Biochemistry and Chemical Biology
Structural Biology and Molecular Biophysics

Structural foundation for the role of enterococcal PrgB in conjugation, biofilm formation and virulence

Reviewed Preprint

Revised by authors after peer review.

About eLife's process

Reviewed preprint version 2
October 4, 2023 (this version)

Reviewed preprint version 1
March 3, 2023

Posted to bioRxiv
November 21, 2022

Sent for peer review
November 1, 2022

Wei-Sheng Sun, Lena Lassinantti, Michael Järvå, Andreas Schmitt, Josy ter Beek, Ronnie P-A Berntsson

Department of Medical Biochemistry and Biophysics, Umeå University, SE-90187 Umeå, Sweden • Wallenberg Centre for Molecular Medicine, Umeå University, Umeå, Sweden

Abstract

Type 4 Secretion Systems are a main driver for the spread of antibiotic resistance genes and virulence factors in bacteria. In Gram-positives, these secretion systems often rely on surface adhesins to enhance cellular aggregation and mating pair formation. One of the best studied adhesins is PrgB from the conjugative plasmid pCF10 of *Enterococcus faecalis*, which has been shown to play major roles in conjugation, biofilm formation and importantly also in bacterial virulence. Since *prgB* orthologs exist on a large number of conjugative plasmids in various different species, this makes PrgB a model protein for this widespread virulence factor. After characterizing the polymer adhesin domain of PrgB previously, we here report the structure for almost the entire remainder of PrgB, which reveals that PrgB contains four immunoglobulin-like domains. Based on this new insight, we re-evaluate previously studied variants and present new *in vivo* data where specific domains or conserved residues have been removed. For the first time, we can show a decoupling of cellular aggregation from biofilm formation and conjugation in *prgB* mutant phenotypes. Based on the presented data, we propose a new functional model to explain how PrgB mediates its different functions. We hypothesize that the Ig-like domains act as a rigid stalk that presents the polymer adhesin domain at the right distance from the cell wall.

eLife assessment

This study presents **valuable** structural data for the bacterial adhesin PrgB, an atypical microbial cell surface-anchored polypeptide that binds DNA. There is **convincing** support for the claims regarding the overall function and importance of individual domains, which integrate a wide range of new and previously published experimental data. The structure-based model of PrgB molecular activity will be impactful in the field of bacterial adhesins, conjugation, and biofilm formation, especially because it focuses on a clinically relevant Gram-positive pathogen, whereas most work in the field has been focused on Gram-negative model systems.

Introduction

Enterococcus faecalis (*E. faecalis*) is one of the leading causes of hospital acquired infections, such as urinary tract infections and endocarditis[1, 2]. These infections are difficult to treat as *E. faecalis* has the tendency to form biofilms and is often resistant to various antibiotics. They are also notorious for spreading antibiotic resistance and other fitness advantages by transfer of mobile genetic elements (MGEs), which can be located on conjugative plasmids or in the chromosome[3, 4]. Conjugative plasmids usually also encode a Type 4 Secretion System (T4SS) that mediates its transfer, via conjugation, from a donor cell to a recipient cell[5–8]. However, conjugative plasmids and their T4SS have almost exclusively been studied in Gram-negative model systems[8].

One of the few well-characterized Gram-positive conjugative plasmids is pCF10 from *E. faecalis*[6, 9, 10]. This conjugative plasmid contains a ~27 kbp operon that is tightly regulated by the P_Q promoter[11–14] and that encodes all proteins needed for conjugation. This operon also encodes three cell-wall anchored proteins: PrgA, PrgB, and PrgC. PrgA is a conjugation regulator that provides surface exclusion to prevent unwanted conjugation. We have previously shown that PrgA consists of a protease domain that is presented far away from the cell wall via a long stalk and that it's likely mediating the proteolytic cleavage of PrgB[15, 16]. PrgC is a virulence factor, but its function and structure remain unknown[17]. PrgB is the main adhesin produced by pCF10 and has been studied for well over 3 decades. This protein is around 140 kDa in size and possesses an N-terminal signal sequence and a C-terminal LPXTG cell wall anchor motif[18]. PrgB, which is indicated to function as a dimer *in vivo*[19], distributes over the entire surface of the cell-wall and increases cellular aggregation, biofilm formation and the efficiency of plasmid transfer[20, 21]. Several mammalian infection model systems have shown that PrgB is a strong virulence factor[18, 22–27]. One reason for this virulence is that PrgB mediates biofilm formation in an extracellular DNA (eDNA) dependent manner[17]. Homologs of PrgB have been identified in many other conjugative plasmids[16], suggesting that PrgB-like proteins confer important roles in a large number of bacterial species[28–30].

PrgB was initially identified as one of the driving forces in cellular aggregation[20]. To understand how it mediated this process, previous research has tried to identify the various protein domains that are present in PrgB and evaluate their function(s). Two RGD (Arg-Gly-Asp) motifs were identified (see **figure 1A**) and found to be important for vegetation and biofilm formation in the host tissue environment[18, 24]. The N-terminal half of PrgB was found to be required for aggregation and to bind lipoteichoic acid (LTA)[30–32], which is a major constituent of the cell-wall in Gram-positive bacteria. In 2018, we solved the structure of $PrgB_{246\text{-}558}$ and showed that it has a lectin-like fold that was most similar to adhesins from various oral Streptococci. As these adhesins are known to bind various polymers, we subsequently referred to $PrgB_{246\text{-}558}$ as the polymer adhesin domain[16]. We have shown that this domain can bind both LTA and eDNA in a competitive manner. Bound eDNA is thereby strongly compacted as it is wrapped around the domain's positively charged surface[19]. We therefore proposed that PrgB could use eDNA to promote cell-to-cell contacts, as an alternative to direct binding to the LTA from a recipient cell (**Fig. 1B**). As all described polymer adhesin domains, PrgB has a central ridge with a conserved cation binding site[16, 33]. In the homologous GbpC, from *Streptococcus mutans*, this site has been suggested to bind glucans[34]. However, no interaction with glucans has been observed for PrgB or any other homologs[35]. Thus, the importance of this conserved motif remains an open question.

The polymer adhesin domain plays an important role in the function of PrgB, but it only accounts for around a quarter of the entire protein. Here we present the structure of almost the entire remainder of PrgB. This allows us to put the large amount of available phenotypic data into a structural context and explain a lot of previous observations. We also constructed several new

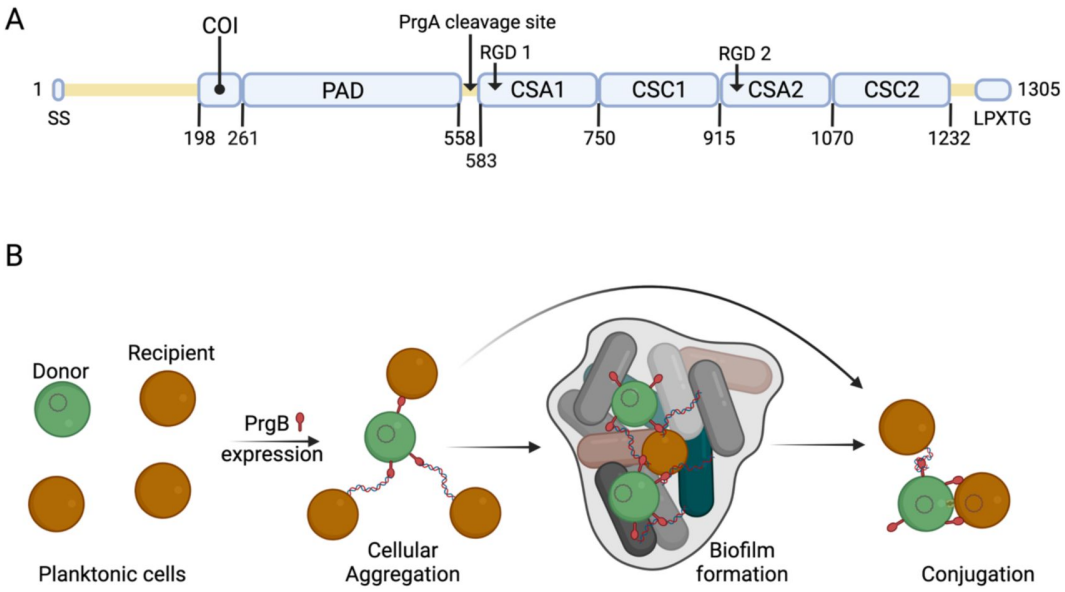

Figure 1.

Schematic overview of PrgB domain organization and function. A) Updated schematic overview of the domain organization of PrgB. SS: Signal sequence, COI: Coiled-coil, PAD: Polymer adhesin domain, CSA: adhesin isopeptide-forming adherence domain, CSC: cell-surface antigen C-terminal domain, LPXTG: cell wall anchor sequence. PrgA cleavage site is located between the polymer adhesin domain and the first Ig-like domain, and has the sequence IFNYGNPKEP. B) In a setting with a donor cell (green) and multiple recipient cells (brown), PrgB is produced and sits on the cell wall. There it enhances cellular aggregation and/or biofilm formation, either by directly binding LTA from the cell-wall of a recipient or by binding first to eDNA. PrgB compacts the eDNA, and thereby likely pulls the recipient cells closer. Once close enough, PrgB binds to the LTA of the recipient bacteria and facilitates mating-pair formation and conjugation.

mutants of *prgB* that better fitted the found domain organization and analyzed their *in vivo* effects on cellular aggregation, biofilm formation and conjugation efficiency. Based on our findings, we conclude with an updated model of how PrgB mediates its different functions.

Results

$PrgB_{584-1233}$ contains 4 immunoglobulin-like domains

Previous bioinformatics and structural analysis of PrgB proposed that PrgB consisted of 3 domains; the previously crystallized polymer adhesin domain responsible for eDNA and LTA binding, and two domains with RGD (Arg-Gly-Asp) motifs[19]. However, when we reanalyzed the PrgB sequence with the new structure-prediction tools that are available, such as AlphaFold[36], it became clear that this model was partially incorrect. PrgB seems to consist of an N-terminal disordered region (residues 35-197), followed by a newly identified coiled-coil (COI) domain (residues 198-257), the previously crystallized polymer adhesin domain (residues 261-558), a linker region (residues 559-582), 4 immunoglobulin (Ig)-like domains (residues 583-1232) and finally the C-terminal disordered region containing the LPXTG motif (residues 1263-1305) that gets anchored to the cell wall (**Fig. 1A**). The Ig-like domains seem to come in pairs of two slightly different structures, denoted as CSA1-CSC1 (first pair) and CSA2-CSC2 (second pair), as named in InterPro (CSA from IPR026345; adhesin isopeptide-forming adherence domain, and CSC from IPR032300; cell-surface antigen C-terminal). To verify this updated domain organization of PrgB, we set out to experimentally determine its structure using a combination of X-ray crystallography and cryo-EM methods.

As described in Schmitt *et al*[19], we were not able to produce full-length PrgB in *E. coli*, but instead produced and purified $PrgB_{188-1233}$. This version of the protein only lacks the disordered N-terminal region and the LPXTG anchor and elutes from size exclusion chromatography in two peaks corresponding to dimeric and a monomeric PrgB. PrgB is in a monomer-dimer equilibrium and the dimer has been described as the main biologically functional unit *in vivo*[7]. The monomeric fraction (**Fig. 1 – figure supplement 1**) was successfully used for crystallization trials. Crystals belonging to space group $P2_12_12_1$ appeared after 8-12 weeks, diffracted to 1.85 Å and contained 2 molecules in the asymmetric unit. The crystallographic phase problem was solved using molecular replacement with SspB (PDB: 2WOY) as a search model. Surprisingly, the resulting electron density lacked the previously crystallized polymer adhesin domain of PrgB, and instead only contained residues 584-1233. There is also no space in the crystal packing to allow for a flexible polymer adhesin domain, so it was likely cleaved off in the crystallization drop before the crystals were formed. The modelled protein indeed consists of four Immunoglobulin (Ig)-like domains: two CSA and two CSC domains (**Fig. 2A**). Previous bioinformatics analysis showed that various homologous adhesin proteins contain different numbers of Ig-like domains[16], but a DALI[37] analysis of $PrgB_{584-1233}$ showed that there is no previously solved structure in the PDB that contains four of these Ig-domains coupled together. There are, however, homologous structures available with either 2 or 3 Ig-domains. The closest structural homologs are the C-terminal parts of Antigen I/II proteins from oral Streptococci, e.g. the surface protein AspA from *Streptococcus pyogenes*[38] (PDB code: 4OFQ), which has 3 Ig-domains and an r.m.s.d. to the Ig-like domains from PrgB of 3.2 Å over 337 residues, or the BspA protein from *Streptococcus agalactiae*[39](PDB code: 4ZLP) which has 2 Ig-domains and an r.m.s.d. of 3.3 Å over 334 residues (see Supplementary file 3 for an overview of the highest ranked DALI hits). The r.m.s.d. decreases to ca. 1 Å if the individual Ig-like domains are superimposed upon each other.

In each of the Ig-like domains of PrgB, isopeptide bonds are formed between a lysine and an asparagine, a bond which is further stabilized by an aspartic acid (**Fig. 2B**). This feature is also present in various homologous Ig-like domains from Antigen I/II proteins[38], [40], [41]. There is density in the conserved metal binding site of the CSA2 domain that has been suggested to bind

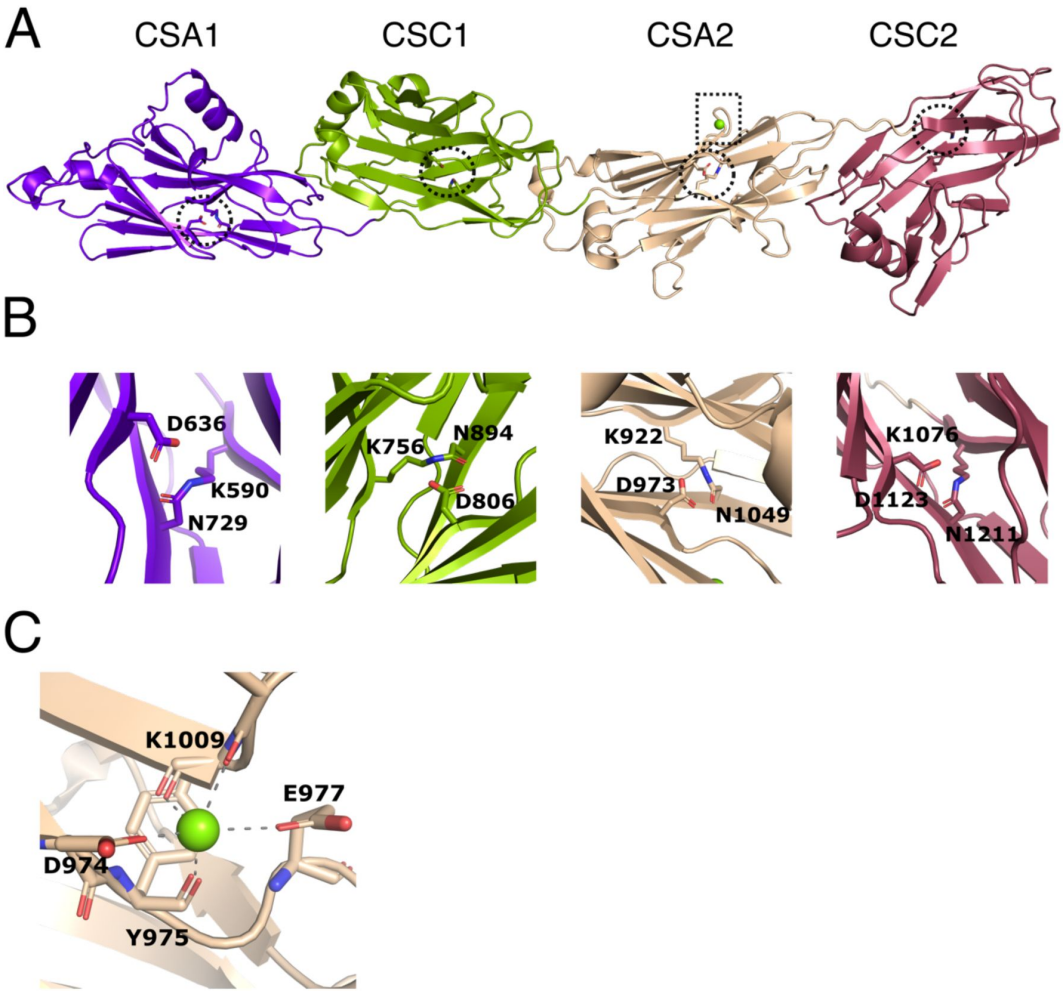

Figure 2.

Structure of Ig-like domains of PrgB. A) The Ig-like domains are arranged as tandem pairs (CSA1-CSC1 and CSA2-CSC2). B) Each Ig-like domain has an internal isopeptide bond (indicated by the striped circle in panel A) strengthening the structural integrity of the domain. Each of the four isopeptide bonds is between a lysine and an asparagine and further stabilized by an aspartic acid residue (highlighted residues are shown in stick representation). C) A conserved metal binding site is situated in the CSA2 domain (highlighted by a striped box in panel A), here modelled with a Mg^{2+} (green sphere).

Ca^{2+} in the Antigen I/II proteins. Refinement of our structure indicated that Mg^{2+} was the best fit to the density (**Fig. 2C**). The homologous C2 domains from AspA[38], Pas[42], SpaP[40] and SspB[41], each have an extra structural feature termed the BAR (SspB adherence region) domain, which mediates adherence in these proteins. This BAR domain is absent in PrgB (**Fig. 2 – figure supplement 1**).

As we didn't manage to crystallize $PrgB_{188\text{-}1233}$ without the loss of the polymer adhesin domain, we tried to determine its structure via cryo-EM and single particle analysis. Two datasets of PrgB, one with and one without ssDNA (120 bases), were collected. This yielded relatively low-resolution volumes (8 and 11 Å, respectively). See **Fig. 2 – figure supplement 2** for an overview of the processing. Docking in the X-ray structures of the stalk domain ($PrgB_{584\text{-}1233}$) and the previously solved polymer adhesin domain (PDB code: 6EVU)[19] into the volumes weakly indicated that the polymer adhesin domain could be interacting with the stalk domain in the absence of DNA (**Fig. 2 – figure supplement 3**). Therefore, we set out to test whether the polymer adhesin domain binds to the Ig-like domains from the stalk domain ($PrgB_{584\text{-}1233}$) *in vitro*. However, neither size exclusion chromatography nor native PAGE indicated any binding of the polymer adhesin domain to the stalk domain *in vitro* (**Fig. 2 – figure supplement 4**).

In vivo assays

Based on the new structural insights for PrgB, we decided to study the importance of the newly defined coiled-coil domain and the Ig-like domains. This was done by complementing *E. faecalis* OG1RF pCF10Δ*prgB* with different *prgB* mutants (from a nisin-inducible plasmid) and characterizing their phenotypes in cellular aggregation, biofilm formation, and plasmid transfer efficiency. In line with previous experiments[17], complementing pCF10Δ*prgB* with exogenous PrgB from the pMSP3545S vector rescues all phenotypes. Aggregation and biofilm formation, both measured after overnight incubation, are even slightly increased as compared to wild-type pCF10 (**Fig. 3A-B**, column 1 and 3), possibly due to a slightly increased production of PrgB (Fig. 3 – figure supplement 2B, lane 1 and 3).

We found that expression of PrgB without the newly identified coiled-coil domain could not rescue the aggregation phenotype of the deletion strain (**Fig. 3A**, column 4). Deletion of either the CSA1 or the CSC2 domain did not affect PrgB-mediated aggregation (**Fig. 3A**, column 8 and 9). However, complementation with PrgBΔCSA2-CSC2 could only partially rescue the aggregation phenotype of the *E. faecalis* OG1RF pCF10Δ*prgB* strain (**Fig. 3A**, column 7) and no rescue at all was seen in the PrgB variants with both the CSA1 and CSC1 domain deleted or without any Ig-like domains (CSA1-CSC1-CSA2-CSC2) (**Fig. 3A**, column 5, and 6).

As expected[17], deletion of *prgB* also leads to a large decrease in biofilm formation (**Fig. 3B**). In line with our observations from the aggregation assays, PrgB with a deletion of either the coiled-coil domain or more than a single Ig-like domain could not rescue the *E. faecalis* OG1RF pCF10Δ*prgB* biofilm formation phenotype. Only exogenous expression of $PrgB_{\Delta CSC2}$ can rescue the level of biofilm formation, but only to the level found in OG1RF pCF10, not to the level of exogenously expressed wild-type PrgB (**Fig 3B**, column 1, 3 and 9). Intriguingly, the expression of exogenous $PrgB_{\Delta CSA1}$ in the OG1RF pCF10Δ*prgB* background did not restore biofilm formation, while it did restore the aggregation phenotype (**Fig. 3A** and **3B**, column 8).

The PrgB variants that failed to rescue the aggregation phenotype of the OG1RF pCF10Δ*prgB* strain ($PrgB_{\Delta COI}$, $PrgB_{\Delta CSA1\text{-}CSC2}$, $PrgB_{\Delta CSA1\text{-}CSC1}$ and $PrgB_{\Delta CSA2\text{-}CSC2}$) were also found to have a decreased plasmid transfer efficiency in the conjugation assay (**Fig. 3C**, column 4-7). However, $PrgB_{\Delta CSA1}$ and $PrgB_{\Delta CSC2}$ could only partially rescue the conjugation rate of the OG1RF pCF10Δ*prgB* strain (**Fig. 3C**, column 8-9), while they did rescue the aggregation phenotype.

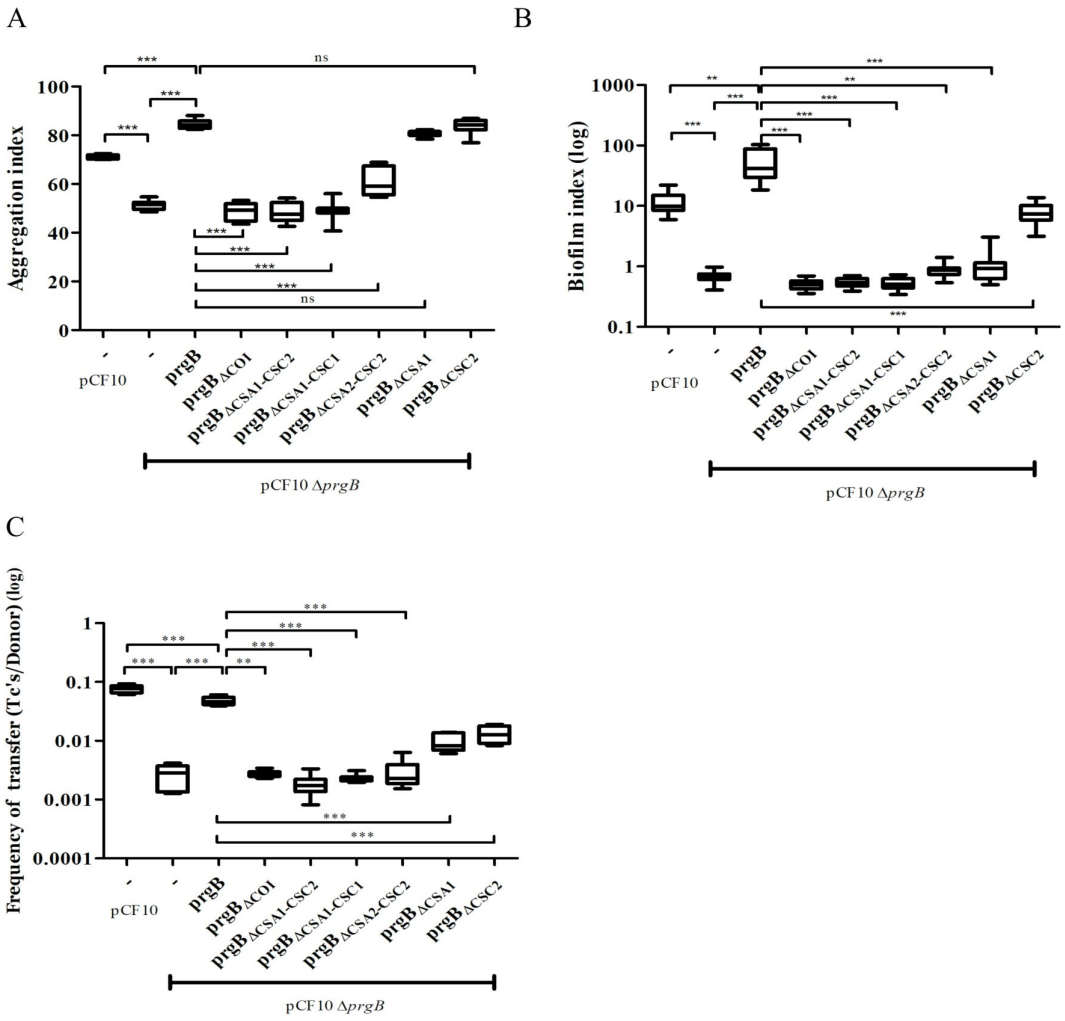

Figure 3.

In vivo phenotypes of PrgB variants. PrgB variants are expressed from the pMSP3545S-MCS vector in the OG1RF pCF10Δ*prgB* background for phenotypic analysis with three assays: A) cellular aggregation B) biofilm formation and C) conjugation assays. For all assays OG1RF pCF10 carrying an empty vector serves as positive control and OG1RF pCF10Δ*prgB* with an empty vector as negative control. The value of each column represents the average of three independent experiments and the error bars represent the standard error of the mean (SEM).

To determine whether all PrgB variants were properly expressed, translocated and linked to the cell wall, we probed the protein levels in the cell wall fraction by Western Bot after 1 hr induction (corresponding to a time point relevant for the conjugation assay) and after overnight incubation (corresponding to a time point relevant to the aggregation and biofilm assays). All PrgB variants were present in the cell wall extract after 1 hr induction (**Fig. 3 – figure supplement 1A**), indicating that the production and folding of them are normal. However, the protein levels of $PrgB_{\Delta COI}$, $PrgB_{\Delta CSA1\text{-}CSC2}$, $PrgB_{\Delta CSA1\text{-}CSC1}$, and $PrgB_{\Delta CSA2\text{-}CSC2}$ were reduced after overnight incubation, as compared to wild type (**Fig. 3 – figure supplement 1B**, lane 5-8 compared to lane 3). This indicates that the protein stability is decreased when these domains are missing, which could explain the observed aggregation deficiency after overnight incubation. Intriguingly, even though $PrgB_{\Delta CSA1}$ and $PrgB_{\Delta CSC2}$ are relatively stable (**Fig. 3 – figure supplement 1B**, lane 9-10) and complemented the aggregation phenotype, they still could not fully rescue the biofilm phenotype. This suggests a functional loss of in these two variants.

The conserved binding cleft in the polymer adhesin domain is essential for conjugation and biofilm formation, but not for aggregation

To investigate the role of the binding cleft in the polymer adhesin domain, we introduced single, double, and triple mutations to alter its conserved residues. The resulting *prgB* mutants were exogenously expressed in the background of pCF10Δ*prgB* for functional complementation; or in the background of wild-type pCF10 to test for any potential dominant negative effects as previously observed for $PrgB_{\Delta 246\text{-}558}$ (deletion of the polymer adhesin domain)[19]. Notably, all the resulting PrgB variants fully restored the aggregation phenotype of pCF10Δ*prgB* (**Fig. 4A**), but did not rescue the defective biofilm formation, nor the reduced conjugation efficiency (**Fig. 4B** and **C**, column 2-6). No dominant negative effects on aggregation or biofilm formation could be observed when these variants were expressed in the wild-type pCF10 background (**Fig. 4A** and **B**, column 7-10), although slightly reduced conjugation rates were observed (**Fig. 4C**, column 7-10). We have previously shown that the polymer adhesin domain binds eDNA and that this binding correlates with both biofilm formation and conjugation efficiency. Therefore, we wanted to test whether eDNA binding was affected in these PrgB variants. To do so, we purified both wild-type $PrgB_{188\text{-}1235}$ and the double binding cleft variant $PrgB_{188\text{-}1233}$:S442A,N444A to compare their DNA binding affinities via electrophoretic mobility shift assays (EMSA). The results indicate that the introduced changes in PrgB did not affect its ability to bind eDNA, as the DNA binding affinities of $PrgB_{188\text{-}1233}$ and $PrgB_{188\text{-}1233}$:S442A,N444A were the same within experimental error (**Fig. 4 – figure supplement 1**).

Discussion

The presented data provides important insights for a widespread virulence factor in Gram-positive bacteria, since genes encoding PrgB homologs exist on a large number of conjugative plasmids[16]. In this study, we expand our structural knowledge of PrgB beyond the polymer adhesin domain, to now encompass almost the entire protein.

The crystal structure of $PrgB_{583\text{-}1233}$ shows that this part of PrgB consists of four tandemly arranged immunoglobulin (Ig)-like domains. These Ig-like domains show a high degree of structural homology to Streptococcal surface proteins, usually found in the oral cavity[16]. These homologous proteins have been indicated to bind various molecules, such as fimbria, collagen and salivary agglutinin (SAG, also designated as glycoprotein 340), and these binding interactions have been shown to be vital for their function[41–45]. However, we have not found any evidence that the Ig-like domains of PrgB bind to a specific substrate in our own experiments, nor have we found this in other reports. PrgB also does not contain a BAR domain, which is crucial for stable

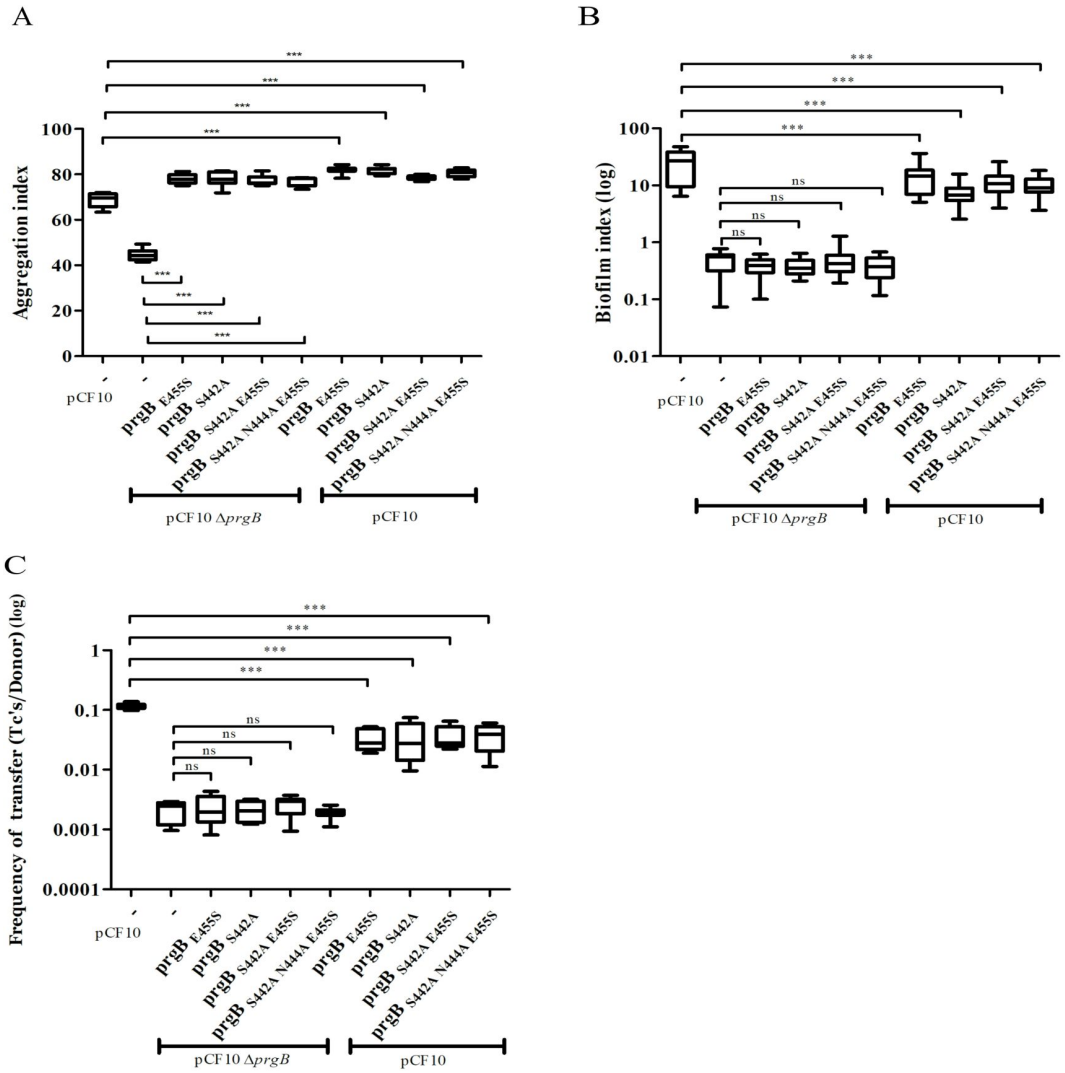

Figure 4.

In vivo phenotypes of PrgB variants with point mutation(s) in the conserved site of the polymer adhesin domain. PrgB variants are expressed from the pMSP3545S-MCS vector in the OG1RF pCF10Δ*prgB* or OG1RF pCF10 strain and analyzed in: A) cellular aggregation, B) biofilm formation, and C) conjugation assays. OG1RF pCF10 carrying the empty vector serves as positive control and OG1RF pCF10Δ*prgB* with the empty vector as negative control. The height of each column represents the average of three independent experiments and the error bars indicate the standard error of the mean (SEM).

interactions between *e.g.* SspB from *S. gordonni* and Mfa-1 of *P. gingivalis*[46], [47]. However, since Ig-like domains are known to bind a large variety of ligands[48], we don't exclude the possibility that ligands for the Ig-like domains in PrgB will be found in the future. The only ligands that have been verified to interact with PrgB so far are eDNA and LTA, which have high affinity to the polymer adhesin domain[19].

The crystal structure of the Ig-like domains from PrgB, $PrgB_{583\text{-}1233}$, was complemented by single particle analysis of $PrgB_{188\text{-}1233}$ via cryo-EM (**Fig. 2 – figure supplement 3**). Despite the low resolution of the EM volumes, we tried to dock in the high-resolution crystal structures of the Ig-like domains and the polymer adhesin domain. The model of PrgB from cryo-EM indicated that the polymer adhesin domain might interact with the Ig-like domains in the absence of substrate, which could have implications for the function or regulation of PrgB. To test this hypothesis, we assayed the interaction between the polymer adhesin domain and the separately purified Ig-like stalk domain *in vitro*. The results did not show any interaction between these domains. However, in the full-length protein these two domains are attached via a linker region, which make their local concentration very high. Thus, we cannot exclude that these two domains can interact, but in that case the dissociation constant must be high (at least high µM range) and the interaction is thus unlikely to be physiologically relevant.

We have now obtained a structural basis to interpret almost all phenotypic data that is available for PrgB. Unfortunately, most of the mutants that were previously described did not correlate well with the newly determined domain boundaries. Therefore, we decided to create specific deletion mutants that were based on the new PrgB structure to determine the role of the various domains. At the N-terminus of PrgB, before the polymer adhesin domain, there is a predicted coiled-coil region (**Fig. 1A**) that we wanted to investigate. To our surprise, the expression of $PrgB_{\Delta COI}$ could not rescue any of the aggregation, biofilm formation or conjugation phenotypes from a *prgB* deletion strain. However, in our experiments $PrgB_{\Delta COI}$ was unstable with substantially decreased amounts present in the cell-wall extract collected after overnight induction as compared to 1 hour induction (**Fig. 3 – figure supplement 1**), indicating that the coiled-coil region is important for protein production and/or stability. In line with this hypothesis, Alphafold predicts this coiled-coil domain to interact with the linker between the polymer adhesin domain and the Ig-like domain (**Fig. 2 – figure supplement 5**). Since the linker region contains the sequence that PrgA recognizes for cleavage of PrgB[15], the lack of the coiled-coil domain and its potential shielding effect could explain the decreased stability of $PrgB_{\Delta COI}$. Deletion of all Ig-like domains ($PrgB_{\Delta CSA1\text{-}\Delta CSC2}$) renders the protein incapable to support aggregation, biofilm formation and conjugation. However, exogenous expression of *prgB* with single Ig-like domain deletions, $prgB_{\Delta CSA1}$ and $prgB_{\Delta CSC2}$ in the pCF10Δ*prgB* background, restore cellular aggregation, while they do not rescue biofilm formation and conjugation (**Fig. 3**). This was unexpected, as the polymer adhesin domain is predicted to mediate all the functions that were tested in these assays: aggregation, biofilm formation and conjugation[17], [19]. Our results, however, indicate that it is important to have all Ig-like domains present and properly folded. The cellular aggregation assays seem to indicate that PrgB can function when at least 3 Ig-like domains are present, as expression of both *prgB*ΔCSA1 and *prgB*ΔCSC2 can complement pCF10Δ*prgB*, but unfortunately this assay is not suitable to detect small changes. Based on the biofilm formation and conjugation efficiency assays, which are more sensitive, we therefore conclude that even removing a single Ig-like domain strongly decreases the function of PrgB. We hypothesize that these Ig-like domains are required to present the polymer adhesin domain at the right distance from the cell.

The conserved Ser-Asn-Glu site in the negatively charged cleft of the polymer adhesin domain intrigued us, as its function is unknown. Any changes that we made in these conserved residues resulted in PrgB variants that didn't facilitate biofilm formation or conjugation (**Fig. 4**). Surprisingly the same PrgB variants did fully support cellular aggregation. This is thought-provoking, since various literature has shown that the PrgB functions in cellular aggregation, biofilm formation and conjugation are strongly correlated. However, even a single point mutation

in this conserved site produced a PrgB variant that completely separates the cellular clumping phenotype from biofilm formation and conjugation. *In vitro* experiments showed that these point mutations did not impair PrgB binding to eDNA (**Figure 4 – figure supplement 1**).

A similar phenotypic pattern was observed with $PrgB_{\Delta CSA1}$ and $PrgB_{\Delta CSC2}$, which both could fully rescue aggregation but not biofilm formation. Our data therefore strongly indicates that PrgB has additional role(s), besides mediating cellular aggregation, to further support conjugation and biofilm formation. We therefore propose that PrgB performs at least one, currently unknown, additional function besides binding to eDNA and/or LTA from the cell wall of the recipient cell. It is highly likely that at least one of these additional functions is mediated by the conserved site in the polymer adhesin domain.

As described in the introduction, there is a plethora of *prgB* mutants made (see **Fig. 5 – figure supplement 1**) and phenotypically analyzed, predominantly by the group of Prof. Gary Dunny. In supplementary file 4 we provide a summary of all *prgB* mutants that we have identified in the literature and our brief reinterpretation based on our new structural knowledge. Below, we will discuss a selected number of these mutants in detail.

As expected, most mutants that have an insertion or a mutation in the polymer adhesin domain (**Fig. 5** – figure supplement 1) show a loss of protein function. Shortly, insertions at amino acids (a.a.) 358 and 359 are on the surface of the protein and likely leads to steric clashes that prevent the domain from binding to eDNA and LTA, whereas insertions at a.a. 439, 473, 517, and 546 are all in secondary structures that are central parts of the polymer adhesin domain and therefore are likely to disrupt folding of this domain.

The RGD motifs in PrgB were previously proposed to be involved in integrin binding and to promote adherence to human neutrophils, as well as internalization in cultured intestinal epithelial cells[26, 49]. The new domain classification showed that these two RGD motifs are in CSA1 and CSA2, respectively (**Fig. 1A**). The structure that we determined shows that these two sequence motifs are not surface exposed, but instead play an important role in the structural integrity of the interfaces between CSA1 and CSC1, and between CSA2 and CSC2 (**Fig. 5**). Mutations in these motifs would thus very likely destabilize the folding of the tandem Ig-domains, which would explain the decreased PrgB biofilm formation observed in these strains[18, 24]. Thus, our new data strongly argues against the previously proposed direct binding interaction between the RGD sequences and integrins of host origin.

Deletion of residues 993-1138 leads to a loss of about half of both the CSA2 and CSC2 domains. Therefore, we were surprised to find that this mutant was described to behave like wild type in aggregation assays[32]. Possibly the remaining parts of CSA2 and CSC2 form a (unfolded) linker region of similar length to the tandem CSA2-CSC2 structure, which allows PrgB to retain its function of promoting aggregation. Similarly, $PrgB_{\Delta 668\text{-}1138}$, corresponding to a complete removal of CSC1 and CSA2 and approximately half of both CSA1 and CSC2 domains, could still support PrgB mediated aggregation. Potentially the remaining residues (ca 180 amino acids) could also form an unfolded linker region allowing for the variant to retain its aggregation phenotype. For $PrgB_{\Delta 993\text{-}1138}$ and $PrgB_{\Delta 668\text{-}1138}$, unfortunately no biofilm formation or conjugation efficiency were reported, but based on our results it is likely that those capabilities would have been severely compromised.

Taking past and present results into account combined with our new structural insights, we propose a new mechanistic model for the function of PrgB. Our *in vivo* data show that removal of one of the Ig-like domains does not largely affect PrgB function. Previous data also indicates that parts of the Ig-like domains can be deleted without affecting the function of PrgB. We therefore hypothesize that the Ig-like domains provide two important features to the protein. First, a rigid stalk that is needed to present the polymer adhesin domain at the correct distance from the cell

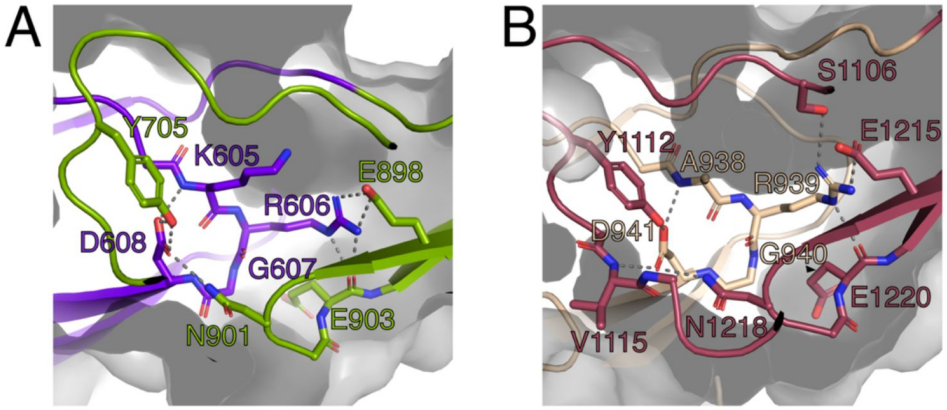

Figure 5.

The RGD motifs play a role in the structural integrity of PrgB. A) Close up view of the RGD motif in the CSA1 domain (purple). Most of the important interactions, mainly hydrogen bonds, formed by this motif are to residues on the CSC1 domain (green). B) Close up view of the RGD motif in the CSA2 domain (sand colored). Most of the important interactions, mainly hydrogen bonds, formed by this motif are to residues on the CSC2 domain (dark red). Potential hydrogen bonds are marked by lines and the surface of the protein indicated by transparent gray.

wall. Second, providing the correct structural positioning for PrgA-mediated cleavage of the polymer adhesin domain, as the potential protease domain in PrgA is also presented away from the cell on a ~40 nm long stalk [15, 16]. The exact distance of the polymer adhesin domain from the cell wall does not seem important for LTA-binding, since aggregation can still take place even with partially disrupted Ig-like domains. However, the length of the stalk may be very important when it comes to facilitating both biofilm formation and conjugation. This indicates that effective mating-pair formation in conjugation may require a PrgB-mediated function distinct from pure aggregation, something that is further shown by the findings of the mutations in the conserved site in the polymer adhesin domain.

Besides providing a structural basis to explain about 30 years of work on PrgB, we here also uncovered that the conserved Ser-Asn-Glu site in the polymer adhesin domain likely provides additional functionality to PrgB that is needed for optimal biofilm formation and conjugation, but that does not affect cellular aggregation. To fully investigate the function of this conserved site in PrgB and other homologous virulence factors from Gram-positive bacteria, remains an exciting question to address in future research.

Materials and Methods

Bacterial strains and growth conditions

See Supplementary File 1 for a full list of all strains, plasmids and oligonucleotides used. *Escherichia coli* Top10 was used in molecular cloning and grown in Lysogeny broth (LB). *E. coli* BL21 (DE3) was used for recombinant protein expression and grown in Terrific Broth (TB). The *E. faecalis* strains were cultured in Brain-Heart infusion broth (BHI) or Tryptic Soy broth without dextrose (TSB-D) as indicated in each assay. Concentrations of antibiotics for *E. coli* selection were as follows: ampicillin (100 µg/ml), kanamycin (50 µg/ml), spectinomycin (50 µg/ml), and erythromycin (150 µg/ml). In *E. faecalis* cultures, antibiotics were used as the following concentrations: tetracycline (10 µg/ml), fusidic acid (25 µg/ml), erythromycin (20 µg/ml for chromosome-encoded resistance; 100 µg/ml for plasmid-encoded resistance), spectinomycin (250 µg/ml for chromosome-encoded resistance; 1000 µg/ml for plasmid-encoded resistance), streptomycin (1000 µg/ml). Plasmids were transformed to *E. coli* by heat-shock transformation, whereas *E. faecalis* strains were transformed by electroporation[50].

Cloning and mutagenesis

To insert a multiple cloning site in pMSP3545S, DNA oligos of MCS_fwd and MCS_rev were resuspended in miliQ to 100 µM, mixed 1:1, denatured at 95°C for 15 min and slowly cooled to room temperature. The annealed MCS oligo was ligated into the gel-purified backbone fragment from pMSP3545S-*prgK* vector[51] digested with NcoI and XbaI (removing the *prgK* insert). This was done in a 30 µL ligation mixture with 90 ng of the digested vector and a 9 times molar access of insert. 5 µL of this ligation mixture was transformed into Top10 competent cells and plated on LB agar plates with spectinomycin (50 µg/ml), and erythromycin (150 µg/ml). The constructed plasmid was analyzed by restriction digestion and sequenced to confirm the correct insertion of the multiple cloning site and is further called pMSP3545S-MCS.

pMSP3545S-*prgB* was constructed by PCR of *prgB* from pCF10 with the NcoI-prgB-F and BamHI-stop-prgB-R primer pair and placement into the pMSP3545S-MCS vector via NcoI/BamHI restriction, gel-purification of the correct DNA fragments and subsequent ligation. The constructed pMSP3545S-*prgB* was subsequently used as a template for mutagenesis creating pMSP3545S-*prgB* deletion or point mutation variants. This was carried out with inverse PCR (iPCR) using partially overlapping primer pairs. The iPCR products were gel-purified with the DNA clean-up kit and digested with DpnI to remove residual template plasmid DNA. The processed iPCR products were then transformed to Top10 competent cells. For over-expression and protein purification, pET-

prgB$_{246\text{-}558}$ and pET-*prgB*$_{188\text{-}1233}$ was transformed into *E. coli* BL21(DE3) as previously described[15]. *prgB*$_{188\text{-}1233}$ and *prgB*$_{580\text{-}1233}$ DNA fragments were amplified with primers mentioned in Supplementry File 1, and cloned into pINIT vector and then sub-cloned into the p7XC3GH vector via the FX cloning system[52]. Mutations were introduced to p7XC3GH-*prgB*$_{188\text{-}1233}$ with the same iPCR approach to obtain the derivative plasmid of *prgB*$_{188\text{-}1233:\ S442A,\ N444A}$. All constructed plasmids were screened by PCR and verified by sequencing.

Protein purification and crystallization

PrgB was produced as previously described[19]. Briefly, PrgB$_{188\text{-}1233}$ was expressed with an N-terminal hexa-histidine tag from pET-*prgB*$_{188\text{-}1233}$ or a C-terminal deca-histidine and GFP tag from p7XC3GH-*prgB* in *E. coli* BL21(DE3). The cells were grown at 37 °C in TB medium until they reached an OD_{600} of 1.5. Then the temperature was lowered to 18 °C and protein production was induced by adding 0.5 mM IPTG. Cells were grown for 16 hours before harvesting by centrifugation. Cells were resuspended in 20 mM HEPES/NaOH (pH 7.0), 300 mM NaCl, 30 mM Imidazole and 0.02 mg/ml DNase I and broken with a Constant cell disruptor at 4 °C and 25 kPsi (Constant Systems). The cell lysate was clarified by centrifugation for 30 minutes at 30,000 x g, 4°C and incubated at 4 °C with Ni-NTA (Macherey-Nagel). The Ni-NTA column was washed with 10 column volumes of 20 mM HEPES/NaOH (pH 7.0), 300 mM NaCl, 30-50 mM imidazole and bound proteins were eluted from the column with the same buffer supplemented with 500 mM Imidazole. The histidine affinity tags and the GFP when present, were cleaved off from the purified protein fractions by incubation with TEV protease (for pET-*prgB*) or Prescission protease (for p7XC3GH-*prgB*) in a 1:100 ratio for 20 h at 4 °C. The cleaved proteins were loaded on a Superdex-200 Increase 10/300 GL column (Cytiva) equilibrated in 20 mM HEPES/NaOH (pH 7.0) and 150 mM NaCl. The elution profile showed two peaks corresponding to a PrgB dimer and monomer, as previously reported[19]. These two peak fractions were pooled separately and concentrated on an Amicon Ultra Centrifugal Filter with a 30 kDa cut-off (Merck-Millipore). 10% glycerol was added to the concentrated protein fractions, which were subsequently flash frozen in liquid nitrogen and stored at –80 °C.

Native gel electrophoresis

Elution fractions from size-exclusion chromatography were mixed with native gel sample buffer (Invitrogen) and loaded on a Novex 4-20% Tris-Glycine gel (Invitrogen). Staining was carried out with InstantBlue Protein Stain (VWR).

Structure determination via X-ray crystallography

Purified PrgB$_{188\text{-}1233}$ from the monomeric fraction, with a protein concentration of 15 mg/mL, were thawed and used in crystallization trials. Crystals of PrgB$_{188\text{-}1233}$ were grown in 8-12 weeks, at 20 °C by sitting drop vapor diffusion in a condition with 0.2 M $CaCl_2$ and 20% PEG 3350 and a protein to reservoir ratio of 1:1 in the drop. Crystals were flash cooled in liquid nitrogen. X-ray diffraction data of PrgB$_{188\text{-}1233}$ was collected on ID23-1 at the ESRF, France. The data was processed using XDS[53]. The PrgB$_{188\text{-}1234}$ crystals belong to the monoclinic space group $P2_12_12_1$ and contain two molecules in the asymmetric unit. The crystallographic phase problem was solved using molecular replacement using PHASER, using the SspB homology model of the Ig-domains as search models (PDB: 2WOY)[41]. Further building of the model was conducted in COOT[54]. The structure was refined to 1.85 Å with crystallographic R_{work} and R_{free} values of 20.9/24.6 using Refmac5 and PHENIX refine[55], [56]. The final PrgB$_{188\text{-}1233}$ model consists of residues 584-1234, and was validated using MolProbity[57]. Atomic coordinates and structure factors have been deposited in the Protein Data Bank (PDB Code: 8BEG).

Sample preparation for electron microscopy

$PrgB_{188-1233}$ dimer fractions were thawed on ice and loaded on a Superdex 200 10/300 GL gel filtration column (GE Healthcare) equilibrated in 20 mM HEPES/NaOH pH 7.0 and 150 mM NaCl. Protein peak fractions, corresponding to the dimer, were diluted to 0.1-0.3 mg/mL and for the DNA-bound structures 120 bp ssDNA (Table S1) was added in 1: 1.2 molar ratio (protein:DNA) and samples were incubated for 15 minutes on ice. For both apo and DNA-bound samples, 4 µl of sample was applied to glow discharged Quantifoil 300 mesh 1.2/1.3 (Quantifoil) grids at 4 °C and 90-100% humidity, blotted for 1 s with blot force –5, and plunge-frozen in liquid ethane using a Vitrobot Mark IV (Thermo Fisher Scientific).

Cryo-EM data collection

Cryo-EM data were collected on an FEI Titan Krios transmission electron microscope (Thermo Fisher Scientific), operated at 300 keV that was equipped with a K2 direct electron detector. Data was collected by the AFIS method using the EPU software V2.8.0 (Thermo scientific) at a nominal magnification of 165,000x (0.82 Å pixel size). Data collection parameters are listed in Supplementary file 2, and the general workflow showing representative micrographs, 2D and 3D classes are shown in **Fig. 2 – figure supplement 2**. A total number of 2787 movie stacks were collected for apo PrgB and 1670 for DNA-bound PrgB.

Cryo-EM data processing

Cryo-EM data of apo PrgB and DNA-bound PrgB were processed in the same way, but separately using cryoSPARC (v3.2.0-3.3.1)[58]. Beam-induced motion was corrected using standard settings, where start frame 1 was excluded, followed by per-micrograph contrast transfer function (CTF) estimation. For apo PrgB a subset of 500 micrographs were picked using the blob picking tool with a 100-200 Å particle diameter, followed by extraction of 170,826 particles with a box size of 384 pix. Picked particles were subjected to consecutive rounds of 2D classifications. Subsequently, representative 2D classes were selected as input for picking of the full dataset, using the template picker tool. PrgB with DNA was directly picked using blob picker with a 100-300 Å particle diameter and standard settings and extracted with 384 pix. PrgB without and with DNA were then separately subjected to 2D classifications resulting in a final number of 283,630 and 163,566 particles respectively. Particles from selected classes were combined and used in ab-initio reconstruction. The initial volume was then subjected to homogenous 3D refinement and the resolution was calculated using the gold standard Fourier shell correlation (FSC threshold, 0.143) and found to be 8 Å and 11 Å for the apo and DNA-bound structure, respectively. The volumes of apo and DNA-bound PrgB have been deposited in the Electron Microscopy Data Bank (EMDB Codes: EMD-16001 and EMD-16002).

The adhesion domain (PDB Code: 6EVU)[19] and stalk domain (PDB Code: 8BEG) were initially docked into the EM volume using Chimera[59] and subsequently run through Namdinator[60] using 10 Å resolution and standard settings. The output was then fitted in the EM volume in Chimera (v 1.15rc[59]), where figures also were generated.

Aggregation assay

E. faecalis strains were inoculated in BHI medium with the indicated antibiotics and cultured overnight at 37°C. Overnight cultures were diluted in a 1:100 ratio in TSB-D with the indicated antibiotics, 10 ng/mL cCF10, and 50 ng/ mL nisin and dispensed into polystyrene cuvettes (Sarstedt) in 0.9 mL triplicates. These were incubated for 24 h at 37°C without agitation. Afterwards, the optical density of each sample was determined at 600 nm both before (OD_{sup}) and after (OD_{mix}) vigorously mixing of the bacterial culture by pipetting. The autoaggregation percentage was then calculated as follows: $100 \times [1 - (OD_{sup}/OD_{mix})]$[17, 30].

Biofilm assay

E. faecalis strains were inoculated in BHI with the indicated antibiotics and kept 16 h at 37°C. The next morning, they were diluted in a 1:100 ratio in TSB-D with the indicated antibiotics, 10 ng/mL cCF10, and 50 ng/mL nisin. 200 µl fractions were dispensed into a 96-well micro-titer plate (Costar) with 8 replicates per strain. 200 µL TSB-D fractions were used as blanks. The 96-well plate was then incubated aerobically at 37°C without agitation in a humidified chamber for 24 h. The suspension was transferred to another 96-well plate to determine the optical density at 600 nm (OD_{600}). The plate containing the biofilm was washed with distilled water three times and then left to air dry at room temp for 2.5 h. The biofilm was stained with 100 µl 0.1% (w/v) safranin (Sigma) at room temp for 20 min, then washed three times with distilled water and left to air dry at room temperature. Afterwards the absorbance was determined using a plate reader (BMG Labtech) at 450 nm. Biofilm production was calculated as an index of safranin staining of the cell biomass divided by absorbance of its optical density (OD_{450}/OD_{600})[61].

Analysis of the protein levels of PrgB variants in cell wall extracts

For the 1-hour time-point, samples of each strain were diluted in a 1:25 ratio in TSB-D with the indicated antibiotics, cultured at 37°C for 2h, until optical density 0.6 was reached, and then induced with 10 ng/mL cCF10, and 50 ng/ mL nisin for 1 hour. Overnight samples were harvested after induction and incubation overnight as described in the aggregation assay. The cell pellets were treated with lysozyme buffer (10 mM Tris, pH 8.0, 1 mM EDTA, 25% sucrose, 15 mg/ml lysozyme) for 30 min at 37°C. The lysozyme-treated bacterial cells were spun down at 13,000 *xg*, 4 °C for 5 minutes. The supernatant, containing the cell wall extract, was mixed with protein loading dye and boiled at 100°C for 12 min. Subsequently, the samples were run on a 8% SDS-PAGE, transferred to Western blot and probed with the PrgB antibody produced in rabbit[62–64].

Conjugation assay

Donor (OG1RF pCF10 pMSP3545S derivative strains) and recipient (OG1ES) strains were inoculated in BHI with the indicated antibiotics and incubated overnight at 37 °C with agitation. Overnight cultured strains were refreshed in BHI without antibiotics in a 1:10 ratio, and donor strains were induced with 50 ng/mL nisin (Sigma). All strains were then incubated at 37 °C for 1h without agitation. Afterwards each of the donor strains was mixed with the recipient cells in ratio of 1:10 and mated at 37 °C statically for 3.5 h. These mixtures were then serially diluted with BHI and plated out in triplicates on BHI agar plates supplemented with tetracycline and spectinomycin (to select for donor cells), or with tetracycline, erythromycin, and streptomycin (to select for transconjugants). Plates were incubated at 37 °C for 48 h, counted and enumerated for colony forming units (CFU). The plasmid transfer rate was determined as CFU of transconjugant over CFU of donor (Tc's/Donors)[19].

Electrophoretic Mobility Shift Assay (EMSA)

EMSA was carried out in the same way as previously described[19]. 0.1 to 20 µM PrgB (wild type and variants) were mixed with 50 nM 100 bp long double stranded DNA[19]. Samples were incubated for 1 hour at 20 °C before loading them onto a 6% TBE-based native acrylamide gel for electrophoresis for 90 min at 50 V and 6 °C. Gels were subsequently stained with 3× GelRed (Biotium) in distilled water for 30 minutes and imaged with a Chemidoc system (BioRad). Quantification of the DNA bands after imaging was done in ImageLab (BioRad).

Statistical analysis

All *in vivo* data is from three independent experiments and was plotted and analyzed using GraphPad Prism (version 5.0) (GraphPad Software). The indicated error is the standard deviation over 3 biologically independent replicates. Statistical significance between the PrgB variants were

analyzed with One-way Anova, with * indicating $p < 0.05$, * * indicating $p < 0.01$, and * * * indicating $p < 0.001$.

Material availability

All data generated or analyzed during this study are included in this published article and its supplemental information. All structural data has been deposited in the Protein Data Bank (PDB) and the Electron Microscopy Data Base (EMDB) and is publicly available. DOIs are listed in the key resources table. Any additional information required to reanalyze the data reported in this paper, *e.g.* bacterial strains, is available from the corresponding author upon request.

Acknowledgements

The authors would like to thank Prof. Gary Dunny for very fruitful discussions about the results and Dr. Karim Rafie and Annika Breidenstein for discussions about EM data processing. We acknowledge MAX IV Laboratory for time on Beamline BioMax under Proposal 20180236. Research conducted at MAX IV, a Swedish national user facility, is supported by the Swedish Research council under contract 2018-07152, the Swedish Governmental Agency for Innovation Systems under contract 2018-04969, and Formas under contract 2019-02496. We also acknowledge the synchrotrons Swiss Light Source (Paul Scherrer Institute, Switzerland) for time at beamline PX1 and the ESRF (France) for time at beamlines ID23 and ID30. The EM data was collected at the Umeå Core Facility for Electron Microscopy, a node of the Cryo-EM Swedish National Facility, funded by the Knut and Alice Wallenberg, Family Erling Persson and Kempe Foundations, SciLifeLab, Stockholm University and Umeå University. This work was supported by grants from the Swedish Research Council (2016-03599), Knut and Alice Wallenberg Foundation, Kempestiftelserna (SMK-1762 & SMK-1869) and Carl-Tryggers stiftelse (CTS 18:39) to R.P-A.B.

CRediT statement

Wei-Sheng Sun: Conceptualization, Investigation, Writing – Original Draft, Writing – Revision. Lena Lassinantti: Investigation, Writing – Original Draft, Writing – Revision. Michael Järvå: Conceptualization, Investigation, Andreas Schmitt: Investigation. Josy ter Beek: Investigation, Writing – Original Draft, Writing – Revision. Ronnie P-A Berntsson: Conceptualization, Writing – Original Draft, Writing – Revision, Supervision, Funding acquisition.

Tables

Supplementary files

Supplementary File 1. Strains, plasmids and oligonucleotides

Supplementary File 2. Cryo-EM data collection and refinement

Supplementary File 3. Top hits of Dali search based on PDB of PrgB Ig-like domains

Supplementary File 4. Structural interpretation of previously observed PrgB phenotypes

Data collection summary	$PrgB_{584-1233}$
Space group	$P2_12_12_1$
Cell dimensions	
a, b, c (Å)	76.0, 87.5, 224.8
α, β, γ (°)	90, 90, 90
Resolution (Å)	47.3 – 1.84 (1.91 – 1.84)
Completeness (%)	99.6 (99.5)
R_{meas} (%)	0.03 (0.99)
I/σ (I)	12.1 (1.0)
CC(1/2)*	1 (0.67)
Redundancy	2.0 (2.0)
No. unique reflections	130301 (12877)
Refinement summary	
Resolution (Å)	47.3 – 1.84
R_{work} (%)	20.9
R_{free} (%)	24.5
Number of atoms	
protein	10080
water	816
other ligands	2
B-factors	
protein	44.9
water	47.7
other ligands	48.3
r.m.s. deviations	
Bond lengths (Å)	0.006
Bond angles (°)	0.79
Ramachandran statistics	
outliers (%)	0
allowed (%)	1.5
favored (%)	98.5

Table 1.

X-ray data collection and refinement statistics. Values within parenthesis correspond to the highest resolution shell.

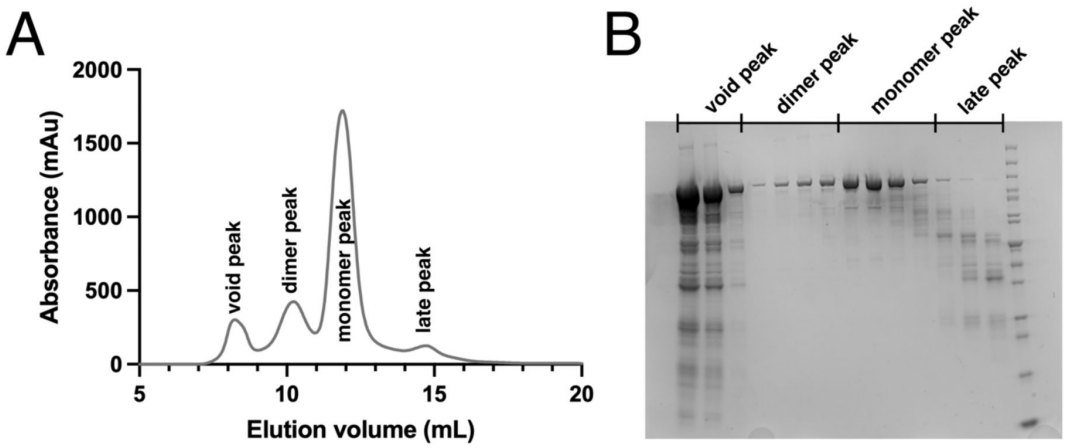

Figure 1 – figure supplement 1.

Purification of PrgB.

A) Representative chromatogram from size exclusion chromatography on a Superdex 200 Increase 10/300 GL column. 4 peaks are observed, corresponding to the void, PrgB dimer, PrgB monomer and a late peak. B) SDS-PAGE of the SEC elution fractions indicated in panel A.

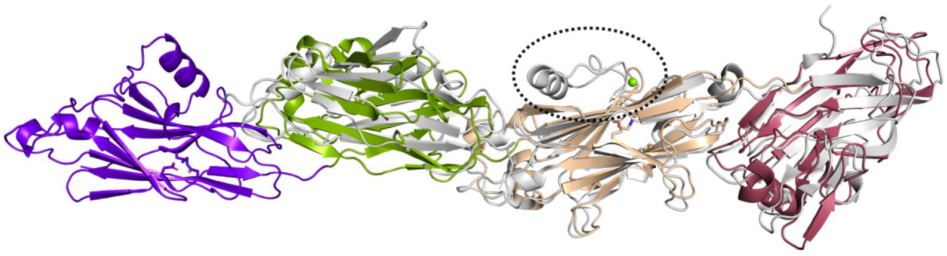

Figure 2 – figure supplement 1.

Structure of PrgB compared to the C-terminal domain of Pas (gray) (PDB Code: 6E3F). The BAR motif from Pas, missing in PrgB, is highlighted by the striped circle.

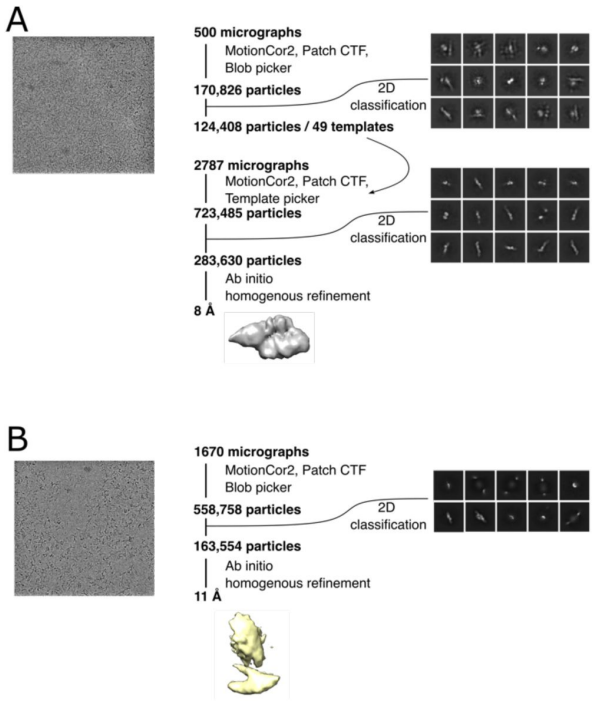

Figure 2 – figure supplement 2.

Cryo-EM data processing scheme. Data collection and processing for the apo PrgB dataset (A) and PrgB bound to DNA (B). For both, representative micrographs and 2D classes used in the processing are shown.

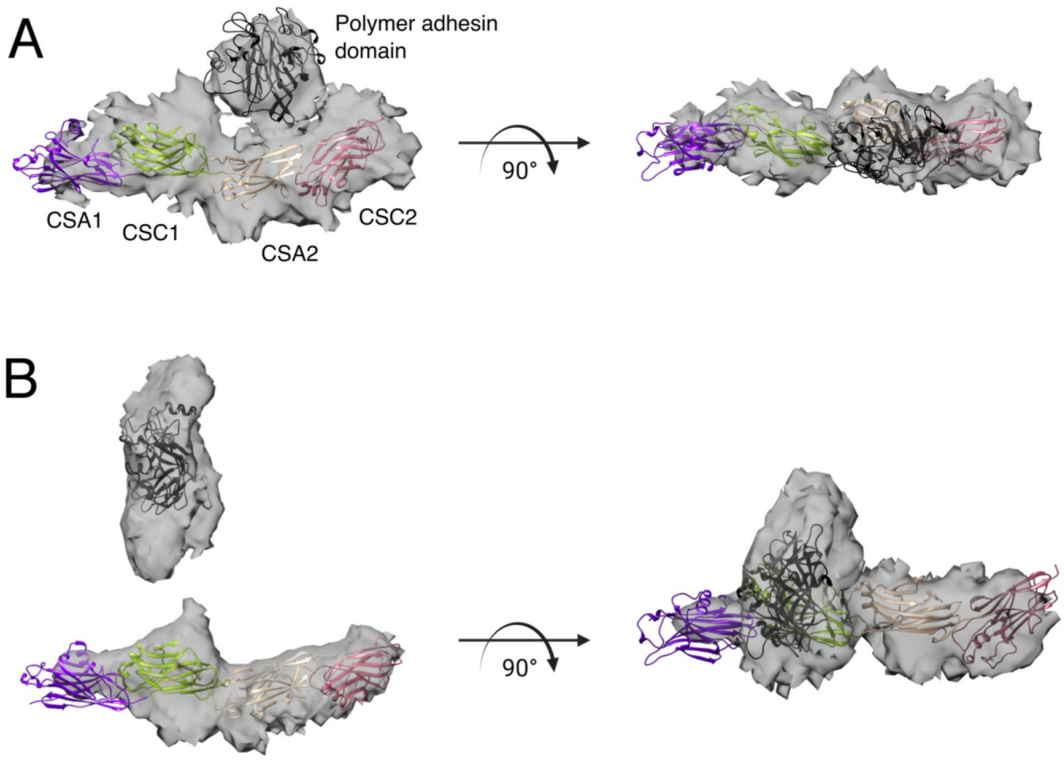

Figure 2 – figure supplement 3.

Crystal structures of the PrgB polymer adhesin domain and Ig-like domains in cartoon representation docked into the volumes acquired by single particle cryo-EM. A) The PrgB188-1233 model shows the polymer adhesin domain folded back onto the Ig-like domains. B) In the DNA-bound PrgB188-1233 model the polymer adhesin domain has undergone a large conformation change away from the Ig-like domain.

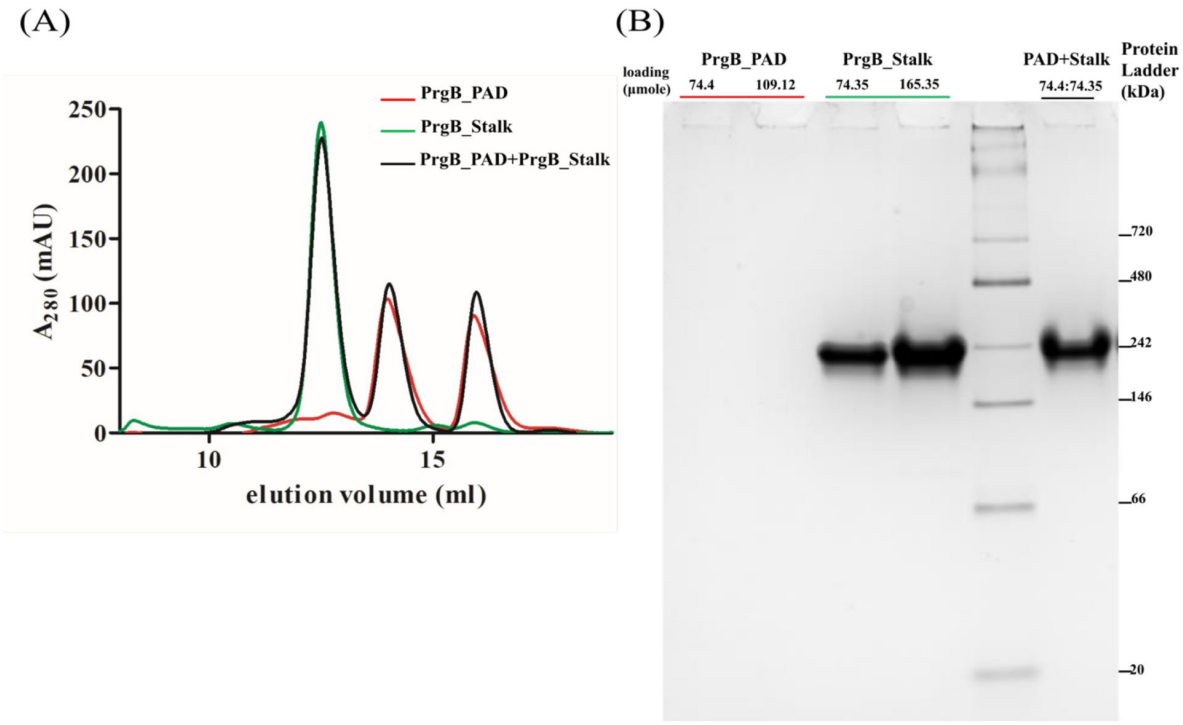

Figure 2 – figure supplement 4.

Analysis of potential interaction of the Polymer adhesin domain and stalk-like domains of PrgB. (A) Size-exclusion chromatography of PrgB_PAD (red), PrgB_Stalk (green), and mixture (black). No shifted peaks are observed in the mixture. (B) Native PAGE electrophoresis of PrgB_PAD, PrgB_Stalk, and mixture in 4-20% native gradient gel. PrgB_PAD has a strongly positively charged surface, and does not enter the gel on its own. No shift is observed when mixed with the Stalk domain.

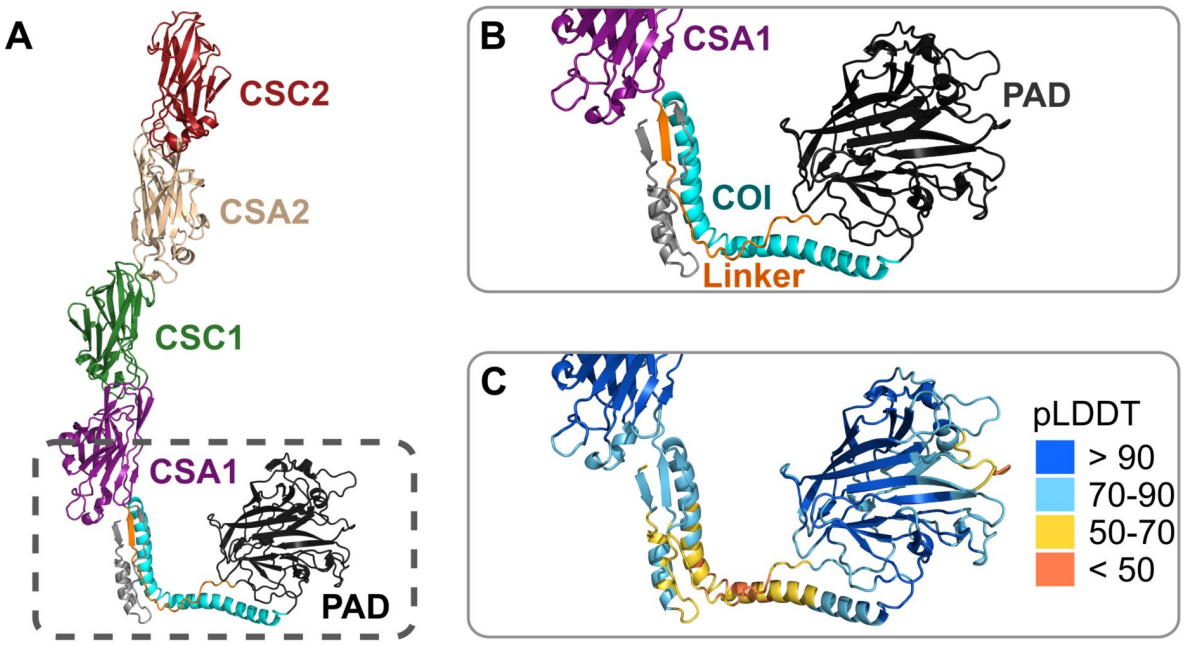

Figure 2 – figure supplement 5.

AlphaFold model of PrgB. A) Model of PrgB with N-proximal residues 157-194 in grey, the coiled-coil (COI, residue 195-260) domain in cyan, the polymer adhesin domain (PAD, residue 261-558) in black, the following linker region (residue 559-583) in orange, CSA1 (residue 584-750) in purple, CSC1 in green, CSA2 in beige and CSC2 in red. Both N– and C-termini (residues 1-156 and 1233-1305) are removed for clarity as they were predicted to be largely unfolded. B) The linker region between the polymer adhesin domain and the CSA1 domain, as indicated in panel A. C) The same region as shown in panel B with the residues colored by their confidence measure (pLDDT score). Regions with a pLDDT score above 90 are expected to be modelled with high accuracy, regions with a score between 70-90 are expected to have a good prediction for the backbone, while regions with a pLDDT score between 50-70 are off low confidence. A pLDDT score < 50 is a reasonable strong predictor of disorder.

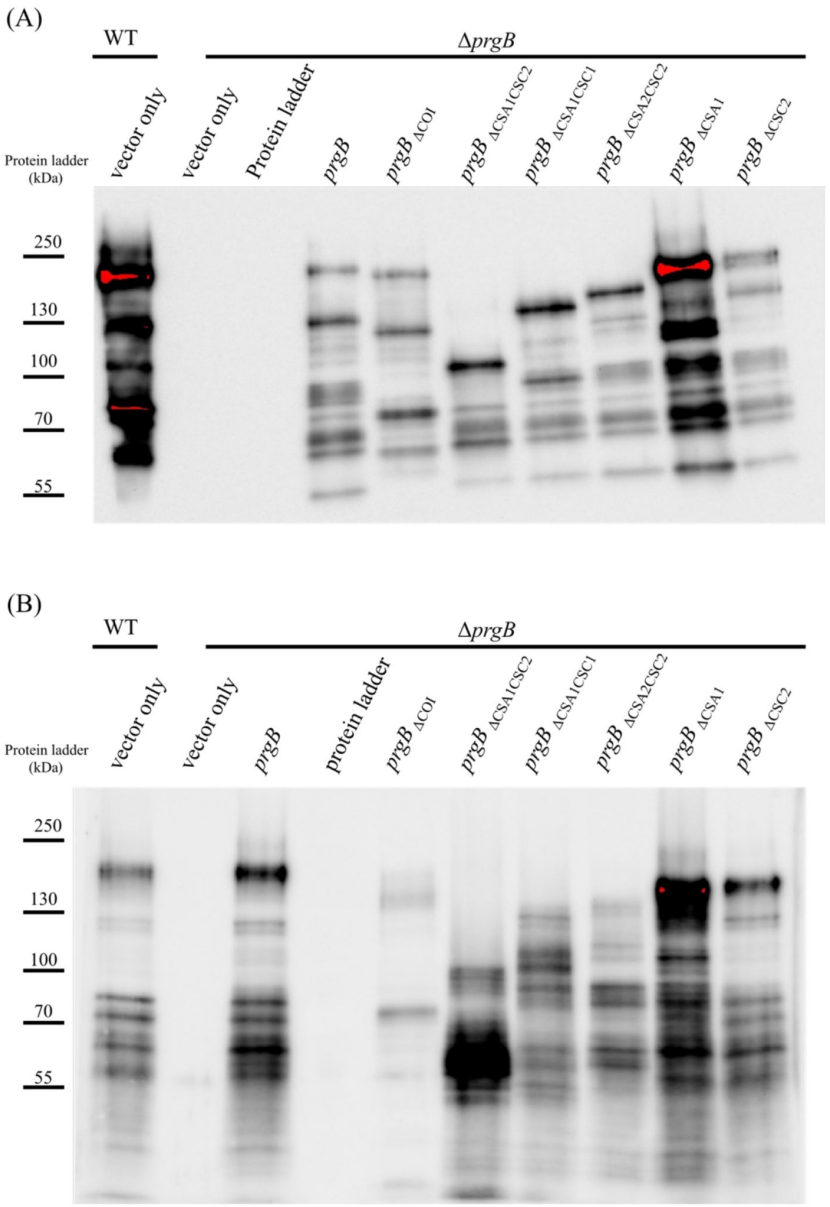

Figure 3 – figure supplement 1.

Western-blot showing the expression levels of the PrgB variants in *E. faecalis* cell wall extracts. (A) Samples collected after 1 hour induction with cCF10 and nisin. (B) Samples collected after overnight induction with cCF10 and nisin. Estimated molecular weight of PrgB variants: WT (140 kDa); ΔCOI (137 kDa); ΔCSA1CSC2 (72 kDa); ΔCSA1CSC1 (103 kDa); ΔCSA2CSC2 (105 kDa); ΔCSA1 (121 kDa); ΔCSA2 (122 kDa).

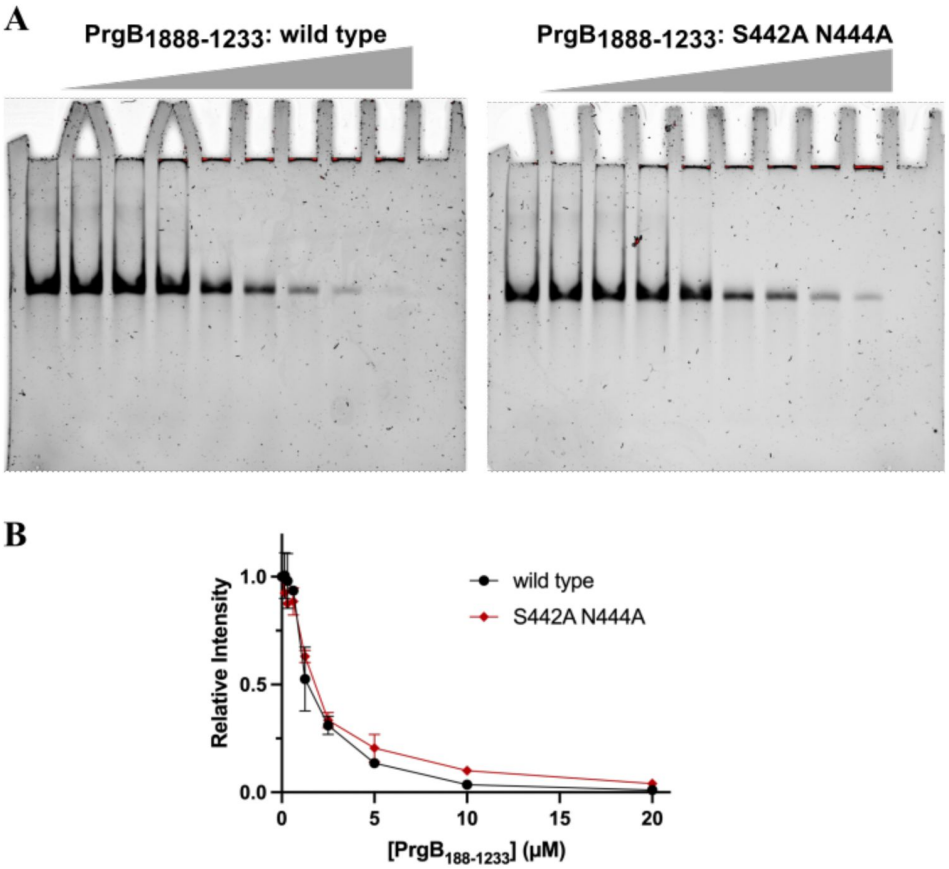

Figure 4 – figure supplement 1.

Comparison of the DNA-binding affinity of the wild-type polymer adhesion domain from PrgB ($PrgB_{188\text{-}1234\ WT}$) and its S442A & N444A variant ($PrgB_{188\text{-}1234\ S442A\ N444A}$). A) For this mobility shift assay, protein was mixed and incubated with 50 nM DNA_{100} and applied to native gels. From left to right, DNA mixtures with 0, 0.16, 0.31, 0.63, 1.25, 2.5, 5, 10 or 20 mM protein were loaded and the final lane is a negative control with only 20 mM protein (and no DNA). B) Plot of the relative intensities of the unbound (not upshifted) DNA bands normalized against the protein-free condition (Lane 1). The error is the standard deviation (N = 2).

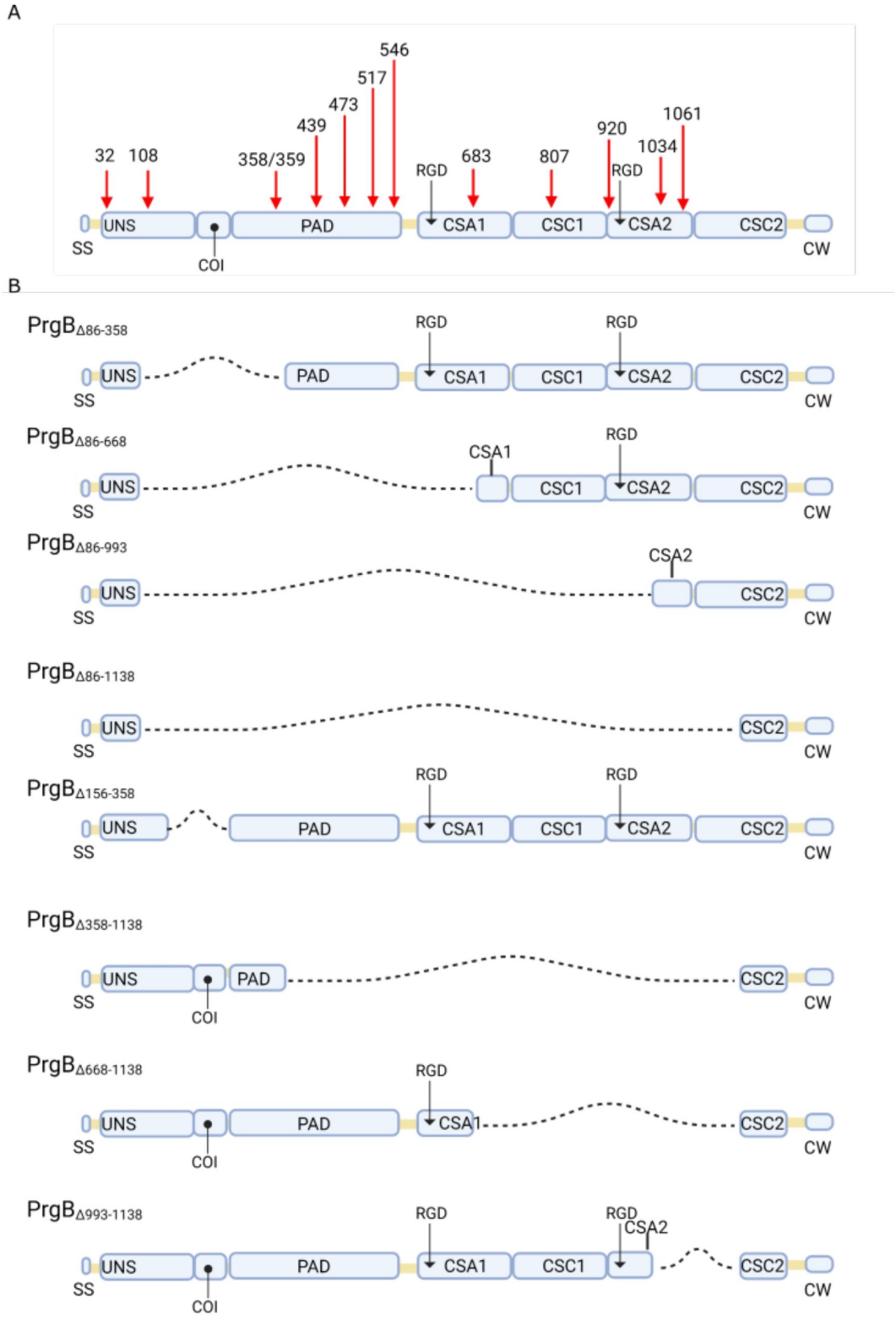

Figure 5 – figure supplement 1.

Schematic overview of previously described *prgB* mutants. A) Red arrows indicate the sites of transposon insertions that cause defective phenotypes in aggregation, biotic tissue biofilm formation, or conjugation. B) Diagrams illustrating the specific deletion mutants discussed in the main text. Dashed lines indicate deleted regions. SS, signal sequence; UNS, unstructured region; PAD, polymer adhesin domain; COI, coiled-coil domain; CSA, cell surface antigen; CSC, cell surface antigen C-terminus; CW, LPXTG-containing cell wall anchor sequence.

References

1. Gilmore M. S., Lebreton F., van Schaik W (2013) **Genomic transition of enterococci from gut commensals to leading causes of multidrug-resistant hospital infection in the antibiotic era** *Curr. Opin. Microbiol* **16**:10–16

2. Hidron A. I., Edwards J. R., Patel J., Horan T. C., Sievert D. M., Pollock D. A., et al. (2008) **NHSN annual update: antimicrobial-resistant pathogens associated with healthcare-associated infections: annual summary of data reported to the National Healthcare Safety Network at the Centers for Disease Control and Prevention, 2006-2007** *Infect. Control Hosp. Epidemiol* **29**:996–1011

3. Palmer K. L., Kos V. N., Gilmore M. S (2010) **Horizontal gene transfer and the genomics of enterococcal antibiotic resistance** *Current Opinion in Microbiology*

4. Palmer K. L., Gilmore M. S (2010) **Multidrug-resistant enterococci lack CRISPR-cas** *MBio* **1**

5. Trokter M., Waksman G (2018) **Translocation through the conjugative Type 4 secretion system requires unfolding of its protein substrate** *J. Appl. Microbiol* **200**:e00615–17

6. Cabezón E., Ripoll-Rozada J., Peña A., de la Cruz F., Arechaga I (2014) **Towards an integrated model of bacterial conjugation** *FEMS Microbiol. Rev*

7. Koraimann G (2018) **Spread and Persistence of Virulence and Antibiotic Resistance Genes: A Ride on the F Plasmid Conjugation Module** *EcoSal Plus* **8**

8. Grohmann E., Christie P. J., Waksman G., Backert S (2018) **Type IV secretion in Gram-negative and Gram-positive bacteria** *Mol. Microbiol* **107**:455–471

9. Hirt H., Manias D. A., Bryan E. M., Klein J. R., Marklund J. K., Staddon J. H., et al. (2005) **Characterization of the pheromone response of the Enterococcus faecalis conjugative plasmid pCF10: complete sequence and comparative analysis of the transcriptional and phenotypic responses of pCF10-containing cells to pheromone induction** *J. Appl. Microbiol* **187**:1044–1054

10. Dunny G. M., Berntsson R. P.-A (2016) **Enterococcal sex pheromones: Evolutionary pathways to complex, two-signal systems** *J. Bacteriol* **198**:1556–1562

11. Chandler J. R., Hirt H., Dunny G. M (2005) **A paracrine peptide sex pheromone also acts as an autocrine signal to induce plasmid transfer and virulence factor expression in vivo** *Proc. Natl. Acad. Sci. U. S. A* **102**:15617–15622

12. Dunny G. M., Berntsson R. P. A (2016) **Enterococcal sex pheromones: Evolutionary pathways to complex, two-signal systems** *J. Bacteriol* **198**:1556–1562

13. Dunny G. M (2013) **Enterococcal Sex Pheromones: Signaling, Social Behavior, and Evolution** *Annu. Rev. Genet* **47**:457–482

14. Lassinantti L., Camacho M. I., Erickson R. J. B., Willett J. L. E., De Lay N. R., ter Beek J., et al. (2021) **Enterococcal PrgU Provides Additional Regulation of Pheromone-Inducible Conjugative Plasmids** *mSphere* **6**:e00264–21

15. Schmitt A., Hirt H., Järvå M. A., Sun W. S., ter Beek J., Dunny G. M., et al. (2020) **Enterococcal PrgA Extends Far Outside the Cell and Provides Surface Exclusion to Protect against Unwanted Conjugation** *J. Mol. Biol* **432**:5681–5695

16. Järvå M. A., Hirt H., Dunny G. M., Berntsson R. P (2020) **Polymer Adhesin Domains in Gram-Positive Cell Surface Proteins** *Front. Microbiol* **11**

17. Bhatty M., Cruz M. R., Frank K. L., Gomez J. A. L., Andrade F., Garsin D. A., et al. (2015) **Enterococcus faecalis pCF10-encoded surface proteins PrgA, PrgB (aggregation substance) and PrgC contribute to plasmid transfer, biofilm formation and virulence** *Mol. Microbiol* **95**:660–677

18. Chuang O. N., Schlievert P. M., Wells C. L., Manias D. A., Tripp T. J., Dunny G. M (2009) **Multiple functional domains of Enterococcus faecalis aggregation substance Asc10 contribute to endocarditis virulence** *Infect. Immun* **77**:539–548

19. Schmitt A., Jiang K., Camacho M. I., Jonna V. R., Hofer A., Westerlund F., et al. (2018) **PrgB promotes aggregation, biofilm formation, and conjugation through DNA binding and compaction** *Mol. Microbiol* **109**:291–305

20. Olmsted S. B., Kao S. M., van Putte L. J., Gallo J. C., Dunny G. M (1991) **Role of the pheromone-inducible surface protein Asc10 in mating aggregate formation and conjugal transfer of the Enterococcus faecalis plasmid pCF10** *J. Appl. Microbiol* **173**:7665–7672

21. Olmsted S. B., Erlandsen S. L., Dunny G. M., Wells C. L (1993) **High-resolution visualization by field emission scanning electron microscopy of Enterococcus faecalis surface proteins encoded by the pheromone-inducible conjugative plasmid pCF10** *J. Appl. Microbiol* **175**:6229–6237

22. Süssmuth S. D., Muscholl-Silberhorn A., Wirth R., Susa M., Marre R., Rozdzinski E (2000) **Aggregation substance promotes adherence, phagocytosis, and intracellular survival of Enterococcus faecalis within human macrophages and suppresses respiratory burst** *Infect. Immun* **68**:4900–4906

23. Schlievert P. M., Chuang-Smith O. N., Peterson M. L., Cook L. C. C., Dunny G. M (2010) **Enterococcus faecalis Endocarditis Severity in Rabbits Is Reduced by IgG Fabs Interfering with Aggregation Substance** *PLoS One* **5**

24. Chuang-Smith O. N., Wells C. L., Henry-Stanley M. J., Dunny G. M (2010) **Acceleration of Enterococcus faecalis biofilm formation by aggregation substance expression in an ex vivo model of cardiac valve colonization** *PLoS One* **5**

25. Rakita R. M., Vanek N. N., Jacques-Palaz K., Mee M., Mariscalco M. M., Dunny G. M., et al. (1999) **Enterococcus faecalis bearing aggregation substance is resistant to killing by human neutrophils despite phagocytosis and neutrophil activation** *Infect. Immun* **67**:6067–6075

26. Vanek N. N., Simon S. I., Jacques-Palaz K., Mariscalco M. M., Dunny G. M., Rakita R. M (1999) **Enterococcus faecalis aggregation substance promotes opsonin-independent binding to human neutrophils via a complement receptor type 3-mediated mechanism** *FEMS Immunol. Med. Microbiol* **26**:49–60

27. Hirt H., Greenwood-Quaintance K. E., Barnes A. M. T., Karau M. J., Till L. M., Palzer E., et al. (2022) **Dynamics of plasmid-mediated niche invasion, immunity to invasion, and pheromone-inducible conjugation in the murine gastrointestinal tract** *Nat. Commun* **13**

28. Muscholl A., Galli D., Wanner G., Wirth R. (1993) **Sex pheromone plasmid pAD1-encoded aggregation substance of Enterococcus faecalis is positively regulated in trans by traE1** *Eur. J. Biochem* **214**

29. Galli D., Friesenegger A., Wirth R (1992) **Transcriptional control of sex-pheromone-inducible genes on plasmid pAD1 of Enterococcus faecalis and sequence analysis of a third structural gene for (pPD1 –encoded) aggregation substance** *Mol. Microbiol* **6**:1297–1308

30. Waters C. M., Dunny G. M (2001) **Analysis of functional domains of the Enterococcus faecalis pheromone-induced surface protein aggregation substance** *J. Appl. Microbiol* **183**:5659–5667

31. Waters C. M., Wells C. L., Dunny G. M (2003) **The Aggregation Domain of Aggregation Substance, Not the RGD Motifs, Is Critical for Efficient Internalization by HT-29 Enterocytes** *Infect. Immun* **71**:5682–5689

32. Waters C. M., Hirt H., McCormick J. K., Schlievert P. M., Wells C. L., Dunny G. M (2004) **An amino-terminal domain of Enterococcus faecalis aggregation substance is required for aggregation, bacterial internalization by epithelial cells and binding to lipoteichoic acid** *Mol. Microbiol* **52**:1159–1171

33. Kohler V., Keller W., Grohmann E (2018) **Enterococcus adhesin PrgB facilitates type IV secretion by condensation of extracellular DNA** *Mol. Microbiol* **109**:263–267

34. Mieher J. L., Larson M. R., Schormann N., Purushotham S., Wu R., Rajashankar K. R., et al. (2018) **Glucan binding protein C of Streptococcus mutans mediates both sucrose-independent and sucrose-dependent adherence** *Infect. Immun* **86**:1–18

35. Forsgren N., Lamont R. J., Persson K (2009) **Crystal structure of the variable domain of the Streptococcus gordonii surface protein SspB** *Protein Sci* **18**:1896–1905

36. Jumper J., Evans R., Pritzel A., Green T., Figurnov M., Ronneberger O., et al. (2021) **Highly accurate protein structure prediction with AlphaFold** *Nature* **596**

37. Holm L (2020) **DALI and the persistence of protein shape** *Protein Sci* **29**:128–140

38. Hall M., Nylander A., Jenkinson H. F., Persson K (2014) **Structure of the C-terminal domain of AspA (antigen I/II-family) protein from Streptococcus pyogenes** *FEBS Open Bio* **4**:283–289

39. Rego S., Heal T. J., Pidwill G. R., Till M., Robson A., Lamont R. J., et al. (2016) **Structural and functional analysis of cell wall-anchored polypeptide adhesin BspA in Streptococcus agalactiae** *J. Biol. Chem* **291**

40. Larson M. R., Rajashankar K. R., Crowley P. J., Kelly C., Mitchell T. J., Brady L. J., et al. (2011) **Crystal Structure of the C-terminal Region of Streptococcus mutans Antigen I/II and Characterization of Salivary Agglutinin Adherence Domains** *J. Biol. Chem* **286**:21657–21666

41. Forsgren N., Lamont R. J., Persson K (2010) **Two intramolecular isopeptide bonds are identified in the crystal structure of the Streptococcus gordonii SspB C-terminal domain** *J. Mol. Biol* **397**:740–751

42. Mieher J. L., Schormann N., Wu R., Patel M., Purushotham S., Wu H., et al. (2021) **Structure-function characterization of Streptococcus intermedius surface antigen pas** *J. Bacteriol* **203**

43. Brady L. J., Maddocks S. E., Larson M. R., Forsgren N., Persson K., Deivanayagam C. C., et al. (2010) **The changing faces of Streptococcus antigen I/II polypeptide family adhesins** *Mol. Microbiol* **77**:276–286

44. Maddocks S. E., Wright C. J., Nobbs A. H., Brittan J. L., Franklin L., Strömberg N., et al. (2011) **Streptococcus pyogenes antigen I/II-family polypeptide AspA shows differential ligand-binding properties and mediates biofilm formation** *Mol. Microbiol* **81**

45. Franklin L., Nobbs A. H., Bricio-Moreno L., Wright C. J., Maddocks S. E., Sahota J. S., et al. (2013) **The AgI/II Family Adhesin AspA Is Required for Respiratory Infection by Streptococcus pyogenes** *PLoS One* **8**

46. Purushotham S., Deivanayagam C (2014) **The Calcium-induced Conformation and Glycosylation of Scavenger-rich Cysteine Repeat (SRCR) Domains of Glycoprotein 340 Influence the High Affinity Interaction with Antigen I/II Homologs** *J. Biol. Chem* **289**:21877–21887

47. Daep C. A., Novak E. A., Lamont R. J., Demuth D. R (2011) **Structural dissection and in vivo effectiveness of a peptide inhibitor of porphyromonas gingivalis adherence to streptococcus gordonii** *Infect. Immun* **79**

48. Halaby D. M., Mornon J. P. E (1998) **The immunoglobulin superfamily: An insight on its tissular, species, and functional diversity** *J. Mol. Evol* **46**

49. Olmsted S. B., Dunny G. M., Erlandsen S. L., Wells C. L (1994) **A Plasmid-Encoded Surface Protein On Enterococcus Faecalis Augments Its Internalization By Cultured Intestinal Epithelial Cells** *J. Infect. Dis* **170**:1549–1556

50. Bae T., Kozlowicz B., Dunny G. M (2002) **Two targets in pCF10 DNA for PrgX binding: their role in production of Qa and prgX mRNA and in regulation of pheromone-inducible conjugation** *J. Mol. Biol* **315**:995–1007

51. Laverde Gomez J. A., Bhatty M., Christie P. J. (2014) **PrgK, a Multidomain Peptidoglycan Hydrolase, Is Essential for Conjugative Transfer of the Pheromone-Responsive Plasmid pCF10** *J. Appl. Microbiol* **196**:527–539

52. Geertsma E. R., Dutzler R (2011) **A versatile and efficient high-throughput cloning tool for structural biology** *Biochemistry* **50**:3272–3278

53. Kabsch W. (2010) **XDS. Acta Crystallogr** *Sect. D* **66**:125–132

54. Emsley P., Lohkamp B., Scott W. G., Cowtan K. (2010) **Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. 66, 486–501 (2010).** *Features and development of Coot* **66**:486–501

55. Winn M. D., Isupov M. N., Murshudov G. N (2001) **Winn, M. D., Isupov, M. N. & Murshudov, G. N. Use of TLS parameters to model anisotropic displacements in macromolecular refinement. 57, 122–133 (2001).** *Use of TLS parameters to model anisotropic displacements in macromolecular refinement* **57**:122–133

56. Adams P. D., Grosse-Kunstleve R. W., Hung L. W., Ioerger T. R., McCoy A. J., Moriarty N. W., et al. (2002) **Adams, P. D., Grosse-Kunstleve, R. W., Hung, L. W., Ioerger, T. R., McCoy, A. J., Moriarty, N. W., et al. PHENIX: building new software for automated crystallographic structure determination. 58, 1948–1954 (2002).** *PHENIX: building new software for automated crystallographic structure determination* **58**:1948–1954

57. Chen V. B., Arendall W. B., Headd J. J., Keedy D. A., Immormino R. M., Kapral G. J., et al. (2010) **Chen, V. B., Arendall, W. B., Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., et al. MolProbity: all-atom structure validation for macromolecular crystallography. 66, 12–21 (2010).** *MolProbity: all-atom structure validation for macromolecular crystallography* **66**:12–21

58. Punjani A., Brubaker M. A., Fleet D. J (2017) **Building Proteins in a Day: Efficient 3D Molecular Structure Estimation with Electron Cryomicroscopy** *IEEE Trans. Pattern Anal. Mach. Intell* **39**

59. Pettersen E. F., Goddard T. D., Huang C. C., Couch G. S., Greenblatt D. M., Meng E. C., et al. (2004) **UCSF Chimera – A visualization system for exploratory research and analysis** *J. Comput. Chem* **25**

60. Kidmose R. T., Juhl J., Nissen P., Boesen T., Karlsen J. L., Pedersen B. P (2019) **Namdinator – Automatic molecular dynamics flexible fitting of structural models into cryo-EM and crystallography experimental maps** *IUCrJ* **6**

61. Willett J. L. E., Ji M. M., Dunny G. M (2019) **Exploiting biofilm phenotypes for functional characterization of hypothetical genes in Enterococcus faecalis** *npj Biofilms Microbiomes* **5**

62. Christie P. J., Kao S. M., Adsit J. C., Dunny G. M (1988) **Cloning and expression of genes encoding pheromone-inducible antigens of Enterococcus (Streptococcus) faecalis** *J. Appl. Microbiol* **170**:5161–5168

63. Chen Y., Staddon J. H., Dunny G. M (2007) **Specificity determinants of conjugative DNA processing in the Enterococcus faecalis plasmid pCF10 and the Lactococcus lactis plasmid pRS01** *Mol. Microbiol* **63**:1549–1564

64. Chen Y., Zhang X., Manias D., Yeo H.-J., Dunny G. M., Christie P. J (2008) **Enterococcus faecalis PcfC, a spatially localized substrate receptor for type IV secretion of the pCF10 transfer intermediate** *J. Appl. Microbiol* **190**:3632–3645

Article and author information

Wei-Sheng Sun
Department of Medical Biochemistry and Biophysics, Umeå University, SE-90187 Umeå, Sweden, Wallenberg Centre for Molecular Medicine, Umeå University, Umeå, Sweden
ORCID iD: 0000-0001-9738-8862

Lena Lassinantti
Department of Medical Biochemistry and Biophysics, Umeå University, SE-90187 Umeå, Sweden
ORCID iD: 0000-0001-5470-591X

Michael Järvå
Department of Medical Biochemistry and Biophysics, Umeå University, SE-90187 Umeå, Sweden
ORCID iD: 0000-0002-8519-5622

Andreas Schmitt
Department of Medical Biochemistry and Biophysics, Umeå University, SE-90187 Umeå, Sweden

Josy ter Beek
Department of Medical Biochemistry and Biophysics, Umeå University, SE-90187 Umeå, Sweden, Wallenberg Centre for Molecular Medicine, Umeå University, Umeå, Sweden
ORCID iD: 0000-0003-4165-9277

Ronnie P-A Berntsson
Department of Medical Biochemistry and Biophysics, Umeå University, SE-90187 Umeå, Sweden, Wallenberg Centre for Molecular Medicine, Umeå University, Umeå, Sweden
For correspondence: ronnie.berntsson@umu.se
ORCID iD: 0000-0001-6848-322X

Copyright

Editors

Reviewing Editor
Michael Eisen
University of California, Berkeley, United States of America

Senior Editor
Michael Eisen
University of California, Berkeley, United States of America

Reviewer #1 (Public Review):

Sun and colleagues outline structural and mechanistic studies of the bacterial adhesin PrgB, an atypical microbial cell surface-anchored polypeptide that binds DNA. The manuscript includes a crystal structure of the Ig-like domains of PrgB, cryo-EM structures of the majority of the intact polypeptide in DNA-bound and free forms, and an assessment of the phenotypes of E. faecalis strains expressing various PrgB mutants. Generally, the study has been conducted with a good level of rigor, and there is consistency in the findings. Initial concerns about inferences initially made from low-resolution Cryo-EM structures have been addressed experimentally and the manuscript correspondingly updated.

Reviewer #2 (Public Review):

Having previously solved the X-ray crystallographic structure of the polymer adhesin domain (PAD) of PrgB from E. faecalis, the authors looked to build on that work by crystallizing a

nearly full-length construct of PrgB. Though they were successful in their crystallization endeavors, the crystal contained only what was previously thought to be two domains with RGD motifs. The authors' high-resolution structure shows that in fact the C-terminal portion of PrgB is made up of four immunoglobulin-like domains. The authors then set out to collect single-particle cryoEM data in a bid to obtain a full-length structure of PrgB, both in the presence and absence of ssDNA. The authors were only able to obtain quite low-resolution data, which they fit their crystal structures into. The authors then used these structures to inform the design of novel deletion mutants and point mutations, as well as to rationalize years of phenotypic data from other published mutants.

The X-ray crystallographic structure is beautiful and in combination with their in vivo data allowed them to propose a model where PrgB positions cells at an appropriate distance for conjugation. The in vivo experiments appear to be done well and the authors' discovery that the Ser-Asn-Glu is not important for generalized aggregation but has an additional yet unknown role in conjugation and biofilm formation is exciting and well supported by their data.

[Editors' note: In response to reviews of a previous version of this manuscript, the authors have carried out additional experiments that have strengthened the already convincing aspects of the work. We commend the authors for responding to questions raised by the reviewers about the inference of interactions of in vivo importance inferred from low-resolution cryo-EM studies by carrying out and reporting on additional experiments that fail to confirm their initial speculative model. The current work is stronger and more convincing as a result.]

Author Response

The following is the authors' response to the original reviews.

First of all, we would like to again thank the reviewers for their work. We appreciate the constructive review comments and useful suggestions to further improve our article. With those comments in mind, we have now revised our manuscript. Please see below for a point-by-point response (our responses in green) to all comments.

> ***Reviewer #1 (Recommendations For The Authors):***
>
> *Sun and colleagues outline structural and mechanistic studies of the bacterial adhesin PrgB, an atypical microbial cell surface-anchored polypeptide that binds DNA. The manuscript includes a crystal structure of the Ig-like domains of PrgB, cryo-EM structures of the majority of the intact polypeptide in DNA-bound and free forms, and an assessment of the phenotypes of E. faecalis strains expressing various PrgB mutants.*
>
> *Generally, the study has been conducted with a good level of rigor, and there is consistency in the findings. However, I do have some specific technical concerns relating to the study that necessitate the undertaking of additional experiments. These are summarized as follows:*
>
> 1. *Recombinant PrgB188-1233 produced in the study purifies as a mixture of monomeric and dimeric species separatable by SEC. There is very limited discussion in the text re. the significance and/or implications of this. Is it feasible that the dimeric form is biologically relevant in the context of the in vivo situation? Or alternatively, is this simply an artifact of protein production?*

Experimental data that we published in 2018 indeed indicates that the dimer is relevant in the in vivo situation. We did not discuss this here since this was discussed in detail in the

previous paper: Schmitt et al, 2018. We have now added a bit more information on this in the results section, highlighting this, so that it is clearer to the reader (lines 114-116).

> 1. *The authors see no evidence of the adhesive domain of PrgB in their PX structure highlighting that this must have been cleaved during crystallisation. Is this claim supported by an inspection of the crystal packing? It could be that this region of the protein is dynamic within the context of the crystal and is thus not observed. This should be clarified in the text either way.*

The crystal packing does not provide any space for the PAD. We have added this to the results section. We have added a sentence describing this in lines 122-124.

> 1. *The Cryo-EM structures reported are both at ~10-angstrom resolution. Are the authors truly confident in the placement of their crystal structures on these maps? Visual inspection indicates that their positioning of the PrgB domains into the EM envelopes is somewhat questionable. The authors need to provide some quantitative measures of the quality of their domain fitting. The narrative of the manuscript very much hinges on this being correct.*

This is something that the other reviewer also commented on. The fitting of the crystal structures in the maps are indeed not optimal, but was the best we could do with the available data. In line with point #6, we have now constructed new protein variants of the stalk domain (the four Ig-like domains) alone, and have assayed it's interaction with the PAD in vitro using native gels and size exclusion chromatography. The outcome of these experiments is that the two domains do not interact in any substantial way on their own. Thus, the added experiments do not support the hypothesis that the PAD interacts with the Ig-like domains, at least not without the local high concentration provided by the linker region in the in vivo situation.

To account for these new experiments, we have moved the cryo-EM structure to the supplement, and rewritten this part of the manuscript to say that the cryo-EM data indicated that there might be an interaction, but that we have not been able to verify this in vitro, indicating that if the interaction at all exists it must have a low affinity and is likely not physiologically relevant. In line with this, we have also further modified the text throughout the manuscript to account for this.

> 1. *The manuscript would be significantly strengthened if the authors could include confirmatory hydrodynamic data in support of the observed conformational reorganization of PrgB in the presence of DNA. SAXS analysis of the DNA-free and bound complexes would be ideal for this and would also help address the issues raised above in pt 3.*

To analyze PrgB radius with and without DNA, we tried both SEC-MALS and DLS experiments. It proved difficult to obtain precise and reproducible values, but the initial data indicated that no large changes were observed upon DNA binding. As we could also not measure specific interaction between the PAD and the stalk in vitro, we did not perform SAXS experiments. As mentioned in the response to point #3, we have modified the results and discussion regarding the potential interaction of th PAD and Stalk domains.

> 1. *The authors present binding studies of various PrgB mutant-expressing strains. A number of the mutations generated delete significant portions of the polypeptide. Can the authors confirm that these mutant proteins are correctly folded despite the introduced mutations? It could be that loss of function is simply a consequence of mutation-induced misfolding. I would like to see some confirmatory data (CD, SEC, etc.) in support of the foldedness of the mutant proteins.*

We cannot completely rule out that the folding of some of the variants is affected in E. faecalis. However, CD or SEC experiments would only give indications of the contrary if the overall fold had been majorly affected in an in vitro situation where the protein is not anchored to the E. faecalis cell wall.

To alleviate this valid concern, we probed if all variants are correctly exported and linked to the cell-wall. Therefore we have now extracted the cell wall of E. faecalis producing wild-type or variant PrgB and performed Western blot . The results of the Western blot with cell wall extract largely matches the whole cell experiments that were in the initial manuscript. If a protein variant was largely misfolded, it would likely not be targeted and linked to the cell-wall, nor would it be stable in vivo. We have added this new data as a new fig 3 – figure supplement 1 and on lines 201-214

> 1. *The authors suggest a direct interaction between the PAD and the stalk domains in PrgB. The discussion of this is very generic and no evidence to support this is provided other than the 10-angstrom resolution EM map. If they believe this to be the case, then additional evidence should be provided.*

Answer: As mentioned previously, we have now performed additional in vitro experiments to probe this potential interaction, but conclude that this indication from the EM data is likely not a real high affinity interaction. In line with this, we have modified the results and discussion regarding this point, see also response to point #3 and 4.

> ***Reviewer #2 (Recommendations For The Authors):***
>
> *As currently presented, I don't feel that the cryoEM data support the authors' proposed model, largely because the fit of the crystal structures to the EM volumes does not seem entirely reasonable for the apo- dataset and because the EM volume for the ssDNA bound dataset is not even contiguous. For me to believe the model as it is currently built, I would want to see a dataset with the PAD deleted, showing that its proposed density disappears, or a dataset with a PAD-specific antibody as a fiducial marker. It would be nice to see some goodness of fit metric with a comparison to other crystal structures fit such low-resolution data as well. At the very least, the authors must include the standard cryoEM workflow supplementary figure showing representative micrographs, 2Ds, and 3Ds along with particle numbers.*

In line with the comments raised by reviewer #1, we have now added more experiments where we have analyzed the potential interaction between PAD and the stalk domain. From this new data, it looks like they do not interact with any substantial affinity, at least not on their own without any linker region holding them together, and that this interaction if it all exist likely is not physiologically relevant. The cryo-EM data has been moved to the supplement as we agree with both reviewers that the resolution, and the fitted model, is not good enough to draw any hard conclusions. The standard table for the cryoEM workflow was present as supplementary table 2, where eg particle numbers etc are described, but we have now also added a new supplementary fig 2 – figure supplement 2 that shows the EM

processing workflow, including representative micrographs, 2D and 3D classes. We debated whether we should remove the EM data, but decided against it in line of transparency and to explain why the interaction studies with the PAD and stalk domains were performed.

> *The X-ray crystallographic structure is very nice, but I was a bit surprised by the R factors in Table 1. After downloading the structure factors and coordinates from the PDB (thank you for depositing before submission!) I was able to see quite a few positive peaks in the difference map that could probably use some cleaning up. I realize I may just be a bit of a masochist when it comes to adding/deleting waters and moving around side chains to get things just right, but for such lovely data, I would have liked to see the model polished up a bit more. I was going to say that the isopeptide bond should be modelled, but I can see from a cursory Google that the authors did in fact try to find a way to model this and that it is indeed a bit of a pain.*

The model refinement proved surprisingly recalcitrant with regards to the remaining difference density, so we took the decision to only model what was solidly there (which leads to slightly higher R factors). We did indeed try to model the isopeptide bond, but we did not find a good way to do so (despite trying quite extensively), and ended up determining them as a linker in the PDB file, so that the bond shows up when one opens the structure in eg. Pymol.

> *For protein production/purification in general I would have liked to see actual traces for the gel filtration and pure protein on a gel in a supplementary figure. I strongly believe that this type of information is so critical for future researchers looking to replicate or build upon published work so that they have some sense that what they are doing is working in the way it should be.*

We have now added a supplementary figure (as new Fig. 1 – figure supplement 1) that shows SEC and SDS-PAGE for the purification of PrgB188-1233.

> *Finally, I think for the in vivo data it only makes sense to show the reader whether any or all the differences measured across your different mutants are statistically significant. Having done the graphing and analysis in GraphPad this should be a simple thing to achieve.*

We have now added statistical test (One way Anova) that show the statistical significance between the mutants, and show that in Fig 3 and Fig 4.

> *Overall, I think it's a very nice paper and while I feel that the cryoEM data in its current form doesn't support the model of occlusion from PrgA, I also don't think that removing the cryoEM data and that specific mechanistic idea from the paper detracts from its overall message and impact.*

Thank you for those comments.