PART I Bacterial adhesins and adhesive structures

CHAPTER 1

Surface protein adhesins of staphylococci

Timothy J. Foster

1.1 INTRODUCTION

Staphylococcus aureus is primarily an extracellular pathogen. In order to initiate infection it adheres to components of the host extracellular matrix (ECM). Adherence is mediated by surface protein adhesins called MSCRAMMs (microbial surface components recognizing adhesive matrix molecules) (Patti *et al.*, 1994a). In most cases the MSCRAMMs are covalently bound to peptidoglycan in the cell wall. However, there are several examples of MSCRAMMs that are non-covalently associated with the wall. Coagulase-negative staphylococci also express MSCRAMMs. This chapter will discuss the mechanisms of attachment of proteins to the cell wall and will review the properties of MSCRAMM proteins that have been characterized at the molecular level.

1.2 ANCHORING OF PROTEINS TO THE CELL WALL

Cell-wall-anchored proteins that are covalently bound to peptidoglycan are recognizable by a motif located at the C-terminus (Navarre and Schneewind, 1999). This comprises the sequence LPXTG (Leu-Pro-X-Thr-Gly) followed by hydrophobic residues that span the cytoplasmic membrane and by several positively charged residues. The positively charged residues are required to hold the protein transiently in the membrane during secretion through the Sec secretome (Schneewind *et al.*, 1993). The LPXTG sequence is recognized by an enzyme called sortase that cleaves LPXTG between the Thr and Gly residues (Navarre and Schneewind, 1994; Ton-That *et al.*, 1999). The carboxyl group of the Thr is joined to the amino group of the branch peptide in nascent peptidoglycan. In the case of *S. aureus*, linkage occurs to the NH_2 group of the fifth Gly residue of the branch peptide, which would otherwise form the interpeptide bridge of cross-linked peptidoglycan

(Ton-That and Schneewind, 1999; Ton-That *et al.*, 1999). The wall-anchored protein becomes joined to the lipid-linked intermediate prior to its incorporation into peptidoglycan (Ton-That *et al.*, 1997). The covalently linked protein can be released from the cell only by enzymatic degradation of peptidoglycan. The glycine endopeptidase lysostaphin releases proteins of homogeneous size whereas muramidases release proteins of heterogeneous size owing to varying amounts of peptidoglycan attached to the C-terminus (Schneewind *et al.*, 1993).

The enzyme that catalyses the sorting reaction is sortase. The protein has an N-terminal hydrophobic domain that provides anchorage to the outer face of the cytoplasmic membrane (Mazmanian *et al.*, 1999; Ton-That *et al.*, 1999). Indeed it is likely that sortase is closely associated with the secretome because it must capture proteins as they are in the process of being secreted through the Sec pathway. Interestingly, a sortase-defective mutant can grow normally *in vitro*, which shows that sortase is not an essential enzyme. The mutant is defective in the expression of several LPXTG-anchored surface proteins and has reduced virulence in murine infection models (Mazmanian *et al.*, 2000). This indicates that sortase could be a novel target for antimicrobial agents.

1.3 CELL-WALL-ASSOCIATED PROTEINS

1.3.1 Protein A

Protein A (Spa) is the archetypal cell-wall-anchored protein of *S. aureus*. It is known primarily for its ability to bind the Fc region of immunoglobulin (Ig) G. Its structural organization is somewhat different from that of other surface proteins in that the N-terminal signal sequence is followed by tandem repeats of five homologous IgG binding domains (Fig. 1.1; Uhlén *et al.*, 1984). Each is composed of an approximately 60 amino acid residue unit that forms three α-helices (Starovasnik *et al.*, 1996). The structure of the subdomain B in complex with the Fc region of IgG subclass 1 has been solved by X-ray analysis of a co-crystal (Deisenhofer, 1981). The binding between the two molecules involves nine amino acid residues in the IgG fragment and 11 amino acid residues in the protein A domain (Gouda *et al.*, 1998). The binding characteristics and specificity of the Spa–IgG interaction have been analysed in great detail (Langone, 1982).

One important role of Spa in staphylococcal infections is that it is antiphagocytic. By binding to protein A on the bacterial surface, the Fc region of IgG is not available for recognition by the Fc receptor on polymorphonuclear leukocytes (PMNLs) (Gemmell *et al.*, 1990). A protein A-defective mutant

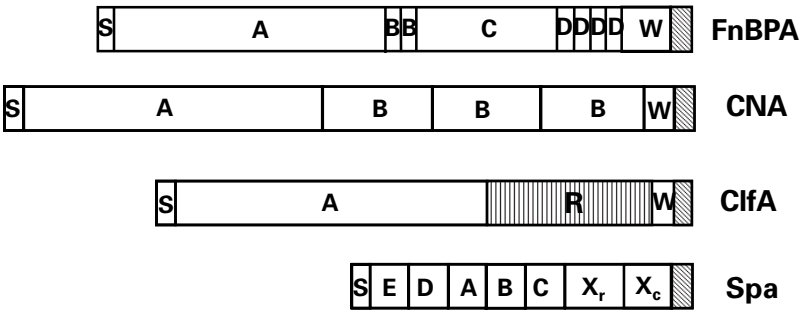

Figure 1.1. Organization of surface proteins of *S. aureus*. The domain organization of the fibronectin-binding protein A (FnBPA), the collagen-binding protein CNA, the fibrinogen-binding protein (ClfA) and protein A (Spa). The signal sequences (S) are removed during secretion across the cytoplasmic membrane. Region X_r of protein A is a proline-rich octapeptide repeat that spans the cell wall. Region X_c is a non-repeated (constant) region. Each protein has common features at the C-terminus indicated by the cross-hatched box (LPXTG motif, hydrophobic region, and positively charged residues). Regions W and R are peptidoglycan-spanning regions. (From Foster and Höök, 2000; with permission from American Society for Microbiology Press.)

was more avidly phagocytosed by PMNLs in the presence of normal serum opsonins than was the wild type, and the mutant was less virulent in murine infection models (Patel *et al.*, 1987). The recent observation that Spa can mediate adherence of bacteria to von Willebrand factor, an extracellular matrix protein important in normal haemostasis, suggests that protein A may have an additional role in the infection process (Hartlieb *et al.*, 2000).

1.3.2 Fibronectin-binding proteins

Fibronectin (Fn) is a dimeric glycoprotein that occurs in a soluble form in body fluids and in a fibrillar form in the ECM (Hynes, 1993). A primary function of insoluble Fn is to act as a substratum for the adhesion of cells mediated by integrin receptors that bind to specific sites in the central part of Fn (Yamada, 1989). The primary binding site for staphylococcal Fn-binding protein is in the 29 kDa N-terminal domain, which is composed of five type I modules (Sottile *et al.*, 1991; Potts and Campbell, 1994).

Most strains of *S. aureus* express two related Fn-binding proteins FnBPA and FnBPB, which are encoded by closely linked genes (Signás *et al.*, 1989; Jönsson *et al.*, 1991). One survey of 163 isolates comprising carriage strains, as well as strains from invasive disease and orthopaedic device-related infection, found that 77% had both *fnbA* and *fnbB* genes, while 23% had only *fnbA*

(Peacock *et al.*, 2000). Strains from orthopaedic infections adhered to Fn at a significantly higher level than did carriage strains or strains from non-device-related infections.

FnBPA and FnBPB have a structural organization similar to that of FnBPs from streptococci (Fig. 1.1; Joh *et al.*, 1994; Patti *et al.*, 1994a; Foster and Höök, 1998). The primary ligand-binding domain (D), which is almost identical in FnBPA and FnBPB, is located very close to the cell-wall-spanning domains (region W) and is composed of three to five repeats of an approximately 40 residue motif. Synthetic peptides mimicking repeated units effectively inhibit Fn binding to bacteria and bacterial attachment to immobilized Fn (Raja *et al.*, 1990).

Studies with peptides and recombinant proteins expressing combinations of different FnBP D repeats indicate that the major interaction occurs between FnBP repeat D3 and Fn modules 4 and 5, but that other discrete sequences within the D region bind to different Fn type I module pairs (Joh *et al.*, 1998). The ligand-binding domain of FnBP reacts simultaneously at multiple sites with Fn (Fig. 1.2). A consensus Fn-binding motif is present with each D repeat (McGavin *et al.*, 1991, 1993). The interaction between the MSCRAMM and Fn involves structural rearrangements in the D repeat region. The ligand-binding D repeat region has an unordered structure and acquires a defined conformation upon binding to the rigid type I modules of Fn (House-Pompeo *et al.*, 1996). This conformational change is accompanied by the formation of neo-epitopes called ligand-induced binding site (LIBS) epitopes that can be demonstrated by monoclonal antibodies and by antibodies isolated from patients recovering from staphylococcal infection (Speziale *et al.*, 1996). The antibodies that recognize the neo-epitopes do not interfere with the MSCRAMM–Fn interaction but rather stabilize the FnBP–ligand complex and appear to promote Fn binding. The immunodominant non-LIBS epitopes in D1–D3 are confined to repeats D1 and D2 (Sun *et al.*, 1997). They are very similar to each other and bind Fn with lower affinity than does D3. The region in D3 corresponding to epitopes in D1–D2 contains the Fn-binding consensus but is otherwise divergent (8/21 different residues). Antibodies in polyclonal sera preferentially react with the low affinity D1–D2 domains and block Fn binding to bacteria by no more than 50%.

Staphylococcus aureus can invade cultured fibroblasts, endothelial and epithelial cells, and Fn plays a major role (Dziewanowska *et al.*, 1999; Lammers *et al.*, 1999; Peacock *et al.*, 1999a; Sinha *et al.*, 1999; Fowler *et al.*, 2000). Bacteria either recruit soluble Fn or bind to Fn bound to the surface of host cells. Bacteria bind Fn via the type I modules at the N-terminus. Fn

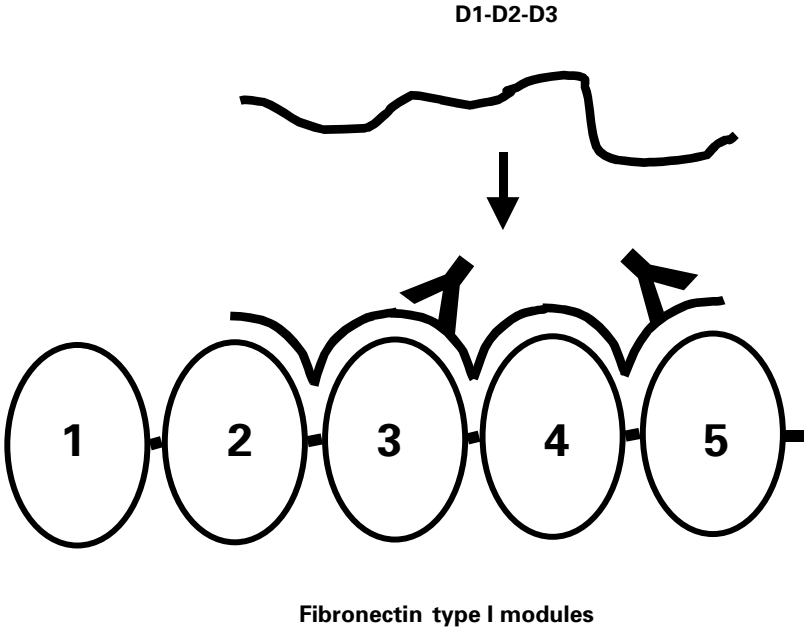

Figure 1.2. Interaction of the ligand-binding region of Fn-binding proteins with Fn. The wavy line represents the ligand-binding D1–D2–D3 repeats of FnBPs, which do not have secondary structure. The protein interacts with the type I modules of Fn and takes on a discernible secondary structure with the formation of neo-epitopes (ligand-induced binding site epitopes). (From Foster and Höök, 2000; with permission from American Society for Microbiology Press.)

is bound to the $\alpha_5\beta_1$ integrin on the surface of the host cell at the central Fn RGD (Arg-Gly-Asp) motif-bearing module. Thus Fn forms a bridge between the bacterial FnBP adhesin and the mammalian cell integrin (Fig. 1.3). This results in stimulation of phagocytosis and bacteria become internalized. FnBP-defective mutants of *S. aureus* are not taken up, non-invasive bacteria that acquire FnBP expression become invasive, and bacterial internalization is blocked by soluble recombinant D repeat regions of FnBP and by anti-integrin function-blocking antibodies. The importance of internalization *in vivo* is unclear, but it could be involved in bacterial escape from the bloodstream and invasion of internal organs, in the initiation of invasive endocarditis, and in bacterial persistence.

FnBPs are considered to be important virulence factors in the initiation

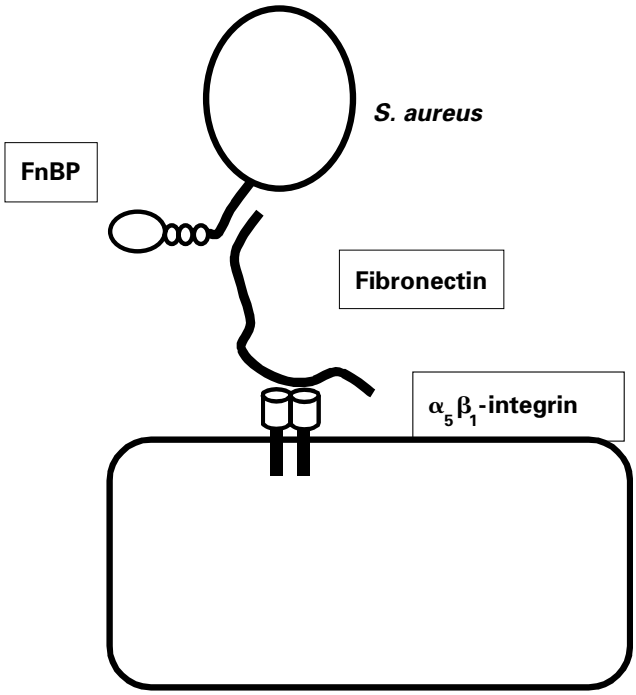

Figure 1.3. Role of Fn in promoting bacterial attachment to mammalian cells. The N-terminal type I modules of Fn are bound to the D1–D2–D3 region of FnBP attached to the cell surface of *S. aureus*. The same molecule of Fn binds to the $\alpha_5\beta_1$ integrin via the centrally located RGD motif. This stimulates actin rearrangement and bacterial internalization.

of foreign body infection. They promote bacterial adherence to immobilized Fn *in vitro*, and to implanted biomaterial that has been in long-term contact with the host such as plastic coverslips implanted subcutaneously in guinea pigs (Greene *et al.*, 1995) and titanium screws implanted in the iliac bone of guinea pigs (Fischer *et al.*, 1996). In contrast, fibrinogen appears to be the major adhesion-promoting factor in the conditioning layer on biomaterial that had been in short-term contact with the host (Vaudaux *et al.*, 1995; François *et al.*, 2000). There are conflicting data concerning the ability of FnBPs to promote infection in experimental animals. One study with a FnBP-defective mutant of strain 879 indicated that the bacterial MSCRAMM is important in promoting bacterial adhesion to damaged heart valve tissue in the rat model for endocarditis (Kuypers and Proctor, 1989), while another report with mutants of strain 8325–4 found no effect (Flock *et al.*, 1996).

These contradictory data may reflect different bacterial strains (8325–4 is known to express FnBPs poorly *in vitro*) or differences in performing the infection models.

1.3.3 Fibrinogen-binding proteins

Fibrinogen is a large protein of M_r 340 000. It is composed of three polypeptide chains (α, β, γ) that are extensively linked by disulphide bonds to form an elongated dimeric structure (Ruggeri, 1993). It is the most abundant ligand for the integrin $\alpha_{IIb}/\beta 3$ (glycoprotein gpIIb/IIIa) on the surface of platelets. The binding of Fg to the integrin receptor on activated platelets results in platelet aggregation and the formation of platelet–fibrin thrombi (Hawiger, 1995). The C-terminal sequences of the α-, β- and γ-chains form independently folded globular domains. The three dimensional structure of the γ-chain module is known (Spraggon *et al.*, 1997; Doolittle *et al.*, 1998)

Until recently it was thought that the ability of *S. aureus* to adhere to Fg-coated substrates and to form clumps in a solution containing Fg (e.g. plasma) was solely due to the clumping factor ClfA (McDevitt *et al.*, 1994, 1995). It is now known that *S. aureus* can express other Fg-binding adhesins: the ClfB protein, which is related to ClfA (Ní Eidhin *et al.*, 1998); and the Fg-binding proteins, which can also interact with Fg via their A domains (Wann *et al.*, 2000). The *clfA* and *clfB* genes are not allelic variants but are distinct genes. They are not closely linked, in contrast to the *fnbA* and *fnbB* genes.

The structural organization of ClfA and ClfB is very similar (Fig. 1.4). The surface-exposed approximately 500 residue ligand-binding A domains are linked to the cell wall via the R domain, which comprises mainly the Ser-Asp dipeptide repeats. The R domain appears to serve as a flexible stalk, allowing the presentation of the A domain on the surface for ligand interactions (Hartford *et al.*, 1997). FnBPA and FnBPB also possess N-terminal A domains that have sequence similarity with the A domains of ClfA and ClfB, and in the case of FnBPA promote binding to Fg (Wann *et al.*, 2000). Otherwise, the FnBP and Clf proteins have a completely different structural organization, apart from the typical cell-wall-anchoring signals at the extreme C-terminus.

Although the structural organization of ClfA and ClfB is similar, the amino acid sequences of the ligand-binding A domains are only 27% identical. ClfA and ClfB bind to different sites in Fg. ClfA recognizes the flexible peptide that extends from the γ-module at the C-terminus of the γ-chain (Fig. 1.5; McDevitt *et al.*, 1997) whereas ClfB binds to the α-chain (Ní Eidhin *et al.*, 1998). The A domain of FnBPA is approximately 25% identical with that of ClfA and binds to the same site in the Fg γ-chain (Wann *et al.*, 2000).

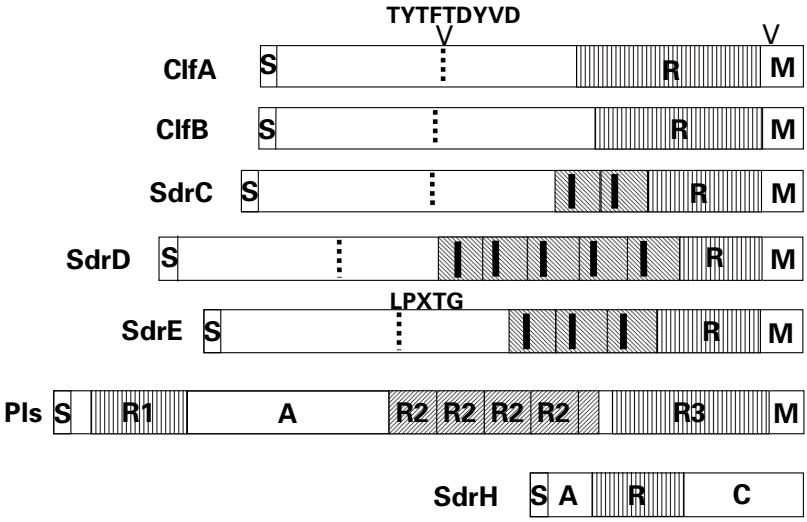

Figure 1.4. The Sdr multigene protein family in staphylococci. The domain organization of different members of the Sdr protein family is shown. The approximately 500 residue A domain of ClfA, ClfB, SdrC, SdrD, SdrE, SdrF and SdrG have a conserved sequence TYTFTDYVD. SdrF and SdrG of *S. epidermidis* have an organization similar to that of the SdrC, SdrD and SdrE proteins of *S. aureus* but are not shown. The A domain of Pls has no sequence similarity to the A domains of the Clf–Sdr proteins, which are related by about 25–30%.

The ClfA domain recognizes a site located at the extreme C-terminus of the γ-chain of Fg (residues 399–411). These residues form a flexible extension from the globular γ-module (Fig. 1.5; Spraggon *et al.*, 1997: Doolittle *et al.*, 1998). A 17 amino acid residue synthetic peptide corresponding to the Fg γ-chain residues 399–411 binds to a recombinant form of the A domain in an interaction that is inhibited by Ca^{2+} (O'Connell *et al.*, 1998). The A domain contains a motif that is reminiscent of a Ca^{2+}-binding EF-hand, and site-specific mutations in this motif resulted in a protein with lower affinity for the γ-chain peptide and less sensitivity to Ca^{2+}. Thus ClfA exhibits fibrinogen-binding characteristics similar to those of the platelet integrin α_{IIb}/β_3 (O'Connell *et al.*, 1998). Both bind to the same site in Fg in interactions that are affected by Ca^{2+}. The recombinant form of the ClfA A domain is a potent inhibitor of Fg-dependent platelet aggregation (McDevitt *et al.*, 1997). This could be a bacterial defence mechanism to prevent release of antimicrobial peptides during degranulation (Yeaman *et al.*, 1992) that might occur during platelet aggregation in the vicinity of colonizing bacteria.

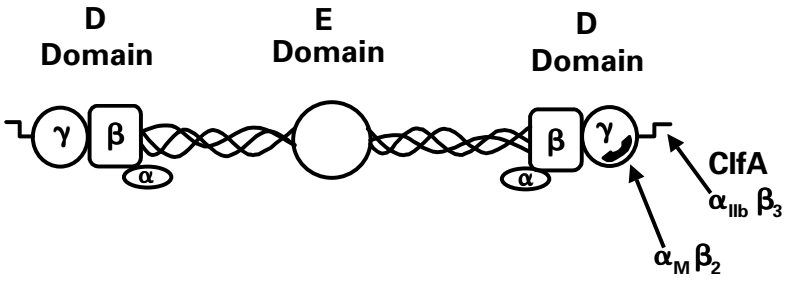

Figure 1.5. Structure of fibrinogen. Schematic diagram showing the structural organization of fibrinogen. The globular D domains comprise the C-terminal residues of the α-, β- and γ-chains. The C-terminus of the γ-chain protrudes from the globular γ-module. Binding sites for ClfA and integrins are shown. The E domain contains the N-terminal residues of α-, β- and γ-chains cross-linked by disulphide bridges. (From Foster and Höök, 2000; with permission from American Society for Microbiology Press.)

The binding of ClfA to fibrinogen is progressively inhibited by Ca^{2+} in the range 1–10 mM (O'Connell *et al.*, 1998). The concentration of free Ca^{2+} in blood plasma is 1.3 mM and is closely regulated at the threshold of the inhibitory range, although concentrations can vary more widely in extracellular spaces (Brown *et al.*, 1995). However, at platelet-rich thrombi, and possibly on the surface of freshly implanted biomaterial, the Ca^{2+} concentration appears to be considerably lower and may allow ClfA to bind fibrinogen. Thus, as bacteria circulate in plasma, they will tend to adhere to Fg/platelet-containing coagulation sites.

The ClfB region A binds to the Fg α-chain (Ní Eidhin *et al.*, 1998) but the precise binding site has not been defined. ClfB-promoted binding to immobilized Fg is also inhibited by millimolar concentrations of Ca^{2+}. The ClfB protein is expressed maximally during the early part of the exponential phase of growth (Ní Eidhin *et al.*, 1998; McAleese *et al.*, 2001). Transcription (and hence translation) terminates before the culture reaches stationary phase, so ClfB molecules become diluted amongst the progeny cells during the remaining cell divisions and some are released into the culture supernatant by cell wall turnover. Also, some of the ClfB protein is cleaved by the metalloprotease aureolysin (McAleese *et al.*, 2001). Protease cleavage results in loss of an N-terminal domain and the protein loses its ability to bind Fg. In contrast, ClfA is expressed abundantly on cells from the stationary phase of growth. It is also cleaved at a site similar to that of ClfA but does not lose ligand-binding activity (McDevitt *et al.*, 1995).

The ClfA protein is the primary adhesin of *S. aureus* for promoting