

# A role for glycosylated serine-rich repeat proteins in Gram-positive bacterial pathogenesis

A. Lizcano\*, C.J. Sanchez\* and C.J. Orihuela

Department of Microbiology and Immunology, The University of Texas Health Science Center at San Antonio, San Antonio, TX, USA

**Correspondence:** Carlos J. Orihuela, Associate Professor, Department of Microbiology and Immunology, The University of Texas Health Science Center at San Antonio, 7703 Floyd Curl Drive, San Antonio, TX 78229-3900, USA Tel.: +1 210 567 3973; fax: +1 210 567 6612; E-mail: orihuela@uthscsa.edu

\*These two authors are co-first authors.

**Keywords:** adhesion; biofilm; Gram-positive bacteria; pathogenesis; serine-rich repeat proteins

**Accepted 3 May 2012**

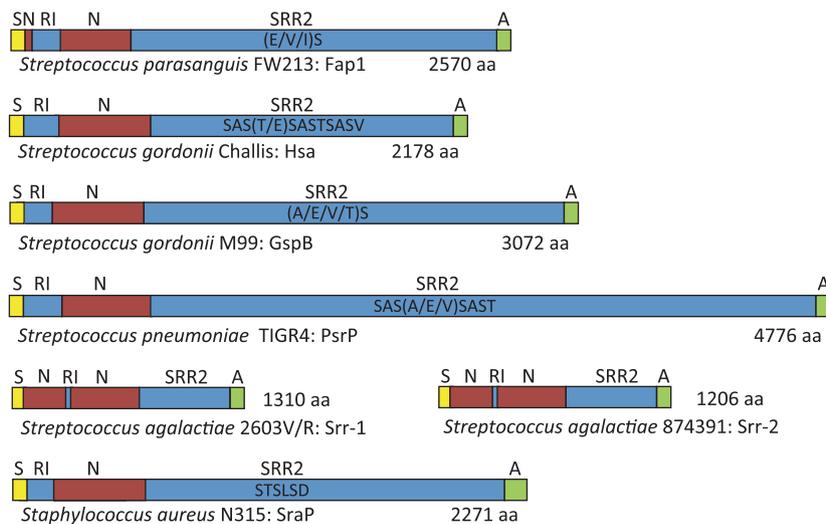
DOI: 10.1111/j.2041-1014.2012.00653.x

## SUMMARY

**Bacterial attachment to host surfaces is a pivotal event in the biological and infectious processes of both commensal and pathogenic bacteria, respectively. Serine-rich repeat proteins (SRRPs) are a family of adhesins in Gram-positive bacteria that mediate attachment to a variety of host and bacterial surfaces. As such, they contribute towards a wide-range of diseases including sub-acute bacterial endocarditis, community-acquired pneumonia, and meningitis. SRRPs are unique in that they are glycosylated, require a non-canonical Sec-translocase for transport, and are largely composed of a domain containing hundreds of alternating serine residues. These serine-rich repeats are thought to extend a unique non-repeat (NR) domain outward away from the bacterial surface to mediate adhesion. So far, NR domains have been determined to bind to sialic acid moieties, keratins, or other NR domains of a similar SRRP. This review summarizes how this important family of bacterial adhesins mediates bacterial attachment to host and bacterial cells, contributes to disease pathogenesis, and might be targeted for pharmacological intervention or used as novel protective vaccine antigens. This review also highlights recent structural findings on the NR domains of these proteins.**

## INTRODUCTION

Serine-rich repeat proteins (SRRPs) are a family of large, surface-exposed proteins exclusive to Gram-positive bacteria and are predominately found within the oral streptococci (Wu *et al.*, 1998; Wu & Fives-Taylor, 1999; Froeliger & Fives-Taylor, 2001; Tettelin *et al.*, 2001; Baba *et al.*, 2002; Bensing & Sullam, 2002; Glaser *et al.*, 2002; Stephenson *et al.*, 2002; Takahashi *et al.*, 2002; Zhang *et al.*, 2003; Handley *et al.*, 2005; Lauer *et al.*, 2005; Takeuchi *et al.*, 2005; Vickerman *et al.*, 2007; Xu *et al.*, 2007; Rigel & Braunstein, 2008; Zhou *et al.*, 2008; Zhou & Wu, 2009). Scanning electron microscopy suggests that SRRPs form large stalks or fimbriae-like structures, similar to the recently characterized Gram-positive pilus (Lauer *et al.*, 2005), that extend outward from the bacterial surface to mediate adhesion (Wu *et al.*, 1998; Wu & Fives-Taylor, 1999). Defining features of SRRPs include two serine-repeat regions (SRR), one of which comprises the majority of the protein, one or two species-unique non-repeat regions (NR domain) near the N-terminal domain that mediate adhesion, and a C-terminal cell wall anchor domain (Fig. 1). The SRRs are typically composed of alternating serine residues, most frequently separated by an alanine, valine, or threonine. Importantly, there are O-linked glycoconjugates attached to the SRRs, this modification occurs in the cytoplasm and is required for protein stability (Bensing *et al.*, 2004; Takamatsu *et al.*, 2004a). Export of SRRPs onto the bacterial



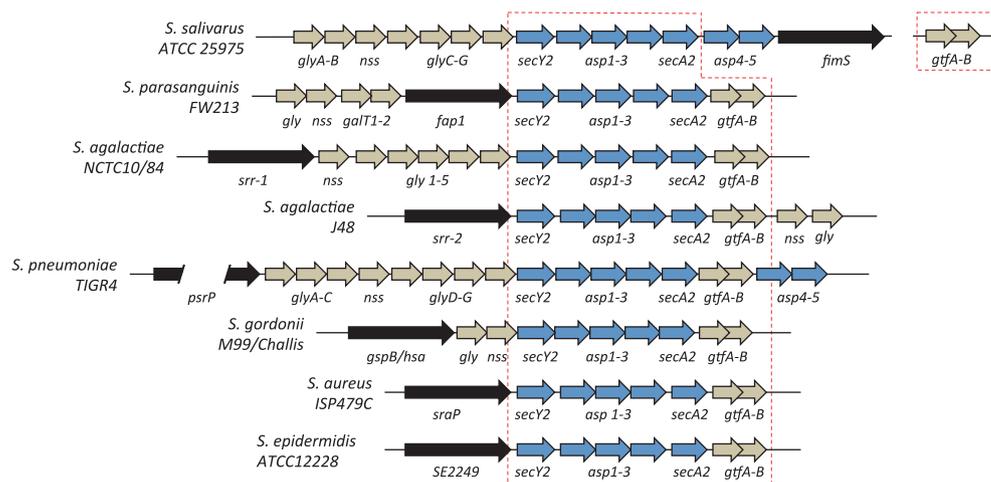
**Figure 1** Structural domain organization of select serine-rich repeat proteins (SRRPs). SRRPs are composed of an N-terminal signal peptide (S); most often a single, but sometimes two, unique non-repeat domains (NR) that mediate adhesion; two serine-rich repeat domains (SRR) flanking the single or second N domain (SRR1 and SRR2); and a C-terminal cell wall anchor domain (A). The SRRs are composed of alternating serine residues most frequently, an alanine, valine, or threonine residue. The large number of repeats in the SRR2 domain is responsible for the large size of most SRRPs. Domains are not drawn to scale.

surface occurs through a dedicated non-canonical Sec translocase, Sec-Y2A2, following recognition of an extended atypical signal sequence peptide at the N-terminus (Bensing & Sullam, 2002, 2009, 2010; Bensing *et al.*, 2004, 2005, 2007; Takamatsu *et al.*, 2004a,b, 2005b; Rigel & Braunstein, 2008; Siboo *et al.*, 2008; Mistou *et al.*, 2009; Yen *et al.*, 2011).

Genes encoding each SRRP, its specific glycosyltransferases, and the required accessory components of an alternate *secY2A2* translocase are found on genomic islands present in the chromosome. These islands vary considerably in size, with the largest being *psrP-secY2A2* in *Streptococcus pneumoniae*, at 37-kb (Bensing & Sullam, 2002; Takahashi *et al.*, 2002, 2004; Chen *et al.*, 2004, 2006; Handley *et al.*, 2005; Siboo *et al.*, 2005; Obert *et al.*, 2006; Seifert *et al.*, 2006; Wu *et al.*, 2007a; van Sorge *et al.*, 2009); in at least one case, the SRRP locus is considered to be a pathogenicity island (Obert *et al.*, 2006). The core genetic component of these loci, found in all strains that carry an SRRP, includes two genes (most commonly called *GtfA* and *GtfB*) encoding glycosyltransferases involved in the first step of SRRP glycosylation, which is required for protein stability, and five genes encoding components of the alternate *secY2A2* complex necessary for SRRP transport [*secA2*, *secY2*, *Asp1–3* (also known as *gap1–3*); the latter may also be involved in glycosylation (Peng *et al.*, 2008a,b)] (Fig. 2) (Bensing *et al.*, 2004; Wu *et al.*, 2007a; Bu *et al.*, 2008; Li *et al.*, 2008; Siboo *et al.*, 2008; Zhou & Wu, 2009; Seepersaud *et al.*, 2010; Zhou *et al.*, 2010, 2011, 2012; Zhu

*et al.*, 2011). Species and strain-specific variation does occur and can include the presence of additional glycosyltransferases, as many as five in the *psrP-secY2A2* island in *S. pneumoniae*, or additional *secY2A2* accessory components, as in the Fap-1 locus in *Streptococcus parasanguinis*. Consequently, differences in the complexity of SRRP glycosylation are presumed to occur and most likely impact the individual activities of these proteins. SRRP biogenesis and glycosylation will not be expanded upon here because an excellent and extensive review on these topics was written recently by Zhou & Wu (2009).

After the initial description of Fap1 (fimbriae-associated protein) by Wu *et al.* (1998) in *S. parasanguinis*, SRRPs were identified in numerous medically relevant streptococci. This includes assorted oral streptococci, the neonatal pathogen *Streptococcus agalactiae*, and the respiratory tract pathogen *S. pneumoniae* (Plummer *et al.*, 2005; Seifert *et al.*, 2006; Xiong *et al.*, 2008; Mistou *et al.*, 2009; Nobbs *et al.*, 2009; Shivshankar *et al.*, 2009). The SRRPs have also been identified in *Staphylococcus aureus*, *Staphylococcus saprophyticus* and *Staphylococcus epidermidis* (Zhang *et al.*, 2003; Takeuchi *et al.*, 2005), as well as in other less virulent bacteria including *Lactococcus johnsonii* (Zhou & Wu, 2009). Importantly, it is now understood that SRRPs function both as intra- and inter-species adhesins, facilitating colonization of host tissue through the formation of biofilms, and that they are involved in the development of tissue infections such as oral infectious diseases, infective endocarditis, pneumonia, and



**Figure 2** Genetic organization of loci encoding serine-rich repeat proteins (SRRPs) in streptococcal and staphylococcal species. SRRPs are encoded on genomic islands present in the chromosome. In addition to the SRRP (black arrow), these loci also contain numerous genes encoding enzymes responsible for the O-linked glycosylation of the protein (tan arrows) and for transport of the protein (blue arrows). No obvious regulatory elements are present. All SRRP loci have a core set of eight genes consisting of an SRRP and *secY2*, *secA2*, *asp1-3* and *gtfA-B* (red dashed line).

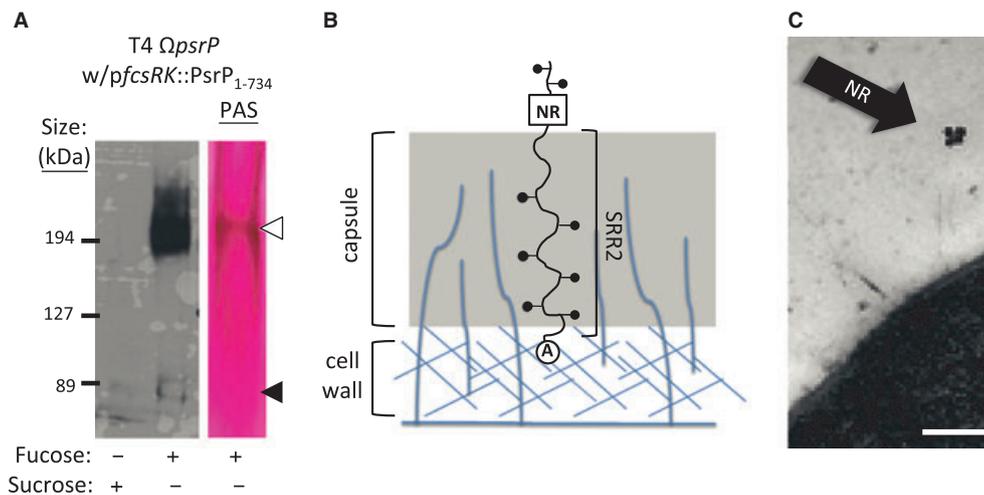
meningitis (Handley *et al.*, 2005; van Sorge *et al.*, 2009; Sanchez *et al.*, 2010). The focus of this review will be to summarize how this important family of bacterial adhesins mediates bacterial attachment to host and bacterial cells, contributes to disease pathogenesis, and might be targeted for pharmacological intervention or used as novel protective vaccine antigens.

### MODEL FOR SRRP FUNCTION

The number and amino acid composition of serine repeats in the SRR2 region is highly variable and is the primary reason why these proteins are so diverse in size. For example, in *S. pneumoniae* strain TIGR4, PsrP is 4776 amino acids long with 539 repeats of the amino acid sequence SAS[A/E/V]SAS[E/T] in its SRR2 domain; it has a predicted non-glycosylated molecular weight of 512 kDa. In contrast, Srr-1 of *S. agalactiae* strain A909 is 1259 amino acids in size with an SRR2 domain containing only 103 SASAS[T/M] repeats. Because these proteins are glycosylated they migrate at sizes much larger than their predicted molecular weight when separated by sodium dodecyl sulfate–polyacrylamide gel electrophoresis (Fig. 3A). Based on the conserved C-terminal location of the cell wall anchor domain for SRRPs and images of Fap1 taken by transmission electron microscopy (Wu

*et al.*, 1998; Wu & Fives-Taylor, 1999), it has been proposed that the primary function of the SRR2 domain is to extend the NR domain, which has adhesive properties (discussed below), beyond the cell wall and other surface components to be able to mediate attachment (Fig. 3B) (Ramboarina *et al.*, 2010; Pyburn *et al.*, 2011). The best experimental evidence supporting this hypothesis is with PsrP of *S. pneumoniae*, where complementation of a *psrP* deletion mutant with a truncated version having only 33 of its normal 539 SRRs in its SRR2 domain was unable to restore bacterial adhesion to lung cells. However, the same truncated construct restored adhesion if transformed into an unencapsulated *psrP*-deficient mutant. This indicates that the length of the SRR2 domain affects its ability to avoid obstruction by the capsular polysaccharide (Shivshankar *et al.*, 2009). Likewise, immunogold electron microscopy of the NR domain of PsrP found the presence of gold particles located distally from the bacterial cell wall surface (Fig. 3C).

Such a model suggests that the number of repeats in the SRR2 domains may have evolved in a manner that optimally matches the span of extracellular components such as capsule for each species and strain. This model also suggests that the exposed SRR1 domain may have an unknown functional role mediated by its attached glycoconjugates. Finally, SRRPs



**Figure 3** Serine-rich repeat proteins (SRRPs) are glycosylated and extend the non-repeat (NR) domains outward from the bacterial surface. (A) Western blot and carbohydrate staining of whole cell lysates of a *Streptococcus pneumoniae* mutant deleted of its chromosomal PsrP but complemented with a plasmid (pfcsRK::PsrP<sub>1-734</sub>) that expresses a truncated version of the protein with only 33 SRR repeats when grown in THY media supplemented with 1% fucose. Carbohydrate staining used a Periodic Acid-Schiff (PAS) stain. Note the presence of a faint fucose-inducible band at 80 kDa that is the predicted unglycosylated size of the recombinant protein (black triangle), and a stronger band at 210 kDa (white triangle). Only the 210-kDa band tested positive for carbohydrates by PAS, suggesting that PsrP is glycosylated. Taken from Shivshankar *et al.* (2009). (B) Hypothetical model showing how SRRPs might employ the SRR2 domain to extend the NR domain outward away from the cell through surface components such that it can mediate adhesion. (C) Immunogold electron microscopy using antibodies against the NR domain of PsrP. Note that the gold particles are found distal to the bacterial surface, supporting the model shown in (B). These findings were first described in Shivshankar *et al.* (2009).

might serve as 'tent-poles' and stabilize/interact with capsule or other loosely attached surface components on the bacterial surface. The last two proposed functions may explain why loci encoding SRRPs such as Fap1 encode additional glycosyltransferases and SRRPs can be glycosylated in a manner that exceeds the minimum required for protein stability. Although these concepts are merely speculative, they are areas of extreme interest because of the possible contribution of SRRPs to Gram-positive biology and pathogenesis.

#### THE NR DOMAINS OF SRRPS ARE HIGHLY VARIABLE BUT MEDIATE ATTACHMENT TO EITHER CONJUGATED SIALIC ACID OR KERATIN

The overall domain structures of SRRPs are conserved between species, but the individual NR domains are highly diverse with limited to no homology when compared by BLAST analysis (Zhou & Wu, 2009). The NR domains also vary considerably in size with the largest, 1231 amino acids, and smallest, 129 amino acids, known versions present in the same strain of *L. johnsonii* strain. In general, most

NR domains are 250–500 amino acids in length with a net charge that is either strongly acidic or basic (Zhou & Wu, 2009). Currently, it is not known if the charge of the NR domain directs the binding properties of the protein.

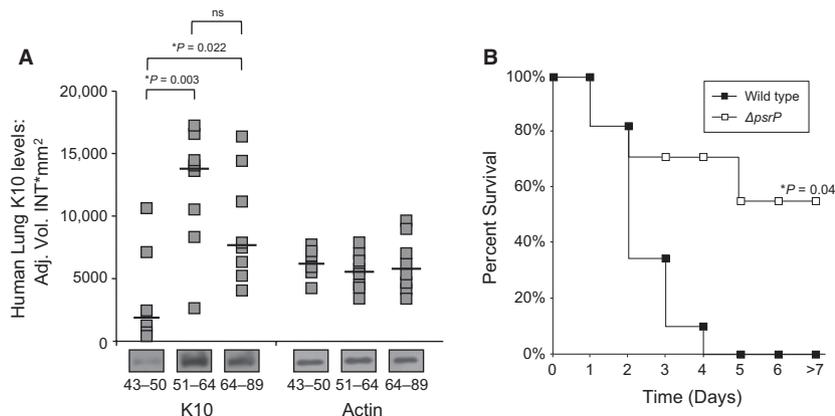
The NR domain in GspB of *Streptococcus gordonii* strain M99 has been shown to bind only to sialyl-T antigen [NeuAc $\alpha$ (2-3)Gal $\beta$ (1-3)GalNAc]. In contrast, the NR in Hsa of *S. gordonii* strain Challis binds both  $\alpha$ (2-3) sialyl-lactosamine [NeuAc $\alpha$ (2-3)Gal $\beta$ (1-4)GlcNAc] and sialyl-T antigen (Takamatsu *et al.*, 2005a). Importantly, the ability of these NR domains to attach to glycoconjugates enables them to mediate bacterial attachment to cell surfaces, primarily platelet membranes and dental surfaces that carry sialyl-lactosamine and sialyl-T antigen respectively. This binding, in some instances, has been clearly shown to contribute to disease. For example, GspB and Hsa has have been shown to bind to salivary mucin MG2 and salivary agglutinin (Takamatsu *et al.*, 2006), and an *S. parasanguinis* mutant deficient in Fap1 fails to attach to saliva-coated hydroxyapatite, an *in vitro* tooth surface model (Fachon-Kalweit *et al.*, 1985; Wu & Fives-Taylor, 1999; Stephenson *et al.*, 2002). In

addition, attachment of GspB and Hsa to sialyl-T antigen on platelet membrane GPIIb $\alpha$  is thought to play a major role in the ability of these oral bacteria to colonize heart valves and cause endocarditis. Indeed, when wild-type *S. gordonii* and *Staphylococcus aureus* are injected intravenously through a catheter into rats, they are both able to form aortic vegetations with platelets in the heart; whereas mutants deficient in GspB and SraP, respectively, had an absence of aortic vegetations (Siboo *et al.*, 2005; Xiong *et al.*, 2008).

For *S. pneumoniae*, the pathogenicity island encoding PsrP, *psrP-secY2A2*, has been shown to be positively correlated with the propensity of a strain to cause pneumonia (Obert *et al.*, 2006). One reason for this is that the NR domain of PsrP binds to surface exposed keratin 10, which is present on lung epithelial cells (Shivshankar *et al.*, 2009). Interestingly, PsrP is distinct from GspB in that binding of its NR to keratin 10 is independent of any lectin properties its NR domain might have. In fact, treatment of lung cells with glycosidases enhanced bacterial adhesion to lung cells, and recombinant NR was able to bind unglycosylated recombinant keratin 10 (Shivshankar *et al.*, 2009). More recently, we have demonstrated that keratin 10 is elevated on lung cells that have experienced DNA damage and are senescent, such as in older individuals (Fig. 4A). Such non-replicating cells are commonly found in the lung of elderly

individuals, smokers, or those who have developed chronic obstructive pulmonary disease (Tsuji *et al.*, 2010). We have also demonstrated that the increased susceptibility of aged mice to *S. pneumoniae* is in part PsrP-dependent (Shivshankar *et al.*, 2011) (Fig. 4B). Hence, enhanced susceptibility to pneumococcal pneumonia for those who smoke and are >65 years of age (two recognized at-risk populations), is partly the result of an increase in PsrP/keratin 10 interactions.

*Streptococcus agalactiae*, which colonizes the genital tract of women, is the leading cause of neonatal sepsis (Glaser *et al.*, 2002); neonates become infected during passage through the birth canal. Consistent with colonization pattern, Srr-1 from *S. agalactiae* binds to keratin 4, which is present on the stratified epithelia lining the vaginal tract (Leube *et al.*, 1988). In addition, incubation of HEp-2 epithelial cells with the NR domain of Srr-1 blocked *S. agalactiae* adherence to cells in a dose-dependent fashion (Samen *et al.*, 2007). However, it is not clear if Srr-1 binding to keratin 4 is glycoconjugate dependent or independent. Separately, studies with an Srr-1 deletion mutant in a mouse model implicated this protein in the ability of *S. agalactiae* to penetrate the blood–brain barrier (van Sorge *et al.*, 2009). As keratin is not expressed by endothelial cells that compose the vasculature (Gu & Coulombe, 2007), one must speculate that this host–cell interaction

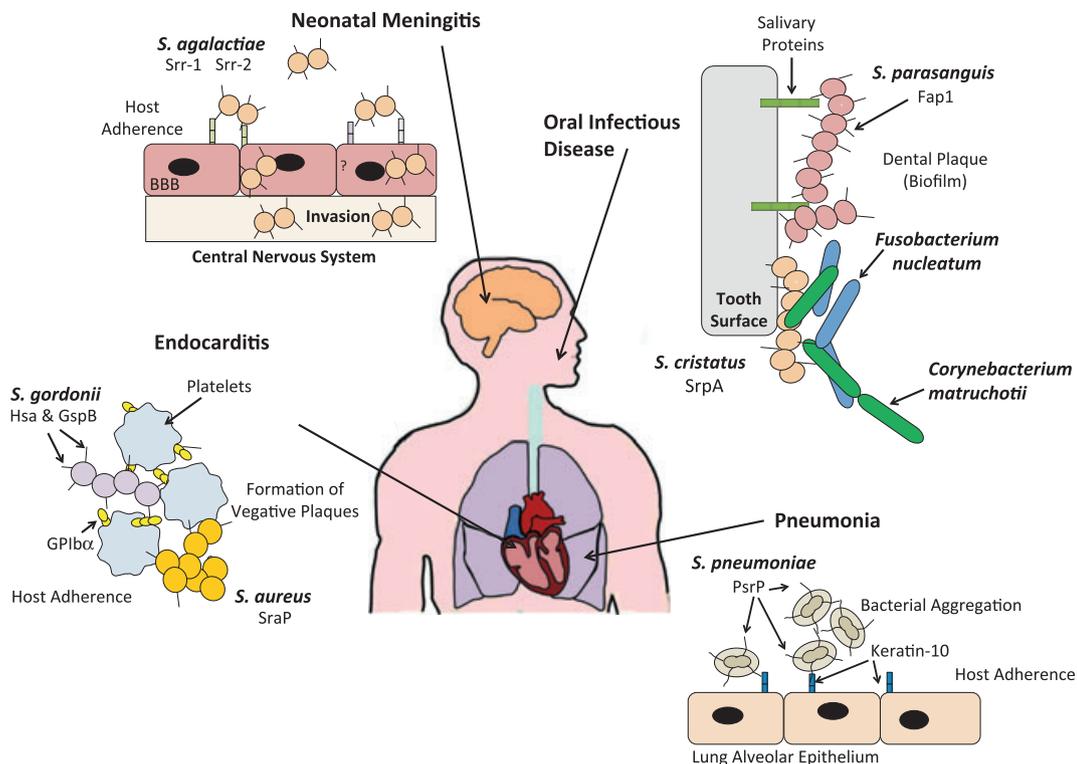


**Figure 4** Aging enhances susceptibility to pneumococcal pneumonia in a K10-dependent manner. (A) Densitometric analyses and representative Western blot bands for K10 in human lung biopsies obtained from young (43–50 years;  $n = 6$ ), mature (51–64 years;  $n = 8$ ), and aged (64–89 years;  $n = 8$ ) humans. Asterisks denote a statistically significant difference when using one-way analysis of variance (Duncan's Method;  $P < 0.05$ ). (B) Kaplan–Meier plot showing percent survival of aged BALB/c mice intranasally challenged with  $10^5$  CFU of wild-type *Streptococcus pneumoniae* (Wild-type:  $n = 8$ ) or an isogenic mutant deficient in the serine-rich repeat protein (SRRP) PsrP ( $\Delta psrP$ :  $n = 11$ ). Statistical analysis was performed using a Kaplan–Meier log-rank test. These findings were first described in Shivshankar *et al.* (2011).

might be limited to the cuboidal epithelial cells of the choroid plexus or that another host ligand is involved. Interestingly, serotype III multi-locus sequence type clone ST-17 of *S. agalactiae*, which is associated with a propensity for invasive disease and meningitis, has been found to carry an SRRP named Srr-2; although the ligands for Srr-2 also remain unknown (Seifert *et al.*, 2006; van Sorge *et al.*, 2009). Interestingly, *S. pneumoniae* is also thought to penetrate the blood–brain barrier (BBB) through the choroid plexus, although a role for PsrP in pneumococcal meningitis has not been described. Figure 5 summarizes the functions of SRRPs that have so far been demonstrated to play a role in human disease.

### SRRPS MEDATE INTRA-SPECIES AGGREGATION

Initial studies with Fap1 demonstrated that the protein was required for *S. parasanguinis* attachment to and robust biofilm formation on glass slides or saliva-coated hydroxyapatite (Wu *et al.*, 1998, 2007b). Studies with SrpA of *Streptococcus cristatus* determined that the protein was required for the formation of corn-cob structures with *Corynebacterium matruchoi* and *Fusobacterium nucleatum* (Handley *et al.*, 2005) (Fig. 5). Only recently has it been fully appreciated that SRRPs might serve as intra- or inter-species adhesins for non-dental plaque Gram-positive bacteria.



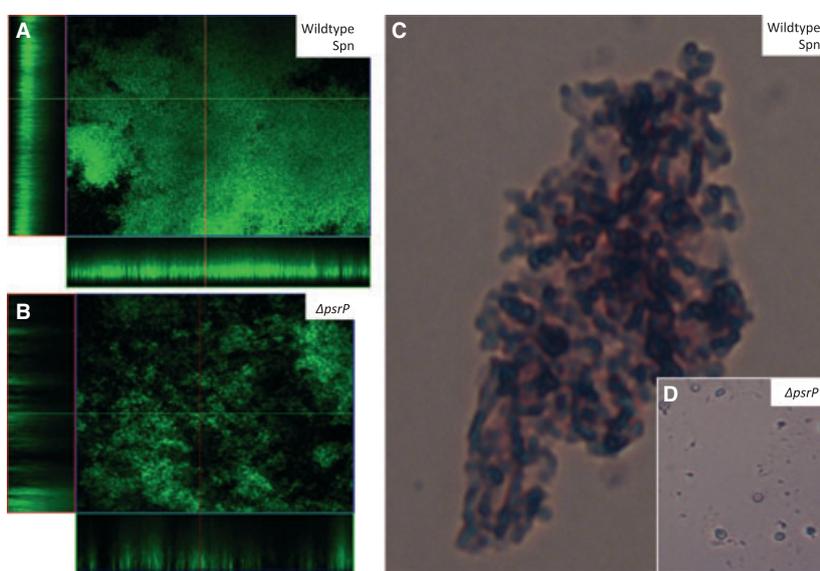
**Figure 5** Serine-rich repeat proteins (SRRPs) expressed in medically relevant bacteria play a role in pathogenesis. PsrP in *Streptococcus pneumoniae* is an intra-species adhesin that initiates attachment to the host by binding to keratin 10 expressed on lung alveolar epithelium. Likewise, it can also mediate attachment to another PsrP molecule expressed on another pneumococcus, facilitating bacterial aggregation, leading to pneumonia. GspB and Hsa are present on *Streptococcus gordonii* M99 and challis, respectively. These SRRPs bind to sialic acid moieties present on the platelet receptor GPIIb $\alpha$  forming vegetative plaques causing infective endocarditis. SraP present in *Staphylococcus aureus* mediates direct binding to platelets also forming vegetative plaques, and leading to infective endocarditis. Neonatal sepsis and meningitis are strongly associated with the presence of Srr-2 in *Streptococcus agalactiae*. The receptor or ligand of Srr-2 has not been identified, but it is suggested that expression of Srr-2 accelerates disease onset in neonates by facilitating transmigration of the BBB. Periodontal disease is associated with the expression of Fap1 from *Streptococcus parasanguinis*. Fap1 binds to glycosylated salivary proteins forming biofilms to cause caries. SrpA in *Streptococcus cristatus* is required for the formation of corn-cob structures with *Corynebacterium matruchoi* and *Fusobacterium nucleatum*, mediating intra-species adhesion seen in dental plaque.

For example, we have demonstrated that amino acids 122–166 of PsrP in the NR domain mediate a bacterial self-interaction that contributes to robust biofilm formation *in vitro* and bacterial aggregation *in vivo* (Fig. 6) (Sanchez *et al.*, 2010). The affinity of recombinant PsrP NR for other pneumococci is dependent on the presence of the NR domain of PsrP on the bacterial surface, suggesting dimerization of the NR. Similar to keratin 10 binding, which is mediated through a different stretch of amino acids in NR, the intra-bacterial interaction was found to be glycoconjugate independent (Shivshankar *et al.*, 2009).

We subsequently determined that SraP (*Staphylococcus aureus*) and GspB (*S. gordonii*) also promote bacterial interactions during *in vitro* biofilm growth. Unlike the NR domain of PsrP, the NRs of these SRRPs have an affinity not only for bacteria carrying the SRRP, but also for cell lysates from other bacteria, whether they carry an SRRP or not (Sanchez *et al.*, 2010). Hence, as was originally demonstrated with *S. cristatus* SrpA, SRRPs most likely have the ability to serve as inter-species adhesins. Although untested, one possible explanation for the limited interactions of PsrP vs. the other SRRPs lies in its cognate ligand that is a specific amino acid motif,

whereas the ligand for GspB is a sialic acid moiety. Hence, the NR domains of GspB and SraP may mediate interactions with any surface components that carry the target glycoconjugate. Furthermore, the possibility exists that these NR domains may bind to glycoconjugates that are found on the SRRP itself, for example on the extended SRR1 domains of another SRRP molecule.

Importantly, the fact that SRRPs can mediate host surface binding and bacterial aggregation clarifies the role that these proteins play during pathogenesis. SRRP-mediated interactions facilitate formation of dental plaque, which is recognized to be a multi-species biofilm. One can also imagine that SRRPs on *Staphylococcus aureus* or the oral streptococci facilitate attachment to platelets and cells in an endocarditic lesion, while at the same time mediating adhesion to other bacteria. Similarly, one could envision a microcolony of the pneumococcus in the lungs or middle ear with some bacteria attached to host cells via PsrPNR::K10 interactions and other bacteria attached to these bacteria through PsrPNR::PsrPNR interactions. In all, because SRRPs seem to be involved in bacterial aggregation and biofilm formation, they may play a role in any aspect of disease that involves biofilms. Consistent with this idea,



**Figure 6** PsrP in *Streptococcus pneumoniae* mediates robust biofilms *in vitro* and pneumococcal aggregates *in vivo*. Micrographs of (A) TIGR4 (wildtype) and (B) T4  $\Delta psrP$  biofilms grown in a flow cell under once-through flow conditions for 3 days. Bacteria were visualized with Live/Dead BacLight stain using an inverted confocal laser scanning microscope at 400 $\times$  magnification. Representative micrographs of (C) TIGR4 and (D) T4  $\Delta psrP$  Gram-stained bacteria from bronchial alveolar lavage samples from mice. These findings were first described in Sanchez *et al.* (2010).

microarray studies and proteomic studies show that PsrP is one of the few pneumococcal virulence determinants that is upregulated during pneumococcal biofilm formation (Sanchez *et al.*, 2011b).

### TRANSCRIPTIONAL REGULATION OF SRRPS

Other than the finding that PsrP is upregulated during biofilm growth (Sanchez *et al.*, 2011b), nothing else is known about the regulation of SRRPs or their accessory genes. As no obvious transcription factors are encoded in any of the SRRP loci, this suggests that the genes may be regulated by *trans*-acting elements or non-coding RNAs embedded within the loci. Future studies are warranted to determine how these large, and most likely highly metabolically expensive, determinants are transcriptionally regulated. Importantly, it is also unknown if SRRPs are also post-translationally modified on the bacterial surface.

### THE NR DOMAINS OF FAP1 AND GSPB HAVE BEEN CRYSTALLIZED AND CONTAIN DISTINCT DOMAINS

To gain insight into the mechanism of NR-mediated adhesion in the SRRPs, high-resolution X-ray crystal structures of two NRs, Fap1 and GspB, have been determined. In addition, the Srr-1 keratin 4-binding domain from *S. agalactiae* has been expressed, purified and crystallized. However, further crystallization trials are needed to obtain diffraction-quality crystals (Sundaresan *et al.*, 2011). Analysis of the NR domains of Fap1 from *S. parasanguinis* and GspB from *S. gordonii* revealed that the predicted binding regions of these adhesins contain different modular organizations (Ramboarina *et al.*, 2010; Garnett *et al.*, 2011, 2012). The GspB NR structure folds into a linear elongated rod, containing three independent subdomains. Interestingly, all three subdomains are predominated by  $\beta$ -strands as a secondary structure. The first subdomain has a folding topology similar to the *Staphylococcus aureus* Microbial Surface Components Recognizing Adhesive Matrix Molecule (MSCRAMM) surface adhesin protein CNA; therefore, it has been referred to as the CnaA subdomain. The second subdomain identified in GspB NR is very similar in topology to the sialic acid-binding immunoglobulin-like lectin 5 (Siglec-5) structure, mostly present on leukocytes. Siglecs are a family of

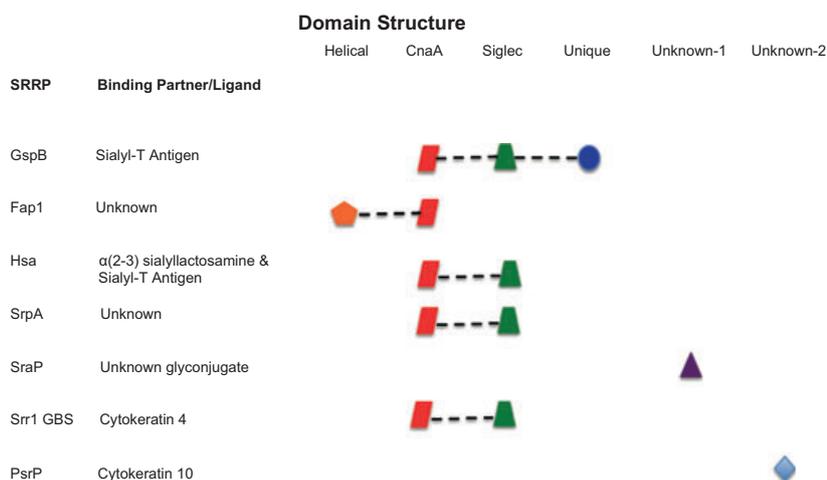
structurally related transmembrane proteins shown to be responsible for recognizing sialylated glycoconjugates (Pyburn *et al.*, 2011). Accordingly, it is this domain that mediates GspB NR binding to sialyl-T antigen (Takamatsu *et al.*, 2005a; Pyburn *et al.*, 2011). The third domain located on the C-terminal part of GspB NR did not have identity with any known protein structures. Therefore, this subdomain has been termed the unique subdomain (Pyburn *et al.*, 2011).

After a similar investigation, the structure of Fap1 NR has been shown to contain two distinct subdomains, NR $\alpha$  and NR $\beta$ , with different secondary structures (Ramboarina *et al.*, 2010; Garnett *et al.*, 2011, 2012). NR $\alpha$  is composed of three-helix bundles packed against each other in an up-down-up arrangement. This subdomain is very similar to the domain of the giant extracellular matrix-protein Ebh (ECM-binding protein homolog) from staphylococci (Tanaka *et al.*, 2008; Ramboarina *et al.*, 2010). Interestingly, Ebh also forms a bridge between the cytoplasm and cell wall and is involved in tolerance to transient hyperosmotic pressure (Kuroda *et al.*, 2008). Whether SRRP function in a similar role remains unexplored. The structure of the NR $\beta$  contains folding topology similar to the CnaA subdomain of GspB and MSCRAMMs with a similar secondary structure predominated by  $\beta$ -strands.

Although the GspB-CnaA and Fap1-NR $\beta$  subdomains have similar structures, they have distinct binding affinities. Only the Siglec domain in GspB NR binds to sialyl-T antigen, whereas in the Fap1 NR both subdomains have been shown to be important for binding saliva-coated hydroxyapatite; although the specific ligand for Fap1 remains unknown. Importantly, the Fap1 NR $\alpha$  structure is unique among the other NRs based on the sequence alignments. Alternatively, the CnaA subdomain in GspB and Fap-1 is likely to be found in other SRRPs (Fig. 7) (Pyburn *et al.*, 2011). It therefore seems possible that neutralizing antibodies could individually target distinct functional aspects of these NR domains, or alternatively, small molecules might be designed to target the CnaA domain.

### SRRPS AS NOVEL VACCINE CANDIDATES

The pivotal role of SRRPs during disease opens the possibility that their neutralization with antibody might confer protection. *In vitro* studies have demonstrated



**Figure 7** Modular organization of the non-repeat (NR) regions within the serine-rich repeat protein (SRRP) family. Subdomains of the NR region organized according to structural similarities. Similarities are based on crystal structure and predicted sequence identity after BLAST or CLUSTALW sequence alignments (Altschul *et al.*, 1990; Chenna *et al.*, 2003; Pyburn *et al.*, 2011; Garnett *et al.*, 2012). The middle column lists the known ligands for these NR. Figure adapted from Pyburn *et al.* (2011).

that monoclonal and polyclonal antibodies against Fap1 block the ability of the bacteria to form biofilms on glass slides. Interestingly, these monoclonal antibodies are specific to the glycoconjugates present on Fap1, suggesting that in addition to the adhesive property of the NR domain, the glycosylated SRR2 might also function as an adhesin (Stephenson *et al.*, 2002). In such an instance the SRRP might be functioning in a similar way to mucin. For *S. pneumoniae* polyclonal antibodies against the NR domain have been shown to block bacterial adhesion to A549 cells, inhibit biofilm formation, and protect mice following both passive and active immunization (Rose *et al.*, 2008; Shivshankar *et al.*, 2009; Sanchez *et al.*, 2010). Notably, individuals who have recovered from invasive pneumococcal disease have detectable antibody against PsrP, providing direct evidence that this protein is produced during human infection (Sanchez *et al.*, 2011a). It remains unclear as to which part of PsrP elicits a protective humoral response.

As a result of the modular domain structure of the NR domains, it seems possible that individual epitopes of the protein could be targeted. For example, antibodies against the MSCRAMM component or the Siglec domain would presumably neutralize these individual activities. Generation of monoclonal antibodies against these subdomains could be used to confirm the individual activity of these proteins and be tested as therapeutic agents. Unfortunately,

SRRPs are not universally present in the Gram-positive bacteria that cause infectious disease and can have distinct forms even within the same species. This suggests that immunization with SRRP fragments would not stand alone as a vaccine antigen and would most likely need to be part of a multi-valent protein vaccine.

## CONCLUSION

SRRPs are an important family of conserved surface proteins involved in Gram-positive pathogenesis. While evidence now strongly supports their role as intra- and inter-species adhesins, numerous questions remain unanswered. The most pressing of these include the role of the glycoconjugate forms of these proteins, how they are transcriptionally regulated, and full characterization of their cognate ligands. Given their important role in pathogenesis, SRRPs might be targeted for intervention. If so, information collected from structural analyses will be critical for the design of either antigens that elicit neutralizing antibodies or small molecules that bind specific subdomains.

## ACKNOWLEDGEMENTS

AL and CJS received support from NIDCR TE-32 DE 16318. CJO is supported by NIH AI078972. We

thank Dr David Kolodrubetz for his helpful review of the manuscript.

## REFERENCES

- Altschul, S.F., Gish, W., Miller, W., Myers, E.W. and Lipman, D.J. (1990) Basic local alignment search tool. *J Mol Biol* **215**: 403–410.
- Baba, T., Takeuchi, F., Kuroda, M. *et al.* (2002) Genome and virulence determinants of high virulence community-acquired MRSA. *Lancet* **359**: 1819–1827.
- Bensing, B.A. and Sullam, P.M. (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.
- Bensing, B.A. and Sullam, P.M. (2009) Characterization of *Streptococcus gordonii* SecA2 as a paralogue of SecA. *J Bacteriol* **191**: 3482–3491.
- Bensing, B.A. and Sullam, P.M. (2010) Transport of preproteins by the accessory Sec system requires a specific domain adjacent to the signal peptide. *J Bacteriol* **192**: 4223–4232.
- Bensing, B.A., Gibson, B.W. and Sullam, P.M. (2004) The *Streptococcus gordonii* platelet binding protein GspB undergoes glycosylation independently of export. *J Bacteriol* **186**: 638–645.
- Bensing, B.A., Takamatsu, D. and Sullam, P.M. (2005) Determinants of the streptococcal surface glycoprotein GspB that facilitate export by the accessory Sec system. *Mol Microbiol* **58**: 1468–1481.
- Bensing, B.A., Siboo, I.R. and Sullam, P.M. (2007) Glycine residues in the hydrophobic core of the GspB signal sequence route export toward the accessory Sec pathway. *J Bacteriol* **189**: 3846–3854.
- Bu, S., Li, Y., Zhou, M. *et al.* (2008) Interaction between two putative glycosyltransferases is required for glycosylation of a serine-rich streptococcal adhesin. *J Bacteriol* **190**: 1256–1266.
- Chen, Q., Wu, H. and Fives-Taylor, P.M. (2004) Investigating the role of secA2 in secretion and glycosylation of a fimbrial adhesin in *Streptococcus parasanguis* FW213. *Mol Microbiol* **53**: 843–856.
- Chen, Q., Wu, H., Kumar, R., Peng, Z. and Fives-Taylor, P.M. (2006) SecA2 is distinct from SecA in immunogenic specificity, subcellular distribution and requirement for membrane anchoring in *Streptococcus parasanguis*. *FEMS Microbiol Lett* **264**: 174–181.
- Chenna, R., Sugawara, H., Koike, T. *et al.* (2003) Multiple sequence alignment with the clustal series of programs. *Nucleic Acids Res* **31**: 3497–3500.
- Fachon-Kalweit, S., Elder, B.L. and Fives-Taylor, P. (1985) Antibodies that bind to fimbriae block adhesion of *Streptococcus sanguis* to saliva-coated hydroxyapatite. *Infect Immun* **48**: 617–624.
- Froeliger, E.H. and Fives-Taylor, P. (2001) *Streptococcus parasanguis* fimbria-associated adhesin fap1 is required for biofilm formation. *Infect Immun* **69**: 2512–2519.
- Garnett, J.A., Ramboarina, S., Lee, W.C., Tagliaferri, C., Wu, W. and Matthews, S. (2011) Crystallization and initial crystallographic analysis of the *Streptococcus parasanguinis* FW213 Fap1-NRalpha adhesive domain at pH 5.0. *Acta Crystallogr Sect F Struct Biol Cryst Commun* **67**: 274–276.
- Garnett, J.A., Simpson, P.J., Taylor, J. *et al.* (2012) Structural insight into the role of *Streptococcus parasanguinis* Fap1 within oral biofilm formation. *Biochem Biophys Res Commun* **417**: 421–426.
- Glaser, P., Rusniok, C., Buchrieser, C. *et al.* (2002) Genome sequence of *Streptococcus agalactiae*, a pathogen causing invasive neonatal disease. *Mol Microbiol* **45**: 1499–1513.
- Gu, L.H. and Coulombe, P.A. (2007) Keratin function in skin epithelia: a broadening palette with surprising shades. *Curr Opin Cell Biol* **19**: 13–23.
- Handley, P.S., Correia, F.F., Russell, K., Rosan, B. and DiRienzo, J.M. (2005) Association of a novel high molecular weight, serine-rich protein (SprA) with fibril-mediated adhesion of the oral biofilm bacterium *Streptococcus cristatus*. *Oral Microbiol Immunol* **20**: 131–140.
- Kuroda, M., Tanaka, Y., Aoki, R., Shu, D., Tsumoto, K. and Ohta, T. (2008) *Staphylococcus aureus* giant protein Ebh is involved in tolerance to transient hyperosmotic pressure. *Biochem Biophys Res Commun* **374**: 237–241.
- Lauer, P., Rinaudo, C.D., Soriani, M. *et al.* (2005) Genome analysis reveals pili in Group B Streptococcus. *Science* **309**: 105.
- Leube, R.E., Bader, B.L., Bosch, F.X., Zimbelmann, R., Achtstaetter, T. and Franke, W.W. (1988) Molecular characterization and expression of the stratification-related cytokeratins 4 and 15. *J Cell Biol* **106**: 1249–1261.
- Li, Y., Chen, Y., Huang, X. *et al.* (2008) A conserved domain of previously unknown function in Gap1 mediates protein-protein interaction and is required for biogenesis of a serine-rich streptococcal adhesin. *Mol Microbiol* **70**: 1094–1104.
- Mistou, M.Y., Dramsi, S., Brega, S., Poyart, C. and Trieu-Cuot, P. (2009) Molecular dissection of the secA2 locus

- of Group B Streptococcus reveals that glycosylation of the Srr1 LPXTG protein is required for full virulence. *J Bacteriol* **191**: 4195–4206.
- Nobbs, A.H., Lamont, R.J. and Jenkinson, H.F. (2009) Streptococcus adherence and colonization. *Microbiol Mol Biol Rev* **73**: 407–450.
- Obert, C., Sublett, J., Kaushal, D. *et al.* (2006) Identification of a candidate *Streptococcus pneumoniae* core genome and regions of diversity correlated with invasive pneumococcal disease. *Infect Immun* **74**: 4766–4777.
- Peng, Z., Fives-Taylor, P., Ruiz, T. *et al.* (2008a) Identification of critical residues in Gap3 of *Streptococcus parasanguinis* involved in Fap1 glycosylation, fimbrial formation and *in vitro* adhesion. *BMC Microbiol* **8**: 52.
- Peng, Z., Wu, H., Ruiz, T. *et al.* (2008b) Role of gap3 in Fap1 glycosylation, stability, *in vitro* adhesion, and fimbrial and biofilm formation of *Streptococcus parasanguinis*. *Oral Microbiol Immunol* **23**: 70–78.
- Plummer, C., Wu, H., Kerrigan, S.W., Meade, G., Cox, D. and Ian Douglas, C.W. (2005) A serine-rich glycoprotein of *Streptococcus sanguis* mediates adhesion to platelets via GPIb. *Br J Haematol* **129**: 101–109.
- Pyburn, T.M., Bensing, B.A., Xiong, Y.Q. *et al.* (2011) A structural model for binding of the serine-rich repeat adhesin GspB to host carbohydrate receptors. *PLoS Pathog* **7**: e1002112.
- Ramboarina, S., Garnett, J.A., Zhou, M. *et al.* (2010) Structural insights into serine-rich fimbriae from Gram-positive bacteria. *J Biol Chem* **285**: 32446–32457.
- Rigel, N.W. and Braunstein, M. (2008) A new twist on an old pathway – accessory Sec systems. *Mol Microbiol* **69**: 291–302.
- Rose, L., Shivshankar, P., Hinojosa, E., Rodriguez, A., Sanchez, C.J. and Orihuela, C.J. (2008) Antibodies against PsrP, a novel *Streptococcus pneumoniae* adhesin, block adhesion and protect mice against pneumococcal challenge. *J Infect Dis* **198**: 375–383.
- Samen, U., Eikmanns, B.J., Reinscheid, D.J. and Borges, F. (2007) The surface protein Srr-1 of *Streptococcus agalactiae* binds human keratin 4 and promotes adherence to epithelial HEp-2 cells. *Infect Immun* **75**: 5405–5414.
- Sanchez, C.J., Shivshankar, P., Stol, K. *et al.* (2010) The pneumococcal serine-rich repeat protein is an intra-species bacterial adhesin that promotes bacterial aggregation *in vivo* and in biofilms. *PLoS Pathog* **6**: e1001044.
- Sanchez, C.J., Hurtgen, B.J., Lizcano, A., Shivshankar, P., Cole, G.T. and Orihuela, C.J. (2011a) Biofilm and planktonic pneumococci demonstrate disparate immunoreactivity to human convalescent sera. *BMC Microbiol* **11**: 245.
- Sanchez, C.J., Kumar, N., Lizcano, A. *et al.* (2011b) *Streptococcus pneumoniae* in biofilms are unable to cause invasive disease due to altered virulence determinant production. *PLoS ONE* **6**: e28738.
- Seepersaud, R., Bensing, B.A., Yen, Y.T. and Sullam, P.M. (2010) Asp3 mediates multiple protein-protein interactions within the accessory Sec system of *Streptococcus gordonii*. *Mol Microbiol* **78**: 490–505.
- Seifert, K.N., Adderson, E.E., Whiting, A.A., Bohnsack, J.F., Crowley, P.J. and Brady, L.J. (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.
- Shivshankar, P., Sanchez, C., Rose, L.F. and Orihuela, C.J. (2009) The *Streptococcus pneumoniae* adhesin PsrP binds to keratin 10 on lung cells. *Mol Microbiol* **73**: 663–679.
- Shivshankar, P., Boyd, A.R., Le Saux, C.J., Yeh, I.T. and Orihuela, C.J. (2011) Cellular senescence increases expression of bacterial ligands in the lungs and is positively correlated with increased susceptibility to pneumococcal pneumonia. *Aging Cell* **10**: 798–806.
- Siboo, I.R., Chambers, H.F. and Sullam, P.M. (2005) Role of SraP, a Serine-Rich Surface Protein of *Staphylococcus aureus*, in binding to human platelets. *Infect Immun* **73**: 2273–2280.
- Siboo, I.R., Chaffin, D.O., Rubens, C.E. and Sullam, P.M. (2008) Characterization of the accessory Sec system of *Staphylococcus aureus*. *J Bacteriol* **190**: 6188–6196.
- van Sorge, N.M., Quach, D., Gurney, M.A., Sullam, P.M., Nizet, V. and Doran, K.S. (2009) The Group B Streptococcal serine-rich repeat 1 glycoprotein mediates penetration of the blood-brain barrier. *J Infect Dis* **199**: 1479–1487.
- Stephenson, A.E., Wu, H., Novak, J., Tomana, M., Mintz, K. and Fives-Taylor, P. (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.
- Sundaresan, R., Samen, U. and Ponnuraj, K. (2011) Expression, purification, crystallization and preliminary X-ray diffraction studies of the human keratin 4-binding domain of serine-rich repeat protein 1 from *Streptococcus agalactiae*. *Acta Crystallogr Sect F Struct Biol Cryst Commun* **67**: 1582–1585.
- Takahashi, Y., Konishi, K., Cisar, J.O. and Yoshikawa, M. (2002) Identification and characterization of hsa, the gene encoding the sialic acid-binding adhesin of *Streptococcus gordonii* DL1. *Infect Immun* **70**: 1209–1218.

- Takahashi, Y., Yajima, A., Cisar, J.O. and Konishi, K. (2004) Functional analysis of the *Streptococcus gordonii* DL1 sialic acid-binding adhesin and its essential role in bacterial binding to platelets. *Infect Immun* **72**: 3876–3882.
- Takamatsu, D., Bensing, B.A. and Sullam, P.M. (2004a) Four proteins encoded in the *gspB-secY2A2* operon of *Streptococcus gordonii* mediate the intracellular glycosylation of the platelet-binding protein GspB. *J Bacteriol* **186**: 7100–7111.
- Takamatsu, D., Bensing, B.A. and Sullam, P.M. (2004b) Genes in the accessory *sec* locus of *Streptococcus gordonii* have three functionally distinct effects on the expression of the platelet-binding protein GspB. *Mol Microbiol* **52**: 189–203.
- Takamatsu, D., Bensing, B.A., Cheng, H. *et al.* (2005a) Binding of the *Streptococcus gordonii* surface glycoproteins GspB and Hsa to specific carbohydrate structures on platelet membrane glycoprotein Ibalpha. *Mol Microbiol* **58**: 380–392.
- Takamatsu, D., Bensing, B.A. and Sullam, P.M. (2005b) Two additional components of the accessory *sec* system mediating export of the *Streptococcus gordonii* platelet-binding protein GspB. *J Bacteriol* **187**: 3878–3883.
- Takamatsu, D., Bensing, B.A., Prakobphol, A., Fisher, S.J. and Sullam, P.M. (2006) Binding of the streptococcal surface glycoproteins GspB and Hsa to human salivary proteins. *Infect Immun* **74**: 1933–1940.
- Takeuchi, F., Watanabe, S., Baba, T. *et al.* (2005) Whole-genome sequencing of staphylococcus haemolyticus uncovers the extreme plasticity of its genome and the evolution of human-colonizing staphylococcal species. *J Bacteriol* **187**: 7292–7308.
- Tanaka, Y., Sakamoto, S., Kuroda, M. *et al.* (2008) A helical string of alternately connected three-helix bundles for the cell wall-associated adhesion protein Ebh from *Staphylococcus aureus*. *Structure* **16**: 488–496.
- Tettelin, H., Nelson, K.E., Paulsen, I.T. *et al.* (2001) Complete genome sequence of a virulent isolate of *Streptococcus pneumoniae*. *Science* **293**: 498–506.
- Tsuji, T., Aoshiba, K. and Nagai, A. (2010) Alveolar cell senescence exacerbates pulmonary inflammation in patients with chronic obstructive pulmonary disease. *Respiration* **80**: 59–70.
- Vickerman, M.M., Iobst, S., Jesionowski, A.M. and Gill, S.R. (2007) Genome-wide transcriptional changes in *Streptococcus gordonii* in response to competence signaling peptide. *J Bacteriol* **189**: 7799–7807.
- Wu, H. and Fives-Taylor, P.M. (1999) Identification of dipeptide repeats and a cell wall sorting signal in the fimbriae-associated adhesin, Fap1, of *Streptococcus parasanguis*. *Mol Microbiol* **34**: 1070–1081.
- Wu, H., Mintz, K.P., Ladha, M. and Fives-Taylor, P.M. (1998) Isolation and characterization of Fap1, a fimbriae-associated adhesin of *Streptococcus parasanguis* FW213. *Mol Microbiol* **28**: 487–500.
- Wu, H., Bu, S., Newell, P., Chen, Q. and Fives-Taylor, P. (2007a) Two gene determinants are differentially involved in the biogenesis of Fap1 precursors in *Streptococcus parasanguis*. *J Bacteriol* **189**: 1390–1398.
- Wu, H., Zeng, M. and Fives-Taylor, P. (2007b) The glycan moieties and the N-terminal polypeptide backbone of a fimbria-associated adhesin, Fap1, play distinct roles in the biofilm development of *Streptococcus parasanguinis*. *Infect Immun* **75**: 2181–2188.
- Xiong, Y.Q., Bensing, B.A., Bayer, A.S., Chambers, H.F. and Sullam, P.M. (2008) Role of the serine-rich surface glycoprotein GspB of *Streptococcus gordonii* in the pathogenesis of infective endocarditis. *Microb Pathog* **45**: 297–301.
- Xu, P., Alves, J.M., Kitten, T. *et al.* (2007) Genome of the opportunistic pathogen *Streptococcus sanguinis*. *J Bacteriol* **189**: 3166–3175.
- Yen, Y.T., Seepersaud, R., Bensing, B.A. and Sullam, P.M. (2011) Asp2 and Asp3 interact directly with GspB, the export substrate of the *Streptococcus gordonii* accessory Sec System. *J Bacteriol* **193**: 3165–3174.
- Zhang, Y.Q., Ren, S.X., Li, H.L. *et al.* (2003) Genome-based analysis of virulence genes in a non-biofilm-forming *Staphylococcus epidermidis* strain (ATCC 12228). *Mol Microbiol* **49**: 1577–1593.
- Zhou, M. and Wu, H. (2009) Glycosylation and biogenesis of a family of serine-rich bacterial adhesins. *Microbiology* **155**: 317–327.
- Zhou, M., Peng, Z., Fives-Taylor, P. and Wu, H. (2008) A conserved C-terminal 13-amino-acid motif of Gap1 is required for Gap1 function and necessary for the biogenesis of a serine-rich glycoprotein of *Streptococcus parasanguinis*. *Infect Immun* **76**: 5624–5631.
- Zhou, M., Zhu, F., Dong, S., Pritchard, D.G. and Wu, H. (2010) A novel glucosyltransferase is required for glycosylation of a serine-rich adhesin and biofilm formation by *Streptococcus parasanguinis*. *J Biol Chem* **285**: 12140–12148.
- Zhou, M., Zhang, H., Zhu, F. and Wu, H. (2011) Canonical SecA associates with an accessory secretory protein complex involved in biogenesis of a streptococcal serine-rich repeat glycoprotein. *J Bacteriol* **193**: 6560–6566.

Zhou, M., Zhu, F., Li, Y., Zhang, H. and Wu, H. (2012) Gap1 functions as a molecular chaperone to stabilize its interactive partner Gap3 during biogenesis of serine-rich repeat bacterial adhesin. *Mol Microbiol* **83**: 866–878.

Zhu, F., Erlandsen, H., Ding, L. *et al.* (2011) Structural and functional analysis of a new subfamily of glycosyltransferases required for glycosylation of serine-rich streptococcal adhesins. *J Biol Chem* **286**: 27048–27057.

Copyright of Molecular Oral Microbiology is the property of Wiley-Blackwell and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.