Published in final edited form as:

Microbiol Spectr. 2019 July ; 7(4): . doi:10.1128/microbiolspec.GPP3-0012-2018.

Surface Proteins on Gram-Positive Bacteria

VINCENT A. FISCHETTI

Laboratory of Bacterial Pathogenesis and Immunology, Rockefeller University, New York, NY 10065, vaf@mail.rockefeller.edu

Abstract

As arms, legs, hair, and fur are used in higher species for their survival in the environment, surface appendages are used by bacteria for similar purposes. Surface molecules in bacteria range from complex structures such as flagella that propel the organism in aqueous environments, to less sophisticated polysaccharides and proteins. All of these molecules serve to benefit the organism for survival in a hostile environment, such as the waters of a rushing stream, the blood of an infected animal, the surface of an object, or the surface of a mucosal epithelium. Although it was previously believed that bacteria were simple single-cell organisms with little complexity, it is now apparent that they are highly evolved, advanced particles that possess a wide array of surface molecules that serve to manipulate the organism in its environment. For human pathogens, surface molecules have been finely tuned to allow adherence and colonization of host surfaces, invasion of cells, evasion of the host's immune response, and persistence in infected tissues.

In an effort to emphasize the complexity of bacterial surface molecules and their use in the everyday life of the bacterium, this chapter will focus on those surface proteins found on gram-positive bacteria. For an extensive review of the subject, see references (1, 2).

GRAM-POSITIVE CELL WALL

From electron microscopic analysis of the gram-positive cell envelope (see chapter 1, this volume) and a number of elegant chemical, structural and immunological analyses, a picture of the gram-positive cell wall has emerged (Fig. 1). The structure differs significantly from the gram-negative cell wall in two ways: (i) the presence of a thicker and more cross-linked peptidoglycan and (ii) the lack of an outer membrane. Because of these differences, surface molecules on gram-positive organisms vary from those on gram-negative organisms, requiring specialized systems to transport and anchor molecules through the outer membrane(3, 4). In general, surface proteins in gram-positive bacteria can be separated into three categories: (i) those that anchor at their C-terminal ends (through an LPxTG motif), (ii) those that bind by way of charge or hydrophobic interactions, and (iii) those that bind via their N-terminal region (lipoproteins) (Fig. 1).

C-TERMINAL-ANCHORED PROTEINS

Decades ago, relatively harsh treatments of extraction were used to remove surface proteins from gram-positive bacteria, resulting in the isolation and characterization of molecular fragments rather than complete molecules. It was not until the early 1980s that the genes for surface proteins on gram-positive bacteria were first published (protein A from

Staphylococcus aureus (5) and M protein from *Streptococcus pyogenes* (6)). Although at first glance these molecules had little in common, it was soon realized that the C-terminal end of the protein A sequence contained a sequence error. Upon correction (5), it became apparent that this region of protein A exhibited very close sequence homologies with the M protein molecule. As other protein sequences from gram-positive bacteria were published in the ensuing years, it was obvious that a common theme existed within the structure of these C-terminal-anchored molecules. Since that time hundreds of proteins from gram-positive bacteria that anchor through their C-terminal region, based on the presence of an LPxTG motif (7), have been reported (see Table 1 for a short representative list). Because of the number of molecules reported and their varied functions, the M protein molecule will be used as a model for those surface protein anchored via an LPxTG motif.

M Protein Structure

Of the C-terminal-anchored proteins, the M6 protein may be considered an archetypical molecule since it was the first gram-positive surface protein to be completely sequenced (6). A unique characteristic of this molecule is the presence of regions with sequence repeats and a conserved anchor. As seen in Fig. 2, the M6 protein is composed of four repeat blocks, each differing in size and sequence(6). The A-repeats are each composed of 14 amino acids in which the central blocks are identical and the end blocks slightly diverge from the central consensus repeats. The B-repeats, composed of 25 amino acids each, are arranged in the same way. The C-repeats, composed of 2.5 blocks of 42 amino acids, are not as similar to each other as are the A- and B-repeats. There are also four short D-repeats that show some homology. These repeat segments make up the central helical rod region of the M molecule because of the high helical potential ascribed to the amino acids found within this segment, as determined by conformational analysis (24, 25).

A close examination of the sequence within the M molecule revealed a repeating seven-residue periodicity of non-polar amino acids, a basic characteristic of α -helical coiled-coil proteins such as mammalian tropomyosin. In general, α -helical coiled-coil proteins are constructed from a reiterating seven-residue amino acid pattern (a-b-c-d-e-f-g)_n in which residues in positions a and d are hydrophobic and form the core residues in the coiled-coil, and the intervening residues are primarily helix-promoting. The arrangement of amino acids of the M6 protein within the seven-residue pattern is shown in Fig. 3. The seven-residue periodicity from amino acids 12 through 362 strongly suggests that this region is in a coiled-coil conformation and forms the helical central rod region of the molecule. Discontinuities in the heptad pattern (seen especially in the B-repeat region), which have been found in the sequence of other M proteins and other coiled-coil molecules, probably account for the flexibility of the M molecules observed in electron micrographs(25). Based on these irregularities in the heptad pattern, the central rod region is divided into three subregions that correlate with the A-, B-, and C-repeat blocks(24).

Subsequent sequence data on a variety of M protein molecules revealed that while all M proteins retain a common structural organization, they vary in the N-terminal and central domains while keeping the conserved C-terminal region (8, 21, 26–39)(Fig 4) (Trends Permission). One of the striking features observed in these various M molecules is their size

variation (see below) which is affected by the presence, number and size of the sequence repeat regions (particularly the A and B repeats) and the length of the N-terminal non-helical region. Based on the 9 M proteins in Figure 4, the A repeats are more the exception than the rule. The biological meaning of this is unclear, since all the M types depicted are found on clinical isolates of *S. pyogenes*. Though they may vary somewhat in number, the C-repeats are present in all M proteins and are surface exposed. This conserved sequence among different M molecules has vaccine implications which is detailed in chapter 10 on vaccines in this volume). For a different perspective on the M molecule please refer to (40)

Electron and Fluorescent Microscopy

Although early electron micrographs of the surface of gram-positive bacteria revealed certain structures, it was not known whether they were proteinaceous or polysaccharides. It was not until Swanson et al.(41) published the first electron micrographs of the M protein on the surface of group A streptococci that it was realized that the molecule was an elongated structure. Although other proteins were also reported to be on the streptococcal surface, they were not apparent in these electron micrographs. Thus, direct visualization of surface molecules is limited to molecules with certain physicochemical characteristics. In the case of the M protein, it is the α -helical coiled-coil structure that allowed its visualization by electron microscopy (24).

In experiments designed to answer questions about the synthesis and placement of M protein on the cell wall, Swanson et al.(41) found that after trypsinization to remove existing M protein on living streptococci newly synthesized M protein was first seen by electron microscopy on the cell in the position of the newly forming septum. In similar experiments designed after the classical experiments of Cole and Hahn(42) using fluorescein-labeled anti-M antibodies, trypsinized streptococci placed in fresh medium for 10 min revealed M protein first at the newly forming septum (Fig 5). On organisms examined after 40 min or more of incubation, M protein was not observed in the position of the old wall, suggesting that the M molecule is produced only where the new cell wall is synthesized. This confirmed the observation of Swanson et al.(41).

In the pioneering work of Carlsson et al in 2006, they showed that the sequence of the signal peptide which contains an YSIRK-G/S motif (SP_{+YSIRK}) directs the molecule to the division septum, while proteins lacking this motif (SP_{-YSIRK}), like that found in streptococcal protein F (also called SfbI) are directed to the poles (43). Modern microscopic techniques were used to better visualize this sorting phenomenon on the bacterial surface (44). The data show that M protein rapidly appears and accumulates at the septum during bacterial growth while SfbI slowly gathers at the poles where the peptidoglycan is older and more heavily crosslinked (Figure 6 Permission). Data showing decreased M protein (but not SfbI) expression using methicillin or the division regulator DivIVA strongly suggests a close connection between the regulation of cell division and protein anchoring.

C-Terminal Anchor Region

An examination of the C-terminal end of those surface molecules that anchor at that end revealed, without exception, that all have a similar arrangement of amino acids. Up to seven

charged amino acids are found at the C terminus and are composed of a mixture of both negatively and positively charged residues. Immediately N-terminal to this short charged region is a segment of 15 to 22 predominantly hydrophobic amino acids sufficient to span the cytoplasmic membrane of the bacterium, placing the charged end in the cytoplasm (Fig. 2). In all these proteins, the sequences found in the hydrophobic and charged regions are not necessarily the same, but the chemical characteristics of the amino acids used to compose them are conserved. Beginning about eight amino acids N-terminal from the hydrophobic domain is a heptapeptide with the consensus sequence LPxTG, which is extraordinarily conserved among C-terminal-anchored proteins examined (45). Although several amino acid substitutions are seen in position 3 of the heptapeptide (predominantly A, Q, E, T, N, D, K, and L), positions 1, 2, 4, and 5 are nearly completely conserved (~100% in the L and P positions and >85% in the T and G positions). This conservation is also maintained at the DNA level. The preservation of this heptapeptide and the high homology within the hydrophobic and charged regions suggest that the method of anchoring these molecules within the bacterial cell is also highly conserved, and the LPxTG motif may serve as an enzyme recognition sequence.

Thus, C-terminal-anchored surface proteins in gram positive bacteria (which could number >20 different molecules in a single organism) are synthesized and exported at the septum, where new cell wall is also being produced and translocated to the surface (24, 42). Therefore, to coordinate the export and anchoring of all these proteins, the C-terminal hydrophobic domain and charged tail function as a temporary stop, precisely positioning the LPxTG motif at the outer surface of the cytoplasmic membrane (Fig 7). This motif is then cleaved, ultimately resulting in the attachment of the surface-exposed segment of the protein to a cellular substrate(46). For an extensive review see (47). This idea is supported by studies indicating that the C-terminal hydrophobic domain and charged tail are missing from the streptococcal M protein extracted from the streptococcal cell wall(46, 48). Studies also suggest that the substrate in the peptidoglycan to which the protein is linked is the cross bridge (e.g., glycine in *S. aureus*, alanine in *S. pyogenes*) of the pentapeptide(49).

Enzymes That Cleave the LPxTG Motif

Sortase—Sortase is a membrane-bound transpeptidase found in gram-positive bacteria(2, 50–56), with the capacity to both cleave the LPxTG motif and anchor the newly formed C terminus to the nascent cross bridge in the peptidoglycan(53, 57) (for a detailed review see reference 69). Mutants in the sortase gene (*SrtA*) have been shown to be defective in their ability to surface display proteins(50–52, 58). Because surface proteins are essential for the infectivity and survival of gram-positive bacteria in the host, SrtA mutants have been shown to be attenuated in colonization experiments(20) and animal infection models(58). A genome-wide analysis of sortases revealed that sortase-like proteins were present in almost all gram-positive bacterial species, and that in every gram-positive bacterium in which one sortase-like protein was identified, at least one other was also found, and in most cases several others(50, 59). Further analysis indicates that genes encoding sortase-like proteins are usually found to be clustered along with genes encoding their probable substrates. Because of their importance in surface protein display, sortase molecules could be a target for antibiotic development. However, unlike current antibiotics, which affect bacterial

viability, antibiotics that block the action of sortases would compromise the pathogenicity of a pathogen, yielding a new class of anti-infectives.

LPxTGase—Recently a new enzyme was identified from *S. pyogenes* that differs significantly from sortase but also cleaves within the LPxTG motif of surface proteins (60, 61). This unique enzyme, termed LPxTGase, has no aromatic amino acids, is rich in alanine (particularly D-Ala(62)), is heavily glycosylated, and is 30% composed of uncommon amino acids, suggesting that it may be constructed independently of ribosomes. The enzyme, which is also found in *S. aureus*, is also important in cell surface protein display as well as cell wall assembly. Because the cleavage activity of LPxTGase is at least two orders of magnitude greater than sortase (57), LPxTGase may play an essential role in the rapid display of thousands of surface protein molecules during a short (20 to 30 min) bacterial division cycle.

Wall-Associated Region

Immediately N-terminal from the LPxTG motif is the wall-associated region, which spans about 50 to 125 amino acid residues and is found in nearly all the proteins analyzed. Although this region does not exhibit a high degree of sequence identity among the known surface proteins, it is characterized by a high percentage of proline/glycine and threonine/serine residues. For some proteins (such as the M protein) the concentration of proline/glycine is significantly higher than that of threonine/serine, while in others this relationship is either reversed or nearly equal. Because of its proximity to the hydrophobic domain (which is located in the cell membrane), this region would be positioned within the peptidoglycan layer of the cell wall (48) (Fig. 1). The reason for the presence of these particular amino acids at this location is unknown. One hypothesis suggests that the prolines and glycines, with their ability to initiate turns and loops within proteins, allow the peptidoglycan to become more easily cross-linked around these folds, thus stabilizing the molecule within the cell wall (48). The function of the threonines and serines within this region is not immediately apparent, despite the fact that these amino acids are substrates for glycosylation.

Surface-Exposed Region

As the C-terminal region is characteristically conserved among the C-terminal-anchored surface proteins, the surface-exposed region is characteristically unique. Despite their differences, this region falls into three distinct groups: (i) those with several sequence repeat blocks, (ii) those with repeat blocks located close to and within the cell wall region, and (iii) those without any repeat sequences (Fig 8). Streptococcal M protein and protein G may be considered molecules representative of those containing several sequence repeats. However, the most commonly found structures are those with repeats located close to the cell wall and an extended nonrepetitive region N-terminal to this segment.

Conformational analysis of 12 representative surface molecules by the algorithm of Garnier et al.(63) revealed that, in general, those proteins containing repeat sequences were predominantly helical within the region containing repeat segments (Fig 9). Conversely, regions and molecules without repeat blocks were predominantly composed of amino acids exhibiting high β -sheet, β -turn, and/or random coil potential. Thus, within most of these

molecules, the presence of repeat segments usually predicts the location of a helical domain. Perhaps the pressure to maintain repeat blocks is to preserve the helix potential within specific regions of these molecules, the presence of which may determine an extended protein structure, as has been shown for the M protein (25).

Seven-Residue Periodicity

As shown above, most of the surface molecules containing repeat sequences were found to be α -helical in those regions composed of repeats, while molecules without repeat blocks exhibited a high degree of β -sheet, β -turn, and random coil. However, little information was available to indicate that, except for M6 protein, any of these surface proteins are also in a coiled-coil conformation. To attempt to answer this question, an algorithm (Matcher) was developed(45) and used to determine if a seven-residue periodicity also exists in those molecules containing α -helical regions. When this algorithm was applied to representative sequences, extended regions of seven-residue periodicity were found in Arp4, FcRA, protein H, PAc, and spaP proteins (Fig 9). This strongly suggested that these molecules may attain a coiled-coil conformation within these α -helical segments. It is apparent that only those proteins that contain sequence repeats and exhibit high α -helical potential show the presence of a seven-residue pattern of hydrophobic amino acids. On the other hand, even though protein G exhibits extensive helix potential, only a limited portion was found to contain the heptad pattern. Interestingly, however, the PAc protein from *Streptococcus mutans* has only one small segment between residues 120 and 520 that displays repeat blocks, and this segment exhibits strong helix potential. This is precisely the region that was also found to contain a regular seven-residue heptad pattern. Taken together, these data strongly suggest that the coiled-coil conformation may also be a common characteristic of surface proteins found on gram-positive organisms.

Size Variation

To fully appreciate the diversity among certain surface proteins on gram-positive organisms, it is important to understand a property of the M protein molecule, namely, size variation. Using an M Protein-specific but broadly cross-reactive monoclonal antibody as a probe, the size of the M protein, extracted by solubilizing the cell wall with phage lysin from a number of streptococcal strains, was examined by Western blot. M protein derived from 20 different serotypes exhibited variations in size ranging from 41 to 80 kDa in molecular mass, depending on the strain(24). Subsequent sequencing of the M protein genes (emm) from other serotypes revealed that though the overall arrangement of the M molecules was the same, differences in the overall size of the repeats, and the presence or absence of certain repeat blocks resulted in observed size variation (see Fig 4). This variation in size may be explained by the long reiterated DNA sequence repeats likely to serve as substrates for recombinational events or replicative slippage, generating deletions and duplications within the emm gene leading to different size M proteins(64). Based on these specific characteristics, M proteins were grouped into four distinct constructions, called emm-patterns A, B, C and D (65). Interestingly, it was discovered that emm-patterns A, B and C were predominantly associated with pharyngeal infections, and D-pattern strains were associated with skin infections (40, 65).

Sequence analysis of the emm6 gene from small streptococcal outbreaks (i.e., classroom or school) revealed that they are clonally related and not the result of separate acquisitions of unrelated M6 strains during the course of the infection (data not shown). The observation that size variants occur among separate clinical isolates within the same outbreak suggests that the size mutant constituted the major organism in the streptococcal population at the time of isolation and that a given size mutant has a selective advantage under those clinical conditions(66). Perhaps serological pressure as a result of a local immune response forces the appearance of the size variant.

Although sequence repeats have been reported in several surface proteins on gram-positive bacteria, size variation has not been reported or tested for all these proteins. However, in addition to M protein, size variation has been observed in PspA from *Streptococcus pneumoniae* (67), HagA from *Porphyromonas gingivalis* (68), protein A from *S. aureus* (69), and Sfb1 from *S. pyogenes* (70).

Multifunctional Proteins

Of the >200 proteins identified on the gram-positive bacterial surface, the great majority contain domains that bind molecules found in body secretions (Table 1), including immunoglobulins (IgA and IgG), albumin, fibronectin, and fibrinogen. Other surface proteins have been identified because of the presence of enzymatic domains. Based on current data, it is clear that the majority of these proteins are multifunctional, in which the identified function is limited to only one segment of the molecule. In most cases, the function of the other domains have not been identified. For example, as described above, it appears that many surface proteins contain a repeat region located close to the cell wall and a nonrepetitive segment N-terminal to it (Fig 8). In most cases the binding domain (e.g., fibronectin, immunoglobulin) has been localized to the repeat segment while the N-terminal region exhibits no known function. In the case of streptococcal opacity factor, both domains have been defined(71); the repeat region binds fibronectin while the N-terminal domain has the catalytic function that cleaves ApoA1 of high-density lipoprotein, resulting in the opacity of serum. In the case of highly repetitive proteins such as M protein and protein G, each repeat domain has a specific function. In M protein, the A-repeats bind albumin, mucin, and C4 binding protein, while the B-repeats bind fibrinogen(72–76), and the C-repeats as well as the hypervariable region bind factor H of complement (77, 78). The C repeats also bind keratinocytes(79). In protein G, one repeat domain binds IgG while the second binds albumin. The multifunctional characteristic of surface proteins is not limited to molecules anchored via their C-terminal region. Proteins bound through hydrophobic/charge interactions and N-terminally anchored proteins also exhibit this characteristic (see below).

Thus, in general, surface proteins on most gram-positive organisms are multifunctional molecules with at least two and in some cases three or more independent functions. Given that up to 25 or more of these surface proteins may ultimately be found on a single cell surface, the potential complexity associated with the bacterial surface becomes enormous.

PROTEINS ANCHORED BY CHARGE AND/OR HYDROPHOBIC INTERACTIONS

Although the great majority of proteins identified on the surface of gram-positive bacteria anchor through their C termini, a few have been identified to bind through less-defined charge and/or hydrophobic interactions (Table 1).

Surface Glycolytic Enzymes

An analysis of the individual proteins identified in a cell wall extract of group A streptococci, it was discovered that some proteins were enzymes normally found in the glycolytic pathway located in the cytosol(80). In all, five enzymes have been identified (triosphosphate isomerase, glyceraldehyde-3-phosphate dehydrogenase [GAPDH], phosphoglycerate kinase, phosphoglycerate mutase, and α -enolase) that form a short contiguous segment of the glycolytic pathway. Interestingly, they are a complex of enzymes involved in the production of ATP as well as other functions (80–83). Thus, given the right substrate, streptococci have the capacity to produce ATP on the cell's surface, which further increases surface complexity. These glycolytic enzymes have been identified on the surface of nearly all streptococcal groups, and certain enzymes have been found on some other pathogens such as *Candida albicans*, trypanosomes, and schistosomes (Table 2). Support for the specific translocation of these enzymes to the cell surface instead of nonspecific release through cell lysis comes from the inability to identify cytoplasmic markers or other glycolytic enzymes in the growth medium or on the cell surface.

Two of the best-characterized surface glycolytic enzymes on the streptococcus are GAPDH and α -enolase (79, 81). Like other surface proteins on gram-positive bacteria, GAPDH is a multifunctional protein with binding affinities for fibronectin and lysozyme as well as for cytoskeletal proteins such as actin and myosin. GAPDH has also been shown to have ADP-ribosylating activity in addition to its GAPDH activity. α -Enolase is also multifunctional in its ability to specifically also bind plasmin(ogen) (84). It is the major plasmin-binding protein for streptococci. Evidence has shown that there is one gene for each of these enzymes, which are produced in the glycolytic cytosol. A proportion (roughly 30 to 40%) is then translocated to the surface. Because these molecules are not synthesized with an N-terminal signal sequence, it is unclear how they are transported to the cell surface. Also, because they do not contain an apparent anchor motif, it is not known how they are attached to the cell surface. Because they can be removed with chaotropic agents, it is reasonable to assume that they are bound through charge and/or hydrophobic interactions(85). Pancholi and coworkers produced a GAPDH mutant that could not be secreted to determine the importance of this molecule's export to the cell surface. Results revealed downregulation of most genes involved in *S. pyogenes* virulence and genes belonging to carbohydrate and amino acid metabolism and upregulation of those related to lipid metabolism. It also affected virulence in a mouse model, indicating that the GAPDH export process is also involved in regulating *S. pyogenes* virulence.

Other Proteins

PspA, a surface protein and a virulence determinant for *S. pneumoniae* (86), has many of the structural characteristics found in M protein, such as sequence repeats and α -helical coiled-coil conformation, but it does not contain a C-terminal-anchor motif. It has been shown that after surface translocation, PspA anchoring requires charge interactions between the membrane-associated choline-containing lipoteichoic acid and the C-terminal repeat region of the PspA molecule (87)(Fig. 1).

N-TERMINAL-ANCHORED LIPOPROTEINS

Although lipoproteins (proteins containing a lipid covalently linked to an N-terminal cysteine) have been described in gram-negative bacteria, they have also been identified in gram-positive bacteria (for reviews, see references (88, 89). Although most are predicted to be lipoproteins by genetic analysis based on the consensus sequence for lipoproteins(90), they usually comprise molecules that are not surface-exposed but located between the cytoplasmic membrane and the peptidoglycan, equivalent to the periplasm of gram-negatives. These proteins are involved in transport systems such as the *mal* and *ami* operons of *S. pneumoniae* and the Msm system of *S. mutans*. A few of these lipoproteins have been identified by their ability to act as adhesins, allowing these organisms to adhere to a variety of substrates. SsaB, a lipoprotein from *Streptococcus sanguis* that shares sequence homology with other streptococcal adhesins, has been implicated in the interaction with a salivary receptor and the coaggregation with *Actinomyces naeslundii*. Many of these so-called adhesins have also been found to participate in various transport systems for the streptococcus. Lipoproteins have also been implicated in the spore cell cycle of *Bacillus subtilis* (91), conjugation in *Enterococcus faecalis* (92)), and enzymatic functions (93, 94).

Thus, it is expected that these proteins will be located predominantly on the surface of the cell membrane, with the majority of the molecule located within the peptidoglycan (Fig. 1). In some cases it is possible that certain domains will be surface exposed, allowing them to function as adhesins or even surface enzymes. The *Mycobacterium tuberculosis* cell wall consists predominantly of mycolic acids, which is atypical, but its hydrophobicity could serve as an anchoring point for lipoproteins (95).

CONCLUSION

Based on current information, the surface of the gram-positive bacterial cell wall is highly complex (Fig 10) and could even be considered to be a distinct organelle, because of its composed of proteins with specific binding functions and enzymatic activity combined with the ability to generate energy. Considering the fact that there could be more than 20 different proteins on a given bacterial cell surface (96), each with the potential of up to three functional domains, more than 60 independent activities could potentially be present on the cell surface, much more than had ever been previously anticipated.

Because extensive cytoplasmic domains are not present within the surface proteins thus far identified in gram-positive bacteria, it is unlikely that the binding of these molecules to specific ligands induces a cytoplasmic signal to activate a gene product. It is more likely that

binding initiates a conformational signal on the bacterial cell surface to perform a specific function. For example, streptococci usually infect the pharynx, particularly the tonsils, through contact with contaminated saliva. Upon entering the oral cavity the organism first encounters the mucus that coats the mucosal epithelium. Not being a motile organism, it must navigate the mucus to reach the tonsillar epithelial surface to initiate infection. Soluble components found in the mucus, such as IgA, IgG, albumin, fibronectin, etc., are able to interact with their respective binding proteins on the streptococcal surface. This binding may initiate a set of conformational events within these proteins to drive the organism through the mucus to the epithelial surface. Perhaps the energy required for this is derived from the surface glycolytic enzymes necessary to generate ATP, or an enzyme is activated to solubilize the mucin. This must all occur quickly or the organism will be swept into the gut and eliminated by ciliary action. Thus, the molecules necessary to initiate infection are poised and ready on the bacterial surface. Therefore, before arriving on the mucosal epithelium the organism may be considered an inert particle with a wide array of surface proteins capable of driving the organism to the mucosal surface. When stable contact is made with the epithelial surface, the organism has time to initiate the infection process. Attempting to sort out when and how these proteins function during the infection process will be the challenge for future studies. With this understanding, however, new strategies could be developed in controlling diseases by this class of bacteria.

REFERENCES

1. Navarre WW, Schneewind O. 1999 Surface proteins of gram-positive bacteria and mechanisms of their targeting to the cell wall envelope. *Microbiol.Molec.Biol.Reviews* 63:174–229.
2. Dramsi S, Bierne H. 2017 Spatial Organization of Cell Wall-Anchored Proteins at the Surface of Gram-Positive Bacteria. *Curr Top Microbiol Immunol* 404:177–201. [PubMed: 27025379]
3. Thanassi DG, Saulino ET, Lombardo MJ, Roth R, Heuser J, Hultgren SJ. 1998 The PapC usher forms an oligomeric channel: implications for pilus biogenesis across the outer membrane. *Proc.Natl.Acad.Sci* 95:3146–3151. [PubMed: 9501230]
4. Hospenthal MK, Costa TRD, Waksman G. 2017 A comprehensive guide to pilus biogenesis in Gram-negative bacteria. *Nat Rev Microbiol* 15:365–379. [PubMed: 28496159]
5. Uhlen M, Guss B, Nilsson B, Gotz F, Lindberg M. 1984 Expression of the gene encoding protein A in *Staphylococcus aureus* and coagulase-negative staphylococci. *J.Bacteriol* 159:713–719. [PubMed: 6086584]
6. Hollingshead SK, Fischetti VA, Scott JR. 1986 Complete nucleotide sequence of type 6 M protein of the group A streptococcus: repetitive structure and membrane anchor. *J.Biol.Chem* 261:1677–1686. [PubMed: 3511046]
7. Fischetti VA, Pancholi V, Schneewind O. 1990 Conservation of a hexapeptide sequence in the anchor region of surface proteins of Gram-positive cocci. *Molec.Microbiol* 4:1603–1605. [PubMed: 2287281]
8. Miller L, Gray L, Beachey EH, Kehoe MA. 1988 Antigenic variation among group A streptococcal M proteins: Nucleotide sequence of the serotype 5 M protein gene and its relationship with genes encoding types 6 and 24 M proteins. *J.Biol.Chem* 263:5668–5673. [PubMed: 3281944]
9. O'Toole P, Stenberg L, Rissler M, Lindahl G. 1992 Two major classes in the M protein family in group A streptococci. *Proc.Natl.Acad.Sci.U.S.A* 89:8661–8665. [PubMed: 1528877]
10. Gomi H, Hozumi T, Hattori S, Tagawa C, Kishimoto F, Bjorck L. 1990 The gene sequence and some properties of protein H. *J.Immunol* 144:4046–4052. [PubMed: 2332638]
11. Schneewind O, Jones KF, Fischetti VA. 1990 Sequence and structural characteristics of the trypsin-resistant T6 surface protein of group A streptococci. *Journal of Bacteriology* 172:3310–3317.

12. Talay SR, Valentin-Weigand P, Jerlstrom PG, Timmis KN, Chhatwal GS. 1992 Fibronectin-binding protein of *Streptococcus pyogenes*: sequence of the binding domain involved in adherence of streptococci to epithelial cells. *Infect.Immun* 60:3837–3844. [PubMed: 1386839]
13. Johnsson H, Lindmark H, Guss B. 1995 A protein G-related cell surface protein in *Streptococcus zooepidemicus* *Infect.Immun* 63:2968–2975. [PubMed: 7542636]
14. Kastern W, Sjobring U, Bjorck L. 1992 Structure of peptostreptococcal protein L and identification of a repeated immunoglobulin light chain-binding domain. *J.Biol.Chem* 267 (18):12820–12825. [PubMed: 1618782]
15. Heden LO, Frithz E, Lindahl G. 1991 Molecular characterization of an IgA receptor from Group B *Streptococci*: sequence of the gene, identification of a proline-rich region with unique structure and isolation of N-terminal fragments with IgA-binding capacity. *Eur.J.Immunol* 21:1481–1490. [PubMed: 2044657]
16. Lindgren PE, McGavin MJ, Signas C, Guss B, Gurusiddappa S, Hook M, Lindberg M. 1993 Two different genes coding for fibronectin-binding proteins from *Streptococcus dysgalactiae*. The complete nucleotide sequences and characterization of the binding domain. *Eur.J.Biochem* 214:819–827. [PubMed: 8319691]
17. Timoney JF, Artiushin SC, Boschwitz JS. 1997 Comparison of the sequences and functions of *Streptococcus equi* M-like proteins SeM and SzPSe. *Infect.Immun* 65:3600–3605. [PubMed: 9284125]
18. Olsson A, Eliasson M, Guss B, Nilsson B, Hellman U, Lindberg M, Uhlen M. 1987 Structure and evolution of the repetitive gene encoding streptococcal protein G. *Eur.J.Biochem* 168:319–324. [PubMed: 3665928]
19. Smith HE, Vecht U, Gielkens ALJ, Smits MA. 1992 Cloning and nucleotide sequence of the gene encoding the 136-kilodalton surface protein (muramidase-released protein) of *Streptococcus suis* type 2. *Infect.Immun* 60:2361–2367. [PubMed: 1587602]
20. Kelly C, Evans P, Bergmeier L, Lee SF, Progulski FA, Harris AC, Aitken A, Bleiweis AS, Lerner T. 1990 Sequence analysis of a cloned streptococcal surface antigen I/II. *FEBS Lett* 258:127–132.
21. Mouw AR, Beachey EH, Burdett V. 1988 Molecular evolution of streptococcal M protein: Cloning and nucleotide sequence of type 24 M protein gene and relation to other genes of *Streptococcus pyogenes* *J.Bacteriol* 170:676–684.
22. Bessen DE, Fischetti VA. 1992 Nucleotide sequences of two adjacent M and M-like protein genes of group A streptococci: different RNA transcript levels and identification of a unique IgA-binding protein. *Infect.Immun* 60:124–135. [PubMed: 1370269]
23. Shimoji Y, Ogawa Y, Osaki M, Kabeya H, Maruyama S, Mikami T, Sekizaki T. 2003 Adhesive surface proteins of *Erysipelothrix rhusiopathiae* bind to polystyrene, fibronectin, and type I and IV collagens. *Journal of Bacteriology* 185:2739–2748. [PubMed: 12700253]
24. Fischetti VA, Parry DAD, Trus BL, Hollingshead SK, Scott JR, Manjula BN. 1988 Conformational characteristics of the complete sequence of group A streptococcal M6 protein. *Proteins: Struct.Func.Genet* 3:60–69.
25. Phillips GN, Flicker PF, Cohen C, Manjula BN, Fischetti VA. 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. [PubMed: 7029524]
26. Haanes EJ, Cleary PP 1989 Identification of a divergent M protein gene and an M protein related gene family in serotype 49 *Streptococcus pyogenes* *J.Bacteriol* 171:6397–6408.
27. Nilson BHK, Frick IM, Akesson P, Forsen S, Bjorck L, Akerstrom B, Wikstrom M. 1995 Structure and stability of protein H and the M1 protein from *Streptococcus pyogenes*. Implications for other surface proteins of gram-positive bacteria. *Biochemistry* 34:13688–13698. [PubMed: 7577960]
28. Manjula BN, Acharya AS, Mische SM, Fairwell T, Fischetti VA. 1984 The complete amino acid sequence of a biologically active 197-residue fragment of M protein from type 5 group A streptococci. *J.Biol.Chem* 259:3686–3693. [PubMed: 6368549]
29. Manjula BN, Trus BL, Fischetti VA. 1985 Presence of two distinct regions in the coiled-coil structure of the streptococcal Pep M5 protein: relationship to mammalian coiled-coil proteins and implications to its biological properties. *Proc.Natl.Acad.Sci.U.S.A* 82:1064–1068. [PubMed: 3856248]

30. Manjula BN, Fischetti VA. 1986 Sequence homology of group A streptococcal PepM5 protein with other coiled-coil proteins. *Biochem.Biophys.Res.Commun* 140:684–690. [PubMed: 3535794]
31. Hollingshead SK, Fischetti VA, Scott JR. 1987 A highly conserved region present in transcripts encoding heterologous M proteins of group A streptococcus. *Infect.Immun* 55:3237–3239. [PubMed: 2445686]
32. Robbins JC, Spanier JG, Jones SJ, Simpson WJ, Cleary PP 1987 Streptococcus pyogenes type 12 M protein regulation by upstream sequences. *J.Bacteriol* 169:5633–5640. [PubMed: 2445730]
33. Berge A, Sjobring U. 1993 PAM, a novel plasminogen-binding protein from Streptococcus pyogenes *J.Biol.Chem* 268(34):25417–25424.
34. Cedervall T, Johansson MU, Akerstrom B. 1997 Coiled-coil structure of group A streptococcal M proteins. Different temperature stability of class A and C proteins by hydrophobic-nophydrophobic amino acid substitutions at heptad positions a and d. *Biochemistry* 36:4987–4994. [PubMed: 9125521]
35. Heden LO, Lindahl G. 1993 Conserved and variable regions in protein Arp, the IgA receptor of Streptococcus pyogenes *J.Gen.Microbiol* 139:2067–2074.
36. Stenberg L, O'Toole PW, Mestecky J, Lindahl G. 1994 Molecular characterization of protein Sir, a streptococcal cell surface protein that binds both immunoglobulin A and immunoglobulin G. *J.Biol.Chem* 269(18):13458–13464. [PubMed: 8175778]
37. Manjula BN, Khandke KM, Fairwell T, Relf WA, Sriprakash KS. 1991 Heptad motifs within the distal subdomain of the coiled-coil rod region of M protein from rheumatic fever and nephritis associated serotypes of group A streptococci are distinct from each other: nucleotide sequence of the M57 gene and relation of the deduced amino acid sequence of other M proteins. *J.Protein Chem* 10:369–383. [PubMed: 1781883]
38. Fischetti VA, Jones KF, Hollingshead SK, Scott JR. 1988 Structure, function, and genetics of streptococcal M protein. *Rev Infect Dis* 10 Suppl 2:S356–359. [PubMed: 3055203]
39. Gubbe K, Misselwitz R, Welfle K, Reichardt W, Schmidt KH, Welfle H. 1997 C repeats of the streptococcal M1 protein achieve the human serum albumin binding ability by flanking regions which stabilize the coiled-coil conformation. *Biochemistry* 36:8107–8113. [PubMed: 9201959]
40. Smeesters PR, McMillan DJ, Sriprakash KS. 2010 The streptococcal M protein: a highly versatile molecule. *Trends Microbiol* 18:275–282. [PubMed: 20347595]
41. Swanson J, Hsu KC, Gotschlich EC. 1969 Electron microscopic studies on streptococci. I. M antigen. *J.Exp.Med* 130:1063–1091. [PubMed: 5347694]
42. Cole RM, Hahn JJ. 1962 Cell wall replication in Streptococcus pyogenes *Science* 135:722–724. [PubMed: 13880442]
43. Carlsson F, Stahlhammar-Carlemalm M, Flardh K, Sandin C, Carlemalm E, Lindahl G. 2006 Signal sequence directs localized secretion of bacterial surface proteins. *Nature* 442:943–946. [PubMed: 16929299]
44. Raz A, Talay SR, Fischetti VA. 2012 Cellular aspects of the distinct M protein and SfbI anchoring pathways in Streptococcus pyogenes. *Mol Microbiol* 84:631–647. [PubMed: 22512736]
45. Fischetti VA, Landau GM, Sellers PH, Schmidt JP. 1993 Identifying periodic occurrences of a template with applications to protein structure. *Inf.Process.Letters* 45:11–18.
46. Schneewind O, Model P, Fischetti VA. 1992 Sorting of protein A to the staphylococcal cell wall. *Cell* 70:267–281. [PubMed: 1638631]
47. Mazmanian SK, Ton-That H, Schneewind O. 2001 Sortase-catalysed anchoring of surface proteins to the cell wall of Staphylococcus aureus. *Mol.Microbiol* 40:1049–1057. [PubMed: 11401711]
48. Pancholi V, Fischetti VA. 1988 Isolation and characterization of the cell-associated region of group A streptococcal M6 protein. *J.Bacteriol* 170:2618–2624. [PubMed: 2453502]
49. Schneewind O, Fowler A, Faull KF. 1995 Structure of the cell wall anchor of surface proteins in Staphylococcus aureus *Science* 268:103–106. [PubMed: 7701329]
50. Barnett TC, Scott JR. 2002 Differential recognition of surface proteins in Streptococcus pyogenes by two sortase gene homologs. *J Bacteriol* 184:2181–2191. [PubMed: 11914350]
51. Bierne H, Mazmanian SK, Trost M, Pucciarelli MG, Liu G, Dehoux P, the European Listeria Genome C, Jansch L, Garcia-Del Portillo F, Schneewind O, Cossart P 2002 Inactivation of the srtA

- gene in *Listeria monocytogenes* inhibits anchoring of surface proteins and affects virulence. *Molec.Microbiol* 43:869–881. [PubMed: 11929538]
52. Kharat AS, Tomasz A. 2003 Inactivation of the *srtA* gene affects localization of surface proteins and decreases adhesion of *Streptococcus pneumoniae* to human pharyngeal cells in vitro. *Infection and Immunity* 71:2758–2765. [PubMed: 12704150]
 53. Mazmanian SK, Liu G, Ton-That H, Schneewind O. 1999 *Staphylococcus aureus* Sortase, an enzyme that anchors surface proteins to the cell wall. *Science* 285:760–762. [PubMed: 10427003]
 54. Osaki M, Takamatsu D, Shimoji Y, Sekizaki T. 2002 Characterization of *Streptococcus suis* genes encoding proteins homologous to sortase of gram-positive bacteria. *Journal of Bacteriology* 184:971–982. [PubMed: 11807057]
 55. Marraffini LA, Schneewind O. 2006 Targeting proteins to the cell wall of sporulating *Bacillus anthracis*. *Mol Microbiol* 62:1402–1417. [PubMed: 17074072]
 56. raz a, Fischetti VA. 2008 Sortase A localizes to distinct foci on the *Streptococcus pyogenes* membrane. *Proc.Natl.Acad.Sci.U.S.A* 105:18549–18554. [PubMed: 19017791]
 57. Ton-That H, Liu G, Mazmanian SK, Faull KF, Schneewind O. 1999 Purification and characterization of sortase, the transpeptidase that cleaves surface proteins of *Staphylococcus aureus* at the LPXTG motif. *Proc.Natl.Acad.Sci.U.S.A* 96:12424–12429. [PubMed: 10535938]
 58. Mazmanian SK, Liu G, Jensen ER, Lenoy E, Schneewind O. 2000 *Staphylococcus aureus* sortase mutants defective in the display of surface proteins and in the pathogenesis of animal infections. *PNAS* 97:5510–5515. [PubMed: 10805806]
 59. Pallen MJ, Lam AC, Antonio M, Dunbar K. 2001 An embarrassment of sortases—a richness of substrates? *Trends Microbiol* 9:97–102. [PubMed: 11239768]
 60. Lee SG, Pancholi V, Fischetti VA. 2002 Characterization of a unique glycosylated anchor endopeptidase that cleaves the LPXTG sequence motif of cell surface proteins of Gram-positive bacteria, p. 46912–46922, vol. 277.
 61. Lee SG, Fischetti VA. 2006 Purification and characterization of LPXTGase from *Staphylococcus aureus* : the amino acid composition mirrors that found in the peptidoglycan. *Journal of Bacteriology* 188:389–398. [PubMed: 16385028]
 62. Lee SG, Fischetti VA. 2003 Presence of D-Alanine in an Endopeptidase from *Streptococcus pyogenes* p. 46649–46653, vol. 278.
 63. Garnier J, Osguthorpe DJ, Robson B. 1978 Analysis of the accuracy and implications of simple methods for predicting the secondary structure of globular proteins. *J.Mol.Biol* 120:97–120. [PubMed: 642007]
 64. Hollingshead SK, Fischetti VA, Scott JR. 1987 Size variation in group A streptococcal M protein is generated by homologous recombination between intragenic repeats. *Mol.Gen.Genet* 207:196–203. [PubMed: 3039291]
 65. Bessen DE, Fiorentino TR, Hollingshead SK. 1997 Molecular markers for throat and skin isolates of group A streptococci. *Adv Exp Med Biol* 418:537–543. [PubMed: 9331709]
 66. Jones KF, Hollingshead SK, Scott JR, Fischetti VA. 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. [PubMed: 2460864]
 67. Waltman WD, McDaniel LS, Gray BM, Briles DE. 1990 Variation in the molecular weight of PspA (pneumococcal surface protein A) among *Streptococcus pneumoniae*. *Microbial Pathogenesis* 8:61–69.
 68. Kozarov E, Whitlock J, Dong H, Carrasco E, Progulsk-Fox A. 1998 The number of direct repeats in *hagA* is variable among *Porphyromonas gingivalis* strains. *Infect.Immun* 66:4721–4725. [PubMed: 9746569]
 69. Cheung AL, Fischetti VA. 1988 Variation in the expression of cell wall proteins of *Staphylococcus aureus* grown on solid and liquid media. *Infect.Immun* 56:1061–1065. [PubMed: 3281899]
 70. Rocha CL, Fischetti VA. 1999 Identification and characterization of a novel fibronectin-binding protein on the surface of group A streptococci. *Infection and Immunity* 67.
 71. Rakonjac JV, Robbins JC, Fischetti VA. 1995 DNA sequence of the serum opacity factor of group A streptococci: identification of a fibronectin-binding repeat domain. *Infect.Immun* 63:622–631. [PubMed: 7822031]

72. Blom AM, Berggard K, Webb JH, Lindahl G, Villoutreix BO, Dahlback B. 2000 Human C4b-binding protein has overlapping, but not identical, binding sites for C4b and streptococcal M proteins. *J.Immunol* 164:5328–5336. [PubMed: 10799895]
73. Horstmann RK, Sievertsen HJ, Leippe M, Fischetti VA. 1992 Role of fibrinogen in complement inhibition by streptococcal M protein. *Infect.Immun* 60:5036–5041. [PubMed: 1452335]
74. Johnsson E, Thern A, Dahlback B, Heden LO, Wikstrom M, Lindahl G. 1996 A highly variable region in members of the streptococcal M protein family binds the human complement regulator C4BP. *J.Immunol*:3021–3029. [PubMed: 8816411]
75. Ryan PA, Pancholi V, Fischetti VA. 2001 Group A streptococci bind to mucin and human pharyngeal cells through sialic acid-containing receptors. *Infect.Immun* 69:7402–7412. [PubMed: 11705914]
76. Ryc M, Beachy EH, Whitnack E. 1989 Ultrastructural localization of the fibrinogen-binding domain of streptococcal M protein. *Infect.Immun* 57(8):2397–2404. [PubMed: 2473035]
77. Horstmann RD, Sievertsen HJ, Knobloch J, Fischetti VA. 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. [PubMed: 2964038]
78. Gustafsson MC, Lannergard J, Nilsson OR, Kristensen BM, Olsen JE, Harris CL, Ufret-Vincenty RL, Stalhammar-Carlemalm M, Lindahl G. 2013 Factor H binds to the hypervariable region of many *Streptococcus pyogenes* M proteins but does not promote phagocytosis resistance or acute virulence. *PLoS Pathog* 9:e1003323. [PubMed: 23637608]
79. Okada N, Liszewski MK, Atkinson JP, Caparon M. 1995 Membrane cofactor protein (CD46) is a keratinocyte receptor for the M protein of group A streptococcus. *Proc.Natl.Acad.Sci.U.S.A* 92:2489–2493. [PubMed: 7708671]
80. Pancholi V, Fischetti VA. 1992 A major surface protein on group A streptococci is a glyceraldehyde-3-phosphate dehydrogenase with multiple binding activity. *J.Exp.Med* 176:415–426. [PubMed: 1500854]
81. Pancholi V 2001 Multifunctional $\ddot{\alpha}$ -enolase: its role in diseases. *CMLS* 58:902–920. [PubMed: 11497239]
82. Pancholi V, Fischetti VA. 1997 Identification of a glycolytic enzyme complex on the surface of group A streptococci. *ASM General Meeting*:35–35.
83. Fontan PA, Pancholi V, Nociari MM, Fischetti VA. 2000 Antibodies to streptococcal surface enolase react with human $\ddot{\alpha}$ -enolase:Implications in poststreptococcal sequelae. *J.Infect.Dis* 182:1712–1721. [PubMed: 11069244]
84. Pancholi V, Fischetti VA. 1998 $\ddot{\alpha}$ -enolase, a novel strong plasmin(ogen) binding protein on the surface of pathogenic streptococci. *The Journal of Biological Chemistry* 273:14503–14515. [PubMed: 9603964]
85. Derbise A, Song YP, Parikh S, Fischetti VA, Pancholi V. 2004 Role of the C-Terminal lysine residues of streptococcal surface enolase in Glu-and Lys-plasminogen-binding activities of group A streptococci, p. 94–105, vol. 72.
86. Wu HY, Nahm MH, Guo Y, Russel MW, Briles DE. 1997 Intranasal immunization of mice with PspA (pneumococcal surface protein A) can prevent intranasal carriage, pulmonary infection, and sepsis with *Streptococcus pneumoniae*. *J.Infect.Dis* 175:839–846. [PubMed: 9086139]
87. Yother J, White JM. 1994 Novel surface attachment mechanism of the *Streptococcus pneumoniae* protein PspA. *J.Bacteriol* 176(10):2976–2985. [PubMed: 7910604]
88. Nakayama H, Kurokawa K, Lee BL. 2012 Lipoproteins in bacteria: structures and biosynthetic pathways. *FEBS J* 279:4247–4268. [PubMed: 23094979]
89. Hutchings MI, Palmer T, Harrington DJ, Sutcliffe IC. 2009 Lipoprotein biogenesis in Gram-positive bacteria: knowing when to hold ‘em, knowing when to fold ‘em. *Trends Microbiol* 17:13–21. [PubMed: 19059780]
90. Sutcliffe IC, Harrington DJ. 2002 Pattern searches for the identification of putative lipoprotein genes in Gram-positive bacterial genomes, p. 2065–2077, vol. 148.
91. Errington J, Appleby L, Daniel RA, Goodfellow H, Partridge SR, Yudkin MD. 1992 Structure and function of the spoIIIJ gene of *Bacillus subtilis*: a vegetatively expressed gene that is essential for

- epsilon G activity at an intermediate stage of sporulation. *J.Gen.Microbiol* 138:2609–2618. [PubMed: 1487728]
92. Tanimoto K, An FY, Clewell DB. 1993 Characterization of the traC determinant of the *Enterococcus faecalis* hemolysin-bacteriocin plasmid pAD1: binding of sex pheromone. *J.Bacteriol* 175:5260–5264. [PubMed: 8349566]
93. Lansing M, Lellig S, Mausolf A, Martini I, Crescenzi F, O'Regan M, Prehm P. 1993 Hyaluronate synthases: cloning and sequencing of the gene from *Streptococcus* sp. *Biochemical Journal* 289:179–184. [PubMed: 8424756]
94. Rothe B, Roggentin P, Frank R, Blocker H, Schauer R. 1989 Cloning, sequencing and expression of sialidase gene from *Clostridium sordellii* *J.Gen.Microbiol* 135:3087–3096. [PubMed: 2693593]
95. Kovacs-Simon A, Titball RW, Michell SL. 2011 Lipoproteins of bacterial pathogens. *Infect Immun* 79:548–561. [PubMed: 20974828]
96. Koller T, Nelson D, Nakata M, Kreutzer M, Fischetti VA, Glocker MO, Podbielski A, Kreikemeyer B. 2008 PlyC, a novel bacteriophage lysin for compartment-dependent proteomics of group A streptococci. *Proteomics* 8:140–148. [PubMed: 18095374]

FIGURE 1:

Major surface structures of the cell wall of gram-positive bacteria. Linked to the surface of the peptidoglycan, many gram-positive organisms have polysaccharide structures that in some cases are used for their immunological classification. Surface proteins are linked by three mechanisms. (i) Lipoproteins have a lipid linked through a cysteine at the N terminus. (ii) C-terminal-anchored proteins are attached and stabilized in the peptidoglycan through a C-terminal complex containing an LPxTG motif. (Most surface proteins are anchored in this way.) (iii) Certain surface proteins are attached through hydrophobic and/or charge interactions to the cell surface. (Some proteins are bound ionically to the lipoteichoic acid.) The teichoic acids (TA) are a common feature of the gram-positive cell wall. TA is usually composed of a repeating carbohydrate-phosphate polymer linked through a phosphodiester linkage to the peptidoglycan. Lipoteichoic acid (LTA) is composed of a similar polymer linked to the cytoplasmic membrane through a fatty acid (see chapter 19, this volume).

FIGURE 2:

Characteristics of the complete M6 protein sequence. Blocks A, B, C, and D designate the location of the sequence repeat blocks. Block C3 is half the size of blocks C1 and C2. Shaded blocks indicate those in which the sequence diverges from the central consensus sequence. Pro/Gly denotes the proline- and glycine-rich region likely located in the peptidoglycan. Membrane anchor is a 19-hydrophobic-amino-acid region adjacent to a 6-amino-acid charged tail. Pepsin identifies the position of the pepsin-sensitive site after amino acid 228. The C-terminal end is located within the cell wall and membrane.

	1	Arg	Val	Phe	Pro	Arg	Gly	Thr		
		Val	Glu	Asn	Pro	Gly	Thr	Leu		NON HELICAL DOMAIN
12	a	Asp	Lys	Lys	Ala	Arg	Glu	Leu		
128	b	Leu	Asn	Met	Leu	Asp	Val	Glu		
135	c	Asn	Ser	Lys	Leu	Gln	Ala	Asn		
139	d	Asn	Asp	Lys	Leu	Thr	Thr	Glu		
146	e	Asn	Lys	Asn	Leu	Thr	Thr	Gln		
153	f	Asn	Lys	Asn	Leu	Thr	Asp	Gln		
160	g	Asn	Lys	Asn	Leu	Thr	Thr	Glu		
167		Asn	Lys	Asn	Leu	Thr	Asp	Gln		
174		Asn	Lys	Asn	Leu	Thr	Thr	Glu		
181		Glu	Asn	Arg	Leu	Thr	Thr	Glu		
188		Asn	Lys	Gly	Leu	Thr	Lys	Lys		
195		Leu	Ser	Glu	Ala	Glu	Glu	Glu		
102		Ala	Asn	Lys	Glu	Arg	Glu	Asn		
109		Lys	Glu	Ala	Ile	Gly	Thr	Leu		
117		Lys	Lys	Thr	Val	Asp	Glu	Thr		
124		Lys	Lys	Thr	Val	Lys	Asp	Lys		
131		Lys	Lys	Thr	Val	Lys	Asp	Lys		
135		Lys	Lys	Thr	Val	Lys	Asp	Lys		
142		Lys	Lys	Thr	Ile	Gly	Thr	Leu		
149		Lys	Lys	Thr	Leu	Asp	Glu	Thr		
156		Lys	Lys	Thr	Val	Lys	Asp	Lys		
160		Lys	Lys	Thr	Glu	Gln	Glu	Ser		
167		Lys	Lys	Thr	Ile	Gly	Thr	Leu		
174		Lys	Lys	Thr	Leu	Asp	Glu	Thr		
181		Lys	Lys	Thr	Val	Lys	Asp	Lys		
185		Lys	Lys	Thr	Glu	Gln	Glu	Ser		
192		Lys	Lys	Thr	Ile	Gly	Thr	Lys		
199		Lys	Lys	Thr	Leu	Asp	Glu	Thr		
206		Lys	Lys	Thr	Val	Lys	Asp	Lys		
210		Lys	Lys	Thr	Glu	Gln	Glu	Ser		
217		Lys	Lys	Thr	Leu	Gly	Ala	Leu		
224		Lys	Lys	Thr	Ala	Lys	Lys	Lys		
231		Lys	Lys	Thr	Val	Lys	Asp	Lys		
234		Lys	Lys	Thr	Ser	Glu	Ala	Ser		
241		Lys	Lys	Thr	Leu	Arg	Arg	Asp		
248		Lys	Lys	Thr	Val	Arg	Arg	Asp		
255		Lys	Lys	Thr	Gln	Val	Glu	Lys		
262		Lys	Lys	Thr	Leu	Thr	Ala	Glu		
269		Lys	Lys	Thr	Val	Lys	Asp	Lys		
276		Lys	Lys	Thr	Ile	Ser	Asp	Lys		
283		Lys	Lys	Thr	Gly	Leu	Arg	Asp		
290		Lys	Lys	Thr	Ala	Ser	Arg	Glu		
297		Lys	Lys	Thr	Gln	Val	Glu	Lys		
304		Lys	Lys	Thr	Ala	Asn	Ser	Lys		
311		Lys	Lys	Thr	Ala	Leu	Glu	Lys		
318		Lys	Lys	Thr	Glu	Leu	Glu	Ser		
325		Lys	Lys	Thr	Ala	Thr	Glu	Lys		
332		Lys	Lys	Thr	Glu	Leu	Gln	Glu		
337		Lys	Lys	Thr	Ala	Leu	Glu	Glu		
343		Lys	Lys	Thr	Ala	Leu	Lys	Gln		
350		Lys	Lys	Thr	Ala	Glu	Ala	Glu		
357		Lys	Lys	Thr	Ala	Leu	Arg	Ala		
363		Gly	Lys	Ala	Ser	Asp	Ser	Gln	Thr	
371		Pro	Asp	Ala	Lys	Val	Val	Ala		
375		Pro	Gly	Ala	Lys	Gln	Ala	Ala		
381		Pro	Gly	Ala	Lys	Gln	Ala	Ala		
387		Pro	Gln	Ala	Lys	Thr	Lys	Ala		
393		Pro	Asn	Gln	Asn	Lys	Lys	Ala		
399		Pro	Met	Lys	Glu	Thr	Lys	Arg	Gln	Leu
408		Pro	Ser	Thr	Gly	Glu	Thr	Ala	Asn	Pro
417		Phe	Phe	Thr	Ala	Ala	Ala	Leu		
424		Thr	Val	Met	Ala	Thr	Ala	Gly		
431		Val	Ala	Ala	Val	Val	Val	Ala		
438		Lys	Arg	Lys	Glu	Glu	Asn			

FIGURE 3:

Complete M6 protein sequence arranged to highlight the seven-residue periodicity found in the helical central rod region. Region assignments are based on sequence and conformational analyses. Arrangement of the sequence is based on the position of amino acids in a seven-residue periodicity designated by letters a through g beginning at residue 12 and continuing through residue 362, with interruptions at residues 109, 131, 156, 181, 206, 231, and 337. Alignment from residue 363 to 416 is used essentially to highlight the regularity of the position of prolines in the sequence. No periodicity is found from residue 417 to the end. Three major regions are indicated (nonhelical, helical, and anchor). The pepsin-sensitive site is between Ala-228 and Lys-229.

FIGURE 4:

Figured reused from (26) with permission. Schematic representation of the organization of nine well-described M proteins showing the presence and location of the sequence repeat regions and pattern arrangement.

FIGURE 5:

Fluorescein-labeled anti-M6 antibody analysis of the appearance of M6 protein on streptococcal cell walls. M6 streptococci were treated with trypsin to remove surface M protein. Cells, reincubated at 37°C, were removed at intervals, fixed, and stained with fluorescein-labeled anti-M6 antibody. (a) After 10 min of incubation the M protein is located within a thin band at the position of the newly forming septum. Magnification, $\times 5,000$. (b) Location of the M protein after 40 min of incubation. Note that no fluorescein label is seen in the position of the old wall. Magnification, $\times 6,000$.

FIGURE 6.

A model representation of the anchoring of M protein and SfbI. The regeneration of surface proteins during two division cycles is presented. M protein (red) is anchored exclusively to newly synthesized peptidoglycan at the septum (sites of active anchoring labelled yellow). Anchoring of M protein at daughter septa begins before the mother septum is completely closed, resulting in simultaneous anchoring at both locations. Following two generations, M protein is anchored to all newly synthesized, but not pre-existing, peptidoglycan. SfbI (green) is anchored over time in patches to peripheral peptidoglycan, with some preference to the poles. Following two generations, the oldest poles (I) display the most intense SfbI labelling, while one-generation-old poles show less labelling (II), and newly formed poles (III) show little SfbI labelling.

FIGURE 7:

Computer-generated model of the C-terminal end of the M protein sequence (residues 371 to 441). A comparable region is found in all C-terminal-anchored surface proteins from gram-positive bacteria (see Table 1). The predicted location of this segment of the molecule is shown in the cytoplasm, membrane, and peptidoglycan. The space between the membrane and peptidoglycan (wall region) may be considered the “periplasm” of the gram-positive bacterium. The figure was generated on a Steller computer using the Quanta 2.1A program for energy minimization.

FIGURE 8:

Characteristics of C-terminal-anchored surface proteins. These proteins fall into three basic categories: those with multiple repeats, single repeats, or no repeats. Proteins with a single repeat located close to the cell wall are the most common. Anchor is the region of the molecule located within the cell wall carbohydrate and peptidoglycan.

FIGURE 9:

Conformational characteristics of the surface molecules from gram-positive bacteria. The sequences were analyzed by the Garnier-Robson algorithm supplied with the EuGene protein analysis package. The location within these molecules of regions exhibiting random coil (C), β -turn (T), β -sheet (S), and α -helix (H) are designated. Sequences were derived from the references listed in Table 1. Shaded areas are those containing a seven-residue periodicity based on the Matcher program. Proteins T6 and wapA exhibit no extended helical regions or seven-residue periodicity.

FIGURE 10:

Appearance of the surface of gram-positive bacteria exhibiting a wide array of protein molecules. Each molecule depicted may have thousands of identical copies densely packed on the surface.

TABLE 1

Sequenced surface proteins from gram-positive bacteria linked to the cell wall through an LPxTG motif. A representative list of molecules from a range of gram-positive pathogens (6, 17, 18, 45, 84–96)

NAME/GENE	FUNCTION/NAME	ORGANISM	LPXTG	ACCESS#
1. emm6	M protein	<i>S. pyogenes</i>	LPSTG	M11338
2. emm5	M protein	<i>S. pyogenes</i>	LPSTG	M20374
3. Mrp4	IgG/fibrinogen binding	<i>S. pyogenes</i>	LPSTG	M87831
4. Prot H	Human IgG Fc binding	<i>S. pyogenes</i>	LPSTG	M29398
5. T6	Protease resistant protein	<i>S. pyogenes</i>	LPSTG	M32978
6. sof22	Serum opacity factor	<i>S. pyogenes</i>	LPASG	U02290
7. Sfb	Fibronectin binding	<i>S. pyogenes</i>	LPATG	X67947
8. ZAG	Binds γ_2 M, Alb, IgG	<i>S. zooepidemicus</i>	LPTTG	U02290
9. Prot L	Light chain binding	<i>P. magnus</i>	LPKAG	M86697
10. bac	IgA binding protein	Gr. B strep	LPYTG	X58470
11. fnbA	Fibronectin binding	<i>S. dysgalactiae</i>	LPQTG	Z22150
12. SzPSe	M-like	<i>S. equi</i>	LPQTS	U73163
13. Prot G	IgG binding protein	Gr. G strep	LPTTG	X06173
14. MRP	Surface protein*	<i>S. suis</i>	LPNTG	X64450
15. spaP	Surface protein	<i>S. mutans</i>	LPNTG	X17390
17. fruA	Fructosidase	<i>S. mutans</i>	LPDTG	L03358
19. Prot A	IgG binding protein	<i>S. aureus</i>	LPETG	X00342
23. inlA	Internalization protein	<i>L. monocytogenes</i>	LPTTG	M67471
26. Hyal1	Hyaluronidase	<i>S. pneumoniae</i>	LPQTG	L20670
27. InlC2	Internalin-like	<i>L. monocytogenes</i>	LPTAG	U77368
28. RSP	Adhesin	<i>E. rhusiopathiae</i>	LPN/QTG	AB052682

TABLE 2:

Glycolytic enzymes found on the surface of microorganisms

Enzyme ^a	Bacteria	Fungus	Parasite
GAPDH	Streptococci <i>S. pneumoniae</i> <i>S. aureus</i> <i>Mycobacterium avium</i>	<i>Candida albicans</i>	Schistosome Trypanosome
PGK	Streptococci	<i>C. albicans</i>	Trypanosome
TPI	Streptococci		Trypanosome
PGM	Streptococci		
α -Enolase	Streptococci	<i>C. albicans</i>	

^aGAPDH, glyceraldehyde-3-phosphate dehydrogenase; PDK, phosphoglycerate kinase; TPI, triphosphosphate isomerase; PGM, phosphoglycerate mutase.