Three Wzy polymerases are specific for particular forms of an internal linkage in otherwise identical O units

Yaoqin Hong,† Vincent A. Morcilla, Michael A. Liu, Elsa L. M. Russell and Peter R. Reeves

School of Molecular Bioscience, Building D17, University of Sydney, NSW 2006, Australia

The Wzx/Wzy-dependent pathway is the predominant pathway for O-antigen production in Gram-negative bacteria. The O-antigen repeat unit (O unit) is an oligosaccharide that is assembled at the cytoplasmic face of the membrane on undecaprenyl pyrophosphate. Wzx then flips it to the periplasmic face for polymerization by Wzy, which adds an O unit to the reducing end of a growing O-unit polymer in each round of polymerization. Wzx and Wzy both exhibit enormous sequence diversity. It has recently been shown that, contrary to earlier reports, the efficiency of diverse Wzx forms can be significantly reduced by minor structural variations to their native O-unit substrate. However, details of Wzy substrate specificity remain unexplored. The closely related galactose-initiated Salmonella O antigens present a rare opportunity to address these matters. The D1 and D2 O units differ only in an internal mannose–rhamnose linkage, and D3 expresses both in the same chain. D1 and D2 polymerases were shown to be specific for O units with their respective α or β configuration for the internal mannose–rhamnose linkage. The Wzy encoded by D3 gene cluster polymerizes only D1 O units, and deleting the gene does not eliminate polymeric O antigen, both observations indicating the presence of an additional wzy gene. The levels of Wzx and Wzy substrate specificity will affect the ease with which new O units can evolve, and also our ability to modify O antigens, capsules or secreted polysaccharides by glyco-engineering, to generate novel polysaccharides, as the Wzx/Wzy-dependent pathway is responsible for much of the diversity.

INTRODUCTION

O antigen is the outermost component of LPS and in many bacteria is exposed on the cell surface. Each O antigen is composed of a variable number of repeats of a three- to eight-sugar repeat unit, called the O unit. The O units are structurally highly variable in most species, with variation present in the sugars, their order and their linkages. For example there are over 186 O-antigen forms in Escherichia coli (including Shigella) and 46 in Salmonella enterica (Liu et al., 2014; Wang et al., 2010). Most of the genes for synthesis and processing of the O unit are found in a gene cluster, which in S. enterica is between the galF and gnd genes (Liu et al., 2014). Most O antigens are synthesized by the Wzx/Wzy-dependent pathway. The several steps of this pathway have been reviewed many times since its discovery, and readers are referred to some recent reviews (Islam & Lam, 2014; Kalynych et al., 2014; Reeves & Cunneen, 2009) for details of the pathway outlined below. Synthesis begins with assembly of O units on the lipid carrier undecaprenyl phosphate (Und-P), at the cytoplasmic face of the inner membrane. In the S. enterica O antigens used in this study, the first sugar is added as a galactose phosphate (Gal-P) to give undecaprenyl pyrophosphate (Und-PP)-linked Gal. The remaining sugars are subsequently transferred from nucleotide–sugar precursors by glycosyltransferases (GTs). Completed O units are flipped across the inner membrane by a Wzx translocase to the periplasm, where they can be polymerized by the O-antigen polymerase, Wzy (Fig. 1), with the chain length of the polymer being determined by Wzz, to give a specific modal pattern (Fig. 1). Both single O units and polymeric O antigen are subsequently incorporated into LPS by the WaaL ligase, and the completed LPS is exported to the outer membrane (Silhavy et al., 2010).
Wzy polymerase transfers O units to a growing O-antigen chain, also on Und-PP. The polymerization reaction resembles that of GTs in generating a glycosidic linkage (Islam & Lam, 2014), in this case by addition of a complete O unit to a growing polymer on Und-PP. The new O unit itself is taken from an Und-PP linked O-unit donor, and Robbins et al. (1967) showed that the addition is at the reducing end of the polymer. However despite these differences, the final transferase reaction is to form a glycosidic linkage as in the GT reaction. The two signature proteins of the pathway, Wzx and Wzy, share an interesting characteristic feature, namely the presence of enormous sequence diversity, including in S. enterica (Liu et al., 2014), despite each having a conserved secondary structure and membrane topology (Islam & Lam, 2014). Both are very hydrophobic with multiple trans-membrane segments (TMSs). Wzy has two major periplasmic loops, which are between TMS 5/6 and TMS 9/10 in the topology maps for Shigella flexneri and Pseudomonas aeruginosa PAO1, with 12 and 14 TMSs, respectively, and similar topology from TMS 1 to TMS 12 (Daniels et al., 1998; Islam et al., 2010). Both of these loops have been shown to be important for Wzy function in the two examples discussed (Islam et al., 2011, 2013; Nath & Morona, 2015; Nath et al., 2015). The diversity of Wzy forms is exemplified by the fact that the 36 Wzy polymerases for N-acetylglucosamine/N-acetylgalactosamine initiated S. enterica O antigens can be divided into 35 distinct homology groups using TribeMCL (Liu et al., 2014), a method that performs rapid and accurate assignment of proteins into families using the Markov cluster algorithm (Enright et al., 2002).

We have recently reported that Wzx translocases can translocate only their native O units at optimal efficiency (Hong et al., 2012; Hong & Reeves, 2014; Liu et al., 2015). Wzy polymerizes the Und-PP-linked O units to produce long-chain O antigen, and also exhibits enormous sequence variation. Some of this reflects the variation in the linkages made by different Wzy forms when linking adjacent O units, but we hypothesize that the variation in Wzy is also related to variation in other parts of the acceptor O unit.

Here we examine the Wzy polymerases of S. enterica group D1, D2 and D3 O antigens (Fig. 2), all members of a set of eight related O-antigen structures and gene clusters that have been well characterized (Reeves et al., 2013). Group C2 has a rather different structure and its own wzy gene: the seven other groups have one of the three unrelated wzy genes (Reeves et al., 2013), which are the subject of this study. The genes are named wzya, wzye and wzyd3, after the groups where they were first observed. It is interesting that this set of eight gene clusters has four wzy forms, in contrast to the situation for the 37 gene clusters for the S. enterica N-acetylglucosamine/N-acetylgalactosamine-initiated O antigens, which includes 36 wzy homology groups.

The D1, D2 and D3 O units have the same three sugars in the same order, and also the same tyvelose (Tyv) sidebranch sugar. The only structural difference between the O units is in the intra-O-unit linkage between the mannose (Man) and rhamnose (Rha) residues, which is α(1→4) in group D1 but β(1→4) in group D2, with both present in group D3 (Nghiem & Staub, 1975; Nghiem et al., 1982, 1992). The D1 and D2 O-antigen clusters also differ in the presence of wbaU or wbaO, respectively, that encode the GTs that give respectively either the α(1→4) or the β(1→4) intra-O-unit Man–Rha linkage (Fig. 2). Interestingly, the D3 gene cluster has only wbaU and it was proposed that there must be a wbaO gene elsewhere on the chromosome (Curd et al., 1998). The chromosomal location of the wzy genes also varies: wzya (D2 and E) and wzyd3 (D3 and B2) are found in the gene cluster, while the wzyb gene (B1

![Fig. 1. Group D O-antigen assembly and processing steps. Note that for each round of polymerization, the new O unit is added to the reducing end of the Und-PP-linked O-unit polymer (Bray & Robbins, 1967; Robbins et al., 1967).](image-url)
and D1) is found outside of the gene cluster, at the rfc locus (Fig. 2). Thus this study involves three different Wzy polymerases, for two O-unit substrates that differ only in the intra-O-unit Man–Rha linkage. WzyB makes an \( \alpha(1\rightarrow2) \) inter-O-unit linkage, but WzyE and WzyD3 both make the same \( \alpha(1\rightarrow6) \) inter-O-unit linkage.

We took advantage of the structural and genetic relatedness of the \( S. \text{ enterica} \) groups D1, D2 and D3 O antigens to generate constructs and demonstrate that each of the Wzy polymerases has a strong specificity for O units containing either the \( \alpha(1\rightarrow4) \) (WzyB and WzyD3) or the \( \beta(1\rightarrow4) \) (WzyE) intra-O-unit Man–Rha linkage.

**METHODS**

**Bacterial strains and plasmids.** Strains and plasmids used in this study are described in Table 1. Our D1 base strain, P9560, with the \( \alpha(1\rightarrow4) \) Man–Rha linkage, was derived from LT2, a group B1 strain, by replacing its \( abe \) gene with \( prt–tyv \) genes, so that it expresses the D1 O antigen. We used LT2 because genetic manipulation of our \( S. \text{ enterica} \) D1 strain was technically very difficult. P9560 also has its \( wzy \) gene replaced by an \( rpsL–kan \) cassette and has a \( galE \) deletion. The \( galE \) deletion allows control of O-unit synthesis, which requires addition of Gal to the medium, while removal of the \( wzy \) gene allows both the native and the heterologous \( wzy \) genes to be present in the same context, in this case on a plasmid. Our D2 base strain, with the \( \beta(1\rightarrow4) \) Man–Rha linkage, was to have been M2967, derived from D2 strain M388, also with a \( galE \) deletion and its \( wzy \) gene replaced by an \( rpsL–kan \) cassette. However, we found that the M2893 strain and its derivative M2967 are unable to use exogenous Gal for O-antigen synthesis, and are presumed to have an uncharacterized mutation that affects this. We therefore used M2987 as our base strain, which is M2967 carrying plasmid pPR2204, with an inducible \( galE \) gene to allow O-unit synthesis to be turned on as needed.

The O antigen produced by each construct was confirmed by testing strains P9528 and P9550 (the constructed D1 strain from which P9560 was derived, by \( wzy \) deletion) for agglutination appropriate factor-specific sera (see Table S2, available in the online Supplementary material). In particular note that the O4 serum reacted only with group B1 and B2 strains and the O9 serum reacted with all D1, D2 and D3 strains, including the group-B-derived D1 strain P9550. The results for the O46 and O27 sera are discussed in the Results.

**Media and reagents.** Nutrient broth (per litre: 5 g NaCl, 5 g yeast extract and 10 g bacteriological peptone) was used for all bacterial growth in this study. Media were supplemented with antibiotics as required at the following concentrations: ampicillin (25 \( \mu \)g ml\(^{-1}\)), kanamycin (25 \( \mu \)g ml\(^{-1}\)), chloramphenicol (25 \( \mu \)g ml\(^{-1}\)) and streptomycin (800 \( \mu \)g ml\(^{-1}\)).

---

**Fig. 2.** O-unit structures, gene clusters and Wzy proteins of \( S. \text{ enterica} \) groups D1, D2 and D3. The D1 and D2 O units are identical, except for the linkages between Man and Rha, i.e. \( \alpha(1\rightarrow4) \) for D1, and \( \beta(1\rightarrow4) \) for D2. D3 is known to produce both types of O units (Reeves et al., 2013). The gene clusters are colour coded, and the GT genes are highlighted in bold. D1, D2 and D3 gene clusters are also highly related, with the other most significant difference being the presence of a \( wbaU \) gene in D1 and D3 [responsible for the \( \alpha(1\rightarrow4) \) Man–Rha linkage], and a \( wbaO \) gene in D2 [for the \( \beta(1\rightarrow4) \) linkage]. The internal Man–Rha linkage difference between the D1 and D2 O-unit forms is indicated by a brown arrow. Different \( wzy \) genes are found in the three \( S. \text{ enterica} \) groups, i.e. \( wzyB \) in D1 (located in the rfc locus elsewhere in the genome), and \( wzyE \) and \( wzyD3 \) (both in the gene cluster) in D2 and D3, respectively. A red arrow indicates the C2 position of Man used for WzyB, and blue arrows indicate the C6 position for WzyD3 and WzyE polymerization. The \( wzyB \) and \( wzyD3 \) genes are also present in groups B1 and B2, which differ from D1 and D3 in the presence of abequose (Abe) in place of tyvelose (Tyv) in the structure and the Abe gene in place of \( p rt \) and \( t yv \) genes in D1 and D3. The \( wzyE \) gene is also present in group E, which has no dideoxyhexose in the structure and none of the genes associated with that. This figure is modified from Reeves et al. (2013).
Gene replacements. Gene replacement was performed as described by Datsenko & Wanner (2000). Plasmid pKD46, which carries the lambda recombinase system (Datsenko & Wanner, 2000), was used to transform strains prior to replacement of the targeted gene by either an \textit{rpsL–kan} or a \textit{prt–tyv} cassette as described earlier (Hong et al., 2012). Gene replacement of \textit{abe} with \textit{prt–tyv} was performed by first inserting the \textit{rpsL–kan} cassette just downstream of the \textit{abe} gene, which was removed when the \textit{prt–tyv} sequence was inserted using streptomycin selection. Primers for gene replacement contained 42-bp sequences flanking the targeted site and an additional 20-bp homology sequence for the PCR amplification (Table S1).

Cloning and transformation. The oligonucleotides used in cloning are detailed in Table S1. PCR-amplified genes were cloned into plasmid vectors, and introduced into the bacterial strains by electro-transformation, using the Gene Pulser Electroporation System (Bio-Rad) as described previously (Hong & Reeves, 2014).

### Table 1. Strains and plasmids used in the study

<table>
<thead>
<tr>
<th>Strain/plasmid</th>
<th>Parent strain</th>
<th>Genotype/plasmid description</th>
<th>Source/reference</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Strain</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P9528</td>
<td>CL4419</td>
<td>S. enterica group B1 serovar</td>
<td>Hong et al. (2012)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Typhimurium strain LT2,</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>\textit{hsdL trp32 nml}</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>\textit{flaA66, rpsL–xylT404 livE452 metE551}</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>\textit{metA22 hsdΔA galE}</td>
<td></td>
</tr>
<tr>
<td>P9550</td>
<td>P9528</td>
<td>P9528, \textit{abe} gene replaced by</td>
<td>This study</td>
</tr>
<tr>
<td></td>
<td></td>
<td>\textit{prt–tyv} from \textit{S. enterica} group D1 Dublin</td>
<td></td>
</tr>
<tr>
<td>P9560</td>
<td>P9550</td>
<td>P9550, \textit{wzy} gene replaced by \textit{rpsL–kan}</td>
<td>This study</td>
</tr>
<tr>
<td>P9561</td>
<td>P9560</td>
<td>P9560, pPR2198</td>
<td>This study</td>
</tr>
<tr>
<td>P9562</td>
<td>P9560</td>
<td>P9560, pPR2202</td>
<td>This study</td>
</tr>
<tr>
<td>P9576</td>
<td>P9560</td>
<td>P9560, pPR2200</td>
<td>This study</td>
</tr>
<tr>
<td>M2893</td>
<td>M388</td>
<td>S. enterica group D2 Strasbourg, \textit{Δ}galE</td>
<td>Hong et al. (2012)</td>
</tr>
<tr>
<td>M2986</td>
<td>M2893</td>
<td>M2893, pPR2204</td>
<td>This study</td>
</tr>
<tr>
<td>M2967</td>
<td>M2893</td>
<td>M2893, \textit{wzy} gene replaced by \textit{rpsL–kan}</td>
<td>This study</td>
</tr>
<tr>
<td>M2987</td>
<td>M2967</td>
<td>M2967, pPR2204</td>
<td>This study</td>
</tr>
<tr>
<td>M2990</td>
<td>M2987</td>
<td>M2987, pPR2198</td>
<td>This study</td>
</tr>
<tr>
<td>M2995</td>
<td>M2987</td>
<td>M2987, pPR2202</td>
<td>This study</td>
</tr>
<tr>
<td>M2992</td>
<td>M2987</td>
<td>M2987, pPR2200</td>
<td>This study</td>
</tr>
<tr>
<td>M1167</td>
<td>M1167</td>
<td>S. enterica group D3 Zuerich</td>
<td>Institut Pasteur</td>
</tr>
<tr>
<td>M2975</td>
<td>M1167</td>
<td>M1167, \textit{Δ}galE</td>
<td>This study</td>
</tr>
<tr>
<td>M2978</td>
<td>M2975</td>
<td>M2975, \textit{wzy} gene replaced by \textit{rpsL–kan}</td>
<td>This study</td>
</tr>
<tr>
<td><strong>Plasmid</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pKD46</td>
<td></td>
<td>Lambda recombinase genes (α, β, γ) controlled by</td>
<td>Datsenko &amp; Wanner (2000)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>\textit{arabinose inducible promoter}, \textit{P}_{arab} temperature-sensitive \textit{oriR101} replicon, ampicillin resistance</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>\textit{replicon, ampicillin and chloramphenicol resistance}</td>
<td></td>
</tr>
<tr>
<td>pPR2176</td>
<td></td>
<td>\textit{rpsL–kan} cloned into \textit{pGEMT}</td>
<td>Hong et al. (2012)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(\textit{rpsL} gene obtained from \textit{E. coli} K-12 MG1655), ampicillin and \textit{kanamycin resistance}</td>
<td></td>
</tr>
<tr>
<td>pPR2204</td>
<td></td>
<td>\textit{galE} gene from \textit{S. enterica} group B1 LT2</td>
<td>This study</td>
</tr>
<tr>
<td></td>
<td></td>
<td>cloned into \textit{pWQ552}, ampicillin resistance</td>
<td></td>
</tr>
<tr>
<td>pPR2198</td>
<td></td>
<td>\textit{wzy} from \textit{S. enterica} group B1 LT2</td>
<td>This study</td>
</tr>
<tr>
<td></td>
<td></td>
<td>cloned into \textit{pWQ572}, chloramphenicol resistance</td>
<td></td>
</tr>
<tr>
<td>pPR2200</td>
<td></td>
<td>\textit{wzy} from \textit{S. enterica} group D2 Strasbourg</td>
<td>This study</td>
</tr>
<tr>
<td></td>
<td></td>
<td>cloned into \textit{pWQ572}, chloramphenicol resistance</td>
<td></td>
</tr>
<tr>
<td>pPR2202</td>
<td></td>
<td>\textit{wzy} from \textit{S. enterica} group D3 Zuerich</td>
<td>This study</td>
</tr>
<tr>
<td></td>
<td></td>
<td>strain M1167 cloned into \textit{pWQ572}, chloramphenicol resistance</td>
<td></td>
</tr>
</tbody>
</table>
LPS extraction and visualization. Cultures (10 ml) were prepared by growth at 37 °C in nutrient broth. When the OD600 reached ~0.2, tetracycline was added to strains carrying either the pWQ552- or the pWQ572-based vectors (King et al., 2014; Larue et al., 2011), to induce gene expression, as described earlier (Hong & Reeves, 2014). After another 30 min Gal was added to strains carrying either the pWQ552- or the pWQ572-based vectors (King et al., 2014; Larue et al., 2011), to induce gene expression, as described earlier (Hong & Reeves, 2014). After another 30 min Gal was added to a final concentration of 0.25 %, and growth was continued for another 60 min. LPS was then prepared generally as described by McGrath & Osborn (1991), but with modifications. Briefly, cells were washed in 1× PBS, and the LPS was extracted using 45 % hot phenol. After cooling, sodium acetate was added to the aqueous phase to 10 mM final concentration, and LPS was precipitated by addition of ice-cold ethanol in a 35 : 65 ratio, and then held at −20 °C for 30 min. The precipitated LPS was collected by centrifugation (16 000 g), and treated with DNase (500 mg ml⁻¹) and RNase (500 mg ml⁻¹) for 60 min at 37 °C, and then with proteinase K (1.5 mg ml⁻¹) for another 60 min at 56 °C. Samples derived from ~2.5 × 10⁸ cells were electrophoresed on 13 % acrylamide Tricine-SDS-PAGE gels. LPS was visualized by silver staining as described by Brown et al. (1991).

RESULTS

Experimental set-up

We constructed strains M2987 and P9560, for use in testing the ability of different Wzy polymerases to use the group D2 and D1 O units, respectively. Both strains have a deletion of the wzy gene and testing is done by addition of a cloned wzy gene. It is known that mutations in O-antigen genes can be deleterious to the cell (Hong et al., 2012; Mäkelä & Stocker, 1984; Rick & Osborn, 1972; Rick et al., 1988; Yuasa et al., 1969), and we therefore provided for controlled synthesis of uridine diphosphate Gal, which is required to initiate O-antigen synthesis (see Methods).

Wzy can discriminate between O units possessing identical sugars but different internal sugar linkages

We first compared the ability of WzyE, WzyB and WzyD₃ to polymerize the group D1 O unit with the α(1→4) intra-O-unit Man–Rha linkage, by adding the respective wzy plasmids to strain P9560. Fig. 3(a) shows an SDS-PAGE gel of LPS from the three strains. Lane 1 has LPS from strain P9550, which carries the chromosomal wzyB gene, as is normal for group D1, and has the expected normal profile with bands for LPS having up to about 40 O units, with the usual modal distribution of chain lengths. Lane 2 with LPS from strain P9560, which has no wzy gene, has only the lipid A-core band and a band for LPS with a single O unit. Lane 4, with LPS from the strain with the wzyB gene on a plasmid, has the same profile as for strain P9550, showing full complementation of the wzy deletion. Lanes 3 and 5 have LPS from strains with the wzyE and wzyD₃ genes, respectively, on a plasmid. The wzyD₃ gene gives high-level complementation of the wzy mutation, but wzyE does not complement the loss at all. Note that complementation by wzyB and wzyD₃ gives LPS with different modal patterns for O-antigen chain length. WzyB gives a wild-type-like pattern (Fig. 3a, lane 4), while WzyD₃ (Fig. 3a, lane 5) generates a rather narrow-range modal value at the top of the ‘wild-type’ molecular mass range (Fig. 3a, lane 1). This could be due to different outcomes from
interaction of the two Wzy polymerases with the Wzz protein, which would be consistent with the recent finding that mutational changes in the Wzy polymerase can modify Wzz-regulated modal chain lengths (Islam et al., 2013; Nath & Morona, 2015; Nath et al., 2015). We also observed that the LPS with D1 O units polymerized by WzyD3 ran faster on SDS-PAGE than those polymerized by WzyB (Fig. 3a, compare lane 4 with lane 5). Our D1 strain was derived from group B strain LT2 (see Methods), which is known to have C6 of the Gal residue glucosylated (Hellerqvist et al., 1969). This glucosylation is now known to occur in the periplasm (Takeshita & Mäkelä, 1971), potentially after polymerization, and it could be that glucosylation is inhibited by the non-native D3 polymerization linkage on the Gal residue, which may account for the altered migration pattern. An alternative explanation could be related to the difference in conformation of the two O-antigen structures.

We next compared the ability of WzyE, WzyB, and WzyD3 to polymerize the group D2 O unit with the \( \beta(1\rightarrow4) \) intra-O-unit Man–Rha linkage, by adding the respective \( \textit{wzy} \) plasmids to strain M2987 in which the \( \textit{wzy} \)E gene has been deleted. The results are shown Fig. 3(b). Lane 1 has LPS from strain M2986, which carries the \( \textit{wzy} \)E gene, as is normal for group D2, but the \( \textit{wzy} \)E plasmid could only partially restore the O-antigen phenotype (Fig. 3b, lane 3), despite many attempts (data not shown). The \( \textit{wzy} \)E gene is upstream of all the GT genes involved in O-unit synthesis and the \( \textit{wzy} \)E deletion in M2987 might well have a polar effect on expression, which could account for this phenomenon. On the other hand, addition of either \( \textit{wzy} \)B or \( \textit{wzy} \)D3 clones (Fig. 3b, lanes 4 and 5 respectively) gives no observable activity, as the LPS phenotype resembles that of the \( \Delta \textit{wzy} \)E strain (Fig. 3b, lane 2).

In summary, the results show that the configuration of the Man–Rha linkage is a determining factor for successful O-antigen polymerization, with WzyB and WzyD3 being active only if the \( \alpha(1\rightarrow4) \) intra-O-unit linkage is present, and WzyE active only if the \( \beta(1\rightarrow4) \) intra-O-unit linkage is present.

**Two functional wzy genes present in \textit{S. enterica} group D3**

We have shown that WzyD3 can only polymerize O units with the \( \alpha(1\rightarrow4) \) intra-O-unit Man–Rha linkage. However, it is known that D3 strains produce long-chain O antigen with both \( \alpha \) and \( \beta \) Man–Rha configurations in the same chain, although the gene cluster has only the single \( \textit{wzy} \)D3 gene, and a \( \textit{wbaU} \) gene but no \( \textit{wbaO} \)-like gene (Reeves et al., 2013). This suggests that there is another functional \( \textit{wzy} \) gene elsewhere in the chromosome, in addition to the anticipated \( \textit{wbaO} \)-like gene. We examined this possibility by construction of a D3 strain in which the \( \textit{wzy} \)D3 gene has been removed by allelic replacement. The \( \Delta \textit{wzy} \)D3 strain still produces polymeric O antigen, although the amount is dramatically reduced compared with the parent strain (Fig. 4, lane 2). This indicates that there is a second \( \textit{wzy} \) polymerase gene in the D3 strain, although it appears to be less active than the known \( \textit{wzy} \)D3 gene, even without competition from the Wzy encoded in the gene cluster, there is less than 50% of the amount of polymer than found in the parent, based on stain intensity. The polymer in the \( \Delta \textit{wzy} \)D3 strain reacts with O9 serum (Methods and Table S2), confirming that it...
contains Tyv and hence is a group D O antigen. We also used O27 and O46 sera that should distinguish the two D3 O antigens. Unfortunately, neither of them worked with the D3 strains (Table S2), possibly due to the presence of side-branch glucose in the Gal moiety (factor 1), which is known to affect binding of the two antibodies (Nghiêm, 1971, 1977; Nghiêm et al., 1982, 1992). However, despite the failures with the O46 and O27 sera, we conclude that the Δwzy strain is producing the D2 O-antigen form. It reacts with the O9 serum, which virtually rules out anything but the expected group D2-like O antigen. Note that the faster migration pattern for the ΔwzyD3 LPS compared with that of the wild-type on SDS-PAGE might be related to the conformational difference in the Man–Rha linkage.

**DISCUSSION**

We evaluated the specificity of the three Wzy polymerases found in the *S. enterica* group D1, D2 and D3 O-antigen clusters, for their O-unit substrates. It has long been known that groups B1 and D1 (previously B and D) have the same polymerase gene (*wzyB*) at the rfc locus, which clearly acts on both O units (Mäkelä, 1965; Nurminen et al., 1971). Also, Nyman et al. (1979) foreshadowed the specificity of WzyE by using phage resistance patterns to infer that a group B1 strain carrying a wzyE gene in place of its normal wzyB gene lacked polymeric O antigen. Note that WzyE polymerases for group E and group D2 are nearly identical, showing that WzyE activity is not much affected by the presence or absence of the Tyv side branch and that the absence of an O-unit side branch has no detectable effect on WzyE polymerization (Hong et al., 2012), itself foreshadowed by Osborn & Weiner (1968).

However, we are not aware of studies that have investigated the specificity of bacterial Wzy polymerases, or the more numerous GTs, as in this study. The group D1 and D2 O units differ only in the α(1→4) or β(1→4) configuration of the intra-O-unit Man–Rha linkages within their O units. The group D3 structure includes both O units, so D3 strains can synthesize and polymerize both. It was previously thought that the WzyD3 polymerase should present a broad enough specificity to recognize both α- and β-mannosyl-configured oligosaccharides (Nghiêm et al., 1992). However, we find that our cloned *wzyD3* gene confers only the ability to polymerize O units with the α(1→4) intra-O-unit Man–Rha linkage. This was unexpected as there is only one wzy gene in the D3 gene cluster, and this finding shows that D3 strains require separate Wzy forms for two O units that differ only in the configuration of the internal linkage between the Man and Rha sugars. This presence of an additional Wzy form is supported by showing that deletion of the *wzyD3* gene does not fully eliminate polymerization.

We also find that the WzyE and WzyB polymerases can use only one of the two O units. The more interesting comparison is between the WzyE and WzyD3 polymerases, which both make an α(1→6) inter-O-unit linkage, but are specific for O units with the α(1→4) or β(1→4) intra-O-unit Man–Rha linkage, respectively.

Wzy polymerases generate a glycosyl linkage that is indistinguishable from a linkage made by a GT. It is clear that GT specificity must always be sufficient to ensure that only the reported structure is made, implying that the sugars can only be added in the order required for that structure. Thus, each GT must distinguish the intermediate that it normally uses as an acceptor substrate, from other intermediates, all of which will be present. As an example, the *S. enterica* C2 O unit has two adjacent Man residues and two mannosyltransferase genes, of which *wbaZ* is required for transfer of the first Man residue to its Und-PP-linked Gal substrate, and *wbaW* for transfer of the second to the Und-PP-linked Gal–Man product (Liu et al., 1993). That work was done only to establish the function of the two proteins, and no other aspects of specificity were studied. There is also clearly specificity in all of the GTs for the donor sugar, as each adds a specific sugar. However, we are not aware of any work on the detailed specificity of the acceptor component.

More is known of the specificity of oligosaccharyltransferases, such as WaaL ligases, which can add either a single O unit or the O-antigen polymer to lipid A-core oligosaccharide. They carry out a similar reaction to Wzy, but can have very low specificity for their donor component except for the carrier lipid. There are five WaaL forms in *Escherichia coli* that correspond to the five LPS core structures (Whitfield, 2006), and between them they can transfer O units and O-antigen polymers for the 186 serovars (including *Shigella*). The K-12 form of WaaL is assumed to be the ligase for all of the O antigens that have been expressed in K-12, and has also been shown to transfer a range of other saccharides that are made on Und-PP in various constructs (Bastin et al., 1991, 1993; Lee et al., 1992; Manning et al., 1986; Neal et al., 1993). The same applies to PglL of *Neisseria* and PgL of *Campylobacter*, which transfer a specific structure to target proteins to form glycoproteins, but will work with a wide range of other structures, built on Und-PP, which is the carrier lipid for the native structures (Faridmoayer et al., 2008; Wacker et al., 2006).

The finding that a single difference in an intra O-unit sugar linkage can influence Wzy activity has important implications. The presence of long-chain O antigen is important for pathogenesis and colonization for many Gram-negative pathogens by evasion of host immunity (Duerr et al., 2009; Murray et al., 2003, 2005). Thus, the outcome of having an ineffective Wzy polymerase, resulting from either the gain or the loss of genes, would be selection for wzy mutations that favour improved polymerization activity. As for its Wzx counterparts (Hong & Reeves, 2014), an implication of Wzy substrate selectivity is that it can impede establishment of genetic changes that could lead to structural changes in the O unit, and at the same time provides
selection for mutations in wzy to accommodate changes in the structure that do survive. The levels of Wzx and Wzy specificity will not only affect the ease with which new O units can evolve, but also our ability to modify O antigens, by glyco-engineering, to generate novel polysaccharides. These effects will also apply to capsules or secreted polysaccharides using the Wzx/Wzy-dependent pathway, and it should be noted that this includes almost all with a main chain consists of more than two types of sugars.

REFERENCES


Hong, Y. & Reeves, P. R. (2014). Diversity of O-antigen repeat unit structures can account for the substantial sequence variation of Wzx translocases. J Bacteriol 196, 1713–1722.


Edited by: D. Hood