Structural basis for translocation of a biofilm-supporting exopolysaccharide across the bacterial outer membrane

Yan Wang\textsuperscript{1,2,*}, Archana AndolePannuri\textsuperscript{3*}, Dongchun Ni\textsuperscript{4}, Haizhen Zhou\textsuperscript{1}, Xiou Cao\textsuperscript{5}, Xiaomei Lu\textsuperscript{6}, Tony Romeo\textsuperscript{3,7} & Yihua Huang\textsuperscript{1,7}

\textsuperscript{1}National Laboratory of Biomacromolecules, Institute of Biophysics, Chinese Academy of Sciences, Beijing 100101, China. \textsuperscript{2}University of Chinese Academy of Sciences, Beijing 100049, China. \textsuperscript{3}Department of Microbiology and Cell Science, Institute of Food and Agricultural Sciences, University of Florida, Gainesville, FL 32611-0700, USA. \textsuperscript{4}Department of Cardiovascular Diseases, Tianjin Xiqing Hospital, Tianjin 300380, China. \textsuperscript{5}School of Life Sciences, Peking University, Beijing 100871, China. \textsuperscript{6}Dongguan Institute of Pediatrics, the Eighth People’s Hospital of Dongguan, Dongguan 523325, Guangdong Province, China. \textsuperscript{7}Correspondence should be address to Y.H. (yihuahuang@sun5.ibp.ac.cn) or T.R. (tromeo@ufl.edu); Contact phone number: 86-10-64888789. *These two authors contributed equally to this work.

Running title: Structure and mechanism of dPNAG secretion by PgaA

Key words: PgaA, Poly-β-1,6-N-acetyl-D-glucosamine, outer membrane protein, biofilm, exopolysaccharide

The partially de-N-acetylated poly-β-1,6-N-acetyl-D-glucosamine (dPNAG) polymer serves as an intercellular biofilm adhesin that plays an essential role for the development and maintenance of integrity of biofilms of diverse bacterial species. Translocation of dPNAG across the bacterial outer membrane is mediated by a tetraricopeptide repeat (TPR)-containing outer membrane protein, PgaA. To understand the molecular basis of dPNAG translocation, we determined the crystal structure of the C-terminal transmembrane domain of PgaA (residues 513-807). The structure reveals that PgaA forms a 16-strand transmembrane β-barrel, closed by four loops on the extracellular surface. Half of the interior surface of the barrel that lies parallel to the translocation pathway is electronegative, suggesting that the corresponding negatively-charged residues may assist the secretion of the positively-charged PgaB-bound dPNAG polymer. In vivo complementation assays in a pgaA-deletion bacterial strain showed that a cluster of negatively charged residues proximal to the periplasm is necessary for biofilm formation. Biochemical analyses further revealed that the TPR domain of PgaA binds directly to the N-deacetylase PgaB and is critical for biofilm formation. Our studies support a model in which the positively-charged PgaB-bound dPNAG polymer is delivered to PgaA through the PgaA-PgaB interaction and is further targeted to the β-barrel lumen of PgaA potentially via a charge complementarity mechanism, thus priming the translocation of dPNAG across the bacterial outer membrane.

Bacterial biofilms are complex surface-attached microbial communities that cause severe medical problems because once established, they are notoriously difficult to eradicate\textsuperscript{4}. A hallmark of bacterial biofilm is that the bacteria are encased within a self-produced hydrated polymeric matrix that is typically composed of proteinaceous adhesion factors, exopolysaccharides (EPS) and DNA\textsuperscript{4-7}. A variety of exopolysaccharides, including poly-β-1,6-N-acetyl-D-glucosamine (PNAG), colonic acid, cellulose, alginate and the Pel and Psl polysaccharides of \textit{Pseudomonas aeruginosa} affect the development and structural integrity of various bacterial biofilms\textsuperscript{8-13}. Of these, the partially de-N-acetylated poly-β-1,6-N-acetyl-D-glucosamine (dPNAG) exopolysaccharides are...
required for the maintenance of biofilm structural stability in a diverse array of human pathogens, including *Yersinia pestis*, *Staphylococcus epidermidis*, *Burkholderia* spp., *Acinetobacter baumannii*, *Aggregatibacter* spp., *Klebsiella pneumoniae* and pathogenic *Escherichia coli* strains, where it may provide protection against biocides and immune killing.\(^{11,14-21}\)

In *E. coli*, four genes of the pgaABCD operon encode proteins required for the synthesis and translocation of dPNAG across the bacterial double-layer membrane.\(^{14,17}\) PgaA, a tetratricopeptide repeat (TPR)-containing outer membrane protein, forms the secretion pore in the outer membrane and translocates the dPNAG polymer from the periplasm to the cell surface.\(^{14,17}\) It is thought that the N-terminal TPR motifs of PgaA participate in protein-protein interactions in the periplasm, while the C-terminus of PgaA forms a transmembrane β-barrel pore that is dedicated for the secretion of the partially de-N-acetylated PNAG (dPNAG) polymers across the outer membrane. De-N-acetylation of PNAG is achieved by the putative lipoprotein PgaB. The partially de-N-acetylated PNAG is the functional form of the exopolysaccharide.\(^{14,17}\) Cell-bound PNAG, including the surface-exposed and periplasmic polymer, contains 3-5% de-N-acetylated glucosamine residues,\(^{14,17}\) while the polymer that has been released from cells into the spent medium contains more than this.\(^{22}\) The positively charged glucosamine residues help to retain the polymer on the bacterial surface by interaction with the negatively charged surface component lipopolysaccharides.\(^{22}\) But in addition, they are essential for secretion of the polymer as catalytically inactive PgaB proteins are unable to support biofilm formation and similar to the loss of PgaA, causing retention of PNAG polymers in the periplasm.\(^{14,23,24}\) PgaC, which is predicted to contain multiple transmembrane helices and a cytosolic domain homologous to family 2 glycosyltransferases (GT), is necessary for the synthesis PNAG. Finally, PgaD, a small membrane protein with two predicted transmembrane helices, assists the GT in polymerizing PNAG and is involved in the regulation of PNAG synthesis. The two inner membrane proteins PgaC and PgaD form a complex that functions as a novel type of c-di-GMP receptor, wherein ligand binding to the two proteins stabilizes their interaction and allosterically activates PNAG synthesis.\(^{25}\) The expression of the *E. coli* pgaABCD operon is tightly regulated. For example, the RNA binding protein CsrA (carbon storage regulator A) negatively regulates pgaA translation, pgaABCD transcription and mRNA stability.\(^{16,26,27}\)

Unlike various relatively well-characterized bacterial protein secretion machineries, mechanisms of exopolysaccharide secretion by outer-membrane secretins remain largely elusive. Recently, structural studies have advanced our understanding of several types of exopolysaccharide secretion across the bacterial outer membrane. For instance, Keisti et al. and Tan J et al. determined the crystal structure of AlgE, and Whitney et al. solved the crystal structure of AlgK. In combination with functional analyses, they proposed a mechanism for alginate secretion across the bacterial outer membrane.\(^{28-30}\) Little et al. reported the crystal structures of PgaB and its complex with β-1,6-(GlcNAc)\(_6\), and disclosed the catalytic mechanism of the N-deacetylase PgaB on PNAG polymers and the manner of periplasmic translocation of PNAG.\(^{23,24}\) Despite this progress, many questions regarding how dPNAG, an essential biofilm-supporting exopolysaccharide, is polymerized and translocated across the bacterial outer membrane remain to be answered. Here we present the crystal structure of the PgaA C-terminal transmembrane domain. Structure-based functional studies indicate that a characteristic electronegative patch that lies inside of the secretion pore, in proximity to the periplasm, is important for the translocation of the positively-charged dPNAG. In addition, the N-terminal TPR domain of PgaA directly interacts with exopolysaccharide N-deacetylase PgaB, which may play an important role in bringing the PgaB-bound dPNAG polymers proximal to the secretion pore. These findings offer to explain how the PgaA protein is structured specifically for secreting dPNAG and why fully acetylated PNAG cannot exit from the periplasm to serve as a biofilm-supporting exopolysaccharide.

**Experimental procedures**
Protein expression purification and crystallization DNA coding sequence of the full-length PgaA protein was amplified by polymerase chain reaction (PCR) using *E. coli* strain K-12 gemonic DNA (*E. coli* strain MG1655, ATCC) as template, and the PCR fragments were subsequently cloned into pBAD22 vector under control of the arabinose promoter. The expressed PgaA protein contains a C-terminal 6-histidine tag for affinity purification. Initially, we crystallized the full-length PgaA, but the crystals did not diffract beyond ~10 Å. To obtain high-quality crystals for structural determination, the purified full-length PgaA protein was subject to limited proteolysis with trypsin at a molar ratio of 100:1 (protein/enzyme) for 1 h at room temperature. Limited proteolysis produced a stable fragment that corresponds to the C-terminal transmembrane domain of PgaA with a molecular mass of 35 kDa (Figs 1A and 1B). N-terminal sequencing indicated that N-terminus of the fragment starts from residue Ala511 of PgaA. Subsequently, DNA coding sequence of this fragment was amplified by polymerase chain reaction (PCR) using *E. coli* strain MG1655 as template, and the PCR fragments were subsequently cloned into pBAD22 vector for over-expression. T

For purification, cells of *E. coli* strain sf100 were induced to express the PgaA C-terminal domain by adding 0.2% arabinose for 3 h at 37 °C and lysed on ice in phosphate-buffered saline (PBS) buffer using a French press high-pressure instrument (JN-02C, JNBio Ltd.) at 16,000 psi. Membranes were collected by centrifugation for 40 min at 39,000g, and solubilized in 1.0% (w/v) lauryldimethylamine oxide (LDAO, Anitrace) in PBS buffer. After centrifugation for 30 min at 39,000g, the supernatant solution was loaded onto a 10 ml Ni-NTA column (metal-chelating Sepharose, Amersham Pharmacia Biotech). The column was washed with 10 volumes of 0.5% LDAO in PBS buffer in the presence of 20 mM imidazole. Detergent exchange was performed by washing the column with 5 volumes of 1.0% n-octyl-β-D-glucopyranoside (OG, Anatrace) in a buffer containing 20 mM Tris, pH 8.0 and 150 mM NaCl. The PgaA β-barrel protein was eluted with a buffer containing 20 mM Tris, pH 8.0, 150 mM NaCl, 250 mM imidazole, 5% glycerol and 0.8% OG. Further purification was carried out by gel filtration on Superdex-200 HR 10/30 column (GE Healthcare Life Sciences) in buffer containing 20 mM Tris, pH 8.0, 150 mM NaCl and 0.8% OG. The PgaA β-barrel protein was concentrated to 10 mg ml⁻¹ with a Centricon-30 concentrator (Amicon; molecular mass cutoff, 30 kDa).

To improve the phasing power of selenomethionine (Se-Met) single-wavelength anomalous dispersion (SAD), two point mutations (Leu610Met and Ile743Met) were introduced into the PgaA β-barrel by overlapping PCR. Se-Met-substituted PgaA β-barrel protein was expressed in *E. coli* strain sf100 cells by inhibition of the methionine biosynthesis pathway. The cells were grown in M9 minimal medium with 1.5% (v/v) glycerol as carbon sources, and induced with 1% arabinose. Purification procedures were as described for the native protein.

Crystallization of the PgaA barrel was performed by hanging-drop vapor diffusion at 16 °C by mixing 1 µl each of protein (10 mg ml⁻¹) and reservoir solutions. Initial crystallization conditions were found using a broad screen. After optimization, the best crystals were obtained in a condition that contains 0.1 M NaOAc, pH 4.5, 31% MPD, 0.4 M NaCl, and 0.3 M NDSB-195. The crystals appeared overnight and grew to maximum dimensions of 70 × 70 × 90 µm within two weeks.

Data collection, structure determination and refinement Crystals were flash frozen in liquid nitrogen directly from the drop. Native datasets were collected at 100 K at a wavelength of 1.000 Å at beamline BL17U of Shanghai Synchrotron Radiation Facility (SSRF, China), and Se-Met datasets were collected at the selenium peak (λ = 0.9795 Å) at beamline BL1A of Japanese KEK synchrotron. The data were indexed and scaled with HKL2000 package. 31. The crystals belong to space group P6₁2₁2₁, and diffract to 2.8 Å and 3.4 Å for native and Se-Met crystals, respectively. Five out of six selenium sites were located using PHENIX. 32 After solvent flattening, electron density maps were generated with useful phases to ~3.5 Å. An initial model was manually built
with the program COOT, using selenium sites and large amino-acid side chains to determine the registry. The model was improved by an iterative process, using solvent flattening and model building. This model was used in the subsequent refinement of the native dataset, giving an R factor of 21.5% (R_{free} of 26.8%). Residues 567-572 of PgaA in the best-refined structure of the PgaA were not included, presumably because of conformational flexibility. The final refined model was analyzed with pro-check of CCP4 program suite and detailed statistics of crystallographic analyses are shown in Table 1. All figures shown in the paper were rendered with Pymol, and the coordinate file was deposited in PDB under ID code 4Y25.

Quantitative biofilm assay The wild-type E. coli strain MG1655 (MG) produces little or no biofilm because under laboratory conditions, the RNA binding protein negatively regulates pgaA translation, pgaABCD transcription and mRNA stability. MGcsrA::kan_ApgaA strains (E. coli strain MG1655 with both genes csrA and pgaA disrupted) were transformed with the pgaA expression plasmid. For these experiments, the wild-type protein and the various mutants contain a C-terminal 6-histidine tag. Overnight LB cultures (37 °C, 250 rpm) were inoculated 1:100 into LB medium containing 100 µg/ml ampicillin and 0.1% L-arabinose, and incubated in a 96-well microtiter plate (total volume of 200 µl per each well) at 26 °C for 24 h without shaking. Growth of planktonic cells after 24 h was determined by absorbance at 600 nm. To measure biofilm formation, planktonic cells were removed by discarding the medium and the wells were rinsed with water three times. Bound cells were stained with 1.25 volumes (250 µl) of crystal violet solution (Fisher Scientific; Cat # 255-960B) for 2 min and the wells were rinsed again three times with water to remove unbound and excess dye. Bound dye was solubilized with 1.25 volumes of 33% acetic acid and absorbance at 630 nm was determined using a microtiter plate reader. Background staining was corrected by subtracting the crystal violet bound to un-inoculated controls. Each experiment was conducted at least twice in replicates of three.

Isolation and immunoblotting of dPNAG Cultures were grown in LB medium with or without 100 µg/ml ampicillin and 0.1% L-arabinose at 26°C, 250 rpm for 24 h, cells were collected by centrifugation (10,000 rpm for 5 min) and dPNAG was isolated from cell-bound and spent medium fractions and detected following the procedures described earlier with the following modifications: The aqueous extracts of dPNAG were concentrated using Amicon Ultra Centrifugal filter (3,000 Da cutoff, Millipore). dPNAG concentrates were applied onto a nitrocellulose membrane (PROTRAN®, Whatman GmbH) and allowed to dry at 37 °C for 1 h. The membrane was blocked for 1 h in 5% BSA in PBS-0.5% Tween, washed twice for 15 min with PBS-0.5% Tween and treated for 1 h at room temperature with α-dPNAG monoclonal antibody (24 µg/ml in 1% BSA PBS-0.5% Tween) that was raised against E. coli dPNAG. The membrane was washed five times (15,5,5,5,5 min) with PBS-0.5% Tween and treated with horseradish-peroxidase-conjugated anti-murine IgM antibody (Sigma-Aldrich A8786; 1:10,000 diluted in 1% BSA PBS-0.5% Tween) for 1 h at room temperature. The membrane was washed again (15,5,5,5,5 min) and the signal detected by Chemiluminescence using Western Lighting Plus-ECL (PerkinElmer) and imaged using ChemiDoc.

Western blotting analysis of PgaA protein expression in bacteria To compare relative protein expression levels of pgaA mutants that have significantly reduced biofilm production to that of the wild-type strain, plasmids harboring wild-type plasmid or various pgaA mutants (all contain a C-terminal 6-histidine tag) were transformed into E. coli MGcsrA::kan_ApgaA strain, and spread on LB agar plates that contained 100 µg/ml ampicillin, 50 µg/ml kanamycin and 0.1 mg/ml L-arabinose. After incubation at 26 °C for 24 h, colonies were scraped off of the plates and the cell density was adjusted to O.D.600 of 1.0 by addition of 1×PBS buffer. Resuspended cells were homogenized by sonication and then centrifuged at 40,000g for 1 h to collect total membrane fractions. Inner membranes were solubilized by 1.0 % N-Laurylsarcosine and discarded after a second round of

Downloaded from http://www.jbc.org/ by guest on November 19, 2017
centrifugation at 40,000 g for 1 h. After that, outer membrane pellets were further solubilized using 1.0% LDAO. 20 μl outer membranes for each PgaA mutant were withdrawn and mixed with 10 μl SDS-loading dye. Subsequently, the sample was equally allocated into two Eppendorf tubes (each 15 μl), kept at room temperature or heated at 95 °C for 5 min prior to SDS-PAGE analysis. After electrophoresis, the proteins were transferred to a PVDF membrane and blocked using TBST buffer (20 mM Tris-HCl, pH 8.0, 150 mM NaCl, 0.05% Tween 20) that contained 8% skim milk for 2 h. The PVDF membrane was then incubated with anti-His antibody (1:3000) (TIANGEN) at room temperature for 2 h. PVDF membranes were subsequently washed with TBST buffer twice and further incubated with horseradish-peroxidase-conjugated secondary antibody (1:5000) (Trans™) at room temperature for 2 h. PVDF membranes were exposed using enhanced chemiluminescence reagents (EasySee Western Blot Kit, Trans™).

Protein co-expression and His-tag pulldown assay DNA coding sequences of the E. coli full-length PgaA, the TPR domain of PgaA (residues 1-512, including the signal peptide and a C-terminal 6-histidine tag) and the full-length PgaB (with or without a C-terminal 6-histidine tag) were cloned in pQLink vector, respectively. To co-express PgaA and PgaB, plasmids (pQLink-PgaA and pQLink-PgaB-His) were utilized to construct plasmid pQLink-PgaA-PgaB-His that is capable of co-expressing both PgaA and PgaB-His proteins. Similarly, to co-express PgaB and TPR-His, plasmids (pQLink-PgaB and pQLink-TPR-His) were utilized to construct plasmid pQLink-PgaB-TPR-His.

To test whether PgaB binds PgaA, three plasmids (pQlink-PgaA, pQlink-PgaB-His and pQlink-PgaA-PgaB-His) were individually transformed into E. coli BL21(DE3). Cells that carry each of the three plasmids were induced with 0.4 mM IPTG for 3 h at 37 °C, and lysed on ice in phosphate-buffered saline (PBS) buffer by sonication. Membranes were collected by centrifugation for 40 min at 40,000 g, and solubilized in 1.0% (w/w) n-dodecyl-β-D-maltoside (DDM, Anatrace) in PBS buffer. After centrifugation for 30 min at 40,000 g, the extract was loaded onto a 1 ml Ni column. The column was washed with 5 volumes of 0.5% DDM in PBS buffer in the presence of 20 mM imidazole. The proteins bound to the Ni-NTA beads were eluted with a 1 ml buffer containing 20 mM Tris, pH 8.0, 150 mM NaCl, 250 mM imidazole, and 0.1% DDM. The eluted samples (15 μl) were subject to 12% SDS-PAGE analysis. Detection of the interaction between TPR and PgaB proteins was performed in a similar procedure as described for PgaA and PgaB-His.

Results

Crystal structure of PgaA(513-807) Based on prediction, the full-length PgaA contains 8 periplasm-localized tetratricopeptide repeats at the N-terminus and a C-terminal β-barrel domain located in the outer membrane (Fig. 1A). TPR domains are known primarily to be involved in protein-protein interactions, while the β-barrel domain of PgaA may serve as the portal and substrate-specificity determinant for dPNAG secretion. Our initial attempts to crystallize the full-length PgaA failed to produce high-resolution crystals for structural determination. To this end, we performed limited proteolysis on the full-length PgaA protein. Limited proteolysis with trypsin produced a stable fragment that corresponds to the transmembrane β-barrel domain of PgaA (residues 513-807) (Fig. 1B). Subsequently, we cloned the coding sequence of PgaA (residues 513-807) in addition to a C-terminal 6-histidine tag into pBAD22 vector with an N-terminal signal peptide coding sequence from E. coli BamA for efficiently targeting the PgaA transmembrane domain to the outer membrane for over-expression. PgaA(513-807) was crystallized in space group P6122 with one molecule in the asymmetric unit. Both native and seleno-methionine-incorporated PgaA(513-807) crystals were grown in the presence of the nonionic detergent n-octyl-β-D-glucoside (OG) and the structure was solved by SeMet single-wavelength anomalous dispersion (SAD) by introducing two extra Met residues into the protein sequence (L610M and I743M). Refinement produced final models with good geometry and R factor of 21.5% (Rfree factor of 26.8%) for the native data (Table 1). In the final refined model, residues 567–572 of PgaA (part of the extracellular loop 2) for the native structure
were not included because there was no interpretable electron density present for these residues.

PgaA (513-807) forms a 16-strand β-barrel that spans the outer membrane with antiparallel and interlocking B1 and B16 strands (Fig. 1C). Unlike AlgE, an 18-strand outer membrane protein that is responsible for exopolysaccharide alginate secretion, no narrow pore constriction is found in the β-barrel of PgaA and no extracellular or periplasmic loops occlude the lumen (Figs 1C and 2A)²⁹. However, on the extracellular side, the PgaA β-barrel is mainly closed by four long extracellular loops eL1, eL4, eL6 and eL8 (Fig. 1C). All of the periplasmic loops of the PgaA β-barrel are relatively short, a common feature found in most β-barrel outer membrane proteins. In search of similar structures using DALI, the β-barrel domain of PgaA was found to be related to a large number of outer membrane proteins such as TamA³⁸, FhaC³⁹ and some outer membrane porins with top Z-scores ranging from 14 to 16.4. Superposition of PgaA β-barrel with that of TamA, FhaC and outer membrane porin OmpG produced root mean square deviations (RMSDs) 3.2 Å for 278 Ca atoms, 3.2 Å for 280 Ca atoms and 3.5 Å for 265 Ca atoms, respectively. Apparently, these three families of outer membrane proteins have significantly different overall β-barrel shape and extracellular loop conformations in comparison to those of the PgaA β-barrel (Fig. 2B).

Electrostatic potential surface representation of PgaA β-barrel revealed that the extracellular surface of PgaA β-barrel is primarily electronegative (Fig. 3A). Remarkably, half of the interior surface of PgaA β-barrel along the substrate translocation pathway is electronegative, whereas the opposite half is mainly neutral with several dispersed positively-charged residues located close to the periplasmic space (Fig. 3B).

A cluster of negatively-charged residues, inside of the β-barrel lumen and proximal to the periplasm, is important for biofilm formation. In many bacterial outer membrane proteins, long extracellular loops play important roles in regulating substrate translocation. To investigate the functional roles of the extracellular loops of PgaA in dPNAG secretion, a pgaA-deletion mutant (E. coli. Strain MG_csrA::kan_ΔpgaA) was complemented with plasmids that express various mutant forms of PgaA. As reported previously, the pgaA-deletion mutant does not secrete dPNAG and therefore produces very low levels of biofilm, but was unaffected in its growth rate. When the pgaA gene was introduced on a plasmid, high levels of biofilm production were restored (Fig. 4A). Using this complementation assay, we evaluated the production of biofilm by several loop-deletion pgaA mutants. Surprisingly, when we tested for complementation by various extracellular loop-deletion mutants pgaA_ΔeL1, pgaA_ΔeL2, pgaA_ΔeL4, pgaA_ΔeL6 and pgaA_ΔeL8, all of them were found to restore biofilm formation (Fig. 4A). We also tested biofilm production by several pgaA mutants pgaA_ΔeL1_2, pgaA_ΔeL2_4 and pgaA_ΔeL6_8, which have two loops deleted. Similarly, these mutants also restored biofilm formation (Fig. 4A). Finally, we examined the ability of these complemented strains to secrete dPNAG using an established approach. In all cases, we found that plasmids encoding the wild type PgaA protein or proteins containing single or dual deletions of the extracellular loops were able to complement the pgaA deletion (Figs. 4A and 4B). Taken together, these studies strongly suggest that, unlike alginate secretion, the extracellular loops of PgaA β-barrel seem not to be important for biofilm formation.

In search of the elements of PgaA β-barrel that are important for dPNAG secretion, we next tested whether the charged residues inside of the β-barrel lumen play roles in biofilm production. In the interior of the β-barrel, 12 negatively-charged residues and 5 positively-charged residues (not counting the charged residues located on the extracellular loops and short periplasmic loops) that constitute a negatively-charged patch and a positively-charged patch, respectively (Fig. 3B). We mutated each of the negatively-charged residues (Glu or Asp) to the positively-charged residue Lys, and mutated the positively-charged residues (Arg or Lys) to negatively-charged residue Asp. Subsequently, biofilm production by all of these single-point mutants was evaluated using the complementation assay. Of the 12 negative-to-positive mutations, E741K, D777K,
E800K and D802K significantly affected biofilm formation; each of them caused about 20-fold reduction of biofilm production in comparison to the control strain (MG_csrA::kan ΔpgaA_pgaA), while the other 5 positive-to-negative mutation caused only modestly reduced biofilm production (Fig. 4C). To examine whether biofilm production correlates with dPNAG secretion, we grew a series of strains in cultures shaking at 26 °C, 250 rpm for 24 h, which allows secreted dPNAG to be released into the spent medium, and detected this polymer by immunoblotting. The cell-bound dPNAG was also isolated and detected to determine whether an effect on secretion of the polymer by a strain might be due to a defect in dPNAG synthesis. Consistent with previous observation, the parent strain (MG_csrA::kan) was efficient in secreting most of its dPNAG into the spent medium, while the MG_csrA::kan ΔpgaA mutant failed to secrete this polymer (Fig. 4D, lanes 1 and 2). Expression of the wild-type pgaA gene from a plasmid was able to complement the observed ΔpgaA defect (Fig. 4D, lane 3), but the resulting strain retained a greater amount of the polymer in the cell-bound fraction than the parent strain, suggesting that the PgaA protein expressed from the pgaABCD genomic locus may function more effectively than the one from a multicopy plasmid. The pgaA mutant proteins with lysine substituted at the 4 critical residues of the electronegative patch of the PgaA channel (lanes 4-7) were all found to be defective for secretion, although they were capable of synthesizing and accumulating dPNAG in the cell-bound fraction. Two mutants with substitution of acidic residues that did not affect biofilm formation (Fig. 4C) were also able to secrete dPNAG (Fig. 4D, lanes 8 and 9). These studies suggest that biofilm production reflected the amount of dPNAG secretion into the spent medium for both wild-type strain and pgaA mutants. Furthermore, Western blot analysis revealed that all four point mutants of PgaA (E741K, D777K, E800K and D802K) that significantly decreased biofilm production were expressed at levels similar to or higher than the wild-type protein (Fig. 4E). In addition, both the wild-type PgaA and PgaA mutant proteins exhibited heat-modifiable mobility difference on SDS-PAGE, suggesting that all of these proteins are folded and reside in the outer membrane (Fig. 4E). These experiments argue that the reduced biofilm production of these mutants is not caused by decreased expression of the mutant protein, but by a decreased ability to secrete dPNAG across the outer membrane. Intriguingly, all of the negative-to-positive mutants that significantly reduced biofilm production were clustered at a position that is proximal to the periplasm and they together form an electronegative patch inside the β-barrel (Fig. 4F). Thus, it is likely that this electronegative patch may make initial contacts with the positively-charged dPNAG polymer and thereby help to usher the polymer into the secretion pore of PgaA.

Both the TPR domain and β-barrel pore of PgaA are required and specific for dPNAG secretion. Of several types of exopolysaccharide outer-membrane secretins, the alginate secretin is among the best characterized. It mainly consists of two proteins: an 18-strand β-barrel outer-membrane protein, AlgE, which forms the secretion pore, and a TPR-containing lipoprotein, AlgK, which facilitates AlgE biogenesis and potentially interacts with other periplasmic components of the secretion machinery. In addition to AlgE, functional studies demonstrated that AlgK is also required for alginate secretion. Distinct from the alginate secretin, PgaA contains two functional domains with an N-terminal TPR domain that might have functions similar to AlgK and a C-terminal transmembrane secretion pore. To test whether the TPR domain of PgaA is required for dPNAG secretion, we studied the biofilm production of several TPR-deletion PgaA mutants that have different helices of the TPR motifs deleted (Fig. 5A). As shown in Figs 5B and 5C, all of the tested TPR-deleted pgaA mutants significantly decreased biofilm production and dPNAG secretion in the spent medium, although their expression levels were close to or even higher than that of the wild-type protein (Fig. 5D). Furthermore, we investigated biofilm production by two fusion proteins by replacing the two domains of PgaA with either AlgK or AlgE. Likewise, AlgK_PgaA-β-barrel fusion protein that was expressed in the bacterial outer membrane also did not support biofilm production and dPNAG secretion (Figs 5B, 5C and 5D, PgaA-TPR_AlgE fusion protein seemed not to be expressed in the bacteria).
together, these studies show that the TPR domain of PgaA is important for biofilm production and suggest that both the TPR domain and the β-barrel domain of PgaA are required and specific for dPNAG secretion.

*The N-terminal TPR domain of PgaA directly interacts with PgaB* Distinct from the outer-membrane alginate secretin that mainly consists of AlgK and AlgE, the dPNAG secretin includes an extra lipoprotein, PgaB, which is also required for dPNAG secretion and biofilm production. To further understand how the PNAG secretin is assembled and why the TPR domain of PgaA is required for biofilm production, we studied the interactions between PgaA and PgaB. To this end, DNA coding sequences of both PgaA and PgaB in addition to a C-terminal 6-histidine tag were cloned in pQLink vector, respectively. Subsequently, a plasmid pQLink-PgaA-PgaB-His was constructed, which is capable of co-expressing both PgaA and PgaB proteins upon induction in bacteria. Plasmids pQLink-PgaA, pQLink-PgaB-His and pQLink-PgaA-PgaB-His were transformed into *E. coli* BL21(DE3) competent cells for protein over-expression, respectively. As shown in Fig. 5E, His-tagged PgaB protein was capable of pulling down a protein that has a molecular mass corresponding to that of PgaA as revealed by SDS-PAGE analysis. Protein identification using MALDI-TOF mass spectrometry (KRATOS Analytical, Shimadzu corporation, Japan) confirmed that the suspected protein is PgaA. Furthermore, we found that the PgaB and PgaA proteins from the pull-down mixture co-migrate on size exclusion column (Fig. 5F), suggesting that PgaA and PgaB form a stable complex. As both the TPR domain of PgaA and the PgaB protein are localized in periplasm, next, we tested whether the TPR domain of PgaA alone is responsible for the interaction. When the TPR domain of PgaA and PgaB were co-expressed in the bacteria, the TPR domain was capable of pulling down PgaB, indicating that the TPR domain of PgaA directly binds PgaB (Fig. 5E). Taken together, these experiments show the N-terminal TPR domain of PgaA interacts with PgaB and suggest that this interaction may play important roles in the assembly of the PNAG secretin by bringing the PgaB-bound dPNAG polymer proximal to the PgaA secretion pore.

**Discussion**

Similar to various bacterial protein secretion systems, outer-membrane polysaccharide secretins differ in composition of the components and the manner of polysaccharide translocation. For instance, the export of the tightly surface-associated capsular polysaccharides across outer membrane occurs via a novel type of octameric α-helical pore formed by Wza and its ortholog KpsD, while the exopolysaccharide alginate secretin AlgE forms an 18-strand β-barrel pore with an arginine-rich highly electropositive pore constriction acting as a selectivity filter for the negatively-charged alginate polymer. Apart from this, the extracellular loop L2 and a long periplasmic loop T8 of the AlgE β-barrel seem to play important roles in regulating alginate secretion (Fig. 2A). Here, we showed that the transmembrane domain of the dPNAG secretin PgaA forms a 16-strand β barrel. Distinct from alginate secretion by AlgE, in which both extracellular L2 and long periplasmic loop T8 occlude the barrel lumen, the extracellular loops that close the lumen of PgaA do not appear to occlude the lumen or play an important role in substrate secretion. Rather, a cluster of negatively-charged residues inside of the PgaA barrel lumen and in proximity to the periplasm is important for dPNAG secretion. This finding correlates well with the essential role of the lipoprotein PgaB in dPNAG secretion, which functions as a PNAG deacetylase, conferring a positive charge to the polymer by partial de-N-acetylation. We propose that complementary charge interactions between the secretion pore and the exopolysaccharide substrate are important for the secretion of both the negatively-charged alginate and the positively-charged dPNAG across the bacterial outer membrane. Our current hypothesis is also consistent with the observation that an electronegative patch on the N-terminal domain of PgaB interacts with the positively-charged glucosamine as demonstrated by molecular simulation studies. Such unusual pore electrostatics has been previously observed in other outer membrane proteins and has been proposed to have consequences for solute transport.
translocation or substrate recognition. For example, Im et al. found that the electrostatic interactions between ions and the charge distribution of the OmpF channel play an important role in governing ion permeation and selectivity\(^{44}\). An additional example of this type of pore feature that may support dPNAG translocation is the electropositive extracellular surface features of PorB from pathogenic *Neisseria meningitides* may mediate Toll-like receptor recognition through a nonspecific electrostatic attraction\(^{45,46}\).

In our structure, the PgaA barrel pore is closed on the extracellular surface by several loops, whereas the PgaA barrel is fully open on the periplasmic side, with a dimension of approximately 25 Å × 35 Å in diameter. Given the linear partially de-N-acetylated poly-β-1,6-N-acetyl-D-glucosamine (dPNAG) polymer is about 10 Å in width\(^{23}\), the lumen of PgaA barrel is large enough to allow the passage of dPNAG polymer through the pore. However, it is unclear how the binding of dPNAG to PgaA elicits the dislocation of the extracellular loops that allows dPNAG to enter into the extracellular space. One possibility could be that the electronegative attributes of the PgaA extracellular surface and/or the lipopolysaccharides present in the outer leaflet of the outer membrane could facilitate this process\(^{22}\).

We also provided evidence that the TPR motifs of PgaA are important for the secretion of dPNAG. In agreement with this, the TPR-containing protein AlgK is also required for exopolysaccharide alginate secretion\(^{26}\). However, the β-barrel of PgaA alone or a fusion protein composed of the AlgK and the PgaA β-barrel were not functional, suggesting that both the TPR domain and the β-barrel of PgaA are necessary and specific for dPNAG secretion. The observed direct interaction between the TPR domain of PgaA and PgaB indicates that this interaction may position the PgaB-bound dPNAG in proximity to the PgaA secretion pore as a prerequisite for dPNAG interaction. However, it remains to be seen if the de-N-acetylated PNAG polymer directly binds to the PgaA β-barrel and/or the TPR domain of PgaA, and how they exactly interact in the process of translocation. Further biochemical characterization of these interactions and structural studies of PgaA in complex with dPNAG and PgaB should help to clarify these issues.

In summary, our crystal structural studies and structure-based functional analysis provide important clues to explain how dPNAG is translocated across the bacterial outer membrane. Improving our understanding of the mechanism underlying dPNAG synthesis and translocation is of great biomedical interest given that it is critical for biofilm production by several important human pathogens\(^{47,48}\).

**Acknowledgments** We thank Yan Zhao, and staff members of beamline BL1A of Japanese KEK synchrotron and Shanghai Synchrotron Radiation Facility (China) synchrotron facilities for their assistance in collecting diffraction data. This work was supported by National Natural Science Foundation of China Grants 31470743 and 31170698 (to Y.H.), Ministry of Science and Technology (China) “973” Project Grants 2013CB910603 and 2012CB917302 (to Y.H.), Chinese Academy of Sciences Grant XDB08020302 (to Y.H.) and National Institutes of Health (U.S.) grants GM059969 and GM066794 (to T.R.) and University of Florida CHRIS project FLA-MCS-004949 (to T.R.).

The authors declare that no conflicts of interest exist.

References


**Data deposition** The coordinates and diffraction data have been deposited in the Protein Data Bank with the accession code 4Y25 for PgaA (513-807).
Figure legends

Figure 1. Molecular structure of PgaA (513-807). (A) Schematic structure of the full-length PgaA. Residues 1-33 constitute a signal peptide. Eight TPR motifs are denoted as R1-R8. The C-terminus of PgaA (residues 513-807) forms a transmembrane β barrel. (B) The full-length PgaA protein purified from membrane fraction and its trypsin-digested product. Both the full-length PgaA and PgaA β-barrel exhibit heat-modifiable mobility to certain degree on 12% SDS-PAGE. (C) Ribbon representation of the structure of PgaA (513-807). The PgaA β-barrel consists of 16 β-strands, closed on the extracellular surface by four long extracellular loops eL1, eL4, eL6 and eL8.

Figure 2. Structural comparison between PgaA (513-807) and other representative β-barrel outer membrane proteins. (A) Ribbon representation of the structure of AlgE (PDB: 3RBH). AlgE β-barrel consists of 18 β-strands and the lumen is occluded by two functionally important loops: a long periplasmic loop 8 (T8, in red) and an extracellular loop 2 (L2, in blue). (B) Structural overlay between PgaA β-barrel and that of TamA (PDB: 4C00), FhaC (PDB: 3NJT) or OmpG (PDB: 2F1C).

Figure 3. Electrostatic potential surface representation of PgaA (513-807). (A) Electrostatic potential surface representation showing that the extracellular surface of PgaA is primarily electronegative. (B) A slab view showing the electrostatic properties inside of the lumen of PgaA. Surface colored according to electrostatic potential, ranging from blue (+70 kT/e) to red (-70 kT/e), where kT is thermal energy and e is the elementary charge. The characteristic negatively-charged patch and the positively-charged patch inside of the β barrel are indicated with dotted lines.

Figure 4. A cluster of negatively-charged residues inside of the barrel and in proximity to periplasm is important for dPNAG secretion. (A) Extracellular loops of PgaA are not important for biofilm production. Wild-type E. coli strain MG1655 (denoted as MG) produced low levels of biofilm due to repression by CsrA (A630<0.1); the csrA-disrupted E. coli strain (denoted as MG_csrA::kan) produced 20-fold more biofilm (A630=2.0); both the pgaA-deleted E. coli strain (denoted as MG ΔpgaA) and E. coli strain MG1655 with both csrA- and pgaA- disrupted (denoted as MG_csrA::kan ΔpgaA) produced low levels of biofilm (A630<0.1). When MG_csrA::kan ΔpgaA was complemented with a plasmid that expresses PgaA (MG_csrA::kan ΔpgaA_pgaA), biofilm production was restored. None of the loop-deletion mutants affected biofilm production. (B) Extracellular loops of PgaA are not important for secretion of dPNA. Immunoblotting of secreted and cell-bound dPNA from MG1655 csrA::kan Δ pgaA strain complemented with plasmids expressing wild-type pgaA or mutant pgaA proteins with extracellular loop deletions. Secreted and cell-bound dPNA samples were blotted onto a nitrocellulose membrane and detected with α-PNAG monoclonal antibody as described14. Rows A - C: dPNA derived from 30, 6 and 2 μl, respectively, of spent medium. Rows D - F: Cell-bound dPNA from 40, 4 and 1 ml of culture. Dilutions of purified dPNA (bottom row) served as a standard for comparison. (C) Twelve negatively-charged residues (E519, D541, E583, D696, E727, E737, E741, D759, D777, E785, E800 and D802) and five positively-charged residues (R609, R621, R652, K660 and R686) inside the lumen were mutated to Lys and Asp, respectively, and their effects on biofilm production were assayed. Each experiment was conducted at least twice in replicates of three. Error bars denote standard deviations from the means. (D) Immunoblotting of secreted and cell-bound dPNA from MG_csrA:: kan ΔpgaA strains expressing wild-type PgaA or single-point PgaA mutants. Duplicate samples represent dPNA extracted from two independent cultures. (E) Nickel-column purification and western blot analysis of the expression levels of both the wild-type and mutant proteins from bacterial outer membrane. Both the wild-type and mutant proteins exhibited apparent heat-modifiable mobility, suggesting that they are all folded and resided in the outer membrane. (F) Residues E741, D777, E800 and D802 of PgaA form a negatively-charged residue cluster in proximity to the periplasm.
Figure 5. The TPR domain of PgaA directly binds PgaB and is necessary for biofilm production. (A) Schematic structures of the full-length PgaA, TPR-deletion mutant PgaA proteins and two fusion proteins (AlgK_PgaA-β and PgaA-TPR_AlgE). (B) TPR domain of PgaA is required for biofilm production. TPR repeats-deletion mutants ΔTPR(R1-R8) (Mr, 35.1kDa), ΔTPR(R1-R2) (Mr, 76.3kDa), ΔTPR(R3-R4) (Mr, 73.5kDa) and ΔTPR(R6-R8) (Mr, 73.8kDa) all significantly affected biofilm production; two fusion proteins AlgK_PgaA-β (Mr, 76.3kDa) and PgaA-TPR_AlgE (Mr, 89.3kDa) are unable to restore biofilm production. Each experiment was conducted at least twice in replicates of three. Error bars denote standard deviations from the means. (C) Immunoblotting of secreted and cell-bound dPNAG from MG_csrA::kan ΔpgaA strain expressing wild-type PgaA or PgaA with TPR deletions or PgaA fusion proteins. (D) Nickel-column purification and western blot analysis of the expression levels of the wild-type, various TPR-deletion PgaA mutant proteins and fusion protein AlgK_PgaA-β-barrel from bacterial outer membranes. All the proteins exhibited apparent heat-modifiable mobility, suggesting that they are all folded and resided in the outer membranes. (E) The TPR domain of PgaA directly binds PgaB. When co-expressed in bacteria, PgaB-His is capable of pulling down PgaA (left panel), and TPR-His is capable of pulling down PgaB (right panel). Pulldown mixtures were resolved on 12% SDS-PAGE gels and were stained with Coomassie blue dye. (F) Size exclusion chromatographic profile of the pull-down mixture of PgaB-His and PgaA on a Superdex 200 10/30 column. PgaB-His (a lipoprotein) alone is not stable and forms aggregates in a gel filtration buffer that contains 20 mM Tris, pH 8.0, 150 mM NaCl and 0.06% DDM, so excess PgaB-His protein was eluted in a void volume (about 8.0 ml). PgaB-His and PgaA seem to form a 1:1 stable complex as resolved by 12% SDS-PAGE analysis.
Table 1. Data collection and refinement statistics

<table>
<thead>
<tr>
<th>Statistics</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Data collection</strong></td>
<td></td>
</tr>
<tr>
<td>Protein crystal</td>
<td>Native, Se-Met (L610M+I743M)</td>
</tr>
<tr>
<td>Beamline</td>
<td>BL17U (SSRF), BL1A (KEK)</td>
</tr>
<tr>
<td>Wavelength (Å)</td>
<td>1.00000, 0.9795</td>
</tr>
<tr>
<td>Space group</td>
<td>P6_22, P6_22</td>
</tr>
<tr>
<td>Cell dimensions</td>
<td></td>
</tr>
<tr>
<td>(a, b, c) (Å)</td>
<td>144.0, 144.0, 90.6, 144.4, 144.4, 88.4</td>
</tr>
<tr>
<td>(\alpha, \beta, \gamma) (°)</td>
<td>90.0, 90.0, 120.0, 90.0, 90.0, 120.0</td>
</tr>
<tr>
<td>Resolution (Å)</td>
<td>50-2.8 (2.9-2.8), 50-3.4 (3.5-3.4)</td>
</tr>
<tr>
<td>(R_{sym}) (%)</td>
<td>6.9 (&gt;100), 13.5 (83.8)</td>
</tr>
<tr>
<td>Completeness (%)</td>
<td>94.4 (91.8), 100.0 (100.0)</td>
</tr>
<tr>
<td>(I/\sigma(I))</td>
<td>28.9 (1.9), 28.9 (4.5)</td>
</tr>
<tr>
<td>Unique reflections (n)</td>
<td>13,051 (1,227), 8,026 (771)</td>
</tr>
<tr>
<td>Redundancy</td>
<td>10.9 (10.5), 21.1 (21.8)</td>
</tr>
<tr>
<td><strong>Refinement</strong></td>
<td></td>
</tr>
<tr>
<td>Resolution (Å)</td>
<td>32.3-2.8</td>
</tr>
<tr>
<td>Reflections (n)</td>
<td>13043</td>
</tr>
<tr>
<td>(R_{work} / R_{free}) (%)</td>
<td>21.5 / 26.8</td>
</tr>
<tr>
<td>Number of atoms (n)</td>
<td></td>
</tr>
<tr>
<td>Protein</td>
<td>2402</td>
</tr>
<tr>
<td>Solvent</td>
<td>12</td>
</tr>
<tr>
<td>B-factor (Å²)</td>
<td>90.59</td>
</tr>
<tr>
<td><strong>RMSD</strong></td>
<td></td>
</tr>
<tr>
<td>Bond length (Å)</td>
<td>0.010</td>
</tr>
<tr>
<td>Bond angles (°)</td>
<td>1.418</td>
</tr>
<tr>
<td>Ramachandran plots</td>
<td></td>
</tr>
<tr>
<td>In preferred regions</td>
<td>267 (93.36%)</td>
</tr>
<tr>
<td>In Allowed regions</td>
<td>19 (6.64%)</td>
</tr>
<tr>
<td>Outliers</td>
<td>0 (0.00%)</td>
</tr>
</tbody>
</table>

Values in parentheses refer to the highest resolution shell.

\(R_{free}\) was calculated in identical manner to \(R_{work}\), using a 5% test set of randomly selected reflections that were omitted from the refinement.
Figure 3

(A) (Extracellular view)

(B) Negatively-charged patch

(Side view) Positively-charged patch
Structural Basis for Translocation of a Biofilm-supporting Exopolysaccharide Across the Bacterial Outer Membrane
Yan Wang, Archana AndolePannuri, Dongchun Ni, Haizhen Zhou, Xiou Cao, Xiaomei Lu, Tony Romeo and Yihua Huang

*J. Biol. Chem.* published online March 8, 2016

Access the most updated version of this article at doi: 10.1074/jbc.M115.711762

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts