THE JOURNAL OF BIOLOGICAL CHEMISTRY VOL. 282, NO. 37, pp. 26786–26792, September 14, 2007

Glycosyltransferases Involved in Biosynthesis of the Outer Core Region of *Escherichia coli* Lipopolysaccharides Exhibit Broader Substrate Specificities Than Is Predicted from Lipopolysaccharide Structures*[S]

Received for publication, May 18, 2007, and in revised form, July 9, 2007 Published, JBC Papers in Press, July 13, 2007, DOI 10.1074/jbc.M704131200

Michael D. Leipold[‡], **Evgeny Vinogradov**[§], **and Chris Whitfield**[‡1]

From the [‡]*Department of Molecular and Cellular Biology, University of Guelph, Guelph, Ontario N1G 2W1, Canada and the* [§]*National Research Council, Ottawa, Ontario K1A 0R6, Canada*

The *waaJ*, *waaT*, and *waaR* genes encode α-1,2-glycosyltransferases involved in synthesis of the outer core region of the lipopolysaccharide of *Escherichia coli*. They belong to the glycosyltransferase CAZy family 8, characterized by the GT-A fold, DXD motifs, and by retention of configuration at the anomeric carbon of the donor sugar. Each enzyme adds a hexose residue at the same stage of core oligosaccharide backbone extension. However, they differ in the epimers for their donor nucleotide sugars, and in their acceptor residues. WaaJ is a UDP-glucose:(galactosyl) LPS α-1,2-glucosyltransferase, whereas WaaR and WaaT have UDP-glucose:(glucosyl) LPS α-1,2-glucosyltransferase and UDP-galactose:(glucosyl) LPS α-1,2-galactosyltransferase activities, respectively. The objective of this work was to examine their ability to utilize alternate donors and acceptors. When expressed in the heterologous host, each enzyme was able to extend the alternate LPS acceptor *in vivo* but they retained their natural donor specificity. *In vitro* assays were then performed to test the effect of substituting the epimeric donor sugar on incorporation efficiency with the natural LPS acceptor of the enzyme. Although each enzyme could utilize the alternate donor epimer, activity was compromised because of significant decreases in k_{cat} and corresponding increases in K_m(donor). Finally, *in vitro* assays were performed to probe acceptor preference in the absence of the cellular machinery. The results were enzyme-dependent: while an alternate acceptor had no significant effect on the kinetic behavior of His_6-WaaT, His_6-WaaJ showed a significantly decreased k_{cat} and increased K_m(acceptor). These results illustrate the differences in behavior between closely related glycosyltransferase enzymes involved in the synthesis of similar glycoconjugates and have implications for glycoengineering applications.

Bacteria produce a variety of glycoconjugates. The diversity in their structures is afforded by an unparalleled range of glycosyltransferase enzymes that transfer sugars from activated donor substrates to acceptor substrates. Bacterial enzymes have provided some influential models to assess glycosyltransferase structure and function because of the relative ease of their manipulation. One source of glycosyltransferase diversity is lipopolysaccharide (LPS)[2] assembly, and the focus of this study are the enzymes involved in biosynthesis of the core oligosaccharide region (core OS). The outer leaflet of the Gram-negative outer membrane contains LPS as a major component. LPS is comprised of three structural domains: lipid A, core OS, and O antigen (1). *Escherichia coli* isolates produce one of five core OS types: K-12, R1, R2, R3, and R4 (reviewed in Ref. 2), and there are at least two core OS types in *Salmonella* isolates (3). The backbone of the inner (lipid A proximal) core OS is typically conserved, and the various core types primarily arise from differences in inner core substitution and the structure of the part of the outer core, which provides the attachment site for O antigen. The genetic basis for these differences has been described (3, 4).

However, moving from the experimentally derived polysaccharide structure to the assignment of a specific glycosyltransferase involved in a particular linkage is not straightforward. Even when candidate genes have been identified, their DNA sequences alone cannot predict which donor or acceptor substrates will be used by the glycosyltransferase. Experiments involving *in vivo* complementation of chromosomal insertion mutants and subsequent PAGE analysis of LPS patterns do not necessarily directly address which gene product is producing the LPS alteration, nor do they identify the donor substrate used (5). Of the known and predicted outer core OS glycosyltransferases from *E. coli*, only WaaJ has undergone *in vitro* characterization with both native substrates and purified enzyme (6).

Bacterial core OS biosynthesis requires the concerted action of specific glycosyltransferase enzymes. The outer core OS bio-

* This work was supported by funding (to C. W.) from the Natural Sciences and Engineering Research Council of Canada. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

[S] The on-line version of this article (available at http://www.jbc.org) contains supplemental Tables S1–S3 and Figs. S1 and S2.

[1] Recipient of a Tier 1 Canada Research Chair. To whom correspondence should be addressed: Dept. of Molecular and Cellular Biology, University of Guelph, Guelph, Ontario N1G 2W1, Canada. Tel.: 519-824-4120, ext. 53361; Fax: 519-837-1802; E-mail: cwhitfie@uoguelph.ca.

[2] The abbreviations used are: LPS, lipopolysaccharide; CAZy, carbohydrate-active enzymes; core OS, core oligosaccharide; Kdo, 3-deoxy-D-manno-octulosonic acid; UDP-Glc, uridine-5′-diphosphoglucose; Gal, galactose; Glc, glucose; GlcNAc, *N*-acetylglucosamine; GlcN, glucosamine; EtN, ethanolamine; Amp, ampicillin; Gm, gentamycin; CE, capillary electrophoresis; MS, mass spectrometry; NMR, nuclear magnetic resonance; KDN, 2-keto-3-deoxy-D-glycero-D-galacto-nononic acid.

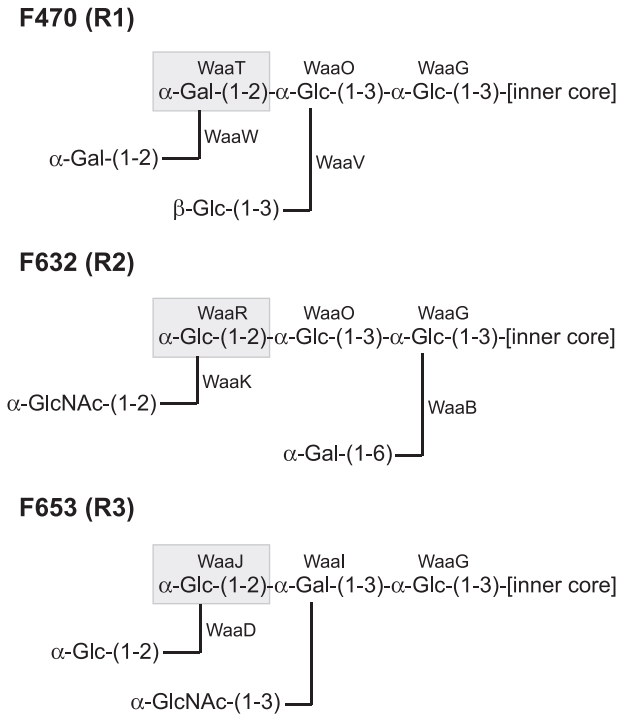

FIGURE 1. **Schematic structures of the outer core regions of *E. coli* R1 (23), R2 (24), and R3 (25) LPS.** When known, the Waa* glycosyltransferase responsible for each linkage is shown. The sugar added by WaaT, WaaR, or WaaJ is highlighted in *gray*.

synthesis glycosyltransferases in *E. coli* provide an interesting collection of related (predicted) enzymes to examine principles of substrate (UDP-sugar) and linkage specificity (Fig. 1). WaaJ catalyzes the addition of an α-1,2-linked Glc to the outer core OS in R3 *E. coli* (7) and *Salmonella enterica* serovar Typhimurium (5), and the kinetic properties of the *E. coli* enzyme have been investigated (6). The related WaaT enzyme adds an α-1,2-linked Gal to the outer core OS in R1 and R4 *E. coli* (8), whereas WaaR is required for the addition of an α-1,2-glucose to the outer core OS in K-12 and R2 *E. coli* (8). While WaaJ, WaaR, and WaaT all belong to glycosyltransferase CAZy family 8 and catalyze the same α-1,2 linkage of the donor sugar to their lipid-linked acceptors, they utilize different UDP-sugar donor substrates and have different terminal sugars on their acceptor LPS. However the predicted activities for WaaR and WaaT have not been examined directly. Here we examine the donor and acceptor specificity of these glycosyltransferases enzymes in a combination of *in vivo* and *in vitro* approaches to generate a better understanding of their substrate specificities and the potential for their manipulation in glycoengineering.

EXPERIMENTAL PROCEDURES

Bacterial Strains and Plasmids—The prototypes for the R1, R2, and R3 core OSs are *E. coli* F470, F632, and F653 respectively; all are rough mutants, *i.e.* lacking O antigen (9). CWG350 (*waaJ:aacC1*) is a derivative of F653 (7). CWG309 (*waaT:aacC1*) and CWG308 (*waaO:aacC1*) are derivatives of F470 (8). These mutants have been described previously and are marked by the gentamycin-resistance cassette (*aacC1*). *E. coli* TOP10 cells F^- *mcr*A Δ(*mrr-hsd*RMS-*mcr*BC) f80Δ*lac*ZM15 Δ*lac*X74 *deoR recA*1 *araD*139 Δ(*ara-leu*)7697 *galU galK rpsL* (Str^r) *endA*1 *nupG* were purchased from Invitrogen and used for expression of the glycosyltransferases.

Plasmid pWQ272 is a pBAD18-derivative (10) containing the coding sequence for an N-terminally hexahistidine-tagged His_6-WaaJ fusion protein and was previously reported (6). Plasmid pWQ269 is a derivative of pBAD18 engineered to contain a plasmid-encoded ribosome binding site and N-terminal His_6-tag 5′ of the multicloning site in an organization identical to pWQ272 (supplemental Table S1). The *waaR* and *waaT* genes were amplified by PCR and cloned into pWQ269. Plasmid pWQ270 contains the coding sequence for His_6-WaaR amplified from *E. coli* F632 genomic DNA (purified using the InstaGene kit, Bio-Rad), while pWQ271 contains the coding sequence for His_6-WaaT from pWQ905 (8). All oligonucleotides were synthesized by Sigma Genosys and are listed in supplemental Table S1. The sequences of all of the constructs were confirmed to be error-free by sequencing at the Guelph Molecular Supercenter (University of Guelph).

Complementation Experiments to Assess in Vivo Activity of His_6-Waa Derivatives*—Function of the various glycosyltransferases was established by electrotransformation of *E. coli* CWG350 (*waaJ:aacC1*), CWG309 (*waaT:aacC1*), and CWG308 (*waaO:aacC1*) with plasmids encoding the appropriate Waa* glycosyltransferase. Cultures of transformed bacteria were grown overnight at 37 °C in LB containing 100 μg/ml ampicillin, and 0.1-ml aliquots were then used to inoculate 5-ml cultures of the same medium supplemented with 0.02% L-arabinose to induce expression from the pBAD promoter in pBAD18 (10). After growth at 37 °C for 5 h, SDS-proteinase K whole cell lysate samples were made following the procedure of Hitchcock and Brown (11). LPS molecular species in these samples were then separated by electrophoresis using 4–12% gradient NuPage gels (Invitrogen). Electrophoresis was carried out at 150 V for 75 min. The gels were silver-stained using standard methods (12). The extent of complementation was determined by scanning the gels using a Bio-Rad GS-800 Calibrated Densitometer and determining the relative amounts of the two major bands with QuantityOne software.

Production and Purification of LPS and Core Oligosaccharides—The LPS (3453 g/mol calculated molecular weight CWG350 LPS; 3368 g/mol calculated molecular weight CWG309 LPS) was purified from F653, CWG350, CWG309, CWG350 (pWQ271), CWG350 (pWQ270), and CWG309 (pWQ272). The LPS was extracted from cells harvested from 6-liter cultures and isolated by phenol/chloroform/petroleum ether method (13), as previously described (7). The isolated LPS was frozen and lyophilized. Working stocks for *in vitro* assays were stored as 2 mg/ml or 5 mg/ml aqueous solutions at −20 °C.

Lipid A was removed by treating LPS (100 mg) in 5 ml of 2% AcOH at 100 °C for 3 h. Lipid A was removed as a precipitate by centrifugation, and soluble products were separated on a Sephadex G-50 column (2.5 × 95 cm) eluted in pyridinium/acetate buffer, pH 4.5 (4 ml of pyridine and 10 ml of AcOH in 1 liter of water). The eluate was monitored using a refractive index detector. The samples were then filtered through a SepPak C18 column (Waters) in water. Anion-exchange chromatography was performed on a 5-ml Hitrap Q column (Amersham Bio-

sciences) in water for 10 min, then in a linear gradient of 0 to 1 M NaCl over 60 min with UV detection at 220 nm.

To obtain a cleaner NMR spectrum, core oligosaccharide from CWG309 (pWQ272) (5 mg) was dephosphorylated by 48% aqueous hydrofluoric acid (0.1 ml) for 20 h at 4 °C. The HF was removed under a stream of nitrogen, and the product was then desalted by gel chromatography on a Sephadex G-15 column (1.6 × 80 cm) column using the pyridinium acetate buffer, pH 4.5 (4 ml of pyridine and 10 ml of AcOH in 1 liter of water) as eluant. The eluant was monitored by a refractive index detector and collected fractions were then lyophilized before use.

Compositional and Methylation Analysis—For compositional analysis, oligosaccharides were hydrolyzed in 4 M CF_3CO_2H (120 °C, 3 h), and monosaccharides were converted into the alditol acetate derivatives. The products were analyzed by gas-liquid chromatography (GC) on an Agilent 6850 chromatograph equipped with DB-17 (30 m × 0.25 mm) fused-silica column using a temperature gradient of 180 °C (2 min) to 240 °C at 2 °C/min. Methylation analysis was performed using Ciucanu-Kerek procedure (14). Methylated products were hydrolyzed, and the monosaccharides were converted to 1*d*-alditol acetates by conventional methods and analyzed by GC-MS. GC-MS was performed on Varian Saturn 2000 system equipped with an ion-trap mass spectral detector using the same column.

NMR Spectroscopy—NMR spectra were recorded at 25 °C in D_2O on a Varian UNITY INOVA 600 instrument using acetone as reference (^1H, 2.225 ppm and ^{13}C, 31.45 ppm). Varian standard programs COSY, NOESY (mixing time of 300 ms), TOCSY (spinlock time 120 ms), HSQC, and gHMBC (evolution delay of 100 ms) were used with digital resolution in F2 dimension <2 Hz/point for proton-proton correlations. Spectra were assigned using the computer program Pronto. The chemical shift data is presented in supplemental Table S2 according to the labeling scheme in supplemental Fig. S1.

Mass Spectrometry—CE-MS spectra were acquired using a 4000 QTrap mass spectrometer (Applied Biosystems/Sciex, Concord, ON, Canada) with CE injection system (Prince Technologies, Netherlands). CE separation was obtained on a 90-cm length of bare fused-silica capillary (365 μm OD × 50 μm ID) with CE-MS coupling using a liquid sheath-flow interface and isopropyl alcohol:methanol (2:1) as the sheath liquid. An aqueous buffer consisting of 30 mM morpholine was used for all experiments in the negative-ion mode. The MS data are presented in supplemental Table S3.

Overexpression, Localization, and Purification of Waa Proteins*—Overexpression and cellular location of the His_6-WaaT and His_6-WaaR enzymes was monitored by Western immunoblotting, essentially as described previously for His_6-WaaJ (6). Cell-free lysates of *E. coli* CWG309 containing pWQ271 and *E. coli* TOP10 containing pWQ270 were separated by ultracentrifugation. The soluble fraction was collected and the membrane pellet was washed twice with 2 ml of 50 mM Tris-HCl, pH 7.5. The fraction volumes were adjusted to facilitate direct comparison of the amount of membrane protein corresponding to a given amount of soluble protein. Protein samples were separated by SDS-PAGE and transferred onto polyvinylidene difluoride membrane. The Western immunoblots were developed using HisProbe H3 mouse anti-His_6 primary antibody (Santa Cruz Biotechnology, Santa Cruz, CA), and goat anti-mouse alkaline phosphatase-conjugated secondary antibody (Jackson ImmunoResearch Laboratories, Inc., Montréal, QB). Nitro blue tetrazolium from Sigma and 5-bromo-4-chloro-3-indolylphosphate from Roche were used as substrates to develop the Western blots. The bands were quantified by densitometry using a Bio-Rad GS-800 Calibrated Densitometer with QuantityOne software.

The purification of His_6-WaaJ from overexpression in *E. coli* TOP10 cells has been described elsewhere (6). Purification of His_6-WaaT was done in a similar fashion using Ni^{2+}-NTA affinity chromatography, except batch binding was done directly using cell-free lysate, without prior removal of membrane material. In addition, 150 mM NaCl and 7.5 mM imidazole were included in the lysis buffer during sonication and in initial washing of the affinity column. The purified proteins were dispensed in 0.2-ml aliquots for single use, and total protein concentration was determined using the Bio-Rad Protein Assay with bovine serum albumin as the standard. His_6-WaaJ and His_6-WaaT were stable in this form at 4 °C for periods of up to 2 weeks. Storage at −20 °C offered no additional stability and enzyme samples thawed after storage at −80 °C showed significant loss of activity. It should be noted that the level of overexpression of His_6-WaaT was lower than His_6-WaaJ. Yields of purified protein were also compromised by the tendency of His_6-WaaT to adsorb to filters used in concentration.

In Vitro Determination of the Activity of His_6-Waa Enzymes*—The activity of the various constructs was determined as previously described (6). The reaction is based on the transfer of radioactivity from [^{14}C-Glc]UDP (Perkin Elmer, 200.0 mCi/mmol) or [^{14}C-Gal]UDP (Perkin Elmer, 258.0 mCi/mmol) donor to an acceptor comprising the LPS isolated from *E. coli* CWG350 (*waaJ:aacC1*) or CWG309 (*waaT:aacC1*). Final reaction conditions after the addition of enzyme were: 1–750 μM UDP-sugar donor, 2–750 μM LPS, 100 mM Tris pH 7.5, 0.4 mM EDTA, 5 mM $MgCl_2$, and 200–900 nM His_6-WaaJ or His_6-WaaT in a final volume of 0.1 ml. The His_6-WaaJ or His_6-WaaT concentration was adjusted to ensure that it was always at least 5-fold below the lowest substrate concentration. The rates at each substrate concentration were then fit to the Michaelis-Menten equation to determine k_{cat} and K_m. These values should be considered as "apparent" because of the stopped nature of the assay.

RESULTS

Comparison of WaaJ, WaaR, and WaaT Sequences—Previous work from our laboratory established the *in vivo* and *in vitro* properties of WT WaaJ (6). The purpose of this work is to expand that analysis to other CAZy family 8 glycosyltransferases involved in *E. coli* LPS biosynthesis, particularly ones employing alternate donor sugars and terminal acceptor linkages.

As might be anticipated for *E. coli* LPS core oligosaccharide biosynthesis proteins that all belong to glycosyltransferase family 8, WaaJ, WaaR, and WaaT share significant primary sequence identity (~40%) and similarity (~60%) and several highly conserved regions (supplemental Fig. S2). The sequences

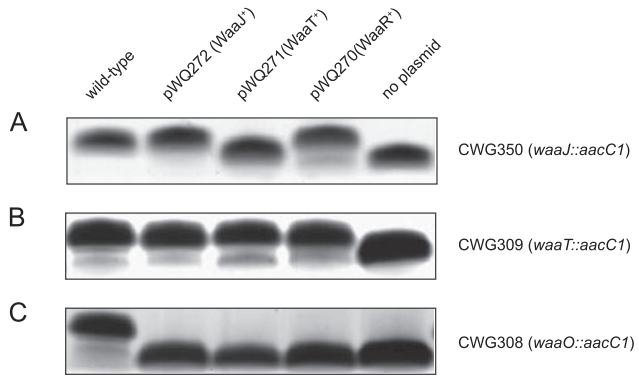

FIGURE 2. ***In vivo* complementation activity of His_6-Waa* enzymes.** The LPS samples from proteinase K-digested whole cell lysates were separated by PAGE and visualized by silver staining. Plasmids encoding the various derivatives were used to transform. *Panel A*, *E. coli* CWG350 *waaJ*::*aacC1* cells, with F653 as wild-type reference; *panel B*, *E. coli* CWG309 *waaT*::*aacC1* cells, with F470 as wild-type reference; and *panel C*, *E. coli* CWG308 *waaO*::*aacC1* cells with F470 as wild-type reference.

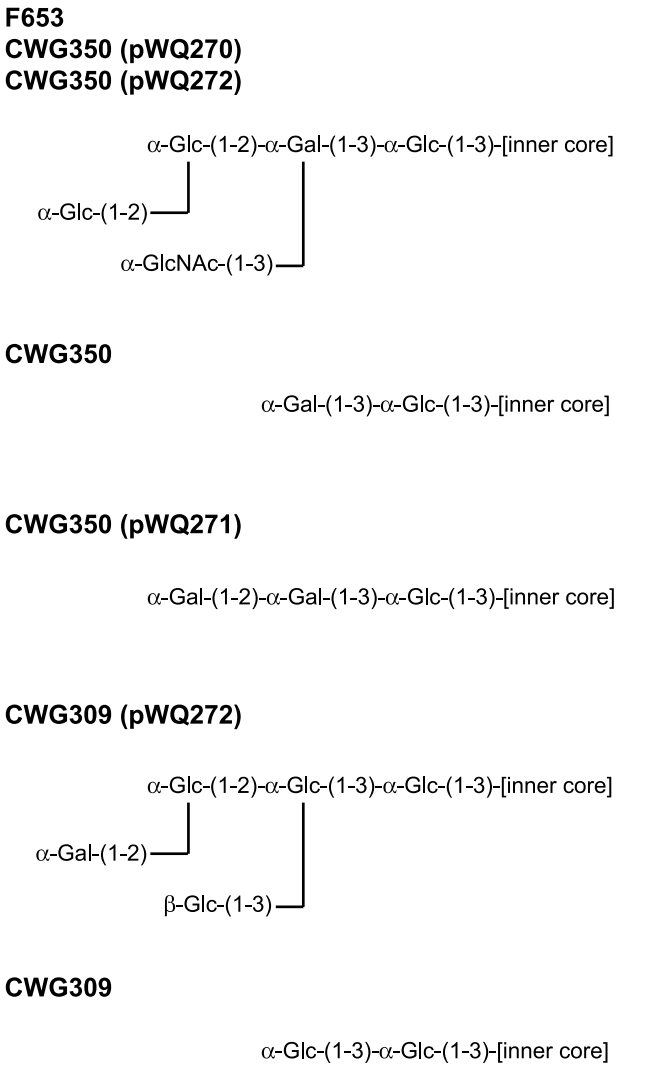

FIGURE 3. **Schematic arrangement and linkages of sugars for the outer core region of O-deacylated *E. coli* LPS as determined by NMR.** The CWG309 structure is reproduced from Ref. 8.

of these proteins predict no transmembrane helices. However, their nascent lipid A-core OS acceptor is membrane-associated, indicating that a membrane association might be beneficial for activity. His_6-WaaT was found to be 75% membrane-associated, and His_6-WaaR was found to be 64% membrane-associated (data not shown), compared with 55% for His_6-WaaJ (6). As described for WaaJ (6), the WaaT and WaaR sequences could be threaded onto the LgtC crystal structure with good alignment of secondary structure (data not shown), but the difference in primary sequences limited the usefulness of this information and precluded any meaningful prediction of residues that might dictate specificity.

Function of His_6-Waa Derivatives in Vivo*—To investigate whether His_6-WaaT, His_6-WaaJ, and His_6-WaaR could function to extend *waaJ:aacC1* or *waaT:aacC1* LPS *in vivo*, plasmids encoding the relevant protein were transformed into CWG350 (*waaJ:aacC1*) and CWG309 (*waaT:aacC1*) and the resulting LPS profile was investigated by silver-stained PAGE. As shown in Fig. 2, WaaR was able to extend the LPS acceptor from CWG350 (*waaJ:aacC1*) to generate a product that comigrated with the wild-type LPS, despite the fact that the native acceptor residue for WaaR is a 1,3-linked Glc, rather than the 1,3-linked Gal provided by CWG350 LPS. The CWG350 LPS acceptor lacks three hexose residues (see below) compared with the wild type and migrates significantly faster, thus WaaR is able to form a product that can then be further extended by additional glycosyltransferases. WaaT was also able to extend the 1,3-linked Gal acceptor but the product remained smaller than the wild type; its migration was consistent with the absence of a single residue. Each of the glycosyltransferases was able to extend the 1,3-linked Glc acceptor provided by CWG309 (*waaT:aacC1*) LPS to restore a wild-type profile, indicating that WaaW (Fig. 1) could utilize their products to add the final core OS residue in the F470 (R1 core OS) background.

To further probe the acceptor specificities, each of these enzymes were expressed in CWG308 (*waaO:aacC1*) cells. Interestingly, no core OS extension occurred in this mutant (Fig. 2*C*). Despite the observation (above), indicating that WaaJ, WaaR, and WaaT did not discriminate between 1,3-linked Gal or 1,3-linked Glc acceptors, the enzymes are apparently sensitive to the position of the acceptor residue in the context of the core OS structure.

While silver-stained PAGE LPS profiling is a facile method to test the ability of a protein to extend a given LPS, it does not give any information about the identity of the sugars added. To identify the sugars added in the above LPS extensions, LPS was isolated from CWG350, CWG350 (pWQ271), CWG350 (pWQ270), CWG350 (pWQ272), and CWG309 (pWQ272). The structures of the LPS molecules were determined by NMR and MS methods (supplemental Tables S2 and S3). The resulting structures are shown graphically in Fig. 3. The CWG350 LPS contains a Gal- (1, 2)-Glc outer core and lacks the 2 terminal residues expected from the *waaJ* mutation (Fig. 1). Also missing is the side-chain GlcNAc residue. This result was consistent with the PAGE profile with respect to the size of the products but differs from the CWG350 structure obtained previously, where traces of GlcNAc were reported (7). The NMR analysis of the CWG350 core OS was repeated and no evidence of the GlcNAc residue could be detected; there was no detect-

TABLE 1

Kinetic parameters for His_6-WaaJ and His_6-WaaT glycosyltransferases

Michaelis-Menten parameters were calculated for the glycosyltransferase activity of His_6-WaaJ and His_6-WaaT using the indicated donor nucleotide sugars and purified *E. coli* LPS acceptor at 37 °C and pH 7.5. The values should be considered as "apparent" because of the stopped nature of the assay.

Protein	Acceptor LPS	Donor	k_{cat}[a]	K_m(LPS)[b]	K_m(donor)[c]
			min^{-1}	μM	μM
WaaJ	CWG350 (*waaJ*::*aacC1*)	UDP-Glc	28 ± 4[d]	11 ± 3[d]	32 ± 8[d]
	CWG350 (*waaJ*::*aacC1*)	UDP-Gal	4.1 ± 0.3	27 ± 7	200 ± 30
	CWG309 (*waaT*::*aacC1*)	UDP-Glc	8.2 ± 0.5	35 ± 3	40 ± 5
WaaT	CWG350 (*waaJ*::*aacC1*)	UDP-Gal	3.7 ± 0.5	25 ± 8	36 ± 9
	CWG309 (*waaT*::*aacC1*)	UDP-Gal	4.6 ± 0.2	37 ± 3	42 ± 7
	CWG309 (*waaT*::*aacC1*)	UDP-Glc	ND[e]	ND[e]	> 300[e]

[a] Data were calculated as the average and propagated error of the k_{cat} values as determined from the acceptor and donor Michaelis-Menten fits.
[b] Values were determined at saturating concentrations of UDP-sugar donor (150 μM).
[c] Values were determined at saturating concentrations of LPS acceptor (150 μM).
[d] Values reproduced from Ref. 6.
[e] Limitations of protein concentration and reduced catalytic activity prevented full characterization. See "Results" for details.

able peak corresponding to an N-acetyl group. Moreover, the MS data also reflected a core OS species lacking GlcNAc. The reason for this disparity between the structures of CWG350 is unknown and the results shown below indicate that the currently unidentified transferase required for GlcNAc addition (7) is still active in the strains used here.

In CWG350 cells expressing either WaaR or WaaJ, the authentic LPS structure was restored. WaaR is normally a UDP-glucose:(glucosyl) LPS α-1,2-glucosyltransferase. However, when active in the context of CWG350, it shows UDP-glucose:(galactosyl) LPS α-1,2-glucosyltransferase activity. As expected, the LPS molecule resulting from WaaR or WaaJ activity serves as an acceptor for WaaD and the still unidentified GlcNAc transferase to complete the core OS structure. In contrast, WaaT added only a single residue to CWG350 LPS. While WaaT is a UDP-galactose:(glucosyl) LPS α-1,2-galactosyltransferase in its wildtype background, in CWG350 it exhibits UDP-galactose:(galactosyl) LPS α-1,2-galactosyltransferase activity. It therefore retains its normal donor specificity but can effectively utilize a different acceptor. The resulting product is not further extended by either WaaD or the GlcNAc transferase, indicating that these transferases are either sensitive to perturbations in acceptor structure, or lose critical protein-protein interactions when the precise combination of glycosyltransferase enzymes changes.

CWG309 cells expressing WaaJ were able to form a complete core OS. However, Glc was incorporated as the second-to-last hexose in the core OS backbone. WaaJ is normally a UDP-glucose:(galactosyl) LPS α-1,2-glucosyltransferase. Thus, in the CWG309 background, WaaJ shows UDP-glucose:(glucosyl) LPS α-1,2-glucosyltransferase activity. This demonstrates that, like WaaT and WaaR, it also retains its normal donor specificity in a heterologous background. The published structure for CWG309 core OS lacks the β-Glc sidechain, suggesting this residue is added late in the assembly process (8). The full extension of the CWG309 (pWQ272) LPS product implies that WaaV and WaaW are unaffected by the local changes in their acceptor structure. This differs from the situation described above for WaaD acting in CWG350 (Fig. 1).

In Vitro Kinetic Behavior of His_6-WaaJ and His_6-WaaT Proteins—The *in vivo* results demonstrate two basic principles: (i) His_6-WaaT, His_6-WaaR, and His_6-WaaJ have a marked preference for their specific donor substrates; (ii) these enzymes appear to have a relaxed specificity for alternate LPS acceptor residues, providing the overall size of the acceptor is conserved. His_6-WaaJ and His_6-WaaT were therefore selected for *in vitro* analysis to more precisely determine the effect of alternate donor and acceptor substrates on their kinetic properties. Those results are summarized in Table 1.

The results of His_6-WaaJ assays with UDP-Glc and its native (CWG350 (*waaJ:aacC1*)) LPS acceptor were reported previously (6). As expected, changing the donor substrate to UDP-Gal resulted in significant drop in enzyme activity, reflected in a 6-fold decrease in k_{cat} and 6-fold increase in K_m(donor). Interestingly, the K_m(LPS) also more than doubled, despite the fact that the identity of the LPS in each experimental series was identical. Therefore, changing the donor altered not only the K_m(donor) and k_{cat}, but, surprisingly, also the parameters for the acceptor, K_m(LPS). In all cases, the enzyme preparations were free of an UDP-Gal-4-epimerase enzyme (15) that might produce the alternate donor *in situ* and generate misleading results (data not shown).

A similar trend was obtained with His_6-WaaT using the CWG309 (*waaT:aacC1*) LPS acceptor but different sugar nucleotide donors. His_6-WaaT readily used UDP-Gal as the donor sugar, as would be expected since this is the physiological donor in its wildtype background. In contrast, there was a significant decrease in k_{cat} when UDP-Glc was used instead. Unfortunately, full kinetic characterization was not possible for this reaction condition. The lower overexpression of His_6-WaaT and its tendency to adsorb to membrane filters limited the ability to reach high stock concentrations of the protein. The decrease in the activity of His_6-WaaT with UDP-Glc was so severe that at the protein concentrations reached, the signal was only slightly (but reproducibly) above background at the highest substrate concentrations tested. Therefore, no k_{cat} or K_m(LPS) are reported. Nevertheless, the value for K_m(UDP-Glc) can be estimated to be greater than 300 μM, because no radioactive incorporation was detected at donor substrate concentrations lower than this value.

The effect of changing the LPS acceptor on His_6-WaaJ activity was studied with UDP-Glc and purified CWG309 (*waaT:aacC1*) LPS. The K_m(donor) was identical and while the k_{cat} did decrease 3-fold, it still remains twice the value obtained with the physiological acceptor and the alternate (UDP-Gal) donor. The K_m(LPS) increased 3-fold relative to its physiological LPS acceptor. Based on the NMR data and LPS PAGE profile of the LPS products, these changes are apparently insufficient to pre-

vent His_6-WaaJ from fully processing CWG309 (*waaT:aacC1*) LPS acceptor *in vivo*.

The effect of changing the LPS acceptor on His_6-WaaT activity was studied with UDP-Gal and purified CWG350 (*waaJ*::*aacC1*) LPS. In contrast to the His_6-WaaJ results, none of the kinetic parameters of His_6-WaaT changed markedly when it was examined in assays containing UDP-Gal donor and CWG350 (*waaJ*::*aacC1*) acceptor.

DISCUSSION

Glycoconjugates serve roles as diverse as structural components, energy storage, antibiotic diversity, and immunological recognition. The specific sugars used and the linkages formed in the glycoconjugates are keys to their function and, due to their biological importance, there has been increasing interest in the commercial scale synthesis of defined complex carbohydrate structures. Unfortunately, their reactivity and structural and linkage diversity has meant generalized solid phase chemical synthesis for carbohydrates is not as mature a technology as the corresponding methods developed for nucleic acids and peptides (reviewed in Ref. 16). Therefore, chemical and enzymatic synthesis has been used to create natural carbohydrates (reviewed in Ref. 17). An understanding of the molecular bases for glycosyltransferase specificity is crucial for engineering glycoconjugates of medical or industrial importance because the inherent specificity potentially limits what can be achieved in enzymatic syntheses.

The group of related glycosyltransferase enzymes from CAZy family 8 provide an excellent opportunity to investigate issues involving substrate specificity. The *in vitro* kinetic results demonstrate that WaaJ and WaaT have a high selectivity for the correct donor sugar nucleotide. In reactions containing the C4-epimer of the natural donor, k_{cat} is decreased and there are corresponding increases not only in the K_m(donor) but also, to a lesser extent, in the K_m for the native LPS acceptor. This selectivity is born out by the structure of the *in vivo* product; the preferred donor substrate is used regardless of the nature of the available acceptor. LgtC has also exhibited a strong preference *in vitro* for the correct donor epimer, exhibiting a 3400-fold decrease in k_{cat} when UDP-Glc was used instead of UDP-Gal (18). A rare exception to this rule, CstII has been shown *in vivo* and *in vitro* to utilize KDN as a donor in addition to sialic acid, albeit at much reduced efficiency (19).

The selectivity for acceptor substrates is not as predictable. Crystal structures and a proposed reaction mechanism are available for one CAZy family 8 representative, the LgtC The ordered bi-bi kinetics of LgtC indicates that the donor sugar nucleotide binds to the protein first, followed by the acceptor substrate. Additionally, the crystal structure shows the donor enclosed in a binding pocket (20). If the closing of the donor binding pocket creates the binding site for the acceptor, then any perturbation in this conformation could potentially affect interaction with the acceptor and therefore the K_m(LPS). This could explain the different K_m(LPS) values obtained from His_6-WaaJ using UDP-Gal *versus* UDP-Glc donors.

Despite the relationships within the CAZy family 8, sequence comparisons do not shed light onto the molecular bases for substrate specificity. The amino acid sequences of the CAZy family 8 glycosyltransferases WaaR, WaaT, and WaaJ only share ~40% identity, most of which is located in regions of high conservation such as the DXD motifs. While mutagenesis of these regions in these proteins has not been performed, this degree of conservation argues that the roles are also conserved. These motifs have been investigated in LgtC, a galactosyltransferase involved in the biosynthesis of the lipooligosaccharide from *Neisseria meningitidis* (20–22). The first DXD motif is predicted to be involved in the binding of the divalent metal ion, while the first aspartate (Asp^{188}) of the second DXD motif is involved in interactions with the galactose ring of the UDP-Gal donor. In the first crystal structure, these are direct contacts (20), while in the second crystal structure, these are indirect contacts (18). The true role of the second aspartate (Asp^{190}) in the second DXD motif is also unclear: in the first structure, it is rotated several Angstroms away from the anomeric carbon, leading to the proposal of an S_Ni reaction mechanism involving Gln^{189} (20). However, this second aspartate was identified by mass spectrometry as the amino acid covalently linked to the donor sugar when UDP-Glc or UDP-Gal was incubated with the Q189E mutant enzyme (18). These crystal structures were solved with donor analogs, rather than physiological donor and acceptor substrates. Therefore the differing results can be resolved if there is a significant reorganization of the active site during catalysis, allowing Asp^{190} to attack the anomeric carbon.

His_6-WaaJ exhibited lower *in vitro* activity with CWG309 (*waaT*::*aacC1*) LPS than with its native acceptor but was still able to catalyze full extension of both acceptors *in vivo*, based on observations from PAGE and NMR studies. In contrast, His_6-WaaT showed similar *in vitro* activity against CWG350 (*waaJ*::*aacC1*) LPS and its native acceptor. The activity of these two family 8 enzymes using non-native acceptors is perhaps surprising but may reflect the genetic backgrounds in which these enzymes have evolved. WaaT is never naturally found in *E. coli* containing the WaaI enzyme (which adds Gal to the growing acceptor; see Fig. 1 and Ref. 4). Similarly, WaaJ is never found in *E. coli* containing the corresponding WaaO enzyme (which adds Glc; Fig. 1). Because these enzymes see only one acceptor type, the need for discrimination between acceptors is limited. In contrast, the enzymes occupy an environment where both potential donors (UDP-Glc and UDP-Gal) are present, so the need for stringent selection of donor is clear. The fact that the enzymes incorporated their correct donor sugars even in non-natural cellular backgrounds containing different LPS acceptor structures demonstrates the power of this donor specificity.

Interestingly, neither His_6-WaaJ nor His_6-WaaT could extent the LPS in CWG308 (*waaO*::*aacC1*). This may reflect an apparent requirement for an acceptor of minimal size. This is also supported by the observation that His_6-WaaJ did not elongate FCHASE-Gal or FCHASE-Lac synthetic acceptors (6). However, the potential impact of altered conformation in the shorter acceptor resulting from its proximity to the inner core (heptose residue) must also be taken into account.

The results presented here for a group of related CAZy family 8 glycosyltransferases indicate that the enzymes may have more flexibility for both donor and acceptor than anticipated, when studied *in vitro*. However, the extent of this flexibility is limited

in vivo due to the strict preference for a specific donor. These observations have significance for both *in vitro* and *in vivo* glycoengineering applications. Furthermore, the flexibility of these glycosyltransferase enzymes seen *in vivo* experiments suggests a need for caution in interpreting genetic "complementation" data if the product analyses are confined to LPS PAGE profiles. While a residue may be added at a particular point in the extension process to support product completion, the identity of the added residue may only be established by further structural analysis.

Acknowledgment—We thank Jacek Stupak for recording of mass spectra.

REFERENCES

1. Raetz, C. R. H., and Whitfield, C. (2002) *Annu. Rev. Biochem.* **71,** 635–700
2. Amor, K., Heinrichs, D. E., Frirdich, E., Ziebell, K., Johnson, R. P., and Whitfield, C. (2000) *Infect. Immun.* **68,** 1116–1124
3. Kaniuk, N. A., Monteiro, M. A., Parker, C. T., and Whitfield, C. (2002) *Mol. Microbiol.* **46,** 1305–1318
4. Heinrichs, D. E., Yethon, J. A., and Whitfield, C. (1998) *Mol. Microbiol.* **30,** 221–232
5. Kadam, S. K., Rehemtulla, A., and Sanderson, K. E. (1985) *J. Bacteriol.* **161,** 277–284
6. Leipold, M. D., Kaniuk, N. A., and Whitfield, C. (2007) *J. Biol. Chem.* **282,** 1257–1264
7. Kaniuk, N. A., Vinogradov, E., Li, J., Monteiro, M. A., and Whitfield, C. (2004) *J. Biol. Chem.* **279,** 31237–31250
8. Heinrichs, D. E., Yethon, J. A., Amor, P. A., and Whitfield, C. (1998) *J. Biol. Chem.* **273,** 29497–29505
9. Schmidt, G., Fromme, I., and Mayer, H. (1970) *Eur. J. Biochem.* **14,** 357–366
10. Guzman, L.-M., Belin, D., Carson, M. J., and Beckwith, J. (1995) *J. Bacteriol.* **177,** 4121–4130
11. Hitchcock, P. J., and Brown, T. M. (1983) *J. Bacteriol.* **154,** 269–277
12. Tsai, C. M., and Frasch, C. E. (1982) *Anal. Biochem.* **119,** 115–119
13. Galanos, C., Luderitz, O., and Westphal, O. (1969) *Eur. J. Biochem.* **9,** 245–249
14. Ciucanu, I., and Kerek, F. (1984) *Carbohydr. Res.* **131,** 209–217
15. Wilson, D. B., and Hogness, D. S. (1964) *J. Biol. Chem.* **239,** 2469–2481
16. Seeberger, P. H., and Werz, D. B. (2005) *Nat. Rev. Drug Discov.* **4,** 751–763
17. Thibodeaux, C. J., Melancon, C. E., and Liu, H.-W. (2007) *Nature* **446,** 1008–1016
18. Lairson, L. L., Chiu, C. P. C., Ly, H. D., He, S., Wakarchuk, W. W., Strynadka, N. C. J., and Withers, S. G. (2004) *J. Biol. Chem.* **279,** 28339–28344
19. Aharoni, A., Thieme, K., Chiu, C. P. C., Buchini, S., Lairson, L. L., Chen, H., Strynadka, N. C. J., Wakarchuk, W. W., and Withers, S. G. (2006) *Nat. Methods* **3,** 609–614
20. Persson, K., Ly, H. D., Dieckelmann, M., Wakarchuk, W. W., Withers, S. G., and Strynadka, N. C. J. (2001) *Nat. Struct. Biol.* **8,** 166–175
21. Gotschlich, E. C. (1994) *J. Exp. Med.* **180,** 2181–2190
22. Yang, Q. L., and Gotschlich, E. C. (1996) *J. Exp. Med.* **183,** 323–327
23. Jansson, P.-E., Lindberg, A. A., Lindberg, B., and Wollin, R. (1981) *Eur. J. Biochem.* **115,** 571–577
24. Heinrichs, D. E., Monteiro, M. A., Perry, M. B., and Whitfield, C. (1998) *J. Biol. Chem.* **273,** 8849–8859
25. Muller-Loennies, S., Lindner, B., and Brade, H. (2002) *Eur. J. Biochem.* **269,** 5982–5991

Glycosyltransferases Involved in Biosynthesis of the Outer Core Region of *Escherichia coli* Lipopolysaccharides Exhibit Broader Substrate Specificities Than Is Predicted from Lipopolysaccharide Structures

Michael D. Leipold, Evgeny Vinogradov and Chris Whitfield

J. Biol. Chem. 2007, 282:26786-26792.
doi: 10.1074/jbc.M704131200 originally published online July 13, 2007
