Unique Regions of the Polysaccharide Copolymerase Wzz$_2$ from *Pseudomonas aeruginosa* Are Essential for O-Specific Antigen Chain Length Control

Steven M. Huszczynski,[a] Chelsea Coumoundouros,[a] Phi Pham,[a] Joseph S. Lam,[a] Cezar M. Khursigara[a]

[a]Department of Molecular and Cellular Biology, University of Guelph, Guelph, Ontario, Canada

ABSTRACT The outer leaflet of the outer membrane of nearly all Gram-negative bacteria contains lipopolysaccharide (LPS). The distal end of LPS may be capped with O antigen, a long polysaccharide that can range from a few to hundreds of sugars in length. The chain length of the polysaccharide has many implications for bacterial survival and consequently is tightly controlled. In the Wzx/Wzy-dependent route of O antigen synthesis, one or more Wzz proteins determine the chain length via an unknown mechanism. To gain insight into this mechanism, we identified and characterized important regions of two Wzz proteins in *Pseudomonas aeruginosa* serotype O13, which confer the production of "long" (Wzz$_1$) and "very long" (Wzz$_2$) chain lengths, respectively. We found that compared to Wzz$_1$, Wzz$_2$ has distinct amino acid insertions in the central α-helices (ins$_{\alpha6}$ and ins$_{\alpha7}$) and in membrane-distal (ins$_{L4}$) and -proximal (ins$_{IL}$) loops. When these regions were deleted in Wzz$_2$, the mutant proteins conferred drastically shortened chain lengths. Within these regions we identified several conserved amino acid residues that were then targeted for site-directed mutagenesis. Our results implicate an RTE motif in loop 4 and a "hot spot" of charged and polar residues in ins$_{\alpha7}$ in the function of Wzz$_2$. We present evidence that the functionally important residues of ins$_{\alpha7}$ are likely involved in stabilizing Wzz through coiled-coil interactions.

IMPORTANCE O antigen is an important virulence factor presented on the cell surface of Gram-negative bacteria that is critical for bacterial physiology and pathogenesis. However, some aspects of O antigen biosynthesis, such as the mechanisms for determining polysaccharide chain length, are poorly understood. In this study, we identified unique regions in the O antigen chain length regulators (termed Wzz) of the problematic opportunistic pathogen *Pseudomonas aeruginosa*. We show that these regions are critical for determining O antigen chain length, which provides new insight into the model of the Wzz mechanism. Ultimately, our work adds knowledge toward understanding an important step in the biosynthesis of this virulence factor, which is applicable to a wide range of Gram-negative pathogens.

KEYWORDS chain length, O antigen, *Pseudomonas aeruginosa*, lipopolysaccharide, polysaccharide co-polymerase

The dominant molecule in the outer leaflet of the outer membrane of nearly all Gram-negative bacteria is lipopolysaccharide (LPS), which plays integral roles in interactions between bacteria and their environment. LPS is a tripartite molecule, but minimally it can be found on the bacterial cell envelope comprised of two domains: lipid A, which anchors the glycolipid in the outer membrane, and a short core oligosaccharide. A third moiety, termed O antigen, may then cap the molecule by covalent linkage to the core. O antigen is characterized by its heterogeneity in the composition of its sugar constituents and in the length of the polymer, even within the

Address correspondence to Joseph S. Lam, jlam@uoguelph.ca, or Cezar M. Khursigara, ckhursig@uoguelph.ca.

same species. Various lengths of O antigen polymer have been shown to benefit bacteria in numerous ways, including resistance to serum (1–4), bacteriophage infection (5), and bile acid (6, 7). Further, regulation of O antigen chains to specific lengths is typically required for full virulence of many organisms (1, 2, 7–9), and the length of the O antigen chain is also an important consideration in the development of effective glycoconjugate vaccines against diseases caused by bacterial pathogens (10). Therefore, understanding the mechanisms by which bacteria regulate O antigen chain length is central to understanding bacterial survival and treatment of infections.

The synthesis of O antigen follows one of four known routes: an ABC transporter-dependent pathway, a synthase-dependent pathway, a Wzk-dependent pathway, or a Wzx/Wzy-dependent pathway (11). In the Wzx/Wzy-dependent pathway, O antigen is synthesized from short oligosaccharide building blocks, often termed O repeats or O units. O units are built on an undecaprenyl phosphate (Und-P) moiety at the inner face of the inner membrane, flipped to the periplasmic face by Wzx, and then polymerized into long chains by Wzy (12). Polymerization of O units by Wzy is relatively inefficient unless accompanied by another class of proteins called polysaccharide copolymerases (PCPs), which are designated Wzz according to standard nomenclature (13, 14). A given Wzz protein determines a range of polysaccharide lengths that is referred to as the modality, or modal lengths, and seems to be inherent to each Wzz protein, such that heterologous expression of a Wzz in a recipient strain will result in the production of LPSs of lengths similar to those of the strain that is the source of the gene (2, 10, 15–18). Remarkably, the lengths of the O antigen conferred by different Wzz proteins can range from a few to >100 O units, and many bacteria express more than one Wzz, resulting in the production of various populations of O antigen chains in a wide range of sizes (8, 19, 20).

The topology of Wzz proteins consists of an N- and C-terminal transmembrane helix with a large periplasmic domain between them. The structures of the periplasmic domains of several Wzz proteins have been solved, including WzzB and FepE from *Salmonella enterica* serovar Typhimurium, WzzB from *Shigella flexneri*, and WzzE from *Escherichia coli* (21–23). These structural studies have revealed that Wzz proteins share conserved protomeric and oligomeric structures; the periplasmic domain contains a membrane-proximal α/β base domain followed by a long α-helix that extends approximately 100 Å into the periplasm and leads into a flexible loop (loop 4) at the apex of the protein. This loop is followed by several helical and coiled structures that back to the membrane, leading into the second transmembrane helix (21). The protomers of Wzz form large oligomers that resemble a bell (see Fig. S1 in the supplemental material), although the number of subunits within this oligomer varies in the different published reports, apparently due to the methods used for structural determination (21, 23, 24). Recently, the first structural information for the transmembrane domains (TMDs) of WzzB in *S.* Typhimurium was reported, and the authors determined that the membrane helices of the same protomer form helical bundles but do not interact with the bundles of other protomers. This results in the TMDs resembling an open claw in the oligomer (25). The authors proposed that Wzy might interact with Wzz between these TMDs, which is supported by biochemical and genetic evidence of Wzy-Wzz interaction (26–29).

Even though the aforementioned structural data have provided insight into Wzz architecture, the details of how O antigen chain length is regulated by this protein are unknown. The most recent model suggests that chain length is determined by a combination of so-called "stopwatch" and "ruler" mechanisms (25). The proposed interaction of Wzy and Wzz via their TMD is the stopwatch: when this association breaks, polymerization rapidly terminates. The likelihood of dissociation would decrease as the polymer increases in length since this interaction would be stabilized by Wzz interactions with the polysaccharide. The ruler is defined by a given capacity of Wzz to bind polysaccharide; i.e., when this capacity is reached, polymerization ceases. While the ruler appears to be dependent on the combined contributions of positions throughout the Wzz protein, specific regions or amino acids have been implicated in

playing a greater role in the regulation of O antigen chain length. First, studies on two Wzz proteins of *S. flexneri* and *E. coli* (WzzB and FepE, respectively) revealed that insertion of small amino acid linkers or specific amino acid changes in a small loop at the base of the proteins could increase or decrease the chain length conferred by Wzz (30, 31). Second, studies on FepE indicated that deletions and specific amino acid replacements in loop 4 (situated at the top of the protein) drastically reduced O antigen chain length (22, 31, 32). However, since there is low sequence similarity of Wzz between species, whether these regions are important in other Wzz proteins is not readily apparent. Further, biochemical characterization of Wzz in other systems is necessary for a comprehensive understanding of Wzz-mediated chain length regulation of O antigen.

In this study, we investigated O antigen chain length regulation in the important opportunistic pathogen *Pseudomonas aeruginosa*. *P. aeruginosa* can colonize individuals with compromised host defenses, causing life-threatening infections. It is also the etiological agent of chronic lung dysfunction and death in patients with cystic fibrosis (33). One O antigen glycoform produced by *P. aeruginosa* is the O-specific antigen (OSA), which is the basis of the classification of this bacterium into 20 unique serotypes depending on the sugar identities and linkages in the polysaccharide chain (34). In those serotypes studied to date, OSA may be regulated to long and very long chain lengths by two Wzz proteins, Wzz$_1$ and Wzz$_2$, respectively (4, 17, 20, 35). Importantly, the long chains endow *P. aeruginosa* with resistance to serum, and the loss of both Wzz$_1$- and Wzz$_2$-mediated O antigen chain length regulation results in attenuation of the bacterium in a mouse model of pneumonia (4, 9). A previous report by Kintz and Goldberg showed that expressing variants of Wzz$_2$ with mutations at amino acid residue 321 in serotype O11 altered the OSA chain length, indicating that this position is important for Wzz$_2$ function (36). However, to date, further biochemical characterization of Wzz$_1$ and Wzz$_2$ is lacking. Here we report on the discovery of key sequence/structural elements involved in Wzz$_1$ and Wzz$_2$ function in a previously uncharacterized serotype, O13. Specifically, we discovered that the sequence of Wzz$_2$ contains unique insertions and deletions compared to the sequence of Wzz$_1$ and that these anomalies are critical to the function of Wzz$_2$. We also identified conserved amino acids in these areas of the protein that, when mutated, have varied effects on the production of Wzz$_2$-mediated chain lengths. We provide evidence that these functionally important amino acids are responsible for maintaining the stability of Wzz$_2$ but are not necessary for oligomerization. Finally, our results lead us to suggest a role for each of the described regions in the mechanism of Wzz.

RESULTS

Confirmation of the role of Wzz$_1$ and Wzz$_2$ in serotype O13. The IATS serotype O13 strain was chosen for this study because we found that this strain produces an LPS banding profile that consistently stains well by silver staining. This is critical when investigating changes to the Wzz-mediated OSA chain lengths, which may be subtle or drastic. Another advantage of using serotype O13 is that all O antigen bands can be attributed to OSA since this serotype does not produce common polysaccharide antigen, which is another O antigen produced by a separate pathway in some strains (37). Finally, probing the role of Wzz function in LPS biosynthesis in O13 offered an opportunity to investigate *P. aeruginosa* OSA chain length regulation in a serotype where this had yet to be examined. We first characterized the function of Wzz$_1$ and Wzz$_2$ in O13 by creating clean/in-frame deletions of each of the corresponding genes. When *wzz$_1$* was deleted from the O13 chromosome, the LPS banding pattern in silver-stained SDS-PAGE gels showed that the bands corresponding to long chain lengths of OSA were completely abrogated, leaving only two distinct very long OSA bands (Fig. 1A). Deleting *wzz$_2$* caused the mutant bacteria to produce LPS totally lacking these very long OSA bands, whereas the long chain length ladder-like bands were not affected. In the case of the double mutant in which *wzz$_1$* and *wzz$_2$* were both deleted, the bacteria produced unregulated OSA chain lengths. These results are
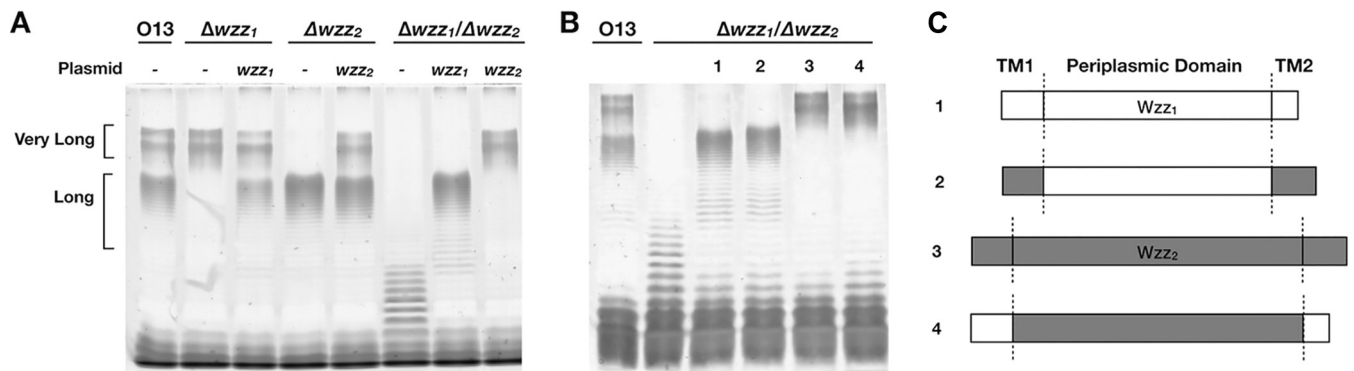
FIG 1 Characterization of the Wzz proteins of *P. aeruginosa* serotype O13. (A) Gene deletions of *wzz_1* and *wzz_2* affect OSA chain length. The results indicate that Wzz_1 regulates the long-chain OSA, while Wzz_2 regulates the very-long-chain OSA in serotype O13. (B) Reciprocal swaps of the periplasmic and TM/cytosolic domains of Wzz_1 and Wzz_2. Lane 1, Wzz_1; lane 2, Z2-Z1(prp)-Z2 (Wzz_2 TM/cytosolic domains with Wzz_1 periplasmic domain); lane 3, Wzz_2; lane 4, Z1-Z2(prp)-Z1 (Wzz_1 TM/cytosolic domains with Wzz_2 periplasmic domain). (C) The constructs were expressed in the O13 Δwzz_1 Δwzz_2 double-mutant background. The results indicate that the periplasmic domain is the main determinant of chain length, while the TM/cytosolic regions play a minor role.

consistent with previous studies of Wzz proteins in *P. aeruginosa*, i.e., Wzz_1 regulates the long chains and Wzz_2 regulates the very long chains.

A previous study by the Cygler group on two Wzz proteins from *E. coli* O157, namely, WzzB and FepE, showed that the periplasmic domain, but not the transmembrane (TM)/cytosolic domain, of these two proteins was responsible for the observed differences in the LPS chain lengths they confer (32). Hence, we aimed to test the hypothesis that the periplasmic domain of the Wzz proteins in *P. aeruginosa* O13 plays a similar role in LPS chain length regulation. To test this, we constructed reciprocal swaps of the periplasmic and TM/cytosolic domains of Wzz_1 and Wzz_2 and expressed them in the O13 Δwzz_1 Δwzz_2 double mutant (Fig. 1B). When a chimera of the Wzz_1 periplasmic domain and the Wzz_2 TM/cytosolic regions was expressed, the resulting OSA chains produced long LPS chain lengths that were nearly identical to the LPS banding pattern regulated by wild-type Wzz_1. However, a very slight increase in the average LPS chain length was observed. Similarly, a construct of the chimera expressing the Wzz_2 periplasmic domain fused with the Wzz_1 TM/cytosolic regions imparted an OSA modality similar to that conferred by Wzz_2, except for a slight decrease in chain length. These experiments verified that the periplasmic domains of the *P. aeruginosa* Wzz proteins play a predominant role in determining O antigen chain length. However, the TMS/cytosolic regions also play a role in determining O antigen chain length, albeit to a lesser extent.

Wzz_2 contains unique insertions and deletions compared to Wzz_1. Due to the obvious role of the periplasmic domains in determining chain length, we sought to identify any sequence features in this particular region that might be unique to either Wzz_1 or Wzz_2. Wzz_2 of O13 is 93 amino acids (aa) longer than Wzz_1 (442 aa versus 349 aa), and alignment of these two protein sequences identified several gaps (see Fig. S2 in the supplemental material). Further, the level of sequence identity between the two proteins is low, at only 29%. To accurately determine the locations of the gaps, we needed to analyze many homologs of these proteins in a multiple-sequence alignment (MSA). We compiled a diverse set of homologs of Wzz_1 and Wzz_2 from the genus *Pseudomonas* (Table S3) by subjecting the amino acid sequences to a BLAST search. Intriguingly, results from the MSA clearly showed that the sequences of the Wzz proteins included in the analysis could be separated into two groups with clearly defined insertions and deletions between them. These indels were at very different positions compared to the alignment involving only the sequences of Wzz_1 and Wzz_2. The first group, containing Wzz_2, had several insertions and one deletion compared to the second group, containing Wzz_1 (Fig. S3). The largest insertions in the Wzz_2 sequence were at aa 191 to 223 and aa 313 to 344. Two smaller insertions were also present at aa 141 to 144 and aa 271 to 279, and a small deletion (4 amino acids) was localized to Wzz_2 N70/V71, which aligned with Wzz_1 N72/E77 (Fig. 2).
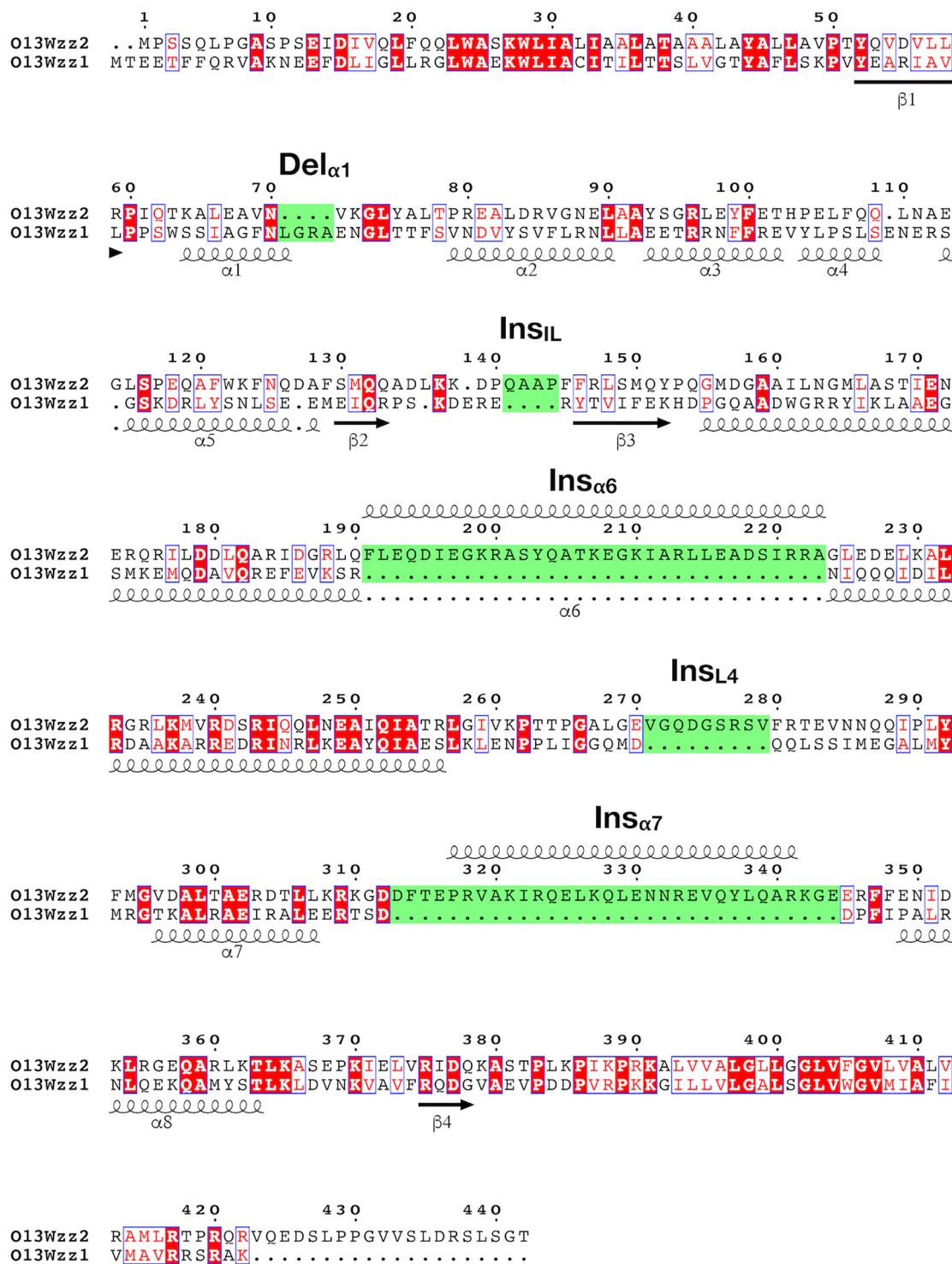
FIG 2 Alignment of Wzz$_1$ and Wzz$_2$ protein sequences reveals distinct insertions and deletions. Sequence alignment of Wzz$_1$ and Wzz$_2$. The alignment was extracted from a larger alignment of Wzz$_1$, Wzz$_2$, and their homologs in various *Pseudomonas* spp. The secondary structure based on the Wzz$_1$ homology model is displayed below the alignment, and the predicted structures of regions exclusively found in Wzz$_2$ are displayed above the alignment. The relevant insertions and deletions are highlighted in green.

To determine how these indels contribute to the overall protein structure, we created models of both Wzz$_1$ and Wzz$_2$ using the protein structure prediction program I-TASSER (38). As anticipated, the server generated an overall structure of Wzz$_1$ that was nearly identical to the top template, FepE (PDB code 3B8M; root mean square deviation
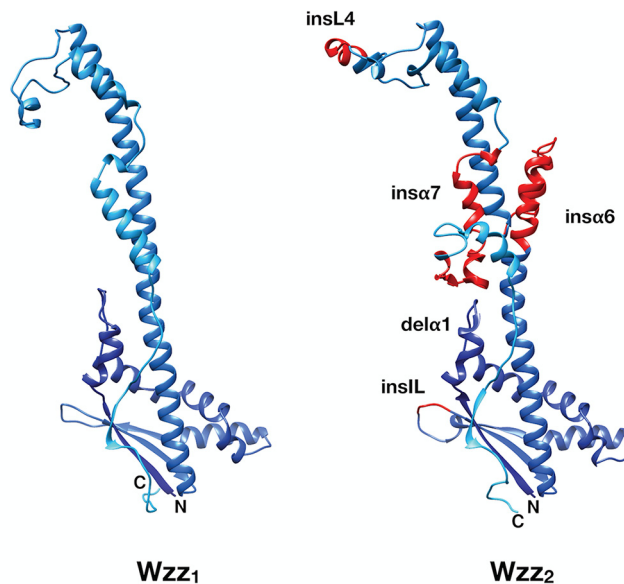
FIG 3 I-TASSER structural prediction of Wzz$_1$ and Wzz$_2$. The models are shaded dark to light blue from the N to the C terminus. The predicted structure and locations of the Wzz$_2$ insertions ("ins") and deletions ("del") are indicated in red.

[RMSD] = 0.81) (Fig. 3). Using this structural model, we were able to map the start of each large insertion of Wzz$_2$ (aa 191 to 223 and aa 313 to 344) to the long central α-helix (α6) and the coiled region between α7 and α8, which runs antiparallel to α6. In the model of the Wzz$_2$ protomer generated by I-TASSER, these insertions were shown to protrude from the central axis (Fig. 3). Since the published structures of Wzz revealed that it is oligomeric, the protrusions would most likely disrupt interactions between neighbouring protein subunits. Further, this model does not agree with the conserved protomeric structure described previously (21) (Fig. S1). Therefore, we deemed this model to be not representative of the true structure of this region. Instead, we propose that the insertions in α6 and between α7/8 would extend the Wzz$_2$ oligomers along the vertical axis. This interpretation is rational because the insertion sequences in question are essentially the same size (33 aa versus 32 aa). Presumably, any extension of the long central helix would require an equal extension in the opposite direction to maintain the overall symmetry of the Wzz structure. The insertions at aa 141 to 144 and aa 271 to 279 were localized to the loop between β2 and β3 (within the internal cavity of the protein) and L4 (at the apex of the protomer), respectively. Finally, the 4-aa deletion was within α1 (Fig. 3). Here we refer to these indels as ins$_{\alpha6}$, ins$_{\alpha7}$, ins$_{IL}$, ins$_{L4}$, ins$_{\alpha1}$, and del$_{\alpha1}$ accordingly.

Deletion of ins$_{IL}$, ins$_{L4}$, or ins$_{\alpha6/7}$ alters the chain length conferred by Wzz$_2$. To investigate the role played by ins$_{IL}$, ins$_{L4}$, and ins$_{\alpha6/7}$ in OSA chain length regulation, we prepared constructs with various deletions of these regions in Wzz$_2$ and observed changes to the LPS banding profiles when each of these proteins was expressed in the double-knockout mutant O13 Δwzz_1 Δwzz_2. Deletion of the four amino acids of ins$_{IL}$ caused a decrease in the LPS chain length that is normally imparted by Wzz$_2$ (Fig. 4A). Moreover, this deletion caused a change to the overall banding pattern of OSA in SDS-PAGE analysis. Rather than the usual bimodal banding pattern conferred by Wzz$_2$, Wzz$_2\Delta$ins$_{IL}$ conferred an LPS banding pattern showing equal distribution between the minimum and maximum chain length bands. Interestingly, this broad pattern is more similar to the pattern conferred by Wzz$_1$, which lacks this insertion. A previous study by Papadopolous and Morona showed that two different mutants with 5-aa insertions in the analogous region in WzzB increased the chain length (30). This led us to test the effect on OSA chain length when ins$_{IL}$ was introduced into Wzz$_1$. When expressed in the O13 Δwzz_1 Δwzz_2 background, this bacterial strain produced an OSA banding pattern that was similar to that imparted by Wzz$_1$ but was shifted toward shorter chain lengths.
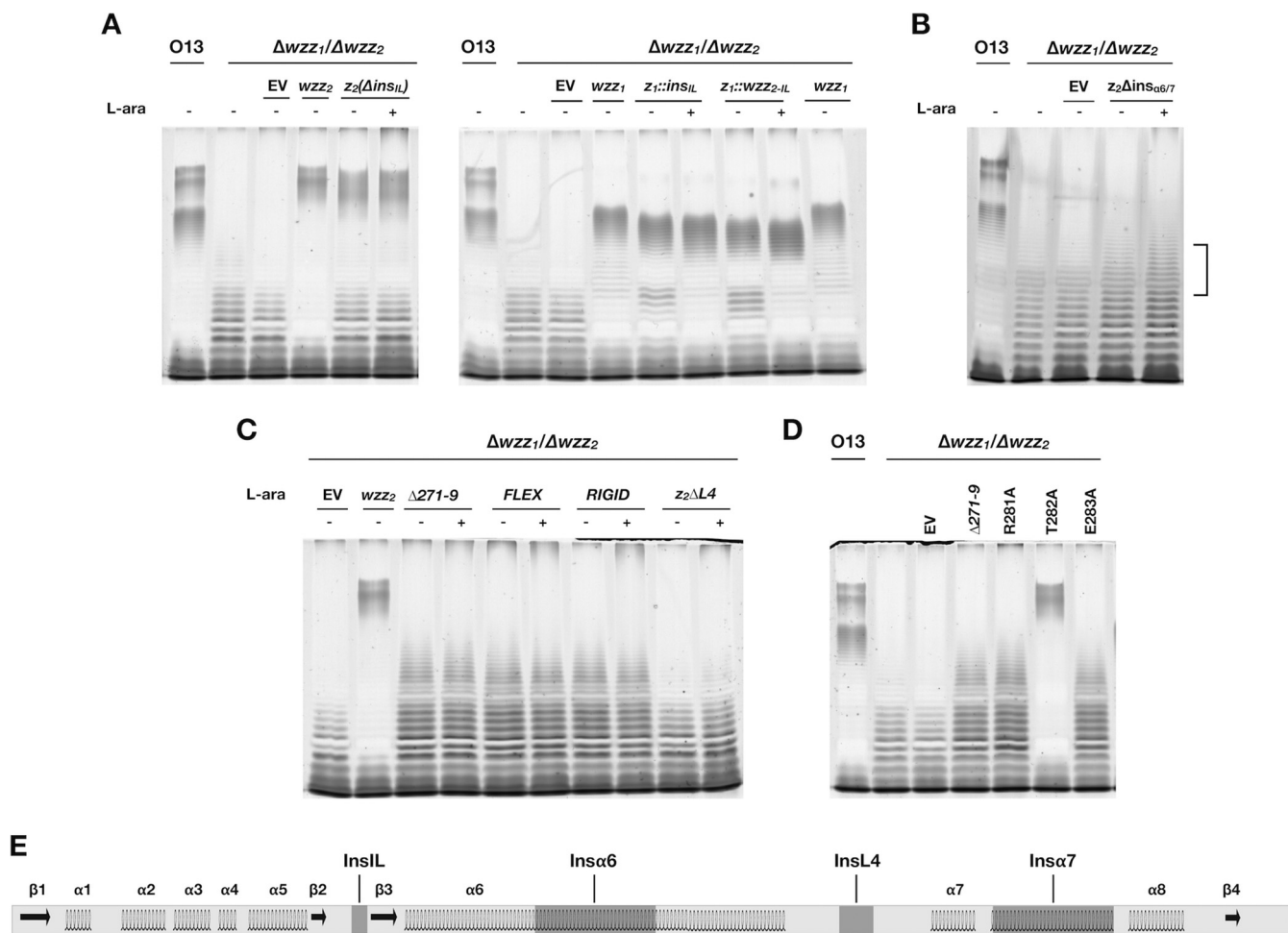
FIG 4 Manipulation of the indels of Wzz$_2$ modulates chain length determination. The mutants were expressed in the O13Δ*wzz*$_1$ Δ*wzz*$_2$ double mutant in the absence (−) or presence (+) of L-arabinose. (A) Deletion of ins$_{IL}$ in Wzz$_2$ [$z_2(\Delta ins_{IL})$] results in a shorter, broader chain length pattern. Insertion of ins$_{IL}$ into Wzz$_1$ ($z_1::ins_{IL}$) reduces chain length. Replacement of the entire Wzz$_1$ internal loop with that of Wzz$_2$ ($z_1::wzz_{2-IL}$) similarly reduces chain length. (B) Deletion of ins$_{\alpha6/7}$ in Wzz$_2$ ($z_2\Delta ins_{\alpha6/7}$) results in severely reduced chain length determining activity (bracket). (C) Deletion of residues 271 to 279 of ins$_{L4}$ in Wzz$_2$ (Δ*271-9*) significantly reduces the OSA chain length. Replacing this region with flexible (*FLEX*) or rigid (*RIGID*) peptide sequences has a similar effect. Deletion of L4 (Δ*L4*) results in the greatest reduction in chain length-determining activity. (D) L4 R281A and E283A variants result in a very short chain length phenotype similar to that observed when ins$_{L4}$ is deleted. (E) Schematic representation of the periplasmic domain of Wzz$_2$. The location of each region investigated is in dark gray. EV, empty vector.

We hypothesized that a more drastic change in this region could have a greater effect on chain length; this prompted us to construct a recombinant protein whereby the entire loop of Wzz$_1$ was swapped with the loop from Wzz$_2$. This construct resulted in an effect on Wzz$_1$-mediated LPS chain length regulation similar to that of the 4-aa insertion (Fig. 4A). This indicates that this particular region is more amenable to changes in amino acid sequence without significantly perturbing Wzz$_1$ function and that an increase in the length of this loop alone does not translate to an increase in OSA chain length. In Wzz$_2$, a small change to this loop appears to have a more drastic effect on LPS chain length regulation.

As described above, the location of ins$_{\alpha6}$ and ins$_{\alpha7}$ suggests that both of these two insertions needed to be present to maintain symmetry of the protomeric structure of Wzz$_2$. This led us to delete both of these insertions simultaneously, resulting in the construct Wzz$_2$Δins$_{\alpha6/7}$. When this construct was expressed in the O13 Δ*wzz*$_1$ Δ*wzz*$_2$ double mutant, it produced a significantly reduced, or "short," chain length (Fig. 4B). This indicates that these two insertions are critical for mediating very long OSA chain lengths. We attempted a similar experiment in which we inserted ins$_{\alpha6}$ and ins$_{\alpha7}$ into Wzz$_1$ to test if this would cause Wzz$_1$ to confer longer chain lengths. We were unable

to create a functional protein, which could be due to the low sequence identity of Wzz_1 and Wzz_2 (29%) (data not shown). Previous researchers have also been unsuccessful in constructing chimeric proteins from dissimilar Wzz proteins of the same strain (32), which highlights the difficulty of this particular experiment. We also tested the importance of $del_{\alpha 1}$ in Wzz_2 function. Since this region contains a 4-aa deletion between N70 and V71 compared to Wzz_1, we prepared a construct with the corresponding amino acids from Wzz_1 inserted into Wzz_2 (Wzz_2N70::LGRA). This construct did not exhibit any chain length-determining activity when expressed in the double-knockout mutant (data not shown). Comparison of the location of $del_{\alpha 1}$ to the FepE structure suggests that this region may be involved in contact with either $\alpha 6$ or $\alpha 8$ and that insertion of amino acids at this site disrupts the protein's fold (21). With this thought in mind, we made the construct Wzz_2(N70::LGRAΔins$_{\alpha 6/7}$), but this protein also did not confer any chain length-regulating activity (data not shown). This suggests that other contacts may be disrupted, but these were difficult to predict from our sequence alignment and protein model alone, which hindered further investigation of this region.

Next we investigated the role of ins_{L4}, which extends L4 of Wzz_2 by nine amino acids compared to Wzz_1. We prepared a Wzz_2 construct with a deletion of aa 271 to 279 ($Wzz_2\Delta$271-9) and expressed it in O13 Δwzz_1 Δwzz_2. The resulting OSA banding profile from LPS prepared from this recombinant bacterium was drastically shorter than that conferred by the wild-type Wzz_2. Interestingly, $Wzz_2\Delta$271-9 regulated the OSA chains to a maximum length similar to those imparted by Wzz_1, which does not contain this insertion (Fig. 4C). However, the most intense bands produced by the $Wzz_2\Delta$271-9 construct were only a few O units long, or "very short." In contrast, a construct of Wzz_2 with a full deletion of L4 appeared to exhibit almost negligible chain length-determining activity, as shown by some faint banding that was not present in the double mutant expressing the empty plasmid alone (Fig. 4C). These results are consistent with those reported by Kalynych et al. wherein a deletion of residues 256 to 273 in L4 of FepE caused a significant decrease in the LPS chain length regulated by that protein (32). In another study, expression of an engineered FepE that had residues 258 to 265 replaced with a short GSG flexible linker conferred an LPS banding pattern that was similar to the one displayed by the deletion construct (22). In those experiments, the length of the GSG linker did not match the number of amino acids that it was replacing. Based on the aforementioned observations, we hypothesized that the length of L4 could be important for determining chain length. Hence, we replaced aa 271 to 279 of Wzz_2 with a flexible linker with an equal number of amino acids (sequence: GGGGSGGGG) (39). We also tested the importance of flexibility in this region by inserting a physically more rigid peptide at this site with the sequence AAEAAAKAA (39). Remarkably, when either construct was expressed in O13 Δwzz_1 Δwzz_2, the chain length was indistinguishable from that imparted by $Wzz_2\Delta$271-9 (Fig. 4C). This suggests that the specific amino acids in this region are important for determining chain length, rather than the flexibility and length of the loop alone.

Conserved residues in ins_{L4} and $ins_{\alpha 6/\alpha 7}$ are critical for determining OSA chain length. Since we determined that ins_{IL}, ins_{L4}, and $ins_{\alpha 6/\alpha 7}$ were important for Wzz_2 function, we next targeted individual amino acids in these regions for site-directed mutagenesis to investigate the effects on Wzz_2-mediated chain length regulation. Guided by the MSA of Wzz homologs, we searched for conserved charged or polar amino acids since these classes of amino acids were previously suggested to be involved in determining chain length (22). We found that ins_{IL} had very little conservation among the Wzz_2-like group and were unable to identify any obvious candidate amino acids to target. Similarly, there was little conservation in ins_{L4} except for an Arg-Thr-Glu motif at positions 281 to 283 (Fig. S4). We switched each of these amino acids to alanine to investigate their role in determining chain length. The T282A variant had no appreciable effect on the chain length-determining activity of Wzz_2 when expressed in the O13 Δwzz_2 mutant background. However, replacement of R281 or E283 with alanine caused a loss of very long OSA chains and the appearance of a new modality that ranged from very short to short and was centered on OSA that was only

a few O units long (data not shown). This phenotype appeared very similar to that of the Wzz$_2$Δ271-9 construct. To clearly compare the three mutations, we analyzed the OSA profiles conferred by each construct in the O13ΔΔwzz$_1$ Δwzz$_2$ background. This analysis revealed that Wzz$_2$ with either mutation R281A or E283A conferred chain lengths that were identical to those conferred by Wzz$_2$Δ271-9 (Fig. 4D).

In ins$_{\alpha6}$ and ins$_{\alpha7}$ we identified several charged, polar, and aromatic amino acids that were highly conserved among the Wzz$_2$-like group (Fig. S4). Interestingly, ins$_{\alpha7}$ contained many more such conserved residues than did ins$_{\alpha6}$ (9 amino acids versus 3 amino acids). We targeted these residues for alanine replacement and analyzed the resulting LPS profiles when expressed in the O13 Δwzz$_2$ mutant. Ten of the 12 amino acids targeted in this manner affected the ability of Wzz$_2$ to regulate the length of the very long OSA chains. Specifically, the following amino acid changes abolished chain length control by Wzz$_2$ when expressed at basal levels from the pBAD promoter of pHERD20T: Y203A, E216A, D312A, F314A, N332A, R341A, and E344A (Fig. S5). Interestingly, two particular amino acid changes altered the Wzz$_2$-mediated banding profile: the R318A and R333A variants both caused a decrease in the average chain length conferred by Wzz$_2$. The R333A variant also caused a loss of the precise double banding pattern characteristic of the wild-type Wzz$_2$ and instead resulted in a broad population of chain lengths. When the expression of these single-amino-acid variants was increased by the addition of 0.1% L-arabinose to the growth medium, the resulting banding profiles were similar to those for the cognate uninduced samples except that we detected small amounts of very long OSA in cultures expressing the D312A and F314A variants (Fig. 5A). Further, in these two site-specific mutants, we also observed an increase in the amount of very short OSA. The results are summarized in Table 1. Due to the conservation of these functionally important residues, we tested the capacity of like amino acids at these positions to restore Wzz$_2$ function. The D312E and F314Y mutations were the only conservative amino acid switches that resulted in wild-type Wzz$_2$-mediated chain lengths when expressed at basal levels, whereas the E216D, R318K, N332Q, R333K, R341K, and E344D mutations did not restore Wzz$_2$ function (Fig. S5). The R318K variant had a chain length profile similar to that of the R318A mutant, wherein a broader and shorter population of very long chain lengths was produced. Surprisingly, the R333K variant was more detrimental to Wzz$_2$ than the more drastic R333A mutation. When induced with 0.1% arabinose, some very long OSA was detected in cultures expressing Wzz$_2$ N332Q and R333K variants (Fig. 5B). Once again, in the LPS preparations from the induced culture, we noticed that many Wzz$_2$ variants had an increased production of very short OSA, namely, the Y203F, D312E, F314Y, N332Q, R333K, R341K, and E344D variants (Fig. 5B and Table 1). Intriguingly, this resulted in a novel phenotype for the D312E, F314Y, N332Q, and R333K variants since they clearly mediated the simultaneous production of both very short and very long OSA. Together, these results point to the importance of a specific amino acid chemistry at certain positions and identifies ins$_{\alpha7}$ as a hot spot for amino acids critical to Wzz$_2$ function. We have also described a novel Wzz$_2$ phenotype: specific amino acid changes result in the production of two populations of OSA of vastly different lengths.

Wzz$_2$ mutants are expressed at reduced levels but are able to form higher-order oligomers. We hypothesized that some of the described mutations may affect the expression of Wzz$_2$. To test this, we prepared N-terminally FLAG-tagged versions of these proteins to allow for their detection by Western immunoblotting. To ensure that the FLAG tag had no effects on the phenotypes conferred by the Wzz$_2$ site-specific mutants, we reanalyzed the LPS profiles of O13 Δwzz$_2$ expressing these FLAG-tagged variants. The chain lengths conferred by these constructs were indistinguishable from those conferred by the corresponding untagged mutants (Fig. S6). Immunoblots of whole-cell lysates revealed that Wzz$_2$ harboring single mutations D312A, F314A, N332A, R341A, and E344A were expressed at levels much lower than that of the wild-type control, while proteins with single mutations R318A and R333A were expressed at levels similar to those of wild-type Wzz$_2$ (Fig. 6). Negligible amounts of Wzz$_2$ with mutations Y203A and E216A were detected. To test if any of these constructs were
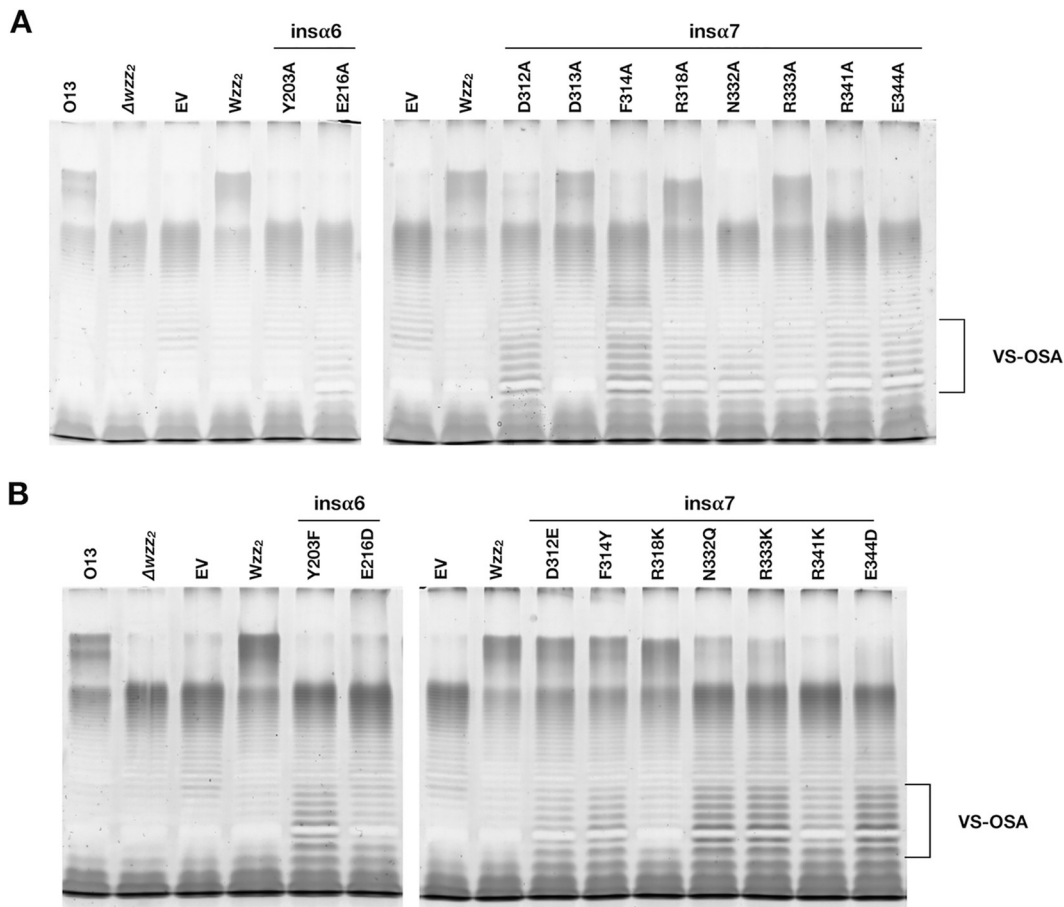
FIG 5 Replacement of conserved residues in ins$_{\alpha6/7}$ affects chain length regulation in Wzz$_2$. (A) Effect of replacing conserved amino acids with alanine. All variants were expressed in the O13 ΔWzz$_2$ background. The R318A and R333A variants alter the Wzz$_2$-mediated banding pattern, while the Y203A, E216A, N332A, R341A, and E344A mutations abolish chain length control. The D312A and F314A variants produce increased amounts of very short OSA and some very long OSA. (B) Effect of substituting functionally important residues with like amino acids in Wzz$_2$. The Y203F, R341K, and E344D variants produce very short OSA, while the D312E, F314Y, N332Q, and R333K variants simultaneously mediated the production of both very short (VS) and very long OSA.

defective in the ability to oligomerize, we performed *in vivo* formaldehyde cross-linking. In samples expressing *FLAG-wzz$_2$*, several bands appeared after cross-linking that were not present in the untreated or plasmid-only controls. The two most intense bands had apparent molecular masses of 53 kDa and 143 kDa, which corresponded closely with the expected size of the Wzz$_2$ monomer and trimer (expected masses of 50 kDa and 150 kDa, respectively). Two less intense bands that migrated in the SDS-PAGE gel above the 245-kDa molecular mass standard had apparent molecular masses of 263 kDa and 281 kDa and likely correspond to pentamers and hexamers (expected masses of 250 kDa and 300 kDa, respectively). As expected, the cross-linked bands of constructs that expressed poorly were less intense than the wild-type protein. However, we were still able to detect banding profiles similar to those of FLAG-Wzz$_2$ in the cross-linked samples of the mutant proteins (except for the Y203A and E216A mutants). In all samples we detected a band consistent with a Wzz$_2$ trimer, but the bands corresponding to the pentamer and hexamer were detected only in the F314A, R318A, and R333A mutants. Overall, this suggests that these mutants were capable of forming higher-order oligomers.

DISCUSSION

In this study, we have thoroughly characterized important regions of the OSA chain length regulator, Wzz$_2$, from *P. aeruginosa* serotype O13 by constructing deletions, chimeras, and single amino acid changes guided by the alignment of a diverse set of

TABLE 1 Summary of LPS profiles conferred by Wzz₂ insα6/7 point mutants[a]

Construct	Location of mutation	Production of:		
		Very short OSA	Very long OSA	Protein
Wzz₂	NA	−	+	+
Y203A variant	insα6	−	−	−
E216A variant	insα6	−	−	−
D312A variant	insα7	+	+/−	+/−
F314A variant	insα7	+	+/−	+/−
R318A variant	insα7	−	+	+
N332A variant	insα7	−	−	−
R333A variant	insα7	−	+	+
R341A variant	insα7	−	−	+/−
E344A variant	insα7	−	−	−
D313A variant	insα7	−	+	NA
Y203F variant	insα6	+	−	NA
E216D variant	insα6	−	−	NA
D312E variant	insα7	+	+	NA
F314Y variant	insα7	+	+	NA
R318K variant	insα7	−	+	NA
N332Q variant	insα7	+	+/−	NA
R333K variant	insα7	+	+/−	NA
R341K variant	insα7	+	−	NA
E344D variant	insα7	+	−	NA

[a]+, appreciable production; +/− intermediate production; −, no appreciable production; NA, not applicable.

Wzz sequences. We established that the differences in the chain lengths imparted by Wzz₁ and Wzz₂ are mostly due to the variations in their periplasmic domains. However, our results showed that chain length was partially dependent on the TM/cytosolic domain (Fig. 1B), which is in contrast to a previous study performed by Kalynych et al. (32). These authors recombinantly fused the periplasmic region of FepE of *E. coli* O157 to the TM/cytosolic domain of WzzB of the same strain, which resulted in the production of O antigen chains identical to those mediated by the wild-type FepE. This led
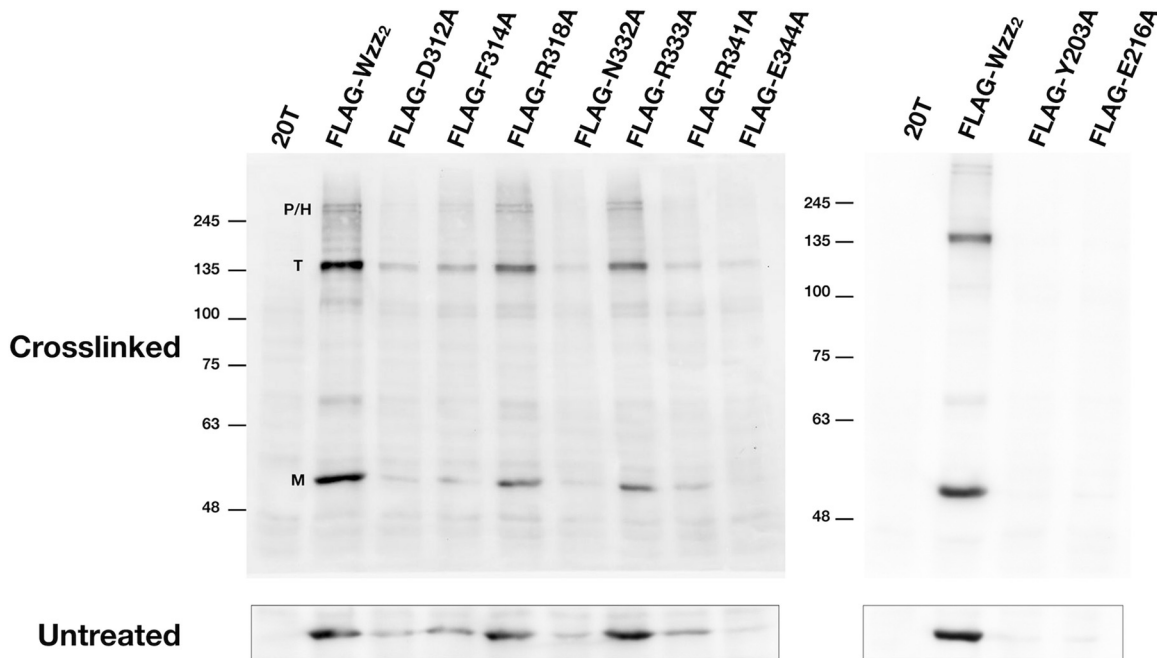
FIG 6 Stability of Wzz₂ insα6/7 mutants determined by whole-cell protein expression and *in vivo* cross-linking. (Bottom) Whole-cell lysates of O13Δ*wzz₂* expressing FLAG-tagged variants of Wzz₂ were subjected to SDS-PAGE and immunoblotting with anti-FLAG antibody and detected using chemiluminescence. Several variants express at levels lower than for the wild-type control. (Top) Corresponding whole-cell cross-linked samples of FLAG-tagged Wzz₂ variants. Bands were detected that approximately correspond to the monomeric (M), trimeric (T), and penta-/hexameric (P/H) forms of FLAG-Wzz₂. Numbers on the left are molecular masses, in kilodaltons.

them to conclude that the periplasmic region solely determines the chain length in this protein. It was recently suggested that the space between the TMDs of Wzz protomers could be a binding pocket for Wzy (25). We postulate that the variations in chain length we observed could be due to differential affinity of Wzy to the TMDs of Wzz_1 and Wzz_2, which affects the time until release of Wzy, i.e., the molecular stopwatch.

Our analysis of Wzz_1 and Wzz_2 homologs from the genus *Pseudomonas* has led to the observation that there are distinct sequence insertions in Wzz_2 compared to Wzz_1, specifically in the loop between $\beta2$ and $\beta3$, in L4, and in the two antiparallel α-helical regions. The results of our study indicate that the loop between $\beta2$ and $\beta3$ has only a minor role in chain length regulation and that Wzz can accommodate drastic changes to the amino acid sequence at this site. In fact, attempts to replace specific regions of Wzz_1 with the cognate segments from Wzz_2 to create a functional chimera were unsuccessful except for the constructs presented here, $Wzz_1::ins_{IL}$ and $Wzz_1::Wzz_{2-IL}$. Indeed, there is a lack of conservation in this loop, which supports the notion that the specific amino acid identities of this region are not entirely essential to Wzz_1 or Wzz_2 function. This is in accordance with the study of Papadopolous and Morona (30), in which the insertion of random 5-aa linkers into the $\beta2$-$\beta3$ loop of $WzzB_{SF}$ resulted in an increase in the chain length conferred by this protein, even though mutations in the α/β-base domain of Wzz proteins typically result in a knockout or abrogated chain length phenotype. In that study, the length of the $\beta2$-$\beta3$ loop was suggested to control the Wzz barrel size and thus partly determine chain length; i.e., a larger loop increases barrel size and consequently the chain length. We found that deletion of amino acids in this loop in Wzz_2 did in fact cause a decrease in chain length, which would be consistent with their hypothesis. However, the reciprocal experiment of inserting amino acids into Wzz_1 decreased rather than increased the chain length conferred by this protein. Importantly, the experimental designs of our study and the study of Papadopoulous and Morona are comparable; the insertion we made in Wzz_1 was at the predicted end of the loop just before the start of $\beta3$, which is the same position as one of the insertions (i131) that was reported to increase the chain length imparted by $WzzB_{SF}$ (30). Notably, the sequence of our insertion (QAAP) and the sequence of i131 (LRPQP) both terminate in proline. We analyzed the solved crystal structures of WzzB and FepE to investigate if the $\beta2$-$\beta3$ loop made any significant inter- or intraprotomer contacts that might explain how mutations in this region could affect chain length, but this exercise did not reveal any obvious interactions. Based on the full-length cryo-electron microscopy (cryo-EM) structure of $WzzB_{SF}$, Collins et al. hypothesized that Wzy and/or the polysaccharide could be contained within the barrel of Wzz (25). Therefore, we suggest that this loop is more likely to be involved with interaction with the polysaccharide chain or Wzy and that variations in this sequence alter the affinity for either of these molecules, rather than affecting the size of the Wzz barrel.

We were surprised to observe that the Wzz_2 constructs with a deletion of aa 271 to 279 of L4 or replacement with either a flexible or rigid peptide sequence produced similar chain lengths that ranged from very short to long. Since experimental evidence indicates L4 is highly flexible (21–23, 32), we expected replacement with a sterically hindered peptide sequence (i.e., AAEAAAKAA) to have a greater effect on Wzz function than a more flexible one (i.e., GGGGSGGGG). One possibility is that we removed important amino acids from region 271 to 279 when constructing these replacements. However, we did not find any obvious residues to target for site-directed mutagenesis to explore this hypothesis. Another explanation is that manipulation of this area caused a displacement of a required orientation of the adjacent residues R281 and E283, which we showed had similar chain length phenotypes when individually mutated to alanine compared to the flexible or rigid amino acid insertions. To explore possible roles for these residues, we draw upon previous mutagenesis and structural studies. One report identified residue D268 in L4 of FepE to be important for chain length control since several amino acid changes at this position resulted in altered O antigen lengths (31). This residue also aligns with another residue, D226, in $WzzB_{SF}$ (31). Analysis of the solved crystal structures of FepE (PDB code 3B8M) and $WzzB_{SF}$ (PDB codes 4ZM1 and

4E2H) indicates that these aspartate residues are involved in putative salt bridge interactions with lysine residues in α7 (FepE) and α3 and α4 (WzzB). Interestingly, these interactions appear to be disrupted in the structure of a FepE L4 variant conferring shorter chain lengths (PDB code 4E2L) (see Fig. S7 in the supplemental material). Our analysis of these structures indicates that the charged residues of L4 have a role in mediating inter- and intraprotomer contacts with other charged residues in the helical regions of Wzz proteins. Therefore, we predict that this is the role of R281 and E283 in Wzz$_2$. However, without an experimentally derived structure of Wzz$_2$, these potential interactions are difficult to discern from sequence alone.

Wzz$_1$ is comparable in size to FepE, which is the largest Wzz for which structural information is available. Since Wzz$_2$ is larger than these two proteins, our results suggest how the structure of Wzz$_2$ might differ from the previously characterized chain length regulators. The most notable difference is in the extra amino acids in α6 and α7. Hence, we hypothesized that the two large insertions identified in this study, ins$_{\alpha6}$ and ins$_{\alpha7}$, extend the Wzz$_2$ protomer along the central axis. This may be one of the reasons that Wzz$_2$ mediates the production of longer chain lengths than for Wzz$_1$; the extension of the protomer creates an oligomer that can accommodate a longer polysaccharide chain. This is supported by the observation that simultaneous deletions of ins$_{\alpha6}$ and ins$_{\alpha7}$ resulted in a Wzz$_2$ protein that conferred drastically shortened chain lengths. The fact that these two large regions could be deleted and a functional protein could still be obtained also suggests that the majority of the inter- and intraprotomer contacts made by ins$_{\alpha6}$ and ins$_{\alpha7}$ either are to each other or are nonspecific interactions. However, we cannot dismiss the idea that deletion of these sequences may have disrupted specific interactions, which could have also contributed to the shortened chain length phenotype.

Replacement of conserved amino acids in ins$_{\alpha6}$ and ins$_{\alpha7}$ revealed several residues that are absolutely essential to producing Wzz$_2$-regulated chain lengths. We found that most of these variants of Wzz$_2$ were expressed at lower levels than the wild-type protein, implying a role in conferring stability or the proper folding of Wzz$_2$. Interestingly, several of the residues we replaced (R318, N332, R333, R341, E344, and E345) are found within a predicted coiled coil that spans aa 318 to 347. Notably, OSA chain length and Wzz$_2$ protein stability were previously reported by Kintz and Goldberg to be dependent on the residue at position 321, which they identified to be within this predicted coiled coil (36). Similar to our findings, the authors described that various amino acid changes at this site decreased protein expression and/or altered the OSA chain length. However, the only change that abolished chain length control was a complete deletion of this position. This is in contrast to our study: we found that replacement of many residues in ins$_{\alpha7}$ with alanine or even a similar amino acid abrogated Wzz$_2$ function, suggesting that the amino acid identity at these positions is more critical. Consistent with this hypothesis, residues N332, R333, and E344 are found in positions *e* and *g* of the predicted coiled-coil heptad, which often form stabilizing interactions with the corresponding positions of another helix. Importantly, the *e* and *g* positions are typically occupied by charged or polar residues, which is consistent with the physiochemical properties of Asn, Arg, and Glu. In contrast, residue 321 occupies the *c* position, which is typically less essential in the formation of the coiled coil (40). Although the number of predicted coiled coils in Wzz proteins was reported to have a positive correlation with modal length (14), it should be noted that the Wzz structures solved to date do not show that these proteins form extensive coiled-coil interactions. However, when applying the algorithm of the Network Protein Server Analysis (NPS@) online platform to predict coiled-coil domains (41), Wzz$_2$ has the highest score of those Wzz proteins with experimentally determined structures (data not shown). It would be useful to determine the structure of Wzz$_2$ to examine the extent of coiled-coil interactions compared to other PCPs and how the residues from this study are positioned in the Wzz$_2$ oligomer. We also found that residues of ins$_{\alpha6}$ and ins$_{\alpha7}$ that were outside of the predicted coiled-coil region were essential for Wzz$_2$ function, namely, Y203, E216, D312, and F314. The decreased expression of alanine mutants at these positions

suggests that they also have a role in maintaining the stability of Wzz_2 but is likely not due to coiled-coil interactions.

Despite the reduced expression of Wzz_2 variants with mutations in $ins_{\alpha6}$ and $ins_{\alpha7}$, we did not detect any differences in the cross-linked protein profiles of this protein from cultures expressing these mutants, which is consistent with experiments performed with other organisms (31, 32, 36, 42). We suggest that *in vivo,* the mutants described here are able to oligomerize but cannot maintain this state and are degraded faster than the wild-type Wzz_2. However, during the cross-linking experiments, this transient oligomer is covalently captured. Our results also suggest that some mutants are able to maintain a biologically relevant oligomer in a complex with Wzy long enough to regulate chain length. Specifically, the Y203F, D312A, D312E, F314A, F314Y, N332Q, R333K, R341K, and E344D variants produced some very short OSA when overexpressed. We propose that these oligomers cannot withstand the force of the growing polysaccharide chain and disassemble before longer chains can be made. In the D312E, F314Y, N332Q, and R333K variants, we observed the distinct production of both very short and very long chain lengths, suggesting that in these cases, the oligomer might be stabilized once the polysaccharide reaches a certain length, thereby allowing polymerization to proceed.

Based on the data presented here and by others, we propose that the capacity of the Wzz "ruler" is a combination of the following: (i) Wzz oligomer size, (ii) oligomer stability, and (iii) nonspecific interactions between the polysaccharide and inside of the Wzz barrel. As polymerization begins, the polysaccharide would be sequestered within the barrel of Wzz and form nonspecific stabilizing interactions with amino acids of the $\beta2$-$\beta3$ loop and residues facing the inside of the Wzz cavity (30, 31). During polymer elongation, the pressure of containing the polysaccharide builds and increases the likelihood of Wzz oligomer disassembly (Fig. S8). In this model, Wzz proteins with larger cavities would accommodate longer polysaccharides, and those that form stronger inter- and intraprotomer contacts would be able to maintain the oligomeric state for more rounds of polymerization. These stabilizing contacts would be made throughout the Wzz protomer and include the interactions made by L4 and the coiled-coil regions. Future studies to determine the structure of Wzz_2 should be pursued to test the extent of these interactions in this protein.

MATERIALS AND METHODS

Bacterial strains, plasmids, and growth conditions. Bacteria were cultured at 37°C in LB. When necessary, cultures were supplemented with the appropriate antibiotics at the following concentrations: 100 μg/ml of ampicillin and 10 μg/ml of gentamicin for *E. coli* or 300 μg/ml of carbenicillin and 300 μg/ml of gentamicin for *P. aeruginosa*. *P. aeruginosa* IATS serotype O13 (ATCC 33360) was acquired from our lab collection.

General DNA methods and construction of *wzz₁* and *wzz₂* deletions. PCR was performed using either Phusion Hot Start DNA polymerase or DreamTaq Green polymerase (Thermo). PCR and plasmid purification were performed using GeneJet purification kits (Thermo). Restriction enzymes and T4 DNA ligase were purchased from New England BioLabs. All enzymes and kits were used according to the manufacturer's instructions. The *wzz₁* and *wzz₂* deletion constructs were cloned into pEX18Gm by gene splicing by overlap extension (SOEing) (43). Briefly, two amplicons corresponding to the upstream and downstream regions of *wzz₁* or *wzz₂* were amplified in separate reactions. The PCR products were designed to contain a short region of homology to each other to facilitate their annealing in a second PCR, wherein the final construct was amplified. The PCR products were cloned into the EcoRI and XbaI sites of pEX18Gm. *P. aeruginosa* O13 was transformed with pHERD20T plasmids using the electroporation protocol described by King et al. (44). For the construction of point mutations, insertions, and deletions in *wzz₂*-pHERD20T, either the QuikChange mutagenesis protocol (Agilent) or the *in vitro* assembly method was followed (45). All primers used in this study are listed in Table S1 and Table S2 in the supplemental material.

Bioinformatics analyses. Homologs of Wzz_1 and Wzz_2 were identified using blastp (46). The search was limited to the *Pseudomonas* genus (excluding *P. aeruginosa*), and the maximum number of target sequences was set to 1,000. Only hits with a sequence identity to Wzz_1 or Wzz_2 greater than 29% were considered. The list of homologs was manually curated to remove partial sequences and those representing non-Wzz polysaccharide copolymerases. Sequences were aligned with Clustal Omega (47) and visualized with Jalview V2.10.0b1 (48). The final alignment was generated by removing redundant protein sequences (70% identity) based on the Jalview "remove redundancy" algorithm. All alignment images were generated using the ESPript 3.0 server (49).

Visualization of LPS. LPS was prepared from overnight cultures according to the method of Hitchcock and Brown (50). Samples (5 μl) were electrophoresed on 10% SDS-PAGE gels and visualized by the rapid silver staining method of Fomsgaard et al. (51).

***In vivo* cross-linking.** *In vivo* formaldehyde cross-linking was performed as described previously (20), with minor modifications. Overnight cultures induced with 0.1% L-arabinose were washed twice in phosphate-buffered saline (PBS) and equilibrated to the equivalent of a 0.5-ml culture with an optical density at 600 nm (OD$_{600}$) of 0.7. The bacteria were suspended in either 0.5 ml of PBS or 0.5 ml of PBS containing 0.5% formaldehyde and incubated at room temperature for 1 h. Following treatment, bacteria were pelleted and suspended in 50 μl of SDS-PAGE loading dye. For Western immunoblotting, samples (15 μl) were electrophoresed on 8% SDS-PAGE gels and transferred to nitrocellulose. The primary antibody used was monoclonal mouse anti-FLAG, and the secondary antibody was goat anti-mouse antibody conjugated to horseradish peroxidase (HRP). Blots were developed using the Illuminata HRP substrate (EMD Millipore). The molecular mass of the bands was predicted using the Quantity One one-dimensional (1D) analysis software from Bio-Rad.

SUPPLEMENTAL MATERIAL

Supplemental material for this article may be found at https://doi.org/10.1128/JB .00165-19.

SUPPLEMENTAL FILE 1, PDF file, 1.8 MB.

ACKNOWLEDGMENTS

REFERENCES

1. Hong M, Payne SM. 1997. Effect of mutations in *Shigella flexneri* chromosomal and plasmid-encoded lipopolysaccharide genes on invasion and serum resistance. Mol Microbiol 24:779–791. https://doi.org/10 .1046/j.1365-2958.1997.3731744.x.

2. Murray GL, Attridge SR, Morona R. 2006. Altering the length of the lipopolysaccharide O antigen has an impact on the interaction of *Salmonella enterica* serovar Typhimurium with macrophages and complement. J Bacteriol 188:2735–2739. https://doi.org/10.1128/JB.188.7.2735 -2739.2006.

3. Bravo D, Silva C, Carter JA, Hoare A, Alvarez SA, Blondel CJ, Zaldívar M, Valvano MA, Contreras I. 2008. Growth-phase regulation of lipopolysaccharide O-antigen chain length influences serum resistance in serovars of *Salmonella*. J Med Microbiol 57:938–946. https://doi.org/10.1099/jmm .0.47848-0.

4. Kintz E, Scarff JM, DiGiandomenico A, Goldberg JB. 2008. Lipopolysaccharide chain length regulation in *Pseudomonas aeruginosa* serogroup O11 strain PA103. J Bacteriol 190:2709–2716. https://doi.org/ 10.1128/JB.01646-07.

5. Cota I, Sánchez-Romero MA, Hernández SB, Pucciarelli MG, García-Del Portillo F, Casadesús J. 2015. Epigenetic control of Salmonella enterica O-antigen chain length: a tradeoff between virulence and bacteriophage resistance. PLoS Genet 11:e1005667. https://doi.org/ 10.1371/journal.pgen.1005667.

6. May JF, Groisman EA. 2013. Conflicting roles for a cell surface modification in *Salmonella*. Mol Microbiol 88:970–983. https://doi.org/10.1111/ mmi.12236.

7. Crawford RW, Keestra AM, Winter SE, Xavier MN, Tsolis RM, Tolstikov V, Bäumler AJ. 2012. Very long O-antigen chains enhance fitness during Salmonella-induced colitis by increasing bile resistance. PLoS Pathog 8:e1002918. https://doi.org/10.1371/journal.ppat.1002918.

8. Murray GL, Attridge SR, Morona R. 2003. Regulation of Salmonella typhimurium lipopolysaccharide O antigen chain length is required for virulence; identification of FepE as a second Wzz. Mol Microbiol 47: 1395–1406. https://doi.org/10.1046/j.1365-2958.2003.03383.x.

9. Ivanov IE, Kintz EN, Porter LA, Goldberg JB, Burnham NA, Camesano TA. 2011. Relating the physical properties of *Pseudomonas aeruginosa* lipopolysaccharides to virulence by atomic force microscopy. J Bacteriol 193:1259–1266. https://doi.org/10.1128/JB.01308-10.

10. Hegerle N, Bose J, Ramachandran G, Galen JE, Levine MM, Simon R, Tennant SM. 2018. Over-expression of O-polysaccharide chain length regulators in Gram-negative bacteria using the Wzx/Wzy-dependent pathway enhances production of defined modal length O polysaccha-

ride polymers for use as haptens in glycoconjugate vaccines. J Appl Microbiol 125:575–585. https://doi.org/10.1111/jam.13772.

11. Whitfield C, Trent MS. 2014. Biosynthesis and export of bacterial lipopolysaccharides. Annu Rev Biochem 83:99–128. https://doi.org/10.1146/ annurev-biochem-060713-035600.

12. Islam ST, Lam JS. 2014. Synthesis of bacterial polysaccharides via the Wzx/Wzy-dependent pathway. Can J Microbiol 60:697–716. https://doi .org/10.1139/cjm-2014-0595.

13. Paulsen IT, Beness AM, Saier MH. 1997. Computer-based analyses of the protein constituents of transport systems catalysing export of complex carbohydrates in bacteria. Microbiology 143:2685–2699. https://doi.org/ 10.1099/00221287-143-8-2685.

14. Morona R, Van Den Bosch L, Daniels C. 2000. Evaluation of Wzz/MPA1/ MPA2 proteins based on the presence of coiled-coil regions. Microbiology 146:1–4. https://doi.org/10.1099/00221287-146-1-1.

15. Daniels C, Morona R. 1999. Analysis of *Shigella flexneri* Wzz (Rol) function by mutagenesis and cross-linking: Wzz is able to oligomerize. Mol Microbiol 34:181–194. https://doi.org/10.1046/j.1365-2958.1999.01591.x.

16. Klee SR, Tzschaschel BD, Timmis KN, Guzmán CA. 1997. Influence of different rol gene products on the chain length of *Shigella dysenteriae* type 1 lipopolysaccharide O antigen expressed by *Shigella flexneri* carrier strains. J Bacteriol 179:2421–2425. https://doi.org/10.1128/jb.179.7.2421 -2425.1997.

17. Burrows LL, Chow D, Lam JS. 1997. *Pseudomonas aeruginosa* B-band O-antigen chain length is modulated by Wzz (Rol). J Bacteriol 179: 1482–1489. https://doi.org/10.1128/jb.179.5.1482-1489.1997.

18. Batchelor RA, Alifano P, Biffali E, Hull SI, Hull RA. 1992. Nucleotide sequences of the genes regulating O-polysaccharide antigen chain length (rol) from *Escherichia coli* and *Salmonella typhimurium*: protein homology and functional complementation. J Bacteriol 174:5228–5236. https://doi.org/10.1128/jb.174.16.5228-5236.1992.

19. Stevenson G, Kessler A, Reeves P. 1995. A plasmid-borne O-antigen chain length determinant and its relationship to other chain length determinants. FEMS Microbiol Lett 125:23–30. https://doi.org/10.1111/j.1574 -6968.1995.tb07330.x.

20. Daniels C, Griffiths C, Cowles B, Lam JS. 2002. *Pseudomonas aeruginosa* O-antigen chain length is determined before ligation to lipid A core. Environ Microbiol 4:883–897. https://doi.org/10.1046/j.1462-2920.2002 .00288.x.

21. Tocilj A, Munger C, Proteau A, Morona R, Purins L, Ajamian E, Wagner J, Papadopoulos M, Van Den Bosch L, Rubinstein JL, Féthière J, Matte A, Cygler M. 2008. Bacterial polysaccharide co-polymerases share a com-

mon framework for control of polymer length. Nat Struct Mol Biol 15:130–138. https://doi.org/10.1038/nsmb.1374.

22. Kalynych S, Yao D, Magee J, Cygler M. 2012. Structural characterization of closely related O-antigen lipopolysaccharide (LPS) chain length regulators. J Biol Chem 287:15696–15705. https://doi.org/10.1074/jbc.M112.354837.

23. Kalynych S, Cherney M, Bostina M, Rouiller I, Cygler M. 2015. Quaternary structure of WzzB and WzzE polysaccharide copolymerases. Protein Sci 24:58–69. https://doi.org/10.1002/pro.2586.

24. Larue K, Kimber MS, Ford R, Whitfield C. 2009. Biochemical and structural analysis of bacterial O-antigen chain length regulator proteins reveals a conserved quaternary structure. J Biol Chem 284:7395–7403. https://doi.org/10.1074/jbc.M809068200.

25. Collins RF, Kargas V, Clarke BR, Siebert CA, Clare DK, Bond PJ, Whitfield C, Ford RC. 2017. Full-length, oligomeric structure of Wzz determined by cryoelectron microscopy reveals insights into membrane-bound states. Structure 25:806–815. https://doi.org/10.1016/j.str.2017.03.017.

26. Nath P, Morona R. 2015. Detection of Wzy/Wzz interaction in *Shigella flexneri*. Microbiology 161:1797–1805. https://doi.org/10.1099/mic.0.000132.

27. Islam ST, Huszczynski SM, Nugent T, Gold AC, Lam JS. 2013. Conserved-residue mutations in Wzy affect O-antigen polymerization and Wzz-mediated chain-length regulation in Pseudomonas aeruginosa PAO1. Sci Rep 3:3441. https://doi.org/10.1038/srep03441.

28. Taylor VL, Udaskin ML, Islam ST, Lam JS. 2013. The D3 bacteriophage α-polymerase inhibitor (Iap) peptide disrupts O-antigen biosynthesis through mimicry of the chain length regulator Wzz in *Pseudomonas aeruginosa*. J Bacteriol 195:4735–4741. https://doi.org/10.1128/JB.00903-13.

29. Marolda CL, Tatar LD, Alaimo C, Aebi M, Valvano MA. 2006. Interplay of the Wzx translocase and the corresponding polymerase and chain length regulator proteins in the translocation and periplasmic assembly of lipopolysaccharide O antigen. J Bacteriol 188:5124–5135. https://doi.org/10.1128/JB.00461-06.

30. Papadopoulos M, Morona R. 2010. Mutagenesis and chemical cross-linking suggest that Wzz dimer stability and oligomerization affect lipopolysaccharide O-antigen modal chain length control. J Bacteriol 192:3385–3393. https://doi.org/10.1128/JB.01134-09.

31. Tran ENH, Morona R. 2013. Residues located inside the *Escherichia coli* FepE protein oligomer are essential for lipopolysaccharide O-antigen modal chain length regulation. Microbiology 159:701–714. https://doi.org/10.1099/mic.0.065631-0.

32. Kalynych S, Ruan X, Valvano MA, Cygler M. 2011. Structure-guided investigation of lipopolysaccharide O-antigen chain length regulators reveals regions critical for modal length control. J Bacteriol 193:3710–3721. https://doi.org/10.1128/JB.00059-11.

33. Folkesson A, Jelsbak L, Yang L, Johansen HK, Ciofu O, Hoiby N, Molin S. 2012. Adaptation of *Pseudomonas aeruginosa* to the cystic fibrosis airway: an evolutionary perspective. Nat Rev Microbiol 10:841–851. https://doi.org/10.1038/nrmicro2907.

34. Lam JS, Taylor VL, Islam ST, Hao Y, Kocíncová D. 2011. Genetic and functional diversity of Pseudomonas aeruginosa lipopolysaccharide. Front Microbiol 1:118.

35. Bélanger M, Burrows LL, Lam JS. 1999. Functional analysis of genes responsible for the synthesis of the B-band O antigen of *Pseudomonas aeruginosa* serotype O6 lipopolysaccharide. Microbiology 145:3505–3521. https://doi.org/10.1099/00221287-145-12-3505.

36. Kintz EN, Goldberg JB. 2011. Site-directed mutagenesis reveals key residue for O antigen chain length regulation and protein stability in

Pseudomonas aeruginosa Wzz2. J Biol Chem 286:44277–44284. https://doi.org/10.1074/jbc.M111.273979.

37. Lam MYC, McGroarty EJ, Kropinski AM, MacDonald LA, Pedersen SS, Hoiby N, Lam JS. 1989. Occurrence of a common lipopolysaccharide antigen in standard and clinical strains of *Pseudomonas aeruginosa*. J Clin Microbiol 27:962–967.

38. Roy A, Kucukural A, Zhang Y. 2010. I-TASSER: a unified platform for automated protein structure and function prediction. Nat Protoc 5:725–738. https://doi.org/10.1038/nprot.2010.5.

39. Chen X, Zaro JL, Shen WC. 2013. Fusion protein linkers: property, design and functionality. Adv Drug Deliv Rev 65:1357–1369. https://doi.org/10.1016/j.addr.2012.09.039.

40. Mason JM, Arndt KM. 2004. Coiled coil domains: stability, specificity, and biological implications. Chembiochem 5:170–176. https://doi.org/10.1002/cbic.200300781.

41. Combet C, Blanchet C, Geourjon C, Deléage G. 2000. NPS@: network protein sequence analysis. Trends Biochem Sci 25:147–150. https://doi.org/10.1016/S0968-0004(99)01540-6.

42. Purins L, Van Den Bosch L, Richardson V, Morona R. 2008. Coiled-coil regions play a role in the function of the *Shigella flexneri* O-antigen chain length regulator WzzpHS2. Microbiology 154:1104–1116. https://doi.org/10.1099/mic.0.2007/014225-0.

43. Horton RM, Hunt HD, Ho SN, Pullen JK, Pease LR. 1989. Engineering hybrid genes without the use of restriction enzymes: gene splicing by overlap extension. Gene 77:61–68. https://doi.org/10.1016/0378-1119(89)90359-4.

44. King JD, Mulrooney EF, Vinogradov E, Kneidinger B, Mead K, Lam JS. 2008. *lfnA* from *Pseudomonas aeruginosa* O12 and *wbuX* from *Escherichia coli* O145 encode membrane-associated proteins and are required for expression of 2,6-dideoxy-2-acetamidino-ʟ-galactose in lipopolysaccharide O antigen. J Bacteriol 190:1671–1679. https://doi.org/10.1128/JB.01708-07.

45. García-Nafría J, Watson JF, Greger IH. 2016. IVA cloning: a single-tube universal cloning system exploiting bacterial in vivo assembly. Sci Rep 6:27459. https://doi.org/10.1038/srep27459.

46. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. J Mol Biol 215:403–410. https://doi.org/10.1016/S0022-2836(05)80360-2.

47. Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, Lopez R, McWilliam H, Remmert M, Soding J, Thompson JD, Higgins DG. 2011. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Mol Syst Biol 7:539. https://doi.org/10.1038/msb.2011.75.

48. Waterhouse AM, Procter JB, Martin DMA, Clamp M, Barton GJ. 2009. Jalview version 2—a multiple sequence alignment editor and analysis workbench. Bioinformatics 25:1189–1191. https://doi.org/10.1093/bioinformatics/btp033.

49. Robert X, Gouet P. 2014. Deciphering key features in protein structures with the new ENDscript server. Nucleic Acids Res 42:W320–W324. https://doi.org/10.1093/nar/gku316.

50. Hitchcock PJ, Brown TM. 1983. Morphological heterogeneity among *Salmonella* lipopolysaccharide chemotypes in silver-stained polyacrylamide gels. J Bacteriol 154:269–277.

51. Fomsgaard A, Freudenberg MA, Galanos C. 1990. Modification of the silver staining technique to detect lipopolysaccharide in polyacrylamide gels. J Clin Microbiol 28:2627–2631.