SCIENTIFIC REPORTS

OPEN

SUBJECT AREAS:
MEMBRANE PROTEINS
GLYCOBIOLOGY
BACTERIAL PATHOGENESIS
LIPOPOLYSACCHARIDES

Received
29 April 2013

Accepted
20 November 2013

Published
6 December 2013

Correspondence and requests for materials should be addressed to J.S.L. (jlam@uoguelph.ca)

Conserved-residue mutations in Wzy affect O-antigen polymerization and Wzz-mediated chain-length regulation in *Pseudomonas aeruginosa* PAO1

Salim T. Islam[1], Steven M. Huszczynski[1], Timothy Nugent[2], Alexander C. Gold[1] & Joseph S. Lam[1]

[1]Department of Molecular and Cellular Biology, University of Guelph, Guelph, Ontario, N1G 2W1, Canada, [2]Bioinformatics Group, Department of Computer Science, University College London, London, WC1E 6BT, United Kingdom.

O antigen (O-Ag) in many bacteria is synthesized via the Wzx/Wzy-dependent pathway in which Wzy polymerizes lipid-linked O-Ag subunits to modal lengths regulated by Wzz. Characterization of 83 site-directed mutants of Wzy from *Pseudomonas aeruginosa* PAO1 (Wzy_{Pa}) in topologically-mapped periplasmic (PL) and cytoplasmic loops (CL) verified the functional importance of PL3 and PL5, with the former shown to require overall cationic properties. Essential Arg residues in the $RX_{10}G$ motifs of PL3 and PL5 were found to be conserved in putative homologues of Wzy_{Pa}, as was the overall sequence homology between these two periplasmic loops in each protein. Amino acid substitutions in CL6 were found to alter Wzz-mediated O-antigen modality, with evidence suggesting that these changes may perturb the C-terminal Wzy_{Pa} tertiary structure. Together, these data suggest that the catch-and-release mechanism of O-Ag polymerization is widespread among bacteria and that regulation of polymer length is affected by interaction of Wzz with Wzy.

Pseudomonas aeruginosa is a ubiquitous Gram-negative bacterium and an opportunistic pathogen that can infect compromised individuals suffering from various conditions including cancer, AIDS, severe burn wounds, post-operative wounds, and cystic fibrosis[1]. Lipopolysaccharide (LPS) in the outer membrane of the bacterium is important for its persistence in the environment as well as its virulence in infection settings, affecting many bacterial and host phenotypes[2].

LPS is anchored in the outer membrane by lipid A, which is linked through the core oligosaccharide to the O-antigen (O-Ag) capping motif[2]. *P. aeruginosa* PAO1 simultaneously synthesizes two forms of O-Ag. The first is a common polysaccharide antigen (CPA, A-band) composed of repeating D-rhamnose sugars. The second is composed of a repeating trisaccharide unit containing a proximal D-fucosamine sugar and two distal substituted dideoxy-mannuronic acid sugars termed O-specific antigen (OSA, B-band). B-band O-Ag is the immunodominant cell-surface antigen of *P. aeruginosa*, with differences in sugar subunit composition and/or linkage stereospecificity leading to the identification of 20 distinct immunological serotypes of the bacterium[2].

B-band O-Ag biosynthesis is mediated by the Wzx/Wzy-dependent pathway[3] (Supplementary Fig. S1), requiring the activities of a series of integral inner membrane (IM) proteins[4,5] that have homologues involved in the biosynthesis of other cell-surface and secreted bacterial polysaccharides including enterobacterial common antigen (ECA), capsule polysaccharide (CPS), and exopolysaccharide (EPS)[6]. In *P. aeruginosa* PAO1, once the anionic trisaccharide B-band O-Ag repeat units have been synthesized in the cytoplasm on the lipid carrier undecaprenyl pyrophosphate (UndPP), they are translocated across the hydrophobic IM through the cationic lumen of Wzx to the periplasmic leaflet[7–9] via a putative antiport mechanism[10]. Transported O-units are subsequently polymerized by Wzy[11,12], with new O-units added at the reducing terminus of the growing chain[13]. The length of O-unit polymerization is regulated by polysaccharide co-polymerase (PCP) proteins[14], with distinct modal lengths in *P. aeruginosa* PAO1 of 12–16 and 22–30 units (termed "long") specified by Wzz_1[15], and 40–50 units (termed "very long") specified by Wzz_2[16]. Polymerized O-Ag is then ligated to lipid A-core oligosaccharide by WaaL to form a complete LPS molecule[17,18], which is transported to the outer membrane via the Lpt pathway[19].

Evidence to elucidate the mechanism of O-Ag polymerization was lacking until a recent investigation in which our group mapped the IM topologies of Wzx, Wzy, and WaaL, revealing the presence of two large periplasmic

loops (PL3 and PL5) in Wzy of *P. aeruginosa* PAO1 (Wzy_{Pa})[4]. Despite a high level of sequence homology between them, PL3 and PL5 were found to possess distinct cationic and anionic characteristics, respectively. In addition, $RX_{10}G$ motifs were localized to the middle of each PL and shown to contain functionally-important Arg residues that confer a combination of charge and stereochemistry-specific properties. These observations led us to propose a "catch-and-release" mechanism for Wzy_{Pa}-mediated O-Ag polymerization (Supplementary Fig. S2), wherein both PL3 and PL5 were in essence binding the same substrate, i.e. an UndPP-linked O unit, via their respective $RX_{10}G$ motifs, but in different capacities. PL3 was suggested to be the recruitment arm for a new UndPP-linked O unit, while PL5 was proposed as the retention arm of the growing chain composed of single O units polymerized upon themselves. For Wzy_{Pa}, the cationic properties of PL3 would contribute to the binding of an incoming anionic O unit, while the net anionic character of PL5 would be required to facilitate the more transitory interactions required at this site[20]. Both the identification of two PLs of comparable sequence and the presence of key Arg residues therein would provide support for this mechanism in other Wzy proteins.

Furthermore, the proteins in the Wzx/Wzy-dependent assembly pathway have long been proposed to form an IM complex[3]; however, only evidence from cross-complementation experiments exists to indirectly support these proposed interactions[21]. These findings were based on the ability of *wzx* from an *Escherichia coli* ECA pathway to complement a *wzx* deficiency in the O-Ag pathway in the same organism (based on relaxed substrate specificities between the two flippases). However, O-Ag biosynthesis could only be restored by the ECA *wzx* when the ECA *wzy* and *wzz* genes were deleted, suggesting preferred interaction of Wzx with the corresponding Wzy and Wzz from the same pathway[21].

Systematic site-directed mutagenesis of 83 periplasmic and cytoplasmic residue positions spanning the length of the protein was performed to identify residues of functional importance for the polymerization of O-Ag, as assayed by the ability of a construct to restore B-band LPS biosynthesis in a *wzy* chromosomal mutant[12]. Importantly, the periplasmic amino acids shown to be important for Wzy_{Pa} function remained confined to PL3 and PL5, with the former also exhibiting positive-charge dependence; moreover, functionally-essential Arg residues in the $RX_{10}G$ motifs of these loops were found to be conserved in putative Wzy proteins identified in a range of phylogenetically distinct bacteria. Unexpectedly, site-specific mutations in cytoplasmic loop (CL) 6 were found to disrupt Wzy_{Pa}-mediated O-Ag polymerization in the periplasm. B-band O-Ag from cells of a *wzy* chromosomal knockout mutant complemented with these CL6 mutants displayed atypical chain-length modality skewed towards Wzz_2-regulated "very long" polymer lengths (despite the presence of a functional Wzz_1). These results suggest a widespread occurrence of a catch-and-release mechanism for Wzy function in bacteria and provide the first evidence for the interaction of full-length Wzy and Wzz in a homologously-expressed system.

Results

PL3 and PL5 are the important periplasmic sites of Wzy. Based on the experimentally-derived topology map of Wzy_{Pa}[4] (Fig. 1), an initial mutagenesis characterization study had been focused on charged residues within the $RX_{10}G$ tracts shared between PL3 and PL5[20] (Supplementary Fig. S3). In the current investigation, we set out to identify functionally important residues in positions outside of the Wzy_{Pa} $RX_{10}G$ motifs, including the flanking regions of PL3 and PL5 as well as the other identified periplasmic and cytoplasmic loop domains. In total, 89 site-specific substitutions were introduced across 83 distinct loop-exposed positions in Wzy_{Pa}, including all aromatic and charged residues within these domains (Fig. 1). The *P. aeruginosa* PAO1 *wzy* chromosomal mutant strain[12] was supplied *in trans* with wild-type and mutant gene copies, followed by analysis of the respective B-band O-Ag phenotypes via Western immunoblot to examine the ability of a given mutant construct to complement the assembly deficiency. Despite the sizeable lengths and varied amino acid compositions of PL1 and PL2, no residues of functional importance were identified therein. Moreover, the long stretches of amino acids flanking the $RX_{10}G$ motifs in PL3 and PL5 (Supplementary Fig. S3) were found to lack functionally-important residues. An exception was the anionic residue D286 in PL5, for which a

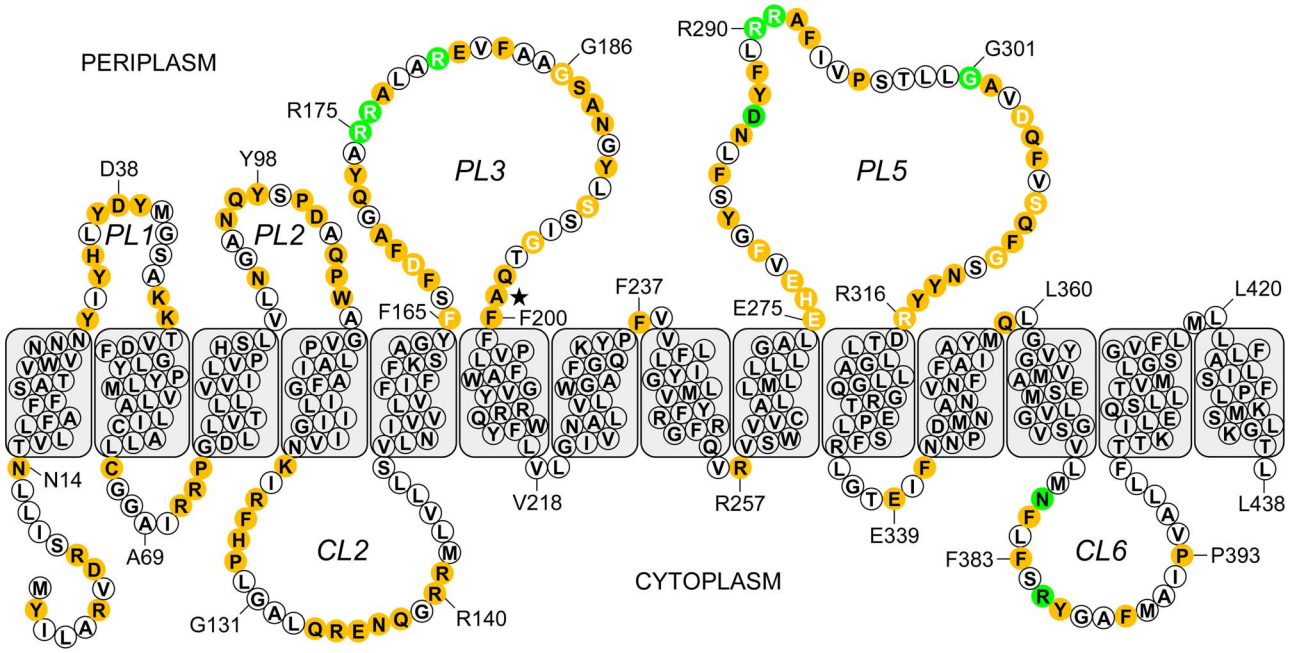

Figure 1 | Wzy_{Pa} topology map indicating residue substitution positions and complementation phenotypes. Highlighted residue colours: *orange*, no defect in O-Ag polymerization; *green*, visible defect in O-Ag polymerization. Text in highlighted residues: *black*, characterized in the current investigation; *white*, previously characterized[20]. ★, residue A199 (site of cationic-to-anionic PL3 charge alterations). $RX_{10}G$ motifs in PL3 and PL5 have been indicated from R175–G186 and R290–G301, respectively.

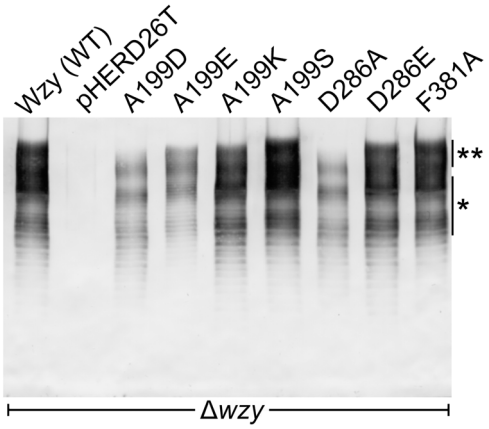

Figure 2 | Western immunoblot analysis of LPS from *P. aeruginosa* PAO1 Δ*wzy* complemented with periplasmic domain Wzy_{Pa} mutants. B-band O-Ag was detected with monoclonal antibody MF15-4[59]. *, "long" heteropolymeric O-Ag bands regulated by Wzz_1; **, "very long" heteropolymeric O-Ag bands regulated by Wzz_2.

structurally-innocuous Ala mutation (Supplementary Fig. S4) resulted in a Wzy_{Pa} variant that could only partially restore B-band O-Ag biosynthesis when used in the complementation assay. This functional defect could be fully complemented by the like-charge substitution D286E, resulting in a Wzy_{Pa} variant that retained WT activity (Fig. 2).

Effects of PL3 charge alterations are consistent with cationic charge-dependence of this loop. Through recruiting the newly-flipped anionic O-Ag subunits polymerized by Wzy_{Pa}, the overall positive charge of PL3 was proposed to play an important role in the catch-and-release mechanism[20] (Supplementary Fig. S2). To further examine the charge dependence of PL3, substitutions were introduced at the base of the loop in order to alter its overall charge properties (Table 1). Position 199 was selected for several reasons. Firstly, it is distant from the previously characterized PL3 $RX_{10}G$ functional site (Fig. 1)[20]. Secondly, non-native residues introduced at this position were not found to alter the predicted tertiary structure of the PL3 peptide (Supplementary Fig. S5). Thirdly, alteration of any Ala-specific stereochemistry by substitution with Ser did not affect function of the protein (Fig. 2). Amino acid substitutions A199D and A199E that altered PL3 charge from net-positive to net-negative (Table 1) diminished Wzy_{Pa} activity (Fig. 2) whereas maintenance of net-positive PL3 charge via substitutions A199S and A199K (Table 1) did not compromise Wzy_{Pa}-mediated polymerization (Fig. 2).

Table 1 | Isoelectric point (pI) values for PL3 mutant peptides

Mutation	Peptide pI	Peptide Charge at Physiological pH	Compromised Function
WT	8.59	+	–
D168A*	9.98	+	no
R175A*	6.07	–	yes
R175K*	8.50	+	no
R180A*	6.07	–	yes
R180K*	8.50	+	no
E181A	9.98	+	no
A199D	6.12	–	yes
A199E	6.18	–	yes
A199K	9.69	+	no
A199S	8.59	+	no

*Characterized in a previous investigation[20].

Putative homologues possess both PL3–PL5-like sequence homology and conserved amino acids essential for Wzy_{Pa} function. To test the hypothesis that characteristics required for the catch-and-release concept could be found in other bacteria using Wzx/Wzy-dependent assembly systems, we initially performed standard BLAST-based protein database searches, but could not uncover any notable hits using the Wzy_{Pa} amino acid sequence as a query. Consequently, we employed a *jackhmmer* search[22] to exploit the increased accuracy and ability of probabilistic inference-based HMMER methods to detect homologues[23,24] from other bacterial genera; this involves building and searching with a global sequence "profile" rather than relying on position-specific residue scores. This approach yielded 30 sequence hits (Supplementary Table S1) for putative homologues from a variety of bacterial families, classes, and phyla (Supplementary Table S2).

Multiple sequence alignment (MSA) output from the *jackhmmer* search was used to generate sequence logos to visually depict amino acid consensus and frequency at each position[25]. This analysis revealed that residues R176 (PL3), R290, and R291 (PL5) were highly conserved (Fig. 3). These three Arg residues were previously shown to be functionally-essential for Wzy_{Pa}, and could not be functionally-substituted with Lys[20]. Together, these data indicated the specific importance of the Arg guanidinium functional group at each of these positions[20]. Given the conserved positioning of these essential residues, the positions of PL3 and PL5 sequence termini from Wzy_{Pa} in the MSA were used as constraints to identify analogous loop sequences in predicted homologues. As seen with PL3 and PL5 from Wzy_{Pa}[20] (Supplementary Fig. S3), the analogous sequences for the various predicted homologues displayed sequence homology to each other (Supplementary Fig. S6). These data support the widespread occurrence of traits required for a catch-and-release mechanism[20] of O-Ag polymerization in other bacteria.

Cytoplasmic amino acid substitutions affect periplasmic O-Ag modality determination. Extensive amino acid substitutions were generated for the cytoplasmic domains of Wzy_{Pa}, with particular emphasis on CL2 and CL6 (Fig. 1). Mutations in CL2 did not display any deleterious effects on Wzy_{Pa} function. In contrast, multiple site-specific substitutions in CL6 were found to affect the B-band O-Ag phenotype (Fig. 4). Mutation N380A and R385A caused decreases in the amount of O-Ag synthesized. More importantly, B-band O-Ag prepared from a *P. aeruginosa* PAO1 *wzy* chromosomal mutant complemented with the N380A and R385A constructs showed a change in the preferred modal length of the polymer. In both cases, high-molecular weight LPS bands corresponding to lengths regulated by the Wzz_2 PCP continued to be made; however, a slight reduction (R385A) or outright loss (N380A) of low-molecular weight LPS bands attributed to Wzz_1-mediated regulation was reproducibly observed (Fig. 4).

To ensure that this shifted-modality phenotype was not simply a result of low gene dosage from the uninduced background expression levels typically used for complementation, the effect of these mutations was compared in the presence and absence of expression induction via addition of 0.1% L-arabinose. For both N380A and R385A, increased expression of the respective Wzy_{Pa} mutant construct resulted in the production of increased amounts of heteropolymeric O-Ag (Supplementary Fig. S7). The initial detection of a minor O-Ag modality shift with the R385A variant was greatly facilitated by the reduced amount of O-Ag produced in the uninduced samples (Fig. 4); as such, overexpression of the construct and the concomitant synthesis of more B-band O-Ag masked any slight modality shifts. For N380A, the drastic nature of the modality shift was reproducible despite the increased expression levels of the construct and the higher amounts of B-band O-Ag produced. Of particular note was the highest-molecular weight band visible in both the uninduced and induced N380A lanes, which became even more prominent in the

latter treatment and was not detected in any of the other samples analyzed (Supplementary Fig. S7). Together, these data indicate an effect of Wzy mutations on the downstream LPS assembly process of chain-length regulation by Wzz proteins.

The novelty of the results from complementation assays with N380A and R385A stems from their cytoplasmic localization in the Wzy_{Pa} topological map; therefore, we sought to provide additional evidence for the cytoplasmic localization of the two residues. Since we have identified a number of putative Wzy_{Pa} homologues, we used the MSA data as input for the FILM3 tool[26] to generate a *de novo* tertiary structure model for Wzy_{Pa}. This approach has recently been demonstrated to utilize correlated rates of amino acid mutations from a MSA to infer residue-residue contacts and successfully generate a representative tertiary structure for a membrane protein sequence of interest[10,26,27]. Due to the minimal number of input sequences available in our MSA, the lowest-energy Wzy_{Pa} model generated had an ambiguous TM-score of 0.42 (as >0.5 indicates a fair model with probably correct fold, while <0.4 indicates a poor model with probably incorrect fold). As such, we were unable to confidently report on the overall TMS packing arrangement in the *de novo* model. However, two important observations were made. Firstly, even with a limited number of input sequences in the MSA, both N380 and R385 were still localized to the same C-terminal CL flanked by helical TMS bundles (Supplementary Fig. S8). Secondly, these data are consistent with the experimentally-derived Wzy_{Pa} topology map (Fig. 1), providing firm support for the cytoplasmic localization of N380 and R385.

Substitutions N380A and R385A destabilize the C-terminus of Wzy_{Pa}. To interpret the N380A- and R385A-mediated changes in O-Ag modality, we hypothesized that the conformation and/or stability of the C-terminal region of Wzy_{Pa} could become disrupted through altered TMS packing caused by the mutations. To explore this further, the potential TMS packing arrangements for wild type (WT) and mutant Wzy_{Pa} variants were analyzed using TM*hit* to predict inter-TMS interactions[28]. As TM*hit* is a support vector machine classification approach, the over-fitting of data was minimized[29]. When compared to WT Wzy_{Pa}, the N380A variant displayed an altered number of total contacts for TMS1, 7, 10, 12, and 14, with reintroduction of a similarly polar side chain (N380Q) able to restore the sums to WT levels (Supplementary Fig. S9). In addition to the change in overall contacts per TMS compared to WT, the N380A mutation also altered the profile for which TMS is predicted to interact with another, a trait that was not observed with the N380Q mutation (Fig. 5). Conversely, changes in inter-TMS interaction profiles (Fig. 5) and the sum of contacts per

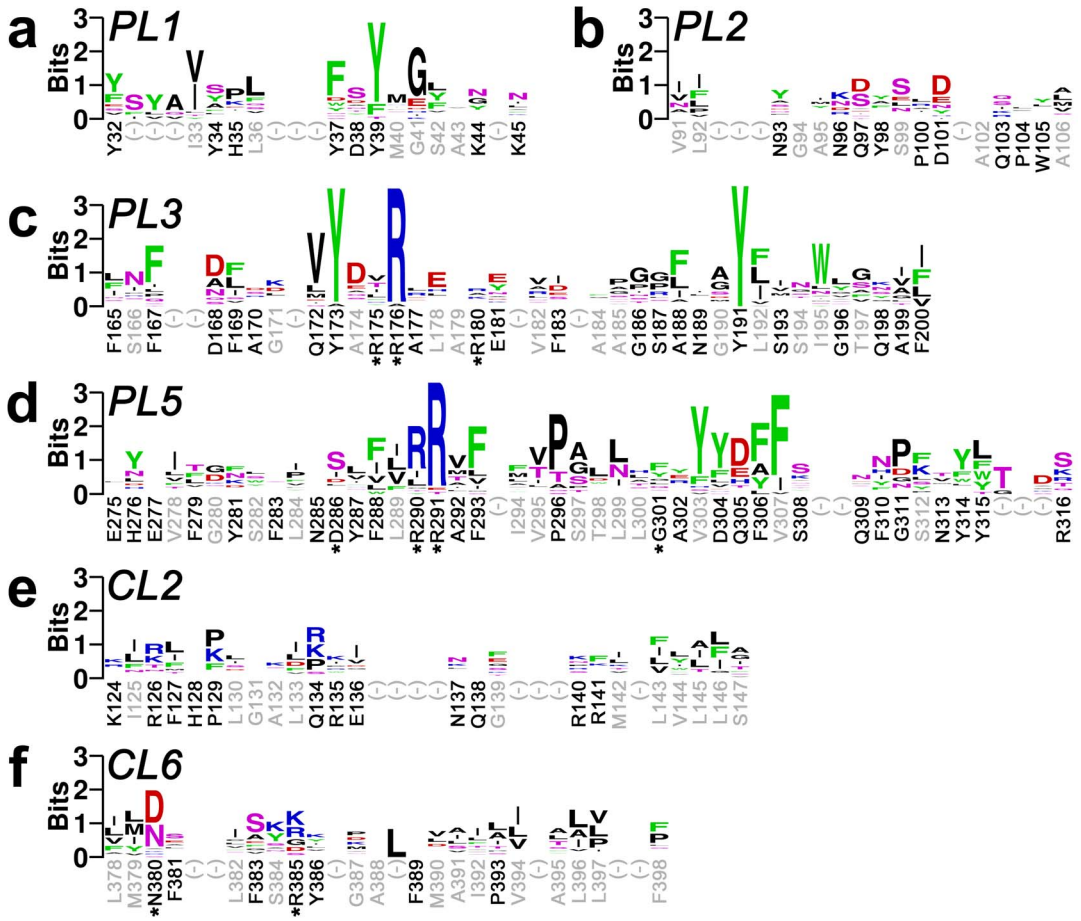

Figure 3 | Periplasmic and cytoplasmic loop sequence logos of residue conservation between Wzy_{Pa} and aligned sequences. The MSA used to generate the sequence logos was obtained from the results of the *jackhmmer* search (Table S1). Logo letter height is proportional to its frequency at a given position in the MSA, with the most common letter displayed on top[25]. Logo residue colours: *blue*, positively-charged (Arg, Lys, His); *red*, negatively-charged (Asp, Glu); *green*, aromatic (Tyr, Phe, Trp); *magenta*, polar (Gln, Ser, Thr). Amino acid identities and numbers along the *x*-axes correspond to the native Wzy_{Pa} sequence, with "(−)" representing a gap in the MSA at specific positions of the Wzy_{Pa} primary structure. *X*-axis legend: *black*, screened for functional importance via site-directed mutagenesis and LPS phenotype complementation; *grey*, has not been screened for functional importance; *, residue with demonstrated functional defect. Conservation of Wzy_{Pa} amino acid sequences are displayed for (a) PL1, (b) PL2, (c) PL3, (d) PL5, (e) CL2, and (f) CL6.

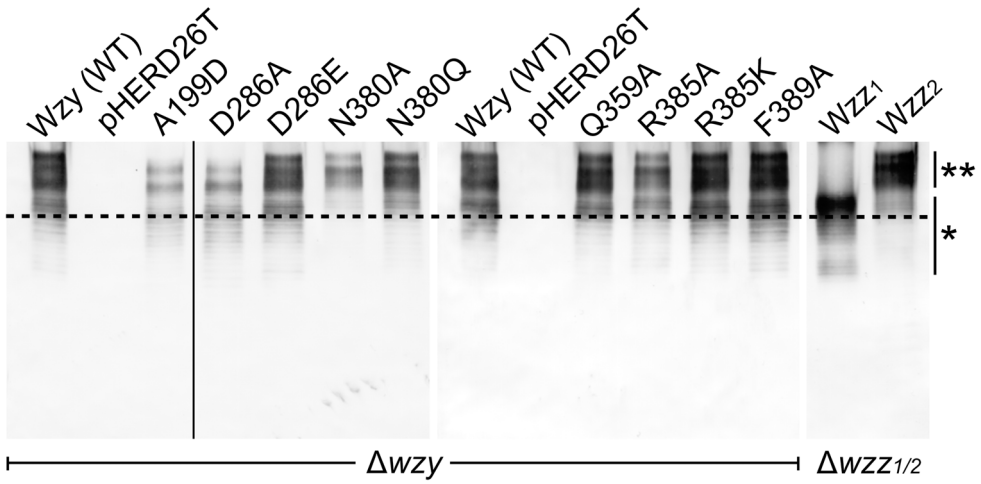

Figure 4 | Western immunoblot analysis of LPS from *P. aeruginosa* PAO1 Δ*wzy* complemented with cytoplasmic domain Wzy_{Pa} mutants. B-band O-Ag was detected with monoclonal antibody MF15-4[59]. O-Ag from the *wzz*$_1$-*wzz*$_2$ double mutant specifically complemented with constructs for either *wzz*$_1$ or *wzz*$_2$ was analyzed to provide a reference for the respective O-Ag modalities conferred by each protein. The horizontal line has been drawn under the most prevalent Wzz_1-specified modality as a reference for shifts in modality caused by mutations in Wzy_{Pa}. All displayed samples were resolved via SDS-PAGE and Western immunoblotted at the same time. Sample lanes from distinct blots are separated by white space. Samples adjoining the vertical black line were separated by one lane on the same blot. *, "long" heteropolymeric O-Ag bands regulated by Wzz_1; **, "very long" heteropolymeric O-Ag bands regulated by Wzz_2.

TMS (Supplementary Fig. S9) were not observed with the R385A substitution, suggesting a more innocuous Wzy_{Pa} structural alteration caused by substitutions at this position. The notion that mutation R385A caused less TMS perturbation than N380A was supported by (i) the higher amounts of B-band O-Ag produced by the former compared to the latter (Supplementary Fig. S7), as well as (ii) only a slight shift of the LPS bands to longer modal lengths for R385A versus the complete loss of shorter-length polymers corresponding to Wzz_1-regulated chain lengths for N380A (Fig. 4). As Wzz_2-mediated regulation was unaffected (particularly in the latter substitution), this indicates that Wzz_1 and Wzz_2 do not engage with Wzy_{Pa} in an identical manner. Moreover, these data support an interaction between Wzy_{Pa} and Wzz_1 that would be largely mediated by the C-terminal TMS of the former.

In addition, we analyzed the propensities of the various Wzy_{Pa} mutant constructs to insert into the membrane. Constructs encoding WT Wzy_{Pa} as well as functionally-compromised mutant variants

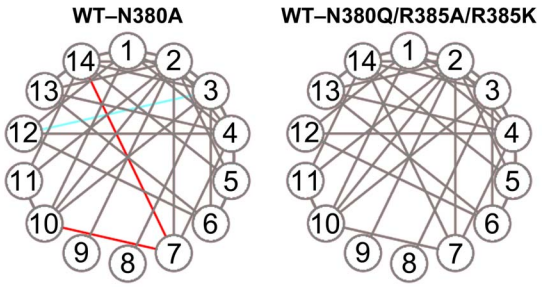

Figure 5 | Inter-TMS contact web predictions for WT and mutant Wzy_{Pa}. Predicted contacts were obtained through analysis of WT and mutant Wzy_{Pa} amino acid sequences with TM*hit*[28], constrained by established TMS boundaries (Fig. 1). The Wzy_{Pa} TMS have been labelled (1–14). Lines connecting TMS are indicative of at least one predicted pair of interacting residues between the TMS involved. The outputs for WT and substituted Wzy_{Pa} variants were overlaid. Colour scheme: *purple*, interactions conserved between WT and mutants; *red*, interactions exclusive to WT; *cyan*, interactions exclusive to a mutant. WT comparison to N380A yielded distinct interaction profiles, while comparison to N380Q, R385A, or R385K did not reveal any differences.

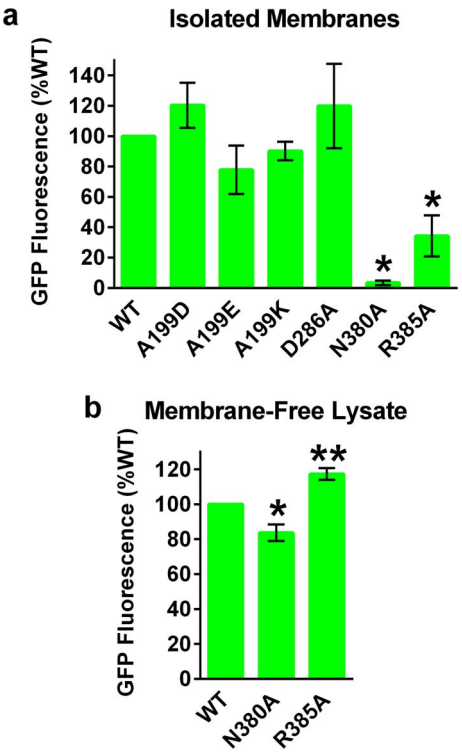

Figure 6 | GFP fluorescence data. (a) Densitometry analysis of monomeric Wzy-GFP-His_8 fluorescence from membrane samples isolated from *P. aeruginosa* PAO1 Δ*wzy* expressing WT and mutant constructs. (b) Densitometry analysis of cleaved GFP-His_8 fluorescence from membrane-free lysates of *P. aeruginosa* PAO1 Δ*wzy* expressing WT and mutant constructs. All samples were analyzed in triplicate ($n = 3$) and displayed ± standard error. Statistical significance ($p \leq 0.05$) was calculated using the Student's *t*-test. *, statistically significant reduction compared to WT. **, statistically significant increase compared to WT. Panel "a": WT–N380A ($p < 0.0001$); WT–R385A ($p = 0.0083$). Panel "b": WT–N380A ($p = 0.0256$); WT–R385A ($p = 0.0068$).

identified in this investigation (A199D, A199E, D286A, N380A, R385A) were first overexpressed with a C-terminal GFP-His_8 tag in the *P. aeruginosa wzy* chromosomal knockout background. Membranes from these samples were isolated and resolved via SDS-PAGE, followed by in-gel fluorescence analysis. Compared to the WT construct, Wzy_{Pa} variants A199D, A199E, and D286A were maintained at equivalent levels in the membrane as detected via C-terminal GFP fusion fluorescence (Fig. 6A), indicating comparable folding and stability of the complete fusion protein. Conversely, mutants N380A and R385A displayed significantly less fluorescence from the encoded C-terminal GFP tag (Fig. 6A).

However, despite the reduced amount of GFP fluorescence in the membrane for N380A and R385A, both mutants were still able to polymerize O-Ag subunits (Fig. 4, Supplementary Fig. S7), albeit at lower efficiency compared to the WT, similar to variants A199D, A199E, and D286A (Fig. 2). As Wzy_{Pa} is an integral IM protein that acts on membrane-embedded substrates, this would indicate that the functional "core" of the protein was still membrane embedded for mutants N380A and R385A. Moreover, the drastic shift in preferred O-Ag modality was still maintained even with induced overexpression of the construct (Supplementary Fig. S7). For these reasons, the C-terminal GFP tag may have become more prone to cleavage due to the disrupted TMS packing described above. Consequently, we compared the levels of GFP fluorescence in the clarified lysates from above (lacking the membrane fraction). Lysates from cells expressing the R385A construct displayed significantly more GFP fluorescence than the WT (Fig. 6b), consistent with an intact but more unstable C-terminus prone to cleavage in this mutant construct. Conversely, N380A displayed less GFP fluorescence than the WT (Fig. 6b). We propose that the N380A substitution could result in misfolding of the C-terminus, possibly affecting lateral exit of the remaining domains into the IM from the Sec translocon. This would affect the ability of GFP to properly fold and/or reduce the overall amount of fusion protein being translated. These results for N380A and R385A support the interpretation that the C-terminus (particularly TMS 13 and 14) of Wzy_{Pa} is disrupted in each of these two mutant variants.

Discussion

Prior to our characterization of Wzy_{Pa}[4], topological mapping of this class of proteins via experimentation had only been carried out on the O-Ag polymerase from *Shigella flexneri* (Wzy_{Sf})[30], as well as on the putative EPS polymerase PssT from *Rhizobium leguminosarum* bv. *trifolii*[31]. In both of these earlier studies, only one sizeable PL domain had been proposed[30,31] for Wzy_{Sf} and PssT, whereas two major PL domains were identified in Wzy_{Pa}[4] (Fig. 1); this major difference was likely a result of the methodologies used in the former two investigations compared to that in the latter[32]. Consequently, modifications to the PssT topology map have recently been proposed to bring it more in line with that of Wzy_{Pa}[33].

Identification of sequence homology between PL3 and PL5 in Wzy_{Pa} has led to the discovery of $RX_{10}G$ amino acid tracts in both loops that were shown to contain essential Arg residues[20] (Supplementary Fig. S1). The R290 of MagA (proposed CPS polymerase) from *Klebsiella pneumoniae* was also present in the middle of a large predicted loop and found to be important for function; however, despite the use of three *in silico* topology prediction algorithms in this study, a consensus number of TMS for MagA was not obtained[34], and as such the periplasmic or cytoplasmic localization of this residue cannot be inferred[32]. As Arg is one of the two most commonly observed amino acids in the binding sites of sugar- and carbohydrate-binding proteins[35,36], its importance in PL3 and PL5 supports a role in O-Ag-binding for both these loops within the context of a catch-and-release mechanism of Wzy polymerase function[20]. Through the use of more sensitive and accurate bioinformatic search techniques that facilitate detection of remote homologues[22–24], we have now identified conserved Arg residues in a diverse phylogenetic range of bacteria (Fig. 3). Interestingly, these Arg residues are in positions that correspond to each of the functionally-essential R176, R290, and R291 side chains in Wzy_{Pa}[20]. Moreover, the putative loops in the homologues (equivalent to PL3 and PL5 in Wzy_{Pa}) in which these Arg residues reside also share considerable sequence homology with one-another in the same protein (Supplementary Fig. S4). The characteristic properties of these periplasmic loops represent two key requirements of the catch-and-release mechanism: (i) the existence of essential Arg residues in both loops to mediate binding of UndPP-linked sugar repeats, and (ii) the presence of two loops displaying comparable amino acid sequences, as one is for recruitment of a single UndPP-linked O-unit, while the other is for retention of the growing chain, which is composed of the same O-unit simply extended upon itself[20].

The observation that functionally-essential periplasmic residues are localized almost exclusively to the $RX_{10}G$ motifs in Wzy_{Pa} (Fig. 1) is in line with the overall lack of sequence conservation in these regions (Fig. 3). By extension, evolutionary retention of the catalytic Arg residues, within the context of organism-specific PL3 and PL5 flanking sequences, is consistent with expected differences in substrate specificity between bacteria conferred by their respective two PL domains; these would have evolved to handle particular chemical structures of organism-specific O-Ag repeat units. At present, a major limitation to substrate specificity characterization in these putative homologues is that the O-Ag compositions and/or structures from the respective bacterial strains have yet to be elucidated. Despite extensive searching, we were only able to find an O-Ag chemical structure[37] for one of the strains (*Photorhabdus luminescens* subsp. *laumondii* TT01) identified in the list of non-redundant Wzy_{Pa} sequence hits; unfortunately, no chemical or topological similarities could be found between this particular O-Ag structure and that of the B-band O-Ag of *P. aeruginosa* PAO1.

To handle the net-negative charge of the *P. aeruginosa* PAO1 B-band O-Ag subunit, the interior of the respective Wzx flippase has been found to be cationic in nature[8]. It is logical that the periplasmic Wzy_{Pa} domain responsible for recruitment of the newly-flipped O-unit would also display this characteristic, for without this initial "catching" step, subsequent polymerization would not occur[20]. The positive-charge dependence of PL3 that we have revealed in this investigation (Supplementary Fig. S3) further strengthens the proposed role of this periplasmic domain in initial binding of UndPP-linked O-unit by the polymerase. The partial loss of Wzy_{Pa} function observed for the D286A substitution is consistent with a decrease in net-negative charge affecting the role of PL5 in the proposed catch-and-release mechanism, as it would alter the affinity of the retention arm for the growing O-Ag chain (bound via the PL5 $RX_{10}G$ motif). However, a catalytic role cannot be ruled out for this residue, as Asp at this position was identified in a subset of the aligned putative Wzy homologues. In the other sequence hits, a conserved Ser residue was present at this position. Either residue can be commonly found in the active site of glycosyltransferases, and may be involved in proton abstraction from the donor substrate to catalyze the formation of a new glycosidic bond[38,39].

The presence of functionally-important amino acids on the cytoplasmic face of the protein was unexpected given the periplasmic nature of O-Ag polymerization and the lack of requirement for cytoplasmic factors such as ATP as an energy source for the protein[11]. It is for this reason that the identification of cytoplasmic Wzy_{Pa} residues (in the same loop) that affect O-Ag polymerization is particularly intriguing. Consistent with their importance, both N380 and R385 were found to be conserved amongst the other identified putative Wzy proteins (Fig. 3). The novelty of observing a reduction in the amount of B-band O-Ag produced (for N380A and R385A) was only surpassed by the surprising finding that the remaining O-Ag was primarily composed of higher-molecular-weight polysaccharide chains displaying Wzz_2-regulated modal lengths (Fig. 4). Despite

the same level of Wzz_1 expression as when all other Wzy_{Pa} mutant constructs were screened, the chain-length regulator had lost its ability to regulate the polymerization of shorter O-Ag chain lengths by Wzy_{Pa}. This would support different interactions coordinating functions of Wzy and Wzz_2 compared to Wzy and Wzz_1.

While using the same precursor substrates (UndPP-linked O units) and the same enzyme (Wzy), Wzz_1 and Wzz_2 display genetic and functional diversity. Typically, *wzz*$_1$ in *P. aeruginosa* and its homologues in other bacteria that produce long O-Ag chains are found in the respective O-Ag biosynthesis gene clusters along with their cognate Wzy proteins. Conversely, *wzz*$_2$ and its homologues that produce very-long O-Ag chains are located elsewhere in the chromosome or on plasmids[16,40–42]. Moreover, long and very-long O-Ag chains in the same bacterium have been shown to be play different roles in infection, virulence, and persistence[40,43–45].

It is important to note that in order for the bacterial cell to elaborate and display any form of polymeric B-band O-Ag on its surface, regardless of chain length, a functional Wzy protein must be present in the inner membrane[12]. Consistent with this, we detected its reaction products with B-band specific monoclonal antibodies[46]. As presented in a previous investigation[20], mutations that affect function of Wzy may simply reduce the functionality of the protein, resulting in an overall reduction in the amount of O-Ag polymerized. Such is the case for the various A199D and A199E substitutions (Fig. 2), as well as the N380A and R385A variants (Fig. 4) characterized herein. However, in most instances of a partial reduction in Wzy_{Pa} functionality from specific residue substitutions, decreased enzyme efficiency resulted in an overall shift in bands towards the lower molecular weight Wzz_1-regulated bands. The only exceptions identified thus far are the N380A and R385A mutants from the same cytoplasmic loop that, despite producing decreased amounts of O-Ag, resulted in chain modality that was shifted away from Wzz_1-regulated lengths (and conversely towards higher lengths regulated by Wzz_2). For these very reasons, if the N380A and R385A mutations significantly contributed to reduced steady-state levels of the proteins in the membrane, O-Ag bands from these samples should still have been shifted downwards towards Wzz_1-regulated lengths as there would simply not have been enough "functional" Wzy present to sustain polymerization to very long Wzz_2-regulated lengths. However, this was not observed. Furthermore, considering that the modality shift away from Wzz_1-regulated lengths was still maintained following induction of expression in particular for the N380A samples, this provides further evidence that steady-state levels of these mutant proteins did not play a significant role in the observed shifted-modality phenotypes.

It is more plausible that the effect on B-band O-Ag biosynthesis caused by substitutions N380A and R385A is a result of tertiary structure disruptions in the C-terminal half of Wzy_{Pa}, as supported by the complementary lines of evidence provided in this investigation (Figs. 3–6, Supplementary Figs. S7–S9). At a minimum, Wzy and Wzz are believed to be localized within close proximity to one another[14]. Since the observed effect of these mutations on Wzy_{Pa} polymerization was to alter the modality conferred by Wzz proteins, this would suggest that a specific interaction between Wzy and Wzz_1 was being affected. This is further supported by the observation that the function of Wzz_2 was unaffected, as a general disruption in proximity should have affected regulation by both Wzz_1 and Wzz_2. Additional evidence is provided by the *in vitro* reconstitution of O-Ag polymerization and chain length regulation using purified Wzy and Wzz, respectively[11]; as the overall structure of Wzz has been evolutionarily selected to extend vertically into the periplasm (and away from the active site of Wzy)[14], it is unlikely that Wzz molecules oriented randomly with respect to Wzy would be able to carry out their native function. As such, the occurrence of a direct interaction between Wzy and Wzz is the likeliest scenario. However, further work is warranted to confirm direct physical interactions between these two proteins.

While a N-terminally-truncated Wzy was still able to polymerize O-Ag to Wzz-specified modal lengths[11], C-terminal polymerase truncations were found to affect PCP-regulated EPS lengths[31]. Such C-terminal truncations were found to affect proposed polymerase–PCP interactions arising from exogenous construct overexpression in a bacterial two-hybrid screen[33], though co-immunoprecipitation of Wzy and Wzz could not be achieved when the target constructs were expressed at chromosomal levels in a native host[47].

Recently, the TMS of Wzz were shown by our group to play a role in the mechanism of seroconversion of *P. aeruginosa* PAO1 from serotype O5 to serotype O16. Previously, the inhibitor of α-polymerase (Iap) peptide (encoded by the lysogenic D3 bacteriophage) had been found to abrogate synthesis of α-linked B-band O-Ag by Wzy_{Pa}, while allowing a phage-encoded Wzy to polymerize the same substrate O-Ag subunits with a β-linkage[48,49]. The propensity of Iap to insert into the IM via a single TMS was subsequently demonstrated, with its inhibitory effect on Wzy_{Pa} shown to be dependent on the presence/absence of Wzz_1 and Wzz_2. The amino acid sequence of the hydrophobic Iap peptide displayed sequence homology with a region in the first TMS of both Wzz proteins, suggesting that Iap mediates shutdown of Wzy_{Pa} function through mimicry of Wzz TMS. This was further supported by the observation that Wzz overexpression was able to overcome the inhibitory effects of Iap on Wzy_{Pa}[50].

Though high-resolution crystallographic data is not yet available for Wzy proteins, interactions of the polymerase with associated Wzz oligomers are likely mediated by a combination of exposed and membrane-embedded domains between the two proteins. While detailed 3D structural data from X-ray crystallography has been obtained for the periplasmic domains of Wzz proteins[51,52], the steric properties conferred by the two TMS of Wzz remain poorly understood. Cryo-electron microscopic analysis of 2D crystals of full-length PCP proteins suggests similar quaternary structures[53,54]; however, the role of TMS in heterotypic interactions of these PCPs remains unknown. As such, TMS domains remain an important avenue of future investigation towards better understanding the interplay between Wzy and Wzz. Together, the added depth of understanding regarding Wzy function we have uncovered as well as the support for its interaction with Wzz identified in this investigation have served to further elucidate the underpinnings of the elongation and length-regulation steps of the Wzx/Wzy-dependent pathway.

Methods

DNA manipulations. QuikChange (Agilent) site-directed mutagenesis was carried out on the previously-created pHERD26T-wzy-GFP plasmid; this fusion encodes Wzy_{Pa} fused to a C-terminal GFP-His_8 moiety that does not affect polymerase function[4,20]. The *wzz*$_1$ and *wzz*$_2$ genes were PCR amplified from *P. aeruginosa* PAO1 genomic DNA and cloned in-frame with the downstream mCherry coding sequence in the pRTL2 vector between the *EcoR*I and *Nhe*I restriction sites to create pRTL2-Wzz_1-mCherry and pRTL2-Wzz_2-mCherry. Fusion gene sequences were PCR amplified and subcloned in the arabinose-inducible pHERD28T backbone[55] between the *EcoR*I and *Xba*I restriction sites to create pHERD28T-Wzz_1-mCherry and pHERD28T-Wzz_2-mCherry. Plasmids were maintained in cells through supplementation of lysogeny broth and agar plates with tetracycline (pHERD26T: 15 μg/mL for *Escherichia coli*, 90 μg/mL for *P. aeruginosa*) or trimethoprim (pHERD28T: 100 μg/mL for *E. coli*, 250 μg/mL for *P. aeruginosa*). Plasmid DNA was isolated with a GenElute (Sigma) or a QIAprep Spin (Qiagen) miniprep kit. Mutagenesis PCR reactions were carried out using KOD Hot Start DNA polymerase (Novagen). The sequences of oligonucleotide primers containing mutagenesis mismatches (Sigma) have been provided in the *Supplementary Information* (Supplementary Table S3).

Complementation analysis. Each mutant construct was assayed for its ability to complement a chromosomal deficiency and restore B-band O-Ag biosynthesis in a *P. aeruginosa* PAO1 *wzy*::Gm^R insertional mutant[12]. To reflect the intrinsically-low expression levels of Wzy[12,56,57], cells expressing the mutant constructs were grown overnight with shaking at 37°C to stationary phase without expression induction[20]. $Wzz_{1/2}$ constructs were expressed overnight in a *P. aeruginosa* PAO1 *wzz*$_1$::Gm^R/Δ*wzz*$_2$ double-mutant background at 37°C[16]. Background expression from pHERD28T-Wzz_1-mCherry was sufficient to restore Wzz_1-specified modal lengths, while overnight induction with 0.1% L-arabinose was required for pHERD28T-Wzz2-mCherry-mediated restoration of Wzz_2-specific B-band O-Ag modality.

LPS was prepared from complemented *P. aeruginosa* PAO1 mutant strains via a method modified from the Hitchcock and Brown protocol[58]. In brief, overnight cultures were equilibrated to 1 mL at an OD_{600} of 0.45. Cells were sedimented in a microfuge (12 000 × *g*, 1 min), followed by aspiration of the supernatant. Each pellet was washed in 1 mL of 0.9% NaCl, followed by sedimentation as described above. Washed pellets were resuspended in 250 μL lysis buffer (2% SDS, 4% β-mercaptoethanol, 10% glycerol, 1 M Tris HCl [pH 6.8], 0.002% bromophenol blue) and boiled for 30 min. Each sample was treated with Proteinase K (120 μg/mL) and incubated at 55°C overnight. Samples (5 μL) were resolved via SDS-PAGE (150 V, 1 h) and transferred to a nitrocellulose membrane via Western blotting (180 mA, 45 min). Blots were blocked (1 h) with 3% bovine serum albumin (BSA) in phosphate-buffered saline (PBS), followed by washing with PBS (5 min). Membranes were incubated overnight with anti-B-band O-Ag mouse monoclonal antibody MF15-4[59], followed by three washes in PBS. Blots were incubated (1 h) with goat anti-mouse secondary antibody conjugated to alkaline phosphatase (Jackson Immunoresearch) (1 : 2000, 3% BSA in PBS), followed by three washes in PBS. Blots were developed in Buffer A (100 mM Tris-HCl [pH 9.5], 100 mM NaCl, 5 mM $MgCl_2$) supplemented with 0.0165% 5-bromo-4-chloro-3-indoyl phosphate (BCIP) and 0.033% nitro blue tetrazolium (NBT). All steps were carried out at room temperature.

Fluorescence analyses. The quantitation of membrane localization for different Wzy-GFP-His_8 mutant constructs was carried out as previously detailed[20]. In brief, overnight cultures of *P. aeruginosa* PAO1 *wzy*::Gm^R induced for construct expression were sedimented via centrifugation (12 000 × *g*, 30 min), followed by lysis using BPER-II protein extraction reagent (Thermo) with 2× Complete protease inhibitor lacking EDTA (Roche). Crude debris and inclusion bodies were removed via sedimentation (12 000 × *g*, 30 min). This supernatant was sedimented in an ultracentrifuge to isolate the membrane fraction (120 000 × *g*, 1.25 h), followed by resuspension of the membrane pellet directly in Laemmli sample buffer. Samples were heated at 37°C (30 min) and resolved on a 12% SDS-PAGE gel. Resolved gels were subsequently scanned for in-gel GFP fluorescence using a Typhoon 9410 imager (GE Healthcare). Densitometry analysis of full-length Wzy-GFP-His_8 in-gel fluorescence was performed using ImageJ software on three biological replicates. All statistical analyses were performed using GraphPad Prism software.

To measure the fluorescence of GFP released from lysed cells (i.e. in the soluble fraction), the clarified supernatant (5 mL) following sedimentation in the ultracentrifuge (described above) was obtained. Clarified supernatant was mixed at a 1 : 1 ratio with 2× Laemmli sample buffer and incubated at 37°C (30 min). Samples (20 μL) were resolved on a 12% SDS-PAGE gel. Resolved gels were scanned as described above using a Typhoon imager. Densitometry analysis of cleaved GFP-His_8 in-gel fluorescence was performed as described above.

Bioinformatic analyses. Detection of remote homologues was accomplished by performing a *jackhmmer*[23] search against the UNIREF100 data bank[60] using three iterations with an E-value inclusion threshold of 10^{-6}, resulting in a final MSA of 30 aligned sequences. Based on the MSA, amino acid sequence logos were generated using WebLogo (http://weblogo.berkeley.edu/logo.cgi)[61]. TM*hit* analysis (L/2 threshold) was performed online (http://bio-cluster.iis.sinica.edu.tw/TMhit/)[28], with topological restraints introduced based on the published TMS boundaries for the Wzy_{Pa} topology map[4]. Isoelectric point values for WT and mutant loop sequences were characterized via the ExPASy portal.

1. Lyczak, J. B., Cannon, C. L. & Pier, G. B. Establishment of *Pseudomonas aeruginosa* infection: lessons from a versatile opportunist. *Microbes Infect.* **2**, 1051–1060 (2000).
2. Lam, J. S., Taylor, V. L., Islam, S. T., Hao, Y. & Kocíncová, D. Genetic and functional diversity of *Pseudomonas aeruginosa* lipopolysaccharide. *Front. Microbiol.* **2**, 1–25 (2011).
3. Whitfield, C. Biosynthesis of lipopolysaccharide O antigens. *Trends Microbiol.* **3**, 178–185 (1995).
4. Islam, S. T., Taylor, V. L., Qi, M. & Lam, J. S. Membrane topology mapping of the O-antigen flippase (Wzx), polymerase (Wzy), and ligase (WaaL) from *Pseudomonas aeruginosa* PAO1 reveals novel domain architectures. *mBio* **1**, e00189–10 (2010).
5. Islam, S. T. & Lam, J. S. Wzx flippase-mediated membrane translocation of sugar polymer precursors in bacteria. *Environ. Microbiol.* **15**, 1001–1015 (2013).
6. Cuthbertson, L., Mainprize, I. L., Naismith, J. H. & Whitfield, C. 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 (2009).
7. Liu, D., Cole, R. A. & Reeves, P. R. An O-antigen processing function for Wzx (RfbX): a promising candidate for O-unit flippase. *J. Bacteriol.* **178**, 2102–2107 (1996).
8. Islam, S. T. *et al.* A cationic lumen in the Wzx flippase mediates anionic O-antigen subunit translocation in *Pseudomonas aeruginosa* PAO1. *Mol. Microbiol.* **84**, 1165–1176 (2012).
9. Burrows, L. L. & Lam, J. S. Effect of *wzx* (*rfbX*) mutations on A-band and B-band lipopolysaccharide biosynthesis in *Pseudomonas aeruginosa* O5. *J. Bacteriol.* **181**, 973–980 (1999).
10. Islam, S. T. *et al.* Proton-dependent gating and proton uptake by Wzx support O-antigen-subunit antiport across the bacterial inner membrane. *mBio* **4**, e00678–13 (2013).
11. Woodward, R. *et al. In vitro* bacterial polysaccharide biosynthesis: defining the functions of Wzy and Wzz. *Nat. Chem. Biol.* **6**, 418–423 (2010).
12. de Kievit, T. R., Dasgupta, T., Schweizer, H. & Lam, J. S. Molecular cloning and characterization of the *rfc* gene of *Pseudomonas aeruginosa* (serotype O5). *Mol. Microbiol.* **16**, 565–574 (1995).
13. Robbins, P. W., Bray, D., Dankert, M. & Wright, A. Direction of chain growth in polysaccharide synthesis: work on a bacterial polysaccharide suggests that elongation can occur at the "reducing" end of growing chains. *Science* **158**, 1536–1542 (1967).
14. Kalynych, S., Valvano, M. A. & Cygler, M. Polysaccharide co-polymerases: the enigmatic conductors of the O-antigen assembly orchestra. *Protein Eng. Des. Sel.* **25**, 797–802 (2012).
15. Burrows, L. L., Chow, D. & Lam, J. S. *Pseudomonas aeruginosa* B-band O-antigen chain length is modulated by Wzz (Ro1). *J. Bacteriol.* **179**, 1482–1489 (1997).
16. Daniels, C., Griffiths, C., Cowles, B. & Lam, J. S. *Pseudomonas aeruginosa* O-antigen chain length is determined before ligation to lipid A core. *Environ. Microbiol.* **4**, 883–897 (2002).
17. Abeyrathne, P. D. & Lam, J. S. WaaL of *Pseudomonas aeruginosa* utilizes ATP in *in vitro* ligation of O antigen onto lipid A-core. *Mol. Microbiol.* **65**, 1345–1359 (2007).
18. Abeyrathne, P. D., Daniels, C., Poon, K. K. H., Matewish, M. J. & Lam, J. S. Functional characterization of WaaL, a ligase associated with linking O-antigen polysaccharide to the core of *Pseudomonas aeruginosa* lipopolysaccharide. *J. Bacteriol.* **187**, 3002–3012 (2005).
19. Silhavy, T. J., Kahne, D. & Walker, S. The bacterial cell envelope. *Cold Spring Harb. Perspect. Biol.* **2**, a000414 (2010).
20. Islam, S. T. *et al.* Dual conserved periplasmic loops possess essential charge characteristics that support a catch-and-release mechanism of O-antigen polymerization by Wzy in *Pseudomonas aeruginosa* PAO1. *J. Biol. Chem.* **286**, 20600–20605 (2011).
21. Marolda, C. L., Tatar, L. D., Alaimo, C., Aebi, M. & Valvano, M. A. 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 (2006).
22. Finn, R. D., Clements, J. & Eddy, S. R. HMMER web server: interactive sequence similarity searching. *Nucl. Acids Res.* **39**, W29–W37 (2011).
23. Eddy, S. R. Accelerated profile HMM searches. *PLoS Comput. Biol.* **7**, e1002195 (2011).
24. Eddy, S. R. A new generation of homology search tools based on probabilistic inference. *Genome Inform.* **23**, 205–211 (2009).
25. Schneider, T. D. & Stephens, R. M. Sequence logos: a new way to display consensus sequences. *Nucleic Acids Res.* **18**, 6097–6100 (1990).
26. Nugent, T. & Jones, D. T. Accurate *de novo* structure prediction of large transmembrane protein domains using fragment-assembly and correlated mutation analysis. *Proc. Natl. Acad. Sci. USA* **109**, E1540–E1547 (2012).
27. Marks, D. S., Hopf, T. A. & Sander, C. Protein structure prediction from sequence variation. *Nat. Biotechnol.* **30**, 1072–1080 (2012).
28. Lo, A. *et al.* Predicting helix-helix interactions from residue contacts in membrane proteins. *Bioinformatics* **25**, 996–1003 (2009).
29. Schneider, G. & Fechner, U. Advances in the prediction of protein targeting signals. *Proteomics* **4**, 1571–1580 (2004).
30. Daniels, C., Vindurampulle, C. & Morona, R. Overexpression and topology of the *Shigella flexneri* O-antigen polymerase (Rfc/Wzy). *Mol. Microbiol.* **28**, 1211–1222 (1998).
31. Mazur, A., Krol, J. E., Marczak, M. & Skorupska, A. Membrane topology of PssT, the transmembrane protein component of the type I exopolysaccharide transport system in *Rhizobium leguminosarum* bv. *trifolii* strain TA1. *J. Bacteriol.* **185**, 2503–2511 (2003).
32. Islam, S. T. & Lam, J. S. Topological mapping methods for α-helical bacterial membrane proteins – an update and a guide. *MicrobiologyOpen* **2**, 350–364 (2013).
33. Marczak, M., Dźwierzyńska, M. & Skorupska, A. Homo- and heterotypic interactions between Pss proteins involved in the exopolysaccharide transport system in *Rhizobium leguminosarum* bv. *trifolii*. *Biol. Chem.* **394**, 541–559 (2013).
34. Lin, T.-L. *et al.* Amino acid substitutions of MagA in *Klebsiella pneumoniae* affect the biosynthesis of the capsular polysaccharide. *PLoS ONE* **7**, e46783 (2012).
35. Elumalai, P., Rajasekaran, M., Liu, H.-L. & Chen, C. Investigation of cation–π interactions in sugar-binding proteins. *Protoplasma* **247**, 13–24 (2010).
36. Malik, A. & Ahmad, S. Sequence and structural features of carbohydrate binding in proteins and assessment of predictability using a neural network. *BMC Struct. Biol.* **7**, 1 (2007).
37. Kondakova, A. N. *et al.* Structure of the O-polysaccharide of *Photorhabdus luminescens* subsp. *laumondii* containing D-glycero-D-manno-heptose and 3,6-dideoxy-3-formamido-D-glucose. *Carbohydr. Res.* **351**, 134–137 (2012).
38. Lairson, L. L., Henrissat, B., Davies, G. J. & Withers, S. G. Glycosyltransferases: structures, functions, and mechanisms. *Annu. Rev. Biochem.* **77**, 521–555 (2008).
39. Breton, C., Fournel-Gigleux, S. & Palcic, M. M. Recent structures, evolution and mechanisms of glycosyltransferases. *Curr. Opin. Struct. Biol.* **22**, 540–549 (2012).

40. Kintz, E., Scarff, J. M., DiGiandomenico, A. & Goldberg, J. B. Lipopolysaccharide O-antigen chain length regulation in *Pseudomonas aeruginosa* serogroup O11 strain PA103. *J. Bacteriol.* **190**, 2709–2716 (2008).
41. Murray, G. L., Attridge, S. R. & Morona, R. 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 (2003).
42. Stevenson, G., Kessler, A. & Reeves, P. R. A plasmid-borne O-antigen chain length determinant and its relationship to other chain length determinants. *FEMS Microbiol. Lett.* **125**, 23–30 (1995).
43. Osawa, K. *et al.* O-antigen chain length modulated by the *wzz* gene in *Escherichia coli* O157 influenced its sensitivities to serum complement. *Microbiol. Immunol.* **57**, 616–623 (2013).
44. Hoare, A. *et al.* The normal chain length distribution of the O antigen is required for the interaction of *Shigella flexneri* 2a with polarized Caco-2 cells. *Biol. Res.* **45**, 21–26 (2012).
45. Morona, R., Daniels, C. & Van Den Bosch, L. Genetic modulation of *Shigella flexneri* 2a lipopolysaccharide O antigen modal chain length reveals that it has been optimized for virulence. *Microbiology* **149**, 925–939 (2003).
46. Lam, J. S., MacDonald, L. A. & Lam, M. Y. Production of monoclonal antibodies against serotype strains of *Pseudomonas aeruginosa*. *Infect. Immun.* **55**, 2854–2856 (1987).
47. Carter, J. A. *et al.* The cellular level of O-antigen polymerase Wzy determines chain length regulation by WzzB and WzzpHS-2 in *Shigella flexneri* 2a. *Microbiology* **155**, 3260–3269 (2009).
48. Newton, G. J. *et al.* Three-component-mediated serotype conversion in *Pseudomonas aeruginosa* by bacteriophage D3. *Mol. Microbiol.* **39**, 1237–1247 (2001).
49. Kaluzny, K., Abeyrathne, P. D. & Lam, J. S. Coexistence of two distinct versions of O-antigen polymerase, Wzy-alpha and Wzy-beta, in *Pseudomonas aeruginosa* serogroup O2 and their contributions to cell surface diversity. *J. Bacteriol.* **189**, 4141–4152 (2007).
50. Taylor, V. L., Udaskin, M. L., Islam, S. T. & Lam, J. S. 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 (2013).
51. Tocilj, A. *et al.* Bacterial polysaccharide co-polymerases share a common framework for control of polymer length. *Nat. Struct. Mol. Biol.* **15**, 130–138 (2008).
52. Kalynych, S., Yao, D., Magee, J. & Cygler, M. Structural characterization of closely related O-antigen lipopolysaccharide (LPS) chain length regulators. *J. Biol. Chem.* **287**, 15696–15705 (2012).
53. Larue, K., Ford, R. C., Willis, L. M. & Whitfield, C. Functional and structural characterization of polysaccharide co-polymerase proteins required for polymer export in ATP-binding cassette transporter-dependent capsule biosynthesis pathways. *J. Biol. Chem.* **286**, 16658–16668 (2011).
54. Larue, K., Kimber, M. S., Ford, R. & Whitfield, C. Biochemical and structural analysis of bacterial O-antigen chain length regulator proteins reveals a conserved quaternary structure. *J. Biol. Chem.* **284**, 7395–7403 (2009).
55. Qiu, D., Damron, F. H., Mima, T., Schweizer, H. P. & Yu, H. D. PBAD-based shuttle vectors for functional analysis of toxic and highly regulated genes in *Pseudomonas* and *Burkholderia* spp. and other bacteria. *Appl. Environ. Microbiol.* **74**, 7422–7426 (2008).
56. Abeyrathne, P. D. & Lam, J. S. Conditions that allow for effective transfer of membrane proteins onto nitrocellulose membrane in Western blots. *Can. J. Microbiol.* **53**, 526–532 (2007).
57. Coyne, M. J. & Goldberg, J. B. Cloning and characterization of the gene (*rfc*) encoding O-antigen polymerase of *Pseudomonas aeruginosa* PAO1. *Gene* **167**, 81–86 (1995).
58. Hitchcock, P. J. & Brown, T. M. Morphological heterogeneity among *Salmonella* lipopolysaccharide chemotypes in silver-stained polyacrylamide gels. *J. Bacteriol.* **154**, 269–277 (1983).
59. Lam, J. S., MacDonald, L. A., Lam, M. Y., Duchesne, L. G. & Southam, G. G. Production and characterization of monoclonal antibodies against serotype strains of *Pseudomonas aeruginosa*. *Infect. Immun.* **55**, 1051–1057 (1987).
60. Magrane, M. UniProt UniProt Knowledgebase: a hub of integrated protein data. *Database* **2011** (2011).
61. Crooks, G. E., Hon, G., Chandonia, J.-M. & Brenner, S. E. WebLogo: a sequence logo generator. *Genome Res.* **14**, 1188–1190 (2004).

Acknowledgments

This work was supported by an operating grant to J.S.L. from the Canadian Institutes of Health Research (CIHR) (Grant MOP-14687). S.T.I. is the recipient of a CIHR Frederick Banting and Charles Best Canada Graduate Scholarship doctoral award, a CIHR Michael Smith Foreign Study award, a Cystic Fibrosis Canada (CFC) doctoral studentship and an Ontario Graduate Scholarship in Science and Technology. S.M.H. is the recipient of an Ontario Graduate Scholarship and a CFC Summer Studentship. T.N. is the recipient of a Special Training Fellowship in Biomedical Informatics from the United Kingdom Medical Research Council. J.S.L. holds a Canada Research Chair in Cystic Fibrosis and Microbial Glycobiology.

Author contributions

S.T.I. conceived and designed the experiments; S.T.I., S.M.H., T.N. and A.C.G. performed experiments; S.T.I., S.M.H., T.N. and J.S.L. analyzed the data; S.T.I. and J.S.L. wrote the paper.

Additional information

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

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

How to cite this article: Islam, S.T., Huszczynski, S.M., Nugent, T., Gold, A.C. & Lam, J.S. Conserved-residue mutations in Wzy affect O-antigen polymerization and Wzz-mediated chain-length regulation in *Pseudomonas aeruginosa* PAO1. *Sci. Rep.* **3**, 3441; DOI:10.1038/srep03441 (2013).