The BR domain of PsrP interacts with extracellular DNA to promote bacterial aggregation; structural insights into pneumococcal biofilm formation

Tim Schulte1, Cecilia Mikaelsson1, Audrey Beaussart2,3, Alexey Kikhney3, Maya Deshmukh1, Sebastian Wolniak1, Anuj Pathak1, Christine Ebel5, Jonas Löfling4, Federico Fogolari6, Birgitta Henriques-Normark4, Yves F. Dufrêne2, Dmitri Svergun3, Per-Åke Nygren7 & Adnane Achour1

The major human pathogen Streptococcus pneumoniae is a leading cause of disease and death worldwide. Pneumococcal biofilm formation within the nasopharynx leads to long-term colonization and persistence within the host. We have previously demonstrated that the capsular surface-associated pneumococcal serine rich repeat protein (PsrP), key factor for biofilm formation, binds to keratin-10 (KRT10) through its microbial surface component recognizing adhesive matrix molecule (MSCRAMM)-related globular binding region domain (BR187–385). Here, we show that BR187–385 also binds to DNA, as demonstrated by electrophoretic mobility shift assays and size exclusion chromatography. Further, heterologous expression of BR187–378 or the longer BR120–378 construct on the surface of a Gram-positive model host bacterium resulted in the formation of cellular aggregates that was significantly enhanced in the presence of DNA. Crystal structure analyses revealed the formation of BR187–385 homo-dimers via an intermolecular β-sheet, resulting in a positively charged concave surface, shaped to accommodate the acidic helical DNA structure. Furthermore, small angle X-ray scattering and circular dichroism studies indicate that the aggregate-enhancing N-terminal region of BR120–166 adopts an extended, non-globular structure. Altogether, our results suggest that PsrP adheres to extracellular DNA in the biofilm matrix and thus promotes pneumococcal biofilm formation.

The human adapted, Gram-positive commensal bacterium Streptococcus pneumoniae (pneumococcus) colonizes the upper respiratory tract in about 10% of healthy adults and up to 60% of children without necessarily causing any symptoms1. The organization of bacterial communities into complex and dynamic biofilm structures allows for growth in a hostile environment, providing protection against anti-microbial agents2,3. Pneumococcal colonization is the precursor for otitis media and sinusitis, as well as severe diseases including pneumonia, septicemia and meningitis4. Pneumococcal biofilms are found on the surface of adenoid and mucosal epithelial tissues in

1Science for Life Laboratory, Department of Medicine Solna, Karolinska Institute, and Department of Infectious Diseases, Karolinska University Hospital, Solna, SE-17176 Stockholm, Sweden. 2Université catholique de Louvain, Institute of Life Sciences, Croix du Sud, 4-5, bte L7.07.06, B-1348 Louvain-la-Neuve, Belgium. 3European Molecular Biology Laboratory (EMBL), Hamburg Outstation, Notkestrasse 85, 22603 Hamburg, Germany. 4Department of Microbiology, Tumor and Cell Biology (MTC), Karolinska Institute; Clinical Microbiology, Karolinska University Hospital Solna, Stockholm, Sweden. 5Institut de Biologie Structurale (IBS), Univ. Grenoble Alpes, CEA, CNRS, 38044 Grenoble, France. 6Dipartimento di Scienze Mediche e Biologiche, Università di Udine, Piazzale Kolbe 4, 33100 Udine - Italy. 7Division of Protein Technology, School of Biotechnology, KTH–Royal, Institute of Technology, Sweden. *Present address: CNRS, LIEC (Laboratoire Interdisciplinaire des Environnements Continentaux), UMR 7360, Vandoeuvre-les-Nancy, F-54501, France. Correspondence and requests for materials should be addressed to T.S. (email: tim.schulte@ki.se) or A.A. (email: adnane.achour@ki.se)
children with recurrent middle-ear infections and otitis media with effusion, highlighting the role of microbial biofilms in disease development.

Recently, studies using *in vitro* and *in vivo* pneumococcal biofilm models have begun to unveil the biogenesis, regulation, structure and composition of pneumococcal biofilms. The exact composition of the pneumococcal biofilm extracellular polymer matrix (EPM) depends largely on the species and environmental conditions. Most often, the EPM is composed of exo-polysaccharides, proteins, nucleic acids and lipids, and its overall structure is stabilized through intermolecular interaction networks, often involving non-catalytic carbohydrate- or eDNA-binding proteins such as lectins and pilins. In pneumococcus, the amount of capsular polysaccharides is highly regulated and significantly reduced during colonization compared to invasion. Non-encapsulated strains are hyper-adhesive to epithelial cells and form biofilms more efficiently than encapsulated strains. On the other hand, they are also less virulent. The increased adhesive phenotype in non-encapsulated strains can possibly be attributed to an increased presentation of adhesion proteins.

The pneumococcal serine rich repeat protein (PsrP), present in 60% of strains capable of causing pneumonia in children, has been identified as a key factor for biofilm formation using *in vitro* and *in vivo* models. A pneumococcal ΔPsrP deletion mutant strain was unable to form biofilm aggregates in an *in vivo* mice colonization model and elicited a significantly enhanced immune response compared to the wild-type strain. PsrP belongs to the serine rich repeat protein (SRRP) family found in Gram-positive bacteria. SRRPs are attached to the capsular surface via a cell wall anchoring domain and display a long, highly repetitive and glycosylated C-terminal serine rich-repeat (SRR) region. The SRR region varies between 400 and 4000 residues in length, extending the functional binding region (BR) domain out of the capsule (Fig. 1A). SRRP BR domains which bind to a broad range of targets including glyco-conjugates and keratins, are variable in sequence and organized into modular domains. Fold topologies similar to the microbial surface component recognizing adhesive structures of self-oligomerizing B-repeat regions of both proteins revealed remarkable "free-standing" β-sheets. Recent studies demonstrated that BR 122–166 is released from PsrP through proteolytic cleavage by the human furin protease, a known maturation factor of other toxins and virulence factors. The discovered molecular mechanisms provide novel structural and mechanistic insights into the role of PsrP during pneumococcal biofilm formation.

**Results and Discussion**

The BR domain comprises a furin protease recognition site and forms irreversibly associated dimers only during heterologous expression. Sequence analysis and biochemical protease cleavage assays revealed that the sequence motif K164RRK168 localized between the KRT10-binding region and the predicted non-globular N-terminal extension is cleaved by the human furin protease (Fig. 1A,B). While wild-type BR120–385 was specifically cleaved by the human furin protease and prone to degradation during purification, a BR120–385 construct with a mutated furin site was not cleaved and significantly more stable compared to wild-type, with no degradation product observed during purification (data not shown). Furin-like proteases activate a large number of pro-protein substrates as well as bacterial and viral pathogenic agents. Thus, we hypothesized that the N-terminal fragment comprising BR120–166 and SRR could be released following proteolysis by furin, similarly to the previously described shedding of the glycoprotein Flo11p from *Saccharomyces cerevisiae* cells into the extracellular matrix of yeast mats.
We have recently determined the crystal structure of the KRT10-binding domain of PsrP (BR\textsubscript{187–385}), revealing a novel MSCRAMM-related DEv-IgG fold resembling a compressed barrel\textsuperscript{20}. The crystals were obtained from a monomeric preparation of BR\textsubscript{187–385} that did not form higher oligomers in solution as assessed by analytical ultracentrifugation (AUC) and small angle X-ray scattering (SAXS) experiments at different concentrations\textsuperscript{20}. However, during purification of the heterologously produced BR\textsubscript{187–385}, higher oligomer states were formed in the final size exclusion chromatography (SEC) step (Fig. 1C). Interestingly, these isolated oligomer populations were stable in solution but not in steady-state equilibrium with the monomer, as demonstrated using AUC and SEC (Figs 1C and S1A, Table S1). Similarly, the larger BR\textsubscript{120–395} construct produced stable higher oligomer species, but at a significantly reduced ratio (Figs 1D and S1B, Table S1).

Since stable oligomer formation of PsrP could be relevant for pneumococcal biofilm formation, crystallization trials were set up for the isolated BR\textsubscript{187–385} dimer in order to obtain a molecular understanding for stable dimer formation. However, crystals were hard to reproduce and those obtained differed drastically. Therefore, a shorter construct was designed by deleting the seven C-terminal residues in BR\textsubscript{187–385} that were not built in the previously determined crystal structure due to missing electron density\textsuperscript{20}. Well-diffracting crystals were obtained of the isolated irreversibly associated BR\textsubscript{187–378} dimer. Analysis of the crystal structure solved by molecular replacement revealed that the residue stretches L\textsubscript{202}–G\textsubscript{315} (region I) and Y\textsubscript{316}–S\textsubscript{377} (region II) are reciprocally exchanged between two separate chains (Fig. S2, Table S2). This reciprocal exchange of entire protein segments is termed 3D-domain swapping and has been first described for diphtheria toxin\textsuperscript{32,33}. A common feature of domain-swapped oligomers is the requirement to pass a high-energy barrier between the non-swapped and swapped states. Such inter-conversions usually do not occur spontaneously, but require harsh and denaturing conditions.

![Figure 1](https://www.nature.com/scientificreports/images/1f1.jpg)
Thus, it should be noted that the irreversibly associated dimers were only obtained following heterologous production of the isolated BR protein constructs, and that, at this stage, we remain uncertain about their formation in vivo.

The putative self-oligomerization region BR\textsubscript{122–166} is non-globular and does not promote self-aggregation. In order to obtain structural information about the longer BR\textsubscript{120–395} construct in solution, comparative small angle X-ray scattering (SAXS) data were collected for both monomeric BR\textsubscript{187–385} and BR\textsubscript{120–395} (Fig. 2A,B and Table S1). The directly accessible and sample-characteristic radii of gyration ($R_g$) and maximal dimensions ($D_{\text{max}}$) of BR\textsubscript{187–385} and BR\textsubscript{120–395} were $20 \pm 1 \, \text{ Å}$ and $29 \pm 2 \, \text{ Å}$ as well as $78 \pm 8 \, \text{ Å}$ and $125 \pm 12 \, \text{ Å}$, respectively, indicating a more extended structure for BR\textsubscript{120–395} (Fig. S3A,B). Indeed, molecular models obtained from the ensemble optimization method (EOM) support the presence of a single globular domain common to BR\textsubscript{187–385} and BR\textsubscript{120–395}, but with a longer highly mobile structural extension for BR\textsubscript{120–395} (Fig. 2C,D). In dimensionless Kratky plots, the curve of BR\textsubscript{187–385} closely resembles the bell-shaped curve expected for a globular protein, while the curve of BR\textsubscript{120–395} is clearly shifted towards the shape of a molecule that comprises a more disordered, random chain (Fig. S3C). This is in line with our interpretation that the longer BR\textsubscript{120–395} construct comprises an extended, highly flexible section in solution compared to BR\textsubscript{187–385}. Additional structural information was obtained from circular dichroism spectra of BR\textsubscript{187–385} and BR\textsubscript{120–395} that also indicated the presence of disordered and helical regions for the additional residues in BR\textsubscript{120–395} (Fig. S3D). Combining the structural information obtained from AUC, SAXS, and CD, we conclude that the N-terminal region in BR\textsubscript{120–395} does not promote aggregation and corresponds to a non-globular, largely disordered structure.

BR\textsubscript{187–385} forms low-affinity homo-dimers through intermolecular β-sheets. Considering that the described domain swap mechanism is unlikely to occur spontaneously in vivo, and that the previously suggested self-aggregating BR\textsubscript{120–166} region does not promote self-oligomerization in our experiments, we looked back into the crystal structure of the isolated BR\textsubscript{187–385} monomer\textsuperscript{20} for hints of a more probable oligomerization mechanism. Indeed, we discovered intermolecular β-sheets between symmetry-related monomeric molecules in the P4\textsubscript{2}2 crystal form (PDB: 3ZGH) and in the P4\textsubscript{2}2 crystal form (PDB: 3ZGI) forming dimers with interface surface
mote cellular aggregation, BR187–378 and BR*120–395 were fused separately to the SasC-derived DUF-domain linker. 

The apparent affinity measured for binding of BR187–385 to DNA was in the low micromolar range. Thus, although we could prevent its detection in the performed SEC, SAXS and AUC experiments (Figs 1C and 2A and S1B)

underlying molecular recognition in protein-protein interactions34, a predicted weak interaction of this interface (Fig. S4B,C). However, although the formation of such edge-to-edge interactions is regarded as a fundamental feature underlying molecular recognition in protein-protein interactions34, a predicted weak interaction of this interface could prevent its detection in the performed SEC, SAXS and AUC experiments (Figs 1C and 2A and S1B)

Single-cell and single-molecule atomic force microscopy have recently been applied to demonstrate that low-affinity (15 mM) homophilic interactions of the fibronectin binding protein A (FnBPA) of Staphylococcus aureus promote cellular aggregation34. Therefore, we also applied single-molecule force spectroscopy to test the capacity of BR187–378 and BR*120–395 to mediate similar low-affinity homophilic bonds. Indeed, homophilic binding events were measured for BR187–378 and BR*120–395 with mean adhesion forces of 70 ± 14 pN and 68 ± 14 pN, respectively (Fig. S5A–C). And as expected for specific bimolecular bonds35, the adhesion forces increased linearly with the logarithm of the loading rate (Fig. S5D).

BR187–385 binds to DNA and forms a positively charged saddle that snugly fits the acidic helical structure of double-stranded DNA. Although the measured weak homophilic binding forces could be sufficient for cellular aggregation, we hypothesized that the highly basic BR187–385 could also bind to acidic biofilm-associated extracellular DNA. Indeed, electrophoretic shift assays (EMSA) demonstrated binding of a randomly chosen 276 bp long DNA (DNA276bp) molecule to BR187–385 with an apparent affinity in the lower micromolar range (Fig. 4A). After incubation of BR187–378 with DNA276bp at a 1:750 molar ratio, the purified high-molecular weight DNA population also comprised bound BR187–378 molecules (Fig. 4B,C). Interestingly, the DNA-bound BR187–385 formed a band at a higher molecular weight in SDS-PAGE, indicating that the DNA was not completely removed even by lauryl-dodecyl sulfate detergent treatment. Based on our previously suggested ligand-binding model, it is tempting to speculate that the negatively charged DNA binds to the extended highly basic groove of the intermolecular β-sheet dimer of BR187–385 (Fig. 4D). Molecular dynamic (MD) simulations also suggested that the docked DNA:BR187–385 complex is flexible but most of the non-nucleotide specific protein-DNA contacts are preserved over a 50 ns simulation period (Movie S1).

Strikingly, the structure of the BR187–385 dimer resembles the molecular saddle of the transcriptional pre-initiation complex-associated TATA-box binding protein (TBP) (Fig. 4E). While the concave under-surface of the TBP saddle is also highly basic, the specificity of the protein:TATATAAA interaction is primarily mediated through hydrophobic interactions36,37. However, it is well established that the specificity of protein-DNA interactions is manifested entirely in the non-electrostatic interaction component, while the electrostatic component typically contributes to the majority of the affinity38. Similar to the interaction of TBP with non-related DNA, the apparent affinity measured for binding of BR187–385 to DNA was in the low micromolar range. Thus, although we cannot exclude the possibility that PsrP recognizes a specific DNA sequence, we here hypothesize that PsrP binds non-specifically to DNA.

BR187–385 promotes DNA-dependent bacterial aggregation. Display of heterologously produced proteins on the surface of the Gram-positive bacterium Staphylococcus carnosus represents an excellent model system to study the functions of surface-associated proteins, e.g. the biofilm-promoting function of the SasC protein of Staphylococcus aureus39,40. In order to assess the capacity of each segment of the BR domain to promote cellular aggregation, BR187–378 and BR187–385 were fused separately to the SasC-derived DUF-domain linker followed by a C-terminal LPxTG motif that becomes covalently attached to the staphylococcal cell wall through Sortase-mediated enzymatic linkage (Fig. 5A). The DUF domain was chosen to substitute the role of the SRR domain of native PsrP, which is to extend the BR region away from the cellular surface for functional accessibility. Dotblot assays of 3C-protease-cleaved surface-released proteins using polyclonal antisera against BR143–156 and BR187–378 confirmed that both BR120–378 and BR187–378 were produced (Fig. 5B). The poly-His tags were detected for all constructs including the DUF control strain.
While bacteria displaying the DUF linker domain alone (Scar-DUF) were mainly observed as single cells or in doublets, cells displaying either BR*120–378 (Scar-BR*120–378) or BR187–378 (Scar-BR187–378) were heavily aggregated (Fig. 5C,D). Using fluorescence phase-contrast microscopy, single bacteria of Scar-BR187–378 and -BR*120–378 could be easily discriminated even within aggregates, while this was more difficult for Scar-DUF that appeared mostly in duplets (Fig. 5C). The origin of the slightly altered morphology of Scar-DUF bacteria under phase-contrast compared to the Scar-BR constructs is not known, but could be related to the absence and presence of the positively charged BR proteins, respectively. While Scar-DUF was not stained when incubated with fluorescently labeled secondary antibodies directed against primary polyclonals against BR 187–385, both Scar-BR*120–378 and Scar-BR187–378 bacteria were stained (Fig. 5C).

Quantitative analysis of the collected microscopy images confirmed our initial qualitative estimation that Scar-BR187–378 and Scar-BR*120–378 were significantly more aggregated compared to Scar-DUF, which was apparent from the particle size histogram plots with median particle sizes (q50) of 11 ± 2 μm², 11 ± 2 μm² and 3.7 ± 0.2 μm², respectively (Fig. 5D,E). A quantile is the fraction at which x% of the data fall below and (100 − x)% fall above that value. While two identical populations give a straight line (Fig. 5G, Scar-DUF in black), upward-shifted curves reveal increased particle sizes for Scar-BR187–378 and Scar-BR*120–378 (Fig. 5G, Scar-BR187–378 and Scar-BR*120–378 in red and blue, respectively). Similar results were
obtained when Scar-DUF, Scar-BR_{187–378} and Scar-BR_{120–378} were grown in M9 minimal medium in the presence and absence of added sheared eDNA derived from herring sperm (Fig. S6). While presence of eDNA did not

Figure 5. Cell-surface display of BR_{187–385} and BR^*_{120–378} promotes eDNA-dependent bacterial aggregation of Staphylococcus carnosus. (A) The domain organization for surface display of BR_{187–378} and BR^*_{120–378} on the surface of S. carnosus. PP - prepeptide from S. hyicus lipase; His_6 - hexahistidine tag; 3C - recognition sequence for protease 3C; X and M - regions from the Spa gene for covalent anchoring of the proteins to the peptidoglycan cell wall. (B) Supernatants of 3C-protease incubated staphylococcal cells were spotted on nitrocellulose membranes. Staining using polyclonal anti-serum against BR_{187–385} confirmed the expression of both BR^*_{120–378} and BR_{187–378} and the absence and presence of signal for the BR_{143–156} region allowed for an unambiguous discrimination between BR^*_{120–378} and BR_{187–378}-expressing cells, respectively. As expected, signals from HRP-coupled anti-His antibodies were detected for all three constructs. (C) The BR_{187–378} and BR^*_{120–378} displaying cells were stained using fluorescently labeled secondary antibodies directed against primary polyclonals against BR_{187–385}. Singlets, duplets and higher aggregates are highlighted in white, bronze and pink, respectively. (D) Phase-like contrast microscopy images revealed the formation of cellular aggregates of S. carnosus displaying BR_{187–378} and BR^*_{120–378}, but not of cells displaying the DUF domain alone. (E) The two-dimensional aggregates sizes of staphylococcal cells were quantitatively analyzed from an area of about 0.05 mm^2, which corresponds to approximately 64 times the area shown in panel D. The derived particle areas are plotted as normalized density histogram distributions, with the Scar-DUF control and the corresponding samples in black and red, respectively. The histogram represents the particle distribution obtained from a single experiment. The results from three independent experiments are summarized in the q-q plot shown in panel G. (F) Incubation of Scar-BR_{187–378} with DNaseI reduced the formation of aggregates, compared to the same sample not incubated with DNasel (panel D). (G) The mean particle area sizes of quantiles with cut-off values at 25%, 50%, 75%, 85%, and 95% were calculated from three independent experiments, and summarized in a quantile-quantile plot (q-q plot). After incubation with DNasel, both q-q plot curves of Scar-BR_{187–378} and Scar-BR_{120–378} are downward-shifted (light and dark green, respectively).
have any effect on Scar-DUF, both Scar-Br_{187–378} and Scar-Br_{120–378} were significantly more aggregated, albeit at a slightly reduced extent compared to bacteria grown in TB. We hypothesize that the observed slight reduction can be attributed to the restricted sequence length of the added eDNA between 15 and 50 bp. From these experiments we conclude that the globular Br_{187–378} domain of PsrP also promotes eDNA-dependent cellular aggregation within pneumococcal biofilms. Future experiments will address the exact contribution of each of the two factors, namely Br dimerization and binding to eDNA for biofilm formation.

Conclusions

PsrP has previously been suggested to promote pneumococcal biofilm formation through residues 122–166 (Br_{122–166}), preceding the globular KRT10-binding MSCRAMM-related Br_{187–378} domain. In this study, small angle X-ray scattering and circular dichroism spectroscopy analysis revealed a non-globular, disordered structure for the Br_{122–166} region that does not promote self-aggregation. Importantly, surface-display of Br_{187–385} demonstrated the capacity of the globular domain to promote eDNA-dependent bacterial aggregation. Structural studies indicate that the saddle-like structure of the Br_{187–385} β-sheet dimer snugly fits the acidic helical double-stranded DNA structure.

Materials and Methods

Cloning, expression and purification of protein constructs. Expression constructs comprising residues 120–395 of PsrP (Br_{120–395}), 187–385 (Br_{187–385}), and 187–378 (Br_{187–378}) all with a N-terminal poly-His (HHHHHHHHH) tag, were cloned into the pET21d (Novagen) expression vector as described previously. The mutated expression construct of Br_{120–395} (Br^{*}_{120–395}) with R165S and R168S substitutions was generated following previously described protocols. All coding sequences of the protein-expression constructs were confirmed by DNA sequencing and are listed in the supplemental information.

Heterologous protein expression in E. coli (T7 express, New England Biolabs) was induced at OD 0.4–0.7 using 400 μM IPTG and performed over night at 25 °C. The poly-Histidine-tagged proteins were purified using immobilized metal affinity (IMAC), cation exchange chromatography (CEC) (HisTrap FF and HiTrap SPFF, GE-Healthcare) and size exclusion chromatography (SEC) on Superdex 75 or 200 columns (GE Healthcare).

Furin cleavage assay. Two units of Furin protease (New England Biolabs, P8077) were added to a total assay volume of 100 μL with a concentration of 0.7 mg/mL Br_{120–395} and mutated Br^{*}_{120–395} in 20 mM HEPES, 500 mM NaCl, 10% glycerol, 1 mM CaCl2, pH 7.5. Hexa-D-arginine amide (Sigma: SCP0148) was used as furin inhibitor at concentrations of 100 μM and 250 μM. The cleavage assays were carried out at room temperature and at 37 °C. Samples for SDS-PAGE were taken at 0 h, 2 h, 4 h and overnight incubation.

Small angle X-ray scattering. Sample preparation and data collection. Data were collected at beamline BM29 at the European Synchrotron Radiation Facility (ESRF, Grenoble) using a Pilatus 1 M detector at a sample to detector distance of 2.9 m and a wavelength of λ = 0.9919 Å, the range of momentum transfer 0.006 < s < 0.45 Å⁻¹ was covered (s = 4πsinθ/λ, where θ is the scattering angle). Prior to data collection, proteins were buffer exchanged into phosphate buffered saline (PBS) with 5% glycerol, pH 7.4. Samples were measured at concentrations between 0.3 and 3 mg/mL using a continuous flow cell capillary.

Data analysis. The forward scattering I(0), the radius of gyration Rg along with the pair distribution function of the particle p(r) and the maximum dimension D_{max} were derived using the automated SAXS data analysis pipeline. Using Ensemble Optimization Method (EOM) analysis of Br_{187–385} and Br_{120–395}, a pool of 10,000 models comprising a rigid domain (residues K206-E375) and flexible N- and C-termini were generated. A subset of the pool was selected using a genetic algorithm such that the calculated averaged scattering of the selected models agreed with the experimental data. The Rg distributions of the selected ensembles were obtained by repeating the selection process multiple times.

Structural analysis. All figures were created using PyMOL version 1.3.0. Interfaces of macromolecular assemblies in the crystal structures were investigated using PBDpISA. The molecular model of the Br_{187–385}-DNA complex was predicted using the non-specific DNA-rigid protein-docking algorithm ParaDock and the electrostatic surface was visualized using APBS. The structure of TBP (PDB: 1YTB) was superimposed onto the Br_{187–385} dimer using the shape similarity algorithm of subcomp in which three-dimensional models are represented as simplified ensembles of points.

Electrophoretic mobility shift assay (EMSA) and isolation of Br_{187–378}–DNA_{276bp} complex using SEC. The 276 bp DNA molecule (DNA_{276bp} sequence listed in supplemental information) was produced and isolated by PCR amplification and gel extraction (QIAEXII gel extraction kit, Qiagen). Br_{187–378} was titrated to DNA at a final concentration of 100 nM in TAE buffer (40 mM Tris, 20 mM acetic acid, and 1 mM EDTA) supplemented with 300 mM NaCl. After an hour incubation time, the samples were mixed with sample loading buffer and run on a 1.5% (w/V) GelGreen-stained (Biotium) agarose gel in TAE buffer.

For SEC analysis, Br_{187–385} and DNA_{276} were incubated at concentrations of 30 μM and 440 nM, respectively. Before mixing, Br_{187–385} and DNA_{276} were prepared in 20 mM Hepes, 200 mM NaCl, 10% glycerol, pH 7.5 and 10 mM Tris at pH 8.5, respectively. The mixed samples were loaded on a Superdex S200 10/300 GL (GE healthcare) equilibrated in 20 mM Hepes, 200 mM NaCl, 10% glycerol, pH 7.5.
Surface display of BR on the surface of \textit{Staphylococcus carnosus}. \textbf{Cellular aggregation assay.} The PCR-amplified PsrP constructs and the gene-synthesized (Eurofins Genomics, Germany) DUF1542 repeat domain of \textit{SacC} (UniProt ID: C7BUR8) were cloned into the staphylococcal display vector ‘pHis3C’ using a sequence and ligature independent cloning (SLIC) method\textsuperscript{51,52}. Cloning sequences of the protein-display constructs were confirmed by DNA sequencing and were listed in the supplemental information. Transformed \textit{S. carnosus} cells were cultured in TB medium with chloramphenicol (10\,\mu g/mL). OD-adjusted cells treated or not-treated with DNaseI were gently re-suspended in PBS, filled into wells of a 96-\(\mu\)-well plate (Ibidi, Germanu) and images were taken using the ZOE Fluorescent Cell Imager (Biorad, USA) at 20x magnification. Bacteria were also grown in M9 derived minimal medium\textsuperscript{53} comprising 1x M9 salts, 2\,mM MgSO\textsubscript{4}, 0.1\,mM CaCl\textsubscript{2}, 1\% glucose, 1\% casaminoacids, 1\,mM Thiamine-HCl and 0.05\,mM nicotinamide, with or without the addition of herring sperm derived degraded DNA (D3159, Sigma Aldrich) at a concentration of 10\,ug/mL that corresponds to approximately 1\,\mu M of 15 bp oligonucleotide fragments. The bacteria were adjusted in OD and imaged in the same way as described previously. Images were analyzed using the software cell profiler (www.cellprofiler.org), the particle parameter values were imported into the R software package for statistical analysis and visualization using density histogram and q-q plots\textsuperscript{54,55}.

\textbf{Dotblot assays and fluorescence imaging.} Polyclonal antibodies against the TEV-cleaved purified BR\textsubscript{187–385} monomer and a synthesized BR\textsubscript{13–156} (RKKPASDYVASVTN) peptide were raised in rabbits and the obtained sera yielded titers of about 75000 and 6000, respectively (Innovagen, Sweden). For fluorescent staining, bacteria were incubated with affinity-purified anti-BR\textsubscript{187–385} antibodies and detected using fluorescein isothiocyanate (FITC)-labeled goat anti-rabbit IgG secondary antibody. Bacteria were fixed in 4\% paraformaldehyde, washed, and visualized using a fluorescence microscope (Leica Leitz DMRBE). For dotblot assays, bac-
teria were incubated with affinity-purified anti-BR187–385 antibodies and detected using fluorescein isothiocyanate (FITC)-labeled goat anti-rabbit IgG (life technologies) secondary antibody. Bacteria were fixed in 4\% paraformaldehyde, washed, and visualized using a fluorescence microscope (Leica Leitz DMRBE). For dotblot assays, surface-displayed proteins were cleaved as previously described\textsuperscript{61} and applied on nitrocellulose filter membranes. Proteins were detected using polyclonal rabbit-antisera against BR187–385, BR143–156 as well as HRP-coupled anti-His antibodies (ab1187, Abcam, UK). The anti-BR antibody stained samples were incubated using HRP-coupled monoclonal anti-rabbit IgG\textsubscript{γ} (A1949, Sigma Aldrich). Membranes were developed using Pierce ECL western blotting substrate (Thermo Scientific). Further details are given in the supplemental material.

\textbf{References}

1. van der Poll, T. & Opal, S. M. Pathogenesis, treatment, and prevention of pneumococcal pneumonia. \textit{Lancet} 374, 1543–1556 (2009).
2. Costerton, J. W., Stewart, P. S. & Greenberg, E. P. Bacterial Biofilms: A Common Cause of Persistent Infections. \textit{Science} 284, 1318–1322 (1999).
3. Hall-Stoodley, L., Costerton, J. W. & Stoodley, P. Bacterial biofilms: from the Natural environment to infectious diseases. \textit{Nat. Rev. Microbiol.} 2, 95–108 (2004).
4. Sinel, B. et al. The fundamental link between bacterial carriage and disease. \textit{Expert Rev. Vaccines} 11, 841–855 (2012).
5. Domenech, M., Garcia, E. & Moscoso, M. Biofilm formation in \textit{Streptococcus pneumoniae}. \textit{Microb. Biotechnol.} 5, 455–465 (2012).
6. Domenech, M., Garcia, E., Prieto, A. & Moscoso, M. Insight into the composition of the intercellular matrix of \textit{Streptococcus pneumoniae} biofilms. \textit{Environ. Microbiol.} 15, 502–516 (2013).
7. Blanchette-Cain, K. et al. \textit{Streptococcus pneumoniae} Biofilm Formation is Strain Dependent, Multifactorial, and Associated with Reduced Invasiveness and Immunoactive during Colonization. \textit{mBio} 4, e00745-13 (2013).
8. Hobley, L., Harkins, C., MacPhee, C. E. & Stanley-Wall, N. R. Giving structure to the biofilm matrix: an overview of individual strategies and emerging common themes. \textit{FEMS Microbiol. Rev.} eva015, doi: 10.1093/femsre/eva015 (2015).
9. Fleming, H.-C. & Wingender, J. The biofilm matrix. \textit{Nat. Rev. Microbiol.} 8, 623–633 (2010).
10. van Schaik, E. J. et al. DNA binding: a novel function of \textit{Pseudomonas aeruginosa} type IV pili. \textit{J. Bacteriol.} 187, 1455–1464 (2005).
11. Camilli, R., Pantosti, A. & Baldassarri, L. Contribution of serotype and genetic background to biofilm formation by \textit{Staphylococcus aureus}. \textit{J. Bacteriol.} 192, 2181–2188 (2010).
12. Pyburn, T. et al. The pneumococcal serine-rich repeat protein is an intra-species bacterial adhesin that promotes bacterial aggregation in vivo and in biofilms. \textit{Plos Pathog.} 6 (2010).
13. Sanchez, C. J. et al. The pneumococcal serine-rich repeat protein is an intra-species bacterial adhesin that promotes bacterial aggregation in vivo and in biofilms. \textit{Plos Pathog.} 6 (2010).
14. Ramboarina, S. et al. Structural insights into serine-rich fimbriae from gram-positive bacteria. \textit{J. Biol. Chem.} 285, 32446–32457 (2010).
15. Sanchez, C. J. et al. The pneumococcal serine-rich repeat protein is an intra-species bacterial adhesin that promotes bacterial aggregation in vivo and in biofilms. \textit{Plos Pathog.} 7, e1002112 (2011).
16. Lizzano, A., Sanchez, C. J. & Orihuela, C. J. A role for glycosylated serine-rich repeat proteins in Gram-positive bacterial pathogenesis. \textit{Mol. Oral Microbiol.} 27, 257–269 (2012).
17. Shrivshankar, P., Sanchez, C., Rose, L. F. & Orihuela, C. J. The \textit{Streptococcus pneumoniae} adhesin PsrP binds to Keratin 10 on lung cells. \textit{Mol. Microbiol.} 73, 663–679 (2009).
18. Shulite, T. et al. The basic keratin 10-binding domain of the virulence-associated pneumococcal serine-rich protein PsrP adopts a novel MSCRAMM fold. \textit{Open Biol.} 4, 130090 (2014).
19. Garnett, J. A. et al. Structural insight into the role of \textit{Streptococcus parasanguinis} Fap1 within oral biofilm formation. \textit{Biochem. Biophys. Res. Commun.} 417, 421–426 (2012).
20. Wu, H., Zeng, M. & Fives-Taylor, P. The glycan moieties and the N-terminal polypeptide backbone of a fimbia-associated adhesin, Fap1, play distinct roles in the biofilm development of \textit{Streptococcus parasanguinis}. \textit{Infect. Immun.} 75, 2181–2188 (2007).
21. Garnett, J. A. & Matthews, S. Interactions in Bacterial Biofilm Development: A Structural Perspective. \textit{Curr. Protein Pept. Sci.} 13, 739–755 (2012).
22. Corrigan, R. M., Rigby, D., Handley, P. & Foster, T. J. The role of \textit{Staphylococcus aureus} surface protein SaG in adherence and biofilm formation. \textit{Microbiology} 153, 2435–2446 (2007).
23. Gruszka, D. T. et al. \textit{Staphylococcus} Biofilm-Forming Protein Has a Contiguous Rod-Like Structure. \textit{Proc. Natl. Acad. Sci.} 109, E1011–E1018 (2012).
24. Rohde, H. et al. Induction of \textit{Staphylococcus epidermidis} biofilm formation via proteolytic processing of the accumulation-associated protein by staphylococcal and host proteases. \textit{Mol. Microbiol.} 55, 1883–1895 (2005).
27. Conrady, D. G., Wilson, J. J. & Herr, A. B. Structural basis for Zn²⁺-dependent intercellular adhesion in staphylococcal biofilms. *Proc. Natl. Acad. Sci. USA* **110**, E202–211 (2013).

28. Herman-Bausier, P., El-Kirat-Chatel, S., Foster, T. J., Geoghegan, J. A. & Dubrèf, Y. F. Staphylococcus aureus Fibronectin-Binding Protein A Mediates Cell-Cell Adhesion through Low-Affinity Homophilic Bonds. *mBio* **6** (2015).

29. Klipfel, K. R., Molloy, S. S., Thomas, G. & Leppa, S. H. Anthrax toxin protective antigen is activated by a cell surface protease with the sequence specificity and catalytic properties of furin. *Proc. Natl. Acad. Sci. USA* **89**, 10277–10281 (1992).

30. Thomas, G. & Furin at the cutting edge: From protein trafficking to embryogenesis and disease. *Nat. Rev. Mol. Cell Biol.* **3**, 753–766 (2002).

31. Karunathil, S. et al. Shedding of the Mucin-Like Floculin Flo11p Reveals a New Aspect of Fungal Adhesion Regulation. *Carr. Biol.* **20**, 1389–1395 (2010).

32. Gronenborn, A. M. Protein acrobatics in pairs—dimerization via domain swapping. *Carr. Opin. Struct. Biol.* **19**, 39–49 (2009).

33. Bennett, M. J. & Eisenberg, D. Bennett structure of monomeric diptheria toxin at 2.3 A resolution. *Protein Sci. Publ. Protein Soc.* **3**, 1464–1475 (1994).

34. Nowick, J. S. & Chung, D. M. Sequence-Selective Molecular Recognition between Sheets. *Angew. Chem. Int. Ed.* **42**, 1765–1768 (2003).

35. Hinterdorfer, P. & Dubrèf, Y. F. Detection and localization of single molecular recognition events using atomic force microscopy. *Nat. Methods* **3**, 347–355 (2006).

36. Kim, Y., Geiger, J. H., Hahn, S. & Sigler, P. Bcrystal structure of a yeast TBP/TATA box complex. *Nature* **365**, 512–520 (1993).

37. Hahn, S., Buratowski, S., Sharp, P. A. & Guarente, L. Yeast TATA-binding protein TFIIID binds to TATA elements with both consensus and nonconsensus DNA sequences. *Proc. Natl. Acad. Sci. USA* **86**, 5718–5722 (1989).

38. Privalov, P. L., Dragan, A. I. & Crane-Robinson, C. Interpreting protein/DNA interactions: distinguishing specific from non-specific and electrostatic from non-electrostatic components. *Nucleic Acids Res.* **39**, 2485–2491 (2011).

39. Samuelson, P., Gunneriusson, E., Nygren, P. A. & Ståhl, S. Display of proteins on bacteria. *J. Biotechnol.* **96**, 129–154 (2002).

40. Schroeder, K. et al. Molecular Characterization of a Novel *Staphylococcus Aureus* Surface Protein (SasC) Involved in Cell Aggregation and Biofilm Accumulation. *Plos One* **4**, e7567 (2009).

41. Li, J. et al. Site-directed mutagenesis by combination of homologous recombination and DpnI digestion of the plasmid template in *Escherichia coli*. *Anal. Biochem.* **373**, 389–391 (2008).

42. Pernot, P. et al. Upgraded ESRF BM29 beamline for SAXS on macromolecules in solution. *J. Synchrotron Radiat.* **20**, 660–664 (2013).

43. Franke, D., Kikhnney, A. G. & Svergun, D. I. Automated acquisition and analysis of small angle X-ray scattering data. *Nucl. Instrum. Methods Phys. Res. Sect. Accel. Spectrometers Detect. Assoc. Equip.* **689**, 52–59 (2012).

44. Bernado, P., Mylomas, E., Petoukhov, M. V., Blackledge, M. & Svergun, D. I. Structural characterization of flexible proteins using small-angle X-ray scattering. *J. Am. Chem. Soc.* **129**, 5656–5664 (2007).

45. Schrödinger, L. *PyMOL Molecular Graphics System*. (2010).

46. Krissinel, E. & Henrick, K. Inference of macromolecular assemblies from crystalline state. *J. Mol. Biol.* **372**, 774–797 (2007).

47. Banitt, I. & Wolfson, H. J. ParaDock: a flexible non-specific DNA—rigid protein docking algorithm. *Nucleic Acids Res.* **39**, e135 (2011).

48. Dolinsky, T. J., Nielsen, J. E., McCammon, J. A. & Baker, N. A. PDB2PQR: an automated pipeline for the setup of Poisson-Boltzmann electrostatics calculations. *Nucleic Acids Res.* **32**, W665–7 (2004).

49. Baker, N. A., Sept, D., Joseph, S., Holst, M. J. & McCammon, J. A. Electrostatics of nanosystems: application to microtubules and the ribosome. *Proc. Natl. Acad. Sci. USA* **98**, 10037–41 (2001).

50. Kozin, M. B. & Svergun, D. I. Automated matching of high- and low-resolution structural models. *J. Appl. Crystallogr.* **34**, 33–41 (2001).

51. Li, M. Z. & Elledge, S. J. Harnessing homologous recombination in vitro to generate recombinant DNA via SLIC. *Nat. Methods* **4**, 251–256 (2007).

52. Kronqvist, N., Löfblom, J., Severa, D., Ståhl, S. & Wernérhus, H. Simplified characterization through site-specific protease-mediated release of affinity proteins selected by staphylococcal display. *FEBS Microbiol. Lett.* **278**, 128–136 (2008).

53. Reed, P. et al. Staphylococcus aureus Survives with a Minimal Peptidoglycan Synthesis Machine but Sacrifices Virulence and Antibiotic Resistance. *Plos Pathog.* **11**, e1004891 (2015).

54. Jones, T. R. et al. CellProfiler Analyst: data exploration and analysis software for complex image-based screens. *BMC Bioinformatics* **9**, 482 (2008).

55. R Core Team. *R: A language and environment for statistical computing*. (R Foundation for Statistical Computing) (2011).
and C.M.; T.S. performed and analyzed the in vitro DNA binding of BR. E.F. performed and analyzed the M.D. simulations, A.B. collected the AFM data that were analyzed by A.B. and Y.F.D.; A.P. took the fluorescence images; C.E. collected and analyzed the AUC data; M.D., S.W. and T.S. performed and analyzed the heterologous surface display experiments.

**Additional Information**

**Accession codes:** Coordinates and structure factors have been deposited in the Protein Data Bank with accession number 5JUI. SAXS data have been deposited in the Small Angle Scattering Biological Databank (sasbdb) with accession numbers SASDAC6 and SASDAE6.

**Supplementary information** accompanies this paper at http://www.nature.com/srep

**Competing financial interests:** The authors declare no competing financial interests.

**How to cite this article:** Schulte, T. et al. The BR domain of PsrP interacts with extracellular DNA to promote bacterial aggregation; structural insights into pneumococcal biofilm formation. *Sci. Rep.* **6**, 32371; doi: 10.1038/srep32371 (2016).

This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/

© The Author(s) 2016