Cell, Vol. 115, 217–228, October 17, 2003, Copyright ©2003 by Cell Press

# A "dock, lock, and latch" Structural Model for a Staphylococcal Adhesin Binding to Fibrinogen

**Karthe Ponnuraj,[1] M. Gabriela Bowden,[2] Stacey Davis,[2] S. Gurusiddappa,[2] Dwight Moore,[1] Damon Choe,[2] Yi Xu,[2] Magnus Hook,[2,*] and Sthanam V. L. Narayana[1,*]**

[1]School of Optometry and Center for Biophysical Sciences and Engineering
University of Alabama at Birmingham
Birmingham, Alabama 35294

[2]Center for Extracellular Matrix Biology
Institute of Biosciences and Technology
Texas A&M University Health Science Center
Houston, Texas 77030

## Summary

**Gram-positive pathogens such as staphylococci contain multiple cell wall-anchored proteins that serve as an interface between the microbe and its environment. Some of these proteins act as adhesins and mediate bacterial attachment to host tissues. SdrG is a cell wall-anchored adhesin from *Staphylococcus epidermidis* that binds to the Bβ chain of human fibrinogen (Fg) and is necessary and sufficient for bacterial attachment to Fg-coated biomaterials. Here, we present the crystal structures of the ligand binding region of SdrG as an apoprotein and in complex with a synthetic peptide analogous to its binding site in Fg. Analysis of the crystal structures, along with mutational studies of both the protein and of the peptide, reveals that SdrG binds to its ligand with a dynamic "dock, lock, and latch" mechanism. We propose that this mechanism represents a general mode of ligand binding for structurally related cell wall-anchored proteins of gram-positive bacteria.**

## Introduction

*Staphylococcus epidermidis* is a major component of the normal skin flora, colonizing the healthy human epidermis and mucous membranes. However, *S. epidermidis* has evolved into one of the leading causes of nosocomial sepsis; this opportunistic pathogen is now the most common organism isolated from nosocomial primary blood stream infections (BSI) (Richards et al., 2000). Alarmingly, antibiotic resistance is emerging among *S. epidermidis* strains and other related gram-positive opportunistic pathogens such as *S. aureus* and *E. faecalis*. The emergence of these multiple antibiotic-resistant strains constitutes an imminent public health crisis. Therefore, we need a better understanding of the pathogenic molecular strategies employed by these organisms to develop new antimicrobial drugs and treatments.

Microbial attachment to the host tissues represents the first critical step in the establishment of most bacterial infections. This process involves adhesins on the surface of the microbe that bind to ligands in the host in "lock and key" interactions. For pathogenic bacteria that do not invade host cells, extracellular matrix proteins are preferred targets for bacterial adhesion. On gram-positive bacteria such as Staphylococci, Streptococci, and Enterococci, the adhesins mediating these interactions have been termed MSCRAMMs (microbial surface components recognizing adhesive matrix molecules) (Patti and Hook, 1994). Multiple MSCRAMMs have been found in many pathogenic gram-positive species, and their amino acid sequence relatedness, similar modular design, and common binding domain organization suggest that they might have arisen from a common ancestor. In general, these structurally related MSCRAMMs contain an N-terminal ~40 amino acid long signal sequence (S), and C-terminal features that are required for sorting the proteins to the cell wall including a proline-rich wall-spanning region (W), the wall anchoring LPXTG motif, a hydrophobic transmembrane region (M), and a positively charged cytoplasmic tail (C). In addition to these C-terminal features, the members of the Sdr family of proteins, including the fibrinogen (Fg) binding MSCRAMMs ClfA and ClfB of *S. aureus*, and SdrG of *S. epidermidis* contain a Ser-Asp repeat region (R region) just outside the cell wall-anchoring region (see the SdrG prototype in Figure 1A). The ligand binding activity of most of these MSCRAMMs has been localized to the N-terminal A regions that are ~500 amino acid long (reviewed by Foster and Hook, 1998). In the case of SdrG, the ligand binding site has been further localized to residues 273–597 (Davis et al., 2001).

Fg appears to be targeted by many pathogens, and staphylococci can express several different proteins that bind to this host component. For instance, the *S. aureus* MSCRAMMs ClfA, ClfB, FnbpA, and FnbpB all recognize Fg. In addition, this organism can secrete several Fg binding proteins including Map/Eap (Jönsson et al., 1995, Palma et al., 1999), Efb (Palma et al., 2001), and Emp (Hussain et al., 2001). During tissue repair, Fg not only controls blood loss by clotting, but it is also part of the host defense against microbial infections. In damaged tissues, the proteolysis of Fg by thrombin releases potent chemotactic fragments such as fibrinopeptide B, attracting leucocytes and fibroblasts to the site of injury (Kay et al., 1974; Richardson et al., 1976; Senior et al., 1986; Skogen et al., 1988). We have shown previously that SdrG's binding site in the Fg Bβ chain overlaps the thrombin cleavage site and that SdrG can interfere with fibrin clot formation by inhibiting the thrombin mediated release of fibrinopeptide B (Davis et al., 2001). It is possible that staphylococci express proteins that bind to Fg to prevent the release of chemotactic elements, thus reducing the influx of phagocytic neutrophils to the infection site and enhancing the bacterial survival in the host.

In this report, we present the crystal structure of the Fg binding domain of SdrG, both as an apoprotein and in complex with a synthetic peptide analogous to the

*Correspondence: narayana@uab.edu (S.V.L.N.), mhook@ibt.tamu.edu (M.H.)

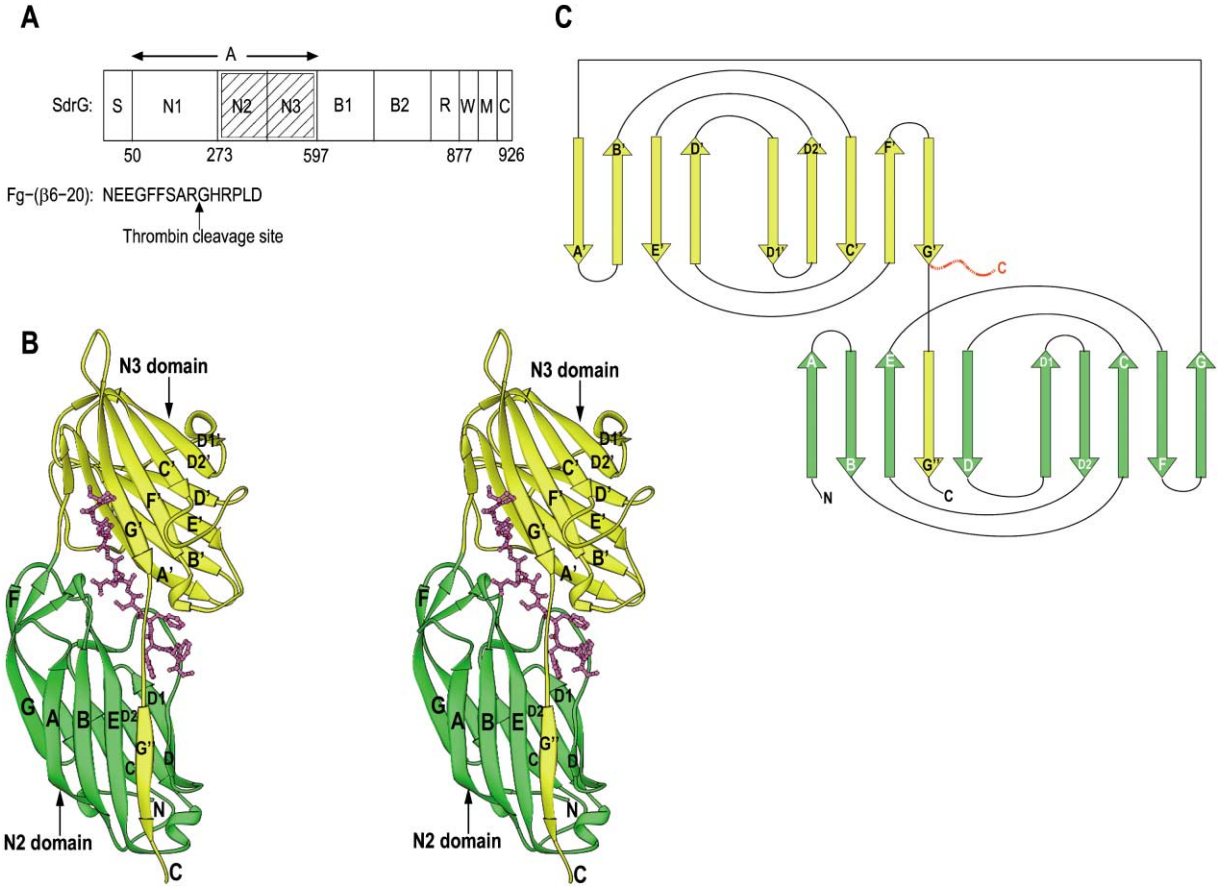


Figure 1. Representation of the rSdrG$_{(276–597)}$-Peptide Complex

(A) Schematic representation of the SdrG molecule (Top): A, N-terminal Fg binding region; N2-N3, subdomains crystallized in the rSdrG$_{(276–597)}$-peptide complex; B1 and B2, repeats of unknown function; R, serine-aspartate repeat region; W, wall-spanning region; M, membrane-spanning region; C, cytoplasmic positively charged tail. The amino acid residue number identifying the boundary between each subdomain is indicated below. The Fg (β6-20) peptide sequence (Below). The arrow indicates the thrombin cleavage site.

(B) Stereo ribbon (Carson, 1997) representation of rSdrG$_{(276–597)}$-Fg (β6-20) peptide analog complex. The peptide is shown in the "ball and stick" model.

(C) Schematic representation of the topology of the rSdrG$_{(276–597)}$ fold. The N2 and N3 domains are shown in green and yellow, respectively. The dotted line at the end of the G′ strand indicates the C terminus of the apo-rSdrG$_{(276–597)}$ structure. In the rSdrG$_{(276–597)}$-peptide complex, the C-terminal strand inserts between strands D and E of the N2 domain.

binding site in its ligand, human Fg. Using a combination of site-directed mutagenesis in SdrG$_{(273–597)}$ and amino acid replacements in the Fg (β6-20) peptide, we have developed and examined a model for the mechanism of the MSCRAMM-ligand interaction. In addition, our analyses suggest that this mechanism is a common mode of ligand binding among structurally related MSCRAMMs from gram-positive bacteria.

## Results and Discussion

### Structure of the SdrG$_{(276–596)}$-Fg Peptide Complex

Initial characterization of the recombinant A-region of SdrG revealed that this protein is readily processed at the N terminus by a contaminating protease, yielding a resistant fragment corresponding to residues 273–597 (Figure 1A). Similar processing was previously observed in vivo for the A-region of ClfB (McAleese et al., 2001) and in vitro for the A-region of ClfA (Deivanayagam et al., 1999). We therefore reengineered a plasmid that would generate a recombinant fusion protein covering residues 273–597 of SdrG with an N-terminal His-tag for purification purposes. This recombinant SdrG$_{(273–597)}$ was crystallized in complex with its ligand, a synthetic peptide designed to contain residues 6–20 of the human Fg β-chain (Figure 1A). The structure of the rSdrG$_{(276–596)}$-peptide complex was determined by multiple isomorphic replacement methods. The solvent flattened electron density map was readily interpretable and allowed a model for the structure of the rSdrG$_{(276–596)}$-peptide complex to be built. The structure has been refined to an R factor of 20.9% ($R_{free}$ of 22.1%) against diffraction data extending to 1.86 Å resolution (Table 1). There are two independent molecules in the crystallographic asymmetric unit of a $P2_12_12_1$ cell and they exhibit an identical structure with an rms deviation of 0.2 Å on all Cα atoms upon superposition.

The folding of the rSdrG$_{(276–597)}$ molecule with a bound peptide is illustrated in Figure 1B. No electron density was observed for the 20 N-terminal residues, which in-

Table 1. Peptide Bound and Free $rSdrG_{(276–597)}$ Crystallographic and Refinement Data

| Parameter | Peptide Bound | Peptide Free |
|---|---|---|
| a (Å) | 89.02 | 44.39 |
| b (Å) | 89.48 | 92.76 |
| c (Å) | 98.04 | 93.92 |
| α (°) | 90 | 83.62 |
| β (°) | 90 | 76.06 |
| γ (°) | 90 | 90.57 |
| Space group | $P2_12_12_1$ | P1 |
| Resolution (Å) | 1.86 | 3.5 |
| Reflections total/unique | 549,956/63,746 | 57,863/13,739 |
| Completeness (%) | 95.7 | 97.3 |
| $R_{sym}$[a] | 6.1 | 8.4 |
| Number of molecules in the asymmetric unit | 2 | 4 |
| **Refinement** | | |
| Resolution range (Å) | 20–1.86 | 20.0–3.5 |
| R factor/$R_{free}$[b] (All data with 0 sigma cutoff) | 0.209/0.221 | 0.250/0.319 |
| Average B value (Å$^2$) | 20.53 | 30.34 |
| Rmsd for bonds (Å) | 0.006 | 0.012 |
| Rmsd for angles (°) | 1.27 | 1.51 |
| Number of nonhydrogen protein atoms | 5003 | 8216 |
| Number of peptide atoms | 170 | |
| Number of water molecules | 438 | |
| Number of $Ca^{2+}$ ions | 2 | |
| PDB code | 1R17 | 1R19 |

[a] $R_{sym}=\Sigma|Ih-<Ih>|/\Sigma Ih$, where $<Ih>$ is the average intensity over symmetry equivalents.
[b] Number of reflections set aside for calculation of the $R_{free}$ value for peptide bound and unbound structures were 1239 and 669, respectively.

clude 12 residues contributed by the vector $His_6$-tag sequence and residues 273–275 of the rSdrG protein. Likewise, no electron density was observed for one residue (two residues in the second molecule of the asymmetric unit) of the rSdrG C terminus. The $rSdrG_{(276–596)}$ polypeptide folds into two distinct domains; N2 and N3. In agreement with earlier terminology (Perkins et al., 2001), the term N1 has been reserved for the proteolytically processed N-terminal segment of the A-region. The topologies of N2 and N3, with 13% residues conserved between them are structurally similar and follow the DEv-IgG (DEvariant-IgG) fold (Figure 1C). The DEv-IgG fold, which represents a variant of the C-type immunoglobulin fold, was recently identified in the crystal structure of the apo-$rClfA_{(221–559)}$, the minimum Fg binding region of the MSCRAMM clumping factor A (ClfA) from *S. aureus* (Deivanayagam et al., 2002). The N2 and N3 domains in the $rSdrG_{(276–596)}$-peptide crystal structure interact with each other and have a combined buried surface area of 1378 Å$^2$. The N-terminal N2-domain has 148 residues (276–424), and is slightly smaller than the N3 domain, which has 170 residues (425–597). The N2 and N3 domains superpose with an rms deviation of 1.8 Å for 48 common Cα atoms and 3.7 Å on all Cα atoms, and the differences are mostly found in the length and conformation of the loops joining the β strands of the two domains.

The two principal β sheets of the N3 domain are composed of β strands A′, B′, E′, and D′ on one face of a "sandwich" and strands D1′, D2′, C′, F′, and G′ on the opposite face (Figure 1C). The insertion of the additional strands D1′ and D2′ represent the unique feature of the DEv-IgG fold compared to the C-type IgG fold. In the N2 domain, strands D1, D2, C, F, and G form one sheet, the second sheet contains an additional strand G″ inserted between the E and the D. This inserted G″ strand is not formed by residues from the N2 domain, but instead, it is formed by residues Ser587-Glu597 that represent the C-terminal extension of the N3 domain (Figure 1C). The inserted G″ strand buries a surface area of 1138 Å$^2$ of the N2 domain.

### Structural Features of the SdrG Binding Site

A cleft of approximately 30 Å in length between the N2 and N3 domains constitutes the Fg binding region of $rSdrG_{(276–596)}$. The Fg-derived peptide fits snuggly into this cleft in an extended conformation, interacting with both the N2 and the N3 domains (Figure 1B). A well-defined electron density was observed for most of the residues of the bound peptide (Figure 2A) except for three N-terminal (Asn6f, Glu7f and Glu8f) and two C-terminal residues (Leu19f and Asp20f). These residues are extended into the solvent environment, having few interactions with the SdrG molecule, and with a weak electron density in the 2fo-fc maps. The superposition of the bound peptide in the two complex structures shows an rms deviation of 0.16 Å (on main chain atoms) indicating that the conformations of the peptide in both structures are nearly identical.

The central region of the cleft is covered by residues Ile581-Ser586, a segment corresponding to the linker connecting the N3 domain G′ strand with strand G″ strand of the N2 domain, giving the binding region a "tunnel-like" appearance (Figure 2B). The peptide residues Phe9af-Ala13f are positioned in this tunnel, appearing as if the peptide is threaded through the $rSdrG_{(276–596)}$ molecule.

An area of 1200 Å$^2$ is buried between $rSdrG_{(276–596)}$ and the peptide, involving 62 contacts with less than 4 Å between the protein and the peptide ligand. For example, there are 8 antiparallel backbone hydrogen bonds between residues Phe10f, Ser12f, Arg14f, and His16f in the peptide and Ser584, Ala582, Thr580, and Asp578 in the "linker" and the N3 domain of $rSdrG_{(276–596)}$ (Figure 2C). In addition, the peptide residues Gly15f, Ala13f, and Phe11f form four more backbone-backbone hydrogen bonds with residues His304, Ala306, and Leu340 of the N2 domain. The backbone atom of Phe11f makes a hydrogen bond with the side-chain of Ser338. The side chains of the peptide amino acids form two hydrogen bonds with $rSdrG_{(276–596)}$: (1) Arg14f (NH1) to Tyr294 (OH) and (2) Ser12f (OG) to Asp339 (OD2). Three Phe residues (Phe9af, Phe10f, and Phe11f) in the N terminus exhibit a large number of hydrophobic interactions with residues from the N2 and the N3 domains and appear like a "bulgy plug" positioned at the end of the tunnel of the ligand binding cleft.

### SdrG-Ligand Binding Specificity

Based on the $rSdrG_{(276–596)}$-peptide complex structure we predicted which amino acids in $rSdrG_{(276–596)}$ and in the peptide are important to stabilize the adhesin-ligand

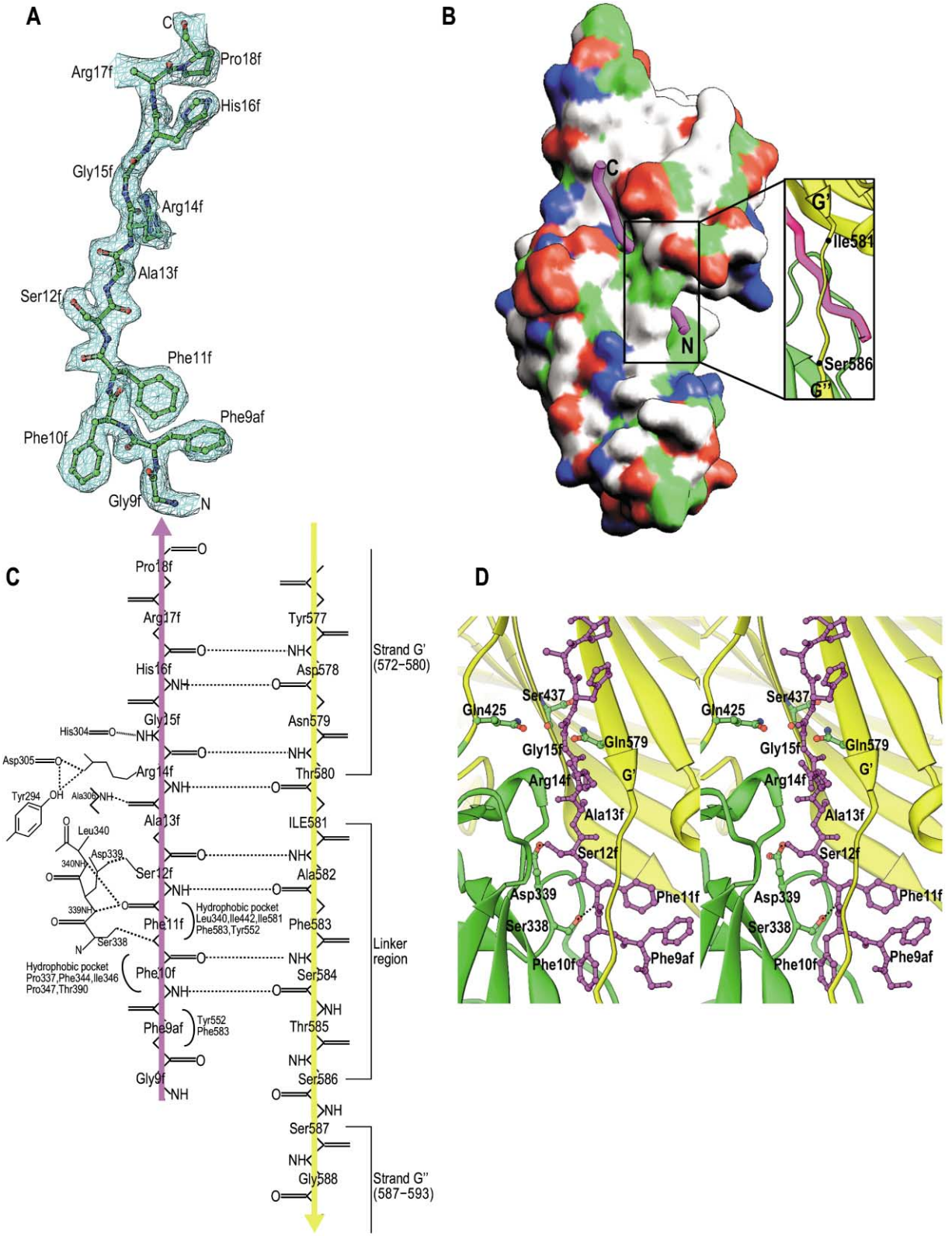


Figure 2. rSdrG$_{(276–597)}$-Peptide Interactions

(A) 2Fo-Fc simulated annealing omit map of the Fg (β6-20) peptide analog. The map is contoured at the level of 1σ. The peptide residues are distinguished from the protein residues by the suffix "f".

(B) Surface representation of rSdrG$_{(276–597)}$ showing the peptide "locked" into the molecule. The surface is color-coded according to nonpolar, polar, acidic, and basic residues that are represented as green, white, red, and blue, respectively. Closer view of the ligand binding tunnel is shown in ribbon representation.

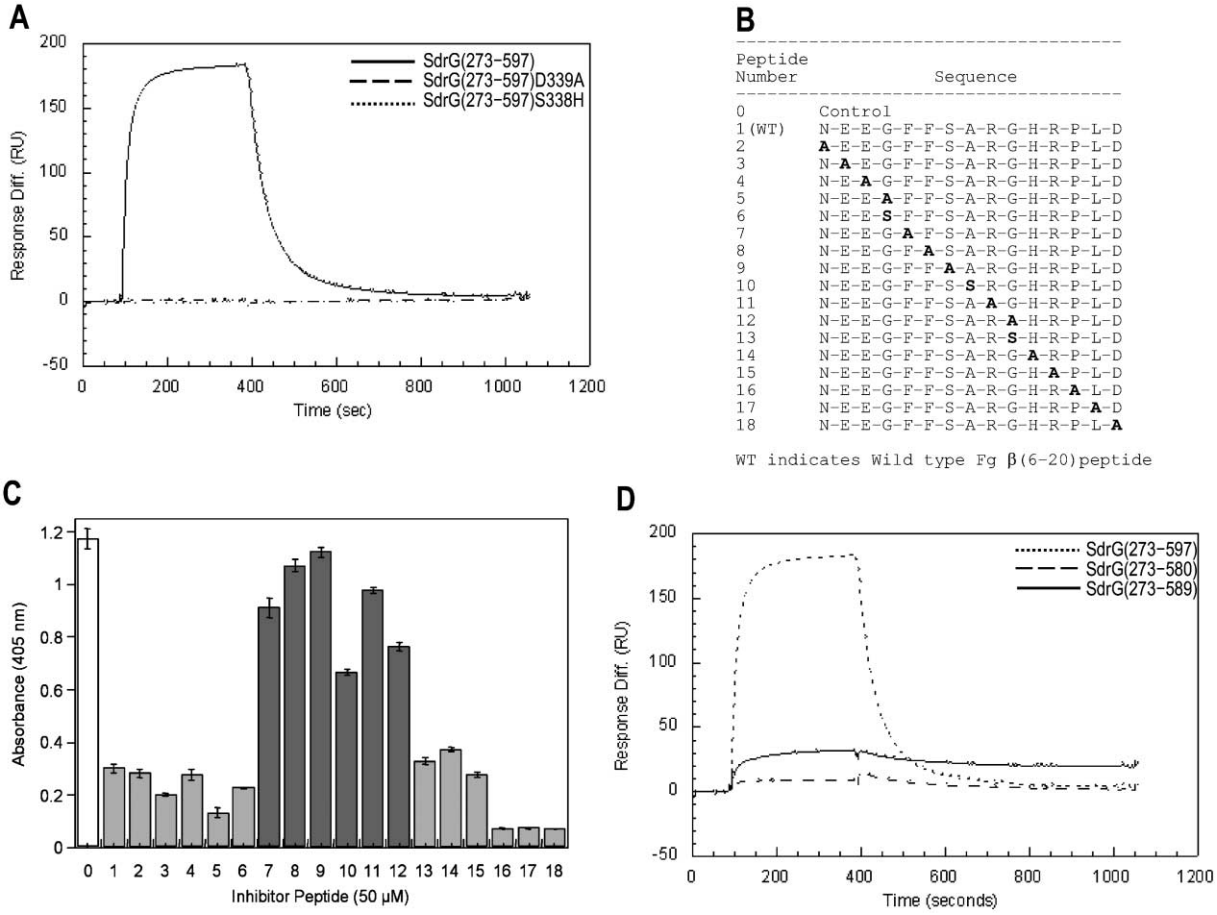


Figure 3. Mutational Analysis of rSdrG$_{(273-597)}$ and β6-20 Peptide

(A) Surface plasmon resonance (SPR) was used to analyze the interaction of wild-type and mutated versions of rSdrG$_{(273-597)}$ with the β1-25 peptide immobilized on a BIAcore sensor chip. Lines represent: (——), SdrG$_{(273-597)}$; (— –), SdrG$_{(273-597)}$D339A; (·····), SdrG$_{(273-597)}$S338H.

(B) Mutation of peptide β6-20 by amino acid replacement. Ala or Ser (**A** or **S**) was used to replace each residue within Fg β(6-20) in a panel of synthetic peptides. The Ala or Ser replacement residues are shown in bold.

(C) Mutant peptide inhibition of rSdrG$_{(273-597)}$ binding to Fg. 50 nM rSdrG$_{(273-597)}$ was preincubated with the different synthetic peptides (10 nM) for 1 hr at RT and then transferred to microtiter wells coated with 1 μg human Fg. After incubation for 1 hr at RT, bound SdrG was detected as described in the Experimental Procedures section. Peptides with residue replacements that affected inhibition (7-12) are shown in dark bars.

(D) The rSdrG$_{(273-597)}$ C terminus is essential for binding to Fg-derived peptides. Surface plasmon resonance shows the binding of rSdrG$_{(273-597)}$ (·· ··), rSdrG$_{(273-589)}$ (——) and rSdrG$_{(273-580)}$ (- - -) to the β1-25 peptide immobilized on a BIAcore sensor chip.

complex. To test these predictions, we constructed two point mutants by site-directed mutagenesis replacing residues Ser338 and Asp339, that interact with Phe11f and Ser12f of the peptide (Figure 2D), with His and Ala respectively. We predicted that, the S338H and D339A mutations would not only disturb the protein-peptide hydrogen bonds, but also that the S338H mutation would introduce a larger side chain within the binding region, occupying the space where the peptide docks, thus preventing ligand binding. We tested the ability of the rSdrG$_{(273-596)}$S338H and rSdrG$_{(273-596)}$D339A to bind to the immobilized β1-25 peptide in the BIAcore (Figure 3A) and to immobilized Fg in an ELISA-type assay (data not shown), with identical results. Both mutations completely abolished the ability of rSdrG$_{(273-596)}$ to bind to human Fg.

A closer inspection of the rSdrG$_{(276-596)}$-peptide complex suggests that the specificity of the peptide for rSdrG$_{(276-596)}$ is determined by the residues Phe10f, Phe11f, Ser12f, Ala13f, Arg14f, and Gly15f (Figure 2D). To further examine which residues within the Fg β-chain sequence (β6-20) are necessary for peptide binding to rSdrG$_{(273-597)}$, we synthesized a panel of mutant peptides (Figure 3B) replacing each residue in the native Fg β-chain sequence with an Ala (or Ser if Ala is present in the native sequence). The ability of the synthesized

(C) Schematic representation of rSdrG$_{(276-597)}$-peptide interactions. The β strands are colored purple and yellow for the peptide and the rSdrG molecule, respectively. Hydrogen bonds are indicated by dotted lines.

(D) Stereo close-up of ligand binding site showing (1) the interactions of residues Ser338 and Asp339 with the peptide (2) the binding pocket of Gly15f.

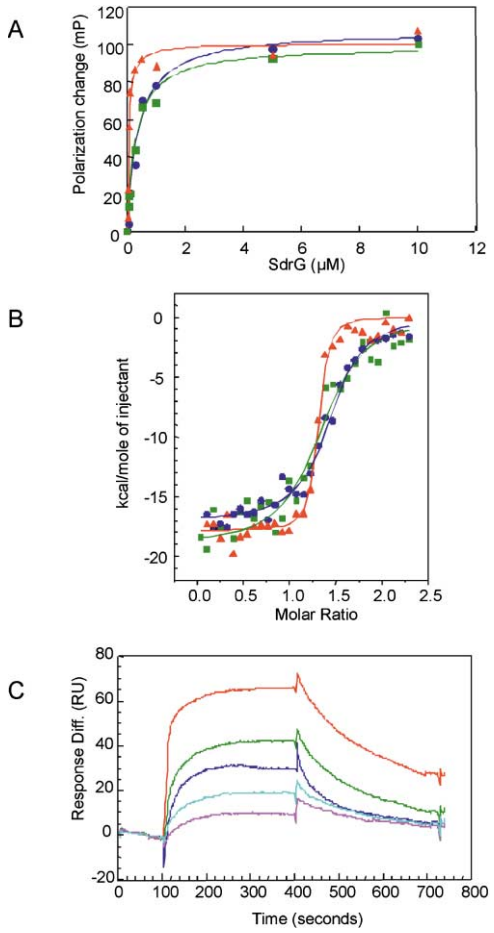


Figure 4. The β6-20 Peptide Is Not the Optimal SdrG $_{(273-597)}$ Ligand

(A) Binding of rSdrG$_{(273-597)}$ to fluorescein-labeled synthetic peptides quantitated by fluorescence polarization. Increasing concentrations of rSdrG$_{(273-597)}$ were incubated with synthetic fluorescein-labeled wild-type and mutant peptides (10 nM). The calculated $K_D$ for each interaction was: rSdrG-(β6-20) [●], 379 nM; rSdrG-(β6-20G15S) [■], 302 nM; rSdrG-[β6-20(F3)] [▲], 49 nM.

(B) Binding of SdrG$_{(273-597)}$ to unlabeled synthetic peptides quantitated by isothermal titration calorimetry. A 20 μM solution of SdrG$_{(273-597)}$ was incubated with an increasing concentration of synthetic peptides. The calculated $K_D$ for each interaction was: rSdrG-(β6-20) [●], 380 nM; rSdrG-(β6-20G15S) [■], 300 nM; rSdrG-[β6-20(F3)] [▲], 50 nM.

(C) Representative profiles of the relative SPR responses for the binding of synthetic peptides to rSdrG$_{(273-597)}$ immobilized on a BIAcore sensor chip. Colors represent: (Red), β6-20(F3); (Green), β6-20(G15S); (Blue), β6-20; (Cyan), β6-20(F11A); (Pink), β1-25 (scrambled).

mutant peptides to bind to rSdrG$_{(273-597)}$ was tested in an inhibition assay. Peptides number 7, 8, 9, 10, 11, and 12 were the least effective in inhibiting the binding of rSdrG$_{(273-597)}$ to immobilized human Fg, indicating that these peptides have a lower affinity for SdrG than the unmodified peptide. Using surface plasmon resonance we confirmed that peptide number 8 (F11A) binds less well to immobilized Fg compared to the native β6-20 peptide and only slightly better than a β1-25 scrambled sequence peptide (Figure 4C). These data indicate that the amino acids Phe10f, Phe11f, Ser12f, Ala13f, Arg14f and Gly15f, which are replaced in peptides number 7-12, significantly contribute to the binding to SdrG (Figure 3C). The peptide residues identified by this method coincides with those shown to contact rSdrG$_{(276-597)}$ in the cocrystal structure. Interestingly, these residues include Arg14f and Gly15f, the thrombin cleavage site of the FgBβ chain.

The crystal structure presented in Figure 1B was expected to show the complex of rSdrG$_{(276-596)}$ with its ligand, a peptide corresponding to residues 6–20 of the human Fg β-chain (Figure 2A). However, in the crystal structure solved, the bound peptide contains an additional Phe residue in between position 9 (Gly9f) and 10 (Phe10f) (Figure 2A) that is not present in the human Fg sequence. Therefore, the peptide present in the cocrystal is an analog of β6-20 containing three consecutive Phe residues [β6-20(F3)] instead of two. This additional Phe residue (Phe9f), along with the two Phe residues present in the native sequence (Phe10f and Phe11f), contributes to the hydrophobic stacking interactions with rSdrG$_{(276-596)}$. It is likely that the β6-20 peptide preparation utilized in the cocrystal complex formation contained a contaminating peptide with an additional Phe residue. Since this was the "selected" peptide in the crystallization process, the contaminating peptide is predicted to have a higher affinity for SdrG compared to a peptide representing the native sequence. We therefore synthesized the β6-20 and β6-20(F3) peptides, purified them, and examined their binding to rSdrG$_{(273-597)}$ by fluorescence polarization (FP) spectroscopy, isothermal titration calorimetry (ITC) and surface plasmon resonance (SPR) (Figures 4A, 4B, and 4C). The results of the FP and ITC experiments indicate that rSdrG$_{(273-597)}$ binds more tightly to the β6-20(F3) peptide compared to the β6-20 peptide, reflected by an almost 8-fold lower $K_D$.

Amino acid modeling at positions Ser12f, Ala13f and Gly15f suggests that larger side-chain residues would not be accommodated at these locations. Closer examination of the binding pockets of these residues revealed that the replacement of Gly15f with Ser would provide additional hydrogen bonds with Gln425, Gln579, and Ser437 in rSdrG$_{(276-596)}$ (Figure 2D). Thus, this variant peptide is predicted to have a higher affinity for SdrG compared to a peptide based on the wild-type Fg sequence. As expected, the G15fS mutation resulted in a peptide that exhibited increased binding affinities when compared to the native β6-20 peptide, both in the FP and the ITC assays (Figure 4B). Taken together, these results demonstrate that the human Fg residues bound by SdrG do not represent the amino acid sequence of the optimal ligand. Similarly, in the case of the *S. aureus* ClfA, we have found synthetic peptide analogs representing amino acid variants of its known ligand, the C terminus of the Fg γ-chain that show higher affinity for the MSCRAMM than the native Fg sequence (M.H., unpublished data). Thus, structurally related MSCRAMMs containing repeated immunoglobulin-like folds such as SdrG and ClfA can accommodate linear peptides with a certain degree of ligand sequence variability.

### SdrG Inhibits the Thrombin Cleavage of the Fg Bβ-Chain: a Structural Analysis

During blood clotting, thrombin cleaves the two Aα chains of Fg rapidly at Arg16f and the two Bβ-chains at

Arg14f more slowly to release fibrinopeptides A and B, respectively. Although the structure of the complex thrombin-fibrinopeptide B has not been reported, the existing structural (Malkowski et al., 1997; Martin et al., 1992, 1996), biochemical and mutational (Le Bonniec et al., 1991) data on complexes of thrombin-fibrinopeptide A and its analogs (Blomback, 1986; Scheraga, 1986; Lord et al., 1990; Binnie and Lord, 1993) show that the N-terminal Aα chain of Fg interacts with thrombin through its residues P10-P4′ (α8-20). In the present rSdrG$_{(276–597)}$-peptide complex structure, residues P7-P4′ (β9-18) of the peptide interact with rSdrG$_{(276–597)}$. Thus, we can now provide the structural explanation for the previously reported (Davis et al., 2001) SdrG interference with thrombin-mediated release of fibrinopeptide B, as the SdrG binding site overlaps with the thrombin cleavage site in the Fg Bβ chain.

In addition, the rSdrG$_{(276–597)}$-peptide complex structure provides an explanation for why SdrG can bind the Fg Bβ chain but not the structurally related Aα chain. In the N-terminal Aα(7-20) (DFLAEGGGV**RG**PRV-thrombin cleavage site shown in bold) the P2, P1 and P1′ positions are occupied by residues Val, Arg, and Gly, compared to Ala, Arg, and Gly in the Fg Bβ chain. These sets of residues are very similar and both easily would fit in the respective pockets of SdrG. However, the P5, P4, and P3 positions are occupied by Gly, Gly, Gly in the Aα chain and Phe, Phe, and Ser residues in the Bβ chain, respectively. These substitutions would result in the elimination of a critical hydrogen bond [Ser12f (OG) with Asp339 (OD2)] and the hydrophobic interactions that Phe10f and Phe11f make with rSdrG$_{(276–596)}$, giving the Aα chain a much lower binding affinity for rSdrG$_{(276–596)}$ compared to that of the Bβ chain.

### Structure of the apo-rSdrG$_{(276–597)}$ Protein

The structure of the apo-rSdrG$_{(273–596)}$ was solved by molecular replacement methods using the structure of the rSdrG$_{(276–596)}$ component of the peptide-MSCRAMM complex as a starting model. The apo-rSdrG$_{(276–596)}$ structure was solved at a 3.5 Å resolution, with four molecules in an asymmetric unit of a P1 cell. Structurally, the four copies are very similar, but each molecule shows several short disordered segments. As a result, the molecular model structures were truncated in up to six places. A high degree of disorder was also observed at the C terminus, where 12 to 14 of the terminal residues could not be modeled.

To compare the structures of rSdrG$_{(276–596)}$ as an apoprotein and in complex with the peptide, we chose the apoprotein structure that showed the best resolved C terminus out of four molecules available in the P1 cell. A structural comparison of rSdrG$_{(276–596)}$ as an apoprotein and in complex with the peptide (Figure 5A) shows an rms deviation of 1.1 Å for 276 Cα atoms and revealed that the major difference between the two forms is at their C termini. In the rSdrG$_{(276–596)}$ peptide-complex structure, the C-terminal region forms a regular β strand (G″) that is inserted between strands E and D of the N2 domain (Figure 5B). In the apo-rSdrG$_{(276–596)}$ structure (Figure 5C), the C terminus does not interact with the N2 domain, it extends into the solvent region. Therefore, the space between the D and E strands in the N2 domain is empty, leaving the binding cleft between the N2 and N3 domains widely open and presumably providing easy access for the ligand to dock within the binding cleft.

The disordered regions in all four molecules of apo-rSdrG$_{(276–596)}$ in the asymmetric unit are better defined in the rSdrG$_{(276–596)}$–peptide complex. For example, the region encompassing amino acids Lys350-Glu355 that is disordered in the apo structure forms the D strand in the complex structure, interacting with strand G″ of the N2 domain through hydrogen bonds. Similarly, Phe373-Glu381 region is stabilized by the hydrogen bonds formed with the N2 domain G″ strand. In addition, the residue Tyr552 of the connecting loop (Tyr545-Tyr552) between E′ and F′ strands makes key van der Waals contacts by stacking against Phe9af and Phe11f residues of the peptide. The structural rearrangements and the direct protein-ligand interactions formed upon peptide binding result in an effectively stabilized rSdrG$_{(276–596)}$ complex compared to its apoform.

### A Dynamic Model for Ligand Binding

The structural comparison of apo-rSdrG$_{(276–596)}$ and the rSdrG$_{(276–596)}$-peptide complex provides the basis for a ligand binding model. The open conformation of the rSdrG$_{(276–596)}$ apoprotein (Figure 5C) contrasts with the closed conformation (Figure 5B) observed for rSdrG$_{(276–596)}$ in the protein-ligand complex. In the protein-ligand complex, the C-terminal segment of the protein, which is an extension of the N3 domain, "locks" on to the "docked" peptide by forming a cover and sequesters it by "latching" on to the neighboring N2 domain. Thus, peptide binding to rSdrG$_{(276–596)}$ appears to involve multiple steps: initially, the peptide docks in the cleft. This binding event is followed by a structural rearrangement at the rSdrG$_{(276–596)}$ C terminus that begins with a deviation in the strand direction at Ile581. Because of this deviation, the strand crosses over the binding cleft and locks the peptide in place. This crossover results in the formation of backbone hydrogen bonds between the bound peptide and the covering segment (Ile581-Ser586) of the adhesin, securing the peptide in the binding pocket. The final step is the complementation of a β sheet in the N2 domain, inserting the C-terminal strand (G″) between strands E and D. This last latching event stabilizes the overall structure.

### Truncation Mutants Support the Binding Model

To examine the proposed dock, lock, and latch model of ligand binding, we constructed two truncation mutants: rSdrG$_{(273–589)}$ contains a midablated version of the N2 G″ C terminus β strand, and rSdrG$_{(273–580)}$ is completely devoid of the N2 G″ C terminus β strand and the connecting residues between the N2 G″ and the N3 G′ strand. We hypothesized that, given the importance of these C-terminal residues in ligand binding and structure stabilization, the rSdrG$_{(273–589)}$ polypeptide would show a decreased ability to bind to Fg, and that rSdrG$_{(273–580)}$ would be completely unable to bind to Fg. Our results show that, as predicted, the rSdrG$_{(273–589)}$ polypeptide has a markedly reduced affinity for the ligand, and that binding of rSdrG$_{(273–580)}$ to the peptide is nearly null when compared to the binding level of the intact rSdrG$_{(273–597)}$ (Figure 3D). These data strongly suggest that our predicted

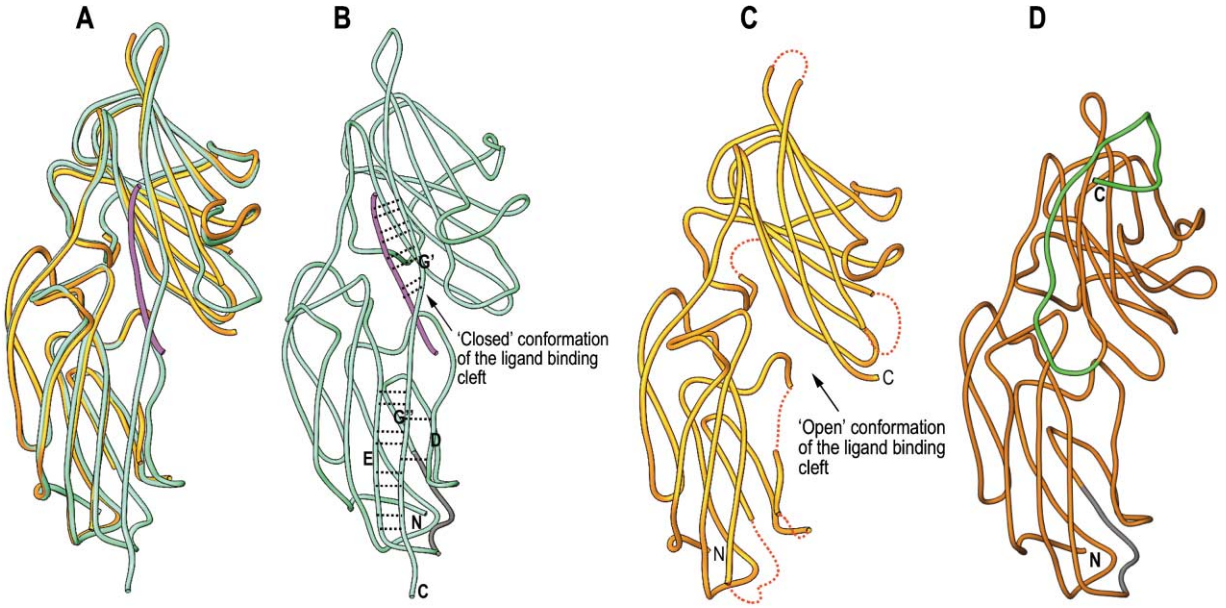


Figure 5. Backbone Representation and Comparison of the apo-rSdrG$_{(273–597)}$, rSdrG$_{(273–597)}$-Peptide Complex and apo-rClfA$_{(221–559)}$ Structures

(A) The superposition of Cα backbones of rSdrG$_{(273–597)}$-peptide (light green) with the apo-rSdrG$_{(273–597)}$ structure (gold).

(B) Representation of the "closed" rSdrG$_{(273–597)}$-peptide complex (light green and purple respectively) and (C) the "open" apo-rSdrG$_{(273–597)}$ conformation.

(D) The apo structure of rClfA$_{(221–559)}$ (light brown). The C terminus of rClfA$_{(221–559)}$ (green) loops back and folds into the N3 domain, partially blocking the proposed ligand binding cleft. The conserved TYTFTDYVD motif is highlighted in gray (B and D).

multistep model of ligand binding and strand complementation is correct.

### A Family of Structurally Related, Cell Wall-Anchored Proteins from Gram-Positive Bacteria

We have recently solved the structures of the ligand binding domains of the *S. aureus* Fg binding MSCRAMMs; rClfA$_{(221–559)}$ (Figure 5D) (Deivanayagam et al., 2002) and rClfB$_{(200–554)}$ (Champion Deivanayagam, Samuel Perkins, Evelyn Walsh, Timothy Foster, M.H., and S.V.L.N., unpublished data). Although these adhesins bind to different chains in Fg, the overall domain arrangement and the topology of the individual domains are strikingly similar. The structural data for rClfA$_{(221–559)}$ and rClfB$_{(200–554)}$ are available only for the apoprotein. However, both apoproteins contain a putative ligand binding cleft between the N2 and N3 domains and an empty latching cleft between the D and E strands of the N2 domain that could accommodate a complementing "guest" strand. In fact, in the rClfB$_{(200–554)}$ crystal structure, this space is occupied by the C-terminal end of a neighboring crystallographically related molecule that forms a G″ like strand and occupies the cleft between the E and D strands of the N2 domain (data not shown). Interestingly, the previously identified conserved sequence motif TYTFTDYVD (McCrea et al., 2000) constitutes the back of the latching cleft in rSdrG$_{(273–597)}$ (strand D2 and a short α-helix connecting it to the E strand in the N2 domain shown in gray color in Figure 5B). This conserved motif is found in a similar location in the rClfA$_{(221–559)}$ and rClfB$_{(200–554)}$ crystal structures. At the moment, we do not understand the precise function of this conserved sequence, but it is tempting to speculate that it is involved in the structural preservation of the latching cleft.

To investigate how widespread the dock, lock, and latch model might be among bacterial cell wall proteins, we searched for MSCRAMM candidates in the genomic sequences from five gram-positive pathogens. In the analyzed genomic sequences, the number of predicted cell wall anchored proteins varied from 8 (*S. mutants)* to 41 (*E. faecalis*) (Table 2). Computer-generated structural models suggest that IgG-like subdomains arranged in tandem are present in 22%–45% of the predicted cell wall anchored proteins. Furthermore, all of the *S. aureus* proteins containing repeated IgG-like subdomains have a TYTFTDYVD-like sequence in one of the repeats. These include the Fg binding domains found in ClfA, ClfB, FnbpA, and FnbpB, the collagen binding MSCRAMMs Cna, the bone sialoprotein binding protein Bbp and the SdrC-E proteins, for which ligands so far have not been identified. A TYTFTDYVD-like motif is also present in SdrG and SdrF from *S. epidermidis*, and in the first N-terminal IgG-like subdomain of two *E. faecalis* cell wall anchored proteins, including the collagen binding MSCRAMM Ace (Jouko Sillanpaa, Y.X., Sreedhar Nallapareddy, Barbara Murray, and M.H., unpublished data). A recent crystal structure determination of the Ace collagen binding domain suggests that this motif forms the base of a cleft into which a hypothetical latching strand could insert (K.P, Y.X., D.C., M.H., and S.V.L.N., unpublished data). Remnants of this motif are also found in several streptococcal proteins. The structures of these proteins have not yet been solved, and we do not know if the identified remnants of the TYTFTDYVD motif signify a latching cleft in one of the predicted β sheets.

Table 2. Summary of Predicted LPXTG-Motif Cell Wall-Anchored Proteins of Gram-Positive Organisms with Repeated Immunoglobulin-Like Fold

| Organism | Number of Predicted LPXTG-Motif Cell Wall Anchored Proteins | Number with Predicted Repeated Ig-Like Fold |
|---|---|---|
| *Enterococcus faecalis* | 41 | 9 |
| *Staphylococcus aureus* | 22 | 9 |
| *Staphylococcus epidermidis* | 11 | 5 |
| *Streptococcus mutans* | 8 | 3 |
| *Streptococcus pneumoniae* | 19 | 6 |

Additional details and methods used to generate this table are available as Supplemental Data (See Supplemental Table S2 available at http://www.cell.com/cgi/content/full/115/2/217/DC1).

Another structural feature conserved in the analyzed proteins is an amino acid pattern in their C-terminal latching strand sequences. Structural analysis of the SdrG latching sequence [SSGQGQG (586–592)] reveals that it is formed by alternating small residues (Ser586, Gly588, Gly590, and Gly592) with their short side chains facing the larger side chains (Tyr and Phe) of the TYTFT DYVD motif present in the opposite β sheet. The crystal structures of apo-rClfB (Champion Deivanayagam, Samuel Perkins, Evelyn Walsh, Timothy Foster, M.H., and S.V.L.N., unpublished data) and the collagen binding MSCRAMMs Ace (K.P., Y.X., D.C., M.H., and S.V.L.N., unpublished data) and Cna (Yinong Zong, Y.X., M.H., and S.V.L.N., unpublished data) show that their latching sequences also exhibit a pattern of alternating small residues. Modeling the C terminus of the apo-rClfA structure (Figure 5D) into the latching cleft of its N2 domain reveals a "GSGSGDG" motif that could occupy the latching cleft, following the pattern found in SdrG and other MSCRAMM structures. In addition, we identified potential latching sequences in putative MSCRAMM proteins from Staphylococci and Enterococci using homology modeling and sequence alignment methods and found that these sequences also follow a pattern of alternating small residues (Table 3). This observation, together with the conservation of a TYTFTDYVD-like motif in the back of the latching cleft suggest that the dock, lock, and latch mechanism is likely to represent a general mode of ligand capture by the A-domains of structurally related cell wall-anchored adhesins of gram-positive bacteria.

The salient features of the dock, lock, and latch model include a ligand binding-induced redirection of the C-terminal end of the N3 domain of SdrG, followed by an insertion of this redirected C-terminal extension between strands E and D in the preceding N2 domain. This molecular rearrangement represents an intramolecular β sheet complementation that is, to our knowledge, unique.

Intermolecular β strand complementation is not an unusual mechanism of protein-protein interaction. In this case, a strand from one protein complements a β sheet in the interacting protein. An example of this type of protein-protein association is the interaction of the PapD-like chaperons with pilus subunits during pili assembly in *E. coli* (Sauer et al. 1999; Choudhury et al., 1999). Recently, the ligand binding mechanism for fibronectin binding MSCRAMMs from gram-positive bacteria was elucidated (Schwarz-Linek et al., 2003). In this case, an unordered repeated motif of the MSCRAMM interacts with the set of type I modules of the N-terminal domain of fibronectin through a multiple β strand complementation event called a β-zipper mechanism. This intermolecular β strand complementation mechanism appears to be

Table 3. Structurally Related Ligand Binding Regions from the A-Domains of Gram-Positive MSCRAMMs Share Conserved Motifs

| MSCRAMM (Accession number) | TYTFTDYVD motif or its variant | Number of residues[a] | Latching Sequence[b] | Method[c] | Organism |
|---|---|---|---|---|---|
| SdrG (AAF72510) | TYTFTDYVD | 207 | **SSGQGQG** | 1a | *S. epidermidis* |
| SdrF (AAF72509) | TYTFTNYVD | 200 | **GSSTAQG** | 4 | *S. epidermidis* |
| ClfA (CAA79304) | IYTFTDYVN | 207 | **GSGSGDG** | 2 | *S. aureus* |
| ClfB (CAA12115) | TFVFTDYVN | 204 | **GGGSADG** | 1b | *S. aureus* |
| SdrC (CAA06650) | TYTFTNYVD | 196 | **GSSTANG** | 4 | *S. aureus* |
| SdrD (CAA06651) | TYTFTDYVD | 212 | **NQSGGAG** | 4 | *S. aureus* |
| SdrE (CAA06652) | TYTFTDYVD | 216 | **GGGDGTV** | 3 | *S. aureus* |
| Bbp (CAB75732) | TYKFTDYVD | 216 | **VVVTVSV** | 3 | *S. aureus* |
| FnbpA (CAA65106) | RYTFTNDIE | 210 | **NKANGNE** | 3 | *S. aureus* |
| FnbpB (CAA44726) | RYTFKEYVQ | 204 | **NNAQGDG** | 3 | *S. aureus* |
| Cna (Q53654) | TITFNDKVE | 200 | **ANAGIEG** | 1b | *S. aureus* |
| Ace (AAG23954) | VFIYKDHVV | 208 | **VEGEASG** | 1b | *E. faecalis* |

[a] Residues in between the TYTFTDYVD motif and the latching sequence.
[b] Alternate residues that are projected into the region between the two β-sheets and face the TYTFTDYVD motif are shown in bold.
[c] Latching sequence identified method: (1a) Crystal structure-this work. (1b) Crystal structures of A-region of collagen binding MSCRAMMS Ace, Cna and Fg- binding MSCRAMM ClfB (Champion Deivanayagam, Y.X., Timothy Foster, M.H., and S.V.L.N., unpublished data; K.P., Y.X., D.C., M.H., S.V.L.N., unpublished data; Yinong Zong, Y.X., M.H., and S.V.L.N., unpublished data). (2) Modeling the C-terminus of rClfA crystal structure. (3) Homology modeling. (4) Pairwise sequence alignment.

unique to fibronectin binding MSCRAMMs. The unordered repeats are so far only found in the C-terminal half of fibronectin binding MSCRAMMs and the N-terminal type I modules of fibronectin are, to this point, the only ligand that accommodates the MSCRAMM unordered repeats.

The dock, lock, and latch model describes a binding mechanism for the N-terminal A domains of the MSCRAMMs. These domains are composed of two or three IgG-like folds and conserved among a large family of related putative MSCRAMMs. In contrast to fibronectin, the ligands that bind to the IgG-like A-domains are linear short peptide sequences. This ligand group is large (and growing) and includes fibrinogen, elastin, keratin, bone sialoprotein, collagen, and laminin.

## Conclusion

The structure of the *S. epidermidis* adhesin SdrG in complex with a Fg-derived peptide presented here provides the atomic details of the interactions between a staphylococcal MSCRAMM and its host ligand molecule. Analysis of the crystal structures of the $SdrG_{(276–596)}$-peptide complex- and the apo-$SdrG_{(276–596)}$ allowed us to propose the dock, lock, and latch mechanism presented in this report.

*S. epidermidis* is an opportunistic pathogen. A common inhabitant of human skin and mucosal membranes, *S. epidermidis* has emerged as a major nosocomial pathogen associated with infected implanted medical devices (Huebner and Goldman, 1999) such as pacemakers, mechanical heart valves, artificial joints, and catheters (Costerton et al., 1999; Donlan, 2001). In biomaterial-centered infections, staphylococci colonize the surfaces of implanted biomaterials that are coated with host proteins such as Fg. These surfaces are often exposed to constant fluid shear forces. Therefore, strong interactions between the pathogen and the biomaterial are essential for bacterial adhesion and colonization. The now proposed dock, lock, and latch model of SdrG binding to Fg provides the molecular basis for the formation of a stable adhesin-ligand complex. This model describes the structural changes that stabilize the overall MSCRAMM-ligand complex, and would tightly secure the ligand in place under fluid shear forces. In addition, the presented structure reveals how SdrG blocks the thrombin cleavage site in Fg Bβ chain, inhibiting the release of fibrinopeptide B. It is possible that staphylococci express proteins that bind to Fg to prevent the release of chemotactic elements such as fibrinopeptide B, thus reducing the influx of phagocytic neutrophils to the infection site and enhancing the probability of bacterial survival in the host. Surprisingly, we found that the native Fg β(6-20) sequence is not the optimal sequence for SdrG binding, suggesting that multiple ligands could be accommodated in the MSCRAMM binding region.

We currently have positively identified the dock, lock, and latch model only for the interaction of SdrG with Fg. However, we have solved additional structures of related ligand binding domains, such as those from ClfA (Deivanayagam et al., 2002), ClfB (Champion Deivanayagam, Samuel Perkins, Evelyn Walsh, Timothy Foster, M.H., and S.V.L.N., unpublished data), Ace (K.P., Y.X., D.C., M.H., and S.V.L.N., unpublished data), and Cna (Yinong Zong, Y.X., M.H., S.V.L.N., unpublished data). All of these MSCRAMMs contain a cleft in one of the IgG-like folds that could accommodate a latching strand. A survey of the available genomes of five pathogenic gram-positive bacteria revealed that all organisms encode putative cell wall anchored proteins predicted to contain repeated, in tandem IgG-like folded domains. Furthermore, variants of the TYTFTDYVD motif, which appears to be associated with the latching cleft and putative latching sequences are also present in some of these cell wall anchored proteins from gram-positive bacteria. We therefore propose that the dock, lock, and latch model of ligand binding is a common theme among these proteins.

## Experimental Procedures

### Purification, Crystallization, and Data Collection

The $rSdrG_{(273–597)}$ polypeptide was expressed in *E. coli* and purified by immobilized metal chelate affinity chromatography, followed by anion-exchange and gel-filtration chromatography, as previously described (Davis et al., 2001). The $rSdrG_{(273–597)}$ protein was mixed with excess peptide (Peptide: protein ratio of 1.5 or greater) and screened for crystallization conditions. Crystals were obtained by hanging drop method using 50% PEG 2000 Mono Methyl Ether (MME), 0.2 M NaCl and 50 mm buffer (Tris, HEPES, MES, Pipes, Cacodylate) between pH 5.5–7.5. The crystals belong to the orthorhombic $P2_12_12_1$ space group and diffracted to 1.86 Å. The crystals of apo-$rSdrG_{(276–597)}$ were obtained at similar conditions, and, however, belong to the triclinic P1 space group. The X-ray diffraction data for both apo- and complex crystals were collected at 100 K using in house X-ray source. DENZO/SCALEPACK package (Otwinowski and Minor, 1997) was used for diffraction data processing.

### Phase Determination, Model Building, and Refinement

#### *$rSdrG_{(276–597)}$-Peptide Complex*

The $rSdrG_{(276–597)}$-peptide complex was solved by multiple isomorphous replacement method using two derivatives such as Trimethyl Lead acetate (TMLA) and platinum. SOLVE (Terwilliger and Berendzen, 1999) program was used for the refinement of heavy atom sites and subsequent phase calculations. The initial solvent flattened electron density map revealed excellent density allowing the construction of a molecular model for both copies of the molecule in the asymmetric unit. QUANTA (Molecular Simulations, Inc) was used for tracing and model building and CNS (Brunger, 1998) for refinement. The refinements were monitored throughout using free R values and the model was refined to final R factor/$R_{free}$ of 0.209/0.221 using all reflections from 20 to 1.86 Å. The side chain density for few residues (in complex I: Lys320, Lys379, Asp474, Glu475, Lys544, Leu594, and Pro596 and in complex II: Lys379, Gln519, Lys549, Glu567, Tyr568, and Leu594) were not clear and they were modified to alanines. Similarly, the peptide residue Phe9af in complex I and Arg17f in complex II were truncated to alanines due to lack of side chain density. Final refinement statistics are summarized in Table 1.

#### *Apo-$rSdrG_{(276–597)}$*

The structure of peptide free $rSdrG_{(276–597)}$ was solved by the molecular replacement method using Amore (Navaza, 1994), refined with CNS and the model built with QUANTA. The crystal structure of $rSdrG_{(276–597)}$-peptide complex (without the bound peptide) was used as the starting model. The first round of refinement, using 20–3.5 Å resolution data reduced R factor to 0.287 and $R_{free}$ to 0.345. A few cycles of further refinement and rebuilding lowered the R factor and $R_{free}$ to 0.250 and 0.319 respectively. A summary of refinement statistics is presented in Table 1.

### Synthetic Peptides

The synthesis and purification procedures of the peptides were described previously (Davis et al., 2001). For the following peptides the amino acid residue numbers are given and the sequence follows:

peptide β1-25, composed of first 25 residues of the N terminus of the Bβ chain of Fg (QGVNDNEEGFFSARGHRPLDKKREE); peptide β6-20 (NEEGFFSARGHRPLD); peptide β6-20(F3) (NEEGFFFSARG HRPLD); a scrambled version of peptide β1-25 (FSERKDLHQGEGN PREFVENDAKGR).

### Surface Plasmon Resonance Spectroscopy

The β1-25 Fg peptide was diluted in 10 mM formate [pH 3.5] and immobilized on a CM5 sensorchip (BIAcore, Uppsala, Sweden). The immobilization level was 430 RU. The recombinant proteins were injected on the peptide-coated surface at 25 μl/min (in HBS, 10 mM HEPES [pH 7.5], 150 mM NaCl, and 25°C) and the binding and dissociation was registered in a BIAcore 3000 instrument. The recombinant $SdrG_{(273–597)}$ was diluted to 50 μg/ml in 10 mM Na Acetate [pH 4.0] and immobilized in a CM5 sensorchip. The immobilization level was 1560 RU. The synthetic peptides, diluted to 0.1 μM in HBS, were injected on the rSdrG-coated surface at 150 μl/min and the binding and dissociation was registered in a BIAcore 3000 instrument.

### Fluorescence Spectroscopy Assays

The binding of rSdrG to immobilized Fg was analyzed in a solid phase assay as previously described (Davis et al., 2001). The binding of rSdrG to fluorescein-labeled N-terminal β chain peptide analogs was quantitated by fluorescence polarization, and this method, along with the peptide synthesis and labeling procedures have been previously described (Davis et al., 2001).

### Isothermal Titration Calorimetry

Isothermal titration calorimetry (ITC) experiments were carried out with a VP-ITC microcalorimeter (MicroCal Inc., Northampton, MA, USA) at 25°C. In a typical experiment the cell contained 20 μM rSdrG and the syringe contained 200 μM peptide. Both solutions were in HBS buffer (10 mM HEPES, 150 mM NaCl, [pH 7.4]) and were degassed at 25°C for 20–30 min. The titration was performed as follows: one preliminary injection of 5 μl and 29 injections of 10 μl with an injection speed of 0.5 μl/sec. The stirring speed was 540 rpm and the delay time between the injections was 5 min. To take into account the heats of dilution, two blank titrations were performed: one injecting peptide into buffer and another injecting buffer into the rSdrG solution. The averaged heats of dilution were subtracted from the main experiment. Data were analyzed using MicroCal Origin software (version 5.0) fitting them to a single binding site model.

### Identification of Putative Cell Wall-Anchored Proteins of Gram-Positive Bacteria

Genome sequences used were: *E. faecalis* V583 (www.tigr.org); *S. aureus* COL (www.tigr.org); N315 and Mu50 (www.ncbi.nlm.nih.gov); *S. aureus* MSSR (www.sanger.ac.uk); *S. epidermidis* RP62A (www.tigr.org); *S. mutans* UA159 (www.ouhsc.edu); and *S. pneumoniae* ATCC BAA-334 (www.tigr.org). For genomes that have not been annotated, a custom-designed algorithm was used to predict open reading frames (ORFs) and translate the ORFs into amino acid sequences. Glimmer 2.0 obtained from TIGR was used for predicting the open reading frames. The open reading frames were translated into amino acid sequences using an in house translation program. The amino acid sequences were formatted into Blast searchable database using formatdb from NCBI. Annotated genomes were formatted using formatdb. LPXTG-motif containing proteins were identified using PHI-blast, obtained from NCBI. The PHI-blast search uses a degenerate LPXTG pattern L-P-X-[TSA]-[GANS], X being any amino acid. The output sequences were further analyzed to select proteins containing typical features of LPXTG-motif containing cell wall-anchored proteins. These include a signal peptide at the N terminus (predicted using SignalP server at www.cbs.dtu.dk/services/SignalP/), the LPXTG-motif close to the C terminus (visual examination) followed by a hydrophobic transmembrane segment (predicted using TMHMM server at www.cbs.dtu.dk/services/TMHMM-2.0/) and several positively charged residues at the C terminus (visual examination). For annotated genomes, the term "LPXTG" or "cell wall" was used in searching the genomes for proteins annotated as cell wall anchored proteins. A detailed account of these studies will be described elsewhere.

### Homology Modeling

Homology modeling was carried out using the WHAT IF server (http://www.cmbi.kun.nl/gv/servers/WIWWWI/). Initially, the pairwise sequence alignment of template and target sequences was performed using the server Protein Information Resource (PIR) (http://pir.georgetown.edu/pirwww/search/pairwise.html). The aligned sequences, along with the template coordinate, were submitted to the WHAT IF server and the homology model were obtained. The insertions in the loop regions were modeled using the server CODA (http://www-cryst.bioc.cam.ac.uk/coda/). The short contacts side chain contacts were manually adjusted and the final model was energy minimized using CNS (Brunger, 1998). The stereochemistry of the model was evaluated with the program PROCHECK (Laskowski et al., 1993) and the residue environment was analyzed with the VERIFY_3D program (Luthy et al., 1992). Qualitative statistics of the models are presented in Supplemental Table S1 available at http://www.cell.com/cgi/content/full/115/2/217/DC1.

### Supplemental Data

Details of bacterial strains, polymerase chain reaction amplification of the *sdrG* gene fragments, site-directed mutagenesis and purification of recombinant proteins are provided in the Supplemental Data, available at available at http://www.cell.com/cgi/content/full/115/2/217/DC1.

### Acknowledgments

We thank Agneta Höök and Xiaowen Liang for their expertise in the development of the BIAcore assays shown in this report. This work was supported by NIAID, National Institutes of Health Grant AI20624 (to M.H.), NASA Cooperative agreement NCC8-246 (to S.V.L.N), NIH Postdoctoral Fellowship Award AI10629 (to M.G.B.), and by a grant from Inhibitex Inc.

Received: April 7, 2003
Revised: September 2, 2003
Accepted: September 23, 2003
Published: October 30, 2003

### References

Binnie, C.G., and Lord, S.T. (1993). The fibrinogen sequences that interact with thrombin. Blood *81*, 3186–3192.

Blomback, B. (1986). Specificity of thrombin and its action on fibrinogen. Ann. N Y Acad. Sci. *485*, 120–123.

Brunger, A.T. (1998). Crystallography and NMR system: a new software system for macromolecular structure determination. Acta Crystallogr. D. Biol. Crystallogr. *45*, 905–921.

Carson, M. (1997). RIBBONS. Methods Enzymol. *277*, 493–505.

Choudhury, D., Thompson, A., Stojanoff, V., Langermann, S., Pinkner, J., Hultgren, S.J., and Knight, S.D. (1999). X-ray structure of the FimC-FimH chaperone-adhesin complex from uropathogenic *Escherichia Coli*. Science *285*, 1061–1066.

Costerton, J.W., Stewart, P.S., and Greenberg, E.P. (1999). Bacterial biofilms: a common cause of persistent infections. Science *284*, 1318–1322.

Davis, S.L., Gurusiddappa, S., McCrea, K.W., Perkins, S., and Hook, M. (2001). SdrG, a fibrinogen-binding bacterial adhesin of the microbial surface components recognizing adhesive matrix molecules subfamily from *Staphylococcus epidermidis*, targets the thrombin cleavage site in the Bβ chain. J. Biol. Chem. *276*, 27799–27805.

Deivanayagam, C.C.S., Perkins, S., Danthuluri, S., Owens, R.T., Bice, T., Nanavathy, T., Foster, T.J., Höök, M., and Narayana, S.V.L. (1999). Crystallization of ClfA and ClfB fragments: the fibrinogen-binding surface proteins of *Staphylococcus aureus*. Acta Crystallogr. D. Biol. Crystallogr. *55*, 554–556.

Deivanayagam, C.C.S., Wann, E.R., Chen, W., Carson, M., Rajashankar, K.R., Hook, M., and Narayana, S.V.L. (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.

Donlan, R.M. (2001). Biofilms and device-associated infections. Emerg. Infect. Dis. *7*, 277–281.

Foster, T.J., and Hook, M. (1998). Surface protein adhesins of *Staphylococcus aureus*. Trends Microbiol. *6*, 484–488.

Huebner, J., and Goldman, D.A. (1999). Coagulase-negative staphylococci: role as pathogens. Annu. Rev. Med. *50*, 223–236.

Hussain, M., Becker, K., Eiff, C.V., Schrenzel, J., Peters, G., and Herrmann, M. (2001). Identification and characterization of a novel 38.5-kilodalton cell surface protein of *Staphylococcus aureus* with extended-spectrum binding activity for extracellular matrix and plasma proteins. J. Bacteriol. *183*, 6778–6786.

Jönsson, K., McDevitti, D., McGavin, M.H., Patti, J.M., and Hook, M. (1995). *Staphylococcus aureus* expresses a major histocompatibility complex class II analog. J. Biol. Chem. *270*, 21457–21460.

Kay, A.B., Pepper, D.S., and McKenzie, R. (1974). The identification of fibrinopeptide B as a chemotactic agent derived from human fibrinogen. Br. J. Haematol. *24*, 669–677.

Laskowski, R.A., MacArthur, M.W., Moss, D.S., and Thorton, J.M. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. J. Appl. Crystallogr. *26*, 283–291.

Le Bonniec, B.F., MacGillvray, R.T.A., and Esmon, C.T. (1991). Thrombin Glu-39 restricts the P′3 specificity to nonacidic residues. J. Biol. Chem. *266*, 13796–13803.

Lord, S.T., Byrd, P.A., Hede, K.L., Wei, C., and Colby, T.J. (1990). Analysis of fibrinogen Aα-fusion proteins. Mutants which inhibit thrombin equivalently are not equally good substrates. J. Biol. Chem. *265*, 838–843.

Luthy, R., Bowie, J.U., and Eisenberg, D. (1992). Assessment of protein models with three-dimensional profiles. Nature *356*, 83–85.

Malkowski, M.G., Martin, P.D., Lord, S.T., and Edwards, B.F. (1997). Crystal structure of fibrinogen-Aα peptide 1–23 (F8Y) bound to bovine thrombin explains why the mutation of Phe-8 to tyrosine strongly inhibits normal cleavage at Arg-16. Biochem. J. *326*, 815–822.

Martin, P.D., Robertson, W., Turk, D., Huber, R., Bode, W., and Edwards, B.F. (1992). The structure of residues 7–16 of the Aα-chain of human fibrinogen bound to bovine thrombin at 2.3-Å resolution. J. Biol. Chem. *267*, 7911–7920.

Martin, P.D., Malkowski, M.G., DiMaio, J., Konishi, Y., Ni, F., and Edwards, B.F.P. (1996). Bovine thrombin complexed with an uncleavable analog of residues 7–19 of fibrinogen A: geometry of the catalytic triad and interactions of the P1′, P2′, and P3′ substrate residues. Biochemistry *35*, 13030–13039.

McAleese, F.M., Walsh, E.J., Sieprawska, M., Potempa, J., and Foster, T.J. (2001). Loss of clumping factor B fibrinogen binding activity by *Staphylococcus aureus* involves cessation of transcription, shedding and cleavage by metalloprotease. J. Biol. Chem. *276*, 29969–29978.

McCrea, K.W., Hartford, O., Davis, S., Ní Eidhin, D., Lina, G., Speziale, P., Foster, T.J., and Höök, M. (2000). The serine-aspartate repeat (Sdr) protein family in *Staphylococcus epidermidis*. Microbiol. *146*, 1535–1546.

Navaza, J. (1994). AmoRe and automated package for molecular replacement. Acta Crystallogr. A *50*, 157–163.

Otwinowski, Z., and Minor, W. (1997). Processing of X-ray diffraction data collected in oscillation mode. Methods Enzymol. *276*, 307–326.

Palma, M., Haggar, A., and Flock, J.-I. (1999). Adherence of Staphylococcus aureus is enhanced by an endogenous secreted protein with broad binding activity. J. Bacteriol. *181*, 2840–2845.

Palma, M., Shannon, O., Quezada, H.C., Berg, A., and Flock, J.-I. (2001). Extracellular fibrinogen-binding protein, Efb, from *Staphylococcus aureus* blocks platelet aggregation due to its binding to the α-chain. J. Biol. Chem. *276*, 31691–31697.

Patti, J.M., and Hook, M. (1994). Microbial adhesins recognizing extracellular matrix macromolecules. Curr. Opin. Cell Biol. *6*, 752–758.

Perkins, S., Walsh, E.J., Deivanayagam, C.C., Narayana, S.V., Foster, T.J., and Hook, M. (2001). Structural organization of the fibrinogen-binding region of the clumping factor B MSCRAMM of *Staphylococcus aureus*. J. Biol. Chem. *276*, 44721–44728.

Richards, M.J., Edwards, J.R., Culver, D.H., and Gaynes, R.P. (2000). Nosocomial infections in combined medical-surgical intensive care units in the United States. Infect. Control Hosp. Epidemiol. *21*, 510–515.

Richardson, D.L., Pepper, D.S., and Kay, A.B. (1976). Chemotaxis for human monocytes by fibrinogen-derived peptides. Br. J. Haematol. *32*, 507–513.

Sauer, P.G., Futterer, K., Pinkner, J.S., Dodson, K.W., Hultgren, S.J., and Waksman, G. (1999). Structural basis of chaperone function and pilus biogenesis. Science *285*, 1058–1061.

Scheraga, H.A. (1986). Chemical basis of thrombin interactions with fibrinogen. Ann. N Y Acad. Sci. *485*, 124–133.

Schwarz-Linek, U., Werner, J.M., Pickford, A.R., Gurusiddappa, S., Kim, J.H., Pilka, E.S., Briggs, J.A., Gough, T.S., Hook, M., Campbell, I.D., and Potts, J.R. (2003). Pathogenic bacteria attach to human fibronectin through a tandem β-zipper. Nature *423*, 177–181.

Senior, R.M., Skogen, W.F., Griffin, G.L., and Wilner, G.D. (1986). Effects of fibrinogen derivatives upon the inflammatory response. Studies with human fibrinopeptide B. J. Clin. Invest. *77*, 1014–1019.

Skogen, W.F., Senior, R.M., Griffin, G.L., and Wilner, G.D. (1988). Fibrinogen-derived peptide Bβ 1–42 is a multidomained neutrophil chemoattractant. Blood *71*, 1475–1479.

Terwilliger, T.C., and Berendzen, J. (1999). Automated structure solution for MIR and MAD. Acta Crystallogr. D. Biol. Crystallogr. *55*, 849–861.

## Accession Numbers

Coordinates have been deposited with the Protein Data Bank (accession numbers 1R17 and 1R19).