcoRI
ori

B.


EmR
pJFP29
(11816 bp)
terminator
EcoRI
nrdD
ori
EcoRI

C.


EmR
pJFP30
(11515 bp)
EcoRI
nrdD
ori
EcoRI

assay as described above. The same overnight culture was prepared for PCR amplification by boiled cell extraction (as described above), or washed for use as template: 1 ml of cells were centrifuged at 18,000 x *g* for 1 min; then washed twice in 1 ml of 10 mM Tris-HCl (pH 8.0), and subsequently suspended in 100 µl of 10 mM Tris-HCl. Boiled cell and whole cell preparations were PCR amplified in reactions containing ARTR-SalI-24495-up and ARTR-SalI-20031-dn primers, 3 µl of template, and 45 µl Supermix in a 50 µl reaction. Cycling conditions were as follows: one cycle of 94 ºC for 5 min; 30 cycles of 94 ºC for 40 sec, 58 ºC for 30 sec, and 72 ºC for 6 min; and a final extension cycle at 72 ºC for 7 min.

***Development of JFP36*:** The *S. sanguinis* suicide plasmid pVA2606 (Fig. 2A) was used as a vector for cloning of the *SSA_0169* (*0169*) locus to generate pJFP34 (Fig. 4A). Primers specific for *0169* flanking DNA, 0169-SalI-up and 0169-SalI-dn were used for PCR amplification with *S. sanguinis* SK36 gDNA as template. Cycling conditions were as follows: 1 cycle at 95 ºC for 5 min; 5 cycles of 94 ºC for 45 sec, 45 ºC for 45 sec, and 72 ºC for 1 min and 40 sec; followed by 25 cycles of 94 ºC for 1 min, 59 ºC for 1 min, and 72 ºC for 1 min and 40 sec; and a final extension step at 72 ºC for 7 min.

The resulting PCR products were combined and purified (Qiagen) and digested by *Sal*I (NEB) for cloning into the *Sal*I digested, shrimp alkaline phosphatase (SAP) (Roche) treated pVA2606 vector. Fifty ng of the digested insert and vector were then combined for ligation (Invitrogen). *E. coli* DH10B was electroporated with the resulting plasmid, pJFP34, and transformants were selected on LA containing Kn and X-gal. Plasmid DNA was extracted as described above (Bio-Rad) and integration of the insert determined by *Sal*I digest.

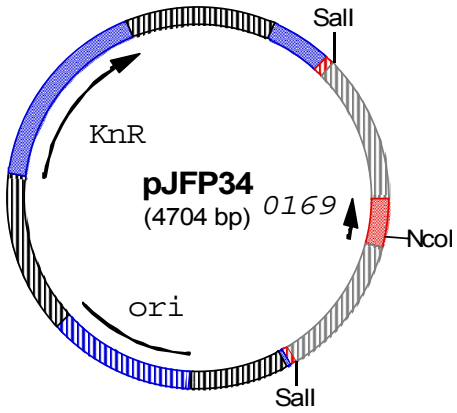
Figure 4. Plasmid maps of pJFP34 and pJFP36

A) *Sal*I restriction sites were incorporated into an amplicon of *S. sanguinis* SK36 by PCR amplification. The PCR product included the *0169* ORF flanked by ~700 bp of upstream and downstream genomic DNA sequence. The amplicon was cloned into the *Sal*I restriction site of pVA2606 to create pJFP34. Blue/white screening identified *E. coli* transformants with the *0169* insert (as a result of insertional inactivation of *lacZα*).

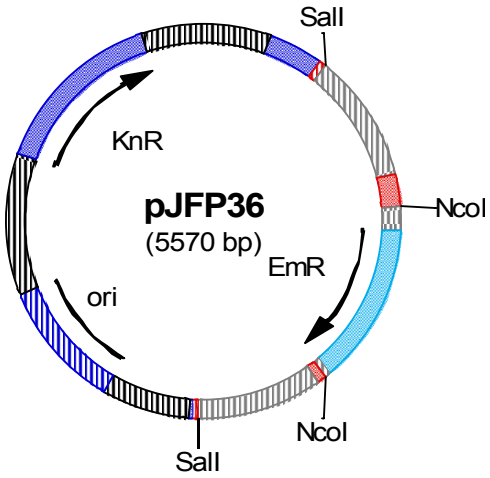
B) A unique *Nco*I site central to *0169* in pJFP34 was adopted for cloning of the erythromycin resistance cassette *pSerm*. The resulting plasmid, pJFP36 was transformed into *S. sanguinis* SK36 for creation of the erythromycin resistant strain, JFP36.

Figure 4.

A. Plasmid map of pJFP34



B. Plasmid map of pJFP36



To develop pJFP36 (Fig. 4B), we adopted the erythromycin resistance cassette, pS*erm,* as a selective marked based on the knowledge that this gene could be expressed in both streptococci and *E. coli*. The *pSerm* construct was kindly provided by Dr. Don Morrison, University of Illinois at Chicago. Primers specific for *p*S*erm*, pSerm-NcoI-up and pSerm-NcoI-dn were used in PCR with Supermix (Invitrogen) for integration of *Nco*I RE sites at the 5' and 3' ends of the *pSerm* construct, with a *pSerm* PCR amplicon as template. PCR cycling conditions were as follows: 1 cycle at 94 ºC for 5 min; 5 cycles at 94 ºC for 30 sec, 40 ºC for 30 sec, and 72 ºC for 1 min; 25 cycles of 94 ºC for 30 sec, 52 ºC for 30 sec, and 72 ºC for 1 min; followed by a final extension step at 72 ºC for 1 min.

The PCR product was purified and digested by *Nco*I for cloning into the pJFP34-encoded *0169* via an *Nco*I restriction site (the *Nco*I site is 39 bp within *0169*). The digested insert and vector were combined at a 1.5 to 1 molar ratio for ligation (Invitrogen). The resulting ligated product was transformed into *E. coli* DH10B, and Em^r, Kn^r clones were selected for further evaluation. Orientation of *0169* and *pSerm* in pJFP36 was determined by *Bss*HII and *Xcm*I digest, respectively. *S. sanguinis* SK36 was transformed with 45 ng pJFP36 and transformants were selected on TSA containing Em. Sequencing with primers 0169-SalI-up, 0169-SalI-dn, Jxn-up, and Jxn-dn confirmed correct construct integration into *S. sanguinis* genomic DNA for creation of JFP36.

***Development of JFP56*:** A linker with *Asc*I and *Not*I sites was integrated into the *Nco*I site of pJFP34 for development of pJFP46 (Fig. 5A). The 0169RE-up and 0169RE-dn oligonucleotides include complementary sequences with, at the 5' end of 0169RE-up an *Nco*I sticky end, followed by a *Not*I site and *Asc*I site. The 0169RE-dn sequence also

includes *Not*I and *Asc*I sites. Complementary base pairing between the two oligonucleotides results in a 5' *Nco*I sticky end overhang for ligation into the *Nco*I site of pJFP34. A base pair change at one end of the oligonucleotide (in bold in Appendix A) abolishes the second *Nco*I site for a series of unique cloning sites including *Asc*I, *Not*I and *Nco*I.

The 0169RE-up and 0169RE-dn oligonucleotides, at 0.5 pmol μl^{-1} were 5' phosphorylated with 0.25 U μl^{-1} T4 DNA Kinase (Invitrogen) in the presence of 0.1 M ATP for 30 min at 37 ºC. The reaction mixtures were then combined and heated at 94 ºC for two min, then put at room temperature for gradual cooling to facilitate annealing between the complementary base pairs of 0169RE-up and 0169RE-dn. The linker was combined in a 6:1 ratio (linker:vector) with *Nco*I digested, SAP treated pJFP34 in the presence of T4 DNA ligase (Invitrogen) for ligation. *E. coli* DH10B was transformed with the ligation product with cells containing the plasmid selected by Kn^r. Clones containing the RE-linker were identified by *Not*I digestion of plasmid DNA. To eliminate the possibility of linker multimerization at the cloning site, plasmids identified as containing the linker were digested with *Not*I, gel electrophoresed, and the correct size linear band extracted from the gel. The band was purified and concentrated as described above (Bio-Rad). Fifteen fmol of the plasmid was then re-ligated in a 20 µl reaction volume. The ligation product was transformed into *E. coli* DH10B cells, and plasmid transformants selected on LA with Kn. Plasmid DNA was extracted from *E. coli* (Bio-Rad), and plasmids containing the linker identified by *Asc*I digestion. Sequence fidelity was confirmed by

sequencing with Jxn-up and Jxn-dn primers. The resulting plasmid, pJFP46 was used for cloning of *tetM* and *aad9* in development of *S. sanguinis* JFP76 and JFP56.

For construction of pJFP56, the *aad9* cassette from pR412 (161) was cloned into the *Not*I and *Nco*I RE sites of pJFP46 (Fig. 5B). The spectinomycin resistance cassette was PCR amplified with Sc-NotI-up and Sc-NcoI-dn primers, with as pR412 template in a reaction containing HIFI Supermix. The PCR cycling conditions used were as follows: one cycle at 95 ºC for 4 min; followed by 30 cycles at 94 ºC for 30 sec, 56 ºC for 30 sec, and 68 ºC for 1.5 min; and a final extension step at 68 ºC for 7 min. The resulting PCR product was purified (Qiagen), and subsequently digested by *Not*I and *Nco*I. The digested PCR product was combined in a 3:1 ratio with the *Not*I and *Nco*I double digested and SAP treated pJFP46 vector for ligation (Invitrogen). *E. coli* DH10B was transformed with the ligation product, and resulting transformants were selected on LA containing Sc. Plasmid DNA was extracted from *E. coli* clones (Qiagen) and digested by *Not*I and *Nco*I for confirmation of the presence of the *aad9* insert in pJFP56. *S. sanguinis* SK36 was transformed with 200 ng pJFP56, and transformants were selected on TSA containing Sc. Sequencing with primers 0169-SalI-F, 0169-SalI-up, Spc638, Spc246, Jxn-up and Jxn-dn validated integration of the correct construct into *S. sanguinis* gDNA.

***Development of JFP76*:** For construction of pJFP76, the *tetM* cassette from pJM133 (88) was cloned into the *Not*I and *Nco*I RE sites of PJFP46 (Fig. 5C). The tetracycline resistance cassette was PCR amplified from pJM133 template by Tet-NotI-up and Tet-NcoI-dn primers, in a reaction containing HIFI Supermix (Invitrogen).

Figure 5. Plasmid maps of pJFP46, pJFP56 and pJFP76

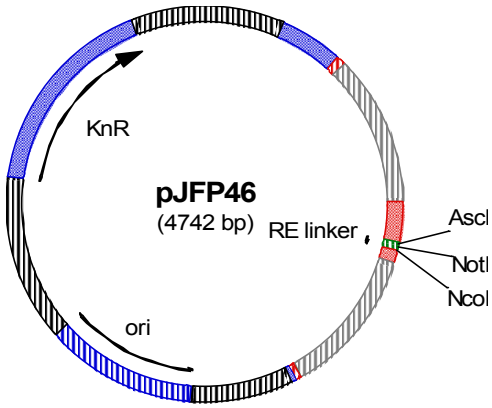
A) The pJFP46 vector was derived from the *E. coli* plasmid pJFP34 by insertion of an *Asc*I, *Not*I, linker in the *0169* insert to create a small, multiple cloning site region for insertion of antibiotic resistance markers.

B) The cloning site linker of pJFP46 was adopted for development of JFP56. The *aad9* gene of pR412 (161) was cloned into the *Not*I and *Nco*I restriction sites, and the resulting plasmid was propagated in *E. coli* prior to transformation into SK36. The *aad9* resistance determinant confers spectinomycin resistance; the resulting SK36 derivative was named JFP56.

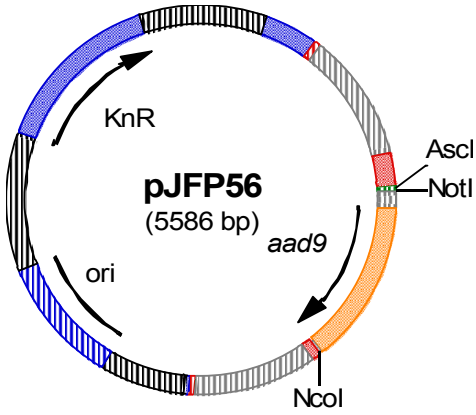
C) The cloning site linker of pJFP46 was also used for development of pJFP76. The *tetM* gene of pJM133 (88) was cloned into the *Not*I and *Nco*I restriction sites, and the resulting plasmid was propagated in *E. coli* DH10B before transformation of SK36 to create a tetracycline resistant derivative of SK36, designated JFP76.

Figure 5.

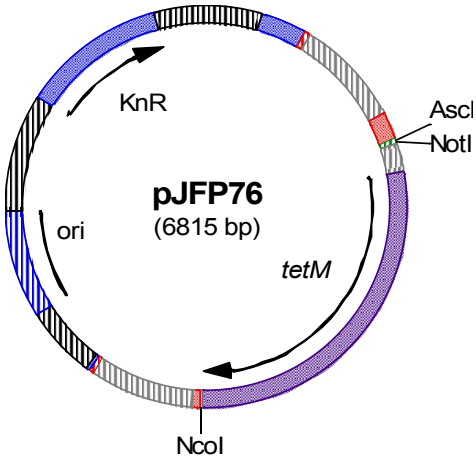
A. Plasmid map of pJFP46



B. Plasmid map of pJFP56



C. Plasmid map of pJFP76



The PCR cycling conditions used were as follows: 1 cycle of 94 ºC for 5 min; followed by 30 cycles of 94 ºC for 40 sec, 45, 49.3 and 50 ºC for 30 sec, and 68 ºC for 2 min and 10 sec; and a final extension step at 68 ºC for 7 min. The resulting PCR product was purified (Qiagen), and subsequently digested by *Nco*I and *Not*I. The digested PCR product was combined in a 3:1 ratio with the *Nco*I and *Not*I double digested and SAP treated pJFP46 vector for ligation (Invitrogen). *E. coli* DH10B was transformed with the ligation product, and resulting transformants were selected on LA containing Tet. Plasmid DNA was extracted from *E. coli* clones (Qiagen) and digested for confirmation of the presence and orientation of the *tetM* insert by *Asc*I and *Hind*III digest, respectively. *S. sanguinis* SK36 was transformed with 200 ng pJFP76 and transformants were selected on TSA containing Tet. Sequencing with primers Jxn-up, Jxn-dn, Tet-NcoI-dn, TetMF1 and TetMF2 validated the correct construct integration into *S. sanguinis* genomic DNA for creation of JFP76.

***Development of pJFP16*:** Plasmid pJFP16 was used for genetic competence comparison of JFP36 and SK36. To construct pJFP16, the mini-transposon containing *cat* and approximately 1.4 kb of flanking DNA was amplified from the STM strain 6-26 using 6-26-BamHI-up and 6-26-BamHI-dn primers, and Supermix (Invitrogen) (198). The PCR product was gel-purified (Bio-Rad) and ligated (Invitrogen) into pVA2606 via *Bam*HI restriction sites. Ligation products were electroporated into *E. coli* DH10B and transformants selected on LA with Kn, Cm and X-gal. Plasmid DNA was extracted as above and orientation of the insert determined by *Eco*RI and *Nde*I digest.

***PCR amplification of ORFs for in vitro transposition*:** PCR was performed in 50 µl reaction volumes with 45 µl of HIFI Supermix (Invitrogen), or 100 µl reaction volumes with 90 µl of HIFI Supermix (Invitrogen), with *S. sanguinis* SK36 gDNA template, and the ORF-specific primer set (Appendix B). The cycling conditions used were as follows: 1 cycle at 94 ºC for five min, followed by 30 cycles of 94 ºC for 40 sec, 53 to 56.5 ºC for 30 sec, and extension at 68 ºC for 3 min. This was followed by a final extension at 68º for 7 min.

***In vitro transposition of PCR amplicons*:** The signature-tagged plasmid library used for *in vitro* transposition was previously described (198). The plasmids were pre-screened by Dr. Todd Kitten to select 40 with signature-tags producing high-signal intensity, and minimal cross-reactivity in dot-blot DNA hybridization. The *in vitro* transposition reaction included 1.2 µg of PCR product template (above), 0.5 µg μl^{-1} of pJFP1-encoded *magellan*2 containing a unique sequence tag (referred to numerically, 1-40), 2 µl of 10 x NEB4 buffer, 100 µg ml^{-1} BSA (New England Biolabs Inc.), and 1µl of 2 mg ml^{-1} MarC9 Tnp in a total volume of 20 µl. Transposition was performed at 30 ºC for 1 h. Immediately following transposition 380 µl 1 x Tris-EDTA (TE) was added, and an the sample extracted with an equal volume of phenol/$CHCl_3$, followed by a second extraction with $CHCl_3$. Organic phase separation was performed in 2 ml Phase-Lock Gel Heavy tubes (5 Prime).

DNA was precipitated from the aqueous phase with the addition of 2 µl oyster glycogen or 1 µl Ethachinmate (Wako), 0.1 volumes of 3 M NaOAc, and 1 ml of EtOH, and suspended in 14 µl of 0.1 x TE. Gaps in the DNA sequence following transposition

were filled by T4 DNA polymerase (Invitrogen). The reaction mixture contained 2 µl of 10 x T4 DNA polymerase buffer, 2 µl of 1 mM dNTPs, 50 µg μl^{-1} of BSA (New England Biolabs), 1 µl of T4 DNA polymerase (1-3 units μl^{-1}), and 14 µl of transposition product. The entire reaction was incubated at 12 ºC for 15 min, and was terminated with the addition of EDTA to 10 mM. The enzyme was then heat inactivated by incubation of the mixture at 75 ºC for 20 min. To remove residual dNTPs and enzyme the transposition products were purified on a column from the MinElute PCR Purification Kit (Qiagen), with DNA eluted from the column in 16 µl 0.1 x TE. Nicks in the DNA phosphate backbone were the sealed by a ligation reaction. For this, 15 µL of eluate was combined with 10 x T4 DNA ligase buffer and 1 µl of T4 DNA ligase (Invitrogen) in a total reaction volume of 20 µl. Ligation was performed at room temperature for 1 h. T4 DNA ligase was subsequently inactivated by incubation of the mixture at 75 ºC for 10 min. Ten µl of the transposition product was then used for transformation of *S. sanguinis* SK36.

***Diagnostic PCR for estimation of transposon location*:** Chloramphenicol resistant clones were selected and inoculated into 2 ml BHI containing Cm, in a 48 well block. The block was covered with an air-pore tape strip and incubated overnight at 6 % O_2, 37 ºC. Once cultures reached turbidity indicative of stationary phase 1 ml was removed and centrifuged in a bench top microcentrifuge at 14k rpm for 1 min. The cells were washed with Tris-HCl (pH 8.0) for use as template in PCR, as described above. For diagnostic PCR primers flanking the ORF were used in conjunction with the *magellan2* inverted repeat primer Mout. For each clone two PCR reactions were performed, one primed by the ORF-up primer and Mout, and a second PCR primed by the ORF-dn primer and Mout. PCR

reactions included 18 µl Platinum Supermix (Invitrogen), 1.2 µl of the washed cells as template, and primer pairs at 0.2 pmol μl^{-1}. The cycling conditions used were: 1 cycle at 94 ºC for 1 min, followed by 30 cycles of 94 ºC for 30 sec, 56 ºC for 30 sec, and 72 ºC for 2 min 30 sec; amplification terminated with a final single extension step at 72 ºC for 7 min.

***Gene SOEing for creation of JFP42 and JFP44*:** Mutagenesis of *srtA* or *srtB* was conducted in a three-step approach. The gene SOEing (105) strategy used is shown in Figure 6. The templates for PCR included *S. sanguinis* SK36 gDNA; pR412 (161); or the kanamycin resistance cassette, *kan*, expressed from a synthetic promoter designed for *S. pneumoniae* (243).

In the first step, antibiotic resistance markers and 5' and 3' regions of the *S. sanguinis* gene were amplified in three separate PCR reactions. pR412 (primer set spcUP (+) and spcDO (-)), or *kan* (primer set DAM303 (+) and DAM347 (-)) was used as template. Cycling conditions for amplification of *aad9* were: 1 cycle of 94 ºC for 5 min, 30 cycles of 94 ºC for 45 sec, 50 ºC for 45 sec, and 68 ºC for 1.5 min, and a final extension cycle at 68 ºC for 7 min. Cycling conditions for amplification of *kan* were 1 cycle of 94 ºC for 5 min, 30 cycles of 94 ºC for 45 sec, 55 ºC for 45 sec, and 68 ºC for 1.3, and a final extension at 68 ºC for 7 min.

For *srtA* and *srtB* constructs the 5' portion of the gene and upstream sequence was amplified with consensus primer 1 (+), and primer 2 (-) designed to include a 5' tail complementary to the antibiotic resistance primer (+). For both constructs the 3' portion of the gene and downstream sequence was amplified with primer 3 (+) which includes a 5' tail complementary to the antibiotic resistance primer (-), and consensus primer 4 (-). The

cycling conditions for srtA1 and srtA2, srtA3 and srtA4 reactions were: 1 cycle at 94 ºC for 5 min, followed by 30 cycles of 94 ºC for 45 sec, 60 ºC for 45 sec, and 68 ºC for 1.75 min, followed by a final extension cycle at 68 ºC for 7 min. The cycling conditions for srtB1 and srtB2, srtB3 and srtB4 reactions were: 1 cycle at 94 ºC for 5 min, followed by 30 cycles of 94 ºC for 45 sec, a gradient between 55 - 65 ºC for 45 sec, and 68 ºC for 1.3 min, followed by a final extension cycle at 68 ºC for 7 min. PCR products from the three reactions were purified from remaining primers and reaction buffer (Qiagen).

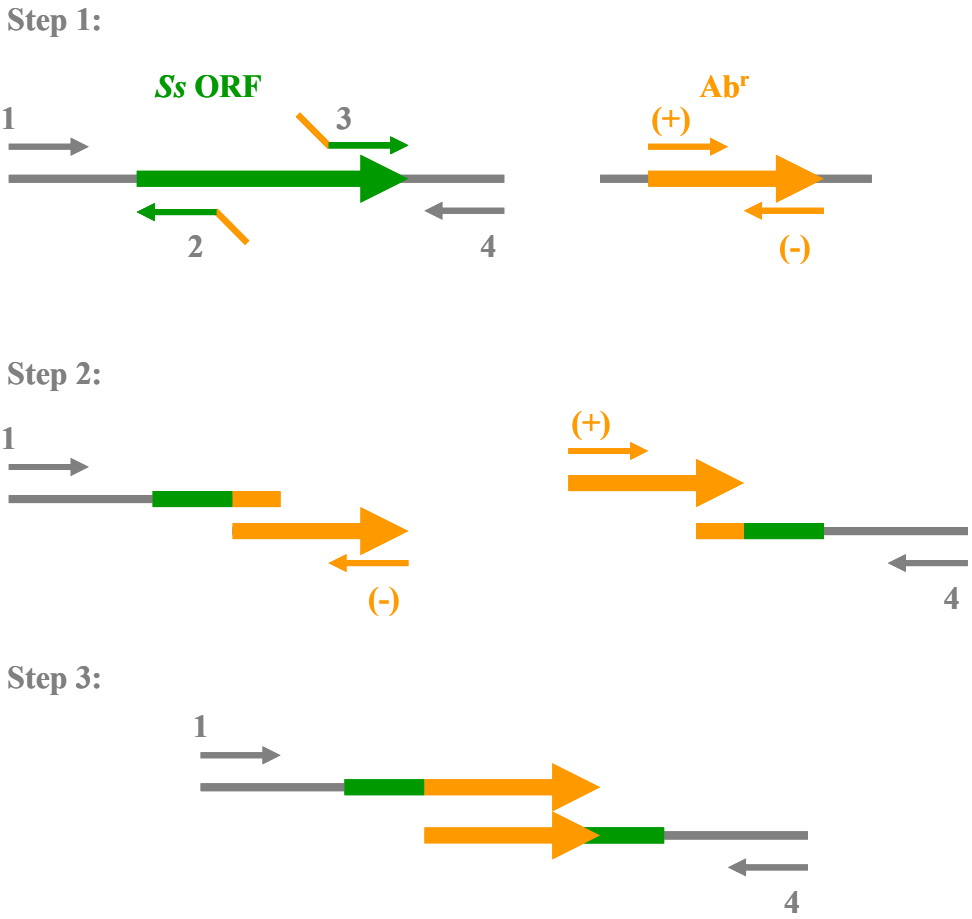
In the second step, 5' or 3' portions of each gene were fused to antibiotic resistance markers in two separate PCR reactions. The cassette *aad9* was fused to the 5' (primer pair srtA1 (+) and spcDO (-)) or 3' (primer pair spcUP (+) and srtA4 (-)) portion of *srtA*. The cassette *kan* was fused to the 5' (primer pair srtB1 (+) and DAM347 (-)) or 3' portion of *srtB* (primer pair DAM303 (+) and srtB4 (-)). The 5' or 3' PCR product and antibiotic cassette PCR product were combined as the PCR template. Amplification conditions for *aad9* fusion were: 1cycle 94 ºC for 45 sec, 30 cycles of 94 ºC for 45 sec, 58 ºC for 45 sec (5':*aad9*) or 54 ºC for 45 sec (*aad9*:3'), and 68 ºC for 2.75 min; followed by a final extension cycle at 68 ºC for 7 min. Amplification conditions for *kan* fusion were: 1 cycle 94 ºC for 45 sec, 30 cycles of 94 ºC for 45 sec, a gradient of 55 to 60 ºC for 45 sec, and 68 ºC for 2.25 min; followed by a final extension at 68 ºC for 7 min. The resulting fusion products were purified as described above.

The third step created a fusion of the entire construct: the 5' portion of each gene, antibiotic resistance cassette, and the 3' portion of each gene. For this the 5':*ab cassette* and 3':*abr cassette* products for each gene were combined as template for PCR. Flanking

Figure 6. Gene SOEing strategy for sortase mutation

The three-step approach used for mutagenesis of *S. sanguinis* (*Ss*) *srtA* and *srtB* is depicted. Through gene SOEing (105), deletion of the majority of the ORF was coupled with insertion of an antibiotic resistance cassette. SOEing was enabled by 5' tails incorporated into primer 2 and primer 3 that were complementary to the (+) and (-) antibiotic resistance (Ab^r) cassette primers, respectively. In the first step the 5' portion and 3' portion of the ORF were amplified, as well as the antibiotic resistance cassette. Amplification of the 5' upstream region of the ORF was primed by pair 1 and 2. Amplification of the 3' downstream region of the ORF was primed by pair 3 and 4. The resistance marker was amplified by the (+) and (-) pair. In the second step the antibiotic resistance cassette was fused to the 5' fragment, or 3' fragment of the gene in a PCR reaction primed by 1 and (-), or (+) and 4, respectively. The resulting products were combined as template in the third, and final amplification reaction primed by the ORF flanking primers 1 and 4.

Figure 6.



consensus primers were used to amplify the entire construct. For *srtA* amplification the primers srtA1 and srtA4 were used, whereas srtB1 and srtB4 were paired for *srtB*. Cycling conditions for *srtA:aad9* fusion were: 1 cycle of 94 ºC for 5 min, followed by 30 cycles of 94 ºC for 40 sec, a gradient of 54, 58, and 63 ºC for 30 sec, and 68 ºC for 3.75 min, followed by a final extension cycle at 68 ºC for 7 min. Cycling conditions for *srtB:kan* fusion were: 1 cycle of 94 ºC for 5 min, followed by 30 cycles of 94 ºC for 40 sec, a gradient of 60, and 65 ºC for 30 sec, and 68 ºC for 3.5 min, followed by a final extension cycle at 68 ºC for 7 min. The complete mutagenic constructs were purified as described above and transformed into *S. sanguinis* SK36. Transformants were selected by antibiotic resistance conferred by the cassette used. DNA sequencing of the *srtA* mutant with primers srtA1, srtA2, srtA3, srtA4, spcDO, spcUP, spc246 and spc638 confirmed sequence fidelity of regions flanking *srtA* in strain JFP42. DNA sequencing of the *srtB* mutant was performed with primers srtB1, srtB2, srtB3, srtB4, DAM303, DAM347, kan306R and kan637F confirmed sequence fidelity in regions flanking *srtB* in strain JFP44.

***Development of JFP47*:** In this complementation strategy, *srtA* downstream of its putative native promoter and RBS, was recombined into the genomic locus used for creation of JFP36, JFP56 and JFP76, *SSA_0169* (Fig. 7B). An extensive gene SOEing (105) strategy was developed for fusion of *srtA* with its predicted native promoter upstream of *gyrA* (Fig. 7A) to the kanamycin resistance cassette *kan*. The fusion construct was flanked by consensus sequences upstream and downstream of *SSA_0169* that would permit homologous recombination with the intended insertion locus. A schematic depicting progressive amplifications is shown in Figure 7C.

Templates for PCR amplification were SK36 gDNA and a *kan* PCR product (243). Successive rounds of amplification resulted in SOEing of the entire construct. Each template was purified (Qiagen) before inclusion in the next PCR reaction. A complementary 5' end to the 0169-2 primer was incorporated in the 0169srtAP fusion primer for SOEing of the upstream *SSA_0169* region to the *srtA* promoter and coding region. A complementary 5' end to the DAM303 primer was incorporated in the kansrtA fusion primer for SOEing of the *kan* resistance marker 3' to *srtA*. A complementary 5' end to the DAM347 primer was incorporated in the kan0169 primer for SOEing of *kan* to the *SSA_0169* downstream region.

Amplification of the *SSA_0169* upstream and downstream regions was performed with the SK36 gDNA template, the 0169-SalI-up [0169-1 in Fig. 7B and 7C, and below] and 0169-2 primer pair, or the kan0169 and 0169-SalI-dn [0169-4 in Fig. 7B and 7C, and below] primer pair, and HIFI Supermix (Invitrogen). The cycling conditions were as follows: 1 cycle of 95 ºC for 5 min; five cycles of 94 ºC for 45 sec, 45 ºC for 45 sec, and 68 ºC for 45 sec; 25 cycles of 94 ºC for 1 min, 56 and 59 ºC for 1 min, and 68 ºC for 45 sec; followed by a final extension cycle of 68 ºC for 7 min. The *kan* cassette was fused to the downstream *SSA_0169* region by combining a previously amplified *kan* PCR product (see construction of JFP44 above) and the kan0169/0169-SalI-dn PCR product as template for amplification with the DAM303 and 0169-SalI-dn primer pair. Cycling conditions were: 1 cycle of 94 ºC for 40 sec; 30 cycles of 94 ºC for 40 sec, 50, 55 and 60 ºC for 30 sec, and 68 ºC for 1.5 min; followed by a final extension cycle at 68 ºC for 7 min.

Figure 7. Construction of JFP47

A) The predicted *srtA* RBS, and *srtA* coding region were fused to the predicted promoter of *srtA* transcription, 5' of *gyrA*. Spacing between the promoter, RBS and start ATG of *gyrA* were conserved, though nucleotides external to the promoter were changed for fidelity to the 24 nucleotides preceding *srtA*.

B) The final PCR amplified fusion construct is shown, along with primer pairs used for sequential PCR amplification in synthesis of the larger amplicon. 5' primer tails enabling gene SOEing are indicated by color coding.

C) The overall PCR strategy for creation of the *srtA*:*kan*:*SSA_0169* fusion is shown. In the first step *kan* and the *SSA_0169* downstream region were PCR amplified in separate reactions primed by the primers shown. The products were combined as template in a SOEing PCR for fusion of the *kan* cassette with the 3' region of *SSA_0169*. A long primer incorporating a 5' tail complementary to *SSA_0169*, the native *gyrA* promoter, and *srtA* sequence was for amplification of *srtA* from SK36 gDNA. The *srtA* amplicon was combined with the *kan*:*SSA_0169* PCR product as template in a reaction for amplification with the primer pairs shown. In a final PCR reaction, the fusion construct was combined with a PCR amplicon of the 5' *SSA_0169* region as template amplified by flanking *0169* primers. The flanking *SSA_0169* ends of the amplicon permitted recombination of the PCR products into the *SSA_0169* locus of JFP42. A Kn^r clone, JFP47 was used in complementation studies.

Figure 7.

A.

Predicted transcription elements of g*yrA* (*SSA_1220*)

Predicted promoter region of *gyrA* — *gyrA* RBS — *gyrA*

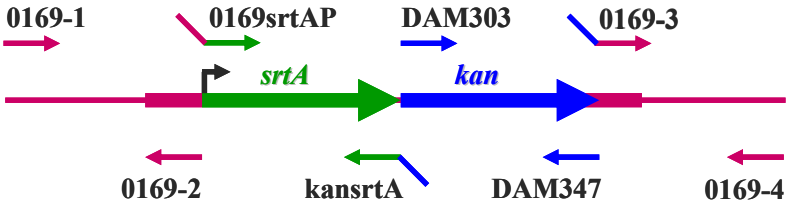
AACGTTTTTACAAACGGTTCAAAATACTGAAAAATTAAAGAAAATGTGGTATAATATAGGATAGTTTAATATG

***gyrA* promoter, *srtA* fusion construct**

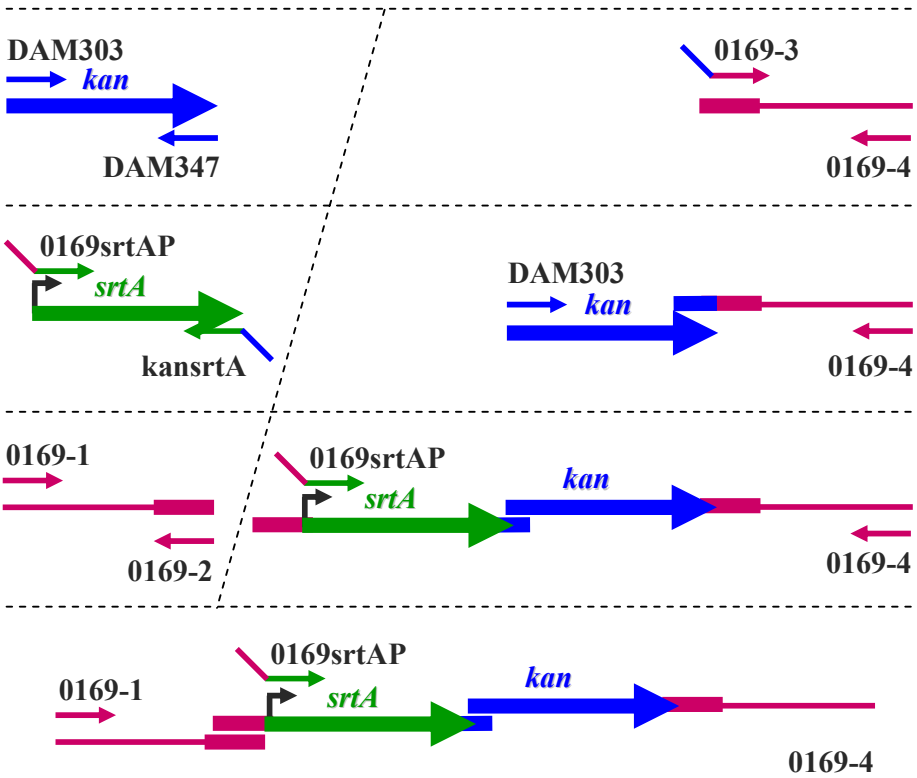
Predicted promoter region of *gyrA* — *srtA* RBS — *srtA*

AACGTTTTTACAAACGGTTCAAAATACTGAAAAATTAAAGAAAATGTCGAATCGTAAAAGAGGACCTCTATATG

B.



C.



Fusion of *srtA* to the *gyrA* promoter was performed using the long fusion primer 0169srtAP and the kansrtA fusion primer with SK36 gDNA as template. The reaction mixture was as described above. Cycling conditions were: 1 cycle at 94 ºC for 5 min; 5 cycles of 94 ºC for 45 sec, 45 ºC for 45 sec, and 68 ºC for 1 min; 25 cycles of 94 ºC for 45 sec, 55 or 60 ºC for 45 sec, and 68 ºC for 7 min; a final extension cycle was performed at 68 ºC for 7 min. The resulting PCR product was fused to the *kan*:*0169* amplicon in a subsequent PCR reaction. The two templates were combined in a reaction as described above, with amplification primed by the 0169srtAP and 0169-SalI-dn primer pair. Cycling conditions were: 1 cycle of 94 ºC for 5 min; 30 cycles of 94 ºC for 45 sec, 55 ºC for 45 sec, and 68 ºC for 2.5 min; followed by a final extension cycle of 68 ºC for 7 min. In a final PCR reaction the entire fusion construct was created by combining the *PsrtA*/*kan*:*0169* amplicon and the upstream *0169* amplicon as template in a reaction primed by the 0169-SalI-up and 0169-SalI-dn primer pair. The templates were combined in reaction as described above. Cycling conditions were: 1 cycle of 94 ºC for 5 min; 30 cycles of 94 ºC for 45 sec, 52, 56.8 or 60 ºC for 45 sec, and 68 ºC for 3.25 min; followed by a final extension at 68 ºC for 7 min.

Ten µl of the PCR product was transformed into competent JFP42 cells (SK36 served as a control for transformation), and transformants were selected on TSA containing Kn. Clones were selected for sequencing of the *SSA_0169* insertion site by primers Jxn-up, Jxn-dn, kan306. Clone JFP47 was selected for further analyses.

***Development of JFP42 (pJFP50)*:** An alternative native promoter of *srtA* was predicted in the 3' sequence of *gyrA*. The 320 bp upstream of *srtA* and the *srtA* coding region were

cloned into the pVA838 shuttle vector. *Eco*RI sites were incorporated into the primer set, srtA-EcoRI-up and srtA-EcoRI-dn, used for PCR amplification of *srtA* from SK36 gDNA. Cycling conditions were as follows: 1 cycle of 94 ºC for 5 min; 30 cycles of 94 ºC for 46 sec, 55 ºC for 45 sec, and 68 ºC for 1.25 min; followed by a final extension cycle at 68 ºC for 7 min. The resulting PCR product was purified (Qiagen), and digested by *Eco*RI for ligation into the *Eco*RI digested, SAP treated pVA838 vector. The digested insert and vector were then combined at a 3 to 1 molar ratio for ligation (Invitrogen). Two µl of the ligation reaction were electroporated into *E. coli* DH10B (Invitrogen). *E. coli* transformants were selected by Em^r on LA plates.

The pVA838 vector encodes both Cm, and Em resistance (Fig. 3A). The *Eco*RI cloning site is within the Cm^r cassette, and therefore clones containing the insert would be rendered Cm^s. Fifty Em^r resistant DH10B clones were replicate plated on LA, followed by LA with Cm at 20 µg μl^{-1} to screen for and recover Cm^s clones. Twelve of the 50 clones were Cm^s. These clones were cultured from the LA plate in LB containing Em, and plasmid DNA prepared (BioRad). The resulting plasmid, termed pJFP50 (Fig. 8) was sent for DNA sequencing with primers srtA-EcoRI-up, pVA838-2939, and pVA838-5331 to confirm *srtA* sequence fidelity. Competent JFP42 was transformed with 200 ng pJFP50, and transformants were selected on TSA containing Em.

***Development of JFP52*:** In an alternative strategy for complementation of JFP42, *srtA* was cloned into the pCM18 *E. coli*-streptococcal shuttle vector (Fig. 9A). *Sph*I RE sites were engineered in the primers, srtA-SphI-up and srtA-SphI-dn, used for PCR amplification of *srtA* from SK36 gDNA. Cycling conditions were: 1 cycle at 94 ºC for 5 min; 30 cycles at

Figure 8. Plasmid map of pJFP50

The 320 bp preceding *srtA* (3' *gyrA*), and the *srtA* coding sequence were cloned into the *Eco*RI restriction site of pVA838. *E. coli* electrotransformants were selected by Em^r, and screened for Cm^s resulting from insertion inactivation of the Cm^r cassette. The plasmid with sequence fidelity designated pJFP50 was transformed into competent JFP42 cells. Em^r transformants were then screened for complementation of *srtA* mutant phenotypes.

Figure 8.

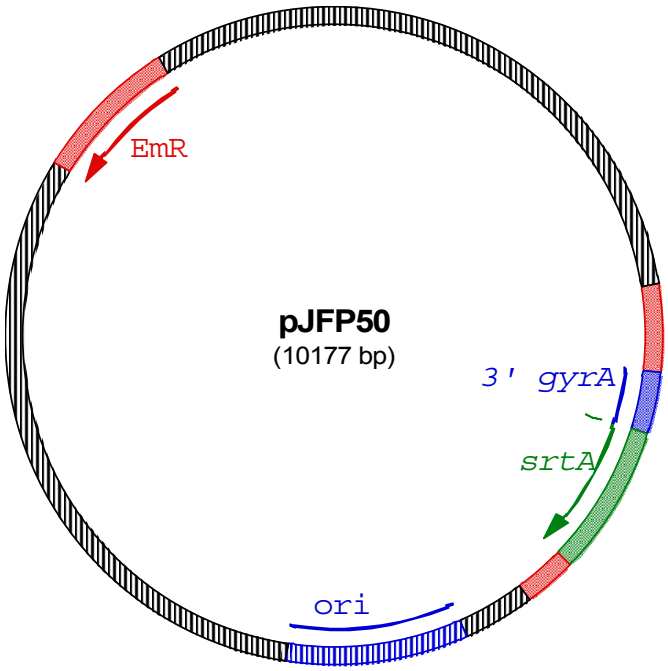


Figure 9. Construction of pJFP52

A) The region of pCM18 used for cloning of *srtA* is shown. The synthetic lactococcal promoter, CP25, and the synthetic ribosomal binding site (RBSII) from pQE70 (Qiagen) drive expression of the *gfp*mut3 green fluorescent protein gene in *E. coli*, as well as *S. gordonii*. Plasmid selection is through an Em^r cassette not shown. This low copy number plasmid (6-9 copies per chromosomal equivalent) exerts a broad host range, owing to its small size (8.2 kb) and the presence of both gram-positive (pAMβ1) and gram-negative (P15A) replicons (90).

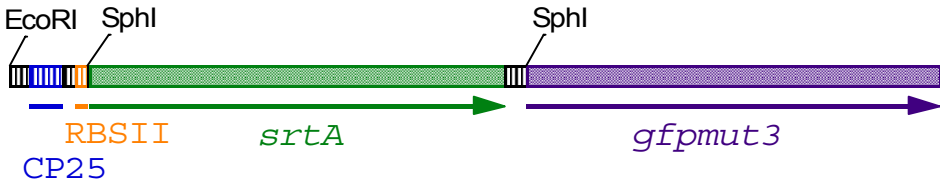
B) The *srtA* gene of *S. sanguinis* was cloned into pCM18 via the *SphI* restriction site occurring in the first codon of *gfp*mut3, thus placing *srtA* under the control of the CP25 promoter. The resulting plasmid, pJFP52 was transformed into *S. sanguinis* JFP42, with Em^r confirming maintenance of the plasmid. Transcription through *srtA* to *gfp*mut3 was evidenced by green fluorescence of JFP42 and SK36 (not shown).

Figure 9.

A.


EcoRI
SphI
RBSII
gfpmut3
CP25

B.


EcoRI
SphI
SphI
RBSII
srtA
gfpmut3
CP25

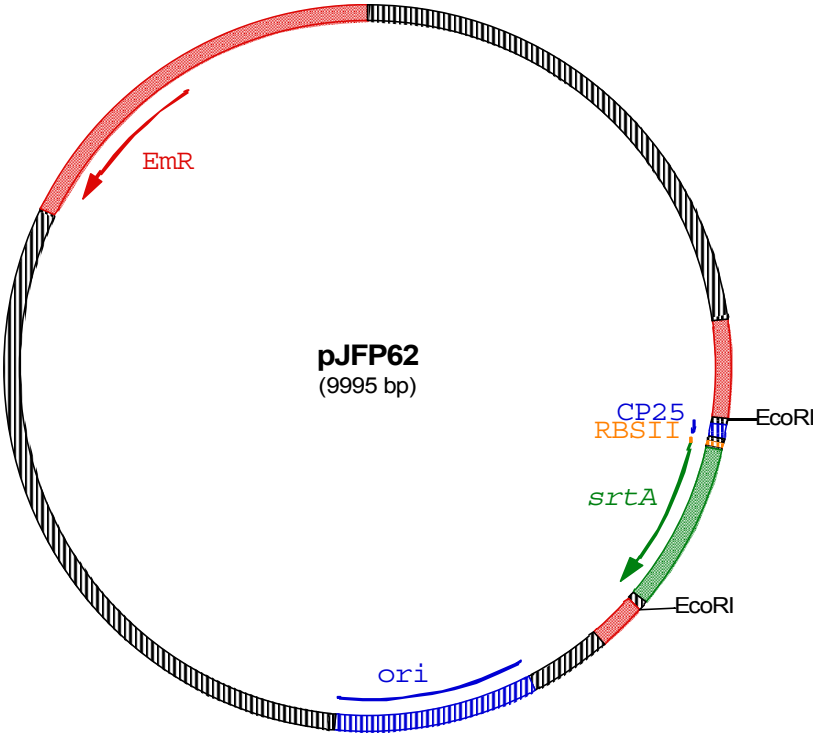
94 ºC for 45 sec, 44 ºC for 45 sec, and 68 ºC for 1 min; followed by a final extension cycle at 68 ºC for 7 min. The PCR product was purified (Qiagen), and then digested with *Sph*I. The digested insert was cloned into *Sph*I linearized, and SAP treated pCM18 (Invitrogen). The resulting ligation product was electroporated into *E. coli* DH10B and transformants were selected by Em^r conferred by pCM18. Plates were viewed on a Dark Reader transilluminator (Clare Chemical Research) for identification of non-fluorescing colonies, which would indicate insertion of *srtA* 5' of *gfp*mut*3*. Plasmid DNA extracted from *E. coli* (Bio-Rad) was digested by *SphI* and *Pvu*I to confirm the presence of *srtA* in pCM18; the plasmid containing the insert was termed pJFP52 (Fig. 9B). Plasmid DNA was sequenced with primers CP25, srtA-SphI-up, and gfpmut. Following confirmation of *srtA* sequence fidelity and proper orientation of *srtA*, 150 ng of plasmid DNA was used for transformation of JFP42. Transformants were selected on TSA containing Em.

***Development of JFP62*:** pCM18 instability in JFP42 resulted in the development of a final complementation construct incorporating features of previous constructs. The *srtA* gene was cloned with the CP25 promoter from pCM18 into the stabile pVA838 shuttle vector (Fig. 3A). The primer set pJFP52srtA-up and pJFP52EcoRI-dn was used for amplification of pJFP52 encoded *srtA*. Cycling conditions were: 1 cycle of 94 ºC for 5 min; 30 cycles of 94 ºC for 40 sec, 50, 56.1 or 60 ºC for 30 sec, and 68 ºC for 1.25 min; followed by a final extension at 68 ºC for 7 min. The PCR product was cloned into the *Eco*RI site of pVA838, and ligation products were electroporated into Electromax DH10B cells as described above. Resulting Em^r clones were screened for Cm^s, as described above. DNA sequencing of the resulting plasmid, pJFP62 (Fig. 10), was completed using primers pVA838-2245,

Figure 10. Plasmid map of pJFP62

The *L. lactis* synthetic promoter CP25, ribosomal binding site (RBSII) and *srtA* gene were cloned from the unstable pCM18 background into the stable pVA838 vector by preexisting or engineered *Eco*RI restriction sites. Integration of the insert at *Eco*RI inactivated the Cm^r marker. This was the basis of screening for the insert in *E. coli*. Plasmid prepared from Em^r, Cm^s *E. coli* was transformed into JFP42. Em^r transformants were further evaluated for complementation of *srtA* mutant phenotypes.

Figure 10.



pVA838-2939, pVA838-5331, and 1219-F. Competent JFP42 cells were transformed with 200 ng pJFP62, and transformants were selected by Em^r. The resulting strain was designated JFP62.

Transformation of *S. sanguinis*.

In vitro transformations of *S. sanguinis* strains were performed as previously described (198). Briefly, strains were cultured in TH-HS to an OD at 660 nm of 0.06 to 0.08. Transforming DNA, 330 µl of cells, and 70 ng synthetic CSP—with the amino acid sequence DLRGVPNPWGWIFGR (79)—were transferred to 0.7-ml microcentrifuge tubes and incubated at 37 °C for 1 h. Cells were plated on TSA with or without selective antibiotics and grown at 37 °C for 48 h under 6 % O_2. The frequency of transformation was defined as the ratio of transformants to total colony-forming units (CFUs). The efficiency of DNA uptake was defined as the number of transformants per µg of transforming DNA.

***In vitro* growth comparisons**

***Soft-agar evaluation of anaerobic growth sensitivity*:** For soft-agar studies (Kurtz 2002) LMP agarose was dissolved in TSB by autoclaving, and then aliquoted in 10 ml volumes into glass Hungate tubes or polystyrene (PS) Falcon tubes (BD). Em, 5 µg $µl^{-1}$ of the oxygen dye indicator resazurin sodium salt (Acros), or 20 µg $µl^{-1}$ of the oxygen indicator, methylene blue were added to select tubes. All tubes were then placed in a 37 ºC incubator inside an anaerobe chamber for 48 h. Tubes were then inoculated with 100 µl of

streptococcal overnight cultures, mixed by inversion, and incubated at ambient temperature in the anaerobe chamber for 30 minutes to permit solidification of the agar. Tubes were then removed from the anaerobe chamber and incubated aerobically with the caps loosened at 37 ºC for 24 h. The depth of growth in cm was measured in each tube with a standard ruler and documented using a digital camera. Values were compared using ANOVA with the Tukey-Kramer multiple-comparisons post-test with $\alpha=0.05$.

***Overnight growth comparisons in a microplate format*:** Stationary phase *S. sanguinis* cultures were diluted 100-fold in BHI, and 200 µl aliquots added to the wells of a 96-well microplate (Greiner). Growth at 37 ºC, at ambient atmosphere was monitored every 10 min over 20 h by measuring the OD at 450 nm. The cells were mixed by orbital shaking (140 rounds/min) with a shaking width of 4 mm for 10 sec before every recording. The experiment was performed in a FLUOstar microplate reader (BMG Technologies).

When noted in the results section a modified procedure was used. Colonies of *S. sanguinis* strains taken from TSA plates containing selective antibiotics were cultured overnight in BHI to stationary phase. Cells were diluted 10-fold in pre-incubated BHI for three additional h of subculture at 37 ºC. Cells were harvested by centrifugation at 3,743 x *g* for 10 min at 4 ºC, and washed once in ice-cold sterile PBS (pH 7.2 – 7.4) before suspension in 7 ml of PBS. The optical density at 660 nm of each strain was adjusted to ~0.8. Cells at this concentration were diluted 200-fold in BHI and 200 µl aliquots monitored in a 96-well plate, as described above.

***In vitro competitive index assays*:** The effect of co-culturing strains *in vitro* was studied by replicating incubation times described above for *in vivo* competitive index (CI) assays

(below). Stationary phase cultures of *S. sanguinis* strains in BHI (sans antibiotics) were diluted 10-fold in fresh BHI and grown for an additional three h at 37 ºC. Cells were washed in PBS as described above, and the optical density of each strain was adjusted to 0.8 at 660 nm. Equal volumes of strains at the same optical density were combined for the "mock" inoculum. From this mixed strain culture, 10 µl was diluted into 10 ml of pre-incubated BHI for growth at 37 ºC, 6 % O_2 overnight. The inoculum was also serially diluted and plated on TSA containing selective antibiotics for enumeration of CFUs for each strain. Following 20.5 – 23 h of co-culture in BHI, cells were serially diluted and plated on TSA containing selective antibiotics. Plates were incubated at 37 ºC, 6 % O_2 for 2 days. The competitive index value was determined by dividing the CFU ratio of the mutant to control after 20 h of co-culture by the CFU ratio of mutant to control inoculated. Significance was determined by a paired t-test of log transformed CI values, with $P < 0.05$ indicating a significant difference in co-culture growth *in vitro*.

Strain evaluation in the rabbit model of IE

A previously described rabbit model of infective endocarditis was adapted for use in this study, with specific-pathogen free New Zealand White rabbits weighing 3 to 3.9 kg as subjects (65, 198). The protocol used was approved by the Institutional Animal Care and Use Committee, and is in compliance with all federal guidelines and institutional policies. To impose pre-disposing valve damage for infective endocarditis, a catheter was inserted along the right carotid artery and past the aortic valve of anesthetized rabbits. The catheter was sutured and remained in place for the duration of the experiment. Two days

post-surgery, streptococci were inoculated into the right lateral ear vein of the rabbits. At pre-determined times rabbits were sacrificed by intravenous injection of euthasol (Virbac AH, Inc.). During necropsy accurate placement of the catheter was confirmed and the heart excised. Apparent vegetations were recovered from the aortic valve and surface of the left ventricle and homogenized in sterile PBS. Multiple vegetation homogenate dilutions were spread on the surface of TSA plates for recovery of bacterial strains.

STM analyses in the rabbit model of IE

***Inoculum preparation*:** Signature-tagged mutants were cultured overnight in a 48-well block in 1.25 ml BHI containing Cm. Cells were mixed by vortexing and pooled into a 50 ml conical tube. From the conical tube, an aliquot was diluted 1:10 into 12.6 ml pre-incubated BHI, and 1.5 ml pooled cells were transferred to 10.5 ml pre-incubated BHIT. The tube containing BHIT was sub-cultured for 2.25 h at 37 ºC, with addition of 0.6 g glycine at 1.5 h post-dilution. Cells in BHIT were then harvested by centrifugation at 3.743k x *g*, 10 min, 4 ºC and the resulting pellet washed in 1 ml cold, sterile dH_2O, and then stored at -20 ºC until DNA extraction was performed (126).

Cells in BHI were subcultured for 3 h at 37 ºC and then harvested by centrifugation at 3.743k x *g*, for 10 min at 4 ºC. The resulting cell pellet was washed in 7 ml cold, sterile PBS (pH 7.2-7.4). The cells were then diluted in PBS to achieve an OD at 660 nm equivalent to ~ 1 x 10^8 cells ml^{-1}. Cells at this concentration were used as the inoculum, with 0.5 ml loaded into 1 ml Leur-lock syringes, fitted with 26 gauge, ½ in. needles. Pre-

catheterized rabbits were injected with the inoculum and infection permitted to proceed to 20 h.

***Amplification of signature tags*:** For each rabbit tested TSA plates containing recovered colonies with 4,500 – 10,000 colonies per plate were suspended in 1 ml BHIT. The suspended colonies were then diluted 60 – 100 fold in fresh BHIT containing Cm for overnight growth. Cells were then diluted 10 fold and sub-cultured for 3 additional hours prior to preparation of chromosomal DNA as previously described (126).

DNA extracted from input pooled strains and output pooled strains was then used as template in PCR with signature tag invariant end primers, PB1 and PB2. PCR components included 1 µg extracted DNA, 90 µl Supermix, and 0.5 pmol μl^{-1} PB1 and PB2 primers in 100 µl total reaction volume. PCR cycling conditions used were: 1 cycle at 95 ºC for 1 min, followed by 25 cycles of 94 ºC for 30 sec, 45 ºC for 1 min, and 72 ºC for 10 sec. PCR terminated in a final extension cycle at 72 ºC for 2 min. PCR products were electrophoresed on 3 % NuSieve agarose gels, purified on Freeze 'N Squeeze columns (BioRad), and ethanol precipitated with a final resuspension of the pellet in 20 µl 0.1x TE. The PCR digoxigenin probe synthesis kit (Roche) was used to DIG-label amplified signature tags in a second round of PCR. The nested primer pair PB11 and PB12, at 0.25 pmol μl^{-1} was combined with 10 x buffer, DIG-labeled dNTPs, unlabeled dNTPs, polymerase and 2 µl of PCR template in a 100 µl reaction volume, as specified by the manufacturer. The PCR cycling protocol was as follows: 1 cycle at 95 ºC for 1 min, followed by 20 cycles of 94 ºC for 30 sec, 37 ºC for 1 min, 72 ºC for 10 sec, and a single final extension at 72 ºC for 2 min. DIG-labeled products were purified (Qiagen), and DNA

eluted in 15 µl 1 x TE was gel electrophoresed on a 3 % NuSieve agarose gel (Cambrex). DNA bands of the input and output amplicons were excised from the gel and prepared for hybridization analysis by the addition of 1 ml of hybridization fluid. The tube was boiled for 5 min to melt the agarose and denature the probe. Hybridization analysis was performed as described below.

***Dot blot hybridization analysis*:** Dot blots were created by transferring 1.1 µg of each STM plasmid (containing tags used) in 22 µM NaOH, and 6X SSC buffer to a MagnaGraph nylon membrane (Micron Separations) with a vacuum manifold (Schleicher & Schuell Bioscience). DNA was then cross-linked to the membrane using a UV Stratalinker (Stratagene). Duplicate membranes were then probed with the input pool and output DNA probes as directed by the Genius chemiluminescent kit protocol (Roche). The resulting chemiluminescent signal was detected by a FluorChem Imager with FluorChem software, V. 3.04A (Alpha Innotech Corporation).

Competitive index analyses in the rabbit model of IE

***Inoculum preparation*:** The *S. sanguinis* colonies from TSA were grown in BHI (without antibiotic unless otherwise specified) overnight. Stationary phase cells of the strains were pooled (unless otherwise indicated) and diluted 1:10 in 14 ml pre-incubated BHI for 3 additional h of sub-culture at 37 ºC. Cells were harvested by centrifugation at 3,743 x *g* for 10 min at 4 ºC, washed once in ice-cold PBS (pH 7.2 – 7.4) and adjusted to an optical density of 0.8 at 660 nm, corresponding to ~5 x 10^7 – 1 x 10^8 CFUs ml^{-1} of each strain.

Five-hundred µl of cells at this optical density were injected into the right lateral ear vein of pre-catheterized rabbits.

When strains were not pooled for co-culture prior to inoculation, they were subcultured separately for 3 h and each strain was subsequently adjusted to an optical density at 660 nm of 0.8, as described above. Equal volumes of strains compared were then combined, and 0.5 ml of the combined strain inoculum injected into previously catheterized rabbits.

***Determination of the CI value*:** The actual number of CFUs rabbits received was determined by plating the inoculum on TSA containing selective antibiotics. Twenty h after infection, rabbits were sacrificed as described above. Vegetations collected from the aortic valve during necropsy were homogenized, diluted and plated. Plates containing the inoculum or vegetation homogenates were grown for 2 days at 37 °C, 6 % O_2. Resulting colonies were enumerated for determination of the competitive index (CI) value, which is defined as the mutant/wild-type ratio of the homogenate divided by the mutant/wild-type ratio in the inoculum. Since colony numbers are often highly skewed log transformed values were used to compute the CI value. Statistical significance was determined by Student's paired t-test, with $P < 0.05$ representing significance.

Previous experiments in our lab revealed that selection via the Cm resistance cassette of STM strains caused reduced colony formation under the normal growth atmosphere of 6 % O_2. A previously described layer-plating technique that is not inhibitory for Cm^r clones was used when comparing STM mutants to JFP36. Briefly, bacterial cells suspended in 12.5 ml TSB with 1 % LMP agarose were poured into a petri dish. The dish

was left at room temperature to allow the agar to harden, and then transferred to a 37 ºC incubator for 2 h. The top layer of 12.5 ml TSA, containing selective antibiotics was then added. As the TSB-LMP agarose layer contains no antibiotic, twice the normal selective antibiotic concentration was added to TSA. The agar was kept at room temperature until completely hardened, and plates were then incubated at 37 ºC, aerobically for 2 days.

In several experiments *S. sanguinis* SK36 derivative strains were compared directly to the parental strain. *S. sanguinis* SK36 lacks a selective marker; therefore, plating in the presence and absence of selective antibiotic (specific for the mutant tested) was used to enumerate mutant CFUs and total CFUs, respectively. SK36 CFUs were determined by subtracting the mutant CFUs from total CFUs.

Individual inoculum analyses in the rabbit model of IE

***Inoculum preparation*:** Colonies of the mutant strain (from TSA with selective antibiotic) or SK36 (from TSA) were inoculated in 5 ml BHI for growth at 37 ºC, 6 % O_2 overnight. Cells of each strain were sub-cultured in pre-incubated BHI, and then adjusted to an optical density at 660 nm of 0.8. Previously catheterized rabbits received 0.5 ml of the single strain inoculum. The inoculum was then diluted and plated on TSA for estimation of the number of bacteria in the inoculum received by rabbits.

***Determination of bacterial recovery*:** Rabbits infected with streptococci were sacrificed at 2 or 20 h post-inoculation. During necropsy surgeons were cautious to only remove material with the characteristic vegetation appearance. Vegetations recovered were weighed, and then homogenized in 2 ml of ice-cold, sterile PBS (pH 7.2 – 7.4). Serial

dilutions of the homogenate were spread on TSA (no antibiotics), and plates were incubated for 2 days at 37 ºC, 6 % O_2. The number of colonies recovered from each rabbit was determined. As colony counts are often skewed the values were log transformed. To account for differences in the amount of vegetation matter collected, the CFU recovered was divided by the gram weight of the homogenized vegetations. Significance was determined by comparing log CFU/g values for each strain in a Welch corrected (equal variance not assumed), unpaired t-test, with $P < 0.5$ indicating a significant difference in bacterial recovery.

qRT-PCR analysis

***RNA extraction for analysis of 0169*:** A stationary phase culture of *S. sanguinis* SK36 was diluted 1:20 into BHI pre-incubated aerobically (225 rpm), under 6 % O_2, or anaerobic conditions (0 % O_2); and grown 5 additional h at 37 °C. RNA was extracted from cells by Trizol (Invitrogen).

***RNA extraction for analysis of srtA*:** Early stationary phase cultures of *S. sanguinis* strains SK36, JFP36, JFP42, and JFP47 in BHI were used for RNA extraction by Trizol (Invitrogen).

***Treatment of RNA for transcript detection*:** RNA samples were purified and double-DNase I treated on RNeasy columns (Qiagen), and RNA quality assessed by the Experion RNA StdSens Chip (Bio-Rad). Reverse transcription was performed with pre-formulated beads from the Ready-To-Go You-Prime First-Strand reverse transcriptase kit (Amersham Biosciences), and primed by the random hexamer primer $pd(N)_6$. For each RNA sample a

control reaction without reverse-transcriptase was performed to monitor for DNA contamination. PCR reactions (10 µl) comprised 5 µl RT^2 Real-Time SYBR Green/Rox PCR master mix (SuperArray), 10 pmol of each forward and reverse primer (detectors are described in Appendix A) and 1 µl template. The cDNA quantity was used for normalization of samples, with a previously described technique (148). Briefly, an aliquot of each RNA sample was combined with RNA hydrolysis solution (1 mM EDTA, 100 mM NaOH) and heated for 20 min at 70 °C before and after reverse transcription. The samples were then neutralized by addition of 6 µl 0.5 M Tris-HCl, pH 6.4. The neutralized samples were mixed with 200-fold diluted RiboGreen (Invitrogen) and fluorescence measured at 485/520 nm excitation/emission wavelengths in a FLUOstar microplate reader (BMG technologies). The correction factor used for normalization was calculated by subtracting fluorescence of samples before reverse transcription from fluorescence values of samples following the reverse transcription reaction. Following quantitative PCR, the transcript abundance value was divided by the correction factor of the corresponding sample. Transcript abundance of each gene was quantified based on standard curves created by serially diluting a PCR amplification product of each gene of interest and using it as template for the real-time reaction. Relative to the standard curve, *n*-fold differences among template quantities for a specific gene could be compared among samples.

Static *in vitro* biofilm formation assay

This assay is based on the ability of streptococci to develop biofilms on solid PS surfaces, and was modified from previously described protocols (190, 76). Early stationary

phase *S. sanguinis* strains were diluted 1:100 in chemically defined biofilm medium (BM) containing 1 % glucose (BM + Glu), or 1 % sucrose (BM + Suc) (151). Aliquots of 0.1 ml were transferred to the wells of a 96-well PS plate for incubation at 37 °C for 20 h at 6 % O_2. Wells with media alone served as negative controls. For biofilm quantification, media and planktonic cells were removed from the wells by decanting and washing twice with dH_2O. Wells were dried and stained with 1 % (wt/vol) crystal violet (CV) for 15 min. Excess dye was removed with three dH_2O washes, and CV staining assessed via release of biofilm-associated dye into 30 % glacial acetic acid. The OD at 600 nm of soluble CV was recorded. Significance was determined by one-way ANOVA with a Tukey-Kramer multiple comparisons post-test, with significance represented by $P < 0.05$.

Polystyrene adhesion screening

The adhesion assay used was modified from previously described PS binding analyses for simultaneous screening of multiple clones. Colonies from transformation plates were inoculated into 1 ml BHI containing the appropriate antibiotics and grown overnight under 6 % O_2, at 37 ºC. Cells were diluted 10-fold into pre-incubated BHI and sub-cultured at 37 ºC for 3 additional h to achieve early stationary phase growth. Cells were then diluted 50-fold into pre-incubated BM + Glu, and 200 µl transferred to the wells of a 96-well microplate (Greiner) in duplicate. Plates were incubated for 4 h at 37 ºC to permit adhesion of bacteria to the PS surface. Non-adherent cells were removed by decanting and a single wash with dH_2O. Adherent cells were stained with 25 µl of 1 % CV for 15 min. Excess dye was removed by three washes with dH_2O, and bound dye released

by the addition of 30 % (v/v) glacial acetic acid. One-hundred and twenty-five µl of the dye was transferred to a new 96-well Greiner plate, and the OD at 660 nm was recorded.

Production of CwFrac antisera

Mutanolysin digestion was used to extract cell wall-associated proteins of SK36 for generation of antisera reactive to the cell surface peptidoglycan-linked proteins. The cell wall fraction of SK36 was prepared from 400 ml of early stationary phase SK36. Cells were harvested by centrifugation at 5k x *g* for 10 min at 4 ºC, and washed twice in 20 ml TE + protease inhibitor cocktail (Pi) (Sigma) (1:100). Each pellet was vigorously suspended in 2.4 ml protoplasting buffer (2 M Tris, 1 mM $MgCl_2$, 0.44 M raffinose) (SEB) + Pi (1:100) using a vortex, and then inverted with 120 µl of mutanolysin (Sigma) (10,000 U ml^{-1}). Tubes were incubated at 37 ºC for 30 min and then centrifuged at 5k x *g* for 20 min at 4 ºC. The supernatant containing mutanolysin released cell wall fragments was collected. This protein fraction was twice dialyzed against 1 L PBS at 4 ºC in Snakeskin tubing (22 mm x 35 ft, 7000 Mw) (Pierce). Protein concentration was then determined by the BCA Protein Assay Kit (Pierce). The cell wall fraction (CwFrac) immunogen was then sent to Covance for production of rabbit polyclonal antisera.

Whole cell ELISA

The procedure used for cell fixation was previously described (103, 189). *S. sanguinis* strains grown overnight in BHI with selective antibiotics were sub-cultured for 3 h in BHI alone. Cells were harvested and adjusted to an optical density at 660 nm of ~0.9

in TNMC (1 mM Tris-HCl [pH 8.0] containing 0.15 M NaCl, 1 mM $MgCl_2$, and 1 mM $CaCl_2$) buffer. Fifty µl of the adjusted cells were transferred to the wells of a 96-well 4HBX Immunolon plate (Thermo electron); 1 mg ml^{-1} BSA (Pierce) was included in several wells as a control for non-specific reactivity. Plates were centrifuged at 800 rpm, for 5 min at 20 ºC. To fix the cells, 0.1 ml TNMC with 0.25 % glutaraldehyde was added to each well, and the plates centrifuged at 800 rpm for 5 min, at 4 ºC. The plates were then incubated at 20 ºC for 20 min. Excess liquid was aspirated from the wells and the cells were then washed twice with 100 µl of blocking buffer (TBS containing 0.2 % Tween-20 and 2.5 % BSA), prior incubation with 100 µl blocking buffer overnight at 4 ºC.

***Detection of SrpA*:** A SrpA mAb provided by Dr. Hui Wu (University of Alabama) was used for detection of SrpA on the surface of whole cells. Following incubation in blocking buffer overnight, 0.05 ml of 1:50 diluted α-SrpA mAb in a blocking buffer was added to each well and plates were incubated at room temperature for 2.5 h. Unassociated mAb was removed by three washes with 0.1 ml TBS-T. The secondary antibody, α-mouse IgG conjugated to alkaline phosphatase (AP) (1:1,000 in blocking buffer) was added to the wells in 0.05 ml volumes. Plates were incubated for 2.5 h, and unassociated secondary antibody removed by three washes in 0.1 ml TBS-T. One-hundred µl of the alkaline phosphatase substrate, p-NPP (KPL) was added to each well. The reaction proceeded for three minutes before addition of 0.1 ml 5 % EDTA. Plates were read at 410 nm on a FLUOStar plate reader. Significance was determined by a Kruskal-Wallis test (nonparametric ANOVA), with $P < 0.05$ representing a significant difference.

***Detection of whole cells by CwFrac*:** Antisera prepared against the cell wall fraction of *S. sanguinis* SK36 (referred to as CwFrac below) was used to demonstrate equal amounts of whole cells following glutaraldehyde fixation. Following incubation with blocking buffer overnight, 0.05 ml of diluted CwFrac in blocking buffer was added to each well. Plates were incubated at room temperature for 2.5 h, and unassociated antibody removed by three washes with 0.1 ml TBS-T. The α-Rabbit IgG AP conjugated secondary antibody, diluted 1:2,000 in blocking buffer was added to the wells in 0.05 ml volumes, and plates were incubated at room temperature for 2.5 h. Following three washes with TBS-T, CwFrac association was detected by p-NPP, as described above.

Two separate protocols were used to demonstrate the specificity and sensitivity of CwFrac for detection of SK36. To determine the antibody dilution permitting specific, non-saturating reactivity with whole cells, serially diluted antisera (1:5,000, 1:10,000, 1:50,000, 1:100,000, and 1:200,000) was reacted against fixed whole cells of SK36 (as well as other strains tested). The optimal dilution of 1:25,000 was then used to demonstrate differential reactivity of the sera reflective of various amounts of cells present in the wells. The optical density of each strain tested was adjusted to ~0.9 at 660 nm, as described above. After adjustment, cells at this optical density were serially diluted across a 96-well plate so that each serial dilution occurred in triplicate. The range of dilutions included undiluted, 2-fold, 4-fold, 8-fold, 16-fold and 32-fold. BSA at 1 mg ml^{-1} served as a negative control for nonspecific reactivity of the immune sera. An identical assay was used with pre-immune sera collected from rabbits prior to challenge with CwFrac. Reactivity to sera was determined in a chromophore assay, as described above.

Fractionation of *S. sanguinis*

S. sanguinis strains cultured overnight in 20 ml BHI were centrifuged at 5k x *g*, for 10 min at 4 ºC. The cell pellet and supernatant were then treated to obtain a supernatant protein fraction, cell wall-associated protein fraction, and a protoplast protein fraction.

To prepare the supernatant fraction, 18 ml of supernatant was collected following centrifugation and then processed for precipitation of secreted proteins. Supernatants were filtered across a 0.2 µm membrane, and 10 % (v/v) TCA was added. Tubes were mixed by vortexing and incubated on ice for 1 h. The mixture was centrifuged at 9.5k x *g* for 15 min at 4 ºC for extraction of TCA precipitated proteins. The resulting pellet was washed twice in 1 ml -20 ºC acetone, and dried in an oven for 5 min to drive off residual acetone. The pellet was resuspended in 0.2 ml PBS + Pi (10:1 (v/v)).

For preparation of the cell wall fraction the cell pellet was washed twice in 1 ml TE Buffer + Pi (10:1 (v/v)) before suspension by vortexing in 0.24 ml SEB + Pi (10:1 (v/v)). The cell suspension was then mixed with 12 µL mutanolysin (Sigma) by inversion and incubated at 37 ºC for 1 h. The suspension was centrifuged at 5k x *g* for 15 min at 4 ºC for collection of the supernatant containing proteins released by mutanolysin.

The remaining protoplast pellet was washed once in 0.5 ml SEB + Pi, and then resuspended in 1 ml PBS + Pi. Protoplasts were lysed by sonication at 80 % for 10 min, followed by bead beating in Lysing Matrix B tubes (BIO 101). Following centrifugation at 15k x *g* for 15 min at 4 ºC, the supernatant containing released proteins was collected to represent the protoplast fraction.

SrpA Immunoblot analysis

Bacterial cell fractions were subjected to SDS-PAGE in 5 % precast Tris-HCl gels (Criterion, Bio-Rad). For immunoblot analysis, proteins were transferred from gels onto a PVDF membrane (Bio-Rad) by electroblotting. Immunoblots were blocked for 1 h in a blocking buffer (TBS containing 0.2 % Tween-20 and 1 % BSA). After blocking, the immunoblot was incubated overnight at 4 ºC with 1:50 diluted α-SrpA primary antibody. AP-conjugated α-mouse IgG (Pierce), diluted 1:7,500 served as the secondary antibody. The immunoblot was incubated with the secondary antibody for 1 h at room temperature. The immunoblot was washed with TBS-T after each incubation step for removal of unbound antibody. Detection of bound AP antibody was by the Western-blue alkaline phosphatase substrate (Pierce). General protein staining of duplicate Tris-HCl gels was performed with SYPRO Ruby (Invitrogen), per the manufacturer's instructions. Protein bands were visualized by UV light in a Fluorchem Imager with Fluorchem software, v.3.04A (Alpha Innotech Corporation).

Hexadecane phase partitioning

Stationary phase overnight cultures were diluted 10-fold into 12.6 ml pre-incubated BHI and sub-cultured for 3 h art 37 ºC. Cells were harvested and washed in cold, sterile 150 mM NaCl once prior to suspension in 8 ml 150 mM NaCl. Cells were then evenly divided between 4 snap-top conical tubes. Two tubes served as non-hexadecane controls. The remaining tubes were supplemented with 333 µl of *n*-hexadecane, followed by

immediate vortexing for 1 min. Afterward cells were incubated at room temperature, undisturbed for 25 min. The optical density at 660 nm was recorded for the aqueous phase of each strain. Duplicate reads in the presence and absence of hexadecane were averaged and the percent affinity for hexadecane of each strain determined by the following equation: % Affinity = 100 x [1- (A/A_0)], where A = the absorbance of cells with *n*-hexadecane, and A_0 = the absorbance of cells without *n*-hexadecane.

Bioinformatics

Oligonucleotide primers used were designed using Oligo (Molecular Biology Insights) and FastPCR (www.biocenter.helsinki.fi/bi/bare-1_html/fastpcr.htm) software. DNA sequences were viewed and aligned using SeqMan II software (DNASTAR Inc). Multiple sequence comparisons were made using ClustalW W (202). Transcriptional terminators were predicted using the Genetics Computer Group program Terminator (29, 30). Plasmid maps were created using Gene Construction Kit® v3.0 (Textco BioSoftware).

***Search algorithms for ORFs encoding putative Cwa proteins*:** Open reading frames (ORFs) predicted by Glimmer (54) from the compiled and contiguous *S. sanguinis* SK36 genome sequence (GenBank accession number CP000387) were searched for proteins containing the characteristic C-terminal LPXTG tripartite motif of Cwa proteins. This cell wall sorting signal included an LPXTG sortase substrate pattern (in which X may be any amino acid), followed by a hydrophobic domain, and at least one positively charged amino acid within 8 residues of the C-terminus. The EMBOSS (http://www.emboss.org) program fuzzpro uses predictive patterns to search protein sequences. The query pattern for

identification of putative *S. sanguinis* Cwa proteins was modified from that published by Janulczyk and Rasmussen: L-P-[SKTAQEHLDN]-[TA]-[GN]-[EDASTV]- X(0,8)-[VIFAGTSML](17)- [RK]-[RK](2), for prediction by fuzzpro of proteins containing the motif terminating within 5 amino acids of the C-terminus (111). Two query patterns were used:

(A): L-P-[SKTAQEHLDN]-[TA]-[GN]-[EDASTV]-X(0,8)-[VIFAGTSML](17)-[RK]-[RK](2)-X(0,5)>

(B): L-P-[SKTAQEHLDN]-[TA]-[GN]-[EDASTV]-X(0,8)-[VIFAGTSML](16,20)-[RK]-[RK](2)-X(0,5)>

In these patterns mandatory residues are distinguished form alternative residues by the bracketing of alternative residues. The gap designated by X (0,8) may be from 0 to 8 amino acids. Parenthetical numbers indicate repetition of the preceding residue. The applied patterns differ in the allowed number of hydrophobic residue repeats [VIFAGTSML] following LPXTG. As more repeats are permitted with pattern (B), greater stringency among hits was enforced by decreasing the number of permitted mismatches from 5 with pattern (A), to 4 with pattern (B).

The SignalP 3.0 server was then used to scan fuzzpro positive ORFs for the presence of a gram-positive signal sequence (15) within 70 amino acid residues of the N-terminus, with both the Neural networks (NN) (D-measure cutoff of 0.45) and Hidden Markov Models (HMM) applied. Proteins positive for both an N-terminal signal peptide by SignalP and a C-terminal tripartite motif were then scanned visually to confirm the presence of an LPXTG pattern. Erroneous positive hits trended toward multiple-pass

integral membrane proteins as suggested by transmembrane helices prediction performed with the TMHMM 2.0 server (http://www.cbs.dtu.dk/services/TMHMM/). For comparison identical searches with the described parameters were performed for ORFs predicted from complete genome sequences of *S. gordonii* strain Challis substrain CH1 (GenBank accession CP000725), *S. mutans* UA159 (GenBank accession AE014133), and the incomplete genome of *S. mitis* NCTC 12261, v11.0 (http://www.tigr.org/).

***Bioinformatics survey of predicted S. sanguinis Cwa proteins*:** Putative *S. sanguinis* Cwa ORFs were annotated by a protein basic local alignment search tool (BLASTP) query of the microbial genome database, with an expect-value (positive hits by chance) cutoff of 1.0. To elucidate similarities and differences between oral streptococcal Cwa proteins, the predicted proteins of *S. gordonii* CH1, *S. mitis* NCTC 12261, and *S. mutans* UA159 were used as queries to search for homologs in the *S. sanguinis* SK36 protein database (http://www.sanguinis.mic.vcu.edu/blast.html).

Cwa proteins unique to *S. sanguinis* SK36 among related streptococcal species, unrelated oral bacterial species, and additional agents of infective endocarditis were identified using the Phylogenetic Profiler application of the Integrated Microbial Genomes (IMG) resources offered by the DOE Joint Genome Institute (http://www.jgi.doe.gov/), with an expect-value cutoff of $1e^{-10}$. Genomic information available through IMG for organisms in (Table 2) was used for identification of unique ORFs of *S. sanguinis* SK36.

Table 2. Subtractive comparison for identification of unique *S. sanguinis* ORFs

The phylogenetic profiler application of IMG, offered by DOE was used to eliminate homologous genes shared by *S. sanguinis* and 54 streptococcal spp., gram-positive pathogens, oral microflora, as well as causative agents of infective endocarditis. The e-value cut-off for extraction of proteins with homologs was $1e^{-10}$. Of the 2272 genes predicted for *S. sanguinis*, 246 did not have significant homologs among genomes compared. Eight Cwa protein genes were identified as being unique to *S. sanguinis*.

Table 2.

Order	Species for comparison	Genes remaining
	S. sanguinis genes prior to comparison	2272
1	*Clostridium perfringens* 13	1399
2	*Corynebacterium diphtheriae* NCTC 13129	1295
3	*Corynebacterium efficiens* YS-314	1258
4	*Corynebacterium jeikeium* K411	1246
5	*Enterococcus faecalis* V583	945
6	*Fusobacterium nucleatum* ssp. *nucleatum* ATCC 25586	900
7	*Lactococcus lactis* ssp. *lactis* Il1403	803
8	*Porphyromonas gingivalis* W83	800
9	*Staphylococcus aureus* RF122	760
10	*Staphylococcus aureus* ssp. *aureus* COL	756
11	*Staphylococcus aureus* ssp. *aureus* MRSA252	752
12	*Staphylococcus aureus* ssp. *aureus* MSSA476	752
13	*Staphylococcus aureus* ssp. *aureus* MW2	752
14	*Staphylococcus aureus* ssp. *aureus* Mu50	752
15	*Staphylococcus aureus* ssp. *aureus* N315	752
16	*Staphylococcus aureus* ssp. *aureus* NCTC 8325	752
17	*Staphylococcus aureus* ssp. *aureus* USA300	751
18	*Staphylococcus epidermidis* ATCC 12228	748
19	*Staphylococcus epidermidis* RP62A	744
20	*Staphylococcus haemolyticus* JCSC1435	739
21	*Staphylococcus saprophyticus* ssp. *saprophyticus* ATCC 15305	733
22	*Streptococcus agalactiae* NEM316	603
23	*Streptococcus mutans* UA159	514
24	*Streptococcus pneumoniae* R6	453
25	*Streptococcus pneumoniae* TIGR4	444
26	*Streptococcus pyogenes* M1 GAS	432
27	*Streptococcus thermophilus* CNRZ1066	412
28	*Streptococcus thermophilus* LMG 18311	407
29	*Treponema denticola* ATCC 35405	399
30	*Treponema pallidum* ssp. *pallidum* Nichols	399
31	*Enterococcus faecium* DO	397
32	*Fusobacterium nucleatum* ssp. *vincentii* ATCC 49256	395

Order	Species for comparison	Genes remaining
33	*Streptococcus suis* 89/1591	384
34	*Corynebacterium glutamicum* ATCC 13032 (Bielefeld)	382
35	*Corynebacterium glutamicum* ATCC 13032 (Kitasato)	382
36	*Lactococcus lactis* ssp. *cremoris* SK11	382
37	*Streptococcus thermophilus* LMD-9	379
38	*Lactococcus lactis* ssp. *cremoris* MG1363	378
39	*Streptococcus pyogenes* Manfredo	378
40	*Corynebacterium glutamicum* R	378
41	*Staphylococcus aureus* ssp. *aureus* JH9	378
42	*Streptococcus suis* 05ZYH33	376
43	*Streptococcus suis* 98HAH33	376
44	*Campylobacter concisus* 13826	374
45	*Campylobacter curvus* 525.92	373
46	*Staphylococcus aureus* ssp. *aureus* JH1	373
47	*Staphylococcus aureus* ssp. *aureus* Mu3	373
48	*Staphylococcus aureus* ssp. *aureus* Newman	373
49	*Streptococcus gordonii* Challis CH1	249
50	*Eubacterium ventriosum* ATCC 27560	247
51	*Actinomyces odontolyticus* ATCC 17982	246
52	*Staphylococcus aureus* ssp. *aureus* USA300_TCH1516	246
53	*Peptostreptococcus micros* ATCC 33270	246
54	*Eubacterium dolichum* DSM 3991	246

Putatively unique *S. sanguinis* SK36 Cwa proteins
SSA_1663
SSA_1632
SSA_1633
SSA_1634
SSA_0146
SSA_0167
SSA_0273
SSA_1023

***Identification of S. sanguinis SK36 sortase homologs*:** *S. sanguinis* SK36 sortase-like ORFs with homology to published sortase sequences were found by BLASTP analysis of predicted proteins of the v.15 contig of *S. sanguinis* SK36. The following forty-five publicly available sortase sequences (www.ncbi.nlm.gov) were queried against predicted ORFs of the nearly complete SK36 genome sequence: GI: 1169933 (*Shewanella oneidensis* MR-1); GI: 3238307 (*S. aureus* ssp. *aureus* COL); GI: 2859503 (*S. aureus* ssp. *aureus* MRSA252); GI: 936904 and GI: 939748 (*Bacillus subtilis* ssp. *subtilis* str. 168); GI: 1201902 and GI: 1201386 (*Enterococcus faecalis* V58); GI: 1199996 (*E. faecalis* V583); GI: 1028429 (*S. mutans* UA159); GI: 2650068, GI: 2650116; GI: 2650498, GI: 2650495, GI: 2650024, and GI: 2650023 (*Corynebacterium diphtheriae* NCTC 13129); GI: 1088124 and GI: 1083901 (*Bacillus anthracis* str. Ames); GI: 2851511 (*B. anthracis* str. Sterne); GI: 1116387 (*Clostridium acetobutylicum* ATCC 824); GI: 111657852, GI: 930404, GI: 930406, and GI: 930405 (*S. pneumoniae* TIGR4); GID: 4962816 (*Streptococcus. pyogenes* str. Manfredo); GI: 901269, GI: 900461, and GI: 900465 (*S. pyogenes* M1 GAS); GI: 891152, GI: 890596, GI: 893453, GI: 893835, GI: 892492, and GI: 892801 (*Bacillus halodurans* C-125); GI: 4916727 (*Clostridium difficile* 630); GI: 3036999 (*Actinomyces naeslundii* T14V); GI: 1099286, GI: 1099285, GI: 1099173, GI: 1102888, GI: 1097914, GI: 1098275, and GI: 1096358 (*Streptomyces coelicolor* A3(2)); GI: 1470914, GI: 1471438 (*Methanothermobacter thermautotrophicus* str. Delta H). *S. sanguinis* SK36 ORFs homologous to query protein sequences were recorded, and identity to one of three sortases classes (A, B, or C) was used for annotation of *S. sanguinis* sortase-like proteins (59). Tmap (EMBOSS) analysis of *S. sanguinis* sortase-like proteins

indicated hydrophobic transmembrane domains directing both extra-cytoplasmic export of the precursor protein, and transmembrane topology of the mature protein (http://www.emboss.org).

***MudPIT detection of Cwa proteins*:** *S. sanguinis* cell wall-associated protein and protoplast fractions were prepared as described above. The protein fractions were subjected to LC/LC/MS–referred to below as MudPIT–with the assistance of Dr. Patricio Manque (Virginia Commonwealth University). Protein samples were suspended in solubilization buffer (7 M urea, 2 M thiourea, 1 % DTT, 2 % Triton X-100, with protease inhibitors; 100 μM PMSF, 100 μM TLCK, 1 μM pepstatin A, 100 μM leupeptine, 5mM EDTA, 50 μM TPCK) and precipitated with a 2D clean-up kit (GE Healthcare). Samples were then reduced by the addition of dithiothreitol, alkylated with iodoacetamide, and digested by trypsin overnight. Resulting peptides were desalted using C8 cartridges (Michrom BioResources) and then subjected to two-dimensional liquid chromatography-MS/MS analyses with a Michrom BioResources Paradigm MS4 multidimensional separation module, a Michrom NanoTrap platform, and an LCQ DecA XP Plus ion trap mass spectrometer. The mass spectrometer was operated in the data dependent mode with the four most abundant peptides selected and fragmented to create tandem mass spectra. Resulting MS/MS spectra were recorded in the profile mode and proteins identified by searching the MS/MS spectra against the *S. sanguinis* SK36 database using Bioworks v3.2. The probability based scoring algorithm of Bioworks v3.2 was used for scoring of peptide and protein hits. Cutoffs for protein identification included only peptides with fully tryptic

termini and cross correlation scores exceeding 1.9 for singly charged peptides, 2.3 for doubly charged peptides, and 3.75 for triply charged peptides.

CHARACTERIZATION OF *S. SANGUINIS* ANAEROBIC GROWTH DEPENDENCY ON THE ANAEROBIC RIBONUCLEOTIDE REDUCTASE

Introduction

Identification of NrdD

In a previously published screen for *S. sanguinis* endocarditis virulence determinants by random STM an apparent avirulent mutant was identified in which the open reading frame *SSA_2230* was interrupted. A competitive index assay was also used to establish a quantitative value for competitiveness of the mutant, 6-26 relative to the parental strain, SK36 in the rabbit model of IE. The mean CI value derived for 6-26, 0.31, confirmed that insertion mutagenesis of *SSA_2230* resulted in a significant decrease in competitiveness *in vivo* ($P < 0.05$) (198). SSA_2230 shares 84 % amino acid sequence similarity to the characterized anaerobic ribonucleotide triphosphate reductase (ARNR), NrdD of *Lactococcus lactis* ssp. *cremoris* MG1363 (accession number: Y_001031637) (260, 261). The sequence similarity between these two proteins and with *E. coli* NrdD sequences resulted in annotation of SSA_2230 as NrdD.

Prokaryotic ribonucleotide reductases

The ribonucleotide reductase (RNR) catalyzes the conversion of the four ribonucleotides, 5' -di (or tri)-phospho -adenosine, -cytidine, -guanosine, and -uridine, to their respective 2' deoxyribonucleotide (dNTP) species to provide substrates for synthesis and repair of DNA. RNR maintenance of cellular deoxyribonucleotide pools is regulated allosterically by both precursors and products of the reaction. Three classes of RNRs have been identified among microorganisms: class I, class II and class III (reviewed in 128). Class I RNRs are further separated into class Ia and class Ib based on sequence differences, and dissimilar allosteric regulation (117). Many bacterial species encode more than one RNR class to permit adaptation of the species to diverse environments. Class I RNRs require oxygen for catalytic activity (203), class II RNRs are active irrespective of the presence of O_2, and the class III RNRs are inactivated in the presence of O_2 and therefore only catalyze the synthesis of dNTPs under anaerobiosis (125).

The classes of RNRs are also distinguished by structural differences, allosteric regulation mechanisms, and by cofactors required for deoxyribonucleotide synthesis; however the general chemistry governing catalysis is conserved. RNRs act by radical chemistry for reduction of the 2' hydroxyl of ribose. For all RNR classes a one of two conserved Cys in the active site of the enzyme is proposed to form a thiyl radical for reaction initiation by substrate oxidation (68). The complete enzymatic turnover in synthesis of dNTPs is reliant upon additional Cys residues, with the number of conserved Cys residues differing between the classes. Class I and II enzymes both

encode five essential Cys residues (128), whereas Class III enzymes encode only the two conserved Cys at the active site (6). While all RNRs act through a free radical mechanism, the generation of the free radical is distinct between the classes.

The class I enzymes, encoded by genes commonly annotated *nrdAB* (class Ia), or *nrdEF* (class Ib) form a α2β2 structure, with the α protein (NrdA/NrdE) conferring both catalytic and allosteric properties, and the β protein (NrdB/NrdF) providing the mechanism for radical generation (116, 128). In the active form the β protein of both class Ia and class Ib enzymes contains a stable tyrosyl radical in proximity to a diiron oxygen cluster (142, 203). Through long-range transfer the tyrosyl radical is passed to the conserved Cys of the active site to generate the requisite thiyl radical. The requirement of oxygen for class I RNR function permits catalysis of dNTPs under both aerobic and microaerophilic conditions, however the reductase is not active in strictly anaerobic atmospheres.

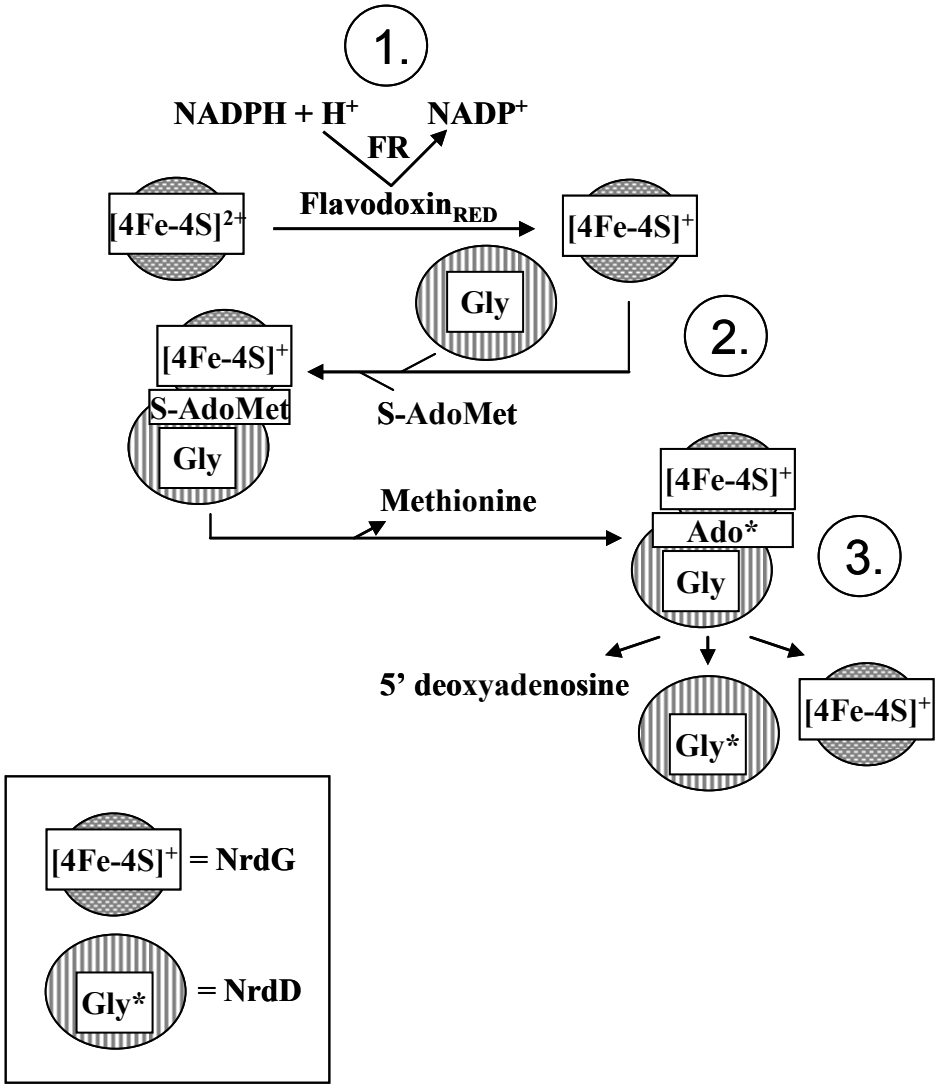
Class II enzymes may consist of a monomeric polypeptide (as is the case in *Lactobacillus leichmannii*; 235), but more commonly exist as homodimeric proteins. Regardless of the protein structure all class II RNRs require adenosylcobalamin, a cobalt containing derivative of vitamin B_{12} for development of the Cys radical in the active site. The thiyl radical of the active site is proposed to be generated directly by the cleavage of 5' deoxyadenosylcobalamin bond to form cobalt and a deoxyadenosyl radical, and therefore a second class I β-like protein is not required (149). This system is independent of O_2, and therefore catalytically active in both aerobic and anaerobic environments.

Class III RNRs, the anaerobic RNRs (ARNR) are destabilized by O_2 with activity restricted to anaerobic conditions. Characterization of the *L. lactis* ssp. *cremoris*, *S. aureus*, and *E. coli* ARNR was complicated by residual activity of class I RNRs of these organisms in standard anaerobic laboratory growth conditions (77, 116, 162). To determine ARNR properties in organisms encoding both class I and III enzymes, inhibition of the class I enzyme is often performed by growth of strains in the presence of hydroxyurea. Class III ARNRs belong to the "Radical S-adenosylmethionine (SAM)" superfamily of proteins that share a common mechanism of free radical generation by the cleavage of SAM (239). ARNRs consist of two proteins, commonly annotated as NrdD and NrdG that form a α2β2 structure (118), with the catalytic and allosteric sites on the *nrdD* encoded, α protein (193), and the smaller *nrdG* encoded β protein conferring radical generation. In contrast to class I RNRs, in which a tyrosyl radical is formed on the activase β protein and then transferred to the α protein active site Cys residue, the free radical of ARNR is generated at a conserved Gly of NrdD (the class III α protein) in a mechanism dependent on an iron-sulfur cluster bound by NrdG. In this process the $[4Fe\text{-}4S]^{2+}$ cluster is reduced to a $[4Fe\text{-}4S]^{+}$ state by reduced flavodoxin (21), which is generated in the bacterial cell by flavodoxin reductase and NADPH. The reduced NrdG forms a tight complex with NrdD, and then associates with SAM (Fig. 11). SAM bound by the reduced NrdG Fe-S is subsequently reductively cleaved to form methionine and a 5'-deoxyadenosyl radical. The deoxyadenosyl radical then transfers its spin to the conserved Gly of NrdD in a step requiring the transfer of one proton and one electron (from Gly) to deoxyadenosyl,

Figure 11. The proposed mechanism for ARNR radical generation

In the first step [1] of ARNR radical generation the $[4Fe\text{-}4S]^{2+}$ cluster of NrdG is reduced to a $[4Fe\text{-}4S]^{+}$ state by reduced flavodoxin, which is generated in the bacterial cell by flavodoxin reductase (FR) and NADPH. In the second step [2] NrdG in this state forms a tight complex with NrdD, and then associates with S-adenosylmethionine (S-AdoMet). S-AdoMet bound by the reduced NrdG Fe-S is subsequently reductively cleaved to form methionine and a 5'-deoxyadenosyl radical. In the third step [3] the deoxyadenosyl radical then transfers its spin to the conserved Gly of NrdD and is then released as 5'deoxadenosine. The figure shown was adapted from (261).

Figure 11.


1.
NADPH + H+
NADP+
FR
Flavodoxin RED
[4Fe-4S]2+
[4Fe-4S]+
Gly
2.
[4Fe-4S]+
S-AdoMet
Gly
S-AdoMet
Methionine
[4Fe-4S]+
Ado*
Gly
3.
5' deoxyadenosine
Gly*
[4Fe-4S]+
[4Fe-4S]+ = NrdG
Gly* = NrdD

this is then released as 5'deoxadenosine. Formation of the glycyl radical of NrdD primes the enzyme for ribonucleoside triphosphate reduction with formate acting as an electron donor (reviewed in 128). The NrdD proteins of both *L. lactis* and *E. coli* harbor the conserved Gly stable free radical at their C-termini. In *L. lactis* the sequence adjacent to the conserved Gly is KRTC**G**YL (260). The O_2 sensitivity of ARNRs is a result of cleavage of NrdD at the Gly radical by O_2 for inactivation of the ARNR (125). Additional conserved features of NrdD include the two active site Cys (6) essential for reduction of ribonucleotide substrates to deoxyribonucleotides, and a region sandwiched between the NrdD homodimers that governs the binding of allosteric effectors for regulation of protein activity (143).

Allosteric regulation of RNR activity

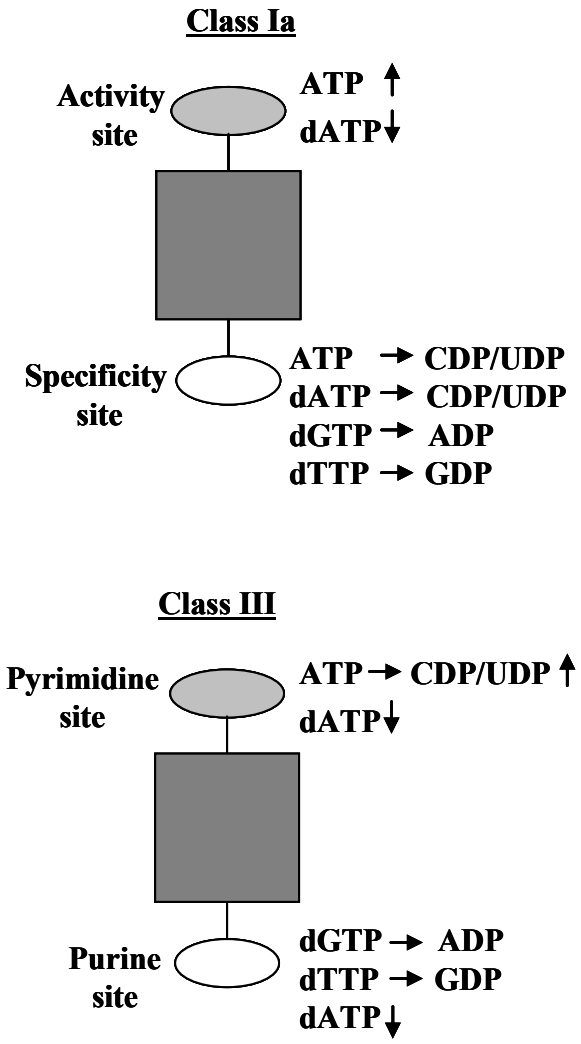
RNRs are allosterically regulated by dNTPs and ATP for synthesis of balanced DNA precursor pools in the cell. These allosteric enzymes contain prototypical distinct sites for substrate binding (the active/catalytic site), and association of the regulatory effector. Binding of the effector subsequently causes conformational changes in the enzyme that affect the substrate binding site and catalytic activity of the protein. Regulation of ARNRs by this mechanism has been studied in *E. coli*, the *Escherichia* bacteriophage T4, and *L. lactis* (5, 70, 260). The specificity of the four RNR substrates is determined by binding of the allosteric effector to the specificity site (S-site), which induces necessary structural changes for substrate binding. The effectors that associate with the S-site in all RNR classes are deoxyribonucleoside 5'triphosphate products

dATP (and in the case of T4 ARNR, ATP), dTTP and dGTP. The general scheme for allosteric regulation, determined for the class Ia RNR for *E. coli* includes the following progressive reduction of ribonucleoside 5' diphosphate (NDP) precursors: binding of dATP at the S-site promotes the binding and reduction of the pyrimidines CDP and UDP; binding of dTTP at the S-site both inhibits the reduction of CDP and UDP and initiates the reduction of GDP; and binding of dGTP inhibits the reduction of GDP and promotes the reduction of ADP (reviewed in 128). In the case of Class I RNRs the substrates reduced are NDPs, and products are deoxyribonucleoside 5'diphosphates (dNDPs). For class II and III RNRs the substrates are 5'triphosphates (NTPs), for production of dNTPs. Class Ia RNRs contain a general activity site (the A site) in addition to the specificity site, which is competitively bound by ATP or dATP to turn on or off the enzyme, respectively (Fig. 12). Two allosteric sites have been identified for *E. coli* and *L. lactis* ARNRs; however only one site was found in the T4 ARNR (5). In both *E. coli* and *L. lactis* experimental evidence indicates that two effector binding sites are exclusively associated with ATP (and dATP); or dATP, dGTP and dTTP (Fig. 12). The site bound exclusively by the ATP effector is referred to as the pyrimidine site, as the association with ATP results in reduction of UTP and CTP substrates. The other effector binding site- bound by dGTP and dTTP is termed the purine site, in that ADP and GDP substrates are reduced. Association with dATP at a threshold concentration at either site results in enzymatic inactivation (70, 260). Through these regulatory mechanisms a balanced supply of dNTPs is created for the replication of DNA.

Figure 12. Effector binding for allosteric regulation of Class Ia and ClassIII RNRs

The schematic shown includes the proposed mechanisms for allosteric regulation of class Ia and III RNRs. The association of effector molecules with specificity site of the class Ia RNR and the pyrimidine and purine sites of the class III ARNR are indicated. The enzymatic activity is induced by association of the class Ia active site with ATP, whereas association with dATP turns the enzyme off. Regulation is different with the class III ARNR, in that a single activity site does not exist. Instead pyrimidine and purine sites have been identified. These are named to indicate the substrate resulting from association with a specific allosteric effector. As for class Ia RNRs, binding of dATP at a certain threshold concentration is thought to result in enzymatic inactivation.

Figure 12.


Class Ia
Activity site
ATP
dATP
Specificity site
ATP → CDP/UDP
dATP → CDP/UDP
dGTP → ADP
dTTP → GDP
Class III
Pyrimidine site
ATP → CDP/UDP
dATP
Purine site
dGTP → ADP
dTTP → GDP
dATP

Transcriptional regulation of RNRs in *E. coli*

Expression analysis of regulation of *E. coli* RNRs indicates they are differentially transcribed in response to O_2 tension, and exposure to agents of oxidative stress. *E. coli* encodes class Ia, class Ib and class III RNRs. Transcription of both aerobic RNRs is induced under aerobic growth conditions; however the capacity for aerobic growth is conferred for the most past by the class Ia RNR (115). In contrast, under oxidative stress expression of the class Ib RNR is enhanced, and it has been proposed that this RNR in *E. coli* responds to DNA damage resulting from oxidative stress (174, 208).The expression of the ARNR operon in *E. coli* is upregulated under anaerobic conditions by the major aerobic/anaerobic shift regulatory protein FNR. FNR directly associates with FNR binding sites of the *nrdDG* promoter to activate transcription of ARNR operon. Deletion of FNR resulted in increased anaerobic expression of *nrdA*, thus indicating that under anaerobiosis FNR inhibits expression of the class Ia RNR. The mechanism of FNR *nrdAB* regulation is thought to be indirect as no putative FNR binding sites appear in the promoter region of the *nrdAB* operon (27). Negative regulation of *E. coli* RNRs is through the NrdR regulatory protein that associates directly with promoters for class Ia, Ib and III RNR genes via NrdR binding sites (262). With *E. coli* as an example it is assumed that in all prokaryotes encoding for RNRs expression is regulated at the level of transcription by both positive and negative regulatory proteins.

Rationale

NrdD is the essential α-component of the ARNR complex formed by the association of an NrdD homodimer with the β-component, NrdG, therefore the anticipated outcome of *nrdD* mutation is sensitivity to entirely reduced atmospheres (125). It was hypothesized that 6-26 would grow only in aerobic atmospheres. Dr. Sehmi Paik demonstrated in a time-course growth experiment that 6-26 grew indistinguishably from SK36 under aerobic conditions (Fig. 13). When the same experiment was conducted under supposed anaerobic conditions 6-26 exhibited a slower growth rate than SK36, but did reach stationary phase within 8 h of subculture (Fig. 13) (196). The observation that anaerobic growth was not entirely attenuated in 6-26 caused Dr. Paik to speculate that 6-26 may have acquired an extragenic suppressor mutation permitting anaerobic growth. Dr. Paik attempted to reconstruct the *nrdD* mutation by PCR amplification of the *nrdD::magellan2* region of 6-26 and cloning of the amplicon in the suicide vector pVA2606, to create pJFP16. SK36 transformed with pJFP16 yielded three Cm^r clones, JFP20, JFP21 and JFP22. These strains were compared to SK36 and 6-26 by surface plating on TSA with or without Cm, and growth under reduced oxygen in the anaerobe jar or in the anaerobe chamber. Dr. Paik found that *S. sanguinis* strains JFP20, JFP21, and JFP22 all exhibited similar growth to SK36 and 6-26 when incubated in the anaerobe jar and in the anaerobe chamber (Fig. 14).

Figure 13. Growth study of SK36 and 6-26

Growth studies were performed for SK36 and the STM mutant, 6-26 in 200 μl volumes, with the optical density at 450 nm recorded at ten minute intervals. In independent experiments SK36 and 6-26 were cultured aerobically (filled symbols), or under an ostensibly anaerobic (reduced oxygen) atmosphere (empty symbols) to compare the growth trends of the two strains as a function of O_2 availability (196).

Figure 13.

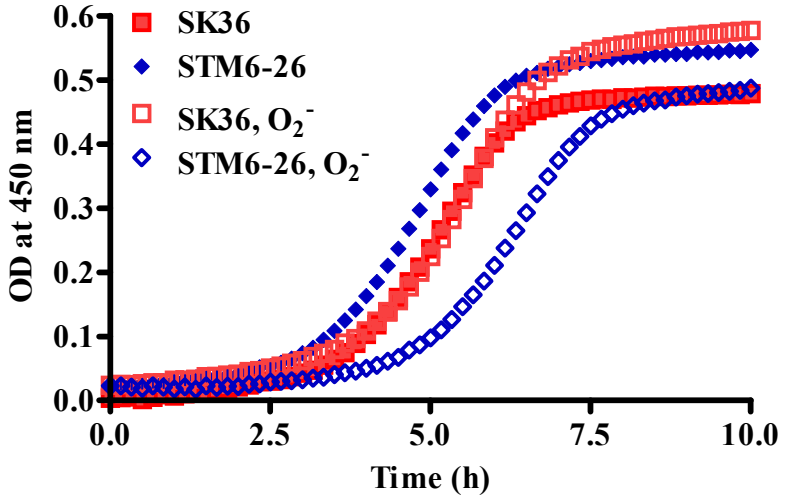
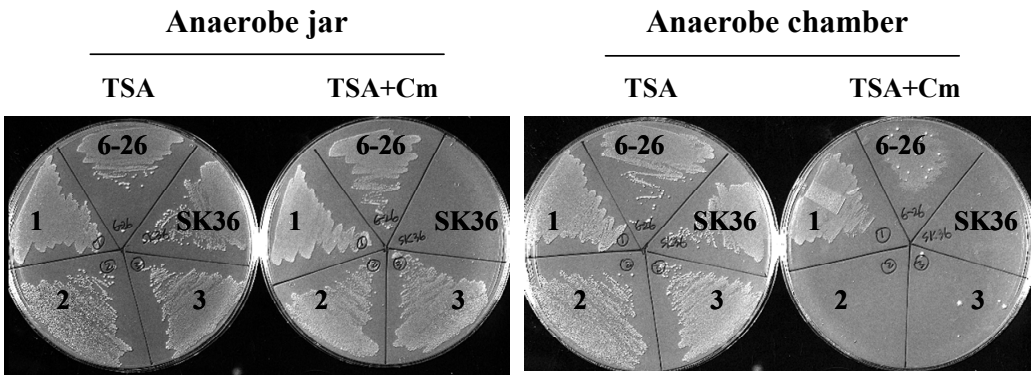

SK36
STM6-26
SK36, O_2^-
STM6-26, O_2^-
OD at 450 nm
0.6
0.5
0.4
0.3
0.2
0.1
0.0
0.0
2.5
5.0
7.5
10.0
Time (h)

Figure 14. Growth of *S. sanguinis* ARNR mutants under different O_2 tensions

The growth of *S. sanguinis* SK36 was compared with that of ARNR mutant strains on agar plates. Cells were plated on either TSA plates or on TSA containing chloramphenicol ('+ Cm'), and grown under reduced O_2 in an anaerobe jar or in an anaerobe chamber for 2 days. SK36, the parental, WT strain; 6-26, the original STM mutant; 1, JFP20; 2, JFP21; 3, JFP22; the JFP20, JFP21, and JFP22 strains were re-created ARNR transposon insertion mutants (196).

Figure 14.


Anaerobe jar
Anaerobe chamber
TSA
TSA+Cm
TSA
TSA+Cm
6-26
SK36
1
2
3
6-26
SK36
1
2
3
6-26
SK36
1
2
3
6-26
SK36
1
2
3

One unexpected outcome of this experiment was that anaerobe chamber growth of 6-26, JFP20, JFP21 and JFP22 was attenuated on TSA with Cm, suggesting that expression of Cm acetyltransferase (CAT) in these strains had an additional negative effect on growth (Fig. 14) (196). As these studies failed to demonstrate a growth defect on TSA under our standard laboratory procedure for imposing anaerobiosis we sought to develop an in-frame deletion of *nrdD* for further analysis of this locus. This strategy was intended to test the possibility of an extragenic suppressor mutation in 6-26, as well as avoidance of polar effects owing to transposon insertion, and to negate apparent growth suppression owing to CAT expression.

Results

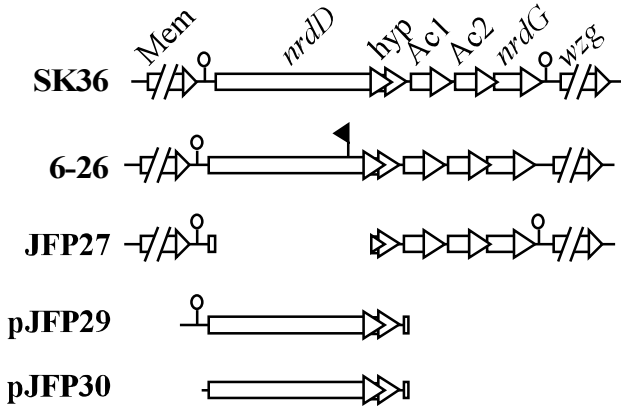
Development of JFP27

The *nrdD* locus of *S. sanguinis* SK36 includes *nrdD* preceded upstream by a putative membrane protein gene (*SSA_2231*), downstream by a hypothetical protein gene (*SSA_2229*) overlapping with the 3' end of *nrdD*, two acetyltransferase genes (*SSA_2228* and *SSA_2227*), and the ARNR β protein encoding gene *nrdG* (*SSA_2226*) (Fig. 15). The strategy for mutagenesis of *nrdD* included deletion of conserved C-terminal residues (KRTC**G**YL) that encompass the site of glycyl radical formation, and maintenance of the 3' end of the gene to avoid mutagenesis of the downstream hypothetical protein gene (*SSA_2229*). Gene SOEing was used to create an in-frame deletion construct with removal of the central 705 aa of the 731 aa sequence, for fusion of the first 12 codons with the last 14 codons of *nrdD*. The plasmid containing this construct was used as transforming DNA with competent SK36 cells. As no selective marker is conferred by the mutagenic construct, *S. sanguinis* colonies on TSA agar were replica plated and grown under a reduced oxygen condition (anaerobe jar) or in an anaerobe chamber for identification of clones incapable of anaerobic growth. All 100 clones screened grew under both conditions, and therefore this screening strategy was unsuccessful. Twenty replica plated clones were selected, and cultured in BHI overnight. A boiled cell preparation of each clone was PCR amplified for size

Figure 15. The *nrdD* locus of *S. sanguinis*

The structure of the *nrdD* locus is indicated for the strains and plasmids listed on the left. Potential terminators are indicated as a circle above a vertical line. The position and orientation of the transposon in strain 6-26 are indicated by the flag. Gene designations are as follows: Mem, putative membrane protein; *nrdD*, anaerobic ribonucleotide reductase gene; hypo, a hypothetical protein gene; Ac1, putative acetyltransferase gene; Ac2, second putative acetyltransferase gene; *nrdG*, anaerobic ribonucleotide reductase small subunit gene; *wzg*, ortholog of putative transcriptional regulator *wzg* of *S. gordonii*.

Figure 15. The *nrdD* locus of *S. sanguinis*



determination by gel electrophoresis for differentiation between strains with an intact *vs*. deleted *nrdD*. Three of the twenty clones screened exhibited a product size indicating the mutation of *nrdD* (data not shown). The clone chosen for further characterization, designated JFP27, exhibited normal growth on agar in anaerobe jars under reduced O_2. Therefore JFP27, 6-26 and SK36 were grown under the same reduced O_2 condition prior to analysis of anaerobic sensitivity.

Optimization of the soft-agar growth assay

The conundrum presented by persistent growth *S. sanguinis nrdD* mutants on agar under assumed anaerobic conditions, meant that an alternative technique was needed for characterization of these strains. In the soft-agar assay chosen, oxygen-reactive dye indicators were mixed into the culture tube for estimation of oxygen gradients in the agar. Preliminary soft-agar assays were conducted in snap-top 14-ml PS round-bottom tubes with TSB-LMP agarose supplemented with resazurin, methylene blue, or no dye indicator at all. Following anaerobe chamber pre-incubation of tubes for 48 h, stationary phase SK36 was added. Incubation at room temperature for 30 minutes following inoculation with SK36 was sufficient for solidification of the agar. Tubes were then removed from the anaerobe chamber; the tops loosened to establish an oxygen gradient within the agar, and grown at 37 ºC for 24 hours. Visual inspection of TSB-LMP agarose tubes containing dye indicated that Falcon tubes were not an ideal choice for this assay, as pink (resazurin) or blue (methylene blue) color was observed throughout the length of the tube (data not shown). The PS tubes used may have been

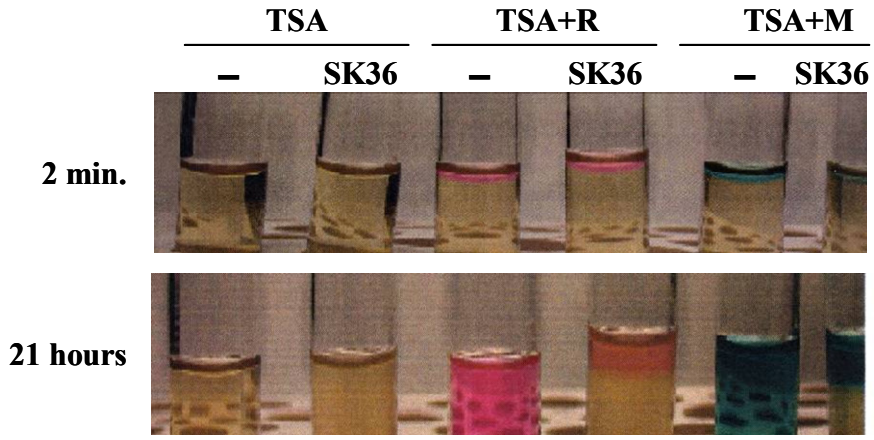
porous to oxygen during aerobic incubation. In support of this, dissection of the agar confirmed that dye reactivity was exclusively at the walls of the tube.

Glass Hungate tubes were then used for comparison of the oxygen dye indicators, resazurin and methylene blue for soft-agar growth of SK36. As described above, TSB-agarose was driven to anaerobiosis by pre-incubation in the anaerobe chamber. Following inoculation of SK36, and agar solidification, the tubes were incubated aerobically. An image taken 2 minutes after removal of tubes from the anaerobe chamber demonstrates the initial reactivity to oxygen at the agar surface (Fig. 16). These tubes lacked any coloration when in the anaerobe chamber. The zones of reactivity increased after aerobic incubation overnight; however they were restricted to the very top of the tubes, suggesting that an anaerobic state was achieved at greater agarose depths (Fig. 16). As shown in Figure 16, the depth of oxygen reactivity was greater in the control tube without SK36, than in the presence of streptococci, thus indicating consumption of oxygen for growth of this facultative anaerobe. In the studies performed comparing resazurin and methylene blue it became apparent that methylene blue was restrictive for SK36 growth under the conditions used. While growth of SK36 was observed in aerobic zones of TSB-LMP agar, no growth occurred in the region of methylene blue reactivity (data not shown). This was not an issue when resazurin was used as a dye indicator. This observation, in addition to the greater reduction potential of resazurin made the magenta indicator a logical choice for further analysis of *S. sanguinis* strains.

Figure 16. Optimization of the soft-agar assay

Two different dye indicators were investigated for oxygen detection in the soft-agar assay, resazurin (+R) and methylene blue (+M). The molten soft-agar was driven to anaerobiosis in an anaerobe chamber. Stationary phase cells of *S. sanguinis* SK36 were inoculated into half the soft-agar tubes (SK36) and images were taken of the tubes 2 minutes after removal from the anaerobe chamber (2 min.) for evidence of dye reactivity. The tubes were then incubated aerobically at 37 ºC for 21 hours to permit growth of SK36. Growth and oxygen densities were documented photographically.

Figure 16.



Comparison of *S. sanguinis* strains by the soft-agar growth assay

Growth comparisons of SK36, JFP27 and 6-26 were performed in glass Hungate tubes in the presence or absence of resazurin, as described above. The parental strain SK36 consistently grew throughout the tube, in aerobic (dye indicated) and anaerobic zones (Fig. 17)**.** In contrast the growth of JFP27 and 6-26 was restricted to oxidized agar, indicated by the pink hue of the oxidized resazurin. This suggests that sensitivity of 6-26 and JFP27 to anaerobic conditions is attributed to the loss of NrdD expression.

To confirm this, we attempted to complement NrdD strains with the *nrdD* gene expressed from its native promoter. Sequence analysis predicted a possible terminator 149 bp upstream of *nrdD* (29, 30). The regions preceding *nrdD* from *S. sanguinis*, two strains each of *S. pyogenes* and *S. pneumoniae*, *S. mutans*, and *S. agalactiae* were compared for conservation of promoter elements among streptococci. With the exception of SK36, all other species shared potential -35, -16, and -10 promoter sequence elements (274) and other potential promoter elements in the 59 bp upstream of *nrdD* (Fig. 18). As a homologous region was not identified for *S. sanguinis* SK36, it appears that an evolutionarily divergent promoter may regulate the expression of *nrdD* in *S. sanguinis*.

The plasmids pJFP29 and pJFP30 contain *nrdD* preceded by 337 or 36 nucleotides, respectively, cloned into the *E. coli-Streptococcus* shuttle vector pVA838. Plasmid pJFP29 was designed to include the possible terminator and the postulated subsequent native promoter of *nrdD* (Fig. 15).

Figure 17. Soft-agar growth of *S. sanguinis nrdD* mutants

S. sanguinis strains SK36, 6-26, and JFP27 were compared by a reduced soft-agar growth assay. Tryptic soy broth soft-agar with or without resazurin (TSA+R, or TSA, respectively) was driven to anaerobiosis in an anaerobe chamber. Stationary phase cell of SK36, 6-26 or JFP27 were then added to the soft-agar and the tubes incubated aerobically to create an oxygen gradient within the soft-agar. The negative control lacking streptococci is shown for indication of the gradient of oxygenated (pink) and reduced (yellow) agar. *S. sanguinis* SK36 exhibited the wild-type phenotype of growth in both oxygenated and reduced zones, whereas the *nrdD* mutants, 6-26 and JFP27 displayed growth restricted to agar containing oxygen.

Figure 17.

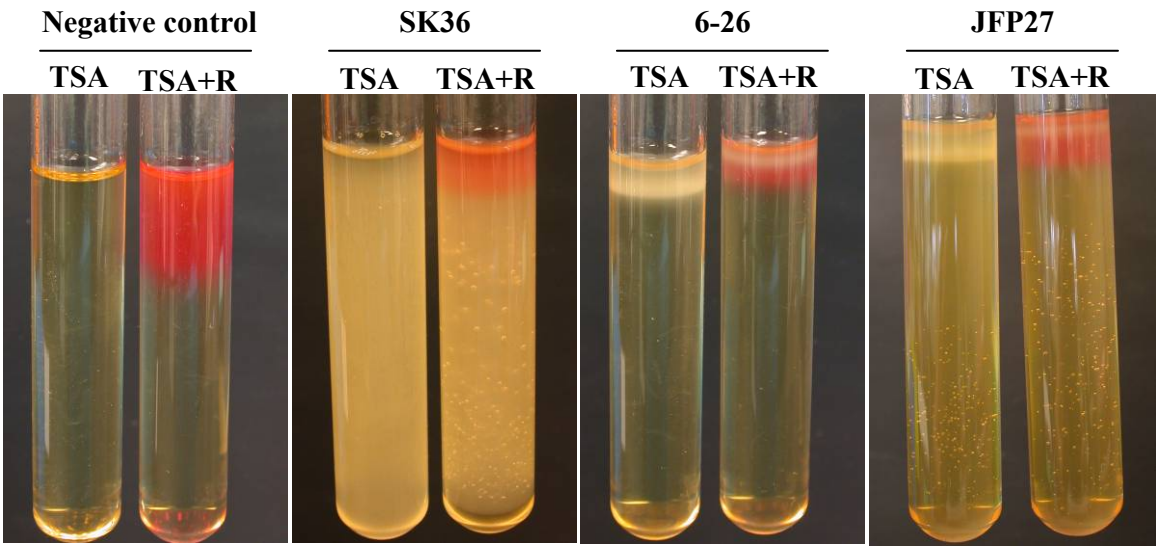

Negative control
SK36
6-26
JFP27
TSA
TSA+R
TSA
TSA+R
TSA
TSA+R
TSA
TSA+R

Figure 18. Identification of putative *nrdD* promoter elements

Regions upstream from the *nrdD* gene in the species indicated were aligned using the GCG program PILEUP, with default parameters. Gray highlighting indicates conserved promoter elements. Potential transcriptional start sites are indicated by a single underline. Start codons for *nrdD* and potential ribosome binding sites are indicated in bold. Sources and accession numbers for the sequences are: Pn R6, *S. pneumoniae* R6 (AE008401); Pn TIGR4, *S. pneumoniae* TIGR4 (AE007335); Py, *S. pyogenes* M1 (AE006630, AE006631); Py 8232, *S. pyogenes* MGAS8232 (AE010118); Sag, *S. agalactiae* 2603V/R (AE014286); Sm UA159: *S. mutans* UA159 (AE015029, AE015030); Ss SK36: *S. sanguinis* SK36 (4806195).

Figure 18.


Pn R6 TGTATTTTTAAAATTTTAGATTTTTAACACAAGATATTGATTTTTCTTTTTAGAGTGGTATAATACTTTTTAGAAAGAACATTTTAGAAAAGAGCATG
Pn TIGR4 TGTATTTTTAAAATTTTAGATTTTTAACACAAGATATTGATTTTTCTTTTTAGAGTGGTATAATACTTTTTAGAAAGAACATTTTAGAAAAGAGCATG
Py M1 AATATATTGTTTCGAGCAAAAACTTGACACAATATATTGATTTTTTTAGTTGTCGTGATATAATAGTTCCGTAAGGAGATTTGATATG~~~~~~~~~~
Py 8232 AATATATTGTTTCGAGCAAAAACTTGACACAATATATTGATTTTTTTAGTTGTCGTGATATAATAGTTCCGTAAGGAGATTTGATATG~~~~~~~~~~
Sag AATATATTGTTCTGAGCAAAAACTTGACACAATATATTGATTTTTTTTCTTTGCATGCTATAATTGTCACTGGAGGATATTTCACATG~~~~~~~~~~
Sm UA159 AATATATTGTTCT.AAATTTTTTTTACCACAATATATTGATTTTTTTCTTTCTTGTGTTATAATAGCTTCTGTAAGGAGAAAGACGTATG~~~~~~~~
Ss SK36 CTGTCGGAGATTCTTGGAATTTCTTAGCAAGCGATAGGTTTAAAGAGACTTGATTTGTCCTAAGGGCAGAGAAAAGGATGTGTTCGTGATG~~~~~~~

To exclude the possibility of *nrdD* being expressed from a plasmid-borne promoter, pJFP30 was created, in which the *nrdD* gene was preceded by 36 bp-enough to contain a ribosome binding site but not a complete promoter (Fig. 15).

S. sanguinis SK36, 6-26 and JFP27 were each transformed with pVA838, pJFP29, and pJFP30. Erythromycin-resistant transformants were screened for the *nrdD* insert by PCR amplification of boiled cell preps of isolated clones with primers pVA838-2939 and pVA838-5331 that flank the insertion site in pVA838. PCR products analyzed by gel electrophoresis confirmed the presence of the plasmid-borne *nrdD* insert in strains transformed with pJFP29 and pJFP30 (data not shown). Complementation of the NrdD mutant phenotype in 6-26 and JFP27 was then evaluated by soft-agar studies. For the control strain, SK36 growth was unaffected by transformation with pVA838, pJFP29 or pJFP30, as all strains grew throughout the tube (typically 7.0 cm in depth) (Fig. 19A and 19B). The *nrdD* mutant strain JFP27 containing pJFP29 exhibited growth that was indistinguishable from that of parent strain, SK36. In contrast pJFP30 failed to complement the growth of JFP27, suggesting that complementation of JFP27 by pJFP29 was due to expression of *nrdD* from its native promoter (Fig. 19A and19B). The complementation observed for JFP27 (pJFP29) additionally confirms that NrdD is essential for growth in an entirely reduced, anaerobic condition.

With STM mutant 6-26, typical dense growth in the reduced agar was not restored by either pJFP29 or pJFP30. The soft-agar growth of 6-26 (pJFP29) and 6-26 (pJFP30) was distinctive, however, in that isolated colonies were observed growing in

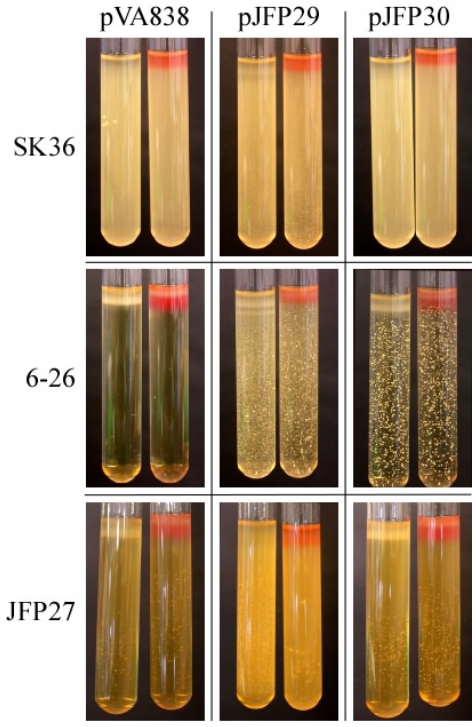
Figure 19. Soft-agar comparison of *S. sanguinis* plasmid transformants

A) *S. sanguinis* strains were inoculated into anaerobic soft-agar tubes and then incubated overnight with exposure to ambient air to create an oxygen gradient. Photographs were taken of the tubes after overnight growth. The strains indicated on the left, containing vector alone (pVA838) or vector plus *nrdD* constructs (pJFP29 and pJFP30) were inoculated into tubes with or without the oxygen indicator dye resazurin (right and left tubes in each pair, respectively).

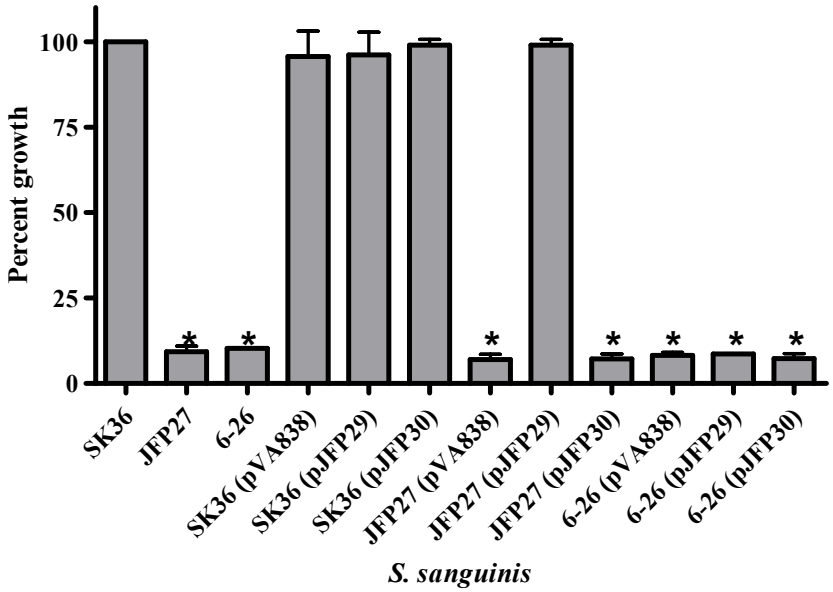
B) Depth of growth in tubes without resazurin was measured in cm with a standard ruler and the percent growth determined by dividing the zone of growth by total agar depth and multiplying by 100. Values are averages ± standard deviations for three samples obtained on at least two separate occasions. Asterisks indicate results significantly different from those for SK36 ($P < 0.001$).

Figure 19.

A.


pVA838
pJFP29
pJFP30
SK36
6-26
JFP27

B.


Percent growth
100
75
50
25
0
SK36
JFP27
6-26
SK36 (pVA838)
SK36 (pJFP29)
SK36 (pJFP30)
JFP27 (pVA838)
JFP27 (pJFP29)
JFP27 (pJFP30)
6-26 (pVA838)
6-26 (pJFP29)
6-26 (pJFP30)
S. sanguinis

the reduced agar (Fig. 19A). (Smaller spots apparent in some of the other tubes are bubbles in the agar rather than colonies. These were easily distinguished by the lenticular disk morphology of the colonies.) This phenotype was not observed for 6-26 (pVA838), suggesting that plasmid-borne *nrdD* caused this aberrant growth. As wild-type turbid growth was not restored in 6-26 containing pJFP29 or pJFP30 we speculated that a specific event occurring in a sub-set of cells resulted in the formation of individual colonies. A plausible explanation of this observation is low-frequency recombination between *nrdD* of pJFP29 or pJFP30 and the corresponding sequence of *nrdD* flanking the transposon at the chromosomal locus of this gene. To further investigate this possibility colonies of 6-26 (pJFP29) and 6-26 (pJFP30) were isolated from the reduced agar by removal of turbid agar at the top of the tube and extraction of colonies from deep in the reduced agar with a sterile pipette tip. The collected colony was cultured in liquid broth overnight, and cells prepared for amplification with chromosome specific primers flanking *nrdD* (ARTR-SalI-22495-F and ARTR-SalI-20031-R) for assessment of the genomic *nrdD* region. Analysis of the PCR products by gel electrophoresis indicated that the transposon had been lost from the genomic *nrdD* in the 6-26 (pJFP30) cells, in that the same size amplicon was observed for SK36 cells (Fig. 20A). In contrast, the size of the 6-26 (pJFP29) PCR product indicated that the transposon was still present in the *nrdD* gene in these cells (Fig. 20A).

Overnight cultures of isolated 6-26 (pJFP29) and 6-26 (pJFP30) colonies were also regrown in soft-agar. Dense growth was observed for both strain variants throughout the tube, a phenotype similar to the wild-type (Fig. 20B). Agar with turbid

growth was extracted from these tubes and grown in liquid culture overnight for preparation of cells for amplification of genomic *nrdD*, as described above. The gel electrophoresis pattern of these PCR products was similar to the isolated colonies in that genomic *nrdD* appeared to be restored in 6-26 (pJFP30), whereas the transposon interrupting *nrdD* was maintained in 6-26 (pJFP29) (Fig. 20A). Therefore the mechanism accounting for restored growth in reduced agar was distinct between these strain variants.

CAT expression does not affect growth in soft-agar

Dr. Sehmi Paik Lee observed that strains 6-26, JFP20, JFP21 and JFP22, expressing CAT from the *magellan2* mini-transposon were hyper-sensitive to growth in an anaerobic atmosphere (196). To determine whether this CAT sensitive phenotype persisted in the soft-agar assay, the growth of a hypervirulent STM mutant, 3-24 was compared to SK36 by both surface plating in the presence and absence of chloramphenicol, and soft-agar culturing in the presence and absence of chloramphenicol. For surface plating the strains were grown in an anaerobe chamber 37 ºC incubator for two days. Soft-agar growth experiments were performed as described above.

Reduced growth on plates containing chloramphenicol was observed for the transposon insertion mutant, 3-24 when compared to lawn formation on plates without the selective antibiotic (Fig. 21A). However, 3-24 exhibited soft-agar growth through

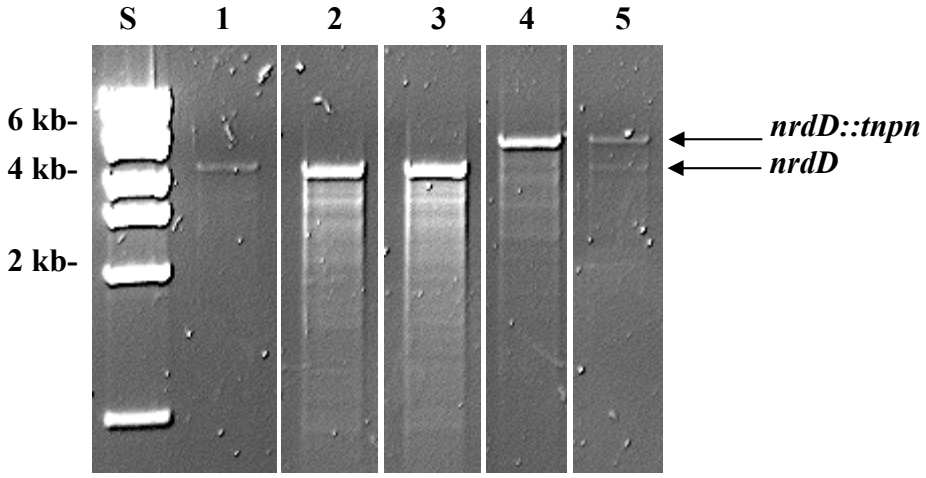
Figure 20. PCR detection of transposon insertion in 6-26

A) Primers flanking *nrdD* at the genomic locus were used for PCR amplification of the *nrdD* gene, with washed cell preparations as template. SK36 was used as a control for reversion at the *nrdD* locus (lane 1), and compared to isolated 6-26 (pJFP30) (lane 2) or 6-26 (pJFP29) (lane 4) colonies. PCR amplification was also performed on washed 6-26 (pJFP30) (lane 3) or 6-26 (pJFP29) (lane 5) cells isolated from turbid agarose. The expected amplicon for SK36 was 4.4 kb, whereas the presence of the transposon within *nrdD* would result in a 5.8 kb product.

B) Isolated 6-26 (pJFP29) and 6-26 (pJFP30) colonies were regrown in reduced soft-agar (+). Control soft-agar tubes for each strain that lack cells, yet containing resazurin are also shown (—).

Figure 20.

A. PCR detection of transposon insertion in 6-26



B. Soft-agar subculture of isolated colonies

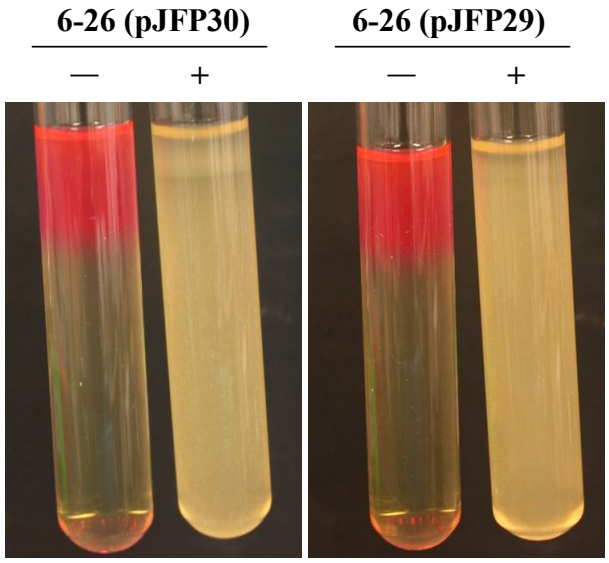
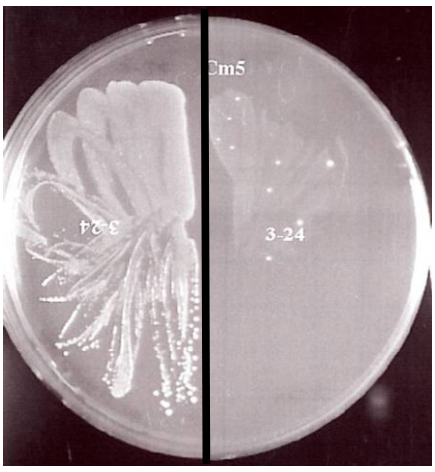


Figure 21. Soft-agar evaluation of 3-24

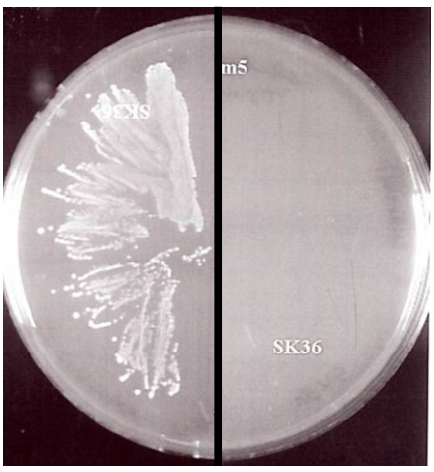
S. sanguinis 3-24 and SK36 were compared for growth on the surface of TSA plates with (TSA+Cm) or without (TSA) chloramphenicol. Plates were incubated at 37 ºC in an anaerobe chamber. 3-24 exhibited decreased growth on a plate containing chloramphenicol, despite the *cat* gene expressed by this strain (A). SK36 grew well on TSA, but did not grow in the presence of chloramphenicol, as anticipated (B). The reduced soft-agar assay was then used for anaerobic culture of 3-24 and SK36 in the presence (TSA+Cm) or absence (TSA) of chloramphenicol. The oxygen dye indicator, resazurin was included in duplicate tubes for visualization of the oxygenated zone of soft-agar (+R). 3-24 grew well in the presence and absence of chloramphenicol when cultured in anaerobic soft-agar (C), with turbidity similar to wild-type SK36 in TSA alone (D). Spots visible in soft-agar tubes are bubbles introduced during inversion to mix cells with the media.

Figure 21.

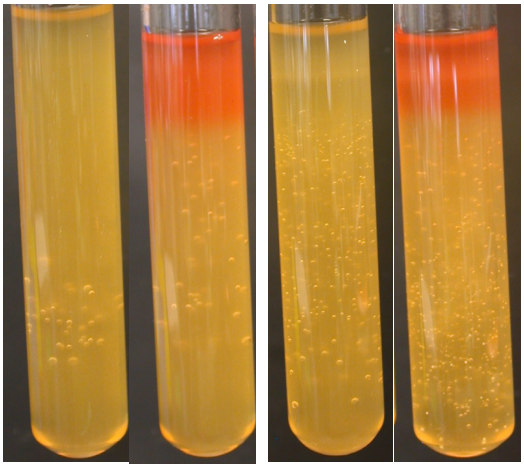
A. Surface plating of 3-24



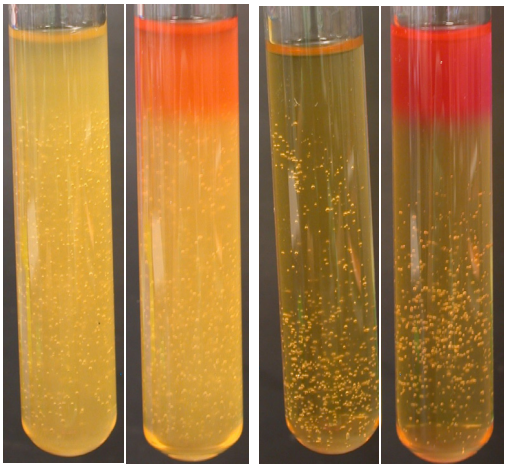
B. Surface plating of SK36



C. Soft-agar comparison of 3-24



D. Soft-agar comparison of SK36



the entire tube in the presence and absence of chloramphenicol, which was indistinguishable from SK36 (Fig. 21C and 21D). These observations suggest that the defect in growth on an agar surface imposed by chloramphenicol selection does not occur when a strain is embedded in agar.

Discussion

S. sanguinis, a facultative anaerobe, possess both class Ib and class III RNRs to permit de novo deoxyribonucleotide synthesis in the diverse atmospheres encountered by this organism. The extreme O_2 sensitivity of class III RNRs mandates deoxyribonucleotide synthesis by class Ib RNRs under growth atmospheres ranging from aerobic to microaerophilic (125). Growth observed for *S. sanguinis* mutants 6-26, JFP20, JFP21, JFP22 and JFP27 under our standard anaerobic conditions suggested that under trace amounts of O_2 the class Ib RNR remains catalytically active. Growth studies of *E. coli* and *S. aureus* ARNR mutants provide additional evidence of residual activity of class I RNRs at low levels of oxygen, with growth defects owing to *nrdDG* mutation most apparent when mutants were cultured in the presence of hydroxyurea, an inhibitor of class I RNRs (77, 162). A soft-agar assay previously used for evaluation of the *P. gingivalis* oxidative stress response was employed for evaluation of anaerobic growth of *S. sanguinis* mutants (246). The oxygen reactive dye resazurin was incorporated into soft-agar to establish the extent of oxygen permeability in the growth medium. Turbid growth of JFP27 and 6-26 was reproducibly restricted to oxygenated zones (as indicated by resazurin), and therefore putatively reliant on the class Ib RNR for dNTP synthesis. The restricted growth observed for 6-26 and JFP27 suggests that transposon insertion mutagenesis or deletion of *nrdD* both result in anaerobic growth inhibition.

Our efforts to complement the mutant phenotypes of these strains suggest a model for the transcriptional organization of the *nrdD* locus in *S. sanguinis*. The genomic locus of *nrdD* in *S. sanguinis* includes *nrdD* preceded upstream by a putative membrane protein gene and downstream by a 3' overlapping hypothetical protein gene, two acetyltransferase genes and *nrdG*, which encodes the β-subunit of the ARNR holoenzyme. Complementation of the in-frame *nrdD* deletion mutant JFP27 by pJFP29 but not pJFP30 suggests that the *nrdD* promoter is contained within the pJFP29 construct. Through nucleotide sequence comparison of the upstream, intergenic region of *S. sanguinis nrdD* to other streptococci it was determined that putative promoter elements conserved for other species do not occur in *S. sanguinis.* Therefore a unique, unidentified promoter may regulate the expression of *nrdD* in *S. sanguinis*.

The failure of pJFP29 to complement the *nrdD* mutation in 6-26 suggests that the transposon mutation was polar on the expression of downstream genes. In 6-26 (pJFP30) occasional recombination between the chromosomal and plasmid borne *nrdD* gene resulted in restoration of the genomic *nrdD* gene and transcription of the remainder of the operon, yielding colony formation in reduced soft-agar. This was likely not observed for JFP27 bearing pJFP29 or pJFP30 because of the limited sequence shared by the genomic and plasmid *nrdD* loci (72 and 260 bp at the 3' and 5' ends, respectively of genomic *nrdD*). Yet isolated colonies in reduced soft-agar were also observed for 6-26 (pJFP29) despite the maintenance of the transposon insertion at the genomic locus. Thus it appears that restoration of growth in 6-26 (pJFP29) results

from an uncharacterized mutation restoring expression of chromosomal genes downstream of *nrdD*.

Analysis of the reduced soft-agar growth of an unrelated STM mutant, 3-24 demonstrated that anaerobic growth sensitivities of STM mutants on chloramphenicol plates are circumvented by this culturing strategy. A layer plating technique was used for competitive index studies of STM mutants, with bacteria in soft-agar overlaid with 1.5 % agarose with antibiotic, which negated the problem of chloramphenicol selection by surface plating (198). The apparent sensitivity of streptococci to chloramphenicol selection under anaerobic conditions has not been previously described, and the mechanism of this growth inhibition is unknown. The observation that *S. sanguinis* selection in chloramphenicol under reduced O_2 is possible when strains are embedded in agar allows for reliable characterization of STM mutants.

The objective of the experiments described here was to characterize whether the *S. sanguinis nrdD* gene was required for growth under anaerobiosis, as previously demonstrated for other bacterial species. Anaerobic growth attenuation of JFP27 and 6-26 in reduced soft-agar indicated that *nrdD* expression was essential. Aside from confirming this *in vitro* phenotype attributed to *nrdD*, isolation of 6-26 as an avirulent strain by STM suggests that regions of the vegetation in IE are anaerobic. Analysis of IE by STM potentially would identify different classes of virulence determinants in strains associated with the vegetation depending on the vegetation microenvironment encountered. At the surface of the vegetation bacteria must be able to colonize the platelet-fibrin matrix and contend with host defenses at the blood interface. Within the

vegetation bacteria potentially encounter reduced levels of nutrients and oxygen as a result of bacterial accruing and progressive fibrin stratification (64). To our knowledge oxygen levels within the vegetation have not been measured. A survey of S. *aureus* STM strains in an endocarditis model identified 5 anaerobic mutants that were attenuated *in vivo* (50). Theses studies indicate that the interior of the vegetation is low in oxygen despite the circulating oxygenated blood at the surface of the vegetation, and that expression of anaerobic specific genes (in the case of *S. sanguinis*, *nrdD*) is required for persistence *in vivo*.

DEVELOPMENT OF GENETIC TOOLS FOR ANALYSIS OF *S. SANGUINIS*

Rationale

Among the viridans group streptococci, *Streptococcus sanguinis* is the most common streptococcal agent of IE (181). An *in vivo* screening for *S. sanguinis* SK36 IE virulence determinants by random signature-tagged mutagenesis identified 5 avirulent mutants with transposon insertions in ORFs encoding housekeeping genes or genes necessary for survival under anaerobic conditions (198). The lack of identification of classical virulence factors, e.g., those facilitating adhesion in colonization, was surprising given the abundance of putative virulence factors predicted from the complete SK36 genome sequence (285). Therefore a need exists for development of genetic tools to facilitate further screening of streptococcal factors required for development of disease.

To screen mutants for defects in the rabbit model of infective endocarditis it is best to inoculate a test strain along with a virulent control strain. In this way factors contributing to inter-rabbit variation including pre-infection vegetation size, physiological differences and immune competency can be accounted for, as both strains encounter the same conditions. The co-administration of strains in this experimental design requires selective markers for distinguishing between strains recovered in a

complex pool comprising both the experimental strain of unknown virulence, and the control. It is commonplace to create marked genomic mutations using antibiotic resistance cassettes; however a conundrum exists as to the best method for selection of a control strain. Using a subtractive approach to estimate the recovery of an unmarked control strain may be misleading due to the assumption that both strains have equal plating efficiency. Therefore an attractive alternative would include exclusive selection of the control strain from a multiple strain pool. A control strain with a selectable marker would also be useful for evaluation of unmarked mutant strains, for example an in-frame deletion mutant. This type of comparison is not feasible with SK36 as a control. In addition, the development of a collection of selective marker constructs would permit us to add a selectable marker to an in-frame deletion mutant, at the same locus as the competitive control. This scenario would permit the detection of phenotypes entirely determined by in-frame deletion of a gene of interest.

Results

Identification of *SSA_0169*

The *S. sanguinis* SK36 genome is more compact than published genomes for other streptococcal species, with an average intergenic region of 115 bp (*vs.* average intergenic region lengths of 130 to 177 bp) (285). Given the proximity among predicted ORFs we sought to identify an insertion site with minimal risk of affecting important genes or operons. A 2.225 kb SK36 chromosomal region including a putative 644 bp stretch of non-coding DNA (shown in brackets in Fig. 22) was cloned into the suicide vector pVA2606 to create pJFP34. A unique *Nco*I site within the predicted non-coding region of the insert was employed for cloning of *pSerm*, an erythromycin resistance cassette that is expressed in *E. coli* and streptococci (47). The cassette used is optimal for our strategy as expression of *pSerm* is not reliant on an external promoter. This cassette was engineered by PCR to fuse the -35 region from the native *ami* operon *S. pneumoniae* promoter, *p*A, proximal to the -10 region and 16 bp upstream of pAMβ1 encoded *erm* (47).

The resulting plasmid containing *pSerm*, pJFP36 was transformed into *S. sanguinis* SK36 to generate JFP36. In the initial annotation of the SK36 genome sequence, no open reading frame (ORF) was predicted for the 644 bp region between

SSA_0168 (*0168*) *and SSA_0170* (*0170*) (Fig. 22). However, after creation of strain JFP36, a subsequent GLIMMER analysis predicted the *SSA_0169* (*0169*) ORF. This ORF encodes a 65 amino acid hypothetical protein that lacks homology to any proteins of known function. The predicted ORF is 303 bp downstream from the *0168* gene, and if co-transcribed, *0169* would be the last gene of the operon. The flanking open reading frames, *1068* and *0170*, are also hypothetical.

Transcriptional analysis of *0169*

As *0169* was predicted to encode a small, non-conserved hypothetical protein we wanted to determine whether expression of this ORF occurs. Quantitative (real-time) reverse transcriptase PCR was used to determine whether *0169* transcript was detectible during *in vitro* culture under various oxygen tensions, including aerobic culture, 6 % O_2, and anaerobic culture. We determined that transcription of *0169* is variable under the conditions tested, with the least expression occurring under reduced oxygen or anaerobic conditions that are typical in laboratory culture (Fig. 23). In contrast, ~80-fold more *0169* transcript was detected in *S. sanguinis* SK36 cultured aerobically relative to anaerobic and 6 % O_2 conditions, indicating that expression of *0169* is enhanced during aerobic growth.

An *in vitro* growth comparison of *S. sanguinis* SK36 and JFP36 was then performed to determine whether interruption of *0169* by *pSerm* affected growth under aerobic conditions. As shown in Figure 24, JFP36 and SK36 grew similarly, suggesting that mutation of *0169* did not influence growth of JFP36.

Figure 22. The *0169* locus of *S. sanguinis* SK36

The *0169* locus is shown. The *0169* gene putatively encodes a 65 aa protein. This ORF was mutated for development of *S. sanguinis* SK36 derivatives selectable by antibiotic resistance determinants. In the initial annotation of the *S. sanguinis* SK36 genome sequence *0169* was not identified, and this cloning region was selected based on an apparent 644 bp non-coding region. However, a subsequent GLIMMER analysis did identify the ORF *0169*, which is flanked on by two hypothetical ORFs, *0168* (upstream) and *1070* (downstream). The region shown in brackets (2.225 kb) was cloned into the *E. coli* vector pVA2606 for creation of pJFP34.

Figure 22.

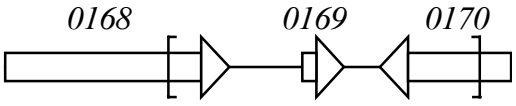


Figure 23. qRT-PCR detection of an *0169* transcript

S. sanguinis SK36 was cultured aerobically (O_2^+), in an anaerobic atmosphere (O_2^-), or in 6 % O_2 for five hours. The amount of RNA transcript of *0169* and 16S rRNA was then determined from RNA extracted from cells grown under the different O_2 tensions. The fold change in transcript abundance relative to the 6 % O_2 condition is shown. RiboGreen fluorescence of nucleic acid was used to normalize samples for differences in the efficiency of reverse transcription. Values depicted represent the mean and SEM 16S rRNA and *0169* transcript levels from three independent samples.

Figure 23.

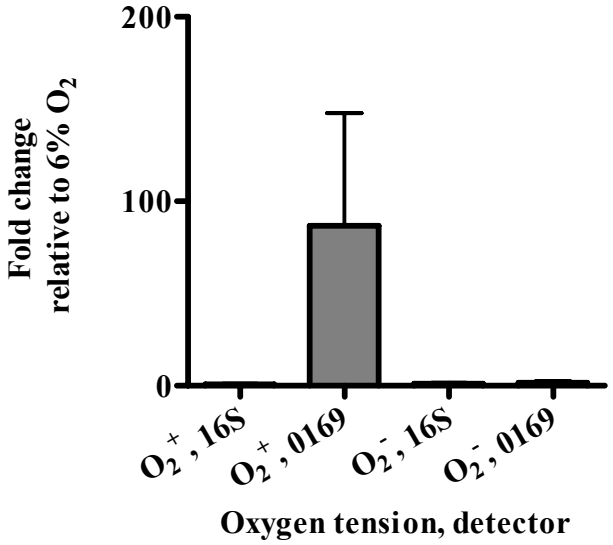
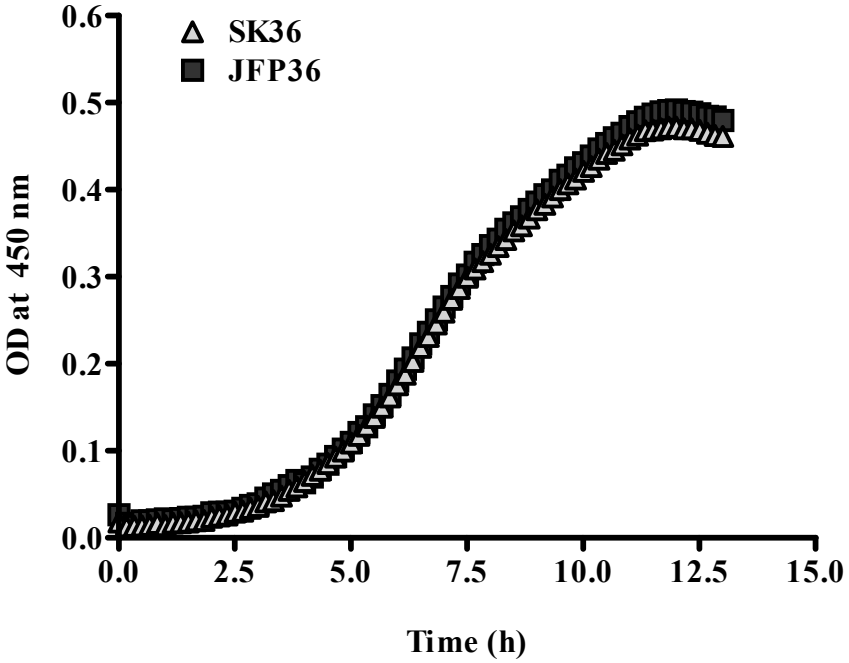

200
100
0
Fold change relative to 6% O_2
O_2^+, 16S
O_2^+, 0169
O_2^-, 16S
O_2^-, 0169
Oxygen tension, detector

Figure 24. Aerobic growth comparison of JFP36 and SK36

S. sanguinis SK36 and JFP36 stationary phase cells were diluted 200 fold in BHI, and the growth of the subcultured cells monitored over a 20 hour period by recording the optical density at 450 nm every 10 minutes. The growth study was performed in a PS 96-well plate, with cells incubated aerobically at 37 ºC. Growth to 13 hours is shown to emphasize the similarity between SK36 and JFP36 in the log growth phase.

Figure 24. Aerobic growth comparison of JFP36 and SK36



Characterization of JFP36 biofilm formation

To investigate the impact of *pSerm* insertion in JFP36 on biofilm development, a previously described semi-quantitative assay was used (76). *S. sanguinis* SK36 served as a positive control in this assay, and its isogenic mutant, VT1614, served as a negative control. In VT1614, *srpA* is interrupted by the kanamycin resistance gene, *aphA-3* (204). The *S. parasanguinis* SrpA homolog, Fap1 was required for wild-type biofilm formation in the presence of 1 % glucose (76). JFP36 biofilm formation was indistinguishable from SK36 (Fig. 25) in the presence of glucose (P= 0.25) and sucrose (P= 0.11), and for both JFP36 and SK36, biofilm formation in 1 % glucose was significantly greater than that of VT1614 ($P < 0.001$). Therefore, mutation of *0169* did not alter biofilm development under the conditions tested.

Genetic Competence of JFP36

The capacity for genetic manipulation is invaluable for studying microbial pathogenesis. *S. sanguinis* is naturally competent and readily transformable when provided extracellular DNA in the presence of CSP and animal serum. One potential future use of JFP36 is the employment of this strain as a genetic background for in-frame deletions in genes of interest. Therefore we sought to compare transformation frequency and efficiency between SK36 and JFP36, to determine whether this application of JFP36 is feasible. The plasmid used as transforming DNA, pJFP16 gives reproducibly high transformation frequencies and efficiencies, and integrates the *cat* gene into the chromosome in single copy. As shown in Table 3, JFP36 was competent

Figure 25. Biofilm comparison of S. sanguinis SK36 and JFP36

S. sanguinis SK36, JFP36 (Δ*0169*), and VT1614 (Δ*srpA*) were grown to early stationary phase in BHI, and then subcultured (1:100) in biofilm medium with 1 % glucose (BM+Glu), or 1 % sucrose (BM+Suc). Strains were grown under static conditions in PS microplates at 37 ºC, 6 % O_2 for 20 hours. Biofilm formation as determined by CV staining was quantified at 600 nm. Assays were performed in triplicate, in three independent experiments, means and SD are shown. ***= $P < 0.0001$

Figure 25.

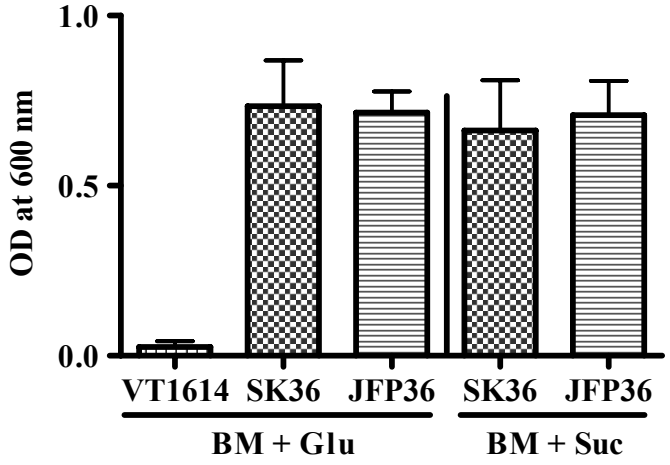


Table 3. Genetic competence comparison of SK36 and JFP36

***S. sanguinis* strain**	**Transformation frequency‡**	**Transformation efficiency¥**
SK36	**27.38 %**	**8.6 x 10^7 CFU μg^{-1} DNA**
JFP36	**27.13 %**	**8.25 x 10^7 CFU μg^{-1} DNA**

‡: The ratio of transformants (Cm^r conferred by homologous recombination of the pJFP16 encoded *nrdD*::*cat*) to total colony forming units
¥: The number of transformants (Cm^r) per µg of pJFP16 transforming DNA

for uptake and recombination of pJFP16 transforming DNA with transformation frequencies and efficiencies similar to wild-type.

Analysis of JFP36 competitiveness *in vivo*.

S. sanguinis JFP36 was then tested for competitiveness in the rabbit model of IE by a competitive index (CI) assay. In this experiment early stationary phase SK36 and JFP36 were co-inoculated into four catheterized rabbits and infection monitored 20 hours post-inoculation. The CI (competitiveness value of JFP36) value was determined by dividing the ratio of JFP36 to SK36 recovered from the rabbit by the ratio of JFP36 to SK36 in the inoculum. A mean CI value of 1.22 (range 0.4 to 1.7) was obtained from the four rabbits. Because a CI value of 1 indicates equal competitiveness between the two strains, this CI value confirms that competitiveness of JFP36 is not significantly different from SK36 (P= 0.52). Therefore JFP36 would serve as a valid control strain in future *in vivo* competitive index assays, and would permit direct discrimination between wild-type and one or more mutant strains on plates containing different antibiotics.

Development of JFP56 and JFP76

Maintenance of competitiveness in JFP36 confirms that manipulation of *0169* is not detrimental *in vivo*. We chose to develop additional novel *S. sanguinis* strains with different antibiotic resistance markers at the *0169* locus to increase the applicability of this tool. In the design of a spectinomycin resistant derivative of SK36, the *aad9*

resistance cassette from pR412 (161) was cloned into a linker central to the *0169* segment of pJFP46 for creation of pJFP56.

Expression of the plasmid pR412 encoded *aad9* is controlled by a strong, synthetic promoter which includes the 5' region of the native *ami* operon pneumococcal promoter, p*A* fused to the 3' region of the *aad9* promoter to permit optimal expression of spectinomycin resistance in *S. pneumoniae* (56) (similar to the promoter of *pSerm*, described above). The pJFP56 insert maintained the synthetic *aad9* promoter adjacent the *aad9* gene. *S. sanguinis* SK36 was subsequently transformed with pJFP56 for identification of the spectinomycin resistant strain, JFP56. Expression of *aad9* at the *0169* locus of *S. sanguinis* was evidenced by successful selection of the JFP56 strain.

A tetracycline resistant *S. sanguinis* strain was also developed by cloning of the *tetM* gene from pJM133 (88) into the linker transecting *0169* in pJFP46 for creation of pJFP76. The class M tetracycline resistance determinant, *tetM* originates from the *Tn*916 composite conjugative transposon of *E. faecalis* (231). In pJM133 *tetM* transcription was controlled by a native *aad9* promoter of the *E. faecalis* plasmid, pD255, *aad9* resistance determinant (205). The promoter controlling expression of *tetM* is constitutively expressed in both *E. coli* and streptococci and therefore ideal for our purposes in this study. *S. sanguinis* SK36 was transformed with the *tetM* containing plasmid, pJFP76, for isolation of the tetracycline resistant strain JFP76. DNA sequencing of *tetM* in the *0169* locus of JFP76 indicated that the actual cassette used shared greater identity with *tetM* of the *S. pneumoniae Tn*1545 determinant (GI:48189) than the *tetM* gene of *Tn916*. The mosaic nature of *tetM* alleles has been previously

investigated by sequence alignment for identification of sequence polymorphisms. It is of interest that allelic variation observed in nature was attributed to differing homologous recombination between two distinct *tetM* alleles of *Tn*916 *S. aureus* and *Tn*1545 of *S. pneumoniae*. The evolution of mosaic *tetM* alleles would have been facilitated by the conveyance of these determinants on mobile conjugative transposon elements (192).

Analysis of JFP56 and JFP76 competitiveness *in vivo* and *in vitro*

An *in vitro* growth study of JFP36, JFP56 and JFP76 was conducted for verification that growth of these strains was similar under our typical laboratory culture conditions. The growth study results shown in Figure 26 confirmed that insertion of different resistance determinants in *0169* had no effect on the growth phenotype of these strains, in that all strains exhibited identical growth patterns in the period studied.

Competitiveness of JFP56 and JFP76 was determined by comparison to the competitive erythromycin resistant strain, JFP36 by the competitive index assay. In past CI analyses our lab has directly compared two strains *in vivo* by co-inoculation into a previously catheterized rabbit (198). In this study we chose to co-inoculate the three strains with different selective makers, JFP36, JFP56 and JFP76 for a direct comparison of strains encountering shared selective pressures *in vivo*. For this purpose the three strains were adjusted to an equal optical density corresponding to ~ 5 x 10^7 to 1 x 10^8 CFUs for co-administration intravenously. The inoculum received by the rabbits was determined to be 6.11 x 10^8 CFUs, with JFP36, JFP56 or JFP76 accounting for 16.9 %,

75.3 % or 7.9 % of the inoculum, respectively. Bacteria were recovered from aortic-valve associated vegetations 20 hours later and plated on TSA containing the appropriate antibiotic for selection of the erythromycin, spectinomycin, or tetracycline resistant strain. The CI values were then determined for JFP56 or JFP76 relative to the control, JFP36 by dividing the ratio of the unknown (JFP56 or JFP76) to JFP36 recovered from the rabbit, by the unknown to JFP36 ratio in the inoculum. The mean CI values determined for JFP56 and JFP76 are depicted and described in Figure 27. JFP76 was determined to be equally competitive to JFP36 with a mean CI value of 0.83 (P= 0.2735), which suggests that JFP76 is also applicable as a competitive control for future studies. The mean CI value for JFP56 of 0.38, suggested that competitiveness of JFP56 was significantly reduced as a result of *aad9* expression at the *0169* locus (P= 0.0470). However, the percentage of JFP56 in the inoculum was much higher than expected and thus indicated that an error in plating may have occurred.

We chose to repeat the direct CI comparison of JFP56 and JFP36 by intravenous co-inoculation of the two strains into catheterized rabbits. In this experiment the rabbits received a total inoculum of 1.28×10^9 CFUs of which 52.3 % was JFP56, and 47.7 % JFP36. The competitiveness of JFP56 was determined by recovery of bacteria associated with aortic valve vegetations at 20 hours post-administration of strains, as described above. The mean CI value for JFP56 relative to JFP36, 0.96, indicated that JFP56 was as competitive as the wild-type strain (P= 0.6932) (Fig. 27).

Figure 26. *In vitro* growth comparison of JFP36, JFP56 and JFP76

Early-stationary phase cultured cells of *S. sanguinis* JFP36, JFP56 and JFP76 were adjusted to an equal optical density at 660 nm, and then subcultured (1:100) in BHI for an overnight growth assay. The growth study was performed in a 96-well PS plate, at 37 ºC, with the optical density at 450 nm recorded every 10 minutes. Growth to 12 hours is shown to emphasize the similarity between the log growth curves of JFP36, JFP56 and JFP76.

Figure 26.

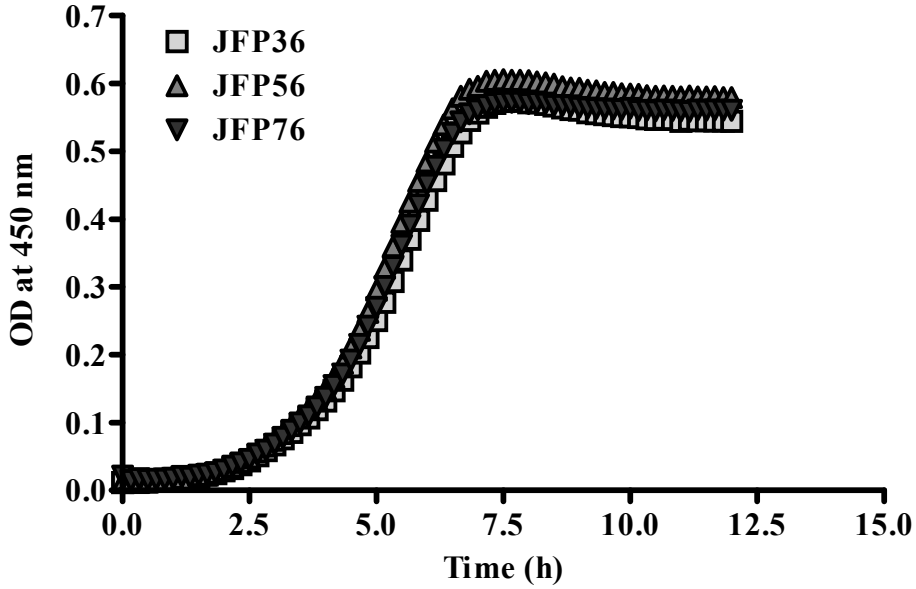
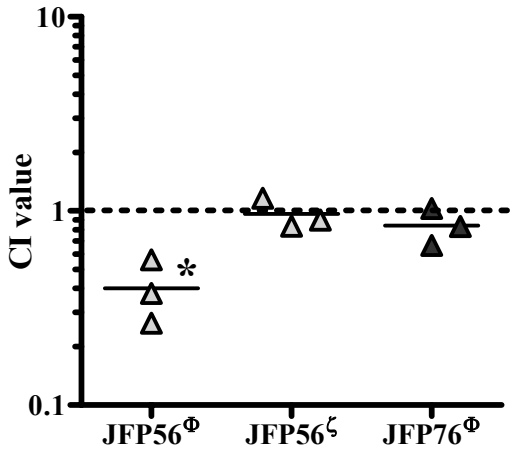

JFP36
JFP56
JFP76
OD at 450 nm
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.0
2.5
5.0
7.5
10.0
12.5
15.0
Time (h)

Figure 27. CI comparison of JFP56 and JFP76 to JFP36

S. sanguinis strains expressing different resistance determinants at the *0169* locus were compared for competitiveness in the rabbit model of IE to determine whether the strains could serve as virulent controls in future experiments. In two separate experiments JFP56, JFP76 and JFP36 were compared by co-inoculation of the three strains (Φ), or co-inoculation of JFP56 and JFP36 only (ζ) into previously catheterized rabbits. Rabbits were sacrificed 20 hours later for recovery of bacterial strains by plating on selective antibiotics. The CI value for the two strains of unknown competitiveness, JFP56 and JFP76 was determined by comparison to the control competitive strain, JFP36 for each rabbit. In the graphical representation the dotted line at 1 indicates the value for equal competitiveness. Each symbol represents a CI value from one rabbit, and the bar indicates the mean CI value. In the triple strain comparison JFP56 was significantly less competitive than JFP36 (*= $P < 0.05$), whereas JFP76 was equally competitive to JFP36. Reevaluation of JFP56 resulted in a CI value indicative of equal competitiveness. The actual mean CI and range of value observed for each strain are shown in the table.

Figure 27.



Strain	Phenotype	CI (Range)
JFP56$^{\Phi}$; JFP56$^{\zeta}$	Less Competitive; Competitive	0.38 (0.26 - 0.56); 0.96 (0.84 - 1.27)
JFP76$^{\Phi}$	Competitive	0.83 (0.66 - 1.03)

The evenly divided inoculum ratio in the JFP56 and JFP36 CI experiment provided these strains an equal opportunity for colonization and growth, and would therefore result in a more reliable CI score. Even combining the CI values from the two experiments evaluating JFP56 does not result in a mean CI value (0.6) significantly different from an equal competitiveness value of 1 (P= 0.0810).

From these experiments we concluded that competitiveness of JFP56 and JFP76 *in vivo* was unaffected by the resistance determinant used, and these strains, in addition to JFP36 could act as controls for screening and identification of mutants attenuated in the rabbit model of IE.

In a final set of experiments an *in vitro* competitiveness assay was used to compare the growth of JFP56 or JFP76 to JFP36 during co-culture. This type of experiment indicates whether *in vivo* findings may be construed by phenotypes unrelated to *in vivo* conditions, instead owing to effects of co-culturing of strains or general differences in the growth rates of strains. For this experiment strains were adjusted to an equal optical density and then co-cultured *in vitro* for the typical infection duration (~20 hours). The *in vitro* CI value was derived by comparing CFUs of strains following co-culture, to CFUs before co-culture ("the inoculum"); with a CI value of 1 suggesting strains grow equally well during co-culture *in vitro*. The mean *in vitro* CI value for JFP56 or JFP76 relative to JFP36 was determined by two independent experiments performed in triplicate. The CI values obtained for JFP56 and JFP76, 1.31 and 1.51, respectively suggest that co-culturing of these strains with JFP36 *in vitro* results in growth significantly greater than JFP36 (P= 0.0086, P= 0.0005, for JFP56,

and JFP76). The observed *in vitro* phenotype was attributed to a greater number of JFP56 or JFP76 CFUs recovered following co-culture, as the ratio of JFP56 or JFP76 to JFP36 in the inoculum was similar to that observed during *in vivo* analysis. The shared core genome of these strains suggests that difference in growth *in vitro*, as determined by CFU formation is an artifact of the selective marker used. The similar *in vivo* CI values for JFP36, JFP56 and JFP76 (discussed above) indicate that additional selective pressures encountered *in vivo* result in comparable growth of these strains in the animal model for a recovery to inoculum ratio closer to 1.

Discussion

We intended to develop a system for integration of selective markers into the chromosome of *S. sanguinis* SK36, while maintaining WT phenotypes of this species. A scan for non-coding sections of the *S. sanguinis* genome sequence prompted the use of a predicted intergenic region separating two opposing, hypothetical ORFs, *0168* and *0170*. Following the development and preliminary characterization of JFP36 an ORF was identified for this intergenic region, *0169*, that was inactivated by insertion of the erythromycin resistance gene. The *0169* ORF putatively encodes a 65 amino acid protein with no homology to previously characterized proteins. In addition, the *0169* locus is non-conserved among closely related streptococcal species (Integrated microbial genomes resource (http://img.jgi.doe.gov/cgi-bin/pub/main.cgi)). Despite the novelty of this ORF, and indications that *0169* was a pseudogene, a *0169* transcript was detectible by qRT-PCR. The transcriptional analysis of *0169* confirmed expression of the ORF was greatly increased during aerobic culture, which is an atypical condition for typical laboratory growth of *S. sanguinis*. The unintended manipulation of *0169* did not have any consequence on relevant *in vitro* phenotypes of *S. sanguinis*, including doubling time during *in vitro* growth, biofilm formation, and competence for DNA transformation. Most important for the objective of this work, JFP36 was equally

competitive to SK36 *in vivo*, and therefore appropriate for use in future competitive index assays.

The proven maintenance of virulence in JFP36 resulted in the development of two additional derivatives of SK36 with unique resistance markers, JFP56 and JFP76. Competitive index comparisons of JFP56 and JFP76 in the rabbit model of IE confirmed both exhibited the virulent phenotype. The *in vitro* growth comparison of JFP36, JFP56 and JFP76 suggested that doubling times of these strains were similar. Contrary to our expectation *in vitro* co-culture of JFP36, JFP56 and JFP76 resulted in different recovery ratios of the three strains, with growth of JFP36 apparently attenuated during co-culture. This phenomenon is most likely caused by the antibiotic determinant gene, as this is the only substantial difference between the strains. However, the effect is independent of selective expression of the resistance marker since no selective antibiotics were present in the liquid growth medium at the time of co-culture. One possible explanation is that expression of the resistance determinant may affect cell viability at greater cell densities indicative of stationary growth phase cells. In the infective vegetation streptococci may reach a cell density of 10^8 to 10^{10} within 20 hours, which is similar to the streptococcal density per ml following overnight culture. However, selective pressures encountered *in vivo* may result in slower growth rates than *in vitro*, so that cells spend less time at stationary phase prior to plating for enumeration of surviving streptococci. If viability for a strain is decreased in stationary phase, then less of that strain would be recovered by plating. Unfortunately the distinguishing host

pressures that might influence growth of these strains can not be replicated in the laboratory.

The genomic tools developed in this chapter will enhance the screening and identification of novel determinants for IE. Since it is often desirable to insert antibiotic resistance markers and other exogenous genes into the chromosome, potential future applications include exploitation of the *0169* chromosomal site for ectopic expression of native genes to allow for complementation of mutations elsewhere in the genome.

A COMPREHENSIVE ANALYSIS OF *S. SANGUINIS* CELL WALL-ASSOCIATED PROTEINS IN EARLY IE

Introduction

Cwa proteins as mediators of pathogenesis

The Cwa protein class of surface exposed proteins was first identified by detection of a common C-terminal LPXTG (aa abbreviations are described in Appendix C) sorting motif among 6 cloned gram-positive surface proteins (72). The prevalence of this protein class in gram-positive species is evidenced by the identification of proteins containing this motif in nearly all published genome sequences (25, 49). Cwa proteins are considered particularly important in pathogenesis as many have been grouped as MSCRAMMs mediating adhesion to adhesive matrix molecules of host tissue and host cells. Previously described streptococcal MSCRAMM adhesins include the *S. pyogenes* M family proteins Emm, Mrp and Enn (reviewed by 185), and viridans streptococcal antigen I/II, and Csh homologs (152, 169). Enzymatic functions have also been ascribed to multiple streptococcal Cwa proteins. *Streptococcus pyogenes* C5a peptidase and the *S. pneumoniae* neuraminidase modulate host immune components and sialylated proteins, respectively to putatively enhance pathogenesis (191, 279). Cwa proteins of *S. mutans*, β-d-fructosidase and dextranase are considered important for nutrient storage and acquisition (24, 48).

Cell surface localization of Cwa proteins

Cwa proteins are typified by two signal sequences that direct, first the extracytoplasmic export of the precursor protein, and second the covalent linkage of the protein to the cell wall. Extracytoplasmic trafficking and export of the precursor Cwa protein is specified by a gram-positive type I N-terminal signal (leader) peptide sequence. The leader peptide includes an N-terminal region with positively charged amino acids, followed by a central hydrophobic region, and then a C-terminal region containing hydrophilic residues (reviewed in 266). The presence of the leader peptide drives Sec-dependent secretion of the precursor protein through a translocase channel. The mediators of Sec secretion have been determined for *E.* coli, and partly characterized in *B. subtilis*. *S. sanguinis* homologs to proteins described below are indicated parenthetically.

In the putative system for translocation in *Bacillus subtilis*, based on *E. coli*, export is considered co-translational. This mode of transport in *E. coli* initiates with the signal recognition particle (SRP: Ffh (SSA_1167) + 4.5S RNA) binding to the leader peptide of the nascent polypeptide as it emerges from the ribosome. Binding of SRP to the polypeptide arrests protein translation until association with FtsY (SSA_1557), and subsequent targeting of the complex to the Sec translocase at the cytoplasmic membrane. Following association with the translocase, translation resumes as the protein is transported across the membrane through a SecYEG heterotrimer (SSA_0127, SSA_2208, and SSA_1604 respectively) translocase channel. The SecA ATPase (SSA_0543) associates with the channel to couple the energy from ATP binding and

hydrolysis to the stepwise translocation of the Cwa protein precursor (reviewed in 288). In *E. coli,* SecD and SecF are large integral membrane proteins that form a heterotrimeric complex with YajC and associate loosely SecYEG in a supramolecular translocase complex (62). In *B. subtilis*, SecD and SecF are fused into a single membrane protein that potentially contributes to the efficiency of protein secretion (26). As described for *S. mutans*, S. *sanguinis* SK36 lacks a SecDF homolog, yet does encode a YajC-like protein (SSA_1792) (91).

As the precursor protein emerges from the translocase, the type I leader peptide is recognized and cleaved by a membrane spanning type I signal peptidase (SSA_0351, SSA_0849). The mature N-terminus of the Cwa protein is then exposed, and the remaining signal peptide degraded. Recent studies in *S. pyogenes* have revealed that the leader peptide also directs the cell wall location of the Cwa protein. For example, the leader peptide of the M protein directs localization at the cell division septum, whereas the protein F leader peptide results in emergence of the protein at the old division pole (39).

Studies concerning the cellular location of the Sec machinery in *S. pyogenes* resulted in the coining of the term “ExPortal” to describe the translocase localization to a single site at the extreme pole of the cell (220). This is distinct from the reported helical pattern assumed by Sec machinery on the cell surface of *B. subtilis* (37). Immunogold labeling suggests that in *S. mutans* SecA and the transpeptidase enzyme, sortase A, required for Cwa protein surface anchoring, are co-localized at a single

position on the cell surface. Therefore the site of protein translocation may also be the site of protein anchoring by sortase A (107).

An accessory *sec* locus with *sec* homologs, *secA2* (SSA_0836) and *secY2* (SSA_0832) was identified in *S. gordonii*. These proteins mediate the selective transport of a single serine-rich Cwa protein, GspB. In concert with specialized transport, GspB contains an unusually long leader peptide (17). The accessory *sec* locus of *S. gordonii* is conserved in *S. sanguinis*, and is adjacent to a serine rich protein in this species as well. Therefore the specialized export of this protein is likely shared between these species.

Following Sec-dependent export, the C-terminal sorting signal of the Cwa protein drives sortase transpeptidase-dependent anchoring of the protein to the cell wall. The sorting signal is ~40 amino acids in length and consists of a tripartite motif with an N-terminal LPXTG-related pattern, followed by a hydrophobic domain, and a positively charged tail at the extreme C-terminus of the protein (reviewed in 255). It is thought that the hydrophobic region and positively charged terminus retain the Cwa protein in the membrane until association with sortase occurs. Evidence of intermediate membrane association is in the observation that mutation of LPXTG abolishes cell wall linkage; however the protein is not secreted from the cell, as has been observed with complete loss of the tripartite motif (227).

Genomic sequence scanning of gram-positive organisms has revealed that almost all encode sortase-like proteins, and many encode multiple sortase paralogs (49, 199). All sortase homologs contain a signature active site motif, LXTC, with Leu, Thr

and Cys residues conserved (109). A comprehensive analysis of 241 bacterial genomes (96 species of bacteria) resulted in the grouping of similar sortase-like proteins (determined by BLAST scores, and confirmed by Hidden Markov Modeling (HMM)). Six sub-family groups were proposed in all, including five occurring in gram-positive organisms: subfamily A sortases (referred to here as the class A sortases, and sortase A), subfamily B sortases (referred to as class B sortases, and sortase B), subfamily 3 sortases (referred to below as sortase C, SrtC), subfamily 4 sortases, and subfamily 5 sortases; as well as a sixth subfamily unique to gram-negative organisms (49). Putative substrates of the sortase subfamilies could be delineated based on proximity to a putative requisite sortase and the observation of distinct cell wall sorting motifs within these groups. The majority of putative cell wall-associated protein substrates rely on a "housekeeping" sortase of the sortase A (SrtA) class for cell wall anchoring. Genes encoding SrtA are ubiquitous among low G+C content gram-positive bacteria (59).

The most well characterized sortase family protein, SrtA of *S. aureus*, catalyzes the covalent linkage of the Cwa protein to the lipid II precursor of peptidoglycan (PG) in a reaction sensitive to the cell wall synthesis inhibitory antibiotics vancomycin and moenomycin (258).

The specific structure of PG differs between species, but several elements are conserved, including the PG glycan strands composed of repetitive disaccharide units, *N*-acetylglucosamine-(β1-4)-*N*-acetylmuramic acid (GlcNAc-MurNAc) (80). In mature PG the glycan polymers are adjoined to peptide chains (in *S. sanguinis* the peptide moieties include $_{L}$-Ala-$_{D}$-iGln-$_{L}$-Lys-$_{D}$-Ala; 19), which are cross-linked by short cell

wall peptide cross-bridges. Biosynthesis of PG initiates in the cytoplasm of the cell with MurNAc and the $_{L}$-$_{D}$-peptide chain associating to form Park's nucleotide. This Park's nucleotide precursor is then linked by a phosphodiester linkage to the membrane bound carrier, undecaprenol phosphate to create the membrane bound lipid I form. Addition of GlcNAc and cross-bridge modification at the ε-amino of $_{L}$-Lys (the crossbridge in *S. sanguinis* is dialanine, Ala_2; 19) generates the lipid II precursor. Lipid II is then translocated across the cytoplasmic membrane.

In the transpeptidation reaction catalyzed by SrtA a covalent linkage occurs between the Cwa protein and the lipid II precursor of PG. The sortase enzyme cleaves the LPXTG pattern between the Thr and Gly residues. The conserved SrtA active site Cys residue captures the cleaved polypeptide as a thioester-acyl enzyme intermediate. Catalysis is completed by the nucleophilic attack of the amino group of (in the case of *S. sanguinis*, Ala_2) the crossbridge of lipid II. Penicillin binding proteins then perform transglycosylation reactions to extend the glycan chain, and transpeptidation reactions that increase crossbridge crosslinking (between the peptide $_{D}$-Ala at position 4 and the amino group of Ala_2) to strengthen the PG (258). These final reactions yield the mature gram-positive cell wall.

Sortase A structure and function

Sortase A is a type II transmembrane protein with an N-terminus rich in hydrophobic residues, which serves as a signal peptide directing extra-cytoplasmic export of the precursor protein; and as a stop-transfer signal for anchoring to the

cytoplasmic membrane. Certain features of the sortase protein sequence are not conserved among other class A sortases, including the signal peptide, membrane anchor and a short linker domain extending from amino acid 26-59 (109). Functional studies of N-terminally truncated *S. aureus* SrtA, $SrtA_{\Delta N59}$ revealed that the first 59 residues were not required for cleavage of a *d*-LPETG-*e* fluorophore (*e*) substrate and, therefore are unlikely to contribute to SrtA function *in vitro*. In contrast residues 60-206 compose the catalytic core of sortase A and are conserved among gram-positive species. A recombinant SrtA protein excluding the non-conserved N-terminal residues (1-59) was used to elucidate the structure of this enzyme by nuclear magnetic resonance (NMR). SrtA forms an 8-stranded β-barrel fold with 2 short helices and several loops (109). The active site of the protein is arranged around the catalytically essential thiol containing side-chain of Cys-184, and includes the conserved residues LXTC. Strands β6 and β 7 of SrtA form the floor of a hydrophobic depression, with walls formed by loops connecting sequential β strands (109). This hydrophobic pocket encompasses the protein catalytic site, and accommodates the hydrophobicity of the sorting signal (LPXTG) substrate (257).

The *in vitro* hydrolytic and transpeptidation activities of sortase are stimulated by Ca^{2+}, and are partially activated by alternative divalent cations Mn^{2+} and Mg^{2+} (182). This dependence on extracellular Ca^{2+} is functionally relevant as *S. aureus* would encounter extracellular calcium concentrations of 1.5 mM or greater during entry into human tissues and roughly 2.5 mM Ca^{2+} in human serum, suggesting that the rate of

sortase catalyzed anchoring of surface proteins would increase during the onset of infection (109, 182).

Seminal studies assessing the contribution of sortase to virulence in an animal model were all performed using *S. aureus* SrtA (163, 254, 257). Expression of SrtA has also been linked with virulence and colonization phenotypes in multiple gram-positive pathogens through *in vivo* and *in vitro* analyses (42, 200, 287). As the defect in previously characterized *srtA* mutants is in cleavage and anchoring of surface proteins, and not protein-secretion preceding the sortase step, protein-lipid II precursors accumulate. A sub-cellular localization experiment detecting SrtA protein substrates revealed that precursor Cwa proteins accumulated in cytoplasmic, membrane and cell wall fractions of an *S. aureus srtA* mutant, whereas Cwa proteins were only detected in the cell wall fraction of wild-type staphylococci (163). In a similar comparison with *S. pyogenes* studying M protein localization, M protein was detected in the cell wall and supernatant fraction of the wild-type strain, and exclusively in the supernatant of a *srtA* mutant derivative strain (12). Though cytoplasmic membrane secretion of SrtA substrates may be intact, without a functional SrtA transpeptidase, SrtA-dependent adhesins exhibit a loss of function *in vitro* (163). SrtA has recently emerged as a protective vaccine candidate in a study evaluating an *S. aureus* polyprotein DNA-vector vaccine combining the extra-membrane portion of SrtA with the fibronectin and fibrinogen binding portions of FnbB and ClfA. The polyprotein vaccine caused a polarized Th1 effect and significantly increased survival of mice challenged intravenously with *S. aureus* (78).

Sortase B attributes

In *S. aureus* and *Bacillus anthracis*, SrtB is required for anchoring a heme-iron scavenging protein, IsdC (157, 165). In both species SrtB and IsdC are encoded in an operon with additional genes composing a putative iron acquisition system. The operon is under the control of the ferric uptake repressor, Fur. Under iron replete conditions Fur inhibits transcription of these genes. In *S. aureus* the SrtB substrate, IsdC contains the sequence NPQTN in place of the typical LPXTG sorting signal (the *isdC* homolog in *B. anthracis* contains NPKTG) (165). Two additional iron surface determinants in the *S. aureus* operon, *isdA* and *isdB* encode proteins contain the SrtA LPXTG sorting signal, and are anchored by SrtA. The susceptibility of these different surface determinants to extracellular proteases revealed that sortase substrates of the *isd* locus are differentially surface exposed (164). Both IsdA and IsdB are susceptible to proteolytic cleavage; whereas IsdC is susceptible to proteinase K degradation only after lysostaphin treatment of the cell wall (159, 164). The model of Isd mediated heme-iron uptake includes IsdA, IsdB and another SrtA substrate IsdH acting as receptors for hemoprotein ligands, followed by passage of heme through the cell wall via IsdC (reviewed in 158).

Biochemical analysis of the anchor structure of IsdC of *S. aureus* revealed that the NPQTN sorting signal is also cleaved between the Thr and Gly residues, and then covalently linked to the Gly_5 crossbridge, as has been described for SrtA substrates. However IsdC was linked to PG fragments with less crosslinking than typically observed for SrtA candidates. Marraffini and Schneewind proposed that this observation, coupled with non-uniform cell surface distribution of the IsdC is attributed

to linking of IsdC to an assembled cell wall fragment, rather than the lipid II precursor of PG (160).

The structures of *S. aureus* and *B. anthracis* SrtB proteins have been determined, and are very similar. Like SrtA, the core of the *S. aureus* SrtB enzyme is formed by an eight stranded β-barrel. The β6/β7 loop, discussed above imparts the substrate specificity of *S. aureus* SrtA and SrtB (18). Evaluation of *L. monocytogenes* and *S. aureus* SrtB in pathogenesis suggests that this protein is not required for establishing infection, but does contribute to persistence in infected tissues (165, 187).

Sortase C attributes

Like SrtB, SrtC is encoded by a subset of gram-positive species, and has been characterized as the mediator of gram-positive surface pilus formation in important gram-positive pathogens including *Actinomyces naeslundii*, *Corynebacterium diphtheriae*, *S. pyogenes*, *Streptococcus agalactiae*, *S. pneumoniae*, and *Enterococcus faecalis* (13, 58, 155, 173, 183, 256). Gram-positive pili are encoded by *cwa* genes typically clustered in an operon with a class C sortase, which is responsible for polymerizing the pilin subunits on the gram-positive cell surface (175, 229). Alignment of the pilus locus Cwa protein substrate sequences led to the identification of a two conserved motifs additional to the N-terminal signal peptide, and C-terminal sorting signal. Pilin protein conserved motifs include a central “pilin motif” that contains a conserved Lys residue in the consensus sequence WxxxVxVYPK (259), and an “E box” with the amino acid residues YxLxETxAPxGY (256). Analysis of the *C. diphtheriae*

SpaA precursor demonstrated that the Lys of the pilin motif was essential for pilin polymerization. In the suggested model for polymerization a *C. diphtheriae* SrtC homolog cleaves between the Thr and Gly residues of the LPXTG pattern to create an acyl enzyme intermediate, as described above. The ε-NH_2 of the conserved Lys then performs nucleophilic attack at the sortase acyl enzyme for crosslinking of two adjacent SpaA subunits by a transpeptide bond (158); repetition of this process would yield the elongated pilus structure.

In species examined thus far Cwa proteins clustered with the cognate SrtC are differentially assembled into the pilus structure. A major pilin subunit constitutes the backbone of the pilin shaft, and contains the pilin motif. The E box mediates association of the pilus backbone with decorative, accessory Cwa subunits also encoded in the pilus locus (256, 259). These accessory pilin proteins may associate with the pilus structure at the pinnacle of the pilus, or along the length of the pilus (229).

In *C. diphtheriae* analysis of sortases mediating Cwa pilus assembly and cell surface anchoring of pilus structures revealed that the major housekeeping sortase (class A sortase) SrtF can catalyze the cell wall anchoring of both pilin monomers and pili, but cannot polymerize pilin subunits. Modulation of SpaA pilin subunit expression revealed that SrtF is critical for pilus anchoring when SpaA is abundant (244). These findings suggest that multiple sortases are involved in gram-positive pilus assembly, and the transpeptidases may be hierarchal in establishing the functional, surface-exposed pilus.

Predictive patterns of Cwa proteins

Essential residues for specific activity of *S. aureus* SrtA were determined using Dnp chromophore-linked A-L-P-X-T-G sorting pattern pentapeptides, with all naturally occurring amino acids (excluding C and W) tested at each underlined position (133). The authors concluded that alternative residues at positions 1, 2, and 4 were minimally tolerated; and G at position 5 essential for catalysis. Therefore the kinetically preferred substrate for *S. aureus* SrtA contains the L-P-X-T-G pattern (133). This specific pattern with few alternative residues permitted has been adapted for identification of novel SrtA substrates within translation products of sequenced genomes of *S. epidermidis* RP62A, and *S. aureus* strains by *in silico* methods (28, 219).

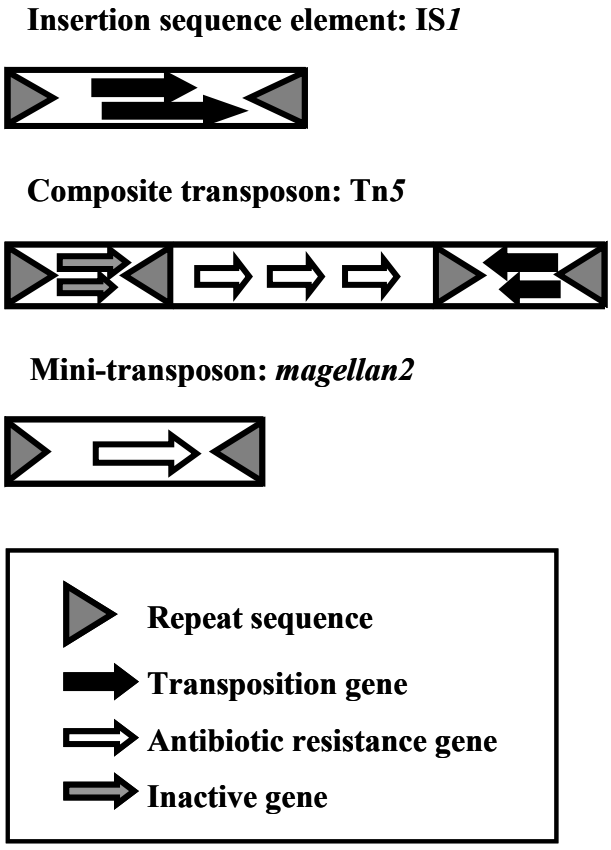
A survey of 65 published cell wall-associated proteins from multiple gram-positive species confirmed the essentiality of L-P at positions 1 and 2, as these were conserved in all identified sequences; however, a variety of amino acids were observed in positions 3 to 6. That sequences analyzed were derived from diverse genetic origins suggests flexibility in positions 5 and 6 that may be species dependent (111). Janulczyk and Rasmussen adapted their observations of conserved residues within the L-P-X-T-G pattern, spacing and relative frequency of residues in the hydrophobic domain, and the percent occurrence and positioning of C-terminal positively charged residues for development of a predictive tripartite pattern less stringent than classical patterns. This recognition of variability within the tripartite pattern resulted in a pattern search tool applicable for prediction of previously uncharacterized proteins in numerous gram-positive species (111).

Transposable elements

Transposons are genetic elements capable of mobilizing between DNA sequences. Transposon elements have been identified in the genomes comprising Eubacteria, Archaea and Eukarya (93). In addition to contributing to natural genomic diversity, transposons also have vast applications in molecular biology. In bacterial species transposons have been adapted for insertion mutagenesis techniques on the genomic scale, as well as for specific analysis of individual genes.

In general, transposons are composed of discrete DNA segments, and the complexity of these regions determines the transposon type. The most simple transposable elements, referred to as insertion sequence elements include gene(s) mediating transposition (mobilization of the transposon) flanked by inverted repeat (IR) sequences- which are recognized by the transposase (Tnp) enzyme during transposition (Fig. 28). In contrast, composite transposons encode additional genes, like antibiotic resistance cassettes. An example composite transposon is Tn*5* (Fig. 28). Tn*5* contains two inversely oriented copies of the IS*50* element separated by genes encoding resistance to Kn, bleomycin, and streptomycin. Conjugative transposons display hybrid properties of transposons, plasmids, and bacteriophages. These transposons can mediate their own transfer from a donor DNA molecule in one bacterial cell to a target molecule in another cell. For example, the conjugative transposon Tn*916* excises from donor chromosomal DNA to form a covalently closed circular molecule that is passed to a recipient cell via conjugal transfer (38). Modified transposons, like the *mariner*-based mini-transposon, *magellan2*, which does not encode Tnp, have also been developed for

Figure 28. Genetic organization of transposons



researcher-controlled transposon mobilization (Fig. 28) (139).

In vitro transposition and *mariner-Himar1* transposons

In the most simplistic format *in vitro* transposition requires only the transposon terminal inverted repeats, purified Tnp, target DNA, and a simple reaction buffer providing essential cations. The efficiency of this technique depends on the insertion site preference of the Tnp, and the catalytic activity of the Tnp used. Native Tnps have evolved for inefficient transposition as a mechanism for maintenance in their resident genome. For example, the wild-type Tnp of Tn*5* does not perform transposition *in vitro*, and is extremely inefficient *in vivo* (213). To circumvent the inactivity of this Tnp, researchers have identified multiple classes of hypermutations that increase catalysis by orders of magnitude. The sites of mutagenesis enhancing transposition have indicated that in part Tnp efficiency is a result of greater protein flexibility for forming DNA contacts, and stabilization of the DNA Tnp complex, once formed (275, 280, 293, 294). Factors affecting *in vitro* transposition of the *Himar1*-type *mariner* transposon have been well characterized. *Himar1* can execute transposition *in vitro* using only its purified Tnp, in the presence of Mg^{2+} or Mn^{2+} (138). The Tnp encoded by these *mariner* elements belong to an extended superfamily of transposases (including Tn*5*) and retroviral integrases characterized by conserved positively charged residues, D,D35E or D,D34D in the catalytic domain of the protein (Fig. 29). As discussed for Tn*5*, the *mariner* Tnps have evolved to be less deleterious in their host via decreased activity. For *in vitro* application Lampe et al. identified hyperactive forms of *Himar1* Tnp by

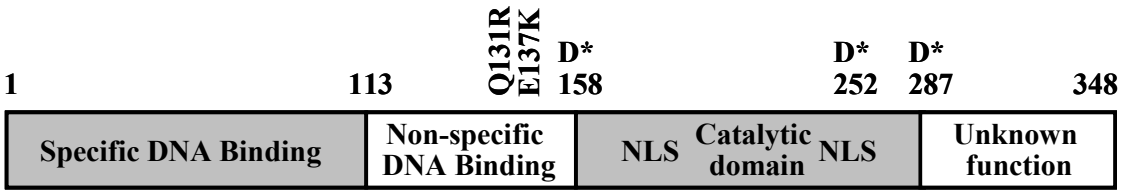
error-prone PCR of the transposase gene (137). A derivative of *Himar1*, termed MarC9 with ~50 fold greater transposase activity was identified with two mutations governing enhanced catalysis (reproduction from reference 137, shown in Fig. 29). The mutations, Q131R and E137K in MarC9 were mapped to the domain of Tnp supposed to mediate nonspecific DNA binding, prompting the authors to speculate that enhanced MarC9 activity may be attributed to a general increase in affinity for DNA(137).

The *mariner* transposon elements are of the short-inverted terminal repeat (ITR)-type. These elements make up a large family of related transposons widely distributed among metazoans (218). More than 13 subfamilies of *mariner* elements have been described; within subfamilies elements have highly conserved sequences with 25-100 % protein sequence identity (217). Despite this sequence similarity, Tnp of the different *mariner* elements almost exclusively act on the ITR of their transposon, in a mechanism that is thought to be entirely sequence specific. For example, the two known active *mariner* elements, *Mos1* and *Himar1* ITRs share 33 % identity, yet their respective Tnps can not transpose the other (140). *Himar1* Tnp has been shown to preferentially transpose shorter transposons bearing the consensus ITR, with transposition frequency decreasing by ~38 % for every 1kb increase in size. The *Himar1* Tnp specifically targets TA dinucleotides for transposon insertion in a manner almost entirely independent of nucleotides flanking the target site. However, *in vitro* studies have revealed that *Himar1* does preferentially transpose into certain sites within a DNA target, suggesting that there is a preference for flanking sequences that are more malleable to the Tnp (139). It is thought that greater malleability of the nucleotide

Figure 29. Schematic of the *Himar1* Tnp derivative MarC9

Functional domains of *Himar1* including the putative DNA binding domains, and catalytic domain, and amino acids separating these regions are shown. D* denotes the putative catalytic residues of the D,D34D catalytic domain. Amino acids changes Q131R and E137K conferring hyperactivity to MarC9 are indicated. NLS indicates nuclear localization signals pertinent to transposition *in vivo* in eukaryotic hosts.

Figure 29.


Q131R
E137K
D*
D*
D*
1
113
158
252
287
348
Specific DNA Binding
Non-specific DNA Binding
NLS
Catalytic domain
NLS
Unknown function

sequence adjacent to the target site enhances recognition of TA nucleotides by Tnp. The TA preference exhibited by *Himar1* Tnp is cation dependent. In the presence of Mg^{2+} the Tnp targets TA sites; however when Mn^{2+} is supplied, insertion site specificity is more lenient and includes G and C nucleotides as well (138).

The *Himar1* element has been advanced as a genetic tool for several reasons alluded to above. The Tnp of *Himar1* displays minimal target site requirements for ensured mutagenesis of the target DNA template. This reaction is made all the more efficient by the application of a hyperactive form of the *Himar1* Tnp resulting in a high frequency of transposition. Also, seminal studies by Robertson and colleagues analyzing *Himar1* transposition *in vitro* confirm the reaction can be entirely researcher controlled with the use of purified Tnp and a *mariner* derived mini-transposable element (not encoding its own Tnp) (138, 139).

Signature-tagged mutagenesis (STM)

The technique of STM was conceived and developed by David Holden for analysis of *Salmonella typhimurium* (95). STM is a transposon mutagenesis system that employs a library of transposons containing unique DNA sequences i.e. “tags”. Traditionally the transposon library has been used for random insertion mutagenesis of bacterial species. The resulting pool of insertion mutants can then be tested in a species-relevant model with recovery/persistence of a specific mutant detected via the unique DNA tag. When applied to *in vivo* studies of bacterial pathogenesis STM permits the high throughput screening of a mixed population of mutants. This type of negative

selection screening identifies factors required for virulence and has the advantage of identifying of previously uncharacterized virulence determinants.

The signature-tags are synthetic oligonucleotides designed with a 40 bp variable central region bordered by invariant ends. The tags were designed to exclude recognition sites of restriction endonucleases used for subsequent cloning of tags into the transposon element, while maintaining enough sequence variability that an identical sequence would occur once in 2 $x10^{17}$ molecules (95).

The initial STM screening studies employed the gram-negative species *S. typhimurium* and a mini-Tn*5* tagged transposon library (95). This mini-Tn*5* transposition system has been subsequently applied to various γ-proteobacteria pathogens, as the transposon is most active in this genetic background (93). STM has additionally been applied to gram-positive pathogens, albeit with alternative transposon delivery systems (Tn*917*, Tn*1545*, IS*1096*) originating from these bacteria (8, 36, 112). The indiscriminate *mariner-Himar1* mini-transposon *magellan2* has also been successfully employed for STM in gram-positive streptococcal species (92, 198).

In the traditional application of STM a suicide plasmid encoded mini-Tn*5* transposon library with unique signature-tags was electroporated into *E. coli* and then transferred into the strain under scrutiny via conjugation. The resulting transconjugants were selected by acquired resistance markers conferred by genomic insertion of the transposon (73, 94, 209, 296). As some organisms are refractory to transposon mutagenesis, and in particular to the Tn*5* system, the versatility of this method was expanded with the application of *in vitro* mutagenesis of target DNA, followed by

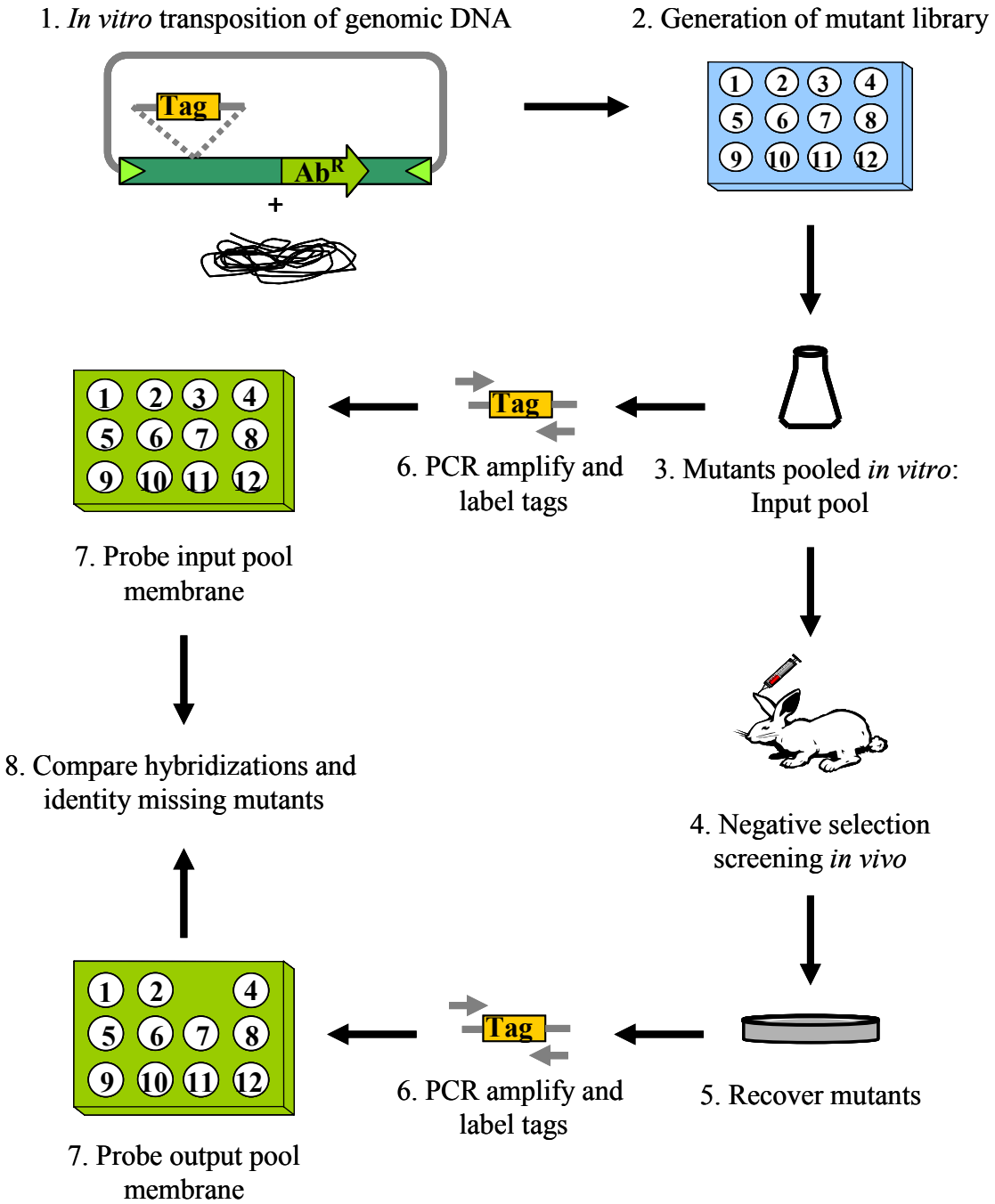
homologous recombination of transposition products for compilation of a mutant library (242). Regardless of the mechanism of transposition, after mutagenesis mutants bearing unique signature-tags are assigned to pools in which each tag is represented only once. During assembly of the mutant libraries bacteria are cultured in liquid media and/or agar plates. In this way, mutants with transposon insertion in genes essential for growth *in vitro* are automatically excluded as they are rendered uncultivable. Following creation of mutant library pools each pool is screened *in vivo*, where high selective pressures are encountered. The model of disease used is dictated by the pathogen under study; however infection must be monitored to find a relevant disease interval permitting colonization, and onset of disease prior to death. After a predetermined time, animals are sacrificed and bacteria recovered. The unique tags of the transposon are then used for detection of distinct strains received by the animal (the "input"), and strains recovered from the animal (the "output").

The tag DNA hybridization detection method depicted in Figure 30 is that originally detailed by David Holden and colleagues (95). For this detection method the mutant library is first used to create hybridization membranes. Colonies may be blotted onto the membrane (colony blots), or dot blots created with PCR amplicons of signature-tag sequences used, or plasmid DNA containing signature-tags used spotted onto the membrane. Differently tagged mutants are pooled into groups prior to inoculation into the animal; a portion of the inoculum is later used to create the input pool probe. After *in vivo* selection viable mutants are recovered and used for preparation of the output pool probe. Probes are developed by PCR amplification and

labeling of signature-tags present in the input and output strain populations. These labeled probes are then hybridized on the dot blot membranes for comparison of the presence and absence of strains in the input and output pool. Mutants not recovered by negative selection yield a signal on the input pool dot blot, lacking in the output pool dot blot.

A few considerations are necessary to derive reliable results from negative selection analysis by STM. To ensure accurate identification of clones, signature-tags are often prescreened for consistent and reproducible detection and lack of cross-hybridization (95). The pool size and complexity of the pools administered should also be considered, as these factors are inversely related to the accuracy of results obtained. As the complexity of the pool tested increases the likelihood that virulent mutants will not de detected also increases (in the instance that they are not present in sufficient numbers), resulting in virulent strains categorized as avirulent. Detection by DNA hybridization is particularly sensitive to the complexity of the pool tested, as pool complexity is inversely related to the quantity of labeled tag for each transposon. Therefore, for accurate detection pool sizes should not exceed the detectable threshold. Another phenomenon to consider is the possibility of a bottleneck occurring during colonization (92). In principle signature-tagged mutagenesis depends on all strains having equal access to colonization for onset of infection. The population of bacteria that survive during colonization must overcome selective pressures encountered *in vivo*. Initial design and optimization of signature-tagged mutagenesis screens typically include experiments evaluating whether a bottleneck is occurring.

Figure 30. STM screening schematic



Achieving consistently reproducible results from the recovery pool confirms that this is not the case. A general requirement for meaningful screening is a large output pool, exceeding 10^4 bacteria to recover the maximal number of tags (73, 95). The parameters described here directly affect the outcome of STM screening.

Rationale

An *in vivo* screening for *S. sanguinis* SK36 IE virulence determinants by random signature-tagged mutagenesis identified 5 avirulent mutants with transposon insertions in ORFs encoding housekeeping genes or genes necessary for survival under anaerobic conditions (198). The lack of identification of classical virulence factors, e.g., those facilitating adhesion in colonization, was surprising given the abundance of putative virulence factors predicted from the complete SK36 genome sequence (285). In the described experiments the signature-tagged mutagenesis technique was modified for a directed application to genes predicted to encode cell wall-associated proteins and sortases. Through this comprehensive analysis we hoped to determine the role of *S. sanguinis* cell wall-associated proteins in early IE.

Results

Identification of putative *S. sanguinis* SK36 Cwa protein ORFs

The genome sequence of *S. sanguinis* SK36 is the largest (2.4 Mb) published for a streptococcal species, and includes the most open reading frames (2,272) (285). Search algorithms incorporating features characteristic of Cwa proteins were used to identify ORFs encoding putative Cwa proteins. The patterns applied include the sorting site LPXTG pattern, followed by a region rich in hydrophobic residues, and capped by positively charged C-terminal residues. Two prediction patterns (in Materials and Methods) were used to screen for putative Cwa proteins among the SK36 ORFs. Thirty-seven positive hits resulted from searching with the more stringent Fuzzpro pattern (B), whereas 14 additional positive hits occurred when searching with pattern (A). N-terminal signal peptide prediction by SignalP differentiated between erroneously identified cytoplasmic proteins, and proteins putatively undergoing Sec-mediated extra-cytoplasmic export. Protein sequences of SK36 ORFs positive by Fuzzpro prediction and predicted to encode an N-terminal signal peptide were then visually inspected for an LPXTG pattern. Three of the 33 remaining ORFs lacking a C-terminal LPXTG pattern were likely multiple-pass integral cytoplasmic membrane proteins, as determined by TMHMM analysis. Using these subtractive prediction methods, 30 putative cell wall-associated proteins were identified for SK36 (Table 4).

Table 4. Predicted *S. sanguinis* Cwa proteins

Putative Cwa proteins were identified by two algorithms (described in Materials and Methods). Proteins positive for algorithm (B) had four or less mismatches, whereas proteins identified by algorithm (A) had five mismatches. Several proteins believed for other reasons to be Cwa proteins were missed by the prediction method used (not identified= NID), due to mismatches exceeding the threshold (¤), or lack of a predicted N-terminal signal peptide (‡). The sorting signal for each ORF is shown, with the LPXTG in bold. The SignalP server was used for prediction of N-terminal signal peptides. For this purpose two algorithm were applied, the neural network (NN) for detection of secreted (cell surface-targeted proteins), and the hidden Markov model (HMM) that assigns a signal-peptide probability value to a given sequence. Numerical values in the NN column are the 'D score'; referred to as the average of the signal peptide algorithm and the predictive cleavage site algorithm score. Proteins with NN cutoff values greater than 0.45 were considered positive for a signal peptide (Y=Yes). The (HMM) probability value is also given, with 1 meaning a sequence is extremely probable for containing a signal peptide (Y=Yes). The TMHMM server was used for prediction of the number of helices (potential transmembrane domains). The number of predicted helices for each protein is given in the THMMM column. Proteins lacking an LPXTG pattern were discarded from the list of putative Cwa proteins. Discarded proteins with greater than two predicted transmembrane domains, and a predicted N-terminal signal peptide are listed are putative integral membrane proteins.

Table 4.

Ss Gene	SignalP NN	SignalP HMM	TMHMM	Fuzzpro	Mismatches
Putative Cwa proteins					
SSA_0146	0.616 Y	1.000 Y	1	**LPHTG**QASNALLTVAGVISALGTAGLSLRSRKKED	2
SSA_0167	0.893 Y	1.000 Y	2	**LPATG**TNQSFLALIGTVMLSVLAFVGFKRKKN	0
SSA_0243	0.811 Y	1.000 Y	0	**LPVTG**EKTSSLGLLGLVMTGLAGIFTFKKRERQ	1
SSA_0227	0.857 Y	1.000 Y	2	**LPKAG**ENASWIMIAAGVALVGLVGFIVFRAKKK	1
SSA_0273	0.617 Y	1.000 Y	1	**LPKTG**SQSSLLTNLLGLGLLGLGAGFLGLKKKS	0
SSA_0303	0.691 Y	1.000 Y	0	**LPETG**AKDAVYLPYLGLAAIIGALGLGKLKGKED	3
SSA_0453	0.416 N	0.999 Y	1	**LPKTG**TSTSLLALLGGFLAFLAGLLTFRKSNN	1
SSA_0565	0.795 Y	1.000 Y	0	**LPKTG**ENSSWLLLAGQMLLLMAMKLFYKKRSLD	3
SSA_0684	0.561 Y	0.996 Y	2	**LPKTG**QDTSVWMLILGFLTFGGAVGLTKRQRD	2
SSA_0805	0.771 Y	1.000 Y	2	**LPKTG**QETTLWLSVAGLAILVGFGSYVFLQKKTR	3
SSA_0829‡	0.318 N	0.000 N	0	**LPQTG**ETESKASILALGLGALGLAFKRRKKKSDSEE	5
SSA_0904¤	0.681 Y	1.000 Y	1	**LPKTG**TEETSYLAASLLAGVSGLGLIGLEKRKKSLKTKR	NID
SSA_0905	0.703 Y	1.000 Y	0	**LPKTG**TEETSYLAASLLAGLSGLGLIGLEKRKKKSED	2
SSA_0906	0.678 Y	1.000 Y	0	**LPKTG**TEETSYLAASLLAGVSGLGLIGLEKRKKKSED	2
SSA_0956	0.708 Y	1.000 Y	0	**LPKTG**SNASSLLIYLGFAALLASAGFRLSKKES	3
SSA_1019	0.787 Y	1.000 Y	1	**LPKTG**SETSIFAIAAGFALILLSALAYRFKKAN	2
SSA_1023	0.696 Y	1.000 Y	0	**LPNTG**NTNSSFLGVAIISLLASIGLLHSKKERKK	3
SSA_1063	0.670 Y	0.999 Y	1	**LPKTG**SSEKAVYSVVGLGLLLAGLGLGISVRKSKEQ	1
SSA_1065	0.910 Y	1.000 Y	2	**LPKTG**DSSSLFLVLLGALFLFGAVYLYDLSSRRL	4
SSA_1112	0.785 Y	1.000 Y	2	**LPKTG**EAQTSTATIGFFGLALAGILGFLGLKEKQKD	1
SSA_1234	0.825 Y	1.000 Y	0	**LPQTG**SESAVALSLVGMVLGFFGLAGIKKSHKED	1
SSA_1301	0.341 N	0.997 Y	0	**LPNTG**SEAERLAIFGMALGAAAFLGTNKRRRRDEDYN	5
SSA_1591	0.804 Y	0.927 Y	1	**LPSTG**EQVTIMGLAGLIILAGVFILLKKTNKN	1
SSA_1632	0.796 Y	1.000 Y	2	**LPNTG**GIGTTILYLIGTSLVLGAGVLFIIKKRADSR	1
SSA_1633	0.833 Y	1.000 Y	2	**LPNTG**GIGTTILYLVSTSLVLGAGVLFVVKKRVSTK	1
SSA_1634	0.879 Y	1.000 Y	2	**LPNTG**GIGTTILYLSGTGLVLGAGILWMLKKRVNKK	2
SSA_1635¤	0.569 Y	0.997 Y	2	**LPETG**GNGTLLYLWGGSVITLGSAIWLIVRNKLYRKFQ	NID
SSA_1663	0.772 Y	1.000 Y	2	**LPFTG**TEISFTLLAFGLILIVGCGVYFGIRFKK	4
SSA_1666	0.254 N	0.999 Y	2	**LPATG**SAESLGLVLAGLVVLALLGLYRKTRLSK	2
SSA_1750	0.814 Y	1.000 Y	2	**LPKTG**QNTSGWAVLGLTLLMFAFTLKKRSRN	2
SSA_2020	0.664 Y	1.000 Y	1	**LPETG**SQETRASFVLAGLLVAGAAGLFLLKKKED	0
SSA_2023	0.472 Y	0.999 Y	1	**LPKTG**NQEDKFLPLAGLAFLGMGLLAGRKKQED	2
SSA_2121	0.884 Y	1.000 Y	2	**LPRTG**QEENLLVTILGFLAALLAGGMLVAKAKCS	3
Putative Integral Membrane Proteins					
SSA_0614	0.619 Y	0.955 Y	10	LWNMNASWLIFIIIGVIALLLSIEKKKDIKER	5
SSA_0871	0.472 Y	1.000 Y	4	DPKLFSPIMIAALFVLGILIGSLGSIISMRRFLKI	4
SSA_1049	0.625 Y	0.111 N	6	LTTQNWGMGSTIGVVLIVAMLFTMWATKERRGR	5

Two proteins not identified by the Fuzzpro pattern searches due to mismatches exceeding the permitted threshold, SSA_0904 and SSA_1635, appeared to be cell wall-associated based on homology and gene location (NID, Table 4). *SSA_0904* encodes CrpA (Csh-related protein A) and exists in an operon with additional Csh-like proteins CrpB (SSA_0905), and CrpC (SSA_0906), both of which were identified as Cwa proteins. A gram-positive N-terminal signal peptide was predicted for this ORF, and SSA_0904 was included in the comprehensive analysis of *S. sanguinis* Cwa proteins *in vivo*.

SSA_1635 is putatively the first gene in a cluster of predicted Cwa protein encoding genes including *SSA_1634*, *SSA_1633*, and *SSA_1632*. ORFs SSA_1632, 1633, and 1634 exhibit amino acid sequence identity ranging from 48 % to 62 % over the entirety of the protein sequences. Identity with SSA_1635 is restricted to the C-terminus of the protein and a unifying LPXTGGXG sorting pattern. The genomic clustering of *SSA_1635-1632* and the proximity of these genes to a gene encoding a putative sortase, *SSA_1631* (sortase C class) (Fig. 33C), suggests these genes encode cell wall-anchored protein substrates recognized by SSA_1631. From this locus SSA_1631, 1632, 1633, and 1634 were included in the comprehensive *in vivo* analysis.

One protein positive by Fuzzpro analysis, SSA_0829, was not predicted by SignalP to include a gram-positive N-terminal signal sequence (Table 4). In the case of SSA_0829, this may be attributed to transport via an alternative accessory Sec-mediated pathway, as described for an SSA_0829 homolog in *S. gordonii*, GspB. This SecA2/SecY2 dependent mode of extracytoplasmic export is reliant upon conserved

Gly residues within a hydrophobic region proximal to the amino terminus of the cytoplasmic, precursor Cwa protein (16, 17). SSA_0829 is a member of the serine-rich glycoprotein family, which includes Fap1 of *S. parasanguinis*, SrpA of *Streptococcus cristatus*, SraP of *S. aureus*, Srr proteins of *S. agalactiae*, and Hsa of *S. gordonii* Challis, in addition to GspB, discussed above (89, 230, 234, 247, 284). In keeping with the repeating serine residues of this protein, SSA_0829 was designated serine-rich protein A, SrpA, by Plummer et al (204).

Identification of unique Cwa proteins of *S. sanguinis*

The Cwa protein repertoire predicted for *S. sanguinis* exceeds predictions for related streptococcal species. Application of identical bioinformatics to *S. gordonii* CH1 subsp. Challis, *S. mutans* UA159, and *S. mitis* NCTC 12261 predicted 23, 6, and 15 cell wall-associated protein encoding ORFs, respectively (Tables 5, 6, and 7). The degree of relatedness between these species is reflected in the number of shared Cwa protein homologs. Eighteen of 33 predicted *S. sanguinis* Cwa proteins have significant homology to putative *S. gordonii* Cwa proteins. These include the previously characterized *S. gordonii* antigen I/II homologs SspA and SspB, cell surface hydrophobicity associated proteins A (CshA) and CshB, and serine-rich protein, Hsa. In addition four *S. sanguinis* uncharacterized proteins with putative enzymatic functions have homologs in *S. gordonii*, and four proteins containing CnaB-like collagen domains are conserved between the two species (Table 5).

Table 5. Predicted Cwa Proteins of *S. gordonii* CH1 (*Sg*)

			SignalP		
***Sg* Gene ID**	**Description**	***Ss***	**NN**	**HMM**	**Fuzzpro**
gi_157151510	Serine protease		0.267 N	0.000 N	**LPNTG**QAGDGMSLLGLFLSLCTFGLIRQRGKEEQL
gi_157151208	Nuclease/phosphatase domain	SSA_1750	0.245 N	0.000 N	**LPATG**EQRGHWVAFGFLVLIFAYPLSRKSKKSSNI
gi_157151122	Collagen binding domain	SSA_1019	0.259 N	0.003 N	**LPKTG**TAESSAFSVLVGFLLLVASVFLYRTKKVN
gi_157150004	Glycosyl hydrolase family		0.253 N	0.000 N	**LPNTA**SQETAFVAEAVLLAALGGLLLAGKKKED
gi_157151439	SspA	SSA_0303	0.561 Y	1.000 Y	**LPKTG**TNDSSYMPYLGLAALVGVLGLGQLKRKEDE
		SSA_0956			
gi_157150597	SspB	SSA_0303	0.548 Y	1.000 Y	**LPKTG**TNDATYMPYLGLAALVGFLGLGLAKRKED
		SSA_0956			
gi_157150159	Serine protease		0.223 N	0.000 N	**LPKTG**QEDQTSLALSLFGITSLALAGMVIHKKREE
gi_157151265	Exo-beta-D-fructosidase	SSA_2023	0.448 N	1.000 Y	**LPKTA**TQEDKFLPLAGLAFLGMGLLAGRKKQED
gi_157151173	Zinc carboxypeptidase family	SSA_2020	0.266 N	0.000 N	**LPKTG**TQEMTEGFKLASILIACSAGLFLLKKKED
gi_157150906	Beta-N-acetylhexosaminidase	SSA_2023	0.445 N	1.000 Y	**LPQTG**TTSAWPITLLGTALALIGLGGRKKRKG
gi_157151454	Cell wall surface protein	SSA_0565	0.865 Y	1.000 Y	**LPNTG**SRLSVWAMMSGILALVSGFALLIWKHKKES
gi_157150739	Cell wall surface protein	SSA_0684	0.610 Y	1.000 Y	**LPNTG**ETSTVLLSMIGFAFAGLVGYVVRKKGKA
gi_157149904	CshA	SSA_0904	0.522 Y	1.000 Y	**LPRTG**SQTSDQTASGLLAAIASLTFFGLANRKKKS
		SSA_0905			
		SSA_0906			
gi_157151344	Collagen binding domain	SSA_0805	0.276 N	0.000 N	**LPKTG**EQAAIWLSVAGVSILVILGGVFVVLRKKNK
gi_157150914	Streptococcal hemagglutinin	SSA_0829	0.240 N	0.000 N	**LPRTG**ESESKASILALGIGALGLAFKKRKKNESEI
gi_157150842	CshB	SSA_0904	0.528 Y	1.000 Y	**LPRTG**SQTSDQTASGLLAAIASLTFFGLANRKKKS
		SSA_0905			
		SSA_0906			
gi_157150937	Cell wall surface protein	SSA_1112	0.538 Y	1.000 Y	**LPNTG**EAQTATATIGFFGLALAGLLGLLGLKEKQK
gi_157150885	5'-nucleotidase	SSA_1234	0.590 Y	1.000 Y	**LPETG**TKESPALLIVTLFTAFLGIFGFKKSQED
gi_157150970	X-prolyl dipeptidylaminopeptidase		0.167 N	0.000 N	**LPETG**QKENSWSIYGMLSLLLAGLWKVISGRKEEK
gi_157150569	Beta-galactosidase	SSA_2023	0.546 Y	1.000 Y	**LPNTG**SQSDQAAITTGLALLGLGASLVAGKRRKEI
gi_157151048	Cna protein B-type domain	SSA_1663	0.211 N	0.000 N	**LPLTG**TELSFTLLALGLTMIVGWGVYYGTRFKN
gi_157150237	Collagen binding domain	SSA_1666	0.222 N	0.003 N	**LPNTG**SSDSKFIFVVGLATLMLLGLVWKKKHS
gi_157150742	Cell wall surface protein		0.459 Y	0.756 Y	**LPNTG**EHANGFAALLGLVLASVTAFFNFRRKKH
gi_157150875	Cell wall surface protein	SSA_0565	0.619 Y	1.000 Y	**LPKTG**TSATMVNEVIIGMILVLMGLLLRRKPKH

The sorting signal for each predicted *Sg* ORF is shown, with the LPXTG in bold. The SignalP server was used for prediction of N-terminal signal peptides by both NN and HMM algorithms (descriptions of values are given in Table 4). The *S. gordonii* protein annotation is listed (270). *S. sanguinis* (*Ss*) SK36 protein homologs were identified by BLASTP analysis.

Table 6. Predicted Cwa Proteins of *S. mutans* UA159 (*Smu*)

			SignalP				
***Smu* Gene ID**	**Description**	***Ss***	**NN**		**HMM**		**Fuzzpro**
gi_24378602	FruA	SSA_2023	0.529	Y	0.989	Y	**LPDTG**DHKTDLSQLGVLAMIGSFLVEIAGYFKKRKD
gi_24379087	SpaP	SSA_0303	0.789	Y	1.000	Y	**LPNTG**VTNNAYMPLLGIIGLVTSFSLLGLKAKKD
gi_24379427	WapA		0.931	Y	1.000	Y	**LPSTG**EQAGLLLTTVGLVIVAVAGVYFYRTRR
gi_24379528	WapE		0.594	Y	1.000	Y	**LPDTG**VQKNNQLALIALGTGLILLSGLLLSKRKSLK
gi_24379801	GbpC		0.823	Y	1.000	Y	**LPHAG**AAKQNGLATLGAISTAFAAATLIAARKKEN
gi_24380381	Dextranase		0.175	N	0.000	N	**LPQTG**DNNETRSNLLKVIGAGALLIGAAGLLSLIKGRKKD

The sorting signal for each predicted *Smu* ORF is shown, with the LPXTG in bold. The SignalP server was used for prediction of N-terminal signal peptides by both NN and HMM algorithms (descriptions of values are given in Table 4). The *S. mutans* protein annotation is listed (3). *S. sanguinis* (*Ss*) SK36 protein homologs were identified by BLASTP analysis.

Table 7. Predicted Cwa proteins of *S. mitis* NCTC 12261 (*Smi*)

			SignalP				
***Smi* Gene ID**	**Description**	***Ss***	**NN**		**HMM**		**Fuzzpro**
SMT0163	Amylopullulanase	SSA_0453	0.791	Y	1.000	Y	**LPNTG**TKNDHKLLFAGISILALLGLGFLLKNKKEN
SMT0272	Glycosyl hydrolase family 85 family		0.113	N	0.000	N	**LPNTG**TVDANEALIAGLASLGLASLALTLKQKKEDED
SMT0737	Cell wall surface protein		0.466	Y	1.000	Y	**LPNTG**TEDHANLVVLGLLGVLSGFGFLAHKKKEVEK
SMT0779	Muramidase-released protein		0.601	Y	1.000	Y	**LPNTG**TEANASLAALGLLGALGGFGLLARKKKED
SMT0780	Muramidase-released protein		0.638	Y	1.000	Y	**LPNTG**TEANASLAAFGLLGALGGFGLLARKKKED
SMT0787	Mucus binding protein precursor (Mub)		0.496	Y	1.000	Y	**LPKTAG**HSSGMAHVLGVFGLIAGFSLVGKAKRDE
SMT0793	Surface protein C (PspC)		0.781	Y	1.000	Y	**LPETA**SHDSILLVVGFLSALSGVLLFKVKKD
SMT0819	CshA	SSA_904 SSA_905 SSA_906	0.456	Y	1.000	Y	**LPNTG**TEDNASLAALGLLGVLSGFGLVARKKKED
SMT0925	ABC transporter efflux protein		0.440	N	0.543	Y	**LPNIG**VLLVFLVALTVLNIVGLRRYRKV
SMT1139	R5 protein		0.532	Y	1.000	Y	**LPNTG**TADNASLAALGLLGVLSGFGLVARKKKED
SMT1189	PspA		0.917	Y	1.000	Y	**LPNTG**GKDNVAIASLGFLGLLLGALPFVKRKN
SMT1291	Cell wall surface protein		0.398	N	0.999	Y	**LPNTG**TESNAALAAFGFVGILSGLSIVLRKKDNE
SMT1537	Bovine homologue of human Hr44		0.294	N	0.000	N	**LPNTG**TEDHASLAALGLLGALSGFGLIARKKREDEE
SMT1642	Cell wall surface protein	SSA_1112	0.123	N	0.000	N	**LPYTG**EAQTSMATLGFFGLALAGLLGGLGLKAKKEEND
SMT0785	Cell wall surface protein	SSA_0829	0.491	Y	1.000	Y	**LPNTG**IQEDRTTGTVGVLSLLGAFGLLFAKKKKDDEEEA

The sorting signal for each predicted *Smi* ORF is shown, with the LPXTG in bold. The SignalP server was used for prediction of N-terminal signal peptides by both NN and HMM algorithms (descriptions of values are given in Table 4). The *S. mitis* protein annotation is listed online (http://www.tigr.org/). *S. sanguinis* (*Ss*) SK36 protein homologs were identified by BLASTP analysis.

Four predicted *S. sanguinis* Cwa proteins are conserved in *S. mitis,* and two conserved between *S. sanguinis* and *S. mutans* (Tables 6 and 7, respectively). A putative alkaline amylopullulanase (SSA_0453) is unique to *S. sanguinis* and *S. mitis* of the four species compared. Comparative analysis suggests the majority of predicted Cwa proteins of *S. sanguinis* (18 of 33) are homologous to predicted proteins of similar spp., suggesting conservation of specific fitness factors during species divergence.

We then sought to determine whether the remaining 15 proteins with no homologs in *S. gordonii*, *S. mutans*, and *S. mitis* were actually unique to *S. sanguinis.* A comparison of the *S. sanguinis* genome to complete genome sequences of 55 other gram-positive spp., related and unrelated oral microbes, and unrelated causative spp. of infective endocarditis indicated that 8 predicted Cwa proteins are unique to *S. sanguinis.* These include the ORFs SSA_1632, SSA_1633 and SSA_1634, described above, in addition to SSA_0146, SSA_0167, and SSA_0273, SSA_1063, and SSA_1023. SSA_1063 has been previously described as the platelet aggregating surface protein, CbpA. SSA_1023 was unique to *S. sanguinis* among the 55 species compared, yet has some sequence similarity to the muramidase released protein (Mrp) of *S. suis*, and von Willebrand factor binding protein (vWf1) of *S. lugdunensis*. Although these analyses are predictive in nature they do suggest multiple mechanisms of acquisition of *S. sanguinis* cell wall associated protein ORFs. While many were likely inherited from a common ancestral species, some—in particular, unique, uncharacterized, hypothetical proteins—were likely acquired by horizontal gene transfer and therefore may be unique to *S. sanguinis* within the viridians group of streptococci.

Identification of putative *S. sanguinis* SK36 sortases

Gram-positive sortases may be divided into sub-families based on conserved residues surrounding the LXTC active site, distinctive transmembrane topology, as well as genomic clustering with putative substrates (49, 59). Three sortase-like proteins of *S. sanguinis* SK36 were identified by homology searches with 45 publicly available sortase sequences. The *S. sanguinis* sortase-like proteins were annotated based on similarity to one of three subfamilies, SrtA, SrtB or SrtC (Table 8). SSA_1219 was homologous to the majority of published SrtA sequences queried, with the exception of the prototypical SrtA of *S. aureus*. As previously documented for this subfamily, within the active site residues (LXTC) X is a Val residue (Fig. 31A). (Iso or Thr are also common.) This protein subfamily is also characterized by an N-terminal signal peptide/transmembrane domain directing type I membrane topology (59). Hydropathy analysis of the *S. sanguinis* SrtA sequence identified a single transmembrane domain, restricted to the N-terminus of the protein (Fig. 31B). Finally, the *SSA_1219* gene encoding SrtA is not clustered with any putative substrate, suggesting that activity of this sortase is not limited to a subset of Cwa proteins (Fig. 31C).

Of the three sortase-like proteins identified, SSA_0022 had the greatest sequence identity to previously published SrtB sequences. These characterized and uncharacterized SrtB-related proteins were grouped based on a conserved Ser residue proximal to the catalytic TLSTC Cys residue (Fig. 32A) (199). A Ser residue also occurs in the LXTC of SSA_0022, suggesting that this protein may be similar in function to other SrtB homologs. Yet the *SSA_0022* genomic locus presents a paradox.

Table 8. Identification of *S. sanguinis* sortase-like ORFs

Sortase A class:	**SSA_1219**	**SSA_0022**	**SSA_1631**
GI:3238307_*S. aureus* subsp. *aureus* COL			●
GI:1201902_*E. faecalis* V583	●		
GI:1028429_*S. mutans* UA159	●		
GI:2650068_*C. diphtheriae* NCTC 13129	○	○	○
GI:1088124_*B. anthracis* str. Ames	●		
GI:111657852_*S. pneumoniae* TIGR4	●		
GI:901269_*S. pyogenes* M1 GAS	●		
GI:1201386_*E. faecalis* V58	●		
Sortase B class:			
GI:2859503_*S. aureus* subsp. *aureus* MRSA252		●	
GI:891152_*B. halodurans* C-125		●	
GI:1083901_*B. anthracis* str. Ames		●	
GI:900461_*S. pyogenes* M1 GAS		●	
GI:4962816_*S. pyogenes* str. Manfredo		●	
GI:4916727 *C. difficile* 630		●	
Sortase C class:			
GI:3036999_A. naeslundii T14V			●
GI:1199996_E. faecalis V583			●
GI:2650116_*C. diphtheriae* NCTC 13129			●
GI:2650498 *C. diphtheriae* NCTC 13129			●
GI:2650495_*C. diphtheriae* NCTC 13129			●
GI:2650024_*C. diphtheriae* NCTC 13129			●
GI:2650023_*C. diphtheriae* NCTC 13129			●
GI:930405_*S. pneumoniae* TIGR4			●
GI:930404_*S. pneumoniae* TIGR4			●
GI:930406_*S. pneumoniae* TIGR4			●
GI:1116387_*C. acetobutylicum* ATCC 824			●
GI:939748_*B. subtilis* subsp. *subtilis* str. 168			●

The sortase GIs used to identify *S. sanguinis* sortase-like proteins are shown. The *S. sanguinis* SK36 protein best matching a given query sequence is indicated (●). Queries lacking a homolog in *S. sanguinis* are also indicated (○).

Figure 31. Characteristics of a SrtA-like protein

A) A ClustalW alignment of SrtA sequences from *S. mutans* UA159 (SMU_1113), *S. pyogenes* M1 GAS (SPY_1154), *S. pneumoniae* TIGR4, *S. sanguinis* SK36*, E. faecalis* (EF_3056), *and B. anthracis* str. Ames highlights conserved residues at the LXTC active site (indicated by asterisks), as well as similar neighboring residues (indicated by periods and colons)

B) A tmap hydropathy prediction of the SrtA sequence highlights a single hydrophobic region at the N-terminus of the protein, suggesting a type I topology of this transmembrane protein.

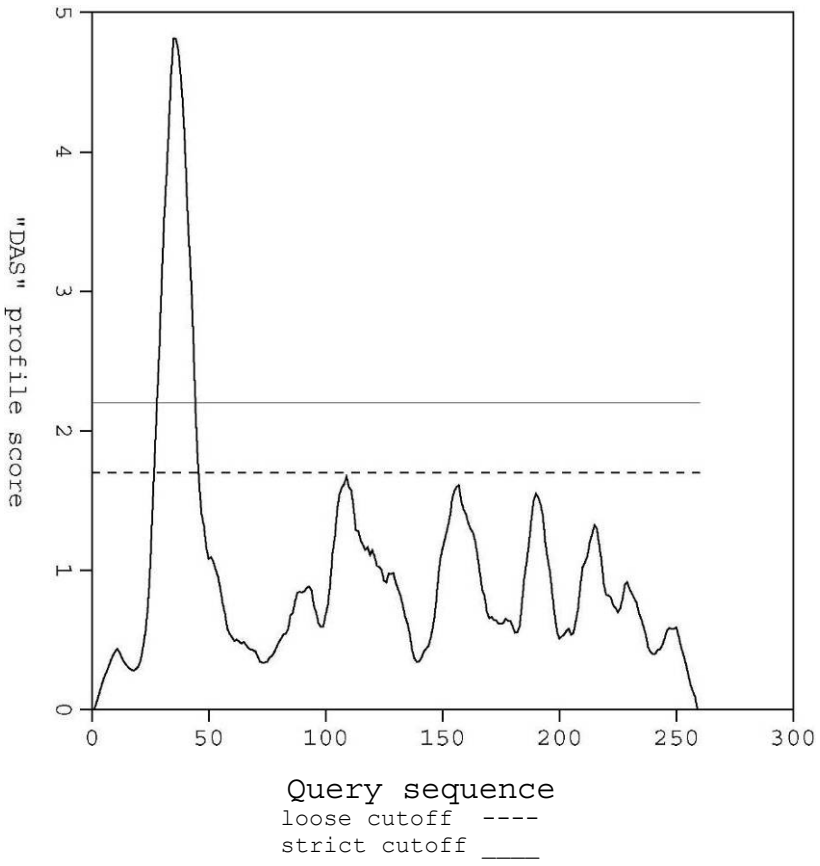
C) The *S. sanguinis* SK36 *srtA* locus is shown. In dark grey are neighboring genes, *ldh* and *gyrA* conserved in *S. gordonii* Challis CH1, *S. mutans* UA159, *S. pyogenes* M1 GAS, *and S. agalactiae* NEM16. *SSA_1218*, putatively encoding *radC* is conserved between *S. sanguinis* and *S. gordonii*. *SSA_1217*, encoding a putative glutamine amidotransferase, is not conserved at this locus. Putative promoter and terminator elements are depicted.

Figure 31.

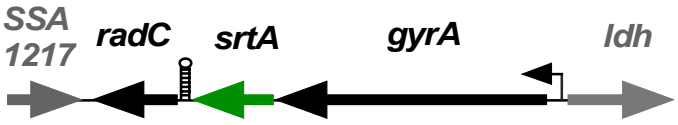
A.

```
GI:1028429_S.mutans         TDKN-KVYTYVISEVKTVTPEHVEVIDNRPGQNEVTLVTCTDAGAT-ART 214
GI:901269_S.pyogenes        TDKE-KIYEYIIKDVFTVAPERVDVIDDTAGLKEVTLVTCTDIEAT-ERI 217
GI:111657852_S.pneumoniae TDKN-KVYTYEIREVKRVTPDRVDEVDDRDGVNEITLVTCEDLAAT-ERI 216
GI:125718040_S.sanguinis  TDKE-QVYTYVITSVETVTPDRTDLIEDTEGVTEITLVTCEDAAAT-NRT 219
GI:1201902_E.faecalis       TDLS-TVYTYKITSVEKIEPTRVELIDDVPGQNMITLITCGDLQAT-TRI 209
GI:1088124_B.anthracis      YDNE-NEYEYAVTGVSEVTPDKWEVVED-HGKDEITLITCVSVKDNSKRY 220
GI:1201386_E.faecalis       NDREGHSFIYKVKEQKHVDMTDTTMLNLTRKPT-LTLITCDQATKTTGRI 235
                             * .   : * :     :       ::       :**:** .   .  *
```

B.



C.



In *Staphylococcus aureus*, *Bacillus* spp., and *Listeria monocytogenes* SrtB is encoded in an iron regulated operon and recognizes a unique sorting signal (NPxyTN) on its cognate substrate, the heme-iron binding protein IsdC (157, 165, 187) (Fig. 32B). The genomic clustering of SrtB with IsdC in *S. aureus*, *B. anthracis*, and *L. monocytogenes* suggests a highly specific activity for this protein. In *S. pyogenes* strains Manfredo and M1 GAS, and *S. sanguinis* SK36 the *srtB* gene is not in an operon with putative Cwa proteins with the NxyTN sorting signal. In the case of *S. pyogenes* Manfredo *srtB* is flanked by genes encoding putative Cwa proteins bearing an LPXTG sorting signal (SPYM50104), the sorting signal LPLAG (SPYM50106a), and a putative novel sorting signal of QVXTG (SPYM50106). A similar genomic arrangement exists in *S. pyogenes* M1 GAS (Fig. 32C). In *S. pyogenes* strain AM3 this sortase transpeptidase, containing catalytic site residues characteristic of SrtB (TLSTC) is annotated as SrtC1. Is has been demonstrated that SrtC1 is required for cell surface anchoring of a proximal ORF with the QVPTGV sorting signal (11). Further investigation of the homologous ORF in *S. pyogenes* M1 GAS, Spy0128, revealed this SrtC1 substrate to be the backbone component of a gram-positive surface pilus, decorated by Spy0130 subunits (175). It appears that in streptococcal species SrtB homologs are required for surface assembly of Cwa proteins with functions unrelated to substrates of analogous sortases in *S. aureus* and *B. anthracis*.

In *S. sanguinis* the *srtB* ORF is not proximal to any genes encoding Cwa proteins with the conventional recognition sequence NxyTN. If streptococcal SrtB homologs recognize a deviant signal sequence, then PcsB (SSA_0019) may be the

Figure 32. Characteristics of a SrtB-like protein

A) SSA_0022 was annotated SrtB due to the conservation of a Ser residue in the LXTC catalytic site. A ClustalW alignment with characteristic SrtB sequences of *S. aureus* subsp. *aureus* MRSA252, *B. anthracis* Ames, as well as SrtB-related proteins of *S. pyogenes* M1 GAS (GI: 900461) and *S. pyogenes* Manfredo (GI: 4962816) reveals similarity in these catalytic site residues.

B) In *S. aureus*, *B. anthracis* and *L. monocytogenes srtB* is within a Fur regulated operon. The candidate SrtB substrate, IsdC is indicated in green. In *L. monocytogenes* SvpA anchoring to PG is dependent on SrtB; a putative additional *L. monocytogenes* SrtB substrate, encoded by *lmo2186* has 33 % identity to IsdC and contains a potential SrtB sorting motif, NKVTN (22, 187). Additional conserved genes at this locus in *S. aureus*, *B. anthracis* and *L. monocytogenes* are indicated by like color-coding, genes not conserved are in black.

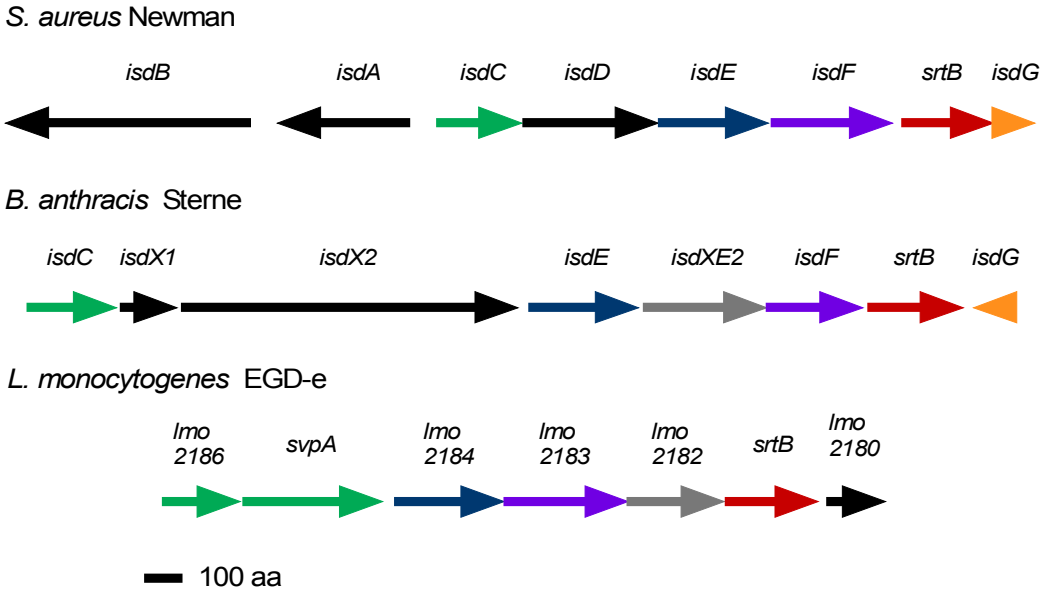
C) In streptococcal species the *srtB* locus is arranged differently. In green are genes putatively coding for cell wall-associated proteins with the LPXTG motif (*spyM50104*, *Spy0130*), or an LPLAG sorting signal (*spyM50106a*). In blue are genes putatively encoding candidate SrtB substrates which contain a sorting signal QXXXG. The anchoring of a SrtB substrate with this sorting pattern has only been demonstrated for *S. pyogenes* AM3 (11).

Figure 32.

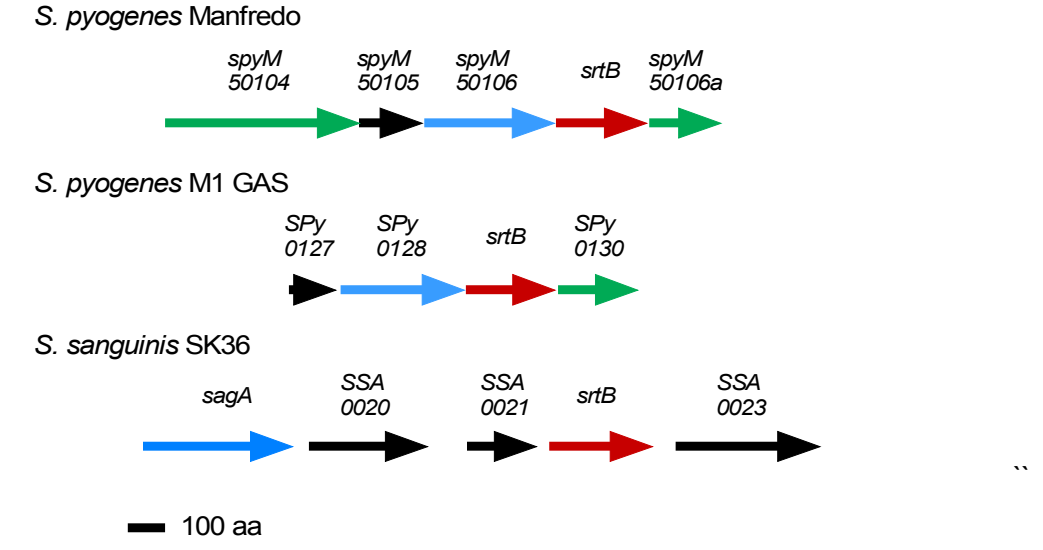
A.

```
GI:2859503_S. aureus          YIRT--DFENDQDYQQFLDETKRKS-VINSDVNVTVKDKIMTLSTCEDAY 227
GI:1083901_B. anthracis       YIET--DFSSDTEYTSFLEKIQEKS-LYKTDTTVTAGDQIVTLSTCDYAL 237
GI:900461_S. pyogenes         LLFNPIDVDISSK-NEFLNHIKQKS-VQYREILTTNESRFVALSTCEDMT 225
GI:4962816_S. pyogenes        LVFNPNAITNQDQQKQLVDYISKRS-KQFKPVKLKHHTKFVAFSTCENFS 233
GI:125716908_S. sanguinis     FYRT--EFKDDEDFTTQLRNIYEAARTKDPEMKIKASDRYLVLSTCREED 246
                                 .   .  . .    :    . :         .   : :.:***
```

B.



C.



candidate substrate for the transpeptidase. This protein includes the residues QPEVG near the C-terminus, and has 44 % sequence identity to the *S. agalactiae* PcsB protein. Altered morphology of *S. agalactiae* and *S. mutans* PcsB-related protein mutants suggests the protein acts as a murein hydrolase to split the division septum between adjacent daughter cells. Streptococci are globular cells that typically divide in approximately the same plane at each cell cycle, producing elongating chains. In contrast *S. mutans sagA* and *S. agalactiae pcsB* mutants exhibit division in multiple planes (44, 212). The loss of PcsB expression has also been linked to enhanced susceptibility to cell wall-targeting antibiotics, as well as antibiotics with intracellular targets, suggesting that additional defects in the cell wall exist beyond proper cell division (212). Likewise an *S. mutans sagA* mutant exhibited greater sensitivity to osmotic stress, increased temperatures, and low pH (44). These observations have resulted in the suggestion that PcsB is a general stress-related protein. Interestingly PcsB is essential in *S. pneumoniae*, but related proteins are not essential for growth of either *S. agalactiae* or *S. mutans* (44, 172, 212). Therefore additional species-specific activities may be attributable to PcsB. *S. pneumoniae* PcsB cell localization analysis by cell fractionation and Western blotting indicated the protein was predominately localized to the plasma membrane or secreted from the cell (172). The exposure of PcsB to extracellular milieu has also been inferred by the protective effect garnered by this protein in pneumococcal vaccine studies (82). Though it remains unconfirmed, it is possible that PcsB may additionally be anchored to the PG of the cell wall during trafficking from the plasma membrane to secretion from the cell. Cell localization,

general function, and the relationship between *pcsB* and the neighboring *srtB* ORF remain uncharacterized in *S. sanguinis*.

SSA_1631 was annotated as SrtC based on translated sequence conservation with the Sortase C-group/subfamily 3 transpeptidases proposed by Dramsi and colleagues (59). Sortase C transpeptidases are characterized by a C-terminal hydrophobic domain that may act as a membrane anchor to drive type-II membrane topology of this protein. Hydropathy analysis of the *S. sanguinis* SrtC amino acid sequence predicted both N- and C-terminal hydrophobic regions (Fig. 33B). As sortases act on proteins at the cell surface it is likely that the N-terminal hydrophobic region drives Sec-dependent export of the enzyme, whereas the C-terminal region determines membrane topology. In addition a conserved proline residue following the LXTC catalytic site links the *S. sanguinis* SrtC to previously described SrtC enzymes (Fig. 33A). As previously mentioned, *SSA_1631* is clustered with the putative Cwa protein encoding genes: *SSA_1635*, *SSA_1634*, *SSA_1633*, and *SSA_1632* (Fig. 33C). In *S. pneumoniae*, *C. diphtheriae*, and *E. faecalis* the Cwa genes flanking SrtC encode gram-positive pilin subunits. The polymerization of pilus subunits is dependent on the adjacent sortase in these species (13, 183, 256).

Patterns predictive of the pilin motif ("WxxxVxVYPK") and the E box ("YxLxETxAPxGY") were screened against the putative *S. sanguinis* Cwa proteins neighboring SrtC. Similar to *E. faecalis* pilus associated proteins EbpA, EbpB, and EbpC, E box determinants were identified in SSA_1635, SSA_1634, SSA_1633, and SSA_1632 with almost perfect fidelity to the E box motif (183).

Figure 33. Characteristics of a SrtC-like protein

A) The *S. sanguinis* SrtC homolog SSA_1631 catalytic site includes conserved residues, indicated by red coloring. The active site encompasses LXTC with a conserved Pro distal to the essential Cys residue. In the case of sortase C group the charge and polarity of X is conserved and is represented by green text.

B) A tmap hydropathy analysis revealed two discrete regions of hydrophobic residues, one at the extreme N-terminus (a signal sequence for extra-cytoplasmic export), and a C-terminal region that may define the membrane topology of the transpeptidase.

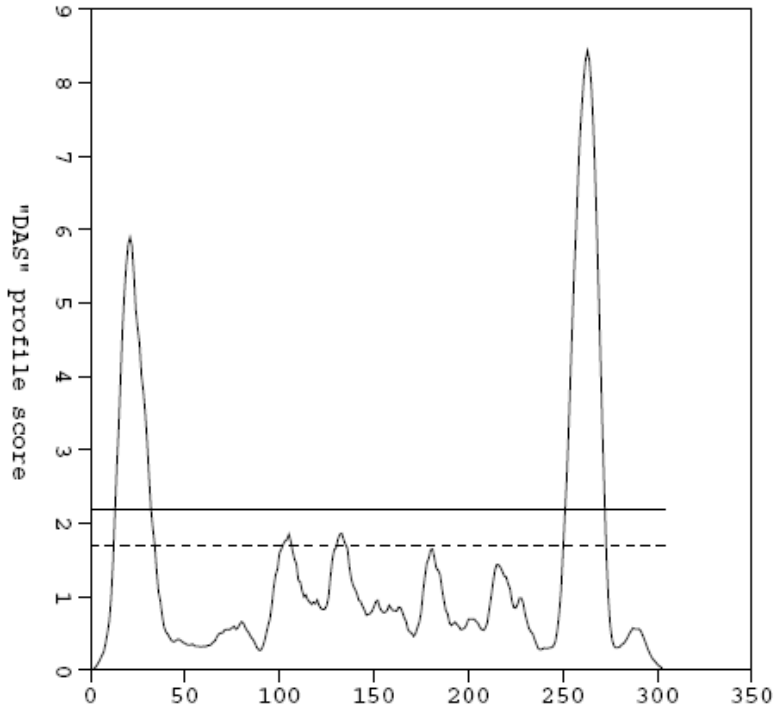
C) In the chromosome of *S. sanguinis* SK36 *srtC* is grouped with potential substrates, *SSA_1635*, *SSA_1634*, *SSA_1633*, and *SSA_1632*.

Figure 33.

A.

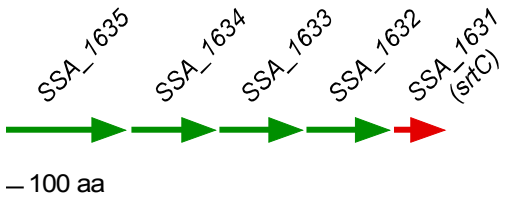
```
GI:125718436 S. sanguinis     VEPEDISSLAIDPDQDYCTLVTCTPYGINSHRLLVRGHRVPNEKNISKSS 236
GI:930404 S. pneumoniae       IEPTNFDDLLIVPGHDYVTLLTCTPYMINTHRLLVRGHRIPYVAEVEEE- 244
GI:930405 S. pneumoniae       VEPNDFEPVLIQHGEDYATLLTCTPYMINSHRLLVRGKRIPYTAPIAER- 212
GI:1199996 E. faecalis        VEPTDTKDLHIESGQDLVTLLTCTPYMINSHRLLVRGHRIPYQPEKAAAG 242
GI:2650498 C. diphtheriae     VLPTETEDLRPEQGKDYITLITCTPYGINTHRLMVRGHQVPLAPEDHEVF 257
GI:2650495 C. diphtheriae     VTPDKTSLLRRTSNKDQVTLITCTPYGINTHRLIITAERVPMDPQGESAF 276
GI:2650116 C. diphtheriae     VLPEDTKLLAPDPNKDQITLITCTPYAVNSHRLLVRAHRVDLDPNDPNLT 249
GI:2650023 C. diphtheriae     IRPAEIDRIQPIPDRDLITLVTCTPYGINTHRLLVTAERVPMEPGEADRA 278
                              : * . . :    ..*  **:***** :*:***:: ..::
```

B.



Query sequence
loose cutoff ----
strict cutoff ___

C.



SSA_1634 contained a putative pilin motif with the least deviations from the consensus pattern (259) (Table 9). These results suggest these putative Cwa proteins may undergo SrtC dependent polymerization in *S. sanguinis* to form a gram-positive surface pilus, as has been previously described for other gram-positive species.

Genes encoding sortases have been identified in almost all gram-positive bacteria (49, 185). Pallen et. al observed that genomes including a sortase enzyme invariably were predicted to encode at least two, and sometimes many more (199). The abundance of sortases and multitude of Cwa precursors suggests that sortases act non-redundantly to the benefit these organisms.

Annotation of identified Cwa protein ORFs

BLASTP and CDD searches enabled annotation of predicted *S. sanguinis* Cwa proteins. Of the 33 proteins identified, 4 did not have adequate similarity to suggest homology, nor conserved domains suggestive of protein function, and are described as hypothetical in Table 10. Proteins described as conserved uncharacterized in Table 10 were detected by MudPIT (described below). The majority of remaining proteins could be grouped functionally into two groups that are not entirely exclusive; putative adhesins, or putative enzymes. Putative enzymatic Cwa proteins include SSA_0243, SSA_1065, SSA_1234, SSA_1591, SSA_1750, and SSA_2023. Proteins with homology to SSA_0453, and SSA_1591 suggest these proteins may confer both enzymatic and adhesion properties.

Table 9. Conserved motifs exist in the pilin-like proteins of SK36

		Features of the major pilin protein			
Strain	Name	CWSS[a]	Pilin motif	E box	References
C. diphtheriae NCTC13129	SpaA	**LP**LT**G**G	WLQDVHVY**PK**	FC**L**V**ET**ATAS**G**Y	(256)
C. diphtheriae NCTC13129	SpaD	**LP**MT**G**G	WNYNVVAY**PK**	FC**L**K**ET**KAPA**G**Y	(256)
C. diphtheriae NCTC13129	SpaH	**LP**LT**G**G	WLYDVNVF**PK**	YV**L**V**ET**EAPT**G**F	(256)
S. pneumoniae TIGR4	RgrB	I**P**K**TG**E	DVVDAHVY**PK**	YY**L**E**ET**KQPA**G**Y	(13)
E. faecalis V583	EbpA	**LP**E**TG**G	KYGEIHYAG**K**	YT**L**T**E**VKAPA**G**H YV**L**T**ET**FTPE**G**Y	(183)
	EbpB	**LP**K**T**NE	SLTHIHLY**PK**	YF**L**E**E**ISAPK**G**Y	(183)
	EbpC	**LPSTG**G	ELAVVHIY**PK**	YI**L**E**E**VKAPNNA YY**L**E**ET**VAPDDY	(183)
S. sanguinis SK36	SSA_1634	**LP**N**TG**G	YNNDVQVYL**K**	YK**L**V**ET**IT**P**K**G**Y	This study
	SSA_1633	**LP**N**TG**G	AGAGFTLYK**K**	YK**L**V**E**STTPA**G**Y	This study
	SSA_1632	**LP**N**TG**G	YVQDVKVSS**K**	YK**L**V**E**STVPS**G**Y	This study
	SSA_1635	**LP**E**TG**G	WNFQEYEK**PK**	YA**L**N**ET**LVDNKT	This study
Consensus sequence[b]		**LPXTG**	WxxxVxVY**PK**	Yx**L**x**ET**xAPx**G**Y	

a. CWSS, cell wall sorting signal
b. Consensus sequences of the sorting signal (226), gram-positive pilin motif (259), and E box motif (256). Residues that are highly conserved are indicated by bold lettering (229).

Cwa proteins with putative enzymatic functions

Putative nucleotide phosphoesterase and 5' nucleotidase activities were predicted for two identified Cwa protein ORFs, SSA_0243 and SSA_1234. The *B. subtilis* YfkN homolog of SSA_0243 (62 % sequence similarity) has tri-functional nucleotide phosphoesterase activity (41). The purified YfkN enzyme can release inorganic phosphate from 2', 3' cyclic nucleotide phosphates (cyclic 2', 3' NMP), which requires both 2', 3' phosphodiesterase activity and 2' or 3' nucleotidase activity. YfkN has an additional lesser activity as a 5' nucleotidase (41). Both domains conferring these functions to YfkN were identified in *S. sanguinis* SSA_0243 (pfam00149: phosphoesterase family including 2', 3' CAMP phosphodiesterase, and pfam02872: 5' nucleotidase C-terminal domain) (156). The multifunctional enzyme YfkN is supposed to have resulted from fusion of two genes conferring the separate functions (41).

SSA_1234 is similar (41 %) to the UshA, 5' nucleotidase of *B. cereus* (228). This nucleotidase was previously demonstrated to have 5', and 3' nucleotidase and phosphotransferase activities. However UshA lacks sugar hydrolase and cyclic phosphodiesterase activities (228). Similar 5' nucleotidases have been reported for many gram-negative species including the NucA protein of *H. influenzae*. One putative biological role of these proteins is hydrolysis of extracellular nucleotides, and subsequent intracellular transport of hydrolytic products for use as substrates of cell wall synthesis (291).

Table 10. *S. sanguinis* Cwa protein annotation

S. sanguinis				
Gene ID	**Gene**	**Ortholog**	**Bacterium**	**Description**
SSA_0227¤				Collagen binding domain; pfam05737 (156)
SSA_0303¤	*sspC*	*sspA, B*	*S. gordonii*	Antigen I/II, Promotes binding to collagen, dentinal tubule invasion(152); coaggregation with *P. gingivalis* (153); binding to sAG (110)
		spaP/pac	*S. mutans*	Adherence to sAG; cariogenicity (52, 152); adherence to fibronectin, collagen and fibrinogen (14); bacteremic virulence
SSA_0453¤¥		*pula*	*S. pyogenes*	Amylopullulanase, hydrolysis of pullulan and starch at neutral pH, independent glycoprotein binding capacity (108)
		spuA	*S. pneumoniae*	Immunogenic; pullulan degradation
SSA_0805¤				Collagen binding domain; pfam05737 (156)
SSA_0829¤	*srpA*	*Has*	*S. gordonii*	Platelet binding and aggregation (122, 204); HUVEC invasion (241); endocarditis virulence (248); binding to sAG (110)
SSA_0904¤	*crpA*	*cshA*	*S. gordonii*	Adherence to *A. naeslundii*, *S. oralis* and immobilized
SSA_0905¤	*crpB*	*cshB*	*S. gordonii*	fibronectin (169); invasion of
SSA_0906¤	*crpC*			HUVEC cells (241); platelet adhesion (*cshA*) (98)
SSA_0956¤	*sspD*	*sspA, B*	*S. gordonii*	Antigen I/II, Promotes binding to collagen, dentinal tubule invasion; coaggregation with *P. gingivalis*; binding to sAG
SSA_1019¤				Collagen binding domain; pfam05737 (156)
SSA_1023¤		*vwb1*	*S. lugdunensis*	Von Willebrand factor (vWf) binding protein (188)
		mrp	*S. suis*	Immunodominant protein expressed during *S. suis* swine infection (252, 283)
SSA_1063¤				Von Willebrand factor type A domain; cd00198 (156)

S. sanguinis				
Gene ID	**Gene**	**Ortholog**	**Bacterium**	**Description**
SSA_1632¤ SSA_1633¤ SSA_1634¤ SSA_1635¤		*rrgA, rrgB*	*S. pneumoniae*	Surface pilus locus; *in vitro* adherence to A549 epithelial cells; colonization and virulence *in vivo*; proinflammatory response mediator (13)
SSA_1663¤	*cbpA*			Platelet aggregation; endocarditis virulence; collagen binding domain; pfam05737 (156)
SSA_1666¤				Collagen binding domain; pfam05737 (156)
SSA_0243¥		*yfkN*	*B. subtilis*	5' nucleotidase, 2', 3' phosphodiesterase and 2' (or 3') nucleotidase for reprocessing extracellular nucleotide phosphates under restrictive growth; binding to HA (41)
SSA_1065¥				Beta-hexosaminidase A; glycosyl hydrolase 3 family NTD (pfam00933); glycosyl hydrolase 3 family CTD (pfam01915); extracellular hydrolysis of (1,3 and 1,4) β-D-glucans (156), (267)
SSA_1234¥		*ushA*	*B. cereus*	Nucleotidase; hydrolysis of 5' and 3' nucleotides (228)
SSA_1591¤¥		*abpB*	*S. gordonii*	Dipeptidase (156)
SSA_1750¥		*ssnA*	*S. suis*	Immunogenic, extracellular nuclease (74)
SSA_2023¥		*fruA*	*S. mutans*	Fructan hydrolase; synthesis of inulin-type β2,1- linked fructose polymers for extracellular storage (23); Cariogenicity (32)
SSA_0565		*prgA*	*E. faecalis*	Surface exclusion protein (61)
SSA_0684		*fibA*	*P. micros*	Fibril-like structure subunit, self-aggregation (131)
SSA_1301				Conserved uncharacterized
SSA_1112				Hypothetical
SSA_0273				Hypothetical
SSA_2020				Hypothetical
SSA_0167				Hypothetical
SSA_0146				Hypothetical

S. sanguinis				
Gene ID	**Gene**	**Ortholog**	**Bacterium**	**Description**
SSA_2121				Conserved uncharacterized

Cwa proteins are grouped by general functions, as putative enzymes (¥), or adhesins (¤). Proteins lacking sufficient sequence identity by BLASTP analysis are listed as hypothetical. Previously hypothetical proteins detected by MudPIT are listed as conserved uncharacterized. Portions of this table were adapted from (285).

SSA_1750 shares 60 % amino acid sequence similarity with the previously described extracellular nuclease of *S. suis*, SsnA, which targets single-stranded, double-stranded and linear DNA for degradation. In *S. suis* SsnA is cell wall anchored and prevalent among porcine isolates from internal sites including joints, brain and organs (74). SsnA is expressed during infection, as has been previously documented for DNA nucleases produced by GAS, and staphylococci (9, 43, 74). The function of SsnA was determined to be Ca^{2+} and Mg^{2+} dependent (74); similarly the pfam03372 Mg^{2+} dependent catalytic domain was identified for SSA_1750 (156). Extracellular nuclease production is elaborated during lethal murine staphylococcal infection, suggesting nuclease production may be a virulence mechanism (43). However nuclease production may be more indirect, possibly for nutrient acquisition at the site of infection or for competition with resident flora, by degradation of extracellular DNA to impact the transformation capacity of other bacteria.

The SSA_1065 amino acid sequence contains both pfam00933 and pfam01915 (156), which comprise the N-terminal and C-terminal domain, respectively of the family 3 glycosyl hydrolases. The function of family 3 glycosyl hydrolases was previously described for germination and elongation of barley (*Hordem vulgare*) in formation of coleoptiles (267). The exohydrolase enzyme, ExoI of barley is a 2-domain globular protein that contains a catalytic site for hydrolysis between the two domains, in addition to a secondary binding site for β1,3 and β1,4 β-D glucans (267). The surface production of bacterial glycosyl hydrolases is thought to aid the organism in degradation of environmental oligosaccharides for liberation of a valuable cellular carbon source (46).

SSA_2023, the FruA (*S. mutans*) homolog likely serves a similar valuable function in nutrient acquisition. In *S. sanguinis* and *S. mutans* fructan hydrolases generate extracellular inulin-type polymers composed mainly of β2,1-linked fructose (23). This function is putatively for storage of extracellular polysaccharide in biofilms that may be catabolized during nutrient deprivation (83, 100). FruA of *S. mutans* has the capacity to release fructose from levan, inulin, and raffinose, and can cleave sucrose to form fructose and glucose. The expression of FruA is tightly controlled and repressed when *S. mutans* is cultured in the presence of hexose sugars (33). Finally, FruA production contributes to dental caries development in the rat model of this disease (32).

SSA_0453 encodes a putative pullulanase, based on conservation of four highly conserved regions that form the catalytic domain of pullulanase enzymes. The pullulanase PulA enzyme of *S. pyogenes* is also referred to as a "strepadhesin", since this protein is associated with binding of *S. pyogenes* to thyroglobulin, submaxillary mucin, fetuin and asialofetuin *in vitro* (108). PulA additionally hydrolyses pullulan and starch at a neutral pH. The dual properties of this protein led Hytonen and colleagues to suggest that PulA may not be as important for nutrient acquisition, as for colonization by degrading host glycoconjugates to reveal specific binding sites. Additional purported functions of PulA include modification of self cell surface carbohydrates, or cell surfaces of other bacteria (108).

SSA_1591 encodes a putative dipeptidase (pfam03577) (156), with 60 % protein sequence similarity to the amylase binding protein B (AbpB) of *S. gordonii*. (Another *S. sanguinis* AbpB homolog, SSA_1593, shares 62 % sequence similarity.) The amylase

binding capacity of streptococci is critical for colonization of the oral cavity, as only animals secreting sufficient amounts of salivary α-amylase are colonized by bacteria capable of binding this protein (224, 253). AbpB has been demonstrated as an important mediator of rat oral cavity colonization in rats consuming a diet void of sucrose. However, AbpB was not required for oral colonization in rats consuming a sucrose diet, and the contribution of AbpB to colonization may be secondary to glucosyltransferase activity (253). AbpB is also predicted to act as a dipeptidase, though this function has not been characterized. In general cell surface exposed dipeptidases may mediate the local extracellular free amino acid concentration and assist in scavenging of proteins around the cell. Additional contributions may include proteolysis of imported peptides, and general cell maintenance by assisting in surface protein turnover (84).

Putative *S. sanguinis* Cwa protein adhesins

Five predicted *S. sanguinis* Cwa proteins contain the domain mediating collagen binding of the *S. aureus* adhesin Cna. Cna of *S. aureus* is an important virulence factor for multiple infections attributed to this organism, including infective endocarditis (99, 214). Cna is an MSCRAMM that contains two major domains, A and B, in addition to features directing extra-cytoplasmic export and cell wall anchoring. The collagen binding function has been localized within the A domain (245), and a trench that occurs within this domain accommodates the collagen triple helix. The A region can be subdivided into three sequential domains (N1, N2, then N3) that differentially contribute to collagen binding. Zong and colleagues deduced the “collagen hug” model of Cna binding to collagen (295).

The authors described progressive associations with collagen leading to secure sequestration of the collagen ligand. To summarize, binding is initiated by hydrophobic interactions between the N1-N2 subdomains and the collagen ligand. The N1 and N2 subdomains then wrap around the collagen triple helix and progressively hug the ligand until an extension of the N2 subdomain, referred to as "the latch" inserts into a trench on the N1 subdomain surface, resulting in secure ligand binding (295). Cwa proteins: SSA_0227, SSA_0805, SSA_1019, SSA_1663, and SSA_1666 are all predicted to contain an A domain-like region permitting collagen adhesion (pfam05737) (156). Additional members of the collagen binding MSCRAMM family have been identified in gram-positive pathogens including *E. faecalis, S. equi*, *S. mutans*, and *E. faecium*, which contain a region resembling the N1N2 domain (141, 184, 215, 223). Thus the proposed "collagen hug" model may be a conserved mode of collagen association in this protein family.

The B domain contains multiple repeats that are not required for collagen adhesion. Instead the B region is thought to provide flexibility and stability to Cna for projection of the A region away from the cell surface (53). Only *S. sanguinis* Cwa proteins containing pfam05737 were annotated as putative collagen binding proteins; of these SSA_0227, SSA_1663, and SSA_1666 also include a Cna B domain-like region (pfam05738) (156). Only the Cna B-like region was identified in Cwa protein SSA_2121; as this region alone is insufficient for collagen binding, SSA_2121 was not annotated as a collagen binding protein. The multitude of collagen adhesins predicted for *S. sanguinis* suggests that this phenotype is important for colonization and persistence in the oral cavity, and may contribute to development of IE.

SSA_1663, annotated as CpbA has also been shown to induce platelet aggregation (98). Expression of *SSA_1663*, and putatively the platelet aggregation phenotype attributed to this protein was demonstrated as a virulence determinant for IE in a past study (96).

The Antigen I/II proteins of oral streptococci are a family of multifunctional adhesins that promote fitness in the oral cavity. The antigen I/II proteins of *S. sanguinis*, SspC and SspD are encoded by *SSA_0303*, and *SSA_0956*, respectively. The well characterized antigen I/II proteins of *S. gordonii*, SspA and SspB have greater sequence identity to SspC of *S. sanguinis*, than SspD. However it remains likely that *S. sanguinis* antigen I/II homologs serve distinct functions, as has been described in *S. gordonii*. Heterologous expression of SspA and SspB in *L. lactis* identified exclusive adhesive characteristics of these proteins. In this heterologous system expression of SspA resulted in enhanced binding to salivary agglutinin glycoprotein over SspB, whereas expression of SspB resulted in enhanced adhesion to collagen I and *C. albicans* (102). Ssp polypeptides of *S. gordonii* additionally mediate attachment to *A. naeslundii*, an organism important in plaque development, and attachment to the periodontal pathogen, *P. gingivalis* (55, 136). In *S. gordonii sspA* and *sspB* are clustered at the same genomic locus, with *sspA* preceding the *sspB* gene. Despite this arrangement the genes are differentially expressed in response to environmental stimuli including pH, temperature and osmolarity (69). In *S. sanguinis sspC* and *sspD* genes are not clustered.

Antigen I/II was first described for *S. mutans*, in which the SpaP/Pac surface protein was first considered to be two distinct antigens; however antigen II was later determined to be a breakdown product of antigen I. This predominant surface protein of *S.*

mutans is required for dentinal tubule invasion (as are SspA and SspB of *S. gordonii*), and enhanced cariogenicity of this species (52, 152). The SpaP/Pac adhesin additionally mediates binding of *S. mutans* to collagen type I, fibronectin and fibrinogen, potentially for promotion of endodontic infection or adherence to fibrin clots (like those occurring in IE) (14). The sequence similarity to these previously characterized proteins suggests that SspC and SspD of *S. sanguinis* also mediate adhesion phenotypes.

SSA_0829 encodes the previously described serine rich protein, SrpA. SrpA of *S. sanguinis* is a glycoprotein that has been shown to associate with the platelet von Willebrand Factor receptor, glycoprotein Ib, in a sialic acid dependent manner, resulting in acceleration of platelet aggregation (122, 204). The *S. parasanguinis* SrpA homolog, Fap1 was required for adhesion to PS and biofilm formation in the presence of glucose (76). An additional SrpA homolog in *S. gordonii*, Hsa, was demonstrated to mediate invasion of HUVEC cells, suggesting a role for this protein in pathogenesis and IE (241). Hsa was shown to contribute to infectivity in the rat model of IE under competitive conditions (248). The particular adhesive property attributed to Hsa and related serine rich proteins is the recognition of sialoglycoconjugates. Binding of this ligand may mediate the multiple functions described for this protein family to date. The contribution of SrpA of *S. sanguinis* to IE was first investigated in the studies described here.

SSA_0904, SSA_0905 and SSA_0906 are homologous to the cell surface hydrophobicity (Csh) proteins A and B of *S. gordonii*. CshA and CshB of *S. gordonii* are antigenically related high molecular weight polypeptides that comprise fibrillar cell surface extensions (168). Both CshA and CshB, as their name implies have been shown to infer

hydrophobicity to *S. gordonii*, and expression of both proteins was shown to be necessary for colonization of the murine oral cavity (170). It has also been demonstrated that CshA and CshB of *S. gordonii* mediate invasion of HUVEC cells, as well as binding to fibronectin, and inter-bacterial adhesion with *A. naeslundii*, and *S. oralis* (169, 241). A CshA-related protein of *S. sanguinis* has also been implicated in aggregation of platelets, a phenotype correlated with virulence in development of IE (98). The *S. sanguinis* Csh-related proteins, SSA_0904, SSA_0905, and SSA_0906 were annotated as Csh-related protein A (CrpA), B (CrpB), and C (CrpC), respectively. In *S. gordonii* the *cshA* and *cshB* genes exist at distinct loci. In contrast the Crp encoding genes in S. *sanguinis* are clustered at the same locus, with *crpA* most 5' and *crpC* most 3'. Promoter and terminator elements were predicted for the intergenic regions separating *crpA*, *crpB*, and *crpC*, indicating these genes may not be co-transcribed. This possibility, in addition to adhesive properties ascribed to *S. sanguinis* Crp proteins require further investigation.

The predicted homolog for SSA_1023 is the von Willebrand factor (vWf) binding protein vWf1 of *S. lugdunensis* (188). vWf is a large glycoprotein occurring in plasma and the storage granules of platelets. vWf mediates hemostasis by binding to exposed matrix molecules, like collagen at sites of vascular endothelium damage to promote subsequent platelet adhesion and aggregation (221). A repetitive region at the C-terminus of the *S. lugdunensis* vWf1 mediates association with vWf. Recurring residues similar to the pattern identified for vWf1 are in the C-terminus of SSA_1023. vWf1 of *S. lugdunensis* is conserved among clinical isolates, suggesting a role for this protein in pathogenesis (188). Specifically vWf1 interaction with vWf at sites of vascular injury may facilitate

colonization of vascular lesions in the development of infection. For *S. sanguinis* and *S. lugdunensis* this may be particularly relevant for progression of IE. A similar repetitive region was identified in the immunodominant, muramidase released protein (Mrp) of *S. suis* (188). SSA_1023 homology with Mrp is restricted to the repeats occurring at the C-terminus of the protein. Mrp is expressed during infection; however it is not required for pathogenesis in porcine infection (236, 252).

A vWf type A domain (cd00198) was identified in SSA_1063 (156). The implications of having a vWf like structure on the surface of *S. sanguinis* are numerous, including enhanced platelet aggregation for colonization of vegetation components or enlargement of the vegetation, as well as colonization of subendothelial extracellular matrix proteins. Also, interaction between SSA_1023 and SSA_1063 might promote aggregation of this species though specific ligand binding. These intriguing possible attributes of a vWf-like streptococcal surface protein require further investigation.

SSA_1632, SSA_1633, SSA_1634, and SSA_1635 were described above in reference to their putative cognitive sortase, SrtC (SSA_1631). The *S. pneumoniae* gram-positive surface pilus was previously characterized (13). As in *S. sanguinis*, the pneumococcal pilin subunit genes, *rrgA*, *rrgB*, and *rrgC* are clustered with sortase enzymes mediating assembly of the pilus structure. Expression of this locus was associated with colonization of A549 lung epithelial cells, as well as competitive colonization of the murine respiratory tract. Expression of the surface pilus was shown to elicit a robust inflammatory response in challenged mice, suggesting that this surface structure has a profound impact on the severity of disease and the host response to infection (13).

Similarly, gram-positive pilus structures described for GBS elicit a protective response against re-infection (144). The *E. faecalis* endocarditis- and biofilm-related pilus locus encodes a surface pilus and pilus related sortase (Bps), and is a determinant of *E. faecalis* biofilm formation and virulence *in vivo* (121, 183). The contribution of a putative *S. sanguinis* surface pilus in IE had not been investigated prior to this study.

SSA_0565 is annotated as a conserved hypothetical protein in GenBank; however the protein sequence has 26 % identity and 42 % similarity to Sec10 of *E. faecalis*. In *E. faecalis* Sec10 (also referred to as PrgA) is encoded by pCF10-borne *prgA*. The components of pCF10 mediate gene transfer via pheromone inducible conjugative plasmids (119). This event requires aggregation of donor (pCF10 encoding) and recipient cells in liquid culture. The aggregation substance mediating mating aggregate formation for genetic transfer, Asc10 (also referred to as PrgB) is also encoded on pCF10 (195). Sec10 acts to promote genetic transfer to naïve cells by preventing conjugation between donor cells (61). *E. faecalis* cell surface exposure of both Sec10 and Asc10 has been demonstrated by SEM (194). Conjugative plasmids are not part of the genome of *S. sanguinis* SK36; however the genetically competent *Streptococcus* may be capable of DNA acquisition by this mechanism. If SSA_0565 has a similar function to Sec10, this ORF may be a remnant of a past conjugative event. In the case of *E. faecalis* conjugative plasmids are of clinical significance due to their role in dissemination of antibiotic resistance cassettes; pCF10 confers tetracycline resistance to recipient cells (60).

A BLASTP search for homologs of SSA_0684 resulted in positive hits for an *S. pneumoniae* FibA subunit-like protein (GI:147930705) (33 % ID, 49 % similarity in

residues 766 – 1082 of SSA_0684), and the FibA-subunit of the fibril-like surface structure of *Peptostreptococcus micros* (GI:4106521) (29 % ID, 39 % similarity in residues 500 – 800 of SSA_0684). *P. micros* is one of the causative agents of periodontal disease (87). Two morphology types have been described for *P. micros*, a smooth and a rough type (265). The rough type morphology is determined by production of FibA-containing fimbriae (131). Interestingly, the rough type variants switch to the smooth morphology when cultured *in vitro* (130). However, both variants produce abscess formation in a murine skin model, and both have been isolated from periodontal patients (129, 264).

Differences between SSA_0684 and FibA occur at the extreme C-terminus of the protein, in the region determining cell surface association. Specifically, the Cp1 family sequence repeats reported for the extreme C-terminus of the 393 aa FibA protein are not conserved in SSA_0684 (273). These repeats are thought to confer carbohydrate binding capabilities to the protein, and in the case of *S. pneumoniae* PspA, this region is involved in anchoring of the proteins to lipoteichoic acid in a choline-mediated interaction (273, 290). The lack of conservation of this region is unsurprising given the identification of an LPXTG pattern at the C-terminus of SSA_0684. Similarity to the remainder of FibA suggests the proteins may share other functional properties that remain to be determined.

Detection of *S. sanguinis* Cwa proteins by MudPIT

Multidimensional protein identification technology (MudPIT) was used for detection of *S. sanguinis* proteins in two cellular fractions, a mutanolysin released cell wall associated protein fraction, and the remaining protoplast fraction. The MudPIT technique

requires proteolytic digestion of the protein pool into fragments that are separated by strong cation exchange chromatography followed by reverse phase chromatography. The peptides are then subjected to tandem mass spectrometry MS, for determination of the mass-to-charge ratio of each fragment. Peptide information derived is searched against an *in silico* proteolytic digestion database of predicted ORFs to identify actual ORFs detected. The MudPIT analysis was performed with the assistance of Dr. Patricio Manque of the Center for the Study of Biological Complexity at Virginia Commonwealth University.

Fourteen of the predicted *S. sanguinis* Cwa proteins ORFs were detected by MudPIT technology (Table 11). Detection of the hypothetical protein SSA_1301 in both protoplast and cell wall (mutanolysin released) fractions, and the hypothetical protein SSA_2121 in the cell wall associated fraction resulted in reclassification of these ORFs as conserved uncharacterized (reflected in Table 10). Antigen I/II homologs SspC and SspD (SSA_0303, and SSA_0956, respectively) were both detected in the cell wall and protoplast fraction suggesting that these proteins were abundant in the cell at the time of sampling for MudPIT analysis. Proteins containing the collagen binding domain pfam05737 (156) were also detected as being simultaneously present. CbpA (SSA_1663) and SSA_0805 were detected in the cell wall associated fraction, whereas SSA_1019 was detected in the protoplast fraction. SSA_1023, which shares homology with the *S. lugdunensis* vWf1, and *S. suis* Mrp was detected in the cell wall fraction. The CshA related proteins, CrpC (SSA_0906) and CrpA (SSA_0904) were both detected as being abundant in the experiment. As for cell wall associated enzymes of *S. sanguinis* SK36, both SSA_0243 (the putative cyclo-nucleotide phosphodiesterase), and SSA_1065 (putative

Table 11. Detection of *S. sanguinis* Cwa proteins by MudPIT

S. sanguinis (*Ss*) Cwa proteins detected as abundant by MudPIT analysis of cell wall-associated proteins and protoplast proteins are described. The E-value given for each protein reflects the probability that the protein was identified incorrectly. A smaller E-value reflects a decreased probability that the protein was identified erroneously (cutoff 1.00E+00). Proteins not listed may have been expressed under the condition tested, but the protein abundance may be under the limit of detection.

Table 11.

***Ss* ORF**	**Description**	**Protoplast**	**Cell wall**
SSA_1666	Collagen binding domain; pfam05737		7.56E-05
SSA_0805	Collagen binding domain; pfam05737		6.50E-06
SSA_1019	Collagen binding domain; pfam05737	1.00E+00	
SSA_1632	Putative pilin subunit		2.69E-08
SSA_0956	SspD	1.00E+00	4.52E-03
SSA_0303	SspC		6.89E-09
SSA_1023	Von-Willebrand factor binding protein		1.38E-03
SSA_0904	CrpA	1.00E+00	
SSA_0906	CrpC	5.24E-03	1.60E-06
SSA_0684	FibA fibril subunit-like protein	7.41E-03	1.07E-04
SSA_1234	5' Nucleotidase		8.56E-06
SSA_1065	Beta hexosaminidase-like protein	2.03E-03	
SSA_0243	Cyclo-nucleotide phosphodiesterase	6.37E-04	1.00E-02
SSA_2121	Conserved uncharacterized		1.52E-04
SSA_1301	Conserved uncharacterized	1.26E-05	1.95E-09

beta-hexosaminidase) were detected in the protoplast fraction. SSA_0243 was also detected in the cell wall fraction. That additional predicted Cwa proteins were not detected does not exclude the possibility of their expression or presence at the time of protein fractionation. This method is particularly sensitive to the concentration of the protein mixture analyzed, proper denaturation of proteins prior to digestion, and detection of proteins that are abundant in the sample.

In vitro transposition to create an STM mutant pool

The 31 predicted cell wall-associated protein encoding ORFs, and 3 sortase-like ORFs described above were subjected to *in vitro* transposition. Each ORF was PCR amplified to create a 2.5 kb product, which was subsequently used as a target for *in vitro* transposition with the pJFP1 encoded *magellan2* mini-transposon containing a unique signature tag. The target amplicon size of 2.5 kb was of conscious design based on previous experiments in the lab indicating that regions of homology ≥ 1kb flanking non-consensus sequence yield greater transformation efficiency (Dr. Todd Kitten, personal communication). With this in mind, primers used for target gene amplification were designed with the preferred insertion site roughly equidistant from either end of the amplicon. Following end-filling and ligation reactions, transposition products were transformed into *S. sanguinis* SK36. A diagnostic PCR reaction coupling ORF-specific flanking primers (up or dn) with a transposon inverted repeat primer (Mout) was used to estimate the site of transposon insertion site in Cm^r transformants. Among the clones screened those with predicted insertion within the ORF, close to the 5' end of the ORF

were selected for further studies. The estimated insertion site for each ORF targeted, as well as the signature tag used is indicated in Table 12. The accuracy of diagnostic PCR for prediction of the transposon insertion location was tested for 7 STM mutants. In all strains analyzed DNA sequencing confirmed that transposon insertion occurred very close to, if not within the predicted interval obtained from diagnostic PCR (Table 12).

The signature tags used in mutant generation were prescreened by Todd Kitten for consistent and specific DNA hybridization. For simplicity each of the 40 tags approved for mutagenic application were assigned a number between 1 and 40. STM mutant strain designations reflect the signature tag number employed, as well as whether the target ORF coded for a Cwa protein (CWA), or a sortase enzyme (SRT).

Successfully mutagenesis of all 34 predicted proteins by the directed STM technique resulted in the creation of a 34-strain mutant library.

In vivo screening of the STM pool

In previously described pilot studies the rabbit model of IE was optimized for evaluation of *S. sanguinis* virulence factors (198). This model of infection was superior to the rat model of IE (179, 180) due to apparent bottlenecking during STM trials in the rat model, not observed in the rabbit model (198). The *S. sanguinis* inoculum size was also previously optimized for the rabbit model. In three inocula tested, 2×10^8, 2×10^9, and 2×10^{10} CFUs, the highest inoculum proved lethal within 24 hours of injection, whereas the other two inoculum sizes had no effect on the observed results. All doses of bacteria administered resulted in infection, and reproducible recovery as detected by dot blot

Table 12. Estimation of mini-transposon insertion location

For each predicted Cwa protein or sortase ORF *in vitro* transposition was performed with a different *magellan2* mini-transposon incorporated signature tag (signature tag #). Following transformation of transposition products, PCR amplification of cells from Cm^r transformants was used for estimation of the transposon insertion site. Mutants were chosen for *in vivo* testing based on the presence of the transposon insertion within the gene of interest and its proximity to the start of the ORF, with preference given to mutants with interruption closer to the beginning of the gene. The estimated distance between the transposon insertion site and the start site of the gene is indicated for each selected mutant. For several mutants, the actual insertion site was determined by DNA sequencing. PCR estimation was quite accurate, with the actual site of insertion very close to, if not within the predicted interval.

Table 12.

SSA gene #	Signature Tag #	ORF size (bp)	Estimated site	Actual site
SSA_1112	1	1575	413 - 445	
SSA_2023	2	4215	879 - 916	
SSA_1663	3	4533	924 - 935	923
SSA_1591	4	2046	549 - 600	
SSA_0805	5	1674	183 - 369	190
SSA_1634	6	1458	187 - 234	
SSA_1633	7	1437	376 - 477	
SSA_1632	8	1437	380 - 425	
SSA_0956	9	4113	327 - 333	
SSA_0227	10	2081	1065 - 1081	1085
SSA_0565	11	2583	452 - 490	478
SSA_1301	12	2559	87 - 181	113
SSA_0303	13	4518	445 - 480	
SSA_0273	14	1413	196 - 204	
SSA_1023	15	2673	883 - 894	
SSA_1019	16	2382	394 - 535	485
SSA_1631	18	885	512 - 567	
SSA_0905	19	5898	1114 - 1157	
SSA_0906	20	8007	1223 - 1254	
SSA_1219	21	753	175 - 279	
SSA_0684	22	3819	366 - 412	
SSA_1234	23	2424	401 - 414	
SSA_0022	24	840	416 - 489	
SSA_1063	25	1350	318 - 372	
SSA_2020	26	3105	808 - 853	
SSA_0167	27	999	311 - 397	
SSA_0146	28	2391	196 - 237	
SSA_0453	29	3699	365 - 435	
SSA_1666	30	1935	560 - 579	569
SSA_1065	31	2790	835 - 885	
SSA_0904	33	8970	824 - 869	
SSA_1750	34	2247	417 - 510	
SSA_2121	38	4683	322 - 362	
SSA_0243	40	2409	364 - 400	

hybridization (198). In the STM screens described below the inoculum administered ranged from 1 x 10^7 CFUs to 2 x 10^8 CFUs.

Three independent *in vivo* screens were conducted for analysis of the STM mutant pool. Each strain was tested in at least two independent experiments, in at least three rabbits in each experiment. The dot blot DNA hybridization technique used for detection of strains inoculated into the rabbits (the input pool), and strains recovered from the rabbits (the output pool) is depicted in Figure 30. Figure 34 presents a summary of strains tested in each experiment, and the results obtained. The first STM experiment was conducted with a portion of the strains composing the final, complete pool. The pool tested included previously unevaluated strains CWA4, CWA9, CWA10, CWA12, CWA13, CWA14 and CWA15. Dot blot hybridization confirmed the recovery of STM mutants CWA4, CWA10, CWA13 and CWA9, suggesting that ORFs mutagenized in these strains are not essential in the rabbit model of IE. The competitiveness of CWA15 was difficult to determine, as detection of this strain among recovery pools was inconsistent (data not shown). CWA12 displayed a strong signal in the input pool, yet weak signal in the recovery pools, suggesting that this strain did not survive negative selection *in vivo*.

The next STM *in vivo* analysis was conducted using the entire 34 strain mutant pool. Of the five rabbits tested in this experiment two had shorter infection duration than others, as a result of apparent animal duress. The study was carried to 20 hours post-inoculation for the remaining rabbits. Despite shorter duration of infection in two of the rabbits, the hybridization patterns of these rabbits were consistent with those experiencing 20 hours of infection. Unfortunately the interpretation of results was confounded by

variability in signature-tag hybridization observed for the input pool dot blot. Because conclusiveness is dependent on presence of the strain in the inoculum, definitive results were obtained only for strains exhibiting a strong signal on the input blot (data not shown). Strains retaining virulence are indicated in Figure 34.

Results representative of the final independent experiment, in which the entire mutant pool was screened, are shown in Figure 35**.** Bacterial recovery from one of the three rabbits tested was below the range yielding reproducible results; therefore only results from rabbits with higher recoveries are included. The input blot of this experiment showed minimal variability with high signature-tag chemiluminescent signal intensity of signature tags for the majority of strains analyzed. Likewise, the majority of strains tested were recovered, indicated by the strong chemiluminescence observed in the output blots. Of 34 mutants screened, 29 retained virulence as detected by signature-tag hybridization (Figure 34).

Relative comparison to the negative control pJFP1 signature tags ('N', Fig. 35) highlights differences among strains. For example, signature tag 12, representing strain CWA12 exhibited stronger signal intensity in the input blot, but negative control signal intensity in the recovery blots. That this tag was not detected in the recovery pool suggested that this strain did not survive negative selection *in vivo*. Results for four strains, SRT21, SRT24, CWA15 and CWA30 were inconclusive due to low signal intensity of signature-tag hybridization on the input pool membrane.

An *in vitro* growth study performed in conjunction with the animal study suggests that poor detection is not entirely reflective of reduced growth *in vitro* (Fig. 36).

Figure 34. Summary of STM screen results

S. sanguinis ORF	Tag	STM Experiment 1	2	3
SSA_1112	1	ND	•	•
SSA_2023	2	ND	•	•
SSA_1663	3	ND	•	•
SSA_1591	4	•	•	•
SSA_0805	5	ND	◘	•
SSA_1634	6	ND	•	•
SSA_1633	7	ND	•	•
SSA_1632	8	ND	•	•
SSA_0956	9	•	◘	•
SSA_0227	10	•	•	•
SSA_0565	11	ND	•	•
SSA_1301	12	○	◘	○
SSA_0303	13	•	•	•
SSA_0273	14	•	•	•
SSA_1023	15	◘	◘	◘
SSA_1019	16	ND	•	•
SSA_1631	18	ND	•	•
SSA_0905	19	ND	◘	•
SSA_0906	20	ND	•	•
SSA_1219	21	ND	◘	◘
SSA_0684	22	ND	•	•
SSA_1234	23	ND	•	•
SSA_0022	24	ND	◘	◘
SSA_1063	25	ND	◘	•
SSA_2020	26	ND	•	•
SSA_0167	27	ND	•	•
SSA_0146	28	ND	•	•
SSA_0453	29	ND	•	•
SSA_1666	30	ND	◘	◘
SSA_1065	31	ND	•	•
SSA_0904	33	ND	◘	•
SSA_1750	34	ND	•	•
SSA_2121	38	ND	•	•
SSA_0243	40	ND	•	•

Key:	
ND	Not determined
•	Competitive
◘	Inconclusive
○	Not competitive

Figure 35. Representative dot blot hybridization results

A) The layout of STM-plasmid containing membranes is shown. Each number represents a Cwa protein mutant bearing that unique signature tag. Tags not used for mutation are negative controls (N), and are indicated by purple boxes.

B) Actual hybridization results are shown for the input pool (at the left), and output pools of bacteria recovered from rabbits. The presence of a strain is indicated by strong chemiluminescent signal intensity. Low signal intensity indicates diminished presence of a strain or poor hybridization for other reasons. Purple boxes indicate background signal intensity of negative control plasmid DNA.

Figure 35.

A. STM dot blot layout:

1	9	N	25	33
2	10	18	26	34
3	11	19	27	N
4	12	20	28	N
5	13	21	29	N
6	14	22	30	38
7	15	23	31	N
8	16	24	N	40

B. Hybridization results:

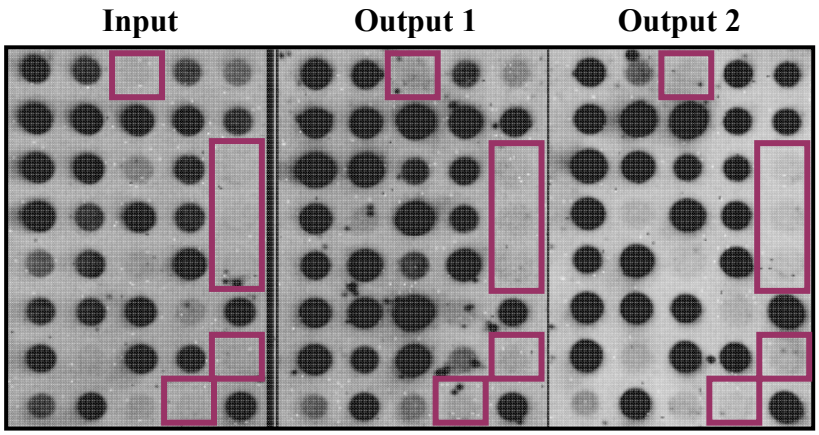
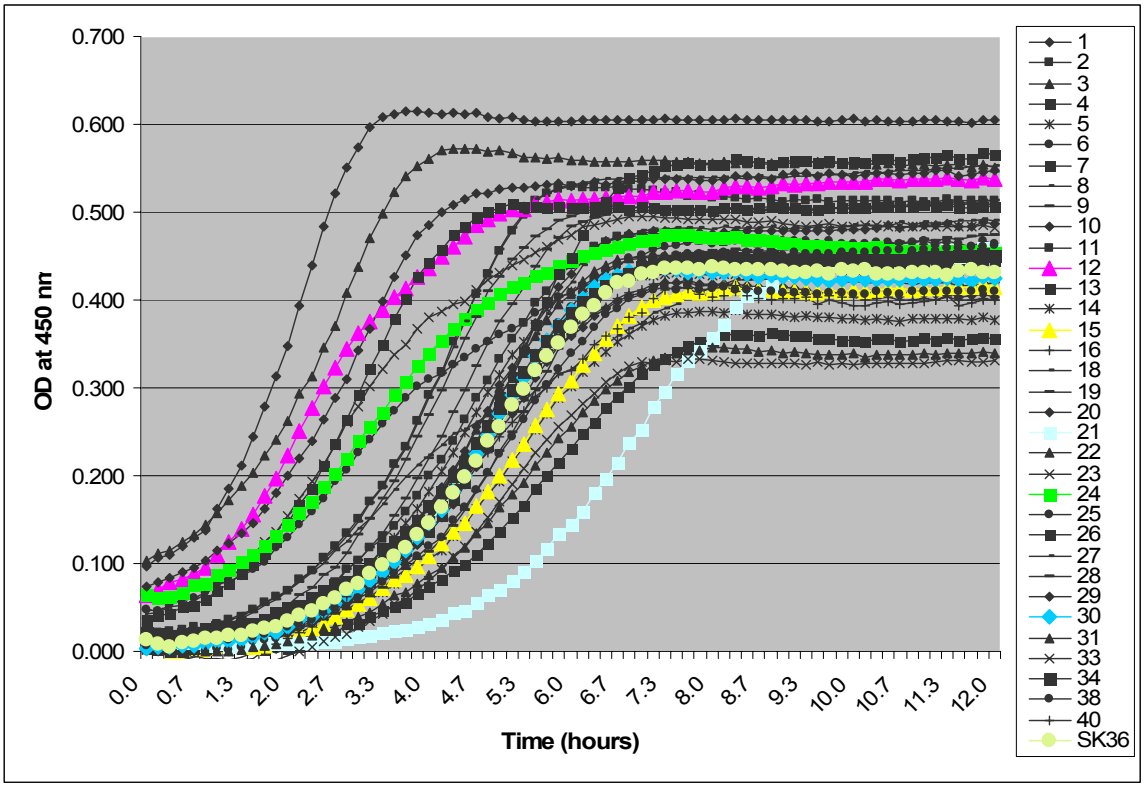


Figure 36. *In vitro* growth of STM strains

STM strains were grown *in vitro* overnight concurrent with the third rabbit experiment. Strains retaining competitiveness *in vivo* are in gray symbols. Strains requiring further analysis, due to apparent avirulence, CWA12, or poor signal intensity in the inoculum, CWA15, SRT21, SRT24, and CWA30 are indicated by colored symbols. Each strain is indicated by its signature-tag number.

Figure 36.


OD at 450 nm
0.700
0.600
0.500
0.400
0.300
0.200
0.100
0.000
0.0
0.7
1.3
2.0
2.7
3.3
4.0
4.7
5.3
6.0
6.7
7.3
8.0
8.7
9.3
10.0
10.7
11.3
12.0
Time (hours)
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
18
19
20
21
22
23
24
25
26
27
28
29
30
31
33
34
38
40
SK36

Strains CWA15, CWA24 and CWA30 grew comparably to the parental strain SK36, whereas SRT21 grew more slowly than wild-type. An alternate explanation for low signal intensity in the input blot is degradation of signature tag DNA impeded DNA:DNA hybridization required for detection. To eliminate this complication strains yielding inconclusive results were further analyzed by competitive index assays.

Re-evaluation of STM strain stability without antibiotic selection

Before further testing of STM mutants we wanted to establish stability of the mutation without antibiotic selection of the transposon. Cells harvested from overnight cultures of each STM strain in BHI (no antibiotic) were used as template in PCR. The region of transposon insertion in each strain was PCR amplified by specific primers 5' and 3' to the ORF interrupted (Appendix B). The presence of the transposon was suggested by the size of the PCR product obtained. Maintenance of the transposon results in an amplicon of 3.5 to 4 kb, whereas absence of the transposon is indicated by a smaller band of 2 to 2.5 kb. For all strains evaluated, a larger band of ~3.5 to 4.0 kb was detected which represents the ORF containing the transposon. For CWA11, CWA12, CWA14 and SRT21 an additional smaller band of was observed (data not shown). The double band pattern observed in these strains may be attributed to non-specificity of primers used, or the presence of more than one genotype. To confirm these results four colonies of strains CWA11, CWA12, CWA14 and SRT21 grown in the presence of Cm were subsequently cultured in liquid BHI with Cm overnight. The stationary phase cells were then subcultured in BHI without Cm for a second overnight period. Cells were harvested from

the plus and minus Cm conditions and used as template for PCR, as described above. The two conditions tested would indicate the dominant allele in the presence and absence of antibiotic selective pressure. PCR products observed for CWA11, CWA12, and CWA14 were identical in the presence and absence of antibiotic (Fig. 37A and 37B), thus confirming the stability of the mutant allele without Cm selection. However the mutant allele was not detected for strain SRT21 in the absence of Cm (Fig. 37B). To correlate these findings with animal experiment data the strains CWA11 and CWA14, with signature-tags 11 and 14, respectively, exhibited strong DNA hybridization signal intensity in both input and output blots (Fig. 35B). CWA12, with signature-tag 12 showed strong signal intensity in the input blot, but not in output blots (Fig. 35B). Transposon maintenance in strain CWA12 independent of Cm selection suggests that low signal intensity of this strain following *in vivo* competition is cannot be attributed to genotypic changes in this strain. In contrast the apparent loss of the mutant allele in SRT21 could explain the poor dot blot hybridization observed for this strain during *in vivo* screening (Fig. 35B, tag 21). For STM screening strains were cultured *in vitro* without antibiotic for 3 hours prior to inoculation, which could provide adequate time for out-growth of the strain lacking the selective marker.

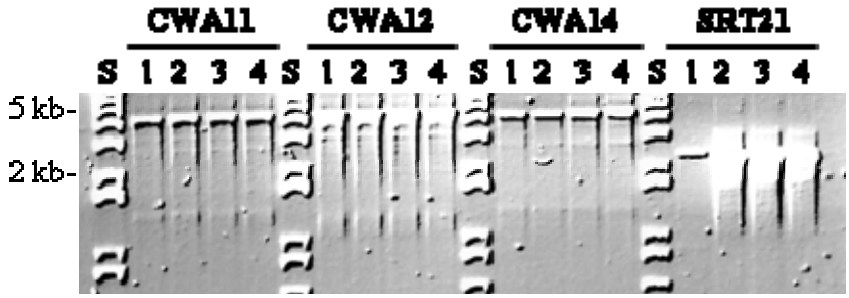
To address this issue a new *srtA* mutant was developed in the SK36 genetic background. A gene SOEing construct was designed for insertion of *aad9* of pR412 between the first 29, and last 23 codons of *srtA*, to couple the deletion of *srtA* with the insertion of a selective marker. The inconclusive *in vivo* results obtained for another sortase STM mutant, SRT24 led to the decision to generate a novel *srtB* mutant as well.

Figure 37. Detection of the transposon in the absence of Cm selection

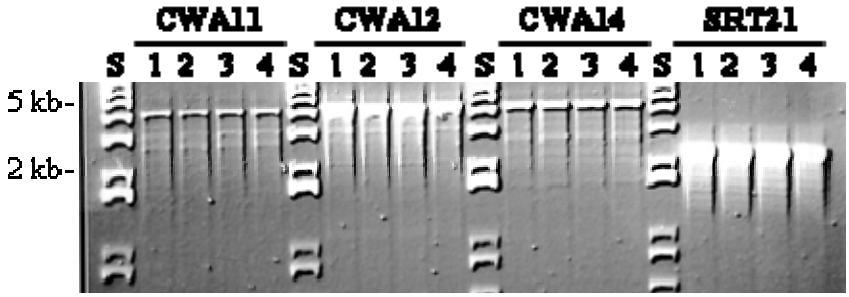
Four colonies (1 to 4) of *S. sanguinis* strains CWA11, CWA12, CWA14 and SRT21 were analyzed by PCR product size determination for the presence of the transposon in strains cultured in the presence of Cm (selective for the transposon) (A). Stationary phase cells in BHI containing Cm were subcultured in the absence of Cm (B). Sizes of PCR products for CWA11, CWA12, and CWA14 were similar (~ 4 kb in size) in the presence and absence of selective antibiotic (A and B), indicating the presence of the transposon in these strains under both growth conditions. In contrast, a smaller dominant band (~ 2 kb) is observed for SRT21 after growth in Cm overnight (A), which becomes the only observable band following subculture in BHI without Cm (B). The amplicon size expected for *srtA::magellan2* was 3.459 kb, the anticipated amplicon for WT *srtA* was 2.077 kb.

Figure 37.

A. PCR amplification of strains from BHI with Cm



B. PCR amplification of strains from BHI



Gene SOEing was used to create a mutant construct by insertion of the kanamycin resistance cassette *kan* between the fist 27 and last 16 codons of *srtB*, thus combining deletion of *srtB* with the insertion of a selective marker for kanamycin resistance. *S. sanguinis* SK36 was transformed with the gene SOEing mutagenic constructs, and transformants selected by spectinomycin or kanamycin resistance conferred by mutation of *srtA* or *srtB*, respectively. The resulting *srtA* mutant was designated JFP42, and the *srtB* mutant designated JFP44.

Competitive index studies

Dot blot hybridization qualitatively differentiated between STM mutants that retained competitiveness in the rabbit model of IE, a single putatively uncompetitive strain (CWA12), and strains of inconclusive competitiveness (CWA15, CWA30, SRT21 and SRT24). These results are merely suggestive of exploitive *vs.* underrepresented strains, and relative attenuation for particular strains is difficult to determine. A competitive index (CI) assay was used to assign a quantitative value to the competitiveness of CWA strains requiring further testing. The previously described Em^r derivative of SK36, JFP36 (Chapter IV) was used as the competitive control strain in these experiments. In the CI algorithm the ratio of mutant/control strain CFUs recovered from rabbits is divided by the ratio of mutant/control strain CFUs administered to rabbits. A CI value of 1 indicated equal competitiveness between the mutant and WT control. A CI value less than one indicated reduced competitiveness of the mutant *vs.* the WT control. A CI value greater than one

suggests that mutant competitiveness exceeds that of the WT control. Statistical analyses were then applied to determine whether the difference in competitiveness was significant

CI analysis of CWA12

To determine the competitiveness of CWA12 relative to the virulent control, JFP36, the two strains were co-inoculated into three catheterized rabbits in two independent studies. Contrary to preliminary evidence from dot blot hybridization implicating CWA12 as a noncompetitive strain, the mean competitive index value obtained for this strain relative to JFP36 was 0.959 (Fig. 38). Therefore CWA12 was equally competitive to the control strain. The conflicting results from these assays may be attributed to reduced plating efficiency of this strain as a function of increased chain length. In dot blot hybridization, signature tag probes are developed from strains in liquid culture for the input pool, *vs. in vivo* recovered colonies subsequently sub-cultured in liquid media for the output pool. If a strain exhibits a poor plating efficiency (forms fewer colonies) at the same optical density as other strains, then it may be construed that the strain is not present, or is diminished in the pool. The design of the CI study would account for differences in plating efficiency of strains, as the ratios of CFUs of strains both inoculated into and recovered from rabbits are compared. When viewed by phase contrast microscopy early stationary phase cultured CWA12 formed much longer chains that the competitive control, JFP36. In addition CWA12 formed roughly 0.5 colonies for every 1 colony of JFP36 at an equal optical density of 0.8. That dot blot hybridization is sufficiently sensitive to detect and translate this difference is very encouraging. Our findings from the CI assay, coupled to observations regarding chain length and plating

efficiency of CWA12 confirm that expression of the ORF mutated in CWA12 is not required for full competitiveness *in vivo*.

CI analysis of CWA15

The competitiveness of CWA15 relative to JFP36 was tested in three rabbits between two CI experiments. The two strains were co-inoculated into catheterized rabbits, with competitiveness assessed 20 hours later. In contrast to the inconclusiveness of dot blot hybridization results, the CI assay definitively showed CWA15 to be significantly more competitive than JFP36 *in vivo*, with a mean CI value of 1.547 (Fig. 38). This result suggests the ORF interrupted in CWA15, *SSA_1023*, is not required for full competitiveness, and in fact expression of *SSA_1023* may impede competitiveness in experimental early infective endocarditis.

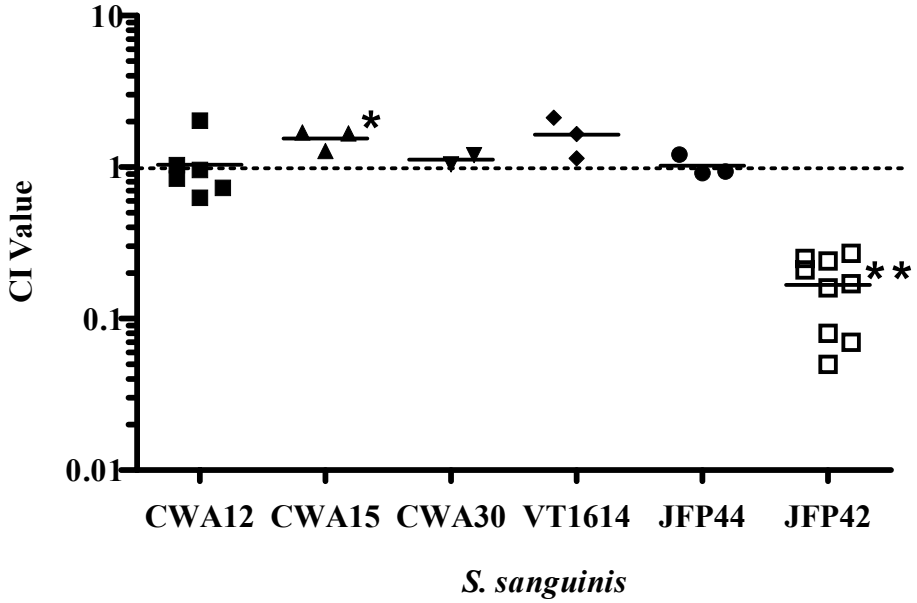
CI Analysis of CWA30

Inconclusiveness of CWA30 in the initial STM screen prompted further investigation of this strain in a CI assay. A CWA30 and JFP36 co-inoculum was tested in two rabbits in a single animal trial. The mean CI value obtained for CWA30, 1.120 suggests that expression of *SSA_1666*, the putative Cwa protein ORF mutated in this strain, does not contribute to competitiveness *in vivo* (Fig. 38). SSA_1666 was annotated as a putative collagen binding protein, based on identification of conserved domains of the Cna collagen binding protein of *S. aureus*, pfam 05737 and 05738 (156). At least four additional putative collagen adhesins and homologs of two characterized collagen adhesins have been identified in *S. sanguinis* SK36. The abundance of Cna-like proteins of *S. sanguinis* suggests that collagen adhesion is important for fitness in the oral niche.

Figure 38. CI analysis of *S. sanguinis* SK36 derivatives

CI comparisons were performed for *S. sanguinis* STM mutants, CWA12, CWA15, and CWA30 relative to the Em^r competitive control JFP36. Isogenic derivatives of SK36, JFP44 and JFP42 were also compared to JFP36. Competitiveness of the *srpA* mutant VT1614 was determined relative to SK36. The scatterplot includes the CI calculated for each animal (symbol), and the mean CI value for the mutant (bar). The dotted line at a CI of one represents the CI value for equal competitiveness. The STM strain CWA15 was significantly more competitive than JFP36 (*= $P = 0.0421$). The *srtA* mutant JFP42 was significantly less competitive than JFP36 (**= $P < 0.0001$). The table includes the CI phenotype for each strain, a description of the mutagenized ORF, the mean CI value for each strain, and the range of CI values obtained.

Figure 38.



Mutant	Phenotype	Description ORF	Mean CI (Range)
CWA12	Competitive	Conserved Uncharacterized	0.959 (0.63 - 2.03)
CWA15	Hyper-competitive	FibA/MRP-like	1.547 (1.28 - 1.69)
CWA30	Competitive	Collagen binding domain (pfam05737)	1.120 (1.04 - 1.20)
VT1614	Competitive	SrpA	1.588 (1.15 - 2.12)
JFP44	Competitive	SrtB	1.011 (0.91 - 1.21)
JFP42	Noncompetitive	SrtA	0.167 (0.05 - 0.27)

Whether this phenotype correlates with development of infective endocarditis has yet to be tested. One complication of such a study would be the abundance of proteins mediating such an interaction, and likelihood of functional redundancy among these proteins.

CI analysis of VT1614

To determine the contribution of SrpA in competitiveness *in vivo* a previously described isogenic *srpA* mutant of SK36, VT1614 (204) was compared to the parental strain in three rabbits in a single CI experiment. At twenty hours of infection bacteria were recovered and plated to determine the CI value of VT1614. The mean CI value for VT1614, 1.58 demonstrates that mutation of *srpA* does not negatively affect competitiveness *in vivo* (Fig. 38). This is particularly interesting given evidence that the *S. gordonii* SrpA homolog, Hsa contributes to competitiveness in the rat model of IE (248). However, the previous evaluation did not account for the number of CFUs administered by computation of a CI value; instead only recovery was expressed as CFUs/mg vegetation.

CI analysis of JFP44

The *SSA_0022*, *srtB* homolog mutant, JFP44 was compared to JFP36 in three rabbits in a single CI experiment. The mean CI derived for JFP44, 1.02, confirms that expression of *srtB* is not required for early infective endocarditis (Fig. 38). The wild-type competitiveness exhibited by JFP44 suggests that proper cell localization of the unknown SrtB transpeptidase substrate(s) is not required for the onset of disease.

CI analysis of JFP42

Initial competitive index analysis of the *srtA* mutant JFP42 in three rabbits revealed the strain to be significantly less competitive that the competitive control strain, JFP36

with a mean CI value of 0.07 (P= 0.0061). This experiment was repeated in two additional independent experiments with three rabbits each. The subsequent studies confirmed that JFP42 showed a significant reduction in competitiveness, with an overall mean CI value of 0.167 ($P < 0.0001$) (Fig. 38). The competitive phenotype of this mutant suggests the expression of *srtA* is required for full competitiveness *in vivo*. SrtA is likely the major, housekeeping sortase in *S. sanguinis*, responsible for anchoring the majority of Cwa proteins. The observed decrease in competitiveness for JFP42 in conjunction with results from the comprehensive analysis of *S. sanguinis* Cwa proteins suggests that development of *S. sanguinis* IE is a multi-factorial process mediated by more than one Cwa protein dependent on SrtA.

In vitro CI comparison of JFP42 and JFP36

An *in vitro* CI assay was used to determine whether JFP42 exhibited a general growth defect *in vitro* owing to mutation of *srtA*. A mutant with a general growth defect would be expected to perform less well in a CI assay, and this experiment would indicate whether the described *in vivo* findings were attributed to attenuated growth rather than reduced colonization efficiency or virulence. For this experiment strains were adjusted to an equal optical density and then co-cultured for the typical infection duration (~20 hours). The *in vitro* CI value was derived by comparing CFUs of strains prior to co-culture to CFUs following co-culture. The mean *in vitro* CI value obtained for JFP42 *vs.* JFP36 from five independent experiments was 0.58. If co-culture had no effect on growth as determined by CFUs, a mean CI value of 1.0 would be expected. As 0.58 (range: 0. 2 –

1.302) is less than 1.0, it appears that JFP42 exhibits a slight growth defect *in vitro* during co-culture with JFP36. However the CI value is not significantly different from a value of equal competitiveness (P= 0.2185), and it is unlikely this *in vitro* effect accounts for the reduced CI value for JFP42 *in vivo*. While this type of assay indicates strain-specific growth deficiencies, it does not replicate the selective pressures encountered by strains during infection.

Individual inoculum comparison of JFP42 and SK36

One possible complication of co-inoculation of strain in the CI assay is phenotypic complementation of the mutant by the presence of the WT strain. For example, if a mutant is capable of adhering to the WT strain (i.e. piggybacking), then it would benefit from adhesins required for colonization of the vegetation expressed by the wild-type. Also, secreted proteins from the WT strain might also prove advantageous to the mutant during co-infection. The possibility of complementation was examined by single-strain infections.

Four catheterized rabbits each were inoculated with JFP42 or SK36, and recovery of bacteria was assessed 20 hours later. In our hands vegetation size is often an indicator of bacterial recovery, with larger recoveries yielding more reliable results obtained from larger vegetations. To control for differences in the size of vegetations recovered, CFUs were normalized by vegetation weight.

In this study four rabbits received 1.84×10^8 CFUs of JFP42, and four rabbits received 6.55×10^7 CFUs of SK36. The mean recovery for rabbits receiving JFP42 was 8.69 log CFU/g vegetation. The mean recovery for rabbits receiving SK36 was 9.03 log

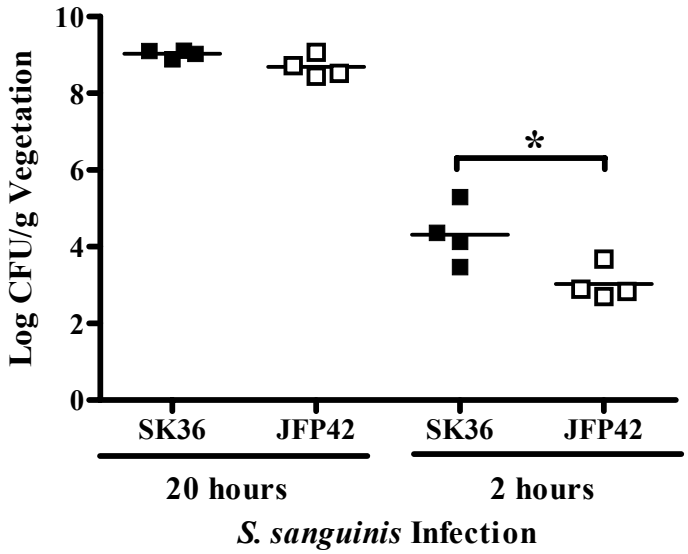
CFU/g vegetation (Fig. 39). The difference in recovery between the two rabbits was not significant (*P*= 0.1036), and therefore virulence of JFP42 is similar to SK36 when monitored 20 hours after infection. This is in contrast to CI studies in which JFP42 was significantly less competitive. The actual defect for JFP42 may be in competitive colonization of the vegetation, as it is apparent that in the absence of competition JFP42 can develop an infection similar to wild-type. An alternative possibility is that JFP42 is attenuated for vegetation colonization, yet grows to levels similar to wild-type in the intervening period between colonization and termination of the experiment.

In an *in vivo* experiment of shortened infection duration (2 hours) we examined whether JFP42 exhibited attenuated colonization of the vegetation in the absence of a competitor. Four catheterized rabbits received 1.35 x 10^7 CFUs JFP42, and four catheterized rabbits received 2 x 10^7 CFUs SK36. Two hours later the infection was terminated, and recovered bacteria were normalized to the vegetation weight collected. At two hours of infection the mean recovery for JFP42 was 3.03 log CFU/g vegetation (range: 2.7 – 3.68 log CFU/g). The mean recovery was greater for SK36, 4.314 log CFU/g vegetation (range: 3.47 – 5.3 log CFU/g) (Fig. 39). This difference is statistically significant (*P*= 0.0422), and suggests that JFP42 is attenuated for colonization of the vegetation relative to the parental strain, SK36. This deficiency is likely exacerbated when JFP42 is competed against a control in an experiment requiring co-inoculation of bacteria. In this scenario the control with an arsenal of Cwa proteins lacking in the *srtA* mutant would be better suited for adherence to or association with vegetation components in the colonization process and would readily out-compete a strain lacking these determinants.

Figure 39. Individual inoculum comparison of JFP42 and SK36

Catheterized rabbits received a single strain inoculum of JFP42 or SK36 and infection proceeded for 20 or 2 hours prior to recovery of bacteria. CFUs obtained were adjusted by the gram vegetation weight recovered. At 20 hours the difference in the mean log CFU/g vegetation recovered from rabbits receiving JFP42 *vs.* SK36 was not quite significant (P = 0.1036). In contrast, at two hours post-infection significantly more CFUs were recovered from rabbits receiving SK36 than rabbits receiving JFP42 (*= P= 0.0422).

Figure 39.


Log CFU/g Vegetation
10
8
6
4
2
0
*
SK36
JFP42
SK36
JFP42
20 hours
2 hours
S. sanguinis Infection

Evaluation of *S. sanguinis* sortase mutant biofilm formation

A semi-quantitative static biofilm formation assay was used to determine the contribution of *S. sanguinis* Cwa proteins to biofilm formation. Sortase mutant derivatives of *S. sanguinis* SK36; JFP42, JFP44, and SRT18, were compared to the parental strain and wild-type competitive control, JFP36. The *srpA* isogenic mutant of SK36, VT1614 was used as a negative control in this experiment (204). The *S. parasanguinis* SrpA homolog, Fap1 was previously shown to be required for wild-type biofilm formation in the presence of 1 % glucose (76). Biofilm formation was measured in the wells of PS plates following 20 hours of static culture in BM supplemented with 1 % sucrose or 1 % glucose.

When cultured in the presence of glucose, biofilm formation as quantified by CV staining was similar for SK36 and its Em^r competitive derivative, JFP36. VT1614 was most attenuated for biofilm formation ($P < 0.001$) relative to wild-type (Fig. 40A**)**. The *srtA* mutant JFP42 also exhibited a significant reduction in biofilm formation, though the amount of biofilm detected for JFP42 exceeded that for VT1614. Biofilm formation of the *srtC* STM mutant, SRT18 was also significantly decreased relative to SK36. In contrast, the *srtB* mutant JFP44 exhibited significantly greater biofilm formation than wild-type, and far exceeded biofilm mass detected for other sortase mutants tested (Fig. 40A). From this experiment it is apparent that Cwa proteins of *S. sanguinis* are required for biofilm formation in glucose supplemented media *in vitro*. This phenotype is most attributable to Cwa proteins dependent on SrtA for surface localization, of which SrpA is critical. However Cwa protein substrates of SRT18 also appear to mediate biofilm formation. In other gram-positive spp. expression of a class A sortase enhances cell surface exposure of

Cwa proteins polymerized by a class C sortase (244). If this is also the case for *S. sanguinis* the phenotype exhibited by SRT18 could be attributed to the lack of pilus assembly, whereas JFP42 would reflect the loss of the pilus from the cell surface. This is the first instance of assessment of SrtB in biofilm formation. The enhanced biofilm formation in this strain suggests that SrtB or SrtB substrates are required for maintenance of WT biofilm development.

Biofilm formation in sucrose-supplemented BM revealed carbon source-dependent intricacies of this process. In the presence of sucrose biofilm formation of VT1614, SK36, and JFP36 was similar; suggesting that under this growth condition SrpA is dispensable (Fig. 40B). Akin to this, SRT18 biofilm formation was similar to wild-type, and therefore pilus assembly may not be required for development of biofilms in sucrose. However, biofilm formation was significantly greater for JFP42 over wild-type, so SrtA-dependent Cwa proteins may impede development of the sucrose biofilm. This phenomenon was also observed for JFP44, with a significant increase in biofilm formation resulting from the loss of *srtB* expression (Fig. 40B).

In the presence of sucrose the glucosyltransferase of *S. sanguinis* generates water soluble glucans (WSG) that are thought to contribute to the synthetic exopolysaccharide glucan matrix in the complex dental plaque biofilm (7, 251). With glucose as a carbon source WSG is not produced, and instead it appears that Cwa proteins are essential. These proteins may enhance primary adhesion of streptococci to the PS surface, and stabilization of the biofilm mass. Results for glucose biofilms indicate that sortases of *S. sanguinis* have drastically different roles in this process. Additional interplay between extracellular

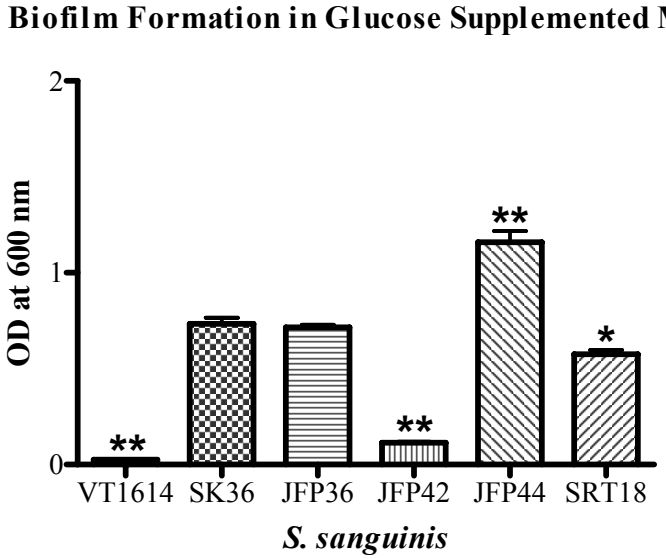
Figure 40. Sortase mutation affects static biofilm formation

A) Biofilm formation of *S. sanguinis* SK36 derivatives was determined in a semi-quantitative static assay. *S. sanguinis* strains were subcultured in BM with 1 % glucose for 20 hours before decanting and washing to remove planktonic cells. The remaining sessile cells were quantified by CV staining. Cell associated CV was suspended in 30 % acetic acid and the optical density recorded at 600 nm. Biofilm formation of JFP42 was diminished and similar to the previously described biofilm deficient mutant, VT1614. SRT18 also formed significantly less biofilm than wild-type. ***= $P < 0.001$, **= $P < 0.01$

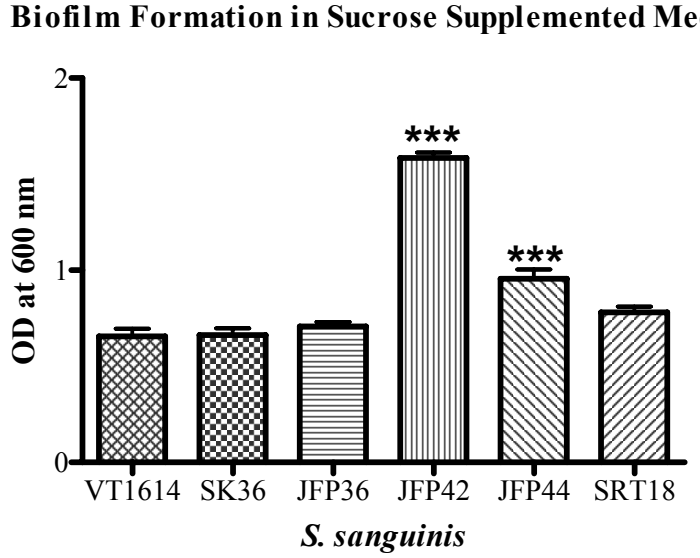
B) *S. sanguinis* strains were subcultured in BM with 1 % sucrose for determination of biofilm formation n as described above. In BM with sucrose JFP42 and JFP44 formed significantly more biofilm than the WT strain, SK36. ***= $P < 0.001$, **= $P < 0.01$

Figure 40.

A.


Biofilm Formation in Glucose Supplemented Media
OD at 600 nm
2
1
0
**
**
**
*
VT1614
SK36
JFP36
JFP42
JFP44
SRT18
S. sanguinis

B.


Biofilm Formation in Sucrose Supplemented Media
OD at 600 nm
2
1
0
***
***
VT1614
SK36
JFP36
JFP42
JFP44
SRT18
S. sanguinis

determinants of exopolysaccharide production and Cwa proteins is suggested by JFP42 biofilms grown in sucrose, when the loss of Cwa proteins caused greater biofilm formation over WT.

S. sanguinis adhesion to PS is decreased for JFP42

A short time-course study was used to determine whether reduced adherence to PS accounted for decreased biofilm formation for JFP42. Late stationary phase *S. sanguinis* strains were sub-cultured in BM with 1 % glucose for 4 hours, and adherent bacteria determined by CV staining. Adhesion to PS was unaffected in JFP36 and JFP44 relative to SK36 (Fig. 41B). Adherence to PS was slightly higher for the SrtC mutant, SRT18, which is surprising given lower biofilm formation detected for this strain at 20 hours. While adhesion to PS is unaffected in this strain, the evolving biofilm post adhesion may not be as stable due to mutation of *srtC*. Attenuated biofilm formation has been previously described for a class C sortase mutant of *E. faecalis* (121). JFP42 was significantly attenuated for adhesion to PS relative to other strains tested (Fig. 41B). This phenotype is likely responsible for the decreased biofilm formation phenotype for JFP42 cultured in glucose supplemented BM (Fig. 40A). Total growth of strains (planktonic and sessile cells) in BM with 1 % glucose was measured at 450 nm prior to removal of non-adherent cells. Growth of JFP42 was significantly less than SK36; however the difference observed suggests that decreased growth does not account for the adhesion phenotype (Fig. 41A).

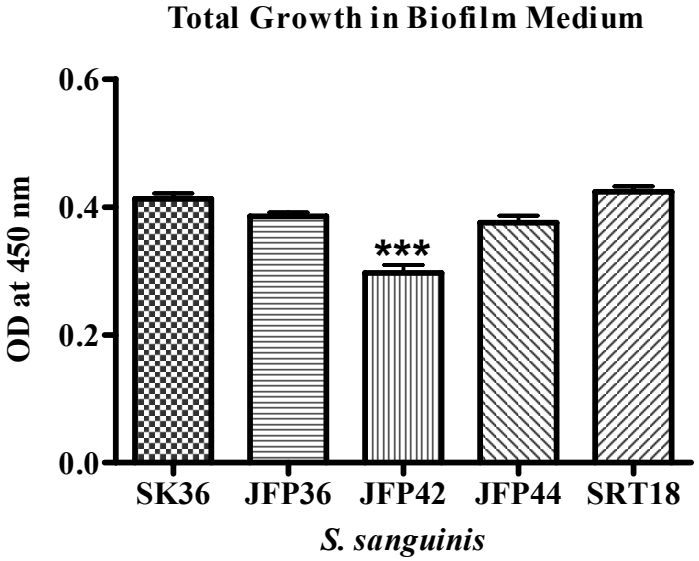
Figure 41. PS adhesion of *S. sanguinis* sortase mutants

A) *S. sanguinis* strains were compared for adhesion to PS by subculture in PS plates for four hours in BM with 1 % glucose. The total growth (both planktonic and sessile cells) in biofilm medium was measured by an optical density reading at 450 nm. Growth of JFP42 was significantly less than SK36; however JFP42 was capable of growth in BM with 1 % glucose. ***= $P < 0.001$

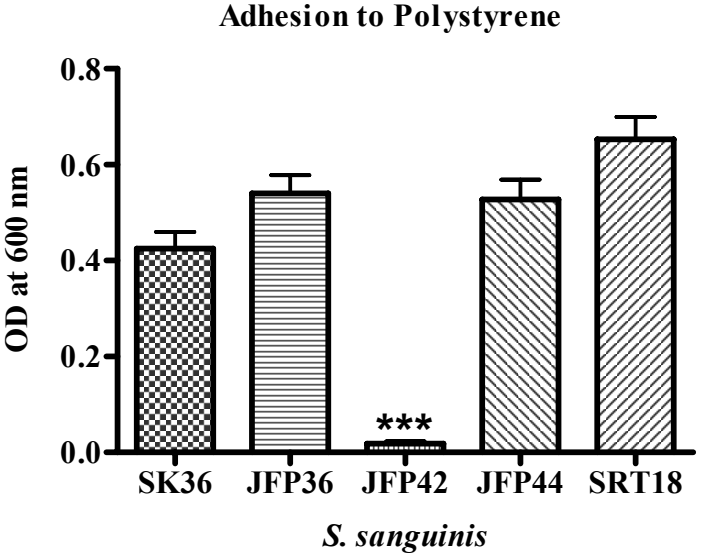
B) Following removal of non-adherent cells, remaining streptococci were determined by CV staining. CV extracted from each well was measured by recording the optical density at 600 nm. The *srtA* mutant, JFP42 exhibited a significant decrease in adherence to PS *vs.* the other strains tested. ***= $P < 0.001$

Figure 41.

A.


Total Growth in Biofilm Medium
OD at 450 nm
0.6
0.4
0.2
0.0
***
SK36
JFP36
JFP42
JFP44
SRT18
S. sanguinis

B.


Adhesion to Polystyrene
OD at 600 nm
0.8
0.6
0.4
0.2
0.0
***
SK36
JFP36
JFP42
JFP44
SRT18
S. sanguinis

Complementation of JFP42

Complementation of the *srtA* mutation was required to confirm that *in vivo* and *in vitro* phenotypes of JFP42 were attributed to loss of *srtA* expression. In *S. sanguinis* and related streptococci, the *srtA* locus is 17 bp downstream of *gyrA*, which encodes DNA gyrase. The proximity of *srtA* and *gyrA*, suggests these genes are co-transcribed. Co-transcription through the intergenic region of *gyrA* and *srtA* was confirmed by RT-PCR (Fig. 42A). *gyrA* is the first gene in this operon, as the proximal gene 5' to *gyrA* is encoded on the opposite strand from *gyrA* and *srtA*. A putative promoter of *gyrA* expression was predicted in the online *S. sanguinis* genome annotation (http://watson.vcu.edu:81/cgi-bin/gbrowse/sanguinis). This promoter element is separated from *gyrA* by a predicted ribosome binding site (RBS). In the first complementation strategy employed (depicted in Fig. 7) the predicted RBS of *srtA*, and the *srtA* coding region were fused to the predicted promoter for the *gyrA* operon by PCR amplification with a long primer including at the 5' end the promoter for *gyrA* and the sequence just upstream from *srtA* at the 3' end. This amplicon was then fused to the *kan* cassette for antibiotic selection, and 5' and 3' flanking regions of *SSA_0169* by multiple step gene splicing by overlap extension (Gene SOEing) PCR amplification. Flanking *SSA_1069* sequences permitted homologous recombination of the PCR product into the gene into which antibiotic resistance cassettes were inserted in the creation of strains JFP36, JFP56, and JFP76 (Chapter IV). Studies of JFP36, JFP56 and JFP76 show that mutagenesis of *SSA_0169* does not affect relevant phenotypes like growth *in vitro*, and competitiveness *in vivo*. Therefore phenotypes observed following recombination of the PCR amplicon into this locus would only reflect restoration of

chromosomally encoded *srtA*. By restoration of *srtA* in single copy we intended to avoid potential effects of complementation by a plasmid encoded *srtA*, including multiple copies of *srtA*, and slower growth due to plasmid replication. In the resulting strain, *S. sanguinis* JFP47, the *srtA* mutation was maintained and selected for through Sc^r, and the wild-type *srtA* was maintained by selection for Kn^r.

A CI assay was used for comparison of competitiveness of JFP47 relative to the competitive control JFP36. Three catheterized rabbits received an inoculum comprising both strains at equal optical densities. JFP47 was significantly less competitive than JFP36, with a mean CI value of 0.12 (range: 0.04 – 0.23). The similarity of the CI value for JFP47 to that obtained for JFP42 (CI of 0.16) demonstrates that competitiveness *in vivo* was not restored by insertion of *srtA* at the *SSA_0169* locus. qRT-PCR was used to determine whether the lack of complementation was attributed to lack of *srtA* expression in JFP47. The primer set used for detection of *srtA* transcript anneals to sequence deleted in creation of JFP42; therefore only transcript of the complete *srtA* gene was measured in JFP47 (Fig. 42B**)**. The *srtA* transcript abundance was compared between SK36, JFP36, JFP42 and the complementation strain JFP47. In JFP36 *srtA* transcript abundance was roughly half of that detected for SK36 (mean= 0.55) (Fig. 42C). The *srtA* transcript in JFP47 was roughly two-thirds less abundant than in SK36 (mean= 0.3) (Fig. 42C). The decreased amount of transcript detected for JFP47 suggests that expression of *srtA* is impaired. This may be attributed to duplication of the *gyrA* promoter and titration of regulatory proteins and transcription factors between the two identical promoter sequences. In addition, SrtA production cannot be assumed from detection of *srtA* transcript.

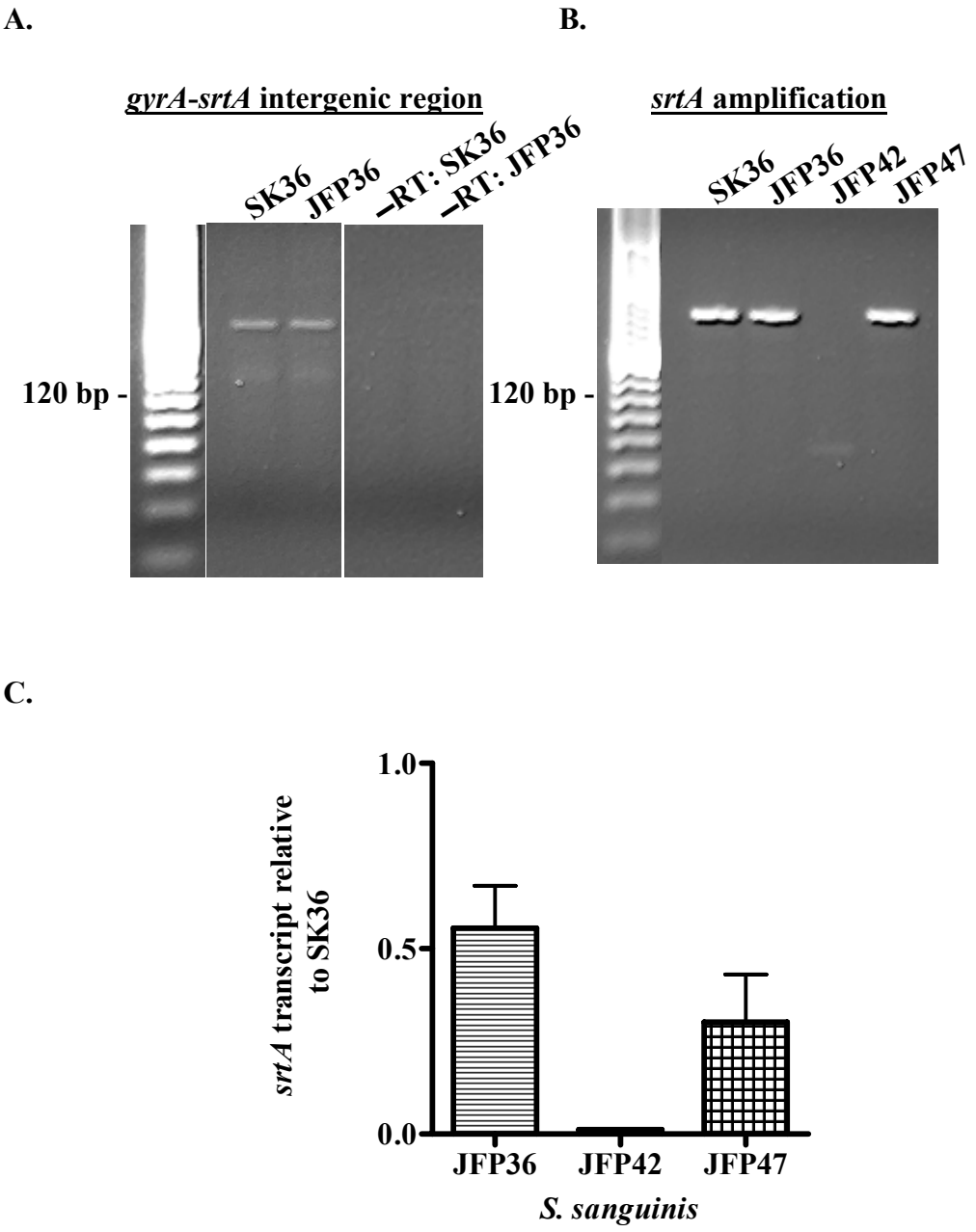
Figure 42. Analysis of *srtA* transcription

A) RT-PCR was used to determine whether *srtA* and *gyrA* are co-transcribed. Primers within the 3' end of *gyrA* and 5' end of *srtA* were used for PCR amplification of the intergenic region between the two genes, with cDNA synthesized from RNA of SK36 and JFP36 as template. As a control for DNA contamination RNA not treated with reverse transcriptase (–RT) was also used as template. A transcript was detected for both SK36 and JFP36 cDNA templates, confirming that *gyrA* and *srtA* are cotranscribed.

B) Primers specific for *srtA* were used to detect *srtA* transcript in cDNA prepared from *S. sanguinis* SK36, JFP36, JFP42, JFP47. The *srtA* transcript was missing from JFP42 (as expected), but present in cDNA synthesized from SK36, JFP36 and JFP47 RNA.

C) The level of transcription of *srtA* was determined in SK36, JFP36, JFP42 and JFP47 by qRT-PCR. Samples were normalized by the total amount of cDNA synthesized to control for differences in reverse transcription efficiency. Transcript level relative to SK36 is shown. The transcript level of *srtA* was one half or one third as abundant in JFP36, or JFP47, respectively.

Figure 42.


A.
B.
gyrA-srtA intergenic region
srtA amplification
SK36
JFP36
−RT: SK36
−RT: JFP36
120 bp -
SK36
JFP36
JFP42
JFP47
120 bp -
C.
srtA transcript relative to SK36
1.0
0.5
0.0
JFP36
JFP42
JFP47
S. sanguinis

Additional phenotypes observed for JFP42 were not restored in JFP47, providing further evidence that functional SrtA is not produced (data not shown). One alternative, albeit less attractive possibility is that mutation of *srtA* in JFP42 resulted in polar effects unrelated to the loss of *srtA* expression.

To negate potential effects of duplicating the *gyrA* promoter, an alternative strategy was used for complementation of JFP42. In *S. mutans* a putative *srtA* promoter occurring within *gyrA* was reported. In this genetic background plasmid-borne *srtA* preceded by the putative promoter sequence caused partial restoration of the *srtA* mutant phenotype of decreased biofilm formation (146). A region corresponding to the predicted promoter for *S. mutans* UA159 was identified in *S. sanguinis*, 187 bp upstream of *srtA*, within the 3' end of *gyrA*. A PCR amplicon of the putative promoter and the *srtA* coding region was cloned into the shuttle vector pVA838, resulting in pJFP50. Clones of JFP42 containing pJFP50 were screened for restoration of the *in vitro srtA* mutant phenotype, adhesion to PS. Adhesion to PS of JFP42 was not restored by pJFP50 in any clone screened, and presence of pJFP50 did not affect the capacity of SK36 to adhere to PS (data not shown). The phenotypes of JFP42 (pJFP50) and JFP47 suggested that complementation of *srtA* was impeded by unsuccessful expression of the ORF. As promoter elements in these constructs were speculative, we sought to design a different complementation construct in which *srtA* would be expressed from a previously characterized promoter. The pCM18 vector was used for this purpose. Strong expression of the green fluorescent protein encoding reporter gene, *gfp*mut3 has been described for the pCM18-encoded *L. lactis* synthetic promoter, CP25. Green fluorescence of *S. gordonii* containing pCM18 confirmed that transcriptional

and translational signals are also recognized in streptococci (90). The promoter was manipulated for expression of *srtA* of *S. sanguinis* SK36 by cloning of *srtA* between the ribosomal binding site (RBSII), and *gfp*mut3 (Fig. 9). The resulting plasmid, pJFP52 was transformed into JFP42. Five clones of *S. sanguinis* JFP42 harboring pJFP52 were screened for restoration of PS adhesion, a phenotype attributed to loss of *srtA* expression. Compared to JFP42, all five clones exhibited restoration of binding equivalent to the WT vector control strain, SK36 (pCM18) (data not shown).

The pCM18 vector was designed for stability in both gram-positive and gram-negative species. Plasmid stability is critical during *in vivo* analysis as antibiotic selection cannot be used to ensure plasmid maintenance. Stability of this complementation plasmid, pFP52 in JFP42 was determined by growth of JFP52 in the absence of antibiotic selection (BHI alone), followed by plating of cells in the presence and absence of the plasmid selective antibiotic Em. A ~8 fold decrease in plating efficiency was observed for JFP52 when plated in the presence of Em *vs*. on Sc plates alone. This difference in CFUs reflects the instability and loss of pJFP52 in the absence of antibiotic selection. Since this phenotype would negatively affect *in vivo* analysis, a final complementation strategy was employed.

In the final scheme for JFP42 complementation, the CP25 promoter and *srtA* coding region were cloned from pCM18 into the stable shuttle vector, pVA838. This strategy addressed both concerns regarding the stability of the vector used, and the capacity for *srtA* expression. The resulting plasmid, pJFP62 was transformed into JFP42, and clones screened for adherence to PS. The complementation strain JFP62 showed adherence to PS

comparable to WT and the WT vector control *vs.* the *srtA* mutant, JFP42 (Fig. 43). With adherence to PS restored, the possibility of co-integration of the plasmid was determined. Genomic DNA preparations of SK36, JFP42, both strains with the vector alone (pCM18), and the complementation strain JFP62 were compared by gel electrophoresis. When compared to the plasmid alone (pJFP62) and the vector control (pCM18) it was clear that the plasmid in JFP62 was maintained as an episomal element, and contained the *srtA* insert (data not shown).

CI analysis was used to compare the competitiveness of JFP62 or JFP42 to the Tet^r competitive control, JFP76. The mean CI value for JFP62 *vs.* JFP76 (0.53) was significantly greater than the mean CI for JFP42 *vs.* JFP76 (0.29) (Fig. 44A). Therefore competitiveness attributable to *srtA* was confirmed via complementation in the JFP42 background. However, JFP42 and JFP62 were both significantly less competitive than the control, JFP76 ($P < 0.001$).

That competitiveness of JFP62 was not completely restored may be attributed to slower growth of this strain as a result of maintenance of the episomal plasmid. *In vitro* growth studies comparing SK36, SK36 (pVA838), JFP42, JFP42 (pVA838), and JFP62 demonstrated that strains harboring pVA838 grew slower (Fig. 44B). In summary, the partial complementation of JFP42 competitiveness by pJFP62-borne *srtA* suggests this phenotype is likely attributed to *srtA*, and not a polar effect of the *srtA* mutation in JFP42.

Figure 43. Expression of *srtA* from pJFP62 restores adherence to PS

The PS adherence assay was used to select a strain complementing the *srtA* mutant phenotype. Adherence was restored to levels similar to WT by expression of *srtA* from pJFP62. The presence of the vector, pVA838 alone did not affect adherence to PS in the SK36 and JFP42 genetic backgrounds.

Figure 43.

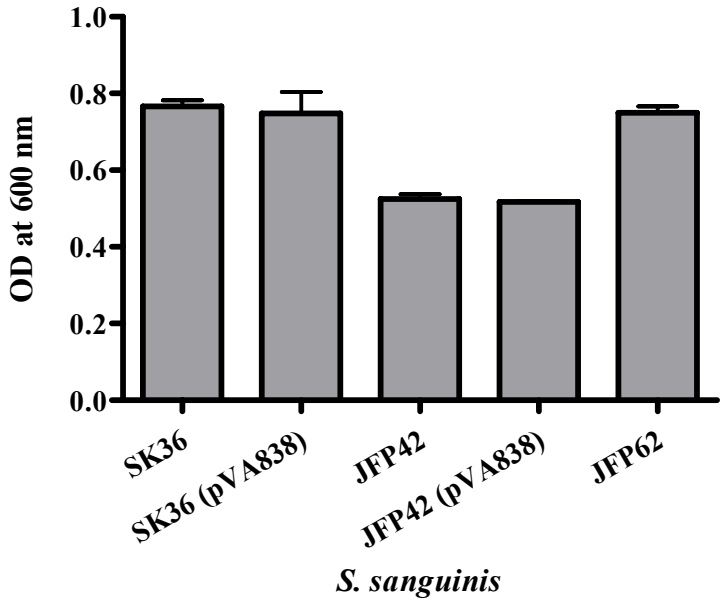

1.0
0.8
0.6
0.4
0.2
0.0
OD at 600 nm
SK36
SK36 (pVA838)
JFP42
JFP42 (pVA838)
JFP62
S. sanguinis

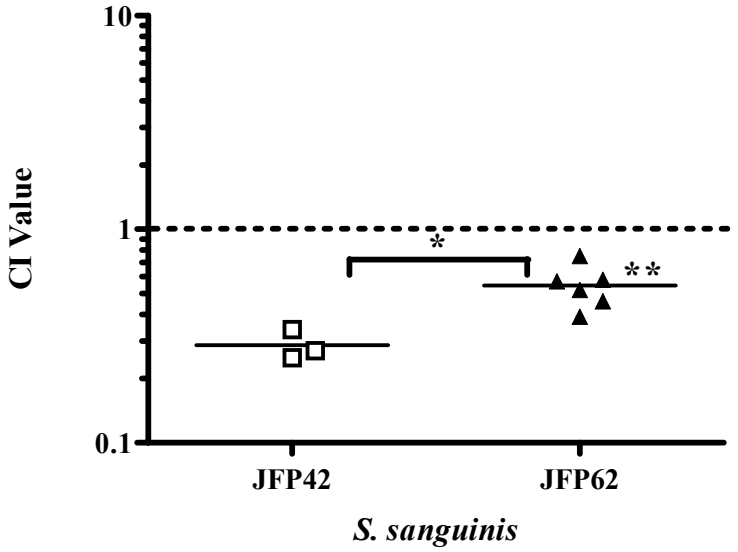
Figure 44. Competitiveness is partially restored in JFP62

A) Competitiveness of JFP42 and JFP62 was determined relative to the virulent control JFP76 (Tet^r SK36) with bacterial recovered at 20 hours post-infection. JFP62 was significantly more competitive than JFP42 (*= *P*= 0.0046), suggesting that reduced competitiveness of JFP42 is attributed to loss of *srtA* expression. However JFP62 was significantly less competitive than the control, JFP76 (**= *P*= 0.001).

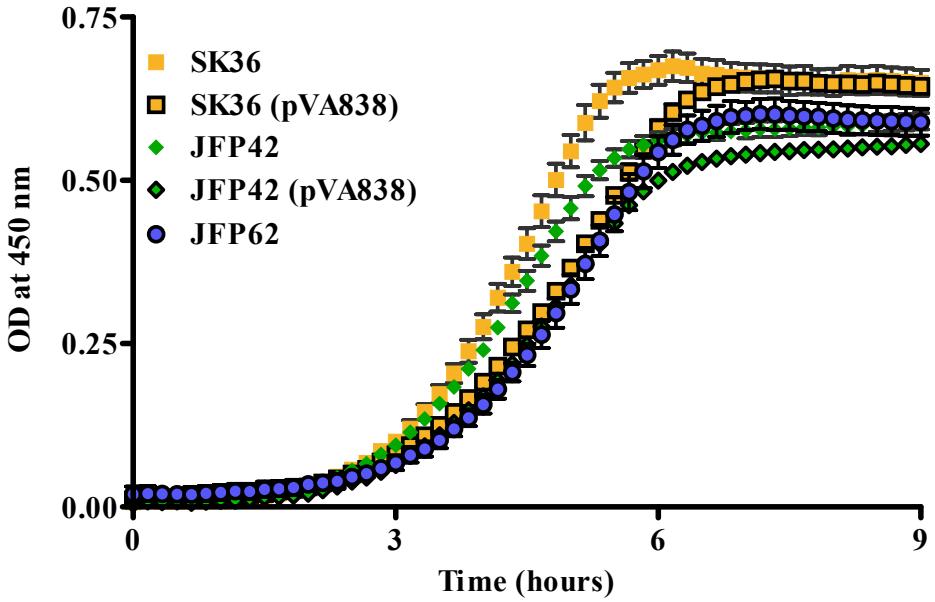
B) The lesser competitiveness of JFP62 is likely attributed to slower growth as a result of maintenance of pJFP62. JFP62 was compared to JFP42 and SK36 harboring pVA838 in an *in vitro* growth study. Strains adjusted to an equal OD were diluted 200-fold and growth monitored overnight. The growth curves of strains harboring pVA838 or its derivative, pJFP62 are prolonged, suggesting that maintenance of the plasmid results in slower growth.

Figure 44.

A.


10
1
0.1
CI Value
*
**
JFP42
JFP62
S. sanguinis

B.


0.75
0.50
0.25
0.00
OD at 450 nm
SK36
SK36 (pVA838)
JFP42
JFP42 (pVA838)
JFP62
0
3
6
9
Time (hours)

In vitro phenotypes related to SrtA

The assumption that *in vivo* and *in vitro* phenotypes of JFP42 were attributed to aberrant cell surface characteristics resulting from the loss of Cwa protein anchoring required further investigation. In past studies the approaches used to study gram-positive cell surface characteristics owing to sortases have included immunogold labeling for visualization of specific proteins by electron microscopy, catalysis of substrates by sortase-anchored enzymes, Western blotting of cell wall protein fractions obtained by digestion with enzymes including lysostaphin, and mutanolysin, as well as flow cytometry of bacterial cells (61, 121, 123, 189). These methods require knowledge of Cwa protein substrates as well as tools for detection.

To demonstrate altered cell surface localization as a result of *srtA* mutation we focused on the most well characterized *S. sanguinis* Cwa protein, SrpA. *In vitro* phenotypes attributed to this protein have been previously described, and the mechanisms of export and glycosylation of the *S. parasanguinis* homolog, Fap1 revealed. A SrpA mAb (generously provided by Dr. Hui Wu, University of Alabama, in Birmingham, AL) was used to detect cell surface exposure of this protein by whole cell ELISA, and by Western immunoblotting of *S. sanguinis* cellular fractions.

CwFrac for general detection of whole cells by ELISA

Polyclonal CwFrac antisera generated against Cwa proteins was used to detect differences in the amount of *S. sanguinis* whole cells present during the whole-cell ELISA. *S. sanguinis* JFP42 exhibits diminished adhesion to PS due to *srtA* mutation; therefore cells

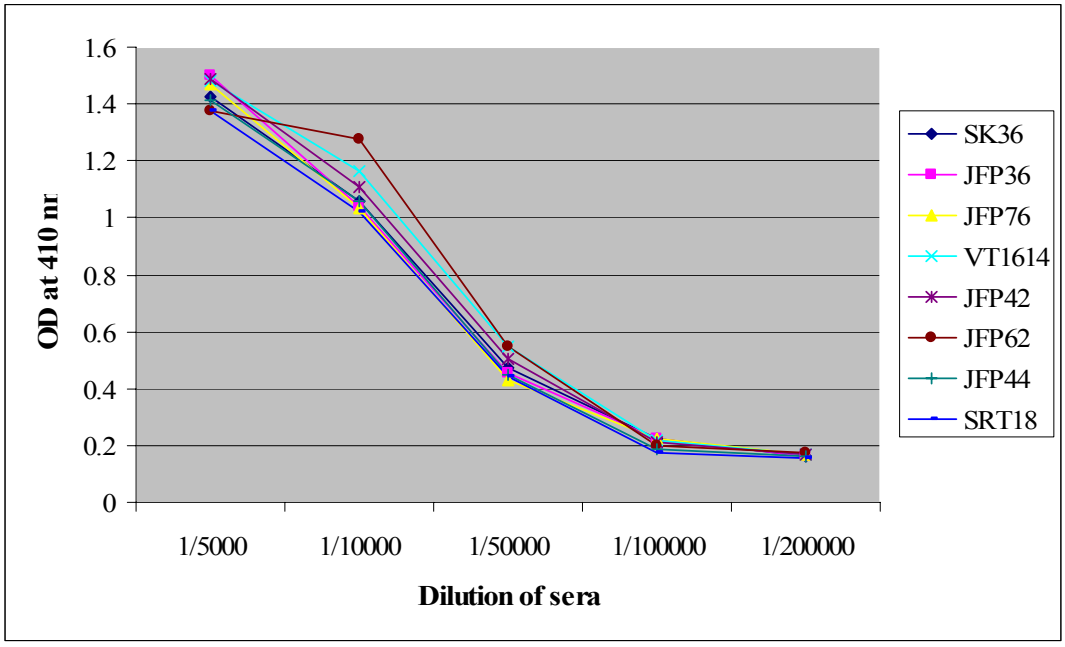
of each strain were fixed to wells by 0.25 % glutaraldehyde by a previously described method to ensure a similar cellular density for each strain (103, 189). In the first experiment reactivity to serially diluted antibody was used to determine the non-saturating dilution of antibody for use. For this purpose CwFrac antisera was serially diluted 5,000 fold, 10,000 fold, 50,000, fold 100,000 fold and 200,000 fold, and reactivity to the different dilutions of antisera determined for *S. sanguinis* strains SK36, JFP36, JFP76, JFP44, SRT18, JFP42 and VT1614. The range of reactivity for each strain over the dilutions used is shown in Figure 45. Near saturation was observed at a 5,000 fold dilution of antisera, whereas reactivity was undetectable at the 100,000 fold dilution of antisera. The linear trend between a 10,000 and 50,000 fold dilution of antisera represents the range of antibody dilution permitting specific and non-saturating detection of *S. sanguinis*. Based on these results the CwFrac antisera dilution of 1:25,000 was chosen for detection of *S. sanguinis* cells.

We then wanted to determine whether ELISA with CwFrac antisera could differentiate between differences in the number of whole cells present. Starting at the cell density used for SrpA ELISA (OD_{660} of ~0.9), cells of each strain were serially diluted two-fold across the plate, and then fixed to the wells by 0.25 % glutaraldehyde. Results showing reactivity to CwFrac (1:25,000) of serially diluted strains are shown in Figure 46A. The stepwise decrease in reactivity with increasing serial dilution proves that CwFrac can be used to detect differences in the amounts of cells present. At the cell density examined by SrpA ELISA similar reactivity to CwFrac was demonstrated for each strain tested (Fig. 46B).

Figure 45. General detection of whole cells by CwFrac

Serial dilutions of CwFrac were tested against glutaraldehyde fixed whole cells of *S. sanguinis* strains shown in a whole cell ELISA. The diminishing reactivity observed as a function of antisera concentration reflects specific association with the *S. sanguinis* cell surface.

Figure 45.


1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
OD at 410 nr
1/5000
1/10000
1/50000
1/100000
1/200000
Dilution of sera
SK36
JFP36
JFP76
VT1614
JFP42
JFP62
JFP44
SRT18

Whole cell ELISA for detection of SrpA

To demonstrate difference in cell surface SrpA localization, *S. sanguinis* strains SK36, JFP36 (Em^r SK36), JFP76 (Tet^r SK36), JFP44 (Δ*srtB*), and SRT18 (Δ*srtC*) were compared to the *srtA* mutant JFP42, and the *srpA* mutant VT1614. The presence of SrpA on whole cells was determined by an ELISA with a SrpA monoclonal antibody. Wild-type levels of cell surface SrpA are demonstrated by SK36 and the competitive controls JFP36 and JFP76 (Fig. 46C). The *srtB* and *srtC* mutants, JFP44 and SRT18 showed similar levels of SrpA to wild-type, confirming that SrpA cell surface localization is not dependent on these sortases. As expected the *srtA* mutant, JFP42 exhibited a significant reduction in cell surface exposed SrpA, that was restored in the complementation strain JFP62. Thus mutation of the putative major housekeeping sortase in *S. sanguinis*, SrtA, resulted in loss of cell surface localization of a Cwa protein substrate, SrpA.

Reactivity of JFP42 was slightly higher than the negative control strain, VT1614, deficient for SrpA production as a result of *srpA* genetic mutation (Fig. 46C). The defect in JFP42 is likely in SrtA-dependent covalent linkage of Cwa protein substrates, and not synthesis or extra-cytoplasmic export of precursor Cwa proteins. The limited amount of SrpA detected on the cell surface of JFP42 is likely residual extra-cytoplasmic, membrane associated protein and not protein covalently linked to PG. In the absence of the anchoring process, this protein is probably shed into the supernatant as previously described for other streptococcal SrtA mutants (189). This fate of SrpA in JFP42 was further investigated by immunoblotting of *S. sanguinis* cell fractions.

Figure 46. Mutation of *srtA* affects cell surface exposure of SrpA

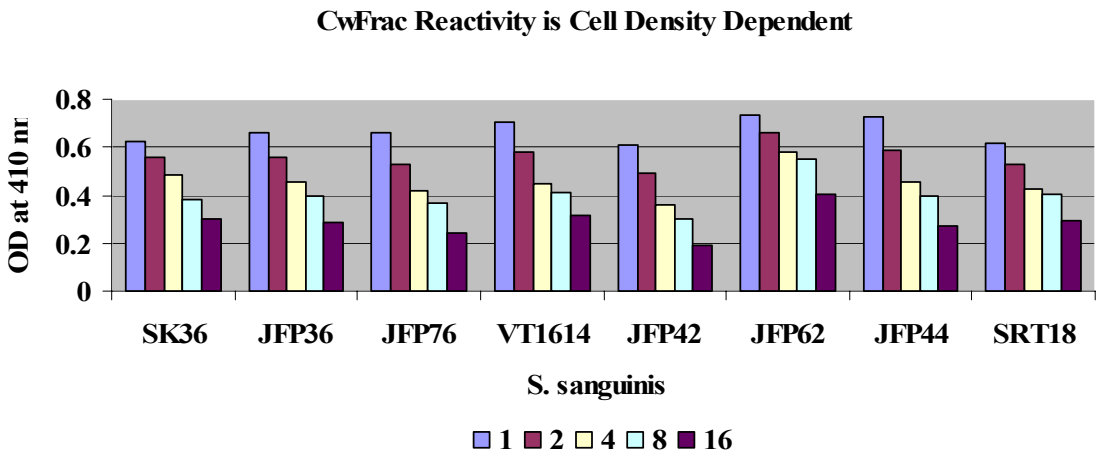
A) To demonstrate that CwFrac could be used for detection of different amounts of fixed cells in an ELISA, serially diluted *S. sanguinis* strains were detected with CwFrac antisera (1:25,000). Undiluted cells at an optical density (OD) at 600 nm of ~ 0.9 (indicated by 1), were diluted 2-fold, 4-fold, 8-fold, and 16-fold to establish the reactivity as the number of cells present decreased. A stepwise decrease in reactivity was observed with each dilution for all strains tested.

B) The CwFrac reactivity for the dilution of cells used for detection of SrpA (undiluted cells at an OD at 660 nm of ~0.9) is shown.

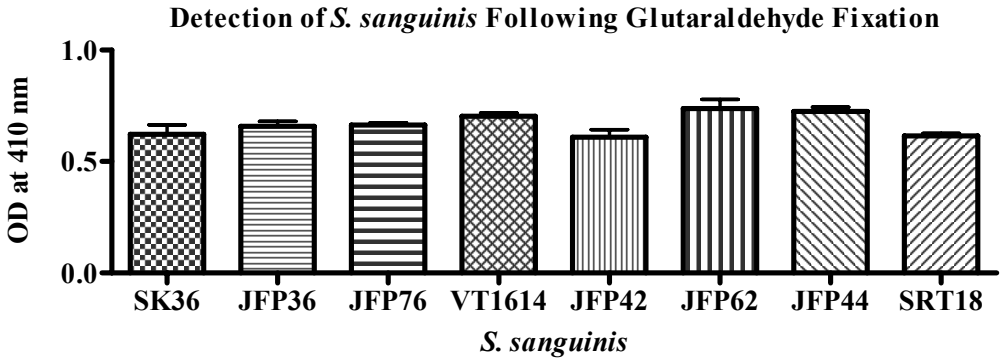
C) The reactivity of *S. sanguinis* whole, fixed cells to the α-SrpA mAb is shown. Mutation of *srtA* resulted in a significant decrease in surface exposed SrpA, comparable to the *srpA* mutant, VT1614. Expression of *srtA* from pJFP62 in the JFP42 background resulted in a restoration of wild-type levels of surface exposed SrpA. The amount of SrpA detected for in JFP44 and SRT18 suggests that cell surface exposure of SrpA is entirely dependent on SrtA. ***= $P < 0.001$

Figure 46.

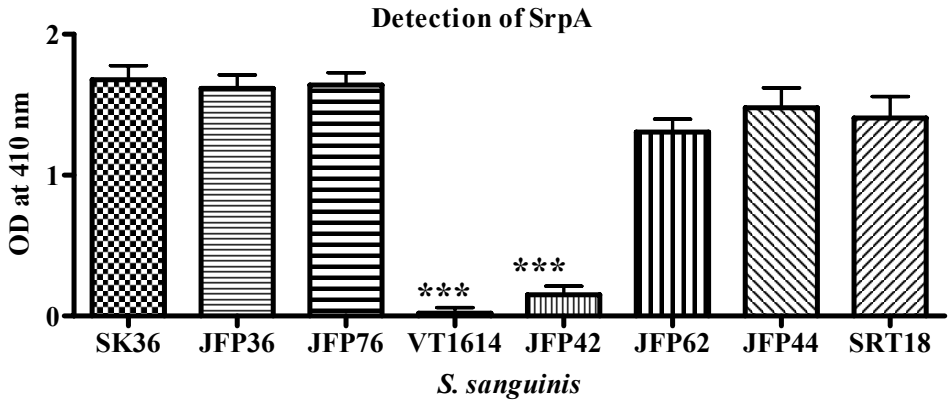
A.



B.



C.



SrpA immunoblot of *S. sanguinis* cellular fractions

To demonstrate the localization of SrpA in JFP42, *S. sanguinis* stationary phase cultures were used for preparation of supernatant containing secreted proteins, mutanolysin extracted cell wall-associated proteins, or proteins interior to the cell wall (protoplast proteins). Immunoblotting of the resulting protein fractions confirmed the presence of SrpA in the protoplast fraction, cell wall fraction and supernatant fraction of SK36 (Fig. 47A). The immunoreactive bands in Figure 47 are typical of this glycosylated protein, and similar to those observed in characterization of orthologous proteins in *S. parasanguinis* and *S. gordonii* (240, 247, 284). The *srpA* gene is predicted to encode a protein of ~150 kDa; however the carbohydrate moieties associated with this protein result in diffuse mobility by SDS-PAGE. SrpA was exclusively detected in the supernatant fraction of JFP42, suggesting that mutation of *srtA* results in the shedding of SrpA from the cell surface (Fig. 47A). This phenotype was partially complemented in JFP62, which exhibited enhanced cell wall association of SrpA.

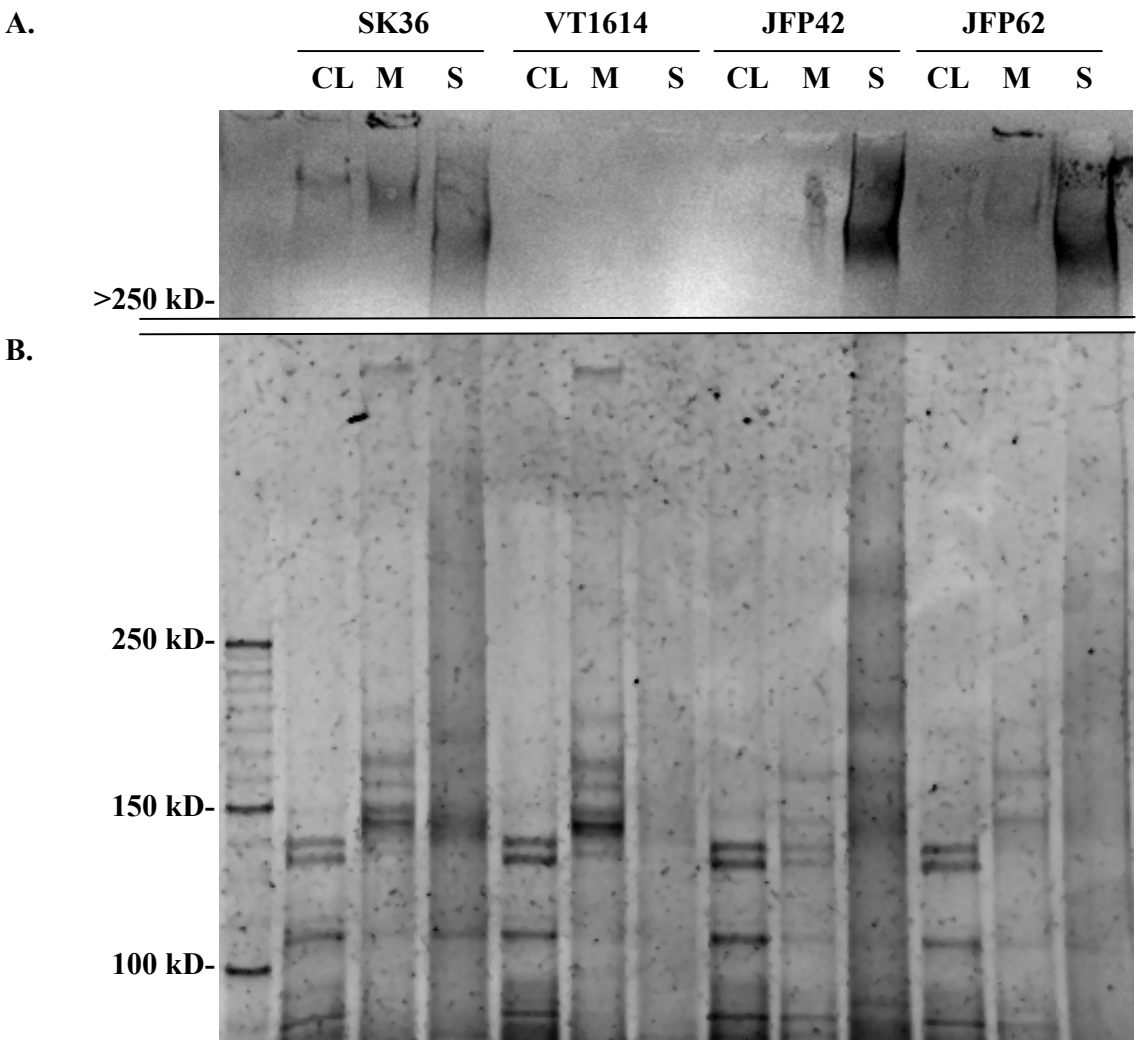
n*-Hexadecane phase partitioning of *S. sanguinis

In addition to the specific detection of a SrtA substrate described above, we also wanted to investigate a general cell surface phenotypes associated with expression of SrtA. One general property previously ascribed to Cwa proteins is hydrophobicity of the cell surface. In *S. gordonii* the property is partially conferred by CshA and CshB. This property is thought to be important for colonization of the tooth surface and maintenance as a member of the oral flora (278). Cell surface hydrophobicity of *S. sanguinis* strains was

Figure 47. Mutation of *srtA* causes extracellular accumulation of SrpA

Stationary phase *S. sanguinis* cells were fractionated for detection of SrpA by immunoblotting. Supernatant associated proteins were precipitated with TCA (S), cell wall-associated proteins were extracted by mutanolysin treatment (M), and protoplast associated proteins were prepared by lytic methods (CL). In panel A the SrpA immunoblotting of the three protein fractions for *S. sanguinis* SK36, VT1614, JFP42, and JFP62 is shown. Mutation of *srtA* in JFP42 resulted in accumulation of SrpA in the supernatant fraction. In panel B the general protein stain SYPRO Ruby was used for observance of dominant protein bands in each fraction. SrpA was not readily detected by SYPRO Ruby, putatively attributed to the extensive glycosylation of this protein.

Figure 47.

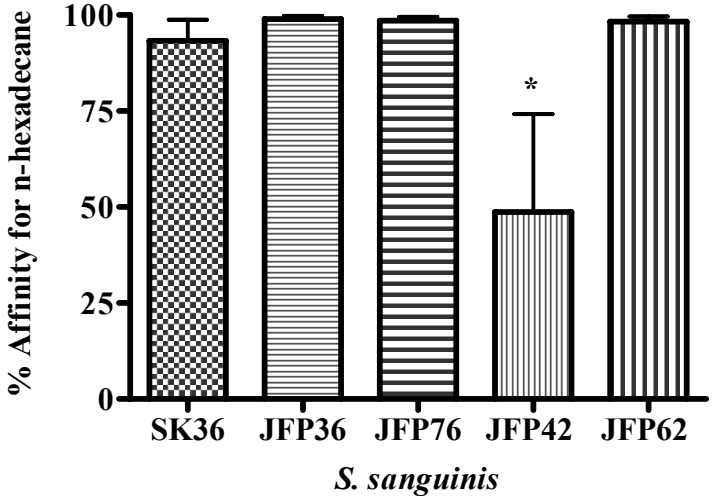

A.
SK36
VT1614
JFP42
JFP62
CL M S CL M S CL M S CL M S
>250 kD-
B.
250 kD-
150 kD-
100 kD-

determined by *n*-Hexadecane phase portioning. The percent affinity for hexadecane of each strain tested is shown in Figure 48. *S. sanguinis* SK36, competitive controls JFP36 and JFP76, sortase mutants JFP44 and SRT18, and the SrtA complemented strain JFP62 all exhibited similar, high affinities for hexadecane indicative of a strongly hydrophobic cell surface. Mutation of *srtA* resulted in significantly decreased affinity for hexadecane for JFP42. The shift in this cell surface property suggests deletion of *srtA* results in the loss of cell wall incorporated Cwa proteins conferring cell surface hydrophobicity. In a previous study the repeated *in vitro* passage of *S. sanguinis* oral isolates resulted in decreased cell surface hydrophobicity, and reduced adherence to hydroxyapatite (278). Therefore the loss of surface anchored Cwa proteins and the general shift of the cell surface of JFP42 to more hydrophilic state may account in part for reduced *in vitro* adherence phenotypes described above.

Figure 48. *n*-Hexadecane phase partitioning of *S. sanguinis* strains

Early stationary phase cells were vortexed with *n*-hexadecane (A0) for 1 minute, or left untreated (A), and then incubated at room temperature for 25 minutes. The optical density of the aqueous phase was then used to determine the percent affinity for hexadecane by the equation: 100 x [1 – (A0/A)]. The percent affinity for hexadecane of JFP42 was significantly reduced (*= $P < 0.05$), relative to JFP36, JFP76, JFP62, JFP44 and SRT18.

Figure 48.


% Affinity for n-hexadecane
100
75
50
25
0
*
SK36
JFP36
JFP76
JFP42
JFP62
S. sanguinis

Discussion

From the first identification of the cell wall-associated protein class of surface proteins it was recognized these proteins act as mediators of gram-positive disease phenotypes (227). The wealth of information currently available points to diverse functions attributed to Cwa proteins, including cell envelope maintenance, essential cofactor acquisition, and adhesion (157, 184, 222). We sought to determine the role of *S. sanguinis* Cwa proteins in IE through a comprehensive analysis of predicted Cwa proteins and sortases of *S. sanguinis* SK36. The studies described here were the first to test the biological relevance of *in vitro* phenotypes previously attributed to *S. sanguinis* Cwa proteins that were thought to be important in development of IE.

In total 33 Cwa proteins were predicted for *S. sanguinis* SK36. The repertoire of predicted Cwa proteins exceeds that of other gram-positive oral streptococci with the same predictive patterns applied. *S. sanguinis* is first and foremost a colonizer of the oral cavity, and member of the oral flora. The genome of *S. sanguinis* SK36, an oral isolate, was surely shaped by selective pressures encountered in this niche. Many Cwa proteins identified in *S. sanguinis* share homology with proteins described for closely related oral streptococci. However the identification of a small number of unique Cwa proteins for *S. sanguinis* suggests that horizontal gene transfer and genetic transformability of this strain further shaped the Cwa proteins encoded. *S. sanguinis* is a prominent streptococcus in dental

plaque (1), and a frequent agent of streptococcal IE (57, 67, 181). The abundance of Cwa proteins predicted for this species may confer a selective advantage to *S. sanguinis* for these phenotypes. Factors mediating *S. sanguinis* colonization of the oral cavity may serve dual functions to promote IE as well. For example, multiple previously described Cwa proteins have been shown to adhere to salivary agglutinin, as well as extracellular matrix proteins (110, 152, 241). In the case of SrpA, and its homolog in *S. gordonii*, Hsa, the mode of interaction for seemingly distinct phenotypes is actually quite similar. The recognition of sialylated glycoconjugates by Hsa has been shown to mediate association with salivary proteins MG2 and SAG, binding to the sialylated GP1bα platelet receptor, and adherence to the α2, 3 linked sialic acid of glycophorin A, an erythrocyte membrane protein (249, 250, 286). Recognition of sialylated glycoproteins by Hsa provides the basis for explanation of how this oral microbe is well suited for development of a cardiac disease.

To elucidate the role of *S. sanguinis* Cwa proteins and sortases in early IE STM was applied to 31 Cwa protein and 3 sortase ORFs predicted by bioinformatics. The resulting pool was screened in the rabbit model of IE. Replicate, independent experiments confirmed that of the 31 identified Cwa protein ORFs, 28 were not required for the development of IE. CI assay of the remaining 3 mutants, CWA12, CWA15 and CWA30 revealed that no single Cwa protein was required for the onset of disease. Rather than refuting *in vitro* phenotypes of *S. sanguinis* Cwa protein mutants considered relevant for disease, reduced competitiveness of the *srtA* mutant indicates *S. sanguinis* Cwa proteins contribute to IE, yet these proteins may perform redundant functions. Previous studies of *S.*

sanguinis platelet interaction identified 3 different Cwa protein mutants attenuated for this phenotype (97, 98). Simultaneous expression of these proteins during IE might negate the mutation of a single ORF, as additional proteins provide the same function. Perhaps the most striking evidence of functional redundancy comes from real-time PCR analysis of Cwa protein transcript levels in an isogenic mutant of *S. gordonii* DL1 lacking expression of antigen I/II homologs. Herzberg and colleagues found that with the loss of SspA and SspB expression, transcript abundance of other cell surface adhesins increased (292). Therefore an additional, unaccounted cellular response may be induced in mutants under selective conditions encountered *in vivo* and *in vitro* for compensation of the introduced mutation.

MudPIT analysis of *S. sanguinis* cell wall and protoplast fractions identified several Cwa proteins with similar putative functions as being simultaneously abundant. Among the proteins detected were three Cwa proteins containing the A domain of Cna (pfam05737). In total 5 putative Cwa proteins of *S. sanguinis* were predicted to include this collagen adhesion domain. Two additional Cwa proteins, SspC and SspD may share this capacity as collagen adhesion has been described for *S. gordonii* antigen I/II homologs (152). Therefore it is possible that *S. sanguinis* encodes 7 proteins promoting collagen binding. This arsenal of putative collagen adhesins suggests that these determinants are very strongly selected in the oral niche. A more detailed analysis of the expression pattern and collagen adhesion potential of predicted adhesins is required to determine whether these proteins are functionally redundant or act cumulatively. The functional redundancy hypothesis does provide further explanation of previous findings in a random STM

analysis of *S. sanguinis* SK36, in which no Cwa protein was identified as a virulence factor (198).

The analysis of an *S. sanguinis srtB* mutant, JFP44, and *srtC* mutant, SRT18 provided additional insight to the unique role of *S. sanguinis* sortases during early IE. In *E. faecalis*, an unrelated common causative agent of IE, expression of SrtC was determined to be required for competitiveness in development of IE (183). Analysis of the *S. pneumoniae* pilus locus has also indicated that pilus polymerization is required for virulence *in vivo* (13). The arrangement and constituents of the *srtC* locus in *S. sanguinis* suggests this protein also acts in polymerization of Cwa protein subunits to create a pilus structure. However STM screening of the *srtC* mutant SRT18 revealed that expression of SrtC is not required for competitiveness *in vivo*. If SrtC of *S. sanguinis* does catalyze pilus assembly, then this function is not required for the onset of disease. Putative pilin subunit mutants CWA6, CWA7 and CWA8 also maintained competitiveness *in vivo*. However the effect of mutation on cell surface structures was not investigated. In *C. diphtheriae*, *S. pyogenes* and *S. pneumoniae* a dominant pilus subunit that composes the pilus shaft has been identified which contains conserved motifs directing pilus polymerization (pilin motif) and association with accessory pilus proteins (E box) (13, 175, 256). All putative pilus proteins identified for *S. sanguinis* were predicted to include the E box and some variation of the pilin motif. It is possible that these proteins are interchangeable in pilus assembly, or compose distinct pilus structures. The presence and composition of a pilus on the surface of *S. sanguinis* requires further investigation.

Expression of *srtB* was determined to be dispensable for development of IE by CI analysis of JFP44. SrtB is conserved in *S. gordonii*; however the Cwa protein substrate(s) for this sortase have yet to be identified. The work conducted in this chapter was restricted to putative Cwa proteins containing the LPXTG pattern. If SrtB of *S. sanguinis* is similar in function to characterized SrtB proteins of other gram-positive spp., it may anchor proteins with a sorting pattern deviating from LPXTG. The potential substrate PcsB, containing the C-terminal pattern, QPEVG and an N-terminal signal peptide (as assessed by SignalP), is encoded upstream of *srtB*. If this ORF is surface anchored by SrtB, the experiments described provide evidence that it is not required for establishment of disease.

As the most pronounced *in vivo* phenotypes were observed for the *srtA* mutant, this strain as well as *srtB* and *srtC* mutants were used for further characterization of cell surface phenotypes. Cell surface hydrophobicity was significantly decreased following mutation of *srtA*. Bacterial cell surface hydrophobicity is commonly used as an indicator for the propensity of bacteria to interact with the surrounding environment. In fact, cell surface hydrophobicity is correlated with co-aggregation between bacterial cells, adherence to hydrophobic surfaces, and adherence to fibrin *in vitro* (101, 168). *S. sanguinis* binding to a highly hydrophobic polystyrene surface was also decreased by mutation of *srtA*. These findings suggest that expression of SrtA and potential Cwa protein substrates of SrtA confer this cell surface property. Cell surface expression of the Cwa protein SrpA was also tested in isogenic mutants of *S. sanguinis* SK36. SrpA was not cell surface exposed following mutation of *srtA*, the putative major sortase of *S. sanguinis* SK36. The defect in Cwa protein substrate surface anchoring coupled with a global shift to a more hydrophilic

cell surface following deletion of *srtA* suggests that reduced competitiveness observed for JFP42 may be attributed to the loss of more than one Cwa protein virulence determinant from the cell surface.

The studies described here established that SrtA is required for full competitiveness in development of IE. This is in keeping with previously described analyses of SrtA homologs in other gram-positive pathogens. In fact, the clearest evidence of Cwa protein contribution to disease has been gleaned from *in vivo* characterization of sortase mutants (113, 135, 145, 200). The 10-fold reduction in competitiveness for JFP42 and persistence of this strain in individual inoculum assays indicates that *S. sanguinis* IE is a process mediated by multiple factors additional to Cwa protein and sortase expression. This is in contrast to a lipoprotein mutant identified for *S. sanguinis* which exhibits a 3,000-fold reduction in competitiveness *in vivo* (Dr. Sankar Das, personal communication).

The observation that SrtA expression is required for normal colonization phenotypes does introduce a possible target for preventative therapies. A combination multi-gene DNA vaccine combining plasmid-encoded clumping factor A (ClfA), fibronectin binding protein A (FbpA), and SrtA was demonstrated to elicit a protective response against *S. aureus* and increased survival in a murine infection model (78). This presents a possibility for advancement of preventative measures against *S. sanguinis* IE.

GENERAL DISCUSSION

The incidence of infective endocarditis has remained unchanged over the past two decades, despite advances in diagnostics methods, antimicrobial and surgical therapies, and complication management (176). The undeterred rate of infection is in part due to changing risk factors for this disease. In the pre-antibiotic era the predominant risk factor for IE was chronic rheumatic heart disease, which is now rare in industrialized countries (171). This risk group has been replaced with persons vulnerable due to high-risk behaviors, or more aged populations prone to pre-disposing cardiac risk factors. The majority of IE cases (~70 % of NVE episodes) are attributed to staphylococci and streptococci. As for the new risk groups, intravenous drug users are most commonly afflicted with *S. aureus* IE, with skin flora as the major source. The incidence of *Streptococcus bovis* as a causative agent of disease increases proportionally with age–an observation potentially attributed to a greater rate of digestive tumors in the elderly (176). These emerging trends are reflective of changing population dynamics in developed countries and imply that IE will remain a health-care concern in the future.

Viridians streptococci are the most common causative agents of streptococcal IE (at least 50 % of community acquired NVE cases not associated with drug use; 282) and multiple reports implicate *S. sanguinis* as the most common causative agent of viridans streptococcal IE (57, 67, 216, 272). American Heart Association guidelines for patients at

risk for IE undergoing dental procedures have shifted drastically in the past 60 years in concert with the recognition that most cases of IE cases are not attributed to dental treatment, but are more likely caused by daily activities. The latest guidelines noted that both dental treatment and daily tasks induce similar densities of bacteria in the blood, yet routine activities are much more likely to cause bacteremic events resulting in IE. The authors suggested that the risk of bacteremia associated with IE would be better prevented by improved oral health care in at-risk patients with underlying heart conditions. In patients most at risk, antibiotic prophylaxis may still be recommended, with the acknowledgement that only a small number of cases are preventable by this strategy (281).

Antimicrobial therapy for treatment of IE is equally complex, with extensive treatment regimens suggested. One complication of this therapy is penetration of antimicrobial compounds within the fibrin-platelet meshwork of the vegetation, which has been demonstrated as refractory to antimicrobials in the past (51). This observation, in addition to the risk of increased antimicrobial resistance among bacterial agents of IE highlights the critical importance of development of prophylactic measures against IE for at-risk groups. Thus far the most promising vaccine candidate for viridans group streptococci is the lipoprotein FimA of *S. parasanguinis*. Expression of FimA was required for development of *S. parasanguinis* IE (34). Administration of a FimA immunogen was found to be protective against *S. parasanguinis* IE, as well as IE caused by additional viridans streptococcal species predicted to encode FimA homologs including *S. mitis*, *S. mutans*, and *S. salivarius* (127, 271). Further molecular characterization of *S. sanguinis* IE determinants is necessary for advancement of our understanding of this disease. The aims

of experiments described here included the characterization of a previously described IE virulence determinant, the development of genetic tools facilitating the screening and identification of virulence factors for IE, and identification of putative adhesins required for development of IE through a comprehensive analysis of *S. sanguinis* Cwa proteins *in vivo*.

In Chapter III the requirement of NrdD for anaerobic growth of *S. sanguinis* SK36 was explored. As documented for *L. lactis*, *S. aureus* and *E. coli*, evaluation of NrdD in *S. sanguinis* was complicated by residual function of the Class Ib RNR under microaerophilic conditions. Instead of supplementing growth media with class I RNR inhibitors, fully reduced soft-agar was used for evaluation of *nrdD* mutant strains. Transposon insertion and in-frame deletion mutagenesis of *nrdD* both resulted in arrested anaerobic growth, confirming the requirement for NrdD expression in the absence of oxygen. In an STM analysis of genes required for *Proteus mirabilis* urinary tract infection an attenuated *nrdD* mutant was also identified. The authors suggested that oxygen-limiting conditions may develop during urinary tract infection as a result of high bacterial cell density (31). This may also be the case in IE, in which bacteria have been documented to grow within the vegetation to reach cell densities of 10^8 to 10^{11} CFU per gram of vegetation (64). The bacteria likely proliferate, protected from circulating host defense mechanisms by layered fibrin-platelet matrices. The aerobic metabolism of the streptococci may generate a progressively reduced condition that can not be relieved by oxygenated blood surrounding the vegetation. This study was not the first to propose anaerobiosis of the infective vegetation; a previous screen of *S. aureus* STM mutants in the rabbit model of IE

demonstrated that anaerobic/microaerophilic mutants were attenuated *in vivo* (50). However, staphylococcal and streptococcal pathologies in IE are quite different in that staphylococcal infection is typically more acute and invasive, whereas streptococcal IE infection often presents as sub-acute, with some features of chronic disease. In addition, the attenuated staphylococcal mutants were not characterized, and therefore the nature of mutations conferring anaerobic sensitivity is not known. *In vitro* characterization of the avirulent *nrdD* mutant (198) provided further insights to streptococcal factors required for persistence *in vivo*. The extreme measures required for demonstration of anaerobic sensitivity of *S. sanguinis nrdD* mutants suggest that anaerobiosis occurs within the infective vegetation–and the time course of *in vivo* infection confirms that this occurs within 20 hours of the onset of bacteremia. Although oxygen levels within the infective vegetation have not been measured, the findings described suggest a restriction of oxygen availability at the site of infection.

The publication of the SK36 *S. sanguinis* genome sequence in 2007 realized the possibility for greater characterization of this species, with this strain serving as the model. Of particular interest to the medical community is identification of *S. sanguinis* virulence determinants that may serve as targets for prophylaxis of IE. In keeping with this goal, in Chapter IV a locus was identified that may be used for genetic manipulation without directly affecting important cellular phenotypes including *in vitro* growth, genetic competence, biofilm formation, and competitiveness in the rabbit model of IE. The plasmid repertoire developed for this purpose, with three different antibiotic resistance determinants, could be expanded to include additional selective markers, or for ectopic

expression of native genes to allow for complementation of mutations elsewhere on the chromosome. Two of the strains developed in Chapter IV, JFP36 and JFP76 were used as controls for competitive index analyses in Chapter V to establish the competitiveness of Cwa protein mutant or sortase mutant strains–with successful discrimination between differing levels of competitiveness in mutants tested. Therefore a genetic tool has been developed which will facilitate the future screening and identification of *S. sanguinis* virulence determinants.

In Chapter V the contribution of *S. sanguinis* Cwa proteins to early experimental IE was investigated. Cwa proteins are attractive vaccine targets as they are surface exposed, and have been demonstrated to interact with host cell surface receptors and extracellular matrix proteins. *S. sanguinis* SK36 putatively encodes 33 Cwa proteins; mutants defective for production of 31 of these proteins were included in a comprehensive STM screen. A *srpA* mutant (204), which lacks the 32nd Cwa protein, was also tested in the rabbit model of IE by competitive index analysis. The *in vivo* experiments described demonstrated that no single Cwa protein was required for development of IE–a finding that is counterintuitive given past characterization of this protein family in animal infection models. Among the predicted Cwa proteins encoded by *S. sanguinis* it is likely that SrtA is responsible for cell surface localization of the majority. Competitive index analyses of a *srtA* mutant determined expression of *srtA* to be required for full competitiveness *in vivo*, with Cwa proteins dependent on SrtA for surface-anchoring putatively mediating initial contacts for colonization of the sterile thrombus. Homology searches of predicted *S. sanguinis* Cwa proteins indicate that several proteins may be grouped by functional similarity, and may be

redundant *in vivo*. This possibility is most clearly evidenced by the *srtA* mutant strain, which would likely be defective for surface localization of the majority of Cwa proteins (including groups of putatively functionally redundant proteins). The *srtA* mutant does provide an opportunity for further characterization of the general contribution of *S. sanguinis* SrtA-dependent Cwa protein substrates to IE. Of particular interest is characterization of streptococcal-host factor interactions, including the assessment of platelet-interaction, adherence to fibrin-platelet matrices, and resistance to phagocytosis.

Expression of SrtA has also been linked to infectivity in *S. aureus* IE, with lesser *srtA* mutant bacterial recovery following 2 days of infection (276). In contrast to *S. aureus*, *srtA* expression in *S. sanguinis* may be important for colonization of the vegetation, but not in replication following colonization, in that bacterial numbers recovered from rabbits infected at 20 hours were not significantly different between the *srtA* mutant and a wild-type strain. These findings imply that SrtA, in addition to unidentified SrtA-dependent Cwa proteins mediating vegetation colonization may serve as candidates for prevention of IE onset.

Gram-positive Cwa proteins have been tested as vaccine candidates in the past, with some success (35, 78, 178, 225). However several complications arise in developing a Cwa protein based preventative strategy for *S. sanguinis* IE. *S. sanguinis* is benign in the oral cavity and may actually be beneficial in this niche through competitive inhibition of oral pathogens (132). An optimal vaccination strategy would balance the advantageous contribution of *S. sanguinis* in the oral flora while reducing the propensity for IE by targeting surface proteins specifically expressed in transient bacteremia. Thus far *S.*

sanguinis proteins uniquely expressed during bacteremia have not been described, and it is possible that few candidates will arise as *S. sanguinis* is an opportunistic agent of IE that evolved primarily for maintenance in the oral cavity. A future direction for prophylaxis research should include expression analysis of *S. sanguinis* genes during bacteremia for identification of proteins that are enhanced under these conditions relative to growth in an oral biofilm. Such a study would require the development of an adequate control condition for relative expression analysis; however it is feasible given the availability of a genomic microarray for *S. sanguinis* SK36.

Additional considerations for prophylactic strategies include identification of vaccine candidates shared among *S. sanguinis* endocarditis isolates to identify those most common in nature, and the use of proteins conserved among common causative agents of viridans IE. Acknowledgement of the natural complexity of resident flora is required for production of an efficacious preventative strategy. In addition, the protection afforded by the previously identified immunogen FimA warrants further investigation (127, 271). Of particular interest is the immunology governing the inhibition of IE by vaccination with FimA. Identification of specific immunological factors that potentiate this response may direct the conception and optimization of future preventative strategies.

Experimental analyses of antibody penetration within the vegetation have shown the fibrin-platelet matrices to form a sheath that limits access of antibody to bacteria within the vegetation (166). Due to the complications with identification of an optimal IE-specific virulence candidate described above, an attractive alternative may include the development of vegetation diffusible small-molecule inhibitors against prokaryotic genes known to be

expressed during IE. One potential target includes NrdD, which has no eukaryotic counterpart. The time course of anaerobiosis in IE would require further monitoring to assess the feasibility of this. If anaerobiosis within the vegetation occurs within hours of colonization, then progressive vegetation enlargement owing to reseeding of bacteria at the site of infection, or outgrowth of bacteria within the vegetation may be avoided. However if anaerobiosis occurs later in infection, after the vegetation has achieved a mass resulting in frequent septic embolization, then efficacy would be diminished.

Conclusions derived from experiments performed here, in combination with the previous random STM screen for *S. sanguinis* virulence determinants present a model of streptococcal transition between initial colonization of the vegetation and growth following colonization. During transient bacteremia the streptococci express and utilize surface-exposed proteins, including SrtA-dependent Cwa proteins, for colonization of the site of endothelial damage in a process that may be dependent on a sterile thrombus. Surface localization of these proteins does not appear to be important for growth following adherence, at 20 hours post colonization. Instead genes encoding for cytoplasmic proteins with housekeeping functions like bacterial cell wall synthesis, generation of nucleic acids and amino acids, and anaerobic condition-specific genes are required for persistence (198). Expression analyses of streptococcal genes at different times post-infection may aid in understanding of this complex and multi-factorial disease, as well as identification of future therapeutic targets.

Literature Cited

1. **Aas, J. A., B. J. Paster, L. N. Stokes, I. Olsen, and F. E. Dewhirst.** 2005. Defining the normal bacterial flora of the oral cavity. J. Clin. Microbiol. **43:**5721-5732.
2. **Abrams, D., G. Derrick, D. Penny, E. Shinebourne, and A. Redington.** 2001. Cardiac complications in children following infection with varicella zoster virus. Cardiol. Young **11:**647-652.
3. **Ajdic, D., W. M. McShan, R. E. McLaughlin, G. Savic, J. Chang, M. B. Carson, C. Primeaux, R. Tian, S. Kenton, H. Jia, S. Lin, Y. Qian, S. Li, H. Zhu, F. Najar, H. Lai, J. White, B. A. Roe, and J. J. Ferretti.** 2002. Genome sequence of *Streptococcus mutans* UA159, a cariogenic dental pathogen. Proc. Natl. Acad. Sci. U.S.A. **99:**14434-14439.
4. **Alter, P., J. Hoeschen, M. Ritter, and B. Maisch.** 2002. Usefulness of cytokines interleukin-6 and interleukin-2R concentrations in diagnosing active infective endocarditis involving native valves. Am. Journal Cardiol. **89:**1400-1404.
5. **Andersson, J., M. Westman, A. Hofer, and B. M. Sjoberg.** 2000. Allosteric regulation of the class III anaerobic ribonucleotide reductase from bacteriophage T4. J. Biol. Chem. **275:**19443-19448.
6. **Andersson, J., M. Westman, M. Sahlin, and B. M. Sjoberg.** 2000. Cysteines involved in radical generation and catalysis of class III anaerobic ribonucleotide reductase. A protein engineering study of bacteriophage T4 NrdD. J. Biol. Chem. **275:**19449-19455.
7. **Ando, T., H. Tsumori, A. Shimamura, Y. Sato, and H. Mukasa.** 2003. Classification of oral streptococci by two-dimensional gel electrophoresis with direct activity stain for glycosyltransferases. Oral Microbiol. Immunol. **18:**171-175.
8. **Autret, N., I. Dubail, P. Trieu-Cuot, P. Berche, and A. Charbit.** 2001. Identification of new genes involved in the virulence of *Listeria monocytogenes* by signature-tagged transposon mutagenesis. Infect. Immun. **69:**2054-2065.
9. **Ayoub, E., and L. Wannamaker.** 1962. Evaluation of the streptococcal deoxyribonuclease B and diphosphopyridine nucleotidase antibody tests in acute rheumatic fever and acute glomerulonephritis. Pediatrics **29:**527-538.
10. **Barnard, J., and M. Stinson.** 1996. The alpha-hemolysin of *Streptococcus gordonii* is hydrogen peroxide. Infect. Immun. **64:**3853-3857.
11. **Barnett, T. C., A. R. Patel, and J. R. Scott.** 2004. A novel sortase, SrtC2, from *Streptococcus pyogenes* anchors a surface protein containing a QVPTGV motif to the cell wall. J. Bacteriol. **186:**5865-5875.

12. **Barnett, T. C., and J. R. Scott.** 2002. Differential recognition of surface proteins in *Streptococcus pyogenes* by two sortase gene homologs. J. Bacteriol. **184:**2181-2191.
13. **Barocchi, M. A., J. Ries, X. Zogaj, C. Hemsley, B. Albiger, A. Kanth, S. Dahlberg, J. Fernebro, M. Moschioni, V. Masignani, K. Hultenby, A. R. Taddei, K. Beiter, F. Wartha, A. von Euler, A. Covacci, D. W. Holden, S. Normark, R. Rappuoli, and B. Henriques-Normark.** 2006. A pneumococcal pilus influences virulence and host inflammatory responses. Proc. Natl. Acad. Sci. U.S.A. **103:**2857-2862.
14. **Beg, A. M., M. N. Jones, T. Miller-Torbert, and R. G. Holt.** 2002. Binding of *Streptococcus mutans* to extracellular matrix molecules and fibrinogen. Biochem. Biophys. Res. Commun. **298:**75-79.
15. **Bendtsen, J. D., H. Nielsen, G. von Heijne, and S. Brunak.** 2004. Improved prediction of signal peptides: SignalP 3.0. J. Mol. Biol. **340:**783-795.
16. **Bensing, B. A., I. R. Siboo, and P. M. Sullam.** 2007. Glycine residues in the hydrophobic core of the GspB signal sequence route export toward the accessory Sec pathway. J. Bacteriol. **189:**3846-3854.
17. **Bensing, B. A., and P. M. Sullam.** 2002. An accessory sec locus of *Streptococcus gordonii* is required for export of the surface protein GspB and for normal levels of binding to human platelets. Mol. Microbiol. **44:**1081-1094.
18. **Bentley, M. L., H. Gaweska, J. M. Kielec, and D. G. McCafferty.** 2007. Engineering the substrate specificity of *Staphylococcus aureus* sortase A: The beta6/beta7 loop from SrtB confers NPQTN recognition to SrtA. J. Biol. Chem. **282:**6571-6581.
19. **Beranova-Giorgianni, S., D. M. Desiderio, and M. J. Pabst.** 1998. Structures of biologically active muramyl peptides from peptidoglycan of *Streptococcus sanguis*. J. Mass. Spectrom. **33:**1182-1191.
20. **Bhatti, S., L. Vilenski, R. Tight, and J. R. A. Smego.** 2005. *Histoplasma* endocarditis: clinical and mycologic features and outcomes. J. Infect. **51:**2-9.
21. **Bianchi, V., R. Eliasson, M. Fontecave, E. Mulliez, D. M. Hoover, R. G. Matthews, and P. Reichard.** 1993. Flavodoxin is required for the activation of the anaerobic ribonucleotide reductase. Biochem. Biophys. Res. Commun. **197:**792-797.
22. **Bierne, H., C. Garandeau, M. G. Pucciarelli, C. Sabet, S. Newton, F. Garcia-del Portillo, P. Cossart, and A. Charbit.** 2004. Sortase B, a new class of sortase in *Listeria monocytogenes*. J. Bacteriol. **186:**1972-1982.
23. **Birkhed, D., K. Rosell, and K. Granath.** 1979. Structure of extracellular water-soluble polysaccharides synthesized from sucrose by oral strains of *Streptococcus mutans*, *Streptococcus salivarius*, *Streptococcus sanguis* and *Actinomyces viscosus*. Arch. Oral. Biol. **24:**53-61.
24. **Birkhed, D., and J. Tanzer.** 1979. Glycogen synthesis pathway in *Streptococcus mutans* strain NCTC 10449S and its glycogen synthesis-defective mutant 805. Arch. Oral. Biol. **24:**67-73.

25. **Boekhorst, J., M. W. H. J. de Been, M. Kleerebezem, and R. J. Siezen.** 2005. Genome-wide detection and analysis of cell wall-bound proteins with LPxTG-like sorting motifs. J. Bacteriol. **187:**4928-4934.
26. **Bolhuis, A., C. P. Broekhuizen, A. Sorokin, M. L. van Roosmalen, G. Venema, S. Bron, W. J. Quax, and J. M. van Dijl.** 1998. SecDF of *Bacillus subtilis*, a molecular siamese twin required for the efficient secretion of proteins. J. Biol. Chem. **273:**21217-21224.
27. **Boston, T., and T. Atlung.** 2003. FNR-mediated oxygen-responsive regulation of the *nrdDG* operon of *Escherichia coli*. J. Bacteriol. **185:**5310-5313.
28. **Bowden, M. G., W. Chen, J. Singvall, Y. Xu, S. J. Peacock, V. Valtulina, P. Speziale, and M. Hook.** 2005. Identification and preliminary characterization of cell wall-anchored proteins of *Staphylococcus epidermidis*. Microbiology **151:**1453-1464.
29. **Brendel, V., and E. N. Trifonov.** 1984. Presented at the Proceedings of the Ninth International CODATA Conference, Jerusalem, Isreal.
30. **Brendel, V., and E. N. Trifonov.** 1984. A computer algorithm for testing potential prokaryotic terminators. Nucleic Acids Res. **12:**4411-4427.
31. **Burall, L. S., J. M. Harro, X. Li, C. V. Lockatell, S. D. Himpsl, J. R. Hebel, D. E. Johnson, and H. L. Mobley.** 2004. *Proteus mirabilis* genes that contribute to pathogenesis of urinary tract infection: identification of 25 signature-tagged mutants attenuated at least 100-fold. Infect. Immun. **72:**2922-2938.
32. **Burne, R. A., Y. Y. M. Chen, D. L. Wexler, H. Kuramitsu, and W. H. Bowen.** 1996. Cariogenicity of *Streptococcus mutans* strains with defects in fructan metabolism assessed in a program-fed specific-pathogen-free rat model. J. Dent. Res. **75:**1572-1577.
33. **Burne, R. A., Z. T. Wen, Y. Y. M. Chen, and J. E. C. Penders.** 1999. Regulation of expression of the fructan hydrolase gene of *Streptococcus mutans* GS-5 by induction and carbon catabolite repression. J. Bacteriol. **181:**2863-2871.
34. **Burnette-Curley, D., V. Wells, H. Viscount, C. L. Munro, J. C. Fenno, P. Fives-Taylor, and F. L. Macrina.** 1995. FimA, a major virulence factor associated with *Streptococcus parasanguis* endocarditis. Infect. Immun. **63:**4669-4674.
35. **Burnie, J. P., W. Brooks, M. Donohoe, S. Hodgetts, A. al-Ghamdi, and R. C. Matthews.** 1996. Defining antibody targets in *Streptococcus oralis* infection. Infect. Immun. **64:**1600-8.
36. **Camacho, L. R., D. Ensergueix, E. Perez, B. Gicquel, and C. Guilhot.** 1999. Identification of a virulence gene cluster of *Mycobacterium tuberculosis* by signature-tagged transposon mutagenesis. Mol. Microbiol. **34:**257-267.
37. **Campo, N., H. Tjalsma, G. Buist, D. Stepniak, M. Meijer, M. Veenhuis, M. Westermann, J. P. Muller, S. Bron, J. Kok, O. P. Kuipers, and J. D. H. Jongbloed.** 2004. Subcellular sites for bacterial protein export. Mol. Microbiol. **53:**1583-1599.
38. **Caparon, M. G., and J. R. Scott.** 1989. Excision and insertion of the conjugative transposon Tn916 involves a novel recombination mechanism. Cell **59:**1027-1034.

39. **Carlsson, F., M. Stalhammar-Carlemalm, K. Flardh, C. Sandin, E. Carlemalm, and G. Lindahl.** 2006. Signal sequence directs localized secretion of bacterial surface proteins. Nature **442:**943-946.
40. **Caufield, P. W., A. P. Dasanayake, Y. Li, Y. Pan, J. Hsu, and J. M. Hardin.** 2000. Natural history of *Streptococcus sanguinis* in the oral cavity of infants: Evidence for a discrete window of infectivity. Infect. Immun. **68:**4018-4023.
41. **Chambert, R., Y. Pereira, and M.-F. Petit-Glatron.** 2003. Purification and characterization of YfkN, a trifunctional nucleotide phosphoesterase secreted by *Bacillus subtilis*. J. Biochem. **134:**655-660.
42. **Chen, S., G. K. Paterson, H. H. Tong, T. J. Mitchell, and T. F. DeMaria.** 2005. Sortase A contributes to pneumococcal nasopharyngeal colonization in the chinchilla model. FEMS Microbiol. Lett. **253:**151-154.
43. **Chesbro, W., and R. Walker.** 1972. Detection of staphylococcal nuclease elaborated during lethal infections in mice. Infect. Immun. **6:**1028-1030.
44. **Chia, J.-S., L. Y. Chang, C.-T. Shun, Y.-Y. Chang, and J.-Y. Chen.** 2001. A 60-kilodalton immunodominant glycoprotein is essential for cell wall integrity and the maintenance of cell shape in *Streptococcus mutans*. Infect. Immun. **69:**6987-6998.
45. **Cicalini, S., G. Forcina, and F. G. De Rosa.** 2001. Infective endocarditis in patients with human immunodeficiency virus infection. J. Infect. **42:**267-271.
46. **Clarke, V. A., N. Platt, and T. D. Butters.** 1995. Cloning and expression of the beta-N-acetylglucosaminidase gene from *Streptococcus pneumoniae*. J. Biol. Chem. **270:**8805-8814.
47. **Claverys, J. P., A. Dintilhac, E. V. Pestova, B. Martin, and D. A. Morrison.** 1995. Construction and evaluation of new drug-resistance cassettes for gene disruption mutagenesis in *Streptococcus pneumoniae*, using an *ami* test platform. Gene **164:**123-128.
48. **Colby, S., G. Whiting, L. Tao, and R. RR.** 1995. Insertional inactivation of the *Streptococcus mutans dexA* (dextranase) gene results in altered adherence and dextran catabolism. Microbiology **141:**2929-2936.
49. **Comfort, D., and R. T. Clubb.** 2004. A comparative genome analysis identifies distinct sorting pathways in gram-positive bacteria. Infect. Immun. **72:**2710-2722.
50. **Coulter, S. N., W. R. Schwan, E. Y. Ng, M. H. Langhorne, H. D. Ritchie, S. Westbrock-Wadman, W. O. Hufnagle, K. R. Folger, A. S. Bayer, and C. K. Stover.** 1998. *Staphylococcus aureus* genetic loci impacting growth and survival in multiple infection environments. Mol. Microbiol. **30:**393-404.
51. **Cremieux, A., B. Maziere, J. Vallois, M. A. Ottaviani, A, H. Raffoul, A. Bouvet, J. Pocidalo, and C. Carbon.** 1989. Evaluation of antibiotic diffusion into cardiac vegetations by quantitative autoradiography. J. Infect. Dis. **159:**938-944.
52. **Crowley, P. J., L. J. Brady, S. M. Michalek, and A. S. Bleiweis.** 1999. Virulence of a spaP mutant of *Streptococcus mutans* in a gnotobiotic rat model. Infect. Immun. **67:**1201-1206.

53. **Deivanayagam, C., R. Rich, M. Carson, R. Owens, S. Danthuluri, T. Bice, M. Höök, and S. Narayana.** 2000. Novel fold and assembly of the repetitive B region of the *Staphylococcus aureus* collagen-binding surface protein. Structure **8:**67-78.
54. **Delcher, A. L., D. Harmon, S. Kasif, O. White, and S. L. Salzberg.** 1999. Improved microbial gene identification with GLIMMER. Nucleic Acids Res. **27:**4636-4641.
55. **Demuth, D., Y. Duan, W. Brooks, A. Holmes, R. McNab, and H. Jenkinson.** 1996. Tandem genes encode cell-surface polypeptides SspA and SspB which mediate adhesion of the oral bacterium *Streptococcus gordonii* to human and bacterial receptors. Mol. Microbiol. **20:**403-413.
56. **Dintilhac, A., G. Alloing, C. Granadel, and J. P. Claverys.** 1997. Competence and virulence of *Streptococcus pneumoniae*: Adc and PsaA mutants exhibit a requirement for Zn and Mn resulting from inactivation of putative ABC metal permeases. Mol. Microbiol. **25:**727-739.
57. **Douglas, C. W., J. Heath, K. K. Hampton, and F. E. Preston.** 1993. Identity of viridans streptococci isolated from cases of infective endocarditis. J. Med. Microbiol. **39:**179-182.
58. **Dramsi, S., E. Caliot, I. Bonne, S. Guadagnini, M.-C. Prevost, M. Kojadinovic, L. Lalioui, C. Poyart, and P. Trieu-Cuot.** 2006. Assembly and role of pili in group B streptococci. Mol. Microbiol. **60:**1401-1413.
59. **Dramsi, S., P. Trieu-Cuot, and H. Bierne.** 2005. Sorting sortases: a nomenclature proposal for the various sortases of Gram-positive bacteria. Res. Microbiol. **156:**289-297.
60. **Dunny, G.** 1990. Genetic functions and cell-cell interactions in the pheromone-inducible plasmid transfer system of *Enterococcus faecalis*. Mol. Microbiol. **4:**689-696.
61. **Dunny, G. M., D. L. Zimmerman, and M. L. Tortorello.** 1985. Induction of surface exclusion (entry exclusion) by *Streptococcus faecalis* sex pheromones: Use of monoclonal antibodies to identify an inducible surface antigen involved in the exclusion process. Proc. Natl. Acad. Sci. U.S.A. **82:**8582-8586.
62. **Duong, F., and W. Wickner.** 1997. The SecDFYajC domain of preprotein translocase controls preprotein movement by regulating SecA membrane cycling. EMBO J. **16:**4871-4879.
63. **Durack, D. T.** 1975. Experimental bacterial endocarditis. IV. Structure and evolution of very early lesions. J. Pathol. **115:**81-89.
64. **Durack, D. T., and P. B. Beeson.** 1972. Experimental bacterial endocarditis. II. Survival of a bacteria in endocardial vegetations. Br. J. Exp. Pathol. **53:**50-53.
65. **Durack, D. T., P. B. Beeson, and R. G. Petersdorf.** 1973. Experimental bacterial endocarditis. III. Production and progress of the disease in rabbits. Br. J. Exp. Pathol. **54:**142-151.
66. **Durack, D. T., A. S. Lukes, and D. K. Bright.** 1994. New criteria for diagnosis of infective endocarditis: utilization of specific echocardiographic findings. Duke Endocarditis Service. Am. J. Med. **96:**200-9.

67. **Dyson, C., R. A. Barnes, and G. A. J. Harrison.** 1999. Infective endocarditis: an epidemiological review of 128 episodes. J. Infect. **38:**87-93.
68. **Eklund, H., U. Uhlin, M. Färnegårdh, D. T. Logan, and P. Nordlund.** 2001. Structure and function of the radical enzyme ribonucleotide reductase. Prog. Biophys. Mol. Biol. **77:**177-268.
69. **El-Sabaeny, A., D. R. Demuth, Y. Park, and R. J. Lamont.** 2000. Environmental conditions modulate the expression of the *sspA* and *sspB* genes in *Streptococcus gordonii*. Microb. Pathog. **29:**101-113.
70. **Eliasson, R., E. Pontis, X. Sun, and P. Reichard.** 1994. Allosteric control of the substrate specificity of the anaerobic ribonucleotide reductase from *Escherichia coli*. J. Biol. Chem. **269:**26052-26057.
71. **Facklam, R.** 2002. What happened to the streptococci: overview of taxonomic and nomenclature changes. Clin. Microbiol. Rev. **15:**613-630.
72. **Fischetti, V., V. Pancholi, and O. Schneewind.** 1990. Conservation of a hexapeptide sequence in the anchor region of surface proteins from gram-positive cocci. Mol. Microbiol. **4:**1603-1605.
73. **Flashner, Y., E. Mamroud, A. Tidhar, R. Ber, M. Aftalion, D. Gur, S. Lazar, A. Zvi, T. Bino, N. Ariel, B. Velan, A. Shafferman, and S. Cohen.** 2004. Generation of *Yersinia pestis* attenuated strains by signature-tagged mutagenesis in search of novel vaccine candidates. Infect. Immun. **72:**908-915.
74. **Fontaine, M. C., J. Perez-Casal, and P. J. Willson.** 2004. Investigation of a novel DNase of *Streptococcus suis* serotype 2. Infect. Immun. **72:**774-781.
75. **Forner, L., T. Larsen, M. Kilian, and P. Holmstrup.** 2006. Incidence of bacteremia after chewing, tooth brushing and scaling in individuals with periodontal inflammation. J. Clin. Periodontol. **33:**401-407.
76. **Froeliger, E. H., and P. Fives-Taylor.** 2001. *Streptococcus parasanguis* fimbria-associated adhesin Fap1 is required for biofilm formation. Infect. Immun. **69:**2512-2519.
77. **Garriga, X., R. Eliasson, E. Torrents, A. Jordan, J. Barbe, I. Gibert, and P. Reichard.** 1996. *nrdD* and *nrdG* genes are essential for strict anaerobic growth of *Escherichia coli*. Biochem. Biophys. Res. Commun. **229:**189-192.
78. **Gaudreau, M.-C., P. Lacasse, and B. G. Talbot.** 2007. Protective immune responses to a multi-gene DNA vaccine against *Staphylococcus aureus*. Vaccine **25:**814-824.
79. **Gaustad, P., and L. S. Havardstein.** 1997. Competence-pheromone in *Streptococcus sanguis*. Identification of the competence gene *comC* and the competence pheromone. Adv. Exp. Med. Biol. **418:**1019-1021.
80. **Ghuysen, J., and J. Strominger.** 1963. Structure of the cell wall of *Staphylococcus aureus*, strain Copenhagen. II. Separation and structure of disaccharides. Biochemistry **2:**1119-1125.
81. **Gibbons, R. J., and J. V. Houte.** 1975. Bacterial adherence in oral microbial ecology. Annu. Rev. Microbiol. **29:**19-42.

82. **Giefing, C., A. L. Meinke, M. Hanner, T. Henics, D. B. Minh, D. Gelbmann, U. Lundberg, B. M. Senn, M. Schunn, A. Habel, B. Henriques-Normark, A. Ortqvist, M. Kalin, A. von Gabain, and E. Nagy.** 2008. Discovery of a novel class of highly conserved vaccine antigens using genomic scale antigenic fingerprinting of pneumococcus with human antibodies. J. Exp. Med. **205:**117-131.
83. **Gold, W., F. B. Preston, M. C. Lache, and H. Blechman.** 1974. Production of levan and dextran in plaque *in vivo*. J. Dent. Res. **53:**442-446.
84. **Goldstein, J. M., D. Nelson, T. Kordula, J. A. Mayo, and J. Travis.** 2002. Extracellular arginine aminopeptidase from *Streptococcus gordonii* FSS2. Infect. Immun. **70:**836-843.
85. **Gong, K., L. Mailloux, and M. C. Herzberg.** 2000. Salivary film expresses a complex, macromolecular binding site for *Streptococcus sanguis*. J. Biol. Chem. **275:**8970-8974.
86. **Guntheroth, W. G.** 1984. How important are dental procedures as a cause of infective endocarditis? Am. J. Cardiol. **54:**797-801.
87. **Haffajee, A. D., and S. S. Socransky.** 1994. Microbial etiological agents of destructive periodontal diseases. Periodontology 2000 **5:**78-111.
88. **Hancock, L. E., and M. Perego.** 2004. Systematic inactivation and phenotypic characterization of two-component signal transduction systems of *Enterococcus faecalis* V583. J. Bacteriol. **186:**7951-7958.
89. **Handley, P. S., F. F. Correia, K. Russell, B. Rosan, and J. M. DiRienzo.** 2005. Association of a novel high molecular weight, serine-rich protein (SrpA) with fibril-mediated adhesion of the oral biofilm bacterium Streptococcus cristatus. Oral Microbiol. Immunol. **20:**131-140.
90. **Hansen, M. C., R. J. Palmer, Jr, C. Udsen, D. C. White, and S. Molin.** 2001. Assessment of GFP fluorescence in cells of *Streptococcus gordonii* under conditions of low pH and low oxygen concentration. Microbiology **147:**1383-1391.
91. **Hasona, A., P. J. Crowley, C. M. Levesque, R. W. Mair, D. G. Cvitkovitch, A. S. Bleiweis, and L. J. Brady.** 2005. Streptococcal viability and diminished stress tolerance in mutants lacking the signal recognition particle pathway or YidC2. Proc. Natl. Acad. Sci. U.S.A. **102:**17466-17471.
92. **Hava, D. L., and A. Camilli.** 2002. Large-scale identification of serotype 4 *Streptococcus pneumoniae* virulence factors. Mol. Microbiol. **45:**1389-406.
93. **Hayes, F.** 2003. Transposon-based strategies for microbial functional genomics and proteomics. Annu. Rev. Genet. **37:**3-29.
94. **Hensel, M.** 1998. Whole genome scan for habitat-specific genes by signature-tagged mutagenesis. Electrophoresis **19:**608-612.
95. **Hensel, M., J. E. Shea, C. Gleeson, M. D. Jones, E. Dalton, and D. W. Holden.** 1995. Simultaneous identification of bacterial virulence genes by negative selection. Science **269:**400-403.
96. **Herzberg, M. C.** 1996. Platelet-streptococcal interactions in endocarditis. Crit. Rev. Oral. Biol. Med. **7:**222-236.

97. **Herzberg, M. C., G. D. MacFarlane, K. Gong, N. N. Armstrong, A. R. Witt, P. R. Erickson, and M. W. Meyer.** 1992. The platelet interactivity phenotype of *Streptococcus sanguis* influences the course of experimental endocarditis. Infect. Immun. **60:**4809-4818.

98. **Herzberg MC, N. A., Tao L, Kilic A, Beckman E, Khammanivong A, Zhang Y.** 2005. Oral streptococci and cardiovascular disease: Searching for the platelet aggregation-associated protein gene and mechanisms of *Streptococcus sanguis*-induced thrombosis. J. Periodontol. **76:**2101-5.

99. **Hienz, S. A., T. Schennings, A. Heimdahl, and J. I. Flock.** 1996. Collagen binding of *Staphylococcus aureus* is a virulence factor in experimental endocarditis. J. Infect. Dis. **174:**83-88.

100. **Higuchi, M., Y. Iwami, T. Yamada, and S. Araya.** 1970. Levan synthesis and accumulation by human dental plaque. Arch. Oral. Biol. **15:**563-567.

101. **Hirt, H., S. L. Erlandsen, and G. M. Dunny.** 2000. Heterologous inducible expression of *Enterococcus faecalis* pCF10 aggregation aubstance Asc10 in *Lactococcus lactis* and *Streptococcus gordonii* contributes to cell hydrophobicity and adhesion to Fibrin. J. Bacteriol. **182:**2299-2306.

102. **Holmes, A. R., C. Gilbert, J. M. Wells, and H. F. Jenkinson.** 1998. Binding properties of *Streptococcus gordonii* SspA and SspB (antigen I/II family) polypeptides expressed on the cell surface of *Lactococcus lactis* MG1363. Infect. Immun. **66:**4633-4639.

103. **Holmes, A. R., P. K. Gopal, and H. F. Jenkinson.** 1995. Adherence of *Candida albicans* to a cell surface polysaccharide receptor on *Streptococcus gordonii*. Infect. Immun. **63:**1827-1834.

104. **Hook, E. W., III, and M. A. Sande.** 1974. Role of the vegetation in experimental *Streptococcus viridans* endocarditis. Infect. Immun. **10:**1433-1438.

105. **Horton, R. M.** 1995. PCR-mediated recombination and mutagenesis. SOEing together tailor-made genes. Mol. Biotechnol. **3:**93-99.

106. **Hoshino, T., T. Fujiwara, and M. Kilian.** 2005. Use of phylogenetic and phenotypic analyses to identify nonhemolytic streptococci isolated from bacteremic patients. J. Clin. Microbiol. **43:**6073-6085.

107. **Hu, P., Z. Bian, M. Fan, M. Huang, and P. Zhang.** 2008. Sec translocase and sortase A are colocalised in a locus in the cytoplasmic membrane of *Streptococcus mutans*. Arch. Oral Biol. **53:**150-154.

108. **Hytonen, J., S. Haataja, and J. Finne.** 2003. *Streptococcus pyogenes* glycoprotein-binding strepadhesin activity is mediated by a surface-associated carbohydrate-degrading enzyme, pullulanase. Infect. Immun. **71:**784-793.

109. **Ilangovan, U., H. Ton-That, J. Iwahara, O. Schneewind, and R. T. Clubb.** 2001. Structure of sortase, the transpeptidase that anchors proteins to the cell wall of *Staphylococcus aureus*. Proc. Natl. Acad. Sci. U.S.A. **98:**6056-6061.

110. **Jakubovics, N. S., S. W. Kerrigan, A. H. Nobbs, N. Stromberg, C. J. van Dolleweerd, D. M. Cox, C. G. Kelly, and H. F. Jenkinson.** 2005. Functions of

cell surface-anchored Antigen I/II family and Hsa polypeptides in interactions of *Streptococcus gordonii* with host receptors. Infect. Immun. **73:**6629-6638.

111. **Janulczyk, R., and M. Rasmussen.** 2001. Improved pattern for genome-based screening identifies novel cell wall-attached proteins in gram-positive bacteria. Infect. Immun. **69:**4019-4026.
112. **Jones, A. L., K. M. Knoll, and C. E. Rubens.** 2000. Identification of *Streptococcus agalactiae* virulence genes in the neonatal rat sepsis model using signature-tagged mutagenesis. Mol. Microbiol. **37:**1444-1455.
113. **Jonsson, I. M., S. K. Mazmanian, O. Schneewind, T. Bremell, and A. Tarkowski.** 2003. The role of *Staphylococcus aureus* sortase A and sortase B in murine arthritis. Microbes Infect. **5:**775-780.
114. **Jönsson, K., C. Signäs, H. Müller, and M. Lindberg.** 1991. Two different genes encode fibronectin binding proteins in *Staphylococcus aureus*. The complete nucleotide sequence and characterization of the second gene. Eur. J. Biochem. **202:**1041-1048.
115. **Jordan, A., I. Gibert, and J. Barbe.** 1994. Cloning and sequencing of the genes from *Salmonella typhimurium* encoding a new bacterial ribonucleotide reductase. J. Bacteriol. **176:**3420-3427.
116. **Jordan, A., E. Pontis, F. Aslund, U. Hellman, I. Gibert, and P. Reichard.** 1996. The ribonucleotide reductase system of *Lactococcus lactis*. J. Biol. Chem. **271:**8779-8785.
117. **Jordan, A., E. Pontis, M. Atta, M. Krook, I. Gibert, J. Barbe, and P. Reichard.** 1994. A second class I ribonucleotide reductase in Enterobacteriaceae: Characterization of the *Salmonella typhimurium* enzyme. Proc. Natl. Acad. Sci. U.S.A. **91:**12892-12896.
118. **Jordan, A., and P. Reichard.** 1998. Ribonucleotide reductases. Annu. Rev. Biochem. **67:**71-98.
119. **Kao, S. M., S. B. Olmsted, A. S. Viksnins, J. C. Gallo, and G. M. Dunny.** 1991. Molecular and genetic analysis of a region of plasmid pCF10 containing positive control genes and structural genes encoding surface proteins involved in pheromone-inducible conjugation in *Enterococcus faecalis*. J. Bacteriol. **173:**7650-7664.
120. **Kawamura, Y., X. G. Hou, F. Sultana, H. Miura, and T. Ezaki.** 1995. Determination of 16S rRNA sequences of *Streptococcus mitis* and *Streptococcus gordonii* and phylogenetic relationships among members of the genus *Streptococcus*. Int. J. Syst. Bacteriol. **45:**406-408.
121. **Kemp, K. D., K. V. Singh, S. R. Nallapareddy, and B. E. Murray.** 2007. Relative contributions of *Enterococcus faecalis* OG1RF sortase-encoding genes, *srtA* and *bps* (*srtC*), to biofilm formation and a murine model of urinary tract infection. Infect. Immun. **75:**5399-5404.
122. **Kerrigan, S. W., I. Douglas, A. Wray, J. Heath, M. F. Byrne, D. Fitzgerald, and D. Cox.** 2002. A role for glycoprotein Ib in *Streptococcus sanguis*-induced platelet aggregation. Blood **100:**509-516.

123. **Kharat, A. S., and A. Tomasz.** 2003. Inactivation of the *srtA* gene affects localization of surface proteins and decreases adhesion of *Streptococcus pneumoniae* to human pharyngeal cells *in vitro*. Infect. Immun. **71:**2758-2765.
124. **Kilian, M., L. Mikkelsen, and J. Henrichsen.** 1989. Taxonomic study of viridans streptococci: description of *Streptococcus gordonii* sp. nov. and emended descriptions of *Streptococcus sanguis* (White and Niven 1946), *Streptococcus oralis* (Bridge and Sneath 1982), and *Streptococcus mitis* (Andrewes and Horder 1906). Int. J. Syst. Bacteriol. **39:**471-484.
125. **King, D. S., and P. Reichard.** 1995. Mass spectrometric determination of the radical scission site in the anaerobic ribonucleotide reductase of *Escherichia coli*. Biochem. Biophys. Res. Commun. **206:**731-735.
126. **Kitten, T., C. L. Munro, S. M. Michalek, and F. L. Macrina.** 2000. Genetic characterization of a *Streptococcus mutans* LraI family operon and role in virulence. Infect. Immun. **68:**4441-4451.
127. **Kitten, T., C. L. Munro, A. Wang, and F. L. Macrina.** 2002. Vaccination with FimA from *Streptococcus parasanguis* protects rats from endocarditis caused by other viridans streptococci. Infect. Immun. **70:**422-425.
128. **Kolberg, M., K. R. Strand, P. Graff, and K. Kristoffer Andersson.** 2004. Structure, function, and mechanism of ribonucleotide reductases. Biochim. Biophys. Acta **1699:**1-34.
129. **Kremer, B., B. Loos, U. van der Velden, A. van Winkelhoff, J. Craandijk, H. Bulthuis, J. Hutter, A. Varoufaki, and T. van Steenbergen.** 2000. *Peptostreptococcus micros* smooth and rough genotypes in periodontitis and gingivitis. J. Periodontol. **71:**209-218.
130. **Kremer, B. H., J. T. Magee, P. J. van Dalen, and T. J. Martijn van Steenbergen.** 1997. Characterization of smooth and rough morphotypes of *Peptostreptococcus micros*. Int. J. Syst. Bacteriol. **47:**363-368.
131. **Kremer, B. H. A., J. J. E. Bijlsma, J. G. Kusters, J. de Graaff, and T. J. M. van Steenbergen.** 1999. Cloning of *fibA*, encoding an immunogenic subunit of the fibril-like surface structure of *Peptostreptococcus micros*. J. Bacteriol. **181:**2485-2491.
132. **Kreth, J., J. Merritt, W. Shi, and F. Qi.** 2005. Competition and coexistence between *Streptococcus mutans* and *Streptococcus sanguinis* in the dental biofilm. J. Bacteriol. **187:**7193-7203.
133. **Kruger, R. G., B. Otvos, B. A. Frankel, M. Bentley, P. Dostal, and D. G. McCafferty.** 2004. Analysis of the substrate specificity of the *Staphylococcus aureus* sortase transpeptidase SrtA. Biochemistry **43:**1541-1551.
134. **Lacks, S. A., and S. S. Springhorn.** 1984. Transfer of recombinant plasmids containing the gene for DpnII DNA methylase into strains of *Streptococcus pneumoniae* that produce DpnI or DpnII restriction endonucleases. J. Bacteriol. **158:**905-909.
135. **Lalioui L, P. E., Dramsi S, Baptista M, Bourgeois N, Doucet-Populaire F, Rusniok C, Zouine M, Glaser P, Kunst F, Poyart C, Trieu-Cuot P.** 2005. The

SrtA sortase of *Streptococcus agalactiae* is required for cell wall anchoring of proteins containing the LPXTG motif, for adhesion to epithelial cells, and for colonization of the mouse intestine. Infect. Immun. **73:**3342-3350.

136. **Lamont, R., S. Gil, D. Demuth, D. Malamud, and B. Rosan.** 1994. Molecules of *Streptococcus gordonii* that bind to *Porphyromonas gingivalis*. Microbiology **140:**867-872.
137. **Lampe, D. J., B. J. Akerley, E. J. Rubin, J. J. Mekalanos, and H. M. Robertson.** 1999. Hyperactive transposase mutants of the *Himar1 mariner* transposon. Proc. Natl. Acad. Sci. U.S.A. **96:**11428-11433.
138. **Lampe, D. J., M. E. Churchill, and H. M. Robertson.** 1996. A purified *mariner* transposase is sufficient to mediate transposition *in vitro* [published erratum appears in EMBO J. 16:4153]. EMBO J. **15:**5470-5479.
139. **Lampe, D. J., T. E. Grant, and H. M. Robertson.** 1998. Factors affecting transposition of the *Himar1 mariner* transposon *in vitro*. Genetics **149:**179-187.
140. **Lampe, D. J., K. K. Walden, and H. M. Robertson.** 2001. Loss of transposase-DNA interaction may underlie the divergence of *mariner* family transposable elements and the ability of more than one mariner to occupy the same genome. Mol. Biol. Evol. **18:**954-961.
141. **Lannergård, J., L. Frykberg, and B. Guss.** 2003. CNE, a collagen-binding protein of *Streptococcus equi*. FEMS Microbiol. Lett. **222:**69-74.
142. **Larsson, A., and B. Sjöberg.** 1986. Identification of the stable free radical tyrosine residue in ribonucleotide reductase. EMBO J. **5:**2037-2040.
143. **Larsson, K. M., J. Andersson, B. M. Sjöberg, P. Nordlund, and D. T. Logan.** 2001. Structural basis for allosteric substrate specificity regulation in anaerobic ribonucleotide reductases. Structure **9:**739-750.
144. **Lauer, P., C. D. Rinaudo, M. Soriani, I. Margarit, D. Maione, R. Rosini, A. R. Taddei, M. Mora, R. Rappuoli, G. Grandi, and J. L. Telford.** 2005. Genome analysis reveals pili in Group B *Streptococcus*. Science **309:**105.
145. **Lee, S. F., and T. L. Boran.** 2003. Roles of sortase in surface expression of the major protein adhesin P1, saliva-induced aggregation and adherence, and cariogenicity of *Streptococcus mutans*. Infect. Immun. **71:**676-681.
146. **Levesque, C. M., E. Voronejskaia, Y.-C. C. Huang, R. W. Mair, R. P. Ellen, and D. G. Cvitkovitch.** 2005. Involvement of sortase anchoring of cell wall proteins in biofilm formation by *Streptococcus mutans*. Infect. Immun. **73:**3773-3777.
147. **Li, J. S., D. J. Sexton, N. Mick, R. Nettles, V. G. Fowler, Jr., T. Ryan, T. Bashore, and G. R. Corey.** 2000. Proposed modifications to the Duke criteria for the diagnosis of infective endocarditis. Clin. Infect. Dis. **30:**633-638.
148. **Libus, J., and H. Storchova.** 2006. Quantification of cDNA generated by reverse transcription of total RNA provides a simple alternative tool for quantitative RT-PCR normalization. Biotechniques **41:**156-164.
149. **Licht, S., G. Gerfen, and J. Stubbe.** 1996. Thiyl radicals in ribonucleotide reductases. Science **271:**477-481.

150. **Loesche, W. J.** 1986. Role of *Streptococcus mutans* in human dental decay. Microbiol. Rev. **50:**353-380.
151. **Loo, C. Y., D. A. Corliss, and N. Ganeshkumar.** 2000. *Streptococcus gordonii* biofilm formation: identification of genes that code for biofilm phenotypes. J. Bacteriol. **182:**1374-1382.
152. **Love, R. M., M. D. McMillan, and H. F. Jenkinson.** 1997. Invasion of dentinal tubules by oral streptococci is associated with collagen recognition mediated by the antigen I/II family of polypeptides. Infect. Immun. **65:**5157-5164.
153. **Love, R. M., M. D. McMillan, Y. Park, and H. F. Jenkinson.** 2000. Coinvasion of dentinal tubules by *Porphyromonas gingivalis* and *Streptococcus gordonii* depends upon binding specificity of streptococcal antigen I/II adhesin. Infect. Immun. **68:**1359-1068.
154. **Macrina, F. L., J. A. Tobian, K. R. Jones, R. P. Evans, and D. B. Clewell.** 1982. A cloning vector able to replicate in *Escherichia coli* and *Streptococcus sanguis*. Gene **19:**345-353.
155. **Manetti, A. G. O., C. Zingaretti, F. Falugi, S. Capo, M. Bombaci, F. Bagnoli, G. Gambellini, G. Bensi, M. Mora, A. M. Edwards, J. M. Musser, E. A. Graviss, J. L. Telford, G. Grandi, and I. Margarit.** 2007. *Streptococcus pyogenes* pili promote pharyngeal cell adhesion and biofilm formation. Mol. Microbiol. **64:**968-983.
156. **Marchler-Bauer, A., and S. H. Bryant.** 2004. CD-Search: protein domain annotations on the fly. Nucl. Acids Res. **32:**W327-331.
157. **Maresso, A. W., T. J. Chapa, and O. Schneewind.** 2006. Surface protein IsdC and sortase B are required for heme-iron scavenging of *Bacillus anthracis*. J. Bacteriol. **188:**8145-8152.
158. **Marraffini, L. A., A. C. DeDent, and O. Schneewind.** 2006. Sortases and the art of anchoring proteins to the envelopes of gram-positive bacteria. Microbiol. Mol. Biol. Rev. **70:**192-221.
159. **Marraffini, L. A., and O. Schneewind.** 2005. Anchor structure of staphylococcal surface proteins: V. Anchor structure of the sortase B substrate IsdC. J. Biol. Chem. **280:**16263-16271.
160. **Marraffini, L. A., H. Ton-That, Y. Zong, S. V. L. Narayana, and O. Schneewind.** 2004. Anchoring of Surface Proteins to the Cell Wall of Staphylococcus aureus: A conserved arginine residue is required for efficient catalysis of sortase A. J. Biol. Chem. **279:**37763-37770.
161. **Martin, B., M. Prudhomme, G. Alloing, C. Granadel, and J. P. Claverys.** 2000. Cross-regulation of competence pheromone production and export in the early control of transformation in *Streptococcus pneumoniae*. Mol. Microbiol. **38:**867-878.
162. **Masalha, M., I. Borovok, R. Schreiber, Y. Aharonowitz, and G. Cohen.** 2001. Analysis of transcription of the *Staphylococcus aureus* aerobic class Ib and anaerobic class III ribonucleotide reductase genes in response to oxygen. J. Bacteriol. **183:**7260-7272.

163. **Mazmanian, S. K., G. Liu, E. R. Jensen, E. Lenoy, and O. Schneewind.** 2000. *Staphylococcus aureus* sortase mutants defective in the display of surface proteins and in the pathogenesis of animal infections. Proc. Natl. Acad. Sci. U.S.A. **97:**5510-5515.

164. **Mazmanian, S. K., E. P. Skaar, A. H. Gaspar, M. Humayun, P. Gornicki, J. Jelenska, A. Joachmiak, D. M. Missiakas, and O. Schneewind.** 2003. Passage of heme-iron across the envelope of *Staphylococcus aureus*. Science **299:**906-909.

165. **Mazmanian, S. K., H. Ton-That, K. Su, and O. Schneewind.** 2002. An iron-regulated sortase anchors a class of surface protein during *Staphylococcus aureus* pathogenesis. Proc. Natl. Acad. Sci. U.S.A. **99:**2293-2298.

166. **McCormick, J. K., H. Hirt, C. M. Waters, T. J. Tripp, G. M. Dunny, and P. M. Schlievert.** 2001. Antibodies to a surface-exposed, N-terminal domain of aggregation substance are not protective in the rabbit model of *Enterococcus faecalis* infective endocarditis. Infect. Immun. **69:**3305-3314.

167. **McDevitt, D., T. Nanavaty, K. House-Pompeo, E. Bell, N. Turner, L. McIntire, T. Foster, and M. Hook.** 1997. Characterization of the interaction between the *Staphylococcus aureus* clumping factor (ClfA) and fibrinogen. Eur. J. Biochem. **247:**416-424.

168. **McNab, R., H. Forbes, P. S. Handley, D. M. Loach, G. W. Tannock, and H. F. Jenkinson.** 1999. Cell wall-anchored CshA polypeptide (259 kilodaltons) in *Streptococcus gordonii* forms surface fibrils that confer hydrophobic and adhesive properties. J. Bacteriol. **181:**3087-3095.

169. **McNab, R., A. R. Holmes, J. M. Clarke, G. W. Tannock, and H. F. Jenkinson.** 1996. Cell surface polypeptide CshA mediates binding of *Streptococcus gordonii* to other oral bacteria and to immobilized fibronectin. Infect. Immun. **64:**4204-4210.

170. **McNab, R., H. F. Jenkinson, D. M. Loach, and G. W. Tannock.** 1994. Cell-surface-associated polypeptides CshA and CshB of high molecular mass are colonization determinants in the oral bacterium *Streptococcus gordonii*. Mol. Microbiol. **14:**743-754.

171. **Millar, B. C., and J. E. Moore.** 2004. Emerging issues in infective endocarditis. Emerg. Infect. Dis. **10:**1110-1116.

172. **Mills, M. F., M. E. Marquart, and L. S. McDaniel.** 2007. Localization of PcsB of *Streptococcus pneumoniae* and its differential expression in response to stress. J. Bacteriol. **189:**4544-4546.

173. **Mishra, A., A. Das, J. O. Cisar, and H. Ton-That.** 2007. Sortase-catalyzed assembly of distinct heteromeric fimbriae in *Actinomyces naeslundii*. J. Bacteriol. **189:**3156-3165.

174. **Monje-Casas, F., J. Jurado, M.-J. Prieto-Alamo, A. Holmgren, and C. Pueyo.** 2001. Expression analysis of the *nrdHIEF* operon from *Escherichia coli*. Conditions that trigger the transcript level *in vivo*. J. Biol. Chem. **276:**18031-18037.

175. **Mora, M., G. Bensi, S. Capo, F. Falugi, C. Zingaretti, A. G. O. Manetti, T. Maggi, A. R. Taddei, G. Grandi, and J. L. Telford.** 2005. Group A

Streptococcus produce pilus-like structures containing protective antigens and Lancefield T antigens. Proc. Natl. Acad. Sci. U.S.A. **102:**15641-15646.

176. **Moreillon, P., and Y. A. Que.** 2004. Infective endocarditis. Lancet **363:**139-49.
177. **Moreillon, P., Y. A. Que, and A. S. Bayer.** 2002. Pathogenesis of streptococcal and staphylococcal endocarditis. Infect. Dis. Clin. North Am. **16:**297-318.
178. **Morsczeck, C., T. Prokhorova, J. Sigh, M. Pfeiffer, M. Bille-Nielsen, J. Petersen, A. Boysen, T. Kofoed, N. Frimodt-Møller, P. Nyborg-Nielsen, and P. Schrotz-King.** 2008. *Streptococcus pneumoniae*: proteomics of surface proteins for vaccine development. Clin. Microbes Infect. **14:**74-81.
179. **Munro, C. L.** 1998. The rat model of endocarditis. Methods Cell Science **20:**203-207.
180. **Munro, C. L., and F. L. Macrina.** 1993. Sucrose-derived exopolysaccharides of *Streptococcus mutans* V403 contribute to infectivity in endocarditis. Mol. Microbiol. **8:**133-142.
181. **Mylonakis, E., and S. B. Calderwood.** 2001. Infective endocarditis in adults. N. Engl. J. Med. **345:**1318-1330.
182. **Naik, M. T., N. Suree, U. Ilangovan, C. K. Liew, W. Thieu, D. O. Campbell, J. J. Clemens, M. E. Jung, and R. T. Clubb.** 2006. *Staphylococcus aureus* sortase A transpeptidase: Calcium promotes sorting signal binding by altering the mobility and structure of an active site loop. J. Biol. Chem. **281:**1817-1826.
183. **Nallapareddy, S., K. Singh, J. Sillanpää, D. Garsin, M. Höök, S. Erlandsen, and B. Murray.** 2006. Endocarditis and biofilm-associated pili of *Enterococcus faecalis*. J. Clin. Invest. **116:**2582-2584.
184. **Nallapareddy, S. R., G. M. Weinstock, and B. E. Murray.** 2003. Clinical isolates of *Enterococcus faecium* exhibit strain-specific collagen binding mediated by Acm, a new member of the MSCRAMM family. Mol. Microbiol. **47:**1733-1747.
185. **Navarre, W. W., and O. Schneewind.** 1999. Surface proteins of gram-positive bacteria and mechanisms of their targeting to the cell wall envelope. Microbiol. Mol. Biol. Rev. **63:**174-229.
186. **Netzer, R. O., E. Zollinger, C. Seiler, and A. Cerny.** 2000. Infective endocarditis: clinical spectrum, presentation and outcome. An analysis of 212 cases 1980-1995. Heart **84:**25-30.
187. **Newton, S. M. C., P. E. Klebba, C. Raynaud, Y. Shao, X. Jiang, I. Dubail, C. Archer, C. Frehel, and A. Charbit.** 2005. The *svpA-srtB* locus of *Listeria monocytogenes*: Fur-mediated iron regulation and effect on virulence. Mol. Microbiol. **55:**927-940.
188. **Nilsson, M., J. Bjerketorp, A. Wiebensjo, A. Ljungh, L. Frykberg, and B. Guss.** 2004. A von Willebrand factor-binding protein from *Staphylococcus lugdunensis*. FEMS Microbiol. Lett. **234:**155-161.
189. **Nobbs, A. H., R. M. Vajna, J. R. Johnson, Y. Zhang, S. L. Erlandsen, M. W. Oli, J. Kreth, L. J. Brady, and M. C. Herzberg.** 2007. Consequences of a sortase A mutation in *Streptococcus gordonii*. Microbiology **153:**4088-4097.

190. **O'Toole, G. A., and R. Kolter.** 1998. Initiation of biofilm formation in *Pseudomonas fluorescens* WCS365 proceeds via multiple, convergent signaling pathways: a genetic analysis. Mol. Microbiol. **28:**449-461.
191. **O'Toole, R., L. Goode, and C. Howe.** 1971. Neuraminidase activity in bacterial meningitis. J. Clin. Invest. **50:**979-985.
192. **Oggioni, M. R., C. G. Dowson, J. M. Smith, R. Provvedi, and G. Pozzi.** 1996. The tetracycline resistance g ene *tet*(M) exhibits mosaic structure. Plasmid **35:**156-163.
193. **Ollagnier, S., E. Mulliez, J. Gaillard, R. Eliasson, M. Fontecave, and P. Reichard.** 1996. The anaerobic *Escherichia coli* ribonucleotide reductase. J. Biol. Chem. **271:**9410-9416.
194. **Olmsted, S. B., S. L. Erlandsen, G. M. Dunny, and C. L. Wells.** 1993. High-resolution visualization by field emission scanning electron microscopy of *Enterococcus faecalis* surface proteins encoded by the pheromone-inducible conjugative plasmid pCF10. J. Bacteriol. **175:**6229-6237.
195. **Olmsted, S. B., S. M. Kao, L. J. van Putte, J. C. Gallo, and G. M. Dunny.** 1991. Role of the pheromone-inducible surface protein Asc10 in mating aggregate formation and conjugal transfer of the *Enterococcus faecalis* plasmid pCF10. J. Bacteriol. **173:**7665-7672.
196. **Paik, S.** 2004. Molecular analysis of virulence determinants for endocarditis in *Streptococcus mutans* and *Streptococcus sanguinis*. Dissertation. Virginia Commonwealth University, Richmond.
197. **Paik, S., A. Brown, C. L. Munro, C. N. Cornelissen, and T. Kitten.** 2003. The *sloABCR* operon of *Streptococcus mutans* encodes an Mn and Fe transport system required for endocarditis virulence and Its Mn-dependent repressor [published erratum appears in J. Bacteriol. 185:7301]. J. Bacteriol. **185:**5967-5975.
198. **Paik, S., L. Senty, S. Das, J. C. Noe, C. L. Munro, and T. Kitten.** 2005. Identification of virulence determinants for endocarditis in *Streptococcus sanguinis* by signature-tagged mutagenesis. Infect. Immun. **73:**6064-6074.
199. **Pallen, M. J., A. C. Lam, M. Antonio, and K. Dunbar.** 2001. An embarrassment of sortases-a richness of substrates? Trends Microbiol. **9:**97-102.
200. **Paterson, G. K., and T. J. Mitchell.** 2006. The role of *Streptococcus pneumoniae* sortase A in colonisation and pathogenesis. Microbes Infect. **8:**145-153.
201. **Pearce, C., G. Bowden, M. Evans, S. Fitzsimmons, J. Johnson, M. Sheridan, R. Wientzen, and M. Cole.** 1995. Identification of pioneer viridans streptococci in the oral cavity of human neonates. J. Med. Microbiol. **42:**67-72.
202. **Pearson, W. R., and D. J. Lipman.** 1988. Improved tools for biological sequence comparison. Proc. Natl. Acad. Sci. U.S.A. **85:**2444-2448.
203. **Petersson, L., A. Graslund, A. Ehrenberg, B. M. Sjoberg, and P. Reichard.** 1980. The iron center in ribonucleotide reductase from *Escherichia coli*. J. Biol. Chem. **255:**6706-6712.

204. **Plummer, C., H. Wu, S. W. Kerrigan, G. Meade, D. Cox, and C. W. I. Douglas.** 2005. A serine-rich glycoprotein of *Streptococcus sanguis* mediates adhesion to platelets via GPIb. Br. J. Haematol. **129:**101-109.
205. **Podbielski, A., B. Spellerberg, M. Woischnik, B. Pohl, and R. Lütticken.** 1996. Novel series of plasmid vectors for gene inactivation and expression analysis in group A streptococci (GAS). Gene **177:**137-147.
206. **Ponnuraj, K., M. Bowden, S. Davis, S. Gurusiddappa, D. Moore, D. Choe, Y. Xu, M. Hook, and S. Narayana.** 2003. A "dock, lock, and latch" structural model for a staphylococcal adhesin binding to fibrinogen. Cell **115:**217-228.
207. **Poveda-Roda, R., Y. Jiménez, E. Carbonell, C. Gavaldá, M. Margaix-Muñoz, and G. Sarrión-Pérez.** 2008. Bacteremia originating in the oral cavity. A review. Med. Oral Patol. Oral Cir. Bucal. **13:**E355-362.
208. **Prieto-Alamo, M.-J., J. Jurado, R. Gallardo-Madueno, F. Monje-Casas, A. Holmgren, and C. Pueyo.** 2000. Transcriptional regulation of glutaredoxin and thioredoxin pathways and related enzymes in response to oxidative stress. J. Biol. Chem. **275:**13398-13405.
209. **Rakhimova, E., A. Munder, L. Wiehlmann, F. Bredenbruch, and B. Tümmler.** 2008. Fitness of isogenic colony morphology variants of *Pseudomonas aeruginosa* in murine airway infection. PLoS ONE **3:**e1685.
210. **Ramirez-Ronda, C. H.** 1978. Adherence of glucan-positive and glucan-negative streptococcal strains to normal and damaged heart valves. J. Clin. Invest. **62:**805-814.
211. **Rawczynska-Englert I, H. T., Dzierzanowska D.** 2000. Evaluation of serum cytokine concentrations in patients with infective endocarditis. J. Heart Valve Dis. **9:**705-709.
212. **Reinscheid, D. J., B. Gottschalk, A. Schubert, B. J. Eikmanns, and G. S. Chhatwal.** 2001. Identification and molecular analysis of PcsB, a protein required for cell wall separation of Group B *Streptococcus*. J. Bacteriol. **183:**1175-1183.
213. **Reznikoff, W. S.** 2003. Tn*5* as a model for understanding DNA transposition. Mol. Microbiol. **47:**1199-206.
214. **Rhem, M. N., E. M. Lech, J. M. Patti, D. McDevitt, M. Hook, D. B. Jones, and K. R. Wilhelmus.** 2000. The collagen-binding adhesin is a virulence factor in *Staphylococcus aureus* keratitis. Infect. Immun. **68:**3776-3779.
215. **Rich, R. L., B. Kreikemeyer, R. T. Owens, S. LaBrenz, S. V. L. Narayana, G. M. Weinstock, B. E. Murray, and M. Hook.** 1999. Ace is a collagen-binding MSCRAMM from *Enterococcus faecalis*. J. Biol. Chem. **274:**26939-26945.
216. **Roberts, R. B., A. G. Krieger, N. L. Schiller, and K. C. Gross.** 1979. Viridans streptococcal endocarditis: the role of various species, including pyridoxal-dependent streptococci. Rev. Infect. Dis. **1:**955-966.
217. **Robertson, H., and E. MacLeod.** 1993. Five major subfamilies of *mariner* transposable elements in insects, including the Mediterranean fruit fly, and related arthropods. Insect Mol. Biol. **2:**125-139.

218. **Robertson, H. M.** 1993. The mariner transposable element is widespread in insects. Nature **362:**241-245.
219. **Roche, F. M., R. Massey, S. J. Peacock, N. P. J. Day, L. Visai, P. Speziale, A. Lam, M. Pallen, and T. J. Foster.** 2003. Characterization of novel LPXTG-containing proteins of *Staphylococcus aureus* identified from genome sequences. Microbiology **149:**643-654.
220. **Rosch, J. W., and M. G. Caparon.** 2005. The ExPortal: an organelle dedicated to the biogenesis of secreted proteins in *Streptococcus pyogenes*. Mol. Microbiol. **58:**959-968.
221. **Ruggeri, Z. M.** 2007. The role of von Willebrand factor in thrombus formation. Thromb. Res. **120:**S5-9.
222. **Salzberg, L. I., and J. D. Helmann.** 2007. An antibiotic-inducible cell wall-associated protein that protects *Bacillus subtilis* from autolysis. J. Bacteriol. **189:**4671-4680.
223. **Sato, Y., K. Okamoto, A. Kagami, Y. Yamamoto, T. Igarashi, and H. Kizaki.** 2004. *Streptococcus mutans* strains harboring collagen-binding adhesin. J. Dent. Res. **83:**534-539.
224. **Scannapieco, F. A., L. Solomon, and R. O. Wadenya.** 1994. Emergence in human dental plaque and host distribution of amylase-binding streptococci. J. Dent. Res. **73:**1627-1635.
225. **Scavone, P., A. Miyoshi, A. Rial, A. Chabalgoity, P. Langella, V. Azevedo, and P. Zunino.** 2007. Intranasal immunisation with recombinant *Lactococcus lactis* displaying either anchored or secreted forms of *Proteus mirabilis* MrpA fimbrial protein confers specific immune response and induces a significant reduction of kidney bacterial colonisation in mice. Microbes Infect. **9:**821-828.
226. **Schneewind, O., D. Mihaylova-Petkov, and P. Model.** 1993. Cell wall sorting signals in surface proteins of gram-positive bacteria. EMBO J. **12:**4803-4811.
227. **Schneewind, O., P. Model, and V. A. Fischetti.** 1992. Sorting of protein A to the staphylococcal cell wall. Cell **70:**267-281.
228. **Schrader, W. P., and J. S. Anderson.** 1978. Membrane-bound nucleotidase of *Bacillus cereus*. J. Bacteriol. **133:**576-583.
229. **Scott, J. R., and D. Zahner.** 2006. Pili with strong attachments: Gram-positive bacteria do it differently. Mol. Microbiol. **62:**320-330.
230. **Seifert, K. N., E. E. Adderson, A. A. Whiting, J. F. Bohnsack, P. J. Crowley, and L. J. Brady.** 2006. A unique serine-rich repeat protein (Srr-2) and novel surface antigen (epsilon) associated with a virulent lineage of serotype III *Streptococcus agalactiae*. Microbiology **152:**1029-1040.
231. **Senghas, E., J. M. Jones, M. Yamamoto, C. Gawron-Burke, and D. B. Clewell.** 1988. Genetic organization of the bacterial conjugative transposon *Tn*916. J. Bacteriol. **170:**245-249.
232. **Shebuski, R. J., and K. S. Kilgore.** 2002. Role of inflammatory mediators in thrombogenesis. J. Pharmacol. Exp. Ther. **300:**729-735.

233. **Shenep, J. L.** 2000. Viridans-group streptococcal infections in immunocompromised hosts. Int. J. Antimicrob. Agents **14:**129-135.
234. **Siboo, I. R., H. F. Chambers, and P. M. Sullam.** 2005. Role of SraP, a serine-rich surface protein of *Staphylococcus aureus*, in binding to human platelets. Infect. Immun. **73:**2273-2280.
235. **Sintchak, M., G. Arjara, B. Kellogg, J. Stubbe, and C. Drennan.** 2002. The crystal structure of class II ribonucleotide reductase reveals how an allosterically regulated monomer mimics a dimer. Nat. Struct. Biol. **9:**293-300.
236. **Smith, H. E., U. Vecht, H. J. Wisselink, N. Stockhofe-Zurwieden, Y. Biermann, and M. A. Smits.** 1996. Mutants of *Streptococcus suis* types 1 and 2 impaired in expression of muramidase-released protein and extracellular protein induce disease in newborn germfree pigs. Infect. Immun. **64:**4409-4412.
237. **Socransky, S., A. Haffajee, M. Cugini, C. Smith, and R. J. Kent.** 1998. Microbial complexes in subgingival plaque. J. Clin. Periodontol. **25:**134-144.
238. **Socransky, S., C. Smith, L. Martin, B. Paster, F. Dewhirst, and A. Levin.** 1994. "Checkerboard" DNA-DNA hybridization. Biotechniques **17:**788-792.
239. **Sofia, H. J., G. Chen, B. G. Hetzler, J. F. Reyes-Spindola, and N. E. Miller.** 2001. Radical SAM, a novel protein superfamily linking unresolved steps in familiar biosynthetic pathways with radical mechanisms: Functional characterization using new analysis and information visualization methods. Nucl. Acids Res. **29:**1097-1106.
240. **Stephenson, A. E., H. Wu, J. Novak, M. Tomana, K. Mintz, and P. Fives-Taylor.** 2002. The Fap1 fimbrial adhesin is a glycoprotein: antibodies specific for the glycan moiety block the adhesion of *Streptococcus parasanguis* in an *in vitro* tooth model. Mol. Microbiol. **43:**147-157.
241. **Stinson, M. W., S. Alder, and S. Kumar.** 2003. Invasion and killing of human endothelial cells by viridans group streptococci. Infect. Immun. **71:**2365-2372.
242. **Sun, Y. H., S. Bakshi, R. Chalmers, and C. M. Tang.** 2000. Functional genomics of *Neisseria meningitidis* pathogenesis. Nat. Med. **6:**1269-1273.
243. **Sung, C. K., H. Li, J. P. Claverys, and D. A. Morrison.** 2001. An *rpsL* cassette, Janus, for gene replacement through negative selection in *Streptococcus pneumoniae*. Appl. Environ. Microbiol. **67:**5190-5196.
244. **Swaminathan, A., A. Mandlik, A. Swierczynski, A. Gaspar, A. Das, and H. Ton-That.** 2007. Housekeeping sortase facilitates the cell wall anchoring of pilus polymers in *Corynebacterium diphtheriae*. Mol. Microbiol. **66:**961-974.
245. **Symersky, J., J. Patti, M. Carson, K. House-Pompeo, M. Teale, D. Moore, L. Jin, A. Schneider, L. DeLucas, H. M, and S. Narayana.** 1997. Structure of the collagen-binding domain from a *Staphylococcus aureus* adhesin. Nat. Struct. Biol. **4:**833-838.
246. **Sztukowska, M., M. Bugno, J. Potempa, J. Travis, and D. M. Kurtz, Jr.** 2002. Role of rubrerythrin in the oxidative stress response of *Porphyromonas gingivalis*. Mol. Microbiol. **44:**479-488.

247. **Takahashi, Y., K. Konishi, J. O. Cisar, and M. Yoshikawa.** 2002. Identification and Characterization of *hsa*, the Gene Encoding the Sialic Acid-Binding Adhesin of *Streptococcus gordonii* DL1. Infect. Immun. **70:**1209-1218.

248. **Takahashi, Y., E. Takashima, K. Shimazu, H. Yagishita, T. Aoba, and K. Konishi.** 2006. Contribution of sialic acid-binding adhesin to pathogenesis of experimental endocarditis caused by *Streptococcus gordonii* DL1. Infect. Immun. **74:**740-743.

249. **Takamatsu, D., B. A. Bensing, H. Cheng, G. A. Jarvis, I. R. Siboo, J. A. Lopez, J. M. Griffiss, and P. M. Sullam.** 2005. Binding of the *Streptococcus gordonii* surface glycoproteins GspB and Hsa to specific carbohydrate structures on platelet membrane glycoprotein Ibalpha. Mol. Microbiol. **58:**380-392.

250. **Takamatsu, D., B. A. Bensing, A. Prakobphol, S. J. Fisher, and P. M. Sullam.** 2006. Binding of the streptococcal surface glycoproteins GspB and Hsa to human salivary proteins. Infect. Immun. **74:**1933-1940.

251. **Tamesada, M., S. Kawabata, T. Fujiwara, and S. Hamada.** 2004. Synergistic effects of streptococcal glucosyltransferases on adhesive biofilm formation. J. Dent. Res. **83:**874-879.

252. **Tan, C., M. Liu, M. Jin, J. Liu, Y. Chen, T. Wu, T. Fu, W. Bei, and H. Chen.** 2008. The key virulence-associated genes of *Streptococcus suis* type 2 are upregulated and differentially expressed *in vivo*. FEMS Microbiol. Lett. **278:**108-114.

253. **Tanzer, J. M., L. Grant, A. Thompson, L. Li, J. D. Rogers, E. M. Haase, and F. A. Scannapieco.** 2003. Amylase-binding proteins A (AbpA) and B (AbpB) differentially affect colonization of rats' teeth by *Streptococcus gordonii*. Microbiology **149:**2653-2660.

254. **Ton-That, H., G. Liu, S. K. Mazmanian, K. F. Faull, and O. Schneewind.** 1999. Purification and characterization of sortase, the transpeptidase that cleaves surface proteins of *Staphylococcus aureus* at the LPXTG motif. Proc. Natl. Acad. Sci. U.S.A. **96:**12424-12429.

255. **Ton-That, H., L. A. Marraffini, and O. Schneewind.** 2004. Protein sorting to the cell wall envelope of Gram-positive bacteria. Biochem. Biophys. Acta **1694:**269-278.

256. **Ton-That, H., L. A. Marraffini, and O. Schneewind.** 2004. Sortases and pilin elements involved in pilus assembly of *Corynebacterium diphtheriae*. Mol. Microbiol. **53:**251-261.

257. **Ton-That, H., S. K. Mazmanian, L. Alksne, and O. Schneewind.** 2002. Anchoring of surface proteins to the cell wall of *Staphylococcus aureus*. Cysteine 184 and histidine 120 of sortase form a thiolate-imidazolium ion pair for catalysis. J. Biol. Chem. **277:**7447-7452.

258. **Ton-That, H., and O. Schneewind.** 1999. Anchor structure of staphylococcal surface proteins. IV. Inhibitors of the cell wall sorting reaction. J. Biol. Chem. **274:**24316-24320.

259. **Ton-That, H., and O. Schneewind.** 2003. Assembly of pili on the surface of *Corynebacterium diphtheriae*. Mol. Microbiol. **50:**1429-1438.
260. **Torrents, E., G. Buist, A. Liu, R. Eliasson, J. Kok, I. Gibert, A. Graslund, and P. Reichard.** 2000. The anaerobic (class III) ribonucleotide reductase from *Lactococcus lactis*. Catalytic properties and allosteric regulation of the pure enzyme system. J. Biol. Chem. **275:**2463-2471.
261. **Torrents, E., R. Eliasson, H. Wolpher, A. Graslund, and P. Reichard.** 2001. The anaerobic ribonucleotide reductase from *Lactococcus lactis*. Interactions between the two proteins NrdD and NrdG. J. Biol. Chem. **276:**33488-33494.
262. **Torrents, E., I. Grinberg, B. Gorovitz-Harris, H. Lundstrom, I. Borovok, Y. Aharonowitz, B.-M. Sjoberg, and G. Cohen.** 2007. NrdR controls differential expression of the *Escherichia coli* ribonucleotide reductase genes. J. Bacteriol. **189:**5012-5021.
263. **Valencia, E., and J. Miró.** 2004. Endocarditis in the setting of HIV infection. AIDS Rev. **6:**97-106.
264. **van Dalen, P., E. van Deutekom-Mulder, J. de Graaff, and T. van Steenbergen.** 1998. Pathogenicity of *Peptostreptococcus micros* morphotypes and *Prevotella* species in pure and mixed culture. J. Med. Microbiol. **47:**135-140.
265. **van Dalen, P., T. van Steenbergen, M. Cowan, H. Busscher, and J. de Graaff.** 1993. Description of two morphotypes of *Peptostreptococcus micros*. Int. J. Syst. Bacteriol. **43:**787-793.
266. **van Wely, K. H. M., J. Swaving, R. Freudl, and A. J. M. Driessen.** 2001. Translocation of proteins across the cell envelope of Gram-positive bacteria. FEMS Microbiol.Rev. **25:**437-454.
267. **Varghese, J., M. Hrmova, and G. Fincher.** 1999. Three-dimensional structure of a barley beta-D-glucan exohydrolase, a family 3 glycosyl hydrolase. Structure **7:**179-190.
268. **Veltrop, M. H., H. Beekhuizen, and J. Thompson.** 1999. Bacterial species- and strain-dependent induction of tissue factor in human vascular endothelial cells. Infect. Immun. **67:**6130-6138.
269. **Veltrop, M. H. A. M., J. Thompson, and H. Beekhuizen.** 2001. Monocytes augment bacterial species- and strain-dependent induction of tissue factor activity in bacterium-infected human vascular endothelial cells. Infect. Immun. **69:**2797-3202.
270. **Vickerman, M. M., S. Iobst, A. M. Jesionowski, and S. R. Gill.** 2007. Genome-wide transcriptional changes in *Streptococcus gordonii* in response to competence signaling peptide. J. Bacteriol. **189:**7799-7807.
271. **Viscount, H. B., C. L. Munro, D. Burnette-Curley, D. L. Peterson, and F. L. Macrina.** 1997. Immunization with FimA protects against *Streptococcus parasanguis* endocarditis in rats. Infect. Immun. **65:**994-1002.
272. **Vlessis, A. A., H. Hovaguimian, J. Jaggers, A. Ahmad, and A. Starr.** 1996. Infective endocarditis: Ten-year review of medical and surgical therapy. Ann. Thorac. Surg. **61:**1217-1222.

273. **von Eichel-Streiber, C., M. Sauerborn, and H. Kuramitsu.** 1992. Evidence for a modular structure of the homologous repetitive C-terminal carbohydrate-binding sites of *Clostridium difficile* toxins and *Streptococcus mutans* glucosyltransferases. J. Bacteriol. **174:**6707-6710.

274. **Voskuil, M. I., and G. H. Chambliss.** 1998. The -16 region of *Bacillus subtilis* and other gram-positive bacterial promoters. Nucleic Acids Res. **26:**3584-90.

275. **Weinreich, M. D., A. Gasch, and W. S. Reznikoff.** 1994. Evidence that the *cis* preference of the Tn5 transposase is caused by nonproductive multimerization. Genes Dev. **8:**2363-2374.

276. **Weiss, W. J., E. Lenoy, T. Murphy, L. Tardio, P. Burgio, S. J. Projan, O. Schneewind, and L. Alksne.** 2004. Effect of *srtA* and *srtB* gene expression on the virulence of *Staphylococcus aureus* in animal models of infection. J. Antimicrob. Chemother. **53:**480-486.

277. **Wells, V. D., C. L. Munro, M. C. Sulavik, D. B. Clewell, and F. L. Macrina.** 1993. Infectivity of a glucan synthesis-defective mutant of *Streptococcus gordonii* (Challis) in a rat endocarditis model. FEMS Microbiol. Lett. **112:**301-305.

278. **Westergren, G., and J. Olsson.** 1983. Hydrophobicity and adherence of oral streptococci after repeated subculture *in vitro*. Infect. Immun. **40:**432-435.

279. **Wexler, D. E., and P. P. Cleary.** 1985. Purification and characteristics of the streptococcal chemotactic factor inactivator. Infect. Immun. **50:**757-764.

280. **Wiegand, T. W., and W. S. Reznikoff.** 1992. Characterization of two hypertransposing Tn5 mutants. J. Bacteriol. **174:**1229-1239.

281. **Wilson, W., K. A. Taubert, M. Gewitz, P. B. Lockhart, L. M. Baddour, M. Levison, A. Bolger, C. H. Cabell, M. Takahashi, R. S. Baltimore, J. W. Newburger, B. L. Strom, L. Y. Tani, M. Gerber, R. O. Bonow, T. Pallasch, S. T. Shulman, A. H. Rowley, J. C. Burns, P. Ferrieri, T. Gardner, D. Goff, and D. T. Durack.** 2008. Prevention of infective endocarditis: Guidelines from the American Heart Association: A guideline from the American Heart Association Rheumatic Fever, Endocarditis and Kawasaki Disease Committee, Council on Cardiovascular Disease in the Young, and the Council on Clinical Cardiology, Council on Cardiovascular Surgery and Anesthesia, and the Quality of Care and Outcomes Research Interdisciplinary Working Group. J. Am. Dent. Assoc. **139:**3S-24.

282. **Wilson, W., K. A. Taubert, M. Gewitz, P. B. Lockhart, L. M. Baddour, M. Levison, A. Bolger, C. H. Cabell, M. Takahashi, R. S. Baltimore, J. W. Newburger, B. L. Strom, L. Y. Tani, M. Gerber, R. O. Bonow, T. Pallasch, S. T. Shulman, A. H. Rowley, J. C. Burns, P. Ferrieri, T. Gardner, D. Goff, and D. T. Durack.** 2007. Prevention of infective endocarditis: Guidelines from the American Heart Association: A guideline from the American Heart Association Rheumatic Fever, Endocarditis and Kawasaki Disease Committee, Council on Cardiovascular Disease in the Young, and the Council on Clinical Cardiology, Council on Cardiovascular Surgery and Anesthesia, and the Quality of Care and

Outcomes Research Interdisciplinary Working Group. J. Am. Dent. Assoc. **138:**739-760.

283. **Wisselink, H., U. Vecht, N. Stockhofe-Zurwieden, and H. Smith.** 2001. Protection of pigs against challenge with virulent *Streptococcus suis* serotype 2 strains by a muramidase-released protein and extracellular factor vaccine. Vet. Rec. **148:**473-477.
284. **Wu, H., K. P. Mintz, M. Ladha, and P. M. Fives-Taylor.** 1998. Isolation and characterization of Fap1, a fimbriae-associated adhesin of *Streptococcus parasanguis* FW213. Mol. Microbiol. **28:**487-500.
285. **Xu, P., J. M. Alves, T. Kitten, A. Brown, Z. Chen, L. S. Ozaki, P. Manque, X. Ge, M. G. Serrano, D. Puiu, S. Hendricks, Y. Wang, M. D. Chaplin, D. Akan, S. Paik, D. L. Peterson, F. L. Macrina, and G. A. Buck.** 2007. Genome of the opportunistic pathogen *Streptococcus sanguinis*. J. Bacteriol. **189:**3166-3175.
286. **Yajima, A., Y. Urano-Tashiro, K. Shimazu, E. Takashima, Y. Takahashi, and K. Konishi.** 2008. Hsa, an adhesin of *Streptococcus gordonii* DL1, binds to alpha2-3-linked sialic acid on glycophorin A of the erythrocyte membrane. Microbiol. Immunol. **52:**69-77.
287. **Yamaguchi, M., Y. Terao, T. Ogawa, T. Takahashi, S. Hamada, and S. Kawabata.** 2006. Role of *Streptococcus sanguinis* sortase A in bacterial colonization. Microbes Infect. **8:**2791-2796.
288. **Yamane, K., K. Bunai, and H. Kakeshita.** 2004. Protein traffic for secretion and related machinery of *Bacillus subtilis*. Biosci. Biotechnol. Biochem. **68:**2007-2023.
289. **Yanisch-Perron, C., J. Vieira, and J. Messing.** 1985. Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. Gene **33:**103-119.
290. **Yother, J., and J. M. White.** 1994. Novel surface attachment mechanism of the *Streptococcus pneumoniae* protein PspA. J. Bacteriol. **176:**2976-2985.
291. **Zagursky, R. J., P. Ooi, K. F. Jones, M. J. Fiske, R. P. Smith, and B. A. Green.** 2000. Identification of a *Haemophilus influenzae* 5'-nucleotidase protein: Cloning of the *nucA* gene and immunogenicity and characterization of the NucA protein. Infect. Immun. **68:**2525-2534.
292. **Zhang, Y., Y. Lei, A. Nobbs, A. Khammanivong, and M. C. Herzberg.** 2005. Inactivation of *Streptococcus gordonii* SspAB alters expression of multiple adhesin genes. Infect. Immun. **73:**3351-3357.
293. **Zhou, M., and W. S. Reznikoff.** 1997. Tn 5 transposase mutants that alter DNA binding specificity. J. Mol. Biol. **271:**362.
294. **Zhou, Y., X. Y. Wan, H. L. Wang, Z. Y. Yan, Y. D. Hou, and D. Y. Jin.** 1997. Bacterial scavengase p20 is structurally and functionally related to peroxiredoxins. Biochem. Biophys. Res. Commun. **233:**848-852.
295. **Zong, Y., Y. Xu, X. Liang, D. Keene, A. Höök, S. Gurusiddappa, M. Höök, and S. Narayana.** 2005. A 'Collagen Hug' model for *Staphylococcus aureus* CNA binding to collagen. EMBO J. **24:**4224-4236.

296. **Zygmunt, M. S., S. D. Hagius, J. V. Walker, and P. H. Elzer.** 2006. Identification of *Brucella melitensis* 16M genes required for bacterial survival in the caprine host. Microbes Infect. **8:**2849-2854.

APPENDIX A

Oligonucleotides for PCR amplification

ARTR-SalI-24495-up ... GCTTTTGGTCGACAATATCGCAGCCAAG
ARTR-23327-dn ..... AGCCATTCATATGCTTGACTGGAGCAGTTTCAAACTTC
ARTR-21212-up ..... GTCAAGCATATGAATGGCT
ARTR-SalI-20031-dn ... TGGTGCGTCGACTGGCTTCGTTGTCTTC
ARTR-EcoRI-up1 ..... AAAACAGTCAGCCGAATTCTATTT
ARTR-EcoRI-up2 ..... AGGTTTAAAGAGAATTCATTTGTCCTAAGGG
ARTR-EcoRI-dn ..... TTCGAATTCTGGTCGTCTCAGCTCC
spcUP ..... CCGCTCTAGAACTAGTGGATCC
spcDO ..... CAATTTTTTTATAATTTTTTTAATCTG
DAM303 ..... AAGGGCCCGTTTGATTTTTAATG
DAM347 ..... CCGAATTCTAGGTACTAAAACAATTCATCCAGTAA
0169-SalI-up ..... CCGGTCGACTGTCTATGGTGGACAGCGTCC
0169-SalI-dn ..... CCGGTCGACTCGTGACGGAAGCAGACTTGG
pSerm-NcoI-up ..... CATGCCATGGCCGGGCCCAAAATTTGTTTGAT
pSerm-NcoI-dn ..... CATGCCATGGAGTCGGCAGCGACTCATAGAAT
1069RE-up ..... CATGGGTGACGCGGCCGCTGCAGGCGCGCCTGCGAACA
1069RE-dn ..... CATGTGTTCGCAGGCGCGCCTGCAGCGGCCGCGTCACC
Sc-NotI-up ..... CCGGCGGCCGCCCGCTCTAGAACTAGTGGATCC
Sc-NcoI-dn ..... TCTCGACCATGGCAATTTTTTTATAATTTTTTTAATCTG
Tet-NotI-Up ..... CAGTCGGCGGCCGCCTAGGTTGATTTTCGTTC
Tet-NcoI-Dn ..... TCTCGACCATGGCTAGGTCACTAAGTTATTTT
TetMF1 ..... TGTGTGTGACGAACTTTACCGA
TetMF2 ..... TCTGCAAAAGATGGCGTACAAGC
6-26-BamHI-up ..... AAGCGGATCCAGTACGCCTGGGCGAAGACC
6-26-BamHI-dn ..... TGGCGGATCCTTTCCCACGCGCAGAGGGAC
srtA1 ..... GTAATAGCGAAACGGACGCAGAAGC
srtA2 ..... GGATCCACTAGTTCTAGAGCGGAAAGATTAAGCCTAGGGCCAC
srtA3 ..... CAGATTAAAAAAATTATAAAAAAATTGGTCGAGTATGACAAGGCT CCT
srtA4 ..... TGAGGGTGTCAAAGAAGTTGGCGGA
srtB1 ..... TCATCCAGTCTTATCTGGTCCTGCT
srtB2 ..... CATTAAAAATCAAACGGGCCCTTGCAAATGCCAACAACAAAGAGC TGT
srtB3 ..... TTACTGGATGAATTGTTTTAGTACCTAGAATTCGGGATTTTCTGGA CAAGCATGGCAAG
srtB4 ..... AATCGTCTCCATACTAGCCGCATAG

Oligonucleotides for PCR amplification

0169-2	CTTCCCAGTTGTAATCCAAGA
0169-3	TTACTGGATGAATTGTTTTAGTACCTAGAATTCGGATATTTCTTATCATCGGGTTT
0169srtAP	TCTTGGATTACAACTGGGAAGGCGATGTAAACGTTTTTACAAACGGTTCAAAATACTGAAAAATTAAAGAAAATGTCGAATCGTAAAAGAGGACCTC
kansrtA	CATTAAAAATCAAACGGGCCCTTCGGAAAAAGAGTCAAAATCAGCAAG
srtA-EcoRI-up	AAGCGAATTCTAGCTAGTGAGTATC
srtA-EcoRI-dn	GTTAGAATTCTTTCTAACCTTTGGAGTG
srtA-SphI-up	GGACGCATGCATGGTACAAGGA
srtA-SphIdn	GAGTGCATGCTAACCTTTGGAGTG
pJFP52srtA-up	CCGGCTCGTATGTTGTGTGG
pJFP52EcoRI-dn	GTCTGAATTCTTACGCATGCTAACCTTTGG

Oligonucleotides for transcript detection

0169-F	ATGAAGCAAAAACCCGATGA
0169-R	TTTTCTGGAAGCTGTTCTGGA
1220-IG-1219 F	TTTACAAGCGTCGAAGCAGA
1220-IG-1219 R	AAGCCTAGGGCCACAAAAAT
1219-F (*srtA*)	GGCATCACAGGGGCTAGTAA
1219-R (*srtA*)	CGCAGCATCCTCACAAGTAA
16S-F	TTCGGATCGTAAAGCTCTGTTG
16S-R	GTAGTTAGCCGTCCCTTTCTGG

Oligonucleotides for screening and sequencing

Jxn-up	GTCCTGCTTTATGCTCGACTG
Jxn-dn	CGCATCCAGTAAAGTGCTGG
Mout	CCGGGGACTTATCAGCCAACC
PB1	GATCTCTACAACCTCAAGC
PB2	TCCATTCTAACCAAG
PB11	TCTACAACCTCAAGC
PB12	TCCATTCTAACCAAG
spc246	TCAATGGTTCAGATACGACGACT
spc638	GCCGTATGATTTTAACTATGGAC
kan306R	AGCCATCATGCCGTTCAAAGTGC
kan637F	CACGGCGACCTGGGAGACAGCAA
CP25	CTTTGGCAGTTTATTCTTGA
gfpmut	TTGTGCCCATTAACATCACC
pVA838-2245	AGGCGTTTAAGGGCACCAAT
pVA838-2939	GTAACACGCCACATCTTG
pVA838-5331	CCAATGGCATCGTAAAGAAC

APPENDIX B

Oligonucleotides for directed signature-tagged mutagenesis

SSA_1112-up.....TGACGATGGCAGCAGTGACTC
SSA_1112-dn.....ACACTCCCCGATACGAGCCTC
SSA_2023-up.....AGTTTCAAATCACGGAGCCGCTGC
SSA_2023-dn.....ACGTAGTAAGTCATGGCCGCA
SSA_1663-up.....TTTGGGAGAAATTGAGGCCGCT
SSA_1663-dn.....ACAAACCCAGTAGGAGCAAGGCTC
SSA_1591-up.....GTGTCAACCGTTGCCAATCAC
SSA_1591-dn.....TCTTTGGATAGCTGGCGGAGC
SSA_0805-up.....GCTGTGACAGCAGTTGGTG
SSA_0805-dn.....TCAGGGTACGCAGGCCCTC
SSA_1634-up.....GACGGTTACACCGGATTCC
SSA_1634-dn.....CTTGTTCGCCCAGCAGCAG
SSA_1633-up.....CGGGACAACACCCCAAAGC
SSA_1633-dn.....AGCTTTGACCTCCGCAGAG
SSA_1632-up.....AGTGCTACCTTAGGAACGACC
SSA_1632-dn.....TCCCTTGACCGCCAACTGG
SSA_0956-up.....CCTTGGATGAACTGTTAGCGAC
SSA_0956-dn.....TGTCAAGACGGCCTGGGGCTC
SSA_0227-up.....CAAGAAGATGACGGAGGAAACG
SSA_0227-dn.....AAAGATGGTTACACACCGAGC
SSA_0565-up.....GCTCAACAGGTCGTCCAAACC
SSA_0565-dn.....TGTGCCTGAGCCAGTTGCTGG
SSA_1301-up.....TGGTCGGACTTTTGACCAAGC
SSA_1301-dn.....TGCTGCTAGAACTGCCTGAGC
SSA_0303-up.....TGATGAAGCCTTGACTGGTGC
SSA_0303-dn.....CCAAGCCTTATCATTGCGGAATTG
SSA_0273-up.....CCCATACTCCGATGCTGGACG
SSA_0273-dn.....AAAGTCCATCTCAAATGGCTCAG
SSA_1023-up.....TCTATGGAGACGGTCCTGACA
SSA_1023-dn.....CAAAGCGAGTCATTTCTAATACTGG
SSA_1019-up.....TCGCTCTTATGCAGACATGC
SSA_1019-dn.....TGAACGGCAGCACCGTCTG
SSA_1631-up.....CGATGATGCTGTTATGGGTGGTCT
SSA_1631-dn.....TCCTACTTTAGATGGCGGTGCT
SSA_0905-up.....CGATTCGGAAACTAAACGTCGGTG
SSA_0905-dn.....CTGCCTCATTAATTGGCACACG
SSA_0906-up.....AGGATTGGATTCACGAGGCGGACTC

Oligonucleotides for directed signature-tagged mutagenesis

SSA_0906-dn.....TGACCGTCACCGAGGCTGGGGTTC
SSA_1219-up.....TTAATGTCCAGCAAGACCGCAGTG
SSA_1219-dn.....CCAGTATGATAGATGTCGCCAG
SSA_0684-up.....ACTTTGAAGTTCGTGAGCGTGCAG
SSA_0684-dn.....TTTAGCCGGCTGTGTCGGGGTGGCT
SSA_1234-up.....TCGCTCAGGATGGTCCTGC
SSA_1234-dn.....TCAAGGCAACTGCTGATTCG
SSA_0022-up.....TTTCACTTTTCGCGGCTGCTG
SSA_0022-dn.....AGCCGACAAGGCCTCTGTTGCGCCA
SSA_1063-up.....AGGTTCCACATCTAACGGACGTG
SSA_1063-dn.....AGGGTAAATACTACAGGGCGGACC
SSA_2020-up.....TTAACGTTGAGCACTTGGCCGA
SSA_2020-dn.....ATCAACCTGACTGGTCCGAGCAGC
SSA_0167-up.....AGCCTCTACTTCTAGCCCCATGG
SSA_0167-dn.....GGCTGATAGAATAAGCAGCTGCAC
SSA_0146-up.....AGGTAAAACGCGACCTGTC
SSA_0146-dn.....AGAGCCGCCATTTGAGGAGCGCT
SSA_0453-up.....TGCTCAATCTGGTCTCCGGTGTGGT
SSA_0453-dn.....GCACCGACAGATCCAGCCAAGCGCA
SSA_1666-up.....TATTGACCAAGGCGGAACTCC
SSA_1666-dn.....ATAGTCCGGATTCAGCGCCAG
SSA_1065-up.....CCCAACCGGTAGCGTTGGTTG
SSA_1065-dn.....TGATACCATCCTTGATGAGCGGCAC
SSA_0904-up.....TCCTTTTAGGTACTGCGGCTACAG
SSA_0904-dn.....ACCTTGAATGCCCTCGCTGCTGGCA
SSA_1750-up.....AGAAGGTAGCTATCAGCCAGC
SSA_1750-dn.....TTGAGACGGTTGGGAGCTGTC
SSA_2121-up.....TTCCTTTATGGTTGCGGCCAG
SSA_2121-dn.....TCCAACCGTTAGCAGCCGTCA
SSA_0243-up.....AGACATAGTCATGGCGAACCTC
SSA_0243-dn.....TGGTTGCTCAATAGGAACCGGT

VITA

Lauren Senty Turner hails from Arlington, VA, where she completed her secondary education at Yorktown High School, graduating in 1998. Between 1998 and 2002 she was enrolled at Virginia Tech. Her molecular biology experience was encouraged and fostered by Dr. Ann Stevens, an Associate Professor in the Department of Biological Sciences. After receiving her Bachelors of Science degree from Virginia Tech in May, 2002 she continued working with Dr. Stevens as a laboratory technician. In the spring of 2003 Lauren embarked on a 2,000 mile journey, hiking the Appalachian Mountain chain of the United States. In this time two events occurred that would shape her future years. The first was meeting her husband, Luke Turner. The second was notification of her acceptance in the doctoral program of the Department of Microbiology and Immunology, at Virginia Commonwealth University (VCU). Lauren matriculated at VCU in August 2003, and began her doctoral research with Dr. Todd Kitten in 2004.

AWARDS AND GRANTS:

2001..... Recipient of Biological Sciences Initiative Grant, Virginia Tech
2006..... ASM Student Travel Grant, 106th ASM General Meeting
2007..... Student Travel Grant, Mid-Atlantic Microbial Pathogenesis Meeting
2007..... AAAS/Science Program for Excellence in Science
2007..... 2nd Place Award, Graduate Research Presentation, VA ASM
2008..... VCU Graduate School Dissertation Assistantship
2008..... 2nd Place Award, AADR/Johnson & Johnson Oral Health Hatton Competition
2008..... ASM Student Travel Grant, 108th ASM General Meeting
2008..... 2nd Place Award, IADR/Unilever Hatton Competition

SOCIETIES:

2005-Present.. American Society of Microbiology (ASM)
2007-Present.. Phi Kappa Phi Honor Society
2007-2008 ..... American Association for the Advancement of Science (AAAS)

PUBLICATIONS:

1. **McDowell, J. V., J. Wolfgang, L. Senty, C. M. Sundy, M. J. Noto, and R. T. Marconi.** 2004. Demonstration of the involvement of outer surface protein E coiled coil structural domains and higher order structural elements in the binding of infection-induced antibody and the complement-regulatory protein, Factor H. J. Immunol. **173**:7471-7480.

2. **Paik, S., L. Senty, S. Das, J. C. Noe, C. L. Munro, and T. Kitten.** 2005. Identification of virulence determinants for endocarditis in *Streptococcus sanguinis* by signature-tagged mutagenesis. Infect. Immun. **73**:6064-6074.

.