



# Evidence for a Widespread Third System for Bacterial Polysaccharide Export across the Outer Membrane Comprising a Composite OPX/ $\beta$ -Barrel Translocon

Johannes Schwabe,<sup>a</sup> María Pérez-Burgos,<sup>a</sup> Marco Herfurth,<sup>a</sup> Timo Glatter,<sup>b</sup>  Lotte Søgaard-Andersen<sup>a</sup>

<sup>a</sup>Department of Ecophysiology, Max Planck Institute for Terrestrial Microbiology, Marburg, Germany

<sup>b</sup>Core Facility for Mass Spectrometry & Proteomics, Max Planck Institute for Terrestrial Microbiology, Marburg, Germany

**ABSTRACT** In Gram-negative bacteria, secreted polysaccharides have multiple critical functions. In Wzx/Wzy- and ABC transporter-dependent pathways, an outer membrane (OM) polysaccharide export (OPX) type translocon exports the polysaccharide across the OM. The paradigm OPX protein Wza of *Escherichia coli* is an octamer in which the eight C-terminal domains form an  $\alpha$ -helical OM pore and the eight copies of the three N-terminal domains (D1 to D3) form a periplasmic cavity. In synthase-dependent pathways, the OM translocon is a 16- to 18-stranded  $\beta$ -barrel protein. In *Myxococcus xanthus*, the secreted polysaccharide EPS (exopolysaccharide) is synthesized in a Wzx/Wzy-dependent pathway. Here, using experiments, phylogenomics, and computational structural biology, we identify and characterize EpsX as an OM 18-stranded  $\beta$ -barrel protein important for EPS synthesis and identify AlgE, a  $\beta$ -barrel translocon of a synthase-dependent pathway, as its closest structural homolog. We also find that EpsY, the OPX protein of the EPS pathway, consists only of the periplasmic D1 and D2 domains and completely lacks the domain for spanning the OM (herein termed a <sup>D1D2</sup>OPX protein). *In vivo*, EpsX and EpsY mutually stabilize each other and interact in *in vivo* pulldown experiments supporting their direct interaction. Based on these observations, we propose that EpsY and EpsX make up and represent a third type of translocon for polysaccharide export across the OM. Specifically, in this composite translocon, EpsX functions as the OM-spanning  $\beta$ -barrel translocon together with the periplasmic <sup>D1D2</sup>OPX protein EpsY. Based on computational genomics, similar composite systems are widespread in Gram-negative bacteria.

**IMPORTANCE** Bacteria secrete a wide variety of polysaccharides that have critical functions in, e.g., fitness, surface colonization, and biofilm formation and in beneficial and pathogenic human-, animal-, and plant-microbe interactions. In Gram-negative bacteria, export of these chemically diverse polysaccharides across the outer membrane depends on two known translocons, i.e., an outer membrane OPX protein in Wzx/Wzy- and ABC transporter-dependent pathways and an outer membrane 16- to 18-stranded  $\beta$ -barrel protein in synthase-dependent pathways. Here, using a combination of experiments in *Myxococcus xanthus*, phylogenomics, and computational structural biology, we provide evidence supporting that a third type of translocon can export polysaccharides across the outer membrane. Specifically, in this translocon, an outer membrane-spanning  $\beta$ -barrel protein functions together with an entirely periplasmic OPX protein that completely lacks the domain for spanning the OM. Computational genomics support that similar composite systems are widespread in Gram-negative bacteria.

**KEYWORDS** beta-barrel proteins, export of polysaccharides, *Myxococcus xanthus*, OM translocon, OPX proteins, synthase-dependent pathway, Wzx/Wzy pathway, capsular polysaccharide, exopolysaccharide

**Editor** Carmen Buchrieser, Institut Pasteur

**Copyright** © 2022 Schwabe et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Lotte Søgaard-Andersen, [sogaard@mpi-marburg.mpg.de](mailto:sogaard@mpi-marburg.mpg.de).

The authors declare no conflict of interest.

**Received** 15 July 2022

**Accepted** 25 July 2022

**Published** 16 August 2022

Most bacteria secrete one or more polysaccharides. These polysaccharides protect against environmental stresses and phage infection, contribute to surface colonization and biofilm formation, have important functions in beneficial and pathogenic human-, animal- and plant-microbe interactions, provide the basis for serotyping and several antibacterial vaccines, and have many applications in the food, pharmaceutical, and medical industries (1–3). Here, we focus on the export of polysaccharides across the outer membrane (OM) in Gram-negative bacteria.

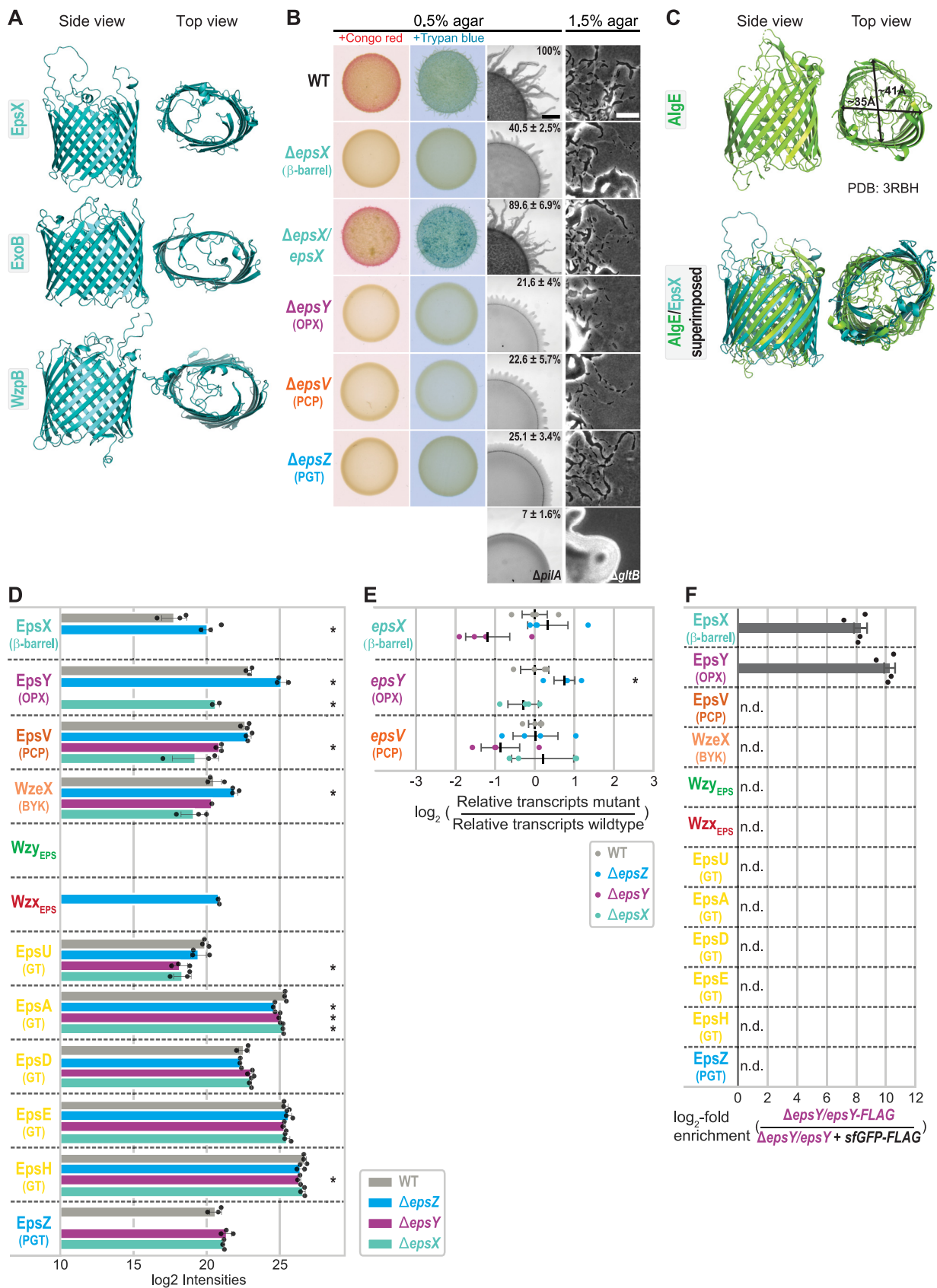
Secreted polysaccharides are large, chemically diverse molecules. Their synthesis hinges on three mechanisms, and their export across the OM in Gram-negative bacteria hinges on two known mechanisms (4, 5). In Wzx/Wzy-dependent pathways, biosynthesis is initiated on the cytoplasmic side of the inner membrane (IM) by a phosphoglycosyltransferase (PGT). Subsequently, glycosyltransferases (GTs) add monosaccharides to generate the repeat unit. The Wzx flippase “flips” individual repeat units across the IM to the periplasm, where the Wzy polymerase polymerizes them. On the periplasmic side, the polysaccharide copolymerase (PCP), an integral IM protein with an extended periplasmic domain, regulates polymerization and polysaccharide transfer across the periplasm to the OM (5, 6). A protein of the OM polysaccharide export (OPX) family exports the polysaccharide across the OM (7, 8). Wza of *Escherichia coli* is the best-studied OPX protein. In contrast to other pore-forming OM proteins, the octameric Wza spans the OM using an  $\alpha$ -helical barrel (7). The periplasmic part of the OPX protein interacts with the oligomeric PCP to establish a conduit for the polysaccharide to reach the OM (6, 8–10). In ABC transporter-dependent pathways, synthesis of the polysaccharide is also initiated on the cytoplasmic side of the IM and fully completed before translocation across the IM by an ABC transporter (5). Transfer across the periplasm involves a PCP, and the polysaccharide is exported across the OM by an OPX protein (5, 6). Synthase-dependent pathways differ considerably from these two pathways and consist of only three core components (4). An IM-embedded synthase with GT activity synthesizes the polysaccharide and, in parallel, facilitates its translocation across the IM. A tetratricopeptide repeat (TPR)-containing protein contributes to polysaccharide transfer across the periplasm. The polysaccharide is exported across the OM through an integral OM 16- to 18-stranded  $\beta$ -barrel protein as described in structural work on PgaA and BcsC of *E. coli* and AlgE of *Pseudomonas aeruginosa* (11–13).

Here, we combine experiments, computational structural biology, and genomics to provide evidence for a novel type of polysaccharide OM export mechanism widespread in Gram-negative bacteria. In this system, a short OPX protein comprises only periplasmic domains and lacks the domain for spanning the OM functions together with an OM  $\beta$ -barrel protein to generate a composite OPX/ $\beta$ -barrel translocon.

## RESULTS AND DISCUSSION

**Myxobacterial gene clusters for secreted polysaccharides encode an OM  $\beta$ -barrel protein.** The Gram-negative deltaproteobacterium *Myxococcus xanthus* secretes three polysaccharides, i.e., exopolysaccharide (EPS), spore coat polysaccharide (SPS), and biosurfactant polysaccharide (BPS), using three dedicated Wzx/Wzy-dependent pathways (see Fig. S1A in the supplemental material) (14). In all three systems, the gene annotated as encoding the OPX protein (EpsY, ExoA, WzaB) is syntenic with a gene encoding a protein of unknown function (EpsX, ExoB, WzpB). This synteny is largely conserved in the orthologous myxobacterial gene clusters (Fig. S1A).

EpsX, ExoB, and WzpB have a type 1 signal peptide based on sequence analysis. AlphaFold structural models (see Materials and Methods) predict with high accuracy that they fold into 18-stranded antiparallel  $\beta$ -barrels with an elliptical shape and a central channel (Fig. 1A; Fig. S2A to C). As expected, the parts of the proteins predicted with low confidence correspond to extracellular loops connecting the antiparallel  $\beta$ -strands (Fig. S2A to C). The three AlphaFold models could readily be superimposed (Fig. S2D), documenting that the proteins have the same fold overall. ExoB is important for SPS biosynthesis by an unknown mechanism (15), and WzpB is an OM protein (16), but it is not known whether it is important for BPS synthesis. Together, these data support that EpsX, ExoB, and WzpB are integral OM 18-stranded  $\beta$ -barrel proteins. To investigate the function of these proteins and their orthologs, we focused on EpsX.



**FIG 1** The 18-stranded  $\beta$ -barrel protein EpsX is an integral part of the EPS pathway. (A) AlphaFold models of EpsX, ExoB, and WzpB. Proteins are oriented based on the N and C termini of OM  $\beta$ -barrel proteins being periplasmic (66). Model ranks 1 are shown. (B) Phenotypic (Continued on next page)

**EpsX is important for EPS biosynthesis.** To investigate the function of EpsX, we generated an in-frame deletion in *epsX* ( $\Delta epsX$ ). Using plate-based colorimetric assays with Congo red or trypan blue as readouts of EPS synthesis, we observed that the wild type (WT) synthesized EPS. The  $\Delta epsX$  mutation, similarly to the  $\Delta epsZ$ ,  $\Delta epsV$ , and  $\Delta epsY$  mutations that inactivate genes for proteins in the EPS pathway (Fig. S1A and B) (17–19), caused strongly reduced EPS synthesis (Fig. 1B). EPS is important for type IV pilus (T4P)-dependent motility (14). Consistently, while the WT formed colonies with long flares at the edge characteristic of T4P-dependent motility and the  $\Delta pilA$  negative control formed smooth-edged colonies, the  $\Delta epsX$  mutant generated colonies with short flares, although these were slightly longer than those previously observed for the  $\Delta epsZ$ ,  $\Delta epsV$ , and  $\Delta epsY$  mutants (17–19) (Fig. 1B). The  $\Delta epsX$  mutant was recently described as having a minor defect in T4P-dependent motility and, in agreement with our observations, to be less impaired in T4P-dependent motility than a  $\Delta epsY$  mutant (20); we attribute the differences to different strain backgrounds and conditions of the motility assay. The  $\Delta epsX$  mutant, similarly to the WT and the  $\Delta epsZ$ ,  $\Delta epsV$ , and  $\Delta epsY$  mutants, displayed the single cells at the colony edge characteristic of gliding motility, while the  $\Delta gltB$  negative control did not (Fig. 1B). To test for polar effects of the  $\Delta epsX$  mutation (Fig. S1B), we performed complementation experiments in which *epsX* was expressed ectopically in the  $\Delta epsX$  mutant. Ectopic expression of *epsX*, as previously shown for *epsZ*, *epsV*, and *epsY* in the respective mutants (17), complemented the defects in EPS biosynthesis and T4P-dependent motility in the  $\Delta epsX$  mutant (Fig. 1B). We conclude that EpsX is important for EPS synthesis and likely an integral component of the EPS pathway.

In the *E. coli* and *Klebsiella pneumoniae* Wzx/Wzy-dependent pathways for capsule biosynthesis, the OM 18-stranded  $\beta$ -barrel protein Wzi is important for cell surface anchoring of the capsule but not for its biosynthesis or its export (21, 22), arguing that Wzi and EpsX have different functions. Wzi and EpsX share 16.8% sequence identity and 22.5% sequence similarity. Moreover, while the  $\beta$ -barrel in the solved structure of Wzi is circular (22) (Fig. S2E), the EpsX  $\beta$ -barrel in the AlphaFold model is elliptical (Fig. 1A). Wzi also contains an N-terminal  $\alpha$ -helical bundle that occludes the periplasmic side of the  $\beta$ -barrel and extracellular loops that fold into and occlude the  $\beta$ -barrel on the extracellular side (22). In contrast, EpsX lacks the N-terminal  $\alpha$ -helical bundle, and the extracellular loops, although modeled with low confidence (Fig. S2A), do not fold into the  $\beta$ -barrel, indicating that EpsX is open to the periplasm and could also, when required, open to the cell exterior. This structural comparison also supports the conclusion that EpsX has a function different from that of Wzi.

#### FIG 1 Legend (Continued)

characterization of  $\Delta epsX$  mutant. Two left columns, cells were placed on 0.5% agar supplemented with 0.5% CTT and Congo red or trypan blue and incubated for 24 h. Two right columns, T4P-dependent motility and gliding motility were tested on 0.5% and 1.5% agar, respectively, supplemented with 0.5% CTT, and images were recorded after 24 h. The  $\Delta pilA$  mutant, which lacks the major pilin of the T4P (67), and the  $\Delta gltB$  mutant, which lacks a component of the gliding motility machinery (23), served as negative controls for T4P-dependent and gliding motility, respectively. In the complementation strain, *epsX* was expressed from the *pilA* promoter on a plasmid integrated in a single copy at the Mx8 *attB* site. Scale bars, 1 mm (left) and 50  $\mu$ m (right). Numbers indicate expansion from the edge of the colony calculated from three biological replicates and relative to that of the WT, where 100% corresponds to 1.4 mm. (C) Comparison of AlgE and EpsX. Upper panel, lateral and top views of the solved structure of AlgE (PDB 3RBH) (11). Arrows indicate the external diameter of the  $\beta$ -barrel. Lower panel, superimposition of the solved structure of AlgE and the EpsX AlphaFold model. EpsX is colored in teal. EpsX aligns to AlgE with a root mean square deviation (RMSD) of 6.035 Å over 1,501 C $_{\alpha}$ . (D) EpsX and EpsY mutually stabilize each other, and EpsY stabilizes EpsV. Protein amounts in whole-cell proteomes of *M. xanthus* strains were determined using LFQ mass spectrometry-based proteomics (see Materials and Methods). Normalized log<sub>2</sub> intensities of Eps proteins in the indicated strains are shown. Missing bars indicate that the proteins were not detected. Data points represent three biological replicates. Error bars, standard deviation (SD) based in these replicates. \*,  $P < 0.05$  (Welch's test). WzeX is important for EPS synthesis and was proposed to act as the BY kinase partner of EpsV (14, 18). (E) RT-qPCR analysis of *epsV*, *epsY*, and *epsX* transcripts levels. Total RNA was isolated from cells grown as panel D. Data are shown as log<sub>2</sub> transcripts in a mutant relative to that of the WT. Individual data points represent four biological replicates with each two technical replicates and are colored according to the strain analyzed. Center marker and error bars represent mean and SD. \*,  $P < 0.05$  (Welch's test). (F) EpsX and EpsY interact. Pulldown experiments on whole-cell lysates of strains expressing EpsY-FLAG or sfGFP-FLAG (negative control). In the EpsY-FLAG strain, *epsY-FLAG* was expressed ectopically from the *pilA* promoter on a plasmid integrated in a single copy at the Mx8 *attB* site. In the sfGFP-FLAG strain, *epsY* was expressed from the *pilA* promoter on a plasmid integrated in a single copy at the Mx8 *attB* site, and *sfGFP-FLAG* was expressed from the *pilA* promoter on a plasmid integrated in a single copy at the 18-19 intergenic locus. Samples from four biological replicates were analyzed by LC-MS (see Material and Methods). Log<sub>2</sub>-fold enrichment of proteins in EpsY-FLAG over sfGFP-FLAG samples was calculated. Columns represent mean log<sub>2</sub>-fold enrichment ( $n = 4$ ). Error bars, SD based on these replicates; n.d., proteins detected in neither EpsY-FLAG nor sfGFP-FLAG samples.

Subsequently, by searching for structural homologs of EpsX using Foldseek (see Materials and Methods), we identified the OM translocon AlgE of the synthase-dependent pathway for alginate export in *P. aeruginosa* (11) as the closest structural homolog. The two proteins share 13.4% sequence identity and 20.1% sequence similarity. Similar to EpsX in the AlphaFold model, the 18-stranded  $\beta$ -barrel of AlgE in the solved structure has an elliptical shape (Fig. 1C). The two proteins could readily be superimposed except for extracellular loops in AlgE that fold into the  $\beta$ -barrel and partially occlude the pore (Fig. 1C).

**EpsX and EpsY mutually stabilize each other and interact.** There are several reported examples in *M. xanthus* of directly interacting OM, periplasmic, and IM proteins that stabilize each other (23–26). Consequently, to identify potential EpsX interaction partners, we performed whole-cell label-free quantitative (LFQ) mass spectrometry-based proteomics (see Materials and Methods) focusing on the Eps proteins.

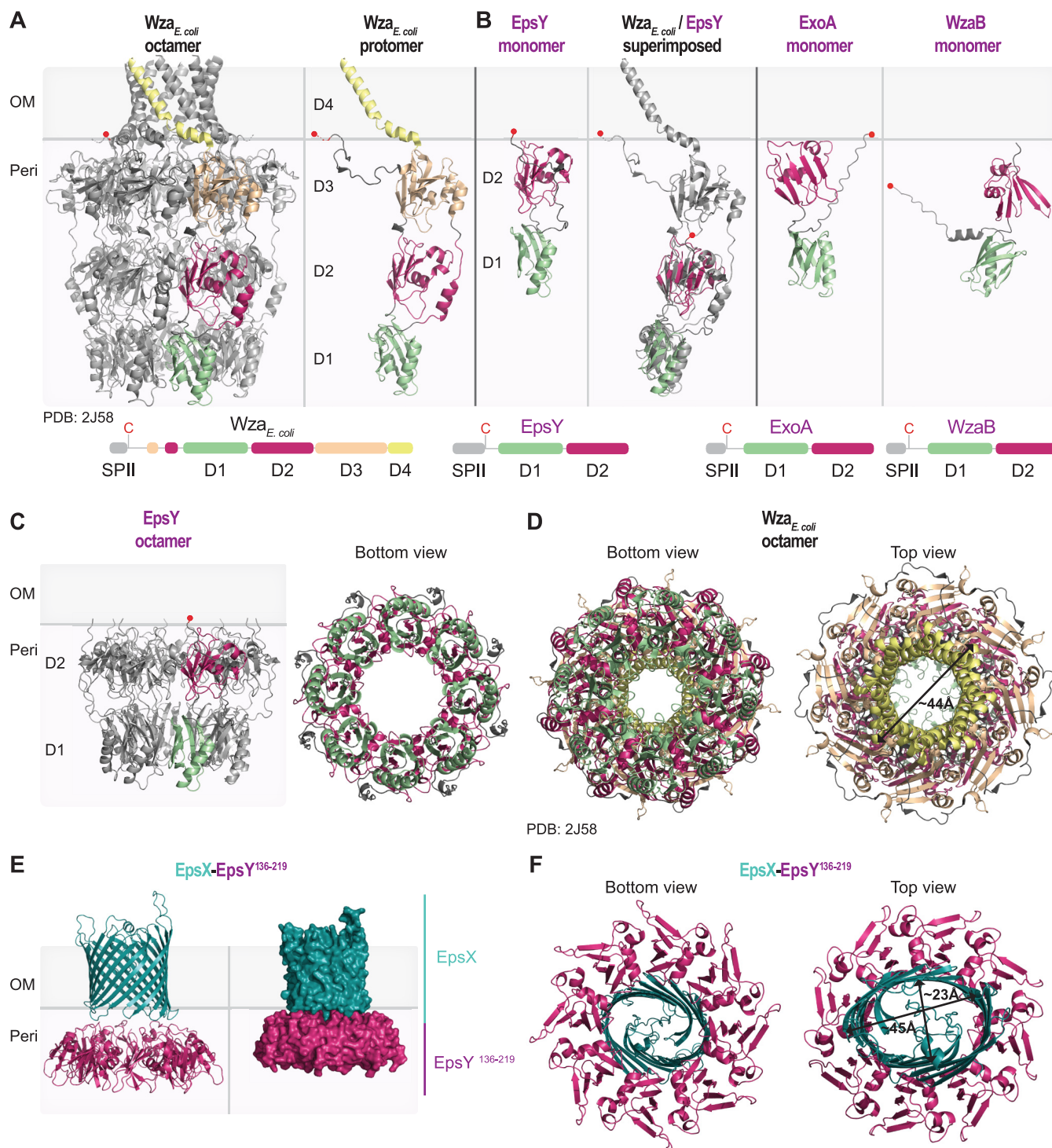
In the WT, we detected all Eps proteins except for the integral IM proteins Wzx<sub>EPS</sub> and Wzy<sub>EPS</sub> (Fig. 1D). In the  $\Delta$ epsX mutant, the OPX protein EpsY accumulated at a significantly reduced level (Fig. 1D). Conversely, in the  $\Delta$ epsY mutant, EpsX was not detected, and accumulation of the PCP EpsV was significantly reduced (Fig. 1D). In contrast, the  $\Delta$ epsZ mutant, which lacks the PGT for initiating EPS synthesis, had slightly decreased (EpsA), increased (Wzx<sub>EPS</sub>, EpsY, and EpsX), or WT levels of the Eps proteins (Fig. 1D). These observations support that the decreased accumulation of EpsY in the  $\Delta$ epsX mutant and of EpsX and EpsV in the  $\Delta$ epsY mutant is not caused by lack of EPS biosynthesis and export *per se*. Some GTs are also slightly but significantly reduced in the  $\Delta$ epsX and  $\Delta$ epsY mutants (Fig. 1D). Here, we focused on the significant accumulation dependency of EpsX, EpsY, and EpsV. Importantly, and in agreement with the previously reported complementation experiments of the  $\Delta$ epsX and  $\Delta$ epsY mutants (Fig. 1B) (17), reverse transcriptase quantitative PCR (RT-qPCR) provided evidence that the changes in EpsX, EpsY, and EpsV levels are independent of transcription in the  $\Delta$ epsX and  $\Delta$ epsY mutants (Fig. 1E). We conclude that the OM  $\beta$ -barrel protein EpsX and the OPX protein EpsY mutually stabilize each other and that EpsY stabilizes the PCP EpsV in the IM. These findings agree with OPX and PCP proteins interacting in the periplasm (8–10). They also support the idea that EpsX, EpsY, and EpsV form a complex that spans the cell envelope.

To address whether EpsX and EpsY interact, we performed pulldown mass spectrometry experiments (see Materials and Methods) using an  $\Delta$ epsY strain ectopically expressing an EpsY protein with a FLAG tag inserted at residue 43 between two periplasmic domains (herein termed EpsY-FLAG) (see details below). As a negative control, we used the  $\Delta$ epsY strain ectopically expressing EpsY without a FLAG tag and a superfolder green fluorescent protein (sfGFP)-FLAG protein. Both FLAG-tagged proteins accumulated, and EpsY-FLAG supported EPS synthesis and T4P-dependent motility, demonstrating that the FLAG tag did not disrupt EpsY's functional interactions with other proteins (Fig. S3A and B). Importantly, among proteins involved in EPS synthesis, only EpsX was highly enriched in the EpsY-FLAG pulldown experiments (Fig. 1F; Fig. S3C and D).

Together, the cooccurrence of EpsX and EpsY, the mutually stabilizing effect of EpsX and EpsY, and the pulldown experiments strongly support that EpsX and EpsY directly interact. The observations that EpsX and EpsY accumulate in the absence of EpsZ (Fig. 1D) and thus stabilize each other in the absence of EPS synthesis suggest that EpsX and EpsY interact independently of EPS synthesis and that this interaction is not transient. We note that the PCP EpsV was not enriched in the EpsY-FLAG pulldown experiments (Fig. 1F), while EpsV still accumulated in the absence of EpsZ (Fig. 1D). We speculate that EpsV was not detected in the pulldown experiments because it might not be efficiently extracted from the IM.

To investigate how EpsX, EpsY, and EpsV interact, we first analyzed EpsY and its homologs.

**Myxobacterial OPX proteins are short and comprise only two periplasmic domains.** In the Wza<sub>*E. coli*</sub> octamer, individual protomers have four structural domains (D1 to D4) (7) (Fig. 2A). The N-terminal domains D1 to D3 are periplasmic, with D1



**FIG 2** Structural characterization of the EpsY<sup>D1D2</sup>OPX protein alone and in complex with EpsX, its partner 18-stranded  $\beta$ -barrel protein. (A) Structure of Wza<sub>E. coli</sub>. Left panel, the solved structure of octameric Wza (PDB 2J58) (7). Right panel, an individual Wza protomer. The four domains of Wza are labeled D1 to D4. Light green, D1; dark pink, D2; light orange, D3; yellow, D4. The acylated N-terminal cysteine is indicated by a red circle and placed at the inner leaflet of the OM. Lower panel, domain organization of Wza. (B) AlphaFold model of EpsY. Left panel, lateral view of EpsY monomer as predicted by AlphaFold. Right panel, superimposition of a Wza protomer from the solved structure (gray) and the EpsY model. The EpsY monomer aligns to the Wza protomer with an RMSD of 3.306 Å over 904 C $\alpha$ . Right panels, AlphaFold models of ExoA and WzaB monomers. In all three AlphaFold models, the two domains are labeled D1 and D2 and colored according to the homologous domains in Wza. The acylated N-terminal cysteine is indicated by a red circle; note that the acylated N-terminal cysteine of WzaB is not modeled “on top” of D2, but the confidence in the relative position of this residue is low (Fig. S3C). Model rank 1 is shown for all structures. Lower panels, domain organization of EpsY, ExoA, and WzaB. SPII, type 2 signal peptide. (C) AlphaFold-Multimer model of octameric EpsY. Left panel, one protomer is colored as described in the legend for panel B. Right panel, bottom view of octameric EpsY with all protomers colored as described for panel B. Model rank 1 is shown. (D) Structure of Wza. All eight protomers are colored as described in the (Continued on next page)

containing the characteristic polysaccharide export sequence (PES) motif. D2 and D3 are structurally related and have a  $\beta$ -grasp fold (27, 28). The C-terminal D4 is an amphipathic  $\alpha$ -helix, which is inserted in the OM. In the octamer, the eight copies of D1 to D3 generate a periplasmic cavity, and the eight C-terminal  $\alpha$ -helices create the  $\alpha$ -helical barrel in the OM (Fig. 2A). Wza has a type 2 signal peptide and is N-terminally acylated (7), which is important for OM integration (27, 29).

Based on sequence analysis and a high-accuracy AlphaFold structural model, EpsY contains only D1 with the PES motif and D2 while lacking D3 and, most strikingly, D4 (Fig. 2B; Fig. S4A). The two EpsY domains could readily be superimposed on the corresponding domains in a Wza protomer (Fig. 2B). Using AlphaFold-Multimer (see Materials and Methods), EpsY with high accuracy could generate an octamer in which D1 and D2 form a stacked, ring-like structure with a central cavity, similar to the Wza counterparts (Fig. 2C; compare Fig. 2A and D; Fig. S5A). The Wza octamer is closed at the periplasmic base of D1 by a so-called tyrosine ring (Fig. 2D); however, this tyrosine ring is lacking in the EpsY octamer (Fig. 2C). As in the monomer model (Fig. 2B), the acylated N-terminal cysteine would be placed on top of D2 (Fig. 2C), facilitating the OM association of octameric EpsY. From here on, we refer to OPX proteins that comprise only D1 and D2 as  $D^{1D2}$ OPX proteins.

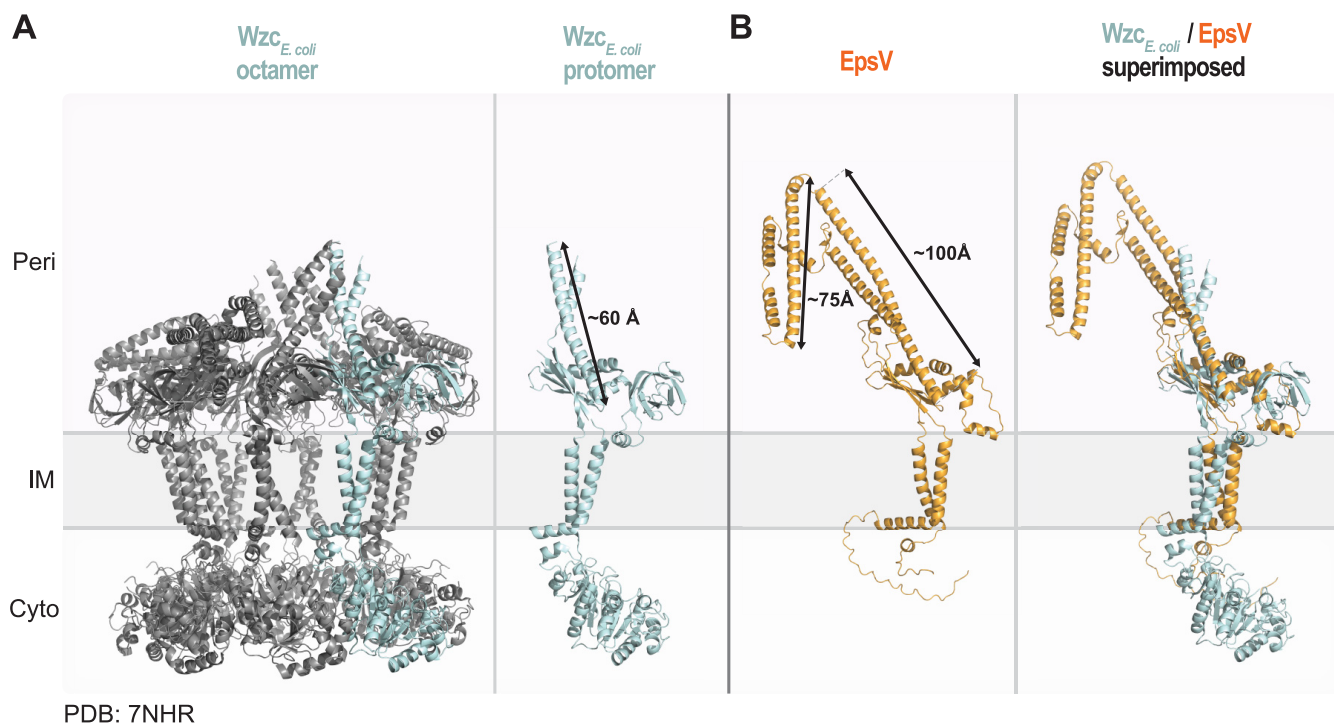
Remarkably, EpsY orthologs in other myxobacteria are also of the  $D^{1D2}$ OPX type (Fig. S6). In addition, the OPX proteins for SPS (ExoA) and BPS (WzaB) export, as well as their orthologs, are of the  $D^{1D2}$ OPX type (Fig. 2B; Fig. S4B and C and S6). Except for one, all these  $D^{1D2}$ OPX proteins have a type 2 signal peptide and are predicted lipoproteins (Fig. S6). Thus, all 48 identified OPX proteins in myxobacteria are  $D^{1D2}$ OPX proteins and, most strikingly, lack D4 for generating the OM  $\alpha$ -helical pore (Fig. S6). Consistent with these findings, a previous classification of OPX proteins (6) defined a subgroup consisting of short OPX proteins that included EpsY and ExoA.

**A  $D^{1D2}$ OPX protein and an 18-stranded  $\beta$ -barrel protein may form a translocon for OM polysaccharide export.**  $D^{1D2}$ OPX proteins lack the domain for spanning the OM and, generally, cooccur with 18-stranded  $\beta$ -barrel proteins (Fig. S1A and S6). Moreover, the 18-stranded  $\beta$ -barrel proteins are structurally similar to the OM translocon AlgE (Fig. 1C), indicating that they can support polysaccharide export across the OM. The  $D^{1D2}$ OPX protein EpsY and the 18-stranded  $\beta$ -barrel protein EpsX mutually stabilize each other and interact in *in vivo* pulldown experiments, supporting that they interact directly. Based on these four lines of evidence, we hypothesized that a  $D^{1D2}$ OPX protein functions together with a  $\beta$ -barrel protein to create a composite  $D^{1D2}$ OPX/ $\beta$ -barrel translocon.

To test this hypothesis, we intended to generate a heterocomplex consisting of eight EpsY molecules and one EpsX molecule using AlphaFold-Multimer. However, this amount of sequence information is computationally highly demanding to analyze. Instead, we generated a heterocomplex consisting of EpsY D2 (amino acids [aa] 136 to 219), which is predicted to be close to the OM (Fig. 2B and C), and full-length EpsX. In a high-accuracy AlphaFold-Multimer structural model, EpsX is placed on top of D2 in the EpsY<sup>136-219</sup> octamer (Fig. 2E; Fig. S5B) in an arrangement strikingly similar to that of the  $\alpha$ -helical barrel “on top” of D3 in the Wza octamer (Fig. 2A). Also, the periplasmic region of EpsX and the OM-facing regions of D2 in the EpsY<sup>136-219</sup> octamer have opposite surface charges (Fig. S5C). Moreover, the diameters of the  $\alpha$ -helical barrel in Wza and the EpsX  $\beta$ -barrel in the EpsX/EpsY<sup>136-219</sup> heterocomplex are similar (Fig. 2D and F). To test the specificity of the AlphaFold-Multimer prediction of the EpsX/EpsY<sup>136-219</sup> complex, we attempted to generate a model of an AlgE/EpsY<sup>136-219</sup> heterocomplex. Even in the best of the five predicted models in which the individual AlgE and EpsY<sup>136-219</sup> proteins are calculated to high accuracy based on predicted local distance difference test (pLDDT) values (see Material and Methods), the AlgE  $\beta$ -barrel did not associate with the octameric ring structure of

## FIG 2 Legend (Continued)

legend for panel A. An arrow indicates the external diameter of the  $\alpha$ -helical pore. In the bottom view, the tyrosine residues that form the so-called tyrosine ring are present in the loops extending into the central channel. (E and F) AlphaFold-Multimer model of a heterocomplex of octameric EpsY<sup>136-219</sup> and an EpsX monomer. In the heterocomplex, the EpsY<sup>136-219</sup> octamer is colored as D2 as described for panel B, and EpsX is colored teal. (E) The right panel is a surface-rendered representation. Model rank 1 is shown. (F) Arrows indicate the diameter of the  $\beta$ -barrel. Model rank 1 is shown.



**FIG 3** Structural characterization of the PCP EpsV. (A) Solved structure of octameric *Wzc*<sub>*E. coli*</sub> (PDB 7NHR) (30). Left panel, the protein is colored in gray, and one protomer is colored in light blue. Right panel, individual *Wzc* protomer in class 2 conformation (30). Note that individual protomers have different conformations in the octamer. An arrow indicates the length of the extended  $\alpha$ -helical stretch. (B) AlphaFold model of EpsV. Arrows indicate the lengths of the  $\alpha$ -helical stretches. Model rank 1 is shown. Right panel, superimposition of a protomer from the solved structure of *Wzc* and the EpsV model. EpsV aligns to *Wzc* with an RMSD of 4.610 Å over 727 C $\alpha$ .

EpsY<sup>136-219</sup>, and the accuracy in the relative positioning of AlgE to octameric EpsY<sup>136-219</sup> is low, as evidenced by the high predicted alignment error (pAE) values (Fig. S5D). Thus, despite EpsX and AlgE being close structural homologs, the low accuracy of the AlgE/EpsY<sup>136-219</sup> model suggests high specificity in the EpsX/EpsY<sup>136-219</sup> heterocomplex, as expected for a functionally relevant protein complex.

Together, this computational approach, the mutually dependent stability of EpsX/EpsY, and the observation that EpsX and EpsY interact in *in vivo* pulldown experiments support that these two proteins form a complex. In this complex, EpsX spans the OM and EpsY is periplasmic and associated with the OM via the N-terminal lipid group and the interaction with EpsX.

**PCPs in *M. xanthus* may have an extended periplasmic domain.** The IM PCP proteins interact with their cognate OPX proteins in the periplasm by using a region rich in  $\alpha$ -helices (8–10, 30). In the case of the *Wzc*-*Wza*<sub>*E. coli*</sub> complex, the PCP *Wzc*<sub>*E. coli*</sub> in the solved structure extends  $\sim 60$  Å from the IM into the periplasm (30) (Fig. 3A), while the periplasmic part of *Wza* extends  $\sim 100$  Å from the OM into the periplasm, with each ring contributing  $\sim 30$  Å (7). This gives an estimated periplasmic height of a *Wzc*-*Wza* complex of  $\sim 160$  Å.

EpsY is important for the stability of the PCP EpsV (Fig. 1D), supporting that the two proteins interact. This raises the question of how the EpsY<sup>D1D2</sup>OPX protein would be able to bridge the periplasm together with its EpsV PCP partner. A high-accuracy AlphaFold model of monomeric EpsV supports that it has two transmembrane  $\alpha$ -helices and a periplasmic domain rich in  $\alpha$ -helices (Fig. 3B; Fig. S7A), similar to other PCPs (6). The periplasmic part of EpsV is composed of two  $\alpha$ -helical stretches with lengths of  $\sim 100$  Å and  $\sim 75$  Å, respectively, and connected by linkers. Depending on the conformation of the  $\alpha$ -helical stretches in EpsV, the EpsY/EpsV complex could, thus, have a height of  $\sim 160$  to 240 Å across the periplasm, supporting that they can jointly span the periplasm.

Similarly, monomeric ExoC of the SPS pathway and WzcB of the BPS pathway have long  $\alpha$ -helical periplasmic regions in high-accuracy AlphaFold models (Fig. S7B and C).



Together, these *in silico* analyses support the idea that <sup>D1D2</sup>OPX proteins function together with a PCP with an extended periplasmic part.

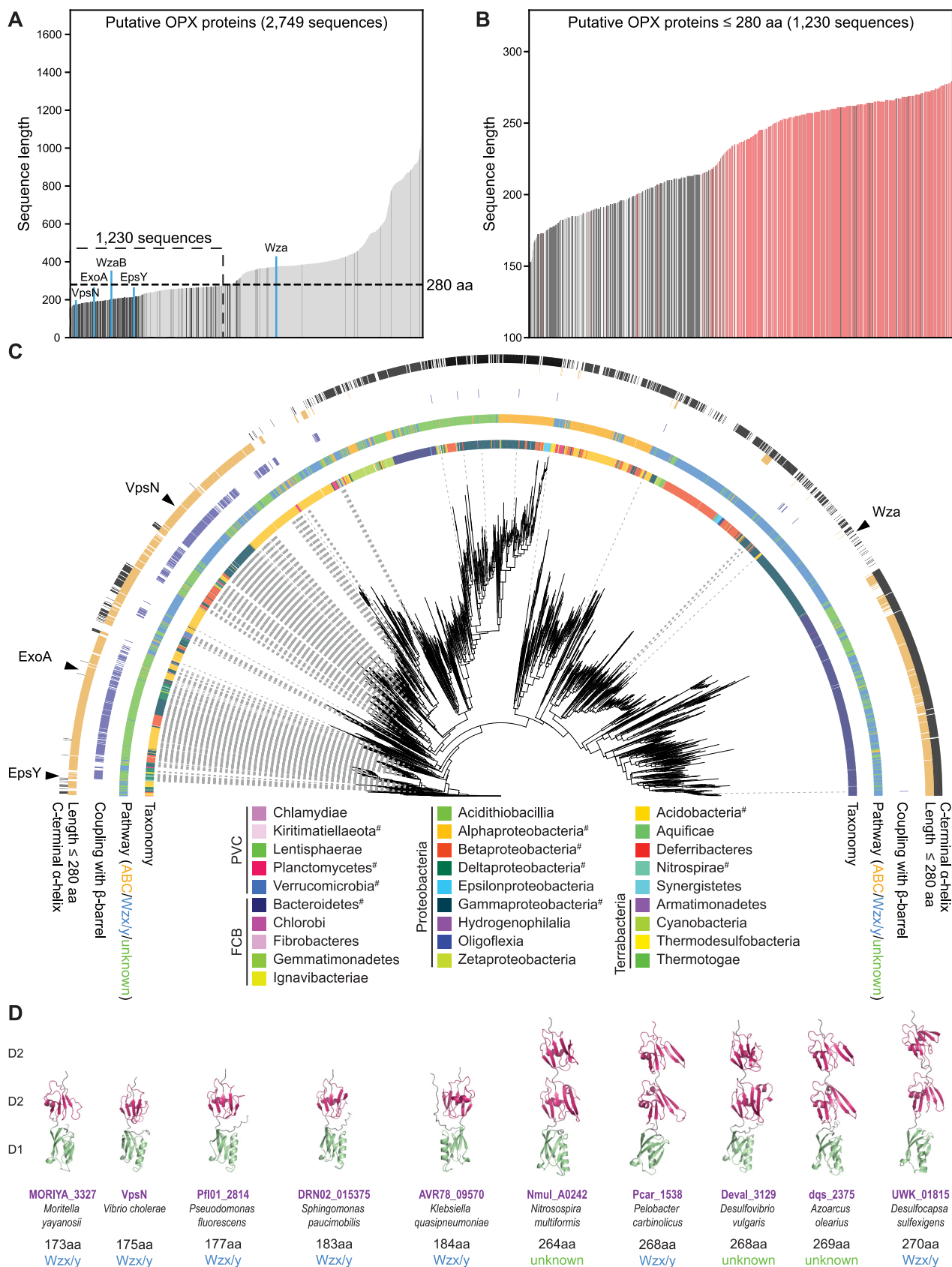
**Coupled genes for short periplasmic OPX proteins and OM  $\beta$ -barrel proteins are widespread in Gram-negative bacteria.** Forty-one of the 48 myxobacterial genes for <sup>D1D2</sup>OPX proteins are syntenic with a gene encoding an 18-stranded  $\beta$ -barrel protein (Fig. S1A and S6). We took advantage of this observation to assess bioinformatically how widespread short, periplasmic OPX proteins are and whether they are coupled with an OM  $\beta$ -barrel protein. Specifically, we identified OPX candidates in 6,607 fully sequenced prokaryotic genomes across 49 phyla using the PES motif (Pfam PF02563) as described previously (6). To substantiate that the identified proteins are part of a polysaccharide biosynthesis pathway, we included only OPX candidate genes with a polysaccharide biosynthesis gene within five genes upstream or downstream (see Materials and Methods). In total, we identified 4,257 OPX proteins and subsequently used 2,749 representative proteins to determine whether they have a C-terminal  $\alpha$ -helix and/or are coupled to an OM  $\beta$ -barrel protein (see Materials and Methods). To determine the coupling of OPX proteins and OM  $\beta$ -barrel proteins, we searched for genes encoding OM  $\beta$ -barrel proteins within five genes of a gene encoding an OPX protein (see Materials and Methods). After verifying the  $\beta$ -barrel fold and OM localization computationally (see Materials and Methods), we identified 486 such OM  $\beta$ -barrel proteins.

The 2,749 OPX proteins range in length from 138 to 1,728 aa (Fig. 4A). Interestingly, the length distribution of the 486 OPX proteins with a coupled  $\beta$ -barrel protein is highly skewed toward shorter OPX proteins (Fig. 4A and B). Based on the size distribution of OPX proteins with a coupled  $\beta$ -barrel protein, and because EpsY, ExoA, and WzaB are 217 aa, 190 aa, and 204 aa, respectively, we empirically defined a size cutoff for short OPX proteins of 280 aa (Fig. 4A and B). We built a phylogenetic tree of all 2,749 OPX proteins based on their PES motif and found these proteins in 28 phyla (Fig. 4C). We then located the branches of the tree that correspond to OPX proteins with a size of  $\leq 280$  aa and found that those with a coupled  $\beta$ -barrel protein are present in 10 phyla (Fig. 4C).

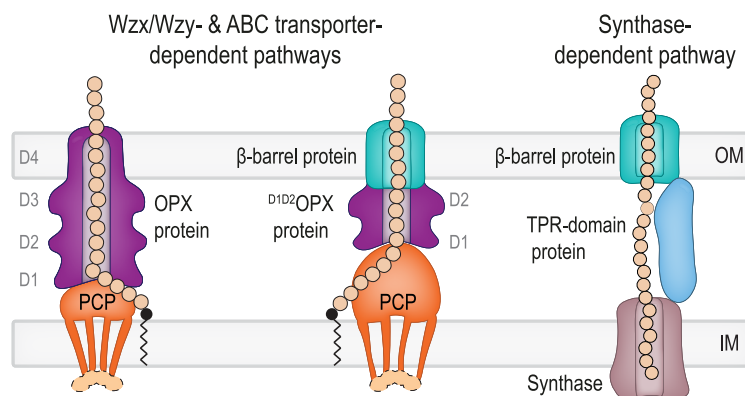
A total of 1,230 OPX proteins are  $\leq 280$  aa, and among these, 453 proteins have a coupled  $\beta$ -barrel protein (Fig. 4B and C). Conversely, 93% of the  $\beta$ -barrel proteins are coupled to an OPX protein of  $\leq 280$  aa. Interestingly, among the 777 OPX proteins of  $\leq 280$  aa and without a coupled  $\beta$ -barrel protein, 535 have a C-terminal  $\alpha$ -helix (Fig. 4B and C) and are found mostly in the *Bacteroidetes* (Fig. 4C). Thus, generally, we observe a dichotomy among the 1,230 OPX proteins of  $\leq 280$  aa. Typically, these OPX proteins have either a coupled  $\beta$ -barrel protein and no C-terminal  $\alpha$ -helix (453 proteins) or no coupled  $\beta$ -barrel protein but a C-terminal  $\alpha$ -helix (535 proteins). For the latter group of OPX proteins, we suggest that they likely function similarly to Wza.

Focusing on the 453 OPX proteins of  $\leq 280$  aa with a coupled  $\beta$ -barrel protein and no C-terminal  $\alpha$ -helix, we found, based on genomic context, that 185 are part of a Wzx/Wzy-dependent pathway, 14 are part of an ABC transporter-dependent pathway, and 287 were not assigned to a pathway because the closest polysaccharide biosynthesis gene encodes a PCP and some of these proteins can be difficult to assign to a specific type of pathway (6). Among the short OPX/ $\beta$ -barrel couples identified, we found VpsN/VpsM of *Vibrio cholerae* (Fig. 4A, C, and D). These two proteins are important for *Vibrio* polysaccharide (VPS) synthesis, biofilm formation, and intestinal colonization in a mouse model (31), thus validating the outlined computational approach. Analysis of the domain architecture of the short OPX proteins is difficult because, except for the PES motif in the D1 domain, their sequences are not well conserved. Therefore, to analyze the domain architecture of the 453 short OPX proteins with a coupled  $\beta$ -barrel protein, we analyzed the domain structure of five proteins of  $\leq 200$  aa and five proteins of  $> 250$  aa. Intriguingly, the five proteins with a length of  $< 200$  aa all had the <sup>D1D2</sup>OPX architecture, while the proteins of  $> 250$  aa had a <sup>D1D2D2</sup>OPX architecture (Fig. 4D; Fig. S8).

We conclude that the coupling between <sup>D1D2</sup>OPX proteins and OM  $\beta$ -barrel proteins is conserved and widespread in Gram-negative bacteria and that this coupling can be extended to include the <sup>D1D2D2</sup>OPX variants.



**FIG 4** Computational genomics of OPX proteins. (A) Length distribution of 2,749 OPX proteins. OPX proteins encoded within a distance of five genes of a gene encoding a  $\beta$ -barrel protein are shown in dark gray, and the remaining OPX proteins are in light gray. OPX proteins discussed in (Continued on next page)



**FIG 5** Three different mechanisms for polysaccharide export across the OM. Left schematic, in classical Wzx/Wzy- and ABC transporter-dependent pathways, polysaccharide transfer across the periplasm and OM is mediated by a complex composed of a PCP and an OPX protein with an  $\alpha$ -helical barrel in the OM. Middle schematic, in pathways in which the OPX protein lacks domain D4 ( $D^{1D2}$ OPX or  $D^{1D2D2}$ OPX proteins), a  $\beta$ -barrel protein constitutes the OM part of the composite OPX/ $\beta$ -barrel protein translocon. The D1 to D4 and D1 to D2 rings of the OPX proteins are indicated. In Wzx/Wzy-dependent pathways, the PCP can be associated with a cytoplasmic BY kinase (light orange). Synthesis, polymerization, and translocation of the polysaccharide across the IM are not shown. Right schematic, in synthase-dependent pathways, transfer across the periplasm and the OM depends on a TPR domain-containing protein and a  $\beta$ -barrel protein translocon in the OM.

**Conclusions.** Here, using experimental approaches and computational analyses, we provide evidence that polysaccharides can be exported across the OM via a composite periplasmic OPX protein/OM  $\beta$ -barrel protein translocon (Fig. 5). The domain architecture of the periplasmic OPX proteins can be  $D^{1D2}$ OPX or  $D^{1D2D2}$ OPX, giving rise to  $D^{1D2}$ OPX/ $\beta$ -barrel translocons and  $D^{1D2D2}$ OPX/ $\beta$ -barrel translocons. The  $D^{1D2}$ OPX/ $\beta$ -barrel translocons likely function with a PCP with extra-long periplasmic  $\alpha$ -helical regions to span the periplasm and compensate for the short OPX protein (Fig. 5). These systems are widespread in Gram-negative bacteria and add to the two established mechanisms for OM export, i.e., the Wza-like OPX protein in Wzx/Wzy- and ABC transporter-dependent pathways and the OM 16- to 18-stranded  $\beta$ -barrel protein in synthase-dependent pathways (Fig. 5). This work provides a framework for future biochemical and structural studies of EpsX and EpsY as well as of other polysaccharide biosynthesis pathways containing a  $D^{1D2}$ OPX or  $D^{1D2D2}$ OPX protein together with a coupled OM  $\beta$ -barrel protein.

The existence of three systems for polysaccharide export across the OM poses the question of their evolutionary origin. Our current analysis does not suggest an evolutionary scenario for their emergence, and this will likely depend on detailed studies of the taxonomic distribution of the three systems.

During the preparation of this article, Saïdi et al. (20) independently identified EpsX, ExoB, and WzpB as cooccurring with EpsY, ExoA, and WzaB and suggested that these proteins could make up a translocon for export of polysaccharides across the OM, thus independently supporting our findings.

## MATERIALS AND METHODS

**Strains and cell growth.** All *M. xanthus* strains are derivatives of the WT strain DK1622 (32) and are listed in Table 1. The plasmids and oligonucleotides used are listed in Table 2 and in Table S1 in the supplemental material. In-frame deletions were generated as described previously (33). The plasmids for

### FIG 4 Legend (Continued)

the text are highlighted in blue. The 1,230 short OPX proteins are indicated based on the upper size limit of  $\leq 280$  aa. (B) Length distribution of OPX proteins of  $\leq 280$  aa. Four hundred fifty-three OPX proteins with a coupled  $\beta$ -barrel protein and no C-terminal  $\alpha$ -helix are in dark gray; 535 OPX proteins with no coupled  $\beta$ -barrel protein but a C-terminal  $\alpha$ -helix are in red. Proteins not matching these criteria are in light gray. (C) Maximum-likelihood phylogenetic tree built from the PES motif of the 2,749 OPX proteins. The inner ring is colored based on NCBI taxonomy (<https://www.ncbi.nlm.nih.gov/taxonomy/>). PVC, FCB, and Terrabacteria are superphyla, and the phylum *Proteobacteria* is divided into classes. #, this phylum contains an OPX protein(s) of  $\leq 280$  aa coupled to a  $\beta$ -barrel protein(s). The second ring indicates the pathway assigned to individual OPX proteins. The third ring indicates, in dark blue, coupling with a  $\beta$ -barrel protein. The fourth ring indicates, in orange, OPX proteins of  $\leq 280$  aa. The fifth ring indicates, in black, whether an OPX protein has a C-terminal  $\alpha$ -helix. (D) AlphaFold models of indicated OPX proteins. Domains are labeled D1 and D2 and colored according to the homologous domains in Wza (Fig. 2A). Model rank 1 is shown for all structures.

**TABLE 1** Strains used in this work

Species and strain	Genotype	Reference or source
<i>M. xanthus</i>		
DK1622	WT	32
DK10410	$\Delta pilA$	67
SA3922	$\Delta gltB$	23
SA7400	$\Delta epsZ$	17
SA7406	$\Delta epsV$	17
SA7408	$\Delta epsY$	17
SA7412	$\Delta epsY attB::pMP030$ ( $P_{pilA} epsY$ )	17
SA11550	$\Delta epsX$	This study
SA11554	$\Delta epsX attB::JSc007$ ( $P_{pilA} epsX$ )	This study
SA11592	$\Delta epsY attB::pJSc053$ ( $P_{pilA} epsY-FLAG$ )	This study
SA11597	$\Delta epsY attB::pMP030$ ( $P_{pilA} epsY$ ) 18-19::pJSc056 ( $P_{pilA} sfGFP-FLAG$ )	This study
<i>E. coli</i> Mach1	$\Delta recA1398 endA1 tonA \Phi 80 \Delta lacM15 \Delta lacX74 hsdR(r_K^- m_K^+)$	Invitrogen

complementation experiments were integrated in a single copy by site-specific recombination into the Mx8 *attB* site and the plasmid for ectopic expression of *sfGFP-FLAG* into the 18-19 site. All plasmids were verified by DNA sequencing, and all strains were verified by PCR. *M. xanthus* was grown at 32°C in 1% CTT broth (1% [wt/vol] Bacto casitone, 10 mM Tris-HCl [pH 8.0], 1 mM  $K_2HPO_4/KH_2PO_4$  [pH 7.6], 8 mM  $MgSO_4$ ) or on 1.5% agar supplemented with 1% CTT and kanamycin (50  $\mu$ g mL<sup>-1</sup>) or oxytetracycline (10  $\mu$ g mL<sup>-1</sup>) when required (34). Plasmids were propagated in *E. coli* Mach1 at 37°C in lysogeny broth (LB) (35) supplemented with kanamycin (50  $\mu$ g mL<sup>-1</sup>) or tetracycline (20  $\mu$ g mL<sup>-1</sup>).

**Plasmid construction.** For pJSc004 (for generation of in-frame deletion of *epsX*), up- and down-stream fragments were amplified from genomic DNA of DK1622 using the primer pairs 7418-A/7418-B and 7418-C/7418-D, respectively. Subsequently, the AB and CD fragments were used as templates for an overlapping PCR with the primer pair 7418-A/7418-D to generate the AD fragment. The AD fragment was digested with KpnI/XbaI and cloned in pBJ114. For pJSc007 (for generation of a strain ectopically expressing *epsX* from the *pilA* promoter), *epsX* was amplified from genomic DNA of DK1622 using 7418-P*pilA*-for and 7418-P*nat*/P*pilA*-rev. Subsequently, the PCR fragment was digested with XbaI/HindIII and cloned into pSW105. For pJSc053 (for generation of a strain ectopically expressing *epsY-FLAG* [FLAG tag fused after aa 43 of EpsY] from the *pilA* promoter), fragment 1 was amplified from genomic DNA using *epsY-PpilA*-for and *epsY-Flag43* rev. Fragment 2 was amplified from genomic DNA using *epsY-PpilA* rev and *epsY-Flag43*-for. Fragments 1 and 2 were used as templates for an overlapping PCR with the primer pair *epsY-PpilA*-for and *epsY-PpilA*-rev. Subsequently, the PCR fragment was digested with XbaI/HindIII and cloned into pSW105. For pJSc056 (for generation of a strain ectopically expressing *sfGFP-FLAG* from the *pilA* promoter, for integration into the 18-19 site), a fragment containing *sfGFP-FLAG* under the control of the *pilA* promoter was obtained from pMAT219 (25) using EcoRI/HindIII restriction sites and cloned into pDS75 (36) (a derivative of pSWU30, where the Mx8 *attB* locus was exchanged with the *MXAN\_18-19* intergenic region) after digesting it with EcoRI/HindIII to remove the insert.

**Detection of EPS.** Colony-based EPS assays were performed as described previously (17). Briefly, exponentially growing cells were harvested (3 min, 6,000 × *g* at room temperature [RT]) and resuspended in 1% CTT to a calculated density of 7 × 10<sup>9</sup> cells mL<sup>-1</sup>, and 20- $\mu$ L aliquots were placed on 0.5% agar plates supplemented with 0.5% CTT and 10 or 20  $\mu$ g mL<sup>-1</sup> of trypan blue or Congo red, respectively. Plates were incubated at 32°C and imaged at 24 h.

**Motility assays.** Motility assays were performed as described previously (37). Briefly, exponentially growing cells were harvested (3 min, 6,000 × *g*, RT) and resuspended in 1% CTT to a density of 7 × 10<sup>9</sup> cells mL<sup>-1</sup>. Five-microliter volumes of cell suspensions were placed on 0.5% and 1.5% agar (Invitrogen) supplemented with 0.5% CTT and incubated at 32°C for 24 h. Cells were imaged using a M205FA stereomicroscope (Leica) and a DMi8 inverted microscope (Leica) equipped with a Hamamatsu ORCA-Flash V2 digital CMOS camera (Hamamatsu Photonics) and DFC9000 GT camera (Leica), respectively.

**Immunoblots.** Immunoblotting was performed as described previously (35). For sample preparation, *M. xanthus* cells were harvested from exponentially growing suspension cultures and resuspended in

**TABLE 2** Plasmids used in this work

Plasmid	Description	Reference
pBJ114	Km <sup>r</sup> <i>galk</i>	68
pSW105	Km <sup>r</sup> $P_{pilA}$	69
pDS75	Tet <sup>r</sup> $P_{pilA}$	36
pJSc004	pBJ114, in-frame deletion construct for <i>epsX</i> ( <i>MXAN_7418</i> ), Km <sup>r</sup>	This study
pJSc007	pSW105, complementation construct for <i>epsX</i> ( <i>MXAN_7418</i> ) expressed from the <i>pilA</i> promoter, Km <sup>r</sup>	This study
pJSc053	pSW105, construct for <i>epsY-FLAG</i> expression from the <i>pilA</i> promoter, Km <sup>r</sup>	This study
pJSc056	pSWU30 for 18-19 site, construct for <i>sfGFP-FLAG</i> expression from the <i>pilA</i> promoter, Tet <sup>r</sup>	This study

SDS lysis buffer. Proteins were loaded from an equal number of cells per sample. As primary antibodies, rabbit polyclonal anti-FLAG (1:2,000; Rockland) and anti-PilC (1:2,000) (38) antibodies were used, with horseradish peroxidase (HRP)-conjugated anti-rabbit immunoglobulin G (1:15,000, Sigma) as the secondary antibody. Immunoblots were developed using Immobilon Forte Western HRP substrate (Millipore) on a LAS-4000 imager (Fujifilm).

**RT-qPCR.** Total RNA from *M. xanthus* cells grown on 1.5% agar supplemented with 1% CTT was extracted using the Monarch total RNA miniprep kit (New England Biolabs [NEB]). Briefly,  $10^9$  cells were scraped off the agar plates, resuspended in 200  $\mu$ L lysis buffer (100 mM Tris-HCl [pH 7.6], 1 mg mL<sup>-1</sup> lysozyme), and incubated at 25°C for 5 min. RNA purification was performed according to the manufacturer's protocol. DNA was removed using Turbo DNase (Thermo Fisher Scientific), and DNase was removed by using the Monarch RNA cleanup kit (50  $\mu$ g; NEB). A LunaScript RT supermix kit (NEB) was used to generate cDNA using 1  $\mu$ g RNA. qPCRs were performed on an Applied Biosystems 7500 real-time PCR system using the Luna universal qPCR mastermix (NEB) with the primers listed in Table S1. Data analysis was performed using the comparative threshold cycle ( $C_T$ ) method (39). *MXAN\_3298*, encoding the elongation factor Tu, served as an internal reference gene as described previously (26).

**Proteomic analysis using DIA-MS.** The total proteome of *M. xanthus* cells grown on 1% CTT–1.5% agar plates was determined as described previously (25). Peptide mixtures were then analyzed using liquid chromatography-mass spectrometry (LC-MS) on an Exploris 480 instrument connected to an Ultimate 3000 RSLCnano system with a ProFlow upgrade and a Nanospray Flex ion source (all Thermo Scientific). Peptide separation was performed on a reverse-phase high-performance liquid chromatography (HPLC) column (75  $\mu$ m by 42 cm) packed in-house with C<sub>18</sub> resin (2.4  $\mu$ m; Dr. Maisch). The following separating gradient was used: 94% solvent A (0.15% formic acid) and 6% solvent B (99.85% acetonitrile, 0.15% formic acid) to 25% solvent B over 95 min and to 35% B for additional 25 min at a flow rate of 300 nL min<sup>-1</sup>. The data-independent acquisition-mass spectrometry (DIA-MS) acquisition method was adapted from reference 40. Briefly, the spray voltage was set to 2.0 kV, the funnel radio frequency level was set at 55, and the heated capillary temperature was set at 275°C. For DIA experiments, full MS resolutions were set to 120,000 at  $m/z$  200 with an automatic gain control (AGC) target of 300% and a maximum injection time (IT) of 50 ms. The mass range was set at 350 to 1,400. The AGC target value for fragment spectra was set at 3000%. Forty-nine windows of 15 Da were used with an overlap of 1 Da. Resolution was set to 15,000, and IT was set to 22 ms. A stepped high-energy collisional dissociation (HCD) collision energy of 25, 27.5, and 30% was used. MS1 data were acquired in profile, and MS2 DIA data were acquired in centroid mode.

Analysis of DIA data was performed using DIA-NN version 1.8 (41) with a UniProt protein database for *M. xanthus*. Full tryptic digestion was allowed with two missed cleavage sites and with oxidized methionines (variable) and carbamidomethylated cysteines (fixed). “Match between runs” and “remove likely interferences” options were enabled. The neural network classifier was set to the single-pass mode, and protein inference was based on “genes.” The quantification strategy was set to any LC (high accuracy). Cross-run normalization was set to “RT-dependent.” Library generation was set to “smart profiling.” The DIA-NN “report” output was used to sum the unique peptide intensities, and identified proteins were filtered out if  $q$  values (false discovery rates) exceeded 0.01. Protein intensities were normalized with the cyclic loess method using the R package NormalizerDE (42).

Pulldown experiments were performed as described previously (25). Briefly, *M. xanthus* cells grown on 1% CTT–1.5% agar plates were collected (80 mg per replicate) and washed in 1 mL phosphate-buffered saline (PBS) (137 mM NaCl, 2.7 mM KCl, 10 mM Na<sub>2</sub>HPO<sub>4</sub>, 1.8 mM KH<sub>2</sub>PO<sub>4</sub>, pH 7.5). Cells were resuspended in 150  $\mu$ L detergent lysis solution consisting of 8% CHAPS-[3-(3-cholamidopropyl)dimethylammonio]-1-propanesulfonate hydrate (wt/vol), 8% 3-(*N,N*-dimethylmyristylammonio)propanesulfonate (wt/vol), 8% sodium lauroyl sarcosinate (SLS) (wt/vol), 0.1% Nonidet P-40 (NP-40) (vol/vol), 40% glycerol (vol/vol), 150  $\mu$ L PBS, and 33  $\mu$ L protease inhibitor concentrate (1 tablet protease inhibitor dissolved in 1 mL PBS; Roche). Lysis was performed for 10 min at RT. The total volume was increased to 5 mL with PBS, 5  $\mu$ L Benzonase (Merck) was added, and incubation proceeded under rotation at 4°C for 60 min. Cell debris was removed by centrifugation (15 min, 10,000  $\times g$  at 4°C). Fifteen microliters of equilibrated magnetic anti-FLAG M2 beads (Sigma) was added to the cleared supernatants and incubated with rotation at 4°C for 60 min. Beads were collected using a magnetic separator and washed twice with 700  $\mu$ L PBS-detergent wash buffer (detergent lysis solution diluted 1:30 in PBS) and five times with 100 mM ammonium bicarbonate (Sigma-Aldrich) to remove detergents. Protein elution from beads was carried out at 90°C for 10 min in the presence of 2% SLS (wt/vol) in 100 mM ammonium bicarbonate. The eluate was reduced and alkylated and subsequently acetone precipitated using an 8 $\times$  volume excess of ice-cold acetone, followed by incubation at –20°C overnight. After centrifugation (10 min, 10,000  $\times g$ ), the protein pellet was washed twice with 200  $\mu$ L cold methanol and air dried. Proteins were reconstituted in 100 mM ammonium bicarbonate and digested in the presence of 1  $\mu$ g trypsin (Serva). Postdigestion, tryptic peptides were desalted using C<sub>18</sub> solid-phase extraction (see reference 25) and analyzed via LC-MS using data-independent acquisition (see above) with the exception that the separating gradient length was reduced to 30 min and DIA raw data were analyzed with DIA-NN with the above-described setting and further evaluated in SafeQuant (43) using median protein intensities for normalization, background value imputation from normal distribution (width, 0.3; down shift, 1.8), and moderate eBayes  $t$ -test statistics as a significance strategy (44). For calculation of enrichment factors in samples over that of the control, identified proteins were filtered out if the peptide count in at least one sample replicate fell below 3. Proteins with an absolute abundance difference of  $\geq 8$  ( $\log_2$ -fold enrichment of  $\geq 3$ ) in samples over that of the control and a  $P$  value of  $\leq 0.001$  ( $-\log_{10} P$  value  $\geq 3$ ) were considered enriched.

**Bioinformatics.** The phylogenetic tree of myxobacteria was prepared as described previously (45) in MEGA-X (46) using the neighbor-joining method (47) and the genome sequences listed in Table S2.

To illustrate the operon structure of the *eps* loci, transcriptome sequencing (RNA-seq) and cappable-seq data from reference 48 were used.

Full-length protein sequences or sequences in which the signal peptide was identified with SignalP 6.0 (49) and removed were used for AlphaFold and AlphaFold-Multimer modeling via ColabFold (50–52) using the Alphafold2\_mmseqs2 notebook with default settings, except that recycles were set to six. Predicted local distance difference test (pLDDT) and predicted alignment error (pAE) graphs of the five models generated by the Alphafold2\_mmseqs2 notebook were made using a custom-made Matlab R2020a (The MathWorks) script. Ranking of the models was performed based on combined pLDDT and pAE values, with the best-ranked models used for further analysis and presentation. Per-residue model accuracy was estimated based on pLDDT values (>90, high accuracy; 70 to 90, generally good accuracy; 50 to 70, low accuracy; <50, should not be interpreted) (50). Relative domain positions were validated by pAE. The pAE graphs indicate the expected position error at residue X if the predicted and true structures were aligned on residue Y; the lower the pAE value, the higher the accuracy of the relative position of residue pairs and, consequently, the relative position of domains/subunits/proteins (50). PyMOL version 2.4.1 (<http://www.pymol.org/pymol>) was used to analyze and visualize the models. Structural alignments were performed using the PyMOL Alignment plugin with default settings. For calculation of surface charges, protein models were prepared using pdb2pqr (53) and electrostatics were calculated via the Adaptive Poisson-Boltzmann Solver (APBS) (53) plugin in PyMOL with default settings. Foldseek was used to identify protein homologs of EpsX in the PDB 100 database (54).

For the cooccurrence analysis of OPX and OM  $\beta$ -barrel proteins, we generated a database of 6,607 fully sequenced prokaryotic genomes that mirrors the KEGG database (55) and comprehensively covers all major prokaryotic phyla. OPX proteins encoded in these genomes were identified with HMMsearch of the HMMER package (v3.3) (56) using the Pfam (57) domain Poly\_export (PF02563) with the Pfam gathering threshold.

Subsequently, we generated a database corresponding to the proteins encoded by five genes up- and downstream of the identified OPX proteins. This database was searched for the Pfam domains of polysaccharide synthesis proteins (ABC2\_membrane [PF01061], ABC\_tran [PF00005], Polysacc\_synt [PF01943], Polysacc\_synt3 [PF13440], Wzt\_C [PF14524], Wzy\_C [PF04932], and Wzz [PF02706]) to (i) corroborate that the OPX protein is encoded in a polysaccharide synthesis gene cluster and (ii) assign a pathway to the OPX protein. PCPs identified with the Wzz domain were additionally analyzed for the presence of the tyrosine kinase domain to assign them as PCP-2. OPX proteins that were not encoded in a polysaccharide synthesis gene cluster were excluded from further analysis. To remove redundancy in the data set, OPX protein sequences were clustered to a 90% identity threshold using Cd-hit (58) and further analysis was carried out on the resulting set of representative sequences.

To generate the phylogenetic tree of the OPX protein sequences, the sequence matching the Poly\_export domain of the OPX proteins was aligned using MUSCLE (59) and a maximum likelihood tree was inferred using IQ-Tree (60) with automated model selection and support calculated with 1,000 ultra-fast bootstrap replicates. The tree was visualized and annotated using iTOL (61).

The secondary structure of the OPX proteins was determined using S4pred (62) and then scanned for an  $\alpha$ -helix within the last 20 C-terminal residues. C-terminal  $\alpha$ -helices were considered positive if they extended over more than nine residues and did not contain more than one gap of two residues predicted to be nonhelical.

$\beta$ -Barrel proteins were initially searched for by using hmmbuild to build a hidden Markov model (HMM) from an alignment of 31 EpsX, ExoB, and MXAN\_1916 homologs generated with MUSCLE. In a second step, proteins from the KEGG orthology group VpsM (K20920) were aligned and used to build a second HMM. Both models were used to search the proteins encoded by five genes up- and downstream of the genes encoding OPX proteins.  $\beta$ -Barrel fold and OM localization of identified proteins was verified using PRED-TMBB2 (63) and PSORTb 3.0 (64).  $\beta$ -Barrel proteins with fewer than 16  $\beta$ -strands were classified as false positive and were not considered further.

**Statistics.** The Welch's *t* test was performed to determine the statistical differences between the samples.

**Data availability.** The data that support the findings of this study are included in the article or in the supplemental material. The proteomics and pulldown data have been deposited in the ProteomeXchange Consortium via the PRIDE (65) partner repository with the data set identifier [PXD035138](https://proteomecentral.proteomexchange.org/protein/PXD035138).

## SUPPLEMENTAL MATERIAL

Supplemental material is available online only.

**FIG S1**, PDF file, 0.5 MB.

**FIG S2**, PDF file, 2.4 MB.

**FIG S3**, PDF file, 1.4 MB.

**FIG S4**, PDF file, 1.6 MB.

**FIG S5**, PDF file, 2.8 MB.

**FIG S6**, PDF file, 0.2 MB.

**FIG S7**, PDF file, 2 MB.

**FIG S8**, PDF file, 1.1 MB.

**TABLE S1**, DOCX file, 0.01 MB.

**TABLE S2**, DOCX file, 0.01 MB.

## ACKNOWLEDGMENTS

We thank Chris Whitfield for many helpful discussions and Dobromir Szadkowski for providing a Matlab script to illustrate AlphaFold model confidence plots.

The Max Planck Society generously supported this work.

We declare no conflicts of interest.

## REFERENCES

- Morris G, Harding S. 2009. Polysaccharides, microbial, p 482–494. In *Encyclopedia of microbiology*. Elsevier. <https://doi.org/10.1016/B978-012373944-5.00135-8>.
- Sutherland IW. 1998. Novel and established applications of microbial polysaccharides. *Trends Biotechnol* 16:41–46. [https://doi.org/10.1016/S0167-7799\(97\)01139-6](https://doi.org/10.1016/S0167-7799(97)01139-6).
- Hall-Stoodley L, Costerton JW, Stoodley P. 2004. Bacterial biofilms: from the natural environment to infectious diseases. *Nat Rev Microbiol* 2: 95–108. <https://doi.org/10.1038/nrmicro821>.
- Low KE, Howell PL. 2018. Gram-negative synthase-dependent exopolysaccharide biosynthetic machines. *Curr Opin Struct Biol* 53:32–44. <https://doi.org/10.1016/j.sbi.2018.05.001>.
- Whitfield C, Wear SS, Sande C. 2020. Assembly of bacterial capsular polysaccharides and exopolysaccharides. *Annu Rev Microbiol* 74:521–543. <https://doi.org/10.1146/annurev-micro-011420-075607>.
- Cuthbertson L, Mainprize IL, Naismith JH, Whitfield C. 2009. Pivotal roles of the outer membrane polysaccharide export and polysaccharide copolymerase protein families in export of extracellular polysaccharides in gram-negative bacteria. *Microbiol Mol Biol Rev* 73:155–177. <https://doi.org/10.1128/MMBR.00024-08>.
- Dong C, Beis K, Nesper J, Brunkan-Lamontagne AL, Clarke BR, Whitfield C, Naismith JH. 2006. Wza the translocon for *E. coli* capsular polysaccharides defines a new class of membrane protein. *Nature* 444:226–229. <https://doi.org/10.1038/nature05267>.
- Nickerson NN, Mainprize IL, Hampton L, Jones ML, Naismith JH, Whitfield C. 2014. Trapped translocation intermediates establish the route for export of capsular polysaccharides across *Escherichia coli* outer membranes. *Proc Natl Acad Sci U S A* 111:8203–8208. <https://doi.org/10.1073/pnas.1400341111>.
- Reid AN, Whitfield C. 2005. Functional analysis of conserved gene products involved in assembly of *Escherichia coli* capsules and exopolysaccharides: evidence for molecular recognition between Wza and Wzc for colanic acid biosynthesis. *J Bacteriol* 187:5470–5481. <https://doi.org/10.1128/JB.187.15.5470-5481.2005>.
- Collins RF, Beis K, Dong C, Botting CH, McDonnell C, Ford RC, Clarke BR, Whitfield C, Naismith JH. 2007. The 3D structure of a periplasm-spanning platform required for assembly of group 1 capsular polysaccharides in *Escherichia coli*. *Proc Natl Acad Sci U S A* 104:2390–2395. <https://doi.org/10.1073/pnas.0607763104>.
- Whitney JC, Hay ID, Li C, Eckford PDW, Robinson H, Amaya MF, Wood LF, Ohman DE, Bear CE, Rehm BH, Howell PL. 2011. Structural basis for alginate secretion across the bacterial outer membrane. *Proc Natl Acad Sci U S A* 108:13083–13088. <https://doi.org/10.1073/pnas.1104984108>.
- Wang Y, Andole Pannuri A, Ni D, Zhou H, Cao X, Lu X, Romeo T, Huang Y. 2016. Structural basis for translocation of a biofilm-supporting exopolysaccharide across the bacterial outer membrane. *J Biol Chem* 291:10046–10057. <https://doi.org/10.1074/jbc.M115.711762>.
- Acheson JF, Derewenda ZS, Zimmer J. 2019. Architecture of the cellulose synthase outer membrane channel and its association with the periplasmic TPR domain. *Structure* 27:1855–1861.e3. <https://doi.org/10.1016/j.str.2019.09.008>.
- Pérez-Burgos M, Søgaard-Andersen L. 2020. Biosynthesis and function of cell-surface polysaccharides in the social bacterium *Myxococcus xanthus*. *Biol Chem* 401:1375–1387. <https://doi.org/10.1515/hsz-2020-0217>.
- Holkenbrink C, Hoiczky E, Kahnt J, Higgs PI. 2014. Synthesis and assembly of a novel glycan layer in *Myxococcus xanthus* spores. *J Biol Chem* 289: 32364–32378. <https://doi.org/10.1074/jbc.M114.595504>.
- Kahnt J, Aguiluz K, Koch J, Treuner-Lange A, Konovalova A, Huntley S, Hoppert M, Søgaard-Andersen L, Hedderich R. 2010. Profiling the outer membrane proteome during growth and development of the social bacterium *Myxococcus xanthus* by selective biotinylation and analyses of outer membrane vesicles. *J Proteome Res* 9:5197–5208. <https://doi.org/10.1021/pr1004983>.
- Pérez-Burgos M, García-Romero I, Jung J, Schander E, Valvano MA, Søgaard-Andersen L. 2020. Characterization of the exopolysaccharide biosynthesis pathway in *Myxococcus xanthus*. *J Bacteriol* 202:e00335-20. <https://doi.org/10.1128/JB.00335-20>.
- Islam ST, Vergara Alvarez I, Saïdi F, Guiseppi A, Vinogradov E, Sharma G, Espinosa L, Morrone C, Brasseur G, Guillemot J-F, Benarouche A, Bridot J-L, Ravicoularamin G, Cagna A, Gauthier C, Singer M, Fierobe H-P, Mignot T, Mauriello EMF. 2020. Modulation of bacterial multicellularity via spatio-specific polysaccharide secretion. *PLoS Biol* 18:e3000728. <https://doi.org/10.1371/journal.pbio.3000728>.
- Lu A, Cho K, Black WP, Duan X-Y, Lux R, Yang Z, Kaplan HB, Zusman DR, Shi W. 2005. Exopolysaccharide biosynthesis genes required for social motility in *Myxococcus xanthus*. *Mol Microbiol* 55:206–220. <https://doi.org/10.1111/j.1365-2958.2004.04369.x>.
- Saïdi F, Mahanta U, Panda A, Jolivet NY, Bitazar R, John G, Martinez M, Mellouk A, Calmettes C, Chang Y-W, Sharma G, Islam ST. 2022. Bacterial outer-membrane polysaccharide export (OPX) proteins occupy three structural classes with selective  $\beta$ -barrel porin requirements for polymer secretion. *bioRxiv*. <https://doi.org/10.1101/2022.02.11.480155>.
- Rahn A, Beis K, Naismith JH, Whitfield C. 2003. A novel outer membrane protein, Wzi, is involved in surface assembly of the *Escherichia coli* K30 group 1 capsule. *J Bacteriol* 185:5882–5890. <https://doi.org/10.1128/JB.185.19.5882-5890.2003>.
- Bushell SR, Mainprize IL, Wear MA, Lou H, Whitfield C, Naismith JH. 2013. Wzi is an outer membrane lectin that underpins group 1 capsule assembly in *Escherichia coli*. *Structure* 21:844–853. <https://doi.org/10.1016/j.str.2013.03.010>.
- Jakobczak B, Keilberg D, Wuichet K, Søgaard-Andersen L. 2015. Contact- and protein transfer-dependent stimulation of assembly of the gliding motility machinery in *Myxococcus xanthus*. *PLoS Genet* 11:e1005341. <https://doi.org/10.1371/journal.pgen.1005341>.
- Friedrich C, Bulyha I, Søgaard-Andersen L. 2014. Outside-in assembly pathway of the type IV pili system in *Myxococcus xanthus*. *J Bacteriol* 196: 378–390. <https://doi.org/10.1128/JB.01094-13>.
- Treuner-Lange A, Chang YW, Glatter T, Herfurth M, Lindow S, Chreifi G, Jensen GJ, Søgaard-Andersen L. 2020. PiliY1 and minor pilins form a complex priming the type IVa pilus in *Myxococcus xanthus*. *Nat Commun* 11: 5054. <https://doi.org/10.1038/s41467-020-18803-z>.
- Herfurth M, Treuner-Lange A, Glatter T, Wittmaack N, Hoiczky E, Pierik AJ, Søgaard-Andersen L. 2022. A noncanonical cytochrome *c* stimulates calcium binding by PiliY1 for type IVa pili formation. *Proc Natl Acad Sci U S A* 119:e2115061119. <https://doi.org/10.1073/pnas.2115061119>.
- Sande C, Bouwman C, Kell E, Nickerson NN, Kapadia SB, Whitfield C. 2019. Structural and functional variation in outer membrane polysaccharide export (OPX) proteins from the two major capsule assembly pathways present in *Escherichia coli*. *J Bacteriol* 201:e00213-19. <https://doi.org/10.1128/JB.00213-19>.
- Sathiyamoorthy K, Mills E, Franzmann TM, Rosenshine I, Saper MA. 2011. The crystal structure of *Escherichia coli* group 4 capsule protein GfcC reveals a domain organization resembling that of Wza. *Biochemistry* 50: 5465–5476. <https://doi.org/10.1021/bi101869h>.
- Nesper J, Hill CMD, Paiment A, Harauz G, Beis K, Naismith JH, Whitfield C. 2003. Translocation of group 1 capsular polysaccharide in *Escherichia coli* serotype K30: structural and functional analysis of the outer membrane lipoprotein Wza. *J Biol Chem* 278:49763–49772. <https://doi.org/10.1074/jbc.M308775200>.
- Yang Y, Liu J, Clarke BR, Seidel L, Bolla JR, Ward PN, Zhang P, Robinson CV, Whitfield C, Naismith JH. 2021. The molecular basis of regulation of bacterial capsule assembly by Wzc. *Nat Commun* 12:4349. <https://doi.org/10.1038/s41467-021-24652-1>.
- Fong JCN, Syed KA, Klose KE, Yildiz FH. 2010. Role of *Vibrio* polysaccharide (*vps*) genes in VPS production, biofilm formation and *Vibrio cholerae* pathogenesis. *Microbiology (Reading)* 156:2757–2769. <https://doi.org/10.1099/mic.0.040196-0>.

32. Kaiser D. 1979. Social gliding is correlated with the presence of pili in *Myxococcus xanthus*. Proc Natl Acad Sci U S A 76:5952–5956. <https://doi.org/10.1073/pnas.76.11.5952>.
33. Shi X, Wegener-Feldbrugge S, Huntley S, Hamann N, Hedderich R, Søgaard-Andersen L. 2008. Bioinformatics and experimental analysis of proteins of two-component systems in *Myxococcus xanthus*. J Bacteriol 190:613–624. <https://doi.org/10.1128/JB.01502-07>.
34. Hodgkin J, Kaiser D. 1977. Cell-to-cell stimulation of movement in non-motile mutants of *Myxococcus*. Proc Natl Acad Sci U S A 74:2938–2942. <https://doi.org/10.1073/pnas.74.7.2938>.
35. Sambrook J, Russell DW. 2001. Molecular cloning: a laboratory manual, 3rd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
36. Schumacher D, Bergeler S, Harms A, Vonck J, Huneke-Vogt S, Frey E, Søgaard-Andersen L. 2017. The PomXYZ proteins self-organize on the bacterial nucleoid to stimulate cell division. Dev Cell 41:299–314.e13. <https://doi.org/10.1016/j.devcel.2017.04.011>.
37. Shi W, Zusman DR. 1993. The two motility systems of *Myxococcus xanthus* show different selective advantages on various surfaces. Proc Natl Acad Sci U S A 90:3378–3382. <https://doi.org/10.1073/pnas.90.8.3378>.
38. Bulyha I, Schmidt C, Lenz P, Jakovljevic V, Höne A, Maier B, Hoppert M, Søgaard-Andersen L. 2009. Regulation of the type IV pili molecular machine by dynamic localization of two motor proteins. Mol Microbiol 74:691–706. <https://doi.org/10.1111/j.1365-2958.2009.06891.x>.
39. Schmittgen TD, Livak KJ. 2008. Analyzing real-time PCR data by the comparative (C(T)) method. Nat Protoc 3:1101–1108. <https://doi.org/10.1038/nprot.2008.73>.
40. Bekker-Jensen DB, Martínez-Val A, Steigerwald S, Rütter P, Fort KL, Arrey TN, Harder A, Makarov A, Olsen JV. 2020. A compact quadrupole-Orbitrap mass spectrometer with FAIMS interface improves proteome coverage in short LC gradients. Mol Cell Proteomics 19:716–729. <https://doi.org/10.1074/mcp.TIR119.001906>.
41. Demichev V, Messner CB, Vernardis SI, Lilley KS, Ralser M. 2020. DIA-NN: neural networks and interference correction enable deep proteome coverage in high throughput. Nat Methods 17:41–44. <https://doi.org/10.1038/s41592-019-0638-x>.
42. Willforss J, Chawade A, Levander F. 2019. NormalizerDE: online tool for improved normalization of omics expression data and high-sensitivity differential expression analysis. J Proteome Res 18:732–740. <https://doi.org/10.1021/acs.jproteome.8b00523>.
43. Glatter T, Ludwig C, Ahrné E, Aebersold R, Heck AJR, Schmidt A. 2012. Large-scale quantitative assessment of different in-solution protein digestion protocols reveals superior cleavage efficiency of tandem Lys-C/Trypsin proteolysis over trypsin digestion. J Proteome Res 11:5145–5156. <https://doi.org/10.1021/pr300273g>.
44. Kammers K, Cole RN, Tiengwe C, Ruczinski I. 2015. Detecting significant changes in protein abundance. EuPA Open Proteom 7:11–19. <https://doi.org/10.1016/j.euprot.2015.02.002>.
45. Pérez-Burgos M, García-Romero I, Valvano MA, Søgaard Andersen L. 2020. Identification of the Wzx flippase, Wzy polymerase and sugar-modifying enzymes for spore coat polysaccharide biosynthesis in *Myxococcus xanthus*. Mol Microbiol 113:1189–1208. <https://doi.org/10.1111/mmi.14486>.
46. Kumar S, Stecher G, Tamura K. 2016. MEGA7: Molecular evolutionary genetics analysis version 7.0 for bigger datasets. Mol Biol Evol 33:1870–1874. <https://doi.org/10.1093/molbev/msw054>.
47. Saitou N, Nei M. 1987. The neighbor-joining method—a new method for reconstructing phylogenetic trees. Mol Biol Evol 4:406–425. <https://doi.org/10.1093/oxfordjournals.molbev.a040454>.
48. Kuzmich S, Blumenkamp P, Meier D, Szadkowski D, Goesmann A, Becker A, Søgaard-Andersen L, Buchrieser C. 2022. CRP-like transcriptional regulator MrpC curbs c-di-GMP and 3',3'-cGAMP nucleotide levels during development in *Myxococcus xanthus*. mBio 13:e00044-22. <https://doi.org/10.1128/mbio.00044-22>.
49. Teufel F, Almagro Armenteros JJ, Johansen AR, Gislason MH, Pihl SI, Tsirogas KD, Winther O, Brunak S, von Heijne G, Nielsen H. 2022. SignalP 6.0 predicts all five types of signal peptides using protein language models. Nat Biotechnol 40:1023–1025. <https://doi.org/10.1038/s41587-021-01156-3>.
50. Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, Tunyasuvunakool K, Bates R, Zidek A, Potapenko A, Bridgland A, Meyer C, Kohl SAA, Ballard AJ, Cowie A, Romera-Paredes B, Nikolov S, Jain R, Adler J, Back T, Petersen S, Reiman D, Clancy E, Zielinski M, Steinegger M, Pacholska M, Berghammer T, Bodenstein S, Silver D, Vinyals O, Senior AW, Kavukcuoglu K, Kohli P, Hassabis D. 2021. Highly accurate protein structure prediction with AlphaFold. Nature 596:583–589. <https://doi.org/10.1038/s41586-021-03819-2>.
51. Evans R, O'Neill M, Pritzel A, Antropova N, Senior A, Green T, Židek A, Bates R, Blackwell S, Yim J, Ronneberger O, Bodenstein S, Zielinski M, Bridgland A, Potapenko A, Cowie A, Tunyasuvunakool K, Jain R, Clancy E, Kohli P, Jumper J, Hassabis D. 2021. Protein complex prediction with AlphaFold-Multimer. bioRxiv. <https://doi.org/10.1101/2021.10.04.463034>.
52. Mirdita M, Schütze K, Moriwaki Y, Heo L, Ovchinnikov S, Steinegger M. 2022. ColabFold: making protein folding accessible to all. Nat Methods 19:679–682. <https://doi.org/10.1038/s41592-022-01488-1>.
53. Jurrus E, Engel D, Star K, Monson K, Brandt J, Felberg LE, Brookes DH, Wilson L, Chen J, Liles K, Chun M, Li P, Gohara DW, Dolinsky T, Konecny R, Koes DR, Nielsen JE, Head-Gordon T, Geng W, Krasny R, Wei G-W, Holst MJ, McCammon JA, Baker NA. 2018. Improvements to the APBS biomolecular solvation software suite. Protein Sci 27:112–128. <https://doi.org/10.1002/pro.3280>.
54. van Kempen M, Kim SS, Tumescheit C, Mirdita M, Söding J, Steinegger M. 2021. Foldseek: fast and accurate protein structure search. bioRxiv. <https://doi.org/10.1101/2022.02.07.479398>.
55. Kanehisa M, Goto S. 2000. KEGG: kyoto encyclopedia of genes and genomes. Nucleic Acids Res 28:27–30. <https://doi.org/10.1093/nar/28.1.27>.
56. Eddy SR. 2011. Accelerated profile HMM searches. PLoS Comput Biol 7:e1002195. <https://doi.org/10.1371/journal.pcbi.1002195>.
57. Mistry J, Chuguransky S, Williams L, Qureshi M, Salazar GA, Sonnhammer ELL, Tosatto SCE, Paladin L, Raj S, Richardson LJ, Finn RD, Bateman A. 2021. Pfam: the protein families database in 2021. Nucleic Acids Res 49:D412–D419. <https://doi.org/10.1093/nar/gkaa913>.
58. Li W, Godzik A. 2006. Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. Bioinformatics 22:1658–1659. <https://doi.org/10.1093/bioinformatics/btl158>.
59. Edgar RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res 32:1792–1797. <https://doi.org/10.1093/nar/gkh340>.
60. Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, von Haeseler A, Lanfear R. 2020. IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. Mol Biol Evol 37:1530–1534. <https://doi.org/10.1093/molbev/msaa015>.
61. Letunic J, Bork P. 2021. Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation. Nucleic Acids Res 49:W293–W296. <https://doi.org/10.1093/nar/gkab301>.
62. Moffat L, Jones DT. 2021. Increasing the accuracy of single sequence prediction methods using a deep semi-supervised learning framework. Bioinformatics 37:3744–3751. <https://doi.org/10.1093/bioinformatics/btab491>.
63. Tsirogas KD, Elofsson A, Bagos PG. 2016. PRED-TMBB2: improved topology prediction and detection of beta-barrel outer membrane proteins. Bioinformatics 32:i665–i671. <https://doi.org/10.1093/bioinformatics/btw444>.
64. Yu NY, Wagner JR, Laird MR, Melli G, Rey S, Lo R, Dao P, Sahinalp SC, Ester M, Foster LJ, Brinkman FS. 2010. PSORTb 3.0: improved protein subcellular localization prediction with refined localization subcategories and predictive capabilities for all prokaryotes. Bioinformatics 26:1608–1615. <https://doi.org/10.1093/bioinformatics/btq249>.
65. Perez-Riverol Y, Bai J, Bandla C, García-Seisdedos D, Hewapathirana S, Kamatchinathan S, Kundu DJ, Prakash A, Frericks-Zipper A, Eisenacher M, Walzer M, Wang S, Brazma A, Vizcaino JA. 2022. The PRIDE database resources in 2022: a hub for mass spectrometry-based proteomics evidences. Nucleic Acids Res 50:D543–D552. <https://doi.org/10.1093/nar/gkab1038>.
66. Schulz GE. 2000. Beta-barrel membrane proteins. Curr Opin Struct Biol 10:443–447. [https://doi.org/10.1016/S0959-440X\(00\)00120-2](https://doi.org/10.1016/S0959-440X(00)00120-2).
67. Wu SS, Kaiser D. 1997. Regulation of expression of the *pilA* gene in *Myxococcus xanthus*. J Bacteriol 179:7748–7758. <https://doi.org/10.1128/jb.179.24.7748-7758.1997>.
68. Julien B, Kaiser AD, Garza A. 2000. Spatial control of cell differentiation in *Myxococcus xanthus*. Proc Natl Acad Sci U S A 97:9098–9103. <https://doi.org/10.1073/pnas.97.16.9098>.
69. Jakovljevic V, Leonardy S, Hoppert M, Søgaard-Andersen L. 2008. PilB and PilT are ATPases acting antagonistically in type IV pilus function in *Myxococcus xanthus*. J Bacteriol 190:2411–2421. <https://doi.org/10.1128/JB.01793-07>.
70. Black WP, Wang L, Davis MY, Yang Z. 2015. The orphan response regulator EpsW is a substrate of the DifE kinase and it regulates exopolysaccharide in *Myxococcus xanthus*. Sci Rep 5:17831. <https://doi.org/10.1038/srep17831>.