

## Review

Correspondence  
Hui Wu  
hwu@uab.edu

# Glycosylation and biogenesis of a family of serine-rich bacterial adhesins

Meixian Zhou and Hui Wu

Department of Pediatric Dentistry, UAB School of Dentistry, University of Alabama at Birmingham, Birmingham, AL 35294, USA

Glycosylation of bacterial proteins is an important process for bacterial physiology and pathophysiology. Both O- and N-linked glycan moieties have been identified in bacterial glycoproteins. The N-linked glycosylation pathways are well established in Gram-negative bacteria. However, the O-linked glycosylation pathways are not well defined due to the complex nature of known O-linked glycoproteins in bacteria. In this review, we examine a new family of serine-rich O-linked glycoproteins which are represented by fimbriae-associated adhesin Fap1 of *Streptococcus parasanguinis* and human platelet-binding protein GspB of *Streptococcus gordonii*. This family of glycoproteins is conserved in streptococcal and staphylococcal species. A gene cluster coding for glycosyltransferases and accessory Sec proteins has been implicated in the protein glycosylation. A two-step glycosylation model is proposed. Two glycosyltransferases interact with each other and catalyse the first step of the protein glycosylation in the cytoplasm; the cross-talk between glycosylation-associated proteins and accessory Sec components mediates the second step of the protein glycosylation, an emerging mechanism for bacterial O-linked protein glycosylation. Dissecting the molecular mechanism of this conserved biosynthetic pathway offers opportunities to develop new therapeutic strategies targeting this previously unrecognized pathway, as serine-rich glycoproteins have been shown to play a role in bacterial pathogenesis.

## Introduction

Protein glycosylation is recognized as one of the major post-translational modifications and plays important roles in cell–cell recognition, adhesion and intracellular sorting (Drickamer & Taylor, 1998). Protein glycosylation is common in eukaryotes. More recently it has become clear that protein glycosylation occurs in all three domains of life (Schmidt *et al.*, 2003). Prokaryotes, including bacteria and archaea, are able to glycosylate proteins (Moens & Vanderleyden, 1997; Upreti *et al.*, 2003). More than 70 bacterial glycoproteins have been reported so far. Some are well characterized in Gram-negative bacteria, and a number of glycosylated proteins have also been documented and studied in Gram-positive bacteria (Upreti *et al.*, 2003). Most characterized glycoproteins are virulence factors of medically important pathogens. The glycan variations of a few bacterial glycoproteins are closely associated with bacterial infection (Banerjee *et al.*, 2002; Hamadeh *et al.*, 1995), suggesting that bacterial protein glycosylation plays an important role in pathogenesis (Moens & Vanderleyden, 1997; Szymanski *et al.*, 2003; Szymanski & Wren, 2005). Thus, the protein glycosylation and biosynthesis pathways may be new targets for therapeutic and prophylactic measures.

The biogenesis of glycoproteins is mediated by multiple genes clustered in so-called ‘glycosylation islands’ (Arora

*et al.*, 2001; Power *et al.*, 2003; Szymanski *et al.*, 2003). There are two basic types of protein glycosylation systems: N-linked glycosylation (oligosaccharides are usually linked to the amide nitrogen of asparagine residues of the peptide chain) and O-linked glycosylation (oligosaccharides are usually linked to the hydroxyl group of serine or threonine residues of peptide chains) (Szymanski & Wren, 2005). These two mechanisms can be differentiated by the mode in which the oligosaccharides are transferred to proteins. In the N-linked protein glycosylation system, an oligosaccharide is preassembled on a lipid carrier before being transferred *en bloc* to selected asparagine residues of acceptor proteins by an oligosaccharyltransferase (Szymanski & Wren, 2005). The N-linked glycosylation pathway is well defined in *Campylobacter jejuni* (Wacker *et al.*, 2002). The classic O-linked protein glycosylation mechanism involves the sequential transfer of glycans directly from nucleotide sugar precursors to acceptor proteins. Bacteria produce a diverse array of O-linked glycoproteins. Recently one O-linked glycosylation pathway was characterized in the Gram-negative bacterium *Neisseria gonorrhoeae* (Hegge *et al.*, 2004), which shares numerous common biosynthetic genes with N-linked glycosylation pathways and does not fit the bill as a classic O-linked glycosylation pathway. The biosynthetic mechanism in most bacteria is still unclear. Here we summarize

recent findings on the structure, genetics and function of serine-rich O-linked glycoproteins of streptococci and staphylococci, and propose a biosynthetic pathway that represents an emerging mechanism for the bacterial protein glycosylation system.

### Identification of serine-rich glycoproteins

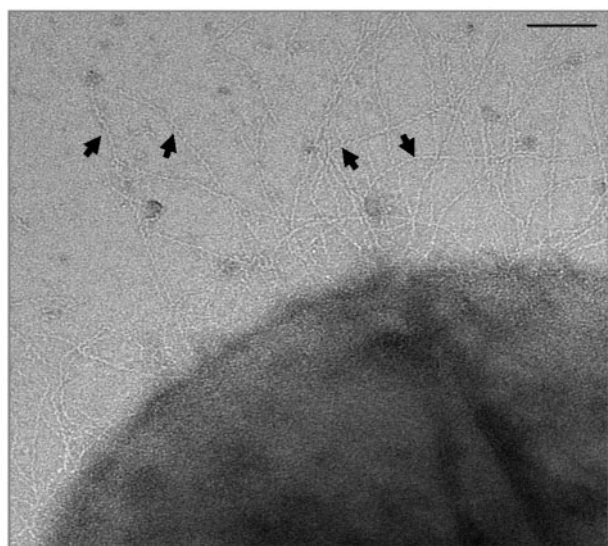
The first serine-rich glycosylated protein was identified from an oral streptococcal bacterium, *Streptococcus parasanguinis* FW213 (Wu *et al.*, 1998; Wu & Fives-Taylor, 1999). *Strep. parasanguinis* expresses peritrichous fimbriae on the cell surface (Fig. 1). Initial screening of nonadherent mutants revealed that they lacked fimbriae (Fachon-Kalweit *et al.*, 1985), suggesting the importance of fimbriae in bacterial adhesion. The protein responsible for the formation of fimbriae was identified by immunological approaches. A series of monoclonal antibodies were produced by immunizing mice with the wild-type FW213. Specific monoclonal antibodies (mAbs) F51, D10 and E42 that react with wild-type bacteria but do not react with the nonadherent bacteria were identified (Fachon-Kalweit *et al.*, 1985; Stephenson *et al.*, 2002) and used to detect bacterial adhesins. Interestingly, all these monoclonal antibodies label the bacterial fimbriae and recognize a 200 kDa protein, suggesting that the 200 kDa protein is a fimbriae-associated protein. Therefore we named the fimbrial protein as Fap1 (fimbriae-associated protein 1). The identity of Fap1 was elusive until the protein was purified in 1998. The purification and characterization of the Fap1 protein helped isolate and clone the full-length

*fap1* gene. Fap1 has been shown to be involved in bacterial fimbrial formation and adhesion, as a knockout mutant failed to form fimbriae and had reduced ability to adhere (Wu *et al.*, 1998). Recently we characterized nonadherent mutants of *Strep. parasanguinis* FW213 generated by random chemical mutagenesis and found that all the mutants had defects in *fap1* or in genes related to Fap1 biogenesis (Zhou *et al.*, 2008). These results demonstrate the importance of Fap1 in bacterial adhesion. The initial cloning and sequencing of the *fap1* gene proved to be challenging because *fap1* possesses extensive repetitive sequences. The sequencing of the large repetitive region was completed by using a transposon mutagenesis strategy which allows us to sequence one 400–600 bp repetitive region at a time and assemble them manually based on the exact transposition insertion sites (Wu & Fives-Taylor, 1999). The unusual repetitive nature of Fap1 was initially thought to be artificial; however, the discovery of Fap1-like proteins in other streptococci, such as *Streptococcus gordonii* strains M99 and DL1 (Bensing & Sullam, 2002; Takahashi *et al.*, 2002; Vickerman *et al.*, 2007), *Streptococcus cristatus* (Handley *et al.*, 2005) and the sequenced genome of *Streptococcus pneumoniae* TIGR4 (Tettelin *et al.*, 2001), has confirmed our findings and stimulated researchers' interest in the area.

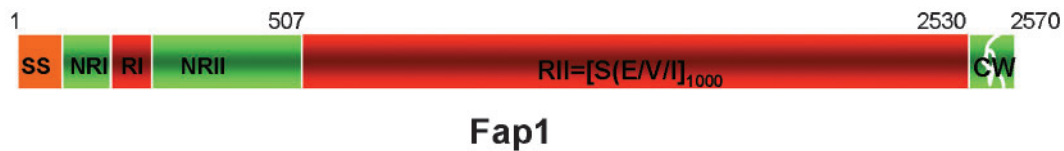
Fap1 was found to be glycosylated, as purified Fap1 stained poorly with the protein stain Coomassie brilliant blue but strongly with the carbohydrate stain periodic acid–Schiff's reagent (Wu *et al.*, 1998). Other serine-rich adhesins shown to be glycosylated include GspB of *Strep. gordonii* (Bensing *et al.*, 2004a), SrpA of *Streptococcus sanguinis* (Plummer *et al.*, 2005) and SraP of *Staphylococcus aureus* (Siboo *et al.*, 2008). The glycosylation status of serine-rich proteins recognized from other Gram-positive bacterial genomes is unknown. Given the fact the predicted serine-rich proteins are localized adjacent to glycosylation and secretion loci similar to the one initially identified from *Strep. gordonii* M99 (Bensing & Sullam, 2002), it is likely that other serine-rich proteins are similarly modified.

### Serine-rich glycoproteins are conserved in oral streptococci and many pathogenic streptococci and staphylococci

The predicted Fap1 protein contains 2570 amino acid residues (Fig. 2). It has a long signal sequence (SS) at the N terminus, and a classic cell wall anchor domain (CW) at the C terminus. Two serine-rich repeat regions (RI and RII) flank a non-repeat region II (NRII). Another non-repeat region (NRI) is located between SS and RI. The two repeat regions contain SESVSESVSI repeats and constitute 80 % of the entire amino acid sequence (Fig. 2). Fap1-like serine-rich proteins are also recognized from many other streptococci, staphylococci and lactobacilli. All serine-rich proteins contain two repeat regions; the amino acid sequences within the two regions are slightly different. The alternate amino acids are quite diverse. For instance,



**Fig. 1.** Electron micrograph of the surface fimbriae of *Strep. parasanguinis*. Wild-type *Strep. parasanguinis* FW213 grown to exponential phase were stained with 1 % uranyl acetate and visualized by electron microscopy. Arrows point to the fimbriae. Bar, 100 nm.



**Fig. 2.** Schematic diagram of Fap1. SS, signal sequence; NRI and NRII, non-repeat regions I and II; RI and RII, repeat regions I and II; CW, cell wall anchor domain.

Fap1 has a large number of glutamic acid (E) residues whereas GspB has alanine (A) residues instead (Table 1). Analysis of the Fap1-like proteins has revealed that all the Fap1-like serine-rich proteins are very acidic, with an isoelectric point (pI) below 6 (Table 2). Interestingly, the NRII regions of different Fap1-like proteins are diverse; they may be either acidic or basic. Most serine-rich proteins, like Fap1, have an acidic NRII region. Only a few Fap1-like proteins contain basic NRII regions. They include *Strep. pneumoniae*, *Strep. gordonii*, *Lactobacillus johnsonii* NCC533 and *Staphylococcus haemolyticus*. Paul Sullam and co-workers have proposed that the pI value of NRII may direct the binding of the serine-rich proteins to host receptors, based on their observations that basic NRII regions of Hsa and GspB of *Strep. gordonii* mediate bacterial adhesion to human platelets whereas the acidic NRII region of *Streptococcus agalactiae* fails to support the adhesion (Takamatsu *et al.*, 2005a, 2006). However, as they have pointed out, SrpA, a serine-rich protein of *Strep. sanguinis* with an acidic NRII, also mediates sialic-acid-dependent bacterial adhesion (Plummer *et al.*, 2005). Further, Fap1, another protein containing an acidic NRII region, does not mediate bacterial binding to human platelets at all. Together these results suggest that more complex mechanisms may be involved in NRII-mediated bacterial adhesion.

We have identified and characterized an eleven-gene cluster located adjacent to the *fap1* locus of *Strep. parasanguinis* FW213. This gene cluster contains genes encoding glycosyltransferases, a nucleotide-sugar synthetase and genes involved in protein secretion (Wu *et al.*, 2007a, b) (Fig. 3). A similar locus was initially identified from *Strep. gordonii* M99 (Bensing & Sullam, 2002) and has subsequently been shown to be important for glycosylation and biogenesis of GspB (Bensing *et al.*, 2004a; Takamatsu *et al.*, 2004a, b, 2005b). A search of completed and ongoing bacterial genome projects led to interesting findings: *fap1*-like gene clusters and related glycosylation and secretion loci were present in the genomes of oral streptococci, *Strep. gordonii* DL1 Challis (Vickerman *et al.*, 2007), *Strep. sanguinis* (Xu *et al.*, 2007) and *Strep. cristatus* (Handley *et al.*, 2005), and in important human pathogens, including *Strep. agalactiae* (Glaser *et al.*, 2002), *Strep. pneumoniae* (Tettelin *et al.*, 2001), *Staphylococcus aureus* (Baba *et al.*, 2002), *Staphylococcus epidermidis* (Zhang *et al.*, 2003) and *Staphylococcus haemolyticus* (Takeuchi *et al.*, 2005). Interestingly, this type of gene cluster is also found in some commensal bacteria, such as *L. johnsonii* NCC533 (NCBI genome site) and *Streptococcus salivarius* (Pombert *et al.*, 2008). Collectively the gene cluster contains two regions. One core region is composed of seven genes: *secY2*, *gap1–3* (designed as *asp1–3* in other organisms), *secA2*, *gtf1*

**Table 1.** Serine repeat sequences in the family of serine-rich proteins

Species and strain (protein name)	Serine repeat sequence	
	RI	RII
<i>Strep. parasanguinis</i> (Fap1)	SESVSSES	SESVSSES
<i>L. johnsonii</i> NCC533 (LJ1711)	SQSESI	SESL(I)SESSES
<i>Strep. sanguinis</i> Sk36 (SrpA)	SASISA	SASTSASTSASV
<i>Strep. cristatus</i> CC5A (SrpA)	SV(A)SV(A)SQ	SASASVSASQ
<i>Strep. agalactiae</i> 2603V/R (Srr-1)	SVSM	SAST(/M)
<i>Strep. agalactiae</i> J48 (Srr-2)	SSS	SESISSSES
<i>Staph. aureus</i> (SraP)	STST	STSLSD
<i>Staph. aureus</i> NCTC 8325 (SAOUHSC_02990)	STS	SVSTS
<i>Staph. epidermidis</i> ATCC 12228 (SE2249)	SSEK	SDSNSASTSL
<i>L. johnsonii</i> NCC533 (LJ_0391)	SLS	SL(/M)SNSVSMSE
<i>Strep. gordonii</i> M99 (GspB)	SAS	SASTSASVSAST(/E)
<i>Strep. gordonii</i> DL1 (Hsa)	SA(/L)S	SAST(/E)SASTSASV
<i>Strep. pneumoniae</i> TIGR4 (PsrP)	SASE(/T)	SASTSASA(/E)
<i>Staph. haemolyticus</i> JCSC1435 (SH0326)	STST	STSTSVSD

**Table 2.** pI values of serine-rich proteins

Group	Species and strain (protein name)	pI	
		Entire protein	NRII
Acidic	<i>Strep. parasanguinis</i> (Fap1)	3.26	4.85 (179–505)
	<i>L. johnsonii</i> NCC533 (LJ1711)	4.23	4.61 (181–1412)
	<i>Strep. sanguinis</i> Sk36 (SrpA)	4.96	5.41 (242–446)
	<i>Strep. cristatus</i> CC5A (SrpA)	5.6	6.21 (256–739)
	<i>Strep. agalactiae</i> 2603V/R (Srr-1)	4.94	4.63 (200–641)
	<i>Strep. agalactiae</i> J48 (Srr-2)	4.21	4.53 (212–531)
	<i>Staph. aureus</i> (SraP)	4.2	5.69 (231–751)
	<i>Staph. aureus</i> NCTC 8325 (SAOUHSC_02990)	4.04	5.61 (229–751)
	<i>Staph. epidermidis</i> ATCC 12228 (SE2249)	4.03	6.00 (263–790)
Basic	<i>L. johnsonii</i> NCC533 (LJ_0391)	4.7	8.13 (232–361)
	<i>Strep. gordonii</i> M99 (GspB)	3.96	9.51 (237–603)
	<i>Strep. gordonii</i> DL1 (Hsa)	4.15	8.74 (221–449)
	<i>Strep. pneumoniae</i> TIGR4 (PsrP)	3.9	9.91 (122–379)
	<i>Staph. haemolyticus</i> JCSC1435(SH0326)	3.58	9.25 (301–496)

and *gtf2* (*gtfA* and *gtfB* homologues); this core region is conserved in every genome that contains Fap1-like serine-rich proteins (Fig. 3). The other region contains putative glycosyltransferases. The number and identity of genes coding for putative glycosyltransferases are quite diverse, as depicted in Fig. 3. For instance, four glycosyltransferase-like genes (*gly*, *nss*, *galT1* and *galT2*) are located upstream of *fap1* whereas the putative glycosyltransferase genes are clustered downstream of the gene coding for some other serine-rich proteins. *Strep. pneumoniae* has large numbers of putative glycosyltransferases, ranging from seven to nine depending on the strain, which may suggest they have the most diverse modifications. The organization of the glycosyltransferases is completely different between *Strep. agalactiae* strains containing the serine-rich proteins Srr-1 and Srr-2. The conservation of the core region may provide the bacteria with a specific mechanism to synthesize surface-anchored adhesins, whereas the diversity of glycosyltransferases allows the bacteria to make adhesion molecules tailored to their specific ecological niches. Identification of the growing number of serine-rich glycoproteins has stimulated interest in this research area, as the role of the conserved gene cluster in the biogenesis of this family of serine-rich proteins is being recognized.

In addition, the search of genome sequences has revealed numerous other *secA2/Y2*-related loci in a wide variety of Gram-positive bacterial pathogens and mycobacterial species. For instance, accessory Sec proteins SecA2 and SecY2 are present in *Bacillus anthracis*, and SecA2 in *Listeria monocytogenes* (Lenz & Portnoy, 2002) and *Mycobacterium tuberculosis* (Rigel & Braunstein, 2008). These findings have revealed a new secretion mechanism. The SecA2/Y2 system plays an important role in protein export of serine-rich glycoproteins (Bensing & Sullam, 2002; Takahashi *et al.*, 2002; Vickerman *et al.*, 2007) and other unrelated substrates (Rigel & Braunstein, 2008). As

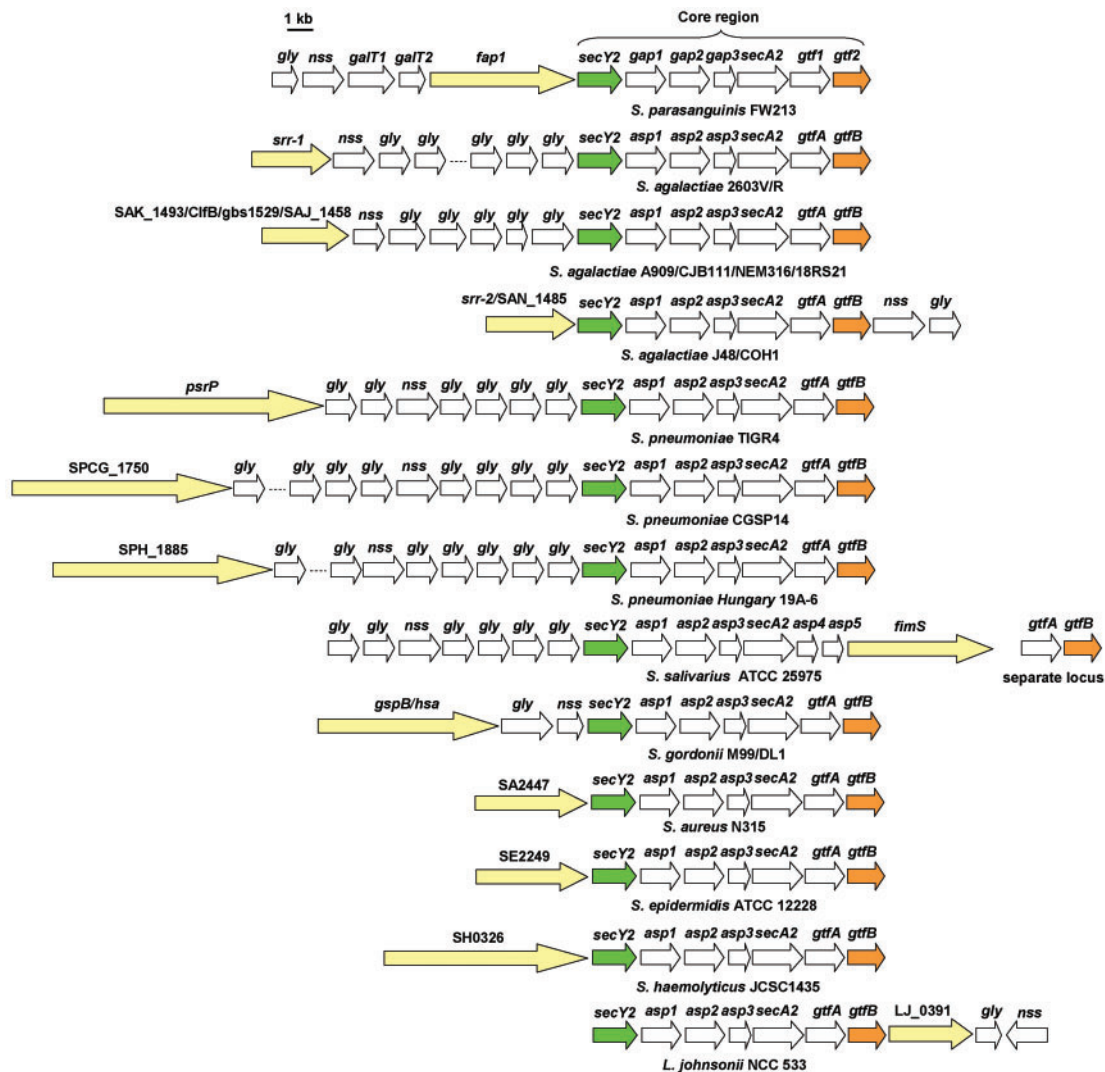
the accessory Sec systems have been reviewed recently, we will not discuss details in this review.

### Serine-rich glycoproteins play important roles in bacterial adhesion, colonization and pathogenesis

Fap1 was identified as an adhesin in *Strep. parasanguinis* as it is responsible for bacterial binding to saliva-coated hydroxylapatite (SHA) (Wu *et al.*, 1998), an *in vitro* tooth model. A 53 kDa salivary receptor has been identified, although the nature of the receptor is unknown (Wu *et al.*, 2007b). Fap1 also mediates bacterial biofilm formation *in vitro* (Froeliger & Fives-Taylor, 2001). Glycosylation of Fap1 is involved in fimbrial assembly, bacterial adhesion and biofilm formation (Peng *et al.*, 2008). It appears that the Fap1 polypeptide backbone is involved in bacterial adhesion whereas the protein glycosylation is responsible for the development of bacterial biofilm (Wu *et al.*, 2007b). Recently we have analysed a variety of non-adherent chemical mutants; strikingly, all the mutants we characterized have defects in Fap1 or Fap1-related genes (Zhou *et al.*, 2008), demonstrating that Fap1 is a key adhesion molecule.

Both GspB and Hsa have been separately identified from *Strep. gordonii* strains M99 and DL1. Hsa was isolated as a polysaccharide receptor (Takahashi *et al.*, 2002) whereas GspB was discovered by its ability to bind to human platelets (Bensing & Sullam, 2002). Hsa was later shown to be involved in binding to human platelets as well (Takahashi *et al.*, 2004). Importantly, inactivation of either *hsa* or *gspB* results in a reduction in *Strep. gordonii*-mediated infective endocarditis in a rat model (Takahashi *et al.*, 2006; Xiong *et al.*, 2008). The interaction of GspB with host components has been studied extensively; GspB not only binds to human platelets but also interacts with a variety of salivary components such as the low-molecular-





**Fig. 3.** Conservation of serine-rich glycoproteins and proteins contributing to their biosynthesis and secretion in streptococci, staphylococci and lactobacilli. DNA sequences of *fap1* and *fap1*-like loci were collected from the related NCBI websites. Genes coding for serine-rich repeat proteins linked to *secY2*–*gtf* loci were designated as follows: *fap1* of *Strep. parasanguinis*; *srr-1*, SAK\_1493 and *srr-2* of *Strep. agalactiae*; *psrP*, SPCG\_1750 and SPH\_1885 of *Strep. pneumoniae* TIGR4, CGSP14 and Hungary 19A-6 strains; *fimS* of *Strep. salivarius* ATCC 25975; *gspB* or *hsa* of *Strep. gordonii*; SA2447 of *Staph. aureus* N315; SE2249 of *Staph. epidermidis* ATCC 12228; SH0326 of *Staph. haemolyticus* JCSC1435; and LJ-0391 of *L. johnsonii* NCC533. Arrows indicate the direction of transcription.

mass salivary mucin MG2 and salivary agglutinin (gp340) (Takamatsu *et al.*, 2006). The binding specificity relies on sialic acid residues of the receptor molecules (Takamatsu *et al.*, 2005a, 2006). A very recent study has demonstrated that Hsa mediates sialic-acid-dependent binding of *Strep. gordonii* to HL-60 cells differentiated into monocytes, granulocytes and macrophages via known phagocyte receptors such as CD11b, CD43 and CD50 (Yajima *et al.*, 2008). All the binding is mediated by basic NRII regions of GspB and Hsa. The ability of *Strep. gordonii* to survive in polymorphonuclear leukocytes following adhesin-mediated phagocytosis has been postulated as an important virulence

determinant for infective endocarditis (Young Lee *et al.*, 2006). Identification of new host receptors for the bacterial adhesin supports the function of Hsa in bacterial infective endocarditis. Hsa provides a selective advantage for *Strep. gordonii* to outcompete *Strep. sanguinis* in binding to SHA *in vitro*, which provides a means for *Strep. gordonii* to persist within the oral cavity, despite the greater natural prevalence of *Strep. sanguinis* in plaque and saliva (Nobbs *et al.*, 2007). Among many adhesins of *Strep. gordonii*, Hsa is a major one that directs the bacterial adhesion to immobilized gp340 and epithelial cells (Jakubovics *et al.*, 2005). The oral streptococcus *Strep. cristatus* possesses

Fap1-like protein SrpA, which was believed to be involved in the assembly of bacterial surface fimbrial structures (Handley *et al.*, 2005); however, its precise function has not been explored. *Strep. sanguinis* has the ability to bind to human platelets; serine-rich glycoprotein SrpA mediates the binding (Plummer *et al.*, 2005) although it has not been possible to demonstrate its involvement in bacterial infective endocarditis (unpublished data). *Strep. pneumoniae* PsrP has been implicated in invasive diseases by epidemiological studies; PsrP deficiency slows the development of bacteraemia and increases the mouse survival rate (Obert *et al.*, 2006). PsrP exhibits a tissue tropism. It is required for bacterial persistence in the lungs but not for colonization in the nasopharynx or replication in the bloodstream during sepsis (Rose *et al.*, 2008). The basic NRII region of PsrP mediates bacterial adhesion to human alveolar type II pneumocytes, A549 cells (Obert *et al.*, 2006; Rose *et al.*, 2008). Two serine-rich proteins, Srr-1 and Srr-2, have been described for *Strep. agalactiae* strains 2603V/R and J48 (Samen *et al.*, 2007; Seifert *et al.*, 2006). Serotype III delta/epsilon-positive strains of *Strep. agalactiae* possess Srr-2. They are highly virulent in an *in vivo* mouse model of neonatal sepsis in comparison with Srr-2-negative serotype III strains (Seifert *et al.*, 2006), suggesting that Srr-2 likely contributes to bacterial pathogenesis. The function of the other serine-rich protein, Srr-1, has also been investigated. Srr-1 can interact with human keratin 4 and promotes adherence to epithelial cells, suggesting that *Strep. agalactiae* has developed a new means to interact with the host (Samen *et al.*, 2007). *Staph. aureus* SraP also binds to human platelets via its N-terminal NRII domain, and mediates the pathogenesis of infective endocarditis in a rabbit model (Siboo *et al.*, 2005). In a search of all 22 sequenced genomes of *Staph. aureus* ([http://www.ncbi.nlm.nih.gov/sites/entrez?Db=genomeprj&Cmd=Search&Term=ToSearch=txid1280\[orgn\]](http://www.ncbi.nlm.nih.gov/sites/entrez?Db=genomeprj&Cmd=Search&Term=ToSearch=txid1280[orgn])), we found that all sequenced strains contain SraP-like serine-rich proteins. Serine-rich glycoproteins Srr-1 and Srr-2 of *Strep. agalactiae* and their putative glycosylation and biosynthetic loci are also found in all eight sequenced genomes. These observations are in sharp contrast to what is documented for *Strep. pneumoniae*. In all 34 sequenced genomes of *Strep. pneumoniae* ([http://www.ncbi.nlm.nih.gov/sites/entrez?Db=genomeprj&Cmd=Search&Term=ToSearch=txid1313\[orgn\]](http://www.ncbi.nlm.nih.gov/sites/entrez?Db=genomeprj&Cmd=Search&Term=ToSearch=txid1313[orgn])), only six genomes have PsrP-like open reading frames. The prevalence of SraP in *Staph. aureus* and Srr-1 and Srr-2 in *Strep. agalactiae* strains highlights their importance.

In summary, the serine-rich proteins are important in bacterial adhesion via their NRII domains; they constitute a new family of bacterial lectins, as many of them bind to sialic-acid-containing surface molecules and play an important role in bacterial pathogenesis. Interestingly, predicted serine-rich proteins also exist in probiotic bacteria such as *Lactobacillus johnsonii* NCC533, *Lactobacillus reuteri* and *Streptococcus salivarius* ATCC 25975, members of the normal microbial community. Therefore, this important family of proteins may be also required for bacterial fitness and normal physiology.

## Glycosylation and biogenesis of serine-rich glycoproteins

Both O- and N-linked glycan moieties have been identified in glycoproteins of Gram-positive bacteria. The platelet-aggregation-associated protein (PAAP) of *Strep. sanguinis* (Erickson & Herzberg, 1993) is a well-defined streptococcal N-linked glycoprotein. Fap1 and its homologues constitute an important family of O-linked streptococcal glycoproteins and present a new biosynthetic mechanism for protein glycosylation (Peng *et al.*, 2008; Stephenson *et al.*, 2002; Wu *et al.*, 2007a, b). It has been reported that accessory Sec proteins SecA2 and SecY2 play an important role in export of serine-rich glycoproteins GspB and Fap1 (Bensing *et al.*, 2005; Chen *et al.*, 2004). As the contribution of glycan moieties and the signal sequence to export of serine-rich glycoproteins has been reviewed recently (Rigel & Braunstein, 2008), we will focus on mechanisms related to glycosylation and maturation of serine-rich proteins.

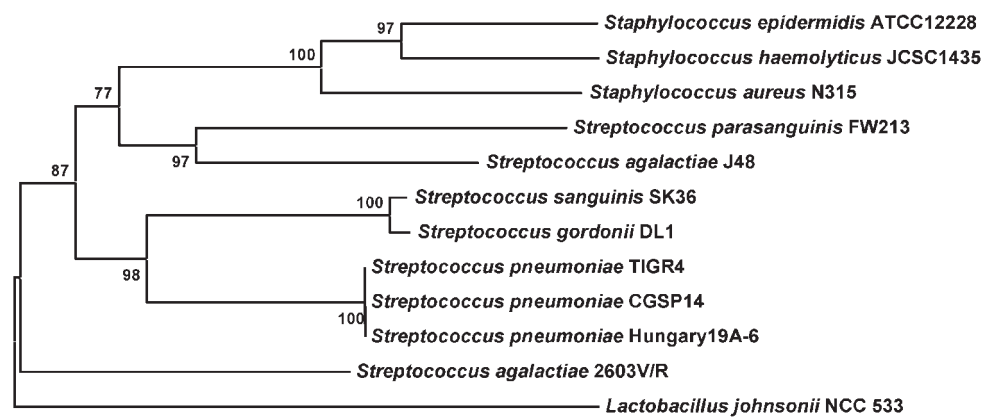
Fap1 is the first protein in the family determined to be glycosylated (Wu *et al.*, 1998). The purified Fap1 protein possesses several monosaccharide residues including glucose, N-acetylglucosamine (GlcNAc), N-acetylgalactosamine and rhamnose (Stephenson *et al.*, 2002). The glycan linkage to the Fap1 peptide backbone is not defined; however, it appears to be O-linked for the following reasons. N-Glycanases fail to release the glycan chains from Fap1 although the N-linked consensus sequence is present in Fap1 (Stephenson *et al.*, 2002). Furthermore, the putative O-linked amino acid residues exist in Fap1, and the glycan chains can be released from a recombinant Fap1 by a condition used to cleave O-linked glycan moieties selectively (Bu *et al.*, 2008). The GspB protein (Bensing *et al.*, 2004b) and other members of this family, SrpA of *Strep. sanguinis* (Plummer *et al.*, 2005) and SraP of *Staph. aureus* (Siboo *et al.*, 2008), all react with the GlcNAc-specific lectin sWGA, suggesting that adding GlcNAc residues to the peptide backbone is common. We have shown that the glycosylation of Fap1 is initiated by addition of a GlcNAc residue by two glycosyltransferases, Gtf1 and Gtf2, of *Strep. parasanguinis*. Inactivation of either *gtf1* or *gtf2* leads to the expression of unglycosylated 360 kDa Fap1 precursor A. Gtf1- or Gtf2-deficient mutants fail to produce mature Fap1. Furthermore, the interaction between Gtf1 and Gtf2 is required for Fap1 glycosylation (Bu *et al.*, 2008). Thus, we conclude that the formation of the Gtf1 and Gtf2 protein complex is important for this first step of Fap1 glycosylation.

What is intriguing though, is that deletion of *gtf1* and *gtf2* homologues *gtfA* and *gtfB* in *Strep. gordonii* abolishes the expression of native GspB although GtfA and GtfB have been shown to be involved in the transfer of the sugar residue to a recombinant GspB of *Strep. gordonii* M99 in the cytoplasm (Takamatsu *et al.*, 2004b). Our work has also demonstrated the intracellular accumulation of glycosylated Fap1 (Chen *et al.*, 2004), which supports the

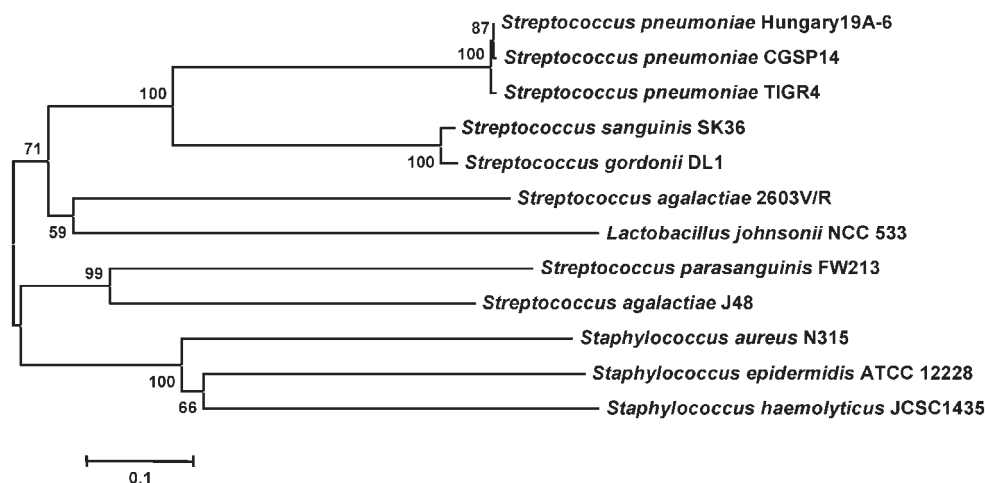
hypothesis that the protein glycosylation occurs in the cytoplasm. Phylogenetic analysis indicates that Gtf1 and Gtf2 homologues are highly conserved in a variety of streptococcal and staphylococcal species (Fig. 4). Furthermore we have shown that the Gtf1 and Gtf2 homologues from *Strep. sanguinis* and *Strep. pneumoniae* can transfer an oligosaccharide containing GlcNAc residues to the recombinant Fap1 molecule (Bu *et al.*, 2008), further highlighting the functional conservation of the Gtf complex. Thus we conclude that the formation of the Gtf1 and Gtf2 complex is required for glycosylation of serine-rich proteins.

A cluster of 11 genes flanking the *fap1* locus has been identified in *Strep. parasanguinis* (Bu *et al.*, 2008; Wu *et al.*,

2007a, b). Mutagenesis of the downstream genes such as *secY2*, *gap1* and *gap3* generated another distinct high-molecular-mass (HMM) 470 kDa Fap1 precursor which migrates slower than the Fap1 precursor A; we designated it Fap1 precursor B (Li *et al.*, 2008; Wu *et al.*, 2007a). The mature 200 kDa Fap1 reacts with all Fap1-specific peptide and glycan antibodies. Interestingly, the 470 kDa Fap1 precursor reacts with peptide-specific antibody (mAb E42) and one glycan-specific antibody, mAb D10, but not with the mature Fap1-specific mAb, F51, suggesting that this Fap1 precursor is incompletely glycosylated or improperly folded. Similar HMM protein precursors are also recognized in mutants of *Strep. gordonii* deficient in *secY2*, *gap1*



**Gtf1 phylogenetic analysis**



**Gtf2 phylogenetic analysis**

**Fig. 4.** Evolutionary relationships of glycosyltransferases Gtf1 and Gtf2. The phylogenetic analysis was performed using MEGA4. The length of the branches reflects the number of amino acid changes between different glycosyltransferases, as indicated by the bar. Bootstrap values are shown at the nodes. Glycosyltransferase sequences were obtained by searching all bacterial genomes and protein databases available from NCBI and TIGR.

and *gap3* homologues (Takamatsu *et al.*, 2005b), indicating that production of such a precursor is common in the biogenesis of serine-rich glycoproteins.

The presence of two distinct Fap1 precursors in the different mutants prompted us to hypothesize a two-step glycosylation pathway in which Fap1 progresses from apparent masses of 360 kDa to 470 kDa to 200 kDa based on SDS-PAGE analysis (Fig. 5). The first glycosylation step is controlled by a two-component glycosyltransferase complex, Gtf1 and Gtf2 (Bu *et al.*, 2008), and occurs intracellularly, while the second step requires several accessory secretion components including SecA2 (Chen *et al.*, 2004, 2006, 2007), SecY2 (Wu *et al.*, 2007a) and glycosylation-associated proteins Gap1 and Gap3 (Li *et al.*, 2008). In support of such a processing pathway, we find that only the 200 kDa protein is readily detectable in mature fimbriae (Wu & Fives-Taylor, 2001).

The expression of three Fap1 proteins, one with undetectable glycosylation (360 kDa), one incompletely glycosylated or improperly folded (470 kDa) and one fully glycosylated (200 kDa), is reminiscent of experimental findings in another oral streptococcus, *Strep. salivarius* (Weerkamp *et al.*, 1986a, b). Two antigens associated with fimbriae have been identified in *Strep. salivarius*, AgB and AgC. Both antigens are predominantly associated with the cell wall in wild-type cells. However, AgC (AgC<sub>in</sub>) accumulates in high amounts in the cytoplasmic fraction of mutants that lack AgC<sub>w</sub>, the cell-wall associated form. This HMM AgC<sub>in</sub> molecule resembles 470 kDa Fap1 precursor accumulated intracellularly in *secY2*, *gap1* and *gap3* mutants of *Strep. parasanguinis*. AgB, AgC<sub>in</sub> and AgC<sub>w</sub> have apparent molecular masses of 380, 488 and ~250 kDa, respectively. Both AgC<sub>in</sub> (488 kDa) and AgC<sub>w</sub> (~250 kDa) are associated with carbohydrate as evidenced by staining



**Fig. 5.** A model for the Fap1 glycosylation and biogenesis pathway. A newly synthesized Fap1 precursor A, a 360 kDa protein coded by the *fap1* locus, is glycosylated to produce an incompletely glycosylated or improperly folded 470 kDa Fap1 precursor B, which acquires additional glycan epitope(s) or proper folding to generate the 200 kDa mature Fap1 protein. The 360 kDa protein is only recognized by peptide-specific antibody mAbE42 whereas the 470 kDa protein reacts with the peptide-specific antibody mAbE42 and one glycan-specific antibody mAbD10. The mature 200 kDa protein is recognized by all the known Fap1-specific antibodies including mAbF51, which recognizes only the mature Fap1. The formation of the Gtf1 and Gtf2 complex is essential for Fap1 glycosylation in the cytoplasm whereas the interaction between accessory Sec components (SecA2 and SecY2) and glycosylation-associated proteins Gap1 and Gap3 is involved in the complete glycosylation process or proper folding as indicated.

with periodic acid–Schiff's reagent, whereas AgB (380 kDa) contains little or no carbohydrate, and may represent the unglycosylated polypeptide backbone. However, the relationships among these three antigens and the corresponding genes coding for these HMM proteins of *Strep. salivarius* have not been identified. These antigens may have a common origin and represent a system analogous to Fap1. Biochemical analysis has revealed the presence of Fap1-like glycoprotein in *Strep. salivarius* (Levesque *et al.*, 2001). Identification of *fap1*-related protein glycosylation and secretion loci in *Strep. salivarius* supports this assumption (Pombert *et al.*, 2008). Together these observations indicate that the proposed two-step glycosylation and secretion model is relevant to the biogenesis of the serine-rich glycoprotein of *Strep. salivarius*.

This two-step model supports the concept that the glycosylation is coupled to the secretion (Chen *et al.*, 2004; Wu *et al.*, 2007a) of the protein. Interestingly, a Gap1 homologue of *Strep. gordonii* has been hypothesized as an accessory secretion protein as it affects mature GspB secretion (Takamatsu *et al.*, 2004a, b). A HMM pre-GspB, but not mature GspB, has also been found in the cell wall fractions isolated from strains carrying mutations in each of the accessory secretory genes (*secY2*, *asp1*, *asp2*, *asp3* and *secA2*). Although the nature of the pre-GspB is unknown, it is possible that this protein precursor is incompletely glycosylated. These results indicate that the accessory Sec system is required not only for the efficient export but also for the proper processing of GspB. The protein glycosylation is likely to contribute to this process. In this regard, GspB export and maturation (glycosylation) also appear to be coupled. Study of the precise structural differences between pre-GspB and mature GspB would help resolve the issue whether the protein glycosylation is directly or indirectly responsible for the processing of GspB. We have evidence supporting the notion that Gap1 is a putative glycosyltransferase and regulates Fap1 glycosylation and maturation (Li *et al.*, 2008). Therefore, we conclude that glycosylation and secretion are coupled. In mycobacteria, Sec-dependent protein secretion mediates O-mannosylation, an analogous process prevalent in eukaryotic O-linked protein glycosylation (VanderVen *et al.*, 2005). Thus, glycosylation of bacterial proteins appears to resemble the analogous process in eukaryotic cells. In eukaryotic cells, the exogenously supplied or endogenously generated nucleotide sugar building blocks are imported into the secretory compartments of cells, including the lumen of the endoplasmic reticulum and the Golgi apparatus. Transfer of a given oligosaccharide to a protein scaffold is controlled by specific localization of glycosyltransferases and mediated by protein secretion (Hounsell *et al.*, 1996). Although the mechanism of prokaryotic glycosylation is not well defined and bacteria do not have specified organelles to carry out this function, it is possible that bacteria have developed an analogous system (such as subcellular localization) to control the function of the protein glycosylation complex. The proposed Fap1



glycosylation pathway is a dynamic process. The conserved sequence features indicate that this coupled glycosylation and secretion action is likely to be a common theme in biogenesis of serine-rich proteins in streptococci and staphylococci. Study of this newly identified biosynthetic pathway for bacterial glycosylation and secretion mechanism may lead to design of new therapeutic antimicrobial agents targeting the pathway.

## Acknowledgements

The work was supported by K22 DE014726, R01DE017954 and R01DE011000 from the National Institute of Dental and Craniofacial Research.

## References

- Arora, S. K., Bangera, M., Lory, S. & Ramphal, R. (2001). A genomic island in *Pseudomonas aeruginosa* carries the determinants of flagellin glycosylation. *Proc Natl Acad Sci U S A* **98**, 9342–9347.
- Baba, T., Takeuchi, F., Kuroda, M., Yuzawa, H., Aoki, K., Oguchi, A., Nagai, Y., Iwama, N., Asano, K. & other authors (2002). Genome and virulence determinants of high virulence community-acquired MRSA. *Lancet* **359**, 1819–1827.
- Banerjee, A., Wang, R., Supernavage, S. L., Ghosh, S. K., Parker, J., Ganesh, N. F., Wang, P. G., Gulati, S. & Rice, P. A. (2002). Implications of phase variation of a gene (*pgtA*) encoding a pilin galactosyl transferase in gonococcal pathogenesis. *J Exp Med* **196**, 147–162.
- Bensing, B. A. & 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., Gibson, B. W. & Sullam, P. M. (2004a). The *Streptococcus gordonii* platelet binding protein GspB undergoes glycosylation independently of export. *J Bacteriol* **186**, 638–645.
- Bensing, B. A., Lopez, J. A. & Sullam, P. M. (2004b). The *Streptococcus gordonii* surface proteins GspB and Hsa mediate binding to sialylated carbohydrate epitopes on the platelet membrane glycoprotein Ib $\alpha$ . *Infect Immun* **72**, 6528–6537.
- Bensing, B. A., Takamatsu, D. & Sullam, P. M. (2005). Determinants of the streptococcal surface glycoprotein GspB that facilitate export by the accessory Sec system. *Mol Microbiol* **58**, 1468–1481.
- Bu, S., Li, Y., Zhou, M., Azadin, P., Zeng, M., Fives-Taylor, P. & Wu, H. (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. & 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. & 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.
- Chen, Q., Sun, B., Wu, H., Peng, Z. & Fives-Taylor, P. M. (2007). Differential roles of individual domains in selection of secretion route of a *Streptococcus parasanguis* serine-rich adhesin, Fap1. *J Bacteriol* **189**, 7610–7617.
- Drickamer, K. & Taylor, M. E. (1998). Evolving views of protein glycosylation. *Trends Biochem Sci* **23**, 321–324.
- Erickson, P. R. & Herzberg, M. C. (1993). Evidence for the covalent linkage of carbohydrate polymers to a glycoprotein from *Streptococcus sanguis*. *J Biol Chem* **268**, 23780–23783.
- Fachon-Kalweit, S., Elder, B. L. & 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. & Fives-Taylor, P. (2001). *Streptococcus parasanguis* fimbria-associated adhesin Fap1 is required for biofilm formation. *Infect Immun* **69**, 2512–2519.
- Glaser, P., Rusniok, C., Buchrieser, C., Chevalier, F., Frangeul, L., Msadek, T., Zouine, M., Couv  , E., Lalioui, L. & other authors (2002). Genome sequence of *Streptococcus agalactiae*, a pathogen causing invasive neonatal disease. *Mol Microbiol* **45**, 1499–1513.
- Hamadeh, R. M., Estabrook, M. M., Zhou, P., Jarvis, G. A. & Griffiss, J. M. (1995). Anti-Gal binds to pili of *Neisseria meningitidis*: the immunoglobulin A isotype blocks complement-mediated killing. *Infect Immun* **63**, 4900–4906.
- Handley, P. S., Correia, F. F., Russell, K., Rosan, B. & DiRienzo, J. M. (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.
- Hegge, F. T., Hitchen, P. G., Aas, F. E., Kristiansen, H., L  vold, C., Egge-Jacobsen, W., Panico, M., Leong, W. Y., Bull, V. & other authors (2004). Unique modifications with phosphocholine and phosphoethanolamine define alternate antigenic forms of *Neisseria gonorrhoeae* type IV pili. *Proc Natl Acad Sci U S A* **101**, 10798–10803.
- Hounsell, E. F., Davies, M. J. & Renouf, D. V. (1996). O-linked protein glycosylation structure and function. *Glycoconj J* **13**, 19–26.
- Jakubovics, N. S., Kerrigan, S. W., Nobbs, A. H., Stromberg, N., van Dolleweerd, C. J., Cox, D. M., Kelly, C. G. & Jenkinson, H. F. (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.
- Lenz, L. L. & Portnoy, D. A. (2002). Identification of a second *Listeria secA* gene associated with protein secretion and the rough phenotype. *Mol Microbiol* **45**, 1043–1056.
- Levesque, C., Vadeboncoeur, C., Chandad, F. & Frenette, M. (2001). *Streptococcus salivarius* fimbriae are composed of a glycoprotein containing a repeated motif assembled into a filamentous nondissociable structure. *J Bacteriol* **183**, 2724–2732.
- Li, Y., Chen, Y., Huang, X., Zhou, M., Wu, R., Dong, S., Pritchard, D. G., Fives-Taylor, P. & Wu, H. (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.
- Moens, S. & Vanderleyden, J. (1997). Glycoproteins in prokaryotes. *Arch Microbiol* **168**, 169–175.
- Nobbs, A. H., Zhang, Y., Khammanivong, A. & Herzberg, M. C. (2007). *Streptococcus gordonii* Hsa environmentally constrains competitive binding by *Streptococcus sanguinis* to saliva-coated hydroxyapatite. *J Bacteriol* **189**, 3106–3114.
- Obert, C., Sublett, J., Kaushal, D., Hinojosa, E., Barton, T., Tuomanen, E. I. & Orihuela, C. J. (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., Wu, H., Ruiz, T., Chen, Q., Zhou, M., Sun, B. & Fives-Taylor, P. (2008). Role of *gap3* in Fap1 glycosylation, stability, in vitro adhesion, and fimbrial and biofilm formation of *Streptococcus parasanguis*. *Oral Microbiol Immunol* **23**, 70–78.

- Plummer, C., Wu, H., Kerrigan, S. W., Meade, G., Cox, D. & Ian Douglas, C. W. (2005). A serine-rich glycoprotein of *Streptococcus sanguis* mediates adhesion to platelets via GPIb. *Br J Haematol* **129**, 101–109.
- Pombert, C. B., Levesque, C. & Frenette, M. (2008). Genetic organization of fimbriae's secretion/glycosylation machinery in *Streptococcus salivarius*. *J Dent Res* **87** (special issue), 3422.
- Power, P. M., Roddam, L. F., Rutter, K., Fitzpatrick, S. Z., Srikhanta, Y. N. & Jennings, M. P. (2003). Genetic characterization of pilin glycosylation and phase variation in *Neisseria meningitidis*. *Mol Microbiol* **49**, 833–847.
- Rigel, N. W. & Braunstein, M. (2008). A new twist on an old pathway – accessory secretion systems. *Mol Microbiol* **69**, 291–302.
- Rose, L., Shivshankar, P., Hinojosa, E., Rodriguez, A., Sanchez, C. J. & 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. & 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.
- Schmidt, M. A., Riley, L. W. & Benz, I. (2003). Sweet new world: glycoproteins in bacterial pathogens. *Trends Microbiol* **11**, 554–561.
- Seifert, K. N., Adderson, E. E., Whiting, A. A., Bohnsack, J. F., Crowley, P. J. & 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.
- Siboo, I. R., Chambers, H. F. & 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. & Sullam, P. M. (2008). Characterization of the accessory Sec system of *Staphylococcus aureus*. *J Bacteriol* **190**, 6188–6196.
- Stephenson, A. E., Wu, H., Novak, J., Tomana, M., Mintz, K. & 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.
- Szymanski, C. M. & Wren, B. W. (2005). Protein glycosylation in bacterial mucosal pathogens. *Nat Rev Microbiol* **3**, 225–237.
- Szymanski, C. M., Logan, S. M., Linton, D. & Wren, B. W. (2003). Campylobacter – a tale of two protein glycosylation systems. *Trends Microbiol* **11**, 233–238.
- Takahashi, Y., Konishi, K., Cisar, J. O. & 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. & 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.
- Takahashi, Y., Takashima, E., Shimazu, K., Yagishita, H., Aoba, T. & Konishi, K. (2006). Contribution of sialic acid-binding adhesin to pathogenesis of experimental endocarditis caused by *Streptococcus gordonii* DL1. *Infect Immun* **74**, 740–743.
- Takamatsu, D., Bensing, B. A. & Sullam, P. M. (2004a). 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. & Sullam, P. M. (2004b). 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., Cheng, H., Jarvis, G. A., Siboo, I. R., Lopez, J. A., Griffiss, J. M. & Sullam, P. M. (2005a). Binding of the *Streptococcus gordonii* surface glycoproteins GspB and Hsa to specific carbohydrate structures on platelet membrane glycoprotein Ib $\alpha$ . *Mol Microbiol* **58**, 380–392.
- Takamatsu, D., Bensing, B. A. & 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. & 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., Yuzawa, H., Ito, T., Morimoto, Y., Kuroda, M., Cui, L., Takahashi, M. & other authors (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.
- Tettelin, H., Nelson, K. E., Paulsen, I. T., Eisen, J. A., Read, T. D., Peterson, S., Heidelberg, J., DeBoy, R. T., Haft, D. H. & other authors (2001). Complete genome sequence of a virulent isolate of *Streptococcus pneumoniae*. *Science* **293**, 498–506.
- Upreti, R. K., Kumar, M. & Shankar, V. (2003). Bacterial glycoproteins: functions, biosynthesis and applications. *Proteomics* **3**, 363–379.
- VanderVen, B. C., Harder, J. D., Crick, D. C. & Belisle, J. T. (2005). Export-mediated assembly of mycobacterial glycoproteins parallels eukaryotic pathways. *Science* **309**, 941–943.
- Vickerman, M. M., Iobst, S., Jesionowski, A. M. & Gill, S. R. (2007). Genome-wide transcriptional changes in *Streptococcus gordonii* in response to competence signaling peptide. *J Bacteriol* **189**, 7799–7807.
- Wacker, M., Linton, D., Hitchen, P. G., Nita-Lazar, M., Haslam, S. M., North, S. J., Panico, M., Morris, H. R., Dell, A. & other authors (2002). N-linked glycosylation in *Campylobacter jejuni* and its functional transfer into *E. coli*. *Science* **298**, 1790–1793.
- Weerkamp, A. H., Handley, P. S., Baars, A. & Slot, J. W. (1986a). Negative staining and immunoelectron microscopy of adhesion-deficient mutants of *Streptococcus salivarius* reveal that the adhesive protein antigens are separate classes of cell surface fibril. *J Bacteriol* **165**, 746–755.
- Weerkamp, A. H., van der Mei, H. C. & Liem, R. S. (1986b). Structural properties of fibrillar proteins isolated from the cell surface and cytoplasm of *Streptococcus salivarius* (K<sup>+</sup>) cells and nonadhesive mutants. *J Bacteriol* **165**, 756–762.
- Wu, H. & 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. & Fives-Taylor, P. M. (2001). Molecular strategies for fimbrial expression and assembly. *Crit Rev Oral Biol Med* **12**, 101–115.
- Wu, H., Mintz, K. P., Ladha, M. & 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. & 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. & 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. & 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., Brown, A., Chen, Z., Ozaki, L. S., Manque, P., Ge, X., Serrano, M. G. & other authors (2007). Genome of the opportunistic pathogen *Streptococcus sanguinis*. *J Bacteriol* **189**, 3166–3175.
- Yajima, A., Urano-Tashiro, Y., Shimazu, K., Takashima, E., Takahashi, Y. & Konishi, K. (2008). Hsa, an adhesin of *Streptococcus gordonii* DL1, binds to  $\alpha$ 2-3-linked sialic acid on glycophorin A of the erythrocyte membrane. *Microbiol Immunol* **52**, 69–77.
- Young Lee, S., Cisar, J. O., Bryant, J. L., Eckhaus, M. A. & Sandberg, A. L. (2006). Resistance of *Streptococcus gordonii* to polymorphonuclear leukocyte killing is a potential virulence determinant of infective endocarditis. *Infect Immun* **74**, 3148–3155.
- Zhang, Y. Q., Ren, S. X., Li, H. L., Wang, Y. X., Fu, G., Yang, J., Qin, Z. Q., Miao, Y. G., Wang, W. Y. & other authors (2003). Genome-based analysis of virulence genes in a non-biofilm-forming *Staphylococcus epidermidis* strain (ATCC 12228). *Mol Microbiol* **49**, 1577–1593.
- Zhou, M., Peng, Z., Fives-Taylor, P. & Wu, H. (2008). A conserved C-terminal thirteen amino acid motif of Gap1 is required for the Gap1 function and necessary for biogenesis of a serine-rich glycoprotein of *Streptococcus parasanguinis*. *Infect Immun* **76**, 5624–5631.