Generic determinants of *Streptococcus* colonization and infection

Angela H. Nobbs¹, Howard F. Jenkinson^{1*}, Dean B. Everett^{2,3}

¹ *School of Oral and Dental Sciences, University of Bristol, Lower Maudlin Street, Bristol BS1 2LY, UK*

² *Department of Clinical Infection, Microbiology and Immunology, Institute of Infection and Global Health, University of Liverpool, The Ronald Ross Building, 8 West Derby Street, Liverpool L69 7BE, UK*

³ *Malawi-Liverpool-Wellcome Trust, Clinical Research Programme, PO Box 30096, Chichiri, Blantyre 3, Malawi*

*Corresponding author; Tel: +44 117 342 4424; Fax: +44 117 342 4313

Email address: howard.jenkinson@bristol.ac.uk

*Corresponding author

Contact details

Angela H. Nobbs

Lecturer in Oral Microbiology

School of Oral and Dental Sciences
University of Bristol
Lower Maudlin Street
Bristol BS1 2LY, UK
Tel: 0117 342 4779
Fax: 0117 342 4313

Howard F. Jenkinson

Professor of Oral Microbiology

School of Oral and Dental Sciences
University of Bristol
Lower Maudlin Street
Bristol BS1 2LY, UK
Tel: 0117 342 4424
Fax: 0117 342 4313

Dean B. Everett

Senior Lecturer in Molecular Microbiology

Malawi-Liverpool-Wellcome Trust
Clinical Research Programme
PO Box 30096
Chichiri, Blantyre 3
Malawi
Tel: +265 1876444
Mob: +265 999016851
Fax: +265 1875774
www.mlw.medcol.mw

Department of Clinical Infection, Microbiology and Immunology
Institute of Infection and Global Health
University of Liverpool
The Ronald Ross Building
8 West Derby Street
Liverpool L69 7BE, UK

ABSTRACT

Bacteria within the genus *Streptococcus* have evolved to become exquisitely adapted to the colonization of humans and animals. These bacteria predominantly live in harmony with their hosts, but all have capacity to cause disease should prevailing conditions allow. Streptococci express a myriad of colonization and virulence attributes that promote their survival at a variety of ecological sites. Many of these factors are surface-expressed adhesins that exhibit conservation at structural or functional levels across the genus. This reflects the importance of adherence interactions with a multitude of host substrata, such as epithelia or extracellular matrix components, to streptococcal survival. Other important factors are more restricted in their distribution, often conferring pathogenic capabilities associated with immune evasion or host tissue destruction. Evidence suggests that dissemination of these streptococcal attributes has frequently been driven by the movement of genetic material via lateral gene transfer, reflecting ecological pressures. Such recombination events have simultaneously facilitated extensive diversification, resulting in distinct tropisms at the species- or strain-level. These generic determinants offer significant potential as targets for combating streptococcal disease. However, this will depend upon better understanding of their mechanistic basis, and refined mapping of their distribution by epidemiological and metagenomic studies.

Contents

1. Introduction
 2. *Streptococcus* genomes
 - 2.1. Core genome
 - 2.2. Lateral gene transfer
 - 2.3. Mosaic genes
 - 2.4. Gene duplication
 3. Colonization factors
 - 3.1. Adhesins
 - 3.2. Matrix-interacting proteins
 - 3.3. Pili
 4. Virulence factors
 - 4.1. Antigenic diversity
 - 4.2. Proteases
 5. Conclusions
- Acknowledgements
- References

1. Introduction

There are over 100 defined species of *Streptococcus* and many of these are natural inhabitants of the human oral cavity and nasopharynx. A commensal microorganism is currently defined as being able to survive at harmony with the host in an arrangement that may be beneficial to one or both. By this definition, these streptococcal inhabitants are commensals, but they do all have the ability to cause some kind of disease, be it superficial or systemic, and so they may be better termed opportunistic pathogens. Since the environmental components present in the human mouth and nasopharynx are complex, the streptococci have evolved a myriad of factors and mechanisms for colonization of these sites. However, the structures and functions of these factors, and their mechanistic properties, have often been to a large degree well-conserved. Their genes have been selectively retained and evolved in the generation of species, and this has been mediated by DNA-mediated transformation, phage-mediated transduction, and conjugal transfer. There is now evidence to suggest that DNA-mediated transformation following the development of competence may be a widely shared mechanism in *Streptococcus* gene transfer (Mashburn-Warren et al., 2010), when it was previously believed that, for example, group A streptococci (GAS) were non-transformable.

In general, *Streptococcus* infections begin by interaction of the bacterial cells with host tissues through adhesin-receptor reactions (Fig.

1). Streptococci express a spectrum of adhesins that can exhibit broad specificity e.g. fibronectin binding, or unique specificity for a molecular ligand, such as sialyl T-antigen (Takamatsu et al., 2005). These adhesins often drive bacteria to the initial colonization site, for example the tooth surface, buccal or lingual mucosae, tonsils etc. that express the available receptors. The outcome of initial adherence is the growth of the bacterial cells, provided environmental conditions are favorable, to form an early biofilm (Nobbs et al., 2009). This may attract different species of bacteria to adhere and form a community, such as in the development of dental plaque (Wright et al., 2013), or stimulate host cellular tissue to become susceptible to bacterial cell invasion (Fig. 1). Intracellular streptococci may then remain in a benign (or latent) state, to re-emerge later (Fig. 1), divide within the tissue to generate localized damage, or become disseminated throughout the body and infect multiple organs. Differential gene expression accompanies these processes, but many of the adhesins and invasins show common structures and properties, and the transcriptional regulators involved share similar structures and recognition motifs.

This article will consider some, but by no means all, of the generic determinants in the genus *Streptococcus* that are involved in mediating colonization and virulence. The focus will be on known factors that are expressed by oral streptococci and that are shared across the genus that promote adhesion to host surfaces, from a genetic and functional standpoint.

2. *Streptococcus* genomes

The streptococci include human pathogens, commensals, animal pathogens, fish pathogens, and non-pathogenic species used within the dairy industry (e.g. *S. thermophilus*). A recent analysis of 46 genome sequences has provided insight into the evolutionary history of the genus. Gene gain/loss analyses revealed a dynamic pattern of genomic evolution. Firstly with a period of gene gain, and then with gene loss, the major groups of the genus diversified. Then, a period of gene expansion led to the origins of the present species. A large proportion of the pan-genome has experienced lateral gene transfer (LGT), but despite this, a number of biochemical characteristics of groups have been retained, suggesting genomic cohesiveness through time (Richards et al., 2014). Proteolysis appears to be a defining feature of the mitis group, and also enrichment for *N*-acetyltransferase activity, which has been associated with resistance to aminoglycoside antibiotics. The salivarius group was enriched for urease, urea metabolism, and transposase activity, suggesting potential for high levels of recombination.

2.1. Core genome

The *Streptococcus* genomes can be best described as consisting of core and accessory genomes. In *S. pneumoniae* (pneumococcus), for

example, the core genome represents those genes that are common to all human pneumococcal isolates. These include key colonization and virulence genes as well as genes that enable niche adaptation and regulation of metabolism in response to alterations in available carbon sources, cations and amino acids, and anaerobic conditions. The core genome size is dependent on the number of isolates used during the analysis, and this gradually decreases with the addition of new genomes until a plateau is reached. To reach the absolute plateau and thus the definitive core genome size for *S. pneumoniae*, a decaying function analysis has to be undertaken (Tettelin et al., 2005).

By fitting a double-exponential decaying function model it has recently been shown that meningitis- versus bacteremia-associated pneumococci share a common core set of genes and that there is no difference in the core genome content from these disease manifestations (unpublished). This supports the requirement for a range of previously described virulence factors for these meningitis and bacteremia causing pneumococci. This high-resolution view of the core genome suggests that despite considerable competency for genetic exchange, the pneumococcus is under considerable pressure to retain key components that are presumably advantageous for colonization and transmission, and are essential for access and survival following invasion. A better understanding of the tropism of the pneumococcus for the blood and the brain will require more detailed analysis of allelic variation and gene expression.

The accessory genome represents the set of genes specific to a particular isolate(s), which are not shared in their entirety across the whole pneumococcal population. These may include, for example, genes associated with antibiotic resistance or novel virulence determinants.

2.2. Lateral gene transfer

Lateral gene transfer (LGT) is a fundamental process in the genome evolution of bacteria (Gogarten and Townsend, 2005). LGT occurs by transformation, transduction or conjugation. It enables bacteria to evolve rapidly through the acquisition of novel genetic determinants, or genetic determinants that are homologous to existing DNA, which were not previously resident within the recipient's genome. The transfer of DNA via homologous recombination (HR) leads to the replacement of a region of the genome of a recipient cell by the corresponding region from the donor cell (Smith et al., 1991).

For pathogenic bacteria, interaction with the human immune system remains a constant battle and leads to some major changes in their genetic makeup. Adaptation of the pneumococcus, for example, to the host environment is facilitated by mutations and by frequent transfers of genetic material between isolates and across bacterial species (Hanage et al., 2009; Croucher et al., 2011; Golubchik et al., 2012). This genetic variability contributes to the considerable redundancy in the range of

tools available to the pneumococcus to target host receptors, overcome the mucosal barrier and survive within the nasopharynx (Jedrzejewski, 2001; Weiser et al., 2003; Bergmann and Hammerschmidt, 2006; Selva et al., 2009).

As such, LGT contributes to the remarkable plasticity of the pneumococcus through this intra- and inter-species genetic exchange. Due to this natural transformability LGT, for example, allows one serotype to belong to multiple genotypes, and equally, a single genotype can express different capsule genes and therefore produce different serotypes. This phenomenon is known as capsular switching (Scott et al., 1996; Coffey et al., 1999). Capsular serotype may be more important than genotype in the ability of pneumococci to cause invasive disease (Brueggermann et al., 2003), but a number of studies have demonstrated the importance of genotypes as well (Enright and Spratt, 1998).

Induction of competence is associated with the formation of distinct pilus structures on the cell surface of pneumococci. It is suggested that the pilus drills a channel across the cell wall, and that this becomes transiently opened by secretion of the pilus. This then allows an entry port for exogenous DNA to gain access to DNA receptors and uptake systems in the cytoplasmic membrane (Balaban et al., 2014).

Although there are many selective pressures upon the pneumococcus, the biggest impact over the last 50 years has been the worldwide consumption of antibiotics. This has driven the acquisition of

antibiotic resistance in the pneumococcus and across many different species of bacteria.

Another example of LGT with implications for streptococcal colonization and pathogenesis involves a section of DNA known as region of difference 2 (RD2). RD2 is an integrative conjugative element (ICE), a form of mobile genetic element with capacity to excise from a donor genome, transfer by conjugation to a recipient bacterium, and integrate onto the recipient genome (Wozniak & Waldor, 2010). RD2 was first shown to be carried by GAS strains belonging to serotype M28, a serotype associated with puerperal sepsis and neonatal infections, and was found to encode 7 cell wall-anchored proteins (Green et al., 2005). These included adhesins R28, which promotes binding to epithelial cells (Stålhammar-Carlemalm et al., 1999), and AspA, an Antigen I/II-family polypeptide associated with biofilm formation and immune evasion (Maddocks et al., 2011; discussed in more detail below). Genome analysis of RD2 implied acquisition by LGT (Green et al., 2005), and RD2-like elements have been found in group B streptococci (GBS), group C streptococci and group G streptococci (Brochet et al., 2008; Sitkiewicz et al., 2011). Moreover, intra- and inter-species dissemination of these elements has been demonstrated in vitro (Sitkiewicz et al., 2011; Puymège et al., 2013). It seems likely then that acquisition of colonization determinants has been driven by dissemination of RD2-like elements across both pathogenic and commensal streptococci.

2.3. *Mosaic genes*

Interspecies recombination results in mosaicism i.e. genes that are composed of alternate blocks of nucleotides derived from a particular donor and its recipient. These blocks may be the result of small fragments that have recombined within a gene or the result of almost the whole gene being replaced by incoming genetic material (Dowson et al., 1989). Whichever outcome, the new mosaic usually encodes a protein with a different activity to that of the original recipient and possibly to that of the donor. Following integration of advantageous mosaics, such as penicillin resistance conferred by alterations within genes encoding penicillin-binding proteins, these become fixed within the population (Coffey et al., 1995). Once fixed, these successful mosaics can then readily spread laterally from species to species.

2.4. *Gene duplication*

Bacteria are generally haploid in nature, possessing a single copy of most genes. Despite this, partial chromosomal duplications, known as merodiploids, allow prokaryotes to evolve new genes. The most common form is tandem duplication (TD), where duplicated regions remain adjacent in the chromosome. TDs are generally thought to form spontaneously by HR between direct repeat sequences such as insertion sequences. These processes can result in tandem duplication of entire

genes, or to duplication of repeat regions that encode the amino acid (aa) repeat blocks present in cell wall-anchored surface proteins, thus contributing to antigenic diversity (see below). In genetic epidemiology, tandem repeats are used to characterize pneumococcal populations. The Multiple-Locus Variable number tandem repeat Analysis (MLVA; <http://www.mlva.net>) uses naturally-occurring variation found among a number of tandem repeated DNA sequences in many different loci within the genome to achieve this.

3. Colonization factors

Although many of the virulence factors associated with streptococci have been well studied, surprisingly less is understood about the factors that determine adhesion and successful colonization. This process is of course fundamental to survival of the bacteria in a multispecies community and to establishment of an infection. The streptococci are generally very successful colonizers; at least 50% of the bacterial species establishing on a clean tooth surface are streptococci (Nyvad and Kilian, 1990), and although different sites in the oral cavity carry distinct microbiomes, streptococci are always the most prevalent (Xu et al., 2014).

Streptococcus species (principally *mitis*, *pseudopneumoniae*, *pneumoniae*) are less abundant in the nasopharynx, where *Moraxella* and *Haemophilus* species tend to dominate, but *Streptococcus* abundance increases in subjects with non-viral pneumonia (Sakwinska et al., 2014). In order for

the streptococci to survive in a complex microbial community they must compete effectively for adhesion sites and nutrients, resist inhibitory molecules such as bacteriocins and toxic metabolites produced by other bacteria, and cope with host defenses; both innate and acquired immune functions. The surface proteins that streptococci express orchestrate adhesion, sense environmental challenges, and assist in protection against host immune defenses. We will consider some of the cell wall-anchored polypeptides produced by streptococci and their structural and functional properties.

3.1. Adhesins

Cell wall-anchored polypeptides are major players in streptococcal adhesion and invasion. Different species carry varying numbers of genes encoding these proteins, recognized by the presence of cell wall anchorage motif LPXTG at the C-terminus (Schneewind and Missiakas, 2014). A characteristic of these proteins is that they contain tandem repetition of aa sequence blocks known as domain repeats. These provide protein diversity and functional complexity, and alterations in the numbers of repeats leads to changes in conformational epitopes (Lin *et al.*, 2012). An adhesin that is produced by oral streptococci, termed Antigen I/II, is important for colonization of the oral cavity by interacting with salivary pellicle, fibronectin, collagen, and other micro-organisms (Brady *et al.*, 2010) (Fig. 2). The structural regions are well conserved in

the polypeptide, but alternative functional properties have been acquired. For example, the Antigen I/II protein produced by M28 serotype *S. pyogenes* appears to have anti-phagocytic properties, in addition to a role in biofilm formation (Maddocks et al., 2011; Franklin et al., 2013). The related protein SspB from *S. gordonii* does not have anti-phagocytic properties. It is possible that this adhesin gene has been acquired by GAS and GBS, providing a competitive phenotype for colonization of mucosal surfaces in the presence of other streptococci expressing these polypeptides (Maddocks et al., 2011; Franklin et al., 2013). The Antigen I/II polypeptide produced by *S. mutans* has been the focus of a dental caries vaccine programme (Sun et al., 2012), but utilizing a colonization factor such as this might result in antibodies that affect colonization of other streptococci that express the same antigenic epitopes (Brady et al., 2010). Modulation of the resident microbiota as a result of clinical intervention is an issue that is recognized and being addressed (Dunne et al., 2013), especially now that it is possible to obtain microbiome data with relative ease.

CshA is an approximately 2,500 aa residue polypeptide produced by a number of different species of mitis streptococci (Elliott et al., 2003). As more and more streptococcal genomes are sequenced we will begin to see better the distribution of this gene family. The *cshA* gene is duplicated in *S. gordonii*, and the paralogous genes *cshA* and *cshB* are at different chromosomal sites. CshA is the structural component of fibrils of 60-80 nm in length that emanate from the streptococcal cell surface and

that mediate binding to fibronectin and to other oral bacteria in early plaque (McNab et al., 1999).

Serine-rich repeat proteins (SRRPs) are expressed by mitis group streptococci, GBS and some serotypes of pneumococci (Fig. 2). The Fap1 SRRP produced by *S. parasanguinis* is a component of surface fimbriae, while the SRRPs from *S. gordonii* and *S. sanguinis* mediate adhesion of bacteria to salivary pellicle (Zhou and Wu, 2011). The SRRPs are also virulence factors because once inside the bloodstream, streptococci utilize these proteins to adhere to platelets via integrin GPIb, with potential to form a thrombus and infective endocarditis (Xiong et al., 2008). The thrombus protects the bacteria from the host immune system and from circulating antibiotics. Binding to platelets also contributes to dissemination (Kahn et al., 2013). The SRRPs have a head and stalk structure, like Antigen I/II and CshA, with the head providing adhesion specificity and the stalk (in SRRPs) being glycosylated. The pneumococcal SRRP designated PsrP mediates adherence to Keratin 10 on the surface of lung cells, and generates biofilm-like structures in the lungs that resist immune clearance (Sanchez et al., 2010).

Glucosyltransferases (GTFs) are secreted enzymes that cleave sucrose and polymerize the glucose into α 1,3- or 1,6-linked glucans. These glucans mediate lectin-like adhesion between streptococcal cells through the activities of cell wall-anchored glucan-binding proteins. The glucan polysaccharides are part responsible for the matrix in dental

plaque (Koo et al., 2009). GTFs are conserved in the catalytic site structure and in substrate binding site sequences. Glucan production is a colonization mechanism for oral streptococci (mitis, mutans, anginosus and salivarius groups). In *S. mutans*, GTFs are essential for initiation and development of dental caries, and are therefore also virulence factors (Koo et al., 2009).

In summary, four families of adhesins have been considered that are different in acquisition profiles across the streptococci, and variously specific in their binding properties. What has determined the pattern of gene distribution is not clear, but we can guess that it relates to environmental niche (e.g. tooth surface; mucosal surface - buccal, lingual, palatal, tonsillar), the nature of intermicrobial competition in the niche, and to the complement of adhesins and colonization factors necessary for growth and survival. It should be noted that much of the information that we have about surface proteins comes from strains that have been cultivated from humans, and almost always from subjects with disease. Common laboratory strains of oral streptococci have often been isolated from carious lesions or from blood, GAS and GBS from hospital patients, and pneumococci from a wide range of ill children and elderly people. We have therefore been working with a relatively limited number of strains and these have not generally been streptococci that simply exist in relative harmony with their hosts. One of the most valuable aspects of microbiome studies and metagenomic sequencing is that there is potentially no bias for disease. With the advent of single cell genomic

sequencing it should be possible to readdress the taxonomic structure of the genus and map the distribution of known colonization and virulence factors across the streptococcal spectrum.

3.2. Matrix-interacting proteins

Fibronectin-binding proteins (Fbps) are produced by all streptococci. There are 7 known in pneumococci (Suits and Boraston, 2013) and at least 8 in GAS. Some of these have been identified as adhesins and invasins, mediating adhesion and uptake of bacteria into epithelial or endothelial cells (Yamaguchi et al., 2012). One mechanism of fibronectin-binding portrayed by these proteins is the structural alignment of the aa residue repeat regions with the type III repeats present in fibronectin, generating a “zipper” that promotes host cell invasion (Norris et al., 2011). Other Fbps display interactions with fibronectin that are not yet understood at the structural level (Mu et al., 2014). The adhesins that also mediate invasion, such as FbaB-type fibronectin-binding proteins (Amelung et al., 2011) are well conserved at the structural level across GAS and GBS, but have not been found in the oral streptococci. Conversely, proteins orthologous to the cell wall anchorless protein Pava produced by pneumococci are found in all streptococci. Pava binds to fibronectin and is essential for pneumococcal infection (Holmes et al., 2001), highlighting the fact that a conserved adhesin protein across the

genus may have alternate functions according to the species (or maybe even strain) in which it is expressed.

Collagen binding proteins (Cbps), or the Cna family of cell wall-anchored proteins, are structurally related and found across the streptococci (Fig. 2). They mediate adhesion to collagen-rich tissues and carry a head region (with collagen-binding function) and C-terminal aa residue repeats. The proteins bind collagen by the “hug” mechanism, in which initial interaction with collagen induces redirection of sub-domains to wrap around the collagen triple helix (Kang et al., 2013). These adhesins can also bind complement C1 complex recognition protein C1q at a collagen-like domain, and thus act as inhibitors of the classical complement pathway, providing a strategy for immune evasion. Recent evidence suggests that the *S. mutans* collagen-binding protein (Cnm) is glycosylated in a threonine-rich region, and that this renders the protein less susceptible to proteolysis (Avilés-Reyes et al., 2014)

α -Enolase is a cytoplasmic and a surface protein in bacteria. In pneumococcus, enolase is held at the cell surface by interaction with the choline-rich cell wall. Enolase binds plasminogen present on the epithelial or endothelial cell surface, thus mediating pneumococcal adhesion to host cells (Bergmann et al., 2013). In GAS, enolase has been shown to enhance plasminogen activation, as well as binding to laminin, fibronectin and collagens (Antikainen et al., 2007). Recombinant enolase from *Streptococcus sobrinus* conferred protection against *S. sobrinus*-induced

dental caries (Dinis et al., 2009). Enolase is thus a well-conserved polypeptide, structurally and functionally, expressed on the surface of streptococci, with an essential intracellular function in glycolysis. It is one of several anchorless cell surface proteins (glyceraldehyde-3-phosphate dehydrogenase is another) that have been considered as GAS vaccine candidates (Henningham et al., 2012). How the polypeptide is exported to the cell surface with no known secretion signal (leader peptide) remains a mystery for now.

3.3 Pili

Pili are filamentous appendages that typically extend 1-3 μM from the cell surface. While the presence of hair-like structures on streptococcal cells has been known since the 1970s, the genes encoding pili were only identified in GBS in 2005 (Lauer et al., 2005). Since then, homologous genes have been found in a number of streptococcal species, including human pathogens GAS (Mora et al., 2005) and *S. pneumoniae* (Bagnoli et al., 2008; Barocchi et al., 2006), animal pathogen *S. suis* (Fittipaldi et al., 2010), and commensal bacteria *S. sanguinis* (Okahashi et al., 2010), *S. mitis*, *S. oralis* (Zähner et al., 2011) and *S. gallolyticus* (Sillanpää et al., 2009; Danne et al., 2011). The genes encoding pili are clustered in loci known as pilus islands (PIs), with numbers of PI variants differing across species. Nine PI variants (otherwise known as FCT types) have been identified in GAS to date (Falugi et al., 2008), while GBS and

S. pneumoniae encode 3 or 2 PI variants respectively (Rosini et al., 2006; Bagnoli et al., 2008). Many of the PI variants are flanked by mobile genetic elements. It has been proposed therefore that PIs amongst streptococci may derive from a common ancestral region (Mandlik et al., 2008; Scott & Zähler, 2006), subsequently spreading as a result of LGT and/or phage transduction.

Specific gene organization varies across each PI variant (comprehensively reviewed by Kreikemeyer et al., 2011), but there are core constituents. First of these is the gene encoding the backbone (Bkb) pilus subunit, which forms the pilus protein shaft. Most PIs then also carry genes encoding 1 or 2 ancillary protein subunits (An1, An2). These often serve as the functional tip or cell wall anchor respectively of the pilus filament (Fig. 3). Notable exceptions to this is the PI of *S. suis*, which lacks any An subunit gene, and PI-2 of *S. pneumoniae*, for which the An1 gene is a pseudogene, in both instances resulting in pili comprising only Bkb subunits (Fig. 3). Within each PI are then 1-3 genes encoding SrtC family sortase transpeptidases. These enzymes covalently link Bkb and An1/2 subunits to form the pilus structure (Fig. 3), with final anchoring to the bacterial cell wall often being mediated by 'housekeeping' sortase A (comprehensively reviewed by Hendrickx et al., 2011). Several PIs also carry a signal peptidase gene. Where investigated, this has been shown to be essential for pilus expression on the cell surface, likely playing a chaperone-type role (Bagnoli et al., 2008; Zähler & Scott, 2008; Nakata et al., 2009; Young et al., 2014). Another

feature common to most PIs is the presence of one or more regulator genes, encoding either standalone regulators or two-component signal transduction systems (TCS). Two regulator families predominate, AraC and RALP, but the specific regulatory networks are complex and exhibit high interspecies variability (Kreikemeyer et al., 2011).

While PIs have diversified across and within streptococcal species, their widescale dissemination implies critical function. Supporting this, pili have been shown to promote attachment by streptococci to a range of substrata. These include epithelial cells of the respiratory (Barocchi et al., 2006; Nelson et al., 2007; Manetti et al., 2007), gastrointestinal (Pezzicoli et al., 2008) or genitourinary tracts (Pezzicoli et al., 2008; Sheen et al., 2011; Jiang et al., 2012), salivary glycoprotein-340 (Edwards et al., 2008), skin-derived keratinocytes (Abbot et al., 2007), and extracellular matrix proteins such as collagen (Kreikemeyer et al., 2005; Hilleringmann et al., 2008; Sillanpää et al., 2009) or fibronectin (Hilleringmann et al., 2008; Okahashi et al., 2010). Pili of GAS, GBS, *S. gallolyticus* and *S. sanguinis* have also been implicated in biofilm formation on biotic and abiotic surfaces (Manetti et al., 2007; Konto-Ghiorgi et al., 2009; Okahashi et al., 2010; Rinaudo et al., 2010; Danne et al., 2011; Kimura et al., 2012). Together, such properties have potential to promote streptococcal colonization of a variety of sites within the host, likely contributing to the promiscuous adhesive capabilities of this genus. Pili may also promote bacterial dissemination, with particular implications for spread of invasive disease (Fig. 1) (Maisey et al., 2007; Pezzicoli et al.,

2008; Orrskog et al., 2012), and there is evidence that pili can protect against host immune defenses (Maisey et al., 2008; Chattopadhyay et al., 2011; Papasergi et al., 2011; Jiang et al., 2012).

One reason for the recent surge in pilus-related studies is the discovery that pilus protein subunits can confer protection against pneumococcal, GBS and GAS disease in animal models, making them putative vaccine targets (Maione et al., 2005; Mora et al., 2005; Gianfaldoni et al., 2007). This offers significant potential, particularly for GAS and GBS, for which no vaccine currently exists. Nonetheless, there are two major challenges to address: PI variation and prevalence. As PIs have spread across streptococci they have diversified, both in terms of allelic variation and overall distribution, likely in response to the distinct environmental conditions encountered by each strain and contributing to their specific tissue tropisms. As a result, it may be difficult to design vaccines that confer sufficient protection against all circulating strains, although vaccines comprising combinations of PI variants may prove successful (Falugi et al., 2008; Margarit et al., 2009; Nuccitelli et al., 2011). For *S. pneumoniae*, only approximately 30% or 16% of clinical isolates have been shown to express PI-1 or PI-2 respectively (Bagnoli et al., 2008; Moschioni et al., 2008). It seems likely then that PI subunits may have to be combined with additional protective antigens to provide broad-spectrum coverage against pneumococcal disease. In Malawi, where there is a high pneumococcal disease burden, there was found to be a low frequency of invasive pilated pneumococci (14%). Nonetheless,

pilus gene diversity was similar to that seen in multiple global pneumococcal lineages, and all common serotypes with pilus were covered by the 13-valent conjugate vaccine based on capsular polysaccharide (Kulohoma et al., 2013).

4. Virulence factors

4.1. Antigenic diversity

The M protein is a fibrillar coiled-coil dimer that is linked to the cell wall in GAS and extends into the environment. The *emm* genes are well conserved, but heterogeneity in the aa sequence at the N-terminal region results in antigenic diversity (and is also the basis of M protein serotyping). *Emm* pattern typing distinguishes distinct chromosomal architectures (patterns A-C, D and E) based on the presence and arrangement of *emm* and *emm*-like genes in the GAS genome (McMillan et al., 2013). Most of the information on M-type protein structure and function is based on M6, an *emm* pattern A-C. The prototype M6 protein contains several blocks of aa residue repeats (denoted A, B, C and D repeats) and these variously interact with human tissue proteins such as fibronectin and fibrinogen, complement factors and immunoglobulins. M protein inhibits phagocytosis in the absence of opsonic antibodies, promotes adhesion to epithelial cells, and helps the bacteria evade innate immune defenses (Smeesters et al., 2010). However, extrapolations of M

protein structure and function based upon M6 are limited, and the group A-C *emm*-type of M6 only accounts for ~20% of *emm*-types. The A repeats (first 50 N-terminal aa residues) and B repeats are more likely to be found in the A-C and D group *emm*-types, but the sequences emanating from different *emm*-types rarely showed extensive sequence homology. The hypervariable A region evades antibody attack through antigenic variation and weak immunogenicity (Lannergård et al., 2011). However, all M proteins carry a C repeat region (1-5 repeats of a 35 aa residue block) and are highly conserved. This has meant that a vaccine based upon a 14 aa residue sequence within the C region has potential to target a wide range of M proteins (Bauer et al., 2012).

4.2. *Proteases*

SpeB is a cysteine protease required for GAS virulence, and a major secreted protein. It has broad range activity against ECM proteins, cytokines, complement components and immunoglobulins, amongst others. It also regulates function of other proteins by degrading or releasing them from the cell surface (Nelson et al., 2011). For example, the surface anchored GAS M1 protein binds fibrinogen and in stationary phase, SpeB cleaved M1 protein and eliminated fibrinogen binding. Inactivation of SpeB function during the shift to invasive disease activates M1-fibrinogen binding, increasing phagocyte resistance and proinflammatory activities (Anderson et al., 2014). SpeB also degrades

host proteins that target bacteria to autophagy, an important innate immune defense. As a result, M1T1 serotype GAS can replicate efficiently in the cytoplasm of epithelial cells, escape autophagy, and establish deep-seated colonization of tissues (Barnett et al., 2013).

SpeB expression is regulated directly by the CovR/S two-component system, and by at least 10 other regulatory proteins (Carroll and Musser, 2011). During infection, mutations in the CovR/S system can accumulate, abrogating SpeB expression, and leading to hypervirulence. In addition, the protease activity of SpeB is modulated by zinc or copper availability, since there are two metal-ion binding sites within the protease and one involves the catalytic site (Chella Krishnan et al., 2014). Because of the central role of SpeB in pathogenesis, a novel vaccine has been developed by combining inactive superantigen SpeA (pyrogenic exotoxin) with SpeB, which elicits neutralizing antibodies in mice (Morefield et al., 2014).

The C5a peptidase of GAS (ScpA) and GBS (ScpB) is a cell wall-anchored multifunctional protein that inhibits neutrophil chemotaxis by cleaving complement factor C5a. ScpB binds directly to epithelial cells and promotes invasion, as does ScpA (Brown et al., 2005). Recently-sequenced genomes of *S. equi* ssp. *zooepidemicus*, *S. iniae* and *S. suis* have revealed many virulence factor genes in common, including C5a peptidase genes (Locke et al., 2008; Baiano & Barnes, 2009; Wei et al., 2013), and the presence of insertion sequence elements in flanking DNA

of C5a peptidase genes for pyogenic streptococci implies spread of this gene by LGT amongst these strains (Franken et al., 2001). Since the C5a peptidase gene is a well-conserved antigen in the respective organisms, it is considered to be a promising vaccine candidate (Sagar et al., 2012).

5. Conclusions

The streptococci present unrelenting arrays of factors that promote their colonization of animal hosts, and some of the organisms express specific factors that encourage virulence. This article has considered some of the generic proteins that are delivered to the cell surface by streptococci that contribute to colonization. Since the sequences and structures have diverged across the genus, it might be possible to target individual polypeptides or specific epitopes to interfere with the colonization processes of those organisms against which we may require protection (for example GAS and GBS). The article has also included two protease virulence factors that are produced by pathogenic streptococci of the pyogenic group. These are considered to be crucial to the pathogenic processes employed by the organisms producing them to evade the immune system and destroy host tissues. It is appreciated that there are many other factors that streptococci produce and employ in infectious processes, but these were selected as generic to some of the designated species.

We have also focused on factors that are currently the subjects of vaccine research. The cell wall-associated proteins described have all been shown to be protective against disease caused by the expressing organisms in animal models of infection. However, antigenic variability in the polypeptides has hindered the development of a human vaccine for GAS, especially based upon M protein epitopes. A pilus vaccine for GBS is in late stages of development and is expected to be effective in protecting humans against 90% of GBS strains across the seven serotypes. The latest pneumococcal 13-conjugate vaccine is proving effective in protecting children against the vaccine capsular serotypes. However, it is shown that the vaccine provides selection for new serotypes not covered by the vaccine (Weinberger et al., 2011; Vila-Corcoles & Ochoa-Gondar, 2013), and that different vaccines will be required for different regions of the World because the endemic pneumococcal infection serotype profiles are currently unique (Feldman and Anderson, 2014). Peptide vaccines based upon GTF or Antigen I/II sequences applied intra-nasally show some promises to be effective in protecting against dental caries (Sun et al., 2012). However, there is much debate as to who such a vaccine should be available to, and the cost of such a vaccine for those societies that are the most susceptible to severe aggressive childhood caries. Overall, the streptococci account for a very significant burden of human disease in the World, yet their close associations with their hosts, developed conjointly over evolution, make controlling their activities most difficult, and very challenging.

Acknowledgements

Research by HFJ is supported by NIH (NIDCR) R01 DE016690, and DBE research is supported by the Gates Foundation OPP1023440.

References

- Abbot, E.L., Smith, W.D., Siou, G.P., Chiriboga, C., Smith, R.J., Wilson, J.A., Hirst, B.H., Kehoe, M.A., 2007. Pili mediate specific adhesion of *Streptococcus pyogenes* to human tonsil and skin. *Cell. Microbiol.* 9, 1822-1833.
- Amelung, S., Nerlich, A., Rohde, M., Spellerberg, B., Cole, J.N., Nizet, V., Chhatwal, G.S., Talay, S.R., 2011. The FbaB-type fibronectin-binding protein of *Streptococcus pyogenes* promotes specific invasion into endothelial cells. *Cell. Microbiol.* 13, 1200-1211.
- Anderson, E.L., Cole, J.N., Olson, J., Ryba, B., Ghosh, P., Nizet, V., 2014. The fibrinogen-binding M1 protein reduces pharyngeal cell adherence and colonization phenotypes of M1T1 group A *Streptococcus*. *J. Biol. Chem.* 289, 3539-3546.
- Antikainen, J., Kuparinen, V., Lähteenmäki, K., Korhonen, T.K., 2007. Enolases from gram-positive bacterial pathogens and commensal lactobacilli share functional similarity in virulence-associated traits. *FEMS Immunol. Med. Microbiol.* 51, 526-534.
- Avilés-Reyes, A., Miller, J.H., Simpson-Haidaris, P.J., Hagen, F.K., Abranches, J., Lemos, J.A., 2014. Modification of *Streptococcus mutans* Cnm by PgfS contributes to adhesion, endothelial cell invasion, and virulence. *J. Bacteriol.* 196, 2789-2797.
- Bagnoli, F., Moschioni, M., Donati, C., Dimitrovska, V., Ferlenghi, I., Facciotti, C., Muzzi, A., Giusti, F., Emolo, C., Sinisi, A., Hillerigmann, M., Pansegrau, W., Censini, S., Rappuoli, R., Covacci, A., Maignani, V., Barocchi, M.A., 2008. A second pilus type in *Streptococcus pneumoniae* is prevalent in emerging serotypes and mediates adhesion to host cells. *J. Bacteriol.* 190, 5480-5492.

- Baiano, J.C., Barnes, A.C., 2009. Towards control of *Streptococcus iniae*. Emerg. Infect. Dis. 15, 1891-1896.
- Balaban, M., Bättig, P., Muschiol, S., Tirier, S.M., Wartha, F., Normark, S., Henriques-Normark, B., 2014. Secretion of a pneumococcal type II secretion system pilus correlates with DNA uptake during transformation. Proc. Natl. Acad. Sci. USA 111, E758-765.
- Barnett, T.C., Liebl, D., Seymour, L.M., Gillen, C.M., Lim, J.Y., LaRock, C.N., Davies, M.R., Schulz, B.L., Nizet, V., Teasdale, R.D., Walker, M.J., 2013. The globally disseminated M1T1 clone of group A *Streptococcus* evades autophagy for intracellular replication. Cell Host Microbe 14, 675-682.
- Barocchi, M.A., Ries, J., Zogaj, X., Hemsley, C., Albiger, B., Kanth, A., Dahlberg, S., Fernebro, J., Moschioni, M., Massignani, V., Hultenby, K., Taddei, A.R., Beiter, K., Wartha, F., von Euler, A., Covacci, A., Holden, D.W., Normark, S., Rappuoli, R., Henriques-Normark, B., 2006. A pneumococcal pilus influences virulence and host inflammatory responses. Proc. Natl. Acad. Sci. USA 103, 2857-2862.
- Bauer, M.J., Georgousakis, M.M., Vu, T., Henningham, A., Hofmann, A., Rettel, M., Hafner, L.M., Sriprakash, K.S., McMillan, D.J., 2012. Evaluation of novel *Streptococcus pyogenes* vaccine candidates incorporating multiple conserved sequences from the C-repeat region of the M-protein. Vaccine 30, 2197-2205.
- Bergmann, S., Hammerschmidt, S., 2006. Versatility of pneumococcal surface proteins. Microbiology 152, 295-303.
- Bergmann, S., Schoenen, H., Hammerschmidt, S., 2013. The interaction between bacterial enolase and plasminogen promotes adherence of *Streptococcus pneumoniae* to epithelial and endothelial cells. Int. J. Med. Microbiol. 303, 452-462.

- Brady, L.J., Maddocks, S.E., Larson, M.R., Forsgren, N., Persson, K., Deivanayagam, C.C., Jenkinson, H.F., 2010. The changing faces of *Streptococcus* antigen I/II polypeptide family adhesins. *Mol. Microbiol.* 77, 276-286.
- Brochet, M., Couve, E., Glaser, P., Guedon, G., Payot, S., 2008. Integrative conjugative elements and related elements are major contributors to the genome diversity of *Streptococcus agalactiae*. *J. Bacteriol.* 190, 6913-6917.
- Brown, C.K., Gu, ZY, Matsuka, Y.V., Purushothaman, S.S. Winter, L.A., Cleary, P.P., Olmsted, S.B., Ohlendorf, D.H., Earhart, C.A., 2005. Structure of the streptococcal cell wall C5a peptidase. *Proc. Natl. Acad. Sci. USA* 102, 18391-18396.
- Brueggemann, A.B., Griffiths, D.T., Meats, E., Peto, .T, Crook, D.W., Spratt, B.G., 2003. Clonal relationships between invasive and carriage *Streptococcus pneumoniae* and serotype- and clone-specific differences in invasive disease potential. *J. Infect. Dis.* 187: 1424-1432.
- Carroll, R.K., Musser, J.M., 2011. From transcription to activation: how group A *Streptococcus*, the flesh-eating pathogen, regulates SpeB cysteine protease production. *Mol. Microbiol.* 81, 588-601.
- Chattopadhyay, D., Carey, A.J., Caliot, E., Webb, R.I., Layton, J.R., Wang, Y., Bohnsack, J.F., Adderson, E.E., Ulett, G.C., 2011. Phylogenetic lineage and pilus protein Spb1/SAN1518 affect opsonin-independent phagocytosis and intracellular survival of Group B *Streptococcus*. *Microbes Infect.* 13, 369-382.
- Chella Krishnan, K., Mukundan, S., Landero Figueroa, J.A., Caruso, J.A., Kotb, M., 2014. Metal-mediated modulation of streptococcal cysteine

- protease activity and its biological implications. *Infect. Immun.* 82, 2992-3001.
- Coffey, T., Daniels, M., Enright, C., Spratt, B., 1999. Serotype 14 variants of the Spanish penicillin-resistant serotype 9 V clone of *Streptococcus pneumoniae* arose by large recombination. *Microbiology* 145, 2023-2031.
- Coffey, T.J., Dowson, C.G., Daniels, M., Spratt, B.G., 1995. Genetics and molecular biology of beta-lactam-resistant pneumococci. *Microb. Drug Resist.* 1, 29-34.
- Croucher, N.J., Harris, S.R., Fraser, C., Quail, M.A., Burton, J., van der Linden, M., McGee, L., von Gottberg, A., Song, J.H., Ko, K.S., Pichon, B., Baker, S., Parry, C.M., Lambertsen, L.M., Shahinas, D., Pillai, D.R., Mitchell, T.J., Dougan, G., Tomasz, A., Klugman, K.P., Parkhill, J., Hanage, W.P., Bentley, S.D., 2011. Rapid pneumococcal evolution in response to clinical interventions. *Science* 331, 430-434.
- Danne, C., Entenza, J.M., Mallet, A., Briandet, R., Debarbouille, M., Nato, F., Glaser, P., Jouvion, G., Moreillon, P., Trieu-Cuot, P., Dramsi, S., 2011. Molecular characterization of a *Streptococcus gallolyticus* genomic island encoding a pilus involved in endocarditis. *J. Infect. Dis.* 204, 1960-1970.
- Dinis, M., Tavares, D., Veiga-Malta, I., Fonseca, A.J., Andrade, E.B., Trigo, G., Ribeiro, A., Videira, A., Cabrita, A.M., Ferreira, P., 2009. Oral therapeutic vaccination with *Streptococcus sobrinus* recombinant enolase confers protection against dental caries in rats. *J. Infect. Dis.* 199, 116-123.
- Dowson, C.G., Hutchison, A., Brannigan, J.A., George, R.C., Hansman, D., Linares, J., Tomasz, A., Smith, J.M., Spratt, B.G., 1989. Horizontal transfer of penicillin-binding protein genes in penicillin-resistant

- clinical isolates of *Streptococcus pneumoniae*. Proc. Natl. Acad. Sci. USA 86, 8842-8846.
- Dunne, E.M., Smith-Vaughan, H.C., Robins-Browne, R.M., Mulholland, E.K., Satzke, C., 2013. Nasopharyngeal microbial interactions in the era of pneumococcal conjugate vaccination. Vaccine 31, 2333-2342.
- Edwards, A.M., Manetti, A.G., Falugi, F., Zingaretti, C., Capo, S., Buccato, S., Bensi, G., Telford, J.L., Margarit, I., Grandi, G., 2008. Scavenger receptor gp340 aggregates group A streptococci by binding pili. Mol. Microbiol. 68, 1378-1394.
- Elliott, D., Harrison, E., Handley, P.S., Ford, S.K., Jaffray, E., Mordan, N., McNab, R., 2003. Prevalence of Csh-like fibrillar surface proteins among mitis group oral streptococci. Oral Microbiol. Immunol. 18, 114-120.
- Enright, M., Spratt, G.A., 1998. Multilocus sequence typing scheme for *Streptococcus pneumoniae*: identification of clones associated with serious invasive disease. Microbiology 144, 3049-3060.
- Falugi, F., Zingaretti, C., Pinto, V., Mariani, M., Amodeo, L., Manetti, A.G., Capo, S., Musser, J.M., Orefici, G., Margarit, I., Telford, J.L., Grandi, G., Mora, M., 2008. Sequence variation in group A *Streptococcus* pili and association of pilus backbone types with Lancefield T serotypes. J. Infect. Dis. 198, 1834-1841.
- Feldman, C., Anderson, R. 2014. Review: Current and new generation pneumococcal vaccines. J. Infect. [Epub ahead of print] doi: 10.1016/j.jinf.2014.06.006
- Fittipaldi, N., Takamatsu, D., de la Cruz Dominguez-Punaro, M., Lecours, M.P., Montpetit, D., Osaki, M., Sekizaki, T., Gottschalk, M., 2010. Mutations in the gene encoding the ancillary pilin subunit of the

Streptococcus suis *srtF* cluster result in pili formed by the major subunit only. PLoS One 5, e8426.

- Franken, C., Haase, G., Brandt, C., Weber-Heynemann, J., Martin, S., Lämmler, C., Podbielski, A., Lütticken, R., Spellerberg, B., 2001. Horizontal gene transfer and host specificity of beta-haemolytic streptococci: the role of a putative composite transposon containing *scpB* and *Imb*. Mol. Microbiol. 41, 925-935.
- Franklin, L., Nobbs, A.H., Bricio-Moreno, L., Wright, C.J., Maddocks, S.E., Sahota, J.S., Ralph, J., O'Connor, M., Jenkinson, H.F., Kadioglu, A., 2013. The AgI/II family adhesin AspA is required for respiratory infection by *Streptococcus pyogenes*. PLoS One 8, e62433.
- Gianfaldoni, C., Censini, S., Hilleringmann, M., Moschioni, M., Facciotti, C., Pansegrau, W., Massignani, V., Covacci, A., Rappuoli, R., Barocchi, M.A., Ruggiero, P., 2007. *Streptococcus pneumoniae* pilus subunits protect mice against lethal challenge. Infect. Immun. 75, 1059-1062.
- Gogarten, J.P., Townsend, J.P., 2005. Horizontal gene transfer, genome innovation and evolution. Nat. Rev. Microbiol. 3, 679-687.
- Golubchik, T., Brueggemann, A.B., Street, T., Gertz, R.E. Jr., Spencer, C.C., Ho, T., Giannoulatou, E., Link-Gelles, R., Harding, R.M., Beall, B., Peto, T.E., Moore, M.R., Donnelly, P., Crook, D.W., Bowden, R., 2012. Pneumococcal genome sequencing tracks a vaccine escape variant formed through a multi-fragment recombination event. Nat. Genet. 44, 352-355.
- Green, N.M., Zhang, S., Porcella, S.F., Nagiec, M.J., Barbian, K.D., Beres, S.B., LeFebvre, R.B., Musser, J.M., 2005. Genome sequence of a serotype M28 strain of group A *Streptococcus*: potential new insights into puerperal sepsis and bacterial disease specificity. J. Infect. Dis. 192, 760-770.

- Hanage, W.P., Fraser, C., Tang, J., Connor, T.R., Corander, J., 2009. Hyper-recombination, diversity, and antibiotic resistance in pneumococcus. *Science* 324, 1454-1457.
- Hendrickx, A.P., Budzik, J.M., Oh, S.Y., Schneewind, O., 2011. Architects at the bacterial surface - sortases and the assembly of pili with isopeptide bonds. *Nat. Rev. Microbiol.* 9, 166-176.
- Henningham, A., Chiarot, E., Gillen, C.M., Cole, J.N., Rohde, M., Fulde, M., Ramachandran, V., Cork, A.J., Hartas, J., Magor, G., Djordjevic, S.P., Cordwell, S.J., Kobe, B., Sriprakash, K.S., Nizet, V., Chhatwal, G.S., Margarit, I.Y.R., Batzloff, M.R., Walker, M.J., 2012. Conserved anchorless surface proteins as group A streptococcal vaccine candidates. *J. Mol. Med.* 90, 1197-1207.
- Hilleringmann, M., Giusti, F., Baudner, B.C., Massignani, V., Covacci, A., Rappuoli, R., Barocchi, M.A., Ferlenghi, I., 2008. Pneumococcal pili are composed of protofilaments exposing adhesive clusters of Rrg A. *PLoS Pathog.* 4, e1000026.
- Holmes, A.R., McNab, R., Millsap, K.W., Rohde, M., Hammerschmidt, S., Mawdsley, J.L., Jenkinson, H.F., 2001. The *pavA* gene of *Streptococcus pneumoniae* encodes a fibronectin-binding protein that is essential for virulence. *Mol. Microbiol.* 41, 1395-1408.
- Jedrzejewski, M.J., 2001. Pneumococcal virulence factors: structure and function. *Microbiol. Mol. Biol. Rev.* 65, 187-207.
- Jiang, S., Park, S.E., Yadav, P., Paoletti, L.C., Wessels, M.R., 2012. Regulation and function of pilus island 1 in group B *Streptococcus*. *J. Bacteriol.* 194, 2479-2490.
- Kahn, F., Hurley, S., Shannon, O., 2013. Platelets promote bacterial dissemination in a mouse model of streptococcal sepsis. *Microbes Infect.* 15, 669-676.

- Kang, M., Ko, Y-P., Liang, X., Ross, C.L., Liu, Q., Murray, B.E., Höök, M., 2013. Collagen-binding microbial surface components recognizing adhesive matrix molecule (MSCRAMM) of gram-positive bacteria inhibit complement activation via the classical pathway. *J. Biol. Chem.* 288, 20520-20531.
- Kimura, K.R., Nakata, M., Sumitomo, T., Kreikemeyer, B., Podbielski, A., Terao, Y., Kawabata, S., 2012. Involvement of T6 pili in biofilm formation by serotype M6 *Streptococcus pyogenes*. *J. Bacteriol.* 194, 804-812.
- Konto-Ghiorghi, Y., Mairey, E., Mallet, A., Dumenil, G., Caliot, E., Trieu-Cuot, P., Dramsi, S., 2009. Dual role for pilus in adherence to epithelial cells and biofilm formation in *Streptococcus agalactiae*. *PLoS Pathog.* 5, e1000422.
- Koo, H., Xiao, J., Klein, M.I., 2009. Extracellular polysaccharide matrix – an often forgotten virulence factor in oral biofilm research. *Int. J. Oral Sci.* 1, 229-234.
- Kreikemeyer, B., Gamez, G., Margarit, I., Giard, J.C., Hammerschmidt, S., Hartke, A., Podbielski, A., 2011. Genomic organization, structure, regulation and pathogenic role of pilus constituents in major pathogenic Streptococci and Enterococci. *Int. J. Med. Microbiol.* 301, 240-251.
- Kreikemeyer, B., Nakata, M., Oehmcke, S., Gschwendtner, C., Normann, J., Podbielski, A., 2005. *Streptococcus pyogenes* collagen type I-binding Cpa surface protein. Expression profile, binding characteristics, biological functions, and potential clinical impact. *J. Biol. Chem.* 280, 33228-33239.
- Kulohoma, B.W., Gray, K., Kamng'ona, A., Cornick, J., Bentley, S.D., Heyderman, R.S., Everett, D.B., 2013. Piliation of invasive

- Streptococcus pneumoniae* isolates in the era before pneumococcal conjugate vaccine introduction in Malawi. Clin. Vaccine Immunol. 20, 1729-1735.
- Lannergård, J., Gustafsson, M.C.U., Waldemarsson, J., Norrby-Teglund, A., Stålhammar-Carlemalm, M., Lindahl, G., 2011. The hypervariable region of *Streptococcus pyogenes* M protein escapes antibody attack by antigenic variation and weak immunogenicity. Cell Host Microbe 10, 147-157.
- Lauer, P., Rinaudo, C.D., Soriani, M., Margarit, I., Maione, D., Rosini, R., Taddei, A.R., Mora, M., Rappuoli, R., Grandi, G., Telford, J.L., 2005. Genome analysis reveals pili in Group B *Streptococcus*. Science 309, 105.
- Lin, I-H., Hsu, M-T., Chang, C-H., 2012. Protein domain repetition is enriched in streptococcal cell-surface proteins. Genomics 100, 370-379.
- Locke, J.B., Aziz, R.K., Vicknair, M.R., Nizet, V., Buchanan, J.T., 2008. *Streptococcus iniae* M-like protein contributes to virulence in fish and is a target for live attenuated vaccine development. PLoS One 3, e2824.
- Maddocks, S.E., Wright, C.J., Nobbs, A.H., Brittan, J.L., Franklin, L., Stromberg, N., Kadioglu, A., Jepson, M.A., Jenkinson, H.F., 2011. *Streptococcus pyogenes* antigen I/II-family polypeptide AspA shows differential ligand-binding properties and mediates biofilm formation. Mol. Microbiol. 81, 1034-1049.
- Maione, D., Margarit, I., Rinaudo, C.D., Massignani, V., Mora, M., Scarselli, M., Tettelin, H., Brettoni, C., Iacobini, E.T., Rosini, R., D'Agostino, N., Miorin, L., Buccato, S., Mariani, M., Galli, G., Nogarotto, R., Nardi-Dei, V., Vegni, F., Fraser, C., Mancuso, G., Teti, G., Madoff, L.C.,

- Paoletti, L.C., Rappuoli, R., Kasper, D.L., Telford, J.L., Grandi, G., 2005. Identification of a universal Group B *Streptococcus* vaccine by multiple genome screen. *Science* 309, 148-150.
- Maisey, H.C., Hensler, M., Nizet, V., Doran, K.S., 2007. Group B streptococcal pilus proteins contribute to adherence to and invasion of brain microvascular endothelial cells. *J. Bacteriol.* 189, 1464-1467.
- Maisey, H.C., Quach, D., Hensler, M.E., Liu, G.Y., Gallo, R.L., Nizet, V., Doran, K.S., 2008. A group B streptococcal pilus protein promotes phagocyte resistance and systemic virulence. *FASEB J.* 22, 1715-1724.
- Mandlik, A., Swierczynski, A., Das, A., Ton-That, H., 2008. Pili in Gram-positive bacteria: assembly, involvement in colonization and biofilm development. *Trends Microbiol.* 16, 33-40.
- Manetti, A.G., Zingaretti, C., Falugi, F., Capo, S., Bombaci, M., Bagnoli, F., Gambellini, G., Bensi, G., Mora, M., Edwards, A.M., Musser, J.M., Graviss, E.A., Telford, J.L., Grandi, G., Margarit, I., 2007. *Streptococcus pyogenes* pili promote pharyngeal cell adhesion and biofilm formation. *Mol. Microbiol.* 64, 968-983.
- Margarit, I., Rinaudo, C.D., Galeotti, C.L., Maione, D., Ghezzi, C., Buttazzoni, E., Rosini, R., Runci, Y., Mora, M., Buccato, S., Pagani, M., Tresoldi, E., Berardi, A., Creti, R., Baker, C.J., Telford, J.L., Grandi, G., 2009. Preventing bacterial infections with pilus-based vaccines: the group B *Streptococcus* paradigm. *J. Infect. Dis.* 199, 108-115.
- Mashburn-Warren, L., Morrison, D.A., Federle, M.J., 2010. A novel double-tryptophan peptide pheromone controls competence in *Streptococcus* spp. via an Rgg regulator. *Mol. Microbiol.* 78, 589-606.

- McMillan, D.J., Dreze, P.-A., Vu, T., Besseb, D.E., Guglielmimi, J., Steer, A.C., Carapetis, J.R., Van Melder, L., Sriprakash, K.S., Smeesters, P.R. for the M Protein Study Group, 2013. Updated model of group A *Streptococcus* M proteins based on a comprehensive worldwide study. Clin. Microbiol. Infect. 19, E222-E229.
- McNab, R., Forbes, H., Handley, P.S., Loach, D.M., Tannock, G.W., Jenkinson, H.F., 1999. Cell wall-anchored polypeptide CshA (259 kilodaltons) in *Streptococcus gordonii* forms surface fibrils that confer hydrophobic and adhesive properties. J. Bacteriol. 181, 3087-3095.
- Mora, M., Bensi, G., Capo, S., Falugi, F., Zingaretti, C., Manetti, A.G., Maggi, T., Taddei, A.R., Grandi, G., Telford, J.L., 2005. Group A *Streptococcus* produce pilus-like structures containing protective antigens and Lancefield T antigens. Proc. Natl. Acad. Sci. USA 102, 15641-15646.
- Morefield, G., Touhey, G., Lu, F., Dunham, S., HogenEsch, H., 2014. Development of a recombinant fusion protein vaccine formulation to protect against *Streptococcus pyogenes*. Vaccine 32, 3810-3815.
- Moschioni, M., Donati, C., Muzzi, A., Massignani, V., Censini, S., Hanage, W.P., Bishop, C.J., Reis, J.N., Normark, S., Henriques-Normark, B., Covacci, A., Rappuoli, R., Barocchi, M.A., 2008. *Streptococcus pneumoniae* contains 3 *rlrA* pilus variants that are clonally related. J. Infect. Dis. 197, 888-896.
- Mu, R., Kim, B.J., Paco, C., Del Rosario, Y., Courtney, H.S., Doran, K.S., 2014. Identification of a group B streptococcal fibronectin binding protein SfbA, that contributes to invasion of brain endothelium and development of meningitis. Infect. Immun. 82, 2276-2286.
- Nakata, M., Koller, T., Moritz, K., Ribardo, D., Jonas, L., McIver, K.S., Sumitomo, T., Terao, Y., Kawabata, S., Podbielski, A., Kreikemeyer,

- B., 2009. Mode of expression and functional characterization of FCT-3 pilus region-encoded proteins in *Streptococcus pyogenes* serotype M49. *Infect. Immun.* 77, 32-44.
- Nelson, A.L., Ries, J., Bagnoli, F., Dahlberg, S., Falker, S., Rounioja, S., Tschop, J., Morfeldt, E., Ferlenghi, I., Hilleringmann, M., Holden, D.W., Rappuoli, R., Normark, S., Barocchi, M.A., Henriques-Normark, B., 2007. RrgA is a pilus-associated adhesin in *Streptococcus pneumoniae*. *Mol. Microbiol.* 66, 329-340.
- Nelson, D.C., Garbe, J., Collin, M., 2011. Cysteine protease SpeB from *Streptococcus pyogenes* – a potent modifier of immunologically important host and bacterial proteins. *Biol. Chem.* 392, 1077-1088.
- Nobbs, A.H., Lamont, R.J., Jenkinson, H.F., 2009. *Streptococcus* adherence and colonization. *Microbiol. Mol. Biol. Rev.* 73, 407-450.
- Norris, N.C., Bingham, R.J., Harris, G., Speakman, A., Jones, R.P., Leech, A., Turkenburg, J.P., Potts, J.R., 2011. Structural and functional analysis of the tandem β -zipper interaction of a streptococcal protein with human fibronectin. *J. Biol. Chem.* 286, 38311-38320.
- Nuccitelli, A., Cozzi, R., Gourlay, L.J., Donnarumma, D., Necchi, F., Norais, N., Telford, J.L., Rappuoli, R., Bolognesi, M., Maione, D., Grandi, G., Rinaudo, C.D., 2011. Structure-based approach to rationally design a chimeric protein for an effective vaccine against Group B *Streptococcus* infections. *Proc. Natl. Acad. Sci. USA* 108, 10278-10283.
- Nyvad, B., Kilian, M., 1990. Comparison of the initial streptococcal microflora on dental enamel in caries-active and in caries-inactive individuals. *Caries Res.* 24, 267-272.
- Okahashi, N., Nakata, M., Sakurai, A., Terao, Y., Hoshino, T., Yamaguchi, M., Isoda, R., Sumitomo, T., Nakano, K., Kawabata, S., Ooshima, T.,

2010. Pili of oral *Streptococcus sanguinis* bind to fibronectin and contribute to cell adhesion. *Biochem. Biophys. Res. Commun.* 391, 1192-1196.
- Orrskog, S., Rounioja, S., Spadafina, T., Gallotta, M., Norman, M., Hentrich, K., Falker, S., Ygberg-Eriksson, S., Hasenberg, M., Johansson, B., Uotila, L.M., Gahmberg, C.G., Barocchi, M., Gunzer, M., Normark, S., Henriques-Normark, B., 2012. Pilus adhesin RrgA interacts with complement receptor 3, thereby affecting macrophage function and systemic pneumococcal disease. *mBio* 4, e00535-00512.
- Papasergi, S., Brega, S., Mistou, M.Y., Firon, A., Oxaran, V., Dover, R., Teti, G., Shai, Y., Trieu-Cuot, P., Dramsi, S., 2011. The GBS PI-2a pilus is required for virulence in mice neonates. *PLoS One* 6, e18747.
- Pezzicoli, A., Santi, I., Lauer, P., Rosini, R., Rinaudo, D., Grandi, G., Telford, J.L., Soriani, M., 2008. Pilus backbone contributes to group B *Streptococcus* paracellular translocation through epithelial cells. *J. Infect. Dis.* 198, 890-898.
- Puymège, A., Bertin, S., Chuzeville, S., Guedon, G., Payot, S., 2013. Conjugative transfer and *cis*-mobilization of a genomic island by an integrative and conjugative element of *Streptococcus agalactiae*. *J. Bacteriol.* 195, 1142-1151.
- Richards V.P., Palmer, S.R., Pavinski Bitar, P.D., Qin, X., Weinstock, G.M., Highlander, S.K., Town, C.D., Burne, R.A., Stanhope, M.J., 2014. Phylogenomics and the dynamic genome evolution of the genus *Streptococcus*. *Genome Biol. Evol.* 6, 741-753.
- Rinaudo, C.D., Rosini, R., Galeotti, C.L., Berti, F., Necchi, F., Reguzzi, V., Ghezzi, C., Telford, J.L., Grandi, G., Maione, D., 2010. Specific

involvement of pilus type 2a in biofilm formation in group B *Streptococcus*. PLoS One 5, e9216.

Rosini, R., Rinaudo, C.D., Soriani, M., Lauer, P., Mora, M., Maione, D., Taddei, A., Santi, I., Ghezzi, C., Brettoni, C., Buccato, S., Margarit, I., Grandi, G., Telford, J.L., 2006. Identification of novel genomic islands coding for antigenic pilus-like structures in *Streptococcus agalactiae*. Mol. Microbiol. 61, 126-141.

Sagar, V., Bergmann, R., Nerlich, A., McMillan, D.J., Nitsche Schmitz, D.P., Chhatwal, G.S., 2012. Variability in the distribution of genes encoding virulence factors and putative extracellular proteins of *Streptococcus pyogenes* in India, a region with high streptococcal disease burden, and implication for development of a regional multisubunit vaccine. Clin. Vaccine Immunol. 19, 1818-1825.

Sakwinska, O., Schmid, V.B., Berger, B., Bruttin, A., Keitel, K., Lepage, M., Moine, D., Bru, C.N., Brussow, H., Gervaix, A., 2014. Nasopharyngeal microbiota in healthy children and pneumonia patients. J. Clin. Microbiol. 52, 1590-1594.

Sanchez, C.J., Shivshankar, P., Stol, K., Trakhtenbroit, S., Sullam, P.M., Sauer, K., Hermans, P.W., Orihuela, C.J., 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.

Schneewind, O., Missiakas, D., 2014. Sec-secretion and sortase-mediated anchoring of proteins in Gram-positive bacteria. Biochim. Biophys. Acta 1843, 1687-1697.

Scott, J.A.G., Hall, A.J., Dagan, R., Dixon, J.M.S., Eykyn, S.J., Fenoll, A., Hortal, M., Jette, L.P., Jorgensen, J.H., Lamothe, F., Latorre, C., Macfarlane, J.T., Shlaes, D.M., Smart, L.E., Taunay, A. 1996.

- Serogroup-specific epidemiology of *Streptococcus pneumoniae* - associations with age, sex, and geography in 7,000 episodes of invasive disease. Clin. Infect. Dis. 22, 973-981.
- Scott, J.R., Zähler, D., 2006. Pili with strong attachments: Gram-positive bacteria do it differently. Mol. Microbiol. 62, 320-330.
- Selva, L., Viana, D., Regev-Yochay, G., Trzcinski, K., Corpa, J.M., Lasa, I., Novick, R.P., Penadés, J.R. (2009). Killing niche competitors by remote-control bacteriophage induction. Proc. Natl. Acad. Sci. USA 106, 1234-1238.
- Sheen, T.R., Jimenez, A., Wang, N.Y., Banerjee, A., van Sorge, N.M., Doran, K.S., 2011. Serine-rich repeat proteins and pili promote *Streptococcus agalactiae* colonization of the vaginal tract. J. Bacteriol. 193, 6834-6842.
- Sillanpää, J., Nallapareddy, S.R., Qin, X., Singh, K.V., Muzny, D.M., Kovar, C.L., Nazareth, L.V., Gibbs, R.A., Ferraro, M.J., Steckelberg, J.M., Weinstock, G.M., Murray, B.E., 2009. A collagen-binding adhesin, Acb, and ten other putative MSCRAMM and pilus family proteins of *Streptococcus gallolyticus* subsp. *gallolyticus* (*Streptococcus bovis* Group, biotype I). J. Bacteriol. 191, 6643-6653.
- Sitkiewicz, I., Green, N.M., Guo, N., Mereghetti, L., Musser, J.M., 2011. Lateral gene transfer of streptococcal ICE element RD2 (region of difference 2) encoding secreted proteins. BMC Microbiol. 11, 65.
- Smeesters, P.R., McMillan, D.J., Sriprakash, K.S., 2010. The streptococcal M protein: a highly versatile molecule. Trends Microbiol. 18, 275-282.
- Smith, J.M., Dowson, C.G., Spratt, B.G., 1991. Localized sex in bacteria. Nature 349, 29-33.

- Stålhammar-Carlemalm, M., Areschoug, T., Larsson, C., Lindahl, G., 1999. The R28 protein of *Streptococcus pyogenes* is related to several group B streptococcal surface proteins, confers protective immunity and promotes binding to human epithelial cells. *Mol. Microbiol.* 33, 208-219.
- Suits, M.D., Boraston, A.B., 2013. Structure of the *Streptococcus pneumoniae* surface protein and adhesin PfbA. *PLoS One* 8, e67190.
- Sun, Y., Shi, W., Yang, J.Y., Zhou, D.H., Chen, Y.Q., Zhang, Y., He, B.X., Zhong, M.H., Li, Y.M., Cao, Y., Xiao, Y., Li, W., Yu, J., Li, Y.H., Fan, M.W., Yan, H.M. 2012. Flagellin-Pac fusion protein is a high-efficiency anti-caries mucosal vaccine. *J. Dent. Res.* 91, 941-947.
- Takamatsu, D., Bensing, B.A., Cheng, H., Jarvis, G.A., Siboo, I.R., López, J.A., Griffiss, J.M., Sullam, P.M., 2005. Binding of the *Streptococcus gordonii* surface glycoproteins GspB and Hsa to specific carbohydrate structures on platelet membrane glycoprotein Ib α . *Mol. Microbiol.* 58, 380-392.
- Tettelin, H., and 46 others, 2005. Genome analysis of multiple pathogenic isolates of *Streptococcus agalactiae*: implications for the microbial "pan-genome". *Proc. Natl. Acad. Sci. USA* 102, 13950-13955.
- Vila-Corcoles, A., Ochoa-Gondar, O., 2013. Preventing pneumococcal disease in the elderly: recent advances in vaccines and implications for clinical practice. *Drugs Aging.* 30, 263-276.
- Wei, Z., Fu, Q., Chen, Y., Li, M., Cong, P., No, D., Liu, X., 2013. *Streptococcus equi* ssp. *zooepidemicus* C5a peptidase, a putative invasin, induces protective immune response in mice. *Res. Vet. Sci.* 95, 444-450.
- Weinberger, D.M., Malley, R., Lipsitch, M. 2011. Serotype replacement in disease after pneumococcal vaccination. *Lancet* 378, 1962-1973.

- Weiser, J.N., Bae, D., Fasching, C., Scamurra, R.W., Ratner, A.J., Janoff, E.N., 2003. Antibody-enhanced pneumococcal adherence requires IgA1 protease. *Proc. Natl. Acad. Sci. USA* 100, 4215-4220.
- Wozniak, R.A., Waldor, M.K., 2010. Integrative and conjugative elements: mosaic mobile genetic elements enabling dynamic lateral gene flow. *Nat. Rev. Microbiol.* 8, 552-563.
- Wright, C.J., Burns, L.H., Jack, A.A., Back, C.R., Dutton, L.C., Nobbs, A.H., Lamont, R.J., Jenkinson, H.F., 2013. Microbial interactions in building of communities. *Mol. Oral Microbiol.* 28, 83-101.
- 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, X., He, J., Xue, J., Wang, Y., Li, K., Zhang, K., Guo, Q., Liu, X., Zhou, Y., Cheng, L., Li, M., Li, Y., Li, Y., Shi, W., Zhou, X., 2014. Oral cavity contains distinct niches with dynamic microbial communities. *Environ. Microbiol.* 16, Epub ahead of print.
- Yamaguchi, M., Terao, Y., Kawabata, S., 2012. Pleiotropic virulence factor – *Streptococcus pyogenes* fibronectin-binding proteins. *Cell. Microbiol.* 15, 503-511.
- Young, P.G., Proft, T., Harris, P.W., Brimble, M.A., Baker, E.N., 2014. Structure and Activity of *Streptococcus pyogenes* SipA: A Signal Peptidase-Like Protein Essential for Pilus Polymerisation. *PLoS One* 9, e99135.
- Zähner, D., Gandhi, A.R., Yi, H., Stephens, D.S., 2011. Mitis group streptococci express variable pilus islet 2 pili. *PLoS One* 6, e25124.

- Zähner, D., Scott, J.R., 2008. SipA is required for pilus formation in *Streptococcus pyogenes* serotype M3. J. Bacteriol. 190, 527-535.
- Zhou, M., Wu, H., 2011. Glycosylation and biogenesis of a family of serine-rich bacterial adhesins. Microbiology 155, 317-327.

Figure legends

Fig. 1. Model for the establishment of longer-term mucosal colonization by streptococci. Following adhesion of bacteria to epithelium, and initial transient colonization, depletion of bacterial cell numbers occurs associated with host immune responses. These include innate factors, such as antimicrobial peptides, agglutinins, mucins, neutrophils and macrophages. A small number of streptococci successfully evade these responses, perhaps associated with internalization by epithelial cells. Bacteria may reside dormant within epithelial cells, or replicate, or become internalized by macrophages within which they survive and become spread systemically. Expression of factors that have anti-phagocytic properties and biofilm-enhancing activities may allow the streptococci to re-emerge to form a biofilm community, from which streptococci may be dispersed to colonize new sites.

Fig. 2. Summary of some generic determinants of colonization and virulence in the *Streptococcus* genus. Cell wall-anchored proteins (CWPs) are shown as linked into the outer layers. The known substrates for the protein families are indicated but not all proteins in a family interact with all substrates. Transcriptional regulators common to many streptococcal species and known to regulate expression of the secreted proteins indicated are shown as rectangular boxes (Rgg [*S. pyogenes*], SloR [*S. mutans*], CovR [*S. agalactiae*]). These regulators are named differently according to

species (see Nobbs et al., 2009). SRRP, serine rich repeat protein; Cna, collagen adhesion; Cbp, collagen binding protein; Fnb, fibronectin binding protein; Fbp, fibronectin binding protein.

Fig. 3. Schematic of pilus filaments found across streptococci. Pilin-specific sortase (yellow) catalyzes the covalent linkage of one pilus protein subunit to another following cleavage of the LPXTG motif. For all streptococci the bulk of the pilus shaft comprises backbone (Bkb) subunits. Some filaments may then also incorporate one or two ancillary subunits (An1, An2), which are typically located at the tip or base of the structure respectively. In most cases, termination of pilus assembly occurs when housekeeping sortase A (pink) catalyzes linkage of the filament to the lipid II precursor of peptidoglycan and thus to the cell wall (depicted for heterotrimeric pilus only).

Figure 1

Figure 2

Figure 3

Monomeric

e.g. *S. suis*;
S. pneumoniae (PI-2)

Heterodimeric

e.g. GAS FCT-1,7,9;
S. gallolyticus; *S. oralis*;
S. mitis; *S. sanguinis*

Heterotrimeric

e.g. GAS FCT-2-6,8;
S. pneumoniae (PI-1); GBS
PI-1, PI-2a, PI-2b;
S. sanguinis