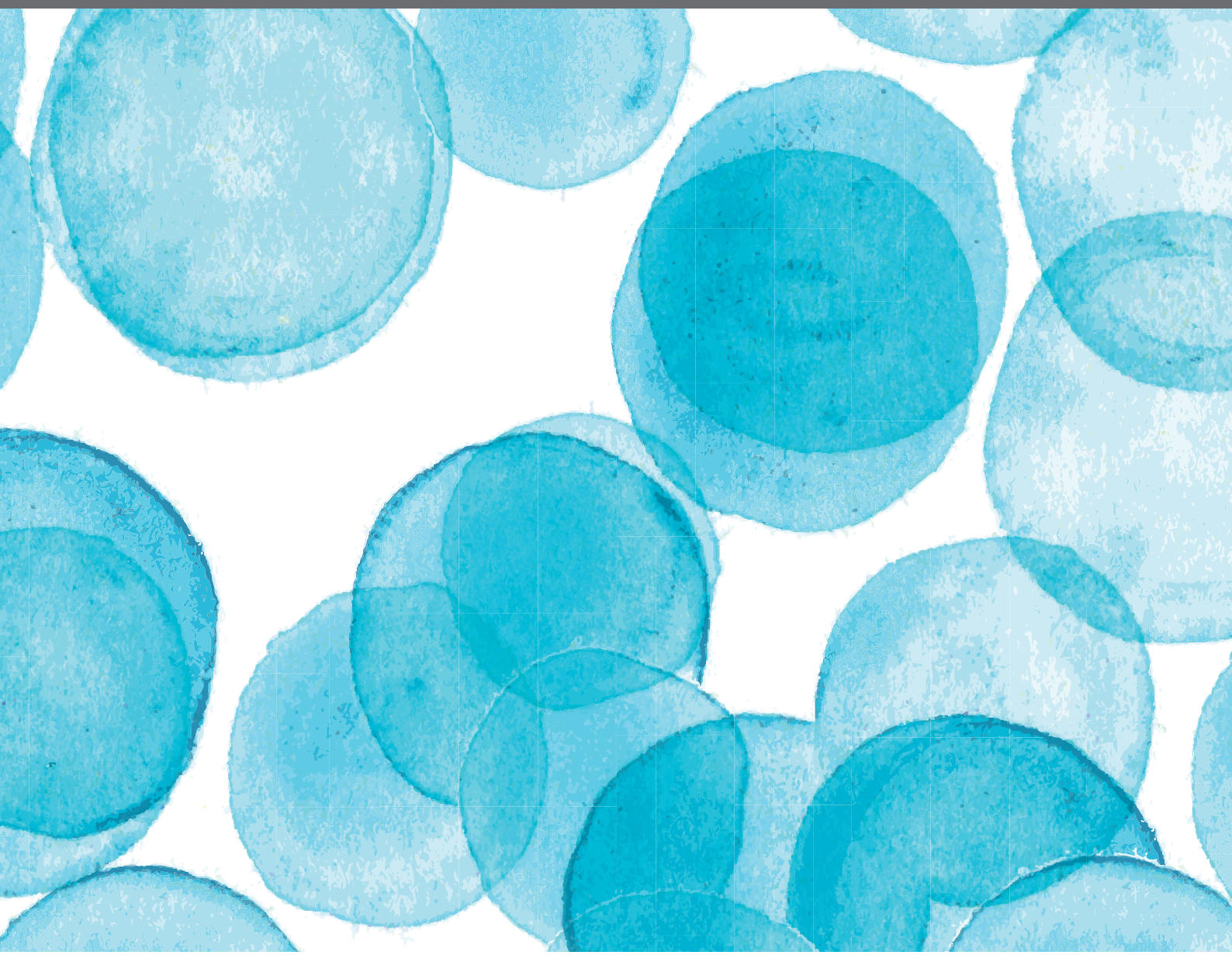
CELL SURFACE PROTEINS OF GRAM-POSITIVE PATHOGENIC BACTERIA

EDITED BY: Magnus Hook and Timothy J. Foster
PUBLISHED IN: Frontiers in Microbiology



frontiers Research Topics

ISSN 1664-8714
ISBN 978-2-88971-046-1
DOI 10.3389/978-2-88971-046-1

About Frontiers

Frontiers is more than just an open-access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers Journal Series

The Frontiers Journal Series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the Frontiers Journal Series operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to Quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews.
Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area! Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: frontiersin.org/about/contact

CELL SURFACE PROTEINS OF GRAM-POSITIVE PATHOGENIC BACTERIA

Topic Editors:
Magnus Hook, Texas A&M Health Science Center, United States
Timothy J. Foster, Trinity College Dublin, Ireland

Citation: Hook, M., Foster, T. J., eds. (2021). Cell Surface Proteins of Gram-positive Pathogenic Bacteria. Lausanne: Frontiers Media SA. doi: 10.3389/978-2-88971-046-1

Table of Contents

frontiers in Microbiology

EDITORIAL
published: 21 May 2021
doi: 10.3389/fmicb.2021.681880

Editorial: Cell Surface Proteins of Gram-Positive Pathogenic Bacteria

Magnus Hook[1] *and Timothy J. Foster*[2]*

[1] Center for Infectious and Inflammatory Diseases, Institute of Biosciences and Technology, Texas A&M Health Science Center, Houston, TX, United States, [2] Microbiology Department, Trinity College Dublin, Dublin, Ireland

Keywords: surface proteins, adhesion, biofilm, immune evasion, iron acquisition, pili, S-layer

Editorial on the Research Topic

Cell Surface Proteins of Gram-Positive Pathogenic Bacteria

The surfaces of Gram-positive bacterial cells are decorated with a diverse array of proteins including a large group that are covalently anchored to peptidoglycan by sortases. Several types of sortase-anchored proteins can be differentiated on the basis of their structures. Notable examples reviewed in this special topic include MSCRAMMs (Arora et al.; Foster; Speziale and Pietrocola), serine rich repeat proteins (Chan et al.), antigen I/II proteins (Manzer et al.), NEAT motif containing proteins (Ellis-Guardiola et al.) and streptococcal fimbriae/pili (Ness and Hilleringmann; Nakata and Kreikemeyer).

Many proteins that are located of the surface of Gram-positive pathogens and not anchored covalently. *Clostridium difficile* and *Bacillus anthracis* are exemplars of bacteria whose outer surfaces comprise crystalline arrays of S layer proteins that are attached non-covalently to cell wall polysaccharides (Ravi and Fioravanti). Lipoproteins are anchored to the outer surface of the cytoplasmic membrane via membrane lipids (Nguyen et al.). Other non-covalently anchored surface proteins including moonlighting cytoplasmic proteins and extracellular vesicles are not reviewed here.

Some articles review specific protein types drawn from a single species for example fibronectin binding proteins (FnBP) of *Staphylococcus aureus* (Speziale and Pietrocola), surface proteins of *Staphylococcus epidermidis* (Foster) and pili of *Streptococcus pneumoniae and Streptococcus pyogenes* (Nakata and Kreikemeyer; Ness and Hilleringmann). Others describe protein families spanning different species/genera including proteins related to the biofilm associated protein BAP of *S. aureus* (Valle et al.), polymer adhesion domain containing proteins (Järvå et al.) and collagen binding proteins (Arora et al.).

An emerging theme is that a single protein can often carry out multiple functions. Proteins that are exposed on the surface of bacterial cells are in direct contact with the host and are subjected to selective pressure to perform functions related to colonization of host tissues and evasion of host defenses. The repertoire of proteins is limited so many have evolved to adopt multiple roles. Amino acid sequence variation can help explain niche and host specialization and why some murine models of infection are limited (Pickering and Fitzgerald).

The defining structural feature of MSCRAMM family is two adjacent IgG-like folded domains that bind ligands by the dock lock latch (DLL) or collagen hug mechanisms (Arora et al.; Foster; Speziale and Pietrocola). An important recent discovery is that binding of MSCRAMMs to their ligands is strengthened by shear forces to the extent that the force required for separation is equivalent to that required to break a covalent bond (Dufrêne and Viljoen). This is reminiscent of catch bonds of pili of uropathogenic *Escherichia coli* binding the urinary tract epithelium.

OPEN ACCESS

Edited by:
Mattias Collin,
Lund University, Sweden

Reviewed by:
Dominique Missiakas,
University of Chicago, United States
Hung Ton-That,
University of California, Los Angeles, United States

***Correspondence:**
Timothy J. Foster
tfoster@tcd.ie

Specialty section:
This article was submitted to Infectious Diseases, a section of the journal Frontiers in Microbiology

Received: *17 March 2021*
Accepted: *23 April 2021*
Published: *21 May 2021*

Citation:
Hook M and Foster TJ (2021) Editorial: Cell Surface Proteins of Gram-Positive Pathogenic Bacteria. Front. Microbiol. 12:681880. doi: 10.3389/fmicb.2021.681880

The N-terminal A domains of FnBPs of *S. aureus* are archetypal MSCRAMMs that bind fibrinogen by DLL in a shear force-enhanced manner (Speziale and Pietrocola; Dufrêne and Viljoen). The A domain also binds plasminogen and histones by non-DLL mechanisms and engages in homophilic interactions that promote biofilm formation. These ancillary functions contribute to pathogenesis and immune evasion. The C-terminal repeat region forms an unfolded flexible stalk that engages fibronectin by a tandem β-zipper mechanism.

The ability to bind collagens is an important feature of many Gram-positive bacterial pathogens. There are two well-known structural classes of collagen binding protein namely the MSCRAMM proteins related to CNA of *S.* aureus that bind collagens by the shear force-enhanced collagen hug mechanism and a subset of M proteins of *S. pyogenes* including those that carry the peptide associated with rheumatic fever in their N-terminal hypervariable domain (Arora et al.). There are three other less well-defined groups including the V region PAD of streptococcal antigen I/II proteins (Manzer et al.).

A structural motif that is widespread among surface proteins has herein been named the polymer adhesin domain (PAD) (Järvå et al.). PADs have been classified into groups based on structural similarities and functions. PADs can bind a diverse array of carbohydrate ligands such as glucans and dextrans, lipoteichoic acids and DNA. They can promote adhesion, cell aggregation, and plasmid conjugation. The ligand binding domain V of the antigen I/ll proteins of oral streptococci are class II PADs that binds protein ligands such as collagen and fibrinogen, promotes adhesion to the surface of epithelial and endothelial cells, and stimulates cellular aggregation and biofilm formation (Manzer et al.).

Surface proteins are involved in both the primary attachment and accumulation phases of biofilm formation (Foster; Speziale and Pietrocola). It has recently been recognized that the cell-cell aggregation and the accumulation phases involve surface proteins forming amyloid fibers in the biofilm matrix. This occurs with the BAP protein of *S. aureus* (Valle et al.), the accumulation associated protein Aap of *S. epidermidis* (Foster) and the SraP protein of *Streptococcus mutans* (Chan et al.).

Some Gram-positive pathogens express filamentous surface appendages called fimbriae/pili that are assembled and anchored to the cell surface by sortases. As pili are extruded from the bacterial cell subunits are covalently linked to each other by sortase-catalyzed isopeptide bonds. Structural analysis of pilins revealed intramolecular isopeptide bonds. P1 pili of *S. pneumoniae* bind a number of host proteins including collagens, they promote biofilm formation, bind receptors on host cells and promote passage of bacteria across the blood-brain barrier in the pathogenesis of meningitis (Ness and Hilleringmann).

Antigenic differences in *S. pyogenes* pili are responsible for the T serotyping scheme. These diverse pili can be categorized into nine different forms based upon heterogeneity of the encoding loci and sequence differences (Nakata and Kreikemeyer). Pili contribute to tropism for the skin or the nasopharynx. This is supported by enhanced expression at temperatures below 37°C that occur at these superficial niches. A novel feature of one of the minor pilins is to harpoon host ligands by formation of covalent thioester bonds.

Several species of streptococci express large serine rich repeat proteins that are heavily glycosylated by cognate glycosyltransferases (Chan et al.). The variable binding region (BR) is sandwiched between serine rich regions at the N-terminus and the longer C-terminal region. The BR of the Srr1 protein of *Streptococcus agalactiae* comprises an MSCRAMM motif that binds fibrinogen and keratin by DLL and contributes to adhesion to endothelial cells and vaginal epithelial cells promoting endocarditis and vaginal colonization, respectively. Both Srr1/2 and PsrP of *S. pneumoniae* play important roles in the transition from the commensal carriage state to invasive infection.

Pathogenic bacteria need a source of iron when growing in the host in order to overcome nutritional immunity. *S. aureus and Staphylococcus lugdunensis* encode iron regulated surface determinant (Isd) systems characterized by surface anchored proteins with NEAT motifs (Ellis-Guardiola et al.). These bind hemoglobin (Hb) and extract bound heme by forcing conformational changes in Hb. It is postulated that Isd proteins form protein wires that relay heme from the cell surface though the peptidoglycan to the membrane transporter.

The special topic also considers some non-covalently anchored surface proteins. All Gram-positive pathogens express lipoproteins that perform a multiplicity of functions (Nguyen et al.). The protein parts assist in nutrient acquisition, have chaperone activity, promote invasion of host cells and conjugation. When released from the bacterium the lipid moiety activates TLR2 pathways in host cells triggering inflammatory responses.

Recent structural studies provide insights into the S-layer and S-layer associated proteins (Ravi and Fioravanti). This includes localizing conserved domains involved in S-layer assembly and folds that promote attachment to cell surface polysaccharides. S-layer proteins are excellent candidates for future translational studies in diagnostics, therapeutics and vaccines.

Surface proteins can confer specialization for different hosts as well as colonization of specific niches within the host (Pickering and Fitzgerald). Staphylococcal IsdB proteins have a higher affinity for human hemoglobin than for the mouse protein which compromizes the study of virulence in murine infection models. Host specificity is also conferred by MSCRAMMs binding to the variable α-chain of fibrinogen (Fg). The ClfB protein of *S. aureus* binds to a human specific sequence and has very low affinity for canine Fg. In contrast the SpsL protein of *Staphylococcus pseudintermedius* binds strongly to the canine Fg α-chain and weakly to the human version. By manipulating a key adhesin/invasin-ligand interaction it is possible to create infection models that more closely resemble the human disease. Internalin A (InlA) binding to E-cadherin is key to invasion of the intestinal epithelium by *Listeria monocytogenes*. Transgenic mice expressing human E cadherin in the intestinal epithelia could be infected orally whereas wild-type mice are resistant. Conversely by altering residues in InlA to strengthen binding to murine E-cadherin it was possible to "murinize" the pathogen.

In conclusion this special topic provides a comprehensive and up-to-date summary of many important aspects of proteins

that are located on the surfaces of Gram-positive bacterial pathogens and provides a useful starting point for anyone who wishes to obtain an overview of this important topic.

AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct and intellectual contribution to the work, and approved it for publication.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Microbiology

MINI REVIEW
published: 25 June 2020
doi: 10.3389/fmicb.2020.01457

Binding Strength of Gram-Positive Bacterial Adhesins

Yves F. Dufrêne and Albertus Viljoen**

Louvain Institute of Biomolecular Science and Technology, Catholic University of Louvain, Louvain-la-Neuve, Belgium

Bacterial pathogens are equipped with specialized surface-exposed proteins that bind strongly to ligands on host tissues and biomaterials. These adhesins play critical roles during infection, especially during the early step of adhesion where the cells are exposed to physical stress. Recent single-molecule experiments have shown that staphylococci interact with their ligands through a wide diversity of mechanosensitive molecular mechanisms. Adhesin–ligand interactions are activated by tensile force and can be ten times stronger than classical non-covalent biological bonds. Overall these studies demonstrate that Gram-positive adhesins feature unusual stress-dependent molecular interactions, which play essential roles during bacterial colonization and dissemination. With an increasing prevalence of multidrug resistant infections caused by *Staphylococcus aureus* and *Staphylococcus epidermidis*, chemotherapeutic targeting of adhesins offers an innovative alternative to antibiotics.

Keywords: Gram-positive bacteria, adhesins, physical stress, force, staphylococcus, atomic force microscopy

OPEN ACCESS

Edited by:
Magnus Hook,
Texas A&M Health Science Center,
United States

Reviewed by:
Jack Christopher Leo,
Nottingham Trent University,
United Kingdom
Mônica Larucci Vieira,
Federal University of Minas Gerais,
Brazil

****Correspondence:***
Yves F. Dufrêne
Yves.Dufrene@uclouvain.be
Albertus Viljoen
Albertus.Viljoen@uclouvain.be

Specialty section:
This article was submitted to Infectious Diseases, a section of the journal Frontiers in Microbiology

Received: *23 April 2020*
Accepted: *04 June 2020*
Published: *25 June 2020*

Citation:
Dufrêne YF and Viljoen A (2020) Binding Strength of Gram-Positive Bacterial Adhesins. Front. Microbiol. 11:1457. doi: 10.3389/fmicb.2020.01457

BACTERIAL ADHESION: FORCE IS THE KEY

Molecular interactions between pathogen adhesins and their ligands play a central role in controlling cell adhesion, the first step leading to infection. Traditional bioassays probe interactions of large ensembles of molecules typically under equilibrium conditions, and provide information on bulk receptor-ligand affinity (dissociation constants). Because most bacterial pathogens are exposed to physical stresses (Berne et al., 2018; Dufrêne and Persat, 2020), it is becoming clear that direct force measurement of adhesin binding strength at non-equilibrium is more relevant than equilibrium methods. During infection, bacteria experience external shear, tension, or compression, meaning understanding how the cells sense and respond to these physical cues has become an important challenge in current mechanobiology (Dufrêne and Persat, 2020). A crucial question is how force controls the adhesive functions of adhesins. The adhesion or unbinding strength is the force at which an adhesin detaches from its ligand with applied load. In most instances, this parameter decreases with applied force (slip bond). However, some adhesion proteins strengthen under tensile load, a behavior called catch bonding (Thomas et al., 2008). The prototypical example of catch bonds is the pilus protein FimH from *Escherichia coli*. This mannose-specific adhesin plays a major role in urinary tract infections by mediating shear-enhanced bacterial adhesion to host cells. Whether catch bonding is a widespread phenomenon among bacterial pathogens remains a

Abbreviations: Clf, clumping factor; Cn, collagen; Cna, Cn adhesin; DLL, dock, lock, and latch; FD, force–distance; Fg, fibrinogen; Fn, fibronectin; FnBP, Fn-binding protein; MSCRAMM, microbial surface components recognizing adhesive matrix molecules; SAG, salivary agglutinin; SCFS, single-cell force spectroscopy; Sdr, serine-aspartate repeat; SMFS, single-molecule force spectroscopy; SpA, surface protein A; vWF, von Willebrand factor.

controversial issue. Here below we show that there is now compelling evidence that Gram-positive bacteria have evolved force-dependent mechanisms to tune cell adhesion. Strong interaction forces provide the cells with a means to firmly adhere to protein-coated surfaces and to resist high shear stress conditions, while weak forces favor cell detachment and the colonization of new sites. Force is thus a critical parameter of the adhesin function and activity.

Atomic force microscopy (AFM) makes it possible to force probe single adhesins on living bacteria, enabling researchers to identify novel binding mechanisms and to understand how they regulate biofilm formation (Xiao and Dufrêne, 2016). SMFS relies on functionalizing an AFM probe to expose a ligand of interest, such as Fn or Fg (**Figure 1A**). Bringing the modified probe in contact with and subsequently retracting it from a living bacterium exposing specific adhesins allows to generate an FD curve (Hinterdorfer and Dufrêne, 2006). From the FD plot, the magnitude of the binding strength (or adhesion force) in piconewtons (pN) can be assessed, together with other biophysical parameters. By varying the pulling speed, dynamic force spectroscopy data are generated (Pfreundschuh et al., 2017). Modeling such data with appropriate theories provides quantitative information on thermodynamic and kinetic parameters of the single-molecule interaction. A variation of SMFS is SCFS where a living cell is attached to the probe, thus enabling the measurement of interaction forces between whole cells and protein-coated surfaces or other cells (**Figure 1B**).

STAPHYLOCOCCAL ADHESINS UNDER TENSION

Staphylococci cause a wide range of infections, including skin and soft tissue infections, bone and joint infections, but also bacteremia and endocarditis. In the latter infections, the bacteria are likely exposed to high fluid shear flow within the vasculature (Guo et al., 1995). The vascular shear flow may allow dissemination of staphylococcal microcolonies, which also have to adhere and resist shear flow (Sherman et al., 2019). The effect of shear on staphylococcal biofilms is thus an important facet of their pathophysiology. Indeed, *Staphylococcus epidermidis* and *Staphylococcus aureus* are notorious for forming biofilms on indwelling medical devices (Otto, 2009). *S. aureus* is equipped with a wide panel of adhesion mechanisms allowing it to evade host immunity (Foster, 2005). Today, highly recalcitrant methicillin-resistant *S. aureus* (MRSA) is regarded as one of the most successful modern pathogens (Turner et al., 2019). Staphylococcal species all express MSCRAMMs, which play crucial roles in adhesion and biofilm formation (Patti et al., 1994; Foster and Höök, 1998; Foster, 2019a,b). Within a few years, AFM has brought fascinating new insights into the molecular mechanisms of MSCRAMMs, showing that force and function are intimately connected in these adhesins.

Fibronectin-Binding Proteins

Undoubtedly, FnBPs have been the most widely investigated MSCRAMMs so far. Fn is a large extracellular matrix glycoprotein that contains three repeat-domain modules (FnI, FnII, and FnIII) and that is non-covalently anchored to cells through its binding to plasma membrane spanning integrins (Mouw et al., 2014). Gram-positive FnBPs bind *via* C-terminal domains containing tandem repeats to canonical (N-terminal FnI_{1-5} modules) and non-canonical ($FnI_6FnII_{1-2}FnI_{7-9}$ and $FnIII_{9,10,12}$ modules) sites in Fn (Hymes and Klaenhammer, 2016). Staphylococcal FnBPs, in particular, play multifunctional roles in adhesion by interacting with several ligands including themselves (Foster, 2016). Early investigations reported a weak ∼60-pN force measured for single Fn–FnBP bonds and a linear increase in unbinding force as a function of the loading rate, i.e., the speed at which force is applied (Bustanji et al., 2003). A positive correlation between contact time and adhesion force was observed (Xu et al., 2008), but conflicted on the specific involvement of FnBPs. Later, it was shown that double knock-out of the genes encoding FnBPA and FnBPB in *S. aureus* extinguished binding to Fn-coated AFM probes, while ectopic expression of these two proteins in *Lactococcus lactis*, conferred Fn-binding in this naturally non-Fn-binding bacterium (Buck et al., 2010). Another study revealed a distinct sawtooth-shaped force signature indicative of unfolding of multiple parallel FnI/FnII domains and a zipper array of Fn–FnBP bonds, supported by the absence of such a signature in isogenic mutants (Lower et al., 2010). Mapping the positions of the FnBP–Fn binding signatures showed that FnBPs on *S. aureus* cells were localized at the cell edges close to the Fn-support suggesting that adhesin clustering is induced in response to a primitive prokaryotic tactile surface sensing mechanism (Lower et al., 2010). In the clinical context, *S. aureus* small colony variants isolated from cystic fibrosis sufferers were demonstrated to sustain strong FnBP–Fn interactions via SigB-dependent high-level FnBP expression (Mitchell et al., 2008). In the same line, bloodstream *S. aureus* isolates from patients with cardiovascular implants formed mechanically strong bonds with Fn, involving cluster bonds of up to 80 proteins in parallel (Casillas-Ituarte et al., 2012). Moreover, isolates from patients with infected devices exhibited significantly longer bond lifetimes with Fn than those from patients with sterile devices, which was accounted for by amino acid polymorphisms in Fn-binding domains of FnBPA (Lower et al., 2011). Amino acid changes within high-affinity Fn-binding repeats in FnBPA in *S. aureus* isolated from patients with persistent bacteremia exhibited increased binding strength with Fn and appeared to impart conformational changes in Fn modulating affinity and unbinding (Xiong et al., 2015). Similarly, amino acid polymorphisms within the structured A domain of *S. aureus* FnBPA altered its binding to the abundant blood circulating glycoprotein, Fg (Casillas-Ituarte et al., 2019).

FnBPA is also engaged in homophilic cell–cell interactions, which originate from multiple low-affinity bonds (force of ∼125 pN) between A domains on neighboring cells (Herman-Bausier et al., 2015). The moderate strength of homophilic bonds may be important for biofilm dissemination, by contributing to cell detachment (isolated cells or cell clusters), therefore favoring colonization of new sites. Similar low-affinity homophilic bonds were also observed for an unrelated protein involved in cell–cell interactions and biofilm formation, the Sdr protein, SdrC

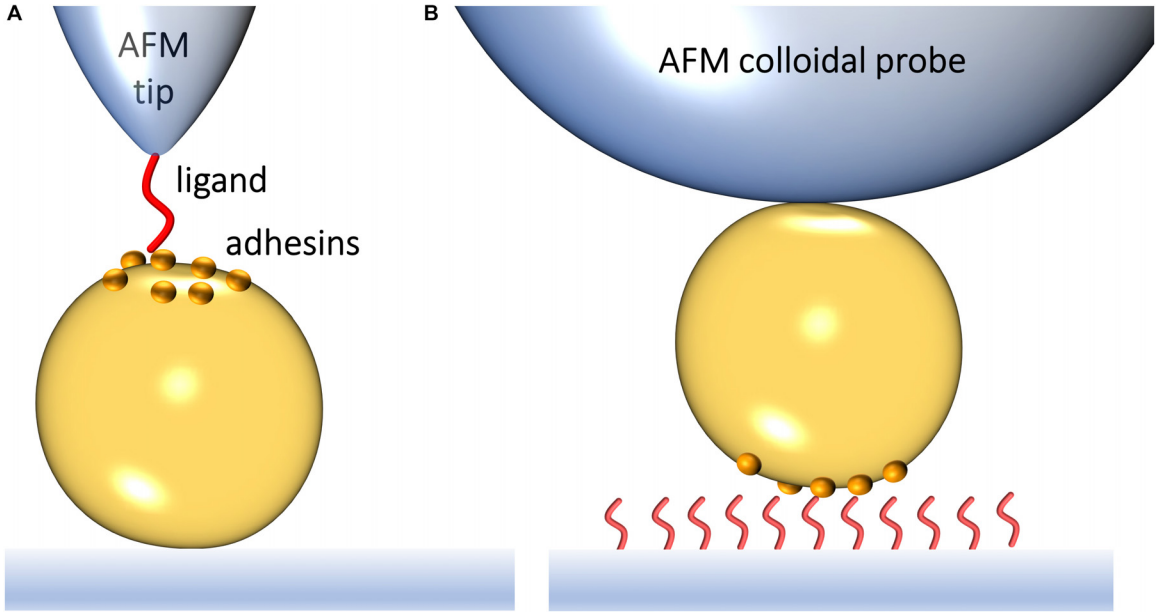


FIGURE 1 | Studying the interaction forces of Gram-positive bacterial adhesins by atomic force microscopy. **(A)** In single-molecule force spectroscopy (SMFS), the AFM probe is functionalized with a specific biomolecule, which can be a purified adhesin or its ligand. Bringing a ligand-functionalized probe into contact with a bacterial surface exposing an adhesin of interest allows capturing the adhesin binding strength upon probe-cell separation. **(B)** In single-cell force spectroscopy (SCFS), a colloidal AFM probe is used to first catch a single bacterium and then probe the strength of interaction between the cell and purified ligands on a substrate or another cell.

(Feuillie et al., 2017). On the other hand, Zn^{2+}-dependent homophilic interactions between pairs of the *S. aureus* surface protein, SasG, resisted much stronger forces (∼500 pN) (Gruszka et al., 2015; Formosa-Dague et al., 2016). The high mechanostability of SasG is likely to be of biological relevance. Under physical stress, protein unfolding may expose cryptic domains, which together with the rod-like shape of the protein will favor strong intercellular adhesion under flow. Looking at cellular invasion, a recent breakthrough is that the FnBPA–Fn complex binds α5β1 integrins with significantly greater strength than Fn alone, favoring invasion (Prystopiuk et al., 2018). The proposed explanation is that binding of FnBPA to Fn allosterically activates integrin binding to Fn, resulting in strong Fn–integrin interactions.

Serine-Aspartate Repeat Proteins and Clumping Factors

A remarkable recent discovery is the ultrastrong forces by which staphylococcal Sdr proteins bind to their ligands (Herman et al., 2013, 2014; Milles et al., 2018). The prototypical example is *S. epidermidis* SdrG, which binds via a "DLL" mechanism to Fg (Ponnuraj et al., 2003; Bowden et al., 2008; Foster et al., 2014). A sequence between the N2 and N3 domains within the N-terminal A region of the adhesin "docks" on to a 14-amino acid sequence within the Fg β-chain N-terminus (**Figure 2A**). Upon stable docking, an extension of the N3 domain folds over the bound Fg peptide, "locking" it in place, and then "latches" on to a β sheet in the N2 domain, stabilizing the SdrG–Fg complex considerably. Single-molecule AFM demonstrated that the SdrG–Fg interaction can sustain forces in the range of 2 nN, the strength of a covalent bond (**Figure 2A**) (Herman et al., 2014; Milles et al., 2018). Molecular dynamics simulations revealed the underlying molecular mechanism (**Figure 2B**). The target peptide, confined in a screw-like manner in the binding pocket of SdrG, distributes forces mainly toward the peptide backbone through an intricate hydrogen bond network (Milles and Gaub, 2020).

In addition to the remarkable strength of the SdrG interaction with Fg, SMFS studies also demonstrated that *S. aureus* clinical isolates exhibited an increased density of SdrG on their surfaces, which correlated with increased adhesion on Fg-coated substrates (Vanzieleghem et al., 2015). SdrG thus appears to be a formidable player in staphylococcal adhesion to Fg-coated medical implants. Examples also exist of Sdr proteins that interact with other extracellular matrix proteins, such as *S. epidermidis* SdrF that is responsible for its binding to Cn, which involved weak as well as strong bonds (Herman-Bausier and Dufrêne, 2016).

Another set of staphylococcal MSCRAMM adhesins that are related to the serine aspartate repeat proteins are the Clfs. It was found that the force of ClfA binding to Fg increases sharply from ∼100 pN under low tension to forces exceeding 1.5 nN under high tension (**Figure 2C**), which was specifically dependent on the C-terminal portion of the Fg γ chain (Herman-Bausier et al., 2018). This shows that ClfA-Fg is activated by mechanical force, reminiscent of a catch bond behavior (Sokurenko et al., 2008). The very strong forces enable the pathogens to resist high shear stress conditions, which often occur during colonization.

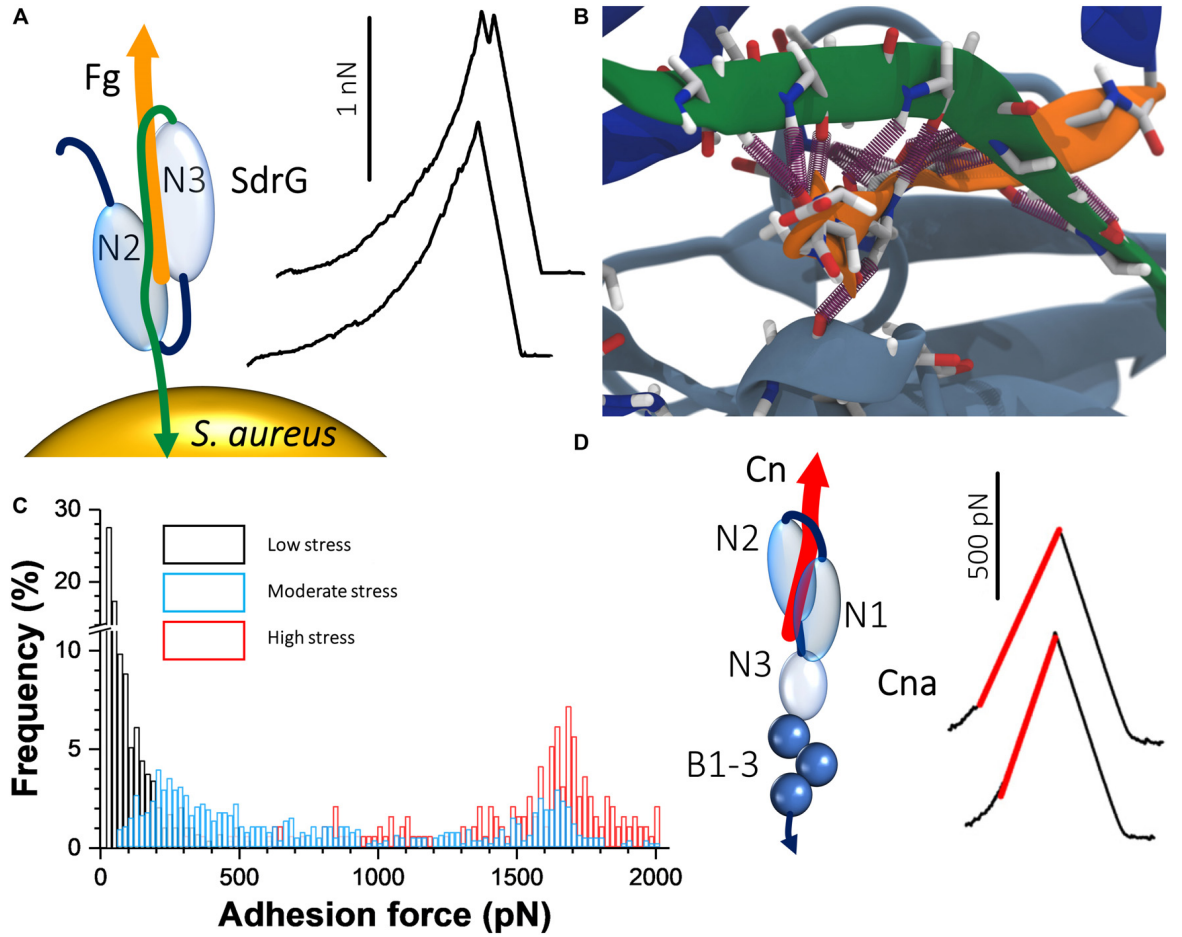


FIGURE 2 | Staphylococcal adhesins can bind their ligands with extremely strong forces. **(A)** The "dock, lock, and latch" mechanism involves docking of Fg on to a sequence between the N2 and N3 domains of SdrG. A C-terminal "locking strand" extension (green) of the N3 domain then folds over the docked Fg, locking it in place and binds to a β sheet within the N2 domain, stabilizing the interaction. On the right are typical force curves for the SdrG–Fg interaction showing that the unbinding force is around 2000 pN. Adapted with permission from Herman et al. (2014). **(B)** The unique mechanical stability of the SdrG–Fg bond is attributed to extensive hydrogen bonding (purple cylinders) formed between the SdrG N2/N3 domains (light and dark blue, respectively) and Fg (orange) and the specific geometry of this molecular interaction under load. A screw-like arrangement of the hydrogen bonds sustains a perfect shear geometry between SdrG and Fg under load. Adapted with permission from Milles et al. (2018). **(C)** Force activation of strong binding between ClfA and Fg. Shown are overlays of histogram plots of adhesion force when the ClfA–Fg bond is subjected to different levels of mechanical stresses by varying the loading rate. Adapted with permission from Herman-Bausier et al. (2018). **(D)** Interaction between collagen (Cn) and single Cna adhesins exposed on living bacteria. The "Cn hug" mechanism involves docking of Cn in a furrow within the Cna N2 domain (left); subsequently, a linker between the N2 and N1 domains folds over the docked Cn and a C-terminal latch sequence of the N2 domain binds within the N1 domain, consequently locking Cn in place. Isopeptide bonds within the C-terminal B repeats of Cna prohibit unfolding under mechanical stress. AFM showed that the Cn–Cna bond behaves like a nanospring when pulled apart, as evidenced by FD curves (right). Adapted with permission from Herman-Bausier et al. (2016).

Analogous to the FnBPA–Fn–α5β1 integrin ternary complex discussed in the previous section, the participation of ClfA in a similar complex involving Fg and the endothelium integrin $\alpha_V\beta3$ was investigated and shown to resist strong forces in the 0.8 nN range (Viela et al., 2019). When stress on the ClfA–Fg–$\alpha_V\beta3$ complex is minimal, the interaction between Fg and the integrin depends on C-terminal RGD sequences of dimeric Fg α chains. However, under mechanical tension, cryptic N-terminal α chain RGD sequences are exposed that increases the stability of the interaction.

Staphylococcus aureus ClfB binds the skin cornified cell envelope proteins cytokeratin and loricrin (Ganesh et al., 2011; Xiang et al., 2012). The force interaction between ClfB and loricrin was investigated and revealed to be mechanically activated, with high forces (1.5 nN range) being sustained under tensile stress (Vitry et al., 2017). Focusing on *S. aureus* adhesion to corneocytes, SCFS studies revealed a dependency on ClfB (Feuillie et al., 2018). Importantly, it was also found that reductions in the levels of natural moisturizing factor in the corneocytes lead to increased adhesion by the bacteria. Interestingly, ClfB binds exclusively to human Fg (Walsh et al., 2008). In the same line, the cell surface protein SpsL from *Staphylococcus pseudintermedius*, which causes disease in dogs, binds with strong forces (up to 2 nN) to canine but not human Fg

(Pickering et al., 2019). Overall, these experiments demonstrate the crucial role of protein mechanics in tuning the adhesive functions of bacterial pathogens. Staphylococci have evolved fascinating force-dependent ligand-binding mechanisms that help the cells to attach firmly to biomaterials under high shear stress, and to detach under low shear stress to colonize new sites.

Collagen Adhesin

The Cn-binding protein Cna plays important roles in bacterium-host adherence and in immune evasion. The strength of Cna–Cn bonds was shown to be very strong (∼1.2 nN; **Figure 2D**) (Herman-Bausier et al., 2016), consistent with the high-affinity "Cn hug" mechanism, a variation of the high-affinity DLL mechanism. The B region of the adhesin was required for strong ligand binding and functioned as a spring capable of sustaining high forces, potentially due to isopeptide bonds that prohibit unfolding. This mechanical response provides a means to project the A region away from the bacterial surface and to maintain bacterial adhesion under conditions of high forces (Deivanayagam et al., 2000). The force interactions between Cna and C1q, a complement component, and another extracellular matrix protein, laminin, were also investigated (Valotteau et al., 2017). The forces in these interactions were considerably smaller than in the Cna–Cn interactions indicating a different binding mechanism to that of the Cn hug. Interestingly, it was observed that at the single-cell level Cna binding to C1q involved at most two bonds, while in the case of laminin, it involved multiple bonds indicating that multivalency or cooperativity could enhance the strength of Cna-mediated adhesion. These results show that Cna is a multifunctional protein capable of binding to different host ligands through mechanisms that differ from the classical Cn hug.

Surface Protein A

Staphylococcus aureus was also found to adhere to endothelial cells under hematogenous shear flow via interactions with the large multimeric glycoprotein vWF (Claes et al., 2014). The forces in the interaction between *S. aureus* (and MRSA) SpA and vWF were recently unraveled (Viela et al., 2019). Like for several staphylococcal adhesins, the SpA–vWF interaction is force activated and very strong, resisting forces in the 2 nN range. Activation of the SpA–vWF interaction under tension may promote adhesion of bacteria to damaged blood vessels.

OTHER GRAM-POSITIVE ADHESINS

Besides the above studies, there have also been reports on non-staphylococcal Gram-positive adhesins. The SpaA pilus subunit from the nasopharyngeal pathogen *Corynebacterium diphtheriae*, was subject of an SMFS study (Echelman et al., 2016). The results revealed how mechanical energy is efficiently dissipated *via* unfolding and refolding of isopeptide bond-delimited polypeptide loops within the CnaA domains of SpaA. The authors posited that CnaA domains may allow pili to withstand severe forces induced by coughing by dissipating energy away as heat, thus supporting *C. diphtheriae* infection. Focusing on *Streptococcus pyogenes* that causes numerous infections among which pharyngitis and tonsillitis, the forces of the pilus-tip adhesin Spy0125 was investigated (Echelman et al., 2017). It was found that in its folded state, a thioester bond within Spy0125 could be cleaved by nucleophiles, but when the adhesin was put under tension resulting in it being mechanically unfolded, cleavage of the thioester bond could no longer be achieved. In the absence of mechanical stress, cleavage and reformation of the thioester bond was in equilibrium. These results thus indicated that the reversible cleavage of the thioester bond may allow the adhesin to circumvent the activity of inflammation-associated molecules that may attack it, allowing interaction with its host ligands, which under mechanical shear stresses is stabilized.

SpaC, the key pilus adhesion protein of the probiotic Gram-positive bacterium *Lactobacillus rhamnosus* GG (LGG) was shown to feature broad specificity. SpaC formed homophilic trans-interactions engaged in bacterial aggregation and specifically bound mucin and Cn (Tripathi et al., 2013). LGG pili exhibit two unique mechanical responses, that is, zipper-like adhesion involving multiple SpaC molecules distributed along the pilus length and nanospring properties enabling pili to resist high forces. These mechanical properties may be a general trait of Gram-positive pili, enabling the cells to adhere under shear stress conditions.

In the context of tooth decay, where pathogens are exposed to salivary sheer flow (Prakobphol et al., 1995), *Streptococcus mutans* is a major cause of dental caries. *S. mutans* adheres to dental-immobilized SAG. The interaction between the *S. mutans* adhesin P1 and SAG was studied by SMFS and relatively weak forces were observed (∼50 pN) (Sullan et al., 2015). However, SCFS revealed much greater forces in this interaction (∼500 pN) that may indicate binding of multiple P1 molecules to SAG glycoproteins, strengthening *S. mutans* adhesion.

TARGETING ADHESINS FOR THERAPY

In a context of increasing drug resistance among Gram-positive pathogens, antiadhesion therapies are attractive because they may supplement waning arsenals of available antibiotics and because they do not target essential processes, they have the added potential advantage of eliciting less and slower resistance acquisition (Krachler and Orth, 2013a,b; Geoghegan et al., 2017; Arciola et al., 2018). Historically, the most widely investigated system is the blocking of the attachment of uropathogenic *E. coli* bacteria to host epithelial cells (Flores-Mireles et al., 2015). Mannosides have been shown to be efficient in inhibiting the adhesion of FimH to host cells (Cusumano et al., 2011). Important lessons were also learned from AFM. Cranberry juice inhibited the fimbriae-mediated adhesion of *E. coli* to solid surfaces and host cells (Pinzón-Arango et al., 2009; Liu et al., 2010; Tao et al., 2011). Similarly, glycofullerenes blocked the force interactions between *E. coli* fimbriae and their carbohydrate receptors (Beaussart et al., 2016).

In the interaction between the *S. aureus* Cna adhesin and Cn, two monoclonal antibodies with competitive

inhibitory activity were identified by screening a collection of monoclonal anti-Cna antibodies (Herman-Bausier et al., 2016). Another monoclonal antibody from the same collection blocked the interaction between Cna and the complement protein C1q as well as the extracellular matrix protein laminin (Valotteau et al., 2017). Also, AFM studies revealed the competitive inhibition of *S. aureus* SdrC homophilic interactions by a peptide derived from β-neurexin (Feuillie et al., 2017) and the inhibition by heparin of *S. epidermidis* clinical isolates' adhesion to Fn (Bustanji et al., 2003). Notably, the resolution of the crystal structure of ClfA in complex with the monoclonal antibody tefibazumab offers the potential for the rational design of antiadhesive antibodies targeting staphylococci (Ganesh et al., 2016), in which AFM could serve as a platform to study structure-activity relationships. Accordingly, AFM may be used to screen novel compounds with antiadhesive properties as well as to decipher their mechanisms of action.

Lastly, AFM is also a valuable tool to unravel how antimicrobials alter the specific and non-specific adhesion forces of bacteria. AFM helped unravel the interplay between staphylococcal adhesion to solid surfaces and cell wall deformation under treatment with cell wall active and non-active antibiotics (Carniello et al., 2018). In mycobacteria, the efficient first line antitubercular, ethambutol, had a strong effect on *Mycobacterium bovis* BCG–Fn interactions (Verbelen and Dufrêne, 2009), while several compounds targeting mycolic acid biosynthesis strongly decreased the magnitudes and frequencies of hydrophobic adhesive forces measured on *M. bovis* BCG or *Mycobacterium abscessus* cells (Alsteens et al., 2008; Viljoen et al., 2020). AFM also disclosed the inhibitory effect of the herbicide 2,4-dichlorophenoxyacetic acid on *E. coli* non-specific adhesion (Bhat et al., 2015, 2018a,b). Interestingly, SMFS studies detected the massive surface exposure of *Candida albicans* Als adhesins after treatment with the antifungal caspofungin (El-Kirat-Chatel et al., 2013). Homophilic interactions between Als adhesins play an important role in cellular aggregation (Dehullu et al., 2019), highlighting the importance of studying the effects of antimicrobials on adhesion.

OUTLOOK

The discoveries discussed here represent an important step forward in our understanding of the molecular mechanisms used by Gram-positive pathogens to mediate cell adhesion and trigger infections. AFM experiments have shown that Gram-positive bacterial adhesins feature a wide range of binding strengths, from ∼50 to more than ∼2000 pN. A remarkable finding of the past years is that some adhesins bind their ligands with extremely strong forces that are activated by mechanical tension, as in catch bonds, an intriguing phenomenon that enables the pathogens to firmly bind to host cells and protein-coated surfaces, and to sustain high shear stress. For years, the only well-investigated catch bond behavior was the mannose-binding pilus-tip adhesin FimH from the Gram-negative *E coli*. We are now starting to realize that Gram-positive staphylococcal adhesins might become a new paradigm in catch bond adhesion. These binding mechanisms represent potential targets to fight infections, and AFM might become a valuable tool for the screening of antiadhesion compounds like small peptides and antibodies.

AUTHOR CONTRIBUTIONS

YD and AV contributed equally to the writing of this manuscript.

FUNDING

Work at UCLouvain was supported by the Excellence of Science-EOS program (Grant #30550343), the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (Grant Agreement No. 693630), the FNRS-WELBIO (Grant No. WELBIO-CR-2015A-05), the National Fund for Scientific Research (FNRS), and the Research Department of the Communauté française de Belgique (Concerted Research Action).

REFERENCES

Alsteens, D., Verbelen, C., Dague, E., Raze, D., Baulard, A. R., and Dufrêne, Y. F. (2008). Organization of the mycobacterial cell wall: a nanoscale view. *Pflugers Arch.* 456, 117–125. doi: 10.1007/s00424-007-0386-0

Arciola, C. R., Campoccia, D., and Montanaro, L. (2018). Implant infections: adhesion, biofilm formation and immune evasion. *Nat. Rev. Microbiol.* 16, 397–409. doi: 10.1038/s41579-018-0019-y

Beaussart, A., Abellán-Flos, M., El-Kirat-Chatel, S., Vincent, S. P., and Dufrêne, Y. F. (2016). Force nanoscopy as a versatile platform for quantifying the activity of antiadhesion compounds targeting bacterial pathogens. *Nano Lett.* 16, 1299–1307. doi: 10.1021/acs.nanolett.5b04689

Berne, C., Ellison, C. K., Ducret, A., and Brun, Y. V. (2018). Bacterial adhesion at the single-cell level. *Nat. Rev. Microbiol.* 16, 616–627. doi: 10.1038/s41579-018-0057-5

Bhat, S. V., Booth, S. C., Vantomme, E. A. N., Afroj, S., Yost, C. K., and Dahms, T. E. S. (2015). Oxidative stress and metabolic perturbations in *Escherichia coli* exposed to sublethal levels of 2,4-dichlorophenoxyacetic acid. *Chemosphere* 135, 453–461. doi: 10.1016/j.chemosphere.2014.12.035

Bhat, S. V., Kamencic, B., Körnig, A., Shahina, Z., and Dahms, T. E. S. (2018a). Exposure to sub-lethal 2,4-Dichlorophenoxyacetic acid arrests cell division and alters cell surface properties in *Escherichia coli*. *Front. Microbiol.* 9:44. doi: 10.3389/fmicb.2018.00044

Bhat, S. V., Sultana, T., Körnig, A., McGrath, S., Shahina, Z., and Dahms, T. E. S. (2018b). Correlative atomic force microscopy quantitative imaging-laser scanning confocal microscopy quantifies the impact of stressors on live cells in real-time. *Sci. Rep.* 8:8305. doi: 10.1038/s41598-018-26433-1

Bowden, M. G., Heuck, A. P., Ponnuraj, K., Kolosova, E., Choe, D., Gurusiddappa, S., et al. (2008). Evidence for the "Dock, Lock, and Latch" ligand binding mechanism of the staphylococcal microbial surface component recognizing adhesive matrix molecules (MSCRAMM) SdrG. *J. Biol. Chem.* 283, 638–647. doi: 10.1074/jbc.M706252200

Buck, A. W., Fowler, V. G., Yongsunthon, R., Liu, J., DiBartola, A. C., Que, Y.-A., et al. (2010). Bonds between fibronectin and fibronectin-binding proteins on *Staphylococcus aureus* and *Lactococcus lactis*. *Langmuir* 26, 10764–10770. doi: 10.1021/la100549u

Bustanji, Y., Arciola, C. R., Conti, M., Mandello, E., Montanaro, L., and Samorí, B. (2003). Dynamics of the interaction between a fibronectin molecule and

a living bacterium under mechanical force. *Proc. Natl. Acad. Sci. U.S.A.* 100, 13292–13297. doi: 10.1073/pnas.1735343100

Carniello, V., Peterson, B. W., Sjollema, J., Busscher, H. J., and van der Mei, H. C. (2018). Surface enhanced fluorescence and nanoscopic cell wall deformation in adhering *Staphylococcus aureus* upon exposure to cell wall active and non-active antibiotics. *Nanoscale* 10, 11123–11133. doi: 10.1039/c8nr01669k

Casillas-Ituarte, N. N., DiBartola, A. C., Broughton, M. J., Pérez-Guzmán, L., Wheeler, R. M., Ibaraki, M., et al. (2019). Fibrinogen binding is affected by amino acid substitutions in C-terminal repeat region of fibronectin binding protein A. *Sci. Rep.* 9:11619. doi: 10.1038/s41598-019-48031-5

Casillas-Ituarte, N. N., Lower, B. H., Lamlertthon, S., Fowler, V. G., and Lower, S. K. (2012). Dissociation rate constants of human fibronectin binding to fibronectin-binding proteins on living *Staphylococcus aureus* isolated from clinical patients. *J. Biol. Chem.* 287, 6693–6701. doi: 10.1074/jbc.M111.285692

Claes, J., Vanassche, T., Peetermans, M., Liesenborghs, L., Vandenbriele, C., Vanhoorelbeke, K., et al. (2014). Adhesion of *Staphylococcus aureus* to the vessel wall under flow is mediated by von Willebrand factor–binding protein. *Blood* 124, 1669–1676. doi: 10.1182/blood-2014-02-558890

Cusumano, C. K., Pinkner, J. S., Han, Z., Greene, S. E., Ford, B. A., Crowley, J. R., et al. (2011). Treatment and prevention of urinary tract infection with orally active FimH inhibitors. *Sci. Transl. Med.* 3:109ra115. doi: 10.1126/scitranslmed.3003021

Dehullu, J., Valotteau, C., Herman-Bausier, P., Garcia-Sherman, M., Mittelviefhaus, M., Vorholt, J. A., et al. (2019). Fluidic force microscopy demonstrates that homophilic adhesion by *Candida albicans* Als proteins is mediated by amyloid bonds between cells. *Nano Lett.* 19, 3846–3853. doi: 10.1021/acs.nanolett.9b01010

Deivanayagam, C. C., Rich, R. L., Carson, M., Owens, R. T., Danthuluri, S., Bice, T., et al. (2000). Novel fold and assembly of the repetitive B region of the *Staphylococcus aureus* collagen-binding surface protein. *Struct. Lond. Engl.* 1993, 67–78. doi: 10.1016/s0969-2126(00)00081-2

Dufrêne, Y. F., and Persat, A. (2020). Mechanomicrobiology: how bacteria sense and respond to forces. *Nat. Rev. Microbiol.* 18, 227–240. doi: 10.1038/s41579-019-0314-2

Echelman, D. J., Alegre-Cebollada, J., Badilla, C. L., Chang, C., Ton-That, H., and Fernández, J. M. (2016). CnaA domains in bacterial pili are efficient dissipaters of large mechanical shocks. *Proc. Natl. Acad. Sci. U.S.A.* 113, 2490–2495. doi: 10.1073/pnas.1522946113

Echelman, D. J., Lee, A. Q., and Fernández, J. M. (2017). Mechanical forces regulate the reactivity of a thioester bond in a bacterial adhesin. *J. Biol. Chem.* 292, 8988–8997. doi: 10.1074/jbc.M117.777466

El-Kirat-Chatel, S., Beaussart, A., Alsteens, D., Jackson, D. N., Lipke, P. N., and Dufrêne, Y. F. (2013). Nanoscale analysis of caspofungin-induced cell surface remodelling in *Candida albicans*. *Nanoscale* 5, 1105–1115. doi: 10.1039/c2nr33215a

Feuillie, C., Formosa-Dague, C., Hays, L. M. C., Vervaeck, O., Derclaye, S., Brennan, M. P., et al. (2017). Molecular interactions and inhibition of the staphylococcal biofilm-forming protein SdrC. *Proc. Natl. Acad. Sci. U.S.A.* 114, 3738–3743. doi: 10.1073/pnas.1616805114

Feuillie, C., Vitry, P., McAleer, M. A., Kezic, S., Irvine, A. D., Geoghegan, J. A., et al. (2018). Adhesion of *Staphylococcus aureus* to corneocytes from atopic dermatitis patients is controlled by natural moisturizing factor levels. *mBio* 9:e01184-18. doi: 10.1128/mBio.01184-18

Flores-Mireles, A. L., Walker, J. N., Caparon, M., and Hultgren, S. J. (2015). Urinary tract infections: epidemiology, mechanisms of infection and treatment options. *Nat. Rev. Microbiol.* 13, 269–284. doi: 10.1038/nrmicro3432

Formosa-Dague, C., Speziale, P., Foster, T. J., Geoghegan, J. A., and Dufrêne, Y. F. (2016). Zinc-dependent mechanical properties of *Staphylococcus aureus* biofilm-forming surface protein SasG. *Proc. Natl. Acad. Sci. U.S.A.* 113, 410–415. doi: 10.1073/pnas.1519265113

Foster, T. J. (2005). Immune evasion by staphylococci. *Nat. Rev. Microbiol.* 3, 948–958. doi: 10.1038/nrmicro1289

Foster, T. J. (2016). The remarkably multifunctional fibronectin binding proteins of *Staphylococcus aureus*. *Eur. J. Clin. Microbiol. Infect. Dis.* 35, 1923–1931. doi: 10.1007/s10096-016-2763-0

Foster, T. J. (2019a). Surface proteins of *Staphylococcus aureus*. *Microbiol. Spectr.* 7:GPP3-0046-2018. doi: 10.1128/microbiolspec.GPP3-0046-2018

Foster, T. J. (2019b). The MSCRAMM family of cell-wall-anchored surface proteins of gram-positive cocci. *Trends Microbiol.* 27, 927–941. doi: 10.1016/j.tim.2019.06.007

Foster, T. J., Geoghegan, J. A., Ganesh, V. K., and Höök, M. (2014). Adhesion, invasion and evasion: the many functions of the surface proteins of *Staphylococcus aureus*. *Nat. Rev. Microbiol.* 12, 49–62. doi: 10.1038/nrmicro3161

Foster, T. J., and Höök, M. (1998). Surface protein adhesins of *Staphylococcus aureus*. *Trends Microbiol.* 6, 484–488. doi: 10.1016/s0966-842x(98)01400-0

Ganesh, V. K., Barbu, E. M., Deivanayagam, C. C. S., Le, B., Anderson, A. S., Matsuka, Y. V., et al. (2011). Structural and biochemical characterization of *Staphylococcus aureus* clumping factor B/ligand interactions. *J. Biol. Chem.* 286, 25963–25972. doi: 10.1074/jbc.M110.217414

Ganesh, V. K., Liang, X., Geoghegan, J. A., Cohen, A. L. V., Venugopalan, N., Foster, T. J., et al. (2016). Lessons from the crystal structure of the *S. aureus* surface protein clumping Factor A in complex with Tefibazumab, an inhibiting monoclonal antibody. *EBioMedicine* 13, 328–338. doi: 10.1016/j.ebiom.2016.09.027

Geoghegan, J. A., Foster, T. J., Speziale, P., and Dufrêne, Y. F. (2017). Live-cell nanoscopy in antiadhesion therapy. *Trends Microbiol.* 25, 512–514. doi: 10.1016/j.tim.2017.04.002

Gruszka, D. T., Whelan, F., Farrance, O. E., Fung, H. K. H., Paci, E., Jeffries, C. M., et al. (2015). Cooperative folding of intrinsically disordered domains drives assembly of a strong elongated protein. *Nat. Commun.* 6:7271. doi: 10.1038/ncomms8271

Guo, Z., Moreau, M., Rickey, D. W., Picot, P. A., and Fenster, A. (1995). Quantitative investigation of in vitro flow using three-dimensional colour Doppler ultrasound. *Ultrasound Med. Biol.* 21, 807–816. doi: 10.1016/0301-5629(95)00007-E

Herman, P., El-Kirat-Chatel, S., Beaussart, A., Geoghegan, J. A., Foster, T. J., and Dufrêne, Y. F. (2014). The binding force of the staphylococcal adhesin SdrG is remarkably strong. *Mol. Microbiol.* 93, 356–368. doi: 10.1111/mmi.12663

Herman, P., El-Kirat-Chatel, S., Beaussart, A., Geoghegan, J. A., Vanzieleghem, T., Foster, T. J., et al. (2013). Forces driving the attachment of *Staphylococcus epidermidis* to fibrinogen-coated surfaces. *Langmuir* 29, 13018–13022. doi: 10.1021/la4029172

Herman-Bausier, P., and Dufrêne, Y. F. (2016). Atomic force microscopy reveals a dual collagen-binding activity for the staphylococcal surface protein SdrF. *Mol. Microbiol.* 99, 611–621. doi: 10.1111/mmi.13254

Herman-Bausier, P., El-Kirat-Chatel, S., Foster, T. J., Geoghegan, J. A., and Dufrêne, Y. F. (2015). *Staphylococcus aureus* fibronectin-binding Protein A mediates cell-cell adhesion through low-affinity homophilic bonds. *mBio* 6:e00413-15. doi: 10.1128/mBio.00413-15

Herman-Bausier, P., Labate, C., Towell, A. M., Derclaye, S., Geoghegan, J. A., and Dufrêne, Y. F. (2018). *Staphylococcus aureus* clumping factor A is a force-sensitive molecular switch that activates bacterial adhesion. *Proc. Natl. Acad. Sci. U.S.A.* 115, 5564–5569. doi: 10.1073/pnas.1718104115

Herman-Bausier, P., Valotteau, C., Pietrocola, G., Rindi, S., Alsteens, D., Foster, T. J., et al. (2016). Mechanical strength and inhibition of the *Staphylococcus aureus* collagen-binding protein Cna. *mBio* 7:e01529-16. doi: 10.1128/mBio.01529-16

Hinterdorfer, P., and Dufrêne, Y. F. (2006). Detection and localization of single molecular recognition events using atomic force microscopy. *Nat. Methods* 3, 347–355. doi: 10.1038/nmeth871

Hymes, J. P., and Klaenhammer, T. R. (2016). Stuck in the middle: fibronectin-Binding proteins in gram-positive bacteria. *Front. Microbiol.* 7:1504. doi: 10.3389/fmicb.2016.01504

Krachler, A. M., and Orth, K. (2013b). Targeting the bacteria–host interface. *Virulence* 4, 284–294. doi: 10.4161/viru.24606

Krachler, A.-M., and Orth, K. (2013a). Made to stick: anti-adhesion therapy for bacterial infections: a major advantage in targeting adhesion is that the body clears invading pathogens instead of killing them. *Microbe Mag.* 8, 286–290. doi: 10.1128/microbe.8.286.1

Liu, Y., Pinzón-Arango, P. A., Gallardo-Moreno, A. M., and Camesano, T. A. (2010). Direct adhesion force measurements between *E. coli* and human uroepithelial cells in cranberry juice cocktail. *Mol. Nutr. Food Res.* 54, 1744–1752. doi: 10.1002/mnfr.200900535

Lower, S. K., Lamlertthon, S., Casillas-Ituarte, N. N., Lins, R. D., Yongsunthon, R., Taylor, E. S., et al. (2011). Polymorphisms in fibronectin binding protein A of *Staphylococcus aureus* are associated with infection of cardiovascular devices. *Proc. Natl. Acad. Sci. U.S.A.* 108, 18372–18377. doi: 10.1073/pnas.1109071108

Lower, S. K., Yongsunthon, R., Casillas-Ituarte, N. N., Taylor, E. S., DiBartola, A. C., Lower, B. H., et al. (2010). A tactile response in *Staphylococcus aureus*. *Biophys. J.* 99, 2803–2811. doi: 10.1016/j.bpj.2010.08.063

Milles, L. F., and Gaub, H. E. (2020). Extreme mechanical stability in protein complexes. *Curr. Opin. Struct. Biol.* 60, 124–130. doi: 10.1016/j.sbi.2019.11.012

Milles, L. F., Schulten, K., Gaub, H. E., and Bernardi, R. C. (2018). Molecular mechanism of extreme mechanostability in a pathogen adhesin. *Science* 359, 1527–1533. doi: 10.1126/science.aar2094

Mitchell, G., Lamontagne, C.-A., Brouillette, E., Grondin, G., Talbot, B. G., Grandbois, M., et al. (2008). *Staphylococcus aureus* SigB activity promotes a strong fibronectin–bacterium interaction which may sustain host tissue colonization by small-colony variants isolated from cystic fibrosis patients. *Mol. Microbiol.* 70, 1540–1555. doi: 10.1111/j.1365-2958.2008.06511.x

Mouw, J. K., Ou, G., and Weaver, V. M. (2014). Extracellular matrix assembly: a multiscale deconstruction. *Nat. Rev. Mol. Cell Biol.* 15, 771–785. doi: 10.1038/nrm3902

Otto, M. (2009). *Staphylococcus epidermidis*–the "accidental" pathogen. *Nat. Rev. Microbiol.* 7, 555–567. doi: 10.1038/nrmicro2182

Patti, J. M., Allen, B. L., McGavin, M. J., and Höök, M. (1994). MSCRAMM-mediated adherence of microorganisms to host tissues. *Annu. Rev. Microbiol.* 48, 585–617. doi: 10.1146/annurev.mi.48.100194.003101

Pfreundschuh, M., Harder, D., Ucurum, Z., Fotiadis, D., and Müller, D. J. (2017). Detecting ligand-binding events and free energy landscape while imaging membrane receptors at subnanometer resolution. *Nano Lett.* 17, 3261–3269. doi: 10.1021/acs.nanolett.7b00941

Pickering, A. C., Vitry, P., Prystopiuk, V., Garcia, B., Höök, M., Schoenebeck, J., et al. (2019). Host-specialized fibrinogen-binding by a bacterial surface protein promotes biofilm formation and innate immune evasion. *PLoS Pathog.* 15:e1007816. doi: 10.1371/journal.ppat.1007816

Pinzón-Arango, P. A., Liu, Y., and Camesano, T. A. (2009). Role of cranberry on bacterial adhesion forces and implications for *Escherichia coli*-uroepithelial cell attachment. *J. Med. Food* 12, 259–270. doi: 10.1089/jmf.2008.0196

Ponnuraj, K., Bowden, M. G., Davis, S., Gurusiddappa, S., Moore, D., Choe, D., et al. (2003). A "dock, lock, and latch" structural model for a staphylococcal adhesin binding to fibrinogen. *Cell* 115, 217–228. doi: 10.1016/S0092-8674(03)00809-2

Prakobphol, A., Burdsal, C. A., and Fisher, S. J. (1995). Quantifying the strength of bacterial adhesive interactions with salivary glycoproteins. *J. Dent. Res.* 74, 1212–1218. doi: 10.1177/00220345950740051101

Prystopiuk, V., Feuillie, C., Herman-Bausier, P., Viela, F., Alsteens, D., Pietrocola, G., et al. (2018). Mechanical forces guiding *Staphylococcus aureus* cellular invasion. *ACS Nano* 12, 3609–3622. doi: 10.1021/acsnano.8b00716

Sherman, E., Bayles, K., Moormeier, D., Endres, J., and Wei, T. (2019). Observations of shear stress effects on *Staphylococcus aureus* biofilm formation. *mSphere* 4:e00372-19. doi: 10.1128/mSphere.00372-19

Sokurenko, E. V., Vogel, V., and Thomas, W. E. (2008). Catch-bond mechanism of force-enhanced adhesion: counterintuitive, elusive, but widespread? *Cell Host Microbe* 4, 314–323. doi: 10.1016/j.chom.2008.09.005

Sullan, R. M. A., Li, J. K., Crowley, P. J., Brady, L. J., and Dufrêne, Y. F. (2015). Binding forces of *Streptococcus mutans* P1 adhesin. *ACS Nano* 9, 1448–1460. doi: 10.1021/nn5058886

Tao, Y., Pinzón-Arango, P. A., Howell, A. B., and Camesano, T. A. (2011). Oral consumption of cranberry juice cocktail inhibits molecular-scale adhesion of clinical uropathogenic *Escherichia coli*. *J. Med. Food* 14, 739–745. doi: 10.1089/jmf.2010.0154

Thomas, W. E., Vogel, V., and Sokurenko, E. (2008). Biophysics of catch bonds. *Annu. Rev. Biophys.* 37, 399–416. doi: 10.1146/annurev.biophys.37.032807.125804

Tripathi, P., Beaussart, A., Alsteens, D., Dupres, V., Claes, I., von Ossowski, I., et al. (2013). Adhesion and nanomechanics of pili from the probiotic *Lactobacillus rhamnosus* GG. *ACS Nano* 7, 3685–3697. doi: 10.1021/nn400705u

Turner, N. A., Sharma-Kuinkel, B. K., Maskarinec, S. A., Eichenberger, E. M., Shah, P. P., Carugati, M., et al. (2019). Methicillin-resistant *Staphylococcus aureus*: an overview of basic and clinical research. *Nat. Rev. Microbiol.* 17, 203–218. doi: 10.1038/s41579-018-0147-4

Valotteau, C., Prystopiuk, V., Pietrocola, G., Rindi, S., Peterle, D., De Filippis, V., et al. (2017). Single-cell and single-molecule analysis unravels the multifunctionality of the *Staphylococcus aureus* collagen-binding protein Cna. *ACS Nano* 11, 2160–2170. doi: 10.1021/acsnano.6b08404

Vanzieleghem, T., Herman-Bausier, P., Dufrene, Y. F., and Mahillon, J. (2015). *Staphylococcus epidermidis* affinity for fibrinogen-coated surfaces correlates with the abundance of the SdrG adhesin on the cell surface. *Langmuir* 31, 4713–4721. doi: 10.1021/acs.langmuir.5b00360

Verbelen, C., and Dufrêne, Y. F. (2009). Direct measurement of *Mycobacterium*-fibronectin interactions. *Integr. Biol. Quant. Biosci. Nano Macro* 1, 296–300. doi: 10.1039/b901396b

Viela, F., Speziale, P., Pietrocola, G., and Dufrêne, Y. F. (2019). Mechanostability of the fibrinogen bridge between staphylococcal surface protein ClfA and endothelial cell integrin αVβ3. *Nano Lett.* 19, 7400–7410. doi: 10.1021/acs.nanolett.9b03080

Viljoen, A., Viela, F., Kremer, L., and Dufrêne, Y. F. (2020). Fast chemical force microscopy demonstrates that glycopeptidolipids define nanodomains of varying hydrophobicity on mycobacteria. *Nanoscale Horiz.* 5, 944–953. doi: 10.1039/C9NH00736A

Vitry, P., Valotteau, C., Feuillie, C., Bernard, S., Alsteens, D., Geoghegan, J. A., et al. (2017). Force-induced strengthening of the interaction between *Staphylococcus aureus* clumping factor B and loricrin. *mBio* 8:e01748-17. doi: 10.1128/mBio.01748-17

Walsh, E. J., Miajlovic, H., Gorkun, O. V., and Foster, T. J. (2008). Identification of the *Staphylococcus aureus* MSCRAMM clumping factor B (ClfB) binding site in the alphaC-domain of human fibrinogen. *Microbiol. Read. Engl.* 154, 550–558. doi: 10.1099/mic.0.2007/010868-0

Xiang, H., Feng, Y., Wang, J., Liu, B., Chen, Y., Liu, L., et al. (2012). Crystal structures reveal the multi-ligand binding mechanism of *Staphylococcus aureus* ClfB. *PLoS Pathog.* 8:e1002751. doi: 10.1371/journal.ppat.1002751

Xiao, J., and Dufrêne, Y. F. (2016). Optical and force nanoscopy in microbiology. *Nat. Microbiol.* 1:16186. doi: 10.1038/nmicrobiol.2016.186

Xiong, Y. Q., Sharma-Kuinkel, B. K., Casillas-Ituarte, N. N., Fowler, V. G., Rude, T., DiBartola, A. C., et al. (2015). Endovascular infections caused by methicillin-resistant *Staphylococcus aureus* are linked to clonal complex-specific alterations in binding and invasion domains of fibronectin-binding protein A as well as the occurrence of fnbB. *Infect. Immun.* 83, 4772–4780. doi: 10.1128/IAI.01074-15

Xu, C.-P., Boks, N. P., de Vries, J., Kaper, H. J., Norde, W., Busscher, H. J., et al. (2008). *Staphylococcus aureus*-fibronectin interactions with and without fibronectin-binding proteins and their role in adhesion and desorption. *Appl. Environ. Microbiol.* 74, 7522–7528. doi: 10.1128/AEM.00948-08

Conflict of Interest: YD is a Research Director at the FNRS.

The remaining author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Microbiology

REVIEW
published: 29 July 2020
doi: 10.3389/fmicb.2020.01829

Surface Proteins of *Staphylococcus epidermidis*

*Timothy J. Foster**

Department of Microbiology, Trinity College Dublin, Dublin, Ireland

Staphylococcus epidermidis is a ubiquitous commensal of human skin. The widespread use of indwelling medical devices in modern medicine provides an opportunity for it to cause infections. Disease causing isolates can come from many different genetic backgrounds. Multiply antibiotic resistant strains have spread globally. *S. epidermidis* has a smaller repertoire of cell wall anchored (CWA) surface proteins than *Staphylococcus aureus*. Nevertheless, these CWA proteins promote adhesion to components of the extracellular matrix including collagen, fibrinogen, and fibronectin and contribute to the formation of biofilm. The A domain of the accumulation associated protein Aap can promote adhesion to unconditioned biomaterial but must be removed proteolytically to allow accumulation to proceed by homophilic Zn^{2+}-dependent interactions. Mature biofilm contains amyloid structures formed by Aap and the small basic protein (Sbp). The latter contributes to the integrity of both protein and polysaccharide biofilm matrices. Several other CWA proteins can also promote *S. epidermidis* biofilm formation.

Keywords: biofilm, amyloid structures, bacterial adhesion, extracellular matrix, homophilic protein-protein interaction, MSCRAMM, phylogenetics

OPEN ACCESS

Edited by:
Mattias Collin,
Lund University, Sweden

Reviewed by:
Michael Otto,
National Institutes of Health (NIH), United States
Alexander Horswill,
University of Colorado Denver, United States

***Correspondence:**
Timothy J. Foster
tfoster@tcd.ie

Specialty section:
This article was submitted to Infectious Diseases, a section of the journal Frontiers in Microbiology

Received: *27 May 2020*
Accepted: *13 July 2020*
Published: *29 July 2020*

Citation:
Foster TJ (2020) Surface Proteins of Staphylococcus epidermidis. Front. Microbiol. 11:1829. doi: 10.3389/fmicb.2020.01829

INTRODUCTION

Staphylococcus epidermidis is a ubiquitous and primarily harmless commensal of human skin compared to the more pathogenic coagulase-positive *Staphylococcus aureus* (Becker et al., 2014). Modern medicine allows the successful treatment of patients with serious and potentially life-threatening illnesses. Indwelling medical devices such as intravenous catheters, prosthetic joints, and heart valves provide an opportunity for *S. epidermidis* to gain access to the body and to cause infections.

The human skin is acidic, dessicated, nutrient poor and has high osmolarity. Shotgun metagenomic whole genome sequencing has allowed the relative abundance of bacteria at different skin sites to be established (Byrd et al., 2018). *S. epidermidis* is a prominent member of the microbiome of both dry and moist skin, as well as areas with sebaceous glands. It is also an important component of the nasal microbiome (Byrd et al., 2018; Liu et al., 2020). There is growing awareness that *S. epidermidis* is not just a benign resident but rather has a proactive role in modulating the host immune system to promote survival of commensals and influence the development of the skin and nasal microbiome (PrabhuDas et al., 2011; Naik et al., 2012, 2015).

Phylogenetic analysis of a large number of commensal strains of *S. epidermidis* isolated from the skin of healthy humans and nosocomial disease-causing isolates indicates that the species comprises two discrete clusters called A/C and B (Meric et al., 2018). Defining the

core genome and the large and growing number of accessory genes indicates that *S. epidermidis* is open to horizontal gene transfer of mobile genetic elements and to chromosomal recombination. Organisms in the A/C cluster are more likely to cause nosocomial infections while the B cluster strains have a higher proportion of commensals.

The two groups are adapted to colonize distinct habitats on the skin and are genetically isolated. A/C strains are adapted to colonize the surface of skin and are fitter under acidic and osmotic stress, they grow well at pH 4.5 typical of macrophage and neutrophil phagosome acidification, and they form biofilm at pH7 (the pH of blood; Espadinha et al., 2019). B strains survive in the microaerophilic environment of deeper skin sites, they are more resistant to bactericidal fatty acids and they metabolize lipids found in sebaceous glands and hair follicles (Espadinha et al., 2019).

S. epidermidis isolated from 15 different body sites ranging from the face to the toe web was subjected to whole genome sequencing. Material from swabs was also analyzed by metagenomic sequencing (Zhou et al., 2020). This revealed that an individual is colonized by multiple lineages whose genomes have diversified by mutation and horizontal gene transfer. A broad representation of phylogenies from both A/C and B groups was found at most sites. Frequent transmission of bacteria between body sites was evident (e.g., face to hand). In contrast, *S. epidermidis* from the toe web showed little diversification suggesting that this is an isolated niche.

Organisms from both clades can cause infection. Genome wide association studies of commensal and disease isolates identified infection-associated genetic sequences in loci associated with biofilm formation, toxicity, inflammation, and resistance to antibiotics (Meric et al., 2018). Whole genes and short sequences carrying allelic variations within genes have been transmitted by chromosomal recombination, while entire SCC*mec* elements encoding β-lactam resistance and potential virulence determinants (Qin et al., 2017; Arora et al., 2020) have been transmitted horizontally. It was concluded that many strains from the two clades can cause disease.

In contrast, three distinct hospital-adapted clones from clonal complex 2 have spread globally (Lee et al., 2018). These strains are resistant to multiple antibiotics including β-lactams and rifampicin. The *rpoB* mutations that confer resistance to rifampicin were analyzed genetically and were found to contribute to insensitivity to the glycopeptide vancomycin. A small proportion of cells within the population expresses a low but clinically significant level resistance to vancomycin and will outgrow the majority in the presence of the drug to compromise the treatment of patients. A toxin expressed by the SCC*mec* element contributes to the pathogenesis of sepsis (Qin et al., 2017).

Biofilm formation on the surface of indwelling medical devices is a major virulence attribute. Two distinct mechanisms of biofilm accumulation have been identified. One requires formation of a polysaccharide glycocalyx and the other involves cell wall anchored (CWA) surface proteins.

S. epidermidis possesses a smaller repertoire of virulence factors compared to *S. aureus*. It lacks the immune evasion proteins and cytolytic toxins that are characteristics of *S. aureus*. It can express several CWA surface proteins including proteins that promote biofilm formation and bind components of the extracellular matrix such as fibronectin, collagen, and fibrinogen (**Table 1**). Their properties will be discussed in the context of their contribution to biofilm formation and the pathogenesis of implanted device-related infections.

CELL WALL ANCHORED SURFACE PROTEINS

Steps in Biofilm Formation

The ability to form biofilm on indwelling medical devices is crucially important and is a major virulence determinant of *S. epidermidis*. The focus of this review is the function of CWA proteins in this process (**Table 1**). Many clinical isolates can express both polysaccharide and protein based mechanisms of biofilm formation depending on *in vitro* growth conditions.

The hallmark of infections associated with indwelling devices is the ability of bacteria to adhere to the implant and to grow as a biofilm (Otto, 2018). Biofilm formation is initiated by attachment of bacteria to the surface of the biomaterial, either

TABLE 1 | Functions of surface-associated proteins.

Protein	Ligands	Role in biofilm formation
SdrF	A and B domains bind collagen	Adhesion to collagen deposited on *ex vivo* biomaterial
SdrG/Fbe	A domain binds fibrinogen	Adhesion to abiotic surfaces Promotes adhesion to conditioned biomaterial
Aap	None reported.	Adhesion to abiotic surfaces *via* A domain.
Embp	G5-E repeats engage in homophilic interactions FIVAR and FIVAR-GA domains bind fibronectin	Cell-cell accumulation. Amyloid formation in mature biofilm. Adhesion to conditioned biomaterial? FIVAR-GA region binds unknown ligand(s) on adjacent cells to promote accumulation
SdrC	Not known	Biofilm formation by unknown mechanism
SesJ	Plasminogen	Not known
SesI	Not known	Might promote adhesion to abiotic surfaces. Otherwise unknown
Geh	Collagen	Pure Geh binds collagen *in vitro*. Bacterial adhesion to immobilized collagen not tested
AtlE	Autolysin	Release of e-DNA promotes adhesion to abiotic surfaces
	Recombinant protein binds vitronectin, fibronectin and fibrinogen *in vitro*	Adhesion to conditioned biomaterial?
Sbp	Secreted protein. Associated with cell wall. Possibly binds Aap	Amyloid formation contributes to biofilm integrity

SdrF, SdrG/Fbe, Aap, Embp, SdrC, SesJ, and SesI are anchored covalently to peptidoglycan by sortase.

to an abiotic surface prior to or at the time of implantation, or to a surface that has been conditioned by deposition of host proteins. Attachment is followed by the accumulation phase where bacteria multiply to form multicellular communities. This requires the cells to stick to each other by mechanisms that promote intercellular adhesion. The biofilm undergoes maturation with the creation of channels by release of some cells in the matrix. In devices that are in contact with the bloodstream detached cells can disseminate.

Investigation of the mechanisms involved in biofilm formation by *S. epidermidis* has been carried out with a small number of strains that could be manipulated genetically. This has mainly involved studies of biofilm formation *in vitro* under static growth conditions and dynamically in flow chambers allowing biofilms to be visualized and quantified by confocal microscopy. *In vitro* studies were complemented with *in vivo* models involving catheter segments implanted subcutaneously or intravenously in rodents.

Adhesion to Abiotic Surfaces

Investigation into the mechanistic basis of biofilm formation by *S. epidermidis* began with the analysis of biofilm defective transposon insertion mutants of strain O-47 (Heilmann et al., 1996). Some mutations knocked out expression of the autolysin AtlE (Heilmann et al., 1997). The mutant was defective in attachment to a polystyrene surface. However, it is unclear if the AtlE protein itself is an adhesin because a null mutation affecting a major enzyme involved in cell wall metabolism results in pleiotropic changes to the cell surface. Point mutations or short in-frame deletions that are defective in adhesion while remaining enzymatically active are required.

The presence of DNase reduced adhesion of both clinical isolates and laboratory strains 1457 and RP62a to plastic and glass surfaces under static and hydrodynamic conditions (Qin et al., 2007). An AtlE mutant of strain 1457 produced much lower levels of extracellular (eDNA). It was concluded that release of eDNA from a small number of cells in the population by the autolytic activity of AtlE promotes bacterial attachment to abiotic surfaces and is important in the primary attachment phase of biofilm formation. Thus, the role of AtlE may be to produce eDNA rather than itself acting as an adhesin.

Adhesion to the abiotic surface plastic biomaterial can also be promoted by the cell wall-associated proteins SdrF (Arrecubieta et al., 2009) and the accumulation associated protein Aap (Conlon et al., 2014). Genomic analysis of *S. epidermidis* revealed the presence of a gene encoding a CWA protein with homology the biofilm associated protein Bap of *S. aureus* (Bhp, Bap homologous protein; Bowden et al., 2005). However, while Bap promotes both attachment to abiotic surfaces and biofilm accumulation (Cucarella et al., 2001), nothing is known about Bhp.

SdrF is a member of the microbial surface components recognizing adhesive matrix molecules (MSCRAMM) family of CWA proteins with an N-terminal A domain linked to B repeats (**Figure 1**). The archetypal MSCRAMM SdrG/Fbe is described below. SdrF promotes adhesion to the highly textured surface of the hydrophobic polymer Dacron used to coat the drive lines of ventricular assist devices (VAD) and could provide *S. epidermidis* with a portal for entering and colonizing the indwelling driveline (Arrecubieta et al., 2009). This was investigated using the non-adhesive surrogate host *Lactococcus lactis* expressing full length SdrF and truncates expressing the A domain or the B domains alone. Both domains promoted adhesion which could be blocked using domain-specific antibodies.

Adhesion to Conditioned Biomaterial

Newly implanted indwelling devices are rapidly coated with host plasma proteins such as fibronectin and fibrinogen (Vaudaux et al., 1995). The conditioning layer of long term implants such as a VAD drive lines also contains collagen (Arrecubieta et al., 2009).

SdrG/Fbe Binds Fibrinogen

The SdrG/Fbe protein is CWA protein of the MSCRAMM family (**Figure 1**; Foster et al., 2014; Foster, 2019). The A domain binds to the C-terminus of the β-chain of fibrinogen in the central E region. X ray crystallography of the A domain in the unbound apo form and in complex with the β-chain fibrinopeptide allowed formulation of the "dock lock latch" mechanism of ligand binding (Ponnuraj et al., 2003). A key feature of the binding mechanism is that it is facilitated by and strengthened by shear stress (Milles et al., 2018). Forces equivalent to those required to break a covalent bond are required to separate SdrG from its bound ligand (Herman et al., 2014).

Binding to immobilized fibrinogen/fibrin in the conditioning layer of an indwelling device is important in initiating device-related infections. *S. epidermidis* HB adhered to immobilized fibrinogen *in vitro* in a SrdG/Fbe-dependent manner (Hartford et al., 2001; Pei and Flock, 2001). Expression of the MSCRAMM was required for the bacterium to colonize the surface of a catheter implanted intravenously in rats (Guo et al., 2007).

SdrF Binding to Collagen

In addition to promoting adhesion to abiotic surfaces, SdrF also promotes bacterial adhesion to immobilized type I collagen (Arrecubieta et al., 2007). This could be important in VAD driveline infections where the conditioning layer contains collagen (Arrecubieta et al., 2009). Adhesion of *S. epidermidis* to *ex vivo* drivelines was in part promoted by SdrF. Atomic force microscopy showed that both the A domain and B repeats could bind collagen (Herman-Bausier and Dufrene, 2016). These interactions were weak in contrast to the strong binding that is characteristic of the collagen binding protein Cna of *S. aureus* binding collagen by the hug mechanism (Zong et al., 2005; Herman-Bausier et al., 2016; Foster, 2019).

Does SesC Promote Binding to Fibrinogen?

Expression of the CWA surface protein SesC ectopically from a multicopy plasmid in a low fibrinogen binding strain of *S. aureus* and in *S. epidermidis* RP62a resulted in a slight but

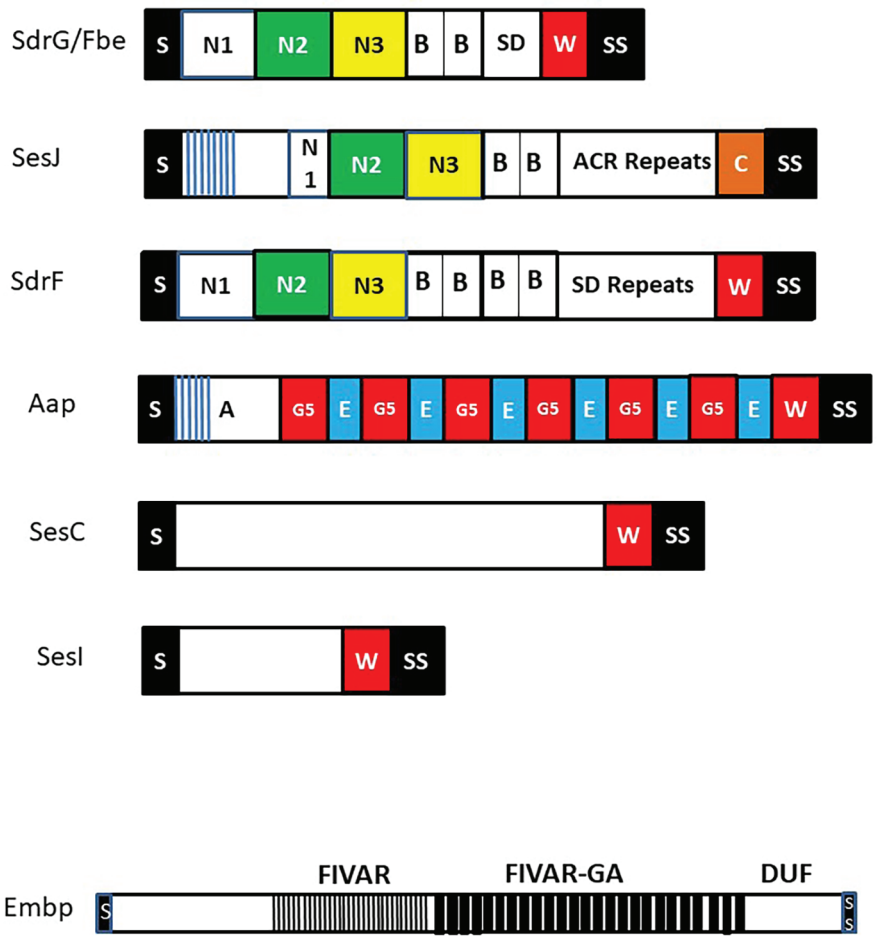


FIGURE 1 | Cell wall anchored (CWA) surface proteins of *Staphylococcus epidermidis*. Schematic diagram of CWA surface proteins that feature in the text. Each protein has a secretory signal (S) at the N-terminus that is removed during secretion and at the C-terminus a sorting signal (SS) that promotes covalent anchorage to cell wall peptidoglycan. Many CWA proteins have a proline-rich cell wall spanning domain W while SesJ has a collagen-like sequence C. Extracellular matrix binding protein (Embp) is shown separately to emphasize the order of magnitude greater size compared to the other proteins. It contains several domains of unknown function (DUF) C-terminal to the found in various architectures (FIVAR) and FIVAR–G-related albumin binding (FIVAR-GA) repeats. Proteins in the microbial surface components recognizing adhesive matrix molecules (MSCRAMM) family are shown at the top with the canonical N2 and N3 subdomains that have potential to engage ligands by the dock lock latch mechanism. Each contains separately folded B repeat domains and unfolded serine asparate (SD) repeats or in the case of SesJ a distinct aspartate containing region (ACR). The SD repeat region of SdrG from strain RP62a is very short. Its length will likely vary in other strains. The N-termini of the A domains of SesJ and Aap contain distinct short N-terminal repeats (blue lines).

significant increase in bacterial adherence to immobilized fibrinogen (Shahrooei et al., 2009). However, recombinant SesC protein did not bind Fg *in vitro*, so the affinity and specificity of the interaction could not be measured. Therefore, the suggestion that SesC promotes binding to Fg must be treated with caution. The role of SesC in biofilm formation is discussed below.

Embp Binds to Fibronectin

S. epidermidis expresses a very large (10,204 residue) CWA surface protein called the extracellular matrix binding protein (Empb; Williams et al., 2002; Christner et al., 2010). A major part of the protein comprises two long repeated domains called found in various architectures (FIVAR, n = 21) and FIVAR–G-related albumin binding (FIVAR-GA, n = 38; **Figure 1**). Overexpression of Embp by *S. epidermidis* promoted bacterial adhesion to immobilized fibronectin. Recombinantly expressed FIVAR and FIVAR-GA modules both bound Fn in solid phase ELISA-type binding assays with similar profiles. Surface plasmon resonance indicated a binding affinity in the nanomolar range. Embp expression was induced by growth in serum, so it is possible that the protein contributes to adhesion to conditioned biomaterial *in vivo*.

Autolysins and Other Surface Associated Proteins

It has been suggested that autolysins, AtlE and Aae, promote *S. epidermidis* adhesion to immobilized plasma proteins vitronectin, fibronectin, and fibrinogen and that they could

be involved in initiating biofilm formation on conditioned biomaterial surfaces (Heilmann et al., 1997, 2003). However, these binding studies were only performed with purified recombinant proteins and must be interpreted with caution. The involvement of autolysins in promoting bacterial adhesion to immobilized ligands was not investigated. An AtlE mutant had reduced virulence in a rat intravenous catheter infection model but this could be attributed to pleiotropic effects and lack of fitness due to loss of the cell wall metabolizing autolysin (Rupp et al., 2001).

Gycerol Ester Hydrolase

The glycerol ester hydrolase (Geh) can bind to collagen *in vitro* (Bowden et al., 2002). However, any role in promoting bacterial adhesion during biomaterial associated infection remains speculative.

Biofilm Accumulation

Following attachment to surfaces bacteria multiply and form multicellular aggregates requiring cells to adhere to each other (Becker et al., 2014; Otto, 2018). Intercellular adhesion can be promoted by the polysaccharide intercellular adhesin (PIA) or by surface proteins. The extracellular matrix of a mature biofilm contains DNA and proteins released from lysed cells. Biofilm formed *in vivo* will also harbor host proteins.

Polysaccharide Intercellular Adhesin

In many strains, the molecule responsible for accumulation phase of biofilm is PIA, a homopolymer comprising at least 130 units of β-1-6-linked N-acetylglucosamine (PNAG; Mack et al., 1996; reviewed by Otto, 2018). The polysaccharide chains are synthesized intracellularly by the integral membrane proteins IcaA and IcaD (Mack et al., 1996; Vuong et al., 2004a). The polysaccharide is transported across the membrane by IcaC. The extracellular polysaccharide is partially (15–20%) deacylated by the IcaD protein. This is essential for biofilm formation because the exposed positively charged NH^{3+} groups allow the molecule to attach by electrostatic interactions to the negatively charged bacterial cell surface (Vuong et al., 2004a). However, the molecules involved in binding PNAG to the cell surface have not been identified. This process is likely to be multifactorial; a mutant defective in wall teichoic acid still carried PNAG on its surface and formed biofilm *in vitro* (Vergara-Irigaray et al., 2008).

The presence of PIA on the cell surface is an important immune evasion mechanism (Le et al., 2018) both because aggregates inhibit engulfment by phagocytes and also because it acts as a capsule and inhibits opsonophagocytosis (Vuong et al., 2004b). Biofilm formation *in vivo* by an Ica^+ strain induced lower levels of proinflammatory cytokines than an Ica^- mutant (Fredheim et al., 2011). Ica strongly activated complement but paradoxically reduced the activation of phagocytes contributing to reduced eradication of biofilm. Inflammation adjacent to an infected catheter insertion site was more severe with a wild type strain (Kristian et al., 2008). Ica also protected against antimicrobial peptides (Vuong et al., 2004b).

The *icaADBC* genes form an operon which is negatively regulated by the IcaR repressor protein. Global transcriptional regulators SigB and SarA, as well as the quorum sensing system LuxS also contribute to the regulation of *ica* expression. The environmental signals that regulate *ica in vivo* are complex and not well understood (O'Gara, 2007).

Accumulation Associated Protein

Device related infections can frequently be caused by strains of *S. epidermidis* that lack the ability to form PNAG/PIA (Kogan et al., 2006; Rohde et al., 2007). Indeed, the *ica* genes are only present in 37% of lineage A–C strains and 4% of group B strains (Espadinha et al., 2019) and one third of disease isolates lack *ica* (Rohde et al., 2007). In contrast, the gene that encodes the CWA accumulation associated protein Aap is widespread in clinical isolates (Kogan et al., 2006; Petrelli et al., 2006; Rohde et al., 2007; Hellmark et al., 2013).

Aap promoted biofilm formation by strain 1457 *in vitro* under static and dynamic conditions and it promoted colonization of a catheter implanted into the jugular vein of rats (Schaeffer et al., 2015). It contributed to biofilm formation *in vivo* either alone in strains that lack *ica* or in strains that also express PIA/PNAG.

The N-terminal A domain of Aap promoted attachment to abiotic surfaces (Conlon et al., 2014). The A domain comprises two subdomains, the N-terminal repeats followed by a lectin-like domain (**Figure 1**). The A domain must be removed either completely or in part by proteolysis allowing the exposed C-terminal G5-E domains to participate in biofilm accumulation (**Figure 2**; Rohde et al., 2005; Paharik et al., 2017). The metalloprotease SepA has two cleavage sites in Aap, one is located between the N-terminal repeats and the lectin-like domain while the other occurs between the A domain and the G5E repeats. Aap isolated from the surface of planktonic cells and from a mature biofilm matrix comprises a mixture of fully and partially cleaved proteins (Rohde et al., 2005; Yarawsky et al., 2020).

The Aap protein of *S. epidermidis* is very similar to the SasG protein of *S. aureus* (Roche et al., 2003). Conclusions drawn from studies with Aap can most likely be applied to SasG and vice versa. Each C-terminal B repeat domain of Aap and SasG comprises two segments called G5 and E (Gruszka et al., 2012, 2015; Conrady et al., 2013). The regions have unusual properties allowing them to fold efficiently to form extended fibrils that are very stable. The E and G5 domains are potentially intrinsically disordered. The minimum unit capable of correct folding is G5-E-G5.

Biofilm accumulation promoted by Aap occurs in two stages following cleavage of the A domain (**Figure 2**). First Zn^{2+}-dependent homophilic interactions occur between exposed G5E repeat domains on adjacent cells that result in twisted rope-like structures (Conrady et al., 2013). This interaction occurs *in vitro* with purified G5-E repeat proteins in solution. Atomic force microscopy studies performed with *S. aureus* expressing SasG showed that considerable force is required to separate cells held together by fully engaged proteins (Formosa-Dague et al., 2016). This interaction could be reversed by addition of the chelator

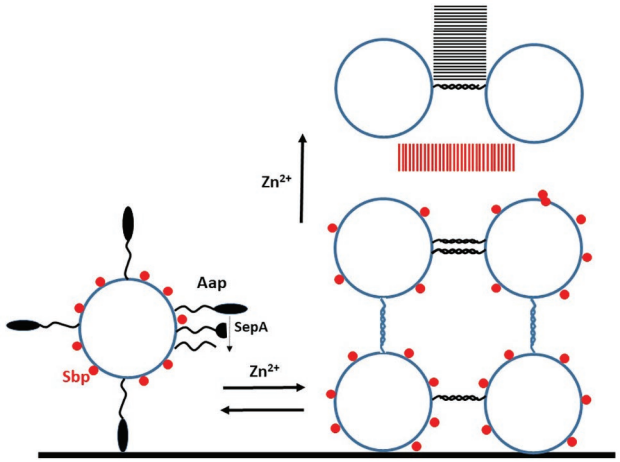


FIGURE 2 | Mechanism of biofilm formation promoted by Aap. Schematic diagram of events that occur during biofilm formation by *S. epidermidis* strains expressing Aap. The intact A domain of the Aap protein can promote attachment to unconditioned (abiotic) biomaterial. In order to participate in the accumulation stages, the A domain must be removed either completely or partially by protease. The small basic protein (Sbp) is associated non-covalently with the cell wall. The first phase of biofilm formation is Zn^{2+}-dependent and reversible by adding a chelator. This involves homophilic interactions between G5-E repeats on adjacent cells. As the biofilm matures Sbp and the Aap proteins form amyloid fibrils (parallel lines), in the case of the latter this is Zn^{2+}-dependent but irreversible. Sbp is also implicated in biofilm formation involving PIA.

DTPA reflecting the Zn^{2+} dependence of recombinant protein-protein interactions *in vitro* and Aap-promoted bacterial biofilm formation.

The second stage is assembly of G5-E repeats into higher order amyloid structures in a Zn^{2+}- dependent fashion. This was investigated by sedimentation analysis (Yarawsky et al., 2020). Amyloids were visualized by transmission electron microscopy both with recombinant G5-E proteins in solution and in a bacterial biofilm matrix formed by strain RP62a *in vitro*. Mass spectroscopy analysis of tryptic digests of purified aggregates isolated from biofilm detected peptides from the lectin domain of region A indicating that complete removal of the A domain is not required for amyloid formation. The amyloids could not be reversed by the chelator DTPA, which is consistent with resistance of established biofilms to dissociation by the chelator.

Small Basic Protein

A small (18 kDa) basic protein called Sbp was recently shown to have an important role in biofilm (Decker et al., 2015). In solution, Sbp is monomeric and partially unfolded. Following agitation recombinant, Sbp forms multimeric protein complexes – amyloid fibrils – that can be stained with the fluorescent dyes thioflavin S and Congo red (Wang et al., 2018).

The role of Sbp in biofilm formation was studied in strain 1457 that forms robust biofilm incorporating both PIA/PNAG and Aap (Decker et al., 2015). Despite being a secreted soluble protein, Sbp was isolated predominantly from the cell wall fraction rather than in the culture supernatant of planktonic bacteria growing under biofilm forming conditions. Confocal microscopy of biofilm showed that Sbp forms large clumps (possibly amyloids) that are unevenly dispersed and concentrated at the surface of the cell aggregates. By studying isogenic mutants, it was shown that Sbp is not required for primary attachment of cells to unconditioned biomaterial but that it is required for mechanically robust tethering of mature biofilm. Biofilm formed statically by an Sbp mutant was more easily washed away. This is consistent with much lower levels of biofilm formed by the mutant under dynamic shear stress. However, *in vivo* a lack of Sbp did not affect the density of biofilm formed on subcutaneously implanted catheters but it should be noted that this model involves injecting bacteria into the lumen of the catheter where it is unlikely that bacteria would be subjected to shear stress. It would be interesting to test the mutant in the rat jugular vein catheter model where bacteria colonize the indwelling device following haematogenous spread.

Contradictory data have been published about the ability of Sbp to interact with the G5-E repeat region of Aap. In solution, no interaction was detected between recombinant Sbp and the minimum folded and functional Aap G5-E-G5 protein (Wang et al., 2018). In contrast, when Sbp was subjected to SDS-PAGE and electroblotting, it showed binding to the soluble full length G5-E repeat protein. In an ELISA based assay, soluble Sbp bound to immobilized G5-E repeats in a dose-dependent and specific manner (Decker et al., 2015). These observations led to the conclusion that Sbp contributes to the biofilm structure by interaction with the G5-E repeats of Aap. Sbp is clearly of importance in the formation of and integrity of biofilms most likely because of its innate ability to form amyloids. Whether or not it interacts with the G5E-repeats of Aap is a lesser consideration.

SesC and Biofilm Formation

The possible role of SesC in biofilm formation by *S. epidermidis* was investigated by testing the inhibitory effect of anti-SesC IgG *in vitro* and *in vivo* and by ectopic expression of the protein in *S. aureus* (Shahrooei et al., 2009, 2012; Khodaparast et al., 2016). Anti-SesC IgG reduced both the primary attachment and accumulation phases of biofilm formation *in vitro*. However, the suggestion that SesC has a direct role in these events must be treated with a degree of caution since the presence of IgG bound to a protein on the bacterial cell surface could have pleiotropic inhibitory effects. The clearest evidence for a direct role was the enhanced protein-dependent biofilm formed *in vitro* and *in vivo* by 8325-4 expressing SesC (Khodaparast et al., 2016). Further studies with *sesC* null mutants are required.

Embp and Biofilm Formation

Overexpression of the giant surface protein Embp *in vitro* by a clinical isolate of *S. epidermidis* that lacks both *ica* and *aap* genes led to clustering of planktonic cells and to biofilm formation (Christner et al., 2010). Embp did not promote primary attachment to abiotic surfaces but rather biofilm accumulation.

The mechanistic basis of aggregation seems to involve heterophilic interactions between the repeated FIVAR-GA domains and ligands on the surface of adjacent cells. Recombinant FIVAR-GA but not FIVAR domains promoted aggregation of cells that lacked the ability to express Embp. Inhibition of Embp-promoted biofilm by antibody directed against both FIVAR-GA or FIVAR domains could be interpreted as the specific blocking of heterophilic interactions in the case of the former and indirect blocking with the latter.

Biofilm Maturation

As the biofilm develops it undergoes maturation to form three-dimensional structures with mushroom-like towers and fluid filled gaps (Otto, 2018). Small amphipathic α-helical peptides with surfactant properties called phenol soluble modulins (PSMs) disrupt non-covalent bonds formed between cells during biofilm development to promote the formation of channels (Le et al., 2019). Bacterial cells that are released from the maturing biofilm *in vivo* enter the bloodstream causing bacteraemia. Mutants of laboratory strain 1457 lacking the ability to express PSMs formed unstructured biofilms under both static and flow conditions *in vitro* that lacked channels and had a greater mass (Wang et al., 2011; Le et al., 2019). In a mouse model of subcutaneous catheter-associated biofilm infection, the *psm* null mutant formed a more substantial biofilm but was disseminated less effectively. This is consistent with the notion that PSMs promote biofilm maturation and structuring.

There is some debate as to whether PSMs contribute to the biofilm matrix by forming amyloid structures (Schwartz et al., 2012). Some of the PSMs expressed by *S. aureus* have the ability to form amyloids *in vitro* (Schwartz et al., 2012) but none of the *S. epidermidis* PSMs displayed that property (Le et al., 2019). Enhanced resistance of wild type *S. aureus* biofilm to degradation by DNase compared to a PSM-defective mutant that was interpreted as being a reflection of PSM amyloids (Schwartz et al., 2012) but could be the result of direct binding of PSMs to DNA (Zheng et al., 2018).

The *psm* genes are directly regulated by the Agr quorum sensing system (Queck et al., 2008), and PSM expression is likely to be induced in the cells that are closely packed together in a biofilm matrix. Blocking Agr has been proposed as an anti-virulence strategy (Dickey et al., 2017). In the case of *S. epidermidis*, this may not influence the formation of a biofilm on an indwelling medical device but could instead reduce bacteraemia, dissemination, and sepsis.

Other Cell Wall Anchored Proteins

Surface proteins that have been shown to have a role in biofilm formation have been discussed in the previous section. Two other CWA proteins expressed by *S. epidermidis* with possible roles in the pathogenesis of bloodstream infection have been investigated.

SesJ

The majority of genes encoding members of the MSCRAMM family of CWA proteins of *S. epidermidis* are present in the core genome (Conlan et al., 2012). A notable exception is *sesJ* which is carried by SCC*mec* or ACME mobile genetic elements (Arora et al., 2020). SesJ was present in about 18% of clinical isolates from bloodstream infections collected in two US hospitals.

SesJ is a chimaeric protein that is a member of a subfamily of MSCRAMMs (**Figure 1**; Arora et al., 2016). Members occur in other CoNS including SdrI in *S. saprophiticus* (Sakinc et al., 2006). The protein has an N-terminal repeat region comprising 15 residues repeated 13–15 times. This is reminiscent of the NTR region of Aap although they lack any sequence similarity. The A region N2 and N3 subdomains have all of the features of a typical MSCRAMM including a TYTFTDYVD motif and a putative latching peptide. However, the N1 subdomain is shorter than that of SdrG and SdrF. The A region occurs in two isoforms that are 95% identical (Arora et al., 2020). This is followed by two B repeats which are similar in sequence and predicted structure to the B repeats of SdrG and SdrF. Linking the A and B regions to the cell wall are aspartate containing repeats, in contrast to the serine aspartate dipeptide repeats of typical MSCRAMMs (Foster et al., 2014; Foster, 2019).

The *sesJ* gene is accompanied by two genes encoding glycosyltransferases (Arora et al., 2020). The gene encoding isoform 1 SesJ is located in an SCC*mec* typeIV element while the isoform 2 gene is located within an ACME element. These were found mainly in ST2, ST5, and ST210 strains.

Studies to identify the ligand(s) recognized by SesJ have been inconclusive (Arora et al., 2020). The recombinant A domain did not bind to fibrinogen, fibronectin, laminin, vitronectin, and several types of collagen. It did, however, bind plasminogen although the involvement of the DLL mechanism was not investigated. *S. aureus* expresses several CWA proteins that bind plasminogen (Pietrocola et al., 2016). The best characterized is FnBPB where the A domain binds *via* ionic bonds between surface located lysine residues and kringle domain 4 of the host protein. DLL is not involved. It is possible that SesJ binds plasminogen by a similar mechanism. Plasminogen captured on the bacterial cell surface could be activated to the potent serine protease plasmin by host plasminogen activators. This could contribute to pathogenesis, for example, by degrading opsonins. A more exhaustive search for potential ligands including those that bind *S. aureus* MSCRAMMs such as cytokeratins, loricrin, and elastin and investigating a role in biofilm formation is warranted.

SesI

The CWA protein SesI is small and has no distinguishing features. The presence of the *sesI* gene is associated with disease causing isolates and less so with commensals isolated from the skin or nares (Qi et al., 2017). The majority of *sesI* carrying strains were from ST2. A *sesI* defective mutant of RP62a was slightly less adhesive to an abiotic surface and planktonic cells formed fewer clumps compared to wild type. However, the density of biofilm formed by the mutant under static growth conditions *in vitro* was no different to that of the wild type. The mutant was not tested in a biofilm infection model. The association of SesI with nosocomial isolates could be coincidental.

DISCUSSION

The genome sequences of more than 1,000 clinical and commensal skin isolates of *S. epidermidis* have recently become available (Meric et al., 2018; Espadinha et al., 2019; Zhou et al., 2020). As more strains were sequenced the core genome shrank and the accessory genome expanded revealing that *S. epidermidis* is open to gene exchange. A detailed genome wide association study (GWAS) failed to determine if any surface protein is associated with disease isolates although polymorphisms within some surface protein genes may be involved (Meric et al., 2018). A study focussing on CWA proteins is warranted.

The genes encoding the SesJ protein and its associated glycosyl transferases are located within SCC mobile genetic elements (Arora et al., 2020). The *sesJ* gene is present in about 18% of disease isolates from several different STs including ST2 that were isolated in two American hospitals. However, the SCC*mec* type IV element carrying *sesJ* is not present in the globally disseminated ST2 strains which instead harbor a SCC*mec* type III element (Lee et al., 2018). The *sesJ* gene was also present in a small collection of commensal isolates. A detailed analysis of clinical and commensal isolates from diverse sources will shed light on any possible association of SesJ with disease although the comprehensive GWAS analysis mentioned above did not report any such link (Meric et al., 2018).

The identification of amyloid fibers as integral components of biofilm matrices is a significant development in understanding the complexities of biofilm formation. Initially the formation of amyloids by PSMs in *S. aureus* was suggested to be of importance (Schwartz et al., 2012) but this proved to be controversial when studies with *S. epidermidis* did not concur (Zheng et al., 2018).

The ubiquitous small basic protein (Sbp) has an important role in biofilm formation (Decker et al., 2015). The Sbp protein is secreted as a soluble monomer and can readily form amyloid fibrils *in vitro* (Wang et al., 2018). Sbp amyloids are integral components of the biofilm matrix formed by cells expressing both PIA/PNAG and Aap (Decker et al., 2015). The Aap protein and its homologue SasG in *S. aureus* trigger biofilm accumulation by homophilic Zn^{2+}-dependent twisted rope interactions (Conrady et al., 2013; Formosa-Dague et al., 2016). As the biofilm matures Aap proteins form amyloid fibrils (Yarawsky et al., 2020). The structures of Sbp and Aap amyloids await elucidation.

Prosthetic bone and joint infections often require surgical intervention because treatment with antibiotics alone is inadequate due to many of the bacteria being in a semi-dormant or persister state. Novel therapies with phage lysins, biofilm matrix degrading enzymes, and quorum sensing blockers have been proposed (reviewed by Otto, 2018). Inhibition of the Agr quorum sensing system might reduce dissemination of cells from the infected device due to reduced expression of PSMs but this may not prevent with the establishment of biofilm (Le et al., 2019).

Monocytes that are associated with biofilm infections are polarized toward an anti-inflammatory state (Le et al., 2018; Yamada et al., 2020). Directed uptake by monocytes of nanoparticles coated with the mitochondrial oxidative phosphate inhibitor oligomycin reprogrammed metabolism toward pro-inflammatory glycolysis (Yamada et al., 2020). These macrophages reduced biofilm density *in vitro* and *in vivo* in a mouse model of biomaterial-associated bone infection. Combined with antibiotics, this led to sterilization of the infected implant. This study was performed with *S. aureus* but it is worthwhile speculating that this treatment will also reduce *S. epidermidis* prosthetic joint infection.

Significant advances have been made in understanding the mechanistic basis of cell-cell accumulation during biofilm formation. The importance of the Sbp in both PIA and Aap promoted biofilm and its formation of amyloids, as well as amyloid formation during the late accumulation phase by Aap indicates an additional layer of complexity. Prevention of biofilm formation on indwelling devices and disruption of established biofilm is a goal for future studies. GWAS studies identified several factors that are associated with disease causing strains but key factors that could be targeted for intervention were not apparent.

AUTHOR CONTRIBUTIONS

The author confirms being the sole contributor of this work and has approved it for publication.

REFERENCES

Arora, S., Li, X., Hillhouse, A., Konganti, K., Little, S. V., Lawhon, S. D., et al. (2020). *Staphylococcus epidermidis* MSCRAMM SesJ is encoded in composite islands. *mBio* 11, e02911–e02919. doi: 10.1128/mBio.02911-19

Arora, S., Uhlemann, A. C., Lowy, F. D., and Hook, M. (2016). A novel MSCRAMM subfamily in coagulase negative *Staphylococcal* species. *Front. Microbiol.* 7:540. doi: 10.3389/fmicb.2016.00540

Arrecubieta, C., Lee, M. H., Macey, A., Foster, T. J., and Lowy, F. D. (2007). SdrF, a *Staphylococcus epidermidis* surface protein, binds type I collagen. *J. Biol. Chem.* 282, 18767–18776. doi: 10.1074/jbc.M610940200

Arrecubieta, C., Toba, F. A., von Bayern, M., Akashi, H., Deng, M. C., Naka, Y., et al. (2009). SdrF, a *Staphylococcus epidermidis* surface protein, contributes to the initiation of ventricular assist device driveline-related infections. *PLoS Pathog.* 5:e1000411. doi: 10.1371/journal.ppat.1000411

Becker, K., Heilmann, C., and Peters, G. (2014). Coagulase-negative *Staphylococci*. *Clin. Microbiol. Rev.* 27, 870–926. doi: 10.1128/CMR.00109-13

Bowden, M. G., Chen, W., Singvall, J., Xu, Y., Peacock, S. J., Valtulina, V., et al. (2005). Identification and preliminary characterization of cell-wall-anchored proteins of *Staphylococcus epidermidis*. *Microbiology* 151, 1453–1464. doi: 10.1099/mic.0.27534-0

Bowden, M. G., Visai, L., Longshaw, C. M., Holland, K. T., Speziale, P., and Hook, M. (2002). Is the GehD lipase from *Staphylococcus epidermidis* a collagen binding adhesin? *J. Biol. Chem.* 277, 43017–43023. doi: 10.1074/jbc.M207921200

Byrd, A. L., Belkaid, Y., and SEGRE, J. A. (2018). The human skin microbiome. *Nat. Rev. Microbiol.* 16, 143–155. doi: 10.1038/nrmicro.2017.157

Christner, M., Franke, G. C., Schommer, N. N., Wendt, U., Wegert, K., Pehle, P., et al. (2010). The giant extracellular matrix-binding protein of *Staphylococcus epidermidis* mediates biofilm accumulation and attachment to fibronectin. *Mol. Microbiol.* 75, 187–207. doi: 10.1111/j.1365-2958.2009.06981.x

Conlan, S., Mijares, L. A., Program, N. C. S., Becker, J., Blakesley, R. W., Bouffard, G. G., et al. (2012). *Staphylococcus epidermidis* pan-genome sequence analysis reveals diversity of skin commensal and hospital infection-associated isolates. *Genome Biol.* 13:R64. doi: 10.1186/gb-2012-13-7-r64

Conlon, B. P., Geoghegan, J. A., Waters, E. M., Mccarthy, H., Rowe, S. E., Davies, J. R., et al. (2014). Role for the A domain of unprocessed accumulation-associated protein (Aap) in the attachment phase of the *Staphylococcus epidermidis* biofilm phenotype. *J. Bacteriol.* 196, 4268–4275. doi: 10.1128/JB.01946-14

Conrady, D. G., Wilson, J. J., and Herr, A. B. (2013). Structural basis for Zn^{2+}-dependent intercellular adhesion in staphylococcal biofilms. *Proc. Natl. Acad. Sci. U. S. A.* 110, E202–E211. doi: 10.1073/pnas.1208134110

Cucarella, C., Solano, C., Valle, J., Amorena, B., Lasa, I., and Penades, J. R. (2001). Bap, a *Staphylococcus aureus* surface protein involved in biofilm formation. *J. Bacteriol.* 183, 2888–2896. doi: 10.1128/JB.183.9.2888-2896.2001

Decker, R., Burdelski, C., Zobiak, M., Buttner, H., Franke, G., Christner, M., et al. (2015). An 18 kDa scaffold protein is critical for *Staphylococcus epidermidis* biofilm formation. *PLoS Pathog.* 11:e1004735. doi: 10.1371/journal.ppat.1004735

Dickey, S. W., Cheung, G. Y. C., and Otto, M. (2017). Different drugs for bad bugs: antivirulence strategies in the age of antibiotic resistance. *Nat. Rev. Drug Discov.* 16, 457–471. doi: 10.1038/nrd.2017.23

Espadinha, D., Sobral, R. G., Mendes, C. I., Meric, G., Sheppard, S. K., Carrico, J. A., et al. (2019). Distinct phenotypic and genomic signatures underlie contrasting pathogenic potential of *Staphylococcus epidermidis* clonal lineages. *Front. Microbiol.* 10:1971. doi: 10.3389/fmicb.2019.01971

Formosa-Dague, C., Speziale, P., Foster, T. J., Geoghegan, J. A., and Dufrene, Y. F. (2016). Zinc-dependent mechanical properties of *Staphylococcus aureus* biofilm-forming surface protein SasG. *Proc. Natl. Acad. Sci. U. S. A.* 113, 410–415. doi: 10.1073/pnas.1519265113

Foster, T. J. (2019). The MSCRAMM family of cell-wall-anchored surface proteins of gram-positive cocci. *Trends Microbiol.* 27, 927–941. doi: 10.1016/j.tim.2019.06.007

Foster, T. J., Geoghegan, J. A., Ganesh, V. K., and Hook, M. (2014). Adhesion, invasion and evasion: the many functions of the surface proteins of *Staphylococcus aureus*. *Nat. Rev. Microbiol.* 12, 49–62. doi: 10.1038/nrmicro3161

Fredheim, E. G., Granslo, H. N., Flaegstad, T., Figenschau, Y., Rohde, H., Sadovskaya, I., et al. (2011). *Staphylococcus epidermidis* polysaccharide intercellular adhesin activates complement. *FEMS Immunol. Med. Microbiol.* 63, 269–280. doi: 10.1111/j.1574-695X.2011.00854.x

Gruszka, D. T., Whelan, F., Farrance, O. E., Fung, H. K., Paci, E., Jeffries, C. M., et al. (2015). Cooperative folding of intrinsically disordered domains drives assembly of a strong elongated protein. *Nat. Commun.* 6:7271. doi: 10.1038/ncomms8271

Gruszka, D. T., Wojdyla, J. A., Bingham, R. J., Turkenburg, J. P., Manfield, I. W., Steward, A., et al. (2012). Staphylococcal biofilm-forming protein has a contiguous rod-like structure. *Proc. Natl. Acad. Sci. U. S. A.* 109, E1011–E1018. doi: 10.1073/pnas.1119456109

Guo, B., Zhao, X., Shi, Y., Zhu, D., and Zhang, Y. (2007). Pathogenic implication of a fibrinogen-binding protein of *Staphylococcus epidermidis* in a rat model of intravascular-catheter-associated infection. *Infect. Immun.* 75, 2991–2995. doi: 10.1128/IAI.01741-06

Hartford, O., O'brien, L., Schofield, K., Wells, J., and Foster, T. J. (2001). The Fbe (SdrG) protein of *Staphylococcus epidermidis* HB promotes bacterial adherence to fibrinogen. *Microbiology* 147, 2545–2552. doi: 10.1099/00221287-147-9-2545

Heilmann, C., Gerke, C., Perdreau-Remington, F., and Gotz, F. (1996). Characterization of Tn917 insertion mutants of *Staphylococcus epidermidis* affected in biofilm formation. *Infect. Immun.* 64, 277–282. doi: 10.1128/IAI.64.1.277-282.1996

Heilmann, C., Hussain, M., Peters, G., and Gotz, F. (1997). Evidence for autolysin-mediated primary attachment of *Staphylococcus epidermidis* to a polystyrene surface. *Mol. Microbiol.* 24, 1013–1024. doi: 10.1046/j.1365-2958.1997.4101774.x

Heilmann, C., Thumm, G., Chhatwal, G. S., Hartleib, J., Uekotter, A., and Peters, G. (2003). Identification and characterization of a novel autolysin (Aae) with adhesive properties from *Staphylococcus epidermidis*. *Microbiology* 149, 2769–2778. doi: 10.1099/mic.0.26527-0

Hellmark, B., Soderquist, B., Unemo, M., and Nilsdotter-Augustinsson, A. (2013). Comparison of *Staphylococcus epidermidis* isolated from prosthetic joint infections and commensal isolates in regard to antibiotic susceptibility, agr type, biofilm production, and epidemiology. *Int. J. Med. Microbiol.* 303, 32–39. doi: 10.1016/j.ijmm.2012.11.001

Herman, P., El-Kirat-Chatel, S., Beaussart, A., Geoghegan, J. A., Foster, T. J., and Dufrene, Y. F. (2014). The binding force of the staphylococcal adhesin SdrG is remarkably strong. *Mol. Microbiol.* 93, 356–368. doi: 10.1111/mmi.12663

Herman-Bausier, P., and Dufrene, Y. F. (2016). Atomic force microscopy reveals a dual collagen-binding activity for the staphylococcal surface protein SdrF. *Mol. Microbiol.* 99, 611–621. doi: 10.1111/mmi.13254

Herman-Bausier, P., Valotteau, C., Pietrocola, G., Rindi, S., Alsteens, D., Foster, T. J., et al. (2016). Mechanical strength and inhibition of the *Staphylococcus aureus* collagen-binding protein Cna. *mBio* 7:e01529-16. doi: 10.1128/mBio.01529-16

Khodaparast, L., Khodaparast, L., Shahrooei, M., Stijlemans, B., Merckx, R., Baatsen, P., et al. (2016). The possible role of *Staphylococcus epidermidis* LPxTG surface protein SesC in biofilm formation. *PLoS One* 11:e0146704. doi: 10.1371/journal.pone.0146704

Kogan, G., Sadovskaya, I., Chaignon, P., Chokr, A., and Jabbouri, S. (2006). Biofilms of clinical strains of *Staphylococcus* that do not contain polysaccharide intercellular adhesin. *FEMS Microbiol. Lett.* 255, 11–16. doi: 10.1111/j.1574-6968.2005.00043.x

Kristian, S. A., Birkenstock, T. A., Sauder, U., Mack, D., Gotz, F., and Landmann, R. (2008). Biofilm formation induces C3a release and protects *Staphylococcus epidermidis* from IgG and complement deposition and from neutrophil-dependent killing. *J. Infect. Dis.* 197, 1028–1035. doi: 10.1086/528992

Le, K. Y., Park, M. D., and Otto, M. (2018). Immune evasion mechanisms of *Staphylococcus epidermidis* biofilm infection. *Front. Microbiol.* 9:359. doi: 10.3389/fmicb.2018.00359

Le, K. Y., Villaruz, A. E., Zheng, Y., He, L., Fisher, E. L., Nguyen, T. H., et al. (2019). Role of phenol-soluble modulins in *Staphylococcus epidermidis* biofilm formation and infection of indwelling medical devices. *J. Mol. Biol.* 431, 3015–3027. doi: 10.1016/j.jmb.2019.03.030

Lee, J. Y. H., Monk, I. R., Goncalves da Silva, A., Seemann, T., Chua, K. Y. L., Kearns, A., et al. (2018). Global spread of three multidrug-resistant lineages of *Staphylococcus epidermidis*. *Nat. Microbiol.* 3, 1175–1185. doi: 10.1038/s41564-018-0230-7

Liu, Q., Liu, Q., Meng, H., Lv, H., Liu, Y., Liu, J., et al. (2020). *Staphylococcus epidermidis* contributes to healthy maturation of the nasal microbiome by stimulating antimicrobial peptide production. *Cell Host Microbe* 27, 68.e5–78.e5. doi: 10.1016/j.chom.2019.11.003

Mack, D., Fischer, W., Krokotsch, A., Leopold, K., Hartmann, R., Egge, H., et al. (1996). The intercellular adhesin involved in biofilm accumulation of *Staphylococcus epidermidis* is a linear beta-1,6-linked glucosaminoglycan: purification and structural analysis. *J. Bacteriol.* 178, 175–183. doi: 10.1128/JB.178.1.175-183.1996

Meric, G., Mageiros, L., Pensar, J., Laabei, M., Yahara, K., Pascoe, B., et al. (2018). Disease-associated genotypes of the commensal skin bacterium *Staphylococcus epidermidis*. *Nat. Commun.* 9:5034. doi: 10.1038/s41467-018-07368-7

Milles, L. F., Schulten, K., Gaub, H. E., and Bernardi, R. C. (2018). Molecular mechanism of extreme mechanostability in a pathogen adhesin. *Science* 359, 1527–1533. doi: 10.1126/science.aar2094

Naik, S., Bouladoux, N., Linehan, J. L., Han, S. J., Harrison, O. J., Wilhelm, C., et al. (2015). Commensal-dendritic-cell interaction specifies a unique protective skin immune signature. *Nature* 520, 104–108. doi: 10.1038/nature14052

Naik, S., Bouladoux, N., Wilhelm, C., Molloy, M. J., Salcedo, R., Kastenmuller, W., et al. (2012). Compartmentalized control of skin immunity by resident commensals. *Science* 337, 1115–1119. doi: 10.1126/science.1225152

O'gara, J. P. (2007). *ica* and beyond: biofilm mechanisms and regulation in *Staphylococcus epidermidis* and *Staphylococcus aureus*. *FEMS Microbiol. Lett.* 270, 179–188. doi: 10.1111/j.1574-6968.2007.00688.x

Otto, M. (2018). Staphylococcal biofilms. *Microbiol. Spectr.* 6:GPP3-0023-2018. doi: 10.1128/microbiolspec.GPP3-0023-2018

Paharik, A. E., Kotasinska, M., Both, A., Hoang, T. N., Buttner, H., Roy, P., et al. (2017). The metalloprotease SepA governs processing of accumulation-associated protein and shapes intercellular adhesive surface properties in *Staphylococcus epidermidis*. *Mol. Microbiol.* 103, 860–874. doi: 10.1111/mmi.13594

Pei, L., and Flock, J. I. (2001). Lack of fbe, the gene for a fibrinogen-binding protein from *Staphylococcus epidermidis*, reduces its adherence to fibrinogen coated surfaces. *Microb. Pathog.* 31, 185–193. doi: 10.1006/mpat.2001.0462

Petrelli, D., Zampaloni, C., D'ercole, S., Prenna, M., Ballarini, P., Ripa, S., et al. (2006). Analysis of different genetic traits and their association with biofilm formation in *Staphylococcus epidermidis* isolates from central venous catheter infections. *Eur. J. Clin. Microbiol. Infect. Dis.* 25, 773–781. doi: 10.1007/s10096-006-0226-8

Pietrocola, G., Nobile, G., Gianotti, V., Zapotoczna, M., Foster, T. J., Geoghegan, J. A., et al. (2016). Molecular interactions of human plasminogen with fibronectin-binding protein B (FnBPB), a fibrinogen/fibronectin-binding protein from *Staphylococcus aureus*. *J. Biol. Chem.* 291, 18148–18162. doi: 10.1074/jbc.M116.731125

Ponnuraj, K., Bowden, M. G., Davis, S., Gurusiddappa, S., Moore, D., Choe, D., et al. (2003). A "dock, lock, and latch" structural model for a staphylococcal adhesin binding to fibrinogen. *Cell* 115, 217–228. doi: 10.1016/S0092-8674(03)00809-2

PrabhuDas, M., Adkins, B., Gans, H., King, C., Levy, O., Ramilo, O., et al. (2011). Challenges in infant immunity: implications for responses to infection and vaccines. *Nat. Immunol.* 12, 189–194. doi: 10.1038/ni0311-189

Qi, X., Jin, Y., Duan, J., Hao, Z., Wang, S., Guo, Y., et al. (2017). SesI may be associated with the invasiveness of *Staphylococcus epidermidis*. *Front. Microbiol.* 8:2574. doi: 10.3389/fmicb.2017.02574

Qin, L., Da, F., Fisher, E. L., Tan, D. C., Nguyen, T. H., Fu, C. L., et al. (2017). Toxin mediates sepsis caused by methicillin-resistant *Staphylococcus epidermidis*. *PLoS Pathog.* 13:e1006153. doi: 10.1371/journal.ppat.1006153

Qin, Z., Ou, Y., Yang, L., Zhu, Y., Tolker-Nielsen, T., Molin, S., et al. (2007). Role of autolysin-mediated DNA release in biofilm formation of *Staphylococcus epidermidis*. *Microbiology* 153, 2083–2092. doi: 10.1099/mic.0.2007/006031-0

Queck, S. Y., Jameson-Lee, M., Villaruz, A. E., Bach, T. H., Khan, B. A., Sturdevant, D. E., et al. (2008). RNAIII-independent target gene control by the agr quorum-sensing system: insight into the evolution of virulence regulation in *Staphylococcus aureus*. *Mol. Cell* 32, 150–158. doi: 10.1016/j.molcel.2008.08.005

Roche, F. M., Meehan, M., and Foster, T. J. (2003). The *Staphylococcus aureus* surface protein SasG and its homologues promote bacterial adherence to human desquamated nasal epithelial cells. *Microbiology* 149, 2759–2767. doi: 10.1099/mic.0.26412-0

Rohde, H., Burandt, E. C., Siemssen, N., Frommelt, L., Burdelski, C., Wurster, S., et al. (2007). Polysaccharide intercellular adhesin or protein factors in biofilm accumulation of *Staphylococcus epidermidis* and *Staphylococcus aureus* isolated from prosthetic hip and knee joint infections. *Biomaterials* 28, 1711–1720. doi: 10.1016/j.biomaterials.2006.11.046

Rohde, H., Burdelski, C., Bartscht, K., Hussain, M., Buck, F., Horstkotte, M. A., et al. (2005). Induction of *Staphylococcus epidermidis* biofilm formation via proteolytic processing of the accumulation-associated protein by staphylococcal and host proteases. *Mol. Microbiol.* 55, 1883–1895. doi: 10.1111/j.1365-2958.2005.04515.x

Rupp, M. E., Fey, P. D., Heilmann, C., and Gotz, F. (2001). Characterization of the importance of *Staphylococcus epidermidis* autolysin and polysaccharide intercellular adhesin in the pathogenesis of intravascular catheter-associated infection in a rat model. *J. Infect. Dis.* 183, 1038–1042. doi: 10.1086/319279

Sakinc, T., Kleine, B., and Gatermann, S. G. (2006). SdrI, a serine-aspartate repeat protein identified in *Staphylococcus saprophyticus* strain 7108, is a collagen-binding protein. *Infect. Immun.* 74, 4615–4623. doi: 10.1128/IAI.01885-05

Schaeffer, C. R., Woods, K. M., Longo, G. M., Kiedrowski, M. R., Paharik, A. E., Buttner, H., et al. (2015). Accumulation-associated protein enhances *Staphylococcus epidermidis* biofilm formation under dynamic conditions and is required for infection in a rat catheter model. *Infect. Immun.* 83, 214–226. doi: 10.1128/IAI.02177-14

Schwartz, K., Syed, A. K., Stephenson, R. E., Rickard, A. H., and Boles, B. R. (2012). Functional amyloids composed of phenol soluble modulins stabilize *Staphylococcus aureus* biofilms. *PLoS Pathog.* 8:e1002744. doi: 10.1371/journal.ppat.1002744

Shahrooei, M., Hira, V., Khodaparast, L., Khodaparast, L., Stijlemans, B., Kucharikova, S., et al. (2012). Vaccination with SesC decreases *Staphylococcus epidermidis* biofilm formation. *Infect. Immun.* 80, 3660–3668. doi: 10.1128/IAI.00104-12

Shahrooei, M., Hira, V., Stijlemans, B., Merckx, R., Hermans, P. W., and Van Eldere, J. (2009). Inhibition of *Staphylococcus epidermidis* biofilm formation by rabbit polyclonal antibodies against the SesC protein. *Infect. Immun.* 77, 3670–3678. doi: 10.1128/IAI.01464-08

Vaudaux, P. E., Francois, P., Proctor, R. A., Mcdevitt, D., Foster, T. J., Albrecht, R. M., et al. (1995). Use of adhesion-defective mutants of *Staphylococcus aureus* to define the role of specific plasma proteins in promoting bacterial adhesion to canine arteriovenous shunts. *Infect. Immun.* 63, 585–590. doi: 10.1128/IAI.63.2.585-590.1995

Vergara-Irigaray, M., Maira-Litran, T., Merino, N., Pier, G. B., Penades, J. R., and Lasa, I. (2008). Wall teichoic acids are dispensable for anchoring the PNAG exopolysaccharide to the *Staphylococcus aureus* cell surface. *Microbiology* 154, 865–877. doi: 10.1099/mic.0.2007/013292-0

Vuong, C., Kocianova, S., Voyich, J. M., Yao, Y., Fischer, E. R., Deleo, F. R., et al. (2004a). A crucial role for exopolysaccharide modification in bacterial biofilm formation, immune evasion, and virulence. *J. Biol. Chem.* 279, 54881–54886. doi: 10.1074/jbc.M411374200

Vuong, C., Voyich, J. M., Fischer, E. R., Braughton, K. R., Whitney, A. R., Deleo, F. R., et al. (2004b). Polysaccharide intercellular adhesin (PIA) protects *Staphylococcus epidermidis* against major components of the human innate immune system. *Cell. Microbiol.* 6, 269–275. doi: 10.1046/j.1462-5822.2004.00367.x

Wang, Y., Jiang, J., Gao, Y., Sun, Y., Dai, J., Wu, Y., et al. (2018). *Staphylococcus epidermidis* small basic protein (Sbp) forms amyloid fibrils, consistent with its function as a scaffolding protein in biofilms. *J. Biol. Chem.* 293, 14296–14311. doi: 10.1074/jbc.RA118.002448

Wang, R., Khan, B. A., Cheung, G. Y., Bach, T. H., Jameson-Lee, M., Kong, K. F., et al. (2011). *Staphylococcus epidermidis* surfactant peptides promote biofilm maturation and dissemination of biofilm-associated infection in mice. *J. Clin. Invest.* 121, 238–248. doi: 10.1172/JCI42520

Williams, R. J., Henderson, B., Sharp, L. J., and Nair, S. P. (2002). Identification of a fibronectin-binding protein from *Staphylococcus epidermidis*. *Infect. Immun.* 70, 6805–6810. doi: 10.1128/IAI.70.12.6805-6810.2002

Yamada, K. J., Heim, C. E., Xi, X., Attri, K. S., Wang, D., Zhang, W., et al. (2020). Monocyte metabolic reprogramming promotes pro-inflammatory activity and *Staphylococcus aureus* biofilm clearance. *PLoS Pathog.* 16:e1008354. doi: 10.1371/journal.ppat.1008354

Yarawsky, A. E., Johns, S. L., Schuck, P., and Herr, A. B. (2020). The biofilm adhesion protein Aap from *Staphylococcus epidermidis* forms zinc-dependent amyloid fibers. *J. Biol. Chem.* 295, 4411–4427. doi: 10.1074/jbc.RA119.010874

Zheng, Y., Joo, H. S., Nair, V., Le, K. Y., and Otto, M. (2018). Do amyloid structures formed by *Staphylococcus aureus* phenol-soluble modulins have a biological function? *Int. J. Med. Microbiol.* 308, 675–682. doi: 10.1016/j.ijmm.2017.08.010

Zhou, W., Spoto, M., Hardy, R., Guan, C., Fleming, E., Larson, P. J., et al. (2020). Host-specific evolutionary and transmission dynamics shape the functional diversification of *Staphylococcus epidermidis* in human skin. *Cell* 180, 454.e18–470.e18. doi: 10.1016/j.cell.2020.01.006

Zong, Y., Xu, Y., Liang, X., Keene, D. R., Hook, A., Gurusiddappa, S., et al. (2005). A 'Collagen Hug' model for *Staphylococcus aureus* CNA binding to collagen. *EMBO J.* 24, 4224–4236. doi: 10.1038/sj.emboj.7600888

Conflict of Interest: The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Microbiology

REVIEW
published: 26 August 2020
doi: 10.3389/fmicb.2020.02054

The Multivalent Role of Fibronectin-Binding Proteins A and B (FnBPA and FnBPB) of *Staphylococcus aureus* in Host Infections

Pietro Speziale and Giampiero Pietrocola*

Department of Molecular Medicine, Unit of Biochemistry, University of Pavia, Pavia, Italy

Staphylococcus aureus, one of the most important human pathogens, is the causative agent of several infectious diseases including sepsis, pneumonia, osteomyelitis, endocarditis and soft tissue infections. This pathogenicity is due to a multitude of virulence factors including several cell wall-anchored proteins (CWA). CWA proteins have modular structures with distinct domains binding different ligands. The majority of *S. aureus* strains express two CWA fibronectin (Fn)-binding adhesins FnBPA and FnBPB (Fn-binding proteins A and B), which are encoded by closely related genes. The N-terminus of FnBPA and FnBPB comprises an A domain which binds ligands such as fibrinogen, elastin and plasminogen. The A domain of FnBPB also interacts with histones and this binding results in the neutralization of the antimicrobial activity of these molecules. The C-terminal moiety of these adhesins comprises a long, intrinsically disordered domain composed of 11/10 fibronectin-binding repeats. These repetitive motifs of FnBPs promote invasion of cells that are not usually phagocytic *via* a mechanism by which they interact with integrin $\alpha_5\beta_1$ through a Fn mediated-bridge. The FnBPA and FnBPB A domains engage in homophilic cell-cell interactions and promote biofilm formation and enhance platelet aggregation. In this review we update the current understanding of the structure and functional properties of FnBPs and emphasize the role they may have in the staphylococcal infections.

Keywords: *Staphylococcus aureus*, FnBPA, FnBPB, adhesin, invasin, extracellular matrix protein, binding mechanism, virulence factor

OPEN ACCESS

Edited by:
Thomas Dandekar, Julius Maximilian University of Würzburg, Germany

Reviewed by:
Jack Christopher Leo, Nottingham Trent University, United Kingdom
Carolyn Brook Ibberson, Georgia Institute of Technology, United States

****Correspondence:***
Pietro Speziale
pspeziale@unipv.it

Specialty section:
This article was submitted to Infectious Diseases, a section of the journal Frontiers in Microbiology

Received: *12 June 2020*
Accepted: *05 August 2020*
Published: *26 August 2020*

Citation:
Speziale P and Pietrocola G (2020) The Multivalent Role of Fibronectin-Binding Proteins A and B (FnBPA and FnBPB) of Staphylococcus aureus in Host Infections. Front. Microbiol. 11:2054. doi: 10.3389/fmicb.2020.02054

INTRODUCTION

Staphylococcus aureus, a formidable pathogen that colonizes 30% of the population asymptomatically (Wertheim et al., 2005), is a major etiological agent of soft tissue infections such as cellulitis and superficial skin disease and a serious cause of abscesses, sepsis, pneumonia and endocarditis (Tong et al., 2015). In addition, to having increasing antibiotic resistance, the bacterium is a master in adapting to its host by avoiding almost every facet of the immune system (de Jong et al., 2019). *S. aureus* can express a broad multitude of virulence factors, most of which are surface proteins covalently anchored to wall peptidoglycan (CWA, cell wall-anchored

proteins). The most studied group of this family of proteins is represented by the microbial surface components recognizing adhesive matrix molecules (MSCRAMMs). All MSCRAMMs share a similar structural organization and are involved in binding specific host ligands (Foster et al., 2014). The fibronectin (Fn)-binding MSCRAMMs FnBPA and FnBPB (fibronectin-binding protein A and B) are among the most intensively studied proteins of *S. aureus*. Since the initial discovery of the Fn binding by *S. aureus* (Kuusela, 1978) and identification of FnBPs (Rydén et al., 1983; Jönsson et al., 1991), studies have been focussed on their biochemical characterization and the mechanism of binding to Fn (Schwarz-Linek et al., 2003; Meenan et al., 2007; Bingham et al., 2008) and fibrinogen (Fbg) (Wann et al., 2000; Stemberk et al., 2014). FnBPs were found to be involved in invasion of a variety of several non-phagocytic cell lines (Sinha et al., 1999; Liang et al., 2016; Prystopiuk et al., 2018). Furthermore, FnBPs were shown to bind to elastin (Roche et al., 2004), plasminogen (Plg) (Pietrocola et al., 2016) and histones (Pietrocola et al., 2019) and to have an important role in the formation of biofilm (O'Neill et al., 2008; Geoghegan et al., 2013; Speziale et al., 2014; Herman-Bausier et al., 2015). This multifactorial capacity to bind several ligands is closely related to the pathogenicity of the bacterium. In this review we focus on the structural and functional aspects of FnBPs.

REGULATION OF FnBPs EXPRESSION

FnBPs, as with many other MSCRAMMs, are predominantly present at a detectable and functional level on the surface of *S. aureus* cells during the exponential phase of growth. This suggests that the transcription of the *fnb* genes is prevalent in exponential phase and that FnBPA and FnBPB may be degraded by protease and be shed in the medium during stationary phase. However, during the stationary phase a sufficient amount of protein remains on the surface of the bacterial cells, as indicated by the ability of bacteria to retain binding to Fn. The *fnb* genes are subjected to control at the transcription level by Agr (accessory gene regulator) and Sar (Staphylococcal accessory regulator) global regulators (Cheung et al., 1992; Novick et al., 1993; Greene et al., 1995; Wolz et al., 2000). The Sar protein activates the synthesis of FnBPs during the exponential phase (Cheung et al., 1992; Giraudo et al., 1997), while Agr down-regulates transcription of genes for Fn-binding proteins post-exponentially (Saravia-Otten et al., 1997).

FnBPs STRUCTURE AND ORGANIZATION

Fn-binding proteins FnBPA and FnBPB are encoded by closely related genes *fnbA* and *fnbB* (Signäs et al., 1989; Jönsson et al., 1991). Analysis of a collection of 163 clinical isolates revealed that the majority of *S. aureus* strains contain both the genes and that some strains, notably those from CC30 and CC45 lineages, do not express FnBPB (for example, the epidemic health care-associated MRSA (HA-MRSA) strain 252) (Peacock et al., 2000; McCarthy and Lindsay, 2010). Both the proteins comprise an N-terminal signal sequence and a sorting signal including a Leu-Pro-X-Thr-Gly (where X stands for any amino acid) motif, a hydrophobic membrane region and a cytosolic tail, at the C-terminus. The most promiscuous ligand-binding region of FnBPA and FnBPB resides in the N-terminal A domain, which is composed of three folded IgG-like subdomains (from N1 to N3). The combination of N2 and N3 subdomains forms a hydrophobic groove that binds ligands. Located between the A domain and the wall spanning region is a long, unstructured Fn-binding repetitive region (FnBR) composed of 11 repeats in FnBPA and 10 in FnBPB.

The Fn-binding repeats bind individually to the N-terminal domain (NTD) of Fn and have considerable sequence variations, which in turn explains the differences in affinity of each repeat for the ligand (**Figure 1**).

LIGANDS OF FnBPA/FnBPB

FnBPA and FnBPB are good examples of promiscuous proteins that exhibit a great capability to interact with several structurally different host extracellular matrix/plasma proteins such as Fbg, Fn and elastin. Recently, these proteins have been shown to bind Plg/plasmin, a blood component that degrades many blood plasma proteins including fibrin clots, and histones.

Fibronectin

Fn is a large glycoprotein composed of two almost identical subunits. It occurs in a soluble form in plasma and other biological fluids and in a fibrillar or reticular multimeric form in the extracellular matrix (ECM), an organized network of secreted macromolecules immobilized in the extracellular space. Functionally, Fn is involved in attachment of cells to surfaces, assembly of ECM, angiogenesis and would repair. The transcript of Fn gene generates by alternative splicing several mRNAs, each encoding similar but not identical monomeric subunits (Henderson et al., 2011).

The subunits of the dimer of both plasma and cellular Fn are disulfide-bonded at the extreme C-terminal ends. The monomers are composed of three different types of modules named type I, II and III repeated units. Sets of adjacent modules in both the chains form specific domains, each involved in binding to various partners (proteins and carbohydrates).

The N-terminal domain (NTD) interacts with ligands such as fibrin, thrombospondin-1, and heparin and many bacterial adhesins, including FnBPA and FnBPB, and is the major site for Fn self-association in the matrix. The NTD contains five type I modules, each of which consists of a double- and a triple-stranded antiparallel β-sheets stacked on top of each other. Each module has four cys residues, which form disulfide bridges that stabilize a double loop structure (Potts et al., 1999; Rudiño-Piñera et al., 2007).

Downstream to this region is the gelatin/collagen-binding domain composed of four type I and two type II modules that also binds to tissue transglutaminase. The type II modules (approximately 45 residues each) consist of two double-stranded

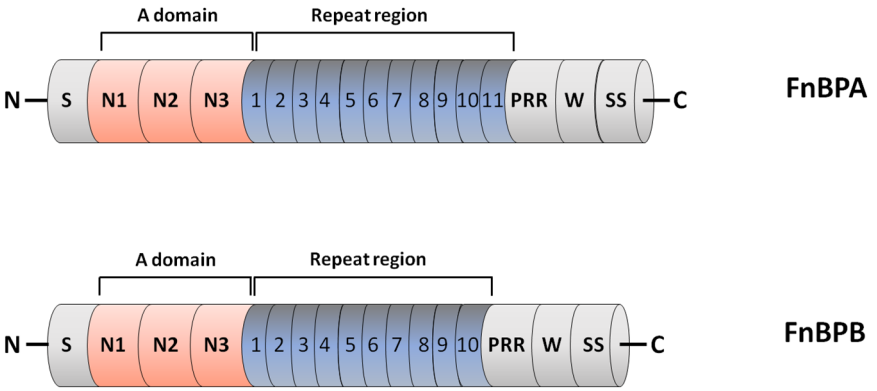


FIGURE 1 | Domain organization of fibronectin-binding protein A and B (FnBPA and FnBPB). The primary translation products of FnBPA and FnBPB proteins contain a signal sequence (S) at the N-terminus and a sorting signal (SS) at the C-terminus. The N-terminal A region of both proteins contains three separately folded subdomains, termed N1, N2 and N3. N2 and N3 form IgG-like folds that combined bind ligands such as fibrinogen by the DLL mechanism. Located distal to the A domain is an unstructured region consisting of tandemly arranged motifs (11 in FnBPA and 10 in FnBPB) that individually bind to the N-terminal domain of fibronectin. Also shown are the locations of proline rich region (PRR) and wall spanning region (W).

anti-parallel β-sheets connected by a loop and are stabilized by two intrachain dsulfide bridges (Sticht et al., 1998). All modules contribute to the optimal interaction with gelatin/collagen (Katagiri et al., 2003).

The large central domain comprises 15 type III repeats and includes sites interacting with the N-terminal assembly domain as well as binding sites for plasma membrane components. Each type III module contains approximately 90 aa residues and is organized in two β-sheets containing three and four strands, respectively (Main et al., 1992; Dickinson et al., 1994). The central domain of Fn, called cell-binding domain (CBD), contains an arginine-glycine-aspartate (RGD) motif in type III module number 10 ($FnIII_{10}$) that directly binds to the integrin $\alpha_5\beta_1$ and a synergy site ($FnIII_9$) that re-enforces binding of Fn to cell surface. Binding of $FnIII_{10}$ motif to the integrin connects the ECM to the cytoskeleton and generates signaling pathways. The region following the CBD includes a second high affinity heparin-binding domain and the C-terminal fibrin-binding domain (**Figure 2A**; Maurer et al., 2015).

Fibrinogen

Fbg is a large, soluble plasma glycoprotein synthesized by hepatocytes. Fbg consists of two monomers, each composed of three non-identical chains, designated Aα, Bβ and γ, linked together by disulfide bonds. Thrombin converts Fbg to fibrin monomer, releasing a 16 residue fibrinopeptide A (FPA) from the Aα chain and a 14 residue fibrinopetide B (FPB) from the Bβ chain. The uncovered N-terminal ends in the α- and β-chains of Fbg promote interactions with complementary sites in the C-terminal domains of neighbouring Fbg molecules, thus triggering the protofibril formation and soft clot. The soft clot is converted to hard clot by activated transglutaminase (FXIIIa), which promotes covalent crosslinks between the α and β chains (Weisel, 2005). The fibrin network is the major protein component of the hemostatic plug. Besides its primary role in clotting, Fbg binds *via* the RGD motifs contained in the α chains to the platelet integrin GPIIb/IIIa and is involved in platelet aggregation and thrombus formation (**Figure 2B**). Fbg is also deposited on implanted medical devices (Plow et al., 2000).

Plasminogen

The circulating mature form of Plg is a 92-kDa zymogen known as Glu-Plg that contains a glutamic acid at the N-terminus. Glu-Plg contains a hairpin-loop structure, the PAN domain, followed by five homologous kringle modules (K1–K5), each containing a typical pattern of four cysteine, and a proteinase catalytic domain. The kringle modules interact with fibrin clots and surface receptors from both eukaryotic and bacterial cells. Proteolytic activation of Plg to plasmin occurs through cleavage of the Arg^{561}–Val^{562} peptide bond in the catalytic domain by the tissue Plg activator (t-PA) and urokinase Plg activator. The cleavage results in the formation of plasmin, the active enzyme that contains the catalytic triad made up of the amino acid residues His^{603}, Asp^{646} and Ser^{741} (**Figure 2C**). Staphylokinase, a product of a gene located on lysogenic bacteriophages in *S. aureus*, can activate Plg when it forms a 1:1 stoichiometric complex with human plasmin (Sanderson-Smith et al., 2012; Verhamme et al., 2015).

Elastin

The protein elastin makes up the highly insoluble amorphous component of elastic fibers in the extracellular space of many tissues such as the lung, dermis and large blood vessels. Elastin is composed largely of glycine (33%), proline (10–13%) and other hydrophobic residues (over 40%). Individual polypeptide chains of soluble elastin, named tropoelastin, assemble into large fibers stabilized by multiple lysine-derived cross links (desmosines) formed by the copper-requiring extracellular enzyme lysil oxidase (Mecham, 2018; Kozel and Mecham, 2019). Tropoelastin contains an N-terminal domain that confers spring-like properties to the molecule

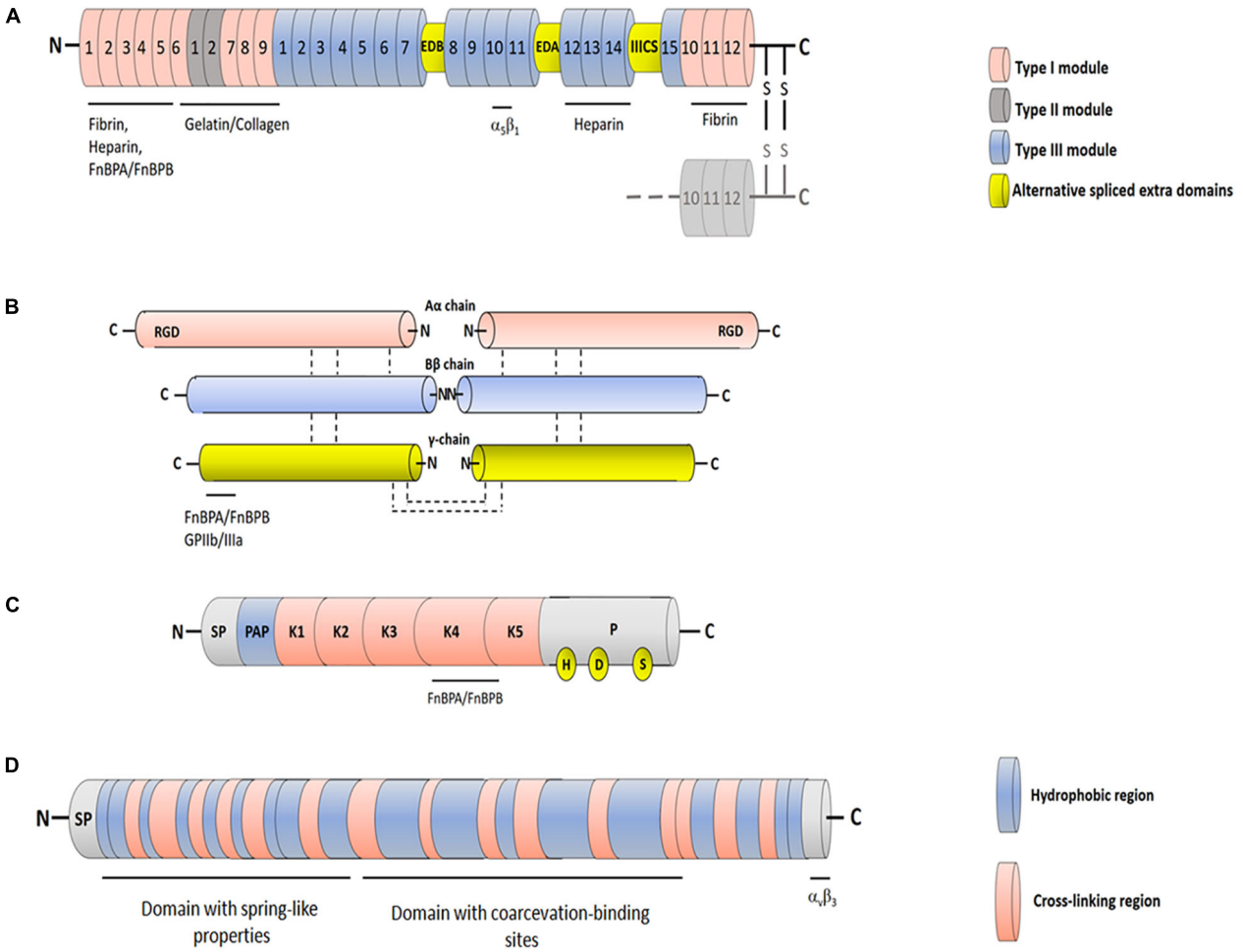


FIGURE 2 | Schematic representation of FnBPs ligands. **(A)** Fibronectin. Each Fn chain consists of three different types of modules with distinct folds: FnI, FnII and FnIII. Sets of adjacent modules form functional protease-stable domains, including the N-terminal domain, the gelatin-binding domain and the cell-binding domain. The cell-binding domain harbors the RGD motif within the 10th FnIII module and "synergy site" in the 9th FnIII module, both involved in interaction with integrin $\alpha_5\beta_1$. Cellular Fn includes two variable proportions of alternatively spliced FnIII modules EDB and EDA (extracellular domains B and A) and one connecting segment (IIICS). Binding sites for different binding partners are indicated. **(B)** Fibrinogen. Fbg is composed of three pairs of non-identical chains named Aα, Bβ and γ. The three chains are connected by disulfide bonds and are in hand to hand conformation forming a central E domain from which fibrinopetides A and B are released by the action of thrombin. The E domain is connected to the C-terminal D domains by coiled-coil region. The C-terminus of the γ chain includes the binding sites for integrin GPIIb/IIIa and FnBPA and FnFPB. **(C)** Plasminogen. Circulating Plg comprises at the N-terminal end a signal peptide (SP) and a preactivation region (PAP) followed by five consecutive disulphide-bonded kringle domains (K1-K5) and a serine protease domain (P). The K4 domain is supposed to be the binding site for FnBPs. The catalytic triad in the protease domain formed by residues His^{603}, Asp^{646} and Ser^{741} is also reported. **(D)** Tropoelastin. Tropoelastin is a protein comprising a signal peptide (SP) followed by alternating hydrophobic and cross-linking regions. The protein is functinctinally composed of an N-terminal domain that confers spring-like properties to the molecule and a C-terminal domain recognizing the $\alpha_v\beta_3$ integrin. The domains are connected by a bridge region that contains sites for the coarcevation process.

and a C-terminal region that mediates binding to $\alpha_v\beta_3$ integrin *via* a GRKRK motif (Bax et al., 2009; Baldock et al., 2011). The central domain contains contact sites for elastin coarcevation, a process observed in specific conditions of temperature, ionic strength and pH and where tropoelastin molecules associate to form elastic fibers (**Figure 2D**; Dyksterhuis et al., 2007).

Histones

Histones are basic proteins classified either as lysine-rich (H1, H2A, H2B) or arginine-rich (H3 and H4). In eukaryotes, histones H2A, H2B, H3 and H4 have a structural role in the formation of the nucleosome units, concurring to the assembly and compactation of chromatin. Together with DNA, histones are also the most abundant proteins in neutrophil extracellular traps (NETs), which are networks of extracellular fibers released by neutrophils to catch and kill infecting bacteria (Sollberger et al., 2018). Histones are also released into the bloodstream during severe sepsis (Xu et al., 2009). Histones have potent antimicrobial activity due to properties similar to cationic antimicrobial peptides. Histones H2A, H3 and H4 have been shown

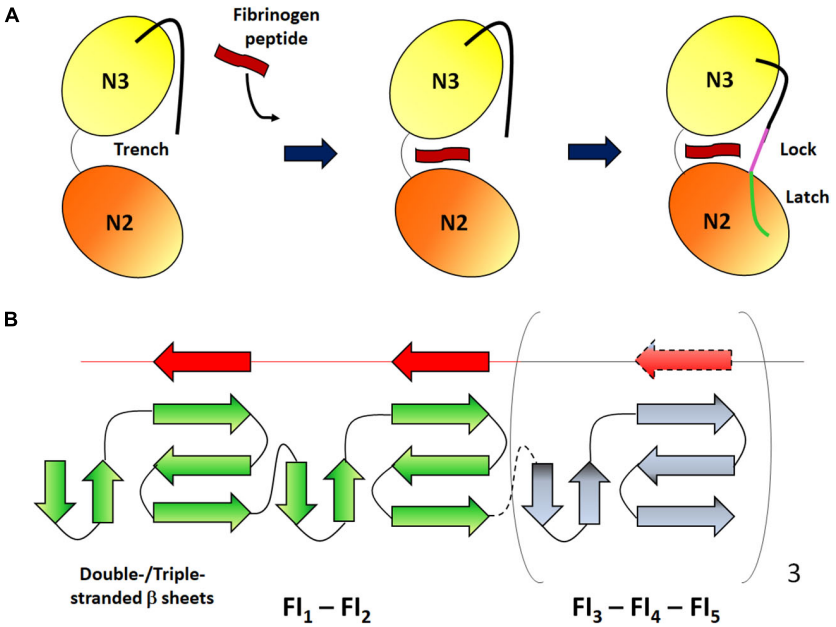


FIGURE 3 | Schematic representation of the binding mechanisms of Fbg and Fn. **(A)** Dock, lock and latch mechanism. The N2 and N3 subdomains of the MSCRAMM form a open trench where a Fbg peptide (red) can be inserted. After ligand binding, the C-terminal unstructured extension of the N3 subdomain undergoes conformational changes so that the segment covers the bound peptide and locks it in place (magenta) and binds to a β-strand in the N2 subdomain (green). Adapted from Geoghegan and Foster, 2017. **(B)** Tandem β-zipper mechanism. The Fn-binding repeats form a disordered region between the A domain and the C-terminal end of FnBPs. Each linear repeat (red) can complex with the type I modules (green) of the N-terminal domain forming additional short β-strands and extending the triple-stranded β-sheet on each module. Adapted from Matthews and Potts (2013).

to have anti-staphylococcal activity (Morita et al., 2013; Pietrocola et al., 2019).

MECHANISMS OF FnBPs INTERACTION WITH LIGANDS

Ligand Binding to A Domain of FnBPs

Sequence homology between FnBPA and clumping factor A (ClfA), a CWA Fbg-binding protein, the observation that the binding site in Fbg is the same (the C-terminal segment of the γ chain) and the finding that both ClfA and FnBPA bind to Fbg with similar affinity, all suggest a very similar mechanism of Fbg binding. Both the proteins interact with Fbg by the DLL (Dock, Lock and Latch) mechanism (Ganesh et al., 2008; Stemberk et al., 2014). In the case of ClfA, binding is initiated by docking of a peptide mimicking the C-terminus of the γ chain (HHLGGAKQAGDV) in the hydrophobic trench formed between the N2 and N3 subdomains. Next, the peptide-MSCRAMM complex induces a reorientation of the C-terminal end of N3 subdomain, so that this segment covers the bound peptide and locks it in place. In the final latching event the short segment inserts within the N2 subdomain and complements a β-sheet in this subdomain (**Figure 3A**; Deivanayagam et al., 2002). Notably, the A domain of ClfA also binds to a site located at the top of subdomain N3 and far away from the Fbg peptide binding trench (Ganesh et al., 2016). Some subtle differences exist between the interactions of FnBPA or ClfA with Fbg. It seems that binding of Fbg peptide to FnBPA does not require the involvement of latch strand residues and that peptide binding marginally affects the conformation of FnBPA. Moreover, it is not clear whether binding to Fbg by FnBPA involves a two-site mechanism as reported for ClfA (Stemberk et al., 2014).

Combining single molecule atomic force microscopy (AFM) and steered molecular dynamics simulations, Milles et al. (2018) revealed that the surface protein SdrG from *Staphylococcus epidermidis* binds to Fbg with a strength of 2 pN, similar to that of the covalent bonds. These authors also found that this mechanostability is related to an intricate hydrogen-bonding network between SdrG and a target Fbg peptide, which is independent of peptide side chains. A similar side-chain independent mechanical stability was observed with ClfB, another Fbg-binding adhesin from *S. aureus* (Milles et al., 2018). It remains to be determined whether this hydrogen bond mechanism may be extended to ClfA and FnBPs.

The A domain of FnBPA and FnBPB also binds Plg by a mechanism that does not involve DLL (Pietrocola et al., 2016). Differently from the Fbg binding, a single subdomain was required for Plg binding to FnBPs: subdomain N2 for FnBPA and subdomain N3 for FnBPB (Herman-Bausier et al., 2017). Notably, preincubation of FnBPB with Fbg increased by tenfold binding of Plg to adhesin expressed on the bacterial surface, suggesting that the conformational change of FnBPB occurring upon Fbg binding induces a better exposure of the Plg binding site and a higher interaction with the MSCRAMM (Herman-Bausier et al., 2017). Capture of Plg from serum and

its conversion to plasmin by host or bacterial activators promotes degradation of opsonins and facilitates spreading of bacteria in infected tissue (Pietrocola et al., 2016).

In addition, FnBPB binds histones, most likely by DLL, with an affinity which is 20-fold higher than for Fbg, indicating that FnBPB will bind preferentially histones when both proteins are incubated with the MSCRAMM. FnBPB could simultaneously bind H3 histone and Plg and activation of bound PLG resulted in cleavage of H3. Thus, FnBPB neutralizes the bactericidal activity of histones by direct capture of histones and prevention of their access to the surface. Alternatively, FnBPB can bind to Plg which, when activated to plasmin, can cleave the sequestered histone. Also FnBPB expression was shown to be responsible for protecting *S. aureus* from the bactericidal effect of the total histones released by neutrophils in the form of NETs (Pietrocola et al., 2019). Finally, the A domain of FnBPB can also bind to Fn by a mechanism that does not involve DLL (Burke et al., 2010).

Moreover, both the A domains of FnBPA and FnBPB promote adhesion of *S. aureus* to tropoelastin (Roche et al., 2004). The A domain of FnBPA binds to tropoelastin with high affinity (Keane et al., 2007a,b). Binding occurs at multiple sites, suggesting that these interactions involve repetitive sequence stretches of tropoelastin. The interaction is stabilized by negative charges on the A domain of FnBPA and lysine residues on tropoelastin (Keane et al., 2007a). The mechanism of elastin binding to FnBPs is unknown.

Binding of the Repetitive Region of FnBPs to Fn

The region that follows the Fbg-binding domain is composed of intrinsically random coil repeats. When bound to Fn, the Fn-binding repeat region acquires an ordered secondary structure. The stoichiometry of FnBP binding to Fn is estimated to be between six to nine Fn molecules per FnBP and at least six repeats bind to the NTD modules with high affinity (Fröman et al., 1987; Meenan et al., 2007). Each of the high affinity Fn-binding repeats binds to NTD by adding a short anti-parallel strand to β-sheets of sequential Fn type I modules. Thus, the intrinsically disordered Fn-binding repeats adopt a conformation determined by the binding to the NTD type I modules. Interactions between the FnBPs and Fn are also stabilized by hydrophobic and hydrophilic bonds. The tandem array of β-strands is called the β-zipper formation (**Figure 3B**; Schwarz-Linek et al., 2003).

AMINO ACID POLYMORPHISM IN FnBPs AND EFFECTS ON Fn AND Fbg BINDING

Seven different *fnbB* allelic variants were identified in both human and bovine isolates. The A domains of the corresponding seven isotypes of the FnBPB are 61–85% identical in sequence and each variant bound to ligands (Fbg, Fn and elastin) with similar affinity (Burke et al., 2010). Likewise, several isotypes of the A domain of FnBPA were identified which are between 66 and 76% identical in amino acid sequence in any pair-wise alignment. The isotypes retain their ligand-binding activity but differ antigenically and exhibit limited immune-cross reactivity (Loughman et al., 2008). Such polymorphisms in FnBPA were confirmed by a comparative study of *S. aureus* strains that originated from humans with infected and uninfected cardiac devices. These studies also found that amino acid substitutions in Fn-binding repeats 5 and 9 of FnBPA isolated from patients with infected cardiac implants confer a higher binding affinity for Fn, suggesting that strains with these substitutions could better impact colonization of damaged tissue or implanted devices (Casillas-Ituarte et al., 2012). Surprisingly, amino acid substitutions in the repeat region of FnBPA also affected binding of FnBPA to Fbg. These results were unexpected since these substitutions are quite distant from the trench binding site in the A domain of FnBPA. The authors suggested that the sequence variation of the C-terminal repetitive sequence might determine its fold back on the A domain so that binding of Fbg to the trench is enhanced (Casillas-Ituarte et al., 2019). Importantly, some substitutions in FnBPA significantly strengthened interactions with Fbg upon applications of a tensile force to the bond (Casillas-Ituarte et al., 2019). This observed strengthening of the Fbg-FnBPA interaction in the presence of a tensile force is reminiscent of catch-bond behavior, a type of non-covalent ligand-receptor interaction that is enhanced by application of a mechanical force in contrast to the traditional slip bonds that detach under increased force (Bartolo et al., 2002).

THE BIOLOGICAL ROLE OF FnBPA/FnBPB IN STAPHYLOCOCCAL INFECTIONS

Adhesion and Invasion

FnBPs contribute to the colonization and infection of the host by *S. aureus via* adhesion to Fn present in the extracellular matrix of the tissues. Following adhesion, *S. aureus* is able to invade a variety of non-professional phagocytic mammalian cells, such as epithelial cells (Dziewanowska et al., 1999), fibroblasts (Usui et al., 1992), endothelial cells (Peacock et al., 1999), osteoblasts (Ahmed et al., 2001) and keratinocytes (Mempel et al., 2002). The significance of bacterial invasion is unclear, but it could be involved in escape from immune defences of the host and protection from antibiotics. Fn-bridging between FnBPs and $\alpha_5\beta_1$ integrin on the host cell surface is sufficient to induce zipper-type uptake of staphylococci (Sinha et al., 1999). A study focussed on the FnBR-dependent role of endothelial cell invasion by *S. aureus* demonstrated that cell invasion can be facilitated by FnBRs with high affinity for Fn, while low affinity FnBRs expressed on the staphylococcal surface mediate adhesion but not the invasion. Furthermore, multiple FnBRs increased the efficiency of invasion without altering the uptake mechanism (Edwards et al., 2010). Small colony variants (SCVs), a subpopulation of *S. aureus* originating by mutations in metabolic genes and adapted to persist viable in the intracellular milieu (Sendi and Proctor, 2009), express high levels of FnBPs, a property which facilitates invasion of the host cells (Vaudaux et al., 2002). Recently, a model for invasion involving FnBPA, Fn and the $\alpha_5\beta_1$ integrin has been proposed. The

dimeric native Fn in solution predominantly exists as a compact globular conformer where the modules ^{9}FnIII-^{10}FnIII located in the central cell-binding domain are occluded and impeded by binding to $\alpha_5\beta_1$ integrin molecules on the surface of cells. Upon binding to the NTD of Fn, FnBPA repeats disrupt specific intramolecular contacts within Fn chains resulting in exposure of the cryptic $\alpha_5\beta_1$ binding sites and facilitation of Fn interaction with and clustering of integrins (Liang et al., 2016). Clustering of integrins promotes a recruitment of proteins such as vinculin, zyxin and tensin, as well as activation of focal adhesion kinases (FAKs) and Src to the site of bacterial attachment. The concerted activity of FAK and Src results in tyrosine phosphorylation of several effectors including cortactin, actin polymerization and the final bacterial invasion (Hauck and Ohlsen, 2006). This model of invasion for FnBPs is reminiscent of the afimbrial adhesin YadA from *Yersinia pseudotuberculosis* (Heise and Dersch, 2006) and the surface protein F1 from *Streptococcus pyogenes* which promote host cell invasion *via* Fn-bridging to integrin $\alpha_5\beta_1$ (Ozeri et al., 1998). Together these findings support the concept that FnBP-mediated invasion mechanism is a pretty common process among bacteria.

FnBPs in Platelet Activation/Aggregation and Thrombus Formation

Binding of *S. aureus* to platelets is a critical factor in the development of infective endocarditis. The process occurs in two stages. Firstly, staphylococci interact with quiescent platelets and promote intracellular signaling *via* the activation of the integrin GPIIb/IIIa. Secondly, activated platelets aggregate and contribute to the formation of thrombi and vegetations on the surface of heart valves. The FnBPs expressed on the surface of *S. aureus* cells from exponential phase of growth are the dominant platelet activating factors. In a study performed with *S. aureus* or *L. lactis* expressing either the Fbg-binding domain or the repetitive Fn-binding domain, Fitzgerald et al. (2006) demonstrated that each region promoted platelet activation, suggesting that FnBPA possesses two different mechanisms to activate resting platelets. Fbg and Fn present in the blood independently bind to the A domain and the repetitive region of FnBPA, respectively, and cross-link the bacterial surface to the low affinity GPIIb/IIIa and trigger activation and subsequent aggregation of the resting platelets. Furthermore, specific antibodies bound to FnBP domains act as additional, essential bridges between the FnBPs and the platelet immunoglobulin Fc receptor FcγRIIa (Fitzgerald et al., 2006). ClfA, another important Fbg-binding protein, expressed by stationary phase cells cooperates to activate and aggregate platelets in a fashion similar to FnBPs (Loughman et al., 2005). Other CWA proteins such as ClfB, SdrE and protein A have also been shown to have a role in the process (O'Brien et al., 2002).

FnBPs Engagement in Homophilic Interaction and Biofilm Formation

FnBPA and FnBPB expressed by clinically relevant HA-MRSA strains promote biofilm development (O'Neill et al., 2008; Vergara-Irigaray et al., 2009). FnBPs are thought to induce biofilm formation by a mechanism based on multiple, Zn^{2+}-dependent, hemophilic, low affinity bonds between the A domains of FnBPA or FnBPB located on adjacent cells. By using atomic force microscopy techniques it was also demonstrated that the mean adhesin force for FnBPA-FnBPA interactions is 190 pN, which is approximately 10 times lower than the binding that occurs between Fbg and its adhesin (2000 pN binding force). The weak force needed to separate neighbouring cells might help cell detachment during the dispersal phase that follows the biofilm formation (Herman-Bausier et al., 2015).

IMMUNOLOGICAL PROPERTIES OF FnBPA/FnBPB

In view of the important role played by FnBPs in pathogenesis it is crucial to define the immunological properties and the potential of these virulence factors in vaccine development. A study performed by Zuo et al. (2014) demonstrated that A domain had a strong immunogenicity and induced an increased survival of mice immunized with the antigen, while the C-terminal Fn-binding region had poor antigenic properties. Moreover, an amino acid segment located between the subdomain N1 and N2 showed immunogenicity and vaccine potential comparable to the full length A domain of FnBPA (Zuo et al., 2014). More specifically, a linear B cell-epitope (IETFNKANNRFSH) in the same subregion of FnBPA was identified and shown to evoke a protective immune response against *S. aureus* infection in immunized mice (Ma et al., 2018). Thus, region A of FnBPs seems to be a promising component for a staphylococcal vaccine.

Reports regarding the immunogenicity of the Fn-binding region of FnBPs are controversial. Immunization of rats with the repeat region of FnBPA induced protection against endocarditis (Rennermalm et al., 2001). Conversely, FnBRs weakly reacted with IgG isolated from sera of patients with staphylococcal endocarditis or region-specific monoclonal antibodies (mAbs) in the absence of Fn, but reacted more strongly in the presence of Fn. Thus, patients'antibodies and mAbs recognize Ligand-Induced Binding Site (LIBS) neo-epitopes formed upon FnBR binding to Fn. Notably, in the presence of Fn high affinity repeats exhibited a better reactivity than low affinity repeats with IgG from patients'sera, suggesting that during *S. aureus* infections patients develop preferentially LIBS antibodies to high affinity FnBRs (Rindi et al., 2006; Meenan et al., 2007). Importantly, LIBS antibodies did not show any antigen neutralizing activity. This observation should be taken into due consideration in the development of an anti-staphylococcal vaccine.

FnBP-LIKE PROTEINS FROM OTHER STAPHYLOCOCCAL SPECIES

Staphylococcus pseudintermedius, a staphylococcal species responsible of canine otitis, pyoderma and surgical wound infections, expresses two cell wall-anchored proteins, SpsD and SpsL, showing a organization and functionality similar

to FnBPA and FnBPB and that are likely to be important in tissue colonization and pathogenesis (Geoghegan et al., 2009; Bannoehr et al., 2012; Richards et al., 2018). SpsD contains a secretory signal sequence at the N-terminus and a C-terminal LPXTG motif required for anchoring the protein to the wall peptidoglycan. The N-terminal A domain, consists of three subdomains N1-N3 and is 40% identical to the A domain of FnBPB of *S. aureus*. Located distal to the A domain are a connecting region (C region) and a repeat region (R region). The N-terminal A domain of SpsD binds Fbg, Fn, elastin and cytokeratin 10. The binding site of the SpsD A region was mapped to residues 395–411 in the fibrinogen γ-chain. SpsD also bound to Gly- and Ser-rich omega loops within the C-terminal tail region of cytokeratin 10 (Pietrocola et al., 2013).

SpsL contains a signal sequence at the N-terminus followed by an A domain organized in three subdomains N1, N2, and N3 and a domain made up of multiple, tandemly identical repeats showing weak homology to the repetitive units of FnBPs of *S. aureus* (Pietrocola et al., 2015). Binding to Fbg occurs *via* the N2-N3 subdomains by a mechanism analogous to the DLL mechanism. There are multiple binding sites in the fibrinogen α-chain C-domain (Pickering et al., 2019). The connecting region of SpsD and the repetitive domain of SpsL represent the main binding sites of Fn binding and are involved in the cell invasion (Pietrocola et al., 2015). The use of LIBS antibodies raised against FnBPA has provided cues to identify in both these regions of SpsD and SpsL minimal binding units sequentially and structurally related to the repeated motifs of FnBPA of *S. aureus* (Viela et al., 2020). This observation can pave the way to get insights on the organization of these regions and for designing peptide analogs with inhibitory potential on MSCRAMM-Fn interaction. Recently, a surface protein from *S. delphini* called *S. delphini* surface protein Y (SdsY) has been described. SdsY shares 68% identity with SpsD from *S. pseudintermedius*, 36–39% identity with FnBPs and possesses the typical structure of Fn-binding proteins, including an N-terminal signal sequence, an A domain, a typical repeated region and a cell wall anchor motif. SdsY is involved in Fn-dependent internalization of bacteria into host osteoblast cell lines (Maali et al., 2020).

The *in silico* screening approach permitted identification of FnBP-like proteins in *S. argenteus* and *S. schweitzeri*, two species closely related to *S. aureus* (Weinman et al., 2020). Structural and functional studies on these proteins are lacking.

FnBPs AS VIRULENCE FACTORS

The role of FnBPs in virulence has been the object of several studies over the years. Since FnBPs are multifunctional, it is reasonable that these proteins play different roles in specific pathogenetic contexts. Importantly, the interpretation of these *in vivo* investigations should be taken with caution due to the number of factors (bacterial strain used, route of infection and infection model) on which the outcome of the infective process depends.

Role of FnBPs in Virulence of *S. aureus* Skin Abscess Infection

Using mouse models of skin abscess formation, it was shown that a *S. aureus* strain lacking all of its cell-wall anchored proteins was less virulent that its wild type strain. Strains specifically lacking FnBPA and FnBPB or other surface CWA protein such as protein A and ClfA also showed a significantly reduced virulence. Conversely, when a model of skin necrosis was studied, the *S. aureus* surface proteins could not be shown to be involved (Kwiecinski et al., 2014). *S. aureus* binds also to skin sections of human patients with atopic dermatitis (AD) (eczema) more efficiently than the biopsies taken from normal or psoriatic skins. In human atopic skin enhanced bacterial binding was primarily due to the increased level and availability of fibronectin present in the stratum corneum and in these conditions FnBPA and FnBPB expressed on the surface of *S. aureus* promoted better adhesion to AD sections compared to healthy ones (Cho et al., 2001).

Involvement of FnBPs in Infective Endocarditis

The role of *S. aureus* ClfA and FnBPA in the colonization of damaged valves in rat with experimental endocarditis was studied by Que et al. (2005) FnBPA-positive lactococci showed increasing bacterial concentrations in vegetations and spleens. The transformants also invaded the adjacent endothelium, possibly due to their capacity to trigger cell invasion. On the contrary, ClfA-positive lactococci were restricted to the vegetations. The authors also investigated the separate roles of the A and the repetitive domains of FnBPA and found that deletion of the A domain did not alter Fn binding and cell invasion, while it eliminated the colonization and infection of the valves *in vivo*. Furthermore, when lactococci expressing FnBPA deleted of the Fbg-binding domain were supplemented *in cis* or *in trans* with the Fbg-binding activity of ClfA acquired the A domain-associated functionality *in vitro* and *in vivo*. Summing up, the Fbg- and Fn- binding actions of the A and repetitive domains of FnBPA could synergically cooperate in the colonization of valves and invasion of endothelium (Que et al., 2005).

FnBPs and Their Role in Bacteremia and Sepsis

Bacteremia due to *S. aureus* can lead to the development of sepsis, a general inflammatory response to infection including fever, weakness, rapid heart rate and breathing rate. The response affects many internal organs such as the kidneys, heart, and lungs. Central venous catheters, surgically implanted materials, and orthopedic prostheses are the most important risk factor for bacteremia and invasive infection caused by *S. aureus* (Jensen et al., 1999) and these prosthetic devices help gain access to the bloodstream (Musher et al., 1994). To spread in the tissues and induce abscess lesions, *S. aureus* needs to leave the vasculature. FnBPA and FnBPB significantly contribute to host cell adherence *in vitro* and *in vivo* (Peacock et al., 1999; Kerdudou et al., 2006) and invasion of endothelial cells (Sinha et al., 1999). Moreover, *S. aureus* invasion of host cells and virulence in sepsis is facilitated by expression of multiple repeats in FnBPA

(Edwards et al., 2010). It is still unclear how staphylococci exit endothelial cells, cross the basement membrane and reach tissues. Possibly, toxins and secreted bacterial proteases play a role at this stage. Importantly, cooperation between FnBPA and FnBPB is indispensable for the induction of severe infection resulting in septic death (Shinji et al., 2011).

Staphylococcal Pneumonia and FnBPs

Until recently, pneumonia caused by *S. aureus* has been regarded as a secondary effect of viral respiratory infections. Now, staphylococcal pneumonia has emerged as an important clinical problem. Several virulence factors have gained attention as potential causative virulence agents of the disease, among others Panton Valentin leukocidin (PVL) (Labandeira-Rey et al., 2007) and the collagen-binding adhesin Cna (de Bentzmann et al., 2004; Otsuka et al., 2006). Using a humanized tracheal xenograft model in the nude mouse and cultures of HAEC (human airway epithelial cells) it has been shown that FnBPs have an important role in the colonization of the airways by *S. aureus* and a potential contributory role in the onset of staphylococcal pneumonia (Mongodin et al., 2002).

Foreign Body Infections and FnBPs

S. aureus is one of the most important causes of implant-associated infections (Arciola et al., 2018). The success of the pathogen in this context depends on rapid and efficient adhesion to virtually all biomaterial surfaces and biofilm development. Staphylococcal attachment to and colonization of the biomedical implants is largely due to the arsenal of surface adhesins expressed by this bacterium and to the rapid coating of the implants with proteins from blood and interstitial fluids, mostly Fbg and Fn (Speziale et al., 2009; Franz et al., 2011). Thus, one should expect that *S. aureus* strains, the majority of which are equipped with FnBPs, will find ideal conditions for adhesion to Fn- and Fbg-coated biomedical implants. Expression of FnBPA and FnBPB represents an additional critical element for the formation of a proteinaceous biofilm, persistence of implant infections and bacterial dissemination to other body sites.

The prompt integration of biomaterials into resident cells in nearby tissues is crucial for the success of implantation of many medical devices (Gristina, 1987) and for preventing bacterial adhesion to abiotic surfaces or biotically coated surfaces (Stones and Krachler, 2016). However, considering the ability of *S. aureus* to invade endothelial cells and osteoblasts, even tissue-integrated biomaterials could become the target for the staphylococcal infection. Thus, biofilm formation and ability to invade host cells, both mediated by FnBPs shelter *S. aureus* cells and encourage persistence of infection in medical devices.

CONCLUSION

In this review we summarize the most recent findings on the structure and function of FnBPs from *S. aureus*.

If structural biology has revealed the mechanisms of FnBPA and FnBPB interacting with Fbg and Fn (DLL and β-zipper mechanisms), the definition of the X-ray crystal structure of FnBPs in complex with the new ligands such as Plg and histones and the binding mechanisms remain to be revealed. Moreover, a more detailed structural analysis of homophilic interactions will be beneficial for understanding the biochemical basis of biofilm formation promoted by FnBPs. These studies would help in designing inhibitors of the various processes mediated by these MSCRAMMs and in controlling infection. Furthermore, it will be of biological importance to test the role of Plg binding to FnBPA and FnBPB in infection models with mice.

However, one should be cautious when studies with murine infection models are performed. In fact, differently from human Plg, the mouse Plg reacts poorly with staphylokinase (Peetermans et al., 2014). Moreover, it should be experimentally established whether FnBPs recognize murine Plg efficiently. Thus, for a full investigation of the role of Plg binding and activation in infection it is worth exploring transgenic mice expressing human Plg.

FnBPA and FnBPB share several structural and functional similarities. Hence, an important question is why *S. aureus* cells express two FnBPs with apparently similar properties. It is possible that expression of each adhesin is differentially regulated and dependent on external factors such as growth conditions. Indeed, a closer examination of these proteins reveals subtle differences both in the structure and function. For example, both proteins bind to Plg *via* different subdomains of region A, but, more intriguingly, only FnBPB protein interact with histones and prevents them from damaging the bacterial membrane. The discovery of new ligands and functions for FnBPs also deserves attention. Such discoveries must be followed by a detailed biochemical and physical investigation and rigorous analysis *in vivo* to prove the biological significance of bacterial interactions with new ligands. Another issue regards the immunological properties of FnBPs. The repetitive region of FnBPs does not appear to be promising for vaccine development, while it remains to be determined whether A region of FnBPs can be used as vaccine. In fact, if some observations indicate that FnBPA antibodies are protective in animal models of infection (Zuo et al., 2014), on the other side, the A region shows a significant polymorphism and antibodies raised against one isoform react weakly with the others (Loughman et al., 2008; Burke et al., 2010). Several CWA proteins including FnBPs are released in the growth medium by the action of endogenous proteases (McGavin et al., 1997) or by mutational events (Grundmeier et al., 2004) and possibly they are involved in important biological functions, for example, modulation of MSCRAMM binding to ligands and interference with events such as cell invasion, biofilm formation and immune evasion (O'Halloran et al., 2015). Therefore, it is worth exploring the role of soluble FnBPs in the regulation of the multivalent activity of these proteins. Finally, since FnBPs are multifactorial and involved in different pathologies, a more clear definition of the role played by these proteins in specific infective processes is key for the development of novel

therapeutic approaches to control the bacterium in diverse disease settings.

AUTHOR CONTRIBUTIONS

PS conceptualized and coordinated the work. PS and GP contributed to this article in terms of data collection and writing and performed critical revision for the work. Both authors contributed to the article and approved the submitted version.

ACKNOWLEDGMENTS

We wish to thank Prof. Timothy J. Foster for fruitful discussion and advice and Angelica Pellegrini for help with computer art.

REFERENCES

Ahmed, S., Meghji, S., Williams, R. J., Henderson, B., Brock, J. H., and Nair, S. P. (2001). *Staphylococcus aureus* fibronectin binding proteins are essential for internalization by osteoblasts but do not account for differences in intracellular levels of bacteria. *Infect. Immun.* 69, 2872–2877. doi: 10.1128/IAI.69.5.2872-2877.2001

Arciola, C. R., Campoccia, D., and Montanaro, L. (2018). Implant infections: adhesion, biofilm formation and immune evasion. *Nat. Rev. Microbiol.* 16, 397–409. doi: 10.1038/s41579-018-0019-y

Baldock, C., Oberhauser, A. F., Ma, L., Lammie, D., Siegler, V., Mithieux, S. M., et al. (2011). Shape of tropoelastin, the highly extensible protein that controls human tissue elasticity. *Proc. Natl. Acad. Sci. U.S.A.* 108, 4322–4327. doi: 10.1073/pnas.1014280108

Bannoehr, J., Brown, J. K., Shaw, D. J., Fitzgerald, R. J., van den Broek, A. H., and Thoday, K. L. (2012). *Staphylococccus pseudintermedius* surface proteins SpsD and SpsO mediate adherence to ex vivo canine corneocytes. *Vet. Dermatol.* 23, 119–124. doi: 10.1111/j.1365-3164.2011.01021.x

Bartolo, D., Derényi, I., and Ajdari, A. (2002). Dynamic response of adhesion complexes: beyond the single-path picture. *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.* 65(5 Pt 1), 051910. doi: 10.1103/PhysRevE.65.051910

Bax, D. V., Rodgers, U. R., Bilek, M. M. M., and Weiss, A. S. (2009). Cell adhesion to tropoelastin is mediated via the C-terminal GRKRK motif and integrin alphaVbeta3. *J. Biol. Chem.* 284, 28616–28623. doi: 10.1074/jbc.M109.017525

Bingham, R. J., Rudiño-Piñera, E., Meenan, N. A. G., Schwarz-Linek, U., Turkenburg, J. P., Höök, M., et al. (2008). Crystal structures of fibronectin-binding sites from *Staphylococcus aureus* Fnbpa in complex with fibronectin domains. *Proc. Natl. Acad. Sci. U.S.A.* 105, 12254–12258. doi: 10.1073/pnas.0803556105

Burke, F. M., McCormack, N., Rindi, S., Speziale, P., and Foster, T. J. (2010). Fibronectin-binding protein B variation in *Staphylococcus aureus*. *BMC Microbiol.* 10:160. doi: 10.1186/1471-2180-10-160

Casillas-Ituarte, N. N., DiBartola, A. C., Broughton, M. J., Pérez-Guzmán, L., Wheeler, R. M., Ibaraki, M., et al. (2019). Fibrinogen binding is affected by amino acid substitutions in C-terminal repeat region of fibronectin binding protein A. *Sci. Rep.* 9:11619. doi: 10.1038/s41598-019-48031-5

Casillas-Ituarte, N. N., Lower, B. H., Lamlertthon, S., Fowler, V. G. Jr., and Lower, S. K. (2012). Dissociation rate constants of human fibronectin binding to fibronectin-binding proteins on living *Staphylococcus aureus* isolated from clinical patients. *J. Biol. Chem.* 287, 6693–6701. doi: 10.1074/jbc.M111.285692

Cheung, A. L., Koomey, J. M., Butler, C. A., Projan, S. J., and Fischetti, V. A. (1992). Regulation of exoprotein expression in *Staphylococcus aureus* by a Locus (Sar) distinct from agr. *Proc. Natl. Acad. Sci. U.S.A.* 89, 6462–6466. doi: 10.1073/pnas.89.14.6462

Cho, S. H., Strickland, I., Boguniewicz, M., and Leung, D. Y. (2001). Fibronectin and fibrinogen contribute to the enhanced binding of *Staphylococcus aureus* to atopic skin. *J. Allergy Clin. Immunol.* 108, 269–274. doi: 10.1067/mai.2001.117455

de Bentzmann, S., Tristan, A., Etienne, J., Brousse, N., Vandenesch, F., and Lina, G. (2004). *Staphylococcus aureus* isolates associated with necrotizing pneumonia bind to basement membrane type I and IV collagens and laminin. *J. Infect. Dis.* 190, 1506–1515. doi: 10.1086/424521

de Jong, N. W. M., van Kessel, K. P. M., and van Strijp, J. A. G. (2019). Immune evasion by *Staphylococcus aureus*. *Microbiol. Spectr.* 7:GPP3-0061-2019. doi: 10.1128/microbiolspec.GPP3-0061-2019

Deivanayagam, C. C. S., Wann, E. R., Chen, W., Carson, M., Rajashankar, K. R., Höök, M., et al. (2002). A novel variant of the immunoglobulin fold in surface adhesins of *Staphylococcus aureus*: crystal structure of the fibrinogen-binding MSCRAMM, clumping factor A. *EMBO J.* 21, 6660–6672. doi: 10.1093/emboj/cdf619

Dickinson, C. D., Veerapandian, B., Dai, X. P., Hamlin, R. C., Xuong, N. H., Ruoslahti, E., et al. (1994). Crystal structure of the tenth type III cell adhesion module of human fibronectin. *J Mol Biol.* 236, 1079–1092. doi: 10.1016/0022-2836(94)90013-2

Dyksterhuis, L. B., Baldock, C., Lammie, D., Wess, T. J., and Weiss, A. S. (2007). Domains 17-27 of tropoelastin contain key regions of contact for coacervation and contain an unusual turn-containing crosslinking domain. *Matrix Biol.* 26, 125–135. doi: 10.1016/j.matbio.2006.10.002

Dziewanowska, K., Patti, J. M., Deobald, C. F., Bayles, K. W., Trumble, W. R., and Bohach, G. A. (1999). Fibronectin binding protein and host cell tyrosine kinase are required for internalization of *Staphylococcus aureus* by epithelial cells. *Infect. Immun.* 67, 4673–4678. doi: 10.1128/iai.67.9.4673-4678.1999

Edwards, A. M., Potts, J. R., Josefsson, E., and Massey, R. C. (2010). *Staphylococcus aureus* host cell invasion and virulence in sepsis is facilitated by the multiple repeats within FnBPA. *PLoS Pathog.* 6:e1000964. doi: 10.1371/journal.ppat.1000964

Fitzgerald, J. R., Loughman, A., Keane, F., Brennan, M., Knobel, M., Higgins, J., et al. (2006). Fibronectin-binding proteins of *Staphylococcus aureus* mediate activation of human platelets via fibrinogen and fibronectin bridges to integrin GPIIb/IIIa and IgG binding to the FcgammaRIIa receptor. *Mol. Microbiol.* 59, 212–230. doi: 10.1111/j.1365-2958.2005.04922.x

Foster, T. J., Geoghegan, J. A., Ganesh, V. K., and Höök, M. (2014). Adhesion, invasion and evasion: the many functions of the surface proteins of *Staphylococcus aureus*. *Nat. Rev. Microbiol.* 12, 49–62. doi: 10.1038/nrmicro3161

Franz, S., Rammelt, S., Scharnweber, D., and Simon, J. C. (2011). Immune responses to implants - a review of the implications for the design of immunomodulatory biomaterials. *Biomaterials* 32, 6692–6709. doi: 10.1016/j.biomaterials.2011.05.078

Fröman, G., Switalski, L. M., Speziale, P., and Höök, M. (1987). Isolation and characterization of a fibronectin receptor from *Staphylococcus aureus*. *J. Biol. Chem.* 262, 6564–6571.

Ganesh, V. K., Liang, X., Geoghegan, J. A., Cohen, A. L. V., Venugopalan, N., Foster, T. J., et al. (2016). Lessons from the crystal structure of the *S. aureus* surface protein clumping factor a in complex with tefibazumab, an inhibiting monoclonal antibody. *EBio Medicine* 13, 328–338. doi: 10.1016/j.ebiom.2016.09.027

Ganesh, V. K., Rivera, J. J., Smeds, E., Ko, Y., Bowden, M. G., Wann, E. R., et al. (2008). A structural model of the *Staphylococcus aureus* ClfA-fibrinogen interaction opens new avenues for the design of anti-staphylococcal therapeutics. *PLoS Pathog.* 4:e1000226. doi: 10.1371/journal.ppat.1000226

Geoghegan, J. A., and Foster, T. J. (2017). Cell wall-anchored surface proteins of *Staphylococcus aureus*: many proteins, multiple functions. *Curr. Top. Microbiol. Immunol.* 409, 95–120. doi: 10.1007/82_2015_5002

Geoghegan, J. A., Monk, I. R., O'Gara, J. P., and Foster, T. J. (2013). Subdomains N2N3 of fibronectin binding protein a mediate *Staphylococcus aureus* biofilm formation and adherence to fibrinogen using distinct mechanisms. *J. Bacteriol.* 195, 2675–2683. doi: 10.1128/JB.02128-12

Geoghegan, J. A., Smith, E. J., Speziale, P., and Foster, T. J. (2009). *Staphylococcus pseudintermedius* expresses surface proteins that closely resemble those from *Staphylococcus aureus*. *Vet. Microbiol.* 138, 345–352. doi: 10.1016/j.vetmic.2009.03.030

Giraudo, A. T., Cheung, A. L., and Nagel, R. (1997). The sae locus of *Staphylococcus aureus* controls exoprotein synthesis at the transcriptional level. *Arch. Microbiol.* 168, 53–58. doi: 10.1007/s002030050469

Greene, C., McDevitt, D., Francois, P., Vaudaux, P. E., Lew, D. P., and Foster, T. J. (1995). Adhesion properties of mutants of *Staphylococcus aureus* defective in fibronectin-binding proteins and studies on the expression of fnb genes. *Mol. Microbiol.* 17, 1143–1152. doi: 10.1111/j.13652958.1995.mmi_17061143.x

Gristina, A. G. (1987). Biomaterial-centered infection: microbial adhesion versus tissue integration. *Science* 237, 1588–1595. doi: 10.1126/science.3629258

Grundmeier, M., Hussain, M., Becker, P., Heilmann, C., Peters, G., and Sinha, B. (2004). Truncation of fibronectin-binding proteins in *Staphylococcus aureus* strain Newman leads to deficient adherence and host cell invasion due to loss of the cell wall anchor function. *Infect. Immun.* 72, 7155–7163. doi: 10.1128/IAI.72.12.7155-7163.2004

Hauck, C. R., and Ohlsen, K. (2006). Sticky connections: extracellular matrix protein recognition and integrin-mediated cellular invasion by *Staphylococcus aureus*. *Curr. Opin. Microbiol.* 9, 5–11. doi: 10.1016/j.mib.2005.12.002

Heise, T., and Dersch, P. (2006). Identification of a domain in *Yersinia virulence* factor YadA that is crucial for extracellular matrix-specific cell adhesion and uptake. *Proc. Natl. Acad. Sci. U.S.A.* 103, 3375–3380. doi: 10.1073/pnas.0507749103

Henderson, B., Nair, S., Pallas, J., and Williams, M. A. (2011). Fibronectin: a multidomain host adhesin targeted by bacterial fibronectin-binding proteins. *FEMS Microbiol. Rev.* 35, 147–200. doi: 10.1111/j.1574-6976.2010.00243.x

Herman-Bausier, P., El-Kirat-Chatel, S., Foster, T. J., Geoghegan, J. A., and Dufrêne, Y. F. (2015). *Staphylococcus aureus* fibronectin-binding protein a mediates cell-cell adhesion through low-affinity homophilic bonds. *mBio* 6, e413–e415. doi: 10.1128/mBio

Herman-Bausier, P., Pietrocola, G., Foster, T. J., Speziale, P., and Dufrêne, Y. F. (2017). Fibrinogen activates the capture of human plasminogen by staphylococcal fibronectin-binding proteins. *mBio* 8:e01067-17. doi: 10.1128/mBio.01067-17

Jensen, A. G., Wachmann, C. H., Poulsen, K. B., Espersen, F., Scheibel, J., Skinhøj, P., et al. (1999). Risk factors for hospital-acquired *Staphylococcus aureus* bacteremia. *Arch. Intern. Med.* 159, 1437–1444. doi: 10.1001/archinte.159.13.1437

Jönsson, K., Signäs, C., Müller, H. P., and Lindberg, M. (1991). Two different genes encode fibronectin binding proteins in *Staphylococcus aureus*. The complete nucleotide sequence and characterization of the second gene. *Eur. J. Biochem.* 202, 1041–1048. doi: 10.1111/j.1432-1033.1991.tb16468.x

Katagiri, Y., Brew, S. A., and Ingham, K. C. (2003). All six modules of the gelatin-binding domain of fibronectin are required for full affinity. *J. Biol. Chem.* 278, 11897–11902. doi: 10.1074/jbc.M212512200

Keane, F. M., Clarke, A. W., Foster, T. J., and Weiss, A. S. (2007a). The N-terminal A domain of *Staphylococcus aureus* fibronectin-binding protein A binds to tropoelastin. *Biochemistry* 46, 7226–7232. doi: 10.1021/bi700454x

Keane, F. M., Loughman, A., Valtulina, V., Brennan, M., Speziale, P., and Foster, T. J. (2007b). Fibrinogen and elastin bind to the same region within the a domain of fibronectin binding protein A, an MSCRAMM of *Staphylococcus aureus*. *Mol. Microbiol.* 63, 711–723. doi: 10.1111/j.1365-2958.2006.05552.x

Kerdudou, S., Laschke, M. W., Sinha, B., Preissner, K. T., Menger, M. D., and Herrmann, M. (2006). Fibronectin binding proteins contribute to the adherence of *Staphylococcus aureus* to intact endothelium in vivo. *Thromb. Haemost.* 96, 183–189. doi: 10.1160/th06-02-0116

Kozel, B. A., and Mecham, R. P. (2019). Elastic fiber ultrastructure and assembly. *Matrix Biol.* 84, 31–40. doi: 10.1016/j.matbio.2019.10.002

Kuusela, P. (1978). Fibronectin binds to *Staphylococcus aureus*. *Nature* 276, 718–720. doi: 10.1038/276718a0

Kwiecinski, J., Jin, T., and Josefsson, E. (2014). Surface proteins of *Staphylococcus aureus* play an important role in experimental skin infection. *APMIS* 122, 1240–1250. doi: 10.1111/apm.12295

Labandeira-Rey, M., Couzon, F., Boisset, S., Brown, E. L., Bes, M., Benito, Y., et al. (2007). *Staphylococcus aureus* Panton-Valentine leukocidin causes necrotizing pneumonia. *Science* 315, 1130–1133. doi: 10.1126/science.1137165

Liang, X., Garcia, B. L., Visai, L., Prabhakaran, S., Meenan, N. A. G., Potts, J. R., et al. (2016). Allosteric regulation of fibronectin/α5β1 interaction by fibronectin-binding MSCRAMMs. *PLoS One* 11:e0159118. doi: 10.1371/journal.pone.0159118

Loughman, A., Fitzgerald, J. R., Brennan, M. P., Higgins, J., Downer, R., Cox, D., et al. (2005). Roles for fibrinogen, immunoglobulin and complement in platelet activation promoted by *Staphylococcus aureus* clumping factor A. *Mol. Microbiol.* 57, 804–818. doi: 10.1111/j.1365-2958.2005.04731.x

Loughman, A., Sweeney, T., Keane, F. M., Pietrocola, G., Speziale, P., and Foster, T. J. (2008). Sequence diversity in the A domain of *Staphylococcus aureus* fibronectin-binding protein A. *BMC Microbiol.* 8:74. doi: 10.1186/1471-2180-8-74

Ma, J., Wei, Y., Zhang, L., Wang, X., Yao, D., Liu, D., et al. (2018). Identification of a novel linear B-cell epitope as a vaccine candidate in the N2N3 subdomain of *Staphylococcus aureus* fibronectin-binding protein A. *J. Med. Microbiol.* 67, 423–431. doi: 10.1099/jmm.0.000633

Maali, Y., Diot, A., Martins-Simões, P., Bes, M., Bouvard, D., Vandenesch, F., et al. (2020). Identification and characterization of *Staphylococcus delphini* internalization pathway in nonprofessional phagocytic cells. *Infect. Immun.* 88:e00002-20. doi: 10.1128/IAI.00002-20

Main, A. L., Harvey, T. S., Baron, M., Boyd, J., and Campbell, I. D. (1992). The three-dimensional structure of the tenth type III module of fibronectin: an insight into RGD-mediated interactions. *Cell* 71, 671–678. doi: 10.1016/0092-8674(92)90600-h

Matthews, J. M., and Potts, J. R. (2013). The tandem β-zipper: modular binding of tandem domains and linear motifs. *FEBS Lett.* 587, 1164–1171. doi: 10.1016/j.febslet.2013.01.002

Maurer, L. M., Ma, W., and Mosher, D. F. (2015). Dynamic structure of plasma fibronectin. *Crit. Rev. Biochem. Mol. Biol.* 51, 213–227. doi: 10.1080/10409238.2016.1184224

McCarthy, A. J., and Lindsay, J. A. (2010). Genetic variation in *Staphylococcus aureus* surface and immune evasion genes is lineage associated: implications for vaccine design and host-pathogen interactions. *BMC Microbiol.* 10:173. doi: 10.1186/1471-2180-10-173

McGavin, M. J., Zahradka, C., Rice, K., and Scott, J. E. (1997). Modification of the *Staphylococcus aureus* fibronectin binding phenotype by V8 protease. *Infect. Immun.* 65, 2621–2628. doi: 10.1128/iai.65.7.2621-2628.1997

Mecham, R. P. (2018). Elastin in lung development and disease pathogenesis. *Matrix Biol.* 73, 6–20. doi: 10.1016/j.matbio.2018.01.005

Meenan, N. A., Visai, L., Valtulina, V., Schwarz-Linek, U., Norris, N. C., Gurusiddappa, S., et al. (2007). The tandem beta-zipper model defines high affinity fibronectin-binding repeats within *Staphylococcus aureus* FnBPA. *J. Biol. Chem.* 282, 25893–25902. doi: 10.1074/jbc.M703063200

Mempel, M., Schnopp, C., Hojka, M., Fesq, H., Weidinger, S., Schaller, M., et al. (2002). Invasion of human keratinocytes by *Staphylococcus aureus* and intracellular bacterial persistence represent haemolysin-independent virulence mechanisms that are followed by features of necrotic and apoptotic keratinocyte cell death. *Br. J. Dermatol.* 146, 943–951. doi: 10.1046/j.1365-2133.2002.04752.x

Milles, L. F., Schulten, K., Gaub, H. E., and Bernardi, R. C. (2018). Molecular mechanism of extreme mechanostability in a pathogen adhesin. *Science* 359, 1527–1533. doi: 10.1126/science.aar2094

Mongodin, E., Bajolet, O., Cutrona, J., Bonnet, N., Dupuit, F., Puchelle, E., et al. (2002). Fibronectin-binding proteins of *Staphylococcus aureus* are involved in adherence to human airway epithelium. *Infect. Immun.* 70, 620–630. doi: 10.1128/iai.70.2.620-630.2002

Morita, S., Tagai, C., Shiraishi, T., Miyaji, K., and Iwamuro, S. (2013). Differential mode of antimicrobial actions of arginine-rich and lysine-rich histones against gram-positive *Staphylococcus aureus*. *Peptides* 48, 75–82. doi: 10.1016/j.peptides.2013.07.025

Musher, D. M., Lamm, N., Darouiche, R. O., Young, E. J., Hamill, R. J., and Landon, G. C. (1994). The current spectrum of *Staphylococcus aureus* infection in a tertiary care hospital. *Medicine* 73, 186–208. doi: 10.1097/00005792-199407000-00002

Novick, R. P., Ross, H. F., Projan, S. J., Kornblum, J., Kreiswirth, B., and Moghazeh, S. (1993). Synthesis of staphylococcal virulence factors is controlled by a regulatory RNA molecule. *EMBO J.* 12, 3967–3975. doi: 10.1002/j.1460-2075.1993.tb06074.x

O'Brien, L., Kerrigan, S. W., Kaw, G., Hogan, M., Penadés, J., Litt, D., et al. (2002). Multiple mechanisms for the activation of human platelet aggregation by *Staphylococcus aureus*: roles for the clumping factors ClfA and ClfB, the serine-aspartate repeat protein SdrE and protein A. *Mol. Microbiol.* 44, 1033–1044. doi: 10.1046/j.1365-2958.2002.02935.x

O'Halloran, D. P., Wynne, K., and Geoghegan, J. A. (2015). Protein A is released into the *Staphylococcus aureus* culture supernatant with an unprocessed sorting signal. *Infect. Immun.* 83, 1598–1609. doi: 10.1128/IAI.03122-14

O'Neill, E., Pozzi, C., Houston, P., Humphreys, H., Robinson, D. A., Loughman, A., et al. (2008). A novel *Staphylococcus aureus* biofilm phenotype mediated by the fibronectin-binding proteins, FnBPA and FnBPB. *J. Bacteriol.* 190, 3835–3850. doi: 10.1128/JB.00167-08

Otsuka, T., Saito, K., Dohmae, S., Takano, T., Higuchi, W., Takizawa, Y., et al. (2006). Key adhesin gene in community-acquired methicillin-resistant *Staphylococcus aureus*. *Biochem. Biophys. Res. Commun.* 346, 1234–1244. doi: 10.1016/j.bbrc.2006.06.038

Ozeri, V., Rosenshine, I., Mosher, D. F., Fassler, R., and Hanski, E. (1998). Role of integrins and fibronectin in the entry of *Streptococcus pyogenes* into cells via protein F1. *Mol. Microbiol.* 30, 625–637. doi: 10.1046/j.1365-2958.1998.01097.x

Peacock, S. J., Day, N. P., Thomas, M. G., Berendt, A. R., and Foster, T. J. (2000). Clinical isolates of *Staphylococcus aureus* exhibit diversity in Fnb genes and adhesion to human fibronectin. *J. Infect.* 41, 23–31. doi: 10.1053/jinf.2000.0657

Peacock, S. J., Foster, T. J., Cameron, B. J., and Berendt, A. R. (1999). Bacterial fibronectin-binding proteins and endothelial cell surface fibronectin mediate adherence of *Staphylococcus aureus* to resting human endothelial cells. *Microbiology* 145, 3477–3486. doi: 10.1099/00221287-145-12-3477

Peetermans, M., Vanassche, T., Liesenborghs, L., Claes, J., Vande Velde, G., Kwiecinksi, J., et al. (2014). Plasminogen activation by staphylokinase enhances local spreading of *S. aureus* in skin infections. *BMC Microbiol.* 14:310. doi: 10.1186/s12866-014-0310-7

Pickering, A. C., Vitry, P., Prystopiuk, V., Garcia, B., Höök, M., Schoenebeck, J., et al. (2019). Host-specialized fibrinogen-binding by a bacterial surface protein promotes biofilm formation and innate immune evasion. *PLoS Pathog.* 15:e1007816. doi: 10.1371/journal.ppat.1007816

Pietrocola, G., Geoghegan, J. A., Rindi, S., Di Poto, A., Missineo, A., Consalvi, V., et al. (2013). Molecular characterization of the multiple interactions of SpsD, a surface protein from *Staphylococcus pseudintermedius*, with host extracellular matrix proteins. *PLoS One* 8:e66901. doi: 10.1371/journal.pone.0066901

Pietrocola, G., Gianotti, V., Richards, A., Nobile, G., Geoghegan, J. A., Rindi, S., et al. (2015). Fibronectin binding proteins SpsD and SpsL both support invasion of canine epithelial cells by *Staphylococcus pseudintermedius*. *Infect. Immun.* 83, 4093–4102. doi: 10.1128/IAI.00542-15

Pietrocola, G., Nobile, G., Alfeo, M. J., Foster, T. J., Geoghegan, J. A., De Filippis, V., et al. (2019). Fibronectin-binding protein B (FnBPB) from *Staphylococcus aureus* protects against the antimicrobial activity of histones. *J. Biol. Chem.* 294, 3588–3602. doi: 10.1074/jbc.RA118.005707

Pietrocola, G., Nobile, G., Gianotti, V., Zapotoczna, M., Foster, T. J., Geoghegan, J. A., et al. (2016). Molecular interactions of human plasminogen with fibronectin-binding protein B (FnBPB), a fibrinogen/fibronectin-binding protein from *Staphylococcus aureus*. *J. Biol. Chem.* 291, 18148–18162. doi: 10.1074/jbc.M116.731125

Plow, E. F., Haas, T. A., Zhang, L., Loftus, J., and Smith, J. W. (2000). Ligand binding to integrins. *J. Biol. Chem.* 275, 21785–21788. doi: 10.1074/jbc.R000003200

Potts, J. R., Bright, J. R., Bolton, D., Pickford, A. R., and Campbell, I. D. (1999). Solution structure of the N-terminal F1 module pair from human fibronectin. *Biochemistry* 38, 8304–8312. doi: 10.1021/bi990202b

Prystopiuk, V., Feuillie, C., Herman-Bausier, P., Viela, F., Alsteens, D., Pietrocola, G., et al. (2018). Mechanical forces guiding *Staphylococcus aureus* cellular invasion. *ACS Nano* 12, 3609–3622. doi: 10.1021/acsnano.8b00716

Que, Y. A., Haefliger, J. A., Piroth, L., François, P., Widmer, E., Entenza, J. M., et al. (2005). Fibrinogen and fibronectin binding cooperate for valve infection and invasion in *Staphylococcus aureus* experimental endocarditis. *J. Exp. Med.* 201, 1627–1635. doi: 10.1084/jem.20050125

Rennermalm, A., Li, Y. H., Bohaufs, L., Jarstrand, C., Brauner, A., Brennan, F. R., et al. (2001). Antibodies against a truncated *Staphylococcus aureus* fibronectin-binding protein protect against dissemination of infection in the rat. *Vaccine* 19, 3376–3383. doi: 10.1016/s0264-410x(01)00080-9

Richards, A. C., O'Shea, M., Beard, P. M., Goncheva, M. I., Tuffs, S. W., Fitzgerald, J. R., et al. (2018). *Staphylococcus pseudintermedius* surface protein L (SpsL) is required for abscess formation in a murine model of cutaneous infection. *Infect. Immun.* 86:e00631-18. doi: 10.1128/IAI.00631-18

Rindi, S., Cicalini, S., Pietrocola, G., Venditti, M., Festa, A., Foster, T. J., et al. (2006). Antibody response in patients with endocarditis caused by *Staphylococcus aureus*. *Eur. J. Clin. Invest.* 36, 536–543. doi: 10.1111/j.1365-2362.2006.01675.x

Roche, F. M., Downer, R., Keane, F., Speziale, P., Park, P. W., and Foster, T. J. (2004). The N-terminal a domain of fibronectin-binding proteins A and B promotes adhesion of *Staphylococcus aureus* to elastin. *J. Biol. Chem.* 279, 38433–38440. doi: 10.1074/jbc.M402122200

Rudiñ;o-Piñera, E., Ravelli, R. B. G., Sheldrick, G. M., Nanao, M. H., Vladimir, V., Korostelev, V. V., et al. (2007). The solution and crystal structures of a module pair from the *Staphylococcus aureus*-binding site of human fibronectin–a tale with a twist. *J. Mol. Biol.* 368, 833–844. doi: 10.1016/j.jmb.2007.02.061

Rydén, C., Rubin, K., Speziale, P., Höök, M., Lindberg, M., and Wadström, T. (1983). Fibronectin receptors from *Staphylococcus aureus*. *J. Biol. Chem.* 258, 3396–3401.

Sanderson-Smith, M. L., De Oliveira, D. M. P., Ranson, M., and McArthur, J. D. (2012). Bacterial plasminogen receptors: mediators of a multifaceted relationship. *J. Biomed. Biotechnol.* 2012:272148. doi: 10.1155/2012/272148

Saravia-Otten, P., Müller, H. P., and Arvidson, S. (1997). Transcription of *Staphylococcus aureus* fibronectin binding protein genes is negatively regulated by Agr and an Agr-independent mechanism. *J. Bacteriol.* 179, 5259–5263. doi: 10.1128/jb.179.17.5259-5263.1997

Schwarz-Linek, U., Werner, J. M., Pickford, A. R., Gurusiddappa, S., Kim, J. H., Pilka, E. S., et al. (2003). Pathogenic bacteria attach to human fibronectin through a tandem beta-zipper. *Nature* 423, 177–181. doi: 10.1038/nature01589

Sendi, P., and Proctor, R. A. (2009). *Staphylococcus aureus* as an intracellular pathogen: the role of small colony variants. *Trends Microbiol.* 17, 54–58. doi: 10.1016/j.tim.2008.11.004

Shinji, H., Yosizawa, Y., Tajima, A., Iwase, T., Sugimoto, S., Seki, K., et al. (2011). Role of fibronectin-binding proteins A and B in vitro cellular infections and in vivo septic infections by *Staphylococcus aureus*. *Infect. Immun.* 79, 2215–2223. doi: 10.1128/IAI.00133-11

Signäs, C., Raucci, G., Jönsson, K., Lindgren, P. E., Anantharamaiah, G. M., and Höök, M. (1989). Nucleotide Sequence Of The Gene For A Fibronectin-Binding Protein From *Staphylococcus aureus*: use of this peptide sequence in the synthesis of biologically active peptides. *Proc. Natl. Acad. Sci. U.S.A.* 86, 699–703. doi: 10.1073/pnas.86.2.699

Sinha, B., François, P. P., Nüsse, O., Foti, M., Hartford, O. M., Vaudaux, P., et al. (1999). Fibronectin-binding protein acts as *Staphylococcus aureus* invasin via fibronectin bridging to integrin alpha5beta1. *Cell Microbiol.* 1, 101–117. doi: 10.1046/j.1462-5822.1999.00011.x

Sollberger, G., Tilley, D. O., and Zychlinsky, A. (2018). Neutrophil extracellular traps: the biology of chromatin externalization. *Dev. Cell* 44, 542–553. doi: 10.1016/j.devcel.2018.01.019

Speziale, P., Pietrocola, G., Foster, T. J., and Geoghegan, J. A. (2014). Protein-based biofilm matrices in Staphylococci. *Front. Cell. Infect. Microbiol.* 4:171. doi: 10.3389/fcimb.2014.00171

Speziale, P., Pietrocola, G., Rindi, S., Provenzano, M., Provenza, G., Di Poto, A., et al. (2009). Structural and functional role of *Staphylococcus aureus* surface components recognizing adhesive matrix molecules of the host. *Future Microbiol.* 4, 1337–1352. doi: 10.2217/fmb.09.102

Stemberk, V., Jones, R. P., Moroz, O., Atkin, K. E., Edwards, A. M., Turkenburg, J. P., et al. (2014). Evidence for steric regulation of fibrinogen binding to *Staphylococcus aureus* fibronectin-binding protein A (FnBPA). *J. Biol. Chem.* 289, 12842–12851. doi: 10.1074/jbc.M113.543546

Sticht, H., Pickford, A. R., Potts, J. R., and Campbell, I. D. (1998). Solution structure of the glycosylated second type 2 module of fibronectin. *J. Mol. Biol.* 276, 177–187. doi: 10.1006/jmbi.1997.1528

Stones, D. H., and Krachler, A. M. (2016). Against the tide: the role of bacterial adhesion in host colonization. *Biochem. Soc. Trans.* 44, 1571–1580. doi: 10.1042/BST20160186

Tong, S. Y., Davis, J. S., Eichenberger, E., Holland, T. L., and Fowler, V. G. Jr. (2015). *Staphylococcus aureus* infections: epidemiology, pathophysiology, clinical manifestations, and management. *Clin. Microbiol. Rev.* 28, 603–661. doi: 10.1128/CMR.00134-14

Usui, A., Murai, M., Seki, K., Sakurada, J., and Masuda, S. (1992). Conspicuous ingestion of *Staphylococcus aureus* organisms by murine fibroblasts in vitro. *Microbiol. Immunol.* 36, 545–550. doi: 10.1111/j.1348-0421.1992.tb02054.x

Vaudaux, P., Francois, P., Bisognano, C., Kelley, W. L., Lew, D. P., Schrenzel, J., et al. (2002). Increased expression of clumping factor and fibronectin-binding proteins by hemB mutants of *Staphylococcus aureus* expressing small colony variant phenotypes. *Infect. Immun.* 70, 5428–5437. doi: 10.1128/iai.70.10.5428-5437.2002

Vergara-Irigaray, M., Valle, J., Merino, N., Latasa, C., García, B., Ruiz de Los Mozos, I., et al. (2009). Relevant role of fibronectin-binding proteins in *Staphylococcus aureus* biofilm-associated foreign-body infections. *Infect. Immun.* 77, 3978–3991. doi: 10.1128/IAI.00616-09

Verhamme, I. M., Panizzi, P. R., and Bock, P. E. (2015). Pathogen activators of plasminogen. *J. Thromb. Haemost.* 13(Suppl. 1), S106–S114. doi: 10.1111/jth.12939

Viela, F., Mathelié-Guinlet, M., Pietrocola, G., Pietro Speziale, P., and Dufrêne, Y. F. (2020). The molecular complex between staphylococcal adhesin SpsD and fibronectin sustains mechanical forces in the nanonewton range. *mBio* 11:e00371-20. doi: 10.1128/mBio.00371-20

Wann, E. R., Gurusiddappa, S., and Höök, M. (2000). The fibronectin-binding MSCRAMM FnbpA of *Staphylococcus aureus* is a bifunctional protein that also binds to fibrinogen. *J. Biol. Chem.* 275, 13863–13871. doi: 10.1074/jbc.275.18.13863

Weinman, A. L., Sullivan, S. G., Vijaykrishna, D., Markey, P., Levy, A., Miller, A., et al. (2020). Epidemiological trends in notified influenza cases in Australia's Northern Territory, 2007-2016. *Influenza Other Respir. Viruses* doi: 10.1111/irv.12757 [Epub ahead of print].

Weisel, J. W. (2005). Fibrinogen and fibrin. *Adv. Protein Chem.* 70, 247–299. doi: 10.1016/S0065-3233(05)70008-5

Wertheim, H. F. L., Melles, D. C., Vos, M. C., van Leeuwen, W., van Belkum, A., Verbrugh, H. A., et al. (2005). The role of nasal carriage in *Staphylococcus aureus* infections. *Lancet Infect. Dis.* 5, 751–762. doi: 10.1016/S1473-3099(05)70295-4

Wolz, C., Pöhlmann-Dietze, P., Steinhuber, A., Chien, Y. T., Manna, A., van Wamel, W., et al. (2000). Agr-independent regulation of fibronectin-binding protein(s) by the regulatory locus sar in *Staphylococcus aureus*. *Mol. Microbiol.* 36, 230–243. doi: 10.1046/j.1365-2958.2000.01853.x

Xu, J., Zhang, X., Pelayo, R., Monestier, M., Ammollo, C. T., Semeraro, F., et al. (2009). Extracellular histones are major mediators of death in sepsis. *Nat. Med.* 15, 1318–1321. doi: 10.1038/nm.2053

Zuo, Q. F., Cai, C. Z., Ding, H. L., Wu, Y., Yang, L. Y., Feng, Q., et al. (2014). Identification of the immunodominant regions of *Staphylococcus aureus* fibronectin-binding protein A. *PLoS One* 9:e95338. doi: 10.1371/journal.pone.0095338

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Microbiology

REVIEW
published: 18 September 2020
doi: 10.3389/fmicb.2020.582582

Lipoproteins in Gram-Positive Bacteria: Abundance, Function, Fitness

*Minh-Thu Nguyen[1], Miki Matsuo[2], Silke Niemann[1], Mathias Herrmann[1] and Friedrich Götz[2]**

[1] Section of Medical and Geographical Infectiology, Institute of Medical Microbiology, University Hospital of Münster, Münster, Germany, [2] Interfaculty Institute of Microbiology and Infection Medicine Tübingen (IMIT), Microbial Genetics, University of Tübingen, Tübingen, Germany

When one thinks of the Gram+ cell wall, the peptidoglycan (PG) scaffold in particular comes to mind. However, the cell wall also consists of many other components, for example those that are covalently linked to the PG: the wall teichoic acid and the cell wall proteins tethered by the sortase. In addition, there are completely different molecules that are anchored in the cytoplasmic membrane and span the cell wall. These are lipoteichoic acids and bacterial lipoproteins (Lpp). The latter are in the focus of this review. Lpp are present in almost all bacteria. They fulfill a wealth of different tasks. They represent the window to the outside world by recognizing nutrients and incorporating them into the bacterial cell via special transport systems. Furthermore, they perform very diverse and special tasks such as acting as chaperonin, as cyclomodulin, contributing to invasion of host cells or uptake of plasmids via conjugation. All these functions are taken over by the protein part. Nevertheless, the lipid part of the Lpp plays an as important role as the protein part. It is the released lipoproteins and derived lipopeptides that massively modulate our immune system and ultimately play an important role in immune tolerance or non-tolerance. All these varied activities of the Lpp are considered in this review article.

Keywords: lipoproteins, gram-positive bacteria, ion transport, host cell invasion, immune system

OPEN ACCESS

Edited by:
Timothy J. Foster, Trinity College Dublin, Ireland

Reviewed by:
Dominique Missiakas, The University of Chicago, United States
Bok-Luel Lee, Pusan National University, South Korea

***Correspondence:**
Friedrich Götz
friedrich.goetz@uni-tuebingen.de

Specialty section:
This article was submitted to Infectious Diseases, a section of the journal Frontiers in Microbiology

Received: *12 July 2020*
Accepted: *27 August 2020*
Published: *18 September 2020*

Citation:
Nguyen M-T, Matsuo M, Niemann S, Herrmann M and Götz F (2020) Lipoproteins in Gram-Positive Bacteria: Abundance, Function, Fitness. Front. Microbiol. 11:582582. doi: 10.3389/fmicb.2020.582582

INTRODUCTION

As far as we know today, bacterial lipoproteins (Lpp) are found in all bacteria. They are anchored in the membrane via their N-terminal lipid structure and are involved in a number of metabolic processes, some of which are vital. Using various computer-assisted algorithms, precursor Lpp can be identified by their signal peptide containing the characteristic "lipobox" (Babu et al., 2006). After cleavage of the signal peptide, Lpp are anchored into the bacterial membrane via a diacylglycerol moiety tethered via a thioether linkage to the N-terminal cysteine, and the free cysteine N-terminus is usually N-acylated (Buddelmeijer, 2015). The first Lpp to be elucidated was the Braun's Lpp, a

Abbreviations: Lpp, Lipoprotein(s); *lpl*, lipoprotein-like lipoprotein genes; Lps, lipopolysaccharides; *lgt*, diacylglyceryl transferase enzyme encoding gene; Lsp, signal peptidase II encoding gene; PGN, peptidoglycan; TLR2: Toll-like receptor 2; *B., Bacillus; C., Clostridium; E., Enterococcus; G., Geobacillus; L., Listeria; M., Mycobacterium; O., Oceanobacillus; S., Staphylococcus* or *Streptococcus*.

murein Lpp of *E. coli* which is anchored with its N-terminal lipid structure in the outer membrane of almost all Gram-negative bacteria (Hantke and Braun, 1973; Braun, 1975). As we now know, Lpp occur in all bacteria, and their structure, function and biosynthesis has been review recently (Sutcliffe and Russell, 1995; Braun and Hantke, 2019). In Gram+ bacteria, the Lpp's lipid moieties are anchored in the outer layer of the cytoplasmic membrane and their protein portions span the cell wall. While the class of proteins covalently bound to the cell wall (peptidoglycan) via sortase consists mostly of adhesins such as immunoglobulin, fibronectin, or collagen binding proteins (Navarre and Schneewind, 1999), most Lpp are involved in nutrient uptake, possess enzymatic activities or fulfill very diverse and specific tasks. The biosynthetic steps from pre-Lpp to mature Lpp include two enzymatic reactions that appear to be highly conserved in all bacteria. These are diacylglyceryl transferase (Lgt) and the signal peptidase II (Lsp). A third modification, namely the acylation of the amino group of the N-terminal cysteine by the apolipoprotein N-acyltransferase (Lnt), occurs mainly in Gram-negative bacteria and the Gram+ actinobacteria. In firmicutes this third reaction is not compulsory and may be catalyzed by a different enzymatic reaction. The Lpp biosynthesis and the crucial role of Lpp maturation for virulence and TLR2 signaling have been recently described in detail (Buddelmeijer, 2015; Nguyen and Götz, 2016).

In this review, we analyze the total Lpp in 14 Gram+ bacteria species from different genera. In the first part, these Lpp are analyzed and classified by their function. In the second part, the diversity of lipid moiety structure in Gram+ Lpp and their impact on host immune modulation are discussed. The third part deals with the tightness of Lpp on the membrane and their release to the environment. Finally, we review the impact of Lpp as vaccine candidates.

FUNCTION OF LPP IN GRAM+ BACTERIA

Here we compared 14 bacterial species from 7 genera, i.e., *Staphylococcus*, *Bacillus*, *Listeria*, *Streptococcus*, *Enterococcus*, *Clostridium*, and *Mycobacterium* (**Table 1**). Most of the species play a role as pathogens. However, we have also included some non-pathogenic species from the *Bacillus* group in which some Lpp have been functionally well studied. The number of Lpp genes varies from species to species but accounts for 1–3% of all genes of a genome, representing considerable proportion of a bacterial genome.

Some genera contain a relatively high, others a comparatively low number of Lpp (20 to 35), such as streptococci, *C. perfringens* and *M. tuberculosis* (**Figure 1**). The high number of different Lpp and also the high expression of certain Lpp such as MntC (SitC)-binding protein of a manganese ABC transporter (Stoll et al., 2005) show that this class of proteins can make up a major part of the cell wall associated proteins. However, their impact on peptidoglycan biosynthesis and structure are barely investigated.

Substrate Binding in ABC Transport System

In **Table 2** the diverse Lpp are summarized and grouped according to their function and occurrence in the selected bacterial species. The full list of Lpp genes in the various Gram+ bacteria with the annotated function is described in **Supplementary Tables S1–S13**. A large proportion of Gram+ Lpp are substrate-binding proteins (SBPs) of ABC transporter systems responsible for the acquisition of multiple nutrients including amino acids and short peptides, sugars, polyamines, and many metal ions. Some of the transporter activities are described below in more detail.

Fe-Acquisition

A relatively high proportion of Lpp is involved in ion transport. Bacterial pathogens in particular employ numerous strategies to acquire the essential nutrient iron since the amount of free iron in the host is extremely low. Several of the Fe-acquisition system involves Lpp. For instance, in *S. aureus* USA300 8 Lpp are associated with iron acquisition or utilization of host-derived heme iron (isd operon) as an iron source (Shahmirzadi et al., 2016). One of the more intensively studied iron transporter systems in *S. aureus* uses the two Lpp (FhuD1 and FhuD2) which are part of the FhuCBG system (Sebulsky et al., 2000). FhuD1 and FhuD2 bind iron(III)-hydroxamate siderophores and present it to the ABC transporter FhuCBG. Another well studied iron transporter system is SirABC which is a more promiscuous importer transporting ferric hydroxamates, ferric enterobactin or ferric citrate and staphylobactin (Dale et al., 2004). Other transport systems initially stirring great interest among bacterial physiology researcher are the twin-arginine translocation (TAT) systems in pathogenic staphylococci. But the reality did not meet fully expectations. In fact, only the existence of one TAT system, namely FepABC could be confirmed comprising of the Lpp FepA domain representing the Fe-binding protein, FepB with the TAT signal peptide functioning as an Fe-dependent peroxidase, and FepC moiety presumably acting as the Fe transporter through the membrane (Biswas et al., 2009).

In staphylococci there is a correlation between the amount of Lpp-dependent Fe-uptake systems and pathogenicity; increasing pathogenic potential appears to be paralleled by an enhanced number of Lpp-dependent Fe-uptake systems (Shahmirzadi et al., 2016). There are only three bacterial groups where Lpp-mediated Fe-transporters are either absent or not yet identified: streptococci and *M. tuberculosis* (**Table 2**). This is astonishing, as iron is an essential nutrient for the growth of most bacteria. The exceptions are certain lactic acid bacteria that do not need iron since they use manganese instead (Götz et al., 1980; Archibald, 1986). But most other bacteria have developed specific iron-transport systems located on the membrane surface to take up iron and iron complexes such as heme or ferrichrome. *S. pneumoniae* uses ferrichrome as an iron source (Yang et al., 2014). To reveal the reason of lacking Lpp-dependent Fe transport system, we blasted the Fe-binding *lpp* genes in streptococci, and found the signal peptide of *lpp* genes modified

TABLE 1 | Number of Lpp of selected Gram-positive bacterial strains.

	Species	Strains	Org code	No of protein genes	No of Lpp	Description	References
Staphylococcus	*Staphylococcus aureus*	USA300	saa	2604	67	Round-shaped, non-spore forming, pathogenic bacterium. USA300 is a globally spread, highly virulent community MRSA strain.	Diep et al., 2006; Nimmo, 2012
Bacillus	*Bacillus subtilis*	subsp. subtilis 168	bsu	4174	63	Non-pathogenic, aerobic, endospore-forming, rod-shaped bacterium, commonly found in soil.	Kunst et al., 1997
	Bacillus cereus	ATCC 14579	bce	5231	96	Rod-shaped, anaerobic, spore forming bacterium. Foodborne pathogenic strain. Ubiquitous environmental distribution.	Ivanova et al., 2003
	Bacillus licheniformis	ATCC 14580	bli	4179	57	Rod-shaped, mesophilic, spore forming soil bacterium. Used in food industry.	Rey et al., 2004
	Geobacillus kaustophilus	HTA426	gka	3540	44	Rod-shaped, aerobic endospore forming bacterium. Isolated from the deep-sea sediment of the Mariana Trench.	Takami et al., 2004
	Oceanobacillus iheyensis	HTE831	oih	3496	99	Rod-shaped, aerobic endosporing forming bacterium. An alkaliphilic and extremely halotolerant Bacillus-related species isolated from a deep-sea sediment collected at a depth of 1050 m on the Iheya Ridge, Japan.	Takami et al., 2002
Listeria	*Listeria monocytogenes*	EGD-e	lmo	2867	53	Foodborne pathogen, non-spore forming bacteria. The serotype 1/2a strain tends to be more sporadic and not associated with epidemics.	Glaser et al., 2001; Donaldson et al., 2009
Streptococcus	*Streptococcus pneumoniae*	TIGR4	spn	2125	27	Pathogenic, non-spore forming, facultative anaerobic and round-shape bacterium. Major cause of bacterial pneumonia and other invasive infection.	Tettelin et al., 2001
	Streptococcus pyogenes	M1 GAS	spy	1693	26	Pathogenic, aerotolerant, non-spore forming, and round shape bacterium, frequently causing skin and soft tissue infections.	Ferretti et al., 2001
	Streptococcus agalactiae	NEM361	san	2094	27	Pathogenic, facultative anaerobic and round-shape bacterium. Group B streptococcus (GBS), frequent colonizer of the urogenital tract and major cause of neonatal sepsis.	Glaser et al., 2002
	Streptococcus mutans	UA159	smu	1960	20	Pathogenic, anaerobic and round shape bacteria. Part of the oropharyngeal flora, major cause for dental caries. Potential for invasive disease (endocarditis).	Ajdic et al., 2002
Enterococcus	*Enterococcus faecalis*	V583	efa	3264	59	Opportunistic pathogenic, anaerobic, non-motile, round-shape bacterium. Vancomycin-resistant clinical isolate from blood cultures from an indexed patient in the United States in 1987.	Paulsen et al., 2003
Clostridium	*Clostridium perfringens*	13	cpe	2723	36	Rod-shaped, anaerobic, spore-forming pathogenic bacterium. Cause of gas gangrene.	Shimizu et al., 2002
Mycobacterium	*Mycobacterium tuberculosis*	CDC1551	mtc	4189	30	Obligate pathogenic, aerobic bacterium, acid-fast. Causative agent of tuberculosis in humans and animals.	Fleischmann et al., 2002

into transmembrane signal peptide. *M. tuberculosis* uses the ESX-3 secretion system (Type VII secretion systems, T7SS) which is essential for siderophore-mediated iron uptake and for heme utilization (Zhang et al., 2020); this system does not involve Lpp. This might explain why *M.t.* does not involve an Lpp in iron acquisition.

In *S. pyogenes* the Lpp MtsA as part of the MtsABC iron transporter system binds primarily $Fe(2+) > Fe(3+) > Cu(2+) > Mn(2+) > Zn(2+)$ (Sun et al., 2008, 2009). Interestingly, MtsA requires bicarbonate as a synergistic anion for stable ferrous binding which is similar to the iron binding in human transferrin. Another Fe-containing Lpp in *S. pyogenes* is SiaA, part of the SiaABC heme transporter (Sun et al., 2010). A typical ABC transport system with Lpp as a substrate binding protein being part of it is shown in **Figure 2A** for the sirABC system in *S. aureus* (Dale et al., 2004).

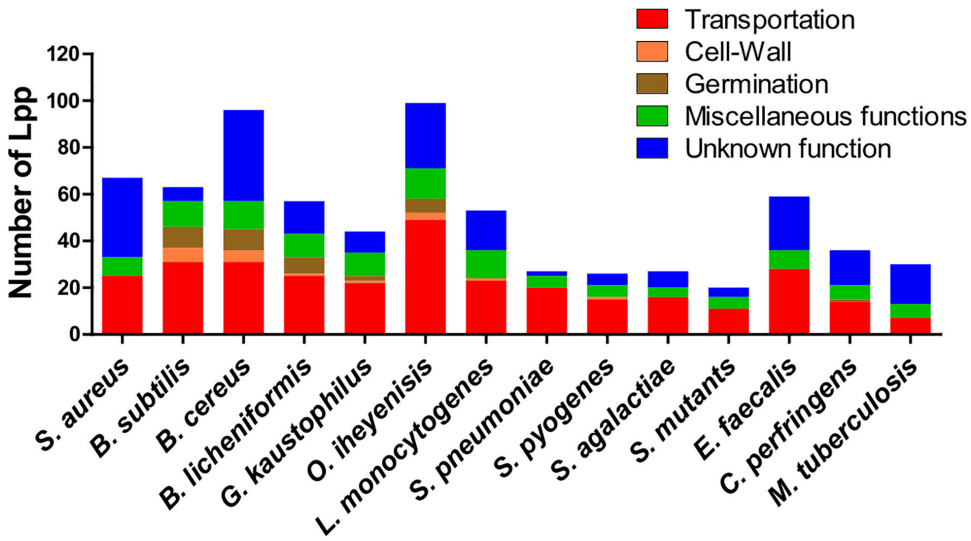


FIGURE 1 | Lipoproteins content in 14 Gram+ bacterial species. Lpp are categorized by their function such as transportation, cell-wall turnover, germination, miscellaneous functions and unknown function. A more detailed information of Lpp in each group are presented in **Supplementary Table S1–S13** for 13 species (*Bacillus subtilis*, *Bacillus cereus*, *Bacillus licheniformis*, *Geobacillus kaustophilus*, *Oceanobacillus iheyenisis*, *Listeria monocytogenes*, *Streptococcus pneumoniae*, *Streptococcus pyogenes*, *Streptococcus agalactiae*, *Streptococcus mutants*, *Enterococcus faecalis*, *Clostridium perfringens*, and *Mycobacterium tuberculosis*). Data for *Staphylococcus aureus* Lpp are taken from Shahmirzadi et al. (2016).

Other Cation Transporters

Other cation transporters mediating the transport of Co, Cu, Mn, Mo, Ni, Zn ions and often involving Lpp have also been studied. Similar to Fe, many of these ions are limited in the infection environment. Therefore, these transporters are very important for survival in an infection, hence also for the pathogenicity of the respective bacteria. Especially Mn plays an important role besides Fe. It is therefore not surprising that one of the most frequently represented Lpp in *S. aureus* is the Mn-binding protein (MntC) (Stoll et al., 2005). It was formerly called SitC because it was thought to be a Fe binding protein, but MntABC turned out to be a Mn transport system (Horsburgh et al., 2002; Diep et al., 2014).

Such manganese transporters with MntA homologs are also prevalent in other bacteria such as bacilli (MntA) or streptococci (PsaA) (Bray et al., 2009; Li et al., 2014). In *S. aureus* three Lpp were annotated as Ni transporters. The *S. aureus* Cnt (Opp1) ABC transporter imports nickel and cobalt in zinc-depleted conditions and contributes to virulence (Remy et al., 2013). The Lpp ModA is part of the molybdate transporter complex (ModABC) (Neubauer et al., 1999). In Group A *Streptococcus* (GAS) the two surface Lpp, AdcA and AdcAII, are involved in zinc uptake and zinc homeostasis with AdcA acting as a zinc-specific importer (Tedde et al., 2016). Deletion of the zinc transporter Lpp AdcAII caused hyperencapsulation of *S. pneumoniae* associated with distinct alleles of the type I restriction-modification system (Durmort et al., 2020).

Anion Transporters

Lipoproteins provide functions as various anion transporters for phosphate, phosphonate or nitrate. In bacilli Lpp are involved in the osmo-regulated ABC transport system OpuA (Kempf et al., 1997). This is a high-affinity uptake system for the osmoprotectant glycine betaine which consists of ATPase (OpuAA), an integral membrane protein (OpuAB), and the Lpp (OpuAC).

Amino Acid and Peptide Transporters

Amino acid and peptide transporters involving Lpp are even more abundant than Fe-transporters. One of the first described oligopeptide permease involving an Lpp is the OppABCD(F) system (**Figure 2B**), in which the OppA represents the Lpp. OppB and OppC are integral membrane proteins, and OppD and OppF are ATP binding proteins that couple ATP hydrolysis to the transport process (Perego et al., 1991). OppA variants are involved in surfactin regulation in *B. subtilis* (Wang et al., 2019). In *Bacillus* there are more such oligopeptide transporters as for example the AppDFABC with AppA corresponding to binding oligopeptide (Koide and Hoch, 1994). In *Staphylococcus aureus* 7 Lpp were identified being involved in amino acid and oligopeptide transport. One of the most important systems seems to be Opp3 (Hiron et al., 2007) in relevance followed by GmpC which is part of an ABC transporter and binds glycyl-methionine (Williams et al., 2004). Methionine is one of the least abundant amino acids in physiological fluids (4 μg ml^{-1}) and many bacteria such as streptococci possess an ABC transporter (MetQNP) for methionine uptake. MetQ represents the Lpp that binds methionine (Basavanna et al., 2013); this transporter is crucial for growth and virulence.

Sugar Transporters

It has been known that sugars are transported by the phosphotransferase system (PTS). Only few studies relating sugar ABC transporters are available. However, screening the Lpp

TABLE 2 | Number of lipoprotein genes categorized as function.

Function	*S. a*	*B. s*	*B. c*	*B. l*	*G. k*	*O. i*	*L. m*	*S. pneu*	*S. pyro*	*S. a*	*S. m*	*E. f*	*C. p*	*M. t*
Transportation	25	31	31	25	22	49	23	20	15	16	11	28	14	7
Fe	8	5	8	5	3	4	4		1			3	1	
Zn, Mo, Mn. . ..	6	5				3	3	2	2	1		3		1
PO4, NO3. . ..	3	2		2	2	3		1	1	2	1	2	1	1
AA and oligopeptide	7	10	20	6	8	23	8	8	3	4	5	13	4	2
Sugar	1	6	2	7	2	6	5	2	3		2	2	3	1
Lipid														2
Unknown transp.		3	1	5	7	10	3	7	5	9	3	5	5	
Cell-Wall		6	5	1	1	3	1		1					
Penicillin-binding protein, transpeptidase		3	2			1								
Carboxypeptidase					1		1							
Polysaccharide deacetylase		1	1											
YqiH		1		1										
LytA		1												
Succinoglycan biosynthesis			1											
Amidase						1								
Ppiase B						1								
Peptidoglycan hydrolase									1					
Capsule biosynthesis			1											
Germination		9	9	7	2	6							1	
Enzymes and foldases	8	11	12	10	10	13	12	5	6	4	5	8	6	6
Membrane insertase YidC	1	1	2	1	2	1	2	1	1	2	2	1		
Peptidylprolyl isomerase PrsA	1		4	1	1	1	2	1	1	2	1	1		
Peptidylproly isomerase B								1	1		1	1	1	
Superoxide dismutase (SOD)		1			1	1								1
SCO family protein			1		1	1								
Thioredoxin family	1							1					1	1
Respiratory chain	1	2		2	1									
Beta lactamase	1													
Proteinase/Peptidase						1		1	1		1	1		1
Hydrolase/Esterase/Lipase		2	2			2	3					1		
Phosphatase			1						1			1	1	1
Nisin, subtilin immunity (LanI)		1												
CamS family sex pheromone	1	1	1			1	1							
Other enzymes	2	3	1	6	4	5	4		1			2	3	2

database we found a number of Lpp genes are responsible for sugar binding in ABC transport system (**Table 2**). It has been shown that the maltodextrin uptake is related to ABC transporter during host colonization by *E. faecalis* (Sauvageot et al., 2017).

Lipid Transporter

Lipid transporters that involve Lpp are rare in bacteria with the exception of mycobacteria. *M. tuberculosis* imports cholesterol across its atypical diderm cell envelope. Four closely related ATP-binding cassette (ABC) transporter-like complexes called Mce are responsible for the import (Wilburn et al., 2018). Three of the Mce proteins represent Lpp (Fenn et al., 2019). Mce proteins are involved in modulating host cell signaling and allow the growth with cholesterol as sole C-source.

Cell Wall

Little is known about Lpp involved in cell wall biosynthesis or degradation. In *Bacillus* spp., *Lactobacillus* spp. and *Staphylococcus* spp. an Lpp named New Lipoprotein C/Protein of 60-kDa (NlpC/P60) has been identified. This is a papain-like cysteine peptidase involved in the catalysis of the N-acetylmuramate-L-alanine or D-γ-glutamyl-meso-diaminopimelate linkages (Anantharaman and Aravind, 2003; Vermassen et al., 2019). In *Bacillus* group, Lpp has been annotated as putative polysaccharide deacetylase, cell elongation specific DD-transpeptidase, and penicillin-binding proteins, respectively (**Supplementary Tables S1–S13**). The gene cluster YqiHIK, with YqiH as an Lpp is involved in hydrolytic activity of the peptidoglycan saccule. Even though YqiHIK is not critical for reestablishing the rod-shaped morphology, the deletion of this operon impaired growth in a defined minimal medium (Fischer and Bremer, 2012). Moreover in *Bacillus* spp. the operon lytRABC with lytA as Lpp regulates the gene of N-acetylmuramoyl-L-alanine amidase (Lazarevic et al., 1992).

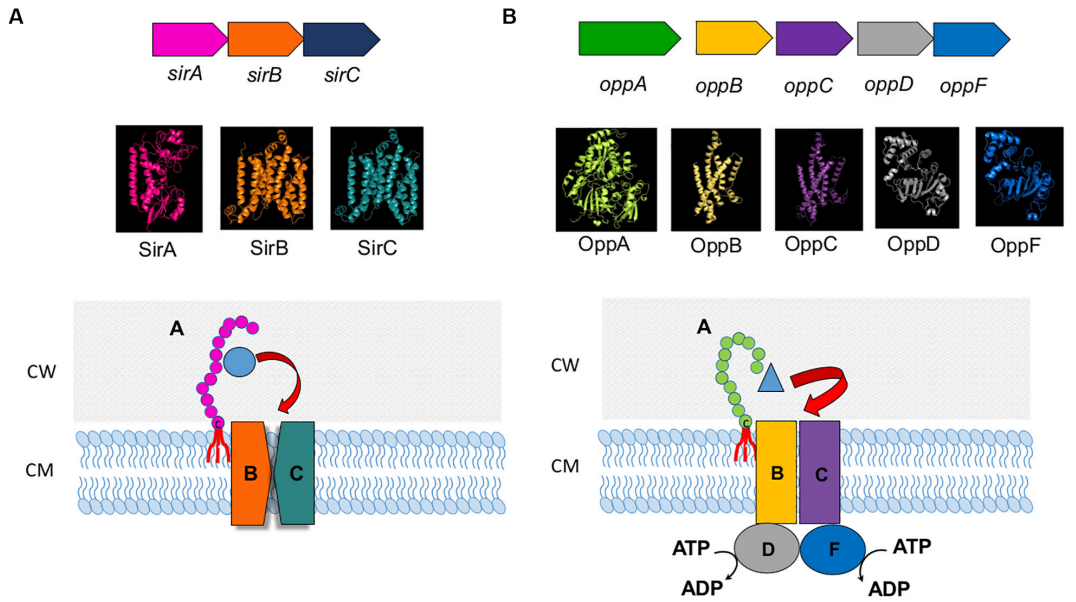


FIGURE 2 | ABC transport system. **(A)** sirABC ion transport system. Upper panel illustrates the operon in *Staphylococcus aureus* (Dale et al., 2004). Middle panel presents 3D models of SirA, SirB, and SirC. 3D crystal model of SirA was downloaded from protein data bank (PDB ID: 3MWF) and visualized in PyMoL software (DeLano, 2009). Phyre2 (Kelley et al., 2015) was used to predict the 3D configuration of SirB and SirC. The Phyre2 predition showed 8 transmembrane motifs in both SirB and SirC ranging from 14 to 324 amino acids in SirB and 16 to 326 amino acids in SirC. There were 2 re-entrant helices which connect the 5th to 6th and 6th to 7th transmembrane motifs in both SirB and SirC. Superimposition of SirB to SirC in PyMoL showed high similarity in the 3D structure of SirB and SirC. The lower panel is the illustration of the SirABC transport system for Fe-staphylobactin (green circle). **(B)** OppABCDF oligopeptide transport system. The upper panel illustrates the operon in *Bacillus subtilis* (Perego et al., 1991). Middle panel presents 3D models of OppA, B, C, D and F. Phyre2 (Kelley et al., 2015) was used to predict the 3D configuration of OppA, OppB, OppC, and OppD. 3D crystal model of OppF was downloaded from protein data bank (SMR ID: P24137) and visualized in PyMoL software (DeLano, 2009). The Phyre2 predition showed 6 transmembrane motifs in both OppB and OppC ranging from 6 to 297 amino acids in OppB and 43 to 291 amino acids in OppC. Superimposition of OppB and OppD to OppC and OppF in PyMoL showed high similarity in the 3D structure of OppB and OppC, and OppD and OppF, respectively. The lower panel is the illustration of the OppABCDF transport system for oligopeptide (blue triangle).

Spore Germination

In group of rod shape bacteria producing endospores, Lpp genes are also involved in the spore germination process for example GerA/D in *Bacillus subtilis* or GerS in *Clostridium difficile* (Pelczar and Setlow, 2008; Li et al., 2014; Diaz et al., 2018). In fact, a *Bacillus anthracis lgt* mutant lacking the lipid moiety germinated inefficiently; in line, the spore displayed attenuated virulence in the mouse infection model (Okugawa et al., 2012).

Enzymes and Foldases

YidC

YidC is probably the most important and also essential Lpp in bacteria. It has been first functionally analyzed in *E. coli* as a membrane insertase for Sec-independent proteins (Samuelson et al., 2000; Serek et al., 2004). The *E. coli* YidC is a 61 kD protein of the inner (cytoplasmic) membrane and is composed of 5–6 transmembrane (TM) helices that contact hydrophobic segments of the substrate proteins. Since YidC also cooperates with the SecYEG translocon it is widely involved in the assembly of many different membrane proteins including proteins that obtain complex membrane topologies (Petriman et al., 2018). YidC homologs are not only present and essential in bacteria but homologs are also found in mitochondria (Oxa1) and thylakoids (Alb3) a fact pointing toward a common evolutionary origin, and also demonstrating the general importance of this cellular process (Dalbey et al., 2014). It is therefore not surprising that YidC homologs are present in all bacteria.

Peptidylprolyl Isomerase PrsA

PrsA homologs are also very abundant in bacteria. In most bacteria it is an essential membrane-bound Lpp that is involved in protein secretion (Kontinen and Sarvas, 1993). Prolyl *cis*/*trans* isomerase (PPIases) PrsA accelerates the folding of proteins containing *cis*-prolines thus abrogating the rate-limiting steps in the folding of proteins containing *cis*-prolines (Gothel and Marahiel, 1999). Numerous reports have demonstrated that *prsA* mutants are affected in protein secretion, and also in virulence in diverse Gram+ pathogens (Vitikainen et al., 2004; Heikkinen et al., 2009; Guo et al., 2013; Chen et al., 2015; Jousselin et al., 2015; Liu et al., 2019).

Respiratory Chain

In *Bacillus* spp. at least one Lpp, i.e., the SCO (Synthesis of Cytochrome c Oxidase) plays a role in the respiratory chain. SCO homologs are found in many bacteria. They have a high affinity for Cu^{2+} and are required for the proper synthesis of cytochrome c oxidase (Xu et al., 2015). Moreover, in *S. aureus* QoxA is an Lpp. QoxA is part of the terminal cytochrome aa3 quinol oxidase encoded by *qoxABCD* (Götz and Mayer, 2013;

Hammer et al., 2013). In *S. pneumonia* two surface-exposed thioredoxin-family lipoproteins, Etrx1 and Etrx2, are crucial for resistance to H_2O_2 oxidative stress (Saleh et al., 2013).

Enzymes

A considerable number of Lpp have been shown to exert enzymatic activity. Many Gram+ pathogens contain two classes of Lpp molecules conferring ß-lactamase activities, a secreted and a membrane bound Lpp. Such Lpp-penicillinases are described in *B. licheniformis*, *B. cereus*, and *S. aureus* (Nielsen and Lampen, 1982a,b). The *S. aureus* Lpp-penicillinase was one of the first Lpp described in a Gram+ bacterium. Other Lpp enzymes have peptidase, protease, lipase, esterase or phosphatase activities. For example, the LppC of *Streptococcus equisimilis* functions as an acid phosphatase (Malke, 1998). In various mycobacteria three Lpp with phosphatase activity (PhoA, SapM) were described (Wolschendorf et al., 2007). Phosphorus is indispensable for the biosynthesis of nucleic acids and phospholipids and for the energy supply of any cell.

Lantibiotic Immunity Lpp (LanI)

Lantibiotics are a subgroup of bacteriocins from Gram+ bacteria consisting of polycyclic peptides containing modified amino acids. The biosynthetic gene cluster of Pep5, Epicidin 280, nisin and subtilin contain an immunity gene that encodes a small Lpp collectively named LanI (Hoffmann et al., 2004; Geiger et al., 2019). The activity of LanI molecules on lantibiotics is highly specific. For example, NisI protects the bacterial cell against nisin, but not against structurally very similar lantibiotics from other species such as subtilin from *B. subtilis*. Structural analysis reveals that the LanI's C-terminal domain binds nisin, thereby preventing nisin from reaching its target molecules (Hacker et al., 2015; Jeong and Ha, 2018).

Sex Pheromones

Enterococcus faecalis, *S. aureus*, *S. epidermidis*, *B. subtilis*, and *L. monocytogenes* produce sex pheromones which are hydrophobic heptameric or octameric peptides derived from the C-terminal Lpp signal sequences (Flannagan and Clewell, 2002; Chandler and Dunny, 2004). These sex pheromones act as signals that facilitate the conjugative transfer of a specific category of plasmids referred to as pheromone-responsive plasmids (Wirth, 1994; Dunny and Leonard, 1997). Eep, a zinc-dependent membrane metalloprotease, has been shown to be involved in processing the Lpp precursor, leading to production of the active pheromone in most but not all pheromone systems (Chandler and Dunny, 2008).

Lpl Are a Special Class of Lpp Involved in Invasion

In *S. aureus* a special class of Lpp has been described that contributes to host cell invasion (Nguyen et al., 2015). These Lpp are called "lipoprotein-like" lipoproteins (Lpl). The corresponding genes are organized in tandem on a pathogenicity island. It was shown that Lpl induce invasion by host cells, the invasion process being mediated by the protein portion and not by the lipid moiety of Lpl (Nguyen et al., 2015). The increased invasion in murine skin and an increased bacterial burden in a murine kidney abscess model suggest that the lpl gene cluster serves as an important virulence factor. The protein part of Lpls' exerts a number of activities on host cells. One activity is that they delay G2/M phase transition in HeLa cells thus acting as cyclomodulins (Nguyen et al., 2016). Unlike the other staphylococcal cyclomodulins, Lpl1 (the model Lpl) shows no cytotoxicity even at high concentrations. A high conservation among Lpl proteins might be suggestive for common function. These observations may stir a speculation that a link exists between the exposure of Lpls on the bacterial surface and the delay in eukaryotic G2/M phase transition delay ultimately influencing host cell invasion processes.

In fact, the underlying mechanism of Lpl-triggered host cell invasion has been recently elucidated (Tribelli et al., 2019). The Lpl1 protein part, without the lipid moiety, binds directly to the isoforms of the human heat shock protein Hsp90-alpha and Hsp90-beta (Tribelli et al., 2019). Although the Hsp90-beta is constitutively expressed, the Hsp90-alpha isoform is heat-inducible and appears to play a major role in Lpl1 interaction. Lpl1-Hsp90 interaction induces F-actin formation, thus triggering an endocytosis-like internalization. The Lpl induced internalization represents a new host cell invasion principle (Tribelli et al., 2019). Pre-incubation of HaCaT cells at 39°C increased both the Hsp90-alpha expression and *S. aureus* invasion. It may be assumed that evolutionary adaption of *S. aureus* to the host has provided for protective mechanism of the microorganism against elevated temperatures in the host, i.e., fever. Above observation, namely the Lpl1/Hsp90-alpha mediated F-actin formation, now allows for the exciting hypothesis that enhanced bacterial invasion might provide a cellular niche for the microorganisms thus protection from host defense. Interestingly, Lpl proteins are mainly found in *S. aureus* and not in other staphylococcal species or other genera (Shahmirzadi et al., 2016).

Taken together, some Lpp are much conserved and present in almost all bacterial species. However, each species, and even each strain, contains its own specific Lpp, depending on its genetic make-up and habitat. The size of Lpp varies from <20 kDa up to >60 kDa which correlates with their function (**Table 3**). Some Lpp with many membrane spanning domains and some external and cytoplasmic loops like YidC are mainly localized in the membrane, while most others penetrate and span the cell wall and are exposed to the surface of the bacteria. Our knowledge about the function of Lpp in bacteria is incomplete and still about 30% of the Lpp have unknown functions, and to present it is unclear whether some of them are at all essential.

LPP LIPID STRUCTURE AND IMMUNITY ACTIVITY

Unlike other proteins, Lpp contains an unusual ether linked S-glyceryl-cysteine residue modified with three fatty acids (N-acyl-S-diacylglyceryl cysteine) at its N-terminus. This unique N-terminal lipid structure anchors Lpp in the bacterial membrane.

TABLE 3 | The correlation between the function and the size of Lpp.

Catalog Lpp by size	Function	Proteins
Large (more than 40 kDa)	**Transportation**	Peptide/oligopeptide transp. (≥50 kDa)
		Sugar transp. (40–50 kDa)
	Cell-wall	penicillin-binding protein
	Miscellaneous	Phosphatase (40 kDa), Copper resistant protein A (50 kDa),
Medium (20–30 kDa)	**Transportation**	Fe, other cation, anion and amino acid transp. (25–35 kDa)
	Germination	Spore germination protein
	Cell wall	Enzymes
	Miscellaneous	YidC, chaperon, foldase, peptidylprolyl isomerase, oxidase, hydrolase, nuclease, regulator of comK, CamS family sex pheromone protein
Small (<20 kDa)	**Cell-wall**	YqiH, LytA,
	Miscellaneous	Thioredoxin family protein, inhibitor, SCO family, cytochrome, superoxide dismutase

In Gram+ bacteria Lpp and/or lipopeptides are the main TLR2 agonists and they play a similar role as the lipopolysaccharides (LPS) in Gram-negative bacteria. In our previous review we raised the impact of Lpp in bacterial virulence and their function as TLR2 ligand (Nguyen and Götz, 2016). Here we concentrate on how Gram+ bacteria alter their lipid structures to modulate the host immune system. Gram-negative bacteria contain tri-acylated Lpp, however, the structure of the lipid moiety of Gram+ Lpp may vary from species to species (**Figure 3**).

N-Acetylation and Lyso-Form in Firmicutes

The enzymes involved in modification and processing of the Lpp in *E. coli* are well studied (Buddelmeijer, 2015). The third and last modification step is the N-acylation of diacylglyceride-cysteine by apolipoprotein N-acyltransferase (Lnt), resulting in mature N-acyl-S-diacylglyceryl-cysteine linked proteins (Buddelmeijer, 2015). While Lgt and Lsp are conserved in all bacterial species, Lnt has only been identified in proteobacteria and actinomycetes. Therefore, it has been assumed that in low GC Gram+ bacteria (e.g., firmicutes) Lpp are only diacylated (Stoll et al., 2005). However, with the development of GC-MS technology (Kurokawa et al., 2012) showed that Lpp from Gram+ bacteria may have different lipid structures. Diacylated Lpp with two O-acylated long-chain fatty acids are described in *L. monocytogenes* (Kurokawa et al., 2012). Triacylated Lpp with the N-terminal cysteine being N-acylated with a long chain fatty acid is described in *S. aureus*, *M. tuberculosis* and in most Gram-negative bacteria (Kurokawa et al., 2009; Tschumi et al., 2009;

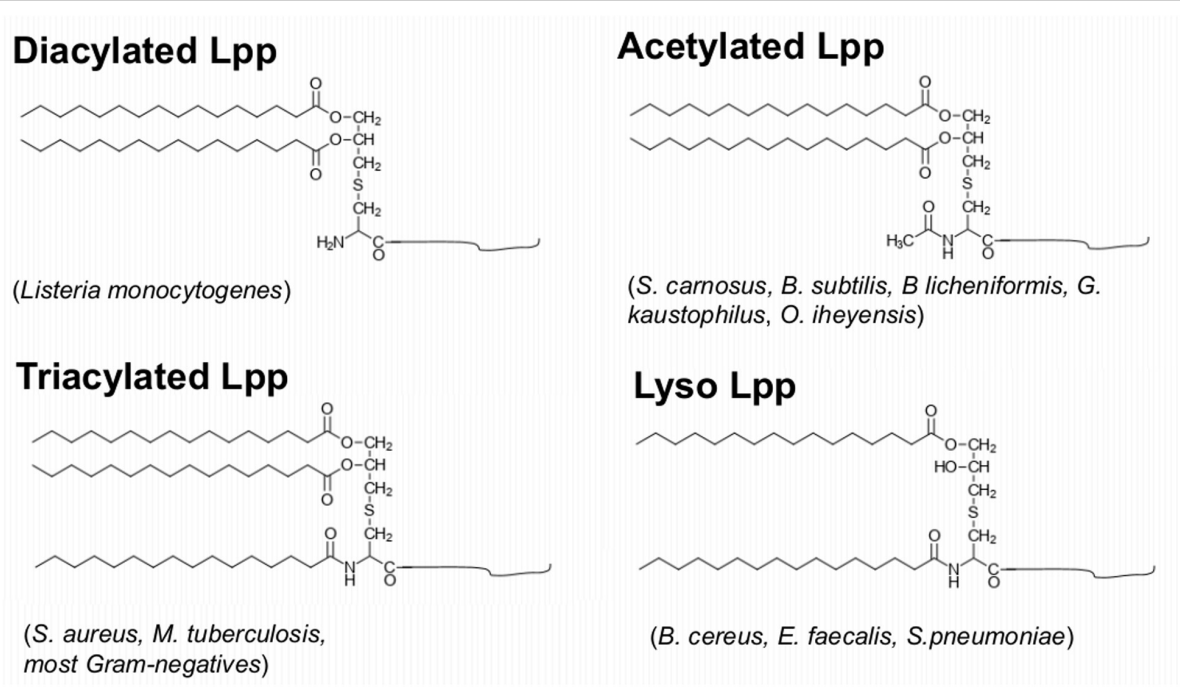


FIGURE 3 | Variation of the Lpp's lipid structures in Gram+ bacteria. Diacylated Lpp contain a modified S-diacyl-glyceryl cysteine which is found for example in *Listeria monocytogenes*; triacylated Lpp contain a modified N-acyl-S-diacyl-glyceryl cysteine which were found in *S. aureus*, *M. tuberculosis*, and most Gram-negatives; N-acetylated Lpp contain N-acetyl-S-diacyl-glyceryl cysteine which were found in *S. carnosus*, *B. subtilis*, *B. licheniformis*, *G. kaustophilus*, *O. iheyensis*; Lyso Lpp contain a modified N-acyl-S-monoacyl cysteine which were found in *B. cereus*, *E. faecalis*, or *S. pneumoniae*.

Buddelmeijer, 2015). An N-acetylated lipid structure was found in *S. carnosus*, *B. subtilis*, *B licheniformis*, *G. kaustophilus*, and *O. iheyensis* (Kurokawa et al., 2012; Kumari et al., 2017). Finally, Lyso-Lpp structures that lack the O-acyl group in A-2 position have been found in *B. cereus*, *E. faecalis* and *S. pneumoniae* (Kurokawa et al., 2012; Armbruster and Meredith, 2017).

These observations raise the important question how firmicutes which do not process a Lnt homolog carry out N-acylation of their Lpp. Recently, it was found that in *S. aureus* two genes/enzymes are involved in N-acylation of Lpp (Gardiner et al., 2020). They are referred to as N-acylation transferase system (Lns). LnsA is an NlpC/P60 superfamily enzyme while LnsB has remote homology to the CAAX protease and to the bacteriocin-processing enzyme (CPBP) family. With either both LnsA and LnsB being necessary, or one enzyme alone being sufficient for N-acylation in *S. aureus,* they convert the Lpp chemotype from the diacylated to the tri-acylated form when heterologously expressed in *L. monocytogenes*.

As shown in **Figure 3** *E. faecalis* and *B. cereus* produce Lyso-Lpp. The Lyso-Lpp is created by a new enzyme, the Lpp intramolecular transacylase (Lit) (Armbruster and Meredith, 2017). The discovery of lipid modifying enzymes in Lpp of Gram+ bacteria is an important step forward in analyzing the consequences of the various lipid structures for the immune modulation. The biosynthesis modification of Lpp in firmicutes is illustrated in **Figure 4**.

Gram+ Bacteria Differ in the Lipid Structure and Immune Modulation

In recent years it has been shown that the lipid structure exerts a tremendous influence on the immune response both of the innate and the adaptive system. A number of studies have been conducted with various staphylococcal species, the majority being skin residents and commensals, while others, like *S. aureus* may cause infections and abscesses. In a study on cutaneous innate immune sensing diacylated and triacylated lipopeptides were compared (Skabytska et al., 2014). In a mouse model diacylated Lpp was shown to induce a much higher inflammatory response than triacylated Lpp. On the other hand through the massive production of interleukin-6 (IL-6), diacylated Lpp induced the activation Gr1(+)CD11b(+) myeloid-derived suppressor cells (MDSCs) which are recruited to the skin. The recruited MDSCs suppress T cell-mediated recall responses such as dermatitis thus dampening the overshooting immune reaction. Therefore, certain Lpp on the skin may counteract inflammation by suppressing immune responses via activation of MDSCs (Skabytska et al., 2014). This study is one of the first reports showing that the different lipid structures of Lpp induce either high or low immune stimulation.

In fact, this insight has recently showed when comparing the immune response of the commensal *S. aureus* and *S. epidermidis* with that of the non-commensal *S. carnosus* species (Nguyen et al., 2017). Structural analysis of the lipid moiety of representatives of the three species revealed that the N-terminus of the lipid moiety is acylated with a long-chain fatty acid (C17) in *S. aureus* and *S. epidermidis*, while it is only acetylated with a short-chain fatty acid (C2) in *S. carnosus*. Consequently, *S. carnosus* triggered a much higher immune response than *S. aureus* and *S. epidermidis*. The difference in immune response is comparable to the immune response elicited by tri- and di-acylated lipopeptides, respectively.

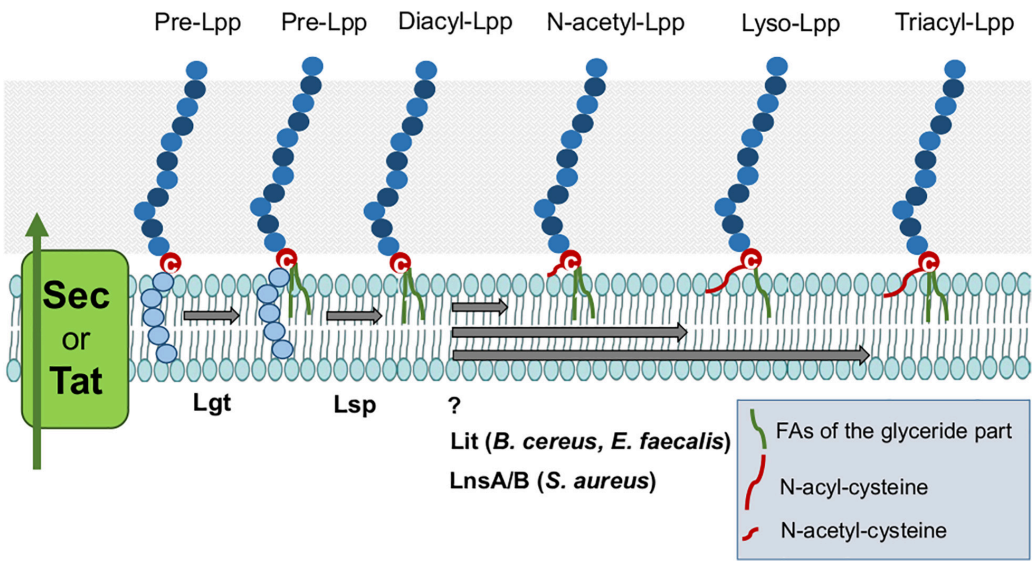


FIGURE 4 | Lipoproteins biosynthesis modification in Firmicutes. After Sec/Tat dependent export, precursor Lpp are located on the cytoplasmic membrane where they can be modified by processing enzymes to yield the mature Lpp. The first enzyme Lgt transfers the diacylglyceryl group from phosphatidylglycerol to the sulfhydryl group of the invariant cysteine residue in the lipobox. The second enzyme Lsp recognizes the diacylglyceryl modified lipo signal peptide and cleaves between the amino acid at position –1 and the lipid-modified cysteine residue at +1, obtaining the diacyl-Lpp. N-acetyl-Lpp are modified by unknown enzyme. Lyso-Lpp are modified by Lit enzyme. And Tri-acyl (N-acyl)-Lpp are processed by LnsAB in *S. aureus*.

The short-chain fatty acid (C2) structure in *S. carnosus* is apparently so small that it behaves similar as a diacylated lipopeptide. This result is in so far important as it also points in favor of the assumption that the Lpp lipid structure is involved in tolerance vs. non-tolerance by our immune system – or otherwise said, through their lipid structures bacterial species may establish their ecological niche as commensals. As *S. carnosus* induces such a fulminant immune reaction it is not surprising that it never has been observed as human commensal.

Although it is well accepted that the di- and tri-acylation of the lipid structure of Lpp modulate our immune system differently, there is no information of how the length of O-acylated fatty acids or the Lyso-Lpp modulate our immune system; this area still requires additional elucidation.

HOW FIRMLY ARE LPP ANCHORED AND HOW ARE THEY RELEASED

There is evidence in Gram-negative bacteria that amino acids next to the cysteine residue (+1) in mature Lpp have an impact on the strength of the anchoring of Lpp in the membrane. In *E. coli*, the amino acid at position +2 determines the localization of Lpp either in the inner or outer membrane (Gennity et al., 1992). In Gram-negative bacteria, the ABC transporter LolCDE complex translocates outer membrane-specific lipoproteins (Lpp) from the inner membrane to the outer membrane. However, Lpp possessing an aspartate (Asp) at position +2 are not translocated because this residue functions as a LolCDE avoidance signal (Fukazawa et al., 2010; Narita and Tokuda, 2011). Since Gram+ bacteria lack the typical outer membrane and the Lol system, Lpp are anchored only in the cytoplasmic membrane by integrating their lipid portion. Investigating whether a similar effect could be observed also in Gram+ bacteria revealed that in *S. aureus* Asp at position +2 plays also a role in withholding Lpp to the cytoplasmic membrane. A replacement of Asp+2 by Ser caused a higher release of the corresponding Lpp (Kumari et al., 2017). It is assumed that Asp in position +2 and adjacent amino acids contribute in tightening the anchoring of Lpp by interaction of the negatively charged Asp with the positively charged Lys-PG (Kumari et al., 2017). In mycobacteria, some Lpp may also be anchored in the mycolic layer which were recently considered as the outer membrane (Bansal-Mutalik and Nikaido, 2014; Becker and Sander, 2016). Even though no Lol homologs could be identified yet, protein-protein interaction studies suggested that two Lpp, LppK and LppI, could be transported by the lipid transport protein LprG (Touchette et al., 2017). Furthermore, the study showed that LprG interacts with the mycolyl-transferase Ag85A, and that a deletion of either lprG or ag85A gene affects growth and mycolylation (Touchette et al., 2017).

Lipoproteins are not exclusively anchored in the bacterial membrane, to a certain extent they are also released in their surrounding environment. The precise mechanism how Lpp are released into the environment is still under investigation. Recently, it was demonstrated that they are released as part of bacterial membrane vesicles (BMVs). One of the first vesicle-like blebbing structure in Gram+ bacteria was reported on the surface of *Bacillus* (Dorward and Garon, 1990). Now we know that such membrane vesicles (MVs) are produced by all domains of life. Bacterial membrane vesicles (BMVs) have been studied in *Clostridium perfringens*, *Listeria monocytogenes*, *Mycobacterium tuberculosis*, *Streptococcus pyogenes*, *Staphylococcus aureus* or *Streptomyces* spp. BMVs carry quite heterogenous containing membranous components, nucleic acids, toxins, enzymes and Lpp which contribute to the manifold activities of such BMVs in microbial physiology and pathogenesis as reviewed by Brown et al., 2015.

Streptococcus pyogenes releases Lpp after treatment of the cells with sublethal concentrations of penicillin in form of Lpp-rich BMVs (Biagini et al., 2015). Penicillin weakens the bacterial cell wall. Interestingly, lipid and proteomic analysis of the vesicles revealed that they were enriched of phosphatidylglycerol and almost exclusively composed of Lpp. This example shows that under antibiotic treatment of a patient BMVs can be formed and may influence the course of infection.

After invasion of macrophages *M. tuberculosis* also forms BMVs loaded with lipoglycans and Lpp (Athman et al., 2015). BMVs are here regarded as the primary means how *M. tuberculosis* exports lipoglycans and Lpp to impair effector functions of infected macrophages and circulate bacterial components beyond the site of infection, in order to regulate immune responses or iron acquisition. In the regulation of BMV production and vesiculogenesis there are various genes involved, including *virR*, the Pst/SenX3-RegX3 signal transduction system and the ESX-5 protein secretion system (White et al., 2018). One of the *M. tuberculosis* extracellular vesicle-associated Lpp (LpqH) can be used as a potential biomarker to distinguish *M. paratuberculosis* infection or vaccination from infection due to *M. tuberculosis* (Palacios et al., 2019).

In *S. aureus* the release of Lpp is boosted by the production of phenol-soluble modulin (PSM) peptides, which act as membrane detergents and surfactants (Hanzelmann et al., 2016). They damage both bacterial and host membranes. The damage of the bacterial cytoplasmic membrane also causes the release of Lpp which enhances the TLR2-mediated innate immune response. Thus, PSM surfactants in *S. aureus*, mainly the alpha PSMs, exert a similar effect as penicillin in *S. pyogenes*. Whenever, the membrane is damaged directly by PSMs or indirectly e.g., by ß-lactam antibiotics, one can expect a release of membrane-bound proteins like Lpp. However, it was also found that PSMs promote the release of BMVs from the cytoplasmic membrane which contains high amounts of Lpp (Schlatterer et al., 2018). There is evidence that the bacterial turgor is the driving force for vesicle budding under hypotonic osmotic conditions. The underlying mechanism of the release of Lpp and BMVs is subject of ongoing research. Many questions arise: for instance, if Lpp are primarily imbedded within the BMV membrane or anchored with their lipid part in the membrane, how can they induce TRL2 response. Overall, the influence of Gram+ BMVs on health and disease has

by far not yet been fully clarified (Liu et al., 2018) leaving room for pertinent and exciting research.

APPLICATION OF LPP AS ADJUVANTS AND VACCINE

Lipoproteins play a role as adjuvant and as antigens for vaccine development. Lpp and their synthetic analogs are strong immune modulators of the early host responses after infection. Synthetic lipopeptides are strong adjuvants for the adaptive immune system. Particularly the water-soluble lipohexapeptide Pam3Cys-Ser-$(Lys)_4$ (PCSL) triggers a high IL-8 induction via TLR2 and constitutes a potent immune adjuvant and induces virus specific CD8(+) T cells in mice when covalently coupled to a synthetic peptide (Schild et al., 1991; Spohn et al., 2004).

Freund's complete adjuvant (FCS) can be replaced by some specific lipopeptides (e.g., PCSL) with the antigen specific IgM response after 7 days, and an ensuing IgG response after 14 days (Reitermann et al., 1989). Among other antigens, inactivated coronavirus strain 800 was also tested (Hofmann et al., 1996). Only after the first booster immunization did the content of specific IgY antibodies in the sera increase significantly, especially in the animals first immunized using PCSL as adjuvants (Hofmann et al., 1996). Although this work dates back almost 25 years, the positive effect of PCSL on the formation of corona virus specific antibodies shows how seminal this study is still today. These results illustrate that synthetic lipopeptides could be extremely valuable for preventive or therapeutic use. Indeed, recently a new water-soluble synthetic Pam3Cys-GDPKHPKSF named XS15 has been designed. Tolerability and immune responses of XS15 together with an antigen were closely monitored (in a single human volunteer). A granuloma formed at the injection site in which activated and functional CD4(+) and CD8(+) effector memory T cells had accumulated and were still detectable after one year. The lipopeptide XS15 was considered a promising adjuvant for tumor peptide vaccination (Rammensee et al., 2019).

Since the protein portion of many Lpp spans the cell wall and is therefore accessible for antibodies, it is obvious to use them for vaccine development. For the therapy of *Enterococcus* infections two metal binding Lpp were tested as vaccine candidates, namely the manganese ABC transporter substrate-binding Lpp (PsaAfm), and the zinc ABC transporter substrate-binding Lpp (AdcAfm) (Romero-Saavedra et al., 2015). The two antigens elicited specific, opsonic and protective antibodies, with an extensive cross-reactivity and serotype-independent coverage. The use of recombinant Lpp with built-in immunostimulating properties for novel subunit vaccine development was reviewed recently (Leng et al., 2015). Promising results were obtained with a dengue subunit vaccine, with a novel subunit vaccine against *Clostridium difficile*-associated diseases and with HPV-based immunotherapeutic vaccines. Against *S. aureus* a cocktail of five conserved antigens was tested containing two Lpp, i.e., the ferric hydroxamate-binding Lpp FhuD2 and the putative Lpp named conserved staphylococcal antigen 1A (Csa1A) as surface-exposed antigens (Bagnoli et al., 2015). These data demonstrated that the rational selection of mixtures of conserved antigens combined with Th1/Th17 adjuvants may lead to promising vaccine formulations against *S. aureus*. Remarkably, a breakthrough was achieved when these cocktail antigens were used together with a TLR7 agonist adsorbed to alum as adjuvants (Bagnoli et al., 2015). Altogether, the potential for the future use of Lpp in vaccination is still by far not fully appreciated even though these exciting molecules provide an extremely promising basis with the respect in vaccine development.

CONCLUSION

A total of 50 years have passed since first discovery of the Braun's Lpp in *E. coli*. In the meantime, we know much more about the biosynthesis of Lpp and their functions. We know that Lpp occur in all Gram+ bacteria. The tasks they perform in the respective species are extremely diverse. Some of the described Lpp are absolutely essential for the life of the bacteria. Many Lpp contribute to fitness and become essential under certain nutrient limiting conditions. Lpp consist of the lipid and the protein part. Both have an important but completely different function and activity. While the protein part can have the most diverse binding and enzymatic activities, the lipid part together with a small peptide residue modulates the immune system. Yet, despite this enormous insight in structure and function of Lpp, there are still large gaps in our knowledge, for example, approximately 30% of Lpp in Gram+ bacteria are still uncharacterized with respect to function. Moreover, the ascertainment that the lipid structure of Gram+ bacteria in particular is so variable has only come to light in recent years, a fact largely owed to the methodical challenges of lipid structure analyses. The discovery of new lipid variants, their activity on the immune system, as well as the search for the modifying enzymes will thus be important aims for research during the upcoming years, and it is strongly expected that ensuing findings will allow for opening novel avenues in the use of Lpp in diagnostic, preventive or therapeutic medicine.

IN HONOR OF VOLKMAR BRAUN THE DISCOVERER OF Lpp

This review article is also in honor of Volkmar Braun at the University Tübingen, who discovered bacterial lipoproteins (Lpp) almost 50 years ago. His discovery of the Lpp was based on accurate scientific work, careful analysis of the results and drawing the right conclusions. He himself describes the first observations which finally led to the discovery of "Braun's lipoprotein" as follows:

"In attempts to isolate membrane proteins from *E. coli*, isolated cell walls were treated with trypsin which caused the fastest decrease of the optical density among all tested proteases. Murein (peptidoglycan) from trypsin-treated cell walls, isolated with hot 4% sodium dodecyl sulfate was free of protein but contained additional lysine and arginine residues which presumably were left-overs from trypsin cleavage (Braun and Rehn, 1969).

The second key observation was the peptide Arg-Lys-Dpm in an acid hydrolysate which showed the attachment site of the protein at Dpm (diaminopimelate) of the murein peptide side chain (Braun and Bosch, 1972). Six amino acids commonly found in proteins were lacking in the protein. This data pointed to a protein of a defined size and composition covalently attached to the murein via the Arg-Lys bond.

Another highlight was the determination of the unique lipid structure covalently linked to the protein (Hantke and Braun, 1973), which is now predicted to be contained in approximately 2.7% of bacterial membrane proteins. Since nearly one million Lpp molecules are evenly distributed over the murein net, one can deduce a model of the Lpp organization in the outer membrane. Many functions were assigned to Lpp, as this review lucidly demonstrates, but most of them are not even touched.

Before the discovery of the major *E. coli* Lpp, the term lipoprotein was used for undefined mixtures of aggregated proteins with all sorts of lipids. No Lpp and even no membrane protein had been sequenced. The Lpp was also the first protein whose heptad repeat sequence fulfilled the prediction of Francis Crick in 1953 of coiled-coil structures."

AUTHOR CONTRIBUTIONS

M-TN and FG wrote the manuscript. MM, SN, and MH contributed to revise the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by the Deutsche Forschungsgemeinschaft (DFG) SFB766/CRC/Transregio 261 to FG, and by infrastructural funding from the Deutsche Forschungsgemeinschaft (DFG), Cluster of Excellence EXC 2124 Controlling Microbes to Fight Infections.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2020.582582/full#supplementary-material

REFERENCES

Ajdic, D., Mcshan, W. M., Mclaughlin, R. E., Savic, G., Chang, J., Carson, M. B., et al. (2002). Genome sequence of *Streptococcus mutans* UA159, a cariogenic dental pathogen. *Proc. Natl. Acad. Sci. U.S.A.* 99, 14434–14439. doi: 10.1073/pnas.172501299

Anantharaman, V., and Aravind, L. (2003). Evolutionary history, structural features and biochemical diversity of the NlpC/P60 superfamily of enzymes. *Genome Biol.* 4:R11.

Archibald, F. (1986). Manganese: its acquisition by and function in the lactic acid bacteria. *Crit. Rev. Microbiol.* 13, 63–109. doi: 10.3109/10408418609108735

Armbruster, K. M., and Meredith, T. C. (2017). Identification of the Lyso-Form N-Acyl intramolecular transferase in Low-GC firmicutes. *J. Bacteriol.* 199:e00099-17.

Athman, J. J., Wang, Y., Mcdonald, D. J., Boom, W. H., Harding, C. V., and Wearsch, P. A. (2015). Bacterial membrane vesicles mediate the release of *Mycobacterium tuberculosis* lipoglycans and lipoproteins from infected macrophages. *J. Immunol.* 195, 1044–1053. doi: 10.4049/jimmunol.1402894

Babu, M. M., Priya, M. L., Selvan, A. T., Madera, M., Gough, J., Aravind, L., et al. (2006). A database of bacterial lipoproteins (DOLOP) with functional assignments to predicted lipoproteins. *J. Bacteriol.* 188, 2761–2773. doi: 10.1128/jb.188.8.2761-2773.2006

Bagnoli, F., Fontana, M. R., Soldaini, E., Mishra, R. P., Fiaschi, L., Cartocci, E., et al. (2015). Vaccine composition formulated with a novel TLR7-dependent adjuvant induces high and broad protection against Staphylococcus aureus. *Proc. Natl. Acad. Sci. U.S.A.* 112, 3680–3685. doi: 10.1073/pnas.1424924112

Bansal-Mutalik, R., and Nikaido, H. (2014). Mycobacterial outer membrane is a lipid bilayer and the inner membrane is unusually rich in diacyl phosphatidylinositol dimannosides. *Proc. Natl. Acad. Sci. U.S.A.* 111, 4958–4963. doi: 10.1073/pnas.1403078111

Basavanna, S., Chimalapati, S., Maqbool, A., Rubbo, B., Yuste, J., Wilson, R. J., et al. (2013). The effects of methionine acquisition and synthesis on *Streptococcus pneumoniae* growth and virulence. *PLoS One* 8:e49638. doi: 10.1371/journal.pone.0049638

Becker, K., and Sander, P. (2016). Mycobacterium tuberculosis lipoproteins in virulence and immunity - fighting with a double-edged sword. *FEBS Lett.* 590, 3800–3819. doi: 10.1002/1873-3468.12273

Biagini, M., Garibaldi, M., Aprea, S., Pezzicoli, A., Doro, F., Becherelli, M., et al. (2015). The human pathogen *Streptococcus pyogenes* releases lipoproteins as lipoprotein-rich membrane vesicles. *Mol. Cell. Proteomics* 14, 2138–2149. doi: 10.1074/mcp.m114.045880

Biswas, L., Biswas, R., Nerz, C., Ohlsen, K., Schlag, M., Schafer, T., et al. (2009). Role of the twin-arginine translocation pathway in Staphylococcus. *J. Bacteriol.* 191, 5921–5929. doi: 10.1128/jb.00642-09

Braun, V. (1975). Covalent lipoprotein from the outer membrane of *Escherichia coli*. *Biochim. Biophys. Acta* 415, 335–377. doi: 10.1016/0304-4157(75)90013-1

Braun, V., and Bosch, V. (1972). Repetitive sequences in the murein-lipoprotein of the cell wall of *Escherichia coli*. *Proc. Natl. Acad. Sci. U.S.A.* 69, 970–974. doi: 10.1073/pnas.69.4.970

Braun, V., and Hantke, K. (2019). Lipoproteins: structure, function, biosynthesis. *Subcell Biochem.* 92, 39–77. doi: 10.1007/978-3-030-18768-2_3

Braun, V., and Rehn, K. (1969). Chemical characterization, spatial distribution and function of a lipoprotein (murein-lipoprotein) of the E. coli cell wall. The specific effect of trypsin on the membrane structure. *Eur. J. Biochem.* 10, 426–438. doi: 10.1111/j.1432-1033.1969.tb00707.x

Bray, B. A., Sutcliffe, I. C., and Harrington, D. J. (2009). Expression of the MtsA lipoprotein of *Streptococcus agalactiae* A909 is regulated by manganese and iron. *Antonie Van Leeuwenhoek* 95, 101–109. doi: 10.1007/s10482-008-9291-6

Brown, L., Wolf, J. M., Prados-Rosales, R., and Casadevall, A. (2015). Through the wall: extracellular vesicles in Gram-positive bacteria, mycobacteria and fungi. *Nat. Rev. Microbiol.* 13, 620–630. doi: 10.1038/nrmicro3480

Buddelmeijer, N. (2015). The molecular mechanism of bacterial lipoprotein modification–how, when and why? *FEMS Microbiol. Rev.* 39, 246–261. doi: 10.1093/femsre/fuu006

Chandler, J. R., and Dunny, G. M. (2004). Enterococcal peptide sex pheromones: synthesis and control of biological activity. *Peptides* 25, 1377–1388. doi: 10.1016/j.peptides.2003.10.020

Chandler, J. R., and Dunny, G. M. (2008). Characterization of the sequence specificity determinants required for processing and control of sex pheromone by the intramembrane protease Eep and the plasmid-encoded protein PrgY. *J. Bacteriol.* 190, 1172–1183. doi: 10.1128/jb.01327-07

Chen, J., Gai, Y., Fu, G., Zhou, W., Zhang, D., and Wen, J. (2015). Enhanced extracellular production of alpha-amylase in Bacillus subtilis by optimization of regulatory elements and over-expression of PrsA lipoprotein. *Biotechnol. Lett.* 37, 899–906. doi: 10.1007/s10529-014-1755-3

Dalbey, R. E., Kuhn, A., Zhu, L., and Kiefer, D. (2014). The membrane insertase YidC. *Biochim. Biophys. Acta* 1843, 1489–1496.

Dale, S. E., Sebulsky, M. T., and Heinrichs, D. E. (2004). Involvement of SirABC in iron-siderophore import in *Staphylococcus aureus*. *J. Bacteriol.* 186, 8356–8362. doi: 10.1128/jb.186.24.8356-8362.2004

DeLano, W. L. (2009). "PyMOL molecular viewer: updates and refinements," in *Proceedings of the 238th ACS National Meeting* Washington, DC, 238.

Diaz, O. R., Sayer, C. V., Popham, D. L., and Shen, A. (2018). Clostridium difficile lipoprotein GerS is required for cortex modification and thus spore germination. *mSphere* 3, 2379–5042.

Diep, B. A., Gill, S. R., Chang, R. F., Phan, T. H., Chen, J. H., Davidson, M. G., et al. (2006). Complete genome sequence of USA300, an epidemic clone of community-acquired meticillin-resistant *Staphylococcus aureus*. *Lancet* 367, 731–739. doi: 10.1016/s0140-6736(06)68231-7

Diep, B. A., Phung, Q., Date, D. Arnott, Bakalarski, C., Xu, M., et al. (2014). Identifying potential therapeutic targets of methicillin-resistant *Staphylococcus aureus* through in vivo proteomic analysis. *J. Infect. Dis.* 209, 1533–1541. doi: 10.1093/infdis/jit662

Donaldson, J. R., Nanduri, B., Burgess, S. C., and Lawrence, M. L. (2009). Comparative proteomic analysis of *Listeria monocytogenes* strains F2365 and EGD. *Appl. Environ. Microbiol.* 75, 366–373. doi: 10.1128/aem.01847-08

Dorward, D. W., and Garon, C. F. (1990). DNA is packaged within membrane-derived vesicles of gram-negative but not gram-positive bacteria. *Appl. Environ. Microbiol.* 56, 1960–1962. doi: 10.1128/aem.56.6.1960-1962.1990

Dunny, G. M., and Leonard, B. A. (1997). Cell-cell communication in gram-positive bacteria. *Annu. Rev. Microbiol.* 51, 527–564. doi: 10.1146/annurev.micro.51.1.527

Durmort, C., Ercoli, G., Ramos-Sevillano, E., Chimalapati, S., Haigh, R. D., De Ste Croix, M., et al. (2020). Deletion of the zinc transporter lipoprotein AdcAII causes hyperencapsulation of streptococcus pneumoniae associated with distinct alleles of the Type I restriction-modification system. *mBio* 11, e00445–20. doi: 10.1128/mBio.00445-20

Fenn, K., Wong, C. T., and Darbari, V. C. (2019). *Mycobacterium tuberculosis* uses mce proteins to interfere with host cell signaling. *Front. Mol. Biosci.* 6:149. doi: 10.3389/fmolb.2019.00149

Ferretti, J. J., Mcshan, W. M., Ajdic, D., Savic, D. J., Savic, G., Lyon, K., et al. (2001). Complete genome sequence of an M1 strain of *Streptococcus pyogenes*. *Proc. Natl. Acad. Sci. U.S.A.* 98, 4658–4663.

Fischer, K. E., and Bremer, E. (2012). Activity of the osmotically regulated yqiHIK promoter from *Bacillus subtilis* is controlled at a distance. *J. Bacteriol.* 194, 5197–5208. doi: 10.1128/jb.01041-12

Flannagan, S. E., and Clewell, D. B. (2002). Identification and characterization of genes encoding sex pheromone cAM373 activity in *Enterococcus faecalis* and *Staphylococcus aureus*. *Mol. Microbiol.* 44, 803–817. doi: 10.1046/j.1365-2958.2002.02922.x

Fleischmann, R. D., Alland, D., Eisen, J. A., Carpenter, L., White, O., and Peterson, J. (2002). Whole-genome comparison of *Mycobacterium tuberculosis* clinical and laboratory strains. *J. Bacteriol.* 184, 5479–5490.

Fukazawa, H., Ikeda, Y., Fukuyama, M., Suzuki, T., Hori, H., Okuda, T., et al. (2010). The resorcylic acid lactone hypothemycin selectively inhibits the mitogen-activated protein kinase kinase-extracellular signal-regulated kinase pathway in cells. *Biol. Pharm. Bull.* 33, 168–173. doi: 10.1248/bpb.33.168

Gardiner, J. H., Komazin, G., Matsuo, M., Cole, K., Götz, F., and Meredith, T. C. (2020). Lipoprotein N-acylation in *Staphylococcus aureus* is catalyzed by a two component acyl transferase system. *mBio* 11, e01619–20. doi: 10.1128/mBio.01619-20

Geiger, C., Korn, S. M., Hasler, M., Peetz, O., Martin, J., Kotter, P., et al. (2019). LanI-mediated lantibiotic immunity in *Bacillus subtilis*: functional analysis. *Appl. Environ. Microbiol.* 85, e00534–19. doi: 10.1128/AEM.00534-19

Gennity, J. M., Kim, H., and Inouye, M. (1992). Structural determinants in addition to the amino-terminal sorting sequence influence membrane localization of *Escherichia coli* lipoproteins. *J. Bacteriol.* 174, 2095–2101. doi: 10.1128/jb.174.7.2095-2101.1992

Glaser, P., Frangeul, L., Buchrieser, C., Rusniok, C., Amend, A., and Baquero, F. (2001). Comparative genomics of Listeria species. *Science* 294, 849–852.

Glaser, P., Rusniok, C., Buchrieser, C., Chevalier, F., Frangeul, L., Msadek, T., et al. (2002). Genome sequence of *Streptococcus agalactiae*, a pathogen causing invasive neonatal disease. *Mol. Microbiol.* 45, 1499–1513. doi: 10.1046/j.1365-2958.2002.03126.x

Gothel, S. F., and Marahiel, M. A. (1999). Peptidyl-prolyl cis-trans isomerases, a superfamily of ubiquitous folding catalysts. *Cell Mol. Life Sci.* 55, 423–436. doi: 10.1007/s000180050299

Götz, F., Elstner, E. F., Sedewitz, B., and Lengfelder, E. (1980). Oxygen utilization by *Lactobacillus plantarum*. II. Superoxide and superoxide dismutation. *Arch. Microbiol.* 125, 215–220. doi: 10.1007/bf00446879

Götz, F., and Mayer, S. (2013). Both terminal oxidases contribute to fitness and virulence during organ-specific *Staphylococcus aureus* colonization. *mBio* 4:e00976-13.

Guo, L., Wu, T., Hu, W., He, X., Sharma, S., Webster, P., et al. (2013). Phenotypic characterization of the foldase homologue PrsA in *Streptococcus mutans*. *Mol. Oral Microbiol.* 28, 154–165. doi: 10.1111/omi.12014

Hacker, C., Christ, N. A., Duchardt-Ferner, E., Korn, S., Gobl, C., Berninger, L., et al. (2015). The solution structure of the lantibiotic immunity protein NisI and its interactions with nisin. *J. Biol. Chem.* 290, 28869–28886. doi: 10.1074/jbc.m115.679969

Hammer, N. D., Reniere, M. L., Cassat, J. E., Zhang, Y., Hirsch, A. O., Indriati Hood, M., et al. (2013). Two heme-dependent terminal oxidases power *Staphylococcus aureus* organ-specific colonization of the vertebrate host. *mBio* 4:e00241-13.

Hantke, K., and Braun, V. (1973). Covalent binding of lipid to protein. Diglyceride and amide-linked fatty acid at the N-terminal end of the murein-lipoprotein of the *Escherichia coli* outer membrane. *Eur. J. Biochem.* 34, 284–296. doi: 10.1111/j.1432-1033.1973.tb02757.x

Hanzelmann, D., Joo, H. S., Franz-Wachtel, M., Hertlein, T., Stevanovic, S., Macek, B., et al. (2016). Toll-like receptor 2 activation depends on lipopeptide shedding by bacterial surfactants. *Nat. Commun.* 7:12304.

Heikkinen, O., Seppala, R., Tossavainen, H., Heikkinen, S., Koskela, H., Permi, P., et al. (2009). Solution structure of the parvulin-type PPIase domain of *Staphylococcus aureus* PrsA–implications for the catalytic mechanism of parvulins. *BMC Struct. Biol.* 9:17. doi: 10.1186/1472-6807-9-17

Hiron, A., Borezee-Durant, E., Piard, J. C., and Juillard, V. (2007). Only one of four oligopeptide transport systems mediates nitrogen nutrition in *Staphylococcus aureus*. *J. Bacteriol.* 189, 5119–5129. doi: 10.1128/JB.00274-07

Hoffmann, A., Schneider, T., Pag, U., and Sahl, H. G. (2004). Localization and functional analysis of PepI, the immunity peptide of Pep5-producing Staphylococcus epidermidis strain 5. *Appl. Environ. Microbiol.* 70, 3263–3271. doi: 10.1128/aem.70.6.3263-3271.2004

Hofmann, A., Erhard, M., Schmidt, P., Bessler, W., Wiesmuller, K.-H., Zinsmeister, P., et al. (1996). Lipopeptide als Adjuvantien bei der immunisierung yon legehennen lipopeptides as adjuvants for the immunisation of laying hens. *ALTEX* 13(Suppl. 96), 26–29.

Horsburgh, M. J., Wharton, S. J., Cox, A. G., Ingham, E., Peacock, S., and Foster, S. J. (2002). MntR modulates expression of the PerR regulon and superoxide resistance in *Staphylococcus aureus* through control of manganese uptake. *Mol. Microbiol.* 44, 1269–1286. doi: 10.1046/j.1365-2958.2002.02944.x

Ivanova, N., Sorokin, A., Anderson, I., Galleron, N., Candelon, B., and Kapatral, V. (2003). The complete genome sequence of the gram-positive bacterium *Bacillus subtilis*. *Nature* 390, 249–256.

Jeong, J. H., and Ha, S. C. (2018). Crystal structure of NisI in a lipid-free form, the nisin immunity protein, from lactococcus lactis. *Antimicrob. Agents Chemother.* 62:AAC.1966-17.

Jousselin, A., Manzano, C., Biette, A., Reed, P., Pinho, M. G., Rosato, A. E., et al. (2015). The *Staphylococcus aureus* chaperone PrsA is a new auxiliary factor of oxacillin resistance affecting penicillin-binding protein 2A. *Antimicrob. Agents Chemother.* 60, 1656–1666. doi: 10.1128/aac.02333-15

Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N., and Sternberg, M. J. E. (2015). The Phyre2 web portal for protein modeling, prediction and analysis. *Nat. Protoc.* 10, 845–858. doi: 10.1038/nprot.2015.053

Kempf, B., Gade, J., and Bremer, E. (1997). Lipoprotein from the osmoregulated ABC transport system OpuA of *Bacillus subtilis*: purification of the glycine betaine binding protein and characterization of a functional lipidless mutant. *J. Bacteriol.* 179, 6213–6220. doi: 10.1128/jb.179.20.6213-6220.1997

Koide, A., and Hoch, J. A. (1994). Identification of a second oligopeptide transport system in *Bacillus subtilis* and determination of its role in sporulation. *Mol. Microbiol.* 13, 417–426. doi: 10.1111/j.1365-2958.1994.tb00436.x

Kontinen, V. P., and Sarvas, M. (1993). The PrsA lipoprotein is essential for protein secretion in *Bacillus subtilis* and sets a limit for high-level secretion. *Mol. Microbiol.* 8, 727–737. doi: 10.1111/j.1365-2958.1993.tb01616.x

Kumari, N., Götz, F., and Nguyen, M. T. (2017). Aspartate tightens the anchoring of staphylococcal lipoproteins to the cytoplasmic membrane. *MicrobiologyOpen* 6:e00525. doi: 10.1002/mbo3.525

Kunst, F., Ogasawara, N., Moszer, I., Albertini, A. M., Alloni, G., Azevedo, V., et al. (1997). The complete genome sequence of the gram-positive bacterium *Bacillus subtilis*. *Nature* 390, 249–256. doi: 10.1038/36786

Kurokawa, K., Kim, M. S., Ichikawa, R., Ryu, K. H., Dohmae, N., Nakayama, H., et al. (2012). Environment-mediated accumulation of diacyl lipoproteins over their triacyl counterparts in *Staphylococcus aureus*. *J. Bacteriol.* 194, 3299–3306. doi: 10.1128/jb.00314-12

Kurokawa, K., Lee, H., Roh, K. B., Asanuma, M., Kim, Y. S., Nakyama, H., et al. (2009). The triacylated ATP binding cluster transporter substrate-binding lipoprotein of *Staphylococcus aureus* functions as a native ligand for the toll-like receptor 2. *J. Biol. Chem.* 284, 8406–8411. doi: 10.1074/jbc.m809618200

Lazarevic, V., Margot, P., Soldo, B., and Karamata, D. (1992). Sequencing and analysis of the *Bacillus subtilis* lytRABC divergon: a regulatory unit encompassing the structural genes of the N-acetylmuramoyl-L-alanine amidase and its modifier. *J. Gen. Microbiol.* 138, 1949–1961. doi: 10.1099/00221287-138-9-1949

Leng, C. H., Liu, S. J., Chen, H. W., and Chong, P. (2015). Recombinant bacterial lipoproteins as vaccine candidates. *Expert Rev. Vaccines* 14, 1623–1632. doi: 10.1586/14760584.2015.1091732

Li, N., Yang, X. Y., Guo, Z., Zhang, J., Cao, K., Han, J., et al. (2014). Varied metal-binding properties of lipoprotein PsaA in *Streptococcus pneumoniae*. *J. Biol. Inorg. Chem.* 19, 829–838. doi: 10.1007/s00775-014-1114-9

Liu, H., Fu, H., Jiang, X., Liao, X., Yue, M., Li, X., et al. (2019). PrsA contributes to *Streptococcus suis* serotype 2 pathogenicity by modulating secretion of selected virulence factors. *Vet. Microbiol.* 236:108375. doi: 10.1016/j.vetmic.2019.07.027

Liu, Y., Defourny, K. A. Y., Smid, E. J., and Abee, T. (2018). Gram-positive bacterial extracellular vesicles and their impact on health and disease. *Front. Microbiol.* 9:1502. doi: 10.3389/fmicb.2018.01502

Malke, H. (1998). Cytoplasmic membrane lipoprotein LppC of *Streptococcus equisimilis* functions as an acid phosphatase. *Appl. Environ. Microbiol.* 64, 2439–2442. doi: 10.1128/aem.64.7.2439-2442.1998

Narita, S., and Tokuda, H. (2011). Overexpression of LolCDE allows deletion of the *Escherichia coli* gene encoding apolipoprotein N-acyltransferase. *J. Bacteriol.* 193, 4832–4840. doi: 10.1128/jb.05013-11

Navarre, W. W., and Schneewind, O. (1999). Surface proteins of gram-positive bacteria and mechanisms of their targeting to the cell wall envelope. *Microbiol. Mol. Biol. Rev.* 63, 174–229. doi: 10.1128/mmbr.63.1.174-229.1999

Neubauer, H., Pantel, I., Lindgren, P. E., and Götz, F. (1999). Characterization of the molybdate transport system ModABC of *Staphylococcus carnosus*. *Arch. Microbiol.* 172, 109–115. doi: 10.1007/s002030050747

Nguyen, M. T., Deplanche, M., Nega, M., Le Loir, Y., Peisl, L., Götz, F., et al. (2016). *Staphylococcus aureus* Lpl lipoproteins delay G2/M phase transition in HeLa cells. *Front. Cell. Infect. Microbiol.* 6:201. doi: 10.3389/fcimb.2016.00201

Nguyen, M. T., and Götz, F. (2016). Lipoproteins of gram-positive bacteria: key players in the immune response and virulence. *Microbiol. Mol. Biol. Rev.* 80, 891–903. doi: 10.1128/mmbr.00028-16

Nguyen, M. T., Kraft, B., Yu, W., Demicrioglu, D. D., Hertlein, T., Burian, M., et al. (2015). The νSaα specific lipoprotein like cluster (lpl) of *S. aureus USA*300 contributes to immune stimulation and invasion in human cells. *PLoS Pathog.* 11:e1004984. doi: 10.1371/journal.ppat.1004984

Nguyen, M. T., Uebele, J., Kumari, N., Nakayama, H., Peter, L., Ticha, O., et al. (2017). Lipid moieties on lipoproteins of commensal and non-commensal staphylococci induce differential immune responses. *Nat. Commun.* 8:2246.

Nielsen, J. B., and Lampen, J. O. (1982a). Glyceride-cysteine lipoproteins and secretion by Gram-positive bacteria. *J. Bacteriol.* 152, 315–322.

Nielsen, J. B., and Lampen, J. O. (1982b). Membrane-bound penicillinases in Gram-positive bacteria. *J. Biol. Chem.* 257, 4490–4495.

Nimmo, G. R. (2012). USA300 abroad: global spread of a virulent strain of community-associated methicillin-resistant *Staphylococcus aureus*. *Clin. Microbiol. Infect.* 18, 725–734. doi: 10.1111/j.1469-0691.2012.03822.x

Okugawa, S., Moayeri, M., Pomerantsev, A. P., Sastalla, I., Crown, D., Gupta, P. K., et al. (2012). Lipoprotein biosynthesis by prolipoprotein diacylglyceryl transferase is required for efficient spore germination and full virulence of *Bacillus anthracis*. *Mol. Microbiol.* 83, 96–109. doi: 10.1111/j.1365-2958.2011.07915.x

Palacios, A., Sampedro, L., Sevilla, I. A., Molina, E., Gil, D., Azkargorta, M., et al. (2019). Mycobacterium tuberculosis extracellular vesicle-associated lipoprotein LpqH as a potential biomarker to distinguish paratuberculosis infection or vaccination from tuberculosis infection. *BMC Vet. Res.* 15:188. doi: 10.1186/s12917-019-1941-6

Paulsen, I. T., Banerjei, L., Myers, G. S., Nelson, K. E., Seshadri, R., and Read, T. D. (2003). Role of mobile DNA in the evolution of vancomycin-resistant *Enterococcus faecalis*. *Science* 299, 2071–2074. doi: 10.1126/science.1080613

Pelczar, P. L., and Setlow, P. (2008). Localization of the germination protein GerD to the inner membrane in *Bacillus subtilis* spores. *J. Bacteriol.* 190, 5635–5641. doi: 10.1128/jb.00670-08

Perego, M., Higgins, C. F., Pearce, S. R., Gallagher, M. P., and Hoch, J. A. (1991). The oligopeptide transport system of *Bacillus subtilis* plays a role in the initiation of sporulation. *Mol. Microbiol.* 5, 173–185. doi: 10.1111/j.1365-2958.1991.tb01838.x

Petriman, N. A., Jauss, B., Hufnagel, A., Franz, L., Sachelaru, I., Drepper, F., et al. (2018). The interaction network of the YidC insertase with the SecYEG translocon. SRP and the SRP receptor FtsY. *Sci. Rep.* 8:578.

Rammensee, H. G., Wiesmüller, K. H., Chandran, P. A., Zelba, H., Rusch, E., Gouttefangeas, C., et al. (2019). A new synthetic toll-like receptor 1/2 ligand is an efficient adjuvant for peptide vaccination in a human volunteer. *J. Immunother. Cancer* 7:307.

Reitermann, A., Metzger, J., Wiesmuller, K. H., Jung, G., and Bessler, W. G. (1989). Lipopeptide derivatives of bacterial lipoprotein constitute potent immune adjuvants combined with or covalently coupled to antigen or hapten. *Biol. Chem. Hoppe Seyler* 370, 343–352. doi: 10.1515/bchm3.1989.370.1.343

Remy, L., Carriere, M., Derre-Bobillot, A., Martini, C., Sanguinetti, M., and Borezee-Durant, E. (2013). The *Staphylococcus aureus* Opp1 ABC transporter imports nickel and cobalt in zinc-depleted conditions and contributes to virulence. *Mol. Microbiol.* 87, 730–743. doi: 10.1111/mmi.12126

Rey, M. W., Ramaiya, P., Nelson, B. A., Brody-Karpin, S. D., Zaretsky, E. J., and Tang, M. (2004). Complete genome sequence of the industrial bacterium *Bacillus licheniformis* and comparisons with closely related Bacillus species. *Genome Biol.* 5:R77.

Romero-Saavedra, F., Laverde, D., Budin-Verneuil, A., Muller, C., Bernay, B., Benachour, A., et al. (2015). Characterization of two metal binding lipoproteins as vaccine candidates for enterococcal infections. *PLoS One* 10:e0136625. doi: 10.1371/journal.pone.0136625

Saleh, M., Bartual, S. G., Abdullah, M. R., Jensch, I., Asmat, T. M., Petruschka, L., et al. (2013). Molecular architecture of *Streptococcus pneumoniae* surface thioredoxin-fold lipoproteins crucial for extracellular oxidative stress resistance and maintenance of virulence. *Embo Mol. Med.* 5, 1852–1870. doi: 10.1002/emmm.201202435

Samuelson, J. C., Chen, M., Jiang, F., Moller, I., Wiedmann, M., Kuhn, A., et al. (2000). YidC mediates membrane protein insertion in bacteria. *Nature* 406, 637–641. doi: 10.1038/35020586

Sauvageot, N., Mokhtari, A., Joyet, P., Budin-Verneuil, A., Blancato, V. S., Repizo, G. D., et al. (2017). *Enterococcus faecalis* uses a phosphotransferase system permease and a host colonization-related ABC transporter for maltodextrin uptake. *J. Bacteriol.* 199, e00878–16. doi: 10.1128/JB.00878-16

Schild, H., Deres, K., Wiesmüller, K. H., Jung, G., and Rammensee, H. G. (1991). Efficiency of peptides and lipopeptides for in vivo priming of virus-specific cytotoxic T cells. *Eur. J. Immunol.* 21, 2649–2654. doi: 10.1002/eji.1830211102

Schlatterer, K., Beck, C., Hanzelmann, D., Lebtig, M., Fehrenbacher, B., Schaller, M., et al. (2018). The mechanism behind bacterial lipoprotein release: phenol-soluble modulins mediate toll-like receptor 2 activation via extracellular vesicle release from *Staphylococcus aureus*. *mBio* 9:e1851-18.

Sebulsky, M. T., Hohnstein, D., Hunter, M. D., and Heinrichs, D. E. (2000). Identification and characterization of a membrane permease involved in iron-hydroxamate transport in *Staphylococcus aureus*. *J. Bacteriol.* 182, 4394–4400. doi: 10.1128/jb.182.16.4394-4400.2000

Serek, J., Bauer-Manz, G., Struhalla, G., Van Den Berg, L., Kiefer, D., Dalbey, R., et al. (2004). *Escherichia coli* YidC is a membrane insertase for Sec-independent proteins. *Embo J.* 23, 294–301. doi: 10.1038/sj.emboj.7600063

Shahmirzadi, S. V., Nguyen, M. T., and Götz, F. (2016). Evaluation of *Staphylococcus aureus* lipoproteins: role in nutritional acquisition and pathogenicity. *Front. Microbiol.* 7:1404. doi: 10.3389/fmicb.2016.01404

Shimizu, T., Ohtani, K., Hirakawa, H., Ohshima, K., Yamashita, A., Shiba, T., et al. (2002). Complete genome sequence of *Clostridium perfringens*, an anaerobic

flesh-eater. *Proc. Natl. Acad. Sci. U.S.A.* 99, 996–1001. doi: 10.1073/pnas.022493799

Skabytska, Y., Wolbing, F., Gunther, C., Koberle, M., Kaesler, S., Chen, K. M., et al. (2014). Cutaneous innate immune sensing of Toll-like receptor 2-6 ligands suppresses T cell immunity by inducing myeloid-derived suppressor cells. *Immunity* 41, 762–775. doi: 10.1016/j.immuni.2014.10.009

Spohn, R., Buwitt-Beckmann, U., Brock, R., Jung, G., Ulmer, A. J., and Wiesmüller, K. H. (2004). Synthetic lipopeptide adjuvants and Toll-like receptor 2–structure-activity relationships. *Vaccine* 22, 2494–2499. doi: 10.1016/j.vaccine.2003.11.074

Stoll, H., Dengjel, J., Nerz, C., and Götz, F. (2005). *Staphylococcus aureus* deficient in lipidation of prelipoproteins is attenuated in growth and immune activation. *Infect. Immun.* 73, 2411–2423. doi: 10.1128/iai.73.4.2411-2423.2005

Sun, X., Baker, H. M., Ge, R., Sun, H., He, Q. Y., and Baker, E. N. (2009). Crystal structure and metal binding properties of the lipoprotein MtsA, responsible for iron transport in *Streptococcus pyogenes*. *Biochemistry* 48, 6184–6190. doi: 10.1021/bi900552c

Sun, X., Ge, R., Chiu, J. F., Sun, H., and He, Q. Y. (2008). Lipoprotein MtsA of MtsABC in Streptococcus pyogenes primarily binds ferrous ion with bicarbonate as a synergistic anion. *FEBS Lett.* 582, 1351–1354. doi: 10.1016/j.febslet.2008.03.020

Sun, X., Ge, R., Zhang, D., Sun, H., and He, Q. Y. (2010). Iron-containing lipoprotein SiaA in SiaABC, the primary heme transporter of *Streptococcus pyogenes*. *J. Biol. Inorg. Chem.* 15, 1265–1273. doi: 10.1007/s00775-010-0684-4

Sutcliffe, I. C., and Russell, R. R. (1995). Lipoproteins of gram-positive bacteria. *J. Bacteriol.* 177, 1123–1128. doi: 10.1128/jb.177.5.1123-1128.1995

Takami, H., Takaki, Y., Chee, G. J., Nishi, S., Shimamura, S., Suzuki, H., et al. (2004). Thermoadaptation trait revealed by the genome sequence of thermophilic Geobacillus kaustophilus. *Nucleic Acids Res.* 32, 6292–6303. doi: 10.1093/nar/gkh970

Takami, H., Takaki, Y., and Uchiyama, I. (2002). Genome sequence of *Oceanobacillus iheyensis* isolated from the Iheya Ridge and its unexpected adaptive capabilities to extreme environments. *Nucleic Acids Res.* 30, 3927–3935. doi: 10.1093/nar/gkf526

Tedde, V., Rosini, R., and Galeotti, C. L. (2016). Zn2+ Uptake in *Streptococcus pyogenes*: characterization of adcA and lmb null mutants. *PLoS One* 11:e0152835. doi: 10.1371/journal.pone.0152835

Tettelin, H., Nelson, K. E., Paulsen, I. T., Eisen, J. A., Read, T. D., and Peterson, S. (2001). Complete genome sequence of a virulent isolate of *Streptococcus pneumoniae*. *Science* 293, 498–506. doi: 10.1126/science.1061217

Touchette, M. H., Van Vlack, E. R., Bai, L., Kim, J., Cognetta, A. B. III, Previti, M. L., et al. (2017). A screen for protein-protein interactions in live mycobacteria reveals a functional link between the virulence-associated lipid transporter LprG and the mycolyltransferase antigen 85A. *ACS Infect. Dis.* 3, 336–348. doi: 10.1021/acsinfecdis.6b00179

Tribelli, P. M., Luqman, A., Nguyen, M. T., Madlung, J., Fan, S. H., Macek, B., et al. (2019). Staphylococcus aureus Lpl protein triggers human host cell invasion via activation of Hsp90 receptor. *Cell Microbiol.* 22:e13111.

Tschumi, A., Nai, C., Auchli, Y., Hunziker, P., Gehrig, P., Keller, P., et al. (2009). Identification of apolipoprotein N-acyltransferase (Lnt) in mycobacteria. *J. Biol. Chem.* 284, 27146–27156. doi: 10.1074/jbc.m109.022715

Vermassen, A., Leroy, S., Talon, R., Provot, C., Popowska, M., and Desvaux, M. (2019). Cell wall hydrolases in bacteria: insight on the diversity of cell wall amidases, glycosidases and peptidases toward peptidoglycan. *Front. Microbiol.* 10:331. doi: 10.3389/fmicb.2019.00331

Vitikainen, M., Lappalainen, I., Seppala, R., Antelmann, H., Boer, H., Taira, S., et al. (2004). Structure-function analysis of PrsA reveals roles for the parvulin-like and flanking N- and C-terminal domains in protein folding and secretion in Bacillus subtilis. *J. Biol. Chem.* 279, 19302–19314. doi: 10.1074/jbc.m400861200

Wang, X., Chen, Z., Feng, H., Chen, X., and Wei, L. (2019). Genetic variants of the oppA gene are involved in metabolic regulation of surfactin in Bacillus subtilis. *Microb. Cell Fact.* 18:141.

White, D. W., Elliott, S. R., Odean, E., Bemis, L. T., and Tischler, A. D. (2018). *Mycobacterium tuberculosis* Pst/SenX3-RegX3 regulates membrane vesicle production independently of ESX-5 activity. *mBio* 9:e00778-18.

Wilburn, K. M., Fieweger, R. A., and Vanderven, B. C. (2018). Cholesterol and fatty acids grease the wheels of *Mycobacterium tuberculosis* pathogenesis. *Pathog. Dis.* 76:fty021.

Williams, W. A., Zhang, R. G., Zhou, M., Joachimiak, G., Gornicki, P., Missiakas, D., et al. (2004). The membrane-associated lipoprotein-9 GmpC from *Staphylococcus aureus* binds the dipeptide GlyMet via side chain interactions. *Biochemistry* 43, 16193–16202. doi: 10.1021/bi048877o

Wirth, R. (1994). The sex pheromone system of *Enterococcus faecalis*. More than just a plasmid-collection mechanism? *Eur. J. Biochem.* 222, 235–246. doi: 10.1111/j.1432-1033.1994.tb18862.x

Wolschendorf, F., Mahfoud, M., and Niederweis, M. (2007). Porins are required for uptake of phosphates by *Mycobacterium smegmatis*. *J. Bacteriol.* 189, 2435–2442. doi: 10.1128/jb.01600-06

Xu, S., Andrews, D., and Hill, B. C. (2015). The affinity of yeast and bacterial SCO proteins for CU(I) and CU(II). A capture and release strategy for copper transfer. *Biochem. Biophys. Rep.* 4, 10–19. doi: 10.1016/j.bbrep.2015.08.010

Yang, X. Y., Sun, B., Zhang, L., Li, N., Han, J., Zhang, J., et al. (2014). Chemical interference with iron transport systems to suppress bacterial growth of *Streptococcus pneumoniae*. *PLoS One* 9:e105953. doi: 10.1371/journal.pone.0105953

Zhang, L., Hendrickson, R. C., Meikle, V., Lefkowitz, E. J., Ioerger, T. R., and Niederweis, M. (2020). Comprehensive analysis of iron utilization by *Mycobacterium tuberculosis*. *PLoS Pathog.* 16:e1008337. doi: 10.1371/journal.ppat.1008337

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Microbiology

MINI REVIEW
published: 29 October 2020
doi: 10.3389/fmicb.2020.594737

The Role of Gram-Positive Surface Proteins in Bacterial Niche- and Host-Specialization

*Amy C. Pickering and J. Ross Fitzgerald**

The Roslin Institute and Edinburgh Infectious Diseases, University of Edinburgh, Easter Bush Campus, Edinburgh, United Kingdom

Gram-positive bacterial pathogens have an array of proteins on their cell surface that mediate interactions with the host environment. In particular, bacterial cell wall-associated (CWA) proteins play key roles in both colonization and pathogenesis. Furthermore, some CWA proteins promote specialization for host-species or mediate colonization of specific anatomical niches within a host. In this mini review, we provide examples of the many ways by which major pathogens, such as Staphylococci, Streptococci and *Listeria monocytogenes,* utilize CWA proteins for both host- and niche-specialization. We describe different biological mechanisms mediated by CWA proteins including: the acquisition of iron from hemoglobin in the bloodstream, adherence to and invasion of host cells, and innate immune evasion through binding to the plasma proteins fibrinogen, immunoglobulin G, and complement. We also discuss the limitations of using animal models for understanding the role of specific CWA proteins in host-specialization and how transformative technologies, such as CRISPR-Cas, offer tremendous potential for developing transgenic models that simulate the host environment of interest. Improved understanding of the role of CWA proteins in niche- or host-specificity will allow the design of new therapeutic approaches which target key host–pathogen interactions underpinning Gram-positive bacterial infections.

Keywords: specialized, niche, host, surface protein, adaptation, bacteria, interaction

OPEN ACCESS

Edited by:
Timothy J. Foster,
Trinity College Dublin, Ireland

Reviewed by:
Maria Rapala-Kozik,
Jagiellonian University, Poland
Delma S. Childers,
University of Aberdeen,
United Kingdom

***Correspondence:**
J. Ross Fitzgerald
Ross.Fitzgerald@ed.ac.uk

Specialty section:
This article was submitted to Infectious Diseases, a section of the journal Frontiers in Microbiology

Received: *14 August 2020*
Accepted: *05 October 2020*
Published: *29 October 2020*

Citation:
Pickering AC and Fitzgerald JR (2020) The Role of Gram-Positive Surface Proteins in Bacterial Niche- and Host-Specialization. Front. Microbiol. 11:594737. doi: 10.3389/fmicb.2020.594737

INTRODUCTION

Bacteria have typically evolved to occupy particular niches within host-species or the environment. For example, *Streptococcus uberis* and some clones of *Staphylococcus aureus* are specialized for infection of the bovine mammary gland and *Listeria monocytogenes* has a tropism for transcytosis of the placenta and intestinal epithelia (Lecuit et al., 2004). Furthermore, bacterial pathogens have adapted to colonize either a single host-species (specialist) or multiple host-species (generalist) (Woolhouse et al., 2001; Bäumler and Fang, 2013). For example, the human pathogen *Streptococcus pyogenes*, associated with infections such as pharyngitis, skin infections, rheumatic fever and necrotizing fasciitis, and the equine pathogen *Streptococcus equi* subsp. *equi*, associated with the respiratory disease strangles, have a very restricted host-specialization. In contrast, the multi-host "generalist" *S. aureus* is associated with an array of different infections in humans, livestock and wild animal species (Sheppard et al., 2018). However, individual clonal lineages or subtypes of *S. aureus* have evolved the capacity to infect distinct host-species following a host-switch event (Lowder et al., 2009; Fitzgerald, 2012; Richardson et al., 2018; Haag et al., 2019). Adaptation to

a novel niche or host-species involves multiple evolutionary mechanisms including mutation, recombination and horizontal gene transfer. Comparative genomic analysis of *S. aureus* from different host-species has identified genetic signatures affecting cell surface proteins, marked by gene acquisition, diversification and loss of function events, suggesting a key role for cell surface proteins in host- and niche- specialization (Herron-Olson et al., 2007; Ben Zakour et al., 2008; Lowder et al., 2009; Guinane et al., 2010; Resch et al., 2013; Spoor et al., 2015). Signatures of adaptation affecting the cell envelope may reveal pathways that could be disrupted therapeutically. For instance, *S. aureus* has undergone several host-switch events from humans into rabbits underpinned by natural adaptive mutations in a gene encoding DltB, a membrane-associated protein involved in D-Ala modification of wall teichoic acids (Viana et al., 2015). These data highlighted the key importance of DltB in host–pathogen interactions, a discovery reinforced by a subsequent study that identified DltB as a novel druggable target (Pasquina et al., 2016).

The aim of this mini review is to highlight the role of cell wall-associated (CWA) proteins of Gram-positive bacteria in overcoming host- and niche-specific barriers to infection. We provide examples of CWA proteins that are either covalently bound to cell wall peptidoglycan, or indirectly attached to the cell wall via non-covalent interactions with wall teichoic acids. Important host- or niche-specific CWA protein interactions underpinning biological mechanisms are described including: acquisition of iron from hemoglobin, adherence to and invasion of host cells, and evasion of the innate immune response (**Figure 1**). Understanding the mechanisms by which bacteria adapt to different niches or host-species can reveal critical host–pathogen interactions that could potentially be targeted to develop novel therapeutic approaches. However, demonstration of the *in vivo* importance of a host-specific CWA protein interaction is often challenging outside of the natural host. Here, we also discuss the potential for the design of transgenic models that express host-specialized surface protein receptors and would facilitate studies into the role of host- or niche-specific interactions in bacterial pathogenesis.

NUTRIENT ACQUISITION OF IRON FROM HEMOGLOBIN

Bacterial adaptation to a particular host-species or niche often requires modification of surface proteins to allow them to interact with distinct ligands or polymorphic host receptors. Iron is an essential nutrient required as a co-factor for amino acid biosynthesis, the TCA cycle, DNA replication, cellular respiration, and electron transport (Sheldon and Heinrichs, 2015). In the host, iron is often sequestered in molecules such as hemoglobin in the blood and lactoferrin in mucosal secretions and milk. During bacterial infection, iron availability is actively decreased by limiting iron uptake from food digestion in a process termed nutritional immunity (Weinberg, 1974). However, bacteria have evolved numerous mechanisms, some of which involve utilization of surface proteins, to overcome this limitation and to successfully compete with the host for iron.

Heme Acquisition – Iron-Regulated Surface-Determinant Pathway

One such mechanism involves the iron-regulated surface-determinant (Isd) pathway present in a number of Gram-positive bacteria, including *S. aureus*, *S. pyogenes*, *Bacillus anthracis*, and *L. monocytogenes* (Sheldon and Heinrichs, 2015). In *S. aureus,* the Isd pathway is involved in acquisition of iron from hemoglobin and myoglobin in the bloodstream, allowing *S. aureus* to rapidly adapt to low iron conditions (Torres et al., 2006). *S. aureus* also responds to low iron conditions by altering the expression of metabolic proteins to generate a more acidic environment, promoting the release of iron from host proteins (Friedman et al., 2006). The Isd pathway is represented by multiple cell envelope-associated proteins encoded in an operon, with the near iron transporter (NEAT) domain 1 of the CWA protein IsdB interacting directly with hemoglobin (Torres et al., 2006). Heme is then transferred to NEAT domain 2 of IsdB and along a chain of Isd proteins concluding in heme binding to the IsdE membrane-associated lipoprotein. This allows IsdF-mediated translocation of the heme iron into the bacterial cytoplasm for use in metabolism (Muryoi et al., 2008; Zhu et al., 2008; Bowden et al., 2014; Gianquinto et al., 2019). Notably, IsdB is a host-specialized CWA protein with enhanced binding affinity for human hemoglobin, in comparison to hemoglobin from other mammalian sources (Pishchany et al., 2010). Strikingly, mice expressing a human hemoglobin protein have increased susceptibility to *S. aureus* infection in comparison to those expressing the native murine hemoglobin (Pishchany et al., 2010). The crystal structure of IsdB, in complex with human hemoglobin, revealed that binding results in conformational changes that displace heme from the β-subunits and subsequently the α-subunits of hemoglobin (Bowden et al., 2018). This IsdB-hemoglobin interface exhibits signatures of positive selection during the evolution of primates, involving both the α- and β-globins, suggesting that the role of IsdB in human-specialization is part of a long evolutionary relationship between *S. aureus* and hemoglobin (Choby et al., 2018). The specialization of IsdB to human hemoglobin, is similarly observed in the closely related *Staphylococcus argenteus* and *Staphylococcus schweitzeri* species, in addition to the more distantly related human-specialized *Staphylococcus lugdunensis* (Zapotoczna et al., 2012; Choby et al., 2018). Recently, a novel iron acquisition system in *S. lugdunensis,* dependent on an ECF-type transporter, has also demonstrated host-restriction based on the inability to lyse non-human erythrocytes to release hemoproteins (Jochim et al., 2020). These findings highlight how the co-evolution of bacterium and host can lead to bacterial adaptation to a key nutrient source in a particular host.

ADHERENCE AND INVASION OF HOST CELLS

In addition to nutrient acquisition, all bacterial pathogens must overcome an array of anatomical, physiological and immunological barriers to colonize and establish infection.

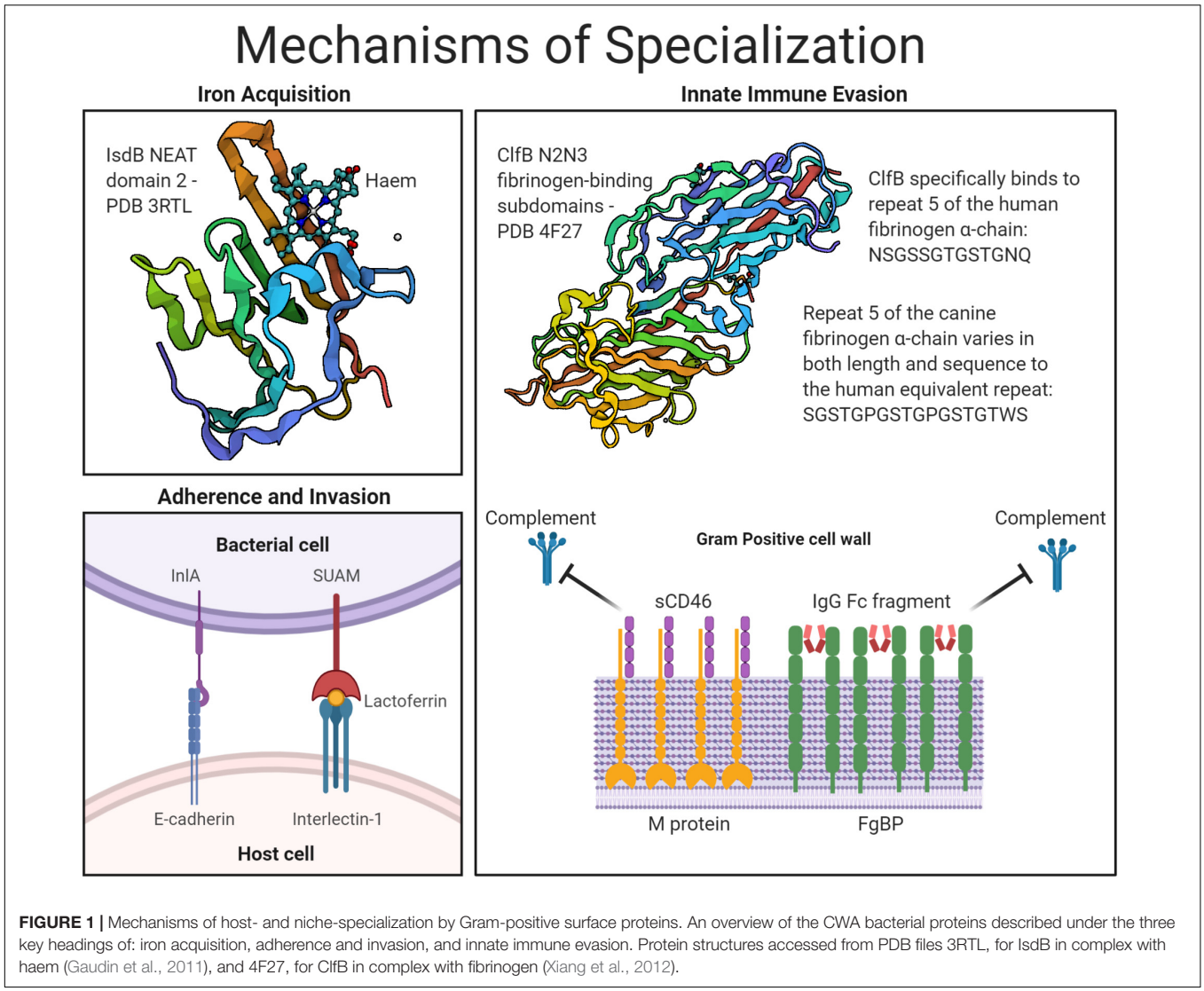


FIGURE 1 | Mechanisms of host- and niche-specialization by Gram-positive surface proteins. An overview of the CWA bacterial proteins described under the three key headings of: iron acquisition, adherence and invasion, and innate immune evasion. Protein structures accessed from PDB files 3RTL, for IsdB in complex with haem (Gaudin et al., 2011), and 4F27, for ClfB in complex with fibrinogen (Xiang et al., 2012).

Adherence to and invasion of host cells is an important strategy for bacterial colonization, immune evasion, and dissemination. Gram-positive bacteria exploit a range of mechanisms to mediate binding and uptake by different cell types (Ribet and Cossart, 2015; Josse et al., 2017). Such interactions may promote establishment in a particular niche found across multiple host-species or may require diversification, to facilitate binding to tissue or cells that have polymorphic ligands or receptors in a distinct host-species.

Internalin A of *Listeria monocytogenes*

A well characterized example of a CWA protein mediating both niche- and host-specialized cellular internalization is the internalin A (InlA) protein of *L. monocytogenes,* a veterinary pathogen that causes encephalitis in ruminants and a highly virulent foodborne pathogen that can cause miscarriage, stillbirth, or premature labor in pregnant women or meningitis in newborns (Williams et al., 2011; Roulo et al., 2014). InlA interacts with the N-terminal region of E-cadherin expressed by epithelial cells, including those present on the placental villous trophoblast barrier, promoting niche-specialization (Mengaud et al., 1996; Schubert et al., 2002; Lecuit et al., 2004). Primarily, this interaction allows *L. monocytogenes,* where E-cadherin is exposed, to pass through the intestinal epithelial barrier via transcytosis either on the luminal surface around intestinal goblet cells or intercellular junctions that form as part of normal gut homeostasis (Pentecost et al., 2006; Nikitas et al., 2011). It has been demonstrated that heterologous expression of InlA confers the capacity for internalization to the invasion-deficient *Listeria innocua,* and that natural premature stop codons in the *inlA* gene of *L. monocytogenes* clinical isolates are associated with a reduced capacity to invade intestinal epithelium and an increased infective dose in experimental models (Gaillard et al., 1991; Van Stelten et al., 2011; Su et al., 2019).

Determination of the crystal structure of the InlA interaction with E-cadherin provided evidence for host-specialization with stronger binding observed to human compared to murine E-cadherin (Lecuit et al., 1999; Schubert et al., 2002). Elegant

mutational experiments demonstrated that a single proline residue at position 16 of E-cadherin was sufficient to alter host-tropism (Lecuit et al., 1999). Presence of Pro16 in E-cadherin, as naturally occurs in humans, rabbits, and guinea pigs, facilitates strong binding, whereas Glu16, as occurs in rodents, leads to a reduced affinity (Lecuit et al., 1999). Of note, rabbits are considered to be the natural hosts of *L. monocytogenes* with cows also encoding Pro16 in bovine E-cadherin (Murray et al., 1926; Zimin et al., 2009). InlA of *L. monocytogenes* has therefore evolved to promote both niche- and host-specialized invasion of intestinal and placental epithelial cells in the natural rabbit host, and conserved residues in human and bovine E-cadherin underpin the capacity for zoonotic infection.

Streptococcus uberis Adhesion Molecule

Streptococcus uberis, a leading cause of bovine mastitis, has also evolved a niche- and host-specialized mechanism of adherence and invasion. Lactoferrin is an iron-sequestering glycoprotein with antimicrobial properties that is highly expressed in milk during bovine mastitis (Harmon et al., 1976; Hagiwara et al., 2003; Chaneton et al., 2008). Due to its abundance and importance in host defense, many bacterial pathogens have evolved mechanisms for interacting with lactoferrin (Valenti and Antonini, 2005; Lu et al., 2020). *S. uberis* encodes two lactoferrin-binding proteins that demonstrate preference for bovine lactoferrin, in comparison to the human variant, and promote evasion of the antimicrobial properties of lactoferrin in the mammary gland niche (Fang and Oliver, 1999; Moshynskyy et al., 2003). In particular, *S. uberis* adhesion molecule (SUAM) uses lactoferrin as a cross-bridge with host interlectin-1 to promote adherence to and internalization of *S. uberis* into bovine mammary epithelial cells (Fang et al., 2000; Moshynskyy et al., 2003; Almeida et al., 2006; Patel et al., 2009; Chen et al., 2011). A mutant *S. uberis* strain deficient in SUAM expression was attenuated for virulence in an experimentally infected mammary gland, suggesting a key role for SUAM in the pathogenesis and niche-specificity of *S. uberis* (Patel et al., 2009; Almeida et al., 2015).

INNATE IMMUNE EVASION

In addition to the acquisition of key nutrients and the subversion of host cells, overcoming the host immune response is critical for the pathogenesis of Gram-positive bacteria. A wide array of innate immune mechanisms have evolved to prevent bacterial infections including the development of blood clots at wound sites and damaged blood vessels, and the activation of the complement cascade leading to intracellular killing by phagocytic cells such as neutrophils, macrophages, and dendritic cells. In response, Gram-positive bacteria can utilize surface proteins to disrupt these immune mechanisms in a niche or host-specialized manner by interacting with host proteins such as the plasma glycoprotein fibrinogen, to inhibit both hemostasis and phagocytosis, and complement factors such as factor H and CD46, to inhibit complement deposition.

Staphylococcal Interactions With Fibrinogen

Fibrinogen plays a central role in hemostasis and contains α- β- and γ-chains as a dimer of trimers (Mosesson, 2005). The ubiquity and abundance of fibrinogen has led a wide range of bacterial pathogens to develop ways of subverting fibrinogen for adherence to host cells, abscess formation, and immune evasion (Ko and Flick, 2016; Thomer et al., 2016). Staphylococcal species, including *S. aureus*, *S. lugdunensis*, and *S. pseudintermedius*, commonly adhere to the C-terminal of the fibrinogen γ-chain to provide a generalized but host-specific interaction, which interferes with fibrinogen-mediated coagulation and platelet aggregation (McDevitt et al., 1997; Geoghegan et al., 2010; Pietrocola et al., 2013). In the case of clumping factor A (ClfA) of *S. aureus*, the fibrinogen γ-chain binding is equivalent for human, canine, feline, murine, and porcine fibrinogen but demonstrates reduced binding to bovine fibrinogen and no detectable ovine fibrinogen-binding, due to a single amino acid substitution in the ovine fibrinogen γ-chain (Geoghegan et al., 2010). This host-specific fibrinogen-interaction of ClfA is essential for *S. aureus*-mediated sepsis, in murine infection models, as well as mediating phagocytosis inhibition and bacterial aggregation, key innate immune evasion strategies of *S. aureus* (Higgins et al., 2006; Flick et al., 2013; Claes et al., 2018).

Additional Staphylococcal CWA proteins demonstrate a more host-specialized fibrinogen interaction with the repeat region of the fibrinogen α-chain, which exhibits both inter- and intraspecies variation (Murakawa et al., 1993; Inaba et al., 2008; Pickering et al., 2019). For example, clumping factor B (ClfB) of *S. aureus* interacts solely with repeat 5 of the human fibrinogen α-chain, promoting human-specialized platelet aggregation, with *S. pseudintermedius* surface protein L (SpsL) exhibiting canine-specialized high-affinity binding to the repeat region of the canine fibrinogen α-chain (Miajlovic et al., 2007; Walsh et al., 2008; Pickering et al., 2019). SpsL is unique in exhibiting both canine-specific binding to the repeat region of the fibrinogen α-chain and a secondary weaker binding interaction that is also observed with human fibrinogen (Bannoehr et al., 2011; Pickering et al., 2019). This canine-specialized fibrinogen interaction promotes *S. pseudintermedius* aggregation and phagocytosis inhibition by neutrophils, demonstrating the role of fibrinogen-binding as a host-specialized immune evasion strategy of *S. pseudintermedius* (Pickering et al., 2019). Adherence to the fibrinogen α-chain is also utilized by the Srr1 and Srr2 glycoproteins of *Streptococcus agalactiae* for host-specialization, with specific binding to human repeats 6-8, demonstrating that this region of fibrinogen is the target of host-specialization for an array of Gram-positive bacteria (Seo et al., 2013).

M protein of *Streptococcus pyogenes*

Streptococcus pyogenes is a major human-specific pathogen that utilizes a range of innate immune evasion strategies to cause acute pharyngitis and impetigo as well as severe invasive infections such as necrotizing fasciitis (Laabei and Ermert, 2019). The most abundant CWA protein presented on the surface of *S. pyogenes* is the M protein, with certain M protein serotypes associated

with specific niches such as the throat or skin (Fischetti, 1989; Cunningham, 2000). The M protein has an array of host ligands that promote both host cell adherence and immune evasion (Fischetti, 1989). The binding of M protein to complement components such as factor H, C4BP and CD46 all lead to complement inhibition (Horstmann et al., 1988; Thern et al., 1995). In the case of CD46 complement regulatory protein, which aids the cleavage of complement factors C3b and C4b on host cells, M protein has been demonstrated to interact in a human-specific manner (Riley-Vargas et al., 2004; Lövkvist et al., 2008). The interaction of M protein with human cellular CD46 mediates binding to keratinocytes and invasion of lung epithelial cells ultimately leading to cell death (Okada et al., 1995; Rezcallah et al., 2005). During epithelial cell apoptosis, soluble CD46 is shed from the host cell leading to reduced killing of *S. pyogenes* in whole blood due to its interaction with M protein (Lövkvist et al., 2008). This immune evasion mechanism is human-specialized as evidenced by increased whole blood survival in transgenic mice expressing human CD46 (Lövkvist et al., 2008). These human-expressing CD46 mice were also more susceptible to experimental infection with increased mortality and bacterial levels during a bloodstream infection and the development of necrotizing fasciitis after subcutaneous injection (Lövkvist et al., 2008; Matsui et al., 2009). *S. pyogenes* has therefore developed a human-specialized immune evasion strategy associated with complement inhibition.

Immunoglobulin Binding by *Streptococcus equi* subsp. *equi*

A final example of host-specific immune evasion is the fibrinogen-binding protein (FgBP) of the equine-specific *S. equi* subsp. *equi,* the cause of the respiratory disease strangles in horses. FgBP is a CWA protein which was originally characterized as a fibrinogen-binding protein with preference for equine Fg and as a protective antigen in a murine infection model (Meehan et al., 1998, 2000). However, it was later identified as an IgG-binding protein of equine, human, rabbit, porcine, and feline IgG but not of murine, rat, goat, sheep, cow or chicken IgG (Meehan et al., 2001). By interacting with the Fc interdomain region of IgG4 and IgG7, FgBP disrupts complement deposition and antibody-mediated activation of the classical complement pathway, leading to increased survival in equine whole blood, representing a host-specific immune evasion strategy of *S. equi* subsp. *equi* (Meehan et al., 2001; Lewis et al., 2008).

MODELS TO STUDY NICHE- AND HOST-SPECIALIZED CELL SURFACE PROTEINS

The examples provided here include structural, biochemical, and molecular evidence for niche- and host-specialized interactions. However, it is often challenging to validate the importance of these host-pathogen interactions in appropriate infection models, especially when the bacteria exhibit a human tropism (Douam et al., 2015). In some cases, an effective model may not exist and the ethical and/or cost implications of performing experimental infections of large animals is prohibitive. Importantly, small animal models may not have ligands or receptors in common with the natural host and would therefore not support the function of the bacterial surface protein in question. Using knockdown and heterologous expression systems in cell lines *in vitro* can provide useful information relating to specific host–pathogen interactions but they cannot address the relevance of the interaction in a complex infection setting *in vivo*. The development of three-dimensional (3D) organoid systems derived from stem cells for infection biology is a highly promising nascent research area that will provide more complex multicellular systems for investigating host–pathogen interactions *in vitro* (Iakobachvili and Peters, 2017). One such system has been developed from murine mammary gland tissue to create a 3D mammary organoid that can be manipulated to induce lactation or involution (Sumbal et al., 2020). A similar system could be used to generate a bovine mammary organoid allowing the *in vitro* examination of the lactoferrin-dependent internalization of *S. uberis*.

Another possibility is to construct transgenic mice that express receptors supporting the function of surface proteins. In addition to the examples already mentioned here for *S. aureus* IsdB and *S. pyogenes* M protein (Lövkvist et al., 2008; Matsui et al., 2009; Pishchany et al., 2010), this approach was carried out to

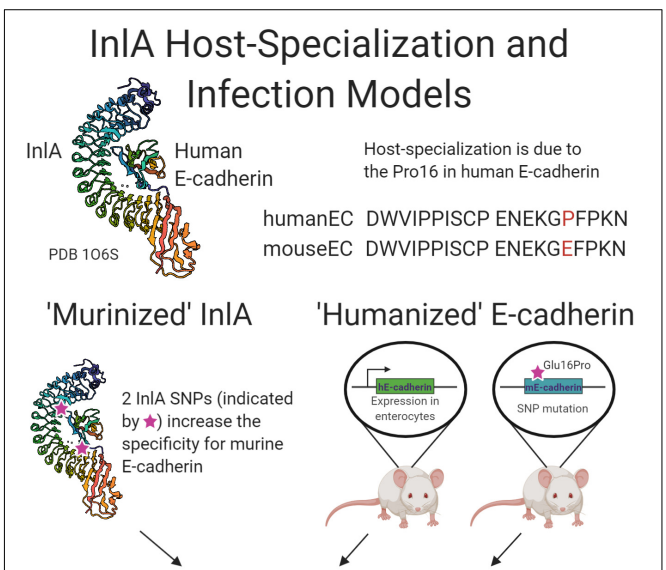


FIGURE 2 | Schematic summary of research examining the interaction of internalin A of *L. monocytogenes* with human E-cadherin. Crystal structure generation of InlA in complex with human E-cadherin, PDB 1O6S (Schubert et al., 2002), allowed the "murinization" of InlA with enhanced binding to murine E-cadherin. Two transgenic murine models were developed that express the human E-cadherin protein either solely in the intestine or throughout the mouse. All three approaches demonstrated the role of InlA in crossing the intestinal epithelial barrier during *L. monocytogenes* experimental infection.

examine the role of *L. monocytogenes* InlA during colonization of the gut (**Figure 2**). Mice were genetically manipulated to either express human E-cadherin from a specialized promoter in enterocytes or a "humanized" Glu16Pro version of the murine E-cadherin in all E-cadherin-expressing cells (Lecuit et al., 2001; Lecuit and Cossart, 2002; Disson et al., 2008, 2009). Both systems demonstrated the requirement of InlA for intestinal crossing of *L. monocytogenes* with the chimeric mice also demonstrating that both InlA and internalin B are required for placental invasion during murine pregnancy (Lecuit et al., 2001; Disson et al., 2008). In the reverse scenario, InlA on the *L. monocytogenes* cell surface has been "murinized" through two single amino acid substitutions, which enhanced binding to murine E-cadherin *in vitro* and led to infection of the murine intestinal epithelium *in vivo* (Wollert et al., 2007). This subsequently led to higher bacterial burden in multiple organs including the liver and spleen, demonstrating the importance of this InlA-dependent route of bacterial dissemination (Wollert et al., 2007). Further utilization of this "murinized" *L. monocytogenes* strain demonstrated that the C57BL/6J mouse cell line has inherent resistance to oral challenge by *Listeria* with more susceptible murine cell lines exhibiting faster dissemination of the mutated strain as well as increased cytokine production (Bergmann et al., 2013). Advances in genome editing tools, including CRISPR-Cas systems, have made transgenesis of rodents a relatively facile approach, allowing the development of new transgenic models that can facilitate investigations into the importance of host-specialized bacterial interactions *in vivo* (Qin et al., 2016). For example, recent studies have employed CRISPR-Cas edited mice expressing the human integrin component CD11b to examine the function of *S. aureus* toxin LukAB *in vivo* (Boguslawski et al., 2020).

CONCLUSION

Surface proteins of Gram-positive bacteria are essential for an array of different host-pathogen interactions. Accordingly, bacterial adaptation to a new host-species or anatomical niche often requires the acquisition of new proteins or diversification of existing proteins to enable binding to polymorphic ligands or receptors. As highlighted by the examples provided in the current review, important advances have been made regarding our understanding of the molecular basis for such interactions *in vitro*. In addition, new approaches for transgenesis and mutagenesis offer great promise for the development of new organoid or whole animal models of infection that allow the dissection of the role in pathogenesis of niche and/or host-specialized surface proteins, many of which have multiple binding tropisms involving several different receptor ligands. Development of appropriate experimental models will not only allow the validation of surface proteins that play important roles *in vivo* but also improve on the current model systems used for the evaluation of novel therapeutic and vaccine candidates. The identification of alternative approaches for control of infection is of particular importance for bacteria that exhibit high levels of antibiotic resistance and are becoming increasingly difficult to treat in a clinical setting.

AUTHOR CONTRIBUTIONS

All authors wrote and edited the manuscript.

FUNDING

JRF is supported by the Biotechnology and Biological Sciences Research Council (UK) Institute Strategic Grant (ISP2) (BBS/E/D/20002173), and a Wellcome Trust collaborative award 201531/Z/16/Z.

ACKNOWLEDGMENTS

The authors would like to thank collaborators and members of the Fitzgerald group past and present for their contributions to research on this topic. Figures were created with BioRender.com.

REFERENCES

Almeida, R. A., Dego, O. K., Headrick, S. I., Lewis, M. J., and Oliver, S. P. (2015). Role of *Streptococcus uberis* adhesion molecule in the pathogenesis of *Streptococcus uberis* mastitis. *Vet. Microbiol.* 179, 332–335. doi: 10.1016/j.vetmic.2015.07.005

Almeida, R. A., Luther, D. A., Park, H.-M., and Oliver, S. P. (2006). Identification, isolation, and partial characterization of a novel *Streptococcus uberis* adhesion molecule (SUAM). *Vet. Microbiol.* 115, 183–191. doi: 10.1016/j.vetmic.2006.02.005

Bannoehr, J., Ben Zakour, N. L., Reglinski, M., Inglis, N. F., Prabhakaran, S., Fossum, E., et al. (2011). Genomic and surface proteomic analysis of the canine pathogen *Staphylococcus pseudintermedius* reveals proteins that mediate adherence to the extracellular matrix. *Infect. Immun.* 79, 3074–3086. doi: 10.1128/iai.00137-11

Bäumler, A., and Fang, F. C. (2013). Host specificity of bacterial pathogens. *Cold Spring Harb. Perspect. Med.* 3:a010041. doi: 10.1101/cshperspect.a010041

Ben Zakour, N. L., Sturdevant, D. E., Even, S., Guinane, C. M., Barbey, C., Alves, P. D., et al. (2008). Genome-wide analysis of ruminant *Staphylococcus aureus* reveals diversification of the core genome. *J. Bacteriol.* 190, 6302–6317. doi: 10.1128/jb.01984-07

Bergmann, S., Beard, P. M., Pasche, B., Lienenklaus, S., Weiss, S., Gahan, C. G. M., et al. (2013). Influence of internalin a murinisation on host resistance to orally acquired listeriosis in mice. *BMC Microbiol.* 13:90. doi: 10.1186/1471-2180-13-90

Boguslawski, K. M., McKeown, A. N., Day, C. J., Lacey, K. A., Tam, K., Vozhilla, N., et al. (2020). Exploiting species specificity to understand the tropism of a human-specific toxin. *Sci. Adv.* 6:eaax7515. doi: 10.1126/sciadv.aax7515

Bowden, C. F. M., Chan, A. C. K., Li, E. J. W., Arrieta, A. L., Eltis, L. D., and Murphy, M. E. P. (2018). Structure–function analyses reveal key features in *Staphylococcus aureus* IsdB-associated unfolding of the heme-binding pocket of human hemoglobin. *J. Biol. Chem.* 293, 177–190. doi: 10.1074/jbc.M117.806562

Bowden, C. F. M., Verstraete, M. M., Eltis, L. D., and Murphy, M. E. P. (2014). Hemoglobin binding and catalytic heme extraction by IsdB near iron transporter domains. *Biochemistry* 53, 2286–2294. doi: 10.1021/bi500230f

Chaneton, L., Tirante, L., Maito, J., Chaves, J., and Bussmann, L. E. (2008). Relationship between milk lactoferrin and etiological agent in the mastitic

bovine mammary gland. *J. Dairy Sci.* 91, 1865–1873. doi: 10.3168/jds.2007-0732

Chen, X., Dego, O. K., Almeida, R. A., Fuller, T. E., Luther, D. A., and Oliver, S. P. (2011). Deletion of sua gene reduces the ability of *Streptococcus uberis* to adhere to and internalize into bovine mammary epithelial cells. *Vet. Microbiol.* 147, 426–434. doi: 10.1016/j.vetmic.2010.07.006

Choby, J. E., Buechi, H. B., Farrand, A. J., Skaar, E. P., and Barber, M. F. (2018). Molecular basis for the evolution of species-specific hemoglobin capture by *Staphylococcus aureus*. *mBio* 9:e01524-18. doi: 10.1128/mBio.01524-18

Claes, J., Ditkowski, B., Liesenborghs, L., Veloso, T. R., Entenza, J. M., Moreillon, P., et al. (2018). Assessment of the dual role of clumping factor a in *S. aureus* adhesion to endothelium in absence and presence of plasma. *Thromb. Haemost.* 118, 1230–1241. doi: 10.1055/s-0038-1660435

Cunningham, M. W. (2000). Pathogenesis of group a streptococcal infections. *Clin. Microbiol. Rev.* 13, 470–511. doi: 10.1128/cmr.13.3.470-511.2000

Disson, O., Grayo, S., Huillet, E., Nikitas, G., Langa-Vives, F., Dussurget, O., et al. (2008). Conjugated action of two species-specific invasion proteins for fetoplacental listeriosis. *Nature* 455, 1114–1118. doi: 10.1038/nature07303

Disson, O., Nikitas, G., Grayo, S., Dussurget, O., Cossart, P., and Lecuit, M. (2009). Modeling human listeriosis in natural and genetically engineered animals. *Nat. Protoc.* 4, 799–810. doi: 10.1038/nprot.2009.66

Douam, F., Gaska, J. M., Winer, B. Y., Ding, Q., von Schaewen, M., and Ploss, A. (2015). Genetic dissection of the host tropism of human-tropic pathogens. *Annu. Rev. Genet.* 49, 21–45. doi: 10.1146/annurev-genet-112414-054823

Fang, W., Almeida, R. A., and Oliver, S. P. (2000). Effects of lactoferrin and milk on adherence of *Streptococcus uberis* to bovine mammary epithelial cells. *Am. J. Vet. Res.* 61, 275–279. doi: 10.2460/ajvr.2000.61.275

Fang, W., and Oliver, S. P. (1999). Identification of lactoferrin-binding proteins in bovine mastitis-causing *Streptococcus uberis*. *FEMS Microbiol. Lett.* 176, 91–96. doi: 10.1111/j.1574-6968.1999.tb13647.x

Fischetti, V. A. (1989). Streptococcal M protein: molecular design and biological behavior. *Clin. Microbiol. Rev.* 2, 285–314. doi: 10.1128/cmr.2.3.285

Fitzgerald, J. R. (2012). Livestock-associated *Staphylococcus aureus*: origin, evolution and public health threat. *Trends Microbiol.* 20, 192–198. doi: 10.1016/j.tim.2012.01.006

Flick, M. J., Du, X., Prasad, J. M., Raghu, H., Palumbo, J. S., Smeds, E., et al. (2013). Genetic elimination of the binding motif on fibrinogen for the *S. aureus* virulence factor ClfA improves host survival in septicemia. *Blood* 121, 1783–1794. doi: 10.1182/blood-2012-09-453894

Friedman, D. B., Stauff, D. L., Pishchany, G., Whitwell, C. W., Torres, V. J., and Skaar, E. P. (2006). *Staphylococcus aureus* redirects central metabolism to increase iron availability. *PLoS Pathog.* 2:e87. doi: 10.1371/journal.ppat.0020087

Gaillard, J. L., Berche, P., Frehel, C., Gouln, E., and Cossart, P. (1991). Entry of *L. monocytogenes* into cells is mediated by internalin, a repeat protein reminiscent of surface antigens from gram-positive cocci. *Cell* 65, 1127–1141. doi: 10.1016/0092-8674(91)90009-N

Gaudin, C. F. M., Grigg, J. C., Arrieta, A. L., and Murphy, M. E. P. (2011). Unique heme-iron coordination by the hemoglobin receptor IsdB of *Staphylococcus aureus*. *Biochemistry* 50, 5443–5452. doi: 10.1021/bi200369p

Geoghegan, J. A., Ganesh, V. K., Smeds, E., Liang, X., Höök, M., and Foster, T. J. (2010). Molecular characterization of the interaction of staphylococcal microbial surface components recognizing adhesive matrix molecules (MSCRAMM) ClfA and Fbl with fibrinogen. *J. Biol. Chem.* 285, 6208–6216. doi: 10.1074/jbc.M109.062208

Gianquinto, E., Moscetti, I., De Bei, O., Campanini, B., Marchetti, M., Luque, F. J., et al. (2019). Interaction of human hemoglobin and semi-hemoglobins with the *Staphylococcus aureus* hemophore IsdB: a kinetic and mechanistic insight. *Sci. Rep.* 9:18629. doi: 10.1038/s41598-019-54970-w

Guinane, C. M., Ben Zakour, N. L., Tormo-Mas, M. A., Weinert, L. A., Lowder, B. V., Cartwright, R. A., et al. (2010). Evolutionary genomics of *Staphylococcus aureus* reveals insights into the origin and molecular basis of ruminant host adaptation. *Genome Biol. Evol.* 2, 454–466. doi: 10.1093/gbe/evq031

Haag, A. F., Fitzgerald, J. R., and Penadés, J. R. (2019). *Staphylococcus aureus* in animals. *Microbiol. Spect.* 7, 1–19. doi: 10.1128/microbiolspec.GPP3-0060-2019

Hagiwara, S., Kawai, K., Anri, A., and Nagahata, H. (2003). Lactoferrin concentrations in milk from normal and subclinical mastitic cows. *J. Vet. Med. Sci.* 65, 319–323. doi: 10.1292/jvms.65.319

Harmon, R. J., Schanbacher, F. L., Ferguson, L. C., and Smith, K. L. (1976). Changes in lactoferrin, immunoglobulin G, bovine serum albumin, and alpha-lactalbumin during acute experimental and natural coliform mastitis in cows. *Infect. Immun.* 13, 533–542.

Herron-Olson, L., Fitzgerald, J. R., Musser, J. M., and Kapur, V. (2007). Molecular correlates of host specialization in *Staphylococcus aureus*. *PLoS One* 2:e1120. doi: 10.1371/journal.pone.0001120

Higgins, J., Loughman, A., Van Kessel, K. P. M., Van Strijp, J. A. G., and Foster, T. J. (2006). Clumping factor A of *Staphylococcus aureus* inhibits phagocytosis by human polymorphonuclear leucocytes. *FEMS Microbiol. Lett.* 258, 290–296. doi: 10.1111/j.1574-6968.2006.00229.x

Horstmann, R. D., Sievertsen, H. J., Knobloch, J., and Fischetti, V. A. (1988). Antiphagocytic activity of streptococcal M protein: selective binding of complement control protein factor H. *Proc. Natl. Acad. Sci. U.S.A.* 85, 1657–1661. doi: 10.1073/pnas.85.5.1657

Iakobachvili, N., and Peters, P. J. (2017). Humans in a dish: the potential of organoids in modeling immunity and infectious diseases. *Front. Microbiol.* 8:2402. doi: 10.3389/fmicb.2017.02402

Inaba, Y., Tsurifune, K., Sasazaki, S., and Mannen, H. (2008). A new tandem repeat in bovine fibrinogen Aα gene. *Anim. Biotechnol.* 19, 122–126. doi: 10.1080/10495390801929122

Jochim, A., Adolf, L., Belikova, D., Schilling, N. A., Setyawati, I., Chin, D., et al. (2020). An ECF-type transporter scavenges heme to overcome iron-limitation in *Staphylococcus lugdunensis*. *eLife* 9:e57322. doi: 10.7554/eLife.57322

Josse, J., Laurent, F., and Diot, A. (2017). Staphylococcal adhesion and host cell invasion: fibronectin-binding and other mechanisms. *Front. Microbiol.* 8:2433. doi: 10.3389/fmicb.2017.02433

Ko, Y.-P., and Flick, M. J. (2016). Fibrinogen is at the interface of host defense and pathogen virulence in *Staphylococcus aureus* infection. *Semin. Thromb. Hemost.* 42, 408–421. doi: 10.1055/s-0036-1579635

Laabei, M., and Ermert, D. (2019). Catch me if you can: *Streptococcus pyogenes* complement evasion strategies. *J. Innate Immun.* 11, 3–12. doi: 10.1159/000492944

Lecuit, M., and Cossart, P. (2002). Genetically-modified-animal models for human infections: the *Listeria paradigm*. *Trends Mol. Med.* 8, 537–542. doi: 10.1016/S1471-4914(02)02413-9

Lecuit, M., Dramsi, S., Gottardi, C., Fedor-Chaiken, M., Gumbiner, B., and Cossart, P. (1999). A single amino acid in E-cadherin responsible for host specificity towards the human pathogen *Listeria monocytogenes*. *EMBO J.* 18, 3956–3963. doi: 10.1093/emboj/18.14.3956

Lecuit, M., Nelson, D. M., Smith, S. D., Khun, H., Huerre, M., Vacher-Lavenu, M.-C., et al. (2004). Targeting and crossing of the human maternofetal barrier by *Listeria monocytogenes*: role of internalin interaction with trophoblast E-cadherin. *Proc. Natl. Acad. Sci. U.S.A.* 101, 6152–6157. doi: 10.1073/pnas.0401434101

Lecuit, M., Vandormael-Pournin, S., Lefort, J., Huerre, M., Gounon, P., Dupuy, C., et al. (2001). A transgenic model for listeriosis: role of internalin in crossing the intestinal barrier. *Science* 292, 1722–1725. doi: 10.1126/science.1059852

Lewis, M. J., Meehan, M., Owen, P., and Woof, J. M. (2008). A common theme in interaction of bacterial immunoglobulin-binding proteins with immunoglobulins illustrated in the equine system. *J. Biol. Chem.* 283, 17615–17623. doi: 10.1074/jbc.M709844200

Lövkvist, L., Sjölinder, H., Wehelie, R., Aro, H., Norrby-Teglund, A., Plant, L., et al. (2008). CD46 contributes to the severity of group a streptococcal infection. *Infect. Immun.* 76, 3951–3958. doi: 10.1128/iai.00109-08

Lowder, B. V., Guinane, C. M., Ben Zakour, N. L., Weinert, L. A., Conway-Morris, A., Cartwright, R. A., et al. (2009). Recent human-to-poultry host jump, adaptation, and pandemic spread of *Staphylococcus aureus*. *Proc. Natl. Acad. Sci. U.S.A.* 106, 19545–19550. doi: 10.1073/pnas.0909285106

Lu, J., Francis, J., Doster, R. S., Haley, K. P., Craft, K. M., Moore, R. E., et al. (2020). Lactoferrin: a critical mediator of both host immune response and antimicrobial activity in response to streptococcal infections. *ACS Infect. Dis.* 6, 1615–1623. doi: 10.1021/acsinfecdis.0c00050

Matsui, H., Sekiya, Y., Nakamura, M., Murayama, S. Y., Yoshida, H., Takahashi, T., et al. (2009). CD46 transgenic mouse model of necrotizing fasciitis caused by

Streptococcus pyogenes infection. *Infect. Immun.* 77, 4806–4814. doi: 10.1128/iai.00577-09

McDevitt, D., Nanavaty, T., House-Pompeo, K., Bell, E., Turner, N., Mcintire, L., et al. (1997). Characterization of the interaction between the *Staphylococcus aureus* clumping factor (ClfA) and fibrinogen. *Eur. J. Biochem.* 247, 416–424. doi: 10.1111/j.1432-1033.1997.00416.x

Meehan, M., Lynagh, Y., Woods, C., and Owen, P. (2001). The fibrinogen-binding protein (FgBP) of *Streptococcus equi* subsp. equi additionally binds IgG and contributes to virulence in a mouse model. *Microbiology* 147, 3311–3322. doi: 10.1099/00221287-147-12-3311

Meehan, M., Muldowney, D. A., Watkins, N. J., and Owen, P. (2000). Localization and characterization of the ligand-binding domain of the fibrinogen-binding protein (FgBP) of *Streptococcus equi* subsp. equi. *Microbiology* 146, 1187–1194. doi: 10.1099/00221287-146-5-1187

Meehan, M., Nowlan, P., and Owen, P. (1998). Affinity purification and characterization of a fibrinogen-binding protein complex which protects mice against lethal challenge with *Streptococcus equi* subsp. equi. *Microbiology* 144, 993–1003. doi: 10.1099/00221287-144-4-993

Mengaud, J., Ohayon, H., Gounon, P., Mège, R.-M., and Cossart, P. (1996). E-Cadherin is the receptor for internalin, a surface protein required for entry of *L. monocytogenes* into epithelial cells. *Cell* 84, 923–932. doi: 10.1016/S0092-8674(00)81070-3

Miajlovic, H., Loughman, A., Brennan, M., Cox, D., and Foster, T. J. (2007). Both complement- and fibrinogen-dependent mechanisms contribute to platelet aggregation mediated by *Staphylococcus aureus* clumping factor B. *Infect. Immun.* 75, 3335–3343. doi: 10.1128/iai.01993-06

Mosesson, M. W. (2005). Fibrinogen and fibrin structure and functions. *J. Thromb. Haemost.* 3, 1894–1904. doi: 10.1111/j.1538-7836.2005.01365.x

Moshynskyy, I., Jiang, M., Fontaine, M. C., Perez-Casal, J., Babiuk, L. A., and Potter, A. A. (2003). Characterization of a bovine lactoferrin binding protein of *Streptococcus uberis*. *Microb. Pathog.* 35, 203–215. doi: 10.1016/S0882-4010(03)00150-5

Murakawa, M., Okamura, T., Kamura, T., Shibuya, T., Harada, M., and Niho, Y. (1993). Diversity of primary structures of the carboxy-terminal regions of mammalian fibrinogen A alpha-chains. Characterization of the partial nucleotide and deduced amino acid sequences in five mammalian species; rhesus monkey, pig, dog, mouse and Syrian hamster. *Thromb. Haemost.* 69, 351–360.

Murray, E. G. D., Webb, R. A., and Swann, M. B. R. (1926). A disease of rabbits characterised by a large mononuclear leucocytosis, caused by a hitherto undescribed bacillus *Bacterium monocytogenes* (n.sp.). *J. Pathol. Bacteriol.* 29, 407–439. doi: 10.1002/path.1700290409

Muryoi, N., Tiedemann, M. T., Pluym, M., Cheung, J., Heinrichs, D. E., and Stillman, M. J. (2008). Demonstration of the iron-regulated surface determinant (Isd) heme transfer pathway in *Staphylococcus aureus*. *J. Biol. Chem.* 283, 28125–28136. doi: 10.1074/jbc.M802171200

Nikitas, G., Deschamps, C., Disson, O., Niault, T., Cossart, P., and Lecuit, M. (2011). Transcytosis of *Listeria monocytogenes* across the intestinal barrier upon specific targeting of goblet cell accessible E-cadherin. *J. Exp. Med.* 208, 2263–2277. doi: 10.1084/jem.20110560

Okada, N., Liszewski, M. K., Atkinson, J. P., and Caparon, M. (1995). Membrane cofactor protein (CD46) is a keratinocyte receptor for the M protein of the group A *Streptococcus*. *Proc. Natl. Acad. Sci. U.S.A.* 92, 2489–2493. doi: 10.1073/pnas.92.7.2489

Pasquina, L., Santa Maria, J. P., McKay Wood, B., Moussa, S. H., Matano, L. M., Santiago, M., et al. (2016). A synthetic lethal approach for compound and target identification in *Staphylococcus aureus*. *Nat. Chem. Biol.* 12, 40–45. doi: 10.1038/nchembio.1967

Patel, D., Almeida, R. A., Dunlap, J. R., and Oliver, S. P. (2009). Bovine lactoferrin serves as a molecular bridge for internalization of *Streptococcus uberis* into bovine mammary epithelial cells. *Vet. Microbiol.* 137, 297–301. doi: 10.1016/j.vetmic.2009.01.013

Pentecost, M., Otto, G., Theriot, J. A., and Amieva, M. R. (2006). Listeria monocytogenes invades the epithelial junctions at sites of cell extrusion. *PLoS Pathog.* 2:e3. doi: 10.1371/journal.ppat.0020003

Pickering, A. C., Vitry, P., Prystopiuk, V., Garcia, B., Höök, M., Schoenebeck, J., et al. (2019). Host-specialized fibrinogen-binding by a bacterial surface protein promotes biofilm formation and innate immune evasion. *PLoS Pathog.* 15:e1007816. doi: 10.1371/journal.ppat.1007816

Pietrocola, G., Geoghegan, J. A., Rindi, S., Di Poto, A., Missineo, A., Consalvi, V., et al. (2013). Molecular characterization of the multiple interactions of SpsD, a surface protein from *Staphylococcus pseudintermedius*, with host extracellular matrix proteins. *PLoS One* 8:e66901. doi: 10.1371/journal.pone.0066901

Pishchany, G., McCoy, A. L., Torres, V. J., Krause, J. C., Crowe, J. E., Fabry, M. E., et al. (2010). Specificity for human hemoglobin enhances *Staphylococcus aureus* infection. *Cell Host Microbe* 8, 544–550. doi: 10.1016/j.chom.2010.11.002

Qin, W., Kutny, P. M., Maser, R. S., Dion, S. L., Lamont, J. D., Zhang, Y., et al. (2016). Generating mouse models using CRISPR-Cas9-mediated genome editing. *Curr. Protoc. Mouse Biol.* 6, 39–66. doi: 10.1002/9780470942390.mo150178

Resch, G., François, P., Morisset, D., Stojanov, M., Bonetti, E. J., Schrenzel, J., et al. (2013). Human-to-bovine jump of *Staphylococcus aureus* CC8 is associated with the loss of a β-hemolysin converting prophage and the acquisition of a new staphylococcal cassette chromosome. *PLoS One* 8:e58187. doi: 10.1371/journal.pone.0058187

Rezcallah, M. S., Hodges, K., Gill, D. B., Atkinson, J. P., Wang, B., and Cleary, P. P. (2005). Engagement of CD46 and α5β1 integrin by group A streptococci is required for efficient invasion of epithelial cells. *Cell. Microbiol.* 7, 645–653. doi: 10.1111/j.1462-5822.2004.00497.x

Ribet, D., and Cossart, P. (2015). How bacterial pathogens colonize their hosts and invade deeper tissues. *Microb. Infect.* 17, 173–183. doi: 10.1016/j.micinf.2015.01.004

Richardson, E. J., Bacigalupe, R., Harrison, E. M., Weinert, L. A., Lycett, S., Vrieling, M., et al. (2018). Gene exchange drives the ecological success of a multi-host bacterial pathogen. *Nat. Ecol. Evol.* 2, 1468–1478. doi: 10.1038/s41559-018-0617-0

Riley-Vargas, R. C., Gill, D. B., Kemper, C., Liszewski, M. K., and Atkinson, J. P. (2004). CD46: expanding beyond complement regulation. *Trends Immunol.* 25, 496–503. doi: 10.1016/j.it.2004.07.004

Roulo, R. M., Fishburn, J. D., Amosu, M., Etchison, A. R., and Smith, M. A. (2014). Dose response of *Listeria monocytogenes* invasion, fetal morbidity, and fetal mortality after oral challenge in pregnant and nonpregnant Mongolian gerbils. *Infect. Immun.* 82, 4834–4841. doi: 10.1128/IAI.01514-14

Schubert, W.-D., Urbanke, C., Ziehm, T., Beier, V., Machner, M. P., Domann, E., et al. (2002). Structure of internalin, a major invasion protein of *Listeria monocytogenes*, in complex with its human receptor E-cadherin. *Cell* 111, 825–836. doi: 10.1016/S0092-8674(02)01136-4

Seo, H. S., Minasov, G., Seepersaud, R., Doran, K. S., Dubrovska, I., Shuvalova, L., et al. (2013). Characterization of fibrinogen binding by glycoproteins Srr1 and Srr2 of *Streptococcus agalactiae*. *J. Biol. Chem.* 288, 35982–35996. doi: 10.1074/jbc.M113.513358

Sheldon, J. R., and Heinrichs, D. E. (2015). Recent developments in understanding the iron acquisition strategies of gram positive pathogens. *FEMS Microbiol. Rev.* 39, 592–630. doi: 10.1093/femsre/fuv009

Sheppard, S. K., Guttman, D. S., and Fitzgerald, J. R. (2018). Population genomics of bacterial host adaptation. *Nat. Rev. Genet.* 19, 549–565. doi: 10.1038/s41576-018-0032-z

Spoor, L. E., Richardson, E., Richards, A. C., Wilson, G. J., Mendonca, C., Gupta, R. K., et al. (2015). Recombination-mediated remodelling of host-pathogen interactions during *Staphylococcus aureus* niche adaptation. *Microb. Genom.* 1:e000036. doi: 10.1099/mgen.0.000036

Su, X., Cao, G., Zhang, J., Pan, H., Zhang, D., Kuang, D., et al. (2019). Characterization of internalin genes in *Listeria monocytogenes* from food and humans, and their association with the invasion of Caco-2 cells. *Gut Pathog.* 11:30. doi: 10.1186/s13099-019-0307-8

Sumbal, J., Chiche, A., Charifou, E., Koledova, Z., and Li, H. (2020). Primary mammary organoid model of lactation and involution. *Front. Cell Dev. Biol.* 8:68. doi: 10.3389/fcell.2020.00068

Thern, A., Stenberg, L., Dahlbäck, B., and Lindahl, G. (1995). Ig-binding surface proteins of *Streptococcus pyogenes* also bind human C4b-binding protein (C4BP), a regulatory component of the complement system. *J. Immunol.* 154, 375–386.

Thomer, L., Schneewind, O., and Missiakas, D. (2016). Pathogenesis of *Staphylococcus aureus* bloodstream infections. *Annu. Rev. Pathol.* 11, 343–364. doi: 10.1146/annurev-pathol-012615-044351

Torres, V. J., Pishchany, G., Humayun, M., Schneewind, O., and Skaar, E. P. (2006). *Staphylococcus aureus* IsdB is a hemoglobin receptor required for heme iron utilization. *J. Bacteriol.* 188, 8421–8429. doi: 10.1128/jb.01335-06

Valenti, P., and Antonini, G. (2005). Lactoferrin: an important host defence against microbial and viral attack. *Cell. Mol. Life Sci.* 62:2576. doi: 10.1007/s00018-005-5372-0

Van Stelten, A., Simpson, J. M., Chen, Y., Scott, V. N., Whiting, R. C., Ross, W. H., et al. (2011). Significant shift in median guinea pig infectious dose shown by an outbreak-associated *Listeria monocytogenes* epidemic clone strain and a strain carrying a premature stop codon mutation in inlA. *Appl. Environ. Microbiol.* 77, 2479–2487. doi: 10.1128/aem.02626-10

Viana, D., Comos, M., McAdam, P. R., Ward, M. J., Selva, L., Guinane, C. M., et al. (2015). A single natural nucleotide mutation alters bacterial pathogen host tropism. *Nat. Genet.* 47, 361–366. doi: 10.1038/ng.3219

Walsh, E. J., Miajlovic, H., Gorkun, O. V., and Foster, T. J. (2008). Identification of the *Staphylococcus aureus* MSCRAMM clumping factor B (ClfB) binding site in the αC-domain of human fibrinogen. *Microbiology* 154, 550–558. doi: 10.1099/mic.0.2007/010868-0

Weinberg, E. D. (1974). Iron and susceptibility to infectious disease. *Science* 184, 952–956. doi: 10.1126/science.184.4140.952

Williams, D., Dunn, S., Richardson, A., Frank, J. F., and Smith, M. A. (2011). Time course of fetal tissue invasion by *Listeria monocytogenes* following an oral inoculation in pregnant guinea pigs. *J. Food Prot.* 74, 248–253. doi: 10.4315/0362-028x.Jfp-10-163

Wollert, T., Pasche, B., Rochon, M., Deppenmeier, S., van den Heuvel, J., Gruber, A. D., et al. (2007). Extending the host range of *Listeria monocytogenes* by rational protein design. *Cell* 129, 891–902. doi: 10.1016/j.cell.2007.03.049

Woolhouse, M. E. J., Taylor, L. H., and Haydon, D. T. (2001). Population biology of multihost pathogens. *Science* 292, 1109–1112. doi: 10.1126/science.1059026

Xiang, H., Feng, Y., Wang, J., Liu, B., Chen, Y., Liu, L., et al. (2012). Crystal structures reveal the multi-ligand binding mechanism of *Staphylococcus aureus* ClfB. *PLoS Pathog.* 8:e1002751. doi: 10.1371/journal.ppat.1002751

Zapotoczna, M., Heilbronner, S., Speziale, P., and Foster, T. J. (2012). Iron-regulated surface determinant (Isd) proteins of *Staphylococcus lugdunensis*. *J. Bacteriol.* 194, 6453–6467. doi: 10.1128/jb.01195-12

Zhu, H., Xie, G., Liu, M., Olson, J. S., Fabian, M., Dooley, D. M., et al. (2008). Pathway for heme uptake from human methemoglobin by the iron-regulated surface determinants system of *Staphylococcus aureus*. *J. Biol. Chem.* 283, 18450–18460. doi: 10.1074/jbc.M801466200

Zimin, A. V., Delcher, A. L., Florea, L., Kelley, D. R., Schatz, M. C., Puiu, D., et al. (2009). A whole-genome assembly of the domestic cow, *Bos taurus*. *Genome Biol.* 10:R42. doi: 10.1186/gb-2009-10-4-r42

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Microbiology

MINI REVIEW
published: 30 October 2020
doi: 10.3389/fmicb.2020.593356

Streptococcal Serine-Rich Repeat Proteins in Colonization and Disease

Jia Mun Chan[1], Andrea Gori[1], Angela H. Nobbs[2] and Robert S. Heyderman[1]*

[1] NIHR Mucosal Pathogens Research Unit, Division of Infection and Immunity, University College London, London, United Kingdom, [2] Bristol Dental School, University of Bristol, Bristol, United Kingdom

Glycosylation of proteins, previously thought to be absent in prokaryotes, is increasingly recognized as important for both bacterial colonization and pathogenesis. For mucosal pathobionts, glycoproteins that function as cell wall-associated adhesins facilitate interactions with mucosal surfaces, permitting persistent adherence, invasion of deeper tissues and transition to disease. This is exemplified by *Streptococcus pneumoniae* and *Streptococcus agalactiae*, which can switch from being relatively harmless members of the mucosal tract microbiota to bona fide pathogens that cause life-threatening diseases. As part of their armamentarium of virulence factors, streptococci encode a family of large, glycosylated serine-rich repeat proteins (SRRPs) that facilitate binding to various tissue types and extracellular matrix proteins. This minireview focuses on the roles of *S. pneumoniae* and *S. agalactiae* SRRPs in persistent colonization and the transition to disease. The potential of utilizing SRRPs as vaccine targets will also be discussed.

Keywords: bacterial glycoproteins, *Streptococcus pneumoniae*, *Streptococcus agalactiae*, serine-rich repeat proteins, pathogenesis, colonization

OPEN ACCESS

Edited by:
Magnus Hook,
Texas A&M Health Science Center, United States

Reviewed by:
Carlos J. Orihuela,
University of Alabama at Birmingham, United States
Jennifer Angeline Gaddy,
Vanderbilt University, United States

****Correspondence:***
Jia Mun Chan
rekgjm0@ucl.ac.uk

Specialty section:
This article was submitted to Infectious Diseases, a section of the journal Frontiers in Microbiology

Received: *10 August 2020*
Accepted: *12 October 2020*
Published: *30 October 2020*

Citation:
Chan JM, Gori A, Nobbs AH and Heyderman RS (2020) Streptococcal Serine-Rich Repeat Proteins in Colonization and Disease. Front. Microbiol. 11:593356. doi: 10.3389/fmicb.2020.593356

INTRODUCTION

It is well-established that glycosylation of eukaryotic proteins is important for regulating cellular processes, including receptor signaling and inflammation (Reily et al., 2019). In contrast, prokaryotic protein glycosylation was thought to be rare until the discovery of general *N*-linked and *O*-linked protein glycosylation pathways in *Campylobacter jejuni, Bacteroides fragilis*, and *Burkholderia cenocepacia*, revealing the role of glycoproteins in bacterial physiology and pathogenesis (Szymanski et al., 1999; Coyne et al., 2013; Mohamed et al., 2019). Many bacterial glycoproteins identified thus far are surface proteins that mediate host-microbe interactions and motility (Tytgat and Lebeer, 2014; Tan et al., 2015; Valguarnera et al., 2016; Eichler and Koomey, 2017; Bhat et al., 2019).

Members of the Gram-positive genera *Streptococcus*, *Staphylococcus,* and *Lactobacillus* express a family of large, glycosylated serine-rich repeat proteins (SRRPs) (Latousakis et al., 2020). They were first characterized as fimbrial-like or platelet-binding proteins in *Streptococcus parasanguinis, Streptococcus cristatus, Streptococcus gordonii,* and *Staphylococcus aureus* (Correia et al., 1997; Wu et al., 1998; Takahashi et al., 2002; Handley et al., 2005; Siboo et al., 2005). SRRPs mediate adhesion to sialic acid, fibrinogen, keratin or as yet unidentified molecules on other bacteria in a strain- or species-dependent manner (Deng et al., 2014; Six et al., 2015; Bensing et al., 2019; Latousakis et al., 2019). The versatility in binding partners suggests that SRRPs facilitate colonization of multiple and diverse niches, which in the case of oral streptococci permits these bacteria to persist in the mouth and to form infective vegetations on damaged heart valves (Yumoto et al., 2019).

Streptococcus pneumoniae and *S. agalactiae* (Group B *Streptococcus*, GBS) are important pathobionts of mucosal surfaces that disproportionately affect young children, the immunocompromised and the elderly. Studies estimate that pneumococci and GBS were respectively responsible for the death of around 300,000 and 90,000 children under 5 in 2015 (Seale et al., 2017; Wahl et al., 2018). Despite their notoriety as deadly pathogens, they are commonly found as asymptomatic colonizers of human mucosal surfaces, and may also cause milder diseases such as otitis media (*S. pneumoniae*) or urinary tract infections *(S. agalactiae)* (Feldman and Anderson, 2019; McLaughlin et al., 2020). Both *S. pneumoniae* and *S. agalactiae* have large accessory genomes, leading to considerable variation in pathogenic potential between serotypes and sequence types (Hiller and Sá-Leão, 2018; Chen, 2019). Genes encoding SRRPs are part of the accessory genome of both pneumococci and *S. agalactiae* (Tettelin et al., 2005; Gámez et al., 2018).

Research on SRRPs thus far has predominantly focused on oral streptococci, where these glycoproteins promote biofilm formation, intra- and interspecies aggregation and development of infective endocarditis (Zhou and Wu, 2009; Lizcano et al., 2012; Zhu et al., 2015; Zhu and Wu, 2016; Latousakis and Juge, 2018; Latousakis et al., 2020). However, there is growing evidence that SRRPs in *S. pneumoniae* and *S. agalactiae* may be important for the transition from asymptomatic carriage to disease. This minireview will focus on the functions of these SRRPs in colonization and disease. We will draw on evidence from studies of oral streptococci and closely related bacteria to compare and contrast the diversity, biogenesis and functions of SRRPs in streptococcal biology. Finally, we will discuss the potential of using SRRPs as vaccine targets for *S. pneumoniae* and *S. agalactiae*.

ARCHITECTURE OF STREPTOCOCCAL SRRPs

Serine-rich repeat proteins are characterized by the presence of (i) an extended N-terminal signal sequence, which facilitates transport of the protein through an accessory secretion system; (ii) two highly glycosylated serine-rich repeat regions (SRR); (iii) at least one non-repeat binding region (BR); and (iv) a C-terminal cell wall anchoring domain carrying a LPxTG motif (**Figure 1A**). Despite architectural conservation, SRRPs share little sequence homology especially in the sequences and lengths of the SRR and BR domains, resulting in significant variation in protein size and binding partners (Zhou and Wu, 2009; Lizcano et al., 2012; Bensing et al., 2019). The streptococcal SRRP preprotein ranges in size from 970 amino acids (*S. agalactiae* Srr1) to over 5000 amino acids (*Streptococcus oralis* subsp. *dentisani* FapC) (Seifert et al., 2006; Ronis et al., 2019).

Protein crystallography and electron microscopy show that SRRPs have a fimbriae-like structure, where the second SRR domain (SRR2) forms an extended stalk structure that projects the globular BR from the cell surface (Wu et al., 1998; Handley et al., 2005; Ramboarina et al., 2010; Lizcano et al., 2012; Six et al., 2015). Expression of a truncated version of the pneumococcal SRRP, PsrP, rescues binding defects of a *psrP* mutation in capsule-null, but not in encapsulated *S. pneumoniae* (Shivshankar et al., 2009). Thus, the length of the SRR2 region is hypothesized to have adapted to extend the BR beyond the polysaccharide capsule in encapsulated streptococci (Shivshankar et al., 2009; Lizcano et al., 2012). The SRRP BRs of *S. parasanguinis* and the gut bacterium *Lactobacillus reuteri* assume different conformations in low and high pH, thereby altering binding affinity (Garnett et al., 2012; Sequeira et al., 2018). It is not known if similar conformation shifts occur with other streptococcal SRRPs.

HETEROGENEITY OF STREPTOCOCCAL SRRPs ENCODED ON GLYCOSYLATION ISLANDS

Streptococcal SRRPs are encoded on putative genomic islands, which we will refer to as glycosylation islands (Takamatsu et al., 2004b). The minimal SRRP locus includes eight genes encoding (i) the SRRP; (ii) an accessory secretion system (*secA2, secY2, asp1, asp2, asp3*; Asp proteins are also known as Gap); and (iii) core glycosyltransferases (known in different species as *gtfAB*, *gtf1/gft2,* and *gtfE/F*) (**Figure 1B**; Zhou and Wu, 2009). Frequently, glycosylation islands include additional accessory genes. While all genes within the minimal locus are typically found on the island, *Streptococcus salivarius gtfAB* homologs are encoded elsewhere on the chromosome (Pombert et al., 2008; Couvigny et al., 2017). Mutations of any gene in the minimal locus typically impairs SRRP function and cell surface presentation (Takamatsu et al., 2004b; Wu et al., 2007; Mistou et al., 2009; Lizcano et al., 2017).

Members of the genus *Staphylococcus* encode the minimal SRRP locus consisting only of the eight genes described above (Siboo et al., 2005). In contrast, streptococcal SRRP loci may harbor up to eight additional genes encoding accessory glycosyltransferases (GTs) (**Figure 1B**; Lizcano et al., 2012). The heterogeneity in the number and types of GTs encoded on the islands suggests significant glycan diversity decorating the glycoproteins. Indeed, mass-spectrometry revealed multiple glycoforms of the *S. agalactiae* SRRP Srr1 (Chaze et al., 2014). *L. reuteri* is also reported to display strain-specific glycosylation of its SRRP (Latousakis et al., 2019). The diversity and identity of glycan structures decorating many SRRPs are still unknown because extensive biochemical and analytical chemistry techniques are needed for such identification (Gloster, 2014).

Some streptococci encode multiple SRRPs while others lack SRRPs entirely. Approximately 50% of pneumococcal strains carry *psrP*, which is more likely to be found in strains isolated from individuals with pneumonia (Selva et al., 2012). *S. agalactiae* strains carry glycosylation islands encoding either Srr1 or Srr2, but not both at once, and the Srr1 island contains five accessory genes more than the Srr2 island (**Figure 1B**; Seifert et al., 2006). Genome-wide association studies revealed that *srr2* and associated GTs are almost exclusively found in hypervirulent strains of *S. agalactiae* (Seifert et al., 2006; Six et al., 2015; Gori et al., 2020). *S. gordonii* strains also express one of two SRRPs named GspB or Hsa. Unlike *S. agalactiae*, the glycosylation

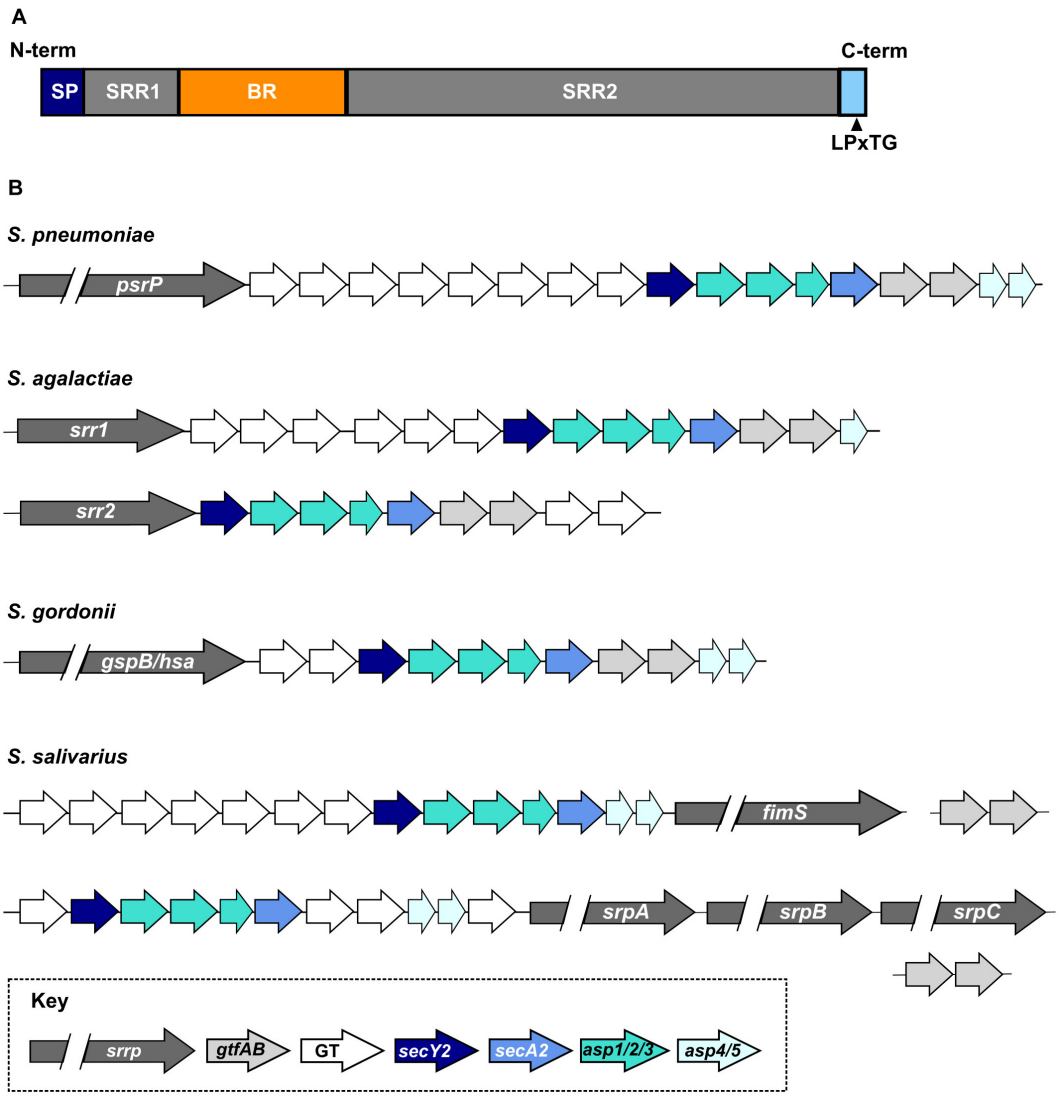


FIGURE 1 | Organization and diversity of streptococcal SRRPs. **(A)** General architecture of the SRRP polypeptide. SP, signal peptide; SRR1/2, serine-rich repeat regions 1 or 2; BR, non-repeat binding region; LPxTG, sortase dependent cell-wall anchor signal. **(B)** Diversity of the SRRP glycosylation loci encoded by a subset of streptococcal species. *S. pneumoniae* carries only one SRRP encoding allele. *S. agalactiae* and *S. gordonii* encode two major SRRP alleles; the glycosylation island associated with each allele differs in *S. agalactiae* but not in *S. gordonii*. *S. salivarius* encodes two major SRRP islands that are associated with one or three SRRP alleles. Homologs of the *S. salivarius gtfA* and *gtfB* are encoded elsewhere on the chromosome. GT refers to accessory glycosyltransferases.

island associated with *gspB* and *hsa* are otherwise identical in content (**Figure 1B**; Bensing and Sullam, 2002; Takahashi et al., 2002). *S. salivarius* and *S. oralis* subsp. *dentisani* strains may carry a glycosylation island associated with three different SRRP alleles (**Figure 1B**; Couvigny et al., 2017; Ronis et al., 2019). Neither *Streptococcus pyogenes*, the causative agent of streptococcal pharyngitis and rheumatic fever, nor *Streptococcus mutans*, a significant contributor to dental caries, encode SRRPs (Mitchell, 2003).

While all SRRPs ultimately mediate adherence, variations in binding partners most likely contribute to the heterogenous colonization sites and disease of streptococcal pathogens. Polymorphisms of *srr1*, *srr2*, *gspB*, and *psrP* result in further variations of protein size and binding specificities (Samen et al., 2007; Zhou and Wu, 2009; Lizcano et al., 2012; Bensing et al., 2019). SRRPs from *S. pneumoniae* and *S. agalactiae* bind directly to the polypeptide backbone of keratin and fibrinogen (Samen et al., 2007; Shivshankar et al., 2009; Seo et al., 2013a). In contrast, SRRPs from oral streptococci tend to bind the glycan moiety, specifically those containing a terminal sialic acid, of glycoproteins such as platelet GPIbα (Bensing et al., 2004, 2016, 2019; Plummer et al., 2005; Singh et al., 2017; Ronis et al., 2019). The lectin-like properties of oral streptococcal SRRPs may permit binding to a larger variety of host structures but with lower affinity or specificity compared to SRRPs from *S. agalactiae* and *S. pneumoniae*.

BIOGENESIS AND EXPORT OF SRRPs

Prokaryotic protein glycosylation typically occurs in an *en bloc* fashion, where the glycan is assembled prior to transfer to a protein, or sequentially, where individual sugars are added to the growing chain on a polypeptide backbone (reviewed in Dell et al., 2010; Tan et al., 2015; Li et al., 2017; Schäffer and Messner, 2017). SRRPs are sequentially glycosylated in the cytoplasm prior to export (Chen et al., 2016, 2018; Zhu et al., 2016). Due to the diversity of accessory GTs encoded in individual glycosylation islands, it is not possible to generalize the pathways from any one species. To further add to the complexity of the glycosylation pathways, studies of *S. parasanguinis* revealed a novel bifunctional GT that sequentially adds two different monosaccharides to the growing glycan chain, and an *in vitro* study of *S. pneumoniae* suggests that multiple pneumococcal GTs can utilize different sugars to create polymorphic glycan decorations (Zhang et al., 2014, 2016; Jiang et al., 2017). The biogenesis of *S. gordonii* and *S. parasanguinis* SRRPs are best studied and have been reviewed in detail; comparatively little is known about *S. pneumoniae* and *S. agalactiae* (Zhou and Wu, 2009; Zhu et al., 2015; Zhu and Wu, 2016; Schäffer and Messner, 2018).

Regulated glycosylation is important for SRRP export and function. The first step in SRRP glycosylation is the covalent attachment of an *N*-acetylglucosamine (GlcNAc) moiety to the *O*-hydroxyl group of Ser or Thr residues on the SRR regions of the protein (*O*-linked glycosylation) by the cooperative action of GtfA and GtfB. With the exception of *S. parasanguinis,* mutation of *gtfA* or *gtfB* homologs leads to a complete loss of SRRP expression or aggregates of insoluble pre-protein in the cytoplasm (Takamatsu et al., 2004a; Wu et al., 2007; Mistou et al., 2009; Lizcano et al., 2017), thereby complicating attempts to dissect the role of glycosylation in SRRP activity. To bypass this problem, many studies either examine the BR in isolation or utilize a truncated version of the protein lacking most of the SRR2 region. While useful, such approaches preclude examination of the role that SRR2 plays in ligand binding and protein conformation.

The accessory secretion system is involved in maturation and transport of SRRP to the cell surface, upon which SRRPs are anchored to the cell wall by the housekeeping sortase (Nobbs et al., 2007; Mistou et al., 2009; Turner et al., 2009; Seepersaud et al., 2010, 2012; Yen et al., 2011). Some streptococcal strains encode additional Asp proteins (Asp4 and Asp5) that share homology with the secretion system components SecE or SecG, respectively (Takamatsu et al., 2005; Braunstein et al., 2019). Disruption of Asp2 prevents *O*-acetylation of the GlcNAc moiety, promoting hyperglycosylation and reducing binding of *S. gordonii* GspB to platelets (Seepersaud et al., 2012, 2017). Parallel work in *S. parasanguinis* supports the model that the accessory secretion system facilitates maturation of the glycoprotein (Wu et al., 2007; Li et al., 2008). The presence of *O*-acetylated GlcNAc in the mature *S. agalactiae* Srr1 coupled with cytoplasmic retention of hyperglycosylated forms of the protein further suggests that differential glycosylation regulates Srr1 secretion and function (Mistou et al., 2009; Chaze et al., 2014).

PNEUMOCOCCAL PsrP FACILITATES LUNG COLONIZATION AND PNEUMONIA

The pneumococcal SRRP, PsrP, facilitates a non-inflammatory, persistent lifestyle in the nasopharynx and lungs (**Figure 2A**). PsrP has a multidomain BR that allows binding to keratin 10, fibrinogen, extracellular DNA, and other PsrP, supporting autoaggregation and biofilm formation (Rose et al., 2008; Shivshankar et al., 2009; Sanchez et al., 2010; Blanchette-Cain et al., 2013; Schulte et al., 2016). Mutation of *psrP* reduces bacterial load in the lungs and blood but not in the nasopharynx of intratracheally infected mice, partly because nasopharyngeal cells do not express keratin 10 (Rose et al., 2008; Shivshankar et al., 2009). Accordingly, immunization of mice against the BR of PsrP reduces pneumococcal burden in the lungs and blood of intranasally infected mice (Rose et al., 2008; Shivshankar et al., 2009). Mutation of *psrP* or *gtfAB* impairs biofilm formation and adhesion to lung cells, suggesting that glycosylation is crucial for PsrP function (Lizcano et al., 2017). Mutation of *psrP* also impairs biofilm formation in murine nasopharynx without altering bacterial numbers in this niche (Blanchette-Cain et al., 2013; Lizcano et al., 2017). Biofilm formation on murine nasal septa co-occurs with sloughing of ciliated cells and exposure of basement membranes, but loss of PsrP actually increases production of inflammatory cytokines such as IL-6 (Blanchette-Cain et al., 2013).

When pneumococci reach the lungs, likely mediated by other virulence factors, PsrP promotes biofilm formation and pneumonia. It is uncertain, however, whether PsrP directly mediates pneumococcal escape from the lungs into the bloodstream, or if invasion is an indirect consequence of high bacterial burden in lung biofilms. Mice infected intraperitoneally with wild type (WT) and *psrP* mutants show similar bacterial burden in the blood (Rose et al., 2008). This observation is partially explained by minimal *psrP* expression in murine blood (Shenoy et al., 2017). PsrP is also minimally expressed during planktonic growth but is significantly upregulated during stationary phase and biofilm formation (Lizcano et al., 2017). Indeed, pneumococci in biofilms are hyperadhesive and less invasive compared to the same strain grown planktonically (Blanchette-Cain et al., 2013), suggesting that PsrP may contribute to a non-invasive, hyperadhesive state that supports persistent colonization (**Figure 2A**). Further investigation of the regulation of PsrP may clarify its contribution to invasion and systemic disease.

Streptococcus agalactiae Srr1 PROMOTES ADHERENCE TO MULTIPLE HOST CELL TYPES

The *S. agalactiae* SRRP, Srr1, binds fibrinogen and keratin 4, allowing *S. agalactiae* to colonize multiple body sites (Samen et al., 2007; Seo et al., 2012). Mutation of *srr1* impairs the ability of *S. agalactiae* to bind brain endothelial cells,

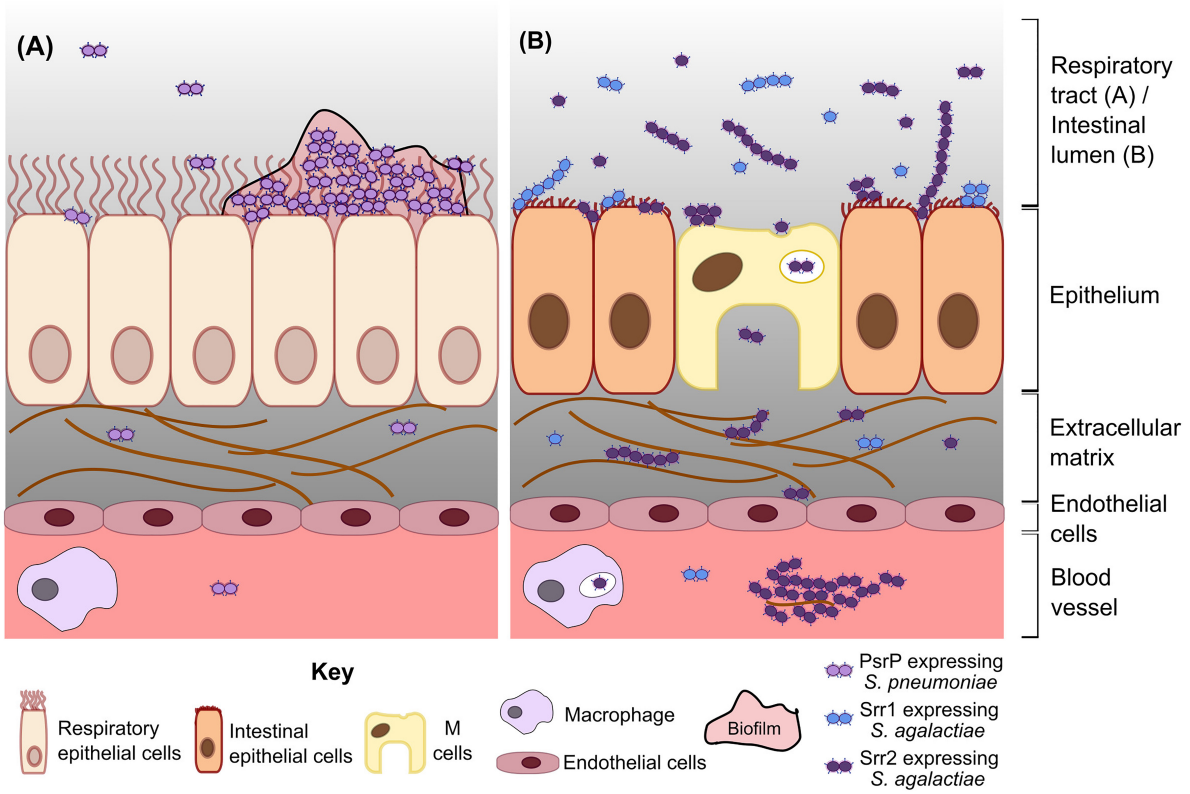


FIGURE 2 | Model of *S. pneumoniae* and *S. agalactiae* SRRP functions in colonization and disease. **(A)** Pneumococcal PsrP mediates biofilm formation and persistent colonization in the respiratory tract, indirectly resulting in invasion of deeper tissue by a subset of bacteria. **(B)** Srr1 and Srr2 expressing *S. agalactiae* (blue and purple diplococci, respectively) adhere to intestinal epithelial cells during colonization. Srr2 promotes invasion to a greater extent than Srr1, partly by mediating bacterial adherence to and transcytosis through M cells. In the bloodstream, Srr2 facilitates persistence by forming large bacterial-plasma aggregates and by increasing survival of phagocytosed bacteria. Internalized *S. agalactiae* may exploit immune cell migration to disseminate and cause other diseases, e.g., meningitis.

laryngeal and lung epithelial cells, intestinal epithelial cells, vaginal and cervical cells, and platelets (Samen et al., 2007; Mistou et al., 2009; van Sorge et al., 2009; Sheen et al., 2011; Seo et al., 2013b; Wang et al., 2014). The ability of Srr1 to bind human platelets and brain endothelial cells is a direct consequence of its capacity to bind fibrinogen on their cell surface, while adherence to the vaginal and cervical epithelium is mediated by binding to both fibrinogen and keratin 4 (Sheen et al., 2011; Seo et al., 2012, 2013b; Wang et al., 2014). *S. agalactiae srr1* mutants are carried at lower density and duration compared to WT in a murine vaginal colonization model, suggesting that Srr1 promotes persistent vaginal colonization (Sheen et al., 2011; Wang et al., 2014). Mutation of *gtfAB* abolishes Srr1 surface expression, while deletion of all six accessory glycosyltransferases (Δ*gtfC-H*) results in a smaller glycoform of Srr1 (Mistou et al., 2009). The Δ*gtfC-H* mutant also displays increased trypsin sensitivity, resulting in loss of Srr1 on the cell surface and decreased adherence to lung and intestinal epithelial cells (Mistou et al., 2009). As such, glycosylation enhances Srr1 stability by resisting proteolytic inactivation, allowing for longer durations of adherence and persistence.

Mutations of *srr1* or associated glycosyltransferases attenuate *S. agalactiae* pathogenesis. When mice or rats are infected with a *srr1* mutant, the rodents show greater survival and lower bacterial counts in the brain and spleen, with fewer cardiac vegetations in an endocarditis model (Mistou et al., 2009; van Sorge et al., 2009; Seo et al., 2012, 2013b). Rat pups infected with the Δ*gtfC-H* mutant show improved survival compared to pups infected with WT *S. agalactiae*, emphasizing the importance of Srr1 glycosylation in pathogenesis (Mistou et al., 2009). Whether Srr1 promotes upper reproductive tract infections or enhances mother-to-child transmission are currently open questions. As mentioned earlier, Srr1 promotes persistent colonization of the vaginal tract and, while frequently asymptomatic, can lead to amnionitis or bacteremia in pregnant women (Seale et al., 2017). Additionally, carriage of *S. agalactiae* by pregnant women increases the risk of transmission and invasive *S. agalactiae* disease in the neonates (Seale et al., 2017). Expression of *srr1* requires the transcription factor Rga, which also regulates the pilin subunit *pilA* (Mistou et al., 2009; Samen et al., 2011). Future studies into the temporal and environmental regulation by Rga in *S. agalactiae* may provide greater insight into the role of Srr1 in health and disease.

Streptococcus agalactiae Srr2 IS ASSOCIATED WITH INCREASED VIRULENCE AND MENINGITIS

Expression of Srr2 enhances the ability of *S. agalactiae* to cause invasive disease, especially meningitis. The *srr2* allele is predominantly carried by hypervirulent strains, which are responsible for most cases of *S. agalactiae*-induced infant meningitis (Seifert et al., 2006). Srr2 facilitates binding to brain endothelial cells, vaginal, cervical and intestinal epithelial cells, and mutation of *srr2* reduces bacterial load in the brain, liver and mesenteric lymph nodes of infected mice (Sheen et al., 2011; Seo et al., 2013a; Wang et al., 2014; Six et al., 2015; Hays et al., 2019). Srr2 is expressed at higher levels and binds fibrinogen with higher affinity than Srr1, further enhancing the adherence ability of hypervirulent strains (Seo et al., 2013a; Six et al., 2015). Unlike *srr1*, strains carrying *srr2* do not encode a *rga* equivalent, and the regulator of Srr2 expression is unknown (Mistou et al., 2009). Nonetheless, Srr2 is expressed during invasive *S. agalactiae* disease; immunohistology from a fatal case demonstrated the presence of Srr2-expressing bacterial aggregates in the infant's brain and liver (Six et al., 2015). Encouragingly, immunization of mice with recombinant Srr2 fragment protected 60–70% of mice infected intravenously with a LD_{80} dose of *S. agalactiae* (Six et al., 2015).

Expression of Srr2 promotes transcytosis through intestinal epithelium and dissemination to other body sites (**Figure 2B**). Srr2 facilitates binding to and transcytosis through murine intestinal M cells, leading to accumulation of the bacteria in mesenteric lymph nodes and translocation to the brain of mice orally gavaged with *S. agalactiae* (Hays et al., 2019). Increased adherence to intestinal cells is also likely to promote persistent colonization in the gastrointestinal tract. *S. agalactiae* expressing *srr2* forms large aggregates in the presence of plasma, which is likely mediated by the ability of Srr2 to bind the plasma proteins fibrinogen, plasminogen, and plasmin (Six et al., 2015). This aggregation increases internalization of the bacteria by macrophages and neutrophils (Six et al., 2015). However, Srr2 also promotes intracellular survival of *S. agalactiae*, potentially allowing a surviving population to hitch hike to distal sites or transmit from person to person (Six et al., 2015). In essence, Srr2 promotes bacterial migration out of the intestinal tract and to other organs, potentially by hijacking of immune cells.

DISCUSSION AND FUTURE DIRECTIONS

As presented here, there is an emerging body of evidence that *S. pneumoniae* and *S. agalactiae* SRRPs facilitate adherence to mucosal surfaces, persistent infections and the transition to disease (**Figure 2**). In general, PsrP promotes persistent colonization and pneumonia through biofilm formation but is largely dispensable during systemic infections caused by *S. pneumoniae.* Meanwhile, Srr1 and Srr2 promote colonization and systemic disease such as meningitis by *S. agalactiae*. Invasiveness may be enhanced in hypervirulent *S. agalactiae* lineages through the increased binding affinity of Srr2 to plasma components, as well as the ability of Srr2 to mediate transcytosis through intestinal M cells and disseminate throughout the body by exploiting immune cells (Six et al., 2015; Hays et al., 2019).

Given their association with colonization and pathogenesis, SRRPs have emerged as potential vaccine targets. Passive and active immunization against PsrP BR and Srr2 N-terminal region ameliorate disease and reduce bacterial burden in infected animals, which are recapitulated in studies of the *S. aureus* SRRP SraP (Zhou et al., 2019). Administration of purified Srr1 and Srr2 BRs shortly before and after inoculation with *S. agalactiae* reduces bacterial burden in a murine vaginal colonization model, suggesting that blocking SRRP-mediated binding may prevent transmission (Sheen et al., 2011). However, the recombinant protein fragments used in these studies were generated in *Escherichia coli* strains and likely not glycosylated. This prompts the question of whether glycosylated forms of SRRP would elicit stronger immune responses. Antibodies generated against native SRRPs from *S. parasanguinis* and *S. gordonii* bind to both peptide and glycan components of the proteins (Bensing and Sullam, 2002; Stephenson et al., 2002; van Sorge et al., 2009). Additionally, highly opsonic antibodies generated during natural MRSA infections specifically recognize the glycosylated domain of the *S. aureus* glycoprotein ClfA (Hazenbos et al., 2013). In recent years, there has been remarkable advancement in engineering novel glycoconjugate vaccines, but much less effort is expended on adopting bacterial glycoproteins as vaccine targets.

Current pneumococcal anticapsular vaccines are effective in reducing invasive pneumococcal disease but are less effective in controlling colonization, particularly in high carriage burden settings (Swarthout et al., 2020). Given that passive immunization against PsrP reduces bacterial burden in murine lungs (Rose et al., 2008), we speculate that adding recombinant PsrP to the vaccine formulation, either as a conjugate protein or alongside the existing vaccines may improve efficacy in reducing pneumococcal carriage and non-invasive pneumonia. Vaccines for *S. agalactiae* are still in development, but adding Srr1 and Srr2 to a multivalent vaccine may also improve efficacy. It is recognized, however, that since SRRPs are not expressed by all *S. pneumoniae* and *S. agalactiae* isolates, an SRRP vaccine component cannot be expected to confer universal protection and therefore would need to be part of a multi-component or glycoconjugate vaccine. Further investigation into the distribution, regulation, and immunogenicity of SRRPs and other similar glycoproteins will better inform of their potential utility as vaccine candidates and may identify new carrier proteins for glycoconjugate vaccines.

AUTHOR CONTRIBUTIONS

JC and RH conceptualized the review. JC planned, wrote, and revised the manuscript. AG, AN, and RH critically read and revised the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This research was funded by the National Institute for Health Research (NIHR) (project reference 16/136/46) using UK aid from the UK Government to support global health research. The views expressed in this publication are those of the author(s) and not necessarily those of the NIHR or the UK government.

REFERENCES

Bensing, B. A., Khedri, Z., Deng, L., Yu, H., Prakobphol, A., Fisher, S. J., et al. (2016). Novel aspects of sialoglycan recognition by the Siglec-like domains of streptococcal SRR glycoproteins. *Glycobiology* 26, 1221–1233. doi: 10.1093/glycob/cww042

Bensing, B. A., Li, L., Yakovenko, O., Wong, M., Barnard, K. N., Iverson, T. M., et al. (2019). Recognition of specific sialoglycan structures by oral streptococci impacts the severity of endocardial infection. *PLoS Pathog.* 15:e1007896. doi: 10.1371/journal.ppat.1007896

Bensing, B. A., López, J., and Sullam, P. M. (2004). The *Streptococcus gordonii* surface proteins GspB and Hsa mediate binding to sialylated carbohydrate epitopes on the platelet membrane glycoprotein Ibα. *Infect. Immun.* 72, 6528–6537. doi: 10.1128/IAI.72.11.6528

Bensing, B. A., and Sullam, P. M. (2002). An accessory *sec* locus of *Streptococcus gordonii* is required for export of the surface protein GspB and for normal levels of binding to human platelets. *Mol. Microbiol.* 44, 1081–1094. doi: 10.1046/j.1365-2958.2002.02949.x

Bhat, A. H., Maity, S., Giri, K., and Ambatipudi, K. (2019). Protein glycosylation: sweet or bitter for bacterial pathogens? *Crit. Rev. Microbiol.* 45, 82–102. doi: 10.1080/1040841X.2018.1547681

Blanchette-Cain, K., Hinojosa, C. A., Babu, R. A. S., Lizcano, A., Gonzalez-Juarbe, N., Munoz-Almagro, C., et al. (2013). *Streptococcus pneumoniae* biofilm formation is strain dependent, multifactorial, and associated with reduced invasiveness and immunoreactivity during colonization. *mBio* 4:e00745-13.

Braunstein, M., Bensing, B. A., and Sullam, P. M. (2019). "The two distinct types of SecA2-dependent export systems," in *Protein Secretion in Bacteria*, eds M. Sandkvist, E. Cascales, and P. J. Christie (Washington, DC: American Society for Microbiology), 29–41. doi: 10.1128/microbiolspec.PSIB-0025-2018

Chaze, T., Guillot, A., Valot, B., Langella, O., Chamot-Rooke, J., Di Guilmi, A.-M., et al. (2014). *O*-glycosylation of the *N*-terminal region of the serine-rich adhesin Srr1 of *Streptococcus agalactiae* explored by mass spectrometry. *Mol. Cell. Proteomics* 13, 2168–2182. doi: 10.1074/mcp.M114.038075

Chen, S. L. (2019). Genomic insights into the distribution and evolution of group B Streptococcus. *Front. Microbiol.* 10:1447. doi: 10.3389/fmicb.2019.01447

Chen, Y., Bensing, B. A., Seepersaud, R., Mi, W., Liao, M., Jeffrey, P. D., et al. (2018). Unraveling the sequence of cytosolic reactions in the export of GspB adhesin from *Streptococcus gordonii*. *J. Biol. Chem.* 293, 5360–5373. doi: 10.1074/jbc.RA117.000963

Chen, Y., Seepersaud, R., Bensing, B. A., Sullam, P. M., and Rapoport, T. A. (2016). Mechanism of a cytosolic *O*-glycosyltransferase essential for the synthesis of a bacterial adhesion protein. *Proc. Natl. Acad. Sci. U.S.A.* 113, E1190–E1199. doi: 10.1073/pnas.1600494113

Correia, F. F., Lamont, R., Bayer, M. E., Rosan, B., and DiRienzo, J. M. (1997). Cloning and sequencing of a mutated locus that affects fimbrial tuft organization and corncob formation in *Streptococcus crista* CC5A. *Int. J. Oral Biol.* 22, 241–248.

Couvigny, B., Lapaque, N., Rigottier-Gois, L., Guillot, A., Chat, S., Meylheuc, T., et al. (2017). Three glycosylated serine-rich repeat proteins play a pivotal role in adhesion and colonization of the pioneer commensal bacterium, *Streptococcus salivarius*. *Environ. Microbiol.* 19, 3579–3594. doi: 10.1111/1462-2920.13853

Coyne, M. J., Fletcher, C. M., Chatzidaki-Livanis, M., Posch, G., Schaffer, C., and Comstock, L. E. (2013). Phylum-wide general protein *O*-glycosylation system of the Bacteroidetes. *Mol. Microbiol.* 88, 772–783. doi: 10.1111/mmi.12220

Dell, A., Galadari, A., Sastre, F., and Hitchen, P. (2010). Similarities and differences in the glycosylation mechanisms in prokaryotes and eukaryotes. *Int. J. Microbiol.* 2010:148178. doi: 10.1155/2010/148178

Deng, L., Bensing, B. A., Thamadilok, S., Yu, H., Lau, K., Chen, X., et al. (2014). Oral streptococci utilize a Siglec-like domain of serine-rich repeat adhesins to preferentially target platelet sialoglycans in human blood. *PLoS Pathog.* 10:e1004540. doi: 10.1371/journal.ppat.1004540

Eichler, J., and Koomey, M. (2017). Sweet new roles for protein glycosylation in prokaryotes. *Trends Microbiol.* 25, 662–672. doi: 10.1016/j.tim.2017.03.001

Feldman, C., and Anderson, R. (2019). Epidemiology, virulence factors and management of the pneumococcus. *F1000Res.* 5:2320. doi: 10.12688/f1000research.9283.1

Gámez, G., Castro, A., Gómez-Mejia, A., Gallego, M., Bedoya, A., Camargo, M., et al. (2018). The variome of pneumococcal virulence factors and regulators. *BMC Genomics* 19:10. doi: 10.1186/s12864-017-4376-0

Garnett, J. A., Simpson, P. J., Taylor, J., Benjamin, S. V., Tagliaferri, C., Cota, E., et al. (2012). Structural insight into the role of *Streptococcus parasanguinis* Fap1 within oral biofilm formation. *Biochem. Biophys. Res. Commun.* 417, 421–426. doi: 10.1016/j.bbrc.2011.11.131

Gloster, T. M. (2014). Advances in understanding glycosyltransferases from a structural perspective. *Curr. Opin. Struct. Biol.* 28, 131–141. doi: 10.1016/j.sbi.2014.08.012

Gori, A., Harisson, O. B., Mlia, E., Nishihara, Y., Chan, J. M., Msefula, J., et al. (2020). Pan-GWAS of *Streptococcus agalactiae* highlights lineage-specific genes associated with virulence and niche adaptation. *mBio* 11:e00728-20.

Handley, P. S., Correia, F. F., Russell, K., Rosan, B., and DiRienzo, J. (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. doi: 10.1111/j.1399-302x.2004.00190.x

Hays, C., Touak, G., Bouaboud, A., Fouet, A., Guignot, J., Poyart, C., et al. (2019). Perinatal hormones favor CC17 group B *Streptococcus* intestinal translocation through M cells and hypervirulence in neonates. *eLife* 8:e48772.

Hazenbos, W. L. W., Kajihara, K. K., Vandlen, R., Morisaki, J. H., Lehar, S. M., Kwakkenbos, M. J., et al. (2013). Novel staphylococcal glycosyltransferases SdgA and SdgB mediate immunogenicity and protection of virulence-associated cell wall proteins. *PLoS Pathog.* 9:e1003653. doi: 10.1371/journal.ppat.1003653

Hiller, N. L., and Sá-Leão, R. (2018). Puzzling over the pneumococcal pangenome. *Front. Microbiol.* 9:2580. doi: 10.3389/fmicb.2018.02580

Jiang, Y.-L., Jin, H., Yang, H.-B., Zhao, R., Wang, S., Chen, Y., et al. (2017). Defining the enzymatic pathway for polymorphic *O*-glycosylation of the pneumococcal serine-rich repeat protein PsrP. *J. Biol. Chem.* 292, 6213–6224. doi: 10.1074/jbc.M116.770446

Latousakis, D., and Juge, N. (2018). How sweet are our gut beneficial bacteria? A focus on protein glycosylation in *Lactobacillus*. *Int. J. Mol. Sci.* 19:136. doi: 10.3390/ijms19010136

Latousakis, D., Mackenzie, D. A., Telatin, A., and Juge, N. (2020). Serine-rich repeat proteins from gut microbes. *Gut Microbes* 11, 102–117. doi: 10.1080/19490976.2019.1602428

Latousakis, D., Nepravishta, R., Rejzek, M., Wegmann, U., Le Gall, G., Kavanaugh, D., et al. (2019). Serine-rich repeat protein adhesins from *Lactobacillus reuteri* display strain specific glycosylation profiles. *Glycobiology* 29, 45–58. doi: 10.1093/glycob/cwy100

Li, H., Debowski, A. W., Liao, T., Tang, H., Nilsson, H.-O., Marshall, B. J., et al. (2017). Understanding protein glycosylation pathways in bacteria. *Future Microbiol.* 12, 59–72. doi: 10.2217/fmb-2016-0166

Li, Y., Chen, Y., Huang, X., Zhou, M., Wu, R., Dong, S., et al. (2008). A conserved domain of previously unknown function in Gap1 mediates protein–protein interaction and is required for biogenesis of a serine-rich streptococcal adhesin. *Mol. Microbiol.* 70, 1094–1104. doi: 10.1111/j.1365-2958.2008.06456.x

Lizcano, A., Babu, R. A. S., Shenoy, A. T., Saville, A. M., Kumar, N., D'Mello, A., et al. (2017). Transcriptional organization of pneumococcal *psrP-secY2A2* and impact of GtfA and GtfB deletion on PsrP-associated virulence properties. *Microbes Infect.* 19, 323–333. doi: 10.1016/j.micinf.2017.04.001

Lizcano, A., Sanchez, C. J., and Orihuela, C. J. (2012). A role for glycosylated serine-rich repeat proteins in Gram-positive bacterial pathogenesis. *Mol. Oral Microbiol.* 27, 257–269. doi: 10.1111/j.2041-1014.2012.00653.x

McLaughlin, J. M., Peyrani, P., Furmanek, S., Khan, F. L., Quinn, A., Jodar, L., et al. (2020). Burden of adults hospitalized with group B streptococcal infection. *J. Infect. Dis.* doi: 10.1093/infdis/jiaa110 [Epub ahead of print].

Mistou, M.-Y., Dramsi, S., Brega, S., Poyart, C., and Trieu-Cuot, P. (2009). Molecular dissection of the *secA2* locus of group B Streptococcus reveals that glycosylation of the Srr1 LPXTG protein is required for full virulence. *J. Bacteriol.* 191, 4195–4206. doi: 10.1128/jb.01673-08

Mitchell, T. J. (2003). The pathogenesis of streptococcal infections: from tooth decay to meningitis. *Nat. Rev. Microbiol.* 1, 219–230. doi: 10.1038/nrmicro771

Mohamed, Y. F., Scott, N. E., Molinaro, A., Creuzenet, C., Ortega, X., Lertmemongkolchai, G., et al. (2019). A general protein *O*-glycosylation machinery conserved in *Burkholderia* species improves bacterial fitness and elicits glycan immunogenicity in humans. *J. Biol. Chem.* 294, 13248–13268. doi: 10.1074/jbc.RA119.009671

Nobbs, A. H., Zhang, Y., Khammanivong, A., and Herzberg, M. C. (2007). *Streptococcus gordonii* Hsa environmentally constrains competitive binding by *Streptococcus sanguinis* to saliva-coated hydroxyapatite. *J. Bacteriol.* 189, 3106–3114. doi: 10.1128/jb.01535-06

Plummer, C., Wu, H., Kerrigan, S. W., Meade, G., Cox, D., and Douglas, C. W. I. (2005). A serine-rich glycoprotein of *Streptococcus sanguis* mediates adhesion to platelets via GPIb. *Br. J. Haematol.* 129, 101–109. doi: 10.1111/j.1365-2141.2005.05421.x

Pombert, C. B., Lévesque, C., and Frenette, M. (2008). Genetic organization of fimbriae's secretion/glycosylation machinery in *Streptococcus salivarius*. *J. Dent. Res.* 87:3422.

Ramboarina, S., Garnett, J. A., Zhou, M., Li, Y., Peng, Z., Taylor, J. D., et al. (2010). Structural insights into serine-rich fimbriae from Gram-positive bacteria. *J. Biol. Chem.* 285, 32446–32457. doi: 10.1074/jbc.M110.128165

Reily, C., Stewart, T. J., Renfrow, M. B., and Novak, J. (2019). Glycosylation in health and disease. *Nat. Rev. Nephrol.* 15, 346–366.

Ronis, A., Brockman, K., Singh, A. K., Gaytán, M. O., Wong, A., McGrath, S., et al. (2019). *Streptococcus oralis* subsp. *dentisani* produces monolateral serine-rich repeat protein fibrils, one of which contributes to saliva binding via sialic acid. *Infect. Immun.* 87:e00406-19.

Rose, L., Shivshankar, P., Hinojosa, E., Rodriguez, A., Sanchez, C. J., and Orihuela, C. J. (2008). Antibodies against PsrP, a novel *Streptococcus pneumoniae* adhesin, block adhesion and protect mice against pneumococcal challenge. *J. Infect. Dis.* 198, 375–383. doi: 10.1086/589775

Samen, U., Eikmanns, B. J., Reinscheid, D. J., and Borges, F. (2007). The surface protein Srr-1 of *Streptococcus agalactiae* binds human keratin 4 and promotes adherence to epithelial HEp-2 cells. *Infect. Immun.* 75, 5405–5414. doi: 10.1128/iai.00717-07

Samen, U., Heinz, B., Boisvert, H., Eikmanns, B. J., Reinscheid, D. J., and Borges, F. (2011). Rga is a regulator of adherence and pilus formation in *Streptococcus agalactiae*. *Microbiology* 157, 2319–2327. doi: 10.1099/mic.0.044933-0

Sanchez, C. J., Shivshankar, P., Stol, K., Trakhtenbroit, S., Sullam, P. M., Sauer, K., et al. (2010). The pneumococcal serine-rich repeat protein is an intra-species bacterial adhesin that promotes bacterial aggregation *in vivo* and in biofilms. *PLoS Pathog.* 6:e1001044. doi: 10.1371/journal.ppat.1001044

Schäffer, C., and Messner, P. (2017). Emerging facets of prokaryotic glycosylation. *FEMS Microbiol. Rev.* 41, 49–91. doi: 10.1093/femsre/fuw036

Schäffer, C., and Messner, P. (2018). Carb loading takes proteins on a ride. *J. Biol. Chem.* 293, 5374–5375. doi: 10.1074/jbc.H118.002411

Schulte, T., Mikaelsson, C., Beaussart, A., Kikhney, A., Deshmukh, M., Wolniak, S., et al. (2016). The BR domain of PsrP interacts with extracellular DNA to promote bacterial aggregation; structural insights into pneumococcal biofilm formation. *Sci. Rep.* 6:32371. doi: 10.1038/srep32371

Seale, A. C., Bianchi-Jassir, F., Russell, N. J., Kohli-Lynch, M., Tann, C. J., Hall, J., et al. (2017). Estimates of the burden of group B Streptococcal disease worldwide for pregnant women, stillbirths, and children. *Clin. Infect. Dis.* 65, S200–S219. doi: 10.1093/cid/cix664

Seepersaud, R., Bensing, B. A., Yen, Y. T., and Sullam, P. M. (2010). Asp3 mediates multiple protein-protein interactions within the accessory Sec system of *Streptococcus gordonii*. *Mol. Microbiol.* 78, 490–505. doi: 10.1111/j.1365-2958.2010.07346.x

Seepersaud, R., Bensing, B. A., Yen, Y. T., and Sullam, P. M. (2012). The accessory Sec protein Asp2 modulates GlcNAc deposition onto the serine-rich repeat glycoprotein GspB. *J. Bacteriol.* 194, 5564–5575. doi: 10.1128/jb.01000-12

Seepersaud, R., Sychantha, D., Bensing, B. A., Clarke, A. J., and Sullam, M. (2017). *O*-acetylation of the serine-rich repeat glycoprotein GspB is coordinated with accessory Sec transport. *PLoS Pathog.* 13:e1006558. doi: 10.1371/journal.ppat.1006558

Seifert, K. N., Adderson, E. E., Whiting, A. A., Bohnsack, J. F., Crowley, P. J., and Brady, L. J. (2006). A unique serine-rich repeat protein (Srr-2) and novel surface antigen (epsilon) associated with a virulent lineage of serotype III *Streptococcus agalactiae*. *Microbiology* 152, 1029–1040. doi: 10.1099/mic.0.28516-0

Selva, L., Ciruela, P., Blanchette, K., Amo, E., Pallares, R., Orihuela, C. J., et al. (2012). Prevalence and clonal distribution of *pcpA*, *psrP* and Pilus-1 among pediatric isolates of *Streptococcus pneumoniae*. *PLoS One* 7:e41587. doi: 10.1371/journal.pone.0041587

Seo, H. S., Minasov, G., Seepersaud, R., Doran, K. S., Dubrovska, I., Shuvalova, L., et al. (2013a). Characterization of fibrinogen binding by glycoproteins Srr1 and Srr2 of *Streptococcus agalactiae*. *J. Biol. Chem.* 288, 35982–35996. doi: 10.1074/jbc.M113.513358

Seo, H. S., Xiong, Y. Q., and Sullam, P. M. (2013b). Role of the serine-rich surface glycoprotein Srr1 of *Streptococcus agalactiae* in the pathogenesis of infective endocarditis. *PLoS One* 8:e64204. doi: 10.1371/journal.pone.0064204

Seo, H. S., Mu, R., Kim, B. J., Doran, K. S., and Sullam, P. M. (2012). Binding of glycoprotein Srr1 of *Streptococcus agalactiae* to fibrinogen promotes attachment to brain endothelium and the development of meningitis. *PLoS Pathog.* 8:e1002947. doi: 10.1371/journal.ppat.1002947

Sequeira, S., Kavanaugh, D., MacKenzie, D. A., Suligoj, T., Walpole, S., Leclaire, C., et al. (2018). Structural basis for the role of serine-rich repeat proteins from *Lactobacillus reuteri* in gut microbe – host interactions. *Proc. Natl. Acad. Sci. U.S.A.* 115, E2706–E2715. doi: 10.1073/pnas.1715016115

Sheen, T. R., Jimenez, A., Wang, N.-Y., Banerjee, A., van Sorge, N. M., and Doran, K. S. (2011). Serine-rich repeat proteins and pili promote *Streptococcus agalactiae* colonization of the vaginal tract. *J. Bacteriol.* 193, 6834–6842. doi: 10.1128/jb.00094-11

Shenoy, A. T., Brissac, T., Gilley, R. P., Kumar, N., Wang, Y., Gonzalez-Juarbe, N., et al. (2017). *Streptococcus pneumoniae* in the heart subvert the host response through biofilm-mediated resident macrophage killing. *PLoS Pathog.* 13:e1006582. doi: 10.1371/journal.ppat.1006582

Shivshankar, P., Sanchez, C., Rose, L. F., and Orihuela, C. J. (2009). The *Streptococcus pneumoniae* adhesin PsrP binds to Keratin 10 on lung cells. *Mol. Microbiol.* 73, 663–679. doi: 10.1111/j.1365-2958.2009.06796.x

Siboo, I. R., Chambers, H. F., and Sullam, P. M. (2005). Role of SraP, a serine-rich surface protein of *Staphylococcus aureus*, in binding to human platelets. *Infect. Immun.* 73, 2273–2280. doi: 10.1128/IAI.73.4.2273

Singh, A. K., Woodiga, S. A., Grau, M. A., and King, S. J. (2017). *Streptococcus oralis* neuraminidase modulates adherence to multiple carbohydrates on platelets. *Infect. Immun.* 85:e00774-16.

Six, A., Bellais, S., Bouaboud, A., Fouet, A., Gabriel, C., Tazi, A., et al. (2015). Srr2, a multifaceted adhesin expressed by ST-17 hypervirulent group B Streptococcus involved in binding to both fibrinogen and plasminogen. *Mol. Microbiol.* 97, 1209–1222. doi: 10.1111/mmi.13097

Stephenson, A. E., Wu, H., Novak, J., Tomana, M., Mintz, K., and Fives-Taylor, P. (2002). The Fap1 fimbrial adhesin is a glycoprotein: antibodies specific for the glycan moiety block the adhesion of *Streptococcus parasanguis* in an *in vitro* tooth model. *Mol. Microbiol.* 43, 147–157. doi: 10.1046/j.1365-2958.2002.02725.x

Swarthout, T. D., Fronterre, C., Lourenço, J., Obolski, U., Gori, A., Bar-Zeev, N., et al. (2020). High residual carriage of vaccine-serotype of *Streptococcus pneumoniae* after introduction of pneumococcal conjugate vaccine in Malawi. *Nat. Commun.* 11:2222.

Szymanski, C. M., Yao, R., Ewing, C. P., Trust, T. J., and Guerry, P. (1999). Evidence for a system of general protein glycosylation in *Campylobacter jejuni*. *Mol. Microbiol.* 32, 1022–1030. doi: 10.1046/j.1365-2958.1999.01415.x

Takahashi, Y., Konishi, K., Cisar, J. O., and Yoshikawa, M. (2002). Identification and characterization of *hsa*, the gene encoding the sialic acid-binding adhesin of *Streptococcus gordonii* DL1. *Infect. Immun.* 70, 1209–1218. doi: 10.1128/IAI.70.3.1209

Takamatsu, D., Bensing, B. A., and Sullam, P. M. (2004a). Four proteins encoded in the *gspB-secY2A2* operon of *Streptococcus gordonii* mediate the intracellular glycosylation of the platelet-binding protein GspB. *J. Bacteriol.* 186, 7100–7111. doi: 10.1128/JB.186.21.7100

Takamatsu, D., Bensing, B. A., and Sullam, P. M. (2004b). Genes in the accessory *sec* locus of *Streptococcus gordonii* have three functionally distinct effects on the expression of the platelet-binding protein GspB. *Mol. Microbiol.* 52, 189–203. doi: 10.1111/j.1365-2958.2004.03978.x

Takamatsu, D., Bensing, B. A., and Sullam, P. M. (2005). Two additional components of the accessory Sec system mediating export of the *Streptococcus gordonii* platelet-binding protein GspB. *J. Bacteriol.* 187, 3878–3883. doi: 10.1128/JB.187.11.3878

Tan, F. Y. Y., Tang, C. M., and Exley, R. M. (2015). Sugar coating: bacterial protein glycosylation and host–microbe interactions. *Trends Biochem. Sci.* 40, 342–350. doi: 10.1016/j.tibs.2015.03.016

Tettelin, H., Masignani, V., Cieslewicz, M. J., Donati, C., Medini, D., Ward, N. L., et al. (2005). Genome analysis of multiple pathogenic isolates of *Streptococcus agalactiae*: implications for the microbial "pan-genome." *Proc. Natl. Acad. Sci. U.S.A.* 102, 13950–13955.

Turner, L. S., Kanamoto, T., Unoki, T., Munro, C. L., Wu, H., and Kitten, T. (2009). Comprehensive evaluation of *Streptococcus sanguinis* cell wall-anchored proteins in early infective endocarditis. *Infect. Immun.* 77, 4966–4975. doi: 10.1128/iai.00760-09

Tytgat, H. L. P., and Lebeer, S. (2014). The sweet tooth of bacteria: common themes in bacterial glycoconjugates. *Microbiol. Mol. Biol. Rev.* 78, 372–417. doi: 10.1128/mmbr.00007-14

Valguarnera, E., Kinsella, R. L., and Feldman, M. F. (2016). Sugar and spice make bacteria not nice: protein glycosylation and its influence in pathogenesis. *J. Mol. Biol.* 428, 3206–3220. doi: 10.1016/j.jmb.2016.04.013

van Sorge, N. M., Quach, D., Gurney, M. A., Sullam, P. M., Nizet, V., and Doran, K. S. (2009). The group B Streptococcal serine-rich repeat 1 glycoprotein mediates penetration of the blood-brain barrier. *J. Infect. Dis.* 199, 1479–1487. doi: 10.1086/598217

Wahl, B., O' Brien, K. L., Greenbaum, A., Majumder, A., Liu, L., Chu, Y., et al. (2018). Burden of *Streptococcus pneumoniae* and *Haemophilus influenzae* type b disease in children in the era of conjugate vaccines: global, regional, and national estimates for 2000-15. *Lancet Glob. Health* 6, e744–e757. doi: 10.1016/S2214-109X(18)30247-X

Wang, N.-Y., Patras, K. A., Seo, H. S., Cavaco, C. K., Rösler, B., Neely, M. N., et al. (2014). Group B streptococcal serine-rich repeat proteins promote interaction with fibrinogen and vaginal colonization. *J. Infect. Dis.* 210, 892–991. doi: 10.1093/infdis/jiu151

Wu, H., Bu, S., Newell, P., Chen, Q., and Fives-Taylor, P. (2007). Two gene determinants are differentially involved in the biogenesis of Fap1 precursors in *Streptococcus parasanguis*. *J. Bacteriol.* 189, 1390–1398. doi: 10.1128/jb.00836-06

Wu, H., Mintz, K. P., Ladha, M., and Fives-Taylor, P. M. (1998). Isolation and characterization of Fap1, a fimbriae-associated adhesin of *Streptococcus parasanguis*. *Mol. Microbiol.* 28, 487–500. doi: 10.1046/j.1365-2958.1998.00805.x

Yen, Y. T., Seepersaud, R., Bensing, B. A., and Sullam, P. M. (2011). Asp2 and Asp3 interact directly with GspB, the export substrate of the *Streptococcus gordonii* accessory Sec system. *J. Bacteriol.* 193, 3165–3174. doi: 10.1128/jb.00057-11

Yumoto, H., Hirota, K., Hirao, K., Ninomiya, M., Murakami, K., Fujii, H., et al. (2019). The pathogenic factors from oral streptococci for systemic diseases. *Int. J. Mol. Sci.* 20:4571. doi: 10.3390/ijms20184571

Zhang, H., Zhou, M., Yang, T., Haslam, S. M., Dell, A., and Wu, H. (2016). New helical binding domain mediates a glycosyltransferase activity of a bifunctional protein. *J. Biol. Chem.* 291, 22106–22117. doi: 10.1074/jbc.M116.731695

Zhang, H., Zhu, F., Yang, T., Ding, L., Zhou, M., Li, J., et al. (2014). The highly conserved domain of unknown function 1792 has a distinct glycosyltransferase fold. *Nat. Commun.* 5:4339. doi: 10.1038/ncomms5339

Zhou, M., and Wu, H. (2009). Glycosylation and biogenesis of a family of serine-rich bacterial adhesins. *Microbiology* 155, 317–327. doi: 10.1099/mic.0.025221-0

Zhou, T.-T., Yue, Y., Zheng, F., Liang, X.-D., Cao, Q.-X., Wang, Y.-W., et al. (2019). Monoclonal antibody against L-lectin module of SraP blocks adhesion and protects mice against *Staphylococcus aureus* challenge. *J. Microbiol. Immunol. Infect.* doi: 10.1016/j.jmii.2019.08.019 [Epub ahead of print].

Zhu, F., and Wu, H. (2016). Insights into bacterial protein glycosylation in human microbiota. *Sci. China Life Sci.* 59, 11–18. doi: 10.1007/s11427-015-4980-7

Zhu, F., Zhang, H., and Wu, H. (2015). Glycosyltransferase-mediated sweet modification in oral streptococci. *J. Dent. Res.* 94, 659–665. doi: 10.1177/0022034515574865

Zhu, F., Zhang, H., Yang, T., Haslam, S. M., Dell, A., and Wu, H. (2016). Engineering and dissecting the glycosylation pathway of a streptococcal serine-rich repeat adhesin. *J. Biol. Chem.* 291, 27354–27363. doi: 10.1074/jbc.M116.752998 doi: 10.1074/jbc.m116.752998

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Microbiology

REVIEW
published: 25 November 2020
doi: 10.3389/fmicb.2020.602305

The Multifaceted Nature of Streptococcal Antigen I/II Proteins in Colonization and Disease Pathogenesis

*Haider S. Manzer[1], Angela H. Nobbs[2] and Kelly S. Doran[1]**

[1] Department of Immunology and Microbiology, University of Colorado Anschutz Medical Campus, Aurora, CO, United States, [2] Bristol Dental School, University of Bristol, Bristol, United Kingdom

Streptococci are Gram-positive bacteria that belong to the natural microbiota of humans and animals. Certain streptococcal species are known as opportunistic pathogens with the potential to cause severe invasive disease. Antigen I/II (AgI/II) family proteins are sortase anchored cell surface adhesins that are nearly ubiquitous across streptococci and contribute to many streptococcal diseases, including dental caries, respiratory tract infections, and meningitis. They appear to be multifunctional adhesins with affinities to various host substrata, acting to mediate attachment to host surfaces and stimulate immune responses from the colonized host. Here we will review the literature including recent work that has demonstrated the multifaceted nature of AgI/II family proteins, focusing on their overlapping and distinct functions and their important contribution to streptococcal colonization and disease.

Keywords: antigen I/II, *Streptococcus*, adhesin, dental caries, respiratory infection, vaginal colonization, meningitis, biofilms

OPEN ACCESS

Edited by:
Magnus Hook, Texas A&M Health Science Center, United States

Reviewed by:
Mark Joseph Walker, The University of Queensland, Australia
Karina Persson, Umeå University, Sweden
Richard Lamont, University of Louisville, United States

****Correspondence:***
Kelly S. Doran
kelly.doran@cuanschutz.edu

Specialty section:
This article was submitted to Infectious Diseases, a section of the journal Frontiers in Microbiology

Received: *03 September 2020*
Accepted: *29 October 2020*
Published: *25 November 2020*

Citation:
Manzer HS, Nobbs AH and Doran KS (2020) The Multifaceted Nature of Streptococcal Antigen I/II Proteins in Colonization and Disease Pathogenesis. Front. Microbiol. 11:602305. doi: 10.3389/fmicb.2020.602305

INTRODUCTION

The genus *Streptococcus* is a heterogeneous group of Gram-positive bacteria that can be part of the natural microbiota. Various streptococci are commonly isolated from the oral cavity, intestines, or female reproductive tracts in healthy humans. These colonizing streptococci are often commensal organisms that persist without causing disease. In some cases, colonization by commensal streptococci can actually benefit the host through mechanisms such as niche competition or direct inhibition of more pathogenic organisms (Abranches et al., 2018; Nobbs and Kreth, 2019); however, there are also many streptococcal species that are opportunistic pathogens with the potential to cause severe invasive disease. Within the viridans group of streptococci, a heterogeneous group consisting of typically commensal α-hemolytic species, exists *Streptococcus mutans*. *S. mutans* is a major agent of dental caries (tooth decay) (Lehner et al., 1975) and infective endocarditis (Hoen et al., 2002). *Streptococcus pyogenes* (Group A *Streptococcus*, GAS) is a major cause of skin and soft tissue infections, as well as respiratory infections (Steer et al., 2012). Similarly, *Streptococcus agalactiae* (Group B *Streptococcus*, GBS) is known to colonize the female reproductive tract and is a leading cause of neonatal sepsis, pneumonia, and meningitis (Edmond et al., 2012). Understanding the shared mechanisms by which these diverse species colonize and cause disease is the first step toward treatment and prevention. One example of a shared virulence factor among these species is the family of multifunctional streptococcal surface anchored adhesins known as antigen I/II (AgI/II) proteins.

Mike Russell et al. and Roy Russell et al. identified the cell wall antigens I and II in *Streptococcus mutans* nearly simultaneously over 40 years ago (Russell and Lehner, 1978; Russell, 1979; Russell et al., 1980). Soon after, it was shown that antigen II was actually a breakdown product of antigen I (Kelly et al., 1989), giving rise to their new classification as antigen I/II proteins. These AgI/II proteins are widely distributed not only among various serotypes within *S. mutans* (Russell and Lehner, 1978; Russell et al., 1980; Ma et al., 1991), but orthologous proteins sharing similar structure and functions also exist within a number of other streptococcal species (Kelly et al., 1989; Jenkinson and Demuth, 1997; Brady et al., 2010). Due in part to the initial discovery of AgI/II proteins within *S. mutans,* their contribution to dental caries has been a major focus of AgI/II related research; however, more recent studies have investigated AgI/II proteins in other species and biological niches as well. The multifunctional nature of these adhesins in multiple niches may directly contribute to the transition toward pathogenesis as the same protein that promotes colonization of a commensal or beneficial organism in one niche may also allow persistence after that organism gains access to new and potentially more susceptible host tissues. Furthermore, discrete locations of various binding sites on AgI/II proteins for interaction with different microorganisms and host molecules ensures maximum protein functionality, which could be a key mechanism for generating diversity in the development of microbial communities. In addition to the cariogenic properties of AgI/II proteins, these studies have revealed diverse functions ranging from adherence to immune modulation, which in turn contribute to streptococcal colonization and disease.

NOMENCLATURE

Proteins belonging to the AgI/II family that have been identified to date are known by various names, many of which are shown in **Table 1**. The majority of these were independently studied and named by different groups prior to the widespread availability of sequencing technologies that eventually revealed that these were, in fact, the same proteins. For example, the single *S. mutans* AgI/II protein is known by at least seven different names. In some cases, the different names represent different proteins coded by unique genes within the same species; BspA-D in GBS, SspA/B in *S. gordonii*, and PAaA/B in *S. criceti* are examples of this. Standardization of the nomenclature in future publications would significantly clarify which proteins are being discussed. A few groups have used the "AgI/II" designation for antigen I/II proteins regardless of the species to which the protein belongs (Sciotti et al., 1997; Chuzeville et al., 2017; Auger et al., 2019).

TABLE 1 | AgI/II family protein names by species.

Species	Protein name(s)	Reference(s)
S. mutans	SpaP/P1/PAc/Sr/MSL-1/AgB/IF	Russell and Lehner, 1978; Russell et al., 1980; Kelly et al., 1989; Demuth et al., 1990b; Ogier et al., 1990; Bleiweis et al., 1992; Jenkinson and Demuth, 1997
S. gordonii	SspA/SspB	Demuth et al., 1996; Brooks et al., 1997; Jenkinson and Demuth, 1997
S. agalactiae	BspA-D/PAc/SSP-5	Tettelin et al., 2005; Rego et al., 2016
S. pyogenes	Spy1352/AspA	Zhang et al., 2006
S. sobrinus	SpaA/PAg	LaPolla et al., 1991
S. sanguinis	SSP-5/PAc	Demuth et al., 1988, 1990a,b
S. intermedius	Pas	Petersen et al., 2001, 2002; Tamura et al., 2001
S. criceti	PAaA/PAaB	Tamura et al., 2001, 2004
S. downei	PAh	Tamura et al., 2008
S. suis	SSP-5/AgI/II	Chuzeville et al., 2017; Auger et al., 2019

The most common names used to refer to AgI/II proteins from different Streptococcus species are listed.

PREVALENCE AND CONSERVATION

AgI/II proteins are nearly ubiquitous within streptococci. In order to examine the relatedness of AgI/II proteins across species, representative genes from those listed in **Table 1** were aligned, and their phylogenetic relationships are shown in **Figure 1A**. Orthologous genes, many of which have yet to be further characterized and may not be expressed at the protein level, were also identified in *S. rattus, S. oralis, S. mitis, S. anginosus, S. constellatus, S. vestibularis,* and *S. parasanguinis* via amplification from *S. mutans* SpaP based PCR probes (Ma et al., 1991). Another potential AgI/II protein was recently identified in *S. salivarius* as well, though it also remains to be characterized (Chaffanel et al., 2018). As AgI/II genes are so widely distributed among streptococci, it is likely that other species that were not listed here also contain AgI/II orthologs that have yet to be identified. For example, while there was previously no evidence for an AgI/II gene within *S. pneumoniae* (Brady et al., 2010), a BLASTP search for homologs of the GAS AgI/II gene AspA within *S. pneumoniae* strains revealed a homologous protein in a newly sequenced strain (NCBI RefSeq: WP_160544519.1). This protein has a similar primary structure with 69% pairwise identity to AspA, with 83% coverage, which is comparable to the variation that exists between AgI/II genes from different species. While this was the only *S. pneumoniae* strain we found containing an AgI/II homolog, its presence nonetheless highlights the widespread distribution of AgI/II proteins and their potential to provide pathogens with a selective advantage in certain niches.

Previous studies comparing sequences across AgI/II genes from multiple species have revealed domain-based variation in conservation (Ma et al., 1991; Brady et al., 2010; Deng et al., 2019). This is shown to be especially true when protein sequences from a larger number of species are compared (**Figure 1B**). Comparisons of AgI/II genes from multiple strains within one species, however, reveal high levels of sequence conservation (Brady et al., 1991; Ma et al., 1991; Do et al., 2010; Chuzeville et al., 2015; Rego et al., 2016). For example, a study comparing 37 *S. mutans* SpaP genes observed non-synonymous polymorphisms in only 4% of sites (Do et al., 2010). This indicates both diversity within the AgI/II family of proteins, as well as specificity, where the AgI/II proteins within each species may have evolved to potentially better interact with niche specific receptors. The fact that AgI/II genes from more invasive streptococcal species cluster separately

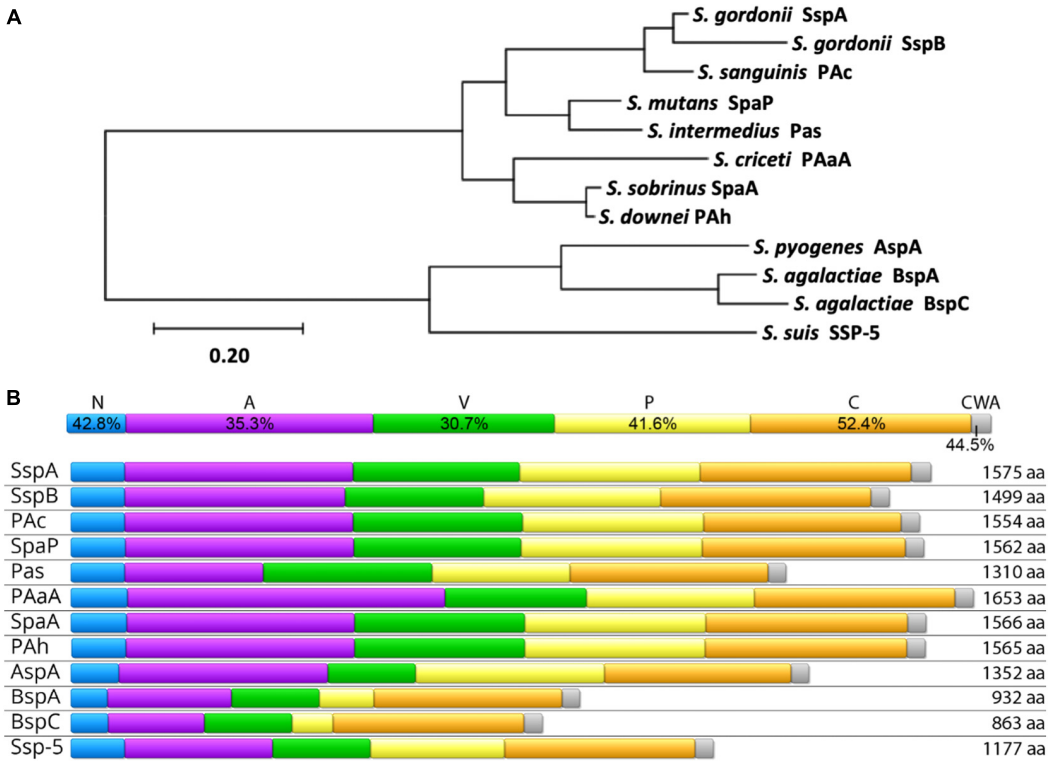


FIGURE 1 | Homology of AgI/II Family Proteins. **(A)** The evolutionary history was inferred by using the Maximum Likelihood method and JTT matrix-based model (Jones et al., 1992) using full-length amino acid sequences of the indicated proteins. The tree with the highest log likelihood (–21701.96) is shown. The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. Evolutionary analyses were conducted in MEGA X (Kumar et al., 2018). **(B)** Conservation of individual domains is shown (top) as the pairwise identity percentage resulting from a MUSCLE protein alignment of each individual domain, as well as the primary structure and amino acid length of AgI/II proteins from select streptococci (bottom). The domains shown are the N-terminal region (N; blue), alanine-rich repeats (A; purple), the variable region (V; green), proline-rich repeats (P; yellow), the C-terminal region (C; orange), and the cell wall anchoring region (CWA; gray). The Uniprot IDs for protein sequences used in both **(A,B)** are as follows: *S. mutans* SpaP (P23504), *S. gordonii* SspA (Q54185), *S. gordonii* SspB (Q54186), *S. agalactiae* BspA (Q8E589), *S. agalactiae* BspC (A0A380IJX7), *S. pyogenes* AspA (Q48S75), *S. intermedius* Pas (Q9KW51), *S. sobrinus* SpaA (Q53414), *S. sanguinis* PAc (F3V086), *S. criceti* PAaA (Q9LBG3), *S. downei* PAh (Q59HN9), and *S. suis* SSP-5 (A0A0Z8CV71).

from the viridans streptococci supports this theory (**Figure 1A**). One factor contributing to the distribution and clustering of AgI/II genes is their ability to be shared via horizontal gene transfer. An example of this is seen with the GAS and GBS AgI/II genes. In GAS, the AgI/II genes are located on the integrative and conjugative element named region of difference 2 (RD2), which is thought to be horizontally shared between GBS, conferring pathogenic traits that support invasive disease in pregnant women and neonates (Zhang et al., 2006; Sitkiewicz et al., 2011; Chuzeville et al., 2015; Deng et al., 2019; Jain et al., 2019). It is especially concerning that this virulence factor is able to spread horizontally since this family of proteins has a wide variety of pathogenic functions, and their presence in streptococcal species occupying unique niches may correlate with the ability to cause disease as well.

STRUCTURE

The primary structure of AgI/II proteins is shown in **Figure 1B**. AgI/II proteins begin with an N-terminal domain containing a signal peptide mediating secretion via the general secretory pathway (Scott and Barnett, 2006). This is adjacent to the A-domain, a region containing multiple alanine-rich repeats (Kelly et al., 1989). The A-domain is connected to a globular domain termed the V-domain based on the variability observed in this region between strains (Brady et al., 1991). The V-domain contains a putative binding pocket, which will be elaborated on in greater detail further on. At the other end of the V-domain is the P-domain, termed for the presence of proline-rich repeats within this region (Kelly et al., 1995). Following the P-domain is the C-terminal region, which consists of two or three globular IgG-like domains. Finally, there is an LPXTG motif for sortase-mediated cell wall anchoring (Kelly et al., 1989). Overall, the AgI/II proteins adopt a cell wall-anchored stalk structure formed by interactions between the intertwined A- and P-domains, which lifts the globular V-domain over 50 nm away from the cell surface in the case of the *S. mutans* SpaP (Larson et al., 2010). The A- and P-domains of SpaP were shown to utilize a unique interlocking between tyrosine residues from heptad repeat motifs in the A-domain with hydrophobic pockets consisting of PxxP motifs within the P-region (Larson et al., 2010). The A- and P-domains

are significantly smaller in some homologs such as the Bsp proteins of GBS (**Figures 1B**, **2**), and it has also been shown that the interaction between A- and P-domains within these homologs is formed in a different manner. The A-domain of BspA contains an asparagine seam whose hydrogens interact with oxygen and nitrogen atoms within the P-domain (Rego et al., 2016). In addition to the A- and P-domains interacting, the tertiary stalk structure of AgI/II proteins is further stabilized by interactions between the N- and C-terminal domains in SpaP (Heim et al., 2014). As the structure of the N-terminal domain in other homologous proteins has not been experimentally determined thus far, the level of conservation of such a locking mechanism remains to be seen.

In addition to projecting the V-domain away from the cell surface to improve interactions with ligands, this stable stalk structure allows for a degree of flexibility that is known to promote bacterial attachment under shear conditions. While this tertiary structure is highly conserved, there is regional variability particular to individual species, especially at the V-domain. Interestingly, the full-length protein sequences used to generate the phylogenetic tree shown in **Figure 1A** cluster into a group of viridans streptococci and a separate group of more traditionally invasive species. Representatives of these two AgI/II protein structures, SpaP from *S. mutans* and BspA from GBS are shown in **Figure 2**. These two species were chosen because each has the largest number of domains with experimentally determined structures published within each cluster. The structural differences between homologs contribute to both overlapping and distinct functions for AgI/II proteins within streptococci. For example, AgI/II proteins from some species are able to bind to sialic acids, while those from other species cannot (Demuth et al., 1990a,b; Soell et al., 1994; Vernier et al., 1996; Chatenay-Rivauday et al., 1998). Structural differences exist even within the phylogenetic clusters, as seen by variations in the published V-domain structure from *S. gordonii* SspB (Forsgren et al., 2009) from that of SpaP. These types of species-dependent variations in AgI/II function will be discussed in greater detail in the following

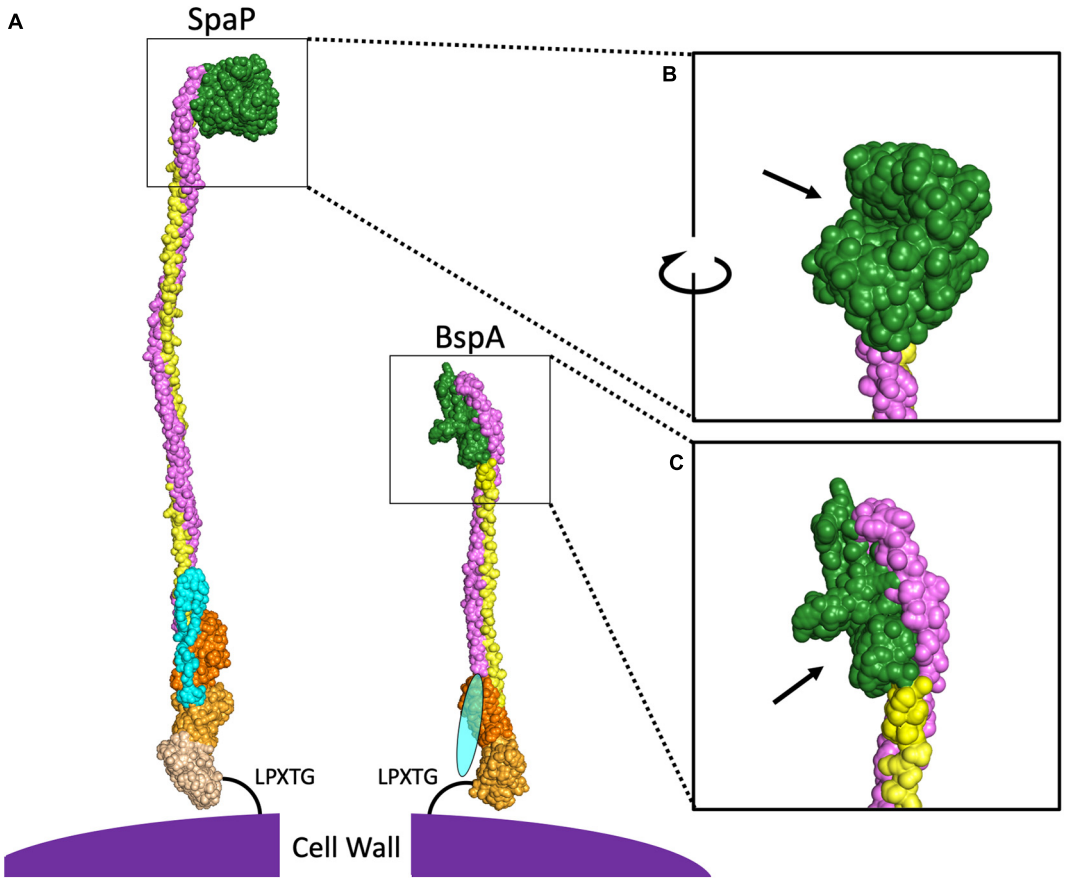


FIGURE 2 | Comparison of SpaP and BspA Protein Structures. **(A)** PyMOL was used to combine the published models for domains of SpaP (N-terminus PDB:4TSH, A3VP1 PDB:3IOX, C-terminus PDB:6E3F) and BspA (V-domain, C-terminus PDB:5DZA). Additional SpaP A and P repeats were modeled from A3VP1, as were the A- and P-domains for BspA. As there is no structure for the N-terminal domain of any AgI/II protein among the cluster of invasive species shown in **Figure 1A**, the BspA N-terminus is depicted by an oval. Light blue indicates the N-terminal domain, purple indicates the A-domain, green indicates the V-domain, yellow indicates the P-domain, and the shades of orange indicate different globular domains at the C-terminus. The V-domain of SpaP **(B)** and BspA **(C)** have been enlarged with arrows pointing to binding pockets.

sections. It is likely that some of these functional variations also depend on structural differences in domains other than the V-domain. While the full structure of SpaP has been mostly solved (Heim et al., 2014, 2015), other than the C-domain of *S. pyogenes* AspA (Hall et al., 2014) and GBS BspA V- and C-domains (Rego et al., 2016), the structures of additional AgI/II domains from the cluster of invasive streptococci have not been experimentally determined. A better understanding of the structural differences between these AgI/II orthologs may reveal key insights regarding their unique mechanisms for colonization or disease progression. Additionally, the structure of *Enterococcus faecalis* Asc10 displays a primary structure akin to that of AgI/II proteins (Waters et al., 2004; Chuang et al., 2009). The region analogous to the AgI/II V-domain of Asc10 was recently shown to have a similar structure to the V-domain of *S. gordonii* SspB (Schmitt et al., 2018). As such, gaining a better understanding of the AgI/II structure may indirectly provide information regarding the mechanisms for pathogenesis utilized by other microorganisms as well.

ADHESIVE FUNCTION

Functional Amyloids and Auto-Aggregation

One well known function of AgI/II proteins is their direct contribution to auto-aggregation of streptococci without intermediate molecules (Demuth et al., 1990a; Jenkinson and Demuth, 1997; Brady et al., 2010; Maddocks et al., 2011; Chuzeville et al., 2017; Deng et al., 2019). In recent years, it has also been shown that fragments of *S. mutans* SpaP interact with each other to form functional amyloid fibers (Oli et al., 2012). This discovery prompted investigations regarding the molecular interactions contributing to this phenomenon. As mentioned earlier, AgI/II proteins are known to fragment, releasing what was previously known as antigen II. These released fragments can interact with surface anchored AgI/II proteins, particularly through the V-domain of the surface anchored protein with either the V- or C-domains of the fragmented protein (Heim et al., 2015; Rego et al., 2016; Tang et al., 2016; Rivière et al., 2020). AgI/II fragments comprised of the C-domain are able to form the functional amyloid fibers *in vitro* (Tang et al., 2016; Barran-Berdon et al., 2020). Furthermore, AgI/II functional amyloids contribute to stabilization of *S. mutans* biofilms (Oli et al., 2012; Tang et al., 2016; Barran-Berdon et al., 2020). Together, this information suggests a critical role of AgI/II protein interactions for the stabilization of streptococcal communities, as well as in the formation of a structural foundation in biofilms that may be taken advantage of by other species. The direct role of AgI/II proteins in polymicrobial biofilms at various tissues will be discussed below.

Extracellular Matrix Interactions

In addition to auto-interactions, AgI/II proteins are able to interact with a large number of host ligands. As a class of multifunction adhesins, AgI/II proteins play a role in streptococcal interactions with extracellular matrix (ECM), or canonical ECM proteins outside of the context of ECM, as seen with some components of the salivary pellicle. These interactions potentially facilitate adherence to both mucosal surfaces such as the lung or vaginal epithelium, as well as non-mucosal surfaces such as endothelial cells or teeth. Thus far, interactions with the fibrous components of the ECM has been a major focus of research, especially collagen, fibrinogen, fibronectin, and laminin. As collagen is a major component of dentin (Linde, 1989), there has been a large interest in collagen-binding proteins and their role in mediating streptococcal access and invasion into tooth surfaces. It has been shown that *S. gordonii* SspA and SspB (Love et al., 1997; Heddle et al., 2003), *S. suis* AgI/II (Chuzeville et al., 2017), and *S. mutans* SpaP (Love et al., 1997; Sciotti et al., 1997; Beg et al., 2002; Petersen et al., 2002; Sullan et al., 2015; Esberg et al., 2017) interact with and promote adherence to surfaces coated with type I collagen, while *S. intermedius* Pas does not (Petersen et al., 2001, 2002). Some of these studies utilized human collagen (Sciotti et al., 1997; Beg et al., 2002; Chuzeville et al., 2017), while others used rat tail collagen (Love et al., 1997; Sullan et al., 2015). This is significant because a study that used type I collagen from both human and rat tail found that *S. mutans* SpaP interacts with rat and not human collagen (Petersen et al., 2002). Differences in experimental set up might explain these contradictory results, but additional investigation is required to definitively determine the role of an AgI/II – collagen interaction within the setting of human disease.

Streptococcal binding of laminin may also play a role in colonization or invasion of various tissues (Singh et al., 2012). Studies have shown that in addition to *S. intermedius* Pas (Petersen et al., 2001), *S. mutans* SpaP (Sciotti et al., 1997; Busscher et al., 2007) also interacts with laminin. Another study saw a minor decrease in the ability of a *S. mutans* SpaP knockout to bind human laminin but as the difference was not significant, the authors concluded that SpaP did not contribute to laminin binding (Beg et al., 2002); thus, additional investigations into the interactions of AgI/II proteins with laminin are still needed. Certain streptococcal species are also major agents of infective endocarditis (IE), and the ability to bind fibronectin and fibrinogen in addition to collagen may contribute to development of this disease (Baddour, 1994). *S. mutans* SpaP (Sciotti et al., 1997; Beg et al., 2002; Petersen et al., 2002; Kelemen et al., 2004; Sullan et al., 2015), *S. intermedius* Pas (Petersen et al., 2001, 2002), and *S. suis* AgI/II (Chuzeville et al., 2017) have all been shown to bind fibronectin. Additionally, *S. mutans* SpaP (Beg et al., 2002; Kelemen et al., 2004), *S. suis* AgI/II (Chuzeville et al., 2017), and GBS BspC (Chuzeville et al., 2015) have been shown to bind fibrinogen. The role of AgI/II proteins in binding of other ECM components such as elastin is still a significant knowledge gap that warrants future investigation.

Additional Host Ligands

Direct AgI/II interactions with host cells are also well documented. These proteins mediate attachment to epithelial (Vernier et al., 1996; Sciotti et al., 1997; Chuzeville et al., 2017; Chaffanel et al., 2018; Pidwill et al., 2018; Jain et al., 2019) and endothelial (Vernier et al., 1996; Al-Okla et al., 1999; Deng et al., 2019; Oho and Nagata, 2019) cells, in addition to monocytes

and fibroblasts (Soell et al., 1994; Chatenay-Rivauday et al., 1998, 2000; Petersen et al., 2001; Hajishengallis et al., 2002; Engels-Deutsch et al., 2003). The diversity of different cell types that streptococci adhere to using AgI/II proteins highlights their dynamic role as adhesins. One ligand for *S. mutans* SpaP that was identified early on is the intermediate filament (IF) protein, keratin (Sciotti et al., 1997). More recently, vimentin, a type III IF protein, was also identified as the GBS BspC receptor mediating adherence to human cerebral microvascular endothelial cells (hCMECs) (Deng et al., 2019). Despite the ubiquitous nature of IF proteins as structural components in all cell types, they remain severely understudied as receptors for AgI/II proteins, and future studies should investigate the role of potential IF-AgI/II interactions in pathogenesis.

Salivary agglutinin (SAG) is perhaps the best documented host receptor for AgI/II proteins and will be discussed in a dedicated section below; however, these investigations have revealed the lectin-like properties of AgI/II proteins. For example, sialic acid (and to a lesser degree, fructose, and mannose) binding by *S. sanguinis* SSP-5 competitively inhibits aggregation and interaction with SAG (Demuth et al., 1990a). Conversely, *S. mutans* SpaP interaction with SAG was inhibited by fucose and lactose, but not by sialic acid (Demuth et al., 1990b). Interestingly, *S. mutans* AgI/II was shown to bind *N*-acetylneuraminic acid (a sialic acid) in addition to fucose on the surface of monocytes and KB cells (Soell et al., 1994; Vernier et al., 1996; Chatenay-Rivauday et al., 1998). Additionally, the V-domain of AgI/II proteins is thought to contain a carbohydrate-binding pocket (Troffer-Charlier et al., 2002; Forsgren et al., 2009; Rego et al., 2016), and another cleft between two of the globular C-terminal domains of SpaP was shown to bind glucose (Larson et al., 2011). Collectively, this information indicates that while members of the AgI/II family of proteins have similar and often overlapping specificity for host factors, many of these interactions may be unique to both the AgI/II protein expressed in certain species, as well as the cell type in which the interaction is taking place.

AgI/II proteins are also able to interact with a number of receptors involved in immune signaling. *S. gordonii* SspA and SspB have been shown to interact with α5β1 integrins on human lung epithelial A549 cells, HEp-2 cells, and coated plates (Nobbs et al., 2007; Andrian et al., 2012). Furthermore, the interaction between SspA/SspB with α5β1 integrins was shown to be mediated via the N-terminal portion of the protein, and this interaction contributed to invasion into HEp-2 cells in addition to adherence (Nobbs et al., 2007). Due to the widespread expression of α5β1 integrins in multiple cell types, it is likely that their interaction with AgI/II proteins contributes to adherence and invasion in other cell types as well, but this has yet to be confirmed.

INTERACTIONS WITH THE IMMUNE SYSTEM

Phagocytosis

AgI/II proteins are able to interact with or evade the host immune system, further impacting disease. One main niche where this occurs is within the bloodstream, as streptococci are often associated with blood-borne systemic diseases. AgI/II proteins have anti-phagocytic functions that may promote survival within the host and development of high levels of bacteremia. An *S. suis* AgI/II deficient mutant was more readily internalized by resident peritoneal macrophages but had no difference from wildtype in intracellular survival (Auger et al., 2019). The knockout strain also displayed increased internalization by dendritic cells, along with decreased intracellular survival within the dendritic cells. Furthermore, the *S. suis* AgI/II knockout was killed more readily when exposed to whole murine blood, collectively resulting in a dramatic difference in both rapid clearance of the knockout strain from murine blood and mouse survival upon intraperitoneal or intravenous infection (Auger et al., 2019). Similarly, deletion of AspA in GAS led to increased killing by the human neutrophil (HL60) and mouse macrophage (J774.2) cell lines, while expressing AspA in *L. lactis* led to decreased killing by the same cell lines (Franklin et al., 2013). Interestingly, heterologous expression of *S. gordonii* SspB within *L. lactis* had no impact on phagocytosis, indicating species specific differences in the sufficiency of AgI/II proteins to impact phagocytosis. This is also supported by the fact that the opposite trend was observed with *S. mutans* PAc, where a PAc knockout and *S. mutans* clinical strains with less AgI/II surface expression exhibited a decreased ability to be phagocytosed by human polymorphonuclear leukocytes (Nakano et al., 2006). Furthermore, the PAc deficient strain was not cleared as efficiently as wildtype *S. mutans* after intravenous injection within a rat model of bacteremia.

Immune Signaling

AgI/II proteins have also been shown to modulate immune responses while in the bloodstream. In the same rat model for bacteremia that was mentioned above, significantly higher concentrations of serum sialic acid [a marker for systemic inflammation (Rajendiran et al., 2014)] were observed after injection with the PAc deficient mutant as compared to the wildtype control (Nakano et al., 2006). In this case, presence of the AgI/II protein appeared to dampen inflammation, although the opposite was observed in other *S. mutans* studies. As mentioned earlier, AgI/II proteins interact with THP-1 monocytes derived from human peripheral blood. *S. mutans* SpaP was able to induce release of the neutrophil chemoattractant IL-8 by THP-1 cells (Engels-Deutsch et al., 2003). SpaP also induced the pro-inflammatory cytokines TNFα and IL-1β from THP-1 cells, which was partially mediated by CD14 and TLR4, but not significantly by TLR2 (Hajishengallis et al., 2002). *S. mutans* Sr was similarly shown to bind human monocytes and induce the release of TNF, IL-1, and IL-6 (Soell et al., 1994). These seemingly contradictory results indicate complex interactions that strike a delicate balance with the host immune system. The mechanistic details regarding the contribution of *S. mutans* AgI/II to immune suppression or activation should be investigated in greater detail.

Despite the controversy involved with SpaP, most other AgI/II proteins appear to induce a pro-inflammatory response by cells found within the blood. *S. intermedius* Pas has been

shown to induce IL-8 from THP-1 cells (Petersen et al., 2001). *S. gordonii* SspA and SspB both also induce TNF, IL-6, IL-10, and IL-12 release by bone marrow derived dendritic cells (Andrian et al., 2012). Intraperitoneal infection of mice with a *S. suis* AgI/II mutant strain led to decreased levels of a large number of pro-inflammatory mediators within the plasma, including IL-6, IL-12p70, IFN-γ, CCL2, CCL3, CCL4, CXCL1, and CXCL2 (Auger et al., 2019). Additionally, TLR2 and TLR4 were shown to contribute to MyD88 dependent induction of TNF, IL-1β, IL-6, and CCL3 in dendritic cells (Auger et al., 2019). TLR2 and TLR4 both interact with CD14 to sense extracellular factors and begin signal transduction (Arroyo-Espliguero et al., 2004). One study showed that THP-1 cells treated with vitamin D_3 (known to increase expression of CD14) were not stimulated to a higher degree by *S. mutans* SpaP and concluded that CD14 was not a receptor for the protein (Chatenay-Rivauday et al., 1998); however, another study soon after showed that monoclonal antibodies specific to CD14 and TLR4 (but not TLR2), both decreased immune activation by SpaP in THP-1 cells (Hajishengallis et al., 2002). This indicates that SpaP does likely interact with CD14, but requires the intracellular signaling domain of TLR4 to induce the expected immune responses. Collectively, this information suggests an overwhelmingly pro-inflammatory response to AgI/II proteins by white blood cells. Still, as there is evidence for immune suppressive or anti-phagocytic characteristics of some of the AgI/II proteins, it remains possible that these AgI/II proteins have similarly complex interactions with the immune system under conditions or niches that have yet to be investigated.

AgI/II proteins also promote inflammatory signaling in other cell types, including epithelial and endothelial cells. As mentioned previously, AgI/II proteins can bind α5β1 integrins, which are known to sense the extracellular environment through adhesive interactions and initiate signaling cascades that alter a variety of host processes (Winograd-Katz et al., 2014). *S. mutans* SpaP is able to bind to α5β1 integrins on human saphenous vein endothelial cells (HSVECs) and osteoblasts, contributing to inflammatory signaling (Al-Okla et al., 1999; Son et al., 2012). AgI/II binding of α5β1 on the HSVEC surface was shown to induce the MAPK signaling pathway and tyrosine phosphorylation of phospholipase Cγ, focal adhesion kinase, and paxillin (Al-Okla et al., 1999). This eventually leads to IL-8 release in a PI3K independent manner. SpaP induction of HSVEC IL-8 release was confirmed by another study as well, along with release of IL-6 (Vernier et al., 1996). *S. gordonii* SspA and SspB also increased IL-6, MCP-1 and IL-8 production in A549 cells, and induction of both IL-6 and MCP-1 could be inhibited with antibodies that blocked β1 integrin. Furthermore, SspA and SspB induced IL-6 and IL-8 within HEp-2 cells, as well as CXCL1 and CXCL2 in lung homogenates (Andrian et al., 2012). Similarly, the GBS AgI/II protein known as BspC was shown to induce inflammatory signaling in hCMECs leading to meningitis-associated inflammation. BspC induced IL-8, IL-1β, and CXCL1 release by hCMECs, and also KC protein and IL-1β within a murine model of hematogenous meningitis (Deng et al., 2019). This is a clear case where the AgI/II mediated inflammatory signaling contributes to disease. Given the large body of data showing that similar immune induction occurs at other body sites, the direct role of AgI/II proteins in inflammatory disease should be investigated in greater detail.

NICHE SPECIFIC INTERACTIONS

Oral Cavity

As mentioned previously, due to the initial discovery of AgI/II proteins in *S. mutans,* the AgI/II proteins of streptococci that colonize the oral cavity are the most well studied. Within the oral cavity, AgI/II proteins are known to interact with one of the host innate defense proteins known as salivary agglutinin (SAG). Initial studies identified that *S. mutans* SpaP (Brady et al., 1992; Hajishengallis et al., 1994; Sciotti et al., 1997; Ahn et al., 2008), *S. sanguinis* PAc (Demuth et al., 1988, 1990a; Lamont et al., 1991), *S. sobrinus* SpaA (Lamont et al., 1991), and *S. gordonii* SspA (Jenkinson et al., 1993) all interacted with fluid-phase SAG, resulting in aggregation. The aggregation acts as a defense mechanism in this case as it promotes clearance of bacteria through actions such as swallowing. It has since been discovered that SAG and the lung scavenger receptor protein known as glycoprotein-340 (gp340) [also known as Deleted in Malignant Brain Tumors 1 or DMBT1 (Mollenhauer et al., 1997)] are the same protein (Prakobphol et al., 2000). In the context of mucosal colonization, the protein has historically been referred to as either SAG or more recently gp340, so it will be simply referred to as SAG/gp340 in this review. Within the oral cavity, SAG/gp340 can exist in fluid phase as a salivary component, or in an adsorbed phase that can mediate bacterial attachment to the tooth surface, thereby contributing to dental plaque biofilm formation (Madsen et al., 2010). Interestingly, it seems that AgI/II proteins have varying interactions with these two phases of SAG/gp340. *S. gordonii* SspA and SspB, *S. intermedius* Pas and *S. mutans* SpaP all contributed to aggregation by fluid-phase SAG/gp340 at comparable levels, but only the AgI/II proteins of *S. gordonii* mediated higher affinity binding to surface-bound SAG/gp340 (Loimaranta et al., 2005; Maddocks et al., 2011). Likewise, while GAS AspA contributes to attachment to immobilized SAG/gp340, it did not cause aggregation by fluid-phase SAG/gp340 (Maddocks et al., 2011). The AgI/II protein of *S. suis* has also been shown to interact with both fluid-phase and immobilized SAG/gp340, contributing to both aggregation and surface adherence, respectively (Chuzeville et al., 2017). While GAS and *S. suis* are not common colonizers of the oral cavity, the ability for their AgI/II proteins to interact with SAG/gp340 highlights the conservation of this particular function.

An early study using monoclonal blocking antibodies revealed that binding to fluid-phase and surface-immobilized SAG/gp340 is mediated by different regions of the AgI/II protein (Brady et al., 1992). This was the first indication that different portions of SAG/gp340 are accessible to AgI/II binding in its two physical states. Multiple studies have investigated this phenomenon in more detail, specifically examining the scavenger-rich cysteine repeat (SRCR) domain of SAG/gp340 that is targeted by

AgI/II proteins (Bikker et al., 2002, 2004). It was shown that immobilized SRCR underwent a conformational change in the presence of calcium that allowed for AgI/II protein interaction with nanomolar affinity, whereas AgI/II protein interacted with fluid-phase SRCR with micromolar affinity (Purushotham and Deivanayagam, 2014). Additionally, SAG/gp340 size variants exist that have altered interactions with AgI/II proteins from different species (Jonasson et al., 2007). Furthermore, alternate AgI/II V-domain segments within a group of *S. mutans* human isolates were shown to confer differential saliva- and SAG/gp340-mediated adherence (Esberg et al., 2012, 2017). Taken together, this information suggests a complex co-evolution of AgI/II proteins and SAG/gp340. While fluid-phase SAG/gp340 normally functions as part of host innate immunity to cause aggregation and clearance of oral streptococci, certain species have evolved to interact with the alternative conformation of immobilized SAG/gp340 on tooth surfaces to promote adherence instead. This provides a mechanism by which both commensal and cariogenic streptococci maintain colonization of the oral cavity. The role of interactions between AgI/II proteins with SAG/gp340 in other diseases will be discussed further on.

Streptococcal AgI/II proteins are able to interact with other microbes known to colonize the oral cavity, contributing to polymicrobial biofilm formation. Often, these interactions stabilize the colonization of pathogens associated with the development of periodontitis or dental caries. While SAG/gp340 has been shown to increase adherence of both *S. mutans* and *S. sobrinus* to immobilized early plaque formers such as *S. sanguinis* and *Actinomyces viscosus*, both species were also shown to adhere to the plaque formers in an AgI/II dependent manner, even without the addition of saliva or purified SAG/gp340 (Lamont et al., 1991). *S. gordonii* SspA and SspB have similarly been shown to interact with *Actinomyces oris* (Back et al., 2015) and the periodontitis-associated pathogen *Porphyromonas gingivalis* (Brooks et al., 1997; Daep et al., 2006). In the case of *P. gingivalis*, it was determined that the minor fimbrial antigen (Mfa1) of *P. gingivalis* interacts with a specific sequence motif found in the C-terminal domain of the SspA/B proteins which is not highly conserved in SpaP (Demuth et al., 2001; Daep et al., 2006; Forsgren et al., 2010). The crystal structure of SspB revealed that this region protrudes from the protein, creating a handle for initial attachment of *P. gingivalis* and inclusion into a biofilm (Forsgren et al., 2010). Peptides mimicking this region have been shown to block the SspB-Mfa1 interaction and inhibit *P. gingivalis* virulence within a murine model for periodontitis (Daep et al., 2011). Furthermore, small molecule inhibitors were recently identified that similarly block the SspB-Mfa1 interaction and attenuate the ability of *P. gingivalis* to cause periodontitis *in vivo* (Roky et al., 2020). *S. gordonii* SspA and SspB, *S. intermedius* Pas, and *S. mutans* SpaP have also all been shown to contribute to aggregation with *Actinomyces naeslundii,* another organism associated with periodontitis (Jakubovics et al., 2005b). Additionally, *S. mutans* SpaP mediates interactions with *Fusobacterium nucleatum* ssp. *polymorphum,* which in turn is known to initiate adherence of other periodontal pathogens (Guo et al., 2017).

AgI/II proteins also contribute to interactions with pathogens associated with dental caries. For example, *Lactobacillus casei* is frequently isolated from dental caries. Co-culturing *L. casei* with *S. mutans* increases the ability of *L. casei* to form biofilms, and this increase was shown to be SpaP dependent (Wen et al., 2010, 2017). Additionally, the fungal pathogen *Candida albicans* is commonly associated with oral streptococcal biofilms, and its presence has recently been shown to be positively correlated with cariogenic traits of *S. mutans* in caries-related biofilms (Falsetta et al., 2014; Bachtiar and Bachtiar, 2018). AgI/II proteins are known to interact with the Agglutinin-like sequence family proteins Als1 and Als3 of *C. albicans*, contributing to the formation of polymicrobial communities. For example, *S. gordonii* SspB interactions with *C. albicans* Als3 have been shown to contribute to biofilm formation (Bamford et al., 2009; Silverman et al., 2010). *S. mutans* SpaP was also shown to be important for incorporation of *C. albicans* within a two-species biofilm, as well as for acid production within the biofilm, which may contribute to the progression of dental caries (Yang et al., 2018).

Despite the body of work indicating the role of AgI/II proteins in polymicrobial biofilm formation with oral pathogens, it has been shown that *S. mutans spaP* is significantly downregulated within two-species biofilms with *S. sanguinis, S. oralis* or *L. casei* (Wen et al., 2010). Very little is known regarding the regulation of AgI/II family proteins, but this indicates that at least within the context of biofilm formation, the regulation of AgI/II proteins is complex and warrants further investigation. Additionally, while it is known that AgI/II proteins contribute to oral disease through adherence or even invasion of dentinal tubules (Love et al., 1997), the role of AgI/II proteins in the immune response within the oral cavity remains largely unexplored.

Respiratory Tract

Streptococci are often colonizers of the nasopharyngeal tract and are a common cause of lung infections. The role of AgI/II proteins in streptococcal persistence within these two niches has been characterized mainly with *S. suis*, *S. gordonii,* and GAS. There is serotype-specific variation in the *S. suis* AgI/II contribution to respiratory infection. The *S. suis* AgI/II protein contributes to adherence to NPTr porcine tracheal cells in a serotype 9 strain, but not in a serotype 2 strain (Chuzeville et al., 2017). A mutant lacking the AgI/II protein in the serotype 9 strain also displayed decreased nasal cavity and tonsil bacterial burden within a porcine model for respiratory infection (Chuzeville et al., 2017). As *S. suis* is also able to cause human infections, it would be of interest to examine this in models for human disease. Studies with *S. gordonii* have used human cell-based models for disease. Some of these were done using HEp-2 cells which were initially thought to be derived from a human laryngeal epithelial carcinoma, however, it has been shown that the cell line was actually the result of HeLa cell contamination (Gartler, 1968). In order to provide historical context, these results will still be mentioned. *S. gordonii* SspA and SspB contribute to adherence and invasion of the HEp-2 cell line (Jakubovics et al., 2005a; Nobbs et al., 2007; Andrian et al., 2012). As mentioned earlier, this was shown to be mediated by recognition of β1

integrin by SspA and SspB (Nobbs et al., 2007). SspA and SspB contribute to *S. gordonii* adherence to a lung epithelial cell line (A549) through β1 integrin as well (Andrian et al., 2012). Additionally, AspA contributes to GAS colonization of the nasopharynx, as seen by its role in adherence to Detroit 562 laryngeal epithelial cells (Franklin et al., 2013). An AspA deficient mutant displayed decreased nasopharyngeal and lung bacterial burden within a murine model of respiratory infection (Franklin et al., 2013). While AspA has also been linked to innate immunity through repression of phagocytic killing (as discussed previously), continued investigation is required to determine its role in immune modulation specifically within the respiratory tract. Additionally, while β1 integrin was shown to be a receptor for SspA and SspB in HEp-2 cells, host respiratory epithelial cell receptors for AgI/II proteins have not been thoroughly characterized; however, since SAG/gp340 is known to be expressed in the lungs, this would be one logical interaction to examine.

Female Reproductive Tract Interactions

Numerous streptococcal species are known to colonize the female reproductive tract. As seen with GBS, this can result in pregnancy-associated complications such as pre-term births. AgI/II proteins are beginning to be implicated in streptococcal colonization of these tissues as well. Recently, the role of the GBS Bsp proteins in colonization of the female reproductive tract has been described. Deletion of BspC from GBS strain 515 resulted in a modest decrease in adherence to vaginal epithelial cell (VECs); however, pretreatment with Bsp antisera resulted in a large decrease in adherence as compared to the preimmune sera control treatment (Pidwill et al., 2018). Additionally, heterologous expression of both BspA and BspC within *Lactococcus lactis* was shown to confer increased adherence to a VEC monolayer (Rego et al., 2016; Pidwill et al., 2018).

The GBS AgI/II proteins are also known to contribute to single species biofilm formation (Chuzeville et al., 2015), but there is evidence that these proteins mediate polymicrobial interactions as well. As mentioned earlier, AgI/II proteins are able to interact with the fungal pathogen *C. albicans,* which is also often co-isolated with GBS from the vaginal tract (Monif and Carson, 1998; Bayó et al., 2002; Cools et al., 2016). BspA and BspC were shown to mediate direct interaction between GBS and *C. albicans* (Rego et al., 2016; Pidwill et al., 2018). Furthermore, adherence of both GBS and *L. lactis* surrogate expression strains to VECs in the presence of *C. albicans* was largely mediated by expression of BspA and BspC. The association between GBS and *C. albicans* is dependent on the Als3 protein of *C. albicans* (Pidwill et al., 2018). Interestingly, Als3 is the same protein that was discussed earlier as being the receptor for AgI/II proteins of oral streptococci, indicating that these interactions are not niche specific. These experiments were all conducted *in vitro*. As such, the role of AgI/II interactions with *C. albicans* within the female reproductive tract *in vivo* remains unexplored.

Other reports of AgI/II interactions within the female reproductive tract are not as direct or as thoroughly studied as the Bsp interaction with vaginal epithelium, as is seen with GAS. Colonization of the genitourinary tract by GAS is a risk factor for puerperal sepsis, and it is possible that the GAS AgI/II protein contributes to colonization of this niche. GAS adherence to vaginal epithelium and colonization of murine vaginal tract was shown to be mediated by genes located on RD2 (Jain et al., 2019). As mentioned earlier, GAS RD2 contains the AgI/II protein AspA. The individual contribution of AspA to vaginal colonization, however, has not yet been shown. Future studies should additionally aim to determine what, if any, impact AgI/II proteins have on streptococcal ascending infection or immune modulation resulting in adverse pregnancy outcomes.

Meningitis

Several streptococcal species are able to penetrate the blood-brain barrier (BBB) in order to cause meningitis, including *S. pneumoniae*, *S. suis* and GBS (Doran et al., 2016). Only in GBS has the role of AgI/II been investigated. A recent study revealed that BspC promoted GBS adherence to hCMECs, a model of the BBB, by interacting with the C-terminus of the intermediate filament protein known as vimentin (Deng et al., 2019). BspC itself induced meningitis-associated inflammation, as discussed in the immune signaling section, and the BspC-vimentin interaction was also necessary for disease progression. A BspC-deficient mutant was unable to cause meningitis within a murine model, while vimentin-deficient mice were protected from infection by wildtype GBS. Furthermore, vimentin was shown to localize to the cell surface in response to interaction with BspC, although the mechanisms governing this reorganization are not yet understood. While *S. suis* AgI/II contributions to systemic infections (discussed within the "Interactions With the Immune System" section) have been previously investigated, their role specifically within central nervous system infections should be examined in future studies.

Infective Endocarditis

Another major example of AgI/II proteins interacting with host endothelial cells is during infective endocarditis (IE), which is characterized by the formation of a bacteria-containing vegetation on a heart valve (Ford and Douglas, 1997). Viridans group streptococci are the second largest cause of IE (Rajani and Klein, 2020) and as such, the role of AgI/II proteins in promoting IE has been widely investigated. The characteristic vegetation in IE is typically composed of fibronectin and fibrinogen, bacteria, and aggregated platelets (Baddour, 1994; Ford and Douglas, 1997). As mentioned previously, many AgI/II proteins are known to interact with both fibrinogen and fibronectin. In addition, *S. mutans* PAc binds platelets and causes their aggregation (Matsumoto-Nakano et al., 2009). *S. gordonii* SspA and SspB have also been shown to bind human erythrocytes (Demuth et al., 1996) and contribute to platelet aggregation, although they do not directly mediate platelet adhesion (Kerrigan et al., 2007).

The role of AgI/II proteins in interacting with endothelial cells is also relevant here, although many of these studies have used cell lines originating from regions distal to the heart. Interactions with brain endothelium were described in the previous section, but SpaP has also been shown to contribute to *S. mutans* adherence to the human saphenous vein endothelial

cell (HSVEC) line (Vernier et al., 1996; Al-Okla et al., 1999). One study used human aortic endothelial cells (HAECs) to model *S. mutans* interactions during IE, and while they did not directly investigate the role of AgI/II proteins, they instead looked at expression of known AgI/II protein receptor DMBT1 (SAG/gp340) in response to infection (Oho and Nagata, 2019). The study found that SAG/gp340 release was upregulated in response to *S. mutans* infection, but this actually decreased adherence and invasion of the HAECs as compared to a SAG/gp340 knockdown cell line. In this case, it seems that the host utilizes the protein in a similar manner to fluid-phase SAG/gp340 to aggregate and promote clearance of *S. mutans*.

Interestingly, despite the role of AgI/II proteins in interacting with platelets, erythrocytes, and endothelial cells, as well as elevated levels of anti-PAc antibody titers in IE patients (Russell et al., 1992), the only study examining AgI/II proteins in an *in vivo* model of IE found no phenotype for an AgI/II-deficient mutant (Ryd et al., 1996). The study found no statistically significant difference in CFUs present in heart valve vegetations between the mutant and wildtype *S. mutans* strains 1-hour post inoculation, although there was a trend toward fewer CFU of the mutant strain. While they observed no difference in bacterial clearance from the blood, CFU counts were only examined for 1 hour post infection, which may have been too early in disease progression. Nonetheless, at 48 hours post infection, they still observed no difference between the wildtype and AgI/II-deficient mutant in CFU counts recovered from heart valve vegetations. Despite these results, due to the numerous differences between AgI/II protein interactions from different species that have been described thus far, revisiting an *in vivo* model for IE to investigate the contribution of AgI/II proteins from other species may provide novel insights into the development of the disease. Additionally, it is also possible that AgI/II proteins mediate IE disease through an altered immune response instead of through direct adherence. A recent study showed that there is a high level of peptide similarity between the *S. mutans* AgI/II protein and cardiovascular autoantigens (Lucchese, 2017). The role of the elevated anti-AgI/II antibody titers (Russell et al., 1992) in recognizing these autoantigens to promote IE should be investigated.

AGI/II PROTEINS AS THERAPEUTIC TARGETS

Given the growing recognition of their role in promoting streptococcal colonization and pathogenesis, the potential for AgI/II proteins to serve as therapeutic targets warrants consideration. Reflecting the timeline for discovery of this protein family, such investigations to date have primarily focused on exploiting the AgI/II protein of *S. mutans* to develop an anti-caries vaccine. Since it was first shown that antibodies against SpaP can impair adherence of *S. mutans* to SAG immobilized on the tooth surface (Hajishengallis et al., 1992), numerous vaccine strategies have been explored, with the ultimate goal of inducing a mucosal immune response that inhibits *S. mutans* colonization of the oral cavity. Approaches include immunization with recombinant SpaP peptides or chimeric proteins incorporating additional *S. mutans* antigens (Lehner et al., 1981; Katz et al., 1993; Yu et al., 1997; Hajishengallis et al., 1998; Zhang et al., 2002; Batista et al., 2017), attenuated *Salmonella enterica* serovar Typhimurium expressing a SpaP epitope (Huang et al., 2001; Jiang et al., 2017), or DNA vaccines (Guo et al., 2004; Jia et al., 2004, 2020; Shi et al., 2012), alongside passive immunotherapy studies using monoclonal antibodies (Lehner et al., 1985, 1992). While most of these studies have been at the preclinical phase using animal models, human studies involving topical application to the tooth surface of anti-SpaP monoclonal antibodies were shown to be effective in preventing colonization by exogenously administered or indigenous *S. mutans* (Ma et al., 1989, 1998). Protection was reported to last for up to 2 years (Ma and Lehner, 1990; Ma et al., 1990), although others failed to replicate such long-term effects (Weintraub et al., 2005), and no follow-up trials have been performed in recent years. Alongside the technical difficulties of vaccine development, another challenge with an anti-caries vaccine is that *S. mutans* is not the only cariogenic bacterium within the oral microbiota, which raises the question of the efficacy of targeting a single bacterium. However, this issue is less of a concern for the development of a vaccine to combat, for example, GBS disease. Sialylated capsular polysaccharide (CPS) has been a major focus of GBS vaccine strategies to date (Kobayashi et al., 2016). However, there are 10 distinct CPS serotypes, as well as non-typeable strains, meaning that any such vaccine design could face the challenge of potential serotype replacement/switching across geographical sites and over time. Vaccines comprising a combination of protein-based antigens have potential to confer broad protection that is serotype-independent, and these are currently under investigation. Based on SpaP, the Bsp proteins or domains thereof might well be anticipated to induce function-blocking antibodies, while their unique structural features offer the potential for the identification of epitopes that allow the selective targeting of GBS while not impacting other streptococci within the resident microbiota at a given ecological niche.

Alongside vaccination strategies, improved knowledge of structure-function relationships across the AgI/II proteins offers the potential for mimetics or small molecule inhibitors to be developed as a therapeutic route. As mentioned, proof-of-concept has been established with use of a peptide to block the interaction of *S. gordonii* AgI/II with *P. gingivalis* Mfa1, resulting in reduced alveolar bone resorption in a murine model of periodontitis (Daep et al., 2011). More recently, this peptide has been incorporated into polymeric nanoparticles to improve peptide dosage and persistence within the oral cavity (Mahmoud et al., 2018), while small molecule inhibitors of the AgI/II-MfaI engagement have also been found to effectively reduce *P. gingivalis* virulence (Roky et al., 2020). Protective effects have also been shown for a peptide (p1025) that blocks *S. mutans* AgI/II adhesion to SAG. In a human trial, topical application of p1025 to the teeth over a 3-week period prevented recolonization by *S. mutans* for up to 4 months but not by another member of the resident oral microbiota, *Actinomyces* (Kelly et al., 1999). With our growing understanding of the importance of the AgI/II protein family amongst pathogenic streptococci, opportunities

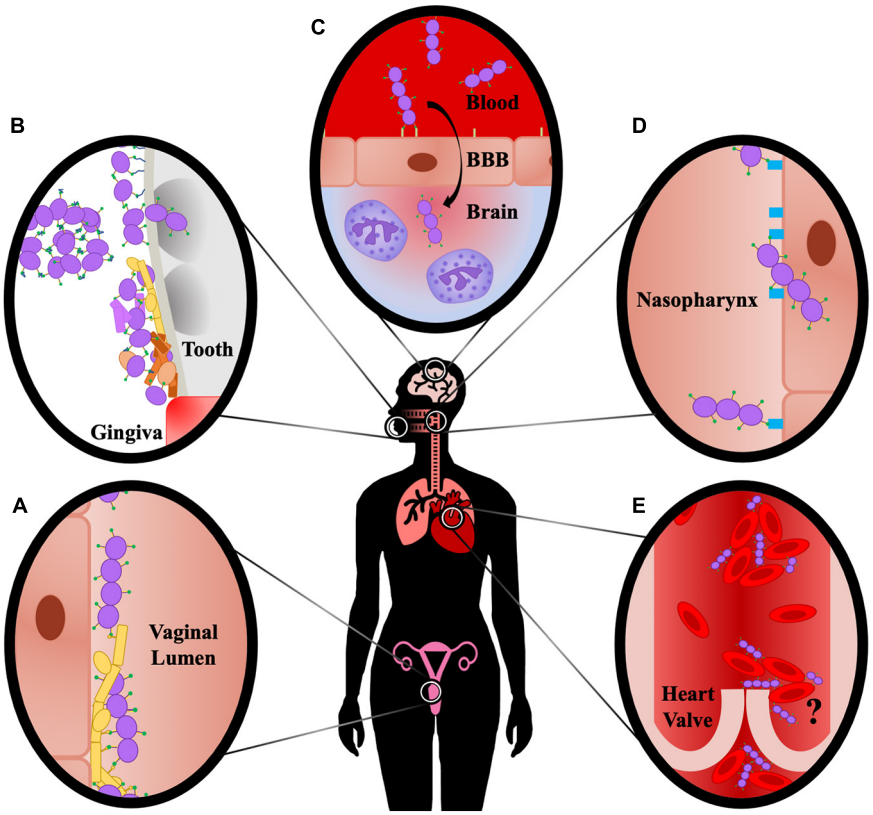


FIGURE 3 | Summary of Site-Specific AgI/II Interactions. **(A)** AgI/II proteins promote direct interactions with vaginal epithelial cells, and interact with Als3 protein of *C. albicans*, promoting fungal and streptococcal adherence to the vaginal epithelium. **(B)** AgI/II proteins bind fluid-phase SAG/gp340, resulting in aggregation and clearance through mechanisms such as swallowing, as well as to immobilized SAG/gp340 on the tooth pellicle, facilitating adherence and colonization. AgI/II proteins mediate direct adherence to tooth components and invasion into dentinal tubules, resulting in caries formation, and also facilitate colonization and polymicrobial biofilm formation with cariogenic pathogens, as well as with pathogens associated with periodontitis. **(C)** AgI/II proteins interact with vimentin on the surface of brain endothelium, and also induce chemokine signaling and neutrophil chemotaxis characteristic of meningitis. **(D)** AgI/II proteins interact with β1 integrin on the nasopharynx epithelium, resulting in adherence and invasion of the cells and consequently increased bacterial burden in the nasopharynx and lungs. **(E)** AgI/II proteins promote aggregation of blood-borne bacteria, as well as attachment to and aggregation of platelets; however, there is currently insufficient evidence to show this promotes increased vegetation formation on heart valves that is characteristic of infective endocarditis, as indicated by the question mark.

for development of therapeutic agents that target them should remain an important area of investigation.

CONCLUSION AND FUTURE DIRECTIONS

The AgI/II family of proteins display both overlapping and diverse functions that vary between streptococcal species as well as the niche within the host (**Figure 3**). The protein structure itself has many unique features, with multiple intramolecular interactions forming an elongated fibrillar stalk that projects the V-domain containing a binding cleft away from the cell surface. This cleft may be used to promote streptococcal colonization of various tissues via both direct cellular adherence or interaction with ECM components. Here we provide a historical perspective of the existing literature and as such, the data presented represent over 40 years of work. During this time, the field has progressed significantly. Some of the pitfalls in the earlier studies that have become clear in hindsight were discussed. In general, validated cell lines, antibodies, recombinant proteins made with the benefit of current structural knowledge, and other modern tools and techniques could provide clarification or new insights into some of the previously identified functions of AgI/II proteins. In addition to the roles for AgI/II proteins in mucosal colonization and blood-borne diseases that were discussed, there are also functions that have yet to be investigated in detail. For example, SspA and SspB protect *S. gordonii* from the antibiotic polymyxin B, as well as the antibacterial peptides nisin and histatin-5 (Andrian et al., 2012). AgI/II proteins may also play a role in streptococcal gut colonization, as the *S. suis* AgI/II was shown to protect against acid stress (Chuzeville et al., 2017), and a potential AgI/II protein identified in *S. salivarius* contributed to adherence to intestinal epithelium (Chaffanel et al., 2018). Each of these potential functions should be further investigated. Additionally, future studies should aim to elucidate the potential for post-translational modifications of AgI/II proteins such as glycosylation to contribute to their functions. Understanding the

diverse nature of these multifunctional adhesins will provide novel insights into the mechanisms by which streptococci and other organisms with structurally similar adhesins cause disease, as well as provide targets for therapeutic intervention.

AUTHOR CONTRIBUTIONS

HM, AN, and KD contributed to writing, reviewing, and editing the manuscript. HM generated the table and figures. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by KD's laboratory NIH/NINDS R01 NS116716 and HM's T32 5T32AI052066-18.

ACKNOWLEDGMENTS

The authors appreciate the contribution of all of the researchers whose work has been reviewed here as well as those whose work may not have been discussed in detail due to space constraints.

REFERENCES

Abranches, J., Zeng, L., Kajfasz, J. K., Palmer, S. R., Chakraborty, B., Wen, Z. T., et al. (2018). Biology of Oral Streptococci. *Microbiol. Spectr.* 6:0042. doi: 10.1128/microbiolspec

Ahn, S. J., Ahn, S. J., Wen, Z. T., Brady, L. J., and Burne, R. A. (2008). Characteristics of biofilm formation by *Streptococcus mutans* in the presence of saliva. *Infect. Immun.* 76, 4259–4268. doi: 10.1128/iai.00422-08

Al-Okla, S., Chatenay-Rivauday, C., Klein, J. P., and Wachsmann, D. (1999). Involvement of alpha5beta1 integrins in interleukin 8 production induced by oral viridans streptococcal protein I/IIf in cultured endothelial cells. *Cell Microbiol.* 1, 157–168. doi: 10.1046/j.1462-5822.1999.00016.x

Andrian, E., Qi, G., Wang, J., Halperin, S. A., and Lee, S. F. (2012). Role of surface proteins SspA and SspB of *Streptococcus gordonii* in innate immunity. *Microbiology* 158, 2099–2106. doi: 10.1099/mic.0.058073-0

Arroyo-Espliguero, R., Avanzas, P., Jeffery, S., and Kaski, J. C. (2004). CD14 and toll-like receptor 4: a link between infection and acute coronary events? *Heart* 90, 983–988. doi: 10.1136/hrt.2002.001297

Auger, J. P., Boa, A. C., Segura, M., and Gottschalk, M. (2019). Antigen I/II Participates in the Interactions of *Streptococcus suis* Serotype 9 With Phagocytes and the Development of Systemic Disease. *Front. Cell Infect. Microbiol.* 9:124. doi: 10.3389/fcimb.2019.00124

Bachtiar, E. W., and Bachtiar, B. M. (2018). Relationship between Candida albicans and *Streptococcus mutans* in early childhood caries, evaluated by quantitative PCR. *F1000Res.* 7:1645. doi: 10.12688/f1000research.16275.2

Back, C. R., Douglas, S. K., Emerson, J. E., Nobbs, A. H., and Jenkinson, H. F. (2015). *Streptococcus gordonii* DL1 adhesin SspB V-region mediates coaggregation via receptor polysaccharide of Actinomyces oris T14V. *Mole. Microbiol.* 30, 411–424. doi: 10.1111/omi.12106

Baddour, L. M. (1994). Virulence factors among gram-positive bacteria in experimental endocarditis. *Infect. Immun.* 62, 2143–2148.

Bamford, C. V., d'Mello, A., Nobbs, A. H., Dutton, L. C., Vickerman, M. M., and Jenkinson, H. F. (2009). *Streptococcus gordonii* modulates Candida albicans biofilm formation through intergeneric communication. *Infect. Immun.* 77, 3696–3704. doi: 10.1128/iai.00438-09

Barran-Berdon, A. L., Ocampo, S., Haider, M., Morales-Aparicio, J., Ottenberg, G., Kendall, A., et al. (2020). Enhanced purification coupled with biophysical analyses shows cross-β structure as a core building block for *Streptococcus mutans* functional amyloids. *Sci. Rep.* 10:5138. doi: 10.1038/s41598-020-62115-7

Batista, M. T., Ferreira, E. L., Pereira, G. S., Stafford, P., Maeda, D., Rodrigues, J. F., et al. (2017). LT adjuvant modulates epitope specificity and improves the efficacy of murine antibodies elicited by sublingual vaccination with the N-terminal domain of *Streptococcus mutans* P1. *Vaccine* 35, 7273–7282. doi: 10.1016/j.vaccine.2017.11.007

Bayó, M., Berlanga, M., and Agut, M. (2002). Vaginal microbiota in healthy pregnant women and prenatal screening of group B streptococci (GBS). *Int. Microbiol.* 5, 87–90. doi: 10.1007/s10123-002-0064-1

Beg, A. M., Jones, M. N., Miller-Torbert, T., and Holt, R. G. (2002). Binding of *Streptococcus mutans* to extracellular matrix molecules and fibrinogen. *Biochem. Biophys. Res. Commun.* 298, 75–79. doi: 10.1016/s0006-291x(02)02390-2

Bikker, F. J., Ligtenberg, A. J., End, C., Renner, M., Blaich, S., Lyer, S., et al. (2004). Bacteria binding by DMBT1/SAG/gp-340 is confined to the VEVLXXXXW motif in its scavenger receptor cysteine-rich domains. *J. Biol. Chem.* 279, 47699–47703. doi: 10.1074/jbc.M406095200

Bikker, F. J., Ligtenberg, A. J., Nazmi, K., Veerman, E. C., van't Hof, W., Bolscher, J. G., et al. (2002). Identification of the bacteria-binding peptide domain on salivary agglutinin (gp-340/DMBT1), a member of the scavenger receptor cysteine-rich superfamily. *J. Biol. Chem.* 277, 32109–32115. doi: 10.1074/jbc.M203788200

Bleiweis, A. S., Oyston, P. C., and Brady, L. J. (1992). Molecular, immunological and functional characterization of the major surface adhesin of *Streptococcus mutans*. *Adv. Exp. Med. Biol.* 327, 229–241. doi: 10.1007/978-1-4615-3410-5_25

Brady, L. J., Crowley, P. J., Ma, J. K., Kelly, C., Lee, S. F., Lehner, T., et al. (1991). Restriction fragment length polymorphisms and sequence variation within the spaP gene of *Streptococcus mutans* serotype c isolates. *Infect. Immun.* 59, 1803–1810.

Brady, L. J., Maddocks, S. E., Larson, M. R., Forsgren, N., Persson, K., Deivanayagam, C. C., et al. (2010). The changing faces of *Streptococcus* antigen I/II polypeptide family adhesins. *Mol. Microbiol.* 77, 276–286. doi: 10.1111/j.1365-2958.2010.07212.x

Brady, L. J., Piacentini, D. A., Crowley, P. J., Oyston, P. C., and Bleiweis, A. S. (1992). Differentiation of salivary agglutinin-mediated adherence and aggregation of mutans streptococci by use of monoclonal antibodies against the major surface adhesin P1. *Infect. Immun.* 60, 1008–1017.

Brooks, W., Demuth, D. R., Gil, S., and Lamont, R. J. (1997). Identification of a *Streptococcus gordonii* SspB domain that mediates adhesion to Porphyromonas gingivalis. *Infect. Immun.* 65, 3753–3758.

Busscher, H. J., van de Belt-Gritter, B., Dijkstra, R. J., Norde, W., Petersen, F. C., Scheie, A. A., et al. (2007). Intermolecular forces and enthalpies in the adhesion of *Streptococcus mutans* and an antigen I/II-deficient mutant to laminin films. *J. Bacteriol.* 189, 2988–2995. doi: 10.1128/JB.01731-06

Chaffanel, F., Charron-Bourgoin, F., Soligot, C., Kebouchi, M., Bertin, S., Payot, S., et al. (2018). Surface proteins involved in the adhesion of *Streptococcus salivarius* to human intestinal epithelial cells. *Appl. Microbiol. Biotechnol.* 102, 2851–2865. doi: 10.1007/s00253-018-8794-y

Chatenay-Rivauday, C., Yamodo, I., Sciotti, M. A., Ogier, J. A., and Klein, J. P. (1998). The A and the extended V N-terminal regions of streptococcal protein I/IIf mediate the production of tumour necrosis factor alpha in the monocyte cell line THP-1. *Mol. Microbiol.* 29, 39–48. doi: 10.1046/j.1365-2958.1998.00881.x

Chatenay-Rivauday, C., Yamodo, I., Sciotti, M. A., Troffer-Charlier, N., Klein, J. P., and Ogier, J. A. (2000). TNF-alpha release by monocytic THP-1 cells through cross-linking of the extended V-region of the oral streptococcal protein I/II. *J. Leukoc. Biol.* 67, 81–89. doi: 10.1002/jlb.67.1.81

Chuang, O. N., Schlievert, P. M., Wells, C. L., Manias, D. A., Tripp, T. J., and Dunny, G. M. (2009). Multiple functional domains of Enterococcus faecalis aggregation substance Asc10 contribute to endocarditis virulence. *Infect. Immun.* 77, 539–548. doi: 10.1128/iai.01034-08

Chuzeville, S., Auger, J. P., Dumesnil, A., Roy, D., Lacouture, S., Fittipaldi, N., et al. (2017). Serotype-specific role of antigen I/II in the initial steps of the pathogenesis of the infection caused by *Streptococcus suis*. *Vet. Res.* 48:39. doi: 10.1186/s13567-017-0443-4

Chuzeville, S., Dramsi, S., Madec, J. Y., Haenni, M., and Payot, S. (2015). Antigen I/II encoded by integrative and conjugative elements of *Streptococcus agalactiae* and role in biofilm formation. *Microb. Pathog.* 88, 1–9. doi: 10.1016/j.micpath.2015.07.018

Cools, P., Jespers, V., Hardy, L., Crucitti, T., Delany-Moretlwe, S., Mwaura, M., et al. (2016). A Multi-Country Cross-Sectional Study of Vaginal Carriage of Group B Streptococci (GBS) and *Escherichia coli* in Resource-Poor Settings: Prevalences and Risk Factors. *PLoS One* 11:e0148052. doi: 10.1371/journal.pone.0148052

Daep, C. A., James, D. M., Lamont, R. J., and Demuth, D. R. (2006). Structural characterization of peptide-mediated inhibition of Porphyromonas gingivalis biofilm formation. *Infect. Immun.* 74, 5756–5762. doi: 10.1128/iai.00813-06

Daep, C. A., Novak, E. A., Lamont, R. J., and Demuth, D. R. (2011). Structural Dissection and In Vivo Effectiveness of a Peptide Inhibitor of Porphyromonas gingivalis Adherence to *Streptococcus gordonii*. *Infect. Immun.* 79, 67–74. doi: 10.1128/iai.00361-10

Demuth, D. R., Davis, C. A., Corner, A. M., Lamont, R. J., Leboy, P. S., and Malamud, D. (1988). Cloning and expression of a *Streptococcus sanguis* surface antigen that interacts with a human salivary agglutinin. *Infect. Immun.* 56, 2484–2490.

Demuth, D. R., Duan, Y., Brooks, W., Holmes, A. R., McNab, R., and Jenkinson, H. F. (1996). Tandem genes encode cell-surface polypeptides SspA and SspB which mediate adhesion of the oral bacterium *Streptococcus gordonii* to human and bacterial receptors. *Mol. Microbiol.* 20, 403–413. doi: 10.1111/j.1365-2958.1996.tb02627.x

Demuth, D. R., Golub, E. E., and Malamud, D. (1990a). Streptococcal-host interactions. Structural and functional analysis of a *Streptococcus sanguis* receptor for a human salivary glycoprotein. *J. Biol. Chem.* 265, 7120–7126.

Demuth, D. R., Irvine, D. C., Costerton, J. W., Cook, G. S., and Lamont, R. J. (2001). Discrete Protein Determinant Directs the Species-Specific Adherence of Porphyromonas gingivalis to Oral Streptococci. *Infect. Immun.* 69, 5736–5741. doi: 10.1128/iai.69.9.5736-5741.2001

Demuth, D. R., Lammey, M. S., Huck, M., Lally, E. T., and Malamud, D. (1990b). Comparison of *Streptococcus mutans* and *Streptococcus sanguis* receptors for human salivary agglutinin. *Microb. Pathog.* 9, 199–211. doi: 10.1016/0882-4010(90)90022-i

Deng, L., Spencer, B. L., Holmes, J. A., Mu, R., Rego, S., Weston, T. A., et al. (2019). The Group B Streptococcal surface antigen I/II protein, BspC, interacts with host vimentin to promote adherence to brain endothelium and inflammation during the pathogenesis of meningitis. *PLoS Pathog.* 15:e1007848. doi: 10.1371/journal.ppat.1007848

Do, T., Gilbert, S. C., Clark, D., Ali, F., Fatturi Parolo, C. C., Maltz, M., et al. (2010). Generation of diversity in *Streptococcus mutans* genes demonstrated by MLST. *PLoS One* 5:e9073. doi: 10.1371/journal.pone.0009073

Doran, K. S., Fulde, M., Gratz, N., Kim, B. J., Nau, R., Prasadarao, N., et al. (2016). Host-pathogen interactions in bacterial meningitis. *Acta Neuropathol.* 131, 185–209. doi: 10.1007/s00401-015-1531-z

Edmond, K. M., Kortsalioudaki, C., Scott, S., Schrag, S. J., Zaidi, A. K., Cousens, S., et al. (2012). Group B streptococcal disease in infants aged younger than 3 months: systematic review and meta-analysis. *Lancet* 379, 547–556. doi: 10.1016/s0140-6736(11)61651-6

Engels-Deutsch, M., Pini, A., Yamashita, Y., Shibata, Y., Haikel, Y., Schöller-Guinard, M., et al. (2003). Insertional inactivation of pac and rmlB genes reduces the release of tumor necrosis factor alpha, interleukin-6, and interleukin-8 induced by *Streptococcus mutans* in monocytic, dental pulp, and periodontal ligament cells. *Infect. Immun.* 71, 5169–5177. doi: 10.1128/iai.71.9.5169-5177.2003

Esberg, A., Lofgren-Burstrom, A., Ohman, U., and Stromberg, N. (2012). Host and bacterial phenotype variation in adhesion of *Streptococcus mutans* to matched human hosts. *Infect. Immun.* 80, 3869–3879. doi: 10.1128/IAI.00435-12

Esberg, A., Sheng, N., Mårell, L., Claesson, R., Persson, K., Borén, T., et al. (2017). *Streptococcus mutans* Adhesin Biotypes that Match and Predict Individual Caries Development. *EBio Med.* 24, 205–215. doi: 10.1016/j.ebiom.2017.09.027

Falsetta, M. L., Klein, M. I., Colonne, P. M., Scott-Anne, K., Gregoire, S., Pai, C. H., et al. (2014). Symbiotic relationship between *Streptococcus mutans* and Candida albicans synergizes virulence of plaque biofilms in vivo. *Infect. Immun.* 82, 1968–1981. doi: 10.1128/iai.00087-14

Ford, I., and Douglas, C. W. (1997). The role of platelets in infective endocarditis. *Platelets* 8, 285–294. doi: 10.1080/09537109777159

Forsgren, N., Lamont, R. J., and Persson, K. (2009). Crystal structure of the variable domain of the *Streptococcus gordonii* surface protein SspB. *Protein. Sci.* 18, 1896–1905. doi: 10.1002/pro.200

Forsgren, N., Lamont, R. J., and Persson, K. (2010). Two intramolecular isopeptide bonds are identified in the crystal structure of the *Streptococcus gordonii* SspB C-terminal domain. *J. Mol. Biol.* 397, 740–751. doi: 10.1016/j.jmb.2010.01.065

Franklin, L., Nobbs, A. H., Bricio-Moreno, L., Wright, C. J., Maddocks, S. E., Sahota, J. S., et al. (2013). The AgI/II family adhesin AspA is required for respiratory infection by *Streptococcus pyogenes*. *PLoS One* 8:e62433. doi: 10.1371/journal.pone.0062433

Gartler, S. M. (1968). Apparent Hela cell contamination of human heteroploid cell lines. *Nature* 217, 750–751. doi: 10.1038/217750a0

Guo, J. H., Jia, R., Fan, M. W., Bian, Z., Chen, Z., and Peng, B. (2004). Construction and immunogenic characterization of a fusion anti-caries DNA vaccine against PAc and glucosyltransferase I of *Streptococcus mutans*. *J. Dent. Res.* 83, 266–270. doi: 10.1177/154405910408300316

Guo, L., Shokeen, B., He, X., Shi, W., and Lux, R. (2017). *Streptococcus mutans* SpaP binds to RadD of *Fusobacterium nucleatum* ssp. polymorphum. *Mole. Microbiol.* 32, 355–364. doi: 10.1111/omi.12177

Hajishengallis, G., Koga, T., and Russell, M. W. (1994). Affinity and specificity of the interactions between *Streptococcus mutans* antigen I/II and salivary components. *J. Dent. Res.* 73, 1493–1502. doi: 10.1177/00220345940730090301

Hajishengallis, G., Martin, M., Sojar, H. T., Sharma, A., Schifferle, R. E., DeNardin, E., et al. (2002). Dependence of Bacterial Protein Adhesins on Toll-Like Receptors for Proinflammatory Cytokine Induction. *Clin. Diagnos. Lab. Immunol.* 9, 403–411. doi: 10.1128/cdli.9.2.403-411.2002

Hajishengallis, G., Nikolova, E., and Russell, M. W. (1992). Inhibition of *Streptococcus mutans* adherence to saliva-coated hydroxyapatite by human secretory immunoglobulin A (S-IgA) antibodies to cell surface protein antigen I/II: reversal by IgA1 protease cleavage. *Infect. Immun.* 60, 5057–5064.

Hajishengallis, G., Russell, M. W., and Michalek, S. M. (1998). Comparison of an adherence domain and a structural region of *Streptococcus mutans* antigen I/II in protective immunity against dental caries in rats after intranasal immunization. *Infect. Immun.* 66, 1740–1743.

Hall, M., Nylander, S., Jenkinson, H. F., and Persson, K. (2014). Structure of the C-terminal domain of AspA (antigen I/II-family) protein from *Streptococcus pyogenes*. *FEBS Open Bio.* 4, 283–289. doi: 10.1016/j.fob.2014.02.012

Heddle, C., Nobbs, A. H., Jakubovics, N. S., Gal, M., Mansell, J. P., Dymock, D., et al. (2003). Host collagen signal induces antigen I/II adhesin and invasin gene expression in oral *Streptococcus gordonii*. *Mol. Microbiol.* 50, 597–607. doi: 10.1046/j.1365-2958.2003.03711.x

Heim, K. P., Crowley, P. J., Long, J. R., Kailasan, S., McKenna, R., and Brady, L. J. (2014). An intramolecular lock facilitates folding and stabilizes the tertiary structure of *Streptococcus mutans* adhesin P1. *Proc. Natl. Acad. Sci. U S A* 111, 15746–15751. doi: 10.1073/pnas.1413018111

Heim, K. P., Sullan, R. M., Crowley, P. J., El-Kirat-Chatel, S., Beaussart, A., Tang, W., et al. (2015). Identification of a supramolecular functional architecture of *Streptococcus mutans* adhesin P1 on the bacterial cell surface. *J. Biol. Chem.* 290, 9002–9019. doi: 10.1074/jbc.M114.626663

Hoen, B., Alla, F., Selton-Suty, C., Beguinot, I., Bouvet, A., Briancon, S., et al. (2002). Changing profile of infective endocarditis: results of a 1-year survey in France. *JAMA* 288, 75–81. doi: 10.1001/jama.288.1.75

Huang, Y., Hajishengallis, G., and Michalek, S. M. (2001). Induction of protective immunity against *Streptococcus mutans* colonization after mucosal immunization with attenuated *Salmonella enterica* serovar typhimurium expressing an S. mutans adhesin under the control of in vivo-inducible nirB promoter. *Infect. Immun.* 69, 2154–2161. doi: 10.1128/IAI.69.4.2154-2161.2001

Jain, I., Sarkar, P., Danger, J. L., Medicielo, J., Roshika, R., Calfee, G., et al. (2019). A Mobile Genetic Element Promotes the Association Between Serotype M28 Group A *Streptococcus* Isolates and Cases of Puerperal Sepsis. *J. Infect. Dis.* 220, 882–891. doi: 10.1093/infdis/jiz195

Jakubovics, N. S., Kerrigan, S. W., Nobbs, A. H., Stromberg, N., van Dolleweerd, C. J., Cox, D. M., et al. (2005a). 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. doi: 10.1128/IAI.73.10.6629-6638.2005

Jakubovics, N. S., Stromberg, N., van Dolleweerd, C. J., Kelly, C. G., and Jenkinson, H. F. (2005b). Differential binding specificities of oral streptococcal antigen I/II

family adhesins for human or bacterial ligands. *Mol. Microbiol.* 55, 1591–1605. doi: 10.1111/j.1365-2958.2005.04495.x

Jenkinson, H. F., and Demuth, D. R. (1997). Structure, function and immunogenicity of streptococcal antigen I/II polypeptides. *Mol. Microbiol.* 23, 183–190. doi: 10.1046/j.1365-2958.1997.2021577.x

Jenkinson, H. F., Terry, S. D., McNab, R., and Tannock, G. W. (1993). Inactivation of the gene encoding surface protein SspA in *Streptococcus gordonii* DL1 affects cell interactions with human salivary agglutinin and oral actinomyces. *Infect. Immun.* 61, 3199–3208.

Jia, R., Guo, J. H., Fan, M. W., Bian, Z., Chen, Z., Peng, B., et al. (2004). Mucosal immunization against dental caries with plasmid DNA encoding pac gene of *Streptococcus mutans* in rats. *Vaccine* 22, 2511–2516. doi: 10.1016/j.vaccine.2004.01.025

Jia, R., Yan, L., and Guo, J. (2020). Enhancing the immunogenicity of a DNA vaccine against *Streptococcus mutans* by attenuating the inhibition of endogenous miR-9. *Vaccine* 38, 1424–1430. doi: 10.1016/j.vaccine.2019.11.083

Jiang, H., Hu, Y., Yang, M., Liu, H., and Jiang, G. (2017). Enhanced immune response to a dual-promoter anti-caries DNA vaccine orally delivered by attenuated *Salmonella* typhimurium. *Immunobiology* 222, 730–737. doi: 10.1016/j.imbio.2017.01.007

Jonasson, A., Eriksson, C., Jenkinson, H. F., Källestål, C., Johansson, I., and Strömberg, N. (2007). Innate immunity glycoprotein gp-340 variants may modulate human susceptibility to dental caries. *BMC Infect. Dis.* 7:57. doi: 10.1186/1471-2334-7-57

Jones, D. T., Taylor, W. R., and Thornton, J. M. (1992). The rapid generation of mutation data matrices from protein sequences. *Comput. Appl. Biosci.* 8, 275–282. doi: 10.1093/bioinformatics/8.3.275

Katz, J., Harmon, C. C., Buckner, G. P., Richardson, G. J., Russell, M. W., and Michalek, S. M. (1993). Protective salivary immunoglobulin A responses against *Streptococcus mutans* infection after intranasal immunization with S. mutans antigen I/II coupled to the B subunit of cholera toxin. *Infect. Immun.* 61, 1964–1971.

Kelemen, L., Rizk, S., Debreczeny, M., Ogier, J., and Szalontai, B. (2004). Streptococcal antigen I/II binds to extracellular proteins through intermolecular beta-sheets. *FEBS Lett.* 566, 190–194. doi: 10.1016/j.febslet.2004.04.029

Kelly, C. G., Todryk, S., Kendal, H. L., Munro, G. H., and Lehner, T. T. - (1995). cell, adhesion, and B-cell epitopes of the cell surface *Streptococcus mutans* protein antigen I/II. *Infect. Immun.* 63, 3649–3658.

Kelly, C. G., Younson, J. S., Hikmat, B. Y., Todryk, S. M., Czisch, M., Haris, P. I., et al. (1999). A synthetic peptide adhesion epitope as a novel antimicrobial agent. *Nat. Biotechnol.* 17, 42–47. doi: 10.1038/5213

Kelly, C., Evans, P., Bergmeier, L., Lee, S. F., Progulske-Fox, A., Harris, A. C., et al. (1989). Sequence analysis of the cloned streptococcal cell surface antigen I/II. *FEBS Lett.* 258, 127–132. doi: 10.1016/0014-5793(89)81632-1

Kerrigan, S. W., Jakubovics, N. S., Keane, C., Maguire, P., Wynne, K., Jenkinson, H. F., et al. (2007). Role of *Streptococcus gordonii* surface proteins SspA/SspB and Hsa in platelet function. *Infect. Immun.* 75, 5740–5747. doi: 10.1128/iai.00909-07

Kobayashi, M., Vekemans, J., Baker, C. J., Ratner, A. J., Le Doare, K., and Schrag, S. J. (2016). Group B *Streptococcus* vaccine development: present status and future considerations, with emphasis on perspectives for low and middle income countries. *F1000Res.* 5:2355. doi: 10.12688/f1000research.9363.1

Kumar, S., Stecher, G., Li, M., Knyaz, C., and Tamura, K. (2018). MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. *Mol. Biol. Evol.* 35, 1547–1549. doi: 10.1093/molbev/msy096

Lamont, R. J., Demuth, D. R., Davis, C. A., Malamud, D., and Rosan, B. (1991). Salivary-agglutinin-mediated adherence of *Streptococcus mutans* to early plaque bacteria. *Infect. Immun.* 59, 3446–3450.

LaPolla, R. J., Haron, J. A., Kelly, C. G., Taylor, W. R., Bohart, C., Hendricks, M., et al. (1991). Sequence and structural analysis of surface protein antigen I/II (SpaA) of *Streptococcus sobrinus*. *Infect. Immun.* 59, 2677–2685.

Larson, M. R., Rajashankar, K. R., Crowley, P. J., Kelly, C., Mitchell, T. J., Brady, L. J., et al. (2011). Crystal structure of the C-terminal region of *Streptococcus mutans* antigen I/II and characterization of salivary agglutinin adherence domains. *J. Biol. Chem.* 286, 21657–21666. doi: 10.1074/jbc.M111.231100

Larson, M. R., Rajashankar, K. R., Patel, M. H., Robinette, R. A., Crowley, P. J., Michalek, S., et al. (2010). Elongated fibrillar structure of a streptococcal adhesin assembled by the high-affinity association of alpha- and PPII-helices. *Proc. Natl. Acad. Sci. U S A* 107, 5983–5988. doi: 10.1073/pnas.0912293107

Lehner, T., Caldwell, J., and Smith, R. (1985). Local passive immunization by monoclonal antibodies against streptococcal antigen I/II in the prevention of dental caries. *Infect. Immun.* 50, 796–799.

Lehner, T., Challacombe, S. J., and Caldwell, J. (1975). Immunological and bacteriological basis for vaccination against dental caries in rhesus monkeys. *Nature* 254, 517–520. doi: 10.1038/254517a0

Lehner, T., Ma, J. K., and Kelly, C. G. (1992). A mechanism of passive immunization with monoclonal antibodies to a 185,000 M(r) streptococcal antigen. *Adv. Exp. Med. Biol.* 327, 151–163. doi: 10.1007/978-1-4615-3410-5_17

Lehner, T., Russell, M. W., Caldwell, J., and Smith, R. (1981). Immunization with purified protein antigens from *Streptococcus mutans* against dental caries in rhesus monkeys. *Infect. Immun.* 34, 407–415.

Linde, A. (1989). Dentin matrix proteins: composition and possible functions in calcification. *Anat. Rec.* 224, 154–166. doi: 10.1002/ar.1092240206

Loimaranta, V., Jakubovics, N. S., Hytönen, J., Finne, J., Jenkinson, H. F., and Strömberg, N. (2005). Fluid- or Surface-Phase Human Salivary Scavenger Protein gp340 Exposes Different Bacterial Recognition Properties. *Infect. Immun.* 73, 2245–2252. doi: 10.1128/iai.73.4.2245-2252.2005

Love, R. M., McMillan, M. D., and Jenkinson, H. F. (1997). Invasion of dentinal tubules by oral streptococci is associated with collagen recognition mediated by the antigen I/II family of polypeptides. *Infect. Immun.* 65, 5157–5164.

Lucchese, A. (2017). *Streptococcus mutans* antigen I/II and autoimmunity in cardiovascular diseases. *Autoimmun. Rev.* 16, 456–460. doi: 10.1016/j.autrev.2017.03.009

Ma, J. K., and Lehner, T. (1990). Prevention of colonization of *Streptococcus mutans* by topical application of monoclonal antibodies in human subjects. *Arch. Oral Biol.* 35(Suppl.), 115S–122S. doi: 10.1016/0003-9969(90)90140-6

Ma, J. K., Hikmat, B. Y., Wycoff, K., Vine, N. D., Chargelegue, D., Yu, L., et al. (1998). Characterization of a recombinant plant monoclonal secretory antibody and preventive immunotherapy in humans. *Nat. Med.* 4, 601–606. doi: 10.1038/nm0598-601

Ma, J. K., Hunjan, M., Smith, R., and Lehner, T. (1989). Specificity of monoclonal antibodies in local passive immunization against *Streptococcus mutans*. *Clin. Exp. Immunol.* 77, 331–337.

Ma, J. K., Hunjan, M., Smith, R., Kelly, C., and Lehner, T. (1990). An investigation into the mechanism of protection by local passive immunization with monoclonal antibodies against *Streptococcus mutans*. *Infect. Immun.* 58, 3407–3414.

Ma, J. K., Kelly, C. G., Munro, G., Whiley, R. A., and Lehner, T. (1991). Conservation of the gene encoding streptococcal antigen I/II in oral streptococci. *Infect. Immun.* 59, 2686–2694.

Maddocks, S. E., Wright, C. J., Nobbs, A. H., Brittan, J. L., Franklin, L., Strömberg, N., et al. (2011). *Streptococcus pyogenes* antigen I/II-family polypeptide AspA shows differential ligand-binding properties and mediates biofilm formation. *Mole. Microbiol.* 81, 1034–1049. doi: 10.1111/j.1365-2958.2011.07749.x

Madsen, J., Mollenhauer, J., and Holmskov, U. (2010). Review: Gp-340/DMBT1 in mucosal innate immunity. *Innate Immun.* 16, 160–167. doi: 10.1177/1753425910368447

Mahmoud, M. Y., Demuth, D. R., and Steinbach-Rankins, J. M. (2018). BAR-encapsulated nanoparticles for the inhibition and disruption of Porphyromonas gingivalis-*Streptococcus gordonii* biofilms. *J. Nanobiotechnol.* 16:69. doi: 10.1186/s12951-018-0396-4

Matsumoto-Nakano, M., Tsuji, M., Inagaki, S., Fujita, K., Nagayama, K., Nomura, R., et al. (2009). Contribution of cell surface protein antigen c of *Streptococcus mutans* to platelet aggregation. *Oral Microbiol. Immunol.* 24, 427–430. doi: 10.1111/j.1399-302X.2009.00521.x

Mollenhauer, J., Wiemann, S., Scheurlen, W., Korn, B., Hayashi, Y., Wilgenbus, K. K., et al. (1997). DMBT1, a new member of the SRCR superfamily, on chromosome 10q25.3-26.1 is deleted in malignant brain tumours. *Nat. Genet.* 17, 32–39. doi: 10.1038/ng0997-32

Monif, G. R., and Carson, H. J. (1998). Female genital tract bacterial coisolates with Candida albicans in patients without clinical vaginitis. *Infect. Dis. Obstet. Gynecol.* 6, 52–56. doi: 10.1002/(sici)1098-099719986:2<52::Aid-idog4<3.0.Co;2-9

Nakano, K., Tsuji, M., Nishimura, K., Nomura, R., and Ooshima, T. (2006). Contribution of cell surface protein antigen PAc of *Streptococcus mutans* to bacteremia. *Microb. Infect.* 8, 114–121. doi: 10.1016/j.micinf.2005.06.005

Nobbs, A. H., Shearer, B. H., Drobni, M., Jepson, M. A., and Jenkinson, H. F. (2007). Adherence and internalization of *Streptococcus gordonii* by epithelial cells involves beta1 integrin recognition by SspA and SspB (antigen I/II family) polypeptides. *Cell Microbiol.* 9, 65–83. doi: 10.1111/j.1462-5822.2006.00768.x

Nobbs, A., and Kreth, J. (2019). Genetics of sanguinis-Group Streptococci in Health and Disease. *Microbiol. Spectr.* 7:0052. doi: 10.1128/microbiolspec.GPP3-0052-2018

Ogier, J. A., Scholler, M., Leproivre, Y., Pini, A., Sommer, P., and Klein, J. P. (1990). Complete nucleotide sequence of the sr gene from *Streptococcus mutans* OMZ 175. *FEMS Microbiol. Lett.* 56, 223–227. doi: 10.1016/0378-1097(90)90155-j

Oho, T., and Nagata, E. (2019). DMBT1 involvement in the human aortic endothelial cell response to *Streptococcus mutans*. *Mole. Microbiol.* 34, 108–117. doi: 10.1111/omi.12257

Oli, M. W., Otoo, H. N., Crowley, P. J., Heim, K. P., Nascimento, M. M., Ramsook, C. B., et al. (2012). Functional amyloid formation by *Streptococcus mutans*. *Microbiology* 158, 2903–2916. doi: 10.1099/mic.0.060855-0

Petersen, F. C., Assev, S., van der Mei, H. C., Busscher, H. J., and Scheie, A. A. (2002). Functional variation of the antigen I/II surface protein in *Streptococcus mutans* and *Streptococcus intermedius*. *Infect. Immun.* 70, 249–256. doi: 10.1128/iai.70.1.249-256.2002

Petersen, F. C., Pasco, S., Ogier, J., Klein, J. P., Assev, S., and Scheie, A. A. (2001). Expression and functional properties of the *Streptococcus intermedius* surface protein antigen I/II. *Infect. Immun.* 69, 4647–4653. doi: 10.1128/IAI.69.7.4647-4653.2001

Pidwill, G. R., Rego, S., Jenkinson, H. F., Lamont, R. J., and Nobbs, A. H. (2018). Coassociation between Group B *Streptococcus* and *Candida albicans* Promotes Interactions with Vaginal Epithelium. *Infect. Immun.* 86:e00669-17. doi: 10.1128/iai.00669-17

Prakobphol, A., Xu, F., Hoang, V. M., Larsson, T., Bergstrom, J., Johansson, I., et al. (2000). Salivary Agglutinin, Which Binds *Streptococcus mutans* and *Helicobacter* pylori, Is the Lung Scavenger Receptor Cysteine-rich Protein gp-340. *J. Biol. Chem.* 275, 39860–39866. doi: 10.1074/jbc.M006928200

Purushotham, S., and Deivanayagam, C. (2014). The calcium-induced conformation and glycosylation of scavenger-rich cysteine repeat (SRCR) domains of glycoprotein 340 influence the high affinity interaction with antigen I/II homologs. *J. Biol. Chem.* 289, 21877–21887. doi: 10.1074/jbc.M114.565507

Rajani, R., and Klein, J. L. (2020). Infective endocarditis: A contemporary update. *Clin. Med.* 20, 31–35. doi: 10.7861/clinmed.cme.20.1.1

Rajendiran, K. S., Ananthanarayanan, R. H., Satheesh, S., and Rajappa, M. (2014). Elevated levels of serum sialic acid and high-sensitivity C-reactive protein: markers of systemic inflammation in patients with chronic heart failure. *Br. J. Biomed. Sci.* 71, 29–32. doi: 10.1080/09674845.2014.11669959

Rego, S., Heal, T. J., Pidwill, G. R., Till, M., Robson, A., Lamont, R. J., et al. (2016). Structural and Functional Analysis of Cell Wall-anchored Polypeptide Adhesin BspA in *Streptococcus agalactiae*. *J. Biol. Chem.* 291, 15985–16000. doi: 10.1074/jbc.M116.726562

Rivière, G., Peng, E. Q., Brotgandel, A., Andring, J. T., Lakshmanan, R. V., Agbandje-McKenna, M., et al. (2020). Characterization of an intermolecular quaternary interaction between discrete segments of the *Streptococcus mutans* adhesin P1 by NMR spectroscopy. *FEBS J.*287, 2597–2611. doi: 10.1111/febs.15158

Roky, M., Tan, J., Sztukowska, M. N., Trent, J. O., and Demuth, D. R. (2020). Identification of small molecule inhibitors targeting Porphyromonas gingivalis interspecies adherence and determination of their in vitro and in vivo efficacy. *Antimicrob. Agent Chemother.* 2020:00884-20. doi: 10.1128/aac.00884-20

Russell, M. W., and Lehner, T. (1978). Characterisation of antigens extracted from cells and culture fluids of *Streptococcus mutans* serotype c. *Arch. Oral. Biol.* 23, 7–15. doi: 10.1016/0003-9969(78)90047-x

Russell, M. W., Bergmeier, L. A., Zanders, E. D., and Lehner, T. (1980). Protein antigens of *Streptococcus mutans*: purification and properties of a double antigen and its protease-resistant component. *Infect. Immun.* 28, 486–493.

Russell, M. W., Wu, H. Y., White, P. L., Kilian, M., and Henrichsen, J. (1992). Serum antibody responses to *Streptococcus mutans* antigens in humans systemically infected with oral streptococci. *Oral Microbiol. Immunol.* 7, 321–325. doi: 10.1111/j.1399-302x.1992.tb00630.x

Russell, R. R. (1979). Wall-associated protein antigens of *Streptococcus mutans*. *J. Gen. Microbiol.* 114, 109–115. doi: 10.1099/00221287-114-1-109

Ryd, M., Schennings, T., Flock, M., Heimdahl, A., and Flock, J. I. (1996). *Streptococcus mutans* major adhesion surface protein, P1 (I/II), does not contribute to attachment to valvular vegetations or to the development of endocarditis in a rat model. *Arch. Oral Biol.* 41, 999–1002. doi: 10.1016/s0003-9969(96)00035-0

Schmitt, A., Jiang, K., Camacho, M. I., Jonna, V. R., Hofer, A., Westerlund, F., et al. (2018). PrgB promotes aggregation, biofilm formation, and conjugation through DNA binding and compaction. *Mol. Microbiol.* 109, 291–305. doi: 10.1111/mmi.13980

Sciotti, M. A., Yamodo, I., Klein, J. P., and Ogier, J. A. (1997). The N-terminal half part of the oral streptococcal antigen I/IIf contains two distinct binding domains. *FEMS Microbiol. Lett.* 153, 439–445. doi: 10.1111/j.1574-6968.1997.tb12608.x

Scott, J. R., and Barnett, T. C. (2006). Surface proteins of gram-positive bacteria and how they get there. *Annu. Rev. Microbiol.* 60, 397–423. doi: 10.1146/annurev.micro.60.080805.142256

Shi, W., Li, Y. H., Liu, F., Yang, J. Y., Zhou, D. H., Chen, Y. Q., et al. (2012). Flagellin enhances saliva IgA response and protection of anti-caries DNA vaccine. *J. Dent. Res.* 91, 249–254. doi: 10.1177/0022034511424283

Silverman, R. J., Nobbs, A. H., Vickerman, M. M., Barbour, M. E., and Jenkinson, H. F. (2010). Interaction of Candida albicans cell wall Als3 protein with *Streptococcus gordonii* SspB adhesin promotes development of mixed-species communities. *Infect. Immun.* 78, 4644–4652. doi: 10.1128/IAI.00685-10

Singh, B., Fleury, C., Jalalvand, F., and Riesbeck, K. (2012). Human pathogens utilize host extracellular matrix proteins laminin and collagen for adhesion and invasion of the host. *FEMS Microbiol. Rev.* 36, 1122–1180. doi: 10.1111/j.1574-6976.2012.00340.x

Sitkiewicz, I., Green, N. M., Guo, N., Mereghetti, L., and Musser, J. M. (2011). Lateral gene transfer of streptococcal ICE element RD2 (region of difference 2) encoding secreted proteins. *BMC Microbiol.* 11:65. doi: 10.1186/1471-2180-11-65

Soell, M., Holveck, F., Schöller, M., Wachsmann, R. D., and Klein, J. P. (1994). Binding of *Streptococcus mutans* SR protein to human monocytes: production of tumor necrosis factor, interleukin 1, and interleukin 6. *Infect. Immun.* 62, 1805–1812.

Son, Y.-O., Jeon, Y.-M., Kim, Y.-S., Park, S.-S., Park, S.-M., Kim, J.-H., et al. (2012). *Streptococcus mutans* GS-5 antigen I/II stimulates cell survival in serum deprived-cultures through PI3K/Akt Pathways. *J. Cell. Biochem.* 113, 1724–1732. doi: 10.1002/jcb.24042

Steer, A. C., Lamagni, T., Curtis, N., and Carapetis, J. R. (2012). Invasive group a streptococcal disease: epidemiology, pathogenesis and management. *Drugs* 72, 1213–1227. doi: 10.2165/11634180-000000000-00000

Sullan, R. M., Li, J. K., Crowley, P. J., Brady, L. J., and Dufrene, Y. F. (2015). Binding forces of *Streptococcus mutans* P1 adhesin. *ACS Nano* 9, 1448–1460. doi: 10.1021/nn5058886

Tamura, H., Kikuchi, T., Shirato, R., and Kato, H. (2001). Cloning and DNA sequencing of the surface protein antigen I/II (PAa) of *Streptococcus cricetus*. *FEMS Microbiol. Lett.* 196, 251–256. doi: 10.1111/j.1574-6968.2001.tb10573.x

Tamura, H., Yamada, A., and Kato, H. (2008). Identification and characterization of an antigen I/II homologous gene, pah, from *Streptococcus downei*. *Curr. Microbiol.* 56, 518–523. doi: 10.1007/s00284-008-9117-9

Tamura, H., Yamada, A., Saito, H., Murai, S., and Kato, H. (2004). Identification of another surface protein antigen I/II gene, paaB, and a putative transcriptional regulator gene, par, from *Streptococcus cricetus*. *Gene Genet. Syst.* 79, 129–137. doi: 10.1266/ggs.79.129

Tang, W., Bhatt, A., Smith, A. N., Crowley, P. J., Brady, L. J., and Long, J. R. (2016). Specific binding of a naturally occurring amyloidogenic fragment of *Streptococcus mutans* adhesin P1 to intact P1 on the cell surface characterized by solid state NMR spectroscopy. *J. Biomol. NMR* 64, 153–164. doi: 10.1007/s10858-016-0017-1

Tettelin, H., Masignani, V., Cieslewicz, M., Donati, C., Medini, D., Ward, N., et al. (2005). Genome analysis of multiple pathogenic isolates of *Streptococcus agalactiae*: Implications for the microbial "pan-genome". *Proc. Natl. Acad. Sci. U S A* 102, 13950–13955. doi: 10.1073/pnas.0506758102

Troffer-Charlier, N., Ogier, J., Moras, D., and Cavarelli, J. (2002). Crystal structure of the V-region of *Streptococcus mutans* antigen I/II at 2.4 A resolution suggests a sugar preformed binding site. *J. Mol. Biol.* 318, 179–188. doi: 10.1016/S0022-2836(02)00025-6

Vernier, A., Diab, M., Soell, M., Haan-Archipoff, G., Beretz, A., Wachsmann, D., et al. (1996). Cytokine production by human epithelial and endothelial cells

following exposure to oral viridans streptococci involves lectin interactions between bacteria and cell surface receptors. *Infect. Immun.* 64, 3016–3022.

Waters, C. M., Hirt, H., McCormick, J. K., Schlievert, P. M., Wells, C. L., and Dunny, G. M. (2004). An amino-terminal domain of Enterococcus faecalis aggregation substance is required for aggregation, bacterial internalization by epithelial cells and binding to lipoteichoic acid. *Mol. Microbiol.* 52, 1159–1171. doi: 10.1111/j.1365-2958.2004.04045.x

Weintraub, J. A., Hilton, J. F., White, J. M., Hoover, C. I., Wycoff, K. L., Yu, L., et al. (2005). Clinical trial of a plant-derived antibody on recolonization of mutans streptococci. *Caries Res.* 39, 241–250. doi: 10.1159/000084805

Wen, Z. T., Liao, S., Bitoun, J. P., De, A., Jorgensen, A., Feng, S., et al. (2017). *Streptococcus mutans* Displays Altered Stress Responses While Enhancing Biofilm Formation by Lactobacillus casei in Mixed-Species Consortium. *Front. Cell. Infect. Microbiol.* 7:524. doi: 10.3389/fcimb.2017.00524

Wen, Z. T., Yates, D., Ahn, S. J., and Burne, R. A. (2010). Biofilm formation and virulence expression by *Streptococcus mutans* are altered when grown in dual-species model. *BMC Microbiol.* 10:111. doi: 10.1186/1471-2180-10-111

Winograd-Katz, S. E., Fässler, R., Geiger, B., and Legate, K. R. (2014). The integrin adhesome: from genes and proteins to human disease. *Nat. Rev. Mole. Cell* Biol. 15, 273–288. doi: 10.1038/nrm3769

Yang, C., Scoffield, J., Wu, R., Deivanayagam, C., Zou, J., and Wu, H. (2018). Antigen I/II mediates interactions between *Streptococcus mutans* and Candida albicans. *Mole. Microbiol.* 33, 283–291. doi: 10.1111/omi.12223

Yu, H., Nakano, Y., Yamashita, Y., Oho, T., and Koga, T. (1997). Effects of antibodies against cell surface protein antigen PAc-glucosyltransferase fusion proteins on glucan synthesis and cell adhesion of *Streptococcus mutans*. *Infect. Immun.* 65, 2292–2298.

Zhang, P., Jespersgaard, C., Lamberty-Mallory, L., Katz, J., Huang, Y., Hajishengallis, G., et al. (2002). Enhanced immunogenicity of a genetic chimeric protein consisting of two virulence antigens of *Streptococcus mutans* and protection against infection. *Infect. Immun.* 70, 6779–6787. doi: 10.1128/iai.70.12.6779-6787.2002

Zhang, S., Green, N. M., Sitkiewicz, I., Lefebvre, R. B., and Musser, J. M. (2006). Identification and characterization of an antigen I/II family protein produced by group A *Streptococcus*. *Infect. Immun.* 74, 4200–4213. doi: 10.1128/IAI.00493-06

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Microbiology

REVIEW
published: 26 November 2020
doi: 10.3389/fmicb.2020.599899

Polymer Adhesin Domains in Gram-Positive Cell Surface Proteins

*Michael A. Järvå[1], Helmut Hirt[2], Gary M. Dunny[2] and Ronnie P.-A. Berntsson[1,3]**

[1] Department of Medical Biochemistry and Biophysics, Umeå University, Umeå, Sweden, [2] Department of Microbiology and Immunology, University of Minnesota, Minneapolis, MN, United States, [3] Wallenberg Centre for Molecular Medicine, Umeå University, Umeå, Sweden

Surface proteins in Gram-positive bacteria are often involved in biofilm formation, host-cell interactions, and surface attachment. Here we review a protein module found in surface proteins that are often encoded on various mobile genetic elements like conjugative plasmids. This module binds to different types of polymers like DNA, lipoteichoic acid and glucans, and is here termed *polymer adhesin domain.* We analyze all proteins that contain a polymer adhesin domain and classify the proteins into distinct classes based on phylogenetic and protein domain analysis. Protein function and ligand binding show class specificity, information that will be useful in determining the function of the large number of so far uncharacterized proteins containing a polymer adhesin domain.

Keywords: adhesin, Gram-positive, conjugation, biofilm, binding

OPEN ACCESS

Edited by:
Magnus Hook,
Texas A&M Health Science Center,
United States

Reviewed by:
Peter Davies,
Queen's University, Canada
Catherine Ayn Brissette,
University of North Dakota,
United States

***Correspondence:**
Ronnie P.-A. Berntsson
ronnie.berntsson@umu.se

Specialty section:
This article was submitted to Infectious Diseases, a section of the journal Frontiers in Microbiology

Received: *28 August 2020*
Accepted: *28 October 2020*
Published: *26 November 2020*

Citation:
Järvå MA, Hirt H, Dunny GM and Berntsson RP-A (2020) Polymer Adhesin Domains in Gram-Positive Cell Surface Proteins. Front. Microbiol. 11:599899. doi: 10.3389/fmicb.2020.599899

BACTERIAL ADHESION IN GRAM-POSITIVE BACTERIA

Bacteria colonize host tissues by adhering to specific surfaces and by establishing bacterial biofilm communities. In Gram-positive bacteria this is often mediated by surface proteins that are anchored to the cell-wall through a sortase-dependent LPxTG-motif (Geoghegan and Foster, 2015; Foster, 2019). Pili and fimbriae are well known examples that form micrometer long filaments that protrude out from the cell allowing easy attachment to targets (Kang and Baker, 2012; Lukaszczyk et al., 2019). Other classes of adhesion proteins exist such as microbial surface components recognizing adhesive matrix molecules (MSCRAMMs), near iron transporter (NEAT) motif family, tandemly repeated three-helical bundles, G5-E domain repeat family, and legume-lectin-like cadherin-like family. For these classes there are multiple thorough reviews and book chapters detailing their structural and functional properties (Foster et al., 2014; Geoghegan and Foster, 2015; Foster, 2019; Dufrêne and Viljoen, 2020). In general, these proteins are comprised of a C-terminal stalk region built from heavily glycosylated disordered regions, coiled-coils, or tandem domain repeats. The N-terminus often feature one or several adhesion modules that are key to the functional part of the protein, and the stalk projects this region away from the peptidoglycan cell-wall, allowing access to the extracellular environment. They interact with key components in the extracellular matrix of their host to facilitate one or several types of pathogenic mechanisms such as surface attachment, host cell internalization, biofilm formation, immune evasion, and/or bactericide/antibiotic resistance.

However, a family of adhesion proteins found in lactic acid bacteria (*Lactobacillales*) cannot readily be categorized into any of the above-mentioned families. In *Streptococcus* these proteins go by the names Glucan-binding protein C, Dextran-binding lectin, and Antigen I/II, and in *Lactococcus* and *Enterococcus* they are called Aggregation substance. In *Streptococcus* these

adhesion proteins promote cariogenicity through attachment to tooth surfaces (Jenkinson and Demuth, 1997; Sato et al., 2002a; Lynch et al., 2013). In *Enterococcus* and *Lactococcus* they are found on conjugative plasmids where they facilitate mating pair formation (Hirt et al., 2000; Waters and Dunny, 2001; Waters et al., 2003; Chuang et al., 2009; Schmitt et al., 2018). Apart from surface adherence and bacterial aggregation these proteins all feature a specific adhesion domain at their N-terminus, which most often is called Glucan-binding domain, Variable domain or Adhesion domain. Despite many similarities between these proteins no studies have yet systematically compared their functions and mechanisms.

As we will discuss in this review, the conserved adhesion module consistently adheres to various types of polymers associated with the extracellular matrix, such as collagen (Love et al., 1997; Holmes et al., 1998; Heddle et al., 2003), extracellular DNA (eDNA; Kohler et al., 2018; Schmitt et al., 2018), lipoteichoic acid (LTA; Waters et al., 2004; Schmitt et al., 2018), or different types of glucans (Sato et al., 2002b; Tamura et al., 2014; Mieher et al., 2018). Therefore, we will throughout this review refer to this adhesion module as the *polymer adhesin domain*.

DOMAIN ARCHITECTURE AND FUNCTION OF PROTEINS WITH A POLYMER ADHESIN DOMAIN

In order to gather a diverse population of polymer adhesin-containing proteins, we utilized the InterPro database (Mitchell et al., 2019). InterPro combines information from numerous other databases which uses various models such as hidden Markov models, scoring matrices, regular expressions, or other profiles that make up identifiable signatures to classify protein families. We gathered all entries that contained an N-proximal polymer adhesin domain (IPR013574) and a C-terminal LPxTG-motif (IPR019931). After removing outdated, deprecated, and fragmented entries the resulting list of 518 proteins was manually cross-referenced for associated literature, which gave us a list of proteins that had, to some extent, been functionally characterized (**Table 1**). One limitation with this approach is that the InterPro database relies on entries in UNIPROT. Thus it does not contain all known sequences in other databases, such as NCBI. In fact, we encountered two proteins in the literature that was not originally captured by the InterPro search. However, the advantage of using it is that all entries are curated. The final list thus corresponds to 21 proteins (**Table 1**).

As we aimed to focus our review on the polymer adhesin domain we calculated a phylogenetic tree using only the polymer adhesin domain sequences rather than full-length protein. This was done to focus on the adhesive relationships of the polymer adhesins without the influence of large sequence and domain variety between the stalk regions. In this analysis we included all proteins with <90% sequence identity (and as low as 15%) of their polymer adhesin domain, plus all 21 reviewed proteins in **Table 1**, resulting in a phylogenetic tree with 131 proteins (**Figure 1A**). This phylogenetic analysis of only the polymer adhesin domain corresponded well with the predicted domain architecture of the 21 reviewed proteins (**Figure 1B**) and allowed us to divide them into five distinct classes (Classes I–V; **Figure 1**). The full phylogenetic analysis also indicates the presence of two additional classes (Classes VI and VII), containing proteins that so far lack any functional data in the literature (yellow and gray clades in **Figure 1A**).

Apart from the polymer adhesin domain and the LPxTG-motif, two other structural domains are found throughout the classes (except Class I proteins); (**i**) The C-terminal cell surface antigen domains (dark blue domain in **Figure 1B**), which are tandemly arranged bacterial immunoglobulins that often feature intramolecular isopeptide bonds and calcium-binding sites (Forsgren et al., 2010; Heim et al., 2014) and (**ii**) The N-terminal scaffold domain (orange domain in **Figure 1B**). It has been proposed that its function is to attach to its own C-terminal region to stabilize the structure and to make the polymer adhesin domain the most matrix exposed feature (Brady et al., 2010; Larson et al., 2010).

All known proteins belonging to these classes contain only one copy of the polymer adhesin domain. The polymer adhesin is often displayed quite a distance out from the cell wall, as there is usually either a coiled-coil stalk and/or one or more immunoglobulin domains between the polymer adhesin domain and the LPxTG cell-wall anchor. Based on the literature, we conclude that the polymer adhesin module likely exerts a core function in most of these proteins. The surrounding immunoglobulin domains likely act as helper modules to provide additional functionality or to display the polymer adhesin domain sufficiently far out from the cell surface. We wanted to investigate whether our classification of the polymer adhesin modules correlates with their variation in function, e.g., which ligands they bind and which pathogenic mechanisms they promote. To address this question, we went through the literature and looked at the available data for proteins associated with Classes I–V (as mentioned previously, Class VI and VII have no associated literature).

STRUCTURE AND FUNCTION OF POLYMER ADHESION CONTAINING PROTEINS

To date, there are six unique polymer adhesin domain structures deposited in the PDB, five of which are described in literature (**Table 1**). These structures originate from Classes I, II, and IV. Despite low sequence similarity (21–34%), the overall fold of the polymer adhesin domain remains remarkedly similar (**Figure 2** and **Table 2**). In all structures, the core fold comprises an antiparallel beta-sandwich of 12–16 strands. On one side of this core, two lobes (made up of highly variable loops and short alpha-helices) create a central ridge. This ridge harbors a cation binding site that is conserved throughout the domain family (**Figure 2**). Due to the differences in primary sequence, the surface charge distribution of the domains also varies, with surfaces ranging from mostly negatively charged to mostly positively charged (**Figure 3**). None of these proteins have so far had their full-length structure determined, but the current

TABLE 1 | Proteins containing a polymer adhesin domain currently described in the literature and their associated meta data and references.

Class	Protein name	Gene(s)	PDB(s)	Accession code	Protein length	Organism	References
I	Glucan-binding protein C (GbpC)	*gbpC*	5UQZ/6CAM	Q8DTF1	583	*Streptococcus mutans*	Sato et al., 1997, 2002a,b; Mieher et al., 2018
I	Glucan-binding protein C (GbpC)	*gbpC*		Q4W7G2	617	*Streptococcus macacae*	Okamoto-Shibayama et al., 2006
I	Dextran-binding lectin A (DblA)	*dblA*		G5EIN8	1093	*Streptococcus criceti*	Tamura et al., 2014
I	Dextran-binding lectin B (DblB)	*dblB*		G5EIN9	1717	*Streptococcus criceti*	Tamura et al., 2014
I	Dextran-binding lectin B (DblB)	*dblB*		A8QYL3/B5BNX9	1425	*Streptococcus sobrinus*	Sato et al., 2009
II	Streptococcal surface protein A (SspA)	*sspA*		Q54185	1575	*Streptococcus gordonii*	Demuth et al., 1996; Holmes et al., 1998; Egland et al., 2001; Heddle et al., 2003; Jakubovics et al., 2005a,b; Giomarelli et al., 2006; Nobbs et al., 2007
II	Streptococcal surface protein B (SspB)	*ssp5/sspB*	2WD6	P16952/Q54186	1500	*Streptococcus gordonii*	Demuth et al., 1988; Duan et al., 1994; Demuth et al., 1996; Holmes et al., 1998; Heddle et al., 2003; Giomarelli et al., 2006; Nobbs et al., 2007; Forsgren et al., 2009; Forsgren et al., 2010
II	Cell-surface protein antigen (SpaP,PA/PAc/P1)	*spaP*		P23504	1562	*Streptococcus mutans serotype c*	Koga et al., 1986; Kelly et al., 1990; Hajishengallis et al., 1994; Heim et al., 2014; Jakubovics et al., 2005b
II	Cell-surface protein antigen (SpaP,PA/PAc/P1)	*spaP, pa, pac*	3IPK/3IOX/1JMM	C9E3B4/A8R5D9/ P11657	1566	*Streptococcus mutans*	Sommer et al., 1987; Oho et al., 1998; Troffer-Charlier et al., 2002; Sato et al., 2002a; Nakano et al., 2006; Busscher et al., 2007; Larson et al., 2010; Heim et al., 2014; Sullan et al., 2015
II	Cell-surface antigen I/II (SpaA)	*spaA*		P21979	1528	*Streptococcus sobrinus*	Kuykindoll and Holt, 1996
II	Cell-surface antigen I/II	*pas*	6E36	Q9KW51	1310	*Streptococcus intermedius*	Petersen et al., 2001; Jakubovics et al., 2005b
III	SAG_1283			A8D815	1631	*Streptococcus dysgalactiae*	Davies et al., 2009
III	Agglutinin receptor I/II			KGI30072.1	1646	*Streptococcus pneumoniae*	Antic et al., 2017
III	Glucan-binding protein C (GbpC)			OYL08640.1	1634	*Streptococcus pneumoniae B1599*	Antic et al., 2017
IV	Aggregation substance (AS)	*prgB, asa1, asp1, asc10*	6EVU/6GED	Q04112	1305	*Enterococcus faecalis (plasmid: pCF10)*	Kreft et al., 1992; Bensing and Dunny, 1993; Rakita et al., 1999; Vanek, 1999; Hirt et al., 2000; Wells et al., 2000; Waters and Dunny, 2001; Isenmann et al., 2002; Waters et al., 2003, 2004; Chuang et al., 2009; Chuang-Smith et al., 2010; Bhatty et al., 2015; Kohler et al., 2018; Schmitt et al., 2018
IV	Aggregation substance (AS)	*prgB, asa1, asp1, asc10*		P17953	1296	*Enterococcus faecalis (plasmid: pAD1/pTEF1)*	Galli et al., 1990, 1992; Chow et al., 1993; Süßmuth et al., 2000; Rozdzinski et al., 2001
V	Sex factor aggregation protein	*cluA*		Q48588	1243	*Lactococcus lactis*	Godon et al., 1994; Gasson et al., 1995; Stentz et al., 2004, 2006; Kojic et al., 2011; Lukić et al., 2012

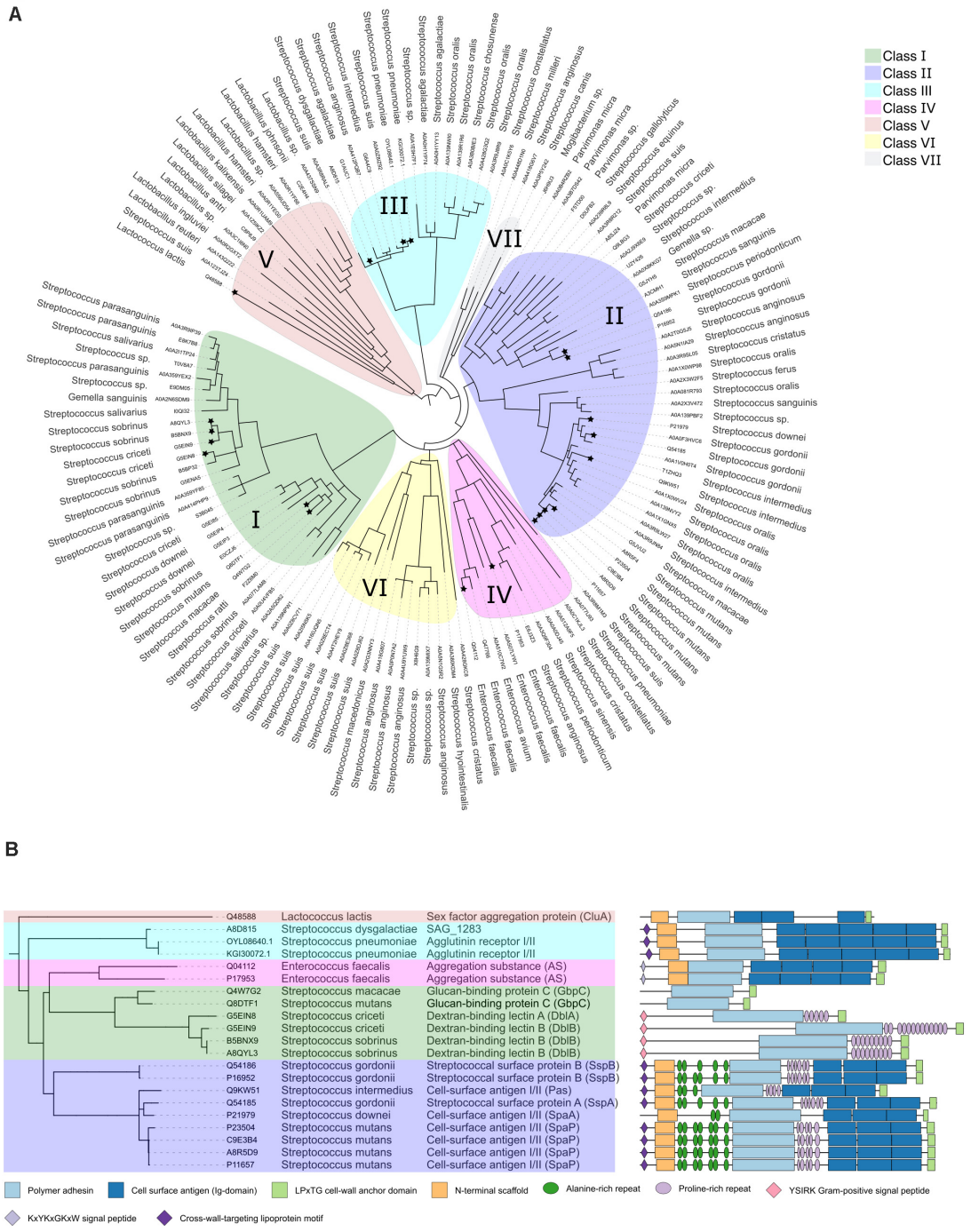
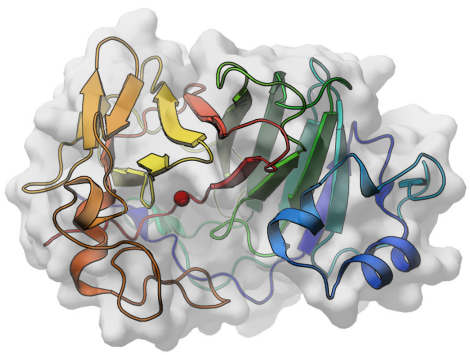
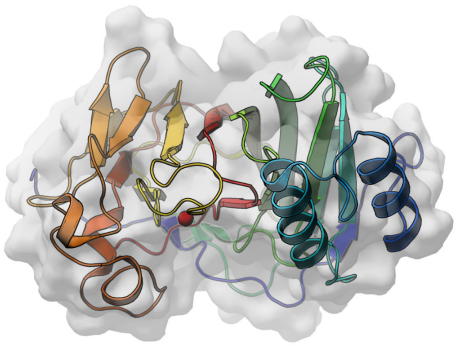


FIGURE 1 | (A) Phylogenetic tree of a representative subset (<90% identity) of the polymer adhesin domain sequences as annotated in InterPro (Mitchell et al., 2019). Alignments done by ClustalO were used to calculate the phylogenetic relationships with PhyML (Dereeper et al., 2008) using 100 bootstraps. Visualization was done with iTol and brances with a bootstrap lower than 0.5 were collapsed. Each node is annotated with its respective UNIPROT accession code (in two instances refseq) and organism name. Clades are color coded, and nodes with associated literature are marked with a star. **(B)** Pruned tree highlighting the 21 reviewed protein entries, using the same class color coding as in panel **(A)**. Here the additional annotations include common protein name, and domain architecture as annotated by InterPro (Mitchell et al., 2019).

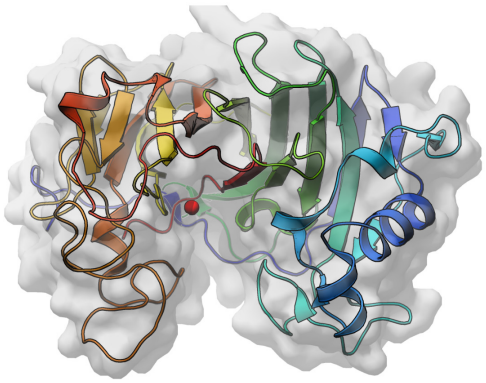
structural evidence points toward that the N-terminal sequence preceding the polymer adhesin domain forms a coiled-coil with the sequence just C-terminal of said domain (Brady et al., 2010; Larson et al., 2010). This means that the tip of the proteins, thus the part extending furthest out from the cell wall, is the polymer adhesin domain.



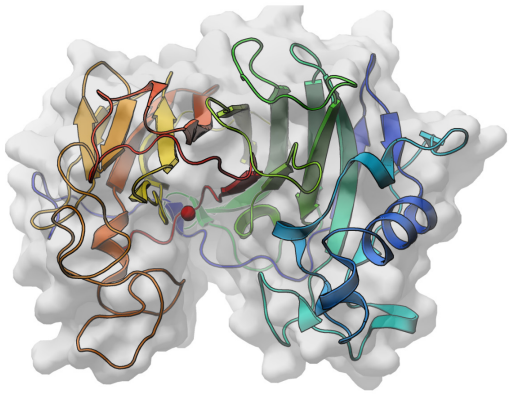
Class I - Glucan-binding protein C (GbpC) *S. mutans*. PDB code: 5UQZ



Class II - Streptococcal surface protein B (SspB) *S. gordonni*. PDB code: 2WD6



Class II - Cell-surface antigen (SpaP) *S. mutans*. PDB code: 1JMM



Class IV - Aggregation substance (PrgB) *E. feacalis*. PDB code: 6GED

FIGURE 2 | All available protein structures of unique proteins from three different classes of polymer adhesin containing proteins. Here drawn as cartoon representations, colored blue to red from the N-terminus and viewed from the same angle, the similarities in overall fold is seen, as well as their conserved cation binding site (red sphere) situated in the middle of the central ridge.

TABLE 2 | Sequence and r.m.s. deviations between the four structurally characterised polymer adhesin domains.

RMSD / % ID	PrgB 6EVU	GbpC 5UQZ	SspB 2WD6	SpaP 1JMM
PrgB 6EVU		2.7Å	2.9Å	2.6Å
GbpC 5UQZ	25%		2.9Å	2.4Å
SspB 2WD6	26%	26%		2.9Å
SpaP 1JMM	21%	28%	34%	

Class I – Glucan-Binding Protein C and Dextran-Binding Lectin

This Class (green clade in **Figure 1A**) is divided into two subgroups: glucan-binding protein C (GbpC) and Dextran-binding lectin (Dbl) proteins. Counter-intuitively, both types bind dextran (branched, primarily α-1,6-glucans) whereas the Dextran-binding lectin proteins also bind to amylose (α-1,4-glucans) and to the α-1,3-branches on dextran (Okamoto-Shibayama et al., 2006). In contrast to the other classes, they do not contain any immunoglobulin domains at their C-terminus. Compared to GbpC, Dbl proteins have longer flanking coils and a long sequence insertion in the middle of the polymer adhesin domain (**Figure 1**). Dbls also feature an N-terminal YSIRK-motif (pink in **Figure 1B**). The YSIRK motif is unique to Streptococci and Staphylococci and enhances the efficiency of protein secretion (Bae and Schneewind, 2003) and enforce spatial positioning to the septal wall (Bierne and Dramsi, 2012).

The affinity of GbpC-polymer adhesins for various lengths of dextrans has been measured by isothermal calorimetry (ITC; Mieher et al., 2018). This enthalpy driven binding is stronger for longer polymers (highest reported affinity was ∼17 μM for a dextran with ∼390 glucose units). Based on these binding experiments, it was estimated that each GbpC-polymer adhesin unit adheres to 11–14 glucose units in a non-cooperative manner. Of the two determined GbpC structures, one is in a apo-state (PDB code: 5UQZ) whereas in the other (PDB code: 6CAM) two glucose molecules are modeled near the cation binding site (Mieher et al., 2018). Removal of a loop region overarching the cation binding site lead to a reduction, but

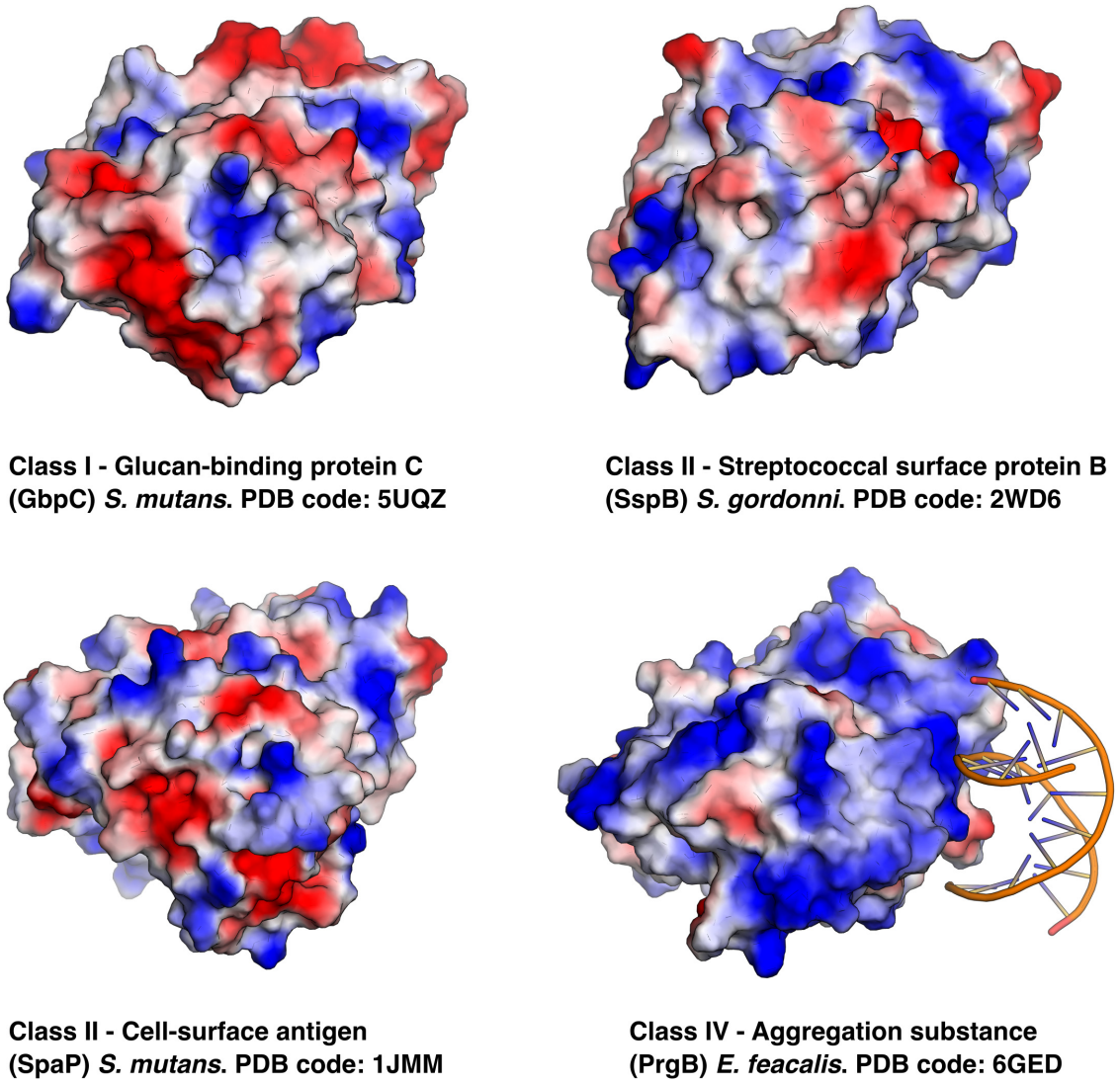


FIGURE 3 | Same proteins as in **Figure 2** but using electrostatic surface representations and in the case of PrgB its eDNA ligand is shown binding to the positively charged surface.

not elimination, in glucose mediated biofilm formation (Mieher et al., 2018). As no other confirmatory experiments have been reported, the validity of this site being a glucan-binding site remains uncertain.

It has been shown that GbpC can interact with the host receptor salivary agglutinin (SAG) with nanomolar affinity and independently of SAGs glycosylation state (Purushotham and Deivanayagam, 2014; Mieher et al., 2018). SAG is also known as gp340 and is expressed from the "*deleted in membrane protein 1*" gene (DMBT1). It is a large extracellular matrix protein which features 13 repeats of the scavenger receptor cysteine rich domain 1 (iSRCR; Reichhardt et al., 2020). GbpC has been shown to interact with these iSRCR domains in a calcium-dependent fashion (Purushotham and Deivanayagam, 2014; Mieher et al., 2018), but since the fold of iSRCR is in itself calcium- dependent (Reichhardt et al., 2020) it is unknown whether or not the calcium-binding site of the polymer adhesin domain is involved in this interaction. Both the interaction to the entire SAG and the individual iSRCR domains are inhibited by the addition of dextran, which could indicate that they compete for the same binding site.

Class II – Cell Surface Antigen I/II

Class II polymer adhesins (purple clade in **Figure 1A**) have been extensively studied for their prominent role in facilitating dental caries. The proteins that are found in this Class are usually named Agglutinin receptor I/II, as they were originally characterized by their ability to bind SAG (Demuth et al., 1988; Demuth et al., 1996). More recently Agglutinin receptor I/II are more commonly referred to as Cell-surface antigen I/II. The corresponding Class II proteins have been shown to interact with collagens (Love et al., 1997; Holmes et al., 1998; Heddle et al.,

2003; Sciotti et al., 2006), fibrinogen (Brady et al., 2010), and laminin (Sciotti et al., 2006) – proteins characterized by long triple coiled-coils. These proteins have also been confirmed to bind to fibronectin (Giomarelli et al., 2006). All these proteins are common extracellular protein components that are often utilized by pathogens as an initial host interaction point (Vercellotti et al., 1985; Schwarz-Linek et al., 2006; Kang et al., 2013). It is mainly *S. mutans*, *S. oralis*, *S. sanguinis,* and *S. gordonii* that contain Class II proteins. *S. mutans* especially is a cause of dental caries (Hamada and Slade, 1980), whereas the other three species can act as opportunistic pathogens.

For three polymer adhesins in this Class the structure is known: SspB (*S. gordonii*; Forsgren et al., 2009), SpaP (*S. mutans*; Troffer-Charlier et al., 2002; Larson et al., 2010), and Pas (*S. intermedius*; **Table 1**). Pas has not been functionally characterized but is 85% identical to SpaP. As these proteins are similar, and it was known that sialic acid can inhibit SpaP binding to SAG (Demuth et al., 1990), it was thought that these proteins bind glucans. However, neither the SspB nor the SpaP polymer adhesin domains bind dextran (Mieher et al., 2018) and even though SspB was extensively tested on glycan arrays, no binding to any glycoconjugates has been observed (Forsgren et al., 2009). Despite the lack of direct interaction with glycans, the polymer adhesin in SpaP is important for both biofilm formation and dextran induced cellular aggregation (Mieher et al., 2018).

Furthermore, the Class II proteins SspB and SpaP can interact with the scavenger receptor cysteine-rich domain 1 (iSRCR) in a calcium-dependent fashion (Purushotham and Deivanayagam, 2014; Mieher et al., 2018) just as the Class I protein GbpC. In contrast to GbpC, however, this interaction is not inhibited by the addition of dextran. SspB and SpaP also have an additional independent iSRCR-interaction site (Larson et al., 2011). This site is located on the first two immunoglobulin domains, which the domain Class I proteins do not have. The second and third of these immunoglobulin domains on both SspB and SpaP bind calcium with submicromolar affinity (Duan et al., 1994; Forsgren et al., 2010; Larson et al., 2011; Nylander et al., 2011).

Class III

Class III adhesins (cyan clade in **Figure 1A**) are most closely related to Class II adhesins in domain architecture (**Figure 1B**). Although not identified as such in Interpro, they do seem to feature similar alanine-rich and proline-rich repeats prior to, and after, the polymer adhesin domain as well. They also contain five Ig-domains rather than the three seen in Class II adhesins. Class III adhesins are found in *Streptococcus* species frequent in the upper respiratory tract of pigs (*S. suis*) and humans (*S. pneumoniae*), and in *S. agalactiae*, which can colonize the intestinal and vaginal microbiota (Barcaite et al., 2008). Two proteins have been shown to increase *S. pneumoniae* adhesion to ocular epithelia (Antic et al., 2017), but that remains the full extent of the known functions of this Class of adhesins. It is interesting to note, however, that bioinformatics analysis of Class III proteins have revealed that they are the result of horizontal gene transfer and that they are found on multiple mobile genetic elements (Davies et al., 2009).

Class IV – Aggregation Substance

Class IV proteins (pink clade in **Figure 1A**) are mostly found in *Enterococcus faecalis* where they are found in sex pheromone responsive conjugative plasmids. Here they facilitate horizontal gene transfer via Type 4 Secretion Systems. Although they are mostly connected to Enterococci, they can spread to other species via conjugation. One protein in this Class has been structurally studied, namely PrgB (Aggregation substance) from the conjugative plasmid pCF10 (Schmitt et al., 2018). PrgB is one of the surface proteins expressed from the *prgQ* operon that encodes for all genes that are needed for conjugation. PrgB aids conjugation via surface attachment, biofilm formation, and cellular mating pair formation (Dunny et al., 1978; Bensing and Dunny, 1993; Bhatty et al., 2015; Schmitt et al., 2018). PrgB-like proteins are encoded by many other T4SS bearing plasmids in Enterococci, such as pAD1 (Galli et al., 1990, 1992; Chow et al., 1993; Süßmuth et al., 2000; Rozdzinski et al., 2001) and pD1 (Schmitt et al., 2020).

Expression of PrgB leads to cellular clumping (Dunny et al., 1978; Bhatty et al., 2015), which is dependent on the polymer adhesin domain binding to eDNA (Kohler et al., 2018; Schmitt et al., 2018). Surprisingly, this interaction does not take place at the conserved ridge with the cation binding site (**Figure 2**). Instead, eDNA binds in a sequence unspecific manner via charge interactions with surface exposed lysines and arginines (Schmitt et al., 2018; **Figure 3**). The same site also binds the core component of the Gram-positive cell-wall, LTA. As LTA is mainly composed of repeating units of ribitol or glycerol phosphate, it has a similar charge distribution to DNA. Compared to other polymer adhesin domain structures, the surface of the domain in PrgB is positively charged (**Figure 3**), enabling it to bind the negatively charged eDNA and LTA. The ability of PrgB to induce cellular aggregation and facilitate biofilm formation is completely dependent on the polymer adhesin domain, as deletion of this domain completely abrogates these functions (Bhatty et al., 2015; Schmitt et al., 2018). In contrast, when all domains between the polymer adhesin domain and the C-terminal LPxTG motif are removed cell aggregation and binding to LTA are still observed (Waters et al., 2004). Like Class II adhesins, PrgB promote adherence to fibronectin (Rozdzinski et al., 2001; Isenmann et al., 2002) and fibrinogen through its polymer adhesin domain (Chuang et al., 2009). Using in-frame deletions it has also been shown that the polymer adhesin domain is responsible for adhering to macrophages (Süßmuth et al., 2000). PrgB and its homologs are strong virulence factors in various infection models, including *C. elegans* (Bhatty et al., 2015) and rabbit endocarditis (Chow et al., 1993; Schlievert et al., 1998), where they play an important role in both vegetation formation and pathogenicity.

Class V – CluA-Like Aggregation Substance

Only one protein in Class V (salmon clade in **Figure 1A**) has been functionally characterized, namely CluA from *Lactococcus lactis*. CluA is functionally homologous to the Class IV protein PrgB, performing highly similar functions related to cellular

aggregation and conjugation (Godon et al., 1994; Stentz et al., 2004). The polymer adhesin domain of CluA has not been explicitly studied, but due to the overall similarity in effect of CluA compared to PrgB, it is likely that the polymer adhesin of CluA works in a similar manner by adhering to cell-wall components and host surfaces.

Class VI and VII

We completely lack studies on any protein from Class VI or VII (yellow and gray clades, respectively, in **Figure 1A**). Domain analysis of proteins from these two classes reveal that they contain mucin (or mucin-like) binding protein domains, indicating that these proteins could be involved in binding to molecules associated with the extracellular matrix.

PROTEIN FUNCTIONALITIES SPAN THE CLASSES

Polymer-Binding Induced Aggregation

The most striking common denominator between the different classes of polymer adhesins is their propensity to bind specific types of polymers; Class I directly interact with positively charged oligosaccharides such as dextran (GbpC) and glucans (Dbl); Class II interact with an array of host proteins with coiled-coil features such as collagen, fibrinogen, and laminin; Class IV bind negatively charged polymers such as eDNA and LTA.

The specific binding site of these polymers has remained largely unknown until the eDNA binding site was elucidated for PrgB recently (Schmitt et al., 2018). Somewhat surprisingly, the PrgB polymer adhesin domain binds DNA on the opposite side of the ridge with the cation binding site (**Figure 3**), which was the proposed glycan binding site. For GbpC, the current literature suggests that the GbpC polymer adhesin domain binds dextran in the cleft containing the cation binding site (Mieher et al., 2018). However, based on the available data of the polymer adhesin domains, we hypothesize a different mechanism for polymer binding by these proteins. In this model, the polymer adhesin domains bind their target polymers via their surface, like the eDNA and LTA binding in PrgB. The driving factor in this adhesion is the avidity effects that naturally occur when polymers accumulate, which has been demonstrated for both GbpC and PrgB (Schmitt et al., 2018).

Post-translational Protease Processing

Several of the polymer adhesin domain containing proteins are known to be post-translationally processed. PrgB from Class IV, is post-translationally cleaved in the unstructured region between

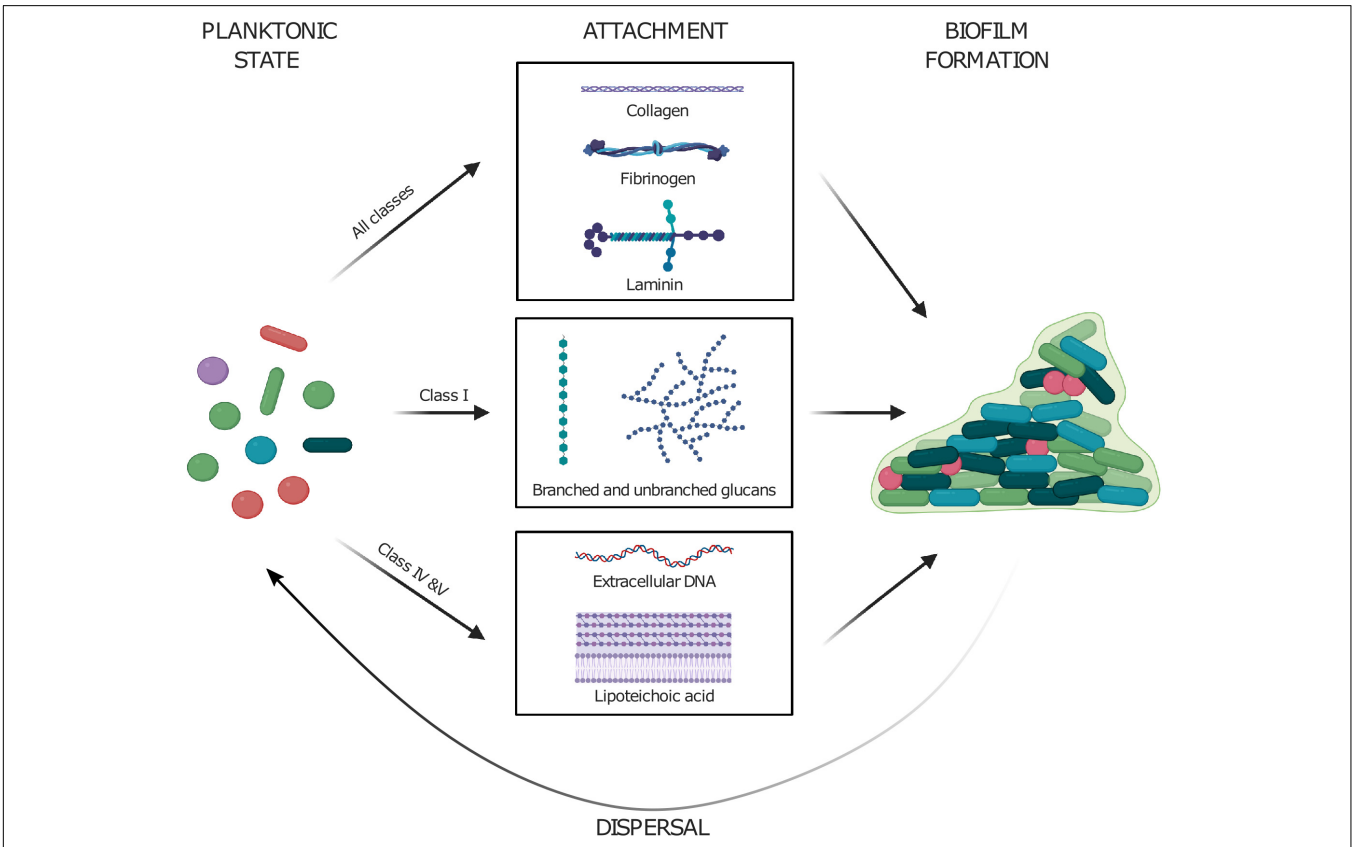


FIGURE 4 | Planktonic cells utilize their polymer adhesin domain to attach to various surfaces by binding various molecules such as (i) coiled-coil proteins in extracellular matrix, (ii) polysaccharides, and (iii) negatively charged polymers such as eDNA and lipoteichoic acid. The attachment to these further drives the formation of cellular biofilms. Enzymatic cleavage of the polymer adhesin domain might aid in the dispersal of cells from mature biofilms. Created with BioRender.

the polymer adhesin-domain and the following immunoglobulin module (Nakayama et al., 1992; Schmitt et al., 2020). This has been linked to another surface protein, PrgA, which is expressed from the same operon (Hirt et al., 2000; Bhatty et al., 2015; Schmitt et al., 2020). The polymer adhesion domain of SpaP from Class II is also known to be enzymatically released from the cell surface (Russell et al., 1980; Sommer et al., 1987), thereby changing the cell's hydrophobicity properties and facilitating biofilm release (Lee, 1992; Lee et al., 1996). Whether this enzymatic release is specific to these two proteins or something that is more common throughout the classes is not yet known. It does point to these proteins being actively degraded to facilitate involved biofilm dispersion, a poorly understood key phase of the biofilm life cycle (Rumbaugh and Sauer, 2020). An analogous feature can be observed in the well-studied Gram-negative RTX-adhesins, which are proteolytically released in response to nutrient restriction (Guo et al., 2019).

Host Receptor Interactions

As these polymer adhesin-containing proteins are cell-wall anchored and surface exposed, they are often used in the interaction with host receptors. As mentioned previously, the polymer adhesin domains from GbpC (Class I), SspB and SpaP (Class II) have strong affinity for SAG and its subdomain iSRCR in a calcium-dependent fashion (Purushotham and Deivanayagam, 2014; Mieher et al., 2018). Whether or not other classes also bind SAG is unknown as this has not been tested. Class II proteins interact with collagen type I, fibrinogen, laminin, and fibronectin (Love et al., 1997; Holmes et al., 1998; Heddle et al., 2003; Giomarelli et al., 2006; Sciotti et al., 2006; Brady et al., 2010). Similarly, *in vivo* studies suggest that PrgB promote adhesion to fibrinogen, fibronectin, thrombospondin, vitronectin, and collagen type I (Rozdzinski et al., 2001; Isenmann et al., 2002). These extracellular matrix proteins are common targets for bacterial adhesion mechanisms (fibronectin in particular) and are thought to be a key component in establishing bacterial infections. However, the exact nature of the interactions of polymer adhesin domain containing proteins and various host receptors can vary widely and exact binding sites have not been established.

A Concluding Model

Each class of polymer adhesin containing protein has developed its own niche and affinity toward different types of ligands, but the overall purpose of its existence appears to be largely conserved. We propose an overall model for polymer adhesin-contribution to life cycle of cellular biofilms in **Figure 4**, where they are involved in transitioning between the phases. In the transition from planktonic cells to attachment, different types of extracellular matrix molecules are recognized. GbpC and Dbl recognize dextran, amylose, and SAG. Antigen I/II recognize SAG and common subendothelial matrix proteins such as fibrinogen. Aggregation substance recognize negatively charged polymers (eDNA and LTA), as well as fibrinogen and similar matrix proteins. Because the polymer adhesion is avidity driven, initial attachment lead to further aggregation as the concentration of polymers and cells continue to increase and promote biofilm formation. They also actively recruit polysaccharides and eDNA which are two major components of bacterial biofilm. Finally, we hypothesize that proteolytic cleavage of the polymer adhesin domain might aid in cell dispersal events in mature biofilms.

OUTLOOK

Albeit more than 518 proteins contain a polymer adhesin domain, we conclude that this family of bacterial adhesins is relatively poorly characterized, since only 21 polymer adhesion domain containing proteins from 5 of the 7 classes have been studied functionally (**Table 1**). Interestingly, we found that these 21 studied proteins naturally fall into the same five separate classes whether they are clustered by polymer adhesin domain sequence identity or by the domain organization of the full-length protein (**Figure 1B**). Furthermore, the type of polymer that the proteins bind might correlate to the different classes, with Class I binding (positively charged) glucans, Class II mainly interacting with coiled-coil proteins and Class IV binding (negatively charged) eDNA and LTA. However, the ligand preferences of the polymer adhesin domain of the other classes have not been conclusively determined to date, so we cannot conclude that each class binds its own kind of polymer.

Class IV PrgB is one of the better studied proteins with a cell-wall anchor and a polymer adhesion domain. However, there are many more proteins with these same two properties that are also encoded from genes on conjugative plasmids in bacilli, plasmids that most often also encode for antibiotic resistance. Like their characterized counterparts, these proteins are very likely to be involved in biofilm formation and adhesion to specific host receptors, and therefore probably also strong virulence factors. Studying these proteins will be useful to further understand virulence and the mechanism of adhesion processes in Gram-positive bacteria.

AUTHOR CONTRIBUTIONS

MJ: conceptualization, performed analysis, and writing the manuscript. HH: conceptualization and validation. GD: conceptualization and aided in writing. RB: conceptualization, analysis, and writing. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by grants from the Swedish Research Council (2016–03599), Knut and Alice Wallenberg Foundation and Kempestiftelserna (SMK-1869) to RB, and NIH grant R35 GM11807 to GD.

ACKNOWLEDGMENTS

The authors thank Dr. Josy ter Beek for critical reading of the manuscript.

REFERENCES

Antic, I., Brothers, K. M., Stolzer, M., Lai, H., Powell, E., Eutsey, R., et al. (2017). Gene acquisition by a distinct phyletic group within streptococcus pneumoniae promotes adhesion to the ocular epithelium. *mSphere* 2, e213–e217.

Bae, T., and Schneewind, O. (2003). The YSIRK-G/S Motif of *Staphylococcal* protein A and its role in efficiency of signal peptide processing. *J. Bacteriol.* 185, 2910–2919. doi: 10.1128/jb.185.9.2910-2919.2003

Barcaite, E., Bartusevicius, A., Tameliene, R., Kliucinskas, M., Maleckiene, L., and Nadisauskiene, R. (2008). Prevalence of maternal group B *Streptococcal* colonisation in european countries. *Acta Obstet. Gynecol. Scand.* 87, 260–271. doi: 10.1080/00016340801908759

Bensing, B. A., and Dunny, G. M. (1993). Cloning and molecular analysis of genes affecting expression of binding substance, the recipient-encoded receptor(s) mediating mating aggregate formation in *Enterococcus faecalis*. *J. Bacteriol.* 175, 7421–7429. doi: 10.1128/jb.175.22.7421-7429.1993

Bhatty, M., Cruz, M. R., Frank, K. L., Gomez, J. A. L., Andrade, F., Garsin, D. A., et al. (2015). *Enterococcus faecalis* pCF10-encoded surface proteins PrgA. PrgB (aggregation substance) and PrgC contribute to plasmid transfer, biofilm formation and virulence. *Mol. Microbiol.* 95, 660–677. doi: 10.1111/mmi.12893

Bierne, H., and Dramsi, S. (2012). Spatial positioning of cell wall-anchored virulence factors in Gram-positive bacteria. *Curr. Opin. Microbiol.* 15, 715–723. doi: 10.1016/j.mib.2012.10.010

Brady, L. J., Maddocks, S. E., Larson, M. R., Forsgren, N., Persson, K., Deivanayagam, C. C., et al. (2010). The changing faces of Streptococcus antigen I/II polypeptide family adhesins. *Mol. Microbiol.* 77, 276–286. doi: 10.1111/j.1365-2958.2010.07212.x

Busscher, H. J., van de Belt-Gritter, B., Dijkstra, R. J. B., Norde, W., Petersen, F. C., Scheie, A. A., et al. (2007). intermolecular forces and enthalpies in the adhesion of *Streptococcus mutans* and an antigen I/II-deficient mutant to laminin films. *J. Bacteriol.* 189, 2988–2995. doi: 10.1128/jb.01731-06

Chow, J. W., Thal, L. A., Perri, M. B., Vazquez, J. A., Donabedian, S. M., Clewell, D. B., et al. (1993). Plasmid-associated hemolysin and aggregation substance production contribute to virulence in experimental *Enterococcal* endocarditis. *Antimicrob. Agents Chemother.* 37, 2474–2477. doi: 10.1128/aac.37.11.2474

Chuang, O. N., Schlievert, P. M., Wells, C. L., Manias, D. A., Tripp, T. J., and Dunny, G. M. (2009). Multiple functional domains of enterococcus faecalis aggregation substance Asc10 contribute to endocarditis virulence. *Infect. Immun.* 77, 539–548. doi: 10.1128/iai.01034-08

Chuang-Smith, O. N., Wells, C. L., Henry-Stanley, M. J., and Dunny, G. M. (2010). Acceleration of *Enterococcus faecalis* biofilm formation by aggregation substance expression in an ex vivo model of cardiac valve colonization. *PLoS One* 5:e15798. doi: 10.1371/journal.pone.0015798

Davies, M. R., Shera, J., Van Domselaar, G. H., Sriprakash, K. S., and McMillan, D. J. A. (2009). Novel integrative conjugative element mediates genetic transfer from group G *Streptococcus* to other β-Hemolytic *Streptococci*. *J. Bacteriol.* 191, 2257–2265. doi: 10.1128/jb.01624-08

Demuth, D. R., Davis, C. A., Corner, A. M., Lamont, R. J., Leboy, P. S., and Malamud, D. (1988). Cloning and expression of a *Streptococcus* sanguis surface antigen that interacts with a human salivary agglutinin. *Infect. Immun.* 56, 2484–2490. doi: 10.1128/iai.56.9.2484-2490.1988

Demuth, D. R., Duan, Y., Brooks, W., Holmes, A. R., McNab, R., and Jenkinson, H. F. (1996). Tandem genes encode cell-surface polypeptides SspA and SspB which mediate adhesion of the oral bacterium *Streptococcus gordonii* to human and bacterial receptors. *Mol. Microbiol.* 20, 403–413. doi: 10.1111/j.1365-2958.1996.tb02627.x

Demuth, D. R., Golub, E. E., and Malamud, D. (1990). *Streptococcal*-host interactions. Structural and functional analysis of a *Streptococcus sanguis* receptor for a human salivary glycoprotein. *J. Biol. Chem.* 265, 7120–7126.

Dereeper, A., Guignon, V., Blanc, G., Audic, S., Buffet, S., Chevenet, F., et al. (2008). Phylogeny.fr: robust phylogenetic analysis for the non-specialist. *Nucleic Acids Res.* 36, W465–W469.

Duan, Y., Fisher, E., Malamud, D., Golub, E., and Demuth, D. R. (1994). Calcium-binding properties of SSP-5, the *Streptococcus gordonii* M5 receptor for salivary agglutinin. *Infect. Immun.* 62, 5220–5226. doi: 10.1128/iai.62.12.5220-5226.1994

Dufrêne, Y. F., and Viljoen, A. (2020). Binding strength of gram-positive bacterial adhesins. *Front. Microbiol.* 11:1457. doi: 10.3389/fmicb.2020.01457

Dunny, G. M., Brown, B. L., and Clewell, D. B. (1978). Induced cell aggregation and mating in *Streptococcus faecalis*: evidence for a bacterial sex pheromone. *Proc. Natl. Acad. Sci. U S A.* 75, 3479–3483. doi: 10.1073/pnas.75.7.3479

Egland, P. G., Dû. L. D. and Kolenbrander, P. E. (2001). Identification of Independent *Streptococcus gordonii* SspA and SspB functions in coaggregation with *Actinomyces naeslundii*. *Infect. Immun.* 69, 7512–7516.

Forsgren, N., Lamont, R. J., and Persson, K. (2009). Crystal structure of the variable domain of the *Streptococcus gordonii* surface protein SspB. *Prot. Sci.* 18, 1896–1905. doi: 10.1002/pro.200

Forsgren, N., Lamont, R. J., and Persson, K. (2010). Two intramolecular isopeptide bonds are identified in the crystal structure of the *Streptococcus gordonii* SspB C-terminal domain. *J. Mol. Biol.* 397, 740–751. doi: 10.1016/j.jmb.2010.01.065

Foster, T. J. (2019). The MSCRAMM Family of cell-wall-anchored surface proteins of gram-positive cocci. *Trends Microbiol.* 27, 927–941. doi: 10.1016/j.tim.2019.06.007

Foster, T. J., Geoghegan, J. A., Ganesh, V. K., and Höök, M. (2014). Adhesion, invasion and evasion: the many functions of the surface proteins of *Staphylococcus aureus*. *Nat. Rev. Microbiol.* 12, 49–62. doi: 10.1038/nrmicro3161

Galli, D., Friesenegger, A., and Wirth, R. (1992). Transcriptional control of sex-pheromone-inducible genes on plasmid pAD1 of *Enterococcus faecalis* and sequence analysis of a third structural gene for (pPD1 -encoded) aggregation substance. *Mol. Microbiol.* 6, 1297–1308. doi: 10.1111/j.1365-2958.1992.tb00851.x

Galli, D., Lottspeich, F., and Wirth, R. (1990). Sequence analysis of *Enterococcus faecalis* aggregation substance encoded by the sex pheromone plasmid pAD1. *Mol. Microbiol.* 4, 895–904. doi: 10.1111/j.1365-2958.1990.tb00662.x

Gasson, M. J., Godon, J. J., Pillidge, C. J., Eaton, T. J., Jury, K., and Shearman, C. A. (1995). Characterization and exploitation of conjugation in *Lactococcus lactis*. *Int. Dairy J.* 5, 757–762. doi: 10.1016/0958-6946(95)00030-5

Geoghegan, J. A., and Foster, T. J. (2015). Cell wall-anchored surface proteins of staphylococcus aureus: many proteins, multiple functions. *Curr. Topics Microbiol. Imm.* 409, 95–120. doi: 10.1007/82_2015_5002

Giomarelli, B., Visai, L., Hijazi, K., Rindi, S., Ponzio, M., Iannelli, F., et al. (2006). Binding of *Streptococcus gordonii* to extracellular matrix proteins. *FEMS Microbiol. Lett.* 265, 172–177.

Godon, J.-J., Jury, K., Shearman, C. A., and Gasson, M. J. (1994). The *Lactococcus lactis* sex-factor aggregation gene cluA. *Mol. Microbiol.* 12, 655–663. doi: 10.1111/j.1365-2958.1994.tb01053.x

Guo, S., Vance, T. D. R., Stevens, C. A., Voets, I., and Davies, P. L. R. T. X. (2019). Adhesins are key bacterial surface megaproteins in the formation of biofilms. *Trends Microbiol.* 27, 453–467. doi: 10.1016/j.tim.2018.12.003

Hajishengallis, G., Koga, T., and Russell, M. W. (1994). Affinity and specificity of the interactions between *Streptococcus mutans* antigen I/II and salivary components. *J. Dent. Res.* 73, 1493–1502. doi: 10.1177/00220345940730090301

Hamada, S., and Slade, H. D. (1980). Biology, immunology, and cariogenicity of *Streptococcus mutans*. *Microbiol. Rev.* 44, 331–384. doi: 10.1128/mmbr.44.2.331-384.1980

Heddle, C., Nobbs, A. H., Jakubovics, N. S., Gal, M., Mansell, J. P., Dymock, D., et al. (2003). Host collagen signal induces antigen I/II adhesin and invasin gene expression in oral *Streptococcus gordonii*. *Mol. Microbiol.* 50, 597–607. doi: 10.1046/j.1365-2958.2003.03711.x

Heim, K. P., Crowley, P. J., Long, J. R., Kailasan, S., McKenna, R., and Brady, L. J. (2014). An intramolecular lock facilitates folding and stabilizes the tertiary structure of *Streptococcus mutans* adhesin P1. *Proc. Natl. Acad. Sci.* 111, 15746–15751. doi: 10.1073/pnas.1413018111

Hirt, H., Erlandsen, S. L., and Dunny, G. M. (2000). Heterologous inducible expression of *Enterococcus faecalis* pCF10 aggregation substance Asc10 in *Lactococcus lactis* and *Streptococcus gordonii* contributes to cell hydrophobicity and adhesion to fibrin. *J. Bacteriol.* 182, 2299–2306. doi: 10.1128/jb.182.8.2299-2306.2000

Holmes, A. R., Gilbert, C., Wells, J. M., and Jenkinson, H. F. (1998). Binding properties of *Streptococcus gordonii* SspA and SspB (antigen I/II family) polypeptides expressed on the cell surface of lactococcus lactis MG1363. *Infect. Immun.* 66, 4633–4639. doi: 10.1128/iai.66.10.4633-4639.1998

Isenmann, R., Schwarz, M., Rozdzinski, E., Christ, C., Schmidt, E., Augat, P., et al. (2002). Interaction of fibronectin and aggregation substance promotes adherence of *Enterococcus faecalis* to human colon. *Dig. Dis. Sci.* 47, 462–468.

Jakubovics, N. S., Kerrigan, S. W., and Nobbs, A. H., Stroümberg, N., van Dolleweerd, C. J., Cox, D. M., et al. (2005a). 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. doi: 10.1128/iai.73.10.6629-6638.2005

Jakubovics, N. S., Strömberg, N., Van Dolleweerd, C. J., Kelly, C. G., and Jenkinson, H. F. (2005b). Differential binding specificities of oral *Streptococcal antigen* I/II family adhesins for human or bacterial ligands. *Mol. Microbiol.* 55, 1591–1605. doi: 10.1111/j.1365-2958.2005.04495.x

Jenkinson, H. F., and Demuth, D. R. (1997). Structure, function and immunogenicity of *Streptococcal antigen* I/II polypeptides. *Mol. Microbiol.* 23, 183–190. doi: 10.1046/j.1365-2958.1997.2021577.x

Kang, H. J., and Baker, E. N. (2012). Structure and assembly of gram-positive bacterial pili: unique covalent polymers. *Curr. Opin. Struct. Biol.* 22, 200–207. doi: 10.1016/j.sbi.2012.01.009

Kang, M., Ko, Y. P., Liang, X., Ross, C. L., Liu, Q., Murray, B. E., et al. (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. doi: 10.1074/jbc.m113.454462

Kelly, C., Evans, P., and Ma, J. K., Bergmeier, L. A., Taylor, W., Brady, L. J., et al. (1990). Sequencing and characterization of the 185 kDa cell surface antigen of *Streptococcus mutans*. *Arch. Oral Biol.* 35, S33–S38.

Koga, T., Asakawa, H., Okahashi, N., and Hamada, S. (1986). Sucrose-dependent cell adherence and cariogenicity of serotype c *Streptococcus mutans*. *Microbiology* 132, 2873–2883. doi: 10.1099/00221287-132-10-2873

Kohler, V., Keller, W., and Grohmann, E. (2018). Enterococcus adhesin PrgB facilitates type IV secretion by condensation of extracellular DNA. *Mol. Microbiol.* 109, 263–267. doi: 10.1111/mmi.13994

Kojic, M., Jovcic, B., Strahinic, I., Begovic, J., Lozo, J., Veljovic, K., et al. (2011). Cloning and expression of a novel *Lactococcal* aggregation factor from *Lactococcus lactis* subsp. lactis BGKP1. *BMC Microbiol.* 11:265. doi: 10.1186/1471-2180-11-265

Kreft, B., Marre, R., Schramm, U., and Wirth, R. (1992). Aggregation substance of *Enterococcus faecalis* mediates adhesion to cultured renal tubular cells. *Infect. Immun.* 60, 25–30. doi: 10.1128/iai.60.1.25-30.1992

Kuykindoll, R. J., and Holt, R. G. (1996). Characterization of a P1-deficient strain of *Streptococcus mutans* that expresses the SpaA protein of *Streptococcus sobrinus*. *Infect. Immun.* 64, 3652–3658. doi: 10.1128/iai.64.9.3652-3658.1996

Larson, M. R., Rajashankar, K. R., Crowley, P. J., Kelly, C., Mitchell, T. J., Brady, L. J., et al. (2011). Crystal structure of the C-terminal region of *Streptococcus mutans* antigen I/II and characterization of salivary agglutinin adherence domains. *J. Biol. Chem.* 286, 21657–21666. doi: 10.1074/jbc.m111.231100

Larson, M. R., Rajashankar, K. R., Patel, M. H., Robinette, R. A., Crowley, P. J., Michalek, S., et al. (2010). Elongated fibrillar structure of a streptococcal adhesin assembled by the high-affinity association of α- and PPII-helices. *Proc. Natl. Acad. Sci. U S A.* 107, 5983–5988. doi: 10.1073/pnas.0912293107

Lee, S. F. (1992). Identification and characterization of a surface protein-releasing activity in *Streptococcus mutans* and other pathogenic streptococci. *Infect. Immun.* 60, 4032–4039. doi: 10.1128/iai.60.10.4032-4039.1992

Lee, S. F., Li, Y. H., and Bowden, G. H. (1996). Detachment of *Streptococcus mutans* biofilm cells by an endogenous enzymatic activity. *Infect. Immun.* 64, 1035–1038. doi: 10.1128/iai.64.3.1035-1038.1996

Love, R. M., McMillan, M. D., and Jenkinson, H. F. (1997). Invasion of dentinal tubules by oral streptococci is associated with collagen recognition mediated by the antigen I/II family of polypeptides. *Infect. Immun.* 65, 5157–5164. doi: 10.1128/iai.65.12.5157-5164.1997

Lukaszczyk, M., Pradhan, B., and Remaut, H. (2019). The biosynthesis and structures of bacterial pili. *Subcell Biochem.* 92, 369–413. doi: 10.1007/978-3-030-18768-2_12 doi: 10.1007/978-3-030-18768-2_12

Lukić, J., Strahinić, I., Jovèić, B., Filipić, B., Topisirović, L., Kojić, M., et al. (2012). Different roles for *Lactococcal* aggregation factor and mucin binding protein in adhesion to gastrointestinal mucosa. *Appl. Environ. Microbiol.* 78, 7993–8000. doi: 10.1128/aem.02141-12

Lynch, D. J., Michalek, S. M., Zhu, M., Drake, D., Qian, F., and Banas, J. A. (2013). Cariogenicity of *Streptococcus mutans* glucan-binding protein deletion mutants. *Oral Health Dent. Manag.* 12, 191–199.

Mieher, J. L., Larson, M. R., Schormann, N., Purushotham, S., Wu, R., Rajashankar, K. R., et al. (2018). Glucan binding protein C of *Streptococcus mutans* mediates both sucrose-independent and sucrose-dependent adherence. *Infect. Immun.* 86, 1–18.

Mitchell, A. L., Attwood, T. K., Babbitt, P. C., Blum, M., Bork, P., Bridge, A., et al. (2019). InterPro in 2019: improving coverage, classification and access to protein sequence annotations. *Nucleic Acids Res.* 47, D351–D360.

Nakano, K., Tsuji, M., Nishimura, K., Nomura, R., and Ooshima, T. (2006). Contribution of cell surface protein antigen PAc of *Streptococcus mutans* to bacteremia. *Microbes Infect.* 8, 114–121. doi: 10.1016/j.micinf.2005.06.005

Nakayama, J., Watarai, H., Isogai, A., Clewell, D. B., and Suzuki, A. C. -(1992). Terminal Identification of AD74, a proteolytic product of *Enterococcus faecalis* aggregation substance: application of liquid chromatography/mass spectrometry. *Biosci. Biotechnol. Biochem.* 56, 127–131. doi: 10.1271/bbb.56.127

Nobbs, A. H., Shearer, B. H., Drobni, M., Jepson, M. A., and Jenkinson, H. F. (2007). Adherence and internalization of *Streptococcus gordonii* by epithelial cells involves β1 integrin recognition by SspA and SspB (antigen I/II family) polypeptides. *Cell. Microbiol.* 9, 65–83. doi: 10.1111/j.1462-5822.2006.00768.x

Nylander, Å, Forsgren, N., and Persson, K. (2011). Structure of the C-terminal domain of the surface antigen SpaP from the caries pathogen *Streptococcus mutans*. *Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun.* 67, 23–26. doi: 10.1107/s174430911004443x

Oho, T., Yu, H., Yamashita, Y., and Koga, T. (1998). Binding of salivary glycoprotein-secretory immunoglobulin a complex to the surface protein antigen of *Streptococcus mutans*. *Infect. Immun.* 66, 115–121. doi: 10.1128/iai.66.1.115-121.1998

Okamoto-Shibayama, K., Sato, Y., Yamamoto, Y., Ohta, K., and Kizaki, H. (2006). Identification of a glucan-binding protein C gene homologue in Streptococcus macacae. *Oral Microbiol. Immunol.* 21, 32–41. doi: 10.1111/j.1399-302x.2005.00251.x

Petersen, F. C., Pasco, S., Ogier, J., Klein, J. P., Assev, S., and Scheie, A. A. (2001). Expression and functional properties of *theStreptococcus* intermedius surface protein antigen I/II. *Infect. Immun.* 69, 4647–4653. doi: 10.1128/iai.69.7.4647-4653.2001

Purushotham, S., and Deivanayagam, C. (2014). The calcium-induced conformation and glycosylation of scavenger-rich cysteine repeat (SRCR) domains of glycoprotein 340 influence the high affinity interaction with antigen I/II homologs. *J. Biol. Chem.* 289, 21877–21887. doi: 10.1074/jbc.m114.565507

Rakita, R. M., Vanek, N. N., Jacques-Palaz, K., Mee, M., Mariscalco, M. M., Dunny, G. M., et al. (1999). *Enterococcus faecalis* bearing aggregation substance is resistant to killing by human neutrophils despite phagocytosis and neutrophil activation. *Infect. Immun.* 67, 6067–6075. doi: 10.1128/iai.67.11.6067-6075.1999

Reichhardt, M. P., Loimaranta, V., Lea, S. M., and Johnson, S. (2020). Structures of SALSA/DMBT1 SRCR domains reveal the conserved ligand-binding mechanism of the ancient SRCR fold. *Life Sci. Alliance* 3:e201900502. doi: 10.26508/lsa.201900502

Rozdzinski, E., Marre, R., Susa, M., Wirth, R., and Muscholl-Silberhorn, A. (2001). Aggregation substance-mediated adherence of *Enterococcus faecalis* to immobilized extracellular matrix proteins. *Microb. Pathog.* 30, 211–220. doi: 10.1006/mpat.2000.0429

Rumbaugh, K. P., and Sauer, K. (2020). Biofilm dispersion. *Nat. Rev. Microbiol.* 18, 571–586. doi: 10.1038/s41579-020-0385-0

Russell, M. W., Bergmeier, L. A., Zanders, E. D., and Lehner, T. (1980). Protein antigens of *Streptococcus mutans*: purification and properties of a double antigen and its protease-resistant component. *Infect. Immun.* 28, 486–493.

Sato, Y., Okamoto, K., and Kizaki, H. (2002a). gbpC and pac gene mutations detected in *Streptococcus mutans* strain GS-5. *Oral Microbiol. Immunol.* 17, 263–266. doi: 10.1034/j.1399-302x.2002.170410.x

Sato, Y., Okamoto-Shibayama, K., Takada, K., Igarashi, T., and Hirasawa, M. (2009). Genes responsible for dextran-dependent aggregation of *Streptococcus sobrinus* strain 6715. *Oral Microbiol. Immunol.* 24, 224–230. doi: 10.1111/j.1399-302x.2008.00499.x

Sato, Y., Senpuku, H., Okamoto, K., Hanada, N., and Kizaki, H. (2002b). *Streptococcus mutans* binding to solid phase dextran mediated by the glucan-binding protein C. *Oral Microbiol. Immunol.* 17, 252–256. doi: 10.1034/j.1399-302x.2002.170408.x

Sato, Y., Yamamoto, Y., and Kizaki, H. (1997). Cloning and sequence analysis of the gbpC gene encoding a novel glucan-binding protein of *Streptococcus mutans*. *Infect. Immun.* 65, 668–675. doi: 10.1128/iai.65.2.668-675.1997

Schlievert, P. M., Gahr, P. J., Assimacopoulos, A. P., Dinges, M. M., Stoehr, J. A., Harmala, J. W., et al. (1998). Aggregation and binding substances enhance pathogenicity in rabbit models of *Enterococcus faecalis* endocarditis. *Infect. Immun.* 66, 218–223. doi: 10.1128/iai.66.1.218-223.1998

Schmitt, A., Hirt, H., Järvå, M. A., Sun, W.-S., Beek, J. T., Dunny, G. M., et al. (2020). *Enterococcal* PrgA extends far outside the cell and provides surface exclusion to protect against unwanted conjugation. *bioRxiv Prepr.* 432, 5681–5695. doi: 10.1101/2020.06.18.156562

Schmitt, A., Jiang, K., Camacho, M. I., Jonna, V. R., Hofer, A., Westerlund, F., et al. (2018). PrgB promotes aggregation, biofilm formation, and conjugation through DNA binding and compaction. *Mol. Microbiol.* 109, 291–305. doi: 10.1111/mmi.13980

Schwarz-Linek, U., Höök, M., and Potts, J. R. (2006). Fibronectin-binding proteins of gram-positive cocci. *Microbes. Infect.* 8, 2291–2298. doi: 10.1016/j.micinf.2006.03.011

Sciotti, M.-A., Yamodo, I., Klein, J.-P., and Ogier, J. A. (2006). The N-terminal half part of the oral *Streptococcal antigen* I/IIf contains two distinct binding domains. *FEMS Microbiol. Lett.* 153, 439–445. doi: 10.1111/j.1574-6968.1997.tb12608.x

Sommer, P., Bruyère, T., Ogier, J. A., Garnier, J. M., Jeltsch, J. M., and Klein, J. P. (1987). Cloning of the saliva-interacting protein gene from *Streptococcus mutans*. *J. Bacteriol.* 169, 5167–5173. doi: 10.1128/jb.169.11.5167-5173.1987

Stentz, R., Gasson, M., and Shearman, C. (2006). The tra domain of the *Lactococcal* CluA surface protein is a unique domain that contributes to sex factor DNA transfer. *J. Bacteriol.* 188, 2106–2114. doi: 10.1128/jb.188.6.2106-2114.2006

Stentz, R., Jury, K., Eaton, T., Parker, M., Narbad, A., Gasson, M., et al. (2004). Controlled expression of CluA in *Lactococcus lactis* and its role in conjugation. *Microbiology* 150, 2503–2512. doi: 10.1099/mic.0.27149-0

Sullan, R. M. A., Li, J. K., Crowley, P. J., Brady, L. J., and Dufrêne, Y. F. (2015). Binding forces of *Streptococcus mutans* P1 Adhesin. *ACS Nano* 9, 1448–1460.

Süßmuth, S. D., Muscholl-Silberhorn, A., Wirth, R., Susa, M., Marre, R., and Rozdzinski, E. (2000). Aggregation substance promotes adherence, phagocytosis, and intracellular survival of enterococcus faecalis within human macrophages and suppresses respiratory burst. *Infect. Immun.* 68, 4900–4906. doi: 10.1128/iai.68.9.4900-4906.2000

Tamura, H., Yamada, A., and Kato, H. (2014). Molecular characterization of the dextran-binding lectin B gene *dblB* of *Streptococcus criceti* in *Streptococcus mutans* strain GS-5 with mutations in both *gbpC* and *spaP* genes. *Genes Genet. Syst.* 89, 41–50. doi: 10.1266/ggs.89.41

Troffer-Charlier, N., Ogier, J., Moras, D., and Cavarelli, J. (2002). Crystal structure of the V-region of *Streptococcus mutans* antigen I/II at 2.4 å resolution suggests a sugar preformed binding site. *J. Mol. Biol.* 318, 179–188. doi: 10.1016/S0022-2836(02)00025-6

Vanek, N. (1999). *Enterococcus faecalis* aggregation substance promotes opsonin-independent binding to human neutrophils via a complement receptor type 3-mediated mechanism. *FEMS Immunol. Med. Microbiol.* 26, 49–60. doi: 10.1016/s0928-8244(99)00120-0

Vercellotti, G. M., McCarthy, J. B., Lindholm, P., Peterson, P. K., Jacob, H. S., and Furcht, L. T. (1985). Extracellular matrix proteins (fibronectin, laminin, and type IV collagen) bind and aggregate bacteria. *Am. J. Pathol.* 120, 13–21.

Waters, C. M., and Dunny, G. M. (2001). Analysis of functional domains of the enterococcus faecalis pheromone-induced surface protein aggregation substance. *J. Appl. Microbiol.* 183, 5659–5667. doi: 10.1128/jb.183.19.5659-5667.2001

Waters, C. M., Hirt, H., McCormick, J. K., Schlievert, P. M., Wells, C. L., and Dunny, G. M. (2004). An amino-terminal domain of enterococcus faecalis aggregation substance is required for aggregation, bacterial internalization by epithelial cells and binding to lipoteichoic acid. *Mol. Microbiol.* 52, 1159–1171. doi: 10.1111/j.1365-2958.2004.04045.x

Waters, C. M., Wells, C. L., and Dunny, G. M. (2003). The aggregation domain of aggregation substance. Not the RGD Motifs, Is Critical for Efficient Internalization by HT-29 Enterocytes. *Infect. Immun.* 71, 5682–5689. doi: 10.1128/iai.71.10.5682-5689.2003

Wells, C. L., Moore, E. A., Hoag, J. A., Hirt, H., Dunny, G. M., and Erlandsen, S. L. (2000). Inducible expression of *Enterococcus faecalis* aggregation substance surface protein facilitates bacterial internalization by cultured enterocytes. *Infect. Immun.* 68, 7190–7194. doi: 10.1128/iai.68.12.7190-7194.2000

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Microbiology

REVIEW
published: 23 December 2020
doi: 10.3389/fmicb.2020.613581

Revisiting Bap Multidomain Protein: More Than Sticking Bacteria Together

*Jaione Valle[1], Xianyang Fang[2] and Iñigo Lasa[3]**

[1]Instituto de Agrobiotecnología, CSIC-Gobierno de Navarra, Mutilva, Spain, [2]Beijing Advanced Innovation Center for Structural Biology, School of Life Sciences, Tsinghua University, Beijing, China, [3]Laboratory of Microbial Pathogenesis, Navarrabiomed-Universidad Pública de Navarra-Departamento de Salud, IDISNA, Pamplona, Spain

One of the major components of the staphylococcal biofilm is surface proteins that assemble as scaffold components of the biofilm matrix. Among the different surface proteins able to contribute to biofilm formation, this review is dedicated to the Biofilm Associated Protein (Bap). Bap is part of the accessory genome of *Staphylococcus aureus* but orthologs of Bap in other staphylococcal species belong to the core genome. When present, Bap promotes adhesion to abiotic surfaces and induces strong intercellular adhesion by self-assembling into amyloid like aggregates in response to the levels of calcium and the pH in the environment. During infection, Bap enhances the adhesion to epithelial cells where it binds directly to the host receptor Gp96 and inhibits the entry of the bacteria into the cells. To perform such diverse range of functions, Bap comprises several domains, and some of them include several motifs associated to distinct functions. Based on the knowledge accumulated with the Bap protein of *S. aureus*, this review aims to summarize the current knowledge of the structure and properties of each domain of Bap and their contribution to Bap functionality.

Keywords: Bap, biofilm, *Staphylococcus*, calcium, amyloid oligomers, adhesion

OPEN ACCESS

Edited by:
Magnus Hook,
Texas A&M Health Science Center,
United States

Reviewed by:
Peter Lipke,
Brooklyn College (CUNY),
United States
Joan A. Geoghegan,
University of Birmingham,
United Kingdom
Andrew Benjamin Herr,
Cincinnati Children's Hospital Medical Center, United States

****Correspondence:***
Iñigo Lasa
ilasa@unavarra.es

Specialty section:
This article was submitted to Infectious Diseases, a section of the journal Frontiers in Microbiology

Received: *02 October 2020*
Accepted: *03 December 2020*
Published: *23 December 2020*

Citation:
Valle J, Fang X and Lasa I (2020) Revisiting Bap Multidomain Protein: More Than Sticking Bacteria Together. Front. Microbiol. 11:613581. doi: 10.3389/fmicb.2020.613581

INTRODUCTION

Biofilm associated protein (Bap) was identified during the analysis of biofilm defective transposon insertion mutants in the bovine mastitis isolate, *Staphylococcus aureus* V329 (Cucarella et al., 2001). *S. aureus* V329 strain had been selected among a collection of bovine mastitis isolates for its strong capacity to produce biofilm in the classical microtiter plate assay. Sequencing of the *bap* gene surrounding region revealed that Bap was contained in a pathogenicity island (SaPIbov2; Ubeda et al., 2003). The finding showing the connection of Bap with the biofilm phenotype was somehow serendipitous. The integrase gene (*sip*) of SaPIbov2 in *S. aureus* V329 carried two-point mutations that abrogated the integrase functionality. As a consequence, SaPIbov2 remained stable in the bacterial chromosome. If the integrase would have been functional, most of the biofilm deficient mutants had been due to the loss of the SaPI element and not to genetic inactivation caused by transposon insertion. The presence of Bap in a pathogenicity island was anticipating that the fitness cost imposed by SaPIbov2 carriage was compensated by the benefit that the SaPI cargo (*bap* gene and the corresponding ABC transporter) conferred during the intramammary gland infection. Initial searches for Bap orthologs revealed

the presence of Bap orthologs in many different bacterial species (Lasa and Penadés, 2006). The recent sequencing of hundreds of genomes from clinical and commensal isolates of different staphylococcal species has confirmed that *bap* gene is widely spread in the genomes of staphylococcal species (Tormo et al., 2005; Potter et al., 2009; Schiffer et al., 2019). Very often, the *bap* genes are carried in mobile genetic elements. However, there is no indication that the *bap* locus is carried on a mobile genetic element in some staphylococcal species such as *Staphylococcus epidermidis*, *Staphylococcus simulans*, *Staphylococcus chromogenes*, and *Staphylococcus warneri*. Bap is a large protein that comprise multiple domains, each of them representing an independent folding unit (**Figure 1**; **Table 1**). Since the discovery of the Bap protein, most studies have been directed to investigate how this protein induces adhesion and biofilm development. Surprisingly, it has been shown that a large part of the protein is not required for adhesion-accumulation biofilm phenotype and that Bap-mediated biofilm development occurs in response to environmental cues. However, there is still very little knowledge about the function of a large part of Bap protein and how different domains of Bap interact to facilitate Bap functionality. In this review, we update the current understanding of the structure and function of each region of Bap. Due to the absence of structural biology data, the regions have been defined arbitrarily based on the specific features of the primary amino acid sequence. Most of the available information has been obtained on the Bap protein of *S. aureus*, but it is assumed that conclusions drawn from studies with this protein can most likely be applied to other Bap orthologs.

MULTIDOMAINS OF THE Bap PROTEIN

N-Terminal Domain

Signal Peptide

The staphylococcal Bap proteins are covalently linked to the cell wall envelope by a mechanism requiring the N-terminal secretory signal peptide and a C-terminal LPXTG motif sorting signal (Ton-That et al., 2000). Amino-acid sequences alignment of signal peptides from different staphylococcal Bap proteins revealed the presence of a consensus motif

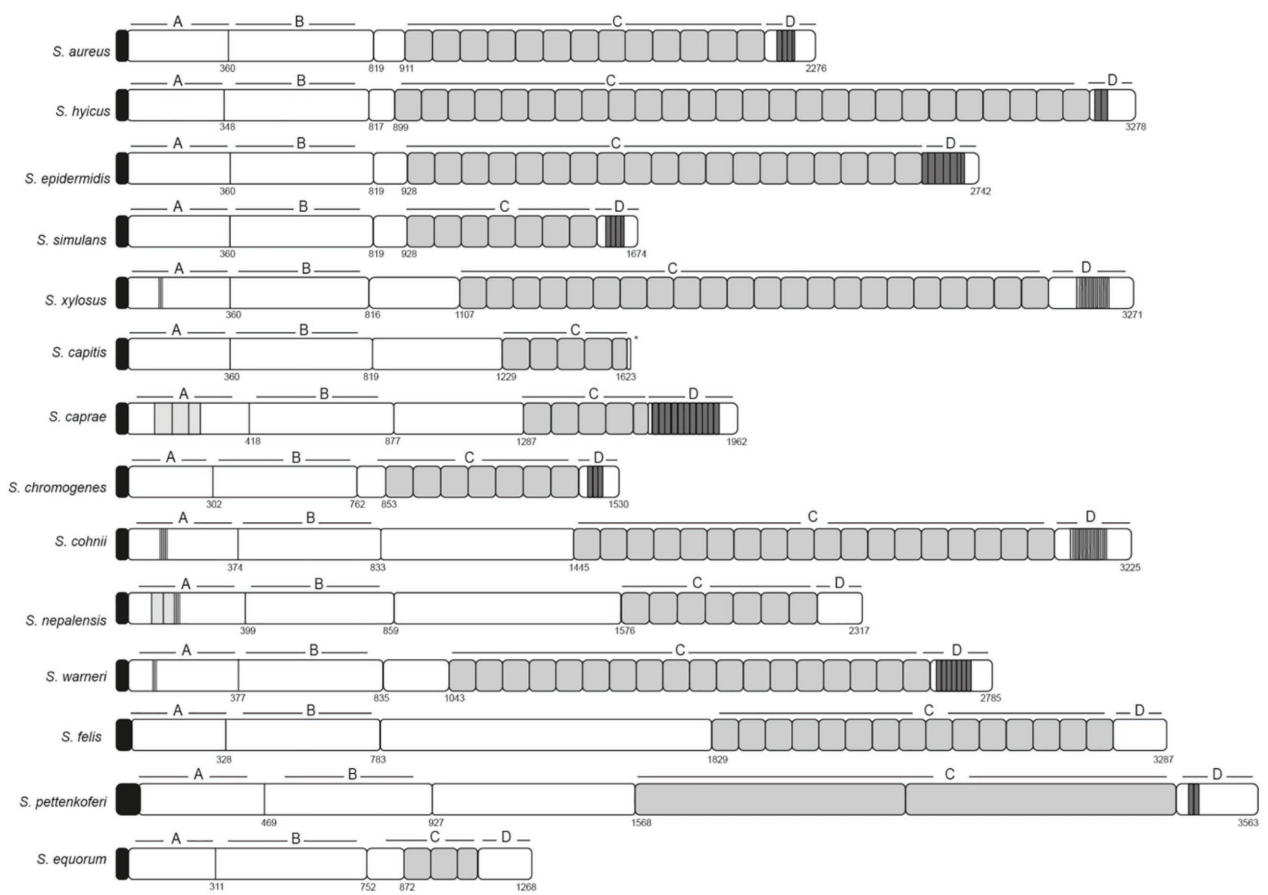


FIGURE 1 | Schematic representation of staphylococcal Bap proteins. Structural organization in multidomains including the N-terminal domain (region A and B), core C-repeats domain and the carboxy-terminal domain (D). Repeats from A, C, and D domain are shown in gray. Signal peptide is shown in black.

TABLE 1 | Staphylococcal Bap protein orthologs.

	Accession number	Length (amino acids)				Number repeats				Coiled-coil (position)	EF-hands (position)
		Protein	A-domain	B-domain	C-domain	A-repeat	C-repeat	SD-repeat	D-repeat		
S. aureus	AAK38834.2	2,276	45–360	361–819	911–2,116	–	14	3.4	–	224–281	727-740; 752-762
S. hyicus	AAY28520.1	3,278	45–348	349–807	899–3,136	–	26	2.1	–	212–269	715-728; 740-750
S. epidermidis	AAY28519.1	2,742	45–360	361–819	928–2,562	–	19	7	–	227–281	727-740; 752-762
S. simulans	AAY28518.1	1,674	45–360	361–819	928–1,530	–	7	3.4	–	227–281	727-740; 752-762
S. xylosus	AAY28517.1	3,271	45–360	361–816	1,107–2,998	2	22	–	17	67–110	723-736; 748-758
S. capitis	WP_080973601	1,623	45–360	361–819	1,229–1,609	–	4	–	–	227–281	727-740; 752-762
S. caprae	WP_170080878	1962	45–418	419–877	1,287–1,667	3	4	11.3	–	284–338	785-798; 810-820
S. chromogenes	AAY28516.1	1,530	45–302	303–761	853–1,370	–	6	3.6	–	169–223	669-682; 694-704
S. cohnii	WP_157947920	3,235	45–374	375–833	1,445–2,992	4	18	–	20	83–113	741-754; 766-776
S. nepalensis	WP_158266100	2,317	45–399	400–859	1,576–2,177	5	7	–	–	94–122, 150–191, 270–321	767-780; 792-802
S. warneri	WP_107536078	2,785	45–376	377–835	1,043–2,590	2	18	7	–	–	743-756; 768-778
S. felis	AVP36288.1	3,287	57–328	329–783	1829–3,118	–	15	–	–	–	–
S. pettenkoferi	ASE37144.1	3,563	76–469	470–927	1,568–3,304	–	2	–	2	353–380	836-840; 861-871
S. equorum	ALM58189	1,268	45–311	312–752	872–1,100	–	3	–	–	138–166	660-673; 685-695

YSIRK/GS located 18–20 residues upstream of the signal peptide cleavage site (**Figure 2**; Dedent et al., 2007, 2008). This motif is conserved in the signal peptides of streptococci and staphylococci, but it is not present in the proteins of *Listeria monocytogenes*, bacilli, clostridia, or nocardia (Rosenstein and Gotz, 2000; Tettelin et al., 2001; Bae and Schneewind, 2003). In *S. aureus*, the YSIRK/GS is present in several cell-wall proteins such as clumping factor A (ClfA), Spa, fibronectin-binding protein B (FnbpB), serine-aspartate repeat protein C (SdrC) and D (SdrD; Dedent et al., 2008). Surface proteins transported through a signal peptide that contained the YSIRK/GS motif are deposited within the peptidoglycan cell wall at discrete foci and distributed in a ring-like or hemispherical manner perpendicular to the cell division plane. On the contrary, proteins secreted *via* signal peptides that lacked the YSIRK/GS motif are distributed in discrete, punctate deposits in the cell wall envelope (Dedent et al., 2007, 2008). Considering that Bap promotes adhesion to biotic and abiotic surfaces as well as intercellular adhesion, it seems reasonable that the protein needs to be evenly distributed on the cell surface to perform efficiently this function. We are not aware of any experimental evidence supporting that the distribution driven by the YSIRK/GS motif affects to Bap functionality, but this is an interesting hypothesis worthy of further study.

Region-A

Region A comprises ~250 to 400 amino acids (depending on the species) following the signal peptide (**Figure 1**). Sequence alignments of region A of Bap proteins from different staphylococcal species range from 98 to 10%. Overall, region A of Bap orthologs share an identity of 37%, being the highest conservation in the last amino acids of the domain. Domain A from *Staphylococcus capitis*, *S. simulans*, and *S. epidermidis* show sequence identity higher than 95% with domain A of Bap from *S. aureus* whereas domain A from *Staphylococcus equorum* shows the lowest sequence identity (10%; **Figure 3**). Region A can contain short repeats ranging from 4 to 56 amino acids, as it is the case for *Staphylococcus xylosus*, *Staphylococcus caprae*, *Staphylococcus cohnii*, *Staphylococcus nepalensis*, and *S. warneri* (**Figures 4A,B**). Secondary structure of region A is dominated by alpha helices (**Figure 4A**). Several informatic programs coincided to identify also coiled-coil motifs in Bap orthologs, except for the ones of *S. warneri* and *Staphylococcus felis*. Coiled-coil motifs are α-helical secondary structures mediating protein-protein interactions and multimerization through the coiling of helices that belongs to the same or different proteins (Parry et al., 2008; Fiumara et al., 2010). Intuitively, the presence of coiled coil secondary structure suggests that Bap-proteins could mediate intercellular adhesion through homophilic interactions between coiled-coil domains of opposing proteins in neighboring cells. Homophylic interactions between proteins of neighbor bacteria have been described for other staphylococcal surface proteins such as Aap and SasG. However, the homophylic interactions occurred between β-sheet–rich G5–E domains of Aap and SasG (Conrady et al., 2008, 2013; Geoghegan et al., 2010; Formosa-Dague et al., 2015). A *S. aureus* strain producing a derivative of Bap protein in which region A was deleted still maintained the capacity to form cell-to-cell aggregates and robust biofilms on polystyrene or on a glass surface under flow culture conditions, indicating that domain A of Bap is dispensable for these functions (Taglialegna et al., 2016b). Domain A could also be involved in the interaction with the host. Bap from *S. aureus* as well as Bap orthologs from *S. epidermidis*, *S. chromogenes*, and *Staphylococcus hyicus* are able to bind the host receptor Gp96 (Valle et al., 2012). The interaction of Bap with GP96 inhibits the entry of the bacteria into the cells by interfering with the fibronectin-binding protein mediated invasion pathway. Further studies are necessary to decipher whether domain A of Bap is involved in such interaction.

FIGURE 2 | Signal peptides of staphylococcal biofilm associated proteins. The N-terminal signal peptides of Bap proteins were aligned at their predicted signal peptidase cleavage sites. The signal peptides harbor the YSIRK/GS motif (red box). The signal cleavage site is indicated by an arrow.

	S. aureus	*S. hyicus*	*S. epidermidis*	*S. simulans*	*S. xylosus*	*S. capitis*	*S. caprae*	*S. chromogenes*	*S. cohnii*	*S. nepalesis*	*S. warneri*	*S. felis*	*S. pettenkoferi*	*S. equorum*
S. aureus		98%	97%	95%	81%	96%	96%	98%	81%	80%	79%	3%	65%	57%
S. hyicus	92%		97%	95%	81%	97%	95%	99%	82%	80%	80%	3%	65%	58%
S. epidermidis	96%	89%		97%	82%	99%	98%	97%	83%	81%	80%	3%	65%	57%
S. simulans	97%	91%	96%		80%	97%	96%	95%	81%	79%	78%	3%	63%	56%
S. xylosus	37%	37%	35%	36%		82%	81%	81%	84%	85%	85%	3%	63%	59%
S. capitis	98%	92%	97%	98%	37%		97%	97%	83%	81%	80%	3%	64%	57%
S. caprae	75%	70%	73%	75%	32%	75%		96%	81%	80%	78%	3%	64%	56%
S. chromogenes	72%	70%	70%	71%	33%	72%	72%		82%	80%	79%	3%	65%	57%
S. cohnii	35%	34%	34%	34%	32%	35%	30%	32%		84%	83%	2%	63%	57%
S. nepalesis	29%	29%	27%	29%	24%	29%	26%	28%	30%		99%	3%	63%	58%
S. warneri	28%	26%	28%	28%	23%	28%	25%	25%	24%	21%		3%	63%	57%
S. felis	20%	17%	20%	20%	18%	19%	20%	18%	18%	17%	19%		2%	2%
S. pettenkoferi	20%	19%	19%	21%	16%	21%	20%	20%	18%	15%	15%	15%		53%
S. equorum	16%	15%	16%	15%	13%	16%	13%	12%	15%	14%	18%	10%	7%	

FIGURE 3 | Pairwise percentage identity matrix between domains A (white boxes) and B (gray boxes) of Bap orthologs from the different staphylococcal species obtained from corresponding multiple sequence alignments done using *Geneious Prime* algorithm.

Region-B

Region B is comprised of the ~450 amino acids following region A (**Figure 1**). Region B is devoid of repetitions (**Figure 5A**). It is highly conserved among Bap orthologs (70% identity; **Figure 3**), and it exhibits similarity with Bap homologs from different bacterial genera, such as Esp protein from *Enterococcus faecalis* (Taglialegna et al., 2020). Region B of *S. aureus* contains two sequences with high similarity to the consensus EF-hand motif (amino acids 729–741 and amino acids 752–764). The EF-hand motif consists of a 12-residue loop flanked on both sides by a 12-residue alpha-helical domain and is involved in binding intracellular calcium (Strynadka and James, 1989). An early study aiming to determine the role of calcium in Bap functionality revealed that concentrations equivalent to those present in the milk (~10 mM) inhibit Bap-mediated bacterial aggregation and biofilm formation (Arrizubieta et al., 2004). This phenotype was dependent on the presence of EF-hand 2 and 3 because point mutations at positions 1, 3, and 12 of such motifs result in a Bap protein that still was able to induce biofilm development but it was unable to sense the presence of calcium in the media. These findings confirmed that EF-hand 2 and 3 constitute a basic stimulus-response coupling mechanism that allows Bap to induce or repress biofilm development depending on changes in the calcium concentration of the environment. With the exception of Bap from *S. felis*, all the staphylococcal Bap proteins present in the databases contain the conserved DxDxDG calcium-binding loop characteristic of the EF-hand motif, suggesting that calcium dependent regulation plays an important role in Bap functionality (**Figure 5B**). As we will see later, other regions of Bap might also be involved in the interaction with calcium.

Another interesting feature of region B is that it is self-sufficient to induce aggregation and biofilm phenotype. A *S. aureus* strain producing a chimeric protein containing the region B of Bap fused to ClfA protein conferred the same capacity to form biofilm than the entire Bap protein. Interestingly, the EF-hand motifs were still functional in the region B-ClfA chimeric protein because the biofilm induced by this protein was sensitive to the presence of calcium in the media and mutation of the EF-hand motifs renders the biofilm induced by the chimeric protein insensitive to the levels of calcium in the media (Taglialegna et al., 2016b). To add another layer of complexity to this region, a detailed analysis of region B sequence revealed the presence of two short stretches, peptide

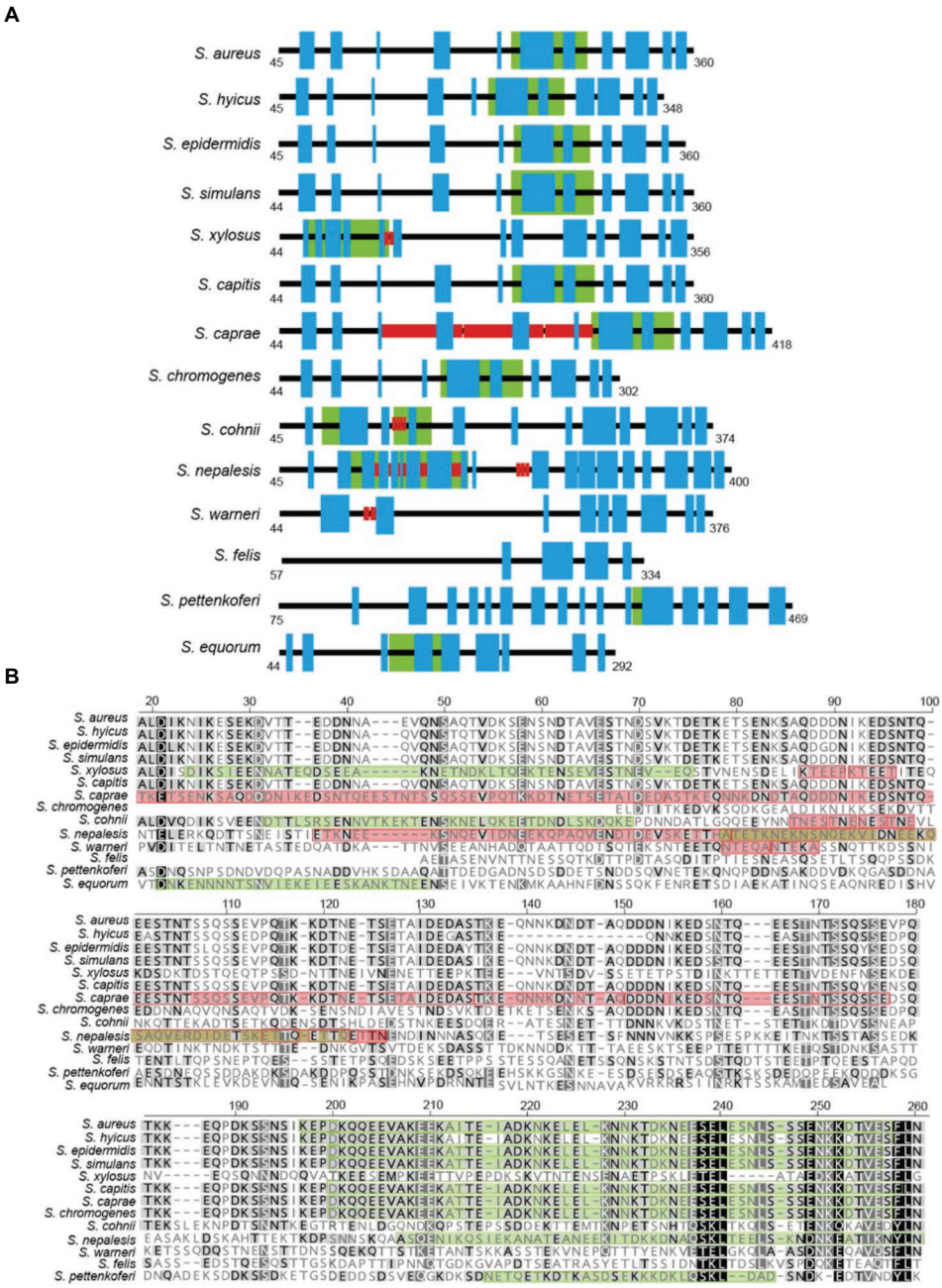


FIGURE 4 | (A) Representative scheme of region A from staphylococcal Bap proteins. Blue boxes indicate the predicted alpha helices secondary structure. Red boxes indicate the presence of short repeats. Green boxes indicate the prediction of coiled-coil motifs. **(B)** Protein sequence alignment of a section of region A from staphylococcal Bap proteins. Alignments were generated using *Geneious Prime* multiple sequence alignment tool. Green box within each diagram indicate predicted coiled-coil domains. Red box within each diagram indicate amino acids repeats. Coiled-coil motif was predicted by J. M. Lupas program (https://npsa-prabi.ibcp.fr/cgi-bin/npsa_automat.pl?page=npsa_lupas.html). Repeats were detected using XSTREAM algorithm (https://amnewmanlab.stanford.edu/xstream).

I $_{487}$TVGNIISNAG$_{496}$ and peptide II $_{579}$GIFSYS$_{584}$, with significant amyloidogenic potential (Taglialegna et al., 2016b; **Figure 5C**). These peptides as well as the purified recombinant Bap B region were able to form amyloid oligomers according to electron microscopy, attenuated total reflectance-Fourier transform infrared spectroscopy (ATR-FTIR), and binding to amyloid-related dyes such as Congo red, ThT, and Proteostat. Experimental evidence suggests that Bap is processed either enzymatically or non-enzymatically at the cell surface and fragments containing the N-terminal domain of the protein are released to the media. Later on, when the pH becomes acidic (pH < 5), the released region B of Bap undergoes a conformational change toward

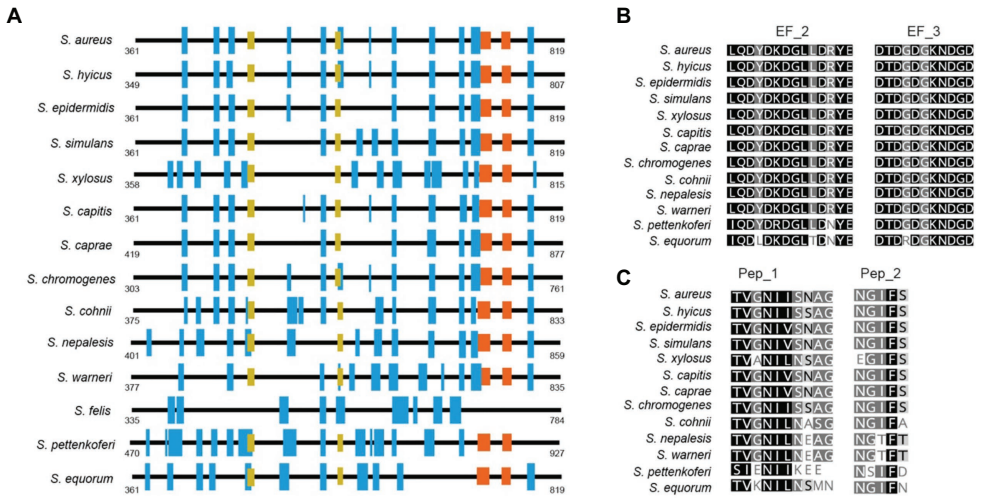


FIGURE 5 | (A) Representative scheme of region B from staphylococcal Bap proteins. Blue boxes show the secondary structure of alpha helices. EF_Hand 2_3 motifs are represented by orange boxes. Yellow boxes show the localization of amyloidogenic peptides 1 (Pep_I) and 2 (Pep_II) present in Bap proteins. **(B)** Alignment between EF-Hand motifs from staphylococci Bap proteins. **(C)** Conservation of amyloidogenic peptides I and II present in Bap of *S. aureus* among staphylococcal Bap orthologs proteins. Conserved residues are shown in black. Secondary structures were predicted at GOR software (https://npsa-prabi.ibcp.fr/cgi-bin/npsa_automat.pl?page=/NPSA/npsa_gor4.html). Coiled-coil motif was predicted by J. M. Lupas program (https://npsa-prabi.ibcp.fr/cgi-bin/npsa_automat.pl?page=npsa_lupas.html). Repeats were detected using XSTREAM algorithm (https://amnewmanlab.stanford.edu/xstream).

the amyloidogenic state resulting in the formation of fibrillar-like structures that mediate biofilm formation. The pH at which the region B of Bap shows aggregation activity is close to the isoelectric point of the region B peptide and therefore, the lack of net charge at this pH favors the conversion of the peptide from its native state (molten globule) to the amyloid-like state. At the same time, binding of calcium to the EF-hand domains of the protein stabilizes its molten-globule-like state, impairing their self-assembly into amyloid structures. Thus, the amyloid conformation of Bap is modulated by at least two environmental conditions, pH and the levels of calcium in the media.

This strategy to mediate multicellular behavior seems to be conserved among the Bap orthologs. Using the curli-dependent amyloid generator (C-DAG) system to detect amyloid aggregates (Sivanathan and Hochschild, 2013), it was shown that the region B from the Bap orthologs of *S. saprophyticus*, *Staphylococcus simiae*, *S. simulans*, *S. xylosus*, *S. epidermidis*, and region B from Esp of *E. faecalis* also displayed a similar amyloid-like properties (Taglialegna et al., 2020). Interestingly, amyloidogenic peptides similar to peptide I and peptide II were not present in the region B of Esp. However, this finding was not totally unexpected because the ability to form amyloid fibrils is a property inherent of the polypeptide chain, rather than a special property of any given sequence element. Other amyloidogenic peptides (sequence STVTVT) have been identified in C repeat region from *S. epidermidis*. Although the peptide is repeated 17 times in Bap protein of *S. epidermidis*, it is only present once in the Bap sequence of *S. aureus* (Lembre:2014jj). Thus, increasing number of evidence suggest that amyloid-like conformation is a general mechanism to mediate biofilm formation of the Bap_like proteins (Martino, 2016; Taglialegna et al., 2016a; Tursi and Tükel, 2018).

Core-Repeats Region

In *S. aureus*, region C is comprised of a core of 14 tandem repeats of 86 amino acids (C-repeats), which are identical even at the nucleotide level (Cucarella et al., 2001). Although the function of this region has not been discerned, it has been proposed that it could have a structural role to maintain the proper protein conformation on the cell surface. The number of C-repeats varies significantly between the different staphylococcal species and even within the same bacterial species (**Figure 1**; **Table 1**; Tormo et al., 2005). Early reports demonstrated that variation in the number of C-repeats is a highly dynamic process because Bap proteins with different C-repeats were generated during the course of a staphylococcal infection most likely through homologous recombination events between identical repeats (Cucarella et al., 2004). Interestingly, the number of repeats had a negligible consequence on the protein functionality, because all the variants of Bap containing different number of C-repeats display similar capacity to induce biofilm development. A later study showed that a Bap protein derivative containing a single repetition (ΔrepBap) have similar capacity to mediate biofilm formation compared to the wild type strain and also, similar capacity to interact with the ligand GP96 protein, a protein expressed on the plasma membrane of different cell types (Valle et al., 2012). However, the interaction of the Bap derivative containing a single C-repeat with GP96 was less efficient inhibiting the entry of the bacteria into the cells probably because it causes lower steric impairment to the fibronectin-binding protein mediated invasion pathway (Dziewanowska et al., 1999; Joh et al., 1999; Fowler et al., 2000).

Homology search identified a Bap-like protein SiiE in *Salmonella*, which contains 53 tandem imperfect repeats that

share significant similarities with the C-repeats of Bap from *S. aureus* (with an average of 25% identity; Latasa et al., 2005). These repeats are predicted to fold as a seven-strand β-sandwich and exhibit similarity to members of HYR (hyaline repeat) family that contain extracellular adhesion modules (Callebaut et al., 2000). Recently, the crystal structure of a SiiE fragment that covers repeats 50–52 of SiiE reveals the bacterial Ig-like (BIg) domain architecture and highlights two types of Ca^{2+}-binding sites. Because the BIg domains of SiiE are highly homologous, the full-length adhesin SiiE is proposed to act as a calcium-rigidified, rod-like, surface-anchored molecule (Griessl et al., 2013). Given the significant similarities in C-repeats between SiiE and Bap, it is tempting to speculate that the structure and function of core C-repeats of Bap proteins are also regulated by Ca^{2+}.

Carboxy-Terminal Domain

The C-terminal domain of the *S. aureus* Bap protein (Region D) comprises three repeats of 18 amino acids and a cell-wall-anchoring region consisting of an LPXTG motif and hydrophobic amino acid segments characteristic of surface proteins covalently anchored to peptidoglycan. Region D repeats are present in many of the staphylococcal Bap proteins (**Figure 1**; **Table 1**). The repeats often contained sequence stretches rich in serine and aspartic acid residues resembling the SD repeats present in staphylococcal surface proteins as the serine-aspartate repeat proteins C, D, and E (SdrC, SdrD, and SdrE) and Clumping factor A (ClfA; Josefsson:1998bq). In other cases, the repeats included short amino acids stretches (3–5 amino acids) that are repeated 17–20 times.

The biological significance of repeats from region D and its implication for biofilm development remains poorly understood. SD repeats might acts as rigid rods to project the N-terminal domain further from the cell surface and certain number of SD repeats are required for functional expression of the ligand-binding domain on the cell surface (Hartford et al., 1997; Foster et al., 2014; Speziale et al., 2014). SD repeats are genetically unstable, and the variation in the number of SD repeats help bacteria to adapt with high efficiency and low cost to environmental changes (Cheng et al., 2012). The SD-repeat region of diverse surface proteins is often glycosylated (Hazenbos et al., 2013). Glycosylation protects protein from the proteolytic attack by human neutrophil. The implication of the SD repeats present in the carboxyl-terminal domain of Bap or the role of the post-translational modification of these repeats in the functionality of Bap awaits further studies.

ROLE OF Bap IN PATHOGENESIS

The prevalence of *bap* in *S. aureus* and coagulase-negative staphylococcal isolates obtained from animal with mastitis is higher than in human isolates suggesting that specific host-dependent pathogenic factors involved in biofilm formation evolved independently in humans and ruminants (Cucarella et al., 2001; Tormo et al., 2005).

In lactating animals, the capacity to form biofilm on the epithelium of the mammary gland has been related to the propensity of staphylococci to produce chronic infections (Baselga et al., 1993, 1994; Cucarella et al., 2004). In this regard, the presence of Bap on the bacterial surface was shown to improve staphylococcal colonization and persistence on the mammary glands of infected animals (Cucarella et al., 2004).

Bap enhances the adhesion but inhibits the entry of *S. aureus* into epithelial cells (Cucarella et al., 2002, 2004). For that, Bap binds directly to the major chaperon Gp96 that is expressed on the surface of different cell types (Valle et al., 2012). The interaction of Bap with Gp96 may cause a steric impairment that hinders the recognition of FnBPs to fibronectin and integrins exposed on the cellular surface of the host cells, thus avoiding entry to host cells (Dziewanowska et al., 1999, 2000; Sinha and Fraunholz, 2010). Contrary to what happens with invasion, increased adhesion of Bap-positive strains was independent of the presence of Gp96 suggesting that Bap interacts with another factor to promote bacterial adhesion to epithelial cells.

In conclusion, Bap promotes adhesion to the epithelial cells of the mammary gland and impair bacterial internalization through the interaction with Gp96. This interaction facilitate the establishment of long-term persistent chronic infections by at least two mechanisms, generation of a microenvironment that acts as a physical barrier for host immune cells and shielding bacterial cells from detection by the immune system (Scherr et al., 2014; Schilcher and Horswill, 2020).

FINAL REMARKS

Future directions of research in Bap protein should considered at least two issues/questions. First, it would be very helpful to obtain a full-length structural model of Bap to understand how the interaction between the different domains or the function of putative active sites hidden at the interface between domains affects Bap biological functions. A structural model of Bap protein will aid to explain the phenotype of mutants generated in previous studies and more importantly in the designing of new mutants to disrupt or alter specific functions. Bap exemplifies the difficulties that exist with large multidomain proteins to close the gap between the increased information about the structure and function of single-domains and the requirement of high-resolution protein structures to understand protein functionality. To solve this problem, several multidomain structure modeling strategies have been recently developed (Xu et al., 2015; Georg et al., 2019; Zhou et al., 2019; Hou et al., 2020). The application of these algorithms for protein structure modeling is particularly challenging in the case of Bap because a global template that covers all individual domains is not available.

A second intriguing question is related with the presence of Bap in bovine mastitis isolates of *S. aureus* and their absence from human clinical isolates. Bap-mediated biofilm seems to be a system specialized for the conditions present in the

mammary gland, where calcium concentration can reach the high values necessary to modulate Bap function (~10 mM). For calcium to serve as a regulator of Bap function, the fluctuations in the local calcium concentration should be higher than the binding affinity of the protein for the cation. Otherwise, Bap-mediated biofilm development will be irreversible, and bacterial propagation from the site of infection will be seriously compromised. It is very likely that only in the mammary gland environment the concentration reaches the high values needed to inhibit the function of Bap. It will be important to determine whether Bap-mediated biofilm is a system specialized for the conditions present in the mammary gland. If so, it will represent an excellent model to explore the real contribution of adhesion and biofilm development during the course of the infection process.

AUTHOR CONTRIBUTIONS

IL and JV conceptualized the content of the article. IL, JV, and XF contributed to writing and revision of the work. All authors contributed to the article and approved the submitted version.

FUNDING

Work in the laboratories of JV and IL is supported by grants RTI2018-096011-B-I00 and BIO2017-83035-R from Spanish Ministry of Science, Innovation and Universities. Work in the laboratory of XF is supported by grants from the National Natural Science Foundation of China (No. 31872712) and the Beijing Advanced Innovation Center for Structural Biology.

REFERENCES

Arrizubieta, M. J., Toledo-Arana, A., Amorena, B., Penadés, J. R., and Lasa, I. (2004). Calcium inhibits bap-dependent multicellular behavior in *Staphylococcus aureus*. *J. Bacteriol.* 186, 7490–7498. doi: 10.1128/jb.186.22.7490-7498.2004

Bae, T., and Schneewind, O. (2003). The YSIRK-G/S motif of staphylococcal protein A and its role in efficiency of signal peptide processing. *J. Bacteriol.* 185, 2910–2919. doi: 10.1128/jb.185.9.2910-2919.2003

Baselga, R., Albizu, I., and Amorena, B. (1994). *Staphylococcus aureus* capsule and slime as virulence factors in ruminant mastitis. *Vet. Microbiol.* 39, 195–204. doi: 10.1016/0378-1135(94)90157-0

Baselga, R., Albizu, I., Cruz, M. D. L., Cacho, E. D., Barberan, M., and Amorena, B. (1993). Phase variation of slime production in *Staphylococcus aureus*: implications in colonization and virulence. *Infect. Immun.* 61, 4857–4862. doi: 10.1128/iai.61.11.4857-4862.1993

Callebaut, I., Mornon, J., Gilgès, D., and Vigon, I. (2000). HYR, an extracellular module involved in cellular adhesion and related to the immunoglobulin-like fold. *Protein Sci.* 9, 1382–1390. doi: 10.1110/ps.9.7.1382

Cheng, J., Xue, H., and Zhao, X. (2012). Variation of serine-aspartate repeats in membrane proteins possibly contributes to staphylococcal microevolution. *PLoS One* 7:e34756. doi: 10.1371/journal.pone.0034756

Conrady, D. G., Brescia, C. C., Horii, K., Weiss, A. A., Hassett, D. J., and Herr, A. B. (2008). A zinc-dependent adhesion module is responsible for intercellular adhesion in staphylococcal biofilms. *Proc. Natl. Acad. Sci.* 105, 19456–19461. doi: 10.1073/pnas.0807717105

Conrady, D. G., Wilson, J. J., and Herr, A. B. (2013). Structural basis for Zn^{2+}-dependent intercellular adhesion in staphylococcal biofilms. *Proc. Natl. Acad. Sci.* 110, E202–E211. doi: 10.1073/pnas.1208134110

Cucarella, C., Solano, C., Valle, J., Amorena, B., Lasa, I., and Penadés, J. R. (2001). Bap, a *Staphylococcus aureus* surface protein involved in biofilm formation. *J. Bacteriol.* 183, 2888–2896. doi: 10.1128/jb.183.9.2888-2896.2001

Cucarella, C., Tormo, M. A., Knecht, E., Amorena, B., Lasa, I., Foster, T. J., et al. (2002). Expression of the biofilm-associated protein interferes with host protein receptors of *Staphylococcus aureus* and alters the infective process. *Infect. Immun.* 70, 3180–3186. doi: 10.1128/iai.70.6.3180-3186.2002

Cucarella, C., Tormo, M. A., Ubeda, C., Trotonda, M. P., Monzón, M., Peris, C., et al. (2004). Role of biofilm-associated protein Bap in the pathogenesis of bovine *Staphylococcus aureus*. *Infect. Immun.* 72, 2177–2185. doi: 10.1128/iai.72.4.2177–2185.2004

Dedent, A., Bae, T., Missiakas, D. M., and Schneewind, O. (2008). Signal peptides direct surface proteins to two distinct envelope locations of *Staphylococcus aureus*. *EMBO J.* 27, 2656–2668. doi: 10.1038/emboj.2008.185

Dedent, A. C., McAdow, M., and Schneewind, O. (2007). Distribution of protein A on the surface of *Staphylococcus aureus*. *J. Bacteriol.* 189, 4473–4484. doi: 10.1128/jb.00227-07

Dziewanowska, K., Carson, A. R., Patti, J. M., Deobald, C. F., Bayles, K. W., and Bohach, G. A. (2000). Staphylococcal fibronectin binding protein interacts with heat shock protein 60 and integrins: role in internalization by epithelial cells. *Infect. Immun.* 68, 6321–6328. doi: 10.1128/iai.68.11.6321-6328.2000

Dziewanowska, K., Patti, J. M., Deobald, C. F., Bayles, K. W., Trumble, W. R., and Bohach, G. A. (1999). Fibronectin binding protein and host cell tyrosine kinase are required for internalization of *Staphylococcus aureus* by epithelial cells. *Infect. Immun.* 67, 4673–4678. doi: 10.1128/iai.67.9.4673-4678.1999

Fiumara, F., Fioriti, L., Kandel, E. R., and Hendrickson, W. A. (2010). Essential role of coiled coils for aggregation and activity of Q/N-rich prions and PolyQ proteins. *Cell* 143, 1121–1135. doi: 10.1016/j.cell.2010.11.042

Formosa-Dague, C., Speziale, P., Foster, T. J., Geoghegan, J. A., and Dufrêne, Y. F. (2015). Zinc-dependent mechanical properties of *Staphylococcus aureus* biofilm-forming surface protein SasG. *Proc. Natl. Acad. Sci.* 113, 410–415. doi: 10.1073/pnas.1519265113

Foster, T. J., Geoghegan, J. A., Ganesh, V. K., and Höök, M. (2014). Adhesion, invasion and evasion: the many functions of the surface proteins of *Staphylococcus aureus*. *Nat. Rev. Microbiol.* 12, 49–62. doi: 10.1038/nrmicro3161

Fowler, T., Wann, E. R., Joh, D., Johansson, S., Foster, T. J., and Höök, M. (2000). Cellular invasion by *Staphylococcus aureus* involves a fibronectin bridge between the bacterial fibronectic-binding MSCRAMMs and host cell β1 integrins. *Eur. J. Cell Biol.* 79, 672–679. doi: 10.1078/0171-9335-00104

Geoghegan, J. A., Corrigan, R. M., Gruszka, D. T., Speziale, P., O'gara, J. P., Potts, J. R., et al. (2010). Role of surface protein SasG in biofilm formation by *Staphylococcus aureus*. *J. Bacteriol.* 192, 5663–5673. doi: 10.1128/jb.00628-10

Georg, J., Lalaouna, D., Hou, S., Lott, S. C., Caldelari, I., Marzi, S., et al. (2019). The power of cooperation: experimental and computational approaches in the functional characterization of bacterial sRNAs. *Mol. Microbiol.* 113, 603–612. doi: 10.1111/mmi.14420

Griessl, M. H., Schmid, B., Kassler, K., Braunsmann, C., Ritter, R., Barlag, B., et al. (2013). Structural insight into the giant Ca^{2+}-binding adhesin siie: implications for the adhesion of *Salmonella enterica* to polarized epithelial cells. *Structure* 21, 741–752. doi: 10.1016/j.str.2013.02.020

Hartford, O., Francois, P., Vaudaux, P., and Foster, T. J. (1997). The dipeptide repeat region of the fibrinogen-binding protein (clumping factor) is required for functional expression of the fibrinogen-binding domain on the *Staphylococcus aureus* cell surface. *Mol. Microbiol.* 25, 1065–1076. doi: 10.1046/j.1365-2958.1997.5291896.x

Hazenbos, W. L. W., Kajihara, K. K., Vandlen, R., Morisaki, J. H., Lehar, S. M., Kwakkenbos, M. J., et al. (2013). Novel staphylococcal glycosyltransferases SdgA and SdgB mediate immunogenicity and protection of virulence-associated cell wall proteins. *PLoS Pathog.* 9:e1003653. doi: 10.1371/journal.ppat.1003653

Hou, J., Adhikari, B., Tanner, J. J., and Cheng, J. (2020). SAXSDom: modeling multidomain protein structures using small-angle X-ray scattering data. *Proteins* 88, 775–787. doi: 10.1002/prot.25865

Joh, D., Wann, E. R., Kreikemeyer, B., Speziale, P., and Hook, M. (1999). Role of fibronectin-binding MSCRAMMs in bacterial adherence and entry into mammalian cells. *Matrix Biol.* 18, 211–223. doi: 10.1016/s0945-053x(99)00025-6

Lasa, I., and Penadés, J. R. (2006). Bap: a family of surface proteins involved in biofilm formation. *Res. Microbiol.* 157, 99–107. doi: 10.1016/j.resmic.2005.11.003

Latasa, C., Roux, A., Toledo-Arana, A., Ghigo, J. -M., Gamazo, C., Penadés, J. R., et al. (2005). BapA, a large secreted protein required for biofilm formation and host colonization of *Salmonella enterica* serovar *Enteritidis*. *Mol. Microbiol.* 58, 1322–1339. doi: 10.1111/j.1365-2958.2005.04907.x

Martino, P. D. (2016). Bap: a new type of functional amyloid. *Trends Microbiol.* 24, 682–684. doi: 10.1016/j.tim.2016.07.004

Parry, D. A. D., Fraser, R. D. B., and Squire, J. M. (2008). Fifty years of coiled-coils and α-helical bundles: a close relationship between sequence and structure. *J. Struct. Biol.* 163, 258–269. doi: 10.1016/j.jsb.2008.01.016

Potter, A., Ceotto, H., Giambiagi-deMarval, M., Santos, K. R. N.dos, Nes, I. F., and Freire Bastos, M. D. C. (2009). The gene bap, involved in biofilm production, is present in *Staphylococcus* spp. strains from nosocomial infections. *J. Microbiol.* 47, 319–326. doi:10.1007/s12275-009-0008-y

Rosenstein, R., and Gotz, F. (2000). Staphylococcal lipases: biochemical and molecular characterization. *Biochimie* 82, 1005–1014. doi: 10.1016/s0300-9084(00)01180-9

Scherr, T. D., Heim, C. E., Morrison, J. M., and Kielian, T. (2014). Hiding in plain sight: interplay between staphylococcal biofilms and host immunity. *Front. Immunol.* 5:37. doi: 10.3389/fimmu.2014.00037

Schiffer, C., Hilgarth, M., Ehrmann, M., and Vogel, R. F. (2019). Bap and cell surface hydrophobicity are important factors in *Staphylococcus xylosus* biofilm formation. *Front. Microbiol.* 10:1387. doi: 10.3389/fmicb.2019.01387

Schilcher, K., and Horswill, A. R. (2020). Staphylococcal biofilm development: structure, regulation, and treatment strategies. *Microbiol. Mol. Biol. Rev.* 84, e00026–e00019. doi: 10.1128/mmbr.00026-19

Sinha, B., and Fraunholz, M. (2010). *Staphylococcus aureus* host cell invasion and post-invasion events. *Int. J. Med. Microbiol.* 300, 170–175. doi: 10.1016/j.ijmm.2009.08.019

Sivanathan, V., and Hochschild, A. (2013). A bacterial export system for generating extracellular amyloid aggregates. *Nat. Protoc.* 8, 1381–1390. doi: 10.1038/nprot.2013.081

Speziale, P., Pietrocola, G., Foster, T. J., and Geoghegan, J. A. (2014). Protein-based biofilm matrices in staphylococci. *Front. Cell. Infect. Microbiol.* 4:171. doi: 10.3389/fcimb.2014.00171

Strynadka, N. C., and James, M. N. (1989). Crystal structures of the helix-loop-helix calcium-binding proteins. *Annu. Rev. Biochem.* 58, 951–998. doi: 10.1146/annurev.bi.58.070189.004511

Taglialegna, A., Lasa, I., and Valle, J. (2016a). Amyloid structures as biofilm matrix scaffolds. *J. Bacteriol.* 198, 2579–2588. doi: 10.1128/jb.00122-16

Taglialegna, A., Matilla-Cuenca, L., Dorado-Morales, P., Navarro, S., Ventura, S., Garnett, J. A., et al. (2020). The biofilm-associated surface protein Esp of *Enterococcus faecalis* forms amyloid-like fibers. *NPJ Biofilms Microbiomes* 6, 15–12. doi: 10.1038/s41522-020-0125-2

Taglialegna, A., Navarro, S., Ventura, S., Garnett, J. A., Matthews, S., Penadés, J. R., et al. (2016b). Staphylococcal Bap proteins build amyloid scaffold biofilm matrices in response to environmental signals. *PLoS Pathog.* 12:e1005711. doi: 10.1371/journal.ppat.1005711

Tettelin, H., Nelson, K. E., Paulsen, I. T., Eisen, J. A., Read, T. D., Peterson, S., et al. (2001). Complete genome sequence of a virulent isolate of *Streptococcus pneumoniae*. *Science* 293, 498–506. doi: 10.1126/science.1061217

Ton-That, H., Mazmanian, S. K., Faull, K. F., and Schneewind, O. (2000). Anchoring of surface proteins to the cell wall of *Staphylococcus aureus*. Sortase catalyzed in vitro transpeptidation reaction using LPXTG peptide and NH(2)-Gly(3) substrates. *J. Biol. Chem.* 275, 9876–9881. doi: 10.1074/jbc.275.13.9876

Tormo, M. A., Knecht, E., Götz, F., Lasa, I., and Penadés, J. R. (2005). Bap-dependent biofilm formation by pathogenic species of *Staphylococcus*: evidence of horizontal gene transfer? *Microbiology* 151, 2465–2475. doi: 10.1099/mic.0.27865-0

Tursi, S. A., and Tükel, Ç. (2018). Curli-containing enteric biofilms inside and out: matrix composition, immune recognition, and disease implications. *Microbiol. Mol. Biol. Rev.* 82, e00028–e00018. doi: 10.1128/mmbr.00028-18

Ubeda, C., Tormo, M. A., Cucarella, C., Trotonda, P., Foster, T. J., Lasa, I., et al. (2003). Sip, an integrase protein with excision, circularization and integration activities, defines a new family of mobile *Staphylococcus aureus* pathogenicity islands. *Mol. Microbiol.* 49, 193–210. doi: 10.1046/j.1365-2958.2003.03577.x

Valle, J., Latasa, C., Gil, C., Toledo-Arana, A., Solano, C., Penadés, J. R., et al. (2012). Bap, a biofilm matrix protein of *Staphylococcus aureus* prevents cellular internalization through binding to GP96 host receptor. *PLoS Pathog.* 8:e1002843. doi: 10.1371/journal.ppat.1002843

Xu, D., Jaroszewski, L., Li, Z., and Godzik, A. (2015). AIDA: ab initio domain assembly for automated multi-domain protein structure prediction and domain-domain interaction prediction. *Bioinformatics* 31, 2098–2105. doi: 10.1093/bioinformatics/btv092

Zhou, X., Hu, J., Zhang, C., Zhang, G., and Zhang, Y. (2019). Assembling multidomain protein structures through analogous global structural alignments. *Proc. Natl. Acad. Sci.* 116, 15930–15938. doi: 10.1073/pnas.1905068116

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Microbiology

MINI REVIEW
published: 06 January 2021
doi: 10.3389/fmicb.2020.607679

NEAr Transporter (NEAT) Domains: Unique Surface Displayed Heme Chaperones That Enable Gram-Positive Bacteria to Capture Heme-Iron From Hemoglobin

Ken Ellis-Guardiola[1,2], *Brendan J. Mahoney*[1,2] *and Robert T. Clubb*[1,2,3]*

[1] *Department of Chemistry and Biochemistry, University of California, Los Angeles, Los Angeles, CA, United States,*
[2] *UCLA-DOE Institute for Genomics and Proteomics, University of California, Los Angeles, Los Angeles, CA, United States,*
[3] *Molecular Biology Institute, University of California, Los Angeles, Los Angeles, CA, United States*

OPEN ACCESS

Edited by:
Magnus Hook,
Texas A&M Health Science Center,
United States

Reviewed by:
Celia Goulding,
University of California, Irvine,
United States
Ann Smith,
University of Missouri–Kansas City,
United States
Anthony J. Clarke,
University of Guelph, Canada
Michael E. P. Murphy,
University of British Columbia,
Canada

***Correspondence:**
Robert T. Clubb
rclubb@mbi.ucla.edu

Specialty section:
This article was submitted to Infectious Diseases, a section of the journal Frontiers in Microbiology

Received: *17 September 2020*
Accepted: *03 December 2020*
Published: *06 January 2021*

Citation:
Ellis-Guardiola K, Mahoney BJ and Clubb RT (2021) NEAr Transporter (NEAT) Domains: Unique Surface Displayed Heme Chaperones That Enable Gram-Positive Bacteria to Capture Heme-Iron From Hemoglobin. Front. Microbiol. 11:607679. doi: 10.3389/fmicb.2020.607679

Iron is an important micronutrient that is required by bacteria to proliferate and to cause disease. Many bacterial pathogens forage iron from human hemoglobin (Hb) during infections, which contains this metal within heme (iron–protoporphyrin IX). Several clinically important pathogenic species within the Firmicutes phylum scavenge heme using surface-displayed or secreted NEAr Transporter (NEAT) domains. In this review, we discuss how these versatile proteins function in the *Staphylococcus aureus* Iron-regulated surface determinant system that scavenges heme-iron from Hb. *S. aureus* NEAT domains function as either Hb receptors or as heme-binding chaperones. *In vitro* studies have shown that heme-binding NEAT domains can rapidly exchange heme amongst one another via transiently forming transfer complexes, leading to the interesting hypothesis that they may form a protein-wire within the peptidoglycan layer through which heme flows from the microbial surface to the membrane. In Hb receptors, recent studies have revealed how dedicated heme- and Hb-binding NEAT domains function synergistically to extract Hb's heme molecules, and how receptor binding to the Hb-haptoglobin complex may block its clearance by macrophages, prolonging microbial access to Hb's iron. The functions of NEAT domains in other Gram-positive bacteria are also reviewed.

Keywords: *Staphylococcus aureus*, heme, hemoglobin, NEAr transporter domains, sortase, iron regulated surface determinant system, iron, pathogen

INTRODUCTION

Nearly all bacterial pathogens require iron to grow because it is an essential metal cofactor that is used by microbial enzymes to mediate cellular metabolism. Its power in biology stems from its ability to toggle between Fe^{II} (ferrous) and Fe^{III} (ferric) oxidation states, with this redox activity playing key roles in metabolic enzymes (oxidases, catalases, peroxidases), electron transfer (iron-sulfur proteins, cytochromes), and DNA synthesis (ribonucleotide reductases). During infections, many pathogens forage iron from human hemoglobin (Hb), as it contains ∼75–80% of the human body's total iron content in the form of heme (iron-protoporphyrin IX) (Parrow et al., 2013; Contreras et al., 2014; Wandersman and Delepelaire, 2014; Ma et al., 2015; Sheldon and Heinrichs, 2015; Choby and Skaar, 2016; Sheldon et al., 2016; Huang and Wilks, 2017; Conroy et al., 2019).

Bacteria gain access to Hb after it is released from senescent erythrocytes that spontaneously lyse or when erythrocytes are actively lysed by bacterial cytotoxins. Dedicated bacterial import systems then scavenge the heme and degrade it to release free iron. These systems have important roles in bacterial pathogenesis and are therefore potential targets for new antimicrobial agents. Microbial import is challenging because heme is lipophilic, prone to aggregation via non-specific interactions, and it can generate damaging reactive oxygen species that are toxic to the cell (Huang and Wilks, 2017; Donegan et al., 2019). Here we review how Gram-positive bacteria within the Firmicutes phylum acquire heme using surface displayed or secreted NEAr Transporter (NEAT) domains (Andrade et al., 2002; Honsa et al., 2014).

The function of NEAT domains in microbial heme scavenging is best understood for *Staphylococcus aureus*, which employs the Iron-regulated surface determinant (Isd) system to capture heme-iron from Hb (**Figure 1A**; Mazmanian et al., 2003; Nobles and Maresso, 2011; Sheldon and Heinrichs, 2015; Conroy et al., 2019). In the Isd system, four proteins containing NEAT domains are covalently attached to the cell wall (IsdA, IsdB, IsdC, and IsdH). IsdA and IsdC each possess a single NEAT domain that binds to heme with high affinity, while the IsdB and IsdH proteins contain multiple NEAT domains that bind to either heme (IsdB-N2, IsdH-N3) or Hb (IsdB-N1, IsdH-N1, IsdH-N2). The NEAT domain-containing Isd proteins are covalently attached to the cell wall by sortase enzymes, either by the housekeeping SrtA sortase (IsdA, IsdB, and IsdH), or by the SrtB sortase (IsdC) (Schneewind and Missiakas, 2012; Jacobitz et al., 2017). The Isd proteins are attached to the peptidoglycan at their C-termini via a peptide bond to the pentaglycine cell wall cross bridge. They are positioned at different depths within the cell wall as evidenced by their susceptibility to proteolytic degradation in whole cells. Their distinct locations are presumably dictated by differences in the number of amino acids that separate the last domain in each protein's primary sequence to their cell wall attached C-termini, with proteins containing longer segments being positioned farther from the membrane (Mazmanian et al., 2003). Positional differences may also depend upon the type of sortase enzyme that is used to attach the protein to the cell wall, since unlike SrtA, the SrtB sortase attaches IsdC to uncrosslinked glycan strands (Marraffini and Schneewind, 2005). Finally, the cell wall locations of the Isd proteins may also depend upon their N-terminal signal peptide sequences, as some cell wall attached Isd proteins contain a YSIRK/GS motif within their signal peptides that primes them for anchoring to the cross wall that forms at the site of cell division (DeDent et al., 2008).

The cell wall-attached NEAT domains are thought to have distinct roles in the process of heme scavenging from Hb (**Figure 1A**). IsdB and IsdH are located toward the cell periphery to capture Hb and remove its heme molecules. Heme then moves to the partially embedded IsdA protein, and then to IsdC, which is located closest to the membrane. IsdC is believed to perform the final step in heme transfer through the cell wall by unloading its heme to IsdE, which is a component of a ATP-binding cassette (ABC)-transporter complex that pumps heme into the cell. Based on their primary sequences, the IsdE and IsdF proteins are believed to function as the ligand binding and permease components of the transporter. The IsdD protein has also been suggested to be part of the transporter complex, but its function is unclear as it does not share significant sequence homology with any protein of known function. The nucleotide binding domain (NBD) component of the transporter that powers heme import has also not been characterized; however, it is possible that FhuC performs this task as it is a promiscuous NBD that has been proposed to function broadly in iron uptake (Beasley et al., 2009). After transport across the membrane, heme is either directly incorporated into bacterial proteins or degraded by the IsdG or IsdI heme oxygenases to release free iron (Wu et al., 2005). Because excess heme is toxic, *S. aureus* also employs a HrtAB complex to export heme, and heme homeostasis is maintained by coordinating heme biosynthesis, import, and export functions (Torres et al., 2007; Stauff and Skaar, 2009). The general features of the Isd heme uptake system are conserved amongst other Firmicutes bacteria, which also employ NEAT domains to acquire heme.

A number of excellent reviews have been written that describe how bacteria acquire heme and employ it as an iron source (Runyen-Janecky, 2013; Contreras et al., 2014; Sheldon and Heinrichs, 2015; Choby and Skaar, 2016; Huang and Wilks, 2017). In this mini-review, we discuss what is known about the structural basis of NEAT domain function in *S. aureus*. We highlight the unique ability of these domains to rapidly exchange heme amongst one another and suggest that it may enable them to form a protein wire within the cell wall through which heme rapidly flows from the microbial surface to the membrane. We also discuss how variations in their primary sequences allow some NEAT domains in *S. aureus* to function as Hb receptors that can rapidly strip heme from Hb, and we briefly review what is known about the functions of NEAT domains in other species of Gram-positive bacteria.

HEME-BINDING NEAT DOMAINS: NOVEL CHAPERONES THAT RAPIDLY TRANSFER HEME

Atomic structures of the heme-binding NEAT domains from *S. aureus* have been determined revealing a unique mode of ligand binding (IsdA, IsdB-N2, IsdC, and IsdH-N3) (**Figure 1B**; Grigg et al., 2007, 2011; Sharp et al., 2007; Villareal et al., 2008; Watanabe et al., 2008; Gaudin et al., 2011; Moriwaki et al., 2011; Vu et al., 2013). Structures of the IsdA- and IsdC-heme complexes were among the first to be determined, and revealed that NEAT domains adopt a conserved fold that binds heme within a hydrophobic cleft that is located at the end of its β-barrel structure (Grigg et al., 2007; Sharp et al., 2007). One face of heme's protoporphyrin ring lies flat on the surface formed by residues in strands β7 and β8, whereas the other face is contacted by a "lip region" that forms a 3_{10} helix. Subsequent studies showed that heme-binding NEAT domains typically harbor a conserved serine and YXXXY sequence motifs that are located within the lip region and strand β8, respectively (hereafter called the S/YXXXY motif) (Honsa et al., 2014). The tyrosine residues are located in

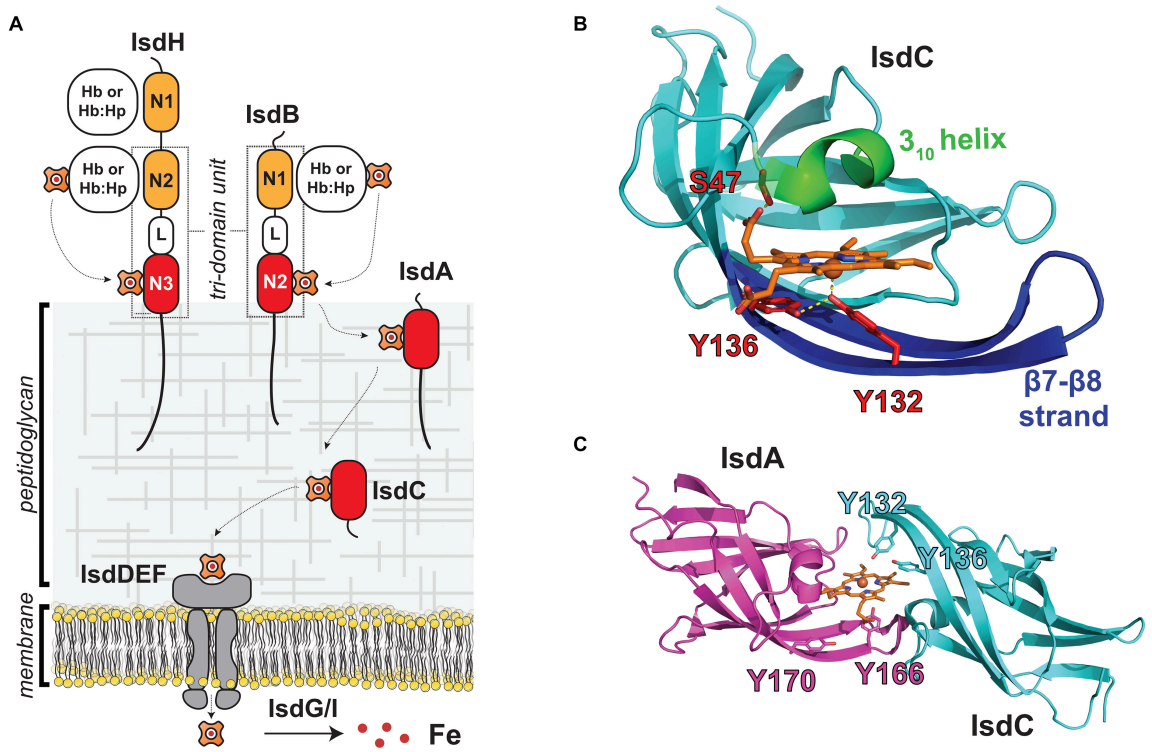


FIGURE 1 | NEAT domains act as heme chaperones on the surface of *S. aureus*. **(A)** Schematic of the *S. aureus* Isd heme acquisition system that uses NEAT domains to bind heme and Hb. The IsdB and IsdH proteins are Hb receptors and contain multiple NEAT domains that bind to either heme (red) or Hb (orange) and are bridged by a helical linker (L) domain. NEAT domains in the IsdA and IsdC proteins shuttle heme across the cell wall for import into the cell where it is degraded. Proteins containing NEAT domains are attached to the cell wall by SrtA (IsdA, IsdB, IsdC, and IsdH) or SrtB (IsdC) sortase enzymes. **(B)** Structure of the IsdC NEAT domain bound to pentacoordinate heme [PDB: 2O6P, (Sharp et al., 2007)]. Heme-binding residues in the S/YXXXY motif are labeled in red, the conserved 3_{10} helix is indicated in green, and the β7/β8 strands are shown in blue. Important hydrogen bonding and axial coordination interactions shown with yellow dashes **(C)** NEAT domains in IsdA [purple, PDB: 2ITF, (Grigg et al., 2007)] and IsdC [cyan, PDB: 2O6P, (Sharp et al., 2007)] transiently associate with one another to rapidly exchange heme. Model of the heme "handclasp" complex consistent with NMR, mass spectrometry and X-ray crystallography data. Other heme-binding NEAT domains also associate with one another to rapidly transfer. Their unique ability to rapidly exchange heme suggests that NEAT domains within the cell wall form a protein wire that transfers heme from the cell surface to the membrane.

strand β8 and have important functions in the pentacoordinate ligation of heme iron (**Figure 1B**, red). The first tyrosine directly coordinates the metal (primary tyrosine), while the second (secondary tyrosine) forms a hydrogen bond to the primary tyrosine to stabilize its positioning and its anionic phenolate state, which provides selectivity for the cationic ferric (Fe^{III}) form of heme (Moriwaki et al., 2011). The serine in the motif is located in the lip region (**Figure 1B**, green) on the opposite face and donates a hydrogen bond to heme's propionate group. A similar mode of heme-binding has been observed in the structures of NEAT domains from other species of bacteria (Honsa et al., 2014). All of the heme-binding NEAT domains in *S. aureus* contain a S/YXXXY motif, but some domains in other bacterial species lack the full complement of motif residues yet can bind heme using distinct axial ligands (Malmirchegini et al., 2014).

The NEAT domains in *S. aureus* have evolved the interesting ability to transiently associate with one another to transfer heme from one NEAT domain to another (**Figure 1C**; Liu et al., 2008; Villareal et al., 2011; Abe et al., 2012). Heme transfer between the domains occurs very rapidly, ∼70,000-times faster than the rate at which each domain spontaneously releases heme into the solvent. This novel ability led to the intriguing idea that NEAT domains function as heme chaperones within the cell wall, forming a protein wire through which heme is rapidly transferred from the cell surface to the membrane. Liu et al. (2008) were the first to discover that NEAT domains associate with one another to rapidly transfer heme by measuring the rate of heme transfer between the IsdA and IsdC domains using UV-Vis stopped-flow experiments. The extremely rapid kinetics of this process could only be explained by a mechanism in which heme is transferred via an IsdA-IsdC transfer complex. Subsequent studies demonstrated that other NEAT domains within the Isd system also associate with one another to form protein-protein complexes through which heme is rapidly transferred (Zhu et al., 2008b; Villareal et al., 2011; Abe et al., 2012). Transfer between the IsdA and IsdC NEAT domains occurs via an ultra-low affinity complex that forms fleetingly in solution (the IsdA-IsdC transfer complex forms with a K_D > ∼5 mM) (**Figure 1C**). NMR experiments revealed that the proteins transfer heme via a transiently forming pseudo-symmetrical "handclasp" complex

in which the 3_{10}-helix of the holo-IsdA donor associates with the loop connecting strands β7 and β8 (β7/β8 loop) in the apo-IsdC acceptor, and vice-versa (Villareal et al., 2011). This basic binding mode was later confirmed using photo-crosslinking and mass spectrometry experiments (Abe et al., 2012). Independently, using molecular docking and atomic structures of isolated NEAT domains, Grigg et al. (2011) also proposed that rapid transfer occurs via a "handclasp" complex. Molecular simulations have provided insight into the IsdA to IsdC heme transfer mechanism, which was modeled to occur via an intermediate in which the tyrosine ligands from each NEAT domain simultaneously form axial linkages to heme's iron atom such that it is hexacoordinate (Moriwaki et al., 2015). In the intermediate, the metal is coordinated by the primary and secondary tyrosines in the donor and acceptor, respectively, further explaining the functions of these residues in the S/YXXXY motif. Heme flow across the cell wall is driven by a thermodynamic gradient in *S. aureus*, as the heme-binding NEAT domains in the Isd-system exhibit progressively higher affinities for heme the more closely they are positioned to the membrane (Tiedemann and Stillman, 2012; Tiedemann et al., 2012); the affinity increases from IsdB (380 ± 60 nM) (Gaudin et al., 2011), to IsdH (34 ± 8 nM), to IsdA (14 ± 4 nM), to IsdC (6.5 ± 1.4 nM) (Moriwaki et al., 2013). The molecular origins of these affinity differences have been deduced using molecular dynamics simulations with theoretical free energy calculations and *in vitro* isothermal titration calorimetry experiments (Moriwaki et al., 2013).

In vitro experiments indicate that IsdC has a unique role in the Isd-system (**Figure 1A**), accepting heme from upstream NEAT domains in the network and then efficiently transferring it to the IsdE component of the IsdEF importer complex. Transfer to IsdE is thought to occur via a structurally unique IsdC-IsdE transfer complex, since unlike NEAT domains, IsdE adopts a bilobed structural topology in which Met and His residues are the axial ligands such that the iron atom is 6-coordinate (Pluym et al., 2007). Evidence for IsdC-IsdE complex formation comes from mass spectrometry studies that demonstrated that only IsdC efficiently transfers heme to IsdE (Tiedemann and Stillman, 2012) and kinetics measurements that revealed that transfer to IsdE occurs ∼4–10 times faster than the rate at which IsdC releases heme into the solvent (Zhu et al., 2008b). Later studies determined that IsdC transfers heme to IsdE at a rate 2.5 times slower ($k = 9.6 \times 10^{-3}\ \mu M^{-1}s^{-1}$) than IsdA to IsdC ($2.3 \times 10^{-2}\ \mu M^{-1}s^{-1}$), indicating that a buildup in IsdC-bound heme is expected at times of heme exposure (Tiedemann et al., 2012). Using crosslinking methods, Abe et al. (2012) obtained direct evidence for complex formation and discovered that a longer β7/β8 strand within IsdC is an important determinant for selective heme transfer to IsdE. IsdC's location within the cell wall is also presumably critical for its function because it is the only component within the heme transfer network that is attached to the cell wall by SrtB, which is thought to attach IsdC to uncrosslinked glycan chains near the cell membrane (Marraffini and Schneewind, 2005). Moreover, studies of IsdC protein from *Staphylococcus lugdunensis* have shown that it employs a specialized peptidoglycan hydrolase (IsdP, a N-acetylmuramoyl-l-alanine amidase) that affects anchoring of IsdC and consequently the ability of this microbe to use Hb as an iron source (Farrand et al., 2015). Close sequence homologs of IsdC are found in bacterial species within the class Bacilli and are frequently anchored to the cell wall by SrtB enzymes, suggesting that they have similar functions in transferring heme to membrane-associated transporters. *In vitro* evidence for transfer selectivity to ABC-transporter complexes has also been obtained for the IsdC-like proteins in *Streptococcus pyogenes* (called Shp) and *Bacillus anthracis* (Liu and Lei, 2005; Nygaard et al., 2006; Zhu et al., 2008a; Fabian et al., 2009; Balderas et al., 2012).

DO NEAT DOMAINS IN *S. aureus* FORM A PROTEIN WIRE THAT MOVES HEME THROUGH THE CELL WALL?

Based on their spatial positioning within the cell wall and their ability to rapidly transfer heme *in vitro*, it is tempting to speculate that the *S. aureus* NEAT domains don't simply bind and release heme, but instead physically interact with one another within the cell wall to rapidly transfer heme from the cell surface to the membrane. Why else would the domains have evolved the rare ability to associate and rapidly transfer heme? Although mechanistically attractive, the only direct evidence that NEAT domains interact with one another within the cell wall comes from a cell fractionation and pull-down experiment that demonstrated interactions between IsdA and IsdB (Pishchany et al., 2009). However, these studies need to be revisited, as strong interactions between these proteins have never been demonstrated *in vitro*. Given the presumed structural constraints imposed by the peptidoglycan, it is also unclear how NEAT domains could move within the cell wall to donate and accept heme by forming "handclasp" complexes (**Figure 1C**; Grigg et al., 2011; Villareal et al., 2011; Abe et al., 2012; Zajdowicz et al., 2012). However, some movement may be possible as the domains are connected to the cell wall by C-terminal polypeptide segments of varying lengths that are presumably unstructured, and the estimated size of the pores that permeate the peptidoglycan are larger than the diameter of the NEAT domains; the diameters of the NEAT domains are 40–50 Å, while pore diameters are estimated to be 50–500 Å depending on the degree of crosslinking and growth stage of the cell (Touhami et al., 2004; Meroueh et al., 2006). Furthermore, NEAT domains can also transfer heme via homotypic interactions (e.g., IsdC to IsdC heme transfer) (Abe et al., 2012), suggesting that a combination of only a few transfer events between like and unlike domains would be sufficient to move heme across the expanse of the cell wall that is 400–1000 Å thick (Geoghegan and Foster, 2017).

Hb RECEPTORS: NEAT DOMAINS CAN WORK TOGETHER TO BIND Hb AND EXTRACT ITS HEME

IsdB and IsdH (originally called HarA) contain unique Hb-binding NEAT domains that enable them to capture Hb on the

microbial surface (Dryla et al., 2003; Torres et al., 2006). In fact, they possess two types of functionally distinct NEAT domains: a canonical heme-binding NEAT domain at their C-termini and one or more Hb-binding domains (Pilpa et al., 2009). The first Hb-binding NEAT domain structures were determined by NMR, with crystal structures of the Hb-bound IsdH-N1 and IsdH-N2 proteins following (Pilpa et al., 2006; Dryla et al., 2007; Krishna Kumar et al., 2011; Dickson et al., 2014; Kumar et al., 2014). More recently, structures of the Hb-binding N1 domain of IsdB in both its apo and Hb-bound forms have been elucidated (Fonner et al., 2014; Bowden et al., 2018). Hb-binding domains are distinguished by the presence of a (F/Y)YH(Y/F) aromatic motif in the lip region and do not contain a S/YXXXY motif. Hb contains α- and β-globin chains that each bind to heme. Structures of isolated IsdB and IsdH NEAT domains free and bound to tetrameric Hb have revealed that residues within their aromatic motifs undergo a disordered-to-ordered transition upon binding, forming a 3_{10} helix that interacts with the A- and E-helices in the globin (Pilpa et al., 2006; Krishna Kumar et al., 2011; Kumar et al., 2014; Macdonald et al., 2019). Heme-binding NEAT domains feature a similarly located binding surface to engage heme. Although the α- and β-globin chains in Hb are structurally similar, subtle variations in the sequence of the aromatic motif enable the IsdH-N1 domain to selectively interact with α-globin, while the IsdH-N2 and IsdB-N1 domains are more promiscuous and bind to both the α- and β-globin subunits of Hb (Dickson et al., 2014). In addition, selectivity for the α chain can be engineered by introducing the appropriate amino acid substitutions into the aromatic motif (Dickson et al., 2015). Interestingly, the residues within the A-helix that form the interface with IsdB and IsdH are important determinants for defining *S. aureus*' host range. In addition, Hb residues at this interface have rapidly changed during primate evolution, consistent with it being the focal point for repeated evolutionary conflicts in the battle for iron during host-pathogen interactions (Pishchany et al., 2010; Choby et al., 2018).

The NEAT domains within the IsdB and IsdH Hb receptors work together to actively extract heme from Hb (Spirig et al., 2013). During infections, *S. aureus* encounters the oxidized (ferric) form of Hb (called methemoglobin, metHb), which spontaneously releases heme at a very slow rate (Hargrove et al., 1994). Spirig et al. (2013) discovered that both the IsdB and IsdH proteins contain a conserved tri-domain unit and demonstrated that in IsdH it actively removes heme from metHb. In IsdH, the tri-domain unit (called IsdH^{N2N3}) is formed by its second (IsdH-N2) and third (IsdH-N3) NEAT domains, which are joined by a helical linker (L) domain (**Figure 2A**; Spirig et al., 2013). The NEAT domains have distinct functions: IsdH-N2 binds Hb, while the C-terminal IsdH-N3 domain binds to heme (Pilpa et al., 2009). The domains function synergistically and must be part of the same polypeptide in order to effectively induce heme release from Hb, which occurs ∼1,250-fold faster than the rate at which Hb spontaneously releases heme into the solvent (Bowden et al., 2014; Sjodt et al., 2018). Crystal structures of the IsdH^{N2N3} complex with Hb reveal that it engages the globin chains within Hb via two distinct binding interfaces (**Figure 2A**; Dickson et al., 2014, 2015; Mikkelsen et al., 2020). The IsdH-N2 domain interacts with the A- and E-helices of the globin, while the linker and IsdH-N3 domains engage the distally positioned heme pocket on the globin via a Hb-LN3 interface through which heme is transferred. Interestingly, Hb's heme-contacting F-helix is distorted within the Hb-LN3 interface by two distinct subsites on the receptor that destabilize Hb-heme interactions and form a bridge through which heme moves from Hb to the IsdH-N3 domain (Ellis-Guardiola et al., 2020). NMR studies indicate that prior to engaging Hb, the tri-domain unit adopts an elongated and semi-flexible state in which the linker and IsdH-N3 form a rigid structure that reorient relative to the IsdH-N2 domain (Sjodt et al., 2016). Surprisingly, these inter-domain motions persist in the IsdH^{N2N3}:Hb complex, enabling the IsdH-N3 domain to transiently engage Hb to remove its heme while remaining tethered to Hb via the IsdH-N2 domain (Ellis-Guardiola et al., 2020). IsdB also extracts heme from Hb using an analogous tri-domain unit that distorts Hb's F-helix, which is consistent with IsdH and IsdB sharing significant primary sequence homology (64% sequence identity) (Zhu et al., 2008b; Bowden et al., 2014; **Figure 1A**). Notably, residues in IsdH's functionally important subsites are conserved in IsdB and both proteins bind and distort Hb in a similar manner, suggesting that they use a generally similar mechanism to extract heme (Fonner et al., 2014; Pishchany et al., 2014; Ellis-Guardiola et al., 2020). However, some differences have been observed, as IsdB is capable of productively extracting heme when its NEAT domains are combined in *trans,* while IsdH must have its full tri-domain unit present as an intact polypeptide to facilitate transfer (Spirig et al., 2013; Bowden et al., 2014). Excitingly, a recent structure of IsdB bound to Hb may have visualized a later step in the heme transfer reaction in which Hb's Fe^{III}-His87 bond is broken and the heme molecule is partially transferred to the receptor (Bowden et al., 2018). In the Hb-IsdB crystal structure, the native Hb heme coordination environment (axial coordination by His87 and H_2O) is disrupted and replaced by a hemichrome exhibiting bis-His metal coordination with the His58 and His89 imidazole rings that are positioned at the edge of Hb's heme-binding pocket. It is yet to be determined if this hemichrome represents a *bona fide* transfer intermediate or an off-pathway structure that was generated during the prolonged crystallization process.

Increasing evidence suggests that IsdH prolongs microbial access to iron during infections by slowing the rate at which Hb is removed from the blood (Saederup et al., 2016; Mikkelsen et al., 2020). In normal conditions, Hb released from erythrocytes as a result of cellular lysis is bound by haptoglobin (Hp), a highly abundant human glycoprotein that is present in blood plasma. The Hb:Hp complex is then removed from circulation by macrophages via CD163-receptor mediated endocytosis, which limits both the potential damaging effects caused by heme redox chemistry and microbial access to iron (Kristiansen et al., 2001). Unlike IsdB, IsdH elaborates its tri-domain heme extraction unit with an additional N-terminal IsdH-N1 Hb-binding domain (**Figure 1A**). Initial binding studies using commercially sourced Hp led to the erroneous conclusion that the IsdH-N1 domain could bind to both Hp and Hb (Dryla et al., 2003, 2007; Pilpa et al., 2009). However, more recent studies using purified Hb and Hp proteins have demonstrated that IsdH only interacts with

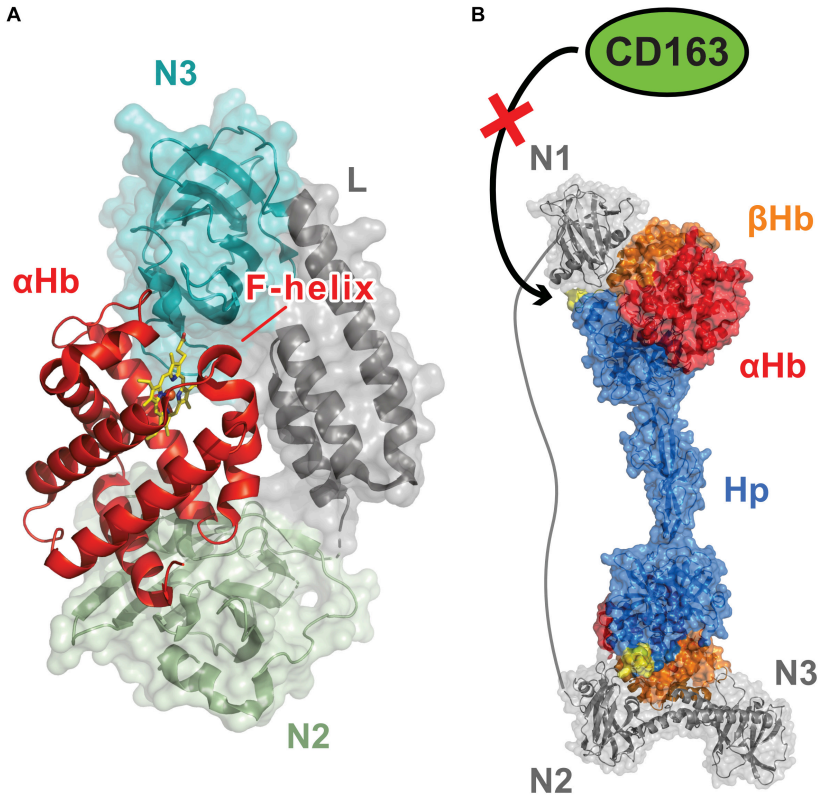


FIGURE 2 | Dedicated Hemoglobin (Hb)-binding NEAT domains enable *S. aureus* to capture Hb and extract its heme molecules. **(A)** Structure of the *S. aureus* IsdH protein bound to Hb [PDB: 4XS0, (Dickson et al., 2015)]. IsdH and IsdB are Hb receptors that extract heme using a conserved tri-domain unit (shown in **Figure 1A**). Each unit engages a single globin chain and contains a dedicated Hb-binding N-terminal NEAT domain that binds to the globin's A-helix and a C-terminal NEAT domain then removes the heme molecule via distortion of the Hb F-helix. The tri-domain unit in IsdH contains Hb-binding (N2, green) and heme-binding (N3, teal) NEAT domains, connected by a helical linker domain (L, gray). Synergistic interactions between the NEAT domains trigger heme release from Hb. The IsdB and IsdH tri-domain units bind to Hb in a similar manner and share related primary sequences. **(B)** Model showing how the multiple NEAT domains within the IsdH receptor may prevent haptoglobin (Hp) mediated removal of Hb from the blood, a nutritional immunity process that limits microbial access to iron [PDB: 4WJG, (Stodkilde et al., 2014) 6TB2, (Mikkelsen et al., 2020)] (Figure adapted from Mikkelsen et al., 2020). The tri-domain unit in IsdH is elaborated with a N-terminal Hb-binding NEAT domain (N1). Interactions between N1 and Hb's β-globin chain in the Hb:Hp complex are thought to disrupt its interactions with the macrophage surface CD163 receptor by occluding the CD163-binding loop of Hp (shown in yellow), thereby preventing the removal of the Hb:Hp complex by receptor mediated endocytosis (Nielsen et al., 2013). IsdB does not contain this accessory N-terminal Hb-binding domain.

Hb alone or within the Hb:Hp complex (Saederup et al., 2016). The results of these studies suggest that when the intact IsdH receptor captures the Hb:Hp complex on the surface of *S. aureus*, the IsdH-N1 domain may sterically occlude interactions between the complex and the CD163 receptor on macrophages (Nielsen et al., 2013). As a result, *S. aureus* has prolonged access to Hb in blood as macrophages are unable to bind and endocytose the Hb:Hp complex (**Figure 2B**). Indeed, this may be the primary function of the extra IsdH-N1 domain, as the intact IsdH receptor that contains the additional IsdH-N1 domain is less efficient at scavenging heme from Hb than IsdH's tri-domain unit. This is presumably because IsdH-N1 binding competes with the extraction unit for binding sites on the globin chains (Mikkelsen et al., 2020). Interestingly, both IsdB and IsdH can't efficiently extract heme from the Hp:Hb complex and recent results have shown that IsdH is unable to distort Hb's F-helix when it binds to the Hb:Hp complex (Bowden et al., 2018; Mikkelsen et al., 2020). This suggests that Hp not only removes Hb from the blood, it also actively prevents *S. aureus* from extracting its heme. *S. aureus* mutants lacking *isdB*, but not *isdH*, exhibit reduced virulence in a murine model of abscess formation (Torres et al., 2006). These differences would seem to suggest that in this particular animal model the ability of *S. aureus* to prolong access to Hb using the IsdH protein is less important for pathogenicity than IsdB's ability to strip heme from Hb.

NEAT DOMAINS IN OTHER SPECIES OF GRAM-POSITIVE BACTERIA

Genes encoding 343 putative NEAT domains have been identified in over 80 species of bacteria that are almost exclusively found in the Firmicutes phylum (Honsa et al., 2014). A detailed analysis of their sequences suggests that all of these domains will be associated with the cell exterior, either by membrane insertion, sortase-mediated covalent attachment to the cell wall,

or secretion, and that nearly half of them will bind to heme (∼48% contain all or most of the residues within the YXXXY motif) (Honsa et al., 2014). In addition to *S. aureus,* NEAT domains have been shown to have a role in heme acquisition in pathogenic *B. anthracis, B. cereus, L. monocytogenes,* and *S. pyogenes* (Bates et al., 2003; Newton et al., 2005; Maresso et al., 2008; Zhu et al., 2008a; Daou et al., 2009; Ouattara et al., 2010; Tarlovsky et al., 2010; Honsa et al., 2011; Balderas et al., 2012). In all instances, these domains function as surface-associated or secreted hemophores, binding heme via the conserved tyrosine linkage. Surprisingly, the genomes of some non-pathogenic soil-dwelling bacteria also encode for domains containing the S/YXXXY motif, which instead of binding heme have been proposed to be involved in chlorophyll uptake. Interestingly, instead of NEAT domains, high G+C Gram-positive bacteria (Actinobacteria) display structurally distinct heme-binding Conserved Region (CR) domains, but it remains unknown whether they function as heme chaperones that rapidly transfer heme.

The results of *in vitro* binding experiments have led to the conclusion that some NEAT domains from *B. anthracis* (IsdX1N and IsdX2N5) and *L. monocytogenes* (Hbp1N) bind to Hb, even though they lack a (F/Y)YH(Y/F) motif (Ekworomadu et al., 2012; Honsa et al., 2013; Malmirchegini et al., 2014). However, recent NMR studies have raised doubts about these findings, because the Hb used in these prior studies was obtained from commercial sources that are now known to contain breakdown products that can lead to erroneous conclusions about Hb binding and Hb-dependent microbial growth (Pishchany et al., 2013; Macdonald et al., 2019). Thus, at present, it appears that only NEAT domains containing a (F/Y)YH(Y/F) motif function as Hb receptors. As NEAT domains are located in extracellular proteins, many might be expected to bind to host proteins other than Hb, or to function as microbial surface components recognizing adhesive matrix molecules. This would seem to make sense, as NEAT domains adopt an immunoglobulin-like fold whose ability to bind to a range of molecules is well documented. Indeed, studies have documented IsdA NEAT domain binding to an array of human proteins (fetuin, asialofetuin, fibrinogen, fibronectin, loricrin, involucrin, and cytokeratin K10) and very recently data demonstrating IsdB interactions with the host protein vitronectin has been reported (Clarke et al., 2004; Pietrocola et al., 2020). However, at present, only structures of NEAT domains bound to either heme or Hb have been determined, so how they recognize these novel binding partners at a molecular level remains to be determined. Finally, as NEAT domains are also widely distributed in many species of non-pathogenic bacteria, these interesting proteins likely perform a range of other functions that have yet to be discovered.

AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct and intellectual contribution to the work, and approved it for publication.

FUNDING

We are thankful for funding support from the National Institutes of Health (AI52217), National Science Foundation (MCB-1716948), and the Department of Energy Office of Science, Office of Biological and Environmental Research Program under Award Number DE-FC02-02ER63421.

REFERENCES

Abe, R., Caaveiro, J. M., Kozuka-Hata, H., Oyama, M., and Tsumoto, K. (2012). Mapping ultra-weak protein-protein interactions between heme transporters of *Staphylococcus aureus*. *J. Biol. Chem.* 287, 16477–16487. doi: 10.1074/jbc.m112.346700

Andrade, M. A., Ciccarelli, F. D., Perez-Iratxeta, C., and Bork, P. (2002). NEAT: a domain duplicated in genes near the components of a putative Fe3+ siderophore transporter from Gram-positive pathogenic bacteria. *Genome Biol.* 3, research0047.1.

Balderas, M. A., Nobles, C. L., Honsa, E. S., Alicki, E. R., and Maresso, A. W. (2012). Hal Is a Bacillus anthracis heme acquisition protein. *J. Bacteriol.* 194, 5513–5521. doi: 10.1128/jb.00685-12

Bates, C. S., Montanez, G. E., Woods, C. R., Vincent, R. M., and Eichenbaum, Z. (2003). Identification and characterization of a *Streptococcus pyogenes* operon involved in binding of hemoproteins and acquisition of iron. *Infect. Immun.* 71, 1042–1055. doi: 10.1128/iai.71.3.1042-1055.2003

Beasley, F. C., Vines, E. D., Grigg, J. C., Zheng, Q., Liu, S., Lajoie, G. A., et al. (2009). Characterization of staphyloferrin A biosynthetic and transport mutants in *Staphylococcus aureus*. *Mol. Microbiol.* 72, 947–963. doi: 10.1111/j.1365-2958.2009.06698.x

Bowden, C. F., Verstraete, M. M., Eltis, L. D., and Murphy, M. E. (2014). Hemoglobin binding and catalytic heme extraction by IsdB near iron transporter domains. *Biochemistry* 53, 2286–2294. doi: 10.1021/bi500230f

Bowden, C. F. M., Chan, A. C. K., Li, E. J. W., Arrieta, A. L., Eltis, L. D., and Murphy, M. E. P. (2018). Structure-function analyses reveal key features in *Staphylococcus aureus* IsdB-associated unfolding of the heme-binding pocket of human hemoglobin. *J. Biol. Chem.* 293, 177–190. doi: 10.1074/jbc.m117.806562

Choby, J. E., Buechi, H. B., Farrand, A. J., Skaar, E. P., and Barber, M. F. (2018). Molecular basis for the evolution of species-specific hemoglobin capture by *Staphylococcus aureus*. *mBio* 9:e01524-18.

Choby, J. E., and Skaar, E. P. (2016). Heme synthesis and acquisition in bacterial pathogens. *J. Mol. Biol.* 428, 3408–3428. doi: 10.1016/j.jmb.2016.03.018

Clarke, S. R., Wiltshire, M. D., and Foster, S. J. (2004). IsdA of *Staphylococcus aureus* is a broad spectrum, iron-regulated adhesin. *Mol. Microbiol.* 51, 1509–1519. doi: 10.1111/j.1365-2958.2003.03938.x

Conroy, B. S., Grigg, J. C., Kolesnikov, M., Morales, L. D., and Murphy, M. E. P. (2019). Staphylococcus aureus heme and siderophore-iron acquisition pathways. *Biometals* 32, 409–424. doi: 10.1007/s10534-019-00188-2

Contreras, H., Chim, N., Credali, A., and Goulding, C. W. (2014). Heme uptake in bacterial pathogens. *Curr. Opin. Chem. Biol.* 19, 34–41. doi: 10.1016/j.cbpa.2013.12.014

Daou, N., Buisson, C., Gohar, M., Vidic, J., Bierne, H., Kallassy, M., et al. (2009). IlsA, a unique surface protein of *Bacillus cereus* required for iron acquisition from heme, hemoglobin and ferritin. *PLoS Pathog.* 5:e1000675. doi: 10.1371/journal.ppat.1000675

DeDent, A., Bae, T., Missiakas, D. M., and Schneewind, O. (2008). Signal peptides direct surface proteins to two distinct envelope locations of *Staphylococcus aureus*. *EMBO J.* 27, 2656–2668. doi: 10.1038/emboj.2008.185

Dickson, C. F., Jacques, D. A., Clubb, R. T., Guss, J. M., and Gell, D. A. (2015). The structure of haemoglobin bound to the haemoglobin receptor IsdH from Staphylococcus aureus shows disruption of the native alpha-globin haem pocket. *Acta Crystallogr. D Biol. Crystallogr.* 71, 1295–1306. doi: 10.1107/s1399004715005817

Dickson, C. F., Kumar, K. K., Jacques, D. A., Malmirchegini, G. R., Spirig, T., Mackay, J. P., et al. (2014). Structure of the hemoglobin-IsdH complex reveals

the molecular basis of iron capture by *Staphylococcus aureus*. *J. Biol. Chem.* 289, 6728–6738. doi: 10.1074/jbc.m113.545566

Donegan, R. K., Moore, C. M., Hanna, D. A., and Reddi, A. R. (2019). Handling heme: the mechanisms underlying the movement of heme within and between cells. *Free Radic. Biol. Med.* 133, 88–100. doi: 10.1016/j.freeradbiomed.2018.08.005

Dryla, A., Gelbmann, D., Von Gabain, A., and Nagy, E. (2003). Identification of a novel iron regulated staphylococcal surface protein with haptoglobin-haemoglobin binding activity. *Mol. Microbiol.* 49, 37–53. doi: 10.1046/j.1365-2958.2003.03542.x

Dryla, A., Hoffmann, B., Gelbmann, D., Giefing, C., Hanner, M., Meinke, A., et al. (2007). High-affinity binding of the staphylococcal HarA protein to haptoglobin and hemoglobin involves a domain with an antiparallel eight-stranded beta-barrel fold. *J. Bacteriol.* 189, 254–264. doi: 10.1128/jb.01366-06

Ekworomadu, M. T., Poor, C. B., Owens, C. P., Balderas, M. A., Fabian, M., Olson, J. S., et al. (2012). Differential function of lip residues in the mechanism and biology of an anthrax hemophore. *PLoS Pathog* 8:e1002559. doi: 10.1371/journal.ppat.1002559

Ellis-Guardiola, K., Clayton, J., Pham, C., Mahoney, B. J., Wereszczynski, J., and Clubb, R. T. (2020). The *Staphylococcus aureus* IsdH receptor forms a dynamic complex with human hemoglobin that triggers heme release via two distinct hot spots. *J. Mol. Biol.* 432, 1064–1082. doi: 10.1016/j.jmb.2019.12.023

Fabian, M., Solomaha, E., Olson, J. S., and Maresso, A. W. (2009). Heme transfer to the bacterial cell envelope occurs via a secreted hemophore in the Gram-positive pathogen *Bacillus anthracis*. *J. Biol. Chem.* 284, 32138–32146. doi: 10.1074/jbc.m109.040915

Farrand, A. J., Haley, K. P., Lareau, N. M., Heilbronner, S., Mclean, J. A., Foster, T., et al. (2015). An iron-regulated autolysin remodels the cell wall to facilitate heme acquisition in *Staphylococcus lugdunensis*. *Infect. Immun.* 83, 3578–3589. doi: 10.1128/iai.00397-15

Fonner, B. A., Tripet, B. P., Eilers, B. J., Stanisich, J., Sullivan-Springhetti, R. K., Moore, R., et al. (2014). Solution structure and molecular determinants of hemoglobin binding of the first NEAT domain of IsdB in *Staphylococcus aureus*. *Biochemistry* 53, 3922–3933. doi: 10.1021/bi5005188

Gaudin, C. F., Grigg, J. C., Arrieta, A. L., and Murphy, M. E. (2011). Unique heme-iron coordination by the hemoglobin receptor IsdB of Staphylococcus aureus. *Biochemistry* 50, 5443–5452. doi: 10.1021/bi200369p

Geoghegan, J. A., and Foster, T. J. (2017). Cell wall-anchored surface proteins of *Staphylococcus aureus*: many proteins, multiple functions. *Curr. Top. Microbiol. Immunol.* 409, 95–120. doi: 10.1007/82_2015_5002

Grigg, J. C., Mao, C. X., and Murphy, M. E. (2011). Iron-coordinating tyrosine is a key determinant of NEAT domain heme transfer. *J. Mol. Biol.* 413, 684–698. doi: 10.1016/j.jmb.2011.08.047

Grigg, J. C., Vermeiren, C. L., Heinrichs, D. E., and Murphy, M. E. (2007). Haem recognition by a *Staphylococcus aureus* NEAT domain. *Mol. Microbiol.* 63, 139–149. doi: 10.1111/j.1365-2958.2006.05502.x

Hargrove, M. S., Singleton, E. W., Quillin, M. L., Ortiz, L. A., Phillips, G. N. Jr., Olson, J. S., et al. (1994). His64(E7)–>Tyr apomyoglobin as a reagent for measuring rates of hemin dissociation. *J. Biol. Chem.* 269, 4207–4214.

Honsa, E. S., Fabian, M., Cardenas, A. M., Olson, J. S., and Maresso, A. W. (2011). The five near-iron transporter (NEAT) domain anthrax hemophore, IsdX2, scavenges heme from hemoglobin and transfers heme to the surface protein IsdC. *J. Biol. Chem.* 286, 33652–33660. doi: 10.1074/jbc.m111.241687

Honsa, E. S., Maresso, A. W., and Highlander, S. K. (2014). Molecular and evolutionary analysis of NEAr-iron Transporter (NEAT) domains. *PLoS One* 9:e104794. doi: 10.1371/journal.pone.0104794

Honsa, E. S., Owens, C. P., Goulding, C. W., and Maresso, A. W. (2013). The near-iron transporter (NEAT) domains of the anthrax hemophore IsdX2 require a critical glutamine to extract heme from methemoglobin. *J. Biol. Chem.* 288, 8479–8490. doi: 10.1074/jbc.m112.430009

Huang, W., and Wilks, A. (2017). Extracellular heme uptake and the challenge of bacterial cell membranes. *Annu. Rev. Biochem.* 86, 799–823. doi: 10.1146/annurev-biochem-060815-014214

Jacobitz, A. W., Kattke, M. D., Wereszczynski, J., and Clubb, R. T. (2017). Sortase transpeptidases: structural biology and catalytic mechanism. *Adv. Protein Chem. Struct. Biol.* 109, 223–264. doi: 10.1016/bs.apcsb.2017.04.008

Krishna Kumar, K., Jacques, D. A., Pishchany, G., Caradoc-Davies, T., Spirig, T., Malmirchegini, G. R., et al. (2011). Structural basis for hemoglobin capture by Staphylococcus aureus cell-surface protein, IsdH. *J. Biol. Chem.* 286, 38439–38447. doi: 10.1074/jbc.m111.287300

Kristiansen, M., Graversen, J. H., Jacobsen, C., Sonne, O., Hoffman, H. J., Law, S. K., et al. (2001). Identification of the haemoglobin scavenger receptor. *Nature* 409, 198–201. doi: 10.1038/35051594

Kumar, K. K., Jacques, D. A., Guss, J. M., and Gell, D. A. (2014). The structure of alpha-haemoglobin in complex with a haemoglobin-binding domain from Staphylococcus aureus reveals the elusive alpha-haemoglobin dimerization interface. *Acta Crystallogr. F Struct. Biol. Commun.* 70, 1032–1037. doi: 10.1107/s2053230x14012175

Liu, M., and Lei, B. (2005). Heme transfer from streptococcal cell surface protein Shp to HtsA of transporter HtsABC. *Infect. Immun.* 73, 5086–5092. doi: 10.1128/iai.73.8.5086-5092.2005

Liu, M., Tanaka, W. N., Zhu, H., Xie, G., Dooley, D. M., and Lei, B. (2008). Direct hemin transfer from IsdA to IsdC in the iron-regulated surface determinant (Isd) heme acquisition system of Staphylococcus aureus. *J. Biol. Chem.* 283, 6668–6676. doi: 10.1074/jbc.m708372200

Ma, L., Terwilliger, A., and Maresso, A. W. (2015). Iron and zinc exploitation during bacterial pathogenesis. *Metallomics* 7, 1541–1554. doi: 10.1039/c5mt00170f

Macdonald, R., Mahoney, B. J., Ellis-Guardiola, K., Maresso, A., and Clubb, R. T. (2019). NMR experiments redefine the hemoglobin binding properties of bacterial NEAr-iron Transporter domains. *Protein Sci.* 28, 1513–1523. doi: 10.1002/pro.3662

Malmirchegini, G. R., Sjodt, M., Shnitkind, S., Sawaya, M. R., Rosinski, J., Newton, S. M., et al. (2014). Novel mechanism of hemin capture by Hbp2, the hemoglobin-binding hemophore from Listeria monocytogenes. *J. Biol. Chem.* 289, 34886–34899. doi: 10.1074/jbc.m114.583013

Maresso, A. W., Garufi, G., and Schneewind, O. (2008). *Bacillus anthracis* secretes proteins that mediate heme acquisition from hemoglobin. *PLoS Pathog.* 4:e1000132. doi: 10.1371/journal.ppat.1000132

Marraffini, L. A., and Schneewind, O. (2005). Anchor structure of staphylococcal surface proteins. V. Anchor structure of the sortase B substrate IsdC. *J. Biol. Chem.* 280, 16263–16271. doi: 10.1074/jbc.m500071200

Mazmanian, S. K., Skaar, E. P., Gaspar, A. H., Humayun, M., Gornicki, P., Jelenska, J., et al. (2003). Passage of heme-iron across the envelope of *Staphylococcus aureus*. *Science* 299, 906–909. doi: 10.1126/science.1081147

Meroueh, S. O., Bencze, K. Z., Hesek, D., Lee, M., Fisher, J. F., Stemmler, T. L., et al. (2006). Three-dimensional structure of the bacterial cell wall peptidoglycan. *Proc. Natl. Acad. Sci. U.S.A.* 103, 4404–4409. doi: 10.1073/pnas.0510182103

Mikkelsen, J. H., Runager, K., and Andersen, C. B. F. (2020). The human protein haptoglobin inhibits IsdH-mediated heme-sequestering by *Staphylococcus aureus*. *J. Biol. Chem.* 295, 1781–1791. doi: 10.1074/jbc.ra119.011612

Moriwaki, Y., Caaveiro, J. M., Tanaka, Y., Tsutsumi, H., Hamachi, I., and Tsumoto, K. (2011). Molecular basis of recognition of antibacterial porphyrins by heme-transporter IsdH-NEAT3 of Staphylococcus aureus. *Biochemistry* 50, 7311–7320. doi: 10.1021/bi200493h

Moriwaki, Y., Terada, T., Caaveiro, J. M., Takaoka, Y., Hamachi, I., Tsumoto, K., et al. (2013). Heme binding mechanism of structurally similar iron-regulated surface determinant near transporter domains of Staphylococcus aureus exhibiting different affinities for heme. *Biochemistry* 52, 8866–8877. doi: 10.1021/bi4008325

Moriwaki, Y., Terada, T., Tsumoto, K., and Shimizu, K. (2015). Rapid heme transfer reactions between NEAr transporter domains of *Staphylococcus aureus*: a theoretical study using QM/MM and MD simulations. *PLoS One* 10:e0145125. doi: 10.1371/journal.pone.0145125

Newton, S. M., Klebba, P. E., Raynaud, C., Shao, Y., Jiang, X., Dubail, I., et al. (2005). The svpA-srtB locus of Listeria monocytogenes: fur-mediated iron regulation and effect on virulence. *Mol. Microbiol.* 55, 927–940. doi: 10.1111/j.1365-2958.2004.04436.x

Nielsen, M. J., Andersen, C. B., and Moestrup, S. K. (2013). CD163 binding to haptoglobin-hemoglobin complexes involves a dual-point electrostatic receptor-ligand pairing. *J. Biol. Chem.* 288, 18834–18841. doi: 10.1074/jbc.m113.471060

Nobles, C. L., and Maresso, A. W. (2011). The theft of host heme by Gram-positive pathogenic bacteria. *Metallomics* 3, 788–796. doi: 10.1039/c1mt00047k

Nygaard, T. K., Blouin, G. C., Liu, M., Fukumura, M., Olson, J. S., Fabian, M., et al. (2006). The mechanism of direct heme transfer from the streptococcal

cell surface protein Shp to HtsA of the HtsABC transporter. *J. Biol. Chem.* 281, 20761–20771. doi: 10.1074/jbc.m601832200

Ouattara, M., Cunha, E. B., Li, X., Huang, Y. S., Dixon, D., and Eichenbaum, Z. (2010). Shr of group A streptococcus is a new type of composite NEAT protein involved in sequestering haem from methaemoglobin. *Mol. Microbiol.* 78, 739–756. doi: 10.1111/j.1365-2958.2010.07367.x

Parrow, N. L., Fleming, R. E., and Minnick, M. F. (2013). Sequestration and scavenging of iron in infection. *Infect. Immun.* 81, 3503–3514. doi: 10.1128/iai.00602-13

Pietrocola, G., Pellegrini, A., Alfeo, M. J., Marchese, L., Foster, T. J., and Speziale, P. (2020). The iron-regulated surface determinant B (IsdB) protein from *Staphylococcus aureus* acts as a receptor for the host protein vitronectin. *J. Biol. Chem.* 295, 10008–10022. doi: 10.1074/jbc.ra120.013510

Pilpa, R. M., Fadeev, E. A., Villareal, V. A., Wong, M. L., Phillips, M., and Clubb, R. T. (2006). Solution structure of the NEAT (NEAr Transporter) domain from IsdH/HarA: the human hemoglobin receptor in *Staphylococcus aureus*. *J. Mol. Biol.* 360, 435–447. doi: 10.1016/j.jmb.2006.05.019

Pilpa, R. M., Robson, S. A., Villareal, V. A., Wong, M. L., Phillips, M., and Clubb, R. T. (2009). Functionally distinct NEAT (NEAr Transporter) domains within the *Staphylococcus aureus* IsdH/HarA protein extract heme from methemoglobin. *J. Biol. Chem.* 284, 1166–1176. doi: 10.1074/jbc.m806007200

Pishchany, G., Dickey, S. E., and Skaar, E. P. (2009). Subcellular localization of the *Staphylococcus aureus* heme iron transport components IsdA and IsdB. *Infect. Immun.* 77, 2624–2634. doi: 10.1128/iai.01531-08

Pishchany, G., Haley, K. P., and Skaar, E. P. (2013). *Staphylococcus aureus* growth using human hemoglobin as an iron source. *JoVE* 72:e50072.

Pishchany, G., Mccoy, A. L., Torres, V. J., Krause, J. C., Crowe, J. E. Jr., Fabry, M. E., et al. (2010). Specificity for human hemoglobin enhances *Staphylococcus aureus* infection. *Cell Host Microbe* 8, 544–550. doi: 10.1016/j.chom.2010.11.002

Pishchany, G., Sheldon, J. R., Dickson, C. F., Alam, M. T., Read, T. D., Gell, D. A., et al. (2014). IsdB-dependent hemoglobin binding is required for acquisition of heme by *Staphylococcus aureus*. *J. Infect. Dis.* 209, 1764–1772. doi: 10.1093/infdis/jit817

Pluym, M., Vermeiren, C. L., Mack, J., Heinrichs, D. E., and Stillman, M. J. (2007). Heme binding properties of *Staphylococcus aureus* IsdE. *Biochemistry* 46, 12777–12787.

Runyen-Janecky, L. J. (2013). Role and regulation of heme iron acquisition in gram-negative pathogens. *Front. Cell Infect. Microbiol.* 3:55. doi: 10.3389/fcimb.2013.00055

Saederup, K. L., Stodkilde, K., Graversen, J. H., Dickson, C. F., Etzerodt, A., Hansen, S. W., et al. (2016). The *Staphylococcus aureus* protein IsdH inhibits host hemoglobin scavenging to promote heme acquisition by the pathogen. *J. Biol. Chem.* 291, 23989–23998. doi: 10.1074/jbc.m116.755934

Schneewind, O., and Missiakas, D. M. (2012). Protein secretion and surface display in Gram-positive bacteria. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 367, 1123–1139. doi: 10.1098/rstb.2011.0210

Sharp, K. H., Schneider, S., Cockayne, A., and Paoli, M. (2007). Crystal structure of the heme-IsdC complex, the central conduit of the Isd iron/heme uptake system in *Staphylococcus aureus*. *J. Biol. Chem.* 282, 10625–10631. doi: 10.1074/jbc.m700234200

Sheldon, J. R., and Heinrichs, D. E. (2015). Recent developments in understanding the iron acquisition strategies of gram positive pathogens. *FEMS Microbiol. Rev.* 39, 592–630. doi: 10.1093/femsre/fuv009

Sheldon, J. R., Laakso, H. A., and Heinrichs, D. E. (2016). Iron acquisition strategies of bacterial pathogens. *Microbiol. Spectr.* 4:2.

Sjodt, M., Macdonald, R., Marshall, J. D., Clayton, J., Olson, J. S., Phillips, M., et al. (2018). Energetics underlying hemin extraction from human hemoglobin by *Staphylococcus aureus*. *J. Biol. Chem.* 293, 6942–6957. doi: 10.1074/jbc.ra117.000803

Sjodt, M., Macdonald, R., Spirig, T., Chan, A. H., Dickson, C. F., Fabian, M., et al. (2016). The PRE-derived NMR model of the 38.8-kDa Tri-Domain IsdH protein from *Staphylococcus aureus* suggests that it adaptively recognizes human hemoglobin. *J. Mol. Biol.* 428, 1107–1129. doi: 10.1016/j.jmb.2015.02.008

Spirig, T., Malmirchegini, G. R., Zhang, J., Robson, S. A., Sjodt, M., Liu, M., et al. (2013). *Staphylococcus aureus* uses a novel multidomain receptor to break apart human hemoglobin and steal its heme. *J. Biol. Chem.* 288, 1065–1078. doi: 10.1074/jbc.m112.419119

Stauff, D. L., and Skaar, E. P. (2009). The heme sensor system of *Staphylococcus aureus*. *Contrib. Microbiol.* 16, 120–135. doi: 10.1159/000219376

Stodkilde, K., Torvund-Jensen, M., Moestrup, S. K., and Andersen, C. B. (2014). Structural basis for trypanosomal haem acquisition and susceptibility to the host innate immune system. *Nat. Commun.* 5:5487.

Tarlovsky, Y., Fabian, M., Solomaha, E., Honsa, E., Olson, J. S., and Maresso, A. W. (2010). A *Bacillus anthracis* S-layer homology protein that binds heme and mediates heme delivery to IsdC. *J. Bacteriol.* 192, 3503–3511. doi: 10.1128/jb.00054-10

Tiedemann, M. T., Heinrichs, D. E., and Stillman, M. J. (2012). Multiprotein heme shuttle pathway in Staphylococcus aureus: iron-regulated surface determinant cog-wheel kinetics. *J. Am. Chem. Soc.* 134, 16578–16585. doi: 10.1021/ja305115y

Tiedemann, M. T., and Stillman, M. J. (2012). Heme binding to the IsdE(M78A; H229A) double mutant: challenging unidirectional heme transfer in the iron-regulated surface determinant protein heme transfer pathway of *Staphylococcus aureus*. *J. Biol. Inorg. Chem.* 17, 995–1007. doi: 10.1007/s00775-012-0914-z

Torres, V. J., Pishchany, G., Humayun, M., Schneewind, O., and Skaar, E. P. (2006). *Staphylococcus aureus* IsdB is a hemoglobin receptor required for heme iron utilization. *J. Bacteriol.* 188, 8421–8429. doi: 10.1128/jb.01335-06

Torres, V. J., Stauff, D. L., Pishchany, G., Bezbradica, J. S., Gordy, L. E., Iturregui, J., et al. (2007). A *Staphylococcus aureus* regulatory system that responds to host heme and modulates virulence. *Cell Host Microbe* 1, 109–119. doi: 10.1016/j.chom.2007.03.001

Touhami, A., Jericho, M. H., and Beveridge, T. J. (2004). Atomic force microscopy of cell growth and division in *Staphylococcus aureus*. *J. Bacteriol.* 186, 3286–3295. doi: 10.1128/jb.186.11.3286-3295.2004

Villareal, V. A., Pilpa, R. M., Robson, S. A., Fadeev, E. A., and Clubb, R. T. (2008). The IsdC protein from Staphylococcus aureus uses a flexible binding pocket to capture heme. *J. Biol. Chem.* 283, 31591–31600. doi: 10.1074/jbc.m801126200

Villareal, V. A., Spirig, T., Robson, S. A., Liu, M., Lei, B., and Clubb, R. T. (2011). Transient weak protein-protein complexes transfer heme across the cell wall of *Staphylococcus aureus*. *J. Am. Chem. Soc.* 133, 14176–14179. doi: 10.1021/ja203805b

Vu, N. T., Moriwaki, Y., Caaveiro, J. M., Terada, T., Tsutsumi, H., Hamachi, I., et al. (2013). Selective binding of antimicrobial porphyrins to the heme-receptor IsdH-NEAT3 of Staphylococcus aureus. *Protein Sci.* 22, 942–953. doi: 10.1002/pro.2276

Wandersman, C., and Delepelaire, P. (2014). Heme-delivering proteins in bacteria. *Handb. Porphyrin Sci.* 26, 191–222. doi: 10.1142/9789814407755_0004

Watanabe, M., Tanaka, Y., Suenaga, A., Kuroda, M., Yao, M., Watanabe, N., et al. (2008). Structural basis for multimeric heme complexation through a specific protein-heme interaction: the case of the third neat domain of IsdH from *Staphylococcus aureus*. *J. Biol. Chem.* 283, 28649–28659. doi: 10.1074/jbc.m803383200

Wu, R., Skaar, E. P., Zhang, R., Joachimiak, G., Gornicki, P., Schneewind, O., et al. (2005). *Staphylococcus aureus* IsdG and IsdI, heme-degrading enzymes with structural similarity to monooxygenases. *J. Biol. Chem.* 280, 2840–2846. doi: 10.1074/jbc.m409526200

Zajdowicz, S., Haller, J. C., Krafft, A. E., Hunsucker, S. W., Mant, C. T., Duncan, M. W., et al. (2012). Purification and structural characterization of siderophore (corynebactin) from Corynebacterium diphtheriae. *PLoS One* 7:e34591. doi: 10.1371/journal.pone.0034591

Zhu, H., Liu, M., and Lei, B. (2008a). The surface protein Shr of *Streptococcus pyogenes* binds heme and transfers it to the streptococcal heme-binding protein Shp. *BMC Microbiol.* 8:15. doi: 10.1186/1471-2180-8-15

Zhu, H., Xie, G., Liu, M., Olson, J. S., Fabian, M., Dooley, D. M., et al. (2008b). Pathway for heme uptake from human methemoglobin by the iron-regulated surface determinants system of *Staphylococcus aureus*. *J. Biol. Chem.* 283, 18450–18460. doi: 10.1074/jbc.m801466200

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Microbiology

REVIEW
published: 05 February 2021
doi: 10.3389/fmicb.2021.628798

Collagen Binding Proteins of Gram-Positive Pathogens

Srishtee Arora, Jay Gordon and Magnus Hook**

Center for Infectious and Inflammatory Diseases, Institute of Biosciences and Technology, Texas A&M Health Science Center, Houston, TX, United States

Collagens are the primary structural components of mammalian extracellular matrices. In addition, collagens regulate tissue development, regeneration and host defense through interaction with specific cellular receptors. Their unique triple helix structure, which requires a glycine residue every third amino acid, is the defining structural feature of collagens. There are 28 genetically distinct collagens in humans. In addition, several other unrelated human proteins contain a collagen domain. Gram-positive bacteria of the genera *Staphylococcus*, *Streptococcus*, *Enterococcus*, and *Bacillus* express cell surface proteins that bind to collagen. These proteins of Gram-positive pathogens are modular proteins that can be classified into different structural families. This review will focus on the different structural families of collagen binding proteins of Gram-positive pathogen. We will describe how these proteins interact with the triple helix in collagens and other host proteins containing a collagenous domain and discuss how these interactions can contribute to the pathogenic processes.

Keywords: Gram-positive bacteria, collagen binding proteins, collagen-like proteins, surface proteins, collagen

OPEN ACCESS

Edited by:
Mattias Collin,
Lund University, Sweden

Reviewed by:
Minh-Thu Nguyen,
University Hospital Münster, Germany
Jack Christopher Leo,
Nottingham Trent University, United Kingdom

****Correspondence:***
Magnus Hook
mhook@tamu.edu
Srishtee Arora
sarora@tamu.edu

Specialty section:
This article was submitted to Infectious Diseases, a section of the journal Frontiers in Microbiology

Received: *12 November 2020*
Accepted: *11 January 2021*
Published: *05 February 2021*

Citation:
Arora S, Gordon J and Hook M (2021) Collagen Binding Proteins of Gram-Positive Pathogens. Front. Microbiol. 12:628798. doi: 10.3389/fmicb.2021.628798

INTRODUCTION

Collagen is the most abundant protein in the human body and an integral component of the extracellular matrix (ECM) (Shoulders and Raines, 2009). The ECM is a complex proteinaceous network that provides structural support to tissues along with the necessary signaling for cell adhesion, migration, and growth as well as for tissue development and regeneration (Frantz et al., 2010). Collagen plays a critical role in the functional integrity of most tissues including bone, skin, tendon, and cartilage (Burgeson and Nimni, 1992; Frantz et al., 2010). Collagen can also be the target of surface-anchored adhesins and other virulence factors produced by both Gram-positive and Gram-negative pathogens (Harrington, 1996; Singh et al., 2012; Zhang et al., 2015; Duarte et al., 2016; Paulsson and Riesbeck, 2018; Vaca et al., 2020). Of these, the cell wall anchored collagen binding proteins in Gram-positive bacteria have been more extensively studied, and will be reviewed here.

There are 28 identified types of collagens in humans (**Table 1**; Ricard-Blum, 2011). Each collagen molecule is formed through the interactions of three protein polypeptides known as α-strands. The α-strands come together to form a canonical right-handed triple helical structure termed the triple helix domain (Kadler et al., 2007; Ricard-Blum, 2011). Triple helices can be formed by association of identical α-strands to form a homotrimer or be composed of different α-strands (heterotrimer) (Ricard-Blum, 2011). The triple helix domain is a flexible rod-shaped structure held together through inter-chain hydrogen bonding (Kadler et al., 2007; Shoulders and Raines, 2009;

Ricard-Blum, 2011). The triple helix is defined by Gly-X-X' amino acid repeats with X and X' commonly representing proline and 4-hydroxyproline, respectively (Shoulders and Raines, 2009). Glycine residues are required every 3rd residue as any other residue would result in steric hindrance and helix destabilization (Theocharis et al., 2016). Collagens also have non-triple helical domains at their N- and C-termini, which are referred here as "non-collagenous" domains. In addition to the conventional collagens, several other mammalian proteins contain collagenous domains (Fraser and Tenner, 2008; Zani et al., 2015; PrabhuDas et al., 2017; Casals et al., 2019).

Bacterial surface proteins contribute to pathogenic processes and play a critical role in mediating adhesion to host cells and tissues, enabling colonization, invasion, and biofilm formation (Foster et al., 2014; Foster, 2019). In addition, binding of bacterial surface proteins to host ligands can lead to evasion of the host defense systems (Foster et al., 2014; Foster, 2019). In Gram-positive bacteria, different classes of surface proteins exist: (1) lipoproteins, (2) proteins covalently anchored to the cell wall, (3) pilus proteins, (4) non-covalently surface-associated proteins, and (5) transmembrane proteins (Desvaux et al., 2006; Fischetti, 2019). Lipoproteins are proteins covalently attached to membrane lipids via their N-terminus (Desvaux et al., 2006). Cell wall anchored proteins and pilus proteins are anchored to the cell wall by the action of enzymes called sortases (Desvaux et al., 2006). Sortases mediate covalent linking of proteins to the peptidoglycan through a transpeptidase reaction, and can also enable assembly of surface pilus and anchor the pilus onto the peptidoglycan layer (Ton-That and Schneewind, 2004; Desvaux et al., 2006; Fischetti, 2019). Lastly, non-covalently surface associated proteins contain cell wall binding domains (Desvaux et al., 2006; Fischetti, 2019).

Bacterial surface proteins are modular multi-domain proteins that can often be grouped into structural families based on their structural similarities. Multiple structurally related families of proteins have been identified in the literature (Waldemarsson et al., 2006; Foster et al., 2014; Frost et al., 2017; Foster, 2019; Taglialegna et al., 2020). Notables examples of structural families in Gram-positive bacteria include the MSCRAMMs (microbial surface components recognizing adhesive matrix molecules) (Foster et al., 2014; Foster, 2019), serine-rich repeat proteins (Lizcano et al., 2012), and M-proteins (Fischetti, 2016). In this review, we will describe collagen binding proteins present on the surface of Gram-positive pathogens that are human pathogens. This review will focus on structural families where more than one protein with structural similarity has been reported to bind collagen directly. Some proteins reported in the literature use fibronectin as a bridging molecule to bind collagen, e.g., streptococcal fibronectin binding protein 1 (SfbI) of *Streptococcus pyogenes* (Dinkla et al., 2003a) and are not covered here.

TYPES OF COLLAGEN

Collagens can be divided into different categories, which include fibrillar collagen, network forming collagen, FACITs (fibril-associated collagens with interrupted triple helices), MACITs (membrane-associated collagens with interrupted triple helices), anchoring fibrils, beaded-filament-forming collagens, and MULTIPLEXIN (multiple triple-helix domains and interruptions). These major classes of collagen will be discussed briefly below (Ricard-Blum, 2011; Theocharis et al., 2016). Collagen structure, chain composition, tissue distribution and functions are listed in **Table 1**.

Fibrillar collagen is the most common type of collagen in humans. Collagen types I, II, III, V, XI, XXIV, and XXVII all have a fibrillary configuration (Kadler et al., 2007; Theocharis et al., 2016). During fibrillogenesis, the protocollagen strands assemble into a triple helical formation called procollagen triple helix, which undergoes cleavage of N- and C-termini to generate tropocollagen triple helix molecule (Kadler et al., 2007; Shoulders and Raines, 2009; Theocharis et al., 2016). Tropocollagen triple helix molecule self-assembles into a D-staggered arrangement with a 67 nm periodicity to form collagen microfibrils [for further details on D-staggering, see Boudko and Bächinger (2016), Kadler (2017), Holmes et al. (2018)]. As a last step, collagen fibrils of diameter 15–500 nm are formed by crosslinking of collagen microfibrils (Kadler et al., 2007; Shoulders and Raines, 2009; Theocharis et al., 2016). Collagen fibrils in turn participate in forming larger structures such as ligaments and tendons (Ricard-Blum, 2011).

Network forming collagens include collagen types IV, VIII, and X, with collagen type IV being the archetype (Theocharis et al., 2016). Collagen type IV is found in the basement membrane along with other molecules such as laminin (Hohenester and Yurchenco, 2013; Theocharis et al., 2016). Unlike fibrillary collagens, the non-collagenous domains of these molecules are not cleaved and are utilized to form tail to tail interactions with other non-collagenous domains of collagen (Sundaramoorthy et al., 2002). Stabilizing tetramers are also formed via N-terminal head to head interactions. Once a mature network is formed, these collagens work to support the surrounding epithelial cell layer (Kadler et al., 2007).

Fibril-associated collagens with interrupted triple helices are relatively short flexible collagens that contain small triple helical regions interrupted by non-collagenous domains (Theocharis et al., 2016). Collagen types IX, XII, XIV, XVI, XIX, XX, XXI, and XXII have been reported as FACITs (Shoulders and Raines, 2009). Their primary role is to connect other collagen types together as well as with various ECM components (Theocharis et al., 2016). Collagen type IX, an archetypal FACIT, is covalently linked to collagen type II present in cartilage (Kadler et al., 2007; Ricard-Blum, 2011; Theocharis et al., 2016) and collagen type XIV binds to type I (Kadler et al., 2007; Ricard-Blum, 2011; Theocharis et al., 2016).

Membrane-associated collagens with interrupted triple helices are transmembrane proteins and contain a short N-terminal cytoplasmic tail, a transmembrane helix, and a collagenous C-terminal extracellular domain. These collagens can act as cellular receptors and facilitate cell adhesion and as soluble collagen in ECM upon cleavage (Ricard-Blum, 2011; Theocharis et al., 2016). Examples of MACITs include collagen types XIII, XXIII, and XXV and these are expressed by several cell types (Kadler et al., 2007; Theocharis et al., 2016).

TABLE 1 | The collagen family.

Collagen type	Classification	Chain composition	Tissue distribution	Function	References
I	Fibril-forming	$\alpha1[I]_2\alpha2[I]$ $\alpha1[I]_3$	Abundant and present in most connective tissues and interstitial membranes	Key structural component	Ricard-Blum, 2011; Henriksen and Karsdal, 2016
II	Fibril-forming	$\alpha1[II]_3$	Cartilage, vitreous humor, intervertebral disk	Tissue integrity and resiliency to stress	Ricard-Blum, 2011; Gudmann and Karsdal, 2016
III	Fibril-forming	$\alpha1[III]_3$	Tissues containing type I collagen, especially embryonic skin and hollow organs like blood vessels, uterus and bowel	Structural component, wound healing, interacts with platelets in blood clotting cascade	Ricard-Blum, 2011; Nielsen and Karsdal, 2016a
IV	Network-forming	$\alpha1[IV]_2$ $\alpha2[IV]$ $\alpha3[IV]$ $\alpha4[IV]$ $\alpha5[IV]$ $\alpha5[IV]_2$ $\alpha6[IV]$	Basement membranes	Barrier between tissue compartments, signaling	Ricard-Blum, 2011; Sand et al., 2016
V	Fibril-forming	$\alpha1[V]_3$ $\alpha1[V]_2$ $\alpha2[V]$ $\alpha1[V]$ $\alpha2[V]$ $\alpha3[V]$	Tissues containing type I collagen	Regulates collagen fibrillogenesis	Ricard-Blum, 2011; Leeming and Karsdal, 2016
VI	Beaded-filament-forming	$\alpha1[VI]$ $\alpha2[VI]$ $\alpha3[VI]$ $\alpha1[VI]$ $\alpha2[VI]$ $\alpha4[VI]$	Most connective tissues	Modulates stiffness and mechanical properties of extracellular matrix, signaling	Ricard-Blum, 2011; Sun and Karsdal, 2016a
VII	Anchoring fibrils	$\alpha1[VII]_2$ $\alpha2[VII]$ $\alpha1[VII]_3$	Many tissues	Stability of extracellular matrix	Ricard-Blum, 2011; Mortensen and Karsdal, 2016
VIII	Network-forming	$\alpha1[VIII]_3$ $\alpha2[VIII]_3$ $\alpha1[VIII]_2$ $\alpha2[VIII]$ $\alpha1[VIII]$ $\alpha2[VIII]_2$	Descemet's membrane, heart, brain, liver, lung, muscles and around chondrocytes in cartilage	Structural component, signaling	Ricard-Blum, 2011; Hansen and Karsdal, 2016
IX	FACIT	$\alpha1[IX]$ $\alpha2[IX]$ $\alpha3[IX]$	Tissues containing type II collagen	Stabilization of the fibrillar collagen network, limits collagen fibril diameter	Ricard-Blum, 2011; He and Karsdal, 2016
X	Network-forming	$\alpha1[X]_3$	Hypertrophic cartilage	Endochondral ossification	Shen, 2005; Ricard-Blum, 2011
XI	Fibril-forming	$\alpha1[XI]$ $\alpha2[XI]$ $\alpha3[XI]$ $\alpha1[XI]$ $\alpha1[V]$ $\alpha3[XI]$	Tissues containing type II collagen	Regulate fibrillogenesis of type II collagen fibrils, nucleator for collagen types I and II fibrillogenesis	Ricard-Blum, 2011; Luo and Karsdal, 2016
XII	FACIT	$\alpha1[XII]_3$	Tissues containing type I collagen	Osteoblast/osteocyte differentiation, skin homeostasis and repair, tendon development, regulation of fibrillogenesis	Ricard-Blum, 2011; Izu et al., 2020; Schönborn et al., 2020
XIII	MACIT	$\alpha1[XIII]_3$	Many tissues but present in low amounts	Plays a role in bone formation, presynaptic and postsynaptic maturation and integrity	Ricard-Blum, 2011; Zainul et al., 2018
XIV	FACIT	$\alpha1[XIV]_3$	Tissues containing type I collagen	Regulates fibrillogenesis by limiting fibril diameter	Ricard-Blum, 2011; Manon-Jensen and Karsdal, 2016
XV	Multiplexin	$\alpha1[XV]_3$	Basement membrane	Crosslinks collagen type I and III fibrils	Ricard-Blum, 2011; Arvanitidis and Karsdal, 2016
XVI	FACIT	$\alpha1[XVI]_3$	Many tissues	Stability of extracellular matrix, signaling	Ricard-Blum, 2011; Sand and Karsdal, 2016

(Continued)

TABLE 1 | Continued

Collagen type	Classification	Chain composition	Tissue distribution	Function	References
XVII	MACIT	$\alpha1[XVII]_3$	Hemidesmosomes	Adhesion of epithelial cells to extracellular matrix, teeth formation	Ricard-Blum, 2011; Sun and Karsdal, 2016b
XVIII	Multiplexin	$\alpha1[XVIII]_3$	Basement membrane	Integrity of basement membrane, inhibit angiogenesis and tumor growth	Ricard-Blum, 2011; Bager and Karsdal, 2016
XIX	FACIT	$\alpha1[XIX]_3$	Basement membrane	Acts as a cross-bridge between collagen fibrils and other extracellular molecules	Ricard-Blum, 2011; Nielsen and Karsdal, 2016b
XX	FACIT	$\alpha1[XX]_3$	Cornea, minor component of multiple connective tissues	Specific role unknown	Ricard-Blum, 2011; Willumsen and Karsdal, 2016
XXI	FACIT	$\alpha1[XXI]_3$	Many tissues	Acts as a cross-bridge between collagen fibrils and other extracellular matrix molecules	Ricard-Blum, 2011; Kehlet and Karsdal, 2016
XXII	FACIT	$\alpha1[XXII]_3$	Tissue junction in skeletal and heart muscle	Plays a role in vascular stability	Ricard-Blum, 2011; Ton et al., 2018
XXIII	MACIT	$\alpha1[XXIII]_3$	Cornea, lung, cartilage, amnion	Induce keratinocyte adhesion and spreading, cancer cell metastasis	Ricard-Blum, 2011; Veit et al., 2011; Spivey et al., 2012
XXIV	Fibril-forming	$\alpha1[XXIV]_3$	Bone, cornea	Regulation of osteoblast differentiation and mineralization	Ricard-Blum, 2011; Wang et al., 2012
XXV	MACIT	$\alpha1[XXV]_3$	Brain, neurons	Fusion of myoblasts into myofibers, regulates intramuscular motor innervation	Ricard-Blum, 2011; Tanaka et al., 2014; Gonçalves et al., 2019
XXVI	–	$\alpha1[XXVI]_3$	Testis, ovary	Development of reproductive tissues	Sato et al., 2002; Ricard-Blum, 2011
XXVII	Fibril-forming	$\alpha1[XXVII]_3$	Cartilage, eye, ear, lung, colon	Structural role in the pericellular extracellular matrix, transition of cartilage to bone	Ricard-Blum, 2011; Luo et al., 2017
XXVIII	–	$\alpha1[XXVIII]_3$	Dorsal root ganglia, peripheral nerves, in low amounts in skin and calvaria	Specific role unknown	Ricard-Blum, 2011; Gebauer et al., 2016

Beaded filament collagens include collagen types VI, XXVI, and XXVIII with type VI being the most studied (Theocharis et al., 2016). Once these collagens are secreted from the cell, they arrange in an anti-parallel fashion to form dimers. Dimers then form tetramers through interactions with other dimers. Next, tetramers connect by their globular domains to form filaments where globular domains appear as beads (Kadler et al., 2007; Theocharis et al., 2016). Beaded filament collagens are found in various connective tissues, e.g., cartilage, bone, tendon, etc. (Fitzgerald et al., 2013).

Multiplexins include collagen types XV and XVIII and have not been studied extensively (Theocharis et al., 2016). They are localized to vascular and epithelial basement membranes and participate in bridging other collagens to underlying structures (Theocharis et al., 2016).

OTHER HOST PROTEINS WITH COLLAGEN-LIKE REGIONS

G-X-X' repeats are the defining feature of the collagen triple helix primary sequence. Proteins with collagen-like regions but not classified as conventional collagen have been identified in mammals and microbes (Pyagay et al., 2005; Tom Tang et al., 2005; Ricard-Blum, 2011; Yu et al., 2014; Casals et al., 2019). Mammalian proteins containing collagen-like domains include membrane proteins (e.g., scavenger receptors) (Ricard-Blum, 2011; Zani et al., 2015; PrabhuDas et al., 2017) and secreted proteins (e.g., human defense collagens) (Ricard-Blum, 2011; Casals et al., 2019).

Human defense collagens include members of the collectin family, ficolins, and C1q and TNF-related proteins

(Casals et al., 2019). Members of the collectin family are surfactant protein A and D, mannan-binding lectin, collectin liver-1, collectin kidney-1, and the heterotrimeric Collectin CL-LK formed by the combination of collectin liver-1 and collectin kidney -1 (Casals et al., 2019). The ficolin family contains three ficolins: M-, H-, and L-ficolin (Casals et al., 2019). The C1q and TNF-related protein family only contains two members: C1q and adiponectin (Casals et al., 2019).

The defense collagens contain a N-terminal segment, a collagen-like region and a globular recognition domain that recognizes pathogen-associated molecular patterns and danger-associated molecular patterns (Fraser and Tenner, 2008; Casals et al., 2018, 2019). These proteins form multimeric structures and play an important role in pathogen clearance (Fraser and Tenner, 2008; Casals et al., 2018, 2019). The collagen-like regions of defense collagens vary in length and contain G-X-X' repeats where X is often a proline, and X' is often a hydroxylysine or a hydroxyproline (Casals et al., 2019). The collagen-like domains in human defense collagens serve two functions: (1) binding to associated proteases responsible for triggering the complement cascade and (2) binding cell receptors involved in clearance of pathogens and dead cells (Casals et al., 2019).

TABLE 2 | Collagen binding CNA-like proteins.

Protein name	Species	Collagen	K_D	References
Cne	*S. equi*	I	125 nM	van Wieringen et al., 2010
		II	50 nM	van Wieringen et al., 2010
		III	100 nM	van Wieringen et al., 2010
Cbm	*S. mutans*	I	ND	Nomura et al., 2012, 2013
		III	ND	Nomura et al., 2013
		IV	ND	Nomura et al., 2013
Acm	*E. faecium*	I	3.8 μM	Nallapareddy et al., 2003
		IV	12.8 μM	Nallapareddy et al., 2003
Ace	*E. faecalis*	I	48 μM	Rich et al., 1999; Ross et al., 2012
		IV	ND	Nallapareddy et al., 2000
Cna	*S. aureus*	I	54 nM	Xu et al., 2004b; Ross et al., 2012
Acb	*S. gallolyticus*	I	45 nM	Sillanpää et al., 2009
		IV	0.3 μM	Sillanpää et al., 2009
		V	0.5 μM	Sillanpää et al., 2009
Cnm	*S. mutans*	I	ND	Sato et al., 2004; Nomura et al., 2013
		III	ND	Nomura et al., 2013
		IV	ND	Nomura et al., 2013
BA0871	*B. anthracis*	I	1.6–3.2 μM	Xu et al., 2004a
BA5258	*B. anthracis*	I	0.6–0.9 μM	Xu et al., 2004a

ND stands for not determined.

COLLAGEN BINDING PROTEINS

CNA-Like MSCRAMMs

In 1985 *Staphylococcus aureus* was reported to bind type I procollagen and soluble collagen type I (Carret et al., 1985; Holderbaum et al., 1985). Later the collagen "receptor" on *S. aureus* was identified as a 135 kDa cell wall-anchored protein and named Collagen Adhesin (CNA) (Speziale et al., 1986). Since then, bioinformatic analyses have identified homologous proteins in other Gram-positive bacteria. These include *Enterococcus faecalis* (Ace) (Rich et al., 1999), *Enterococcus faecium* (Acm) (Nallapareddy et al., 2003), *Streptococcus mutans* (Cnm) (Sato et al., 2004), *Streptococcus equi* (Cne) (Lannergard et al., 2003), *S. mutans* (Cbm) (Nomura et al., 2012), *Bacillus anthracis* (BA0871 and BA5258) (Xu et al., 2004a), *Erysipelothrix rhusiopathiae* (RspA and RspB) (Shimoji et al., 2003), and Acb from *Streptococcus gallolyticus* (Sillanpää et al., 2009; **Table 2**). Amongst these proteins, Cna and Ace are the best-studied members.

With the exception of Acb, all CNA-like proteins are anchored directly to the cell wall. Acb is unique and is a minor pilus protein of *S. gallolyticus* (Sillanpää et al., 2009) but has a predicted CNA-like structure. Furthermore, it shares 50–70% sequence identity with Acm, Cna, and Cne (Sillanpää et al., 2009).

Structure

Collagen Adhesin is the prototype of Collagen-binding MSCRAMMs (Foster et al., 2014; Foster, 2019). CNA like proteins harbor a N-terminal signal sequence, an A-region, a variable number of characteristic B repeats, a C-terminal cell wall and membrane spanning region and a short cytoplasmic tail (**Figure 1**). The ligand-binding A-region of CNA-like proteins is further divided into two or three sub-domains: N1, N2, and N3 (**Figure 1**; Patti et al., 1993; Zong et al., 2005). X-ray crystallography of Cna and Ace N1N2 sub-domains revealed that these domains adopt IgG-like folds called Dev-IgG and are consequently composed of mostly β-sheets (**Figure 2**; Foster et al., 2014; Foster, 2019). The N1 and N2 domains are connected by a rather long (10 aa) hydrophobic linker region, which creates a hole of ∼15 Å between the two domains and provides flexibility in domain orientation (**Figure 2**; Zong et al., 2005). Additionally, proteins in the CNA-like MSCRAMM family have a variable number of B repeats depending upon the protein (Patti et al., 1994a; Kang et al., 2013). One B repeat is ∼180 aa long and is further divided into two ∼90 aa subdomains, D1 and D2. The D subdomains adopt an inverse IgG fold and together B repeats are thought to form a stalk projecting the ligand binding region away from the bacterial cell surface (Deivanayagam et al., 2000).

Binding Mechanism

The truncated N2 domain is the minimum collagen-binding region of CNA, although optimal binding is achieved by the N1N2 segments. The CNA N1N2 segment binds collagen type I with an affinity of 54 nM (Patti et al., 1993, 1995; Zong et al., 2005; Ross et al., 2012). Electron microscopy imaging of rCNA with collagen triple helix monomers revealed that CNA binds collagen at multiple sites, without any obvious preference for a "hot spot". Surface plasmon resonance (SPR) studies of $rCNA_{31-344}$ with synthetic collagen peptides further confirmed its preference for a triple helical structure (Zhang et al., 2015). CNA binds

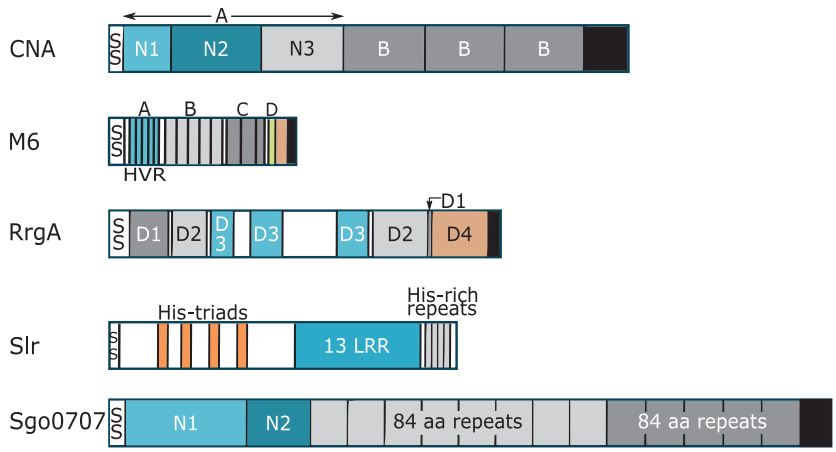


FIGURE 1 | Cartoon representation showing domain organization of collagen binding proteins. Each protein has a signal peptide (SS) at the N-terminus. With the exception of Slr, each protein has a LPXTG motif, and a cell wall-spanning transmembrane region (black). Confirmed and putative collagen-binding domains in the proteins are colored in blue. Where two domains are involved in collagen binding, the first domain is light blue and the second domain is colored darker blue. Proline-glycine rich region of M6 protein is shown in salmon color.

preferentially to cleaved collagen in damaged or inflamed tissues (Madani et al., 2017).

Collagen Adhesin-like proteins bind collagen by a "collagen hug" mechanism where the N1N2 segment "hugs" or wraps around the collagen triple helix molecule (**Figure 2**). A co-crystal of CNA bound in complex with the synthetic collagen peptide $(GPO)_4GPRGRT(GPO)_4$, where O is hydroxyproline, provided the insights into the molecular basis of this model. The collagen hug binding mechanism is initiated when the collagen triple helix interacts with the shallow groove on the CNA N2 domain. This interaction is low affinity, and involves polar and hydrophobic residues (Zong et al., 2005). The initial interaction leads to structural rearrangements within the N1 domain that repositions N1 closer to the N2 domain creating a "tunnel-like" structure. Finally, the C-terminal extension of the N2 domain undergoes structural changes, and inserts into the N1 domain by β-strand complementation thus forming a "latch" (**Figure 2**). The N1 domain of CNA interacts with the middle chain while the N2 domain interacts with the leading and trailing chains of the synthetic collagen peptide. The N1N2 linker region covers the collagen peptide and holds it in place (Zong et al., 2005; Liu et al., 2007).

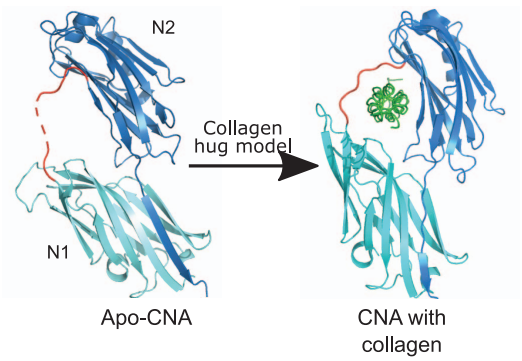


FIGURE 2 | Collagen hug model. The N1N2 domains of CNA in both apo-form [PDB code: 2F68 (Zong et al., 2005)] and with collagen peptide [PDB code: 2F6A (Zong et al., 2005)] are shown. N2 and N1 domains are shown in dark blue and light teal color, respectively. CNA N1N2 domains are connected together by a linker shown in red. Collagen triple helix is shown in green. C-terminal extension of the N2 domain that forms a latch is shown as dark blue β-strand in N1 domain.

The two-step binding mechanism of CNA to collagen was confirmed by atomic force microscopy studies where a moderate force (∼250 pN) was observed for the initial hydrophobic interaction between collagen and the N2 domain of CNA (Herman-Bausier et al., 2016). After binding collagen, a strong force of ∼1.2 nN was observed for the full interaction. Although B-repeats of CNA do not bind collagen directly, they act as a spring and help withstand the high mechanical stress encountered *in vivo* (Herman-Bausier et al., 2016).

Although all members of the CNA-like MSCRAMM family appear to bind collagen by a collagen hug mechanism, the proteins show differences in affinity (**Table 2**) and mechanistic details because of structural variations. For example, CNA has a higher affinity for the collagen triple helix than Ace (Ross et al., 2012). In contrast to the two-step mechanism used by CNA, Ace binds collagen with a rapid association and dissociation rate in a one-step binding mechanism (Rich et al., 1999; Ross et al., 2012).

Virulence

Most proteins in the CNA-like MSCRAMM sub-family have been shown to act as virulence factors in experimental bacterial infections. CNA-like proteins target collagen to enhance adhesion of the bacteria to host tissues in early and later stages of infection. For example, CNA is a critical virulence factor of *S. aureus* in experimental septic arthritis and osteomyelitis models and this role depends on its ability to bind collagen (Patti et al., 1994b; Elasri et al., 2002; Xu et al., 2004b). Although CNA is not required

in the initial targeting of joints, it is critical for hematogenous spread of *S. aureus* leading to bone infections (Elasri et al., 2002). Additionally, more bacteria were isolated from joints of mice infected with collagen binding *cna*$^+$ bacterial strains than those infected with non-collagen binding strains. Most CNA-like proteins also bind to collagen present in vegetations observed in non-bacterial thrombotic endocarditis, thus leading to infective endocarditis (Hienz et al., 1996; Nallapareddy et al., 2008; Singh et al., 2010). Ace and Acm, the enterococcal CNA-like proteins, are important virulence factors in infective endocarditis (Nallapareddy et al., 2008; Singh et al., 2010). The *ace* deletion mutant of *E. faecalis* OG1RF strain showed decreased colonization of heart valves in a mixed-infection rat endocarditis model compared to the wild type strain. Higher bacterial colony forming unit (CFU) counts were recovered from aortic valve vegetations at 4 h in mono endocarditis infection of rats with *ace* expressing *E. faecalis* OG1RF compared to the *ace* deletion mutant, indicating a role in early colonization of heart valves (Singh et al., 2010). Similarly, significantly more wild type (WT) *E. faecium* TX0082 CFUs were recovered from rat vegetations after mixed endocarditis infection compared to *acm* deletion mutant *E. faecalis* TX6051. Furthermore, Acm was also shown to enhance early adherence to heart valves (Nallapareddy et al., 2008). On the other hand, CNA's ability to bind collagen is of limited significance in early stages of attachment to traumatized aortic valves, but like Acm and Ace (Nallapareddy et al., 2008), CNA does contribute to establishment of infection at a 24 h time point in both mono and mixed endocarditis infections of rats with *S. aureus* isolates (Hienz et al., 1996).

Cbm and Cnm are homologous *S. mutans* proteins with 78% identity in their collagen binding domains. The *cnm* gene is sufficient and necessary for primary human coronary artery endothelial cell invasion by *S. mutans* isolates as shown with Δ*cnm S. mutans* clinical isolates as well as *cnm*$^+$ *Lactococcus lactis* (Abranches et al., 2011; Freires et al., 2017). The *cnm* gene also permits invasion of other non-phagocytic cells like human gingival fibroblasts and human oral keratinocytes (Miller et al., 2015). In addition, *cnm*$^+$ *S. mutants* OMZ175 and *cnm*$^+$ *L. lactis* outcompeted Δ*cnm S. mutans* OMZ175 and *L. lactis*, by 10 and 100-fold, respectively, in *ex vivo* bacterial adherence to aortic valve sections. Using a rabbit model of infective endocarditis, it was shown that *cnm L. lactis* mediated attachment to injured endocardium but not to the vegetations (Freires et al., 2017). Similar to Cnm, *cbm*$^+$ *S. mutans* attaches to aortic valves and leads to larger vegetations formed on the impaired heart valve tissue compared to *cbm*$^-$ *S. mutans*.

In addition, collagen-binding proteins have been implicated in various infections. For example, CNA has been implicated in pathogenesis of *S. aureus* keratitis (Rhem et al., 2000) and orthopedic prosthesis infections (Montanaro et al., 1999). Similarly, Cnm has been implicated in *S. mutans* cerebral hemorrhaging (Tonomura et al., 2016) and colonization of dental pulp (Nomura et al., 2016).

M and M-Like Proteins

M-protein, described by Rebecca Lancefield almost a century ago (Lancefield, 1928), is a major cell wall-anchored protein and virulence factor present on the surface of Group A, B, and C streptococci (GAS, GBS, and GCS) (Dinkla et al., 2007; Barroso et al., 2009; Reissmann et al., 2012). There are around ∼250 known M-protein types in GAS based on sequence variation in the first 50 amino acids of the protein. Variations in the M-protein lead to strain-specific immunity and, hence, M-proteins serve as a strain typing marker (Lancefield, 1928). M proteins have multiple functions, including inhibition of phagocytosis and binding to fibrinogen, collagen, complement, and other host proteins (Metzgar and Zampolli, 2011).

Structure

M proteins are multi-domain proteins that adopt an elongated α- helical structure and dimerize to form helical coiled-coil structures, a structure form also seen in mammalian proteins like tropomyosin and myosin (McNamara et al., 2008; Fischetti, 2016). M-protein fibrils are ∼500 Å long and coat the surface of Group A streptococcus (Phillips et al., 1981; Fischetti, 1989). When viewed by transmission electron microscopy, M-protein appears like "fuzz on a tennis ball" (Phillips et al., 1981). All M-proteins contain a signal peptide, a hypervariable region, a less variable central domains and a highly conserved C-terminus (**Figure 1**; Fischetti, 1989).

The prototypic M6 protein consists of a cleavable signal sequence, A repeats, which includes the hypervariable region (HVR), B repeats, C repeats, D-region and a LPXTG motif for sortase mediated anchoring to the cell wall (**Figure 1**). The HVR region is the first 50 amino acids of the mature M protein and shows variation amongst the different M-proteins. The M6 A-repeat region consists of five repeats of 14 amino acids each, where the central repeats are identical and end repeats are slightly divergent (Smeesters et al., 2010; Fischetti, 2016). The B-repeat region contains five repeats, each 25 amino acid long (Fischetti, 1989, 2016). The M6 protein contains two C-repeats where each repeat is 35 residues long (Fischetti, 1989, 2016). C-repeats show higher sequence conservation compared to A- and B-repeats. Lastly, the M6 protein contains four D-repeats, each 7 amino acid long (Fischetti, 1989, 2016). Amongst the A, B, C, and D repeat regions, D-repeats show highest sequence homology to each other for any M protein (Smeesters et al., 2010). Together, A-, B-, C-, and D-repeats form the central helical rod (Fischetti, 1989, 2016, 2019; McNamara et al., 2008).

As observed in tropomyosin and myosin, the coiled-coil nature of a protein molecule comes from heptad repeats, where the first and fourth residues in the register are generally hydrophobic (Fischetti, 1989, 2016; McNamara et al., 2008). Hydrophobic residues form the core of the coiled coil and the remaining residues in the heptad repeats are generally helix promoting (McNamara et al., 2008; Fischetti, 2016). Heptad repeats found in M-proteins are not perfect, which leads to irregularities and instabilities of the coiled-coil region (McNamara et al., 2008; Macheboeuf et al., 2011). McNamara et al. found that destabilizing residues in the coiled-coil region of M1 protein promote conformational dynamics, which is required for binding of M1 protein to fibrinogen (McNamara et al., 2008; Stewart et al., 2016). These irregularities in the heptad repeats also

form the basis for sub-division of the protein into A-, B-, and C-repeats (Fischetti, 2019).

Sequence and structural variations amongst M-proteins are common. Homologous recombination in M-protein leads to differences in the frequency and length of the repeats and an overall variation in size (Fischetti, 1989). As a result, A- and B- repeats are not present in all M-proteins and when present, their sizes can vary. However, all M-proteins contain C-repeats and their total number can vary from two to four (Smeesters et al., 2010). The sequence variations between M-proteins lead to functional differences and hence not all M-proteins possess all the functional capabilities described in the literature.

Binding Mechanism

Amongst the >250 known types of M-proteins, about 20 have been shown to bind collagen (**Table 3**). M-proteins bind directly to the triple helical regions of collagen (Nitsche et al., 2006; Barroso et al., 2009; Dinkla et al., 2009; Bober et al., 2011; Reissmann et al., 2012) with the exception of the M1 protein, which also interacts with the globular domain of collagen type VI (Bober et al., 2010). Rotary shadowing electron microscopy revealed that M3 protein binds collagen type IV at two different sites: one located on cyanogen bromide fragment 3 (CB3) and the other at a site 20 nm away from the 7S domain (Eble et al., 1993). CB3 is a fragment of collagen type IV that maintains its triple helix and is generated after cleavage of collagen with cyanogen bromide (Eble et al., 1993). When expressed on the surface of a heterologous non-collagen binding host (*Streptococcus gordonii* GP1221), M-proteins from GCS and Group G streptococci (GGS) enabled GP 1221 to bind to collagen type IV at the same level as GCS and GGS (Barroso et al., 2009).

Peptide associated with rheumatic fever (PARF) is an eight-residue motif present in the hypervariable A region of some M- and M-like proteins (Dinkla et al., 2007; Barroso et al., 2009; Reissmann et al., 2012). Based on careful examination of multiple M-proteins from 69 isolates, a consensus sequence of the PARF motif was determined to be (A/T/E)XYLXX(L/F)N where charged amino acids are preferred at positions 2, 5, and 6, with at least one of the charged amino acids containing a basic side chain (Barroso et al., 2009; Reissmann et al., 2012). A PARF motif is required for binding of these M-proteins to collagen (Dinkla et al., 2007; Reissmann et al., 2012), as one or two substitutions of the conserved residues in the PARF motif abolishes binding to collagen type IV (Reissmann et al., 2012). However, additional data suggests that the binding of M-proteins to collagen can be more complicated and extends beyond the PARF motif. First, a series of recombinant truncated PARF-containing versions of an M-protein bind collagen with significantly different affinities (Dinkla et al., 2007). A full-length recombinant M-protein of GGS called "fibrinogen-binding protein of G streptococci" (FOG) binds to collagen type IV with a K_D of 6 nM, whereas a truncated FOG protein containing A- and B-repeats binds collagen type IV with 24 times higher K_D and a FOG protein containing the A-region only binds collagen type IV with a 200 fold higher K_D compared to the full length FOG protein (Dinkla et al., 2007). Similarly, a truncated recombinant FOG protein binds collagen type I with a 20 fold higher K_D than the full length recombinant FOG protein (Nitsche et al., 2006). Furthermore, Reissmann et al. (2012) identified M-proteins with PARF motifs that did not bind collagen type IV. Interestingly, M-proteins stG120.1, stG120.0, and stGM220 all contain the same PARF motif but only stG120.1 binds collagen type IV, while all three proteins bind fibrinogen. Moreover, the M1-protein lacks a PARF motif (Reissmann et al., 2012) but still binds to the triple helix of collagen types I and IV (Bober et al., 2011) and globular domains of collagen type VI (Bober et al., 2010).

M-proteins binding to different types of collagens can have different consequences. Binding of M-proteins to collagen type IV leads to aggregation of collagen on the surface of the bacteria (Dinkla et al., 2003b, 2007; Barroso et al., 2009), which is not observed with the interaction of collagen type I to M-protein (Barroso et al., 2009). Expression of M-protein on the surface of a heterologous host leads to collagen type IV aggregation, demonstrating that the M-protein alone is sufficient for collagen aggregation.

TABLE 3 | Collagen binding M-proteins.

M-protein	Species	Collagen	K_D	References
stG4545.0	SDSE	IV	1.8 pM	Barroso et al., 2009
stC2sk.0	SDSE	IV	3.5 pM	Barroso et al., 2009; Reissmann et al., 2012
stC5344	SDSE	IV	920 pM	Barroso et al., 2009; Reissmann et al., 2012
stG2574.0	SDSE	IV	1.2 nM	Barroso et al., 2009; Reissmann et al., 2012
stC-NSRT2.0	SDSE	IV	830 pM	Barroso et al., 2009; Reissmann et al., 2012
stG10.0	SDSE	IV	610 pM	Barroso et al., 2009; Reissmann et al., 2012
FOG (stG11.0)	SDSE	I	80 pM	Nitsche et al., 2006
		IV	6 nM	Dinkla et al., 2007; Barroso et al., 2009
M3	*S. pyogenes*	IV	5 nM	Dinkla et al., 2009; Reissmann et al., 2012
stG97	SDSE	IV	ND	Reissmann et al., 2012
stC6746	SDSE	IV	ND	Reissmann et al., 2012
M31.5	SDSE	IV	0.6 nM	Reissmann et al., 2012
M3.22	*S. pyogenes*	IV	ND	Reissmann et al., 2012
stG211.1	SDSE	IV	ND	Reissmann et al., 2012
stG120.1	SDSE	IV	ND	Reissmann et al., 2012
stG351	SDSE	IV	ND	Reissmann et al., 2012
stCQ343	SDSE	IV	ND	Reissmann et al., 2012
stG211.0	SDSE	IV	ND	Reissmann et al., 2012
stC922	SDSE	IV	ND	Reissmann et al., 2012
M55	*S. pyogenes*	IV	5 nM	Reissmann et al., 2012
M1	*S. pyogenes*	I	54 nM	Bober et al., 2011
		VI	ND	Bober et al., 2010

SDSE stands for S. dysgalactiae subsp. *equisimilis; ND stands for not determined.*

Virulence

M or M-like proteins are major virulence factors of Streptococci and their role in streptococcus pathogenesis have been reported on extensively (Oehmcke et al., 2010; Smeesters et al., 2010; Frost et al., 2017; Fischetti, 2019). In this review article, we will focus on the contribution of the M-protein:collagen interaction to the

pathogenesis of streptococci. Binding of M-proteins to collagen can have two consequences: (1) mediating bacterial adhesion to connective tissues and (2) inducing collagen auto-immunity.

M-protein binding to collagen is important in the colonization of human skin by streptococci (Nitsche et al., 2006). When incubated with human dermis *ex vivo*, higher CFU counts were recovered from a GGS strain expressing FOG protein compared to a FOG-deficient strain. Incubation of the bacteria with collagen type I decreased adherence of the FOG expressing strain to human dermis, thereby also suggesting that the interaction of FOG with collagen type I enables adhesion.

Acute rheumatic fever (ARF) and rheumatic heart disease are antibody-mediated autoimmune sequelae that can develop after a streptococcal infection (Tandon et al., 2013; Carapetis et al., 2016). Binding of M-protein to collagen has been shown to be a relevant factor in developing ARF (Dinkla et al., 2003b, 2007; Barroso et al., 2009). Binding of M or M-like protein to collagen type IV can lead to production of antibodies binding the collagen molecule (Dinkla et al., 2003b, 2007; Barroso et al., 2009). Analysis of mouse sera obtained from immunization with recombinant M or M-like protein led to identification of two distinct antibody populations: anti-collagen type IV antibodies and anti-M protein antibodies. These distinct antibodies did not cross-react with each other (Dinkla et al., 2007), indicating that collagen type IV autoimmunity was not generated through molecular mimicry. In addition, sera of ARF patients contain antibodies that specifically recognize the CB3 region of collagen type IV and the collagen-binding region of the M3 protein (Dinkla et al., 2007, 2009). The N-terminal half of the protein containing the PARF motif is required for generating auto-immunity (Dinkla et al., 2007). Immunization of mice with full-length FOG led to a significantly higher titer of anti-collagen type IV antibodies compared to mice immunized with FOGB2-C2, a region of FOG that does not bind collagen (Dinkla et al., 2007). Similar results have been obtained with other M-proteins (Dinkla et al., 2007; Barroso et al., 2009). While auto-antibodies to collagen type I have not been demonstrated, given the structural similarities between the collagens, anti-collagen type IV antibodies potentially could also react with other collagen types.

EMERGING FAMILIES OF COLLAGEN-BINDING PROTEINS

Numerous collagen-binding proteins of Gram-positive pathogens have been reported in the literature but their mechanisms of collagen binding are unclear. We have identified three emerging families of collagen-binding proteins where, although one or more than one family member binds to collagen, a clear picture of how these proteins bind to collagen is not yet available.

von Willebrand Factor A- Domain Containing Proteins

von Willebrand factor (vWF) is a host glycoprotein found in blood, blood vessel ECM, and platelet α-granules (Manon-Jensen et al., 2016). vWF is a large modular protein that contains two binding sites for collagen located in the A1 and A3 domains. The A3 domain of vWF binds collagen types I and III whereas the A1 domain binds collagen types IV and VI (Manon-Jensen et al., 2016). Crystal structures of both A1 and A3 domains show a central β-sheet composed of six β-strands and flanked on both sides by α-helices (Huizinga et al., 1997; Emsley et al., 1998). These domains are structurally similar to the I-domain of some integrin α-chains, including the collagen-binding α1-, 2-, 10-, and 11- chains. The collagen-binding α-chain integrins also contain a metal ion-dependent adhesion site (MIDAS) important for ligand binding (Lee et al., 1995; Huizinga et al., 1997; Emsley et al., 1998).

Structural homologs of vWF A-domains, called vWA domains, have been found in minor pilus proteins that bind to ECM proteins and host cells. These pilus proteins include RrgA from *Streptococcus pneumoniae* (Izore et al., 2010), GBS104 from *Streptococcus agalactiae* (Krishnan et al., 2013), PilA from *S. agalactiae* (Konto-Ghiorghi et al., 2009; Banerjee et al., 2011), SpaC from *Corynebacterium diphtheriae* (Mandlik et al., 2007), and EbpA from *E. faecalis* (Nielsen et al., 2012). Most structural information about bacterial vWA domains comes from crystal structures of the RrgA and the GBS104 proteins (Izore et al., 2010; Krishnan et al., 2013). RrgA and GBS104 are homologs that share 51% sequence identity with each other and have a similar domain organization (Krishnan et al., 2013). Both proteins contain an N-terminal signal sequence, four D domains named D1, D2, D3, and D4, and a C-terminal sorting signal (**Figure 1**). The primary sequence of both D1 and D2 domains is non-contiguous, and is divided into two regions, one present in the N-terminal half and other present in the C-terminal half of the protein (**Figure 1**). The two regions fold back on each other to form the tertiary structure of the D1 and D2 domains. The D3 domain is inserted in between the two regions encoding the D1 and D2 domains and the D4 region is located distal to the C-terminal half of the D1-D2 domain (**Figure 1**). It is worth noting that while RrgA and GBS104 are structural homologs, other pilus proteins containing vWA domains like PilA have a different overall domain organization (Mandlik et al., 2007; Konto-Ghiorghi et al., 2009; Izore et al., 2010; Banerjee et al., 2011; Nielsen et al., 2012; Krishnan et al., 2013).

The D3 domains of both RrgA and GBS104 adopt a structure similar to the vWF A-domain. These D3 domains of both RrgA and GBS104 consist of a central β-sheet flanked by α-helices on both sides as seen in the vWF A-domain and the integrin I-domain (**Figure 3A**; Huizinga et al., 1997; Emsley et al., 1998; Izore et al., 2010; Krishnan et al., 2013). In addition, both RrgA and GBS104 have two arms inserted into the vWA-domain that are absent in the A-domains of vWF and the I-domain of integrins (Lee et al., 1995; Huizinga et al., 1997; Emsley et al., 1998; Izore et al., 2010; Krishnan et al., 2013). The first arm of RrgA contains two β-hairpins folded together to form an elongated arm (**Figure 3A**). The second arm of RrgA consists mostly of loops along with one short hairpin, two α-helices and loops (**Figure 3A**; Izore et al., 2010; Krishnan et al., 2013). The two inserted arms extend away from the core of the domain and extend the length of the protein (**Figure 3A**; Izore et al., 2010).

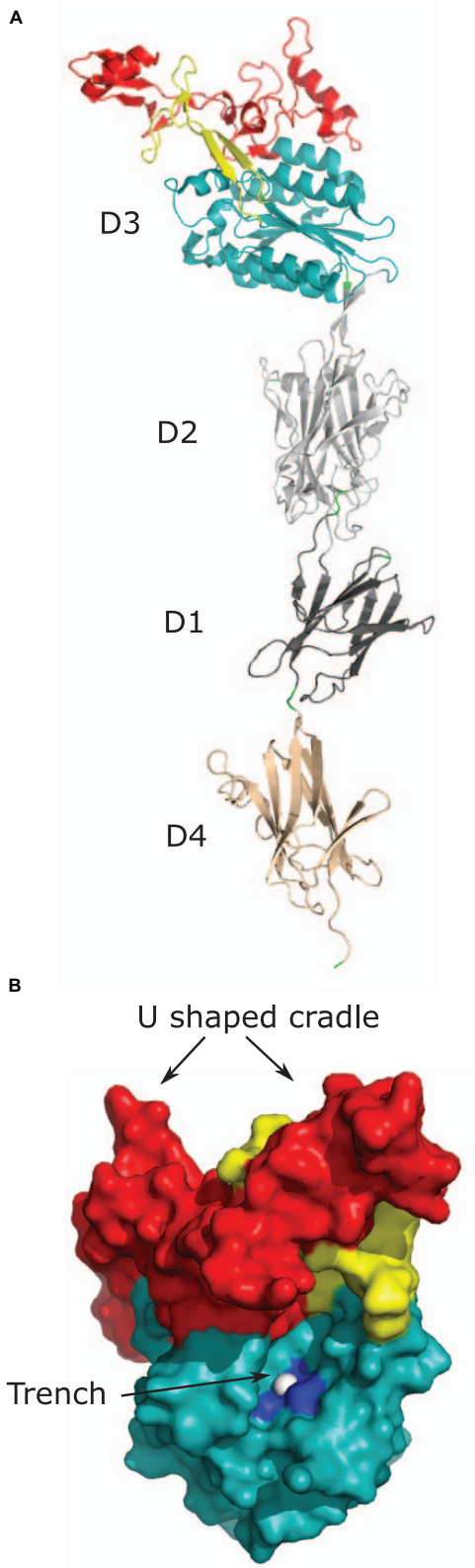


FIGURE 3 | Tip pilus protein RrgA. **(A)** Tip pilus protein RrgA of *S. pneumoniae* [PDB code: 2WW8 (Izore et al., 2010)] contains a D3 *(Continued)*

FIGURE 3 | Continued
domain with vWA fold (shown in teal), D1 (shown in dark gray), and D4 domains (shown in light brown) with an IgG-rev fold, and a D2 domain (shown in light gray) with a Dev-IgG fold. The collagen-binding D3 domain is positioned at the tip and contains two inserted arms in the D3 domain. First inserted arm is shown in yellow and second inserted arm is shown in red. **(B)** Figure shows two predicted collagen binding sites on the surface of D3 domain of RrgA. D3 domain, first inserted arm and second inserted arm are shown in light teal, yellow and red, respectively. The U-shaped cradle on the top of D3 domain is formed by two inserted arms. Magnesium ion present in MIDAS motif is shown as white sphere in the trench and residues interacting with magnesium ion (S234, S236, and D387) are shown in dark blue.

The D3 domain of the two bacterial proteins also contains a MIDAS motif (**Figure 3B**) present in the I-domain of integrins but absent in the vWF A-domains.

Amongst the pilus proteins with a vWA-domain, RrgA, PilA, and GBS104 have been reported to bind collagen. RrgA binds collagen type I, fibronectin, and laminin (Hilleringmann et al., 2008; Moschioni et al., 2010). However, the D3 vWA domain alone was not able to bind ECM proteins (Moschioni et al., 2010). Full-length RrgA protein is required for binding (Moschioni et al., 2010). RrgA binds to collagen type I with weaker force than expected for a ligand:receptor interaction. It has been suggested that the low binding force might help the pilus adhere and detach under physiological flow conditions. However, kinetic data for RrgA and collagen type I is lacking and the suggested consequences of low binding force awaits elucidation (Becke, 2019). Although recombinant PilA has been reported to bind collagen, its role in *S. agalactiae* collagen binding is not clear (Banerjee et al., 2011; Dramsi et al., 2012). Similarly, GBS104 has been reported to bind collagen but the interaction in a solid-phase binding assay is weak and does not reach saturation indicating that the interaction of GBS104 and collagen type I may not be specific or have functional relevance (Krishnan et al., 2013).

Two different binding regions in the vWA domain-containing pilus proteins have been proposed: the vWA-domain with the MIDAS motif and the U-shaped cradle formed by the inserted arms (**Figure 3B**; Izore et al., 2010). Apo-crystal structure of the vWA-domain with the MIDAS motif revealed a trench-like region formed by the two inserted arms and the MIDAS motif present on the central β-sheet (Izore et al., 2010; Krishnan et al., 2013). Based on structural comparison with co-crystals of integrin α2β1 and a synthetic triple helix peptide, the trench-like region has been proposed to be the collagen binding site (**Figure 3B**; Emsley et al., 2000; Izore et al., 2010). The vWA-domain and the integrin I-domain undergoes conformational change during binding events and transition from a closed form to an open form. Participation of the trench-like region and a change in confirmation upon ECM binding was confirmed using an open form of the GBS104-D3 domain stabilized by a disulfide bridge. The open form of the GBS104-D3 domain alone was sufficient for binding to fibronectin, whereas the closed form of the D3 domain showed no binding (Krishnan et al., 2013). The vWA-domains of pilus proteins have considerable variation in their primary sequence, with the most variations in the inserted arms (Konto-Ghiorghi et al., 2009; Izore et al., 2010;

Krishnan et al., 2013). Therefore, despite structural similarities, these pilus proteins have been suggested to bind different ligands with different affinities (Izore et al., 2010; Krishnan et al., 2013). A second binding site is the U-shape cradle formed by the two inserted arms joining together at the tip of the protein (**Figure 3B**). This cradle contains basic residues and has been proposed to bind negatively charged molecules like glycosaminoglycans attached to ECM proteins (Izore et al., 2010). While the vWA domain is critical for virulence (Konto-Ghiorghi et al., 2009; Nielsen et al., 2012), evidence that the vWA domain of RrgA is responsible for binding to collagen is lacking.

Leucine Rich Repeat Containing Proteins

Leucine rich repeats (LRRs) are protein recognition motifs present in eukaryotic proteins with diverse functions (Kobe and Kajava, 2001). Small leucine rich proteoglycans (SLRPs) in mammals are an example of LRR proteins and play important roles in collagen fibrillogenesis (Kalamajski and Oldberg, 2010). LRR containing proteins have been found in some pathogenic bacteria, e.g., *Yersinia pestis*, *Listeria monocytogenes*, plants, animals, and fungi (Kobe and Kajava, 2001). Each repeat is 20–29 aa long and are often present in tandem with multiple LRRs to form an overall curved shape where β-sheets are present on the concave side and α-helices are often on the convex side (Kobe and Kajava, 2001).

Streptococcal leucine rich (Slr) protein is an LRR-containing lipoprotein present on the surface of *S. pyogenes* (Bober et al., 2011). The N-terminal half of the protein contains a 21 aa long signal sequence and 4 histidine triad motifs (**Figure 1**; Bober et al., 2011). The C-terminal half of the protein contains 13 leucine rich regions that form β-sheets, followed by histidine rich repeat sequences (**Figure 1**; Bober et al., 2011). The horseshoe shape of Slr was visible with electron microscopy (Bober et al., 2011). Orthologs of Slr have been identified in *Streptococcus suis* 05ZYH33 (1577), *Streptococcus equi* subsp. *zooepidemicus* H70 (13200), *Streptococcus dysgalactiae* subsp. *equisimilis* GGS_124 (1372), *Streptococcus agalactiae* (Blr), *Streptococcus uberis* 0140J (1212), and *Streptococcus suis* 2 (HtpsC) (Waldemarsson et al., 2006; Plumptre et al., 2012; Li et al., 2015). Slr has been demonstrated to bind collagen directly with a K_D of 12 nM (Waldemarsson et al., 2006). Interestingly, HtpsC has been reported to bind laminin and fibronectin, but HtpsC did not bind collagen type I (Li et al., 2015).

The extracellular matrix protein (Emp) of *S. aureus* is 340 aa long secreted protein with a 26 aa long signal peptide at the N-terminus. Emp binds collagen type I with a K_D of 27 nM. Emp is structurally intriguing, as Emp is not predicted to be multi-domained. When viewed through a transmission electron microscope, the Emp monomer was revealed to form a horseshoe-type structure with an 8 nm diameter. Interestingly, even though it lacks leucine repeats, structure prediction through I-TASSER identified leucine rich repeat proteins as the top ten structural analogs (Geraci et al., 2017).

While molecular details of Slr and Emp binding to collagen have not been studied beyond the confirmation of their interaction, their intriguing overall structural similarities leads the way for postulating an emerging LRR-containing or LRR like protein family that binds collagen. Given that several human LRR proteins [e.g., decorin (Schönherr et al., 1995), fibromodulin (Font et al., 1998)] interact with collagen, it is not surprising that bacterial LRR proteins bind collagen. Additional studies are needed to determine the residues that mediate the interaction and determine similarities of those interactions with host LRR proteins and collagen.

Sgo0707 N1-Domain Containing Proteins

Streptococci express multiple surface proteins that have been reported to bind collagen (Avilés-Reyes et al., 2017). One emerging family of collagen-binding proteins in Streptococci is related to the N1-domain of Sgo0707 protein from *S. gordonii*. The Sgo0707 protein, which has been shown to bind collagen, contains a N-terminal signal sequence, a 419 aa long N-terminal region, eight repeats of 84 aa, five repeats of 88 aa, a unique domain, an LPXTG cell wall sorting signal and a transmembrane helix (**Figure 1**; Nylander et al., 2013). The 419 aa long N-terminal region is divided into two domains: N1 and N2 (**Figure 1**). Both the N1 and N2 domains adopt a β-sandwich with anti-parallel β-sheets (Nylander et al., 2013), where β-sheet 1 contains nine β-strands and β-sheet 2 contains eight strands (**Figure 4A**). The N1 domain also contains two small sub-domains A and B. The N2 domain consists of two β-sheets of five β-strands and a third small sheet of three strands and adopts a DeV-IgG fold also observed in the N1N2 domains of CNA (Nylander et al., 2013).

A search of proteins with similar N1 domains identified the variable domains in two Ag I/II family proteins; SpaP from *S. mutans* and SspB from *S. gordonii*. Both these domains are predicted to adopt a similar structure despite having only 10% sequence identity to N1 of Sgo0707 (Forsgren et al., 2009; Larson et al., 2010; Nylander et al., 2013). These proteins have a different domain organization than Sgo0707, with an N-terminal signal sequence, alanine rich repeats, a variable domain, proline rich repeats, a C-terminal domain, and an LPXTG cell wall sorting signal (Forsgren et al., 2009; Larson et al., 2010). All three proteins form an extended confirmation with the putative collagen binding domain (N-region of Sgo0707 and variable domain of SpaP and SspB) predicted to be located at the tip of the protein (Forsgren et al., 2009; Larson et al., 2010; Nylander et al., 2013).

Docking of the collagen triple helix to the Sgo0707 N1N2 domain identified two different potential binding sites (**Figures 4B,C**). The first binding site is on top of the N1 domain in the open cleft formed by the two subdomains in the N1 domain (**Figure 4B**). This site has a higher negative surface potential compared to the SspB and SpaP proteins, and lacks the metal ion located in the cleft found in both of the Ag I/II proteins (Forsgren et al., 2009; Larson et al., 2010; Nylander et al., 2013). A second putative collagen-binding site is formed by the loops of the N1 domain and a β-sheet of the N2 domain, which together form a concave surface where collagen can dock (**Figure 4C**; Nylander et al., 2013). The concave site consists of mostly non-polar residues (Nylander et al., 2013).

All three proteins (Sgo0707, SspB, and SpaP) have been implicated in collagen binding. Binding of the three proteins to collagen type I was shown in a bacterial adhesion assay using deletion mutants (Love et al., 1997, 2000; Nylander et al., 2013). Deletion of the *sgo0707* gene in *S. gordonii* DL1 decreased

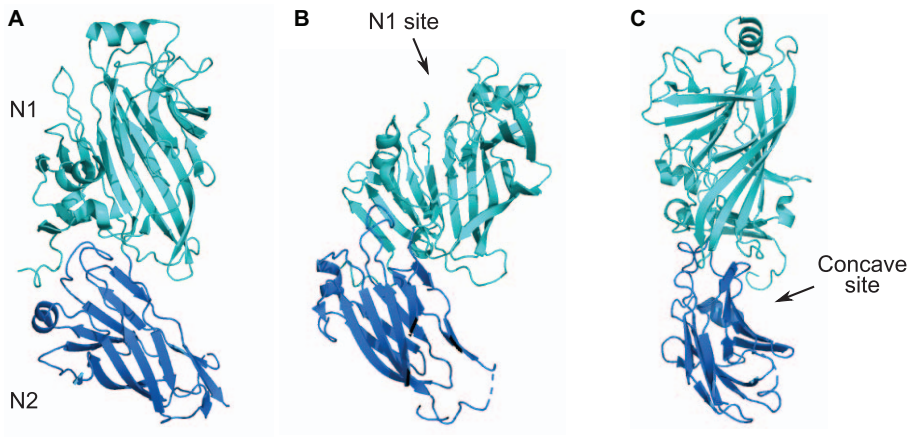


FIGURE 4 | N1N2 domains of Sgo0707. **(A)** N1 and N2 domains of Sgo0707 from *S. gordonii* [PDB code: 4IGB (Nylander et al., 2013)] are shown in light teal and dark blue, respectively. **(B)** Figure shows predicted collagen binding site present in the N1 domain. **(C)** Figure shows second predicted collagen binding site formed the loops of the N1 domain and a β-sheet of the N2 domain. This site forms a concave surface where collagen can dock.

collagen type I binding by 40% compared to the WT strain (Nylander et al., 2013). Similarly, an isogenic deletion mutant of the *sspB* gene in *S. gordonii* and of the *spaP* gene in *S. mutans* showed decreased binding to collagen type I compared to WT strains (Love et al., 1997, 2000). Additionally, binding of *S. gordonii* DL1 to collagen type I in a bacterial adhesion assay was inhibited by recombinant N-region of Sgo0707, thus narrowing down the N-region as the collagen binding partner (Nylander et al., 2013). While the three proteins have been implicated in collagen binding, their direct binding to collagen has not been demonstrated. The three proteins only share structurally similar N1-domains as both SspB and SpaP lack the N2 domain found in Sgo0707 (Love et al., 1997, 2000; Nylander et al., 2013). Do the three proteins bind collagen at the cleft on top of the N1-domain? Further studies are required to narrow down the collagen-binding site in these proteins and to determine if they form a structural family of proteins that bind collagen.

CONCLUDING REMARKS

Gram-positive pathogens utilize their interactions with the ECM for tissue colonization and to establish infections in the host. Molecular insight into these interactions can pave the way for the design of novel anti-infectives. However, studies of collagen-binding proteins in Gram-positive pathogens are in their infancy and do not provide a complete picture of the different binding mechanisms involved. Further structural studies are required to fully understand the molecular basis for the interaction between bacterial collagen-binding proteins and the triple helix of collagen. In particular, the interaction of emerging collagen-binding protein families with collagen needs to be further characterized using biochemical and microbiological techniques to determine which family members bind collagen.

Mammalian proteins containing a collagen-like region play a role in host defense. Collagen-binding host proteins, e.g., the LRR proteoglycan decorin, bind soluble host defense collagens (Krumdieck et al., 1992). Interaction of bacterial collagen-binding proteins with soluble defense collagens can provide an opportunity for pathogens to evade the host immune response. CNA binds C1q, a complement protein and collagen (Kang et al., 2013). The classical complement pathway is initiated upon recognition of pathogen-bound antibodies by the C1 complex, which consists of C1q, C1r, and C1s. C1q protein contains the globular recognition domain and binds pathogen-bound antibodies. C1r and C1s are proteases that are required for the complement cascade. C1r and C1s bind the collagen-like stalk of C1q (Mortensen et al., 2017). CNA uses its interaction with C1q for immune evasion by interfering with the interaction between C1r and C1q and thus deactivating the C1 complex (Kang et al., 2013).

Interaction of collagen-binding bacterial proteins with other host proteins containing collagen like-regions, especially soluble human defense collagens, is an understudied area. While acknowledging that not all collagen-binding proteins will bind soluble defense collagens and vice versa, future studies focusing on the interaction between bacterial collagen binding proteins and host defense collagens will lead to a better understanding of the pathogenic mechanisms utilized by Gram-positive bacteria.

AUTHOR CONTRIBUTIONS

SA and JG wrote the manuscript. SA and MH edited the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by NIH R01 grant AI020624-35.

ACKNOWLEDGMENTS

We would like to thank Dr. Stefan Siwko for his assistance in editing the review article.

REFERENCES

Abranches, J., Miller, J. H., Martinez, A. R., Simpson-Haidaris, P. J., Burne, R. A., and Lemos, J. A. (2011). The collagen-binding protein Cnm is required for Streptococcus mutans adherence to and intracellular invasion of human coronary artery endothelial cells. *Infect. Immun.* 79, 2277–2284. doi: 10.1128/iai.00767-10

Arvanitidis, A., and Karsdal, M. A. (2016). "Type XV Collagen," in *Biochemistry of Collagens, Laminins and Elastin*, Chap. 15, ed. M. A. Karsdal (Cambridge, MA: Academic Press), 97–99.

Avilés-Reyes, A., Miller, J. H., Lemos, J. A., and Abranches, J. (2017). Collagen-binding proteins of Streptococcus mutans and related streptococci. *Mol. Oral. Microbiol.* 32, 89–106. doi: 10.1111/omi.12158

Bager, C. L., and Karsdal, M. A. (2016). "Type XVIII Collagen," in *Biochemistry of Collagens, Laminins and Elastin*, Chap. 18, ed. M. A. Karsdal (Cambridge, MA: Academic Press), 113–121.

Banerjee, A., Kim, B. J., Carmona, E. M., Cutting, A. S., Gurney, M. A., Carlos, C., et al. (2011). Bacterial Pili exploit integrin machinery to promote immune activation and efficient blood-brain barrier penetration. *Nat. Commun.* 2:462. doi: 10.1038/ncomms1474

Barroso, V., Rohde, M., Davies, M. R., Gillen, C. M., Nitsche-Schmitz, D. P., Dinkla, K., et al. (2009). Identification of active variants of PARF in human pathogenic group C and group G streptococci leads to an amended description of its consensus motif. *Int. J. Med. Microbiol.* 299, 547–553. doi: 10.1016/j.ijmm.2009.04.004

Becke, T. (2019). *Streptococcus Pneumoniae TIGR4 Pilus-1 Biomechanical Aspects of Adhesion During Interaction With Host Extracellular Matrix Proteins Fibronectin and Collagen I*. München: Technische Universität München.

Bober, M., Enochsson, C., Collin, M., and Mörgelin, M. (2010). Collagen VI is a subepithelial adhesive target for human respiratory tract pathogens. *J. Innate Immun.* 2, 160–166. doi: 10.1159/000232587

Bober, M., Mörgelin, M., Olin, A. I., von Pawel-Rammingen, U., and Collin, M. (2011). The membrane bound LRR lipoprotein Slr, and the cell wall-anchored M1 protein from *Streptococcus pyogenes* both interact with type I collagen. *PLoS One* 6:e20345. doi: 10.1371/journal.pone.0020345

Boudko, S. P., and Bächinger, H. P. (2016). Structural insight for chain selection and stagger control in collagen. *Sci. Rep.* 6:37831. doi: 10.1038/srep37831

Burgeson, R. E., and Nimni, M. E. (1992). Collagen types. molecular structure and tissue distribution. *Clin. Orthop. Relat. Res.* 282, 250–272.

Carapetis, J. R., Beaton, A., Cunningham, M. W., Guilherme, L., Karthikeyan, G., Mayosi, B. M., et al. (2016). Acute rheumatic fever and rheumatic heart disease. *Nat. Rev. Dis. Primers* 2:15084. doi: 10.1038/nrdp.2015.84

Carret, G., Emonard, H., Fardel, G., Druguet, M., Herbage, D., and Flandrois, J. P. (1985). Gelatin and collagen binding to *Staphylococcus aureus* strains. *Ann. Inst. Pasteur. Microbiol. (1985)* 136a, 241–245.

Casals, C., Campanero-Rhodes, M. A., García-Fojeda, B., and Solís, D. (2018). The role of collectins and galectins in lung innate immune defense. *Front. Immunol.* 9:1998. doi: 10.3389/fimmu.2018.01998

Casals, C., García-Fojeda, B., and Minutti, C. M. (2019). Soluble defense collagens: sweeping up immune threats. *Mol. Immunol.* 112, 291–304. doi: 10.1016/j.molimm.2019.06.007

Deivanayagam, C. C., Rich, R. L., Carson, M., Owens, R. T., Danthuluri, S., Bice, T., et al. (2000). Novel fold and assembly of the repetitive B region of the *Staphylococcus aureus* collagen-binding surface protein. *Structure* 8, 67–78.

Desvaux, M., Dumas, E., Chafsey, I., and Hébraud, M. (2006). Protein cell surface display in Gram-positive bacteria: from single protein to macromolecular protein structure. *FEMS Microbiol. Lett.* 256, 1–15. doi: 10.1111/j.1574-6968.2006.00122.x

Dinkla, K., Nitsche-Schmitz, D. P., Barroso, V., Reissmann, S., Johansson, H. M., Frick, I. M., et al. (2007). Identification of a streptococcal octapeptide motif involved in acute rheumatic fever. *J. Biol. Chem.* 282, 18686–18693. doi: 10.1074/jbc.M701047200

Dinkla, K., Rohde, M., Jansen, W. T., Carapetis, J. R., Chhatwal, G. S., and Talay, S. R. (2003a). *Streptococcus pyogenes* recruits collagen via surface-bound fibronectin: a novel colonization and immune evasion mechanism. *Mol. Microbiol.* 47, 861–869. doi: 10.1046/j.1365-2958.2003.03352.x

Dinkla, K., Rohde, M., Jansen, W. T. M., Kaplan, E. L., Chhatwal, G. S., and Talay, S. R. (2003b). Rheumatic fever-associated *Streptococcus pyogenes* isolates aggregate collagen. *J. Clin. Invest.* 111, 1905–1912. doi: 10.1172/JCI17247

Dinkla, K., Talay, S. R., Mörgelin, M., Graham, R. M. A., Rohde, M., Nitsche-Schmitz, D. P., et al. (2009). Crucial role of the CB3-region of collagen IV in PARF-induced acute rheumatic fever. *PLoS One* 4:e4666. doi: 10.1371/journal.pone.0004666

Dramsi, S., Morello, E., Poyart, C., and Trieu-Cuot, P. (2012). Epidemiologically and clinically relevant Group B Streptococcus isolates do not bind collagen but display enhanced binding to human fibrinogen. *Microbes Infect.* 14, 1044–1048. doi: 10.1016/j.micinf.2012.07.004

Duarte, A. S., Correia, A., and Esteves, A. C. (2016). Bacterial collagenases – A review. *Crit. Rev. Microbiol.* 42, 106–126. doi: 10.3109/1040841X.2014.904270

Eble, J. A., Golbik, R., Mann, K., and Kühn, K. (1993). The alpha 1 beta 1 integrin recognition site of the basement membrane collagen molecule [alpha 1(IV)]2 alpha 2(IV). *Embo J.* 12, 4795–4802.

Elasri, M. O., Thomas, J. R., Skinner, R. A., Blevins, J. S., Beenken, K. E., Nelson, C. L., et al. (2002). *Staphylococcus aureus* collagen adhesin contributes to the pathogenesis of osteomyelitis. *Bone* 30, 275–280.

Emsley, J., Cruz, M., Handin, R., and Liddington, R. (1998). Crystal structure of the von Willebrand Factor A1 domain and implications for the binding of platelet glycoprotein Ib. *J. Biol. Chem.* 273, 10396–10401. doi: 10.1074/jbc.273.17.10396

Emsley, J., Knight, C. G., Farndale, R. W., Barnes, M. J., and Liddington, R. C. (2000). Structural basis of collagen recognition by integrin alpha2beta1. *Cell* 101, 47–56. doi: 10.1016/s0092-8674(00)80622-4

Fischetti, V. A. (1989). Streptococcal M protein: molecular design and biological behavior. *Clin. Microbiol. Rev.* 2, 285–314. doi: 10.1128/cmr.2.3.285

Fischetti, V. A. (2016). *M Protein and Other Surface Proteins on Streptococci*. Oklahoma: University of Oklahoma Helath Sciences Center.

Fischetti, V. A. (2019). Surface proteins on gram-positive bacteria. *Microbiol. Spectr.* 7, GPP3-0012-2018. doi: 10.1128/microbiolspec.GPP3-0012-2018

Fitzgerald, J., Holden, P., and Hansen, U. (2013). The expanded collagen VI family: new chains and new questions. *Connect Tissue Res.* 54, 345–350. doi: 10.3109/03008207.2013.822865

Font, B., Eichenberger, D., Goldschmidt, D., Boutillon, M. M., and Hulmes, D. J. (1998). Structural requirements for fibromodulin binding to collagen and the control of type I collagen fibrillogenesis–critical roles for disulphide bonding and the C-terminal region. *Eur. J. Biochem.* 254, 580–587. doi: 10.1046/j.1432-1327.1998.2540580.x

Forsgren, N., Lamont, R. J., and Persson, K. (2009). Crystal structure of the variable domain of the *Streptococcus gordonii* surface protein SspB. *Protein Sci.* 18, 1896–1905. doi: 10.1002/pro.200

Foster, T. J. (2019). The MSCRAMM family of cell-wall-anchored surface proteins of gram-positive cocci. *Trends Microbiol.* 27, 927–941. doi: 10.1016/j.tim.2019.06.007

Foster, T. J., Geoghegan, J. A., Ganesh, V. K., and Höök, M. (2014). Adhesion, invasion and evasion: the many functions of the surface proteins of *Staphylococcus aureus*. *Nat. Rev. Microbiol.* 12, 49–62. doi: 10.1038/nrmicro3161

Frantz, C., Stewart, K. M., and Weaver, V. M. (2010). The extracellular matrix at a glance. *J. Cell Sci.* 123(Pt 24), 4195–4200. doi: 10.1242/jcs.023820

Fraser, D. A., and Tenner, A. J. (2008). Directing an appropriate immune response: the role of defense collagens and other soluble pattern recognition molecules. *Curr. Drug Targets* 9, 113–122. doi: 10.2174/138945008783502476

Freires, I. A., Avilés-Reyes, A., Kitten, T., Simpson-Haidaris, P. J., Swartz, M., Knight, P. A., et al. (2017). Heterologous expression of *Streptococcus mutans* Cnm in *Lactococcus lactis* promotes intracellular invasion, adhesion to human cardiac tissues and virulence. *Virulence* 8, 18–29. doi: 10.1080/21505594.2016.1195538

Frost, H. R., Sanderson-Smith, M., Walker, M., Botteaux, A., and Smeesters, P. R. (2017). Group A streptococcal M-like proteins: from pathogenesis to vaccine potential. *FEMS Microbiol. Rev.* 42, 193–204. doi: 10.1093/femsre/fux057

Gebauer, J. M., Kobbe, B., Paulsson, M., and Wagener, R. (2016). Structure, evolution and expression of collagen XXVIII: lessons from the zebrafish. *Matrix Biol.* 49, 106–119. doi: 10.1016/j.matbio.2015.07.001

Geraci, J., Neubauer, S., Pöllath, C., Hansen, U., Rizzo, F., Krafft, C., et al. (2017). The *Staphylococcus aureus* extracellular matrix protein (Emp) has a fibrous structure and binds to different extracellular matrices. *Sci. Rep.* 7:13665. doi: 10.1038/s41598-017-14168-4

Gonçalves, T. J. M., Boutillon, F., Lefebvre, S., Goffin, V., Iwatsubo, T., Wakabayashi, T., et al. (2019). Collagen XXV promotes myoblast fusion during myogenic differentiation and muscle formation. *Sci. Rep.* 9:5878. doi: 10.1038/s41598-019-42296-6

Gudmann, N. S., and Karsdal, M. A. (2016). "Type II collagen," in *Biochemistry of Collagens, Laminins and Elastin*, Chap. 2, ed. M. A. Karsdal (Cambridge, MA: Academic Press), 13–20.

Hansen, N. U. B., and Karsdal, M. A. (2016). "Type VIII collagen," in *Biochemistry of Collagens, Laminins and Elastin*, Chap. 8, ed. M. A. Karsdal (Cambridge, MA: Academic Press), 61–65.

Harrington, D. J. (1996). Bacterial collagenases and collagen-degrading enzymes and their potential role in human disease. *Infect. Immun.* 64, 1885–1891. doi: 10.1128/iai.64.6.1885-1891.1996

He, Y., and Karsdal, M. A. (2016). "Type IX collagen," in *Biochemistry of Collagens, Laminins and Elastin*, Chap. 9, ed. M. A. Karsdal (Cambridge, MA: Academic Press), 67–71.

Henriksen, K., and Karsdal, M. A. (2016). "Type I collagen," in *Biochemistry of Collagens, Laminins and Elastin*, Chap. 1, ed. M. A. Karsdal (Cambridge, MA: Academic Press), 1–11.

Herman-Bausier, P., Valotteau, C., Pietrocola, G., Rindi, S., Alsteens, D., Foster, T. J., et al. (2016). Mechanical strength and inhibition of the *Staphylococcus aureus* collagen-binding protein Cna. *mBio* 7:e1529–16. doi: 10.1128/mBio.01529-16

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

Hilleringmann, M., Giusti, F., Baudner, B. C., Masignani, V., Covacci, A., Rappuoli, R., et al. (2008). Pneumococcal Pili are composed of protofilaments exposing adhesive clusters of Rrg A. *PLoS Pathog.* 4:e1000026. doi: 10.1371/journal.ppat.1000026

Hohenester, E., and Yurchenco, P. D. (2013). Laminins in basement membrane assembly. *Cell Adh. Migr.* 7, 56–63. doi: 10.4161/cam.21831

Holderbaum, D., Spech, R. A., and Ehrhart, L. A. (1985). Specific binding of collagen to *Staphylococcus aureus*. *Coll. Relat. Res.* 5, 261–271.

Holmes, D. F., Lu, Y., Starborg, T., and Kadler, K. E. (2018). "Collagen fibril assembly and function," in *Current Topics in Developmental Biology*, Chap. Three, eds E. S. Litscher and P. M. Wassarman (Cambridge, MA: Academic Press), 107–142.

Huizinga, E. G., van der Plas, R. M., Kroon, J., Sixma, J. J., and Gros, P. (1997). Crystal structure of the A3 domain of human von Willebrand factor: implications for collagen binding. *Structure* 5, 1147–1156. doi: 10.1016/s0969-2126(97)00266-9

Izore, T., Contreras-Martel, C., El Mortaji, L., Manzano, C., Terrasse, R., Vernet, T., et al. (2010). Structural basis of host cell recognition by the pilus adhesin from *Streptococcus pneumoniae*. *Structure* 18, 106–115. doi: 10.1016/j.str.2009.10.019

Izu, Y., Adams, S. M., Connizzo, B. K., Beason, D. P., Soslowsky, L. J., Koch, M., et al. (2020). Collagen XII mediated cellular and extracellular mechanisms regulate establishment of tendon structure and function. *Matrix Biol.* doi: 10.1016/j.matbio.2020.10.004 [Epub ahead of print].

Kadler, K. E. (2017). Fell muir lecture: collagen fibril formation *in vitro* and in vivo. *Int. J. Exp. Pathol.* 98, 4–16. doi: 10.1111/iep.12224

Kadler, K. E., Baldock, C., Bella, J., and Boot-Handford, R. P. (2007). Collagens at a glance. *J. Cell Sci.* 120:1955. doi: 10.1242/jcs.03453

Kalamajski, S., and Oldberg, Å (2010). The role of small leucine-rich proteoglycans in collagen fibrillogenesis. *Matrix Biol.* 29, 248–253. doi: 10.1016/j.matbio.2010.01.001

Kang, M., Ko, Y. P., Liang, X., Ross, C. L., Liu, Q., Murray, B. E., et al. (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. doi: 10.1074/jbc.M113.454462

Kehlet, S. N., and Karsdal, M. A. (2016). "Type XXI collagen," in *Biochemistry of Collagens, Laminins and Elastin*, Chap. 21, ed. M. A. Karsdal (Cambridge, MA: Academic Press), 131–133.

Kobe, B., and Kajava, A. V. (2001). The leucine-rich repeat as a protein recognition motif. *Curr. Opin. Struct. Biol.* 11, 725–732. doi: 10.1016/S0959-440X(01)00266-4

Konto-Ghiorghi, Y., Mairey, E., Mallet, A., Duménil, G., Caliot, E., Trieu-Cuot, P., et al. (2009). Dual role for pilus in adherence to epithelial cells and biofilm formation in *Streptococcus agalactiae*. *PLoS Pathog.* 5:e1000422. doi: 10.1371/journal.ppat.1000422

Krishnan, V., Dwivedi, P., Kim, B. J., Samal, A., Macon, K., Ma, X., et al. (2013). Structure of *Streptococcus agalactiae* tip pilin GBS104: a model for GBS pili assembly and host interactions. *Acta Crystallogr. D Biol. Crystallogr.* 69(Pt 6), 1073–1089. doi: 10.1107/s0907444913004642

Krumdieck, R., Höök, M., Rosenberg, L. C., and Volanakis, J. E. (1992). The proteoglycan decorin binds C1q and inhibits the activity of the C1 complex. *J. Immunol.* 149, 3695–3701.

Lancefield, R. C. (1928). The antigenic complex of *Streptococcus haemolyticus* : I. demonstration of a type-specific substance in extracts of *Streptococcus haemolyticus*. *J. Exp. Med.* 47, 91–103. doi: 10.1084/jem.47.1.91

Lannergard, J., Frykberg, L., and Guss, B. (2003). CNE, a collagen-binding protein of *Streptococcus equi*. *FEMS Microbiol. Lett.* 222, 69–74.

Larson, M. R., Rajashankar, K. R., Patel, M. H., Robinette, R. A., Crowley, P. J., Michalek, S., et al. (2010). Elongated fibrillar structure of a streptococcal adhesin assembled by the high-affinity association of alpha- and PPII-helices. *Proc. Natl. Acad. Sci. U.S.A.* 107, 5983–5988. doi: 10.1073/pnas.0912293107

Lee, J. O., Rieu, P., Arnaout, M. A., and Liddington, R. (1995). Crystal structure of the A domain from the alpha subunit of integrin CR3 (CD11b/CD18). *Cell* 80, 631–638. doi: 10.1016/0092-8674(95)90517-0

Leeming, D. J., and Karsdal, M. A. (2016). "Type V collagen," in *Biochemistry of Collagens, Laminins and Elastin*, Chap. 5, ed. M. A. Karsdal (Cambridge, MA: Academic Press), 43–48.

Li, M., Shao, Z.-Q., Guo, Y., Wang, L., Hou, T., Hu, D., et al. (2015). The type II histidine triad protein HtpsC is a novel adhesion with the involvement of *Streptococcus suis* virulence. *Virulence* 6, 631–641. doi: 10.1080/21505594.2015.1056971

Liu, Q., Ponnuraj, K., Xu, Y., Ganesh, V. K., Sillanpaa, J., Murray, B. E., et al. (2007). The *Enterococcus faecalis* MSCRAMM ACE binds its ligand by the Collagen Hug model. *J. Biol. Chem.* 282, 19629–19637. doi: 10.1074/jbc.M611137200

Lizcano, A., Sanchez, C. J., and Orihuela, C. J. (2012). A role for glycosylated serine-rich repeat proteins in Gram-positive bacterial pathogenesis. *Mol. Oral Microbiol.* 27, 257–269. doi: 10.1111/j.2041-1014.2012.00653.x

Love, R. M., McMillan, M. D., and Jenkinson, H. F. (1997). Invasion of dentinal tubules by oral streptococci is associated with collagen recognition mediated by the antigen I/II family of polypeptides. *Infect. Immun.* 65, 5157–5164. doi: 10.1128/iai.65.12.5157-5164.1997

Love, R. M., McMillan, M. D., Park, Y., and Jenkinson, H. F. (2000). Coinvasion of dentinal tubules by *Porphyromonas gingivalis* and *Streptococcus gordonii* depends upon binding specificity of streptococcal antigen I/II adhesin. *Infect. Immun.* 68, 1359–1365. doi: 10.1128/iai.68.3.1359-1365.2000

Luo, Y., Sinkeviciute, D., He, Y., Karsdal, M., Henrotin, Y., Mobasheri, A., et al. (2017). The minor collagens in articular cartilage. *Protein Cell* 8, 560–572. doi: 10.1007/s13238-017-0377-7

Luo, Y. Y., and Karsdal, M. A. (2016). "Type XI collagen," in *Biochemistry of Collagens, Laminins and Elastin*, Chap. 11, ed. M. A. Karsdal (Cambridge, MA: Academic Press), 77–80.

Macheboeuf, P., Buffalo, C., Fu, C.-Y., Zinkernagel, A. S., Cole, J. N., Johnson, J. E., et al. (2011). Streptococcal M1 protein constructs a pathological host fibrinogen network. *Nature* 472, 64–68. doi: 10.1038/nature09967

Madani, A., Garakani, K., and Mofrad, M. R. K. (2017). Molecular mechanics of *Staphylococcus aureus* adhesin, CNA, and the inhibition of bacterial adhesion by stretching collagen. *PLoS One* 12:e0179601. doi: 10.1371/journal.pone.0179601

Mandlik, A., Swierczynski, A., Das, A., and Ton-That, H. (2007). *Corynebacterium diphtheriae* employs specific minor pilins to target human pharyngeal epithelial cells. *Mol. Microbiol.* 64, 111–124. doi: 10.1111/j.1365-2958.2007.05630.x

Manon-Jensen, T., and Karsdal, M. A. (2016). "Type XIV collagen," in *Biochemistry of Collagens, Laminins and Elastin*, Chap. 14, ed. M. A. Karsdal (Cambridge, MA: Academic Press), 93–95.

Manon-Jensen, T., Kjeld, N. G., and Karsdal, M. A. (2016). Collagen-mediated hemostasis. *J. Thromb. Haemost.* 14, 438–448. doi: 10.1111/jth.13249

McNamara, C., Zinkernagel, A. S., Macheboeuf, P., Cunningham, M. W., Nizet, V., and Ghosh, P. (2008). Coiled-coil irregularities and instabilities in group a *Streptococcus* M1 are required for virulence. *Science* 319:1405. doi: 10.1126/science.1154470

Metzgar, D., and Zampolli, A. (2011). The M protein of group A *Streptococcus* is a key virulence factor and a clinically relevant strain identification marker. *Virulence* 2, 402–412. doi: 10.4161/viru.2.5.16342

Miller, J. H., Avilés-Reyes, A., Scott-Anne, K., Gregoire, S., Watson, G. E., Sampson, E., et al. (2015). The collagen binding protein Cnm contributes to oral colonization and cariogenicity of *Streptococcus mutans* OMZ175. *Infect. Immun.* 83:2001. doi: 10.1128/IAI.03022-14

Montanaro, L., Arciola, C. R., Baldassarri, L., and Borsetti, E. (1999). Presence and expression of collagen adhesin gene (cna) and slime production in *Staphylococcus aureus* strains from orthopaedic prosthesis infections. *Biomaterials* 20, 1945–1949. doi: 10.1016/s0142-9612(99)00099-x

Mortensen, J. H., and Karsdal, M. A. (2016). "Type VII collagen," in *Biochemistry of Collagens, Laminins and Elastin*, Chap. 7, ed. M. A. Karsdal (Cambridge, MA: Academic Press), 57–60.

Mortensen, S. A., Sander, B., Jensen, R. K., Pedersen, J. S., Golas, M. M., Jensenius, J. C., et al. (2017). Structure and activation of C1, the complex initiating the classical pathway of the complement cascade. *Proc. Natl. Acad. Sci. U.S.A.* 114, 986–991.

Moschioni, M., Emolo, C., Biagini, M., Maccari, S., Pansegrau, W., Donati, C., et al. (2010). The two variants of the *Streptococcus pneumoniae* pilus 1 RrgA adhesin retain the same function and elicit cross-protection in vivo. *Infect. Immun.* 78, 5033–5042. doi: 10.1128/iai.00601-10

Nallapareddy, S. R., Qin, X., Weinstock, G. M., Höök, M., and Murray, B. E. (2000). *Enterococcus faecalis* adhesin, ace, mediates attachment to extracellular matrix proteins collagen type IV and laminin as well as collagen type I. *Infect. Immun.* 68, 5218–5224. doi: 10.1128/iai.68.9.5218-5224.2000

Nallapareddy, S. R., Singh, K. V., and Murray, B. E. (2008). Contribution of the collagen adhesin acm to pathogenesis of *Enterococcus faecium* in experimental endocarditis. *Infect. Immun.* 76, 4120–4128. doi: 10.1128/iai.00376-08

Nallapareddy, S. R., Weinstock, G. M., and Murray, B. E. (2003). Clinical isolates of *Enterococcus faecium* exhibit strain-specific collagen binding mediated by Acm, a new member of the MSCRAMM family. *Mol. Microbiol.* 47, 1733–1747.

Nielsen, H. V., Guiton, P. S., Kline, K. A., Port, G. C., Pinkner, J. S., Neiers, F., et al. (2012). The metal ion-dependent adhesion site motif of the *Enterococcus faecalis* EbpA pilin mediates pilus function in catheter-associated urinary tract infection. *mBio* 3, e177–12. doi: 10.1128/mBio.00177-12

Nielsen, M. J., and Karsdal, M. A. (2016a). "Type III collagen," in *Biochemistry of Collagens, Laminins and Elastin*, Chap. 3, ed. M. A. Karsdal (Cambridge, MA: Academic Press), 21–30.

Nielsen, S. H., and Karsdal, M. A. (2016b). "Type XIX collagen," in *Biochemistry of Collagens, Laminins and Elastin*, Chap. 19, ed. M. A. Karsdal (Cambridge, MA: Academic Press), 123–125.

Nitsche, D. P., Johansson, H. M., Frick, I. M., and Mörgelin, M. (2006). Streptococcal protein FOG, a novel matrix adhesin interacting with collagen I in vivo. *J. Biol. Chem.* 281, 1670–1679. doi: 10.1074/jbc.M506776200

Nomura, R., Naka, S., Nemoto, H., Inagaki, S., Taniguchi, K., Ooshima, T., et al. (2013). Potential involvement of collagen-binding proteins of *Streptococcus mutans* in infective endocarditis. *Oral Dis.* 19, 387–393. doi: 10.1111/odi.12016

Nomura, R., Nakano, K., Naka, S., Nemoto, H., Masuda, K., Lapirattanakul, J., et al. (2012). Identification and characterization of a collagen-binding protein, Cbm, in *Streptococcus mutans*. *Mol. Oral Microbiol.* 27, 308–323. doi: 10.1111/j.2041-1014.2012.00649.x

Nomura, R., Ogaya, Y., and Nakano, K. (2016). Contribution of the collagen-binding proteins of *Streptococcus mutans* to bacterial colonization of inflamed dental pulp. *PLoS One* 11:e0159613. doi: 10.1371/journal.pone.0159613

Nylander, Å, Svensäter, G., Senadheera, D. B., Cvitkovitch, D. G., Davies, J. R., and Persson, K. (2013). Structural and functional analysis of the N-terminal domain of the *Streptococcus gordonii* adhesin Sgo0707. *PLoS One* 8:e63768. doi: 10.1371/journal.pone.0063768

Oehmcke, S., Shannon, O., Mörgelin, M., and Herwald, H. (2010). Streptococcal M proteins and their role as virulence determinants. *Clin. Chim. Acta* 411, 1172–1180. doi: 10.1016/j.cca.2010.04.032

Patti, J. M., Allen, B. L., McGavin, M. J., and Höök, M. (1994a). MSCRAMM-mediated adherence of microorganisms to host tissues. *Annu. Rev. Microbiol.* 48, 585–617. doi: 10.1146/annurev.mi.48.100194.003101

Patti, J. M., Boles, J. O., and Hook, M. (1993). Identification and biochemical characterization of the ligand binding domain of the collagen adhesin from *Staphylococcus aureus*. *Biochemistry* 32, 11428–11435.

Patti, J. M., Bremell, T., Krajewska-Pietrasik, D., Abdelnour, A., Tarkowski, A., Ryden, C., et al. (1994b). The *Staphylococcus aureus* collagen adhesin is a virulence determinant in experimental septic arthritis. *Infect. Immun.* 62, 152–161.

Patti, J. M., House-Pompeo, K., Boles, J. O., Garza, N., Gurusiddappa, S., and Hook, M. (1995). Critical residues in the ligand-binding site of the *Staphylococcus aureus* collagen-binding adhesin (MSCRAMM). *J. Biol. Chem.* 270, 12005–12011.

Paulsson, M., and Riesbeck, K. (2018). How bacteria hack the matrix and dodge the bullets of immunity. *Eur. Respir. Rev.* 27:180018. doi: 10.1183/16000617.0018-2018

Phillips, G. N. Jr., Flicker, P. F., Cohen, C., Manjula, B. N., and Fischetti, V. A. (1981). Streptococcal M protein: alpha-helical coiled-coil structure and arrangement on the cell surface. *Proc. Natl. Acad. Sci. U.S.A.* 78, 4689–4693. doi: 10.1073/pnas.78.8.4689

Plumptre, C. D., Ogunniyi, A. D., and Paton, J. C. (2012). Polyhistidine triad proteins of pathogenic streptococci. *Trends Microbiol.* 20, 485–493. doi: 10.1016/j.tim.2012.06.004

PrabhuDas, M. R., Baldwin, C. L., Bollyky, P. L., Bowdish, D. M. E., Drickamer, K., Febbraio, M., et al. (2017). A consensus definitive classification of scavenger receptors and their roles in health and disease. *J. Immunol.* 198, 3775–3789. doi: 10.4049/jimmunol.1700373

Pyagay, P., Heroult, M., Wang, Q., Lehnert, W., Belden, J., Liaw, L., et al. (2005). Collagen triple helix repeat containing 1, a novel secreted protein in injured and diseased arteries, inhibits collagen expression and promotes cell migration. *Circ. Res.* 96, 261–268. doi: 10.1161/01.RES.0000154262.07264.12

Reissmann, S., Gillen, C. M., Fulde, M., Bergmann, R., Nerlich, A., Rajkumari, R., et al. (2012). Region specific and worldwide distribution of collagen-binding M proteins with PARF motifs among human pathogenic streptococcal isolates. *PLoS One* 7:e30122. doi: 10.1371/journal.pone.0030122

Rhem, M. N., Lech, E. M., Patti, J. M., McDevitt, D., Höök, M., Jones, D. B., et al. (2000). The collagen-binding adhesin is a virulence factor in *Staphylococcus aureus* keratitis. *Infect. Immun.* 68, 3776–3779. doi: 10.1128/iai.68.6.3776-3779.2000

Ricard-Blum, S. (2011). The collagen family. *Cold Spring Harb. Perspect. Biol.* 3:a004978. doi: 10.1101/cshperspect.a004978

Rich, R. L., Kreikemeyer, B., Owens, R. T., LaBrenz, S., Narayana, S. V., Weinstock, G. M., et al. (1999). Ace is a collagen-binding MSCRAMM from *Enterococcus faecalis*. *J. Biol. Chem.* 274, 26939–26945.

Ross, C. L., Liang, X., Liu, Q., Murray, B. E., Hook, M., and Ganesh, V. K. (2012). Targeted protein engineering provides insights into binding mechanism and affinities of bacterial collagen adhesins. *J. Biol. Chem.* 287, 34856–34865. doi: 10.1074/jbc.M112.371054

Sand, J. M. B., Genovese, F., and Karsdal, M. A. (2016). "Type IV collagen," in *Biochemistry of Collagens, Laminins and Elastin*, Chap. 4, ed. M. A. Karsdal (Cambridge, MA: Academic Press), 31–41.

Sand, J. M. B., and Karsdal, M. A. (2016). "Type XVI collagen," in *Biochemistry of Collagens, Laminins and Elastin*, Chap. 16, ed. M. A. Karsdal (Cambridge, MA: Academic Press), 101–106.

Sato, K., Yomogida, K., Wada, T., Yorihuzi, T., Nishimune, Y., Hosokawa, N., et al. (2002). Type XXVI collagen, a new member of the collagen family, is specifically expressed in the testis and ovary. *J. Biol. Chem.* 277, 37678–37684. doi: 10.1074/jbc.M205347200

Sato, Y., Okamoto, K., Kagami, A., Yamamoto, Y., Igarashi, T., and Kizaki, H. (2004). *Streptococcus mutans* strains harboring collagen-binding adhesin. *J. Dent. Res.* 83, 534–539. doi: 10.1177/154405910408300705

Schönborn, K., Willenborg, S., Schulz, J. N., Imhof, T., Eming, S. A., Quondamatteo, F., et al. (2020). Role of collagen XII in skin homeostasis and repair. *Matrix Biol.* 94, 57–76. doi: 10.1016/j.matbio.2020.08.002

Schönherr, E., Hausser, H., Beavan, L., and Kresse, H. (1995). Decorin-type I collagen interaction. Presence of separate core protein-binding domains. *J. Biol. Chem.* 270, 8877–8883. doi: 10.1074/jbc.270.15.8877

Shen, G. (2005). The role of type X collagen in facilitating and regulating endochondral ossification of articular cartilage. *Orthod Craniofac. Res.* 8, 11–17. doi: 10.1111/j.1601-6343.2004.00308.x

Shimoji, Y., Ogawa, Y., Osaki, M., Kabeya, H., Maruyama, S., Mikami, T., et al. (2003). Adhesive surface proteins of *Erysipelothrix rhusiopathiae* bind

to polystyrene, fibronectin, and type I and IV collagens. *J. Bacteriol.* 185, 2739–2748.

Shoulders, M. D., and Raines, R. T. (2009). Collagen structure and stability. *Annu. Rev. Biochem.* 78, 929–958. doi: 10.1146/annurev.biochem.77.032207.120833

Sillanpää, J., Nallapareddy, S. R., Qin, X., Singh, K. V., Muzny, D. M., Kovar, C. L., et al. (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. doi: 10.1128/jb.00909-09

Singh, B., Fleury, C., Jalalvand, F., and Riesbeck, K. (2012). Human pathogens utilize host extracellular matrix proteins laminin and collagen for adhesion and invasion of the host. *FEMS Microbiol. Rev.* 36, 1122–1180. doi: 10.1111/j.1574-6976.2012.00340.x

Singh, K. V., Nallapareddy, S. R., Sillanpää, J., and Murray, B. E. (2010). Importance of the collagen adhesin ace in pathogenesis and protection against *Enterococcus faecalis* experimental endocarditis. *PLoS Pathog.* 6:e1000716. doi: 10.1371/journal.ppat.1000716

Smeesters, P. R., McMillan, D. J., and Sriprakash, K. S. (2010). The streptococcal M protein: a highly versatile molecule. *Trends Microbiol.* 18, 275–282. doi: 10.1016/j.tim.2010.02.007

Speziale, P., Raucci, G., Visai, L., Switalski, L. M., Timpl, R., and Hook, M. (1986). Binding of collagen to *Staphylococcus aureus* Cowan 1. *J. Bacteriol.* 167, 77–81.

Spivey, K. A., Chung, I., Banyard, J., Adini, I., Feldman, H. A., and Zetter, B. R. (2012). A role for collagen XXIII in cancer cell adhesion, anchorage-independence and metastasis. *Oncogene* 31, 2362–2372. doi: 10.1038/onc.2011.406

Stewart, C. M., Buffalo, C. Z., Valderrama, J. A., Henningham, A., Cole, J. N., Nizet, V., et al. (2016). Coiled-coil destabilizing residues in the group A Streptococcus M1 protein are required for functional interaction. *Proc. Natl. Acad. Sci. U.S.A.* 113, 9515–9520.

Sun, S., and Karsdal, M. A. (2016a). "Type VI collagen," in *Biochemistry of Collagens, Laminins and Elastin*, Chap. 6, ed. M. A. Karsdal (Cambridge, MA: Academic Press), 49–55.

Sun, S., and Karsdal, M. A. (2016b). "Type XVII collagen," in *Biochemistry of Collagens, Laminins and Elastin*, Chap. 17, ed. M. A. Karsdal (Cambridge, MA: Academic Press), 107–111.

Sundaramoorthy, M., Meiyappan, M., Todd, P., and Hudson, B. G. (2002). Crystal structure of NC1 domains. Structural basis for type IV collagen assembly in basement membranes. *J. Biol. Chem.* 277, 31142–31153. doi: 10.1074/jbc.M201740200

Taglialegna, A., Matilla-Cuenca, L., Dorado-Morales, P., Navarro, S., Ventura, S., Garnett, J. A., et al. (2020). The biofilm-associated surface protein Esp of *Enterococcus faecalis* forms amyloid-like fibers. *NPJ Biofilms Microbiomes* 6:15. doi: 10.1038/s41522-020-0125-2

Tanaka, T., Wakabayashi, T., Oizumi, H., Nishio, S., Sato, T., Harada, A., et al. (2014). CLAC-P/Collagen Type XXV is required for the intramuscular innervation of motoneurons during neuromuscular development. *J. Neurosci.* 34:1370. doi: 10.1523/JNEUROSCI.2440-13.2014

Tandon, R., Sharma, M., Chandrashekhar, Y., Kotb, M., Yacoub, M. H., and Narula, J. (2013). Revisiting the pathogenesis of rheumatic fever and carditis. *Nat. Rev. Cardiol.* 10, 171–177. doi: 10.1038/nrcardio.2012.197

Theocharis, A. D., Skandalis, S. S., Gialeli, C., and Karamanos, N. K. (2016). Extracellular matrix structure. *Adv. Drug Deliv. Rev.* 97, 4–27. doi: 10.1016/j.addr.2015.11.001

Tom Tang, Y., Hu, T., Arterburn, M., Boyle, B., Bright, J. M., Palencia, S., et al. (2005). The complete complement of C1q-domain-containing proteins in Homo sapiens. *Genomics* 86, 100–111. doi: 10.1016/j.ygeno.2005.03.001

Ton, Q. V., Leino, D., Mowery, S. A., Bredemeier, N. O., Lafontant, P. J., Lubert, A., et al. (2018). Collagen COL22A1 maintains vascular stability and mutations in COL22A1 are potentially associated with intracranial aneurysms. *Dis. Models Mech.* 11:dmm033654. doi: 10.1242/dmm.033654

Tonomura, S., Ihara, M., Kawano, T., Tanaka, T., Okuno, Y., Saito, S., et al. (2016). Intracerebral hemorrhage and deep microbleeds associated with cnm-positive *Streptococcus mutans*; a hospital cohort study. *Sci. Rep.* 6:20074. doi: 10.1038/srep20074

Ton-That, H., and Schneewind, O. (2004). Assembly of pili in Gram-positive bacteria. *Trends Microbiol.* 12, 228–234. doi: 10.1016/j.tim.2004.03.004

Vaca, D. J., Thibau, A., Schütz, M., Kraiczy, P., Happonen, L., Malmström, J., et al. (2020). Interaction with the host: the role of fibronectin and extracellular matrix proteins in the adhesion of Gram-negative bacteria. *Med. Microbiol. Immunol.* 209, 277–299. doi: 10.1007/s00430-019-00644-3

van Wieringen, T., Kalamajski, S., Lidén, Å, Bihan, D., Guss, B., Heinegård, D., et al. (2010). The streptococcal collagen-binding protein CNE specifically interferes with αVβ3-mediated cellular interactions with triple helical collagen. *J. Biol. Chem.* 285, 35803–35813. doi: 10.1074/jbc.M110.146001

Veit, G., Zwolanek, D., Eckes, B., Niland, S., Käpylä, J., Zweers, M. C., et al. (2011). Collagen XXIII, novel ligand for integrin alpha2beta1 in the epidermis. *J. Biol. Chem.* 286, 27804–27813. doi: 10.1074/jbc.m111.220046

Waldemarsson, J., Areschoug, T., Lindahl, G., and Johnsson, E. (2006). The streptococcal Blr and Slr proteins define a family of surface proteins with leucine-rich repeats: camouflaging by other surface structures. *J. Bacteriol.* 188, 378–388. doi: 10.1128/jb.188.2.378-388.2006

Wang, W., Olson, D., Liang, G., Franceschi, R. T., Li, C., Wang, B., et al. (2012). Collagen XXIV (Col24α1) promotes osteoblastic differentiation and mineralization through TGF-β/Smads signaling pathway. *Int. J. Biol. Sci.* 8, 1310–1322. doi: 10.7150/ijbs.5136

Willumsen, N., and Karsdal, M. A. (2016). "Type XX collagen," in *Biochemistry of Collagens, Laminins and Elastin*, Chap. 20, ed. M. A. Karsdal (Cambridge, MA: Academic Press), 127–129.

Xu, Y., Liang, X., Chen, Y., Koehler, T. M., and Hook, M. (2004a). Identification and biochemical characterization of two novel collagen binding MSCRAMMs of *Bacillus anthracis*. *J. Biol. Chem.* 279, 51760–51768. doi: 10.1074/jbc.M406417200

Xu, Y., Rivas, J. M., Brown, E. L., Liang, X., and Hook, M. (2004b). Virulence potential of the staphylococcal adhesin CNA in experimental arthritis is determined by its affinity for collagen. *J. Infect. Dis.* 189, 2323–2333. doi: 10.1086/420851

Yu, Z., An, B., Ramshaw, J. A., and Brodsky, B. (2014). Bacterial collagen-like proteins that form triple-helical structures. *J. Struct. Biol.* 186, 451–461. doi: 10.1016/j.jsb.2014.01.003

Zainul, Z., Heikkinen, A., Koivisto, H., Rautalahti, I., Kallio, M., Lin, S., et al. (2018). Collagen XIII is required for neuromuscular synapse regeneration and functional recovery after peripheral nerve injury. *J. Neurosci.* 38, 4243–4258.

Zani, I. A., Stephen, S. L., Mughal, N. A., Russell, D., Homer-Vanniasinkam, S., Wheatcroft, S. B., et al. (2015). Scavenger receptor structure and function in health and disease. *Cells* 4, 178–201. doi: 10.3390/cells4020178

Zhang, Y.-Z., Ran, L.-Y., Li, C.-Y., and Chen, X.-L. (2015). Diversity, structures, and collagen-degrading mechanisms of bacterial collagenolytic proteases. *Appl. Environ. Microbiol.* 81, 6098–6107. doi: 10.1128/aem.00883-15

Zong, Y., Xu, Y., Liang, X., Keene, D. R., Hook, A., Gurusiddappa, S., et al. (2005). A 'Collagen Hug' model for *Staphylococcus aureus* CNA binding to collagen. *Embo J.* 24, 4224–4236. doi: 10.1038/sj.emboj.7600888

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Microbiology

MINI REVIEW
published: 09 February 2021
doi: 10.3389/fmicb.2021.615924

Streptococcus pneumoniae Type 1 Pilus – A Multifunctional Tool for Optimized Host Interaction

*Stefan Ness and Markus Hilleringmann**

FG Protein Biochemistry & Cellular Microbiology, Department of Applied Sciences and Mechatronics, Munich University of Applied Sciences, Munich, Germany

Streptococcus pneumoniae represents a major Gram-positive human pathogen causing bacterial pneumonia, otitis media, meningitis, and other invasive diseases. Several pneumococcal isolates show increasing resistance rates against antibacterial agents. A variety of virulence factors promote pneumococcal pathogenicity with varying importance in different stages of host infection. Virulence related hair-like structures ("pili") are complex, surface located protein arrays supporting proper host interaction. In the last two decades different types of pneumococcal pili have been identified: pilus-1 (P1) and pilus-2 (P2) are formed by the catalytic activity of sortases that covalently assemble secreted polypeptide pilin subunits in a defined order and finally anchor the resulting pilus in the peptidoglycan. Within the long pilus fiber the presence of intramolecular isopeptide bonds confer high stability to the sequentially arranged individual pilins. This mini review will focus on *S. pneumoniae* TIGR4 P1 molecular architecture, the subunits it builds and provides insights into P1 sortase-mediated assembly. The complex P1 architecture (anchor-/backbone-/tip-subunits) allows the specific interaction with various target structures facilitating different steps of colonization, invasion and spreading within the host. Optimized pilin subunit confirmation supports P1 function under physiological conditions. Finally, aspects of P1- host interplay are summarized, including recent insights into P1 mechanobiology, which have important implications for P1 mediated pathogenesis.

Keywords: pilus, bacteria, *Streptococcus pneumoniae*, architecture, sortase, virulence, host-interaction, mechanobiology

OPEN ACCESS

Edited by:
Magnus Hook,
Texas A&M Health Science Center, United States

Reviewed by:
Thomas Proft,
The University of Auckland, New Zealand
Marco Rinaldo Oggioni,
University of Leicester, United Kingdom

***Correspondence:**
Markus Hilleringmann
markus.hilleringmann@hm.edu

Specialty section:
This article was submitted to Infectious Diseases, a section of the journal Frontiers in Microbiology

Received: *10 October 2020*
Accepted: *20 January 2021*
Published: *09 February 2021*

Citation:
Ness S and Hilleringmann M (2021) Streptococcus pneumoniae Type 1 Pilus – A Multifunctional Tool for Optimized Host Interaction. Front. Microbiol. 12:615924. doi: 10.3389/fmicb.2021.615924

INTRODUCTION: *Streptococcus pneumoniae* A MAJOR HUMAN PATHOGEN EXPRESSING DIFFERENT TYPES OF PILI

Streptococcus pneumoniae (the pneumococcus) is a human commensal bacterium that can cause lethal diseases like pneumonia, septicemia, and meningitis. As major human pathogen, it provokes high morbidity and mortality rates especially in children and the elderly. Licensed pneumococcal vaccines are not covering all relevant virulent strains and an increasing number of antibiotic resistant isolates makes treatment challenging (Subramanian et al., 2019). Novel, broad-spectrum vaccination strategies and new antibacterials are of utmost importance to combat *S. pneumoniae*.

The switch from a human commensal to an invasive pneumococcal pathogen and its disease causing capacity in various host niches is an area of intense study (Loughran et al., 2019). A multiplicity of differentially regulated cell surface located molecules mediate the complex interplay of *S. pneumoniae* and the human host. This includes several types of high molecular weight protein assemblies, so-called pili, which promote pneumococcal virulence. Although *S. pneumoniae* was first isolated by Pasteur in 1881, pneumococcal pili were discovered only at the beginning of the 21st century. Besides a recently described type IV competence pilus (Laurenceau et al., 2013; Muschiol et al., 2019), pneumococcal isolates express two variants of Gram-positive, sortase cross-linked multi-subunit pili ["pilus-1 (P1) and pilus-2 (P2)"] (Barocchi et al., 2006; Bagnoli et al., 2008). In this review, we focus on the assembly and particular architecture of pneumococcal P1 and derived from that summarize its role as multifunctional host-interaction tool. This indicates a virulence-mediating role of P1 at different phases of pneumococcal diseases and an optimized P1-structure for various host environments.

SUBUNITS AND SORTASE-MEDIATED ASSEMBLY OF PNEUMOCOCCAL PILUS-1

While typical Gram-negative pili are formed by non-covalent interactions between pilins, the covalent assembly of Gram-positive pili is catalyzed by specific sortases involving pilus subunit polymerization and cell wall anchoring of the resulting pilus fiber (Telford et al., 2006; Hendrickx et al., 2011; Hospenthal et al., 2017). *Streptococcus pneumoniae* TIGR4, belonging to the highly invasive pneumococcal serotype 4, is a patient isolate were P1 was initially identified (Barocchi et al., 2006) that represents an important reference strain studying P1 biology. Major components involved in pneumococcal TIGR4 P1 formation are clustered in a defined genetic region [pilus island 1 (PI-1)] that encodes 3 P1 specific class C sortases (SrtC-1, SrtC-2, SrtC-3), 3 P1 subunits (RrgA, RrgB, and RrgC) and a transcriptional regulator (RlrA) (**Figure 1-A1**). Analysis of strain collections indicates that only a subset of pneumococcal isolates expresses P1 (∼ 30%) (Dzaraly et al., 2020). Electron microscopic analysis allowed the initial visualization of pili on the surface of negative stained TIGR4 isolates. Negative staining procedure improves contrast for better visualization of few nm thin pilus filaments. The explicit identification of PI-1 encoded pilins within P1 required specific immuno-labeling strategies and resulted in heterotrimeric P1 working models with RrgB as major pilin and 2 minor pilins (RrgA and RrgC) (Barocchi et al., 2006; LeMieux et al., 2006, 2008; Hilleringmann et al., 2008; Fälker et al., 2008).

Similar to other Gram-positive pilus systems, the current model of P1 formation involves pilus subunit- and Srt-synthesis in the cytoplasm and respective surface localization followed by a biphasic process of pilus-polymerization and -anchoring to the bacterial cell wall (Khare and Narayana, 2017; **Figure 1-A2**): After their cytoplasmic expression, respective N-terminal signal sequences of P1-pilins and -sortases allow Sec-translocation into the exoplasmic zone. A C-terminal hydrophobic stretch functions as membrane anchor and promotes their embedding into the lipid bilayer. Since P1 discovery, a main research focus lies on the respective role of PI-1 sortases for P1 assembly, including molecular details of the catalyzed transpeptidation and Srt-regulation: during P1 polymerization, PI-1 individual class C sortases recognize particular variants of LPXTG motif pentapeptide-cell wall sorting signals (CWSSs) at the C-terminus of individual P1 pilins as described for other Gram-positive pilus systems. The sortase-catalytic activity hydrolyses between the CWSS threonine and glycine residues by an initial nucleophilic attack of the threonine's carbonyl carbon atom via the enzyme's active site cysteine residue (**Figure 1-A2**/I). The Thr C-terminus in the formed sortase-pilin thioacyl intermediate is linked to a conserved lysine ε-amino group within the pilin motif of the next P1 subunit of the growing pilus, RrgB K183 (within RrgB pilin motif WXXXVXVYPK) and finally specific lysine residue of RrgC (structural analysis suggest Lys K142 as respective nucleophilic residue of RrgC, although experimental proof is missing; (**Figure 1-A2**/II–III; LeMieux et al., 2006, 2008; Fälker et al., 2008; Manzano et al., 2008; Neiers et al., 2009a,b; El Mortaji et al., 2012b; Shaik et al., 2014). The data support a gradual selectivity of the three PI-1 SrtC isoforms to individual P1 pilin CWSSs during P1 polymerization (**Figure 1-A2**/II). Differences in CWSSs (RrgA: YPRTG; RrgB: IPQTG; RrgC: VPDTG) and resulting conformational changes might direct binding preferences of respective PI-1 SrtCs. Recent molecular dynamics simulations suggest a key role of the middle amino acid in the five-residue pilin CWSSs for selective and specific sorting signal targeting of the individual SrtCs (Naziga and Wereszczynski, 2017).

The second step in P1 biosynthesis is the covalent linkage of the polymerized P1 to the bacterial cell wall. Deletion of minor pilin RrgC leads to P1 polymers secreted into the supernatant and identifies RrgC as P1 cell wall anchor molecule (Hilleringmann et al., 2009; Shaik et al., 2014). The primary sortase for Lipid II-mediated RrgC anchoring to peptidoglycan is the pneumococcal housekeeping sortase SrtA (Shaik et al., 2014; Naziga and Wereszczynski, 2017). This is in agreement with SrtA pilus-anchoring activity in other Gram-positive pili systems.

Both, PI-1 class C-Srts and SrtA are central elements for P1 formation. Their coordinated activity seems essential for proper P1 assembly. Interestingly, all P1 class C Srts contain a N-terminal lid-region that covers the active catalytic site triad region constituted of His, Arg, and Cys that is missing in the respective housekeeping sortase and suggest a regulatory element of SrtCs activity. Whereas initial results indicate a flexible, "mobile" lid behavior of SrtC-1 in solution (Manzano et al., 2009), recent data propose a rather rigid SrtC-lid condition in the absence of substrate with a potential negative regulatory function (Jacobitz et al., 2016).

Some important questions regarding the mechanistics of Srt-mediated P1 fiber assembly are ambiguous: (i) Precise chronology of P1 formation including the factors determining the starting of the biosynthesis until final anchoring of P1

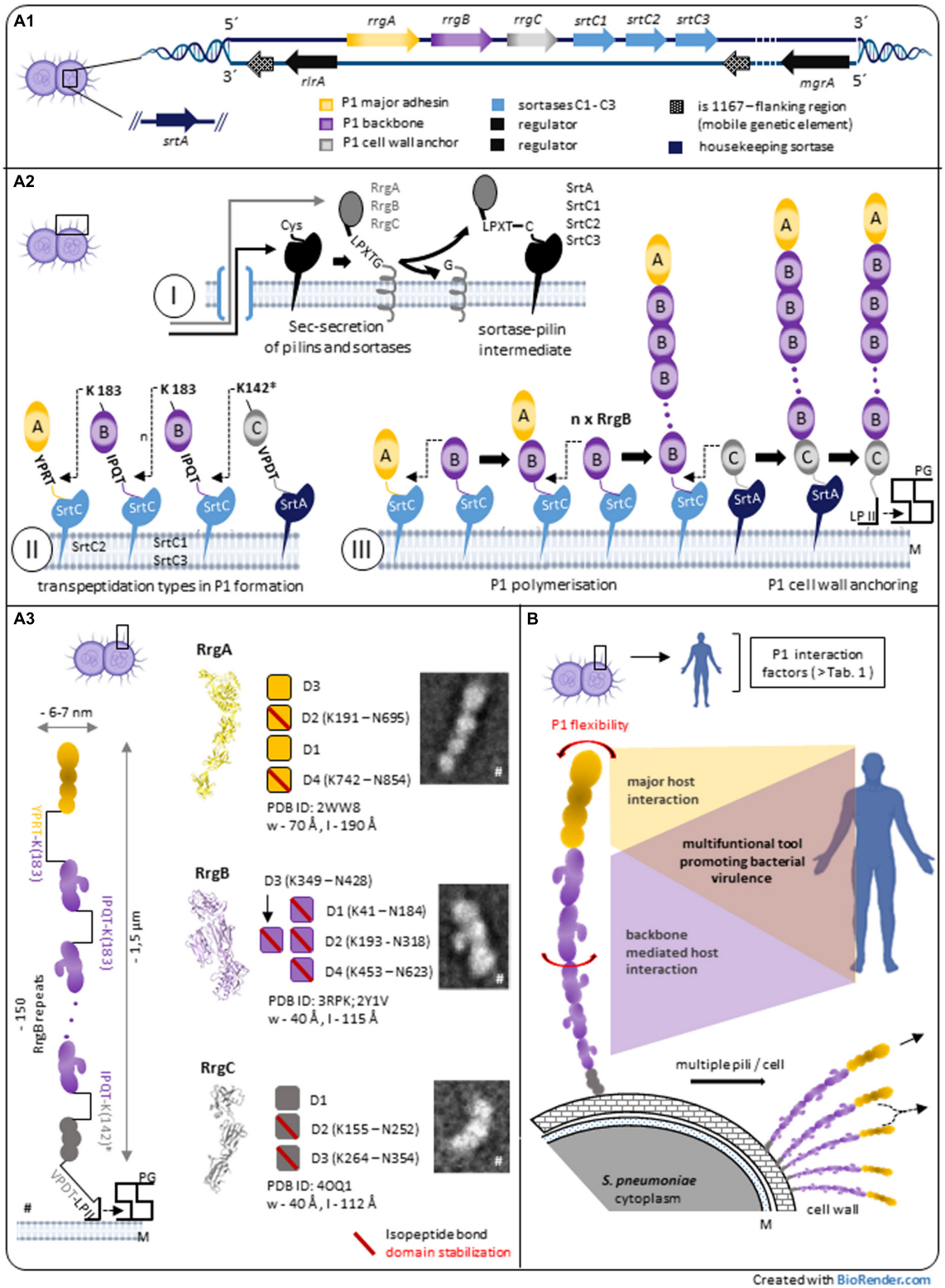


FIGURE 1 | Pilus-1 of *Streptococcus pneumoniae* – assembly, architecture and structure-derived properties. **(A1)** *Streptococcus pneumoniae* TIGR4 pilus-island 1 (PI-1) and related genes involved in P1 biosynthesis. **(A2)** Model of sortase-mediated assembly of pneumococcal pilus-1. *A2-I:* P1-pilins and relevant sortases are expressed in the cytoplasm, exported by the Sec secretion system and surface located via a C-terminal hydrophobic stretch in the bacterial plasma membrane. The initial step of sortase-catalyzed transpeptidation is a pilin CWSS specific hydrolysis between Thr and Gly of particular variants of the LPXTG motif via the enzyme's

(Continued)

FIGURE 1 | Continued
active site Cys resulting in a sortase-pilin thioacyl intermediate. *A2-II:* C-terminal pilin Thr in the activated sortase-pilin complex (A2-I) is linked to a conserved Lys in the next P1 subunit of the growing pilus. P1 heterotrimeric formation is a complex process involving the sequential incorporation of RrgA, RrgB, and RrgC via pilin specific PI-1 class C-sortases. Shown are details of individual P1 sortase-pilin intermediates and respective Lys residues involved in transpeptidation reactions during P1 formation. Simulation results combined with existing biochemical and genetic data of several groups suggest the following PI-1 class C-Srt specificities (Naziga and Wereszczynski, 2017): SrtC1 is the primary Srt for RgB polymerization; in the absence of RrgB, SrtC1 might add RrgA, and RrgC to the pilus. SrtC3 shows a pronounced RrgB crosslinking activity when compared to RrgA and RrgC although less distinctive than SrtC1-RrgB. The first substrate of SrtC2 is RrgA linking the P1 adhesin to the RrgB backbone. Apart from background activity of SrtC-1, housekeeping SrtA is defined as principal RrgC-related sortase. *A2-III:* P1 assembly is a biphasic process of P1 polymerization and final, lipid II-mediated, cell wall anchoring. **(A3)** P1 heterotrimeric architecture and pilin characteristics. P1 is a thin, long filament composed of a proximal RrgC anchor, n-serial RrgB backbone molecules and a distal RrgA adhesin. TEM images and high-resolution structures of P1-pilins (RrgA, RrgB, and RrgC) identify elongated molecules containing intramolecular, isopeptide bond stabilized domains. A distinctive feature of RrgB is a lateral subunit (D3) that is involved in host interaction (#parts adapted from Hilleringmann et al., 2009). **(B)** P1 expressing pneumococci exhibit multiple pilus filaments on the bacterial surface functioning as single fiber or potential P1 aggregations. In addition to P1 subunit stabilization **(A3)**, the entire P1 fiber shows a pronounced bendability. Longitudinal rotation of serial arranged RrgB subunits along the P1 axis suggest additional optimization of P1 architecture. P1 host interaction is complex, mainly mediated by terminal adhesin RrgA. Force induced RrgB host factor interaction indicate a functional role of P1 backbone in addition to its role as stabilized stalk exposing terminal P1 adhesin. *Potential nucleophilic RrgC Lys residue. CWSS, cell wall sorting signal; LP II, lipid II; M, bacterial plasma membrane; PG, peptidoglycan; P1, Pilus-1; Srt, Sortase.

to the cell wall is missing. (ii) Sequence details of P1 pilin incorporation: although major backbone RrgB polymerization was observed in a RrgA deletion strain background (Nelson et al., 2007), the "tip first rule" proposing a first transpeptidation reaction between the terminal pilus adhesin and a pilin backbone molecule [as described for the well studied heterotrimeric SpaA pili of *Corynebacterium diphtheriae* (Ramirez et al., 2020)] seems reasonable for other Gram-positive pili systems, but needs detailed experimental proof in the P1 system. (iii) Regulation of P1 length: based on a model of "pilus chain terminator role" of anchor pilin SpaB in *C. diphtheriae* SpaA pilus (Mandlik et al., 2008), it should be investigated whether the step of RrgC anchor incorporation into the SrtC-activated $RrgB_nRrgA$ fiber has a similar role in defining P1 final length. Importantly proper regulation of P1 RrgB mediated length determines the relative position of major adhesin RrgA and might be crucial for P1 functionality as recently demonstrated in other Gram-positive pilus systems (Chang et al., 2019).

TIGR4 bacterial surface is covered with multiple copies of P1 (Barocchi et al., 2006; Hilleringmann et al., 2009). The regulation of the spatio-temporal P1 distribution, including signal sequence directed cellular export and surface-positioning of pilins and Srts (Spirig et al., 2011), plus the precise interplay between P1 assembly and peptidoglycan metabolism is not understood. Topological studies describe a specific pattern of P1 subunits on the bacterial surface ("symmetric P1 foci") indicating a non-homogenous P1 distribution (Fälker et al., 2008).

In addition to specific control mechanisms focused on proper P1 biosynthesis, signals from general regulatory networks of the complex interplay between *S. pneumoniae* and the human host are supposed to influence P1 expression (Kreikemeyer et al., 2011; Gómez-Mejia et al., 2018). Transcriptional P1 regulation seems to be of major importance: besides a suggested negative regulation by RrgA (Basset et al., 2011), PI-1 was described to be positively controlled by a RlrA feedback loop (Basset et al., 2012). In addition to central *rlrA*, other PI-1 positive regulators were described (Hava et al., 2003; Kreikemeyer et al., 2011). Within the growing list of negative regulators (Basset et al., 2011; Kreikemeyer et al., 2011; Herbert et al., 2015), some of them might require additional validation (Basset et al., 2017). Detailed investigations with extending systems that reflect the *in vivo* situation as closely as possible will help to further clarify the overall context of pneumococcal P1 regulation during host colonization and invasion (Pancotto et al., 2013; Figueira et al., 2014). The described absence of P1 expression below 31°C indicates the importance of host environmental factors on P1 regulation (Basset et al., 2017). Analysis of PI-1 containing isolates identified sub-populations of P1-expressing and – non expressing bacteria. Angelis et al. (2011), Basset et al. (2011) which might be a compromise balancing P1-related advantages like improved human colonization and –potential disadvantages of P1 induced immunogenicity and biosynthesis costs (Binsker et al., 2020; Iovino et al., 2020).

DESIGN OF MULTIMERIC TIGR4 PILUS-1 ARCHITECTURE REFLECTS AN OPTIMIZED INTERACTION TOOL

High resolution analysis of surface digested individual pili allowed a detailed analysis of heterotrimeric P1 architecture and pilin stoichiometry, overcoming interferences and reducing complexity of respective cell bound-P1 study objects. TIGR4 P1 basic structure is a long, only ∼ 6–7 nm-wide filament composed of a multiple repetition of RrgB backbone molecules, with one RrgA adhesin and one RrgC minor pilin at the P1 distal and proximal end, respectively, associated in a head-to-tail covalently linked fashion. Scanning TEM analysis indicated that a 1.5-μm-long pilus comprises approximately 150 RrgB monomers in which a nose-like protrusion in each RrgB subunit defines the polarity of the fiber (Hilleringmann et al., 2009; **Figure 1-A3**). Adjacent P1 filaments on the cell wall surface can tangle or form bundles. Aggregations of different length pili with terminal RrgA, could explain the described P1 coiled-coiled phenotype with surface located RrgA molecules (Hilleringmann et al., 2008). RrgA aggregates were described as better TLR2 agonist *in vitro* when compared to RrgA monomer (Basset et al., 2013). Respectively, clustered P1 RrgA on the bacterial surface

might be better recognized by the innate immune receptor TLR2 *in vivo*.

High resolution crystal structures of all P1-pilins have been resolved (**Figure 1-A3**): major adhesion RrgA (PDB: 2WW8) represents a four domain, elongated molecule, carrying segments of eukaryotic origin important for host interaction, mainly mediated by RrgA D3 domain (Izoré et al., 2010; Moschioni et al., 2010). The backbone RrgB (PDB: 3RPK and 2Y1V) subunit displays a four domain fold (Paterson and Baker, 2011; El Mortaji et al., 2012a). Interestingly a domain (D3) is arranged laterally of the RrgB middle domain (D2) which constitutes a special structural feature among Gram-positive major pilins (Young et al., 2014) and presumes a specific P1 adaptation for optimized interaction (Becke et al., 2019). Reconstruction of P1 architecture by fitting of the P1 fiber with high resolution crystal structure of RrgB D2-D4 molecules and RrgB D1 computer model data showed a rotation along the longitudinal P1 axis between two neighboring RrgB subunits of about 17–22° (Spraggon et al., 2010). A resulting different spatial domain arrangement of sequential RrgB molecules might further improve host interaction via multiple linearly arranged RrgB molecules (**Figure 1-B**), like a RrgB D3 mediated collagen interaction (Becke et al., 2019). Anchor subunit RrgC is the smallest P1 pilin and folds into three independent domains (PDB: 4OQ1) (Shaik et al., 2014). P1 pilins show a characteristic domain-based architecture that incorporated into P1 filament resemble a "beads on a string" macroscopic phenotype. Like described for other Gram-positive pilus systems, the existence of eukaryotic IgG-like subdomain variants within all P1 pilins (Izoré et al., 2010; Paterson and Baker, 2011; Shaik et al., 2014) suggest an adaptation strategy for proper P1 mediated host factor interaction while minimizing adverse host immune system responses by mimicry of eukaryotic immune system elements (Krishnan et al., 2007; Shaik et al., 2014).

Structural analysis of Gram-positive pilins revealed domain stabilization via intramolecular isopeptide bond formations (Kang et al., 2007; Baker et al., 2015). P1 subunits show different numbers of stabilized individual domains [RrgA: 2 (Izoré et al., 2010); RrgB: 4 (El Mortaji et al., 2010, 2012a; Paterson and Baker, 2011), and RrgC: 2 (Shaik et al., 2014); **Figure 1-A3**] that might reflect individual pilin adaptation to generate an optimized P1 multimeric tool able to withstand mechanical perturbations induced during host interaction and supporting P1 assembly. Recent single-molecule force spectroscopy data propose structural concepts to protect covalent bonds of Gram-positive pilins from cleavage under mechanical challenge (Echelman et al., 2016).

In addition to pure pilus mechanical stability and similar to other Gram-positive pili, functional P1 mediated host interaction requires a certain extent of flexibility of the P1-filament to efficiently target host structures (**Figure 1-B**). Macroscopic analysis of P1 phenotype of cell-bound and isolated pili revealed a substantial degree of fiber bendability (Barocchi et al., 2006; Hilleringmann et al., 2009). The respective percental contribution of the flexible linker regions between individual P1 pilins within the fiber and subunit derived flexibility, i.e., interdomain malleability, needs further investigation and might be P1 subunit specific (Spraggon et al., 2010; El Mortaji et al., 2012a).

MULTIFUNCTIONAL PNEUMOCOCCAL TYPE-1 PILI PROMOTE VARIOUS INTERACTION SCENARIOS

The deletion of P1 attenuated the virulence of respective clinical isolates in mouse models of pneumococcal infection (Barocchi et al., 2006). Work by several groups indicate a significant involvement of major adhesin RrgA in various steps of pneumococcal colonization and virulence addressing different host niches, which is described in a recent, comprehensive review (Iovino et al., 2020) and summarized in **Table 1A**). RrgA mediates in a RrgB backbone independent way, adherence to host epithelial cells, potentially promoting initial host colonization (Nelson et al., 2007). Dose-dependent interactions of RrgA with ECM molecules (fibronectin, laminin, and collagen I), not seen for RrgB and RrgC, indicate specific binding to host cell associated molecules (Hilleringmann et al., 2008). In addition to static ELISA approaches, data from single molecule force spectroscopy specify a two-domain binding mechanism for RrgA with Fn, that suggest a P1-mediated bacterial adaptation to keep contact with host tissue surface in shear force environments (Becke et al., 2018). RrgA was found to interact with complement receptor 3 (CR3) that promotes CR3 –mediated uptake of *S. pneumoniae* expressing P1-RrgA by macrophages and results in pneumococcal stimulation of macrophage motility. Additionally, expression of P1 RrgA together with host expression of CR3 affects virulence and spreading of the pathogen from local sites to the bloodstream in mice (Orrskog et al., 2012). Data of a RrgA D3 dependent TLR2-activation and induction of inflammatory TNF-α responses (Basset et al., 2013) suggest further P1-mediated interaction with components of the innate immune system. Overall implications of the RrgA-mediated TLR2-activation of the host immune response, a potential limiting consequence of P1 expression on pneumococcal fitness and virulence, need further investigation. The described more potent activation of TLR2 by oligmeric forms of RrgA might be mediated by aggregated monomeric RrgA of individual P1 fibers on the pneumococcal surface (**Figure 1-B**). Lethal meningitis caused by *S. pneumoniae* requires the penetration of the blood-brain-barrier (BBB) by the bacteria (Iovino et al., 2016b): an interaction of RrgA with polymeric immunoglobulin receptor (pIgR) and platelet endothelial cell adhesion molecule 1 (PECAM-1), two BBB endothelial receptors, was found to promote the entry of bacteria into the brain and meningitis development (Iovino et al., 2017). Interestingly, a subpopulation of spherical single pneumococci, that do not express DivIVA, with surface-located P1-RrgA seem to be favored to cross the BBB. In addition to the expression of RrgA, the small size of these pneumococcal variants might explain easier penetration of the BBB (Iovino et al., 2016a). In addition, recent data imply that meningitis related neuronal death is mediated by an interaction of RrgA and pneumolysin with β-actin

TABLE 1 | *Streptococcus pneumoniae* pilus-1 mediated interactions.

Pilin (PI-1)	Variant	Category*	Major characteristics and target	References
***A:* RrgA**_ PBD: 2WW8	RrgA recombinant; bacteria associated RrgA (P1-variants)	I	RrgA binds to human respiratory epithelial A594 cells (independent of P1-RrgB polymerization); P1-Rrg A mediates colonization of the upper respiratory tract in mice.	Nelson et al., 2007
	recombinant RrgA; purified P1 (TIGR4 WT)	I	Dose-dependent binding of RrgA monomer and isolated pili to ECM components (fibronectin, collagen I, laminin)/ELISA.	Hilleringmann et al., 2008
	recombinant RrgA FL and individual RrgA domains	I°	D3/D4 – domain binding mechanism of RrgA with ECM fibronectin under force/AFM-based single molecule force spectroscopy.	Becke et al., 2018
	RrgA recombinant; bacteria associated RrgA (P1-variants)	I/II	RrgA mediated binding to CR3 and CR3 dependent uptake of RrgA containing pneumococci by macrophages; RrgA-expression promotes systemic pneumococcal spread and virulent in mice expressing CR3.	Orrskog et al., 2012
	RrgA recombinant; bacteria associated RrgA (P1-variants)	I	P1 functions as TLR2 agonist – with major contribution of RrgA D3 – in human epithelial cells. Rekombinant RrgA oligomers show increased TLR 2 activation. RrgA D3 is involved in pneumococcal mediated TLR2-activation, TNF-α induction and virulence in a mouse model of infection.	Basset et al., 2013
	RrgA recombinant; bacteria associated RrgA (P1-variants)	I/III	RrgA-containing P1 facilitate passage of *S. pneumoniae* through the blood-brain barrier (BBB) to cause lethal meningitis. Favored variants passing BBB of mice are spherical, single, P1-RrgA + pneumococci.	Iovino et al., 2016a
			Binding of RrgA to BBB endothelial receptors (PECAM-1 and pIgR) promotes bacterial entry and meningitis development.	Iovino et al., 2017
			Interaction of RrgA and pneumolysin with β-actin stimulate meningitis related neuronal death.	Tabusi et al., 2020
	Bacteria associated RrgA (P1-variants)	I + IV	Potential role of RrgA in biofilm formation / promotion of inter-bacterial interaction.	Muñoz-Elías et al., 2008
	RrgA recombinant; bacteria associated RrgA (P1-variants)	I + n/a	Role of RrgA as lectin targeting different host glycosylation pattern.	Day et al., 2017
***B:* Rrg B**_ PDB: 3RPK/2Y1V	RrgB recombinant; bacteria P1-variants	I + n/a	Role of RrgB as lectin targeting different host glycosylation pattern.	Day et al., 2017
	Recombinant RrgB FL and RrgB ΔD3 variants	I°	RrgB binds to ECM collagen I in a force-dependent manner and depends on the orientation of lateral RrgB D3-domain and the respective position of the collagen fibrils.	(Becke et al., 2019
***C:* RrgC**_ PDB: 4OQ1	RrgC recombinant; bacteria P1-variants	I + n/a	Role of RrgC as lectin targeting different host glycosylation pattern.	Day et al., 2017

Category: I, adhesion to ECM/surfaces/mammalian cells; II, evasion of innate and adaptive immune responses; III, invasion of non-phagocytic host cells by pathogen-directed endocytosis; IV, cell–cell aggregation during biofilm formation.*
°Interaction under mechanical force conditions; n/a, not available.
P1, pilus-1; WT, wild-type; ECM, extracellular matrix; FL, full length; CR3, complement receptor 3; TLR2, Toll-like receptor 2; TNFα, tumor necrosis factor α; PECAM-1, platelet endothelial cell adhesion molecule 1; pIgR, polymeric immunoglobulin receptor; AFM, atomic force microscopy.

of human neurons (Tabusi et al., 2020). The formation of biofilms promotes the persistence of pathogenic bacteria on patient's tissues and overall virulence (Muhammad et al., 2020). *In vitro* screening for biofilm-altered TIGR4 mutants identified RrgA as potential molecule favoring bacterial interaction (Muñoz-Elías et al., 2008).

Polymerized backbone subunits of Gram-positive pili are primarily considered as stabilized and flexible stalk exposing tip based adhesins for proper host interaction. Recent data suggest a specific role for backbone pilins in host interaction, as shown for pili of *Streptococcus pyogenes* (Tsai et al., 2017; Chen et al., 2020). Interestingly, under mechanical load, P1 RrgB was found to strongly interact with human collagen I (**Table 1B**) not measurable using static ELISA (Hilleringmann et al., 2008; Becke et al., 2019). The particular lateral domain D3 of RrgB (**Figure 1-A3**) is essential for the observed exceptionally high binding forces and a force-induced bond strengthening, a described property of bacterial adhesins (Herman-Bausier et al., 2018). This implies a discrete functional role of the P1 backbone in host factor interaction. Specific binding characteristics mediated by numerous linearly arranged RrgB molecules within P1 needs further investigation. Broader experimental approaches better mimicking *in vivo* conditions might reveal novel concepts of host interaction, also mediated by backbone subunits of Gram-positive pili.

A common class of host target structures of many bacterial adhesins are glycoconjugates. Work by Day et al. (2017) describe P1 subunits as new pneumococcal lectins binding several

glycosylation pattern: major P1 adhesin RrgA shows the broadest glycan binding repertoire (namely maltose, cellobiose, α/β linked galactose, blood group A and H antigens) when compared to RrgB and RrgC specificities. Interestingly, screening glycosylation targets of P1 subunits also identified non-human and non-mammalian patterns, assigning an even larger spectrum of potential P1 interactions (Day et al., 2017).

Apart from its potential role as lectin (Day et al., 2017), no further host interaction factor for P1 anchor RrgC is known (**Table 1C**).

The multiplicity of described P1-mediated host interaction scenarios makes P1 a pneumococcal virulence factor with strong impact on the pathogenesis of *S. pneumoniae*. P1 was shown to contribute to initial steps of colonization but also promotes invasion and spreading within the host (**Table 1**). Despite obvious advantages, P1 was identified in only a relatively small proportion of pneumococcal clinical isolates (−30%) belonging to few clonal complexes (Dzaraly et al., 2020) and PI-1 positive strains show a biphasic P1 expression pattern (Angelis et al., 2011). This implies potential limitations for pneumococcal fitness and virulence related to P1 expression: synthesis of high molecular weight architecture of P1 is complex (**Figure 1**) and might be tightly regulated by the bacterial metabolism. Iovino et al. (2020) propose two trade-off mechanisms explaining the relatively low % of P1 expressing strains: a high P1 induced host immunogenicity related to surface exposed P1 filaments that may also prevent recolonization and a low host transmissibility due to P1 adhesive properties. Additionally *S. pneumoniae* produces a variety of surface exposed, virulence related factors promoting various host interactions (Mitchell and Mitchell, 2010; Hilleringmann et al., 2015) that may substitute similar P1 mediated functions (e.g., PavA – fibronectin recognition). Importantly epidemiological data show that strain isolates that contain PI-1 are often associated with successful pneumococcal lineages and antimicrobial resistance pattern (Dzaraly et al., 2020), suggesting P1 as a particular tool among the diversity of pneumococcal virulence factors.

FUTURE DIRECTION AND CONCLUDING REMARKS

One and a half decades after the first description of *S. pneumoniae* P1 a considerable amount of data characterizes these very long, thin and highly stable surface appendages as evolutionary optimized subunit assemblies that promote pneumococcal virulence mediating multifunctional interactions in different host niches. Although the main components essential for P1 formation and the resulting architecture are well described, details of the spatio-temporal interplay during P1 assembly on the bacterial surface and their regulation ("P1-assembly machinery") needs further analysis, applying, e.g., novel high resolution microscopic approaches. In addition, a more complete picture of P1 functional aspects and the *in vivo* relevance demands a greater focus on complementary experimental approaches mimicking host environments [e.g., mechanical force- conditions (Dufrêne and Persat, 2020; Viela et al., 2020)] in addition to suitable *in vivo* models. These data will allow a more profound evaluation of P1 mediated bacterial fitness benefits with related costs of P1 expression and the risk of P1 induced adverse host immune responses. Promising antigenic properties of P1 subunits were demonstrated in several animal studies (Gianfaldoni et al., 2007; Moschioni et al., 2010; Gentile et al., 2011; Harfouche et al., 2012). Although P1 is found in only ∼ 30% of clinical pneumococcal isolates, its complex involvement with bacterial virulence, and data implying an association of P1 with antibiotic resistance and evolving non-vaccine serotypes (Dzaraly et al., 2020) make P1 subunits interesting protein-based vaccine candidates. Rational antigen (AG) design and innovative formulation strategies might enable a future non-serotype-dependent, efficient vaccination against *S. pneumoniae*, containing several protective pneumococcal protein-AGs, potentially in combination with known polysaccharide-conjugate AG designs (Converso et al., 2020). A direct interference with P1 function, like small molecule or antibody-mediated inhibition of PI-1 sortase activity, blocking of specific P1 epitopes involved in host interaction (Amerighi et al., 2016) or destabilizing P1 backbone subunits to reduce P1 functionality in host environment as shown for Spy0128, the major pilin from the Gram-positive human pathogen *S. pyogenes* (Rivas-Pardo et al., 2018) constitute further potential strategies to reduce P1-mediated virulence. In addition, understanding P1 on a molecular level also enabled the design of innovative bioconjugation tools (Bonnet et al., 2017) and demonstrates the many facets of this fascinating structure.

AUTHOR CONTRIBUTIONS

SN and MH wrote the manuscript and performed critical revision of the work. Both authors approved the submitted version.

FUNDING

This work was financially supported through the Open Access Publication fund of the Munich University of Applied Sciences (MUAS).

REFERENCES

Amerighi, F., Valeri, M., Donnarumma, D., Maccari, S., Moschioni, M., Taddei, A., et al. (2016). Identification of a monoclonal antibody against pneumococcal pilus 1 ancillary protein impairing bacterial adhesion to human epithelial cells. *J. Infect. Dis.* 213, 516–522. doi: 10.1093/infdis/jiv461

Angelis, G., de, Moschioni, M., Muzzi, A., Pezzicoli, A., Censini, S., et al. (2011). The *Streptococcus pneumoniae* pilus-1 displays a biphasic expression pattern. *PLoS One* 6:e21269. doi: 10.1371/journal.pone.0021269

Bagnoli, F., Moschioni, M., Donati, C., Dimitrovska, V., Ferlenghi, I., Facciotti, C., et al. (2008). A second pilus type in *Streptococcus pneumoniae* is prevalent in emerging serotypes and mediates adhesion to host cells. *J. Bacteriol.* 190, 5480–5492. doi: 10.1128/jb.00384-08

Baker, E. N., Squire, C. J., and Young, P. G. (2015). Self-generated covalent cross-links in the cell-surface adhesins of Gram-positive bacteria. *Biochem. Soc. Trans.* 43, 787–794. doi: 10.1042/bst20150066

Barocchi, M. A., Ries, J., Zogaj, X., Hemsley, C., Albiger, B., Kanth, A., et al. (2006). A pneumococcal pilus influences virulence and host inflammatory responses. *Proc. Natl. Acad. Sci. U.S.A.* 103, 2857–2862. doi: 10.1073/pnas.0511017103

Basset, A., Herd, M., Daly, R., Dove, S. L., and Malley, R. (2017). The Pneumococcal type 1 Pilus genes are thermoregulated and are repressed by a member of the snf2 protein family. *J. Bacteriol.* 199:e00078-17.

Basset, A., Turner, K. H., Boush, E., Sayeed, S., Dove, S. L., and Malley, R. (2011). Expression of the type 1 pneumococcal pilus is bistable and negatively regulated by the structural component RrgA. *Infect. Immun.* 79, 2974–2983. doi: 10.1128/iai.05117-11

Basset, A., Turner, K. H., Boush, E., Sayeed, S., Dove, S. L., and Malley, R. (2012). An epigenetic switch mediates bistable expression of the type 1 pilus genes in *Streptococcus pneumoniae*. *J. Bacteriol.* 194, 1088–1091. doi: 10.1128/jb.06078-11

Basset, A., Zhang, F., Benes, C., Sayeed, S., Herd, M., Thompson, C., et al. (2013). Toll-like receptor (TLR) 2 mediates inflammatory responses to oligomerized RrgA pneumococcal pilus type 1 protein. *J. Biol. Chem.* 288, 2665–2675. doi: 10.1074/jbc.m112.398875

Becke, T. D., Ness, S., Gürster, R., Schilling, A. F., Di Guilmi, A.-M., Sudhop, S., et al. (2018). Single molecule force spectroscopy reveals two-domain binding mode of pilus-1 tip protein RrgA of *Streptococcus pneumoniae* to Fibronectin. *ACS Nano* 12, 549–558. doi: 10.1021/acsnano.7b07247

Becke, T. D., Ness, S., Kaufmann, B. K., Hartmann, B., Schilling, A. F., Sudhop, S., et al. (2019). Pilus-1 backbone protein RrgB of *Streptococcus pneumoniae* binds collagen I in a force-dependent way. *ACS Nano* 13, 7155–7165. doi: 10.1021/acsnano.9b02587

Binsker, U., Lees, J. A., Hammond, A. J., and Weiser, J. N. (2020). Immune exclusion by naturally acquired secretory IgA against pneumococcal pilus-1. *J. Clin. Invest.* 130, 927–941. doi: 10.1172/jci132005

Bonnet, J., Cartannaz, J., Tourcier, G., Contreras-Martel, C., Kleman, J. P., Morlot, C., et al. (2017). Autocatalytic association of proteins by covalent bond formation: a Bio Molecular Welding toolbox derived from a bacterial adhesin. *Sci. Rep.* 7:43564.

Chang, C., Wu, C., Osipiuk, J., Siegel, S. D., Zhu, S., Liu, X., et al. (2019). Cell-to-cell interaction requires optimal positioning of a pilus tip adhesin modulated by gram-positive transpeptidase enzymes. *Proc. Natl. Acad. Sci. U.S.A.* 116, 18041–18049. doi: 10.1073/pnas.1907733116

Chen, Y.-H., Li, S.-H., Yang, Y.-C., Hsu, S.-H., Nizet, V., and Chang, Y.-C. (2020). T4 Pili promote colonization and immune evasion phenotypes of nonencapsulated M4 *Streptococcus pyogenes*. *mBio* 11:e01580-20.

Converso, T. R., Assoni, L., André, G. O., Darrieux, M., and Leite, L. C. C. (2020). The long search for a serotype independent pneumococcal vaccine. *Expert Rev. Vaccines* 19, 57–70. doi: 10.1080/14760584.2020.1711055

Day, C. J., Paton, A. W., Harvey, R. M., Hartley-Tassell, L. E., Seib, K. L., Tiralongo, J., et al. (2017). Lectin activity of the pneumococcal pilin proteins. *Sci. Rep.* 7:17784.

Dufrêne, Y. F., and Persat, A. (2020). Mechanomicrobiology: how bacteria sense and respond to forces. *Nat. Rev. Microbiol.* 18, 227–240. doi: 10.1038/s41579-019-0314-2

Dzaraly, N. D., Muthanna, A., Mohd Desa, M. N., Taib, N. M., Masri, S. N., Rahman, N. I. A., et al. (2020). Pilus islets and the clonal spread of piliated *Streptococcus pneumoniae*: a review. *Int. J. Med. Microbiol.* 310:151449. doi: 10.1016/j.ijmm.2020.151449

Echelman, D. J., Alegre-Cebollada, J., Badilla, C. L., Chang, C., Ton-That, H., and Fernández, J. M. (2016). CnaA domains in bacterial pili are efficient dissipaters of large mechanical shocks. *Proc. Natl. Acad. Sci. U.S.A.* 113, 2490–2495. doi: 10.1073/pnas.1522946113

El Mortaji, L., Contreras-Martel, C., Moschioni, M., Ferlenghi, I., Manzano, C., Vernet, T., et al. (2012a). The full-length *Streptococcus pneumoniae* major pilin RrgB crystallizes in a fibre-like structure, which presents the D1 isopeptide bond and provides details on the mechanism of pilus polymerization. *Biochem. J.* 441, 833–841. doi: 10.1042/bj20111397

El Mortaji, L., Fenel, D., Vernet, T., and Di Guilmi, A. M. (2012b). Association of RrgA and RrgC into the *Streptococcus pneumoniae* pilus by sortases C-2 and C-3. *Biochemistry* 51, 342–352. doi: 10.1021/bi201591n

El Mortaji, L., Terrasse, R., Dessen, A., Vernet, T., and Di Guilmi, A. M. (2010). Stability and assembly of pilus subunits of *Streptococcus pneumoniae*. *J. Biol. Chem.* 285, 12405–12415. doi: 10.1074/jbc.m109.082776

Fälker, S., Nelson, A. L., Morfeldt, E., Jonas, K., Hultenby, K., Ries, J., et al. (2008). Sortase-mediated assembly and surface topology of adhesive pneumococcal pili. *Mol. Microbiol.* 70, 595–607. doi: 10.1111/j.1365-2958.2008.06396.x

Figueira, M., Moschioni, M., Angelis, G., de, Barocchi, M., Sabharwal, V., et al. (2014). Variation of pneumococcal Pilus-1 expression results in vaccine escape during Experimental Otitis Media EOM. *PLoS One* 9:e83798. doi: 10.1371/journal.pone.0083798

Gentile, M. A., Melchiorre, S., Emolo, C., Moschioni, M., Gianfaldoni, C., Pancotto, L., et al. (2011). Structural and functional characterization of the *Streptococcus pneumoniae* RrgB pilus backbone D1 domain. *J. Biol. Chem.* 286, 14588–14597. doi: 10.1074/jbc.m110.202739

Gianfaldoni, C., Censini, S., Hilleringmann, M., Moschioni, M., Facciotti, C., Pansegrau, W., et al. (2007). *Streptococcus pneumoniae* pilus subunits protect mice against lethal challenge. *Infect. Immun.* 75, 1059–1062. doi: 10.1128/iai.01400-06

Gómez-Mejia, A., Gámez, G., Hirschmann, S., Kluger, V., Rath, H., Böhm, S., et al. (2018). Pneumococcal metabolic adaptation and colonization are regulated by the two-component regulatory system 08. *mSphere* 3:e00165-18.

Harfouche, C., Filippini, S., Gianfaldoni, C., Ruggiero, P., Moschioni, M., Maccari, S., et al. (2012). RrgB321, a fusion protein of the three variants of the pneumococcal pilus backbone RrgB, is protective in vivo and elicits opsonic antibodies. *Infect. Immun.* 80, 451–460. doi: 10.1128/iai.05780-11

Hava, D. L., LeMieux, J., and Camilli, A. (2003). From nose to lung: the regulation behind *Streptococcus pneumoniae* virulence factors. *Mol. Microbiol.* 50, 1103–1110. doi: 10.1046/j.1365-2958.2003.03764.x

Hendrickx, A. P. A., Budzik, J. M., Oh, S.-Y., and Schneewind, O. (2011). Architects at the bacterial surface - sortases and the assembly of pili with isopeptide bonds. *Nat. Rev. Microbiol.* 9, 166–176. doi: 10.1038/nrmicro2520

Herbert, J. A., Mitchell, A. M., and Mitchell, T. J. (2015). A serine-threonine kinase (StkP) regulates expression of the Pneumococcal pilus and modulates bacterial adherence to human epithelial and endothelial cells *in vitro*. *PLoS One* 10:e0127212. doi: 10.1371/journal.pone.0127212

Herman-Bausier, P., Labate, C., Towell, A. M., Derclaye, S., Geoghegan, J. A., and Dufrêne, Y. F. (2018). Staphylococcus aureus clumping factor A is a force-sensitive molecular switch that activates bacterial adhesion. *Proc. Natl. Acad. Sci. U.S.A.* 115, 5564–5569. doi: 10.1073/pnas.1718104115

Hilleringmann, M., Giusti, F., Baudner, B. C., Masignani, V., Covacci, A., Rappuoli, R., et al. (2008). Pneumococcal pili are composed of protofilaments exposing adhesive clusters of Rrg A. *PLoS Pathog.* 4:e1000026. doi: 10.1371/journal.ppat.1000026

Hilleringmann, M., Kohler, S., Gámez, G., and Hammerschmidt, S. (2015). "Pneumococcal pili and adhesins," in *Streptococcus pneumoniae Molecular Mechanisms of Host-Pathogen Interactions*. eds J. Brown, S. Hammerschmidt and C. Orihuela (London: Elsevier).

Hilleringmann, M., Ringler, P., Müller, S. A., Angelis, G., de, Rappuoli, R., et al. (2009). Molecular architecture of *Streptococcus pneumoniae* TIGR4 pili. *EMBO J.* 28, 3921–3930. doi: 10.1038/emboj.2009.360

Hospenthal, M. K., Costa, T. R. D., and Waksman, G. (2017). A comprehensive guide to pilus biogenesis in Gram-negative bacteria. *Nat. Rev. Microbiol.* 15, 365–379. doi: 10.1038/nrmicro.2017.40

Iovino, F., Engelen-Lee, J.-Y., Brouwer, M., van de Beek, D., van der Ende, A., Valls Seron, M., et al. (2017). pIgR and PECAM-1 bind to pneumococcal adhesins RrgA and PspC mediating bacterial brain invasion. *J. Exp. Med.* 214, 1619–1630. doi: 10.1084/jem.20161668

Iovino, F., Hammarlöf, D. L., Garriss, G., Brovall, S., Nannapaneni, P., and Henriques-Normark, B. (2016a). Pneumococcal meningitis is promoted by single cocci expressing pilus adhesin RrgA. *J. Clin. Invest.* 126, 2821–2826. doi: 10.1172/jci84705

Iovino, F., Nannapaneni, P., Henriques-Normark, B., and Normark, S. (2020). The impact of the ancillary pilus-1 protein RrgA of *Streptococcus pneumoniae* on colonization and disease. *Mol. Microbiol.* 113, 650–658. doi: 10.1111/mmi.14451

Iovino, F., Seinen, J., Henriques-Normark, B., and van Dijl, J. M. (2016b). How Does *Streptococcus pneumoniae* invade the brain? *Trends Microbiol.* 24, 307–315. doi: 10.1016/j.tim.2015.12.012

Izoré, T., Contreras-Martel, C., El Mortaji, L., Manzano, C., Terrasse, R., Vernet, T., et al. (2010). [Duplikat] Structural basis of host cell recognition by the pilus adhesin from *Streptococcus pneumoniae*. *Structure* 18, 106–115. doi: 10.1016/j.str.2009.10.019

Jacobitz, A. W., Naziga, E. B., Yi, S. W., McConnell, S. A., Peterson, R., Jung, M. E., et al. (2016). The "Lid" in the *Streptococcus pneumoniae* SrtC1 sortase adopts a rigid structure that regulates substrate access to the active site. *J. Phys. Chem. B* 120, 8302–8312. doi: 10.1021/acs.jpcb.6b01930

Kang, H. J., Coulibaly, F., Clow, F., Proft, T., and Baker, E. N. (2007). Stabilizing isopeptide bonds revealed in gram-positive bacterial pilus structure. *Science* 318, 1625–1628. doi: 10.1126/science.1145806

Khare, B., and Narayana, S. V. L. (2017). Pilus biogenesis of Gram-positive bacteria: roles of sortases and implications for assembly. *Protein Sci.* 26, 1458–1473. doi: 10.1002/pro.3191

Kreikemeyer, B., Gámez, G., Margarit, I., Giard, J.-C., Hammerschmidt, S., Hartke, A., et al. (2011). Genomic organization, structure, regulation and pathogenic role of pilus constituents in major pathogenic *Streptococci* and *Enterococci*. *Int. J. Med. Microbiol.* 301, 240–251. doi: 10.1016/j.ijmm.2010.09.003

Krishnan, V., Gaspar, A. H., Ye, N., Mandlik, A., Ton-That, H., and Narayana, S. V. L. (2007). An IgG-like domain in the minor pilin GBS52 of Streptococcus agalactiae mediates lung epithelial cell adhesion. *Structure* 15, 893–903. doi: 10.1016/j.str.2007.06.015

Laurenceau, R., Péhau-Arnaudet, G., Baconnais, S., Gault, J., Malosse, C., Dujeancourt, A., et al. (2013). A type IV pilus mediates DNA binding during natural transformation in *Streptococcus pneumoniae*. *PLoS Pathog.* 9:e1003473. doi: 10.1371/journal.ppat.1003473

LeMieux, J., Hava, D. L., Basset, A., and Camilli, A. (2006). RrgA and RrgB are components of a multisubunit pilus encoded by the *Streptococcus pneumoniae* rlrA pathogenicity islet. *Infect. Immun.* 74, 2453–2456. doi: 10.1128/iai.74.4.2453-2456.2006

LeMieux, J., Woody, S., and Camilli, A. (2008). Roles of the sortases of *Streptococcus pneumoniae* in assembly of the RlrA pilus. *J. Bacteriol.* 190, 6002–6013. doi: 10.1128/jb.00379-08

Loughran, A. J., Orihuela, C. J., and Tuomanen, E. I. (2019). *Streptococcus pneumoniae*: invasion and inflammation. *Microbiol. Spectr.* 7:10.1128/microbiolspec.GPP3-0004-2018.

Mandlik, A., Das, A., and Ton-That, H. (2008). The molecular switch that activates the cell wall anchoring step of pilus assembly in gram-positive bacteria. *Proc. Natl. Acad. Sci. U.S.A.* 105, 14147–14152. doi: 10.1073/pnas.0806350105

Manzano, C., Contreras-Martel, C., El Mortaji, L., Izoré, T., Fenel, D., Vernet, T., et al. (2008). Sortase-mediated pilus fiber biogenesis in *Streptococcus pneumoniae*. *Structure* 16, 1838–1848. doi: 10.1016/j.str.2008.10.007

Manzano, C., Izoré, T., Job, V., Di Guilmi, A. M., and Dessen, A. (2009). Sortase activity is controlled by a flexible lid in the pilus biogenesis mechanism of gram-positive pathogens. *Biochemistry* 48, 10549–10557. doi: 10.1021/bi901261y

Mitchell, A. M., and Mitchell, T. J. (2010). *Streptococcus pneumoniae*: virulence factors and variation. *Clin. Microbiol. Infect.* 16, 411–418. doi: 10.1111/j.1469-0691.2010.03183.x

Moschioni, M., Emolo, C., Biagini, M., Maccari, S., Pansegrau, W., Donati, C., et al. (2010). The two variants of the *Streptococcus pneumoniae* pilus 1 RrgA adhesin retain the same function and elicit cross-protection in vivo. *Infect. Immun.* 78, 5033–5042. doi: 10.1128/iai.00601-10

Muhammad, M. H., Idris, A. L., Fan, X., Guo, Y., Yu, Y., Jin, X., et al. (2020). Beyond risk: bacterial biofilms and their regulating approaches. *Front. Microbiol.* 11:928. doi: 10.3389/fmicb.2020.00928

Muñoz-Elías, E. J., Marcano, J., and Camilli, A. (2008). Isolation of *Streptococcus pneumoniae* biofilm mutants and their characterization during nasopharyngeal colonization. *Infect. Immun.* 76, 5049–5061. doi: 10.1128/iai.00425-08

Muschiol, S., Aschtgen, M.-S., Nannapaneni, P., and Henriques-Normark, B. (2019). Gram-positive type IV pili and competence. *Microbiol. Spectr.* 7, 129–135.

Naziga, E. B., and Wereszczynski, J. (2017). Molecular mechanisms of the binding and specificity of *Streptococcus pneumoniae* Sortase C enzymes for pilin subunits. *Sci. Rep.* 7:13119.

Neiers, F., Madhurantakam, C., Fälker, S., Manzano, C., Dessen, A., Normark, S., et al. (2009a). Two crystal structures of pneumococcal pilus sortase C provide novel insights into catalysis and substrate specificity. *J. Mol. Biol.* 393, 704–716. doi: 10.1016/j.jmb.2009.08.058

Neiers, F., Madhurantakam, C., Fälker, S., Normark, S., Henriques-Normark, B., and Achour, A. (2009b). Cloning, expression, purification, crystallization and preliminary X-ray analysis of the pilus-associated sortase C from *Streptococcus pneumoniae*. *Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun.* 65, 55–58. doi: 10.1107/s1744309108040025

Nelson, A. L., Ries, J., Bagnoli, F., Dahlberg, S., Fälker, S., Rounioja, S., et al. (2007). RrgA is a pilus-associated adhesin in *Streptococcus pneumoniae*. *Mol. Microbiol.* 66, 329–340. doi: 10.1111/j.1365-2958.2007.05908.x

Orrskog, S., Rounioja, S., Spadafina, T., Gallotta, M., Norman, M., Hentrich, K., et al. (2012). Pilus adhesin RrgA interacts with complement receptor 3, thereby affecting macrophage function and systemic pneumococcal disease. *mBio* 4:e00535-12.

Pancotto, L., Angelis, G., de, Bizzarri, E., Barocchi, M. A., Del Giudice, G., et al. (2013). Expression of the *Streptococcus pneumoniae* pilus-1 undergoes on and off switching during colonization in mice. *Sci. Rep.* 3:2040.

Paterson, N. G., and Baker, E. N. (2011). Structure of the full-length major pilin from *Streptococcus pneumoniae*: implications for isopeptide bond formation in gram-positive bacterial pili. *PLoS One* 6:e22095. doi: 10.1371/journal.pone.0022095

Ramirez, N. A., Das, A., and Ton-That, H. (2020). New paradigms of pilus assembly mechanisms in gram-positive Actinobacteria. *Trends Microbiol.* 28, 999–1009. doi: 10.1016/j.tim.2020.05.008

Rivas-Pardo, J. A., Badilla, C. L., Tapia-Rojo, R., Alonso-Caballero, Á., and Fernández, J. M. (2018). Molecular strategy for blocking isopeptide bond formation in nascent pilin proteins. *Proc. Natl. Acad. Sci. U.S.A.* 115, 9222–9227. doi: 10.1073/pnas.1807689115

Shaik, M. M., Maccagni, A., Tourcier, G., Di Guilmi, A. M., and Dessen, A. (2014). Structural basis of pilus anchoring by the ancillary pilin RrgC of *Streptococcus pneumoniae*. *J. Biol. Chem.* 289, 16988–16997. doi: 10.1074/jbc.m114.555854

Spirig, T., Weiner, E. M., and Clubb, R. T. (2011). Sortase enzymes in Gram-positive bacteria. *Mol. Microbiol.* 82, 1044–1059. doi: 10.1111/j.1365-2958.2011.07887.x

Spraggon, G., Koesema, E., Scarselli, M., Malito, E., Biagini, M., Norais, N., et al. (2010). Supramolecular organization of the repetitive backbone unit of the *Streptococcus pneumoniae* pilus. *PLoS One* 5:e10919. doi: 10.1371/journal.pone.0010919

Subramanian, K., Henriques-Normark, B., and Normark, S. (2019). Emerging concepts in the pathogenesis of the *Streptococcus pneumoniae*: from nasopharyngeal colonizer to intracellular pathogen. *Cell. Microbiol.* 21:e13077.

Tabusi, M., Thorsdottir, S., Lysandrou, M., Narciso, A. R., Minoia, M., Henriques-Normark, B., et al. (2020). Neuronal death in pneumococcal meningitis is triggered by pneumolysin and pilus-1 interactions with β-actin. *bioRxiv* [Preprint]. doi: 10.1101/2020.08.20.258681

Telford, J. L., Barocchi, M. A., Margarit, I., Rappuoli, R., and Grandi, G. (2006). Pili in gram-positive pathogens. *Nat. Rev. Microbiol.* 4, 509–519.

Tsai, J.-Y. C., Loh, J. M. S., Clow, F., Lorenz, N., and Proft, T. (2017). The Group A Streptococcus serotype M2 pilus plays a role in host cell adhesion and immune evasion. *Mol. Microbiol.* 103, 282–298. doi: 10.1111/mmi.13556

Viela, F., Mathelié-Guinlet, M., Viljoen, A., and Dufrêne, Y. F. (2020). What makes bacterial pathogens so sticky? *Mol. Microbiol.* 113, 683–690. doi: 10.1111/mmi.14448

Young, P. G., Moreland, N. J., Loh, J. M., Bell, A., Atatoa Carr, P., Proft, T., et al. (2014). Structural conservation, variability, and immunogenicity of the T6 backbone pilin of serotype M6 *Streptococcus pyogenes*. *Infect. Immun.* 82, 2949–2957.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Microbiology

REVIEW
published: 09 February 2021
doi: 10.3389/fmicb.2021.616508

Genetics, Structure, and Function of Group A Streptococcal Pili

Masanobu Nakata[1]* *and Bernd Kreikemeyer*[2]

[1] *Department of Oral Microbiology, Graduate School of Medical and Dental Sciences, Kagoshima University, Kagoshima, Japan,* [2] *Institute of Medical Microbiology, Virology and Hygiene, University of Rostock, Rostock, Germany*

Streptococcus pyogenes (Group A *Streptococcus*; GAS) is an exclusively human pathogen. This bacterial species is responsible for a large variety of infections, ranging from purulent but mostly self-limiting oropharynx/skin diseases to streptococcal sequelae, including glomerulonephritis and rheumatic fever, as well as life-threatening streptococcal toxic-shock syndrome. GAS displays a wide array of surface proteins, with antigenicity of the M protein and pili utilized for M- and T-serotyping, respectively. Since the discovery of GAS pili in 2005, their genetic features, including regulation of expression, and structural features, including assembly mechanisms and protein conformation, as well as their functional role in GAS pathogenesis have been intensively examined. Moreover, their potential as vaccine antigens has been studied in detail. Pilus biogenesis-related genes are located in a discrete section of the GAS genome encoding fibronectin and collagen binding proteins and trypsin-resistant antigens (FCT region). Based on the heterogeneity of genetic composition and DNA sequences, this region is currently classified into nine distinguishable forms. Pili and fibronectin-binding proteins encoded in the FCT region are known to be correlated with infection sites, such as the skin and throat, possibly contributing to tissue tropism. As also found for pili of other Gram-positive bacterial pathogens, GAS pilin proteins polymerize via isopeptide bonds, while intramolecular isopeptide bonds present in the pilin provide increased resistance to degradation by proteases. As supported by findings showing that the main subunit is primarily responsible for T-serotyping antigenicity, pilus functions and gene expression modes are divergent. GAS pili serve as adhesins for tonsillar tissues and keratinocyte cell lines. Of note, a minor subunit is considered to have a harpoon function by which covalent thioester bonds with host ligands are formed. Additionally, GAS pili participate in biofilm formation and evasion of the immune system in a serotype/strain-specific manner. These multiple functions highlight crucial roles of pili during the onset of GAS infection. This review summarizes the current state of the art regarding GAS pili, including a new mode of host-GAS interaction mediated by pili, along with insights into pilus expression in terms of tissue tropism.

Keywords: *Streptococcus pyogenes*, pili, thermoregulation, T serotyping, FCT region

OPEN ACCESS

Edited by:
Timothy J. Foster,
Trinity College Dublin, Ireland

Reviewed by:
Debra E. Bessen,
New York Medical College, United States
Awdhesh Kalia,
University of Texas MD Anderson Cancer Center, United States
Thomas Proft,
The University of Auckland, New Zealand

***Correspondence:**
Masanobu Nakata
nakata@dent.kagoshima-u.ac.jp

Specialty section:
This article was submitted to Infectious Diseases, a section of the journal Frontiers in Microbiology

Received: *12 October 2020*
Accepted: *18 January 2021*
Published: *09 February 2021*

Citation:
Nakata M and Kreikemeyer B (2021) Genetics, Structure, and Function of Group A Streptococcal Pili. Front. Microbiol. 12:616508. doi: 10.3389/fmicb.2021.616508

INTRODUCTION

Several different types of pathogenic bacteria colonize distinct niches by adhering to host tissues via long filamentous appendages termed pili or fimbriae, which project from the cell surface. Pili are also involved in conjugation, twitching motility, and virulence. Gram-positive bacterial pili were undetected until recently because of their thin structure. Following their discovery in bacteria belonging to the *Actinomyces* and *Corynebacterium* genera (Yanagawa and Honda, 1976; Cisar and Vatter, 1979), and the unraveling of their assembly mechanisms (Ton-That and Schneewind, 2003), pili of pathogenic streptococci, including *Streptococcus agalactiae*, *Streptococcus pyogenes*, and *Streptococcus pneumoniae* have been reported since 2005 (Lauer et al., 2005; Mora et al., 2005; Barocchi et al., 2006). Commensal oral streptococci, such as *Streptococcus sanguinis* and *Streptococcus oralis*, were also shown to produce pili (Okahashi et al., 2010; Zähner et al., 2011). Pili are considered to be physiologically distinctive to typical cell wall-anchored surface proteins regarding biological functions during the course of infection, since covalent linkage of subunits allows pilus proteins to locate not only on the bacterial cell surface but also >1 μm away from the surface, thus providing first contact with host molecules.

Streptococcus pyogenes (Group A *Streptococcus*; GAS) is a human pathogen responsible for a wide variety of human diseases (Walker et al., 2014). The major manifestations of GAS infections are local suppurative inflammation in the upper respiratory tract and skin, i.e., pharyngitis and impetigo. The annual number of pharyngitis cases worldwide has been estimated to be 616 million (Carapetis et al., 2005), while it has been speculated that there are 162 million children affected by impetigo at any one time (Bowen et al., 2015). GAS also causes sequelae, including rheumatic heart disease and acute glomerulonephritis, as well as streptococcal toxic-shock syndrome. The major typing scheme involves M and T serotyping. The former is based on the antigenicity of the M protein encoded by the *emm* gene. The 90 bp DNA sequence encoding the N-terminal variable region of the mature M protein is utilized to classify GAS into more than 240 types, known as *emm* typing (Beall et al., 1996; Sanderson-Smith et al., 2014). T typing is an alternative scheme based on the antigenicity of trypsin-resistant antigens (T antigens) (Griffith, 1934; Lancefield, 1940; Lancefield and Dole, 1946). It is conducted using trypsin-treated GAS cells and hyperimmune rabbit typing serum. The typing serum is raised against trypsin-treated GAS surface proteins, followed by adsorption with undigested GAS cells of different T types. Commercial antisera consists of five types of polyvalent sera and 21 of monovalent sera (Takizawa et al., 1970). A drawback of T serotyping is a lack of specificity. Several different typing sera react with recombinant pilus tip minor subunits (Lizano et al., 2007; Falugi et al., 2008; Nakata et al., 2009). Also, GAS isolates often react with several typing sera, such as T3/13/B3264 (Falugi et al., 2008). Thus, there is a lack of resolution compared to M typing. Moreover, unlike *emm* typing, there is no widespread genotyping method for pilus genes or the FCT region. Mora et al. (2005) reported that the major pilin of the GAS pilus is responsible for the antigenicity of T typing. The mechanism of protease resistance of a major pilin T antigen was uncovered by solving the structure by X-ray crystallography (Kang et al., 2007).

Over the last decade GAS pili have been found to be responsible for several functions, including host cell adherence, biofilm formation, immune evasion, and virulence. In this review, advances in understanding of pilus functions, the mode of pilus expression, perspectives for pilus-based vaccines, and the application of the structure and mechanisms of assembly of pili to biotechnology are summarized.

GENETIC ORGANIZATION OF FCT GENOMIC REGION CONTAINING PILUS-RELATED GENES

The genes that specify pili are located in the FCT genomic region (Bessen and Kalia, 2002). The acronym FCT stands for fibronectin-binding proteins, collagen-binding proteins, and trypsin-resistant antigens. Prior to discovery of pili, the gene encoding the T antigen type 6 (*tee*6) was reported (Schneewind et al., 1990). A BlastN search with *tee*6 and the flanking sequences as the query revealed a short stretch of sequence identity (91% over 70 bp) between the downstream region of *tee*6 and intergenic regions of *spy*0133 and *spy*0135 in the M1 strain SF370. Subsequent comparative analyses of deposited DNA sequences from additional serotypes revealed an approximately 11–16 kb recombinatorial region (Bessen and Kalia, 2002). This region contains genes encoding the fibronectin-binding F1/SfbI (Sela et al., 1993; Talay et al., 1994) and F2/PFBP/FbaB (Jaffe et al., 1996; Rocha and Fischetti, 1999; Terao et al., 2002) proteins, as well as the collagen-binding Cpa protein (Kreikemeyer et al., 2005). The FCT region is located between the conserved genes *hsp*33 and *spy*0136 (genome of SF370), and positioned nearly equidistant from the replication origin (clockwise from *ori*) as the M protein-coding *emm* locus. The heterogeneity of gene content among different GAS strains has allowed classification into nine subtypes designated FCT forms 1–9 (**Figure 1**, hereafter referred to as FCT-1 to FCT-9) (Kratovac et al., 2007; Falugi et al., 2008). To the best of our knowledge, the genome sequences of all except FCT-7 and FCT-8 are available. Generally, the same *emm* type strains share the same FCT form, though some exceptions have been reported (Falugi et al., 2008; Köller et al., 2010; Steemson et al., 2014). This region also encodes three kinds of transcriptional regulators, namely Nra, RofA, and MsmR (Fogg et al., 1994; Podbielski et al., 1999; Nakata et al., 2005). RofA and Nra belong to the RofA-like protein (RALP) regulator family that consists of four members with a mean amino acid sequence identity of 29% (Granok et al., 2000). Among these three regulators, Nra and RofA show an approximately 62% protein identity. The *nra* gene occurs in FCT-3, while other FCT forms contain the *rofA* gene. MsmR, an AraC-type regulator, is specific to FCT-3 and -4, and is always encoded by a gene located adjacent to *prtF*2 family genes (Nakata et al., 2005).

Pilus-related genes constitute an operon and encode one major and one or two minor subunits, at least one pilin-specific SrtB or SrtC type sortase, and the FCT-form specific

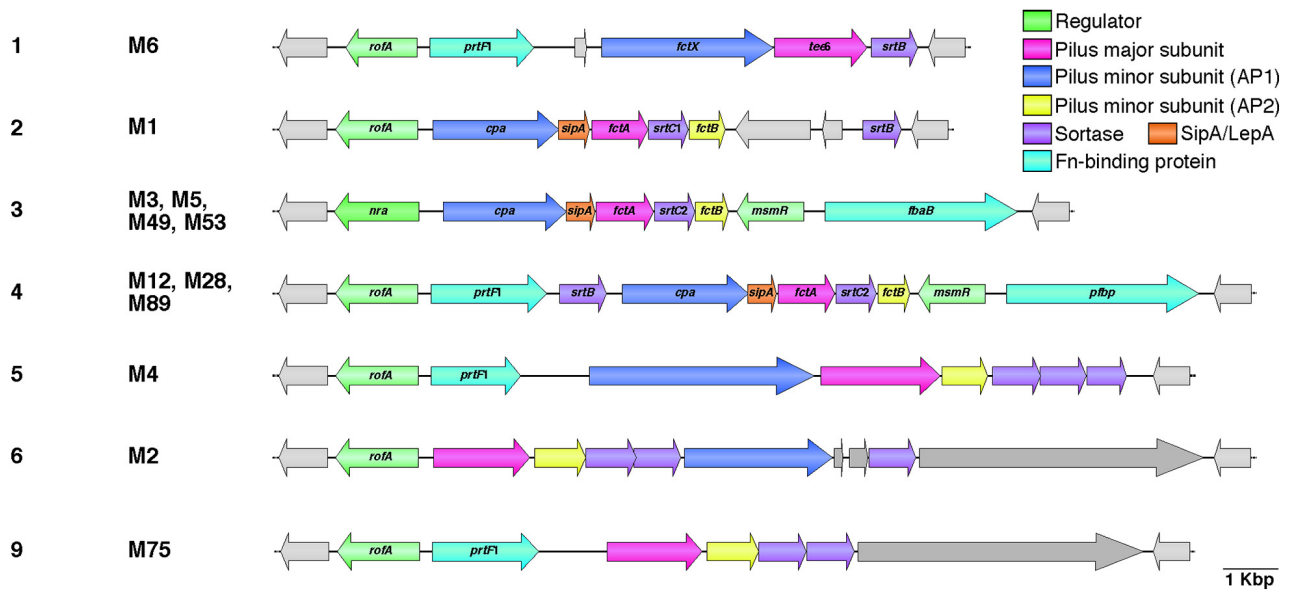


FIGURE 1 | Heterogeneic organization of FCT region. Gene content heterogeneity for the seven FCT forms is shown on the basis of genome sequences and previously reported data (Kratovac et al., 2007; Falugi et al., 2008). Representative M types for each FCT form are also shown. Pilus major and minor subunit (ancillary proteins 1 and 2; AP1 and AP2) genes are shown in pink and blue, respectively. SipA/LepA and sortase genes are colored orange and purple, respectively. Fibronectin-binding protein genes, including *prtF2* family genes (*pfbp* and *fbaB*) and *prtF1*, are shown in green, while transcriptional regulator genes, including *rofA*, *nra*, and *msmR*, are shown in light green. Other genes are light gray in color. Deposited DNA sequences of strains (M type, accession number) used for each FCT form are as follows: form 1, MGAS10394 (M6, NC_006086); form 2, SF370 (M1, NC_002737); form 3, SSI-1 (M3, NC_004606); form 4, A735 (M12, AF447492); form 5, MGAS10750 (M4, NC_008024); form 6, MGAS10270 (M2, NC_008022); and form 9, STAB14018 (M75, CCP014542.1). Genome sequences for FCT forms 7 and 8 are not available.

chaperone SipA/LepA (Barnett and Scott, 2002; Barnett et al., 2004; Zähner and Scott, 2008). Based on primary amino acid sequences, five sortase classes have been defined as SrtA to SrtF (Dramsi et al., 2005; Spirig et al., 2011). Class B sortases are predominant in Firmicutes, and their functions include pilus assembly and cell wall anchoring of proteins involved in iron acquisition (Mazmanian et al., 1999; Mora et al., 2005). Class C sortases are predominant in Firmicutes and Actinobacteria, and specifically function in pilus assembly. Confusingly, the GAS SrtB and SrtC proteins belong to class C and B, respectively, and both function in pilus assembly (Kang et al., 2011; Nakata et al., 2011). The major and minor subunits are often denoted as the backbone protein (bp) and ancillary proteins (ap), respectively. The number of minor ap subunits varies among types, with the tip minor subunit and base subunit usually termed ap1 and ap2, respectively. Falugi et al. analyzed seven different FCT forms and showed that the major subunit bp can be grouped into 15 variants. This was later expanded to 18 variants. There are 14 and 5 variants for the minor subunits ap1 and ap2, respectively (Falugi et al., 2008; Steemson et al., 2014). They also demonstrated that the major subunit bp is mainly responsible for T serotyping specificity.

Among the FCT forms, FCT-3 and FCT-4 share the greatest similarity. Inter-strain recombination of pilus genes between FCT-3 and FCT-4 has been speculated based on phylogenetic analysis and findings showing that an M5 strain possessed a *cpa* gene from FCT-4 (Falugi et al., 2008). Interestingly, the FCT-6 pilus minor subunit genes of several M2 strains show considerable homology to Group B *Streptococcus* pilus island I (PI-1) minor subunit genes, while the FCT region (FCT-1) of the M6 strain D471 has homology with the *rlrA* pathogenicity islet of *S. pneumoniae* (Bessen and Kalia, 2002; Hava and Camilli, 2002; Barocchi et al., 2006; Falugi et al., 2008). Horizontal gene transfer and recombination seems to have occurred between related species.

While *emm* typing is based on the sequence of 5′-end of the *emm* gene, the 3'-ends of *emm* and *emm*-like genes encoding M-like proteins, such as Mrp and Enn (Frost et al., 2018; Frost et al., 2020), are used for *emm* pattern groupings (Bessen et al., 1996). A strong correspondence between three groupings (patterns A-C, D, and E) and infection site preference, i.e., throat or skin, has been shown. Based on epidemiological data, *emm* pattern A-C and pattern D strains are designated "throat specialists" and "skin specialists," respectively, while pattern E strains are designated "generalists." The correlation between *emm* and FCT forms has been emphasized by data showing that 83% of FCT-3 strains harbor *emm* pattern D, whereas 84% of FCT-4 strains harbor *emm* pattern E (Kratovac et al., 2007). This strong linkage between FCT form and *emm* pattern raises the possibility that factors encoded in the FCT region, including pili, have roles in tissue tropism.

ASSEMBLY OF GAS PILI

In Gram-positive bacteria pilus subunits are linked to each other by isopeptide bonds mediated by pilus-specific sortases encoded in pilus gene clusters (Hendrickx et al., 2011). Among pilus types of pathogenic streptococci, the number of pilus-specific sortases varies (**Figure 2**). In GAS, FCT-5 and FCT-6 strains contain multiple class C sortases, as observed for pili of *S. agalactiae* and *S. pneumoniae*, while there is only one class B or class C sortase in FCT-1 to FCT-4. Pilin subunits possess a secretory signal sequence at their N-termini and a C-terminal cell wall sorting signal (CWSS) containing an LPXTG or LPXTG-like motif. This is followed by a stretch of hydrophobic residues and a positive-charged anchor that retains subunit proteins in the membrane during secretion via the Sec apparatus. The pilus-specific sortase cleaves an LPXTG or LPXTG-like motif between the threonine and glycine residues, and subsequently forms an acyl-enzyme intermediate by linking the active cysteine residue to the carboxyl group of the threonine. This intermediate is relieved by nucleophilic attack by the lysine residue side chain in the adjacent pilus subunit, forming isopeptide bonds between adjacent subunits. A series of consecutive reactions elongate pili until the occurrence of a stop signal, namely the incorporation of a minor subunit containing the canonical LPXTG motif (Smith et al., 2010), allowing the assembled pili to be connected to a free amino group of the peptidoglycan layer by

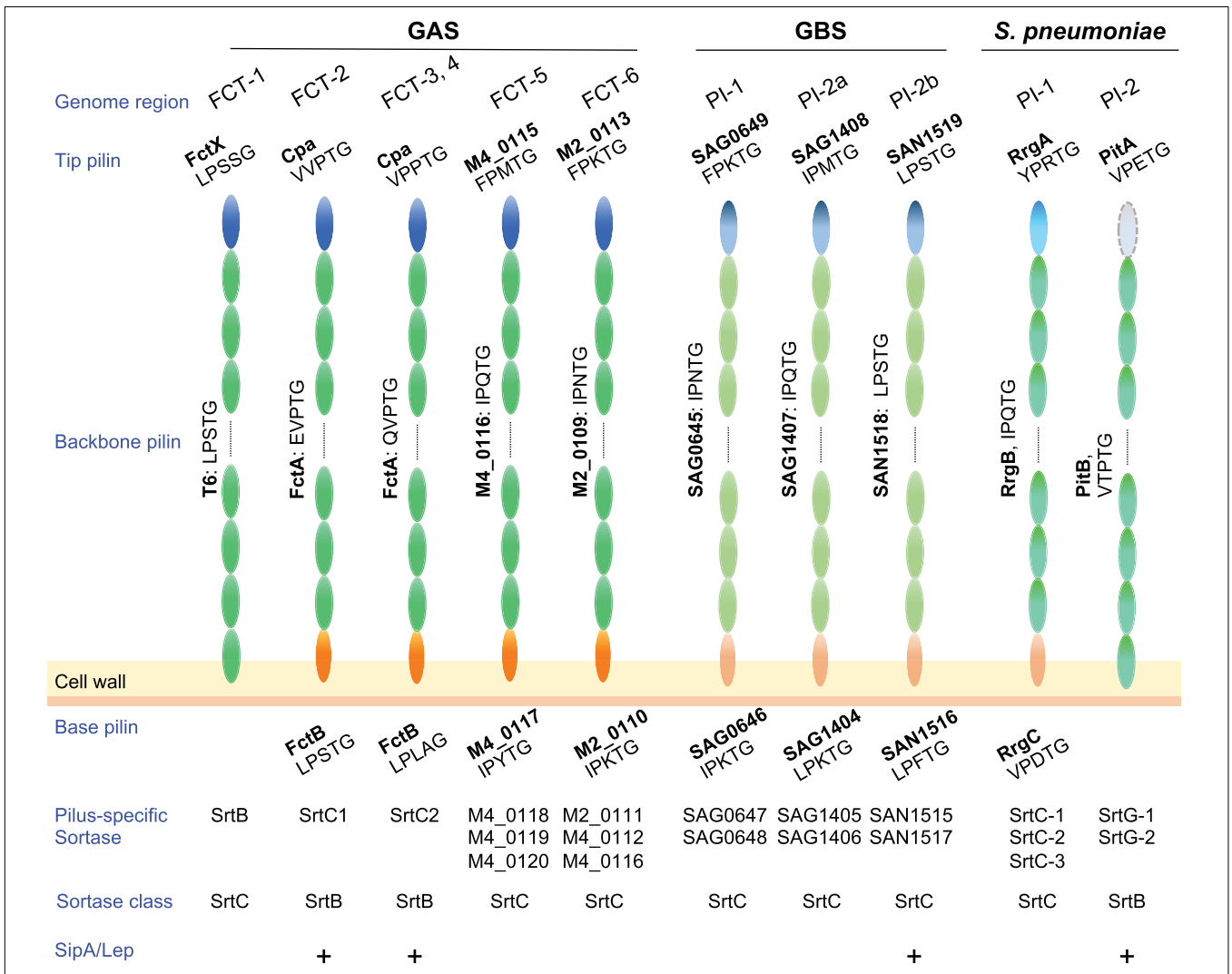


FIGURE 2 | Pilus components and related factors of pathogenic streptococci. Pilus major (backbone pilin) and minor (tip and base pilin) subunits of *Streptococcus pyogenes* (GAS), *Streptococcus agalactiae* (GBS), and *Streptococcus pneumoniae* are depicted. GBS pilus genes are located in three pilus island (PI) types, while *S. pneumoniae* pilus genes are located in two types of pilus islets (PIs). LPXTG or an LPXTG-like motif is shown under each pilin. Related pilus-specific sortases and their class, i.e., sortase class B or C, are also shown. Note that GAS SrtB and SrtC1/2 belongs to class C and B, respectively. A requirement of SipA/LepA for pilus assembly is also shown by "+." Pilin subunits of FCT forms 5 and 6 are presented as gene tag numbers of MGAS10750 (serotype M4, genome accession number NC_008024) and MGAS10270 (serotype M2, genome accession number NC_008022), respectively. Subunits of GBS pili are also shown as tag numbers in 2603V/R (serotype V, genome accession number NC_004116.1) and COH1 (serotype III, genome accession number HG939456).

the housekeeping sortase SrtA (Mazmanian et al., 1999. Thus far, assembly mechanisms have been mainly investigated for FCT-1, FCT-2, and FCT-3 pili in GAS.

Among the nine FCT forms, FCT-2, FCT-3, and FCT-4 pili comprise three components, including the major subunit FctA, and minor subunits Cpa (ap1) and FctB (ap2) (**Figure 2**). The pilus-specific sortase has been named SrtC. It comprises two alleles; SrtC1 in FCT-2 and SrtC2 in FCT-3 and FCT-4 (Barnett et al., 2004; Dramsi et al., 2005; Spirig et al., 2011). In an M3 strain, SrtC2 was shown to be responsible both for linkage between Cpa and FctA, as well as FctA polymerization (Quigley et al., 2009). FctA and Cpa in FCT-3 contain LPXTG-like VPPTG and QVPTG sorting sequences, respectively, with isopeptide bonds formed with K173 of FctA (Quigley et al., 2009). Those authors also reported that Cpa is located exclusively at the pilus tip. The corresponding LPXTG-like sequences of Cpa and FctA in M1 strains (FCT-2) are EVPTG and VVPTG, respectively, with the slight differences in substrate sequences likely attributed to variations in the substrate recognition of two SrtC alleles and two SipA/LepA alleles (**Figure 2**). Covalent linkage between the C-terminal threonine of Cpa and a lysine residue of FctA was also reported in an M1 strain (Smith et al., 2010). For anchoring of pili to the cell wall by SrtA, that report also noted that the minor subunit FctB was incorporated into the base of the Cpa-FctA complex as a stop signal for FctA polymerization. The FctB protein of M1 strains contains a canonical LPXTG motif (LPLAGE in FCT-3), which is a substrate for SrtA. A distinctive feature of FCT-2, FTC-3, and FCT-4 pili is that their assembly requires the signal peptidase I homolog SipA/LepA, the gene for which is also located in the pilus gene operon (Zähner and Scott, 2008; Nakata et al., 2009). Catalytic residues are not conserved in SipA/LepA, and *in vitro* assays using peptide fragments of pili and recombinant SipA/LepA show no peptide cleavage. Thus, SipA/lepA has been suggested to act as a molecular chaperone that coordinates pilus assembly with SrtC (Young et al., 2014b). The signal peptidase I homologs in *S. agalactiae* and *S. pneumoniae* are also required for assembly of pili (**Figure 2**; Bagnoli et al., 2008; Périchon et al., 2019), the genes for which are situated in pilus island II b (PI-2b) and the pathogenicity islet 2 (PI-2), respectively.

Assembly of FCT-1 pili in an M6 strain has been investigated (Nakata et al., 2011). Pili are composed of the major T6 subunit and minor FctX subunit as a tip protein. The CWSSs of T6 and FctX include an LPSTG and LPSSG sequence, respectively (Schneewind et al., 1990; Bessen and Kalia, 2002). The K175 residue of T6 was shown to participate in T6 polymerization as well as linkage of T6 and FctX (Nakata et al., 2011; Young et al., 2014a). The pilus-specific sortase SrtB belongs to the SrtC family and is primarily required for efficient pilus assembly while SrtA is responsible for cell wall anchoring of T6 pili (Nakata et al., 2011). Since deletion of the *srtB* gene does not completely abrogate T6 polymerization or formation of an FctX-T6 complex, as shown by immunoblot assay results, it is likely that SrtA can compensate for the loss of SrtB in pilus assembly to a certain extent (Nakata et al., 2011). Unlike FCT-2 and FCT-3 pili, there is no minor subunit for a stop signal and the mechanism of stopping polymerization remains unknown. Furthermore, several studies have demonstrated that deletion of the gene encoding the pilus tip protein (ap1) decreases the detection level of polymerized major subunits in FCT-1, FCT-2, FCT-3, and FCT-6 pili. This prompted speculation that heterodimer formation between the major and minor tip subunits accelerates polymerization of the major subunits (Lizano et al., 2007; Nakata et al., 2009, 2011; Tsai et al., 2017).

REGULATION OF PILUS GENE EXPRESSION

GAS pilus-related gene expression has been shown to be mediated by RALP family transcriptional regulators, including RofA and Nra, in a serotype- or strain-dependent manner. Previous reports have indicated that both RofA and Nra can function as autoregulators (Podbielski et al., 1999; Granok et al., 2000), and expression of pilus genes in FCT-3 strains is positively or negatively regulated by Nra in a strain-specific manner (Podbielski et al., 1999; Luo et al., 2008). Recent studies of in-frame deletion mutants indicated that Nra acts as a positive regulator in several M3 strains and an M49 strain (Calfee et al., 2018; Nakata et al., 2020). RofA has been reported to be a positive regulator for the protein F1 gene (*prtF*1) (Fogg et al., 1994). Also, the involvement of RofA in pilus gene expression was indirectly shown by replacement of *nra* with *rofA* along with respective upstream promoter regions from an M6 strain in the background of an M53 strain (FCT-3) (Lizano et al., 2008). The replacement resulted in preserved pilus gene expression. On the other hand, deletion of the *rofA* gene reduced pilus gene expression in an M1 strain (Calfee et al., 2018). Thus, RALP family members, such as Nra and RofA, likely promote pilus gene expression. In addition, MsmR, Mga, and RALP3 can influence pilus gene expression in a strain-specific manner (Nakata et al., 2005; Kreikemeyer et al., 2007; Kwinn et al., 2007; Luo et al., 2008).

In vitro induction of pilus gene expression occurs under a variety of culture conditions including low pH and low temperature (Nakata et al., 2009; Manetti et al., 2010). A topic gaining increasing focus is the molecular mechanisms that underlie modulation of pilus gene expression in response to environmental signals, such as temperature and acidity. The level of pilin detection is altered by shifting the culture temperature. Utilizing an M49 strain, we showed increased FctA expression at 30°C as compared to 37°C. Moreover, the expression of FctA appeared to be bistable, as only some cells in *S. pyogenes* chains were FctA-positive in immunofluorescence experiments (Nakata et al., 2009). Furthermore, decreasing the temperature to 25°C induced pilus production by the majority of cells (Nakata et al., 2020). This bistabilty was later characterized in more detail for type 1 pilus genes from pneumococci and shown to depend on the positive regulator RlrA acting in a positive feedback loop on pilus genes (Basset et al., 2011, 2017). Moreover, such bistabilty has clear implications for infections, as pilus-1 was shown to be preferentially expressed during early colonization in animal infection models (Pancotto et al., 2013). It remains to be determined if bistability of *S. pyogenes* pilus expression has any implications for *in vivo* pathogenesis ecology.

Historically, GAS cultures for a T-typing test have been grown at 30°C (Griffith, 1934). Thermosensitive pilus production occurs at the transcriptional level and the expression pattern is restricted to *nra*-positive FCT-3 strains (Nakata et al., 2009, 2020). The underlying mechanism involves post-transcriptional control of *nra* mRNA translation, namely promoting translation at low temperatures (**Figure 3**). Of note, in a study that utilized an M3 and an M49 strain, introduction of silent base substitutions in the chromosome to melt the predicted stem loop structure located 23 bases downstream of the AUG start codon decreased detection of both the Nra protein and pili especially at low temperatures (Nakata et al., 2020). Considering that temperature at the initial infection site is lower than the core body temperature and formation of mRNA stem-loop structure is influenced by temperature, we speculated that the predicted stem-loop structure is an mRNA thermometer within *nra* mRNA. It might form the stable base-pairing at lower temperatures, and could be more susceptible to melting at the core body temperature. Thus, lower temperatures reflecting the initial infection site would promote the *nra* translation and subsequent pilus gene expression, thereby promoting colonization. If temperature increased due to bacterial invasion of tissues and inflammation, pilus production by FCT-3 strains would be halted. How the stem loop promotes translation of *nra* mRNA remains elusive. It possibly involves a "starting block" mechanism, whereby the stem-loop prevents the 30S ribosomal subunit from sliding onto mRNA, optimizing the positioning of 30S rRNA and promoting translation initiation (**Figure 3**; Jagodnik et al., 2017). Although additional experimental confirmation is required to elucidate the mechanism, the existence of a temperature-perception system in the pilus gene transcriptional regulator adds a new level of regulation to virulence factor expression in GAS. Furthermore, Kratovac et al. noted that among 39 *emm* types associated with FCT-3, 32 (88.8%) represented pattern D of skin specialists (Kratovac et al., 2007; Bessen, 2016) raising the possibility of a link between thermosensitive pilus expression of FCT-3 strains and skin tropism. The molecular interactions between pili and host factors in skin have yet to be elucidated. Moreover, whether regulation of pilus gene expression studied *in vitro* matches the *in vivo* situation and thereby could be translated into clinical scenarios remains to be investigated. If pilus expression *per se* contributes to switching *S. pyogenes* lifestyles, it will most likely occur in a serotype- or even strain/isolate-specific manner.

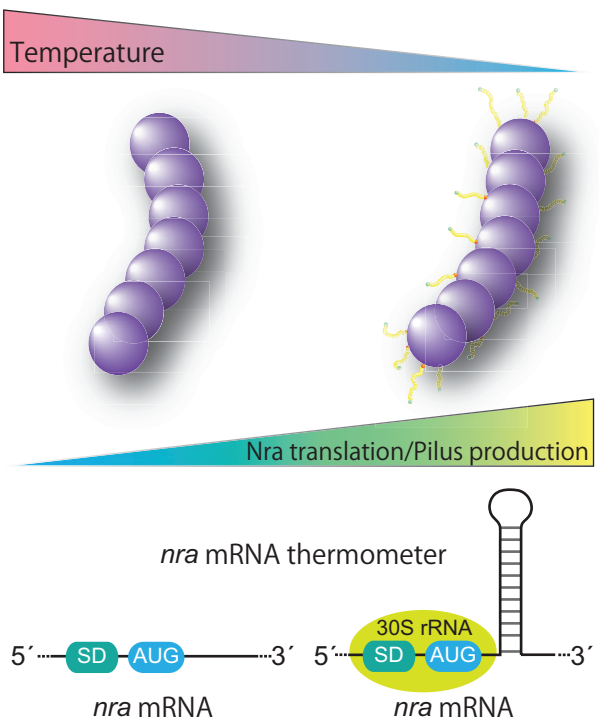


FIGURE 3 | Proposed model for thermoregulation of pilus production from FCT form 3 strains. FCT form 3 *nra*-positive *Streptococcus pyogenes* produces pili in a temperature-dependent manner. The underlying mechanism includes post-transcriptional control of *nra* mRNA translation via a putative stem loop structure in the protein coding region of *nra* mRNA. The putative stem loop structure most likely functions as a thermometer to modulate the translational efficiency of *nra* mRNA by potential interactions with the translation initiation complex. Thermosensitive modulation of pilus production highlights the importance of pili in an initial infection phase and involvement of pili in bacterial fitness in the host.

Upregulation of pilus expression by acid stress is also conducted by two-component systems. Comparisons of genome sequences of M3 (FCT-3) isolates recovered from symptomatic pharyngitis and subsequent asymptomatic carriage in the same patient (at day 63 post infection) revealed three single nucleotide polymorphisms, including one mutation in the sensor kinase LiaS (R135G) of the LiaFSR three-component system (or YvqE of the YvqEC two-component system) (Flores et al., 2015). The mutation was shown to alter the transcriptome. It also resulted in an increased ability to adhere to cultured epithelial cells and to colonize nasopharynx tissues, increased susceptibility to antibiotics targeting cell wall synthesis and decreased virulence in a mouse model of necrotizing fasciitis (Flores et al., 2015). An M3 strain with the R135G substitution in LiaS expressed pili at a higher level when exposed to bacitracin, a condition which is known to promote *liaFSR* expression. It seems that bacitracin did not promote the pilus gene expression in the R135G mutant (Flores et al., 2017). On the other hand, LiaS was demonstrated to play an important role in virulence and to sense acidic conditions, as an *liaS* mutant of an M1 strain had reduced growth at pH 6.0 (Ichikawa et al., 2011). Also, an *liaS* mutant produced less acid during sugar fermentation (Isaka et al., 2016). LiaS also controls pilus production and biofilm formation under acidic conditions. Introduction of the D26N substitution into a predicted extracellular region of LiaS compromised acid production and biofilm formation. This indicates that sensing organic acid by the LiaS extracellular domain is connected to the relationship between pilus expression and pH (Manetti et al., 2010; Isaka et al., 2016). It is not known if LiaS directly regulates pilus gene expression. Deletion of the homologue of *liaS* in *S. pneumoniae* and *S. agalactiae* also showed altered pilus expression (Rosch et al., 2008; Klinzing et al., 2013).

Non-coding RNA (ncRNA) is a crucial element in modulating virulence factor expression (Caldelari et al., 2013). Genome-wide tiling array and differential RNA sequencing analyses of GAS

ncRNAs together with bioinformatic and expression analyses revealed a *bona fide* expression of ncRNA in M1, M3, and M49 strains, including crRNA and trans-activating RNA of the Cas9-CRISPR system (Perez et al., 2009; Raasch et al., 2010; Deltcheva et al., 2011; Patenge et al., 2012, 2015; Tesorero et al., 2013). Among GAS ncRNA *fasX* negatively regulates translation of *cpa* in an M1 strain, *tee*6 in an M6 strain, the gene encoding a minor subunit (ap1) in an M2 strain, and *fctA* in an M28 strain (Liu et al., 2012; Danger et al., 2015). The mechanism involves base pairing of *fasX* with ribosome-binding sites, which leads to a reduction in mRNA stability and translation. *fasX* ncRNA is under the control of the FasBCA two-component system, originally reported to promote streptokinase production and haemolysis, and downregulating fibronectin/fibrinogen binding (Kreikemeyer et al., 2001). The environmental cues leading to expression of *fasBCAX* remain unknown. The levels of *fasX* expression show intraspecies variability as well as differences among FCT-3-associated serotypes including M3. For example, many M3 strains harbor a 4 bp deletion in the *fasC* gene encoding a histidine sensor kinase and consequently *fasX* expression is relatively low, which promotes pilus expression (Perez et al., 2009). However, inter-serotype transcriptome results demonstrated that several M3 strains have lower amounts of pili compared to several M1 and M49 strains due to a lower level of *nra* expression (Calfee et al., 2018). Together with frameshift mutations in *rocA* and *rivR* encoding a pseudo-kinase and transcriptional regulator, respectively (Biswas and Scott, 2003; Roberts et al., 2007), the loss of *fasX* regulation contributes to a selective advantage for M3 strains (Sarkar and Sumby, 2017).

The CovRS (or CsrRS) system is one of the most studied two-component systems and known to regulate approximately 15% of GAS genes (Graham et al., 2002). A prominent feature of this system is downregulation of several different virulence factors and natural mutations of these genes are strongly correlated with the onset of invasive diseases. A mutation of *covR* encoding the response regulator promoted pilus expression in several FCT-3 strains of serotypes M3 and M49, but not in an FCT-2 M1 strain (Kreth et al., 2011; Horstmann et al., 2015; Calfee et al., 2018) reflecting the fact that the intergenic regions of *nra-cpa* and *rofA-cpa* are divergent. Thus, counteractive modulation of pilus expression is governed by Nra and CovRS in large numbers of FCT-3 strains.

As shown by these findings, regulation of pilus expression is governed by various factors, including transcriptional regulators, ncRNA, and the mRNA thermometer. An intricate interplay among those factors shapes expression in response to the extracellular milieu and intracellular metabolic activity of the infected host cell. The FCT form exhibiting specific patterns of pilus expression highlights the importance of pili in regard to adaptation to host environments and their requirement for causing a variety of diseases. Further elucidation of these regulatory mechanisms is warranted.

BIOLOGICAL FUNCTIONS OF GAS PILI

The various functions of the GAS pili are summarized in **Table 1**.

Adhesion to Host Cells

A diverse array of secreted and surface-anchored components can mediate host cell adherence. These include cell wall anchored surface proteins, namely MSCRAMMs (microbial surface components recognizing adhesive matrix molecules, Patti et al., 1994), such as M proteins and a variety of fibronectin-binding, laminin-binding, and collagen-binding proteins. Other surface components are lipoteichoic acid, a hyaluronic acid capsule, and moonlighting proteins, including streptococcal enolase and glyceraldehyde-3-phosphate dehydrogenase (Brouwer et al., 2016). Cultured cell lines such as HEp-2 (HeLa derivative) and A549 (human alveolar adenocarcinoma cell line) have been

TABLE 1 | Involvement of GAS pili in host cell adherence, interactions with host molecules, biofilm formation, and virulence.

FCT form	M type (T type)	Host cell ADHERENCE Cell lines/tissues—adhesin	Interactions between host molecules and pilin	Involvement in biofilm formation*	Effects of pilus gene deletion on virulence in infection models (Host)*
FCT-1	M6 (T6)	A549—FctX	Gp340—FctA	+	Decreased virulence I. p. infection (CD1 mouse)
FCT-2	M1 (T1)	HaCaT, Detroit-562—Cpa Tonsil epithelium	Gp340—FctA	+	Increased virulence S. c. infection (CD1 mouse)
FCT-3	M3 (T3) M49 M53	HaCaT	Human collagen type I—Cpa (M49) Gp340—FctA (M3)	– (M49)	Decreased virulence S. c. infection (Human skin-engrafted SCID mouse) (M53)
FCT-5	M4 (T4)	HaCaT, RPMI 2650	Haptoglobin—SPyM4_0116[a]	+	Decreased virulence S. c. and i.p. infection (CD1 mouse)
FCT-6	M2 (T2)	HaCaT, Detroit-562 - SPyM2_0109[#]	Fibronectin and Fibrinogen—SPyM2_0109[#]		Decreased virulence Infection into the lower left proleg (*Galleria mellonella*)

**I. p., intraperitoneal; S. c., subcutaneous.*
[a] Tag number of MGAS10750 (genome accession number, NC_008024).
[#] Tag number of MGAS10270 (genome accession number, NC_008022).

frequently used to identify and analyze those factors. Abbot et al. (2007) utilized both cell lines and showed that an *fctA* deletion in the M1 strain SF370 had no effect on bacterial adherence. On the other hand, Crotty Alexander et al. reported that an *fctA* deletion in M1 strain 5448 induced a slight though statistically significant decrease in bacterial adherence to HEp-2 cells, although complementation failed to recover completely the reduction in adherence (Crotty Alexander et al., 2010). Importantly, T1 pili promoted adhesion to clinically relevant tissues in the throat and skin. Abbot et al. (2007) clearly demonstrated T1 pili binding to freshly isolated human tonsil tissues and to primary human keratinocytes, as well as to the human keratinocyte cell line HaCaT. Adhesion of M1 strain SF370 to the pharyngeal cell line Detroit 562 also showed T1 pili dependence (Manetti et al., 2007; Smith et al., 2010).

The above-mentioned findings raised questions regarding how T1 pili recognize host cells and which pilin component is responsible for binding. Flow cytometric analyses with recombinant pilin components revealed that recombinant Cpa and FctB, but not FctA, bound to the surface of Detroit 562 cells (Manetti et al., 2007). However, inhibition assays with antisera against each pilus subunit indicated that only anti-Cpa1 serum significantly inhibited bacterial adherence to both HaCaT cells and human tonsil epithelium (Smith et al., 2010). The central region of Cpa1 extending from Asn286 to Pro559 was responsible for bacterial adhesion via pili. Indeed, Cpa in T9 pili was considered to be a molecular harpoon that exerts adhesion via its own amine-reactive thioester bonds (Linke-Winnebeck et al., 2014), suggesting that Cpa plays a central role in cell recognition. It is likely that FctA polymerization allows Cpa to be located away from the cell surface and placed in the vicinity of host cells. This scheme would also be applicable to *cpa* operon-positive FCT-3 and FCT-4 forms. Involvement of either a major or minor subunit of other FCT-form pili in host cell adherence has also been reported. Regarding FCT-1 pili, the minor subunit FctX was shown to contribute to adherence of an M6 strain to A549 cells (Becherelli et al., 2012). In contrast, the tip pilin protein (SPyM2_0113) of an M2 strain (FCT-6) did not promote adhesion to HaCaT or Detroit-562 cells (Tsai et al., 2017). Instead, the major subunit T2 (SPyM2_0109) promoted adherence. Thus, in general, GAS pili can function as primary adhesins during the initial stages of colonization in the upper respiratory tract or skin.

Details regarding pilus receptors remain elusive, though some interactions between tip pilin and host molecules have been reported. The tip pilin Cpa from an M49 strain binds human collagen type I (Kreikemeyer et al., 2005). A high affinity interaction with a Kd in the nanomolar to low micromolar range was measured by ELISA-type assays and surface plasmon resonance (Kreikemeyer et al., 2005). This somehow resembles function of the collagen adhesion Cna of *Staphylococcus aures*. Interestingly, binding of major pilus subunits to host proteins has also been reported. FCT-1, FCT-2, and FCT-3 pili bind the salivary glycoprotein gp340 and the major subunit T2 binds fibronectin and fibrinogen (Edwards et al., 2008; Tsai et al., 2017). FctA plays a major role in interactions between FCT-2 T1 pili, but not FCT-4, and gp340. Binding of gp340 to bacteria mediates bacterial aggregation and it also inhibits bacterial adhesion to Detroit-562 and HeLa cells. Since gp340 binds to secretory IgA and complement C1q (Madsen et al., 2010) bacterial aggregation via gp340 binding may promote bacterial clearance and contribute to innate immunity.

Bacterial Aggregation and Biofilm Formation

Similar to many other pathogenic bacteria, GAS forms micro-colonies and biofilm on both biotic and abiotic surfaces. These actvities have been shown in numerous *in vitro* studies as well as *in vivo* infection models with various hosts, including mice, zebrafish, and chinchillas (Neely et al., 2002; Roberts et al., 2010; Connolly et al., 2011). More importantly, micro-colonies and biofilm-like structures have been found in clinical specimens of human impetigo lesions and tonsil tissues (Akiyama et al., 2003; Roberts et al., 2012). The 3-D structure of bacterial biofilm is defined by sessile cells being encased in a matrix of extracellular polymeric substances comprising proteins, DNA, and a glycocalyx. Biofilm-embedded bacteria exhibit a low growth rate and reduced metabolism, which poses problems when attempting antibiotic therapy (Donlan and Costerton, 2002). Moreover, several clinical studies have noted that the ability of GAS to form biofilm is related to recurrent infection episodes (Roberts et al., 2010; Torretta et al., 2012). Therefore, GAS biofilms are likely to be clinically relevant and therapeutic approaches against them may be effective for infection control. The ability of GAS strains to form biofilm varies and *in vitro* conditions required for biofilm development differ among strains. A coating of matrix or serum proteins can promote biofilm formation on an abiotic surface (Lembke et al., 2006) so adhesins such as MSCRAMMS likely contribute to biofilm formation.

A systemic evaluation of relationships between biofilm formation and FCT forms was conducted using 183 clinical isolates in Germany (Köller et al., 2010). This study showed that all tested FCT-1 strains, including M6 strains, efficiently formed massive biofilms in peptide-rich but carbohydrate-poor C medium, as well as in enrichment medium such as brain heart infusion (BHI) broth. Several FCT-2, FCT-5, and FCT-6 strains also formed biofilms in BHI, though to a lesser extent in C-medium, while two FCT-9 strains formed weak biofilms regardless of the culture medium. Furthermore, FCT-3 and FCT-4 strains showed a widely varying ability to form biofilm in both types of media. These phenotypic variations indicate that the ability to form biofilm is affected, at least in part, by culture conditions and can be roughly grouped by FCT form, although strain specificity occurs within some FCT forms. Several studies have shown a direct role for pili in biofilm formation (Manetti et al., 2007; Becherelli et al., 2012; Kimura et al., 2012; Chen et al., 2020). In the study of Manetti et al., T1 pili promoted *in vitro* biofilm formation by an M1 strain on polylysine-coated glass via aggregation and microcolony formation. On the other hand, biofilm formation of a M49 strain (FCT-3) was not affected by several mutations of pilus-related genes (Nakata et al., 2009). The remarkable ability of M6 strains to form biofilm was shown to be attributable

to T6 pilus production (Kimura et al., 2012). Deletion of the gene encoding either the major (T6) or minor (FctX) subunit decreased biofilm formation, while the same was true when the *srtB* gene encoding the pilus-specific sortase was deleted. Surprisingly, as compared with the parental strain, these mutant showed increased bacterial aggregation. When the entire group of T6 pilus-related genes was ectopically expressed in the M1 strain SF370, biofilm formation was promoted and aggregation inhibited (Kimura et al., 2012). It seems that T1 and T6 pili mediate biofilm formation by different mechanisms. FctA pili mediate biofilm formation by auto-aggregation and microcolony formation, while the T6 pilus functions as an adhesin responsible for initial attachment leading to biofilm formation. However, the contribution of pili to aggregation remains controversial. Becherelli et al. (2012) reported that FctX exhibited homophilic interactions that mediated inter-bacterial contact, thereby mediating aggregation. The phenotypic difference in aggregation of the *fctX* mutant strain could be attributable to differences in experimental conditions. Further analysis is required before drawing a final conclusion. Becherelli et al. (2012) also reported that homophilic interactions were observed for Cpa with serotypes M1 and M3 strains, suggesting a general mechanism of aggregation via homophilic interactions between minor subunits.

Among culture conditions examined, an acidic environment facilitates biofilm formation by specific FCT form strains (Manetti et al., 2010). The authors compared the biofilm forming ability in C-medium at pH 6.4 and 7.5, and found that the lower pH condition was favorable for biofilm development with strains from FCT-2 (M1), FCT-3 (M3 and M5), FCT-5 (M4), and FCT-6 (M2). Pilus production was also upregulated at the lower pH, indicating a pH-dependent relationship between biofilm formation and pilus production. On the other hand, pH levels did not influence the biofilm forming ability of M28 and M89 strains belonging to FCT-4 or M75 strains belonging to FCT-9, though biofilm mass was relatively low. FCT-1 strains (M6 and M109) efficiently formed biofilms under both conditions. Thus, strains with specific FCT forms have the ability to sense the environmental acidity and form biofilm via increased pilus production. Such differential response to environmental signals influences variations in biofilm formation. For more details regarding GAS factors involved in biofilm formation please refer to the review of Fiedler et al. (2015).

Virulence in Infection Models

The relationship between GAS pilin expression and virulence has been examined in strains belonging to FCT-1, FTC-2, FCT-3, FCT-5, and FCT-6. The extent to which each pilus type promotes or attenuates virulence varies, indicating differences among the forms (Lizano et al., 2007; Luo et al., 2008; Nakata et al., 2009; Crotty Alexander et al., 2010; Becherelli et al., 2012; Rouchon et al., 2017; Tsai et al., 2017; Chen et al., 2020). The contribution of FCT-1 T6 pili to pathogenesis was examined in a mouse intraperitoneal infection model, which indicated that the minor subunit FctX contributes to bacterial dissemination to the spleen, lungs, and kidneys, as well as survival in blood (Becherelli et al., 2012). In contrast, FCT-2 T1 pili reduced virulence in a murine subcutaneous infection model and decreased bacterial survival in human blood. Additionally, FCT-2 T1 pili had no influence on neutrophil phagocytosis, complement deposition in human sera, or sensitivity to the cathelicidin-related antimicrobial peptide LL-37. However, FCT-2 T1 pili induced neutrophil IL-8 production, neutrophil endothelial transcytosis, and neutrophil extracellular traps (NETs), thereby promoting entrapment and killing of GAS via NETs. Utilizing a human skin-engrafted SCID mouse line and an M53 skin-tropic strain (FCT-3), Lizano et al. evaluated the role of the Cpa and FctA pilus subunits in superficial skin infection. Deletion of the gene encoding Cpa attenuated virulence, while the *fctA* mutant showed virulence comparable to that of the parent strain (Lizano et al., 2007).

T4 pili of an M4 non-encapsulated strain (FCT-5) promoted adherence to HaCaT cells and human nasal septum RPMI 2,650 cells, survival in human blood, and virulence in both mouse skin and peritoneal infection models (Chen et al., 2020). Other studies have showed that the major subunit of T4 pili sequesters the serum protein haptoglobin to confer M4 GAS resistance to antimicrobial peptides released by neutrophils and platelets (Köhler and Prokop, 1978; Lämmler et al., 1988; Chen et al., 2020). Binding to haptoglobin was not observed for M1 strains. Increased expression of the major subunit gene was also associated with virulence of a non-encapsulated M4 GAS strain in an intraperitoneal mouse infection model (Galloway-Pena et al., 2018).

Virulence of an M2 strain (FCT-6) was examined using a *Galleria mellonella* infection model. Survival of infected *G. mellonella* was decreased by deletion of all pilus-related genes (Tsai et al., 2017). The mutation compromised the ability to survive in both macrophage cell lines and human whole blood. The major pilin subunit bound fibrinogen, and fibrin clot formation in human plasma was partially inhibited in the presence of the recombinant major pilin.

THE STRUCTURES OF PILUS SUBUNITS AND OTHER PROTEINS ENCODED IN FCT REGION

The structures of several pilus proteins have been solved by X-ray crystallography. This has been instrumental in understanding the mechanisms of assembly and the trypsin-resistant property of pili. With pilin structures of other Gram-positive pathogens revealed, it has become evident that major and minor pilins are assembled in a modular fashion with tandem Ig-like domains of CnaA and/or CnaB domains, which are present in the *Staphylococcus aureus* adhesin Cna (Deivanayagam et al., 2002; Zong et al., 2005). Crystal structure analysis of the major subunit FctA from an M1 strain revealed that it is comprised of two immunoglobulin (Ig) folds, each of which contains a CnaB domain (Kang et al., 2007). Crystal packing of FctA showed a head-to-tail orientation, with the side chain of the lysine residue K161 adjacent to the C terminus of the next molecule (Kang et al., 2007). Mass spectrometry analysis of fragmented pili and gene mutagenesis analyses demonstrated that a covalent

linkage is formed between K161 and T311 on adjacent subunits. Those residues are positioned in the omega loop of the CnaB fold and the LPXTG-like sortase recognition motif, respectively (Kang et al., 2007). Linkage between K161 and T311 allows polymerization of FctA by the class B sortase SrtC1 (Spy0129) (Barnett et al., 2004). The canonical YPKN pilin motif is not present in FctA (Ton-That and Schneewind, 2003) and the K161 position is different from that observed in major subunits of other bacterial species such as SpaA of *Corynebacterium diphtheriae* (Kang et al., 2009). The SpaA acceptor lysine in a YPKN pilin motif is located on the last β-strand of the N-terminal domain, close to the junction between domains, whereas the location of the acceptor lysine of FctA is near the top of the N-terminal domain. Thus, it is speculated that the difference in position of the acceptor lysine is correlated with type of pilus-specific sortase, i.e., class B or class C (Kang et al., 2011).

In an earlier study, Kang et al. (2007) also uncovered a striking characteristic feature of Gram-positive pilus subunits, namely formation of intramolecular isopeptide bonds between side chains of lysine and asparagine. Unlike the sortase-mediated linkage between subunits, this isopeptide bond is autocatalytically formed close to the domain boundary by an intramolecular reaction that involves a glutamic acid residue and surrounding aromatic residues. One intramolecular isopeptide bond is formed in each of the two CnaB domains (K36-N168 in the N terminal domain and K179-N303). A lysine residue in the first β-strand is linked to an asparagine residue in the last β-strand, endowing the pilin with thermal stability, resistance to proteolysis and mechanical stress (Kang and Baker, 2009; Alegre-Cebollada et al., 2010). Also, an intramolecular isopeptide bond occurs in the minor pilus subunit Cpa and in other Gram-positive bacterial Antigen I/II family of proteins (Forsgren et al., 2010; Hagan et al., 2010; Larson et al., 2011; Walden et al., 2014).

A minor subunit located at the pilus tip is considered to play a critical role in binding to host cells because of its positional advantage to reach the cell surface. The tip protein in FCT-2, FCT-3, and FCT-4 strains is Cpa (**Figure 2**). The C terminus (carboxyl group of C-terminal threonine residue) of Cpa (Spy0125) is linked to the above-mentioned lysine residue (K161) responsible for intermolecular linkage of FctA. Examination of the crystal structure of the Cpa C-terminal region (SPy0125, N286-T723) from an M1 strain revealed a three-domain structure, two of which contain an intramolecular isopeptide bond, K297-D595 and K610- N715 (Pointon et al., 2010). Moreover, an unusual thioester bond is internally formed between the side chains of a cysteine and a glutamine residue. This type of thioester bond has only been reported in proteins of the immune system, such as complement C3 and C4, complement-like proteins, and α2-macroglobulin (Chu and Pizzo, 1994; Law and Dodds, 1997; Cherry and Silverman, 2006), suggesting potential for covalent binding of Cpa to host factors. As in the case of FctA, intramolecular isopeptide bonds contribute to resistance to proteolysis and thermostability, whereas an alteration affecting the thioester bond had less influence on protein stability (Walden et al., 2014). Prevention of the thioester from Cpa compromised the ability of an M1 strain to bind to HaCaT cells, indicating a direct role in interaction with the host (Pointon et al., 2010). Linke-Winnebeck et al. (2014) reported that the N-terminal domain of Cpa (CpaN) from an *emm* ST6030.1 strain contains an additional thioester bond. X-ray crystallography and mass spectrometry analyses found that CpaN forms a dimer cross-linked by a polyamine spermidine, which was derived from *Escherichia coli* during recombinant protein preparation. They also reported that both thioesters contribute to binding to spermidine. This indicates that the reactive thioester has a preference for amine groups although the mechanism of covalent receptor binding has yet to elucidated.

SfbI/PrtF1 and PrtF2/FbaB are fibronectin binding proteins whose genes are located in the FCT genomic region that contain N-terminal domains homologous to the thioester-containing domain of Cpa. A homology search using the domain revealed that similar thioester domains are also present in diverse Gram-positive bacterial cell wall proteins, suggesting that the reactivity of thioester bonds is exploited by other surface proteins in other pathogens (Linke-Winnebeck et al., 2014; Walden et al., 2015). SfbI/PrtF1 also binds to the A subunit of fibrinogen in a thioester-dependent manner (Walden et al., 2015). Together, these studies provide a paradigm shift in understanding interactions between host and pathogens. It is likely that Gram-positive bacterial adhesins evolved to use covalent binding to host cells. Future studies are needed to determine the host target molecules of pilus adhesins, such as Cpa, which may provide information about GAS tissue tropism and provide the basis for effective therapeutic intervention.

The crystal structure of the minor subunit FctB from a T9 strain (Linke et al., 2010) comprises Ig-like and proline-rich tail domains and has no intramolecular isopeptide bond. The lysine residue responsible for linkage to FctA resides in the final β-strand of the Ig-like domain. The LPXTG motif of FctB in an M1 strain is LPSTG while the LPXTG-like tripartite motif (LPLAG) was found in other serotypes (Janulczyk and Rasmussen, 2001). This motif is recognized by the house-keeping sortase SrtA and promotes cross-linking to an alanine residue in cell wall peptidoglycan. The incorporation of the basal pilin into growing pili halts polymerization of FctA. Thus, all pilus components are connected by isopeptide bonds and finally become anchored to the cell wall.

Although the exact function of SipA/LepA remains unknown some predictions can be made from structural and biochemical data previously reported by Young et al. (2013; 2014b). SPase-I and SipA/LepA have a peptide-binding groove. In the crystal packing of SipA/LepA, the peptide-binding groove of one molecule is associated with the N-terminal peptide chain of the other molecule, indicating an ability of SipA/LepA to bind peptides. Young et al. also performed pull-down assays with recombinant FctA containing extracellular regions of both the signal-peptide and sortase motif, but no association of SipA/LepA and FctA was noted. Additionally, there was no interaction between SipA/LepA and the pilus-specific sortase. The same was true for synthesized peptides encompassing the extracellular region of the signal-peptides of Cpa, FctA, and FctB and the sorting motif region of FctA. The authors speculated that no detectable association was attributable to the

non-physiological octameric structure of recombinant SipA and a possible requirement of the membrane-spanning region for fully functional SipA/LepA, as seen with SPase I (Carlos et al., 2000). Also, they speculated that SipA/LepA might recognize sorting signals of pilus subunits coordinately with SrtC, or provide a scaffold that modifies or deploys pilin proteins for SrtC enzymatic activity. Thus, the interactions between SipA/LepA and pilus-related factors remain unclear and further exploration is needed.

Finally, the crystal structure of SrtC1 (class B family) has also been reported (Kang et al., 2011). SrtC1 has a canonical sortase fold, in which 8-stranded β-sheets mainly in C-terminal regions generate a core β-barrel, with the surface modified with loops and helices. This β-barrel structure has a concave surface that provides an active site comprising the key catalytic residues Cys, His, and Arg. Differences occurred in conformation of β4/β5 and β7/β8 loops between two molecules in the crystal, which suggested a level of flexibility important for SrtC1 function. The catalytic residue His126 located at the start of the β4/β5 loop was positioned differently in the two molecules in the crystal, thus potentially enabling a dual acid/base role by protonating the leaving group in the cleavage reaction and deprotonating the attacking amine in the transfer reaction (Suree et al., 2009). The corresponding region of *B. anthracis* SrtA is also flexible (Weiner et al., 2010). Conformational flexibility also occurs in the β7/β8 loop of *B. anthracis* and *S. aureus* SrtA, which may be a lipid II-binding site (Suree et al., 2009; Weiner et al., 2010). The length and conformation of the β7/β8 loop is highly variable among different sortases, suggesting a role in binding the second substrate. The structure of SrtC1 is closely related to class B sortases from *S. aureus* and *B. anthracis*, which anchor NPQTN motif-containing surface proteins to the cell wall, and those share the same surface loops and helices (Zhang et al., 2004; Zong et al., 2004; Kang et al., 2011). On the other hand, the unique characteristics of pneumococcal pilus-specific class C sortases, including flexible lids and a C-terminal transmembrane region, were not observed in SrtC1. It was concluded that the pilus polymerizing activity is a consequence of the co-evolution of the pilin and the cognate sortase, thus enabling substrate selection (Kang et al., 2011).

It is not clear why FCT-2, FCT-3, and FCT-4 GAS use a class B sortase and SipA/LepA for pilus biogenesis. It may be associated with differences in domain structures (two domains in FctA compared to three to four domains in others) and positions of the nucleophilic lysine residues for intermolecular linkage. Additional biochemical and biophysical analysis of the interactions of pilin, SipA, and SrtC are needed to unravel the exact mechanism of assembly of these pili. Additionally, analysis of other FCT-form pili will provide insights regarding the biological consequences of GAS pilus diversity.

PROSPECTS FOR A PILUS-BASED VACCINE

No vaccines are currently available for GAS. The M protein has been proposed as a primary vaccine candidate since IgG reactive to the hypervariable N-terminal region induces complement deposition and phagocytosis (Jones and Fischetti, 1988). Thus, construction of N-terminal peptide chimeras from multiple M proteins formed 26- and 30-valent M protein-based vaccines (Steer et al., 2009a; Dale et al., 2013). These were designed to provide coverage against strains circulating in developed countries. They also exhibit protection against some strains expressing M proteins that are not included among the targets of the vaccine (Dale et al., 2013). However, efficacy remains uncertain in countries where circulating strains exhibit a high level of diversity (Steer et al., 2009b).

There are 21 known T serotypes. While any protective effects of antibodies in T-type specific serum have not been reported, vaccination with recombinant pilin proteins can be effective, as noted below. T-typing serum is directed to trypsin-digested pili and it is likely that epitopes for typing are not necessarily equivalent to those exposed on the surface of native pili. One possible reason is that different epitopes exposed by trypsinization are responsible for T typing. Development of a pilus-based vaccine would be beneficial since there is less variation of T-antigenicity than observed with M-antigenicity or *emm*-typing, and fewer antigens would provide comparable coverage. Faulgi et al. sequenced *tee* genes of 39 strains representing 23 *emm* types and classified the *tee* genes into 15 clusters, with a sequence identity of greater than 90% within each cluster. The authors suggested that a vaccine containing epitopes from 12 types of T antigens would provide 90% coverage in the United States and EU (Falugi et al., 2008). Thereafter, the *tee* genotype was extended to 18 types and six subtypes, and it is expected that protective epitopes from these 18 T antigens could provide nearly full coverage for globally disseminated strains (Steemson et al., 2014). Individual *tee* alleles are highly stable over time and among geographical locations, further supporting T antigens as suitable vaccine candidates. Moreover, pili protrude from the bacterial surface by up to 2 μm (Mora et al., 2005; Kang et al., 2007), leading to unimpeded accessibility by the immune system and allowing exposure of many epitopes. Recently, whole-genome sequence analyses of 1,454 invasive GAS strains in the United States showed that 1,388 (95.5%) had one of the 21 different pilus (*tee*) types (Chochua et al., 2017).

Immunization with pilus subunit proteins confers protective immunity in mouse infection models (Mora et al., 2005; Loh et al., 2017). When serum reactivity against pilus components was tested using 100 serum samples obtained from children recovering from GAS pharyngitis using a protein array carrying four kinds of major pilin subunits and seven minor subunits, 76 of the samples reacted with at least one pilin protein (Manetti et al., 2007). Also, IgG in five of six serum samples from acute rheumatic fever patients reacted with T6 (Young et al., 2014a). Those studies indicated that pili are produced *in vivo* during infection and elicit specific antibody responses which supports their relevance as vaccine targets. However, since invasive GAS strains of FCT-3 may not produce pili *in vivo* at the inner body temperature of 37°C, a potential drawback of a pilus-based vaccine is lack of effectiveness for a subset of strains. Furthermore, it is not clear whether vaccination with pilus-based

antigens generates autoantibodies, which was problematic with M protein-based vaccines.

An important concern related to development of a type-specific epitope-based vaccine is the potential of the bacterium to generate new epitopes by intragenic recombination leading to loss of opsonizing ability of antibodies directed toward this region (Jones et al., 1988). This has not been reported thus far for pilus genes. However, the FCT region is a recombination hotspot. The possibility that interaction with the immune system induces antigenic drift should be examined.

Since a major site of GAS infection is the mucosal surface of the upper respiratory tract, generation of a mucosal immune response might be important for providing protection against infection (D'Alessandri et al., 1978; Batzloff et al., 2005). The food-grade organism *Lactococcus lactis* has been used as a vehicle to deliver vaccine antigens without adjuvants to the mucosal surface and elicits immune responses in animal models (Robinson et al., 1997). *L. lactis* has also been tested for delivering mucosal vaccines against pili (Buccato et al., 2006). Immunization of rabbits with a heat-killed *L. lactis* strain expressing either FCT-3 or FCT-4 pili via the oral gavage elicited specific antibody responses (Loh et al., 2017). Anti-pilus antibodies inhibited bacterial adhesion and immune serum efficiently promoted opsonophagocytic killing of bacteria. The authors speculated that the T antigen was the most likely target for opsonophagocytic killing. Furthermore, intranasal immunization of mice with a pilus-expressing *L. lactis* strain also improved clearance rates of GAS following nasopharyngeal challenge. These results demonstrate the potential for a pilus-based vaccine to protect against GAS infection.

Exploitation of the pilus biogenesis system and *L. lactis* has been utilized to present non-pilus related antigens (Quigley et al., 2010; Chamcha et al., 2015). The *E. coli* maltose-binding protein (MBP) was fused to the C-terminal region of the pilus tip protein (Cpa) of GAS T3 pili and expressed with pilus genes in *L. lactis* allowing the MBP to be presented on the tip of pili. This strain induced both systemic and mucosal responses against the MBP. Localization of a vaccine antigen on the pilus tip and covalent fixation to the lactococcal cell wall may be an effective strategy to promote exposure of vaccine antigens, though the influence of pilus-biogenesis factors on immunization must be considered.

The group of Thomas Proft developed PilVax, another vaccine platform that uses the GAS pilus and *L. lactis*. Immunogenic peptides were inserted into 3 different loop regions of FctA, resulting in assembly of pili and presentation of multiple peptides on the surface of *L. lactis*. Mouse intranasal immunization was shown to elicit both systemic and mucosal responses (Wagachchi et al., 2018; Clow et al., 2020).

BIOTECHNOLOGICAL APPLICATIONS OF FCT PROTEINS

Covalent linkage between proteins has been employed in therapeutics, biomaterials, diagnostics, and vaccines (Reddington and Howarth, 2015). Several methods have been used to generate stable protein complexes with diverse features related to efficiency, specificity, and stability (Rashidian et al., 2013; Antos et al., 2016; Stevens et al., 2016). For example, a cross-linking method based on an intramolecular isopeptide bond of FctA from an M1 strain has been reported (Zakeri and Howarth, 2010; Abe et al., 2013). FctA was split into two fragments at the final b-strand, with one fragment (pilin-C, residues 18–299) containing the reactive K179 and the other termed isopeptag consisting of 16 amino acids including reactive N303. These two fragments spontaneously formed an isopeptide bond *in vitro* and in *E. coli* as well as on the surface of mammalian cells. The reaction yield and rate were independent of temperature (range 4–37°C) and pH (range 6–8) in several conventional buffer systems, thus highlighting numerous advantages over other methods (Zakeri and Howarth, 2010).

Attempts of applying engineered FctA of an M1 strain as a protein shackle have also been reported (Matsunaga et al., 2013). By utilizing the intramolecular isopeptide linkage of FctA, the molecule could spontaneously polymerize into nanochains under reductive conditions. Later, the Howarth laboratory investigated the feasibility of exploiting fragments of other proteins containing isopeptide bonds. The SpyCatcher-SpyTag system is based on the isopeptide bond formed in the CnaB2 domain (K31-D117) of FbaB of FCT-3 strains (Terao et al., 2002; Hagan et al., 2010; Zakeri et al., 2012). The CnaB2 domain was split into two fragments, a 13-residue peptide from the C-terminal β-strand containing reactive D117 (SpyTag), and the rest of the 138-residue fragment termed SpyCatcher containing the reactive K31 and catalytic E77. Those fragments formed a covalent bond with high affinity under a wide range of conditions (Zakeri et al., 2012). A distinct tag-catcher system, SdyTag-SdyCatcher, has also been engineered from the CnaB domain of a fibronectin-binding protein from *S. dysgalactiae* (Tan et al., 2016). Subsequently, the SpyLigase-SpyTag-KTag system was developed from the SpyCatcher-SpyTag system (Fierer et al., 2014). Briefly, SpyCatcher was split into two factors, the scaffold protein SpyLigase, and the 10-residue peptide KTag containing catalytic E77 and K31, so that the active lysine, aspartic acid, and catalytic glutamic acid residues could be separately distributed into three factors. A mixture of the three factors generated a linkage between SpyTag and KTag, though the formation was dependent on strict buffer conditions and low temperature (Veggiani et al., 2016). The SpyCatcher-SpyTag system was then further modified with a phage display selection as well as more rational design to reach an improved affinity, and applied to the Spy&Go protein purification system with a non-reactive SPyCatcher mutant (SpyDock) for affinity purification of Spy-tagged proteins (Keeble et al., 2019; Khairil Anuar et al., 2019; Keeble and Howarth, 2020).

Similar systems have also been developed using the D4 domain of the pneumococcal pilin RrgA (Izoré et al., 2010; Veggiani et al., 2016). This domain contains an isopeptide bond formed between K742 and N854, which is catalyzed by the adjacent catalytic residue E803. Two fragments, SnoopCatcher and the 12-residue peptide SnoopTag, are generated, which contain N854/E803 and K742, respectively. This system can be simultaneously used with the SpyCatcher-SpyTag system

(Veggiani et al., 2016). The SnoopCatcher-SnoopTag system was also developed into the SnoopLigase-SnoopTagJr-DogTag three-component system (Buldun et al., 2018), and showed a higher level of efficiency and required less strict buffer conditions as compared to the SpyLigase-SpyTag-KTag system. SpyCatcher-SpyTag and related systems have been used in a wide variety of applications, including protein labeling, stable and directional protein display on surfaces and particles, modular covalent assembly with scaffolds of multimeric structures of other proteins, and increasing enzyme resilience by cyclisation (Keeble and Howarth, 2020), as well as for the study of bacterial proteins (Hatlem et al., 2019). These systems can also be used in GAS research, such as generation of modular vaccine antigens, and promoting protein complexes for structural and functional analyses.

An attempt has also been made to introduce an isopeptide bond into a non-isopeptide-containing protein. Kwon et al. introduced lysine, glutamic acid, and asparagine residues, N13K, Q67E, and P117N, respectively, into rational positions in the non-isopeptide-containing CnaB-type fold of FctB, with one more change (V26F) that restricted movement of the engineered lysine residue to bring it closer to N117 and entrap it in a hydrophobic environment. Spontaneous formation of an isopeptide bond was observed and thermal stability was increased. This method for stabilizing IgG-like proteins could be adopted for engineering of antibodies that share similar β-clasp Ig-type domains (Kwon et al., 2017; Young and Baker, 2020).

DISCUSSION

Recent research with streptococcal pili has revealed the diversity of structure, function, and control of expression and revealed their potential as vaccines antigens. The revelation that the major pilin subunit is the T antigen underscored the importance of pili as an epidemiological marker. Combined with epidemiological and evolutionary studies, analyses of the diversity of the genetic organization of the FCT region indicate a relationship of its components with tissue tropism (Bessen, 2016). Furthermore, unexpected finding of an intramolecular isopeptide bond allowed development of tools with a wide range of applications (Keeble and Howarth, 2020). However, important pilus-related issues with regard to clinical and biological consequences await experimental confirmation.

Important questions have arisen related to pilus binding partners. In consideration of the positional advantage over other MSCRAMMs, primary contact with host molecules and interactions with host cells might be initiated by pili. In other words, bacteria may evolve to locate an adhesin at the tip of a long shaft, despite pilus synthesis being an energy-consuming process. This has inspired a hypothesis where host cell adherence determines both host specificity and tissue preference (Bessen, 2016). An especially intriguing question is which host molecules are targets of the thioester-containing domain (TED)-mediated linkage of the pilin adhesin Cpa. Binding partners for some pilin proteins have been reported although it is unclear if those interactions confer tissue specificity. Although TED-mediated binding of the fibronectin-binding protein PrtF1/SfbI to the human fibrinogen Aα subunit has been revealed, notable differences seen between structures of TED from various molecules hint at the presence of target specificity (Walden et al., 2015). The most prominent differences in Cpa proteins from different serotypes lie in the N-terminal region that contains a TED domain, which also raises the possibility of variations in ligands or binding affinity (Kreikemeyer et al., 2005). To clarify the relevance of TED-mediated binding for GAS pathogenesis, further research is needed for identification of binding partners of Cpa as well as other adhesive pilins whose ligands remain unknown.

Despite remarkable advances in structural analyses of pilin and related factors in several different bacterial species it is not well understood how Gram-positive pilus assembly is spatially and temporally organized on the cell wall. Detailed knowledge of the structures of protein complexes, such as the sortase/full-length pilin complex, will be required to gain insight into the functionality of sortases, including substrate specificity. Furthermore, SipA/LepA only exists in FCT-2, FCT-3, and FCT-4, with the class B sortase SrtC and requirement of SipA/LepA for pilus assembly in certain FCT forms remains to be addressed. It has been speculated that the function of SipA is recognition of pilin sorting signals in a coordinated manner with SrtC or that it constitutes a scaffold that positions pilin proteins for optimal sortase transpeptidase activity (Young et al., 2014b). Solving the structure in complex with full-length pilin and detailed mutagenesis analyses are expected to reveal the function.

The regulation of pilus biogenesis is complex and occurs at several different levels. Temperature dependent regulation is important in FCT-3 strains. This is primarily governed by post-transcriptional control of *nra* mRNA translation via a stem loop structure in the coding region (Nakata et al., 2020). The stem loop is positioned proximal to the ribosome border and is considered to be important for promoting translation, although other factors such as its distance from the translation start codon and other mRNA structures around the Shine and Dalgarno sequence may contribute. This type of translational regulation leads to the hypothesis that larger and likely specific subsets of the GAS mRNA repertoire is thermoregulated during the initial stage of infection in the throat and skin, where the temperature is lower than core body temperature. Such information may provide insight into the contribution of pili to tissue tropism. Differences in the mechanisms of controlling pilus gene expression between FCT-3 and other forms may be attributable to variations in pilus functions, such as sensitivity to the human immune system and relative contribution to virulence and fitness in the host. In other FCT forms, *nra* is replaced with *rofA* and molecular epidemiology suggests that these genes have undergone balancing selection (Bessen et al., 2005).

A long-standing goal of GAS research is development of an effective vaccine. Clinical trials of a multivalent M protein vaccine have been conducted (Kotloff et al., 2004). Along with factors

extracted by use of a population-derived sequence approach (Davies et al., 2019) and antigens shown to be effective in animal studies (Azuar et al., 2019), T antigens have also been demonstrated to be viable vaccine antigen candidates. Unlike a multivalent M protein vaccine, there are fewer T type variations. However, comparative structural analyses of three two-domain T antigens (FctA), including T3, T13, and T18, revealed that the overall core structure is conserved and variations are distributed through the entire region (Young et al., 2019). Ideally, a candidate vaccine antigen would comprise a multivalent linkage of whole T antigens or domains. Further comparative crystal structure analyses and examination of pilin regions for antigenicity may lead to refinement of protective epitopes and development of a peptide-based pilin vaccine. Theoretical findings of combinations of a multivalent vaccine with other antigens, such as a family of M-related proteins (Frost et al., 2020), have demonstrated increased vaccine coverage and enhanced effectiveness (Courtney et al., 2017). Such combinations with other vaccine antigens might offer potentiating effects on prophylactic efforts for combatting GAS infections.

AUTHOR CONTRIBUTIONS

Both authors conceived the concept for this review article, participated in writing the manuscript, contributed to reading, editing, and reviewing the manuscript.

FUNDING

The work of MN was financed by JSPS KAKENHI (Grant No. 19K22715). The work of BK was supported by grants from the Deutsche Forschungsgemeinschaft (DFG) and BMBF (Federal Ministry of Education) during the last 15 years.

ACKNOWLEDGMENTS

Both authors are indepted to all of their current and former team members, Bachelor-, Master-, and Ph.D.-students who have contributed to this work in Japan and Germany.

REFERENCES

Abbot, E. L., Smith, W. D., Siou, G. P., Chiriboga, C., Smith, R. J., Wilson, J. A., et al. (2007). Pili mediate specific adhesion of *Streptococcus pyogenes* to human tonsil and skin. *Cell Microbiol.* 9, 1822–1833. doi: 10.1111/j.1462-5822.2007.00918.x

Abe, H., Wakabayashi, R., Yonemura, H., Yamada, S., Goto, M., and Kamiya, N. (2013). Split Spy0128 as a potent scaffold for protein cross-linking and immobilization. *Bioconjug. Chem.* 24, 242–250. doi: 10.1021/bc300606b

Akiyama, H., Morizane, S., Yamasaki, O., Oono, T., and Iwatsuki, K. (2003). Assessment of *Streptococcus pyogenes* microcolony formation in infected skin by confocal laser scanning microscopy. *J. Dermatol. Sci.* 32, 193–199. doi: 10.1016/s0923-1811(03)00096-3

Alegre-Cebollada, J., Badilla, C. L., and Fernández, J. M. (2010). Isopeptide bonds block the mechanical extension of pili in pathogenic *Streptococcus pyogenes*. *J. Biol. Chem.* 285, 11235–11242. doi: 10.1074/jbc.M110.102962

Antos, J. M., Truttmann, M. C., and Ploegh, H. L. (2016). Recent advances in sortase-catalyzed ligation methodology. *Curr. Opin. Struct. Biol.* 38, 111–118. doi: 10.1016/j.sbi.2016.05.021

Azuar, A., Jin, W., Mukaida, S., Hussein, W. M., Toth, I., and Skwarczynski, M. (2019). Recent advances in the development of peptide vaccines and their delivery systems against group A Streptococcus. *Vaccines* 7, 58. doi: 10.3390/vaccines7030058

Bagnoli, F., Moschioni, M., Donati, C., Dimitrovska, V., Ferlenghi, I., Facciotti, C., et al. (2008). A second pilus type in *Streptococcus pneumoniae* is prevalent in emerging serotypes and mediates adhesion to host cells. *J. Bacteriol.* 190, 5480–5492. doi: 10.1128/JB.00384-08

Barnett, T. C., Patel, A. R., and Scott, J. R. (2004). A novel sortase, SrtC2, from *Streptococcus pyogenes* anchors a surface protein containing a QVPTGV motif to the cell wall. *J. Bacteriol.* 186, 5865–5875. doi: 10.1128/JB.186.17.5865-5875.2004

Barnett, T. C., and Scott, J. R. (2002). Differential recognition of surface proteins in *Streptococcus pyogenes* by two sortase gene homologs. *J. Bacteriol.* 184, 2181–2191. doi: 10.1128/jb.184.8.2181-2191.2002

Barocchi, M. A., Ries, J., Zogaj, X., Hemsley, C., Albiger, B., Kanth, A., et al. (2006). A pneumococcal pilus influences virulence and host inflammatory responses. *Proc. Natl. Acad. Sci. U.S.A.* 103, 2857–2862. doi: 10.1073/pnas.0511017103

Basset, A., Herd, M., Daly, R., Dove, S. L., and Malley, R. (2017). The pneumococcal type 1 pilus genes are thermoregulated and are repressed by a member of the Snf2 protein family. *J. Bacteriol.* 199, e00078-17. doi: 10.1128/JB.00078-17

Basset, A., Turner, K. H., Boush, E., Sayeed, S., Dove, S. L., and Malley, R. (2011). Expression of the type 1 pneumococcal pilus is bistable and negatively regulated by the structural component RrgA. *Infect. Immun.* 79, 2974–2983. doi: 10.1128/IAI.05117-11

Batzloff, M. R., Yan, H., Davies, M. R., Hartas, J., Lowell, G. H., White, G., et al. (2005). Toward the development of an antidisease, transmission-blocking intranasal vaccine for group A Streptococcus. *J. Infect. Dis.* 192, 1450–1455. doi: 10.1086/466528

Beall, B., Facklam, R., and Thompson, T. (1996). Sequencing *emm*-specific PCR products for routine and accurate typing of group A Streptococci. *J. Clin. Microbiol.* 34, 953–958. doi: 10.1128/JCM.34.4.953-958.1996

Becherelli, M., Manetti, A. G., Buccato, S., Viciani, E., Ciucchi, L., Mollica, G., et al. (2012). The ancillary protein 1 of *Streptococcus pyogenes* FCT-1 pili mediates cell adhesion and biofilm formation through heterophilic as well as homophilic interactions. *Mol. Microbiol.* 83, 1035–1047. doi: 10.1111/j.1365-2958.2012.07987.x

Bessen, D. E. (2016). Tissue tropisms in group A *Streptococcus*: what virulence factors distinguish pharyngitis from impetigo strains? *Curr. Opin. Infect. Dis.* 29, 295–303. doi: 10.1097/QCO.0000000000000262

Bessen, D. E., and Kalia, A. (2002). Genomic localization of a T serotype locus to a recombinatorial zone encoding extracellular matrix-binding proteins in *Streptococcus pyogenes*. *Infect. Immun.* 70, 1159–1167. doi: 10.1128/iai.70.3.1159-1167.2002

Bessen, D. E., Manoharan, A., Luo, F., Wertz, J. E., and Robinson, D. A. (2005). Evolution of transcription regulatory genes is linked to niche specialization in the bacterial pathogen *Streptococcus pyogenes*. *J. Bacteriol.* 187, 4163–4172. doi: 10.1128/JB.187.12.4163-4172.2005

Bessen, D. E., Sotir, C. M., Readdy, T. L., and Hollingshead, S. K. (1996). Genetic correlates of throat and skin isolates of group A Streptococci. *J. Infect. Dis.* 173, 896–900. doi: 10.1093/infdis/173.4.896

Biswas, I., and Scott, J. R. (2003). Identification of *rocA*, a positive regulator of *covR* expression in the group A Streptococcus. *J. Bacteriol.* 185, 3081–3090. doi: 10.1128/jb.185.10.3081-3090.2003

Bowen, A. C., Mahé, A., Hay, R. J., Andrews, R. M., Steer, A. C., Tong, S. Y., et al. (2015). The global epidemiology of impetigo: a systematic review of the population prevalence of impetigo and pyoderma. *PLoS One* 10:e0136789. doi: 10.1371/journal.pone.0136789

Brouwer, S., Barnett, T. C., Rivera-Hernandez, T., Rohde, M., and Walker, M. J. (2016). *Streptococcus pyogenes* adhesion and colonization. *FEBS Lett.* 590, 3739–3757. doi: 10.1002/1873-3468.12254

Buccato, S., Maione, D., Rinaudo, C. D., Volpini, G., Taddei, A. R., Rosini, R., et al. (2006). Use of *Lactococcus lactis* expressing pili from group B *Streptococcus* as a broad-coverage vaccine against streptococcal disease. *J. Infect. Dis.* 194, 331–340. doi: 10.1086/505433

Buldun, C. M., Jean, J. X., Bedford, M. R., and Howarth, M. (2018). SnoopLigase catalyzes peptide-peptide locking and enables solid-phase conjugate isolation. *J. Am. Chem. Soc.* 140, 3008–3018. doi: 10.1021/jacs.7b13237

Caldelari, I., Chao, Y., Romby, P., and Vogel, J. (2013). RNA-mediated regulation in pathogenic bacteria. *Cold Spring Harb. Perspect. Med.* 3, a010298. doi: 10.1101/cshperspect.a010298

Calfee, G., Danger, J. L., Jain, I., Miller, E. W., Sarkar, P., Tjaden, B., et al. (2018). Identification and characterization of serotype-specific variation in group A Streptococcus pilus expression. *Infect. Immun.* 86, e00792-17. doi: 10.1128/IAI.00792-17

Carapetis, J. R., Steer, A. C., Mulholland, E. K., and Weber, M. (2005). The global burden of group A streptococcal diseases. *Lancet Infect. Dis.* 5, 685–694. doi: 10.1016/S1473-3099(05)70267-X

Carlos, J. L., Paetzel, M., Brubaker, G., Karla, A., Ashwell, C. M., Lively, M. O., et al. (2000). The role of the membrane-spanning domain of type I signal peptidases in substrate cleavage site selection. *J. Biol. Chem.* 275, 38813–38822. doi: 10.1074/jbc.M007093200

Chamcha, V., Jones, A., Quigley, B. R., Scott, J. R., and Amara, R. R. (2015). Oral immunization with a recombinant *Lactococcus lactis*-expressing HIV-1 antigen on group A *Streptococcus* pilus induces strong mucosal immunity in the gut. *J. Immunol.* 195, 5025–5034. doi: 10.4049/jimmunol.1501243

Chen, Y. H., Li, S. H., Yang, Y. C., Hsu, S. H., Nizet, V., and Chang, Y. C. (2020). T4 pili promote colonization and immune evasion phenotypes of nonencapsulated M4 *Streptococcus pyogenes*. *mBio* 11, e01580-20. doi: 10.1128/mBio.01580-20

Cherry, S., and Silverman, N. (2006). Host-pathogen interactions in drosophila: new tricks from an old friend. *Nat. Immunol.* 7, 911–917. doi: 10.1038/ni1388

Chochua, S., Metcalf, B. J., Li, Z., Rivers, J., Mathis, S., Jackson, D., et al. (2017). Population and whole genome sequence based characterization of invasive group A Streptococci recovered in the United States during 2015. *mBio* 8, e01422-17. doi: 10.1128/mBio.01422-17

Chu, C. T., and Pizzo, S. V. (1994). α2-Macroglobulin, complement, and biologic defense: antigens, growth factors, microbial proteases, and receptor ligation. *Lab. Invest.* 71, 792–812.

Cisar, J. O., and Vatter, A. E. (1979). Surface fibrils (fimbriae) of *Actinomyces viscosus* T14V. *Infect. Immun.* 24, 523–531. doi: 10.1128/IAI.24.2.523-531.1979

Clow, F., Peterken, K., Pearson, V., Proft, T., and Radcliff, F. J. (2020). PilVax, a novel *Lactococcus lactis*-based mucosal vaccine platform, stimulates systemic and mucosal immune responses to *Staphylococcus aureus*. *Immunol. Cell Biol.* 98, 369–381. doi: 10.1111/imcb.12325

Connolly, K. L., Braden, A. K., Holder, R. C., and Reid, S. D. (2011). Srv mediated dispersal of streptococcal biofilms through SpeB is observed in CovRS+ strains. *PLoS One* 6:e28640. doi: 10.1371/journal.pone.0028640

Courtney, H. S., Niedermeyer, S. E., Penfound, T. A., Hohn, C. M., Greeley, A., and Dale, J. B. (2017). Trivalent M-related protein as a component of next generation group A streptococcal vaccines. *Clin. Exp. Vaccine Res.* 6, 45–49. doi: 10.7774/cevr.2017.6.1.45

Crotty Alexander, L. E., Maisey, H. C., Timmer, A. M., Rooijakkers, S. H., Gallo, R. L., von Köckritz-Blickwede, M., et al. (2010). M1T1 group A streptococcal pili promote epithelial colonization but diminish systemic virulence through neutrophil extracellular entrapment. *J. Mol. Med.* 88, 371–381.

Dale, J. B., Penfound, T. A., Tamboura, B., Sow, S. O., Nataro, J. P., Tapia, M., et al. (2013). Potential coverage of a multivalent M protein-based group A streptococcal vaccine. *Vaccine.* 31, 1576–1581. doi: 10.1016/j.vaccine.2013.01.019

D'Alessandri, R., Plotkin, G., Kluge, R. M., Wittner, M. K., Fox, E. N., Dorfman, A., et al. (1978). Protective studies with group A streptococcal M protein vaccine. III. Challenge of volunteers after systemic or intranasal immunization with Type 3 or Type 12 group A *Streptococcus*. *J. Infect. Dis.* 138, 712–718. doi: 10.1093/infdis/138.6.712

Danger, J. L., Cao, T. N., Cao, T. H., Sarkar, P., Treviño, J., Pflughoeft, K. J., et al. (2015). The small regulatory RNA FasX enhances group A *Streptococcus* virulence and inhibits pilus expression via serotype-specific targets. *Mol. Microbiol.* 96, 249–262. doi: 10.1111/mmi.12935

Davies, M. R., McIntyre, L., Mutreja, A., Lacey, J. A., Lees, J. A., Towers, R. J., et al. (2019). Atlas of group A streptococcal vaccine candidates compiled using large-scale comparative genomics. *Nat. Genet.* 51, 1035–1043. doi: 10.1038/s41588-019-0417-8

Deivanayagam, C. C., Wann, E. R., Chen, W., Carson, M., Rajashankar, K. R., Höök, M., et al. (2002). A novel variant of the immunoglobulin fold in surface adhesins of *Staphylococcus aureus*: crystal structure of the fibrinogen-binding MSCRAMM, clumping factor A. *EMBO J.* 21, 6660–6672. doi: 10.1093/emboj/cdf619

Deltcheva, E., Chylinski, K., Sharma, C. M., Gonzales, K., Chao, Y., Pirzada, Z. A., et al. (2011). CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. *Nature.* 471, 602–607. doi: 10.1038/nature09886

Donlan, R. M., and Costerton, J. W. (2002). Biofilms: survival mechanisms of clinically relevant microorganisms. *Clin. Microbiol. Rev.* 15, 167–193. doi: 10.1128/cmr.15.2.167-193.2002

Dramsi, S., Trieu-Cuot, P., and Bierne, H. (2005). Sorting sortases: a nomenclature proposal for the various sortases of Gram-positive bacteria. *Res. Microbiol.* 156, 289–297. doi: 10.1016/j.resmic.2004.10.011

Edwards, A. M., Manetti, A. G., Falugi, F., Zingaretti, C., Capo, S., Buccato, S., et al. (2008). Scavenger receptor gp340 aggregates group A Streptococci by binding pili. *Mol. Microbiol.* 68, 1378–1394. doi: 10.1111/j.1365-2958.2008.06220.x

Falugi, F., Zingaretti, C., Pinto, V., Mariani, M., Amodeo, L., Manetti, A. G., et al. (2008). Sequence variation in group A *Streptococcus* pili and association of pilus backbone types with lancefield T serotypes. *J. Iinfect. Dis.* 198, 1834–1841. doi: 10.1086/593176

Fiedler, T., Köller, T., and Kreikemeyer, B. (2015). *Streptococcus pyogenes* biofilms-formation, biology, and clinical relevance. *Front. Cell. Infect. Microbiol.* 5:15. doi: 10.3389/fcimb.2015.00015

Fierer, J. O., Veggiani, G., and Howarth, M. (2014). SpyLigase peptide-peptide ligation polymerizes affibodies to enhance magnetic cancer cell capture. *Proc. Natl. Acad. Sci. U.S.A.* 111, E1176–E1181. doi: 10.1073/pnas.1315776111

Flores, A. R., Jewell, B. E., Yelamanchili, D., Olsen, R. J., and Musser, J. M. (2015). A single amino acid replacement in the sensor kinase LiaS contributes to a carrier phenotype in group A *Streptococcus*. *Infect. Immun.* 83, 4237–4246. doi: 10.1128/IAI.00656-15

Flores, A. R., Olsen, R. J., Cantu, C., Pallister, K. B., Guerra, F. E., Voyich, J. M., et al. (2017). Increased pilus production conferred by a naturally occurring mutation alters host-pathogen interaction in favor of carriage in *Streptococcus pyogenes*. *Infect. Immun.* 85:e00949-16. doi: 10.1128/IAI.00949-16

Fogg, G. C., Gibson, C. M., and Caparon, M. G. (1994). The identification of *rofA*, a positive-acting regulatory component of *prtF* expression: use of an mγδ-based shuttle mutagenesis strategy in *Streptococcus pyogenes*. *Mol. Microbiol.* 11, 671–684. doi: 10.1111/j.1365-2958.1994.tb00345.x

Forsgren, N., Lamont, R. J., and Persson, K. (2010). Two intramolecular isopeptide bonds are identified in the crystal structure of the *Streptococcus gordonii* SspB C-terminal domain. *J. Mol. Biol.* 397, 740–751. doi: 10.1016/j.jmb.2010.01.065

Frost, H. R., Davies, M. R., Delforge, V., Lakhloufi, D., Sanderson-Smith, M., Srinivasan, V., et al. (2020). Analysis of global collection of group A *Streptococcus* genomes reveals that the majority encode a trio of M and M-like proteins. *mSphere* 5, e00806-19. doi: 10.1128/mSphere.00806-19

Frost, H. R., Sanderson-Smith, M., Walker, M., Botteaux, A., and Smeesters, P. R. (2018). Group A streptococcal M-like proteins: From pathogenesis to vaccine potential. *FEMS Microbiol. Rev.* 42, 193–204. doi: 10.1093/femsre/fux057

Galloway-Pena, J., DebRoy, S., Brumlow, C., Li, X., Tran, T. T., Horstmann, N., et al. (2018). Hypervirulent group A *Streptococcus* emergence in an acaspular background is associated with marked remodeling of the bacterial cell surface. *PLoS One* 13:e0207897. doi: 10.1371/journal.pone.0207897

Graham, M. R., Smoot, L. M., Migliaccio, C. A., Virtaneva, K., Sturdevant, D. E., Porcella, S., et al. (2002). Virulence control in group A *Streptococcus* by a two-component gene regulatory system: global expression profiling and *in vivo* infection modeling. *Proc. Natl. Acad. Sci. U.S.A.* 99, 13855–13860. doi: 10.1073/pnas.202353699

Granok, A. B., Parsonage, D., Ross, R. P., and Caparon, M. G. (2000). The RofA binding site in *Streptococcus pyogenes* is utilized in multiple transcriptional pathways. *J. Bacteriol.* 182, 1529–1540. doi: 10.1128/jb.182.6.1529-1540.2000

Griffith, F. (1934). The serological classification of *Streptococcus pyogenes*. *J. Hyg.* 34, 542–584. doi: 10.1017/s0022172400043308

Hagan, R. M., Björnsson, R., McMahon, S. A., Schomburg, B., Braithwaite, V., Bühl, M., et al. (2010). NMR spectroscopic and theoretical analysis of a spontaneously formed Lys-Asp isopeptide bond. *Angew. Chem. Int.* 49, 8421–8425. doi: 10.1002/anie.201004340

Hatlem, D., Trunk, T., Linke, D., and Leo, J. C. (2019). Catching a SPY: Using the SpyCatcher-SpyTag and Related Systems for Labeling and Localizing Bacterial Proteins. *Int. J. Mol. Sci.* 20, 2129. doi: 10.3390/ijms20092129

Hava, D. L., and Camilli, A. (2002). Large-scale identification of serotype 4 *Streptococcus pneumoniae* virulence factors. *Mol. Microbiol.* 45, 1389–1406. doi: 10.1046/j.1365-2958.2002.03106.x

Hendrickx, A. P., Budzik, J. M., Oh, S. Y., and Schneewind, O. (2011). Architects at the bacterial surface – sortases and the assembly of pili with isopeptide bonds. *Nat. Rev. Microbiol.* 9, 166–176. doi: 10.1038/nrmicro2520

Horstmann, N., Sahasrabhojane, P., Saldaña, M., Ajami, N. J., Flores, A. R., Sumby, P., et al. (2015). Characterization of the effect of the histidine kinase CovS on response regulator phosphorylation in group A *Streptococcus*. *Infect. Immun.* 83, 1068–1077. doi: 10.1128/IAI.02659-14

Ichikawa, M., Minami, M., Isaka, M., Tatsuno, I., and Hasegawa, T. (2011). Analysis of two-component sensor proteins involved in the response to acid stimuli in *Streptococcus pyogenes*. *Microbiology* 157, 3187–3194. doi: 10.1099/mic.0.050534-0

Isaka, M., Tatsuno, I., Maeyama, J., Matsui, H., Zhang, Y., and Hasegawa, T. (2016). The YvqE two-component system controls biofilm formation and acid production in *Streptococcus pyogenes*. *APMIS.* 124, 574–585. doi: 10.1111/apm.12538

Izoré, T., Contreras-Martel, C., El Mortaji, L., Manzano, C., Terrasse, R., Vernet, T., et al. (2010). Structural basis of host cell recognition by the pilus adhesin from *Streptococcus pneumoniae*. *Structure.* 18, 106–115. doi: 10.1016/j.str.2009.10.019

Jaffe, J., Natanson-Yaron, S., Caparon, M. G., and Hanski, E. (1996). Protein F2, a novel fibronectin-binding protein from *Streptococcus pyogenes*, possesses two binding domains. *Mol. Microbiol.* 21, 373–384. doi: 10.1046/j.1365-2958.1996.6331356.x

Jagodnik, J., Chiaruttini, C., and Guillier, M. (2017). Stem-loop structures within mRNA coding sequences activate translation initiation and mediate control by small regulatory RNAs. *Mol. Cell* 68, 158e–170e. doi: 10.1016/j.molcel.2017.08.015

Janulczyk, R., and Rasmussen, M. (2001). Improved pattern for genome-based screening identifies novel cell wall-attached proteins in Gram-positive bacteria. *Infect. Immun.* 69, 4019–4026. doi: 10.1128/IAI.69.6.4019-4026.2001

Jones, K. F., and Fischetti, V. A. (1988). The importance of the location of antibody binding on the M6 protein for opsonization and phagocytosis of group A M6 streptococci. *J. Exp. Med.* 167, 1114–1123. doi: 10.1084/jem.167.3.1114

Jones, K. F., Hollingshead, S. K., Scott, J. R., and Fischetti, V. A. (1988). Spontaneous M6 protein size mutants of group A Streptococci display variation in antigenic and opsonogenic epitopes. *Proc. Natl. Acad. Sci. U.S.A.* 85, 8271–8275. doi: 10.1073/pnas.85.21.8271

Kang, H. J., and Baker, E. N. (2009). Intramolecular isopeptide bonds give thermodynamic and proteolytic stability to the major pilin protein of *Streptococcus pyogenes*. *J. Biol. Chem.* 284, 20729–20737. doi: 10.1074/jbc.M109.014514

Kang, H. J., Coulibaly, F., Clow, F., Proft, T., and Baker, E. N. (2007). Stabilizing isopeptide bonds revealed in Gram-positive bacterial pilus structure. *Science.* 318, 1625–1628. doi: 10.1126/science.1145806

Kang, H. J., Coulibaly, F., Proft, T., and Baker, E. N. (2011). Crystal structure of Spy0129, a *Streptococcus pyogenes* class B sortase involved in pilus assembly. *PLoS One* 6:e15969. doi: 10.1371/journal.pone.0015969

Kang, H. J., Paterson, N. G., Gaspar, A. H., Ton-That, H., and Baker, E. N. (2009). The *Corynebacterium diphtheriae* shaft pilin SpaA is built of tandem Ig-like modules with stabilizing isopeptide and disulfide bonds. *Proc. Natl. Acad. Sci. U.S.A.* 106, 16967–16971. doi: 10.1073/pnas.0906826106

Keeble, A. H., and Howarth, M. (2020). Power to the protein: enhancing and combining activities using the Spy toolbox. *Chem. Sci.* 11, 7281–7291. doi: 10.1039/D0SC01878C

Keeble, A. H., Turkki, P., Stokes, S., Khairil Anuar, I., Rahikainen, R., Hytönen, V. P., et al. (2019). Approaching infinite affinity through engineering of peptide-protein interaction. *Proc. Natl. Acad. Sci. U.S.A.* 116, 26523–26533. doi: 10.1073/pnas.1909653116

Khairil Anuar, I., Banerjee, A., Keeble, A. H., Carella, A., Nikov, G. I., and Howarth, M. (2019). SpyGo purification of SpyTag-proteins using pseudo-SpyCatcher to access an oligomerization toolbox. *Nat. Commun.* 10, 1734. doi: 10.1038/s41467-019-09678-w

Kimura, K. R., Nakata, M., Sumitomo, T., Kreikemeyer, B., Podbielski, A., Terao, Y., et al. (2012). Involvement of T6 pili in biofilm formation by serotype M6 *Streptococcus pyogenes*. *J. Bacteriol.* 194, 804–812. doi: 10.1128/JB.06283-11

Klinzing, D. C., Ishmael, N., Hotopp, J., Tettelin, H., Shields, K. R., Madoff, L. C., et al. (2013). The two-component response regulator LiaR regulates cell wall stress responses, pili expression and virulence in group B *Streptococcus*. *Microbiology* 159, 1521–1534. doi: 10.1099/mic.0.064444-0

Köhler, W., and Prokop, O. (1978). Relationship between haptoglobin and *Streptococcus pyogenes* T4 antigens. *Nature* 271, 373. doi: 10.1038/271373a0

Köller, T., Manetti, A. G., Kreikemeyer, B., Lembke, C., Margarit, I., Grandi, G., et al. (2010). Typing of the pilus-protein-encoding FCT region and biofilm formation as novel parameters in epidemiological investigations of *Streptococcus pyogenes* isolates from various infection sites. *J. Med. Microbiol.* 59, 442–452. doi: 10.1099/jmm.0.013581-0

Kotloff, K. L., Corretti, M., Palmer, K., Campbell, J. D., Reddish, M. A., Hu, M. C., et al. (2004). Safety and immunogenicity of a recombinant multivalent group A Streptococcal vaccine in healthy adults: phase 1 trial. *JAMA* 292, 709–715. doi: 10.1001/jama.292.6.709

Kratovac, Z., Manoharan, A., Luo, F., Lizano, S., and Bessen, D. E. (2007). Population genetics and linkage analysis of loci within the FCT region of *Streptococcus pyogenes*. *J. Bacteriol.* 189, 1299–1310. doi: 10.1128/JB.01301-06

Kreikemeyer, B., Boyle, M. D., Buttaro, B. A., Heinemann, M., and Podbielski, A. (2001). Group A streptococcal growth phase-associated virulence factor regulation by a novel operon (Fas) with homologies to two-component-type regulators requires a small RNA molecule. *Mol. Microbiol.* 39, 392–406. doi: 10.1046/j.1365-2958.2001.02226.x

Kreikemeyer, B., Nakata, M., Köller, T., Hildisch, H., Kourakos, V., Standar, K., et al. (2007). The *Streptococcus pyogenes* serotype M49 Nra-Ralp3 transcriptional regulatory network and its control of virulence factor expression from the novel *eno ralp3 epf sagA* pathogenicity region. *Infect. Immun.* 75, 5698–5710. doi: 10.1128/IAI.00175-07

Kreikemeyer, B., Nakata, M., Oehmcke, S., Gschwendtner, C., Normann, J., and 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. doi: 10.1074/jbc.M502896200

Kreth, J., Chen, Z., Ferretti, J., and Malke, H. (2011). Counteractive balancing of transcriptome expression involving CodY and CovRS in *Streptococcus pyogenes*. *J. Bacteriol.* 193, 4153–4165. doi: 10.1128/JB.00061-11

Kwinn, L. A., Khosravi, A., Aziz, R. K., Timmer, A. M., Doran, K. S., Kotb, M., et al. (2007). Genetic characterization and virulence role of the RALP3/LSA locus upstream of the streptolysin s operon in invasive M1T1 group A *Streptococcus*. *J. Bacteriol.* 189, 1322–1329. doi: 10.1128/JB.01256-06

Kwon, H., Young, P. G., Squire, C. J., and Baker, E. N. (2017). Engineering a Lys-Asn isopeptide bond into an immunoglobulin-like protein domain enhances its stability. *Sci. Rep.* 7, 42753. doi: 10.1038/srep42753

Lämmler, C., Guszczynski, T., and Dobryszycka, W. (1988). Further characterization of haptoglobin binding to streptococci of serological group A. *Zentralbl. Bakteriol. Mikrobiol. Hyg. A.* 269, 454–459. doi: 10.1016/s0176-6724(88)80067-1

Lancefield, R. C. (1940). Type-specific antigens, M and T, of matt and glossy variants of group A Hemolytic Streptococci. *J. Exp. Med.* 71, 521–537. doi: 10.1084/jem.71.4.521

Lancefield, R. C., and Dole, V. P. (1946). The properties of t antigens extracted from group A Hemolytic Streptococci. *J. Exp. Med.* 84, 449–471. doi: 10.1084/jem.84.5.449

Larson, M. R., Rajashankar, K. R., Crowley, P. J., Kelly, C., Mitchell, T. J., Brady, L. J., et al. (2011). Crystal structure of the C-terminal region of *Streptococcus mutans* antigen I/II and characterization of salivary agglutinin adherence domains. *J. Biol. Chem.* 286, 21657–21666. doi: 10.1074/jbc.M111.231100

Lauer, P., Rinaudo, C. D., Soriani, M., Margarit, I., Maione, D., Rosini, R., et al. (2005). Genome analysis reveals pili in Group B *Streptococcus*. *Science.* 309, 105. doi: 10.1126/science.1111563

Law, S. K., and Dodds, A. W. (1997). The internal thioester and the covalent binding properties of the complement proteins C3 and C4. *Protein Sci.* 6, 263–274. doi: 10.1002/pro.5560060201

Lembke, C., Podbielski, A., Hidalgo-Grass, C., Jonas, L., Hanski, E., and Kreikemeyer, B. (2006). Characterization of biofilm formation by clinically relevant serotypes of group A Streptococci. *Appl. Environ. Microbiol.* 72, 2864–2875. doi: 10.1128/AEM.72.4.2864-2875.2006

Linke, C., Young, P. G., Kang, H. J., Bunker, R. D., Middleditch, M. J., Caradoc-Davies, T. T., et al. (2010). Crystal structure of the minor pilin FctB reveals determinants of Group A streptococcal pilus anchoring. *J. Biol. Chem.* 285, 20381–20389. doi: 10.1074/jbc.M109.089680

Linke-Winnebeck, C., Paterson, N. G., Young, P. G., Middleditch, M. J., Greenwood, D. R., Witte, G., et al. (2014). Structural model for covalent adhesion of the *Streptococcus pyogenes* pilus through a thioester bond. *J. Biol. Chem.* 289, 177–189. doi: 10.1074/jbc.M113.523761

Liu, Z., Treviño, J., Ramirez-Peña, E., and Sumby, P. (2012). The small regulatory RNA FasX controls pilus expression and adherence in the human bacterial pathogen group A *Streptococcus*. *Mol. Microbiol.* 86, 140–154. doi: 10.1111/j.1365-2958.2012.08178.x

Lizano, S., Luo, F., and Bessen, D. E. (2007). Role of streptococcal T antigens in superficial skin infection. *J. Bacteriol.* 189, 1426–1434. doi: 10.1128/JB.01179-06

Lizano, S., Luo, F., Tengra, F. K., and Bessen, D. E. (2008). Impact of orthologous gene replacement on the circuitry governing pilus gene transcription in Streptococci. *PLoS One* 3:e3450. doi: 10.1371/journal.pone.0003450

Loh, J., Lorenz, N., Tsai, C. J., Khemlani, A., and Proft, T. (2017). Mucosal vaccination with pili from group A Streptococcus expressed on *Lactococcus lactis* generates protective immune responses. *Sci. Rep.* 7, 7174. doi: 10.1038/s41598-017-07602-0

Luo, F., Lizano, S., and Bessen, D. E. (2008). Heterogeneity in the polarity of Nra regulatory effects on streptococcal pilus gene transcription and virulence. *Infect. Immun.* 76, 2490–2497. doi: 10.1128/IAI.01567-07

Madsen, J., Mollenhauer, J., and Holmskov, U. (2010). Review: Gp-340/DMBT1 in mucosal innate immunity. *Innate Immun.* 16, 160–167. doi: 10.1177/1753425910368447

Manetti, A. G., Köller, T., Becherelli, M., Buccato, S., Kreikemeyer, B., Podbielski, A., et al. (2010). Environmental acidification drives *S. pyogenes* pilus expression and microcolony formation on epithelial cells in a FCT-dependent manner. *PLoS One* 5:e13864. doi: 10.1371/journal.pone.0013864

Manetti, A. G., Zingaretti, C., Falugi, F., Capo, S., Bombaci, M., Bagnoli, F., et al. (2007). *Streptococcus pyogenes* pili promote pharyngeal cell adhesion and biofilm formation. *Mol. Microbiol.* 64, 968–983. doi: 10.1111/j.1365-2958.2007.05704.x

Matsunaga, R., Yanaka, S., Nagatoishi, S., and Tsumoto, K. (2013). Hyperthin nanochains composed of self-polymerizing protein shackles. *Nat. Commu.* 4, 2211. doi: 10.1038/ncomms3211

Mazmanian, S. K., Liu, G., Ton-That, H., and Schneewind, O. (1999). *Staphylococcus aureus* sortase, an enzyme that anchors surface proteins to the cell wall. *Science.* 285, 760–763. doi: 10.1126/science.285.5428.760

Mora, M., Bensi, G., Capo, S., Falugi, F., Zingaretti, C., Manetti, A. G., et al. (2005). group A *Streptococcus* produce pilus-like structures containing protective antigens and Lancefield T antigens. *Proc. Natl. Acad. Sci. U.S.A.* 102, 15641–15646. doi: 10.1073/pnas.0507808102

Nakata, M., Kimura, K. R., Sumitomo, T., Wada, S., Sugauchi, A., Oiki, E., et al. (2011). Assembly mechanism of FCT region type 1 pili in serotype M6 *Streptococcus pyogenes*. *J. Biol. Chem.* 286, 37566–37577. doi: 10.1074/jbc.M111.239780

Nakata, M., Köller, T., Moritz, K., Ribardo, D., Jonas, L., McIver, K. S., et al. (2009). Mode of expression and functional characterization of FCT-3 pilus region-encoded proteins in *Streptococcus pyogenes* serotype M49. *Infect. Immun.* 77, 32–44. doi: 10.1128/IAI.00772-08

Nakata, M., Podbielski, A., and Kreikemeyer, B. (2005). MsmR, a specific positive regulator of the *Streptococcus pyogenes* FCT pathogenicity region and cytolysin-mediated translocation system genes. *Mol. Microbiol.* 57, 786–803. doi: 10.1111/j.1365-2958.2005.04730.x

Nakata, M., Sumitomo, T., Patenge, N., Kreikemeyer, B., and Kawabata, S. (2020). Thermosensitive pilus production by FCT type 3 *Streptococcus pyogenes* controlled by Nra regulator translational efficiency. *Mol. Microbiol.* 113, 173–189. doi: 10.1111/mmi.14408

Neely, M. N., Pfeifer, J. D., and Caparon, M. (2002). Streptococcus-zebrafish model of bacterial pathogenesis. *Infect. Immun.* 70, 3904–3914. doi: 10.1128/iai.70.7.3904-3914.2002

Okahashi, N., Nakata, M., Sakurai, A., Terao, Y., Hoshino, T., Yamaguchi, M., et al. (2010). Pili of oral *Streptococcus sanguinis* bind to fibronectin and contribute to cell adhesion. *Biochem. Biophys. Res. Commun.* 391, 1192–1196. doi: 10.1016/j.bbrc.2009.12.029

Pancotto, L., De Angelis, G., Bizzarri, E., Barocchi, M. A., Del Giudice, G., Moschioni, M., et al. (2013). Expression of the *Streptococcus pneumoniae* pilus-1 undergoes on and off switching during colonization in mice. *Sci. Rep.* 3, 2040. doi: 10.1038/srep02040

Patenge, N., Billion, A., Raasch, P., Normann, J., Wisniewska-Kucper, A., Retey, J., et al. (2012). Identification of novel growth phase- and media-dependent small non-coding RNAs in *Streptococcus pyogenes* M49 using intergenic tiling arrays. *BMC Genomics* 13:550. doi: 10.1186/1471-2164-13-550

Patenge, N., Pappesch, R., Khani, A., and Kreikemeyer, B. (2015). Genome-wide analyses of small non-coding RNAs in streptococci. *Front. Genet.* 6:189. doi: 10.3389/fgene.2015.00189

Patti, J. M., Allen, B. L., McGavin, M. J., and Höök, M. (1994). MSCRAMM-mediated adherence of microorganisms to host tissues. *Annu. Rev. Microbiol.* 48, 585–617. doi: 10.1146/annurev.mi.48.100194.003101

Perez, N., Treviño, J., Liu, Z., Ho, S. C., Babitzke, P., and Sumby, P. (2009). A genome-wide analysis of small regulatory RNAs in the human pathogen group A *Streptococcus*. *PLoS One.* 4:e7668. doi: 10.1371/journal.pone.0007668

Périchon, B., Guignot, J., Szili, N., Gao, C., Poyart, C., Trieu-Cuot, P., et al. (2019). Insights into *Streptococcus agalactiae* PI-2b pilus biosynthesis and role in adherence to host cells. *Microbes Infect.* 21, 99–103. doi: 10.1016/j.micinf.2018.10.004

Podbielski, A., Woischnik, M., Leonard, B. A., and Schmidt, K. H. (1999). Characterization of *nra*, a global negative regulator gene in group A Streptococci. *Mol. Microbiol.* 31, 1051–1064. doi: 10.1046/j.1365-2958.1999.01241.x

Pointon, J. A., Smith, W. D., Saalbach, G., Crow, A., Kehoe, M. A., and Banfield, M. J. (2010). A highly unusual thioester bond in a pilus adhesin is required for efficient host cell interaction. *J. Biol. Chem.* 285, 33858–33866. doi: 10.1074/jbc.M110.149385

Quigley, B. R., Hatkoff, M., Thanassi, D. G., Ouattara, M., Eichenbaum, Z., and Scott, J. R. (2010). A foreign protein incorporated on the Tip of T3 pili in *Lactococcus lactis* elicits systemic and mucosal immunity. *Infect. Immun.* 78, 1294–1303. doi: 10.1128/IAI.01037-09

Quigley, B. R., Zähner, D., Hatkoff, M., Thanassi, D. G., and Scott, J. R. (2009). Linkage of T3 and Cpa pilins in the *Streptococcus pyogenes* M3 pilus. *Mol. Microbiol.* 72, 1379–1394. doi: 10.1111/j.1365-2958.2009.06727.x

Raasch, P., Schmitz, U., Patenge, N., Vera, J., Kreikemeyer, B., and Wolkenhauer, O. (2010). Non-coding RNA detection methods combined to improve usability, reproducibility and precision. *BMC Bioinform.* 11:491. doi: 10.1186/1471-2105-11-491

Rashidian, M., Dozier, J. K., and Distefano, M. D. (2013). Enzymatic labeling of proteins: techniques and approaches. *Bioconjug. Chem.* 24, 1277–1294. doi: 10.1021/bc400102w

Reddington, S. C., and Howarth, M. (2015). Secrets of a covalent interaction for biomaterials and biotechnology: SpyTag and SpyCatcher. *Curr. Opin. Chem. Biol.* 29, 94–99. doi: 10.1016/j.cbpa.2015.10.002

Roberts, A. L., Connolly, K. L., Doern, C. D., Holder, R. C., and Reid, S. D. (2010). Loss of the group A *Streptococcus* regulator Srv decreases biofilm formation *in vivo* in an otitis media model of infection. *Infect. Immun.* 78, 4800–4808. doi: 10.1128/IAI.00255-10

Roberts, A. L., Connolly, K. L., Kirse, D. J., Evans, A. K., Poehling, K. A., Peters, T. R., et al. (2012). Detection of group A *Streptococcus* in tonsils from pediatric patients reveals high rate of asymptomatic streptococcal carriage. *BMC Pediatr.* 12:3. doi: 10.1186/1471-2431-12-3

Roberts, S. A., Churchward, G. G., and Scott, J. R. (2007). Unraveling the regulatory network in *Streptococcus pyogenes*: the global response regulator CovR represses *rivR* directly. *J. Bacteriol.* 189, 1459–1463. doi: 10.1128/JB.01026-06

Robinson, K., Chamberlain, L. M., Schofield, K. M., Wells, J. M., and Le Page, R. W. (1997). Oral vaccination of mice against tetanus with recombinant *Lactococcus lactis*. *Nat. Biotechnol.* 15, 653–657. doi: 10.1038/nbt0797-653

Rocha, C. L., and Fischetti, V. A. (1999). Identification and characterization of a novel fibronectin-binding protein on the surface of group A Streptococci. *Infect. Immun.* 67, 2720–2728. doi: 10.1128/IAI.67.6.2720-2728.1999

Rosch, J. W., Mann, B., Thornton, J., Sublett, J., and Tuomanen, E. (2008). Convergence of regulatory networks on the pilus locus of *Streptococcus pneumoniae*. *Infect. Immun.* 76, 3187–3196. doi: 10.1128/IAI.00054-08

Rouchon, C. N., Ly, A. T., Noto, J. P., Luo, F., Lizano, S., and Bessen, D. E. (2017). Incremental contributions of FbaA and other impetigo-associated surface proteins to fitness and virulence of a classical group A streptococcal skin strain. *Infect. Immun.* 85, e00374-17. doi: 10.1128/IAI.00374-17

Sanderson-Smith, M., De Oliveira, D. M., Guglielmini, J., McMillan, D. J., Vu, T., Holien, J. K., et al. (2014). A systematic and functional classification of *Streptococcus pyogenes* that serves as a new tool for molecular typing and vaccine development. *J. Infect. Dis.* 210, 1325–1338. doi: 10.1093/infdis/jiu260

Sarkar, P., and Sumby, P. (2017). Regulatory gene mutation: a driving force behind group A *Streptococcus* strain- and serotype-specific variation. *Mol. Microbiol.* 103, 576–589. doi: 10.1111/mmi.13584

Schneewind, O., Jones, K. F., and Fischetti, V. A. (1990). Sequence and structural characteristics of the trypsin-resistant T6 surface protein of group A Streptococci. *J. Bacteriol.* 172, 3310–3317. doi: 10.1128/jb.172.6.3310-3317.1990

Sela, S., Aviv, A., Tovi, A., Burstein, I., Caparon, M. G., and Hanski, E. (1993). Protein F: an adhesin of *Streptococcus pyogenes* binds fibronectin via two distinct domains. *Mol. Microbiol.* 10, 1049–1055. doi: 10.1111/j.1365-2958.1993.tb00975.x

Smith, W. D., Pointon, J. A., Abbot, E., Kang, H. J., Baker, E. N., Hirst, B. H., et al. (2010). Roles of minor pilin subunits Spy0125 and Spy0130 in the serotype M1 *Streptococcus pyogenes* strain SF370. *J. Bacteriol.* 192, 4651–4659. doi: 10.1128/JB.00071-10

Spirig, T., Weiner, E. M., and Clubb, R. T. (2011). Sortase enzymes in Gram-positive bacteria. *Mol. Microbiol.* 82, 1044–1059. doi: 10.1111/j.1365-2958.2011.07887.x

Steemson, J. D., Moreland, N. J., Williamson, D., Morgan, J., Carter, P. E., and Proft, T. (2014). Survey of the *bp/tee* genes from clinical group A Streptococcus isolates in New Zealand - implications for vaccine development. *J. Med. Microbiol.* 63, 1670–1678. doi: 10.1099/jmm.0.080804-0

Steer, A. C., Batzloff, M. R., Mulholland, K., and Carapetis, J. R. (2009a). Group A streptococcal vaccines: facts versus fantasy. *Curr. Opin. Infecti. Dis.* 22, 544–552. doi: 10.1097/QCO.0b013e328332bbfe

Steer, A. C., Law, I., Matatolu, L., Beall, B. W., and Carapetis, J. R. (2009b). Global *emm* type distribution of group A Streptococci: systematic review and implications for vaccine development. *Lancet Infect. Dis.* 9, 611–616. doi: 10.1016/S1473-3099(09)70178-1

Stevens, A. J., Brown, Z. Z., Shah, N. H., Sekar, G., Cowburn, D., and Muir, T. W. (2016). Design of a Split Intein with Exceptional Protein Splicing Activity. *J. Am. Chem. Soc.* 138, 2162–2165. doi: 10.1021/jacs.5b13528

Suree, N., Liew, C. K., Villareal, V. A., Thieu, W., Fadeev, E. A., Clemens, J. J., et al. (2009). The structure of the *Staphylococcus aureus* sortase-substrate complex reveals how the universally conserved LPXTG sorting signal is recognized. *J. Biol. Chem.* 284, 24465–24477. doi: 10.1074/jbc.M109.022624

Takizawa, K., Akiyama, S., and Miyamoto, Y. (1970). Reexamination and characterization of the T-agglutination complex or pattern of *Streptococcus pyogenes*: preparation of anti-T factor sera. *Jpn. J. Microbiol.* 14, 269–277. doi: 10.1111/j.1348-0421.1970.tb00524.x

Talay, S. R., Valentin-Weigand, P., Timmis, K. N., and Chhatwal, G. S. (1994). Domain structure and conserved epitopes of Sfb protein, the fibronectin-binding adhesin of *Streptococcus pyogenes*. *Mol. Microbiol.* 13, 531–539. doi: 10.1111/j.1365-2958.1994.tb00448.x

Tan, L. L., Hoon, S. S., and Wong, F. T. (2016). Kinetic Controlled Tag-Catcher Interactions for Directed Covalent Protein Assembly. *PLoS One.* 11:e0165074. doi: 10.1371/journal.pone.0165074

Terao, Y., Kawabata, S., Nakata, M., Nakagawa, I., and Hamada, S. (2002). Molecular characterization of a novel fibronectin-binding protein of *Streptococcus pyogenes* strains isolated from toxic shock-like syndrome patients. *J. Biol. Chem.* 277, 47428–47435. doi: 10.1074/jbc.M209133200

Tesorero, R. A., Yu, N., Wright, J. O., Svencionis, J. P., Cheng, Q., Kim, J. H., et al. (2013). Novel regulatory small RNAs in *Streptococcus pyogenes*. *PLoS One.* 8:e64021. doi: 10.1371/journal.pone.0064021

Ton-That, H., and Schneewind, O. (2003). Assembly of pili on the surface of *Corynebacterium diphtheriae*. *Mol. Microbiol.* 50, 1429–1438. doi: 10.1016/j.tim.2004.03.004

Torretta, S., Marchisio, P., Drago, L., Baggi, E., De Vecchi, E., Garavello, W., et al. (2012). Nasopharyngeal biofilm-producing otopathogens in children with nonsevere recurrent acute otitis media. *Otolaryngol. Head Neck Surg.* 146, 991–996. doi: 10.1177/0194599812438169

Tsai, J. C., Loh, J. M., Clow, F., Lorenz, N., and Proft, T. (2017). The group A *Streptococcus* serotype M2 pilus plays a role in host cell adhesion and immune evasion. *Mol. Microbiol.* 103, 282–298. doi: 10.1111/mmi.13556

Veggiani, G., Nakamura, T., Brenner, M. D., Gayet, R. V., Yan, J., Robinson, C. V., et al. (2016). Programmable polyproteams built using twin peptide superglues. *Proc. Natl. Acad. Sci. U.S.A.* 113, 1202–1207. doi: 10.1073/pnas.1519214113

Wagachchi, D., Tsai, J. C., Chalmers, C., Blanchett, S., Loh, J. M. S., and Proft, T. (2018). PilVax - a novel peptide delivery platform for the development of mucosal vaccines. *Sci. Rep.* 8, 2555. doi: 10.1038/s41598-018-20863-7

Walden, M., Crow, A., Nelson, M. D., and Banfield, M. J. (2014). Intramolecular isopeptide but not internal thioester bonds confer proteolytic and significant thermal stability to the *S. pyogenes* pilus adhesin Spy0125. *Proteins* 82, 517–527. doi: 10.1002/prot.24420

Walden, M., Edwards, J. M., Dziewulska, A. M., Bergmann, R., Saalbach, G., Kan, S. Y., et al. (2015). An internal thioester in a pathogen surface protein mediates covalent host binding. *eLife* 4, e06638. doi: 10.7554/eLife.06638

Walker, M. J., Barnett, T. C., McArthur, J. D., Cole, J. N., Gillen, C. M., Henningham, A., et al. (2014). Disease manifestations and pathogenic mechanisms of group A *Streptococcus*. *Clin. Microbiol. Rev.* 27, 264–301. doi: 10.1128/CMR.00101-13

Weiner, E. M., Robson, S., Marohn, M., and Clubb, R. T. (2010). The Sortase A enzyme that attaches proteins to the cell wall of *Bacillus anthracis* contains an unusual active site architecture. *J. Biol. Chem.* 285, 23433–23443. doi: 10.1074/jbc.M110.135434

Yanagawa, R., and Honda, E. (1976). Presence of pili in species of human and animal parasites and pathogens of the genus *Corynebacterium*. *Infect. Immun.* 13, 1293–1295. doi: 10.1128/IAI.13.4.1293-1295.1976

Young, P. G., and Baker, E. N. (2020). Engineering of group A *Streptococcus* isopeptide bonds into immunoglobulin-like protein domains. *Methods Mol. Biol.* 2136, 377–395. doi: 10.1007/978-1-0716-0467-0_30

Young, P. G., Kang, H. J., and Baker, E. N. (2013). An arm-swapped dimer of the *Streptococcus pyogenes* pilin specific assembly factor SipA. *J. Struct. Biol.* 183, 99–104. doi: 10.1016/j.jsb.2013.05.021

Young, P. G., Moreland, N. J., Loh, J. M., Bell, A., Atatoa Carr, P., Proft, T., et al. (2014a). Structural conservation, variability, and immunogenicity of the T6 backbone pilin of serotype M6 *Streptococcus pyogenes*. *Infect. Immun.* 82, 2949–2957. doi: 10.1128/IAI.01706-14

Young, P. G., Proft, T., Harris, P. W., Brimble, M. A., and Baker, E. N. (2014b). Structure and activity of *Streptococcus pyogenes* SipA: a signal peptidase-like protein essential for pilus polymerisation. *PLoS One* 9:e99135. doi: 10.1371/journal.pone.0099135

Young, P. G., Raynes, J. M., Loh, J. M., Proft, T., Baker, E. N., and Moreland, N. J. (2019). group A *Streptococcus* T antigens have a highly conserved structure concealed under a heterogeneous surface that has implications for vaccine design. *Infect. Immun.* 87, e00205-19. doi: 10.1128/IAI.00205-19

Zähner, D., Gandhi, A. R., Yi, H., and Stephens, D. S. (2011). Mitis group streptococci express variable pilus islet 2 pili. *PLoS One* 6:e25124. doi: 10.1371/journal.pone.0025124

Zähner, D., and Scott, J. R. (2008). SipA is required for pilus formation in *Streptococcus pyogenes* serotype M3. *J. Bacteriol.* 190, 527–535. doi: 10.1128/JB.01520-07

Zakeri, B., Fierer, J. O., Celik, E., Chittock, E. C., Schwarz-Linek, U., Moy, V. T., et al. (2012). Peptide tag forming a rapid covalent bond to a protein, through engineering a bacterial adhesin. *Proc. Natl. Acad. Sci. U.S.A.* 109, E690–E697. doi: 10.1073/pnas.1115485109

Zakeri, B., and Howarth, M. (2010). Spontaneous intermolecular amide bond formation between side chains for irreversible peptide targeting. *J. Am. Chem. Soc.* 132, 4526–4527. doi: 10.1021/ja910795a

Zhang, R., Wu, R., Joachimiak, G., Mazmanian, S. K., Missiakas, D. M., Gornicki, P., et al. (2004). Structures of sortase B from *Staphylococcus aureus* and *Bacillus anthracis* reveal catalytic amino acid triad in the active site. *Structure* 12, 1147–1156. doi: 10.1016/j.str.2004.06.001

Zong, Y., Mazmanian, S. K., Schneewind, O., and Narayana, S. V. (2004). The structure of sortase B, a cysteine transpeptidase that tethers surface protein to the *Staphylococcus aureus* cell wall. *Structure* 12, 105–112. doi: 10.1016/j.str.2003.11.021

Zong, Y., Xu, Y., Liang, X., Keene, D. R., Höök, A., Gurusiddappa, S., et al. (2005). A 'Collagen Hug' model for *Staphylococcus aureus* CNA binding to collagen. *EMBO J.* 24, 4224–4236. doi: 10.1038/sj.emboj.7600888

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Microbiology

MINI REVIEW
published: 06 April 2021
doi: 10.3389/fmicb.2021.663468

S-layers: The Proteinaceous Multifunctional Armors of Gram-Positive Pathogens

Janani Ravi[1]* *and Antonella Fioravanti*[2,3]*

[1] *Pathobiology and Diagnostic Investigation, Michigan State University, East Lansing, MI, United States,* [2] *Structural and Molecular Microbiology, VIB-VUB Center for Structural Biology, Brussels, Belgium,* [3] *Structural Biology Brussels, Vrije Universiteit Brussel, Brussels, Belgium*

S-layers are self-assembled crystalline 2D lattices enclosing the cell envelopes of several bacteria and archaea. Despite their abundance, the landscape of S-layer structure and function remains a land of wonder. By virtue of their location, bacterial S-layers have been hypothesized to add structural stability to the cell envelope. In addition, S-layers are implicated in mediating cell-environment and cell-host interactions playing a key role in adhesion, cell growth, and division. Significant strides in the understanding of these bacterial cell envelope components were made possible by recent studies that have provided structural and functional insights on the critical S-layer and S-layer-associated proteins (SLPs and SLAPs), highlighting their roles in pathogenicity and their potential as therapeutic or vaccine targets. In this mini-review, we revisit the sequence-structure-function relationships of S-layers, SLPs, and SLAPs in Gram-positive pathogens, focusing on the best-studied classes, Bacilli (*Bacillus anthracis*) and Clostridia (*Clostridioides difficile*). We delineate the domains and their architectures in archetypal S-layer proteins across Gram-positive genera and reconcile them with experimental findings. Similarly, we highlight a few key "flavors" of SLPs displayed by Gram-positive pathogens to assemble and support the bacterial S-layers. Together, these findings indicate that S-layers are excellent candidates for translational research (developing diagnostics, antibacterial therapeutics, and vaccines) since they display the three crucial characteristics: accessible location at the cell surface, abundance, and unique lineage-specific signatures.

Keywords: gram-positive bacteria, Firmicutes, cell envelope, S-layer, cell surface proteins, pathogenicity, sequence-structure features, molecular evolution

OPEN ACCESS

Edited by:
Magnus Hook, Texas A&M Health Science Center, United States

Reviewed by:
Felix Broecker, Icahn School of Medicine at Mount Sinai, United States
Maria de los Angeles Serradell, Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Argentina

***Correspondence:**
Janani Ravi
janani@msu.edu
Antonella Fioravanti
antonella.fioravanti@vub.be

Specialty section:
This article was submitted to Microbial Physiology and Metabolism, a section of the journal Frontiers in Microbiology

Received: *04 February 2021*
Accepted: *08 March 2021*
Published: *06 April 2021*

Citation:
Ravi J and Fioravanti A (2021) S-layers: The Proteinaceous Multifunctional Armors of Gram-Positive Pathogens. Front. Microbiol. 12:663468. doi: 10.3389/fmicb.2021.663468

INTRODUCTION

Prokaryotes have evolved sophisticated and multi-layered cell envelopes to protect them while allowing selective cell-environment trafficking of nutrients, metabolites, integration of signals, and release of effectors. Despite the enormous diversity observed among prokaryotes and their environmental niches, the most commonly observed outermost cell envelope component is the surface layer (S-layer; Sleytr and Beveridge, 1999; Sara and Sleytr, 2000; Albers and Meyer, 2011; Fagan and Fairweather, 2014; Rodrigues-Oliveira et al., 2017). S-layers are semipermeable (glyco-) protein monolayers formed by S-layer proteins (SLPs) that once released at the cell surface self-assemble into a paracrystalline 2D lattice with defined symmetry that is anchored at the cell

envelope. S-layers demand a high metabolic investment from the organism producing them; they comprise 5–15% of the total cellular protein production, making them amongst the most abundant proteins on Earth (Sara and Sleytr, 2000).

Since their first observation in the 1950s, the biological and biotechnological relevance of S-layers has been of great interest to the scientific community (Sleytr et al., 2014). Nevertheless, half a century later, even though S-layers are found nearly in all major bacterial clades and represent an almost universal feature of the archaeal cell envelope, our knowledge about their structure and function remains patchy. Multiple factors have contributed to this lack of knowledge: (i) the absence of S-layers in classical model organisms such as *Escherichia coli* and *Bacillus subtilis*; (ii) the self-assembling characteristic of SLPs, that has long hampered structural and biochemical studies; and (iii) their low sequence homology, making it challenging to identify S-layer-carrying organisms based on their sequence alone. Function-wise, while the SLPs in archaea are known to maintain cell shape, often as the sole cell-wall components (Albers and Meyer, 2011; Rodrigues-Oliveira et al., 2017), bacterial S-layers are known to carry out multiple functions ranging from adhesive surface to protective and selective barriers (Gerbino et al., 2015).

Several Gram-positive and negative pathogens possess S-layers that play potentially significant roles in their virulence (Carl and Dasch, 1989; Kawai et al., 1998; Mignot et al., 2001; Thompson, 2002; Shimotahira et al., 2013; Rasmussen-Ivey et al., 2016). In this mini-review, we provide a comprehensive overview of the current understanding of S-layer structure, function, and contribution to the pathogenicity of Gram-positive bacteria, focusing on the best-characterized S-layer-carrying human pathogens: *Bacillus anthracis* and *Clostridioides difficile*. We discuss recent breakthroughs on the S-layer structure and function of these two pathogens that emphasize the role of SLPs as promising antimicrobial targets (Kirk et al., 2017; Fioravanti et al., 2019; Oatley et al., 2020; Banerji et al., 2021). Along with highlighting variations seen across Firmicutes and Actinobacteria, this review provides a foundation and context for future studies to fully exploit the potential of SLPs as targets for the development of novel diagnostics, vaccines, and antibacterial therapies.

TWO ARMORS ARE BETTER THAN ONE? THE *B. anthracis* S-LAYER CASE

Bacillus anthracis is the etiological agent of anthrax (Kamal et al., 2011; Goel, 2015; Okinaka and Keim, 2016) and a CDC Category A bioterrorist agent [Centers for Disease Control and Prevention (CDC), 2018]. As part of its immune evasion strategy, this sporulating bacterium, displays a complex and dynamic cell envelope composition (Chateau et al., 2020) that includes switchable S-layers (Mignot et al., 2002).

The Two Switchable S-layers, Sap and EA1

On top of a thick peptidoglycan layer (PG), the bacterium cell surface is enveloped by one of two mutually exclusive S-layers, the Sap or EA1 S-layer, present, respectively, during exponential and stationary growth phase of cells grown in rich medium (Mignot et al., 2002, 2003; Fioravanti et al., 2019). Electron microscopy (EM) observation of single S-layer mutants revealed a clear difference between the 2D arrays: the Sap S-layer forms a continuous array, whereas the EA1 S-layer is organized in patches (Couture-Tosi et al., 2002). The SLPs developmental switch is controlled by growth-phase-specific sigma factors and the two SLPs (Mignot et al., 2002), which contain *C*-terminal domains with DNA-binding activity that independently repress the *eag* promoter (Couture-Tosi et al., 2002; Mignot et al., 2002). To ensure high expression, bacterial SLPs are associated with strong promoters, efficient transcription, and mRNAs with increased stability (∼6–10x in *B. anthracis* SLPs compared to the average bacterial mRNA half-life; Glatron and Rapoport, 1972; Fisher et al., 1988; Emory et al., 1992; Mignot et al., 2002). The S-layer switch also occurs during systemic infection as both proteins are immunogenic during human anthrax infection (Baillie et al., 2003). It is still unclear why *B. anthracis* performs this energetically expensive S-layer remodeling during its life cycle and infection, emphasizing the need to understand environmental and host triggers that induce this switch.

Sequence-Structure Features

The two SLPs contain similar domain architectures that include: an *N*-terminal signal peptide for secretion, an S-layer homology (SLH) domain for cell anchoring, and a *C*-terminal assembly domain (AD) that self-assembles into the S-layer (**Figure 1A**; Mesnage et al., 1999; Candela et al., 2005; Wang et al., 2015). Given their abundance, both SLPs are secreted by an accessory and dedicated secretion system (Nguyen-Mau et al., 2012). Once released at the cell surface, they spontaneously fold and anchor at the cell wall through non-covalent interactions between the SLH domain and the pyruvylated secondary cell wall polysaccharides (SCWP) bound to the PG (Mesnage et al., 2000; Kern et al., 2010; Missiakas and Schneewind, 2017). SCWP is essential for cell growth and division and plays a critical role in bacterial pathogenicity (Oh et al., 2016; Chateau et al., 2018). The SLH-SCWP interaction is considered an ancestral mechanism for SLP anchoring to the cell envelope (Cava et al., 2004), and in Gram-positive bacteria, it is recurrent in cell-wall-anchored proteins (**Figure 2**). *B. anthracis* additionally encodes 22 S-layer-associated proteins (SLAPs, called BSLs in Bacilli) that harbor SLH domains (Kern and Schneewind, 2008). Unlike SLPs, BSLs are minor components of the envelope that do not form paracrystalline arrays but exploit several enzymatic functions participating in different cellular processes [e.g., peptidoglycan metabolism, host adhesion (Kern and Schneewind, 2008; Tarlovsky et al., 2010; Kern et al., 2012)]. Structure determination of the Sap SLH domain revealed that it comprises three SLH motifs that fold in a pseudo-trimer and that conserved positively charged residues sustain its interaction with the terminal PG-anchored pyruvylated-SCWP unit (Blackler et al., 2018; Sychantha et al., 2018). While the Sap and EA1 SLH domains are similar (74%), their AD are divergent (42% similar; 22% identity; Chateau et al., 2020).

In a recent study, we overcame hurdles concerning SLP self-polymerization and monomer stability. Using anti-Sap

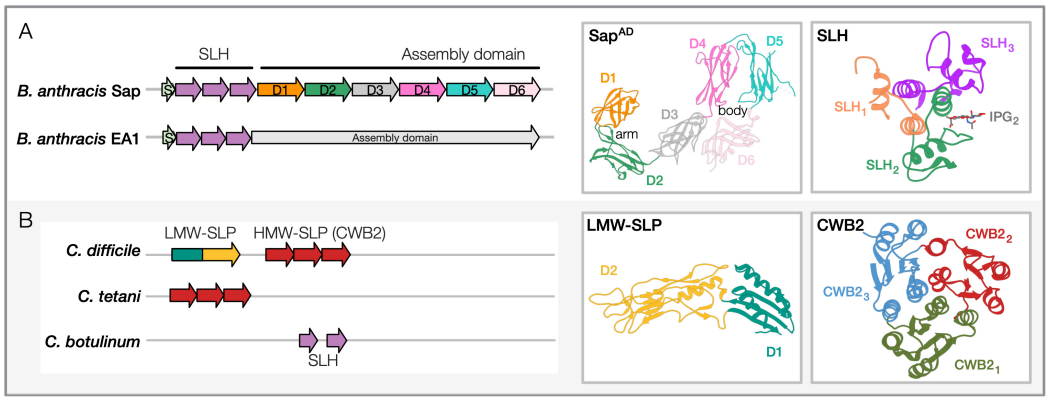


FIGURE 1 | Sequence-structure features of pathogenic Bacilli and Clostridia SLPs. **(A)** *Bacillus anthracis* SLPs, Sap and EA1. Domain architectures: SLH region (containing three SLH domain motifs, Pfam PF00395, ProSiteProfiles PS51272) and the assembly domain are shown for both the representative Bacillus SLPs. PDB structures: Shown for Sap^{AD} and Sap-SLH are shown and labeled [PDB: 6HHU, 6BT4 (Kern et al., 2011; Fioravanti et al., 2019)]. The crystal structure of the SLH domain is in complex with a synthetic SCWP unit (IPG). Accession numbers for proteins shown: Sap, AAT53168.1; EA1, AAP24884.1. Similar domain architectures are observed in *B. cereus*, *B. mycoides*, and *B. thuringiensis* SLPs. **(B)** Clostridia SLPs in *Clostridioides difficile*, *Clostridium tetani*, and *Clostridium botulinum*. Domain architectures: *C. difficile* contains low-molecular-weight (LMW) and high-molecular-weight (HMW) SLP domains (HMW-SLP with three CWB2 domain motifs; LMW-SLP). *C. tetani* contains the cell-wall binding domain (CWB2, Pfam PF04122) as well, while *C. botulinum* carries two SLH domains. PDB structures: Shown for the LMW-SLP and CWB2 of Cwp8 are shown [PDB: 3CVZ, 5J6Q; (Fagan et al., 2009; Usenik et al., 2017)]. Accession numbers for proteins shown: *C. difficile*, WP_078051019.1; *C. tetani*, WP_035111087.1; *C. botulinum*, WP_039307708.1. Domain architectures are marked from cited literature and InterProScan predictions (Jones et al., 2014). Representative PDB structures for Bacillus and Clostridia SLP domains have been redrawn using Phyre2 (Kelley et al., 2015).

nanobodies (Nbs) as crystallization-aid (Muyldermans, 2013), we determined the first complete SLP AD structure of a pathogen, the Sap^{AD} (Fioravanti et al., 2019). Sap represents a novel class of SLPs that folds and assembles in a calcium-independent manner. Sap^{AD} folds into an extensive multi-domain protein consisting of six β-sandwich domains connected by short linkers. In solution, it adopts a flat tile-like supertertiary structure consisting of an "arm" (D1–2) and "body" (D3–6; **Figure 1A**). Interestingly, our recent comparison of the *B. anthracis* Sap^{AD} with the AD from SbsB from *Geobacillus stearothermophilus* revealed that the "arm-and-body" modular architecture is conserved across Bacillales (Fioravanti et al., 2019; **Figure 1A**). This architectural conservation is remarkable given the low average pairwise sequence identity (∼25%) and high variability in the domain ultrastructure. Moreover, the different proteins assemble into unrelated lattices and differ in their need for divalent metal ions for folding and S-layer assembly (Fioravanti et al., 2019). Further structural and functional insights are needed to understand better what governs the selective pressure(s) to maintain this "arm-and-body" architecture, despite the lack of structural conservation within and across domains and monomers within the S-layers.

2D Lattice

Electron microscopy has enabled an in-depth investigation of the lattice topology of native S-layer fragments and *in vitro* recrystallized S-layers (Couture-Tosi et al., 2002; Wang et al., 2015; Fioravanti et al., 2019). EM analysis on native S-layers resulted in low-resolution projection maps for both SLPs, with the density distribution hinting at six/seven domains for Sap and four for EA1 (Couture-Tosi et al., 2002). While recent EM studies on *in vitro* recrystallized S-layers show increased details about the Sap lattice (Fioravanti et al., 2019), the combined literature still does not reveal interdomain contacts responsible for Sap and EA1 S-layer assembly. In addition to the EA1 atomic structure, *in vitro* and on-cell cryo-EM/tomography studies leading to higher resolution density maps will be required to unveil details of the peculiar SLP switching mechanism, the *raison d'être* of the two S-layers, and avenues for therapeutic targeting of *B. anthracis* SLPs.

Sap and EA1 as Vaccines or Antimicrobial Targets

At the outset, the contribution of the S-layer to *B. anthracis* virulence remains unclear. Deletion strains of either SLPs are viable *in vitro* but have never been tested under infection. Interestingly, the *sap* deletion mutant showed cell division defects due to a displacement of BslO, a SLAP that catalyses mother-daughter cell separation, which requires Sap S-layer for correct deposition at nascent cell division sites (Kern et al., 2012). In early 2000, Sap and EA1 were shown to be immunogenic during human infection (Baillie et al., 2003), nominating SLPs as potential vaccine candidates against anthrax. Subsequent studies have shown that immunization using EA1, but not Sap, offers a protective effect in a mouse model of inhalational anthrax (Uchida et al., 2012; Fioravanti et al., 2019).

Recently, Fioravanti et al. (2019) established a direct link between the *B. anthracis* S-layer integrity and its potential as an antimicrobial target. Anti-Sap-Nbs were shown to depolymerize the Sap S-layer *in vitro* and *in vivo*. *In vivo*, the Nbs-mediated disruption of the Sap S-layer resulted in severe morphological defects (wrinkled phenotype) and attenuated growth. The Nbs-induced phenotype was more striking than the *sap* knockout,

suggesting that cells undergoing an acute loss of S-layer cannot adapt by switching to an EA1 S-layer to rescue such defects. These data point to a more critical contribution of S-layers in cell shape maintenance. Moreover, subcutaneous delivery of Sap-inhibitory Nbs cleared *B. anthracis* infection and prevented lethality in a mouse model of anthrax (Fioravanti et al., 2019). Together, these findings represent the first evidence that the disruption of S-layer integrity is a mechanism with therapeutic potential in S-layer-carrying pathogens.

Similarly, in *B. cereus* G9241, the causative agent of anthrax-like disease, mutants incapable of retaining Sap, EA1, and BSLs in the bacterial envelope showed reduced virulence in mice (Wang et al., 2013). Moreover, studies on S-layer distribution among the *B. cereus* group (containing *B. anthracis*) have observed SLPs in all clinical strains but only sporadically in environmental strains, suggesting a correlation between virulence and the presence of an S-layer (Mignot et al., 2001).

THE TWO-TIERED ARMOR: *C. difficile* S-LAYER

Clostridioides difficile is an obligate anaerobic, spore-forming bacterium involved in a broad spectrum of diseases: from mild post-antibiotic diarrhea to severe pseudomembranous colitis, resulting in severe healthcare burden (Rupnik et al., 2009). CDC has designated *C. difficile* as the pre-eminent of five "Urgent Threats" to US healthcare, emphasizing its increasing antibiotic resistance [Centers for Disease Control and Prevention (CDC), 2019]. The *C. difficile* S-layer is shown to play a crucial role in the intestinal colonization step during infection (Calabi et al., 2002), in sporulation, toxin production, and resistance to components of the innate immune system (Kirk et al., 2017), representing an ideal candidate for the development of new therapeutics.

S-layer Composition

The *C. difficile* S-layer represents a rare case where the 2D crystal is made by the assembly of heterodimers (Calabi et al., 2001). The *slpA* gene encodes for a common precursor (Karjalainen et al., 2001), which upon signal peptide removal and cell secretion, undergoes a second cleavage by the cysteine protease, Cwp84 (Kirby et al., 2009), releasing the high-molecular-weight (HMW) and the low-molecular-weight (LMW) SLPs. Together, they form a tightly-associated non-covalent H/L complex that anchors at the cell surface and assembles into the S-layer (Fagan et al., 2009; **Figure 1B**). A recent microscopy-based study revealed novel insights on subcellular SlpA secretion and S-layer growth. While S-layer growth occurs at specific sites that coincide with cell wall synthesis, the SLPs are secreted all over the cytoplasmic membrane, suggesting that there is a reservoir of SLPs within the cell wall ready to be utilized for S-layer growth (Oatley et al., 2020).

Domain Organization and Structure

The SlpA precursor comprises an *N*-terminal signal peptide, the LMW-SLP, and the *C*-terminal HMW-SLP (Calabi et al., 2001; **Figure 1B**). The HMW-SLP is anchored to SCWP anionic polymer PSII by the cell wall binding domain 2 (CWB2; **Figure 1B**), while the LMW-SLP is presented as the outermost component of the *C. difficile* surface, showing a high degree of antigenic variation between strains (Calabi et al., 2001; Willing et al., 2015). CWB2 comprises three tandem motifs (Willing et al., 2015) as seen for SLH (Kern et al., 2011; **Figure 1B**). Despite being similar in sequence, the CWB2 motifs are not redundant; it takes three motifs to ensure the S-layer anchoring to the cell wall (Willing et al., 2015). *C. difficile* encodes an additional 28 CWB2 carrying SLAPs, called the clostridia cell wall proteins (CWPs; Fagan et al., 2011). As with Bacillus *BSLs,* CWPs do not form the S-layer but exploit a variety of enzymatic and host-pathogen interaction functions (Kirby et al., 2009; Bradshaw et al., 2017). Recent structure determination of Cwp8 unveiled the CWB2 domain fold (Usenik et al., 2017; **Figure 1B**). Each CBW2 motif assumes a topoisomerase-primase fold, and together they assemble in a trefoil-like shape (**Figure 1B**). EM studies revealed the presence of a two-tiered S-layer at the cell surface (Cerquetti et al., 2000). Determining the atomic structure of the SlpA heterodimer or in its S-layer form has proven challenging. The crystal structure of an LMW-SLP truncated version was determined (**Figure 1B**); the missing 59 *C*-terminal residues were reported as necessary for heterodimer formation (Fagan et al., 2009). The LMW-SLP assumes a novel fold comprising two domains: D1 contains both the *N*- and *C*-termini of the protein that fold into a sandwiched conformation; D2, likely exposed at the cell surface, presents a novel fold with a high loop content. The loops allow a high-level of sequence variability that promotes host immune system evasion while retaining the overall SLP fold (Fagan et al., 2009; Spigaglia et al., 2011; Merrigan et al., 2013). Small-angle X-ray scattering was used to study the H/L complex. In solution, the two SLPs are arranged in an "end-to-end" complex with presumably the *C*-terminus of LMW-SLP and *N*-terminus of HMW-SLP interacting with each other. A recent preprint describes the SlpA S-layer organization in atomic detail (PDB: 7ACY; Banerji et al., 2021). In this structure, the LMW/HMW SLP-interacting domains are described to fold into a "paper-clip" arrangement, while the three CWB2 motifs of the HMW subunit are organized in a triangular prism. Moreover, the crystallographic structure of the H/L heterodimer could be docked in the EM projection maps obtained on native SlpA S-layer, unveiling important intramolecular interfaces essential for S-layer formation. This work represents a significant advancement for the S-layer and *C. difficile* communities, offering a plethora of possibilities for the design of S-layer-structure-tailored antimicrobials (Banerji et al., 2021).

SlpA and Virulence

SlpA is required for gastrointestinal tissue adherence and is implicated in pathogenicity (Calabi et al., 2002; Merrigan et al., 2013). SLP mutants have been impossible to obtain, suggesting the essentiality of the *slpA* gene. Instead, two rare resistant mutants to diffocin, a bacteriocin that selectively kills *C. difficile* strains, displayed an SLP-null phenotype that presents severe sporulation defects and a significant increase in bacterial

susceptibility to lysozyme and the antimicrobial peptide, LL-37 (Kirk et al., 2017). Interestingly, these mutants are capable of colonizing the intestinal tract of hamsters despite a complete attenuation of virulence. SLPs are also found in several other Clostridia species, including *Clostridium botulinum* (Takumi et al., 1992) and *Clostridium tetani* (Sleytr and Messner, 1983; Takumi et al., 1991). Further characterization and comparative studies are needed to delineate the SLP biology in other Clostridia pathogens.

OTHER GRAM-POSITIVE "FLAVORS" OF SLPS

In addition to the well-characterized Bacilli and Clostridia SLPs, several pathogenic members within Firmicutes (e.g., Paenibacilli, Lactobacilli, and Listeria), as well as Actinobacterial species (e.g., Corynebacteria, Mycobacteria), are known to form S-layers linked to their virulence and pathogenicity (Sleytr and Messner, 1983; Fagan and Fairweather, 2014). In this section, we highlight a few Gram-positive variations, with and without SLH/CWB2 domains (**Figure 2**).

Paenibacilli

Many Paenibacilli, including the etiological agent of the epizootic of honeybees *P. larvae*, possess a functionally proven virulent S-layer made of SlpA (Poppinga et al., 2012). *P. alvei* cells present an S-layer comprised of glycosylated SLP, SpaA. In these SLP homologs, the *N*-terminal SLH domain has dual recognition for SCWP and PG, and is sufficient for *in vivo* cell surface display of foreign proteins at the cell surface (Janesch et al., 2013; **Figure 2**). The SLH domain trimer rearrangement also relieves any S-layer strain caused by cell growth and division (Blackler et al., 2018). Notably, a second SLH-containing protein in *P. alvei*, SlhA, is found to be vital for swarming and biofilm formation (Janesch et al., 2013).

Other Pathogenic Firmicutes

Other notable Firmicutes such as Streptococcus, Staphylococcus, and Listeria have been predicted to carry SLPs (**Figure 2**) that have been indirectly linked to pathogenicity (Navarre and Schneewind, 1999; Camejo et al., 2009; He et al., 2019). For instance, studies involving *L. monocytogenes* virulence factors suggest a role for S-layer glycoproteins in Listeria

Domains / Genera	SLH	SlpA	LMW-SLP	CWB2	Bac-Ig-like	Extracellular	Outer-MP	Other unique domains
Bacilli	✓				✓	✓	✓	
Clostridia	✓		✓	✓		✓		NAM_LA_amidase
Paenibacilli	✓				✓	✓		Lyase, Fibronectin
Listeria						✓		Endotoxin, BLR0621
Streptococci	✓					✓	✓	Glucosaminidase, Mem_lipoprot_site
Staphylococci	✓						✓	Glutamine aminotransferase
Corynebacteria						✓		
Mycobacteria	✓						✓	Haemagluttinin repeat, Protein_CBG03580

FIGURE 2 | Sequence features of SLPs in key Gram-positive pathogenic genera. Previously documented or predicted [using InterProScan (Jones et al., 2014)] sequence-structure features in representative SLPs from select pathogens in Firmicutes and Actinobacteria. Species considered: *B. anthracis*, *B. cereus*, *B. mycoides*, *B. thuringiensis*, *Clostrioides difficile*, *Clostridium tetani*, *Clostridium botulinum*, *P. alvei*, *Listeria monocytogenes*, *L. seeligeri*, *L. booriae*, *L. fleischmannii*, *Streptococcus pneumoniae*, *S. dysgalactiae*, *S. pyogenes*, *Staphylococcus cohnii*, *S. haemolyticus*, *Corynebacterium glutamicum*, *C. aurimucosum*, *C. minutissimum*, *M. intracellulare*, and *M. kansasii*. Abbreviations: *SLH,* S-layer homology domain (Pfam PF00395, ProSiteProfiles PS51272); *LMW-SLP,* Low molecular weight S layer protein *N*-terminal (Pfam PF12211); *CWB2,* Cell wall binding domain 2 (High molecular weight; Pfam PF04122); Extracellular, Region of a membrane-bound protein predicted to be outside the membrane in the extracellular region (Phobius prediction); Outer-MP, Outer membrane protein alpha-related (PANTHER PTHR43308); *Bac-Ig-like,* Bacterial Ig-like domain (Pfam clan CL0159: PF02368, PF13205); Lyase, Hyaluronate lyase/Polysaccharide lyase family 8 (Pfam PF02278, PF02884, PF08124, PANTHER PTHR38481); Fibronectin, Fibronectin type-III domain (Pfam PS50853); *NAM_LA_amidase,* N-acetylmuramoyl-L-alanine amidase-related domain (PANTHER PTHR30032); *Glucosaminidase,* Mannosyl-glycoprotein endo-beta-N-acetylglucosaminidase, Transglutaminase-like superfamily (Pfam PF01832, PF01841); *Mem_lipoprot_site,* Prokaryotic membrane lipoprotein lipid attachment (ProSiteProfiles PS51257); *Glutamine aminotransferase,* Glutamine amidotransferase type 2; Glucosamine-fructose-6-phosphate aminotransferase, isomerizing (Pfam PF13522, ProSiteProfiles: PS51278, PANTHER PTHR10937); and Endotoxin, Delta endotoxin (Pfam PF18449).

virulence (Camejo et al., 2009). Further structural and functional characterizations remain to be performed.

Non-Pathogenic Lactobacillales

One of the rarer symbiotic functional contexts in which S-layers have been reported is in the Lactobacilli and Enterococcus species that adhere to intestinal epithelial cells. Several of these species contain non-glycosylated SLPs with an SlpA domain for cell anchoring instead of the typical SLH domain. While few Lactobacilli species carry multiple copies of SLPs (Åvall-Jääskeläinen et al., 2008; He et al., 2019), probiotic strains carry SLAPs that contain collagen- and fibronectin-binding domains, which are useful to adhere to the extracellular matrix of the intestinal epithelial cells (Bahl et al., 1997; Hymes et al., 2016) and contribute to pathogen exclusion (Martínez et al., 2012). In other species such as *L. crispatus*, SlpB interacts with the bacterial cell wall, and its collagen-binding activity is thought to aid in antigenic variation in adherence (Bahl et al., 1997) and gut colonization (Sun et al., 2017).

Corynebacteria, Mycobacteria

Even actinobacterial species with atypical outer membrane-like structures (known as Mycomembrane) carry SLPs. For instance, few strains of *C. glutamicum* contain a hexagonal S-layer made of PS2 (Bahl et al., 1997; Chami et al., 1997; Houssin et al., 2002; Burkovski, 2013). While PS2's *N*-terminus is responsible for monomer interactions, its *C*-terminal region, especially with a hydrophobic stretch, is needed for cell wall anchoring (Bahl et al., 1997; Bayan et al., 2003). In contrast, S-layers with an oblique arrangement have been reported in Mycobacterial species such as *M. bovis* (Lounatmaa and Brander, 1989). The highly immunogenic nature of a few mycobacterial CWPs suggests that the cell wall antigens are located in the S-layer (Lounatmaa and Brander, 1989).

CONCLUSION

S-layers are paracrystalline protein arrays that are among the most commonly observed cell envelope components in prokaryotes. They are important for cell development, cell-environment, and cell-host interactions (Fagan and Fairweather, 2014). Bacterial SLPs exhibit considerable variation in their composition and structure, as evident from the low sequence similarities across SLP homologs (Sleytr and Messner, 1983; Bahl et al., 1997; Navarre and Schneewind, 1999; Sleytr and Beveridge, 1999; Fagan et al., 2009; Kufel et al., 2017). Typically, Gram-positive SLPs comprise an *N*-terminal signal peptide, a cell wall anchoring domain, and an AD that self-polymerizes into the S-layer (**Figures 1**, **2**). In this mini-review, we discuss recent breakthroughs in S-layer structure and function in two prominent Gram-positive pathogen-containing classes, Bacilli and Clostridia (**Figure 1**; Kirk et al., 2017; Fioravanti et al., 2019; Oatley et al., 2020; Banerji et al., 2021) accentuating the critical role played by S-layer in cell envelope integrity and bacterial pathogenicity. We also highlight notable variations of S-layers in other Firmicutes and Actinobacteria, with the responsible proteins containing lineage-specific SLP (SLH, CWB2, SlpA, and LMW-SLP) and paralogous SLAP (e.g., amidase, fibronectin, adhesin, and endotoxin) domains (**Figure 2**). With no apparent sequence signature, finer evolutionary analyses on SLPs and SLAPs across bacterial and archaeal phyla are required to shed light on their evolutionary origin and function.

In summary, S-layers represent the frontline for host-pathogen interactions playing a significant role in virulence and modulating the host immune response. Because they are abundant pathogen-specific components, exposed on the cell surface, SLPs can be exploited as diagnostic, vaccine, and therapeutic targets. The advent of new experimental and computational technologies will open new avenues to further characterize the currently unresolved sequence-structure-function links in these extraordinary macromolecular scaffolds.

AUTHOR CONTRIBUTIONS

JR and AF contributed equally to the conceptualization and the writing of this manuscript. Both authors contributed to the article and approved the submitted version.

FUNDING

This research was supported by FWO Flanders through project grant number G065220N and the FWO Senior Postdoc Fellowship number 1253121N to AF and start-up funds from Michigan State University to JR.

ACKNOWLEDGMENTS

We are grateful to Sander Van Der Verren, Inge Van Molle, Arne Janssens, Adria Sogues Castrejon, Lauren Sosinski, and Dr. Arjun Krishnan for feedback on the manuscript.

REFERENCES

Albers, S. V., and Meyer, B. H. (2011). The archaeal cell envelope. *Nat. Rev. Microbiol.* 9, 414–426. doi: 10.1038/nrmicro2576

Åvall-Jääskeläinen, S., Hynönen, U., Ilk, N., Pum, D., Sleytr, U. B., and Palva, A. (2008). Identification and characterization of domains responsible for self-assembly and cell wall binding of the surface layer protein of *Lactobacillus brevis* ATCC 8287. *BMC Microbiol.* 8:165. doi: 10.1186/1471-2180-8-165

Bahl, H., Scholz, H., Bayan, N., Chami, M., Leblon, G., Gulik-Krzywicki, T., et al. (1997). IV. Molecular biology of S-layers. *FEMS Microbiol. Rev.* 20, 47–98. doi: 10.1016/s0168-6445(97)00050-8

Baillie, L., Hebdon, R., Flick-Smith, H., and Williamson, D. (2003). Characterisation of the immune response to the UK human anthrax vaccine. *FEMS Immunol. Med. Microbiol.* 36, 83–86. doi: 10.1016/s0928-8244(03)00085-3

Banerji, O., Lanzoni-Mangutchi, P., Vaz, F., Barwinska-Sendra, A., Wilson, J., Kirk, J., et al. (2021). Structure and assembly of the S-layer

determine virulence in *C. difficile*. *Res. Sq.* doi: 10.21203/rs.3.rs-79088/v1

Bayan, N., Houssin, C., Chami, M., and Leblon, G. (2003). Mycomembrane and S-layer: two important structures of *Corynebacterium glutamicum* cell envelope with promising biotechnology applications. *J. Biotechnol.* 104, 55–67. doi: 10.1016/s0168-1656(03)00163-9

Blackler, R. J., López-Guzmán, A., Hager, F. F., Janesch, B., Martinz, G., Gagnon, S. M. L., et al. (2018). Structural basis of cell wall anchoring by SLH domains in *Paenibacillus alvei*. *Nat. Commun.* 9:3120.

Bradshaw, W. J., Kirby, J. M., Roberts, A. K., Shone, C. C., and Acharya, K. R. (2017). Cwp2 from *Clostridium difficile* exhibits an extended three domain fold and cell adhesion in vitro. *FEBS J.* 284, 2886–2898. doi: 10.1111/febs.14157

Burkovski, A. (2013). Cell envelope of corynebacteria: structure and influence on pathogenicity. *ISRN Microbiol.* 2013:935736.

Calabi, E., Calabi, F., Phillips, A. D., and Fairweather, N. F. (2002). Binding of *Clostridium difficile* surface layer proteins to gastrointestinal tissues. *Infect. Immun.* 70, 5770–5778. doi: 10.1128/iai.70.10.5770-5778.2002

Calabi, E., Ward, S., Wren, B., Paxton, T., Panico, M., Morris, H., et al. (2001). Molecular characterization of the surface layer proteins from *Clostridium difficile*. *Mol. Microbiol.* 40, 1187–1199. doi: 10.1046/j.1365-2958.2001.02461.x

Camejo, A., Buchrieser, C., Couvé, E., Carvalho, F., Reis, O., Ferreira, P., et al. (2009). In vivo transcriptional profiling of *Listeria monocytogenes* and mutagenesis identify new virulence factors involved in infection. *PLoS Pathog.* 5:e1000449. doi: 10.1371/journal.ppat.1000449

Candela, T., Mignot, T., Hagnerelle, X., Haustant, M., and Fouet, A. (2005). Genetic analysis of *Bacillus anthracis* Sap S-layer protein crystallization domain. *Microbiology (Read. Engl.)* 151, 1485–1490. doi: 10.1099/mic.0.27832-0

Carl, M., and Dasch, G. A. (1989). The importance of the crystalline surface layer protein antigens of rickettsiae in T-cell immunity. *J. Autoimmun.* 2(Suppl.), 81–91. doi: 10.1016/b978-0-12-252682-4.50012-3

Cava, F., de Pedro, M. A., Schwarz, H., Henne, A., and Berenguer, J. (2004). Binding to pyruvylated compounds as an ancestral mechanism to anchor the outer envelope in primitive bacteria. *Mol. Microbiol.* 52, 677–690. doi: 10.1111/j.1365-2958.2004.04011.x

Centers for Disease Control and Prevention (CDC) (2018). *Bioterrorism Agents/Diseases*. Available online at: https://emergency.cdc.gov/agent/agentlist-category.asp (accessed January, 2021).

Centers for Disease Control and Prevention (CDC) (2019). *Antibiotic Resistance Urgent Threats Report 2019*. Available online at: https://www.cdc.gov/drugresistance/biggest-threats.html#cdiff (accessed January, 2021).

Cerquetti, M., Molinari, A., Sebastianelli, A., Diociaiuti, M., Petruzzelli, R., Capo, C., et al. (2000). Characterization of surface layer proteins from different *Clostridium difficile* clinical isolates. *Microb. Pathog.* 28, 363–372. doi: 10.1006/mpat.2000.0356

Chami, M., Bayan, N., Peyret, J. L., Gulik-Krzywicki, T., Leblon, G., and Shechter, E. (1997). The S-layer protein of *Corynebacterium glutamicum* is anchored to the cell wall by its C-terminal hydrophobic domain. *Mol. Microbiol.* 23, 483–492. doi: 10.1046/j.1365-2958.1997.d01-1868.x

Chateau, A., Lunderberg, J. M., Oh, S. Y., Abshire, T., Friedlander, A., Quinn, C. P., et al. (2018). Galactosylation of the secondary cell wall polysaccharide of *Bacillus anthracis* and its contribution to anthrax pathogenesis. *J. Bacteriol.* 200:e562-17.

Chateau, A., Van der Verren, S. E., Remaut, H., and Fioravanti, A. (2020). The *Bacillus anthracis* cell envelope: composition, physiological role, and clinical relevance. *Microorganisms* 8:1864. doi: 10.3390/microorganisms8121864

Couture-Tosi, E., Delacroix, H., Mignot, T., Mesnage, S., Chami, M., Fouet, A., et al. (2002). Structural analysis and evidence for dynamic emergence of *Bacillus anthracis* S-layer networks. *J. Bacteriol.* 184, 6448–6456. doi: 10.1128/jb.184.23.6448-6456.2002

Emory, S. A., Bouvet, P., and Belasco, J. G. (1992). A 5′-terminal stem-loop structure can stabilize mRNA in *Escherichia coli*. *Genes Dev.* 6, 135–148. doi: 10.1101/gad.6.1.135

Fagan, R. P., Albesa-Jove, D., Qazi, O., Svergun, D. I., Brown, K. A., and Fairweather, N. F. (2009). Structural insights into the molecular organization of the S-layer from Clostridium difficile. *Mol. Microbiol.* 71, 1308–1322.

Fagan, R. P., and Fairweather, N. F. (2014). Biogenesis and functions of bacterial S-layers. *Nat. Rev. Microbiol.* 12, 211–222. doi: 10.1038/nrmicro3213

Fagan, R. P., Janoir, C., Collignon, A., Mastrantonio, P., Poxton, I. R., and Fairweather, N. F. (2011). A proposed nomenclature for cell wall proteins of *Clostridium difficile*. *J. Med. Microbiol.* 60, 1225–1228. doi: 10.1099/jmm.0.028472-0

Fioravanti, A., Van Hauwermeiren, F., Van der Verren, S. E., Jonckheere, W., Goncalves, A., Pardon, E., et al. (2019). Structure of S-layer protein Sap reveals a mechanism for therapeutic intervention in anthrax. *Nat. Microbiol.* 4, 1805–1814. doi: 10.1038/s41564-019-0499-1

Fisher, J. A., Smit, J., and Agabian, N. (1988). Transcriptional analysis of the major surface array gene of *Caulobacter crescentus*. *J. Bacteriol.* 170, 4706–4713. doi: 10.1128/jb.170.10.4706-4713.1988

Gerbino, E., Carasi, P., Mobili, P., Serradell, M. A., and Gomez-Zavaglia, A. (2015). Role of S-layer proteins in bacteria. *World J. Microbiol. Biotechnol.* 31, 1877–1887. doi: 10.1007/s11274-015-1952-9

Glatron, M. F., and Rapoport, G. (1972). Biosynthesis of the parasporal inclusion of *Bacillus thuringiensis*: half-life of its corresponding messenger RNA. *Biochimie* 54, 1291–1301. doi: 10.1016/s0300-9084(72)80070-1

Goel, A. K. (2015). Anthrax: a disease of biowarfare and public health importance. *World J. Clin. Cases* 3, 20–33. doi: 10.12998/wjcc.v3.i1.20

He, Y., Xu, X., Zhang, F., Xu, D., Liu, Z., Tao, X., et al. (2019). Anti-adhesion of probiotic *Enterococcus faecium* WEFA23 against five pathogens and the beneficial effect of its S-layer proteins against *Listeria monocytogenes*. *Can. J. Microbiol.* 65, 175–184. doi: 10.1139/cjm-2018-0031

Houssin, C., Nguyen, D. T., Leblon, G., and Bayan, N. (2002). S-layer protein transport across the cell wall of *Corynebacterium glutamicum*: in vivo kinetics and energy requirements. *FEMS Microbiol. Lett.* 217, 71–79. doi: 10.1111/j.1574-6968.2002.tb11458.x

Hymes, J. P., Johnson, B. R., Barrangou, R., and Klaenhammer, T. R. (2016). Functional analysis of an S-Layer-associated fibronectin-binding protein in *Lactobacillus acidophilus* NCFM. *Appl. Environ. Microbiol.* 82, 2676–2685. doi: 10.1128/aem.00024-16

Janesch, B., Koerdt, A., Messner, P., and Schäffer, C. (2013). The S-layer homology domain-containing protein SlhA from *Paenibacillus alvei* CCM 2051(T) is important for swarming and biofilm formation. *PLoS One* 8:e76566. doi: 10.1371/journal.pone.0076566

Jones, P., Binns, D., Chang, H. Y., Fraser, M., Li, W., McAnulla, C., et al. (2014). InterProScan 5: genome-scale protein function classification. *Bioinformatics* 30, 1236–1240. doi: 10.1093/bioinformatics/btu031

Kamal, S. M., Rashid, A. K. M. M., Bakar, M. A., and Ahad, M. A. (2011). Anthrax: an update. *Asian Pac. J. Trop. Biomed.* 1, 496–501.

Karjalainen, T., Waligora-Dupriet, A. J., Cerquetti, M., Spigaglia, P., Maggioni, A., Mauri, P., et al. (2001). Molecular and genomic analysis of genes encoding surface-anchored proteins from Clostridium difficile. *Infect. Immun.* 69, 3442–3446. doi: 10.1128/iai.69.5.3442-3446.2001

Kawai, E., Akatsuka, H., Idei, A., Shibatani, T., and Omori, K. (1998). *Serratia marcescens* S-layer protein is secreted extracellularly via an ATP-binding cassette exporter, the Lip system. *Mol. Microbiol.* 27, 941–952. doi: 10.1046/j.1365-2958.1998.00739.x

Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N., and Sternberg, M. J. (2015). The Phyre2 web portal for protein modeling, prediction and analysis. *Nat. Protoc.* 10, 845–858. doi: 10.1038/nprot.2015.053

Kern, J., Ryan, C., Faull, K., and Schneewind, O. (2010). *Bacillus anthracis* surface-layer proteins assemble by binding to the secondary cell wall polysaccharide in a manner that requires csaB and tagO. *J. Mol. Biol.* 401, 757–775. doi: 10.1016/j.jmb.2010.06.059

Kern, J., Wilton, R., Zhang, R., Binkowski, T. A., Joachimiak, A., and Schneewind, O. (2011). Structure of surface layer homology (SLH) domains from Bacillus anthracis surface array protein. *J. Biol. Chem.* 286, 26042–26049. doi: 10.1074/jbc.m111.248070

Kern, J. W., and Schneewind, O. (2008). BslA, a pXO1-encoded adhesin of *Bacillus anthracis*. *Mol. Microbiol.* 68, 504–515. doi: 10.1111/j.1365-2958.2008.06169.x

Kern, V. J., Kern, J. W., Theriot, J. A., Schneewind, O., and Missiakas, D. (2012). Surface-layer (S-layer) proteins sap and EA1 govern the binding of the S-layer-associated protein BslO at the cell septa of *Bacillus anthracis*. *J. Bacteriol.* 194, 3833–3840. doi: 10.1128/jb.00402-12

Kirby, J. M., Ahern, H., Roberts, A. K., Kumar, V., Freeman, Z., Acharya, K. R., et al. (2009). Cwp84, a surface-associated cysteine protease, plays a role in

the maturation of the surface layer of Clostridium difficile. *J. Biol. Chem.* 284, 34666–34673. doi: 10.1074/jbc.m109.051177

Kirk, J. A., Gebhart, D., Buckley, A. M., Lok, S., Scholl, D., Douce, G. R., et al. (2017). New class of precision antimicrobials redefines role of *Clostridium difficile* S-layer in virulence and viability. *Sci. Transl. Med.* 9:eaah6813. doi: 10.1126/scitranslmed.aah6813

Kufel, W. D., Devanathan, A. S., Marx, A. H., Weber, D. J., and Daniels, L. M. (2017). Bezlotoxumab: a novel agent for the prevention of recurrent *Clostridium difficile* infection. *Pharmacotherapy* 37, 1298–1308.

Lounatmaa, K., and Brander, E. (1989). Crystalline cell surface layer of *Mycobacterium bovis* BCG. *J. Bacteriol.* 171, 5756–5758. doi: 10.1128/jb.171.10.5756-5758.1989

Martínez, M. G., Prado Acosta, M., Candurra, N. A., and Ruzal, S. M. (2012). S-layer proteins of *Lactobacillus acidophilus* inhibits JUNV infection. *Biochem. Biophys. Res. Commun.* 422, 590–595. doi: 10.1016/j.bbrc.2012.05.031

Merrigan, M. M., Venugopal, A., Roxas, J. L., Anwar, F., Mallozzi, M. J., Roxas, B. A., et al. (2013). Surface-layer protein A (SlpA) is a major contributor to host-cell adherence of *Clostridium difficile*. *PLoS One* 8:e78404. doi: 10.1371/journal.pone.0078404

Mesnage, S., Fontaine, T., Mignot, T., Delepierre, M., Mock, M., and Fouet, A. (2000). Bacterial SLH domain proteins are non-covalently anchored to the cell surface via a conserved mechanism involving wall polysaccharide pyruvylation. *EMBO J.* 19, 4473–4484. doi: 10.1093/emboj/19.17.4473

Mesnage, S., Tosi-Couture, E., and Fouet, A. (1999). Production and cell surface anchoring of functional fusions between the SLH motifs of the *Bacillus anthracis* S-layer proteins and the *Bacillus subtilis* levansucrase. *Mol. Microbiol.* 31, 927–936. doi: 10.1046/j.1365-2958.1999.01232.x

Mignot, T., Denis, B., Couture-Tosi, E., Kolstø, A. B., Mock, M., and Fouet, A. (2001). Distribution of S-layers on the surface of *Bacillus cereus* strains: phylogenetic origin and ecological pressure. *Environ. Microbiol.* 3, 493–501. doi: 10.1046/j.1462-2920.2001.00220.x

Mignot, T., Mesnage, S., Couture-Tosi, E., Mock, M., and Fouet, A. (2002). Developmental switch of S-layer protein synthesis in *Bacillus anthracis*. *Mol. Microbiol.* 43, 1615–1627. doi: 10.1046/j.1365-2958.2002.02852.x

Mignot, T., Mock, M., and Fouet, A. (2003). A plasmid-encoded regulator couples the synthesis of toxins and surface structures in *Bacillus anthracis*. *Mol. Microbiol.* 47, 917–927. doi: 10.1046/j.1365-2958.2003.03345.x

Missiakas, D., and Schneewind, O. (2017). Assembly and function of the *Bacillus anthracis* S-Layer. *Annu. Rev. Microbiol.* 71, 79–98.

Muyldermans, S. (2013). Nanobodies: natural single-domain antibodies. *Annu. Rev. Biochem.* 82, 775–797. doi: 10.1146/annurev-biochem-063011-092449

Navarre, W. W., and Schneewind, O. (1999). Surface proteins of gram-positive bacteria and mechanisms of their targeting to the cell wall envelope. *Microbiol. Mol. Biol. Rev.* 63, 174–229. doi: 10.1128/mmbr.63.1.174-229.1999

Nguyen-Mau, S.-M., Oh, S.-Y., Kern, V. J., Missiakas, D. M., and Schneewind, O. (2012). Secretion genes as determinants of *Bacillus anthracis* chain length. *J. Bacteriol.* 194, 3841–3850. doi: 10.1128/jb.00384-12

Oatley, P., Kirk, J. A., Ma, S., Jones, S., and Fagan, R. P. (2020). Spatial organization of *Clostridium difficile* S-layer biogenesis. *Sci. Rep.* 10:14089.

Oh, S.-Y., Lunderberg, J. M., Chateau, A., Schneewind, O., and Missiakas, D. (2016). Genes required for *Bacillus anthracis* secondary cell wall polysaccharide synthesis. *J. Bacteriol.* 199, e00613–e00616.

Okinaka, R. T., and Keim, P. (2016). The phylogeny of *Bacillus cereus sensu lato*. *Microbiol. Spectr.* 4. doi: 10.1128/microbiolspec.TBS-0012-2012

Poppinga, L., Janesch, B., Fünfhaus, A., Sekot, G., Garcia-Gonzalez, E., Hertlein, G., et al. (2012). Identification and functional analysis of the S-layer protein SplA of *Paenibacillus larvae*, the causative agent of American Foulbrood of honey bees. *PLoS Pathog.* 8:e1002716. doi: 10.1371/journal.ppat.1002716

Rasmussen-Ivey, C. R., Figueras, M. J., McGarey, D., and Liles, M. R. (2016). Virulence factors of *Aeromonas hydrophila*: in the wake of reclassification. *Front. Microbiol.* 7:1337. doi: 10.3389/fmicb.2016.01337

Rodrigues-Oliveira, T., Belmok, A., Vasconcellos, D., Schuster, B., and Kyaw, C. M. (2017). Archaeal S-Layers: overview and current state of the art. *Front. Microbiol.* 8:2597. doi: 10.3389/fmicb.2017.02597

Rupnik, M., Wilcox, M. H., and Gerding, D. N. (2009). *Clostridium difficile* infection: new developments in epidemiology and pathogenesis. *Nat. Rev. Microbiol.* 7, 526–536. doi: 10.1038/nrmicro2164

Sara, M., and Sleytr, U. B. (2000). S-layer proteins. *J. Bacteriol.* 182, 859–868.

Shimotahira, N., Oogai, Y., Kawada-Matsuo, M., Yamada, S., Fukutsuji, K., Nagano, K., et al. (2013). The surface layer of *Tannerella forsythia* contributes to serum resistance and oral bacterial coaggregation. *Infect. Immun.* 81, 1198–1206. doi: 10.1128/iai.00983-12

Sleytr, U. B., and Beveridge, T. J. (1999). Bacterial S-layers. *Trends Microbiol.* 7, 253–260.

Sleytr, U. B., and Messner, P. (1983). Crystalline surface layers on bacteria. *Annu. Rev. Microbiol.* 37, 311–339. doi: 10.1146/annurev.mi.37.100183.001523

Sleytr, U. B., Schuster, B., Egelseer, E.-M., and Pum, D. (2014). S-layers: principles and applications. *FEMS Microbiol. Rev.* 38, 823–864.

Spigaglia, P., Barbanti, F., and Mastrantonio, P. (2011). Surface layer protein A variant of *Clostridium difficile* PCR-ribotype 027. *Emerg. Infect. Dis.* 17, 317–319.

Sun, Z., Li, P., Liu, F., Bian, H., Wang, D., Wang, X., et al. (2017). Synergistic antibacterial mechanism of the *Lactobacillus crispatus* surface layer protein and nisin on *Staphylococcus saprophyticus*. *Sci. Rep.* 7:265.

Sychantha, D., Chapman, R. N., Bamford, N. C., Boons, G. J., Howell, P. L., and Clarke, A. J. (2018). Molecular basis for the attachment of S-layer proteins to the cell wall of *Bacillus anthracis*. *Biochemistry* 57, 1949–1953. doi: 10.1021/acs.biochem.8b00060

Takumi, K., Ichiyanagi, S., Endo, Y., Koga, T., Oka, T., and Natori, Y. (1992). Characterization, self-assembly and reattachment of S layer from *Clostridium botulinum* type *E saroma*. *Tokushima J. Exp. Med.* 39, 101–107.

Takumi, K., Susami, Y., Takeoka, A., Oka, T., and Koga, T. (1991). S layer protein of *Clostridium tetani*: purification and properties. *Microbiol. Immunol.* 35, 569–575.

Tarlovsky, Y., Fabian, M., Solomaha, E., Honsa, E., Olson, J. S., and Maresso, A. W. (2010). A *Bacillus anthracis* S-layer homology protein that binds heme and mediates heme delivery to IsdC. *J. Bacteriol.* 192, 3503–3511. doi: 10.1128/jb.00054-10

Thompson, S. A. (2002). *Campylobacter* surface-layers (S-layers) and immune evasion. *Ann. Periodontol.* 7, 43–53. doi: 10.1902/annals.2002.7.1.43

Uchida, M., Harada, T., Enkhtuya, J., Kusumoto, A., Kobayashi, Y., Chiba, S., et al. (2012). Protective effect of *Bacillus anthracis* surface protein EA1 against anthrax in mice. *Biochem. Biophys. Res. Commun.* 421, 323–328. doi: 10.1016/j.bbrc.2012.04.007

Usenik, A., Renko, M., Mihelič, M., Lindič, N., Borišek, J., Perdih, A., et al. (2017). The CWB2 cell wall-anchoring module is revealed by the crystal structures of the *Clostridium difficile* cell wall proteins Cwp8 and Cwp6. *Structure (London England 1993)* 25, 514–521. doi: 10.1016/j.str.2016.12.018

Wang, X. Y., Wang, D. B., Zhang, Z. P., Bi, L. J., Zhang, J. B., Ding, W., et al. (2015). A S-layer protein of *Bacillus anthracis* as a building block for functional protein arrays by in vitro self-assembly. *Small* 11, 5826–5832. doi: 10.1002/smll.201501413

Wang, Y. T., Oh, S. Y., Hendrickx, A. P., Lunderberg, J. M., and Schneewind, O. (2013). *Bacillus cereus* G9241 S-layer assembly contributes to the pathogenesis of anthrax-like disease in mice. *J. Bacteriol.* 195, 596–605. doi: 10.1128/jb.02005-12

Willing, S. E., Candela, T., Shaw, H. A., Seager, Z., Mesnage, S., Fagan, R. P., et al. (2015). *Clostridium difficile* surface proteins are anchored to the cell wall using CWB2 motifs that recognise the anionic polymer PSII. *Mol. Microbiol.* 96, 596–608. doi: 10.1111/mmi.12958

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Advantages of publishing in Frontiers

OPEN ACCESS
Articles are free to read for greatest visibility and readership

FAST PUBLICATION
Around 90 days from submission to decision

HIGH QUALITY PEER-REVIEW
Rigorous, collaborative, and constructive peer-review

TRANSPARENT PEER-REVIEW
Editors and reviewers acknowledged by name on published articles

Frontiers
Avenue du Tribunal-Fédéral 34
1005 Lausanne | Switzerland

Visit us: www.frontiersin.org
Contact us: frontiersin.org/about/contact

REPRODUCIBILITY OF RESEARCH
Support open data and methods to enhance research reproducibility

DIGITAL PUBLISHING
Articles designed for optimal readership across devices

FOLLOW US
@frontiersin

IMPACT METRICS
Advanced article metrics track visibility across digital media

EXTENSIVE PROMOTION
Marketing and promotion of impactful research

LOOP RESEARCH NETWORK
Our network increases your article's readership