

# SCIENTIFIC REPORTS



OPEN

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

Received: 09 June 2016  
Accepted: 05 August 2016  
Published: 01 September 2016

Tim Schulte<sup>1</sup>, Cecilia Mikaelsson<sup>1</sup>, Audrey Beaussart<sup>2,†</sup>, Alexey Kikhney<sup>3</sup>, Maya Deshmukh<sup>1</sup>, Sebastian Wolniak<sup>1</sup>, Anuj Pathak<sup>4</sup>, Christine Ebel<sup>5</sup>, Jonas Löfling<sup>4</sup>, Federico Fogolari<sup>6</sup>, Birgitta Henriques-Normark<sup>4</sup>, Yves F. Dufrêne<sup>2</sup>, Dmitri Svergun<sup>3</sup>, Per-Åke Nygren<sup>7</sup> & Adnane Achour<sup>1</sup>

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 (BR<sub>187–385</sub>). Here, we show that BR<sub>187–385</sub> also binds to DNA, as demonstrated by electrophoretic mobility shift assays and size exclusion chromatography. Further, heterologous expression of BR<sub>187–378</sub> or the longer BR<sub>120–378</sub> 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 BR<sub>187–385</sub> homo-dimers via an intermolecular  $\beta$ -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 BR<sub>120–166</sub> 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 symptoms<sup>1</sup>. The organization of bacterial communities into complex and dynamic biofilm structures allows for growth in a hostile environment, providing protection against anti-microbial agents<sup>2,3</sup>. Pneumococcal colonization is the precursor for otitis media and sinusitis, as well as severe diseases including pneumonia, septicemia and meningitis<sup>4</sup>. Pneumococcal biofilms are found on the surface of adenoid and mucosal epithelial tissues in

<sup>1</sup>Science for Life Laboratory, Department of Medicine Solna, Karolinska Institute, and Department of Infectious Diseases, Karolinska University Hospital, Solna, SE-17176 Stockholm, Sweden. <sup>2</sup>Université catholique de Louvain, Institute of Life Sciences, Croix du Sud, 4-5, bte L7.07.06, B-1348 Louvain-la-Neuve, Belgium. <sup>3</sup>European Molecular Biology Laboratory (EMBL), Hamburg Outstation, Notkestrasse 85, 22603 Hamburg, Germany. <sup>4</sup>Department of Microbiology, Tumor and Cell Biology (MTC), Karolinska Institute; Clinical Microbiology, Karolinska University Hospital Solna, Stockholm, Sweden. <sup>5</sup>Institut de Biologie Structurale (IBS), Univ. Grenoble Alpes, CEA, CNRS, 38044 Grenoble, France. <sup>6</sup>Dipartimento di Scienze Mediche e Biologiche, Università di Udine, Piazzale Kolbe 4, 33100 Udine - Italy. <sup>7</sup>Division of Protein Technology, School of Biotechnology, KTH-Royal, Institute of Technology, Sweden. <sup>†</sup>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<sup>5</sup>.

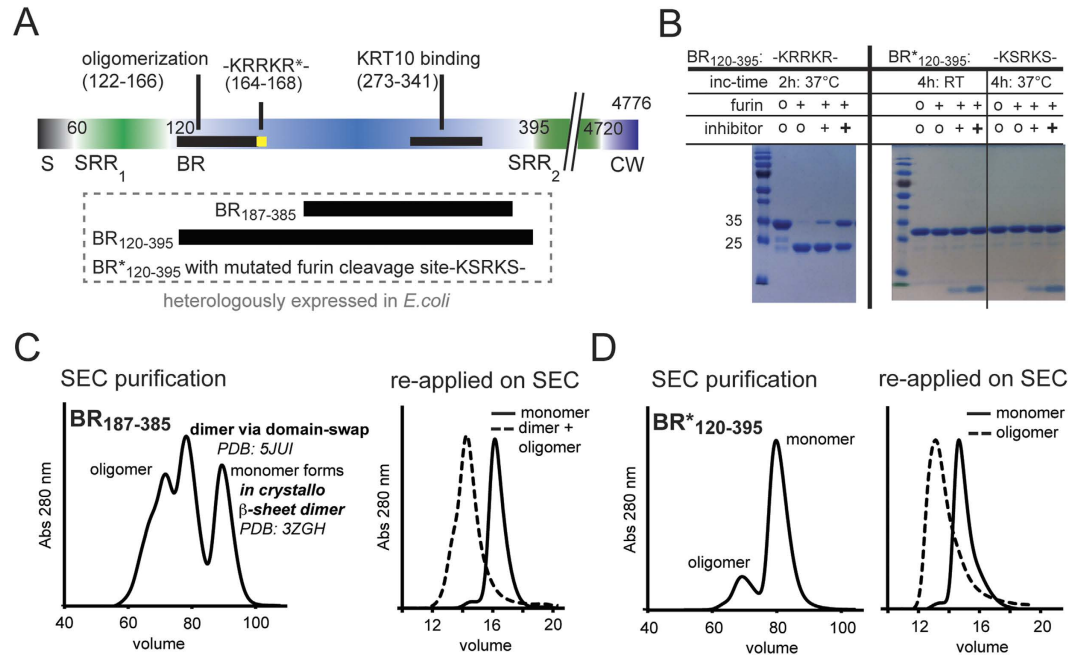
Recently, studies using *in vitro* and *in vivo* pneumococcal biofilm models have begun to unveil the biogenesis, regulation, structure and composition of pneumococcal biofilms<sup>6–8</sup>. The exact composition of the pneumococcal biofilm extracellular polymer matrix (EPM) depends largely on the species and environmental conditions<sup>5,9</sup>. 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<sup>8,10</sup>. In pneumococcus, the amount of capsular polysaccharides is highly regulated and significantly reduced during colonization compared to invasion<sup>1</sup>. 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<sup>5,11–13</sup>. The increased adhesive phenotype in non-encapsulated strains can possibly be attributed to an increased presentation of adhesion proteins<sup>13,14</sup>.

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<sup>7,15</sup>. A pneumococcal  $\Delta$ PsrP 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<sup>7</sup>. 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<sup>15–18</sup> (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<sup>17,18</sup>. Fold topologies similar to the microbial surface component recognizing adhesive molecule (MSCRAMM) CnaA from *Staphylococcus aureus* were determined for domains of the SRRPs GspB and Fap1 from *Streptococcus parasanguinis* and *Streptococcus gordonii*, respectively, and also predicted for domains of four other SRRPs<sup>16,17,19</sup>. Although the fold topology of the BR domain of PsrP compared to the other BR domains could not be predicted due to very low sequence similarity<sup>17</sup>, the recently determined crystal structure of the KRT10-binding domain of PsrP (BR<sub>187–385</sub>) revealed a topology distantly related to the CnaA fold<sup>20</sup>. Interestingly, GspB, Fap1 and PsrP, among several other SRRPs, were demonstrated to mediate intra-species interactions during biofilm formation<sup>15,17,18,21,22</sup>. A molecular model for the oligomerization of the helical subdomain NR $\alpha$  of Fap1 has also been proposed based on higher oligomers found in the crystal lattice and predicted as stable in solution<sup>21</sup>. The suggested molecular model is one of only a few known examples of biofilm-associated protein oligomerizations for which high-resolution structural data exist<sup>23</sup>. Recently, the crystal structures of the *Staphylococcus aureus* surface protein G (SasG) and of the accumulation-associated protein (Aap) from *Staphylococcus epidermidis*, both essential in biofilm formation, have been determined<sup>24–27</sup>. Like PsrP, these homologous multi-domain proteins are covalently attached to the peptidoglycan via their LPxTG motifs. Dimerization of the Aap B-repeat domains requires upstream proteolytic cleavage and a similar mechanism has also been suggested for SasG<sup>24,26</sup>. Crystal structures of self-oligomerizing B-repeat regions of both proteins revealed remarkable “free-standing”  $\beta$ -sheet topologies<sup>25,27</sup>. Based on the Aap B-domain crystal structure, a model for the Zn<sup>2+</sup>-dependent dimerization of the entire B-repeat domains was suggested according to which two molecules are wrapped around each other like entangled fibers in a rope<sup>27</sup>. Although a molecular understanding of the Zn<sup>2+</sup>-dependent dimerization of the MSCRAMM Fibronectin-Binding Protein A (FnBPA) of *S. aureus* is still lacking, atomic force microscopy (AFM) experiments elegantly demonstrated that cellular aggregation can be induced through multiple low-affinity homophilic bonds between copies of a single protein<sup>28</sup>.

The self-oligomerization property of PsrP has previously been suggested to occur through residues 122–166 (BR<sub>122–166</sub>), preceding the KRT10-binding region (BR<sub>187–385</sub>)<sup>15</sup>. However, a molecular mechanism underlying PsrP oligomerization has so far not been provided. In this study, crystal structure analysis and atomic force microscopy experiments revealed homophilic low affinity interactions, through the formation of an intermolecular  $\beta$ -sheet between two BR<sub>187–385</sub> molecules. Electrophoretic mobility shift assays demonstrated that BR<sub>187–385</sub> binds efficiently to DNA. Importantly, heterologous display of BR<sub>187–385</sub> on the surface of *Staphylococcus carnosus* demonstrated the capacity of BR<sub>187–385</sub> to promote eDNA-dependent cellular aggregation. Small-angle X-ray scattering and circular dichroism experiments revealed a disordered and non-globular structure in the region corresponding to BR<sub>122–166</sub>. We here demonstrate that BR<sub>122–166</sub> 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

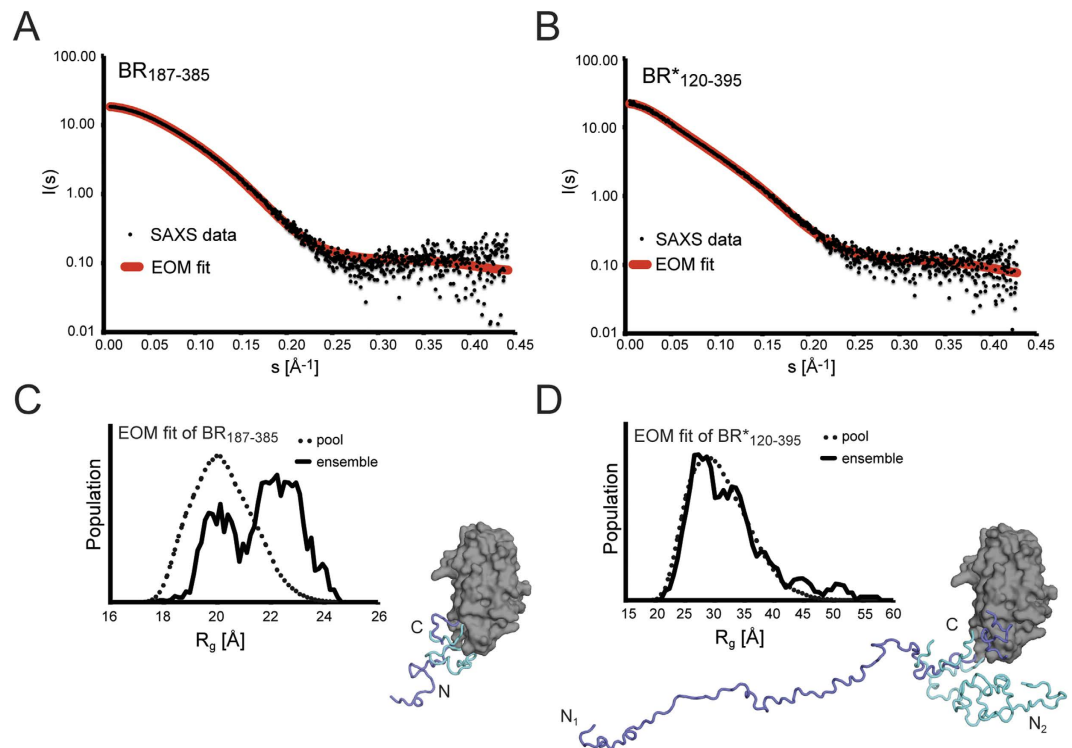
**The BR domain comprises a furin protease recognition site and forms irreversibly associated dimers only during heterologous expression.** Sequence analysis and biochemical protease cleavage assays revealed that the sequence motif K<sub>164</sub>RRKR<sub>168</sub>, localized between the KRT10-binding region<sup>20</sup> and the predicted non-globular N-terminal extension BR<sub>120–166</sub>, is cleaved by the human furin protease (Fig. 1A,B). While wild-type BR<sub>120–395</sub> was specifically cleaved by the human furin protease and prone to degradation during purification, a BR<sub>120–395</sub>\* 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<sup>29,30</sup>. Thus, we hypothesized that the N-terminal fragment comprising BR<sub>120–166</sub> and SRR<sub>1</sub> 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<sup>31</sup>.



**Figure 1. The BR domain comprises a furin protease recognition site and forms irreversibly associated dimers only during heterologous expression.** (A) PsrP is organized into five domains: the N-terminal signal sequence (S) for extracellular translocation of PsrP, two putatively glycosylated serine rich repeat regions (SRR<sub>1</sub> and SRR<sub>2</sub>), the binding region (BR) domain and the cell wall domain (CW). BR harbors two distinct sub-regions for KRT10 binding (black bar, residues 273–341) and self-oligomerization (black bar, residues 122–166), respectively. The sequence K<sub>164</sub>RRKR<sub>168</sub> is recognized by the human Furin protease. The three constructs BR<sub>120–395</sub>, BR<sub>187–385</sub> and BR<sub>187–378</sub> used within this study are displayed below the overall schematic representation of PsrP. In the mutated version of BR<sub>120–395</sub> (BR\*<sub>120–395</sub>), KRRKR was substituted to KSRKS. (B) A protease cleavage assay confirmed that furin recognized the KRRKR motif, but did not cleave in the presence of the furin protease inhibitor or when the KRRKR sequence was substituted to KSRKS. Cleavage was performed at room temperature (RT) and 37 °C for the indicated incubation times. The inhibitor was used at 100 μM and 250 μM, shown as plus signs in regular and bold format, respectively. (C) Distinct populations of BR<sub>187–385</sub> were eluted using Superdex 200 HiLoad 16/600 at elution column volumes (CV) of 89, 78, 71 and 66 mL corresponding to apparent molecular weights of 22, 55 and 90, 150 kDa, as well as higher oligomers. In this study, we show that the irreversibly associated dimer of BR<sub>187–385</sub> is formed through a domain swap mechanism (PDB: 5JUI, Fig. S2) and that a low-affinity β-sheet dimer is created between two symmetry-related molecules in the previously determined crystal structure of the BR<sub>187–385</sub> monomer (PDB: 3ZGH, Fig. 3). (D) BR\*<sub>120–395</sub> was eluted from the same column at CV of 79 and 69 mL corresponding to apparent MW of 50 and 115 kDa. (C,D) When the isolated monomer and oligomer populations of BR<sub>187–385</sub> and BR\*<sub>120–395</sub> were re-applied on analytical Superdex 200 HR10/30 columns, the monomer and oligomer populations were stable and did not interconvert between each other.

We have recently determined the crystal structure of the KRT10-binding domain of PsrP (BR<sub>187–385</sub>), revealing a novel MSCRAMM-related Dev-IgG fold resembling a compressed barrel<sup>20</sup>. The crystals were obtained from a monomeric preparation of BR<sub>187–385</sub> 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<sup>20</sup>. However, during purification of the heterologously produced BR<sub>187–385</sub>, 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<sub>120–395</sub> 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<sub>187–385</sub> dimer in order to obtain a molecular understanding for stable dimer formation. However, crystals were hard to reproduce and those obtained diffracted poorly. Therefore, a shorter construct was designed by deleting the seven C-terminal residues in BR<sub>187–385</sub> that were not built in the previously determined crystal structure due to missing electron density<sup>20</sup>. Well-diffracting crystals were obtained of the isolated irreversibly associated BR<sub>187–378</sub> dimer. Analysis of the crystal structure solved by molecular replacement revealed that the residue stretches L202–G315 (region I) and Y316–S377 (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<sup>32,33</sup>. 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



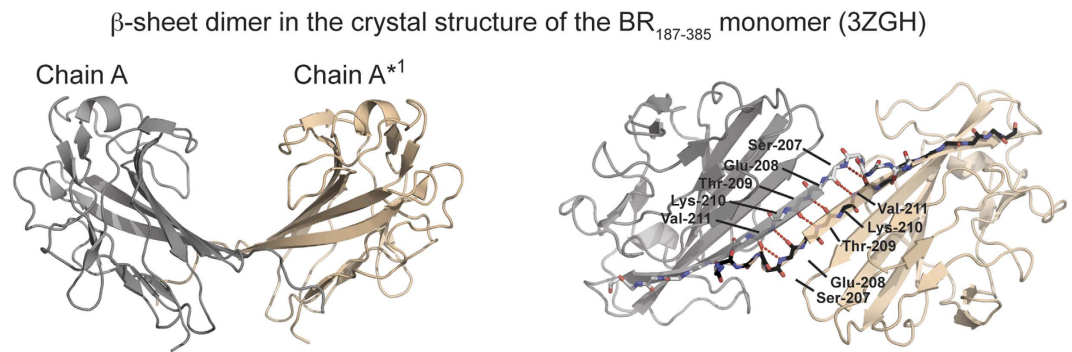
**Figure 2. The N-terminal region BR<sub>120-186</sub> forms a non-globular, flexible structure as revealed by SAXS.**

(A,B) Processed solution scattering data (black dots) from (A) BR<sub>187-385</sub> and (B) BR\*<sub>120-395</sub> fit to the theoretical scattering curves (red lines) from ensembles of 50 monomer models obtained from the ensemble optimization method (EOM) with  $\chi$ -values of 0.9 and 1.7. (C,D) The histograms of the radii of gyration distribution of the pool of 10 000 structures generated by EOM (dotted lines) and the  $R_g$  distribution of the selected ensembles (solid lines). (C) The ensemble of 50 BR<sub>187-385</sub> molecules selected by EOM was separated into two groups. The first group had an  $R_g$  value of around 20 Å and is represented by the molecular model with the extension colored in cyan. The second had an extended  $R_g$  value of 22.5 Å and is represented by the molecular model with the extension colored in blue. (D) The ensemble of 50 BR\*<sub>120-395</sub> molecules had a similar distribution of  $R_g$  values as compared the large pool of molecules with random conformations of the extensions. The two illustrated models displayed in blue and cyan have  $R_g$  values of 52 and 26 Å, respectively.

conditions<sup>32</sup>. 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<sub>122-166</sub> is non-globular and does not promote self-aggregation.** In order to obtain structural information about the longer BR\*<sub>120-395</sub> construct in solution, comparative small angle X-ray scattering (SAXS) data were collected for both monomeric BR<sub>187-385</sub> and BR\*<sub>120-395</sub> (Fig. 2A,B and Table S1). The directly accessible and sample-characteristic radii of gyration ( $R_g$ ) and maximal dimensions ( $D_{max}$ ) of BR<sub>187-385</sub> and BR<sub>120-395</sub> were  $20 \pm 1$  Å and  $29 \pm 2$  Å as well as  $78 \pm 8$  Å and  $125 \pm 12$  Å, respectively, indicating a more extended structure for BR<sub>120-395</sub> (Fig. S3A,B). Indeed, molecular models obtained from the ensemble optimization method (EOM) support the presence of a single globular domain common to BR<sub>187-385</sub> and BR\*<sub>120-395</sub>, but with a longer highly mobile structural extension for BR<sub>120-395</sub> (Fig. 2C,D). In dimensionless Kratky plots, the curve of BR<sub>187-385</sub> closely resembles the bell-shaped curve expected for a globular protein, while the curve of BR<sub>120-395</sub> 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<sub>120-395</sub> construct comprises an extended, highly flexible section in solution compared to BR<sub>187-385</sub>. Additional structural information was obtained from circular dichroism spectra of BR<sub>187-385</sub> and BR\*<sub>120-395</sub> that also indicated the presence of disordered and helical regions for the additional residues in BR\*<sub>120-395</sub> (Fig. S3D). Combining the structural information obtained from AUC, SAXS, and CD, we conclude that the N-terminal region in BR\*<sub>120-395</sub> does not promote aggregation and corresponds to a non-globular, largely disordered structure.

**BR<sub>187-385</sub> forms low-affinity homo-dimers through intermolecular  $\beta$ -sheets.** Considering that the described domain swap mechanism is unlikely to occur spontaneously *in vivo*, and that the previously suggested self-aggregating BR<sub>120-166</sub> region does not promote self-oligomerization in our experiments, we looked back into the crystal structure of the isolated BR<sub>187-385</sub> monomer<sup>20</sup> for hints of a more probable oligomerization mechanism. Indeed, we discovered intermolecular  $\beta$ -sheets between symmetry-related monomeric molecules in the P4<sub>1</sub>22 crystal form (PDB: 3ZGH) and in the P4<sub>3</sub>2<sub>1</sub>2 crystal form (PDB: 3ZGI) forming dimers with interface surface



**Figure 3.** BR<sub>187-385</sub> forms a  $\beta$ -sheet dimer with a saddle-like structure. Intermolecular  $\beta$ -sheets are formed between symmetry mates of BR<sub>187-385</sub> in the P4<sub>1</sub>22 (PDB: 3ZGH) crystal structure with a total interface surface area (ISA) of 600 Å<sup>2</sup>. The sheet is stabilized by eight hydrogen bonds between the backbones of the two A1 strands. The molecules forming intermolecular  $\beta$ -sheets are colored grey and wheat. The view of the right panel was rotated by  $\sim 90^\circ$  compared to the left panel. Hydrogen bond interactions are represented as dashed red lines.

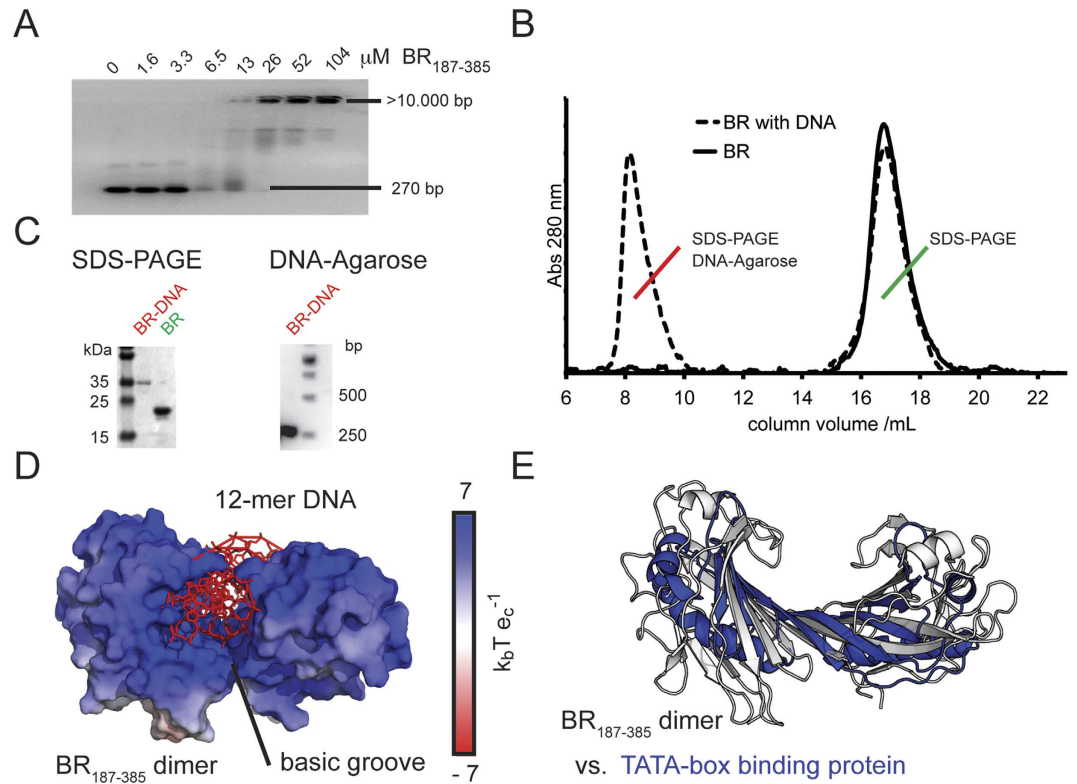
areas of 600 Å<sup>2</sup> and 440 Å<sup>2</sup>, respectively (Figs 3 and S4A). Moreover, intermolecular  $\beta$ -sheets are also observed between symmetry-related domain-swapped dimers with similar interfaces and hydrogen bond patterns (Fig. S4B,C). However, although the formation of such edge-to-edge  $\beta$ -sheets is regarded as a fundamental feature underlying molecular recognition in protein-protein interactions<sup>34</sup>, 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)<sup>20</sup>.

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 aggregation<sup>28</sup>. Therefore, we also applied single-molecule force spectroscopy to test the capacity of BR<sub>187-378</sub> and BR\*<sub>120-395</sub> to mediate similar low-affinity homophilic bonds. Indeed, homophilic binding events were measured for BR<sub>187-378</sub> and BR\*<sub>120-395</sub> with mean adhesion forces of 70  $\pm$  18 pN and 68  $\pm$  14 pN, respectively (Fig. S5A–C). And as expected for specific bimolecular bonds<sup>35</sup>, the adhesion forces increased linearly with the logarithm of the loading rate (Fig. S5D).

**BR<sub>187-385</sub> 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 BR<sub>187-385</sub> could also bind to acidic biofilm-associated extracellular DNA. Indeed, electromobility shift assays (EMSA) demonstrated binding of a randomly chosen 276 bp long DNA (DNA<sub>276bp</sub>) molecule to BR<sub>187-385</sub> with an apparent affinity in the lower micromolar range (Fig. 4A). After incubation of BR<sub>187-378</sub> with DNA<sub>276bp</sub> at a 1:750 molar ratio, the purified high-molecular weight DNA population also comprised bound BR<sub>187-385</sub> molecules (Fig. 4B,C). Interestingly, the DNA-bound BR<sub>187-385</sub> ran 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  $\beta$ -sheet dimer of BR<sub>187-385</sub> (Fig. 4D). Molecular dynamic (MD) simulations also suggested that the docked DNA:BR<sub>187-385</sub> 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 BR<sub>187-385</sub> 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 interactions<sup>36,37</sup>. 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 affinity<sup>38</sup>. Similar to the interaction of TBP with non-related DNA, the apparent affinity measured for binding of BR<sub>187-385</sub> 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.

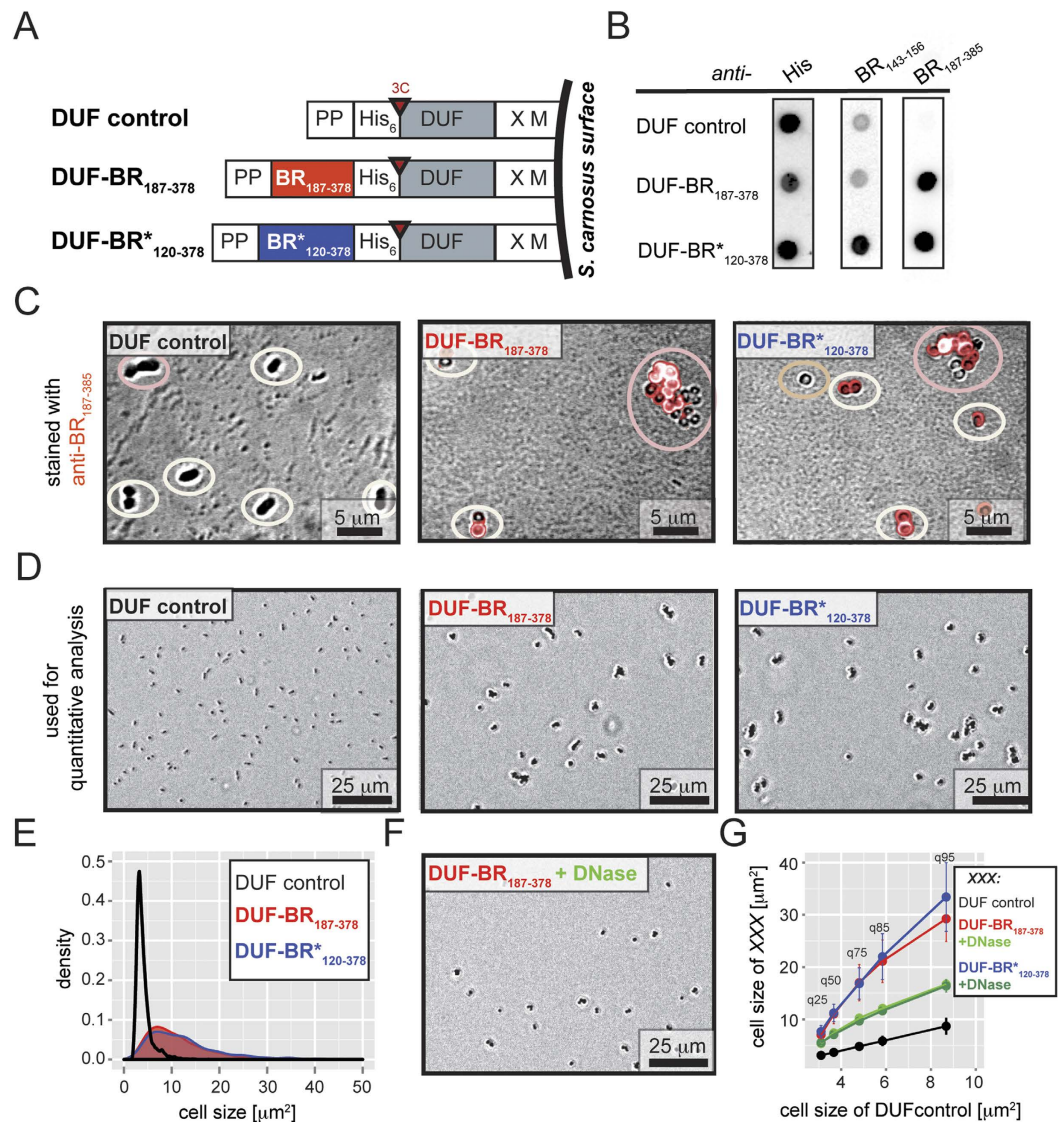
**BR<sub>187-385</sub> promotes DNA-dependent bacterial aggregation.** Display of heterologously produced proteins on the surface of the Gram-positive bacterium *Staphylococcus carnosus* represents an excellent model system to study the functions of surface-associated proteins, e.g. the biofilm-promoting function of the SasC protein of *Staphylococcus aureus*<sup>39,40</sup>. In order to assess the capacity of each segment of the BR domain to promote cellular aggregation, BR<sub>187-378</sub> and BR\*<sub>120-378</sub> 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 BR<sub>143-156</sub> and BR<sub>187-378</sub> confirmed that both BR\*<sub>120-378</sub> and BR<sub>187-378</sub> were produced (Fig. 5B). The poly-His tags were detected for all constructs including the DUF control strain.



**Figure 4. Basic BR<sub>187-385</sub> binds to DNA and forms a saddle-like dimer structure that snugly fits the acidic double stranded DNA helix.** (A) BR<sub>187-385</sub> concentration-dependent complex formation with a non-specific 270 bp DNA molecule (DNA<sub>270bp</sub>) was detected as retarded DNA bands in an electrophoretic mobility shift assay (EMSA). The apparent affinity of the interaction was estimated to around 10  $\mu$ M. (B) After incubation of BR<sub>187-385</sub> with DNA<sub>270bp</sub>, the area of the BR<sub>187-385</sub> monomer peak, obtained at a retention volume of around 17 mL and an absorption wavelength of 280 nm, was slightly reduced. The high-molecular weight population was detected at a retention volume of about 8 mL with an estimated MW of about >1 MDa. DNA<sub>270bp</sub> alone would elute at the same retention volume (not shown). (C) The putative BR<sub>187-385</sub>-DNA complex was isolated from SEC and run on SDS-PAGE (left) and DNA-Agarose (right) gels. In SDS-PAGE the protein band was detected as a slightly larger complex, possibly retarded by the bound negatively charged DNA. The DNA<sub>270bp</sub> that was isolated from SEC was detected at the expected size, and was not retarded due to the low amount of bound BR<sub>187-385</sub>. (D) The intermolecular  $\beta$ -sheet dimer resembles a saddle-like structure with an extended intermolecular  $\beta$ -sheet. The electrostatic potentials are plotted in  $k_B T e_c^{-1}$  with the Boltzmann's constant  $k_B$ , the charge of an electron  $e_c$  at a temperature  $T$  of 298 K. The saddle-like structure snugly fits an acidic helical non-specific DNA molecule, as predicted in a molecular model of the BR<sub>187-385</sub>-DNA complex. (E) Similar to the BR<sub>187-385</sub> dimer, the TATA-box binding protein (TBP, PDB: 1YTB) adopts a saddle-like structure. TBP was superimposed onto the BR<sub>187-387</sub> dimer via shape similarity and represented as white and blue ribbons.

While bacteria displaying the DUF linker domain alone (Scar-DUF) were mainly observed as single cells or in doublets, cells displaying either BR\*<sub>120-378</sub> (Scar-BR\*<sub>120-378</sub>) or BR<sub>187-378</sub> (Scar-BR<sub>187-378</sub>) were heavily aggregated (Fig. 5C,D). Using fluorescence phase-contrast microscopy, single bacteria of Scar-BR<sub>187-378</sub> and -BR\*<sub>120-378</sub> 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<sub>187-385</sub>, both Scar-BR\*<sub>120-378</sub> and Scar-BR<sub>187-378</sub> bacteria were stained (Fig. 5C).

Quantitative analysis of the collected microscopy images confirmed our initial qualitative estimation that Scar-BR<sub>187-378</sub> and Scar-BR\*<sub>120-378</sub> were significantly more aggregated compared to Scar-DUF, which is apparent from the particle size histogram plots with median particle sizes (q50) of  $11 \pm 2 \mu\text{m}^2$ ,  $11 \pm 2 \mu\text{m}^2$  and  $3.7 \pm 0.2 \mu\text{m}^2$ , respectively (Fig. 5D,E). Since the staphylococcal cells were grown in terrific broth (TB) medium comprising yeast extract, we reasoned that the DNA extracted from yeast contributed to the observed cellular aggregates. Indeed, following incubation with the DNA-degrading enzyme DNaseI, aggregates of Scar-BR\*<sub>120-378</sub> and Scar-BR<sub>187-378</sub> were reduced by almost 40% over all listed quantile values (Fig. 5E,G). A quantile is the fraction at which  $x\%$  percent 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-BR<sub>187-378</sub> and Scar-BR\*<sub>120-378</sub> (Fig. 5G, Scar-BR<sub>187-378</sub> and Scar-BR\*<sub>120-378</sub> in red and blue, respectively). Similar results were



**Figure 5. Cell-surface display of BR<sub>187-378</sub> and BR\*<sub>120-378</sub> promotes eDNA-dependent bacterial aggregation of *Staphylococcus carnosus*.** (A) The domain organization for surface display of BR<sub>187-378</sub> and BR\*<sub>120-378</sub> on the surface of *S. carnosus*. PP - propeptide from *S. hyicus* lipase; His<sub>6</sub> - 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<sub>187-385</sub> confirmed the expression of both BR\*<sub>120-378</sub> and BR<sub>187-378</sub>, and the absence and presence of signal for the BR<sub>143-156</sub> region allowed for an unambiguous discrimination between BR\*<sub>120-378</sub>- and BR<sub>187-378</sub>-expressing cells, respectively. As expected, signals from HRP-coupled anti-His antibodies were detected for all three constructs. (C) The BR<sub>187-378</sub> and BR\*<sub>120-378</sub> displaying cells were stained using fluorescently labeled secondary antibodies directed against primary polyclonals against BR<sub>187-385</sub>. 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<sub>187-378</sub> and BR\*<sub>120-378</sub>, 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<sup>2</sup>, 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<sub>187-378</sub> with DNaseI reduced the formation of aggregates, compared to the same sample not incubated with DNaseI (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 DNaseI, both q-q plot curves of Scar-BR<sub>187-378</sub> and Scar-BR\*<sub>120-378</sub> are downward-shifted (light and dark green, respectively).

obtained when Scar-DUF, Scar-BR<sub>187-378</sub> and Scar-BR\*<sub>120-378</sub> 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

have any effect on Scar-DUF, both Scar-BR<sub>187–378</sub> and Scar-BR<sub>120–378</sub> 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<sub>187–378</sub> 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<sub>122–166</sub>), preceding the globular KRT10-binding MSCRAMM-related BR<sub>187–385</sub> domain. In this study, small angle X-ray scattering and circular dichroism spectroscopy analysis revealed a non-globular, disordered structure for the BR<sub>122–166</sub> region that does not promote self-aggregation. Importantly, surface-display of BR<sub>187–385</sub> demonstrated the capacity of the globular domain to promote eDNA-dependent bacterial aggregation. Structural studies indicate that the saddle-like structure of the BR<sub>187–385</sub>  $\beta$ -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<sub>120–395</sub>), 187–385 (BR<sub>187–385</sub>), and 187–378 (BR<sub>187–378</sub>) all with a N-terminal poly-His (HHHHHH) tag, were cloned into the pET21d (Novagen) expression vector as described previously<sup>20</sup>. The mutated expression construct of BR<sub>120–395</sub> (BR\*<sub>120–395</sub>) with R165S and R168S substitutions was generated following previously described protocols<sup>41</sup>. 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  $\mu$ M IPTG and performed overnight 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  $\mu$ L with a concentration of 0.7 mg/mL BR<sub>120–395</sub> and mutated BR\*<sub>120–395</sub> in 20 mM HEPES, 500 mM NaCl, 10% glycerol, 1 mM CaCl<sub>2</sub>, pH 7.5. Hexa-D-arginine amide (Sigma: SCP0148) was used as furin inhibitor at concentrations of 100  $\mu$ M and 250  $\mu$ M. The cleavage assays were carried out at room temperature and at 37 °C. Samples for SDS-PAGE were taken at after 0 h, 2 h, 4 h and overnight incubation.

**Small angle X-ray scattering.** *Sample preparation and data collection.* Data were collected at beam-line BM29 at the European Synchrotron Radiation Facility (ESRF, Grenoble)<sup>42</sup>. Using a Pilatus 1 M detector at a sample to detector distance of 2.9 m and a wavelength of  $\lambda = 0.9919$  Å, the range of momentum transfer  $0.006 < s < 0.45$  Å<sup>-1</sup> was covered ( $s = 4\pi \sin \theta / \lambda$ , where  $2\theta$  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  $R_g$  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<sup>43</sup>. Using Ensemble Optimization Method (EOM)<sup>44</sup> analysis of BR<sub>187–385</sub> and BR<sub>120–395</sub>, 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  $R_g$  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<sup>45</sup>. Interfaces of macromolecular assemblies in the crystal structures were investigated using PBDDePISA<sup>46</sup>. The molecular model of the BR<sub>187–385</sub>-DNA complex was predicted using the non-specific DNA-rigid protein-docking algorithm ParaDock<sup>47</sup>, and the electrostatic surface was visualized using APBS<sup>48,49</sup>. The structure of TBP (PDB: 1YTB) was superimposed onto the BR<sub>187–387</sub> dimer using the shape similarity algorithm of subcomp in which three-dimensional models are represented as simplified ensembles of points<sup>36,50</sup>.

**Electrophoretic mobility shift assay (EMSA) and isolation of BR<sub>187–378</sub>-DNA<sub>276bp</sub> complex using SEC.** The 276 bp DNA molecule (DNA<sub>276</sub> sequence listed in supplemental information) was produced and isolated by PCR amplification and gel extraction (QIAEXII gel extraction kit, Qiagen). BR<sub>187–385</sub> 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<sub>187–385</sub> and DNA<sub>276</sub> were incubated at concentrations of 30  $\mu$ M and 440 nM, respectively. Before mixing, BR<sub>187–385</sub> and DNA<sub>276</sub> 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 *Staphylococcus carnosus*.** *Cellular aggregation assay.* The PCR-amplified PsrP constructs and the gene-synthesized (Eurofins Genomics, Germany) DUF1542 repeat domain of SasC (Uniprot ID: C7BUR8) were cloned into the staphylococcal display vector ‘pHis3C’ using a sequence and ligation independent cloning (SLIC) method<sup>51,52</sup>. Coding sequences of the protein-display constructs were confirmed by DNA sequencing and are listed in the supplemental information. Transformed *S. carnosus* cells were cultured in TB medium with chloramphenicol (10 µg/mL). OD-adjusted cells treated or non-treated with DNaseI were gently re-suspended in PBS, filled into wells of a 96-µ-well plate (Ibidi, Germany) and images were taken using the ZOE Fluorescent Cell Imager (Biorad, USA) at 20x magnification. Bacteria were also grown in M9 derived minimal medium<sup>53</sup> comprising 1x M9 salts, 2 mM MgSO<sub>4</sub>, 0.1 mM CaCl<sub>2</sub>, 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 µg/mL that corresponds to approximately 1 µM 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<sup>54,55</sup>.

*Dotblot assays and fluorescence imaging.* Polyclonal antibodies against the TEV-cleaved purified BR<sub>187–385</sub> monomer and a synthesized BR<sub>143–156</sub> (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<sub>187–385</sub> 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<sup>52</sup> and applied on nitrocellulose filter membranes. Proteins were detected using polyclonal rabbit-antisera against BR<sub>187–385</sub>, BR<sub>143–156</sub> 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 $\gamma$  (A1949, Sigma Aldrich). Membranes were developed using Pierce ECL western blotting substrate (Thermo Scientific). Further details are given in the supplementary material.

## References

- van der Poll, T. & Opal, S. M. Pathogenesis, treatment, and prevention of pneumococcal pneumonia. *Lancet* **374**, 1543–1556 (2009).
- Costerton, J. W., Stewart, P. S. & Greenberg, E. P. Bacterial Biofilms: A Common Cause of Persistent Infections. *Science* **284**, 1318–1322 (1999).
- Hall-Stoodley, L., Costerton, J. W. & Stoodley, P. Bacterial biofilms: from the Natural environment to infectious diseases. *Nat. Rev. Microbiol.* **2**, 95–108 (2004).
- Simell, B. *et al.* The fundamental link between pneumococcal carriage and disease. *Expert Rev. Vaccines* **11**, 841–855 (2012).
- Domenech, M., Garcia, E. & Moscoso, M. Biofilm formation in *Streptococcus pneumoniae*. *Microb. Biotechnol.* **5**, 455–465 (2012).
- Domenech, M., Garcia, E., Prieto, A. & Moscoso, M. Insight into the composition of the intercellular matrix of *Streptococcus pneumoniae* biofilms. *Environ. Microbiol.* **15**, 502–516 (2013).
- Blanchette-Cain, K. *et al.* *Streptococcus pneumoniae* Biofilm Formation is Strain Dependent, Multifactorial, and Associated with Reduced Invasiveness and Immunoreactivity during Colonization. *mBio* **4**, e00745–13 (2013).
- 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. *FEMS Microbiol. Rev.* fuv015, doi: 10.1093/femsre/fuv015 (2015).
- Flemming, H.-C. & Wingender, J. The biofilm matrix. *Nat. Rev. Microbiol.* **8**, 623–633 (2010).
- van Schaik, E. J. *et al.* DNA binding: a novel function of *Pseudomonas aeruginosa* type IV pili. *J. Bacteriol.* **187**, 1455–1464 (2005).
- Camilli, R., Pantosti, A. & Baldassarri, L. Contribution of serotype and genetic background to biofilm formation by *Streptococcus pneumoniae*. *Eur. J. Clin. Microbiol. Infect. Dis. Off. Publ. Eur. Soc. Clin. Microbiol.* **30**, 97–102 (2011).
- Hall-Stoodley, L. *et al.* Characterization of biofilm matrix, degradation by DNase treatment and evidence of capsule downregulation in *Streptococcus pneumoniae* clinical isolates. *BMC Microbiol.* **8**, 173 (2008).
- Sanchez, C. J. *et al.* *Streptococcus pneumoniae* in Biofilms Are Unable to Cause Invasive Disease Due to Altered Virulence Determinant Production. *Plos One* **6**, e28738 (2011).
- Muñoz-Eliás, E. J., Marcano, J. & Camilli, A. Isolation of *Streptococcus pneumoniae* Biofilm Mutants and Their Characterization during Nasopharyngeal Colonization. *Infect. Immun.* **76**, 5049–5061 (2008).
- 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. *Plos Pathog.* **6** (2010).
- Ramboarina, S. *et al.* Structural insights into serine-rich fimbriae from gram-positive bacteria. *J. Biol. Chem.* **285**, 32446–32457 (2010).
- Pyburn, T. M. *et al.* A Structural Model for Binding of the Serine-Rich Repeat Adhesin GspB to Host Carbohydrate Receptors. *Plos Pathog.* **7**, e1002112 (2011).
- Lizcano, A., Sanchez, C. J. & Orihuela, C. J. A role for glycosylated serine-rich repeat proteins in Gram-positive bacterial pathogenesis. *Mol. Oral Microbiol.* **27**, 257–269 (2012).
- Shivshankar, P., Sanchez, C., Rose, L. F. & Orihuela, C. J. The *Streptococcus pneumoniae* adhesin PsrP binds to Keratin 10 on lung cells. *Mol. Microbiol.* **73**, 663–679 (2009).
- Schulte, T. *et al.* The basic keratin 10-binding domain of the virulence-associated pneumococcal serine-rich protein PsrP adopts a novel MSCRAMM fold. *Open Biol.* **4**, 130090 (2014).
- Garnett, J. A. *et al.* Structural insight into the role of *Streptococcus parasanguinis* Fap1 within oral biofilm formation. *Biochem. Biophys. Res. Commun.* **417**, 421–426 (2012).
- Wu, H., Zeng, M. & Fives-Taylor, P. The glycan moieties and the N-terminal polypeptide backbone of a fimbria-associated adhesin, Fap1, play distinct roles in the biofilm development of *Streptococcus parasanguinis*. *Infect. Immun.* **75**, 2181–2188 (2007).
- Garnett, J. A. & Matthews, S. Interactions in Bacterial Biofilm Development: A Structural Perspective. *Curr. Protein Pept. Sci.* **13**, 739–755 (2012).
- Corrigan, R. M., Rigby, D., Handley, P. & Foster, T. J. The role of *Staphylococcus aureus* surface protein SasG in adherence and biofilm formation. *Microbiology* **153**, 2435–2446 (2007).
- Gruszka, D. T. *et al.* Staphylococcal Biofilm-Forming Protein Has a Contiguous Rod-Like Structure. *Proc. Natl. Acad. Sci.* **109**, E1011–E1018 (2012).
- Rohde, H. *et al.* Induction of *Staphylococcus epidermidis* biofilm formation via proteolytic processing of the accumulation-associated protein by staphylococcal and host proteases. *Mol. Microbiol.* **55**, 1883–1895 (2005).

27. Conrady, D. G., Wilson, J. J. & Herr, A. B. Structural basis for Zn<sup>2+</sup>-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. & Dufrène, Y. F. Staphylococcus aureus Fibronectin-Binding Protein A Mediates Cell-Cell Adhesion through Low-Affinity Homophilic Bonds. *mBio* **6** (2015).
29. Klimpel, K. R., Molloy, S. S., Thomas, G. & Leppla, 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.* **89**, 10277–10281 (1992).
30. Thomas, G. Furin at the cutting edge: From protein traffic to embryogenesis and disease. *Nat. Rev. Mol. Cell Biol.* **3**, 753–766 (2002).
31. Karunanithi, S. *et al.* Shedding of the Mucin-Like Flocculin Flo11p Reveals a New Aspect of Fungal Adhesion Regulation. *Curr. Biol.* **20**, 1389–1395 (2010).
32. Gronenborn, A. M. Protein acrobatics in pairs—dimerization via domain swapping. *Curr. Opin. Struct. Biol.* **19**, 39–49 (2009).
33. Bennett, M. J. & Eisenberg, D. Refined structure of monomeric diphtheria toxin at 2.3 Å resolution. *Protein Sci. Publ. Protein Soc.* **3**, 1464–1475 (1994).
34. Nowick, J. S. & Chung, D. M. Sequence-Selective Molecular Recognition between  $\beta$  Sheets. *Angew. Chem. Int. Ed.* **42**, 1765–1768 (2003).
35. Hinterdorfer, P. & Dufrène, 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. B. Crystal 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 TFIID 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**, 2483–2491 (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., Kikhney, 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. Bernadó, P., Mylonas, 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. & Wernérus, H. Simplified characterization through site-specific protease-mediated release of affinity proteins selected by staphylococcal display. *FEMS 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).

## Acknowledgements

We gratefully acknowledge fruitful discussions with Tatyana Sandalova as well as access to synchrotron beamlines BM29 and ID23-1 at ESRF (Grenoble, France). This work was facilitated by the Protein Science Facility at Karolinska Institute/SciLifeLab (<http://psf.ki.se>). We thank John Löfblom (KTH, Stockholm) and Simon Elsässer (MKB, Karolinska Institute, Stockholm) for providing the *S. carnosus* cells and access to the ZOE Fluorescent Cell Imager, respectively. This work was supported by grants from the Swedish Research Council (AA, BHN), the Knut and Alice Wallenberg Foundation (AA, BHN), the Swedish foundation for strategic research (AA and BHN), the Wenner Gren Foundation and (TS), the European Community's Seventh Framework Programme (FP7/2007–2013) under BioStruct-X (grant agreement No. 283570). We gratefully acknowledge access to synchrotron beamlines BM29 and ID23-1 at ESRF (Grenoble, France). We acknowledge the platforms of the Grenoble Instruct center (ISBG; UMS 3518 CNRS-CEA-UGA-EMBL) supported by the French Infrastructure for Integrated Structural Biology Initiative FRISBI (ANR-10-INSB-05-02) and by the Grenoble Alliance for Integrated Structural Cell Biology GRAL (ANR-10-LABX-49-01) within the Grenoble Partnership for Structural Biology (PSB). Work at the Université catholique de Louvain was supported by the National Fund for Scientific Research (FNRS), the FNRS-WELBIO under Grant no. WELBIO-CR-2015A-05, the Federal Office for Scientific, Technical and Cultural Affairs (Interuniversity Poles of Attraction Programme), and the Research Department of the Communauté française de Belgique (Concerted Research Action). Y.F.D. is a Research Director at the FNRS. The funders had no role in study design, data collection and interpretation, or the decision to submit the work for publication.

## Author Contributions

T.S. and A.A. wrote the initial draft that was complemented by contributions from C.M., A.B., A.K., A.P., C.E., F.F., B.H.-N., Y.F.D., D.S., P.-Å.N. and A.A.; T.S., C.M., J.L., M.D. and S.W. performed the cloning work for heterologous expression in *E. coli* and *S. carnosus*; T.S. and C.M. produced the BR constructs and performed the furin cleavage assays as well SEC analysis; C.M. collected the SAXS data that were analyzed by A.K., T.S.

and C.M.; T.S. performed and analyzed the *in vitro* DNA binding of BR, F.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