Acceptor substrate specificity of UDP-Gal: GlcNAc-R β1,3-galactosyltransferase (WbbD) from *Escherichia coli* O7:K1

Inka Brockhausen · John G. Riley · Meileen Joynt · Xiaojing Yang · Walter A. Szarek

Abstract Most of the glycosyltransferases involved in O antigen biosynthesis have not yet been characterized. We recently demonstrated that the *wbbD* gene of the O7 lipopolysaccharide biosynthesis cluster in *E. coli* strain VW187 (O7:K1) encodes WbbD, a UDP-Gal: GlcNAcα-pyrophosphate-lipid β1,3-Gal-transferase (EC 2.4.1., accession number AAC27537) that transfers the second sugar moiety in the assembly of the O7 repeating unit. The enzyme utilizes undecaprenol-pyrophosphate-GlcNAc as a natural acceptor substrate, but can also transfer Gal to GlcNAcα-PO_3-PO_3-$(CH_2)_{11}$-O-phenyl (GlcNAc-PP-PhU). A number of acceptor substrate analogs have now been tested to further characterize the acceptor specificity of WbbD and to determine the roles of the pyrophosphate bond and the lipid moiety in the acceptor substrate. The enzyme was found to have a low activity with a substrate containing only one phosphate group directly α-linked to GlcNAc, and the enzyme was inactive when the phosphate was absent or further removed from the anomeric carbon of GlcNAc. Modifications of the lipid chain yielded substrates with variable activities. GlcNAc derivatives that were inactive as substrates did not inhibit WbbD suggesting that these compounds did not bind to the active site of the enzyme. The specificity of mammalian β4-galactosyltransferase I has been compared to that of WbbD. The results indicate that the bacterial WbbD enzyme has a distinct specificity for GlcNAc-PP-lipid, and that WbbD recognition of its acceptor substrate is very different from that of the ubiquitous mammalian β4-galactosyltransferase I. These studies help to understand mechanisms of O antigen synthesis, to develop methods to synthesize defined oligosaccharide structures and to develop specific O antigen inhibitors.

Key words *E. coli* · O antigen synthesis · Galactosyltransferase · Substrate specificity · Undecaprenol-pyrophosphate · Enzyme assay

Abbreviations

EDTA	ethylene diamine tetraacetic acid
GlcNAc-PP-PhU	GlcNAcα-pyrophosphate-phenylundecyl
HPLC	high pressure liquid chromatography
LPS	lipopolysaccharide
MALDI	matrix-assisted laser desorption ionization
MES	2-*N*-morpholinoethanesulfonate
MS	mass spectrometry
NMR	nuclear magnetic resonance
PBS	phosphate buffered saline
Rha	rhamnose
VioNAc	4-acetamido-4,6-deoxy-D-glucose

I. Brockhausen (✉) · J. G. Riley · M. Joynt · X. Yang
Department of Biochemistry, Department of Medicine,
Division of Rheumatology, and The Arthritis Centre and Human
Mobility Research Centre, Queen's University, Etherington Hall,
Kingston, ON K7L 3N6, Canada
e-mail: brockhau@queensu.ca

W. A. Szarek
Department of Chemistry, Queen's University,
Kingston, ON K7L 3N6, Canada

Introduction

Lipopolysaccharides (LPS) of Gram negative bacteria are important components of the outer leaflet of the bacterial membrane, and are essential for membrane stability and strain viability. The outermost O antigenic polysaccharide of the LPS of *Escherichia coli* can be used to classify these bacteria into almost 200 different serotypes (http://www. casper.organ.su.se/ECODAB/). As one example, the O7 antigen consists of repeating pentasaccharides (repeating units) with the structure 3-VioNAc (4-acetamido-4,6-deoxy-D-glucose)β1–2 (Rha (rhamnose)α1–3) Manα1–4 Galβ1–3 GlcNAcα1- [1]. The O antigens bind bacteriophages, and are also important virulence factors. Inhibitors of O antigen synthesis have the potential of rendering bacteria susceptible to the host immune system.

A number of genes and pathways have been proposed to be involved in the biosynthesis of O antigens [2–4]. Based on sequence similarity, many of these genes have been proposed to encode 'putative' glycosyltransferases, but most of these enzymes have not been assayed and characterized.

The repeating units of the O antigen are assembled on an undecaprenol-phosphate lipid carrier. The enzymes and acceptor substrates for O antigen synthesis are thought to be associated with the inner membrane, while nucleotide-sugar donor substrates for these reactions are synthesized in the cytoplasm. In the case of the O7 antigen biosynthesis in the *E. coli* strain VW187, the first step of the O antigen repeating unit synthesis involves the transfer of GlcNAcα-phosphate from UDP-GlcNAc to undecaprenol-phosphate to form GlcNAc-pyrophosphate-undecaprenol. This first reaction is catalyzed by a GlcNAc-phospho-transferase WecA that is present on the cytoplasmic face of the inner membrane [5]. The second step of O7 synthesis, the addition of a Gal residue to GlcNAc-pyrophosphate-lipid, is catalyzed by WbbD, a β1,3-Gal-transferase that utilizes UDP-Gal as the donor substrate [6, 7]. We characterized WbbD and determined that the enzyme is essential for O antigen synthesis in the VW187 strain. The enzyme requires Mn^{2+}, and detergents were found to decrease the activity. A new synthetic substrate, GlcNAcα-PO_3-PO_3-$(CH_2)_{11}$-O-phenyl (GlcNAc-PP-PhU), which is an analog of GlcNAc-pyrophosphate-undecaprenol, was found to be an excellent acceptor for Gal transfer. The enzyme has been classified as a carbohydrate-active enzyme (CAZy) GT2 family member, EC 2.4.1., accession number AAC27537 (http://www.cazy.org/fam/GT2.html).

After the addition of Gal to GlcNAc-pyrophosphate-lipid, the remaining sugars are added to form the O7 repeating unit. The sequence of these additions and the enzymes and substrates involved have not yet been characterized. After the formation of the repeating unit on undecaprenol-pyrophosphate, the oligosaccharide is most likely flipped across the inner membrane by a flippase Wzx, and the units are polymerized by a polymerase Wzy. The final length of the O antigen is determined by a chain length regulator Wzz. Lipid A is assembled simultaneously on the inner membrane, and is then processed by the addition of inner and outer core oligosaccharides to lipid A [3, 5, 8]. The complete O polysaccharide is then ligated by a putative ligase WaaL to the outer core-lipid A acceptor to form LPS which is translocated to the outer membrane of the cell and is exposed to the exterior of the cell [2, 3]. In a second mechanism, monopolymeric O polysaccharide chains can be assembled in an ATP-binding cassette transporter-dependent pathway where the entire oligosaccharide is synthesized on the cytoplasmic side of the inner membrane, followed by translocation and ligation to the lipid A core oligosaccharide [8].

The Gal–GlcNAc linkage is found in a number of O-antigens and is also a common backbone structure in mammalian glycoproteins and glycolipids. One of our goals is to develop specific inhibitors for O antigen synthesis that do not affect mammalian enzymes. However, the activity of bacterial WbbD has many similarities to the ubiquitous mammalian β4-Gal-transferase activity. For both the bacterial and mammalian enzymes, UDP-Gal is the donor substrate, and GlcNAc derivatives that contain a hydrophobic group work well as acceptor substrates. Both types of enzymes require Mn^{2+} or other divalent metal ions for activity. Mammalian β4Gal-transferase 1 is a ubiquitous and important β4-Gal-transferase; like WbbD, it has a GT-A type fold and a DxD motif that may be involved in catalysis [9, 10]. However, the mammalian Gal-transferase is a member of the CAZy GT7 family (http://www.cazy.org).

In order to further understand the recognition by WbbD of its acceptor substrate, we synthesized a series of undecaprenol-pyrophosphate-lipid analogs as acceptor substrates. We found that WbbD efficiently transfers Gal to GlcNAc when the pyrophosphate linkage is present in the acceptor substrate. The mammalian β1,4-Gal-transferase has a very different substrate specificity and prefers substrates that are inactive with WbbD, such as GlcNAcβ-benzyl. None of the inactive compounds, including a potent inhibitor of mammalian β1,4-Gal-transferase, inhibited WbbD activity. These differences between mammalian and bacterial enzymes could be exploited in the design of specific glycosyltransferase inhibitors that block the synthesis of O antigens. The information on enzyme specificity is also important for our understanding of the mechanisms of O antigen synthesis and to prepare oligosaccharides with defined structures.

Materials and methods

Materials

The synthesis of acceptor substrates has been described previously [7, 10] or will be reported elsewhere [J. Riley *et al.*, in preparation]. Glycopeptides were synthesized by Hans Paulsen, University of Hamburg, Germany. All other reagents were purchased from Sigma Chemical Co., St. Louis, MO, unless otherwise indicated. Compounds were analyzed by electrospray mass spectrometry (ESI-MS) in the negative ion mode, matrix-assisted laser desorption ionization-MS, NMR and high pressure liquid chromatography (HPLC) for purity and to determine the best conditions for enzyme product separation.

Enzyme preparations

Bacterial cultures of *E. coli* VW187 (*E. coli* O7:K1) and a WbbD mutant containing plasmid pCM227 were grown as described [6, 7]. Bacteria from one 100 mL culture flask were washed with phosphate buffered saline (PBS), and resuspended in 12 mL of PBS, pH 7.2/glycerol (9:1), and were stored in aliquots at −20°C. The bacterial cells were sedimented by centrifugation, washed with 1 mL of PBS, pH 7.2, resuspended in 0.5 mL sonication buffer (10 mM Tris/acetate, pH 8.5, 50 mM sucrose, 1.2 mM EDTA (ethylene diamine tetraacetic acid) and sonically ruptured on ice (twice for 15 s with 2 min between, at setting three of a Sonic Dismembrator Model 100 from Fisher Scientific). Sonicates were used immediately for enzyme assays. Protein concentrations were determined with the BioRad (Bradford) protein assay using bovine serum albumin as a standard.

Enzyme assays for Gal transfer to exogenous acceptors

Standard assays for galactosyl transfer to exogenously added substrate were carried out in reaction mixtures of 40 μL total volume, containing the standard substrate 0.2 mM GlcNAc-PP-PhU (compound 30) or other substrate analogs, 10% methanol, 10 μL sonication buffer, 5 mM $MnCl_2$, 75 mM MES buffer, pH 7, 0.5 mM UDP-[^{3}H]Gal (800 to 4,400 cpm/nmol), and 10 μL enzyme homogenate (30–50 μg protein). Controls had no exogenous acceptor substrate. Mixtures were incubated for 10 min at 37°C. The reaction was stopped by the addition of 0.7 mL of ice cold water. For assays containing hydrophobic substrates, the mixture was applied to a regenerated C18 Sep-Pak column, followed by washing with 5 mL water. Enzyme product was eluted with 3 mL methanol. The radioactivity in the

last 1 mL fraction eluted with water and in the first three 1 mL fractions eluted with methanol was determined by scintillation counting. Alternatively, enzyme products of neutral substrates were isolated by anion exchange chromatography using AG1x8 as previously described [10]. All assays were carried out at least in duplicate determinations. Substrates and enzyme products were analyzed by reverse phase HPLC, using a C18 column and acetonitrile/water mixtures and by mass spectrometry [7].

Mammalian β1,4-Gal-transferase I (lactose synthase, Sigma) was assayed as described using 0.2 or 0.5 mM GlcNAcβ-Bn as the acceptor substrate [10]. After incubation for 30 min, AG1x8 chromatography was used for neutral acceptor substrates, or C18 Sep-Pak columns for acceptor substrates containing negatively charged phosphate groups and lipid chains [10]. Homogenates prepared from HT29 cells [11], that contain membrane-bound Gal-transferase I, were tested under the same conditions but in the presence of 0.125% Triton X-100 and without the addition of bovine serum albumin.

Results

Acceptor substrate specificity of WbbD β 3-Gal-transferase

The general catalytic properties of UDP-Gal: GlcNAc-R β3Gal-transferase WbbD have previously been determined [7]. The analysis of the amino acid sequence showed that the enzyme has a short hydrophobic sequence at the C terminus that may help to localize the enzyme to the inner surface of the plasma membrane. The enzyme has been classified into the CAZy glycosyltransferase family GT2. It has a GT-A type fold with a 3-dimensional conservation of a potential nucleotide binding fold (Dr. Christelle Breton, personal communication) and a DxD sequence in the vicinity of this fold (data not shown). The ubiquitous bovine β4-Gal-transferase I has been classified in a different glycosyltransferase family, CAZy GT7, but has many properties in common with WbbD. Both enzymes are inverting transferases, act on GlcNAc-R acceptors, bind UDP-Gal and require divalent metal ion for activity. However, alignment of the two sequences by ClustalW analysis shows only sporadic amino acid identity and no apparent sequence similarity between WbbD and bovine β4-Gal-transferase (Fig. 1).

The acceptor specificity of WbbD β3Gal-transferase was tested using bacterial homogenates from wild type VW187 bacteria and from a *wbbD* deletion mutant, complemented with *wbbD* in the *pCM227* plasmid [7]. During the incubation time of 10 min galactosyl transfer was linear

```
Q03084|Q03084_ECOLX    ------------------------MSDDTPKFSVLMAIYIKDSPLFLSEALQSIYK-32
P08037|B4GT1_BOVIN     MKFREPLLGGSAAMPGASLQRACRLLVAVCALHLGVTLVYYLAGRDLRRLPQLVGVHPPL 60
                                        ::  :  : ::.* :. *:.  *      *  .::


Q03084|Q03084_ECOLX    --------NTVAPDEVIIIRDGKVTSELNSVIDSWRRYLNIKDFTLEKNMGLGAALNFG- 83
P08037|B4GT1_BOVIN     QGSSHGAAAIGQPSGELRLRGVAPPPPLQNSSKPRSRAPSNLDAYSHPGPGPGPGSNLTS 120
                               *.  : :*.   .. *::  ..  *  .  *    .. * *.. *:


Q03084|Q03084_ECOLX    ---LNQCMHDLVIRADSDDINR---TNRFECILDFMTK---NGDVHILSSWVEEFEFNPG 134
P08037|B4GT1_BOVIN     APVPSTTTRSLTACPEESPLLVGPMLIEFNIPVDLKLVEQQNPKVKLGGRYTPMDCISPH 180
                          .   :.*.  .:.. :      .*:  :*:     * .*:: . :.    :.*


Q03084|Q03084_ECOLX    DKGIIKKVPSRNSILKYSK-------NRSPFNH--------PAVAFKKCEIMRVG----- 174
P08037|B4GT1_BOVIN     KVAIIIPFRNRQEHLKYWLYYLHPILQRQQLDYGIYVINQAGESMFNRAKLLNVGFKEAL 240
                       . .**   . .*:. ***           :*. :::            *::.::.**


Q03084|Q03084_ECOLX    -GYG-NEYLYEDYALWLKSLANG----CNGDNIQQVLVDMRFSKETAKRRGGIKYAISEI 228
P08037|B4GT1_BOVIN     KDYDYNCFVFSDVDLIPMNDHNTYRCFSQPRHISVAMDKFGFSLPYVQYFGGVSALSKQQ 300
                        .*. *  :::.*  *   .  *      .:  :*. .: .: **   .:  **:.   .:


Q03084|Q03084_ECOLX    KAQYHFYRANYISYQDFIINIITRIFVR-------------------------------- 256
P08037|B4GT1_BOVIN     FLSINGFPNNYWGWGGEDDDIYNRLAFRGMSVSRPNAVIGKCRMIRHSRDKKNEPNPQRF 360
                        . : :  ** .:.   :* .*: .*


Q03084|Q03084_ECOLX    ---------LLPTSFR--GYIYKKVIRRFL------------ 275
P08037|B4GT1_BOVIN     DRIAHTKETMLSDGLNSLTYMVLEVQRYPLYTKITVDIGTPS 402
                                :*. .:.   *:  :* *   *
```

Fig. 1 Sequence alignment of bacterial and mammalian Gal-transferases. The sequences of *E. coli* WbbD and bovine β4-Gal-transferase I were aligned using multiple sequence alignment ClustalW. The DxD motif is highlighted in *bold*. *Stars* indicate identity and *dots* similarity. The *dotted line (underlined)* indicates a hydrophobic sequence at the N-terminus of the bovine Gal-transferase that helps to target the enzyme to the Golgi membrane, and at the C terminus of the bacterial Gal-transferase that may, in combination with positively charged amino acids, associate the bacterial enzyme with the inner membrane

with time and proportional to protein concentration. The wild type WbbD and plasmid-derived WbbD had similar substrate specificity. The standard acceptor substrate, compound 30 (GlcNAcα-PO$_3$-PO$_3$-(CH$_2$)$_{11}$-OPh), is an effective substrate in enzyme assays (Table 1) [6]. However, elimination of one phosphate group in the substrate (compound 31) decreased the activity by 84% to 87%. This indicates that the pyrophosphate group is required for full activity of WbbD. This was further supported by substitution of the pyrophosphate group with a phosphonate derivative (in compound 29α, GlcNAcα-CO-CH$_2$-PO$_3$-(CH$_2$)$_{11}$-OPh), which was inactive as a substrate. This suggests that charge alone is not sufficient but the position of the phosphate group is important; thus the phosphate needs to be directly linked to GlcNAc to form a productive substrate.

Replacement of both phosphate groups with a malonyl group (in compound 24α, GlcNAcα-CO-CH$_2$-CO-O-(CH$_2$)$_{11}$-OPh), was not successful in yielding an active substrate. Removal of both phosphates having GlcNAc in either α or β-linkage also yielded compounds (20α and 20β, respectively) that were inactive as substrates.

In order to test the role of the lipid moiety, compounds were synthesized that had the GlcNAc–pyrophosphate linkage and an aliphatic chain length of six carbons (compound 7, GlcNAcα-PO$_3$-PO$_3$-(CH$_2$)$_6$-OPh) or 16 carbons (compound 15, GlcNAcα-PO$_3$-PO$_3$-(CH$_2$)$_{16}$-OPh). Both compounds were excellent substrates with compound 15 being the superior substrate (Table 1). Compound 7 showed 19% activity with wild type WbbD and 47% activity with recombinant WbbD from the pCM227 plasmid. Compound 15 showed 112% activity with wild type WbbD and 94% activity with recombinant WbbD. This suggests that the enzyme does not require a specific length of the lipid chain but can act on GlcNAcα-PO$_3$-PO$_3$-(CH$_2$)$_R$-OPh substrates where R can be 6, 11 or 16 carbons. Due to the low solubility of hydrophobic compounds in aqueous solution, it was difficult to determine the kinetic parameters.

After passing through C18 Sep-Pak, substrates and Gal-transferase enzyme products were analyzed by Electrospray ionization-MS (Table 2; Fig. 2). As expected, the molecular weight of the products increased by 162 m/z compared to the respective substrates (compounds 30, 7 and 15) indicating that one Gal residue had been transferred to the acceptors.

A series of GlcNAcα- and GlcNAcβ-derivatives lacking phosphate linkages but having hydrophobic aglycone

Table 1 Specificity of WbbD β3-Gal-transferase activity towards GlcNAc terminating compounds as acceptor substrates

Compound	Acceptor	Activity (%)		
		VW187	pCM227	β4GalT
30	0.2 mM GlcNAcα-PO$_3$-PO$_3$-(CH$_2$)$_{11}$-OPh	100	100	9
7	0.2 mM GlcNAcα-PO$_3$-PO$_3$-(CH$_2$)$_6$-OPh	19	47	18
15	0.2 mM GlcNAcα-PO$_3$-PO$_3$-(CH$_2$)$_{16}$-OPh	112	94	7
31	0.2 mM GlcNAcα-PO$_3$-(CH$_2$)$_{11}$-OPh	13	16	2
20α	0.2 mM GlcNAcα-(CH$_2$)$_{11}$-OPh	<1	<1	7
20β	0.2 mM GlcNAcβ-(CH$_2$)$_{11}$-OPh	<1	<1	14
24α	0.2 mM GlcNAcα-CO-CH$_2$-CO-O-(CH$_2$)$_{11}$-OPh	<1	<1	5
29α	0.2 mM GlcNAcα-CO-CH$_2$-PO$_3$-(CH$_2$)$_{11}$-OPh	<1	<1	4
32	0.5 mM GlcNAcα-Bn	<1[a]		
33	0.5 mM GalNAcα-Bn	<1[a]		
34	2 mM GlcNAcβ-(3-isoquinolinyl)	<1		
35	2 mM GlcNAcβ-(2-naphthyl)	<1		
36	2 mM GlcNBuβ-(2-naphthyl)	<1		
37	2 mM 1-ThioGlcNAcβ-(2-quinolinyl)	<1		
38	2 mM 1-ThioGlcNAcβ-(2-naphthyl)	<1		
39	2 mM GlcNAcβ-[8-(5-chloroquinolinyl)]	<1		
40	2 mM GlcNAcβ-(8-quinolinyl)	<1		
41	2 mM 3-DeoxyGlcNAcα-Bn	<1		
42	2 mM 3-FluoroGlcNAcα-Bn	<1		
43	2 mM 6-ThioGlcNAcβ-Bn	<1		
44	2 mM 2-AzidoGalα-Bn	<1		
45	2 mM GlcNAcβ-Bn	<1[a]		100
46	2 mM 1-Thio-N-butyrylGlcNβ-(2-naphthyl)	<1		
47	0.5 mM TTTVTPTP(GlcNAcβ1-6[GlcNAcβ1-3]GalNAcα-)TG	<1[a]		69[b]

Assays were carried out as described in the "Materials and methods" section using bacterial homogenates of *E. coli* VW187 (VW187) or of a WbbD mutant complemented with the plasmid pCM227 encoding WbbD (pCM227) [7]; 100% activity corresponds to 0.145 μmol/h/mg for VW187, and 0.625 μmol/h/mg for pCM227. Compounds were also tested as substrates for bovine β4-Gal-transferase (β4GalT); 100% activity corresponds to 1.6 μmol/h/mg. The glycopeptide was analyzed for purity on HPLC, using a C18 column and acetonitrile/water=6/94 as the mobile phase at 1 mL/min. Enzyme products having long hydrophobic chains were separated by the C18 Sep-Pak assay. Neutral compounds were also assayed by the AG1x8 method [10]. The synthesis of compounds 1–31 has been published [6] or will be published elsewhere [Riley *et al.*, in preparation]. **T** (in bold) is the attachment site for O-linked oligosaccharide in compound 47

[a] The activity was also analyzed using AG1x8 anion exchange chromatography [10]

[b] The value was taken from reference [10]

groups were also tested as substrates for WbbD. Many of these compounds have previously been shown to be excellent substrates for bovine β4Gal-transferase [10]. In addition, a glycopeptide having two terminal GlcNAc residues (compound 47, TTTVTPTP(GlcNAcβ1-6[GlcNAcβ1-3]GalNAcα-)TG, where the glycosylation site is indicated by **T**), that is an excellent substrate for bovine β4Gal-transferase [10], was tested as a substrate for WbbD at 0.1 to 0.5 mM concentrations in the assay. No WbbD activity was detected with these neutral acceptor substrate analogs, including compounds 20α, 20β, 24α, 20α, GlcNAcα-Bn, GalNAcα-Bn, and glycopeptide 47 (Table 1), as well as many other neutral GlcNAc-derivatives [10, data not shown].

Donor specificity of WbbD

The donor specificity was tested of WbbD in bacterial homogenates of the VW187 strain that potentially contains

a number of different transferase activities. We showed that Gal could be transferred from UDP-Gal to GlcNAcα-PO$_3$-PO$_3$-(CH$_2$)$_{11}$-OPh, but when enzyme was incubated with cytidine-monophosphate *N*-acetylneuraminic acid, UDP-GalNAc, GDP-Man or UDP-GlcNAc no glycosyl transfer was observed. The results indicate that WbbD is specific for UDP-Gal as the donor substrate. It suggests that the sequence of addition of sugars to build the O7 repeating unit is not random. However, a relatively small amount of [^{14}C]Glc was transferred from UDP-[^{14}C]Glc in the standard assay (29.6 nmol/h/mg).

Inhibition of WbbD activity

Substrate analogs that were inactive as acceptors, including GlcNAcα-Bn (at 0.5 mM concentration) and compounds 20α, 20β, 24α and 29α (Table 1), were tested in the standard assay for WbbD to determine whether they can

Table 2 Analysis of WbbD enzyme products by ESI-mass spectrometry

Compound	m/z	Molecular weight	
30.	626	626	Substrate GlcNAc-PO_3-PO_3-$(CH_2)_{11}$-OPh
	788	788	Product Gal–GlcNAc-PO_3-PO_3-$(CH_2)_{11}$-OPh
7.	556	556	Substrate GlcNAc-PO_3-PO_3-$(CH_2)_6$-OPh
	718	718	Product Gal–GlcNAc-PO_3-PO_3-$(CH_2)_6$-OPh
15.	696	696	Substrate GlcNAc-PO_3-PO_3-$(CH_2)_{16}$-OPh
	858	858	Product Gal–GlcNAc-PO_3-PO_3-$(CH_2)_{16}$-OPh

Enzyme assays of WbbD using acceptor substrates compounds 7, 15 and 30 were carried out as described in the "Materials and methods" section. Substrates and enzyme products were eluted with methanol from Sep-Pak columns and analyzed by electrospray ionization mass spectrometry (ESI-MS) in the negative ion mode (see Fig. 2)

bind to the enzyme and inhibit glycosyl transfer to GlcNAc-PP-PhU. In addition, a large series of GlcNAcα-R and GlcNAcβ-R derivatives [10] were tested as inhibitors (data not shown). At concentrations between 0.2 and 2 mM, none of the compounds inhibited Gal transfer by more than 30% in the standard assay. Thus, the pyrophosphate group appears to be required to form a substrate, and the presence of one phosphate group (compound 29α), or a pyrophosphate mimic (compound 24α) did not sufficiently substitute for the pyrophosphate. Compound 46, 1-Thio-N-butyryl-glucosamineβ-2-naphthyl is a potent inhibitor of bovine β1,4-Gal-transferase [10] but was neither a substrate nor an inhibitor of bacterial WbbD activity.

Substrates and inhibitors for mammalian β1,4-Gal-transferase

Mammalian (bovine) β4-Gal-transferase recognizes GlcNAc in β linkage and is poorly active on GlcNAcα-derivatives [10]. The standard substrate in the β4-Gal-transferase assay is GlcNAcβ-Bn (100% activity). The compound was inactive with the bacterial Gal-transferase. Generally, hydrophobic aglycone groups form excellent substrates for β4-Gal-transferase, but the long aliphatic lipid chain with 11 carbons in the substrate was not favourable and inferior to the benzyl group in a substrate. Thus, the α- and β-linked derivatives of GlcNAc, 20α and 20β showed activities of 7% and 14%, respectively, of the activity towards GlcNAcβ-Bn, respectively.

The standard substrate for WbbD Gal-transferase, compound 30, GlcNAcα-PO_3-PO_3-$(CH_2)_{11}$-OPh, at 0.2 mM concentration, was 9% active as a substrate for the mammalian β4-Gal-transferase (Table 1). When the lipid chain was extended to 16 carbons, (compound 15), the activity was similar (7%). However, when the lipid chain was truncated to six carbons in length (compound 7), the activity was 18%. This may indicate that the enzyme does not favour a long aliphatic chain in the acceptor substrate. The remaining compounds having this long lipid chain, with or without phosphate group (compounds 24α, 29α, 31) also showed low activities (Table 1).

In order to rule out that the low activity was due to the low solubility of compounds, assays were carried out in the presence of 0.125% to 0.25% Triton X-100. However, the presence of detergent did not increase the activity towards compound 30. The β4-Gal-transferase activity in HT29 cell homogenates towards GlcNAcβ-Bn and the standard WbbD substrate, compound 30, in the presence of detergent, showed a similar ratio to that determined with the purified enzyme. This indicates that the native membrane-bound enzyme also prefers GlcNAcβ-Bn over the pyrophosphate-lipid-containing substrate.

Neutral compounds that were poor substrates (20α, 20β, 24α and 29α) were assayed as inhibitors of β4-Gal-transferase. Compound 20β, having GlcNAc in β-linkage and the highest activity as a substrate of this series, inhibited the incorporation into GlcNAcβ-Bn by 18.4%, as expected for a poorly binding alternative substrate. The other compounds (20α, 24α, 29α) showed inhibition of less than 5%, and thus did not significantly bind to the acceptor binding site of the enzyme.

Discussion

GlcNAc-phospho-transferase is the first enzyme in the synthesis of the O7 antigen repeating unit; it transfers GlcNAc-phosphate from UDP-GlcNAc [12, 13] to undecaprenol-phosphate (undecaprenol-phosphate), conserving the α-linkage of GlcNAc in UDP-GlcNAc in the enzyme product. The second enzyme of the pathway also utilizes a nucleotide sugar, UDP-Gal, and transfers Gal to the GlcNAcα-pyrophosphate-undecaprenol with the formation of a Galβ1–3GlcNAc linkage. This enzyme, identified as WbbD β3Gal-transferase, therefore, transfers Gal to GlcNAcα-pyrophosphate-$(C_5H_8)_{11}$ that has an unsaturated polyisoprene lipid moiety of 55 carbon length [6]. The natural substrate for O antigen synthesis is thought to be localized to the inner bacterial membrane and oriented towards the cytoplasm, where the transferase reactions occur and where the enzyme is thought to be loosely associated with the membrane. *In vivo*, the lipid moiety may orient the substrate favourably to the enzyme.
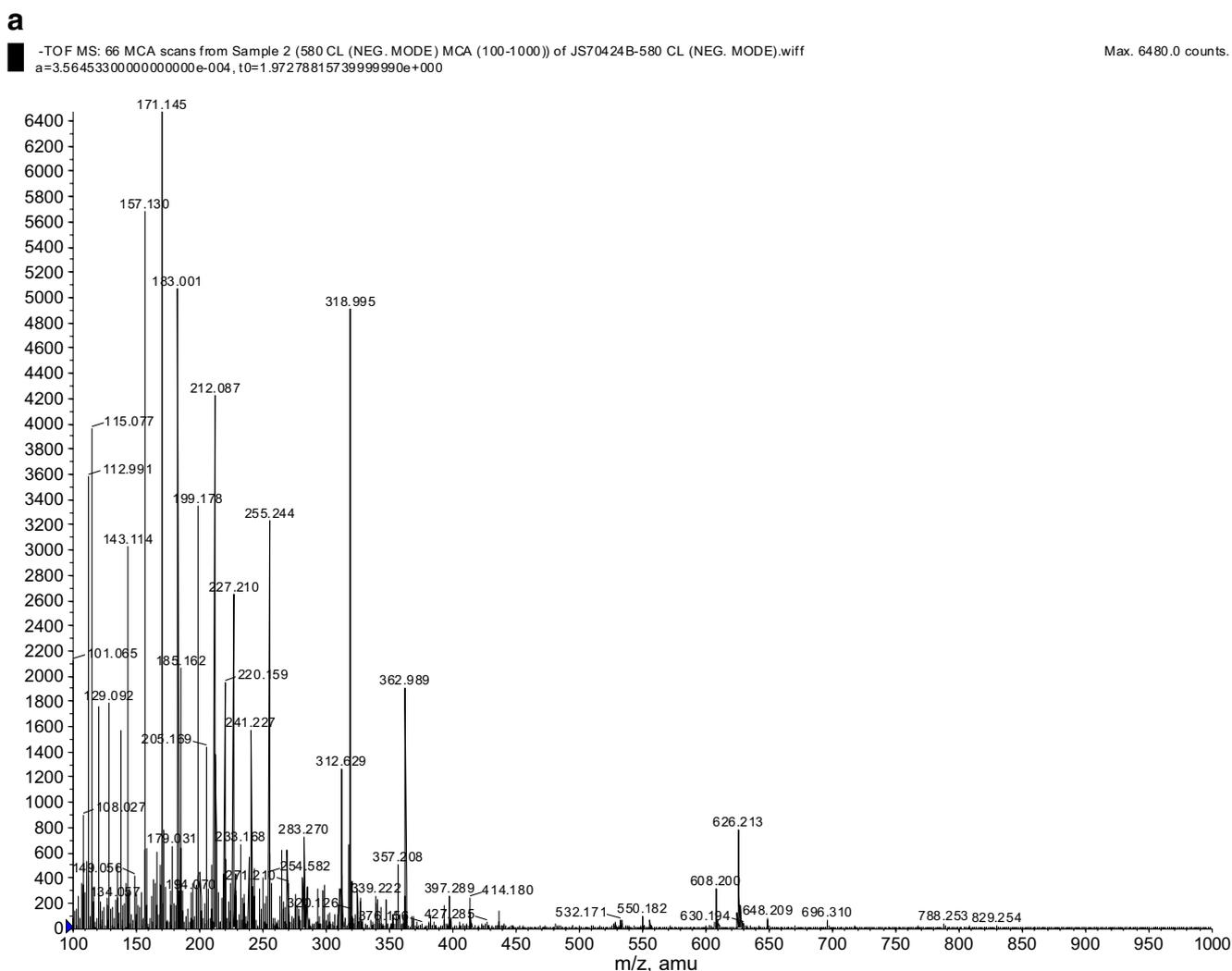
a

Fig. 2 Mass spectra of enzyme substrates and products. Compounds 30, 7 and 15 were used as acceptor substrates in WbbD assays as described in "Materials and methods", using non-radioactive UDP-Gal. After the incubation, mixtures were passed through Sep-Pak C18 columns and the methanol eluates were analyzed by ESI-MS (negative mode). **a** Compound 30 as a substrate has m/z 626,WbbD product has m/z 788 due to the addition of one hexose; **b** compound 7 as a substrate has m/z 556, product has m/z 718; **c** compound 15 as a substrate has m/z 696, product has m/z 858

The polyisoprene lipid requirement for bacterial glycosyltransferase activity has been investigated for the MurG enzyme that synthesizes peptidoglycan [14] as well as for TagA that synthesizes teichoic acid [15]. It was found that MurG can utilize synthetic analogs of the natural lipid-linked substrate, and that these synthetic lipids are in fact better substrates than undecaprenol-phosphate in assays containing 5% to 15% methanol [14]. Similarly, for WbbD, a specific length or structure of the lipid chain of the substrate does not appear to be required. The enzyme also does not require the O-phenyl group in the acceptor since GlcNAcα-PP-$(CH_2)_9$-CH_3 is an active substrate for WbbD [16]. Although the lipid chain has 55 carbons in the natural substrate, synthetic compounds with lipid chains containing 12 to 22 carbons are also compatible with high activity. It

needs to be determined if a lipid chain is at all necessary in the acceptor substrate for WbbD, or if a short chain linked to pyrophosphate suffices. It is also not yet clear how the synthetic water-insoluble substrates are present in the enzyme assays containing 10% methanol. It is possible that they are incorporated into favourable liposomes formed by the bacterial membranes, making them available to the enzyme. The addition of detergents *in vitro* appears to disturb this favourable orientation since several detergents were shown to decrease the activity [7]. Since hydrophobic compounds at high concentrations are not likely to be in solution in the assay mixtures, it was not possible to determine the accurate K_M and V_{max} values. The critical micelle concentrations of the amphipathic compounds active as WbbD substrates are subject of a future study.

b

-TOF MS: 0.500 to 1.050 min from JS70424A-615 CL (NEG. MODE).wiff
a=3.56453300000000000e-004, t0=1.97278815739999990e+000
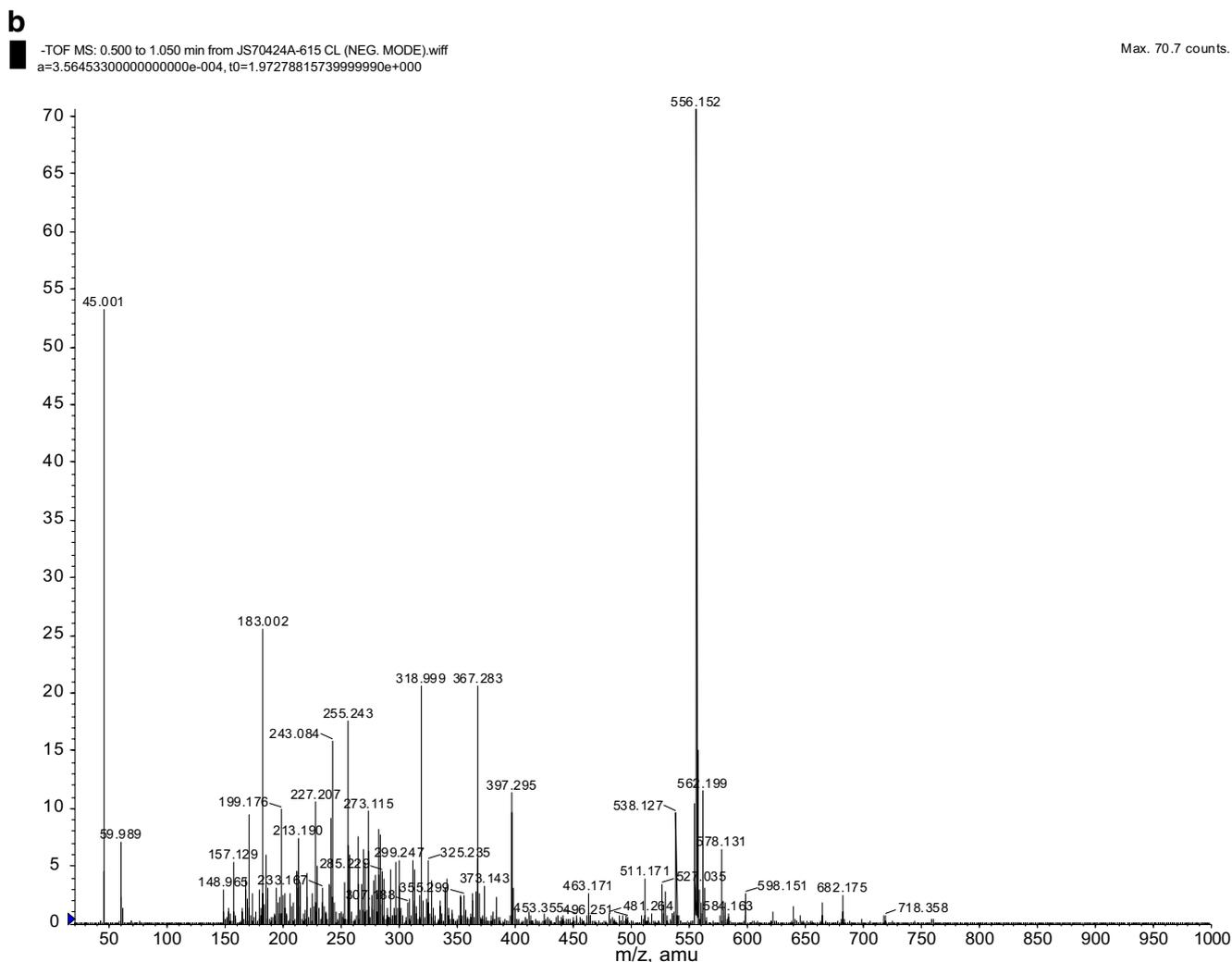
Max. 70.7 counts.

Fig. 2 (continued)

The present work shows that both the natural endogenous enzyme and the recombinant WbbD β1,3-Gal-transferases require at least one phosphate in the acceptor substrate which has to be directly linked to GlcNAc, and are optimally active when GlcNAc is linked to pyrophosphate. The neutral malonate derivative could not substitute for the charged pyrophosphate group. A large number of compounds lacking phosphate were found to be inactive as substrates and as inhibitors, possibly because the interaction between pyrophosphate and enzyme is required for binding or catalysis. Since the enzyme has a distinct requirement for divalent metal ion, it is possible that Mn^{2+} participates in binding negatively charged UDP-Gal and/or the pyrophosphate-containing acceptor substrate. It is likely that positively charged groups of the enzyme are also involved.

WbbD is a member of the CAZy GT2 family with a GT-A fold and a hydrophobic sequence at the C terminus.

Another member of the GT2 family is SpsA, of which the structure has been determined. The N terminal domain of the enzyme contains the nucleotide binding fold while the acceptor binding grove is in the C terminal domain [17]. However, the specific substrates have not yet been defined for SpsA. The role of the C terminal hydrophobic segment has been investigated for the α4-Gal-transferase LgtC from *Neisseria meningitidis* [18], and for WaaJ [19], a α2-Glc-transferase from *E. coli* and *Salmonella*, involved in the addition of the last sugar to the outer core oligosaccharide. Both enzymes are retaining transferases, belong to the CAZy GT8 family and have a GT-A fold. Truncation of the C terminus of LgtC retained the activity but resulted in a decreased ability to aggregate and to associate with membrane fractions. The truncation of WaaJ decreased its association with membranes and produced soluble enzyme, but also decreased or eliminated enzyme activity, depending on the number of amino acids deleted. While the K_M
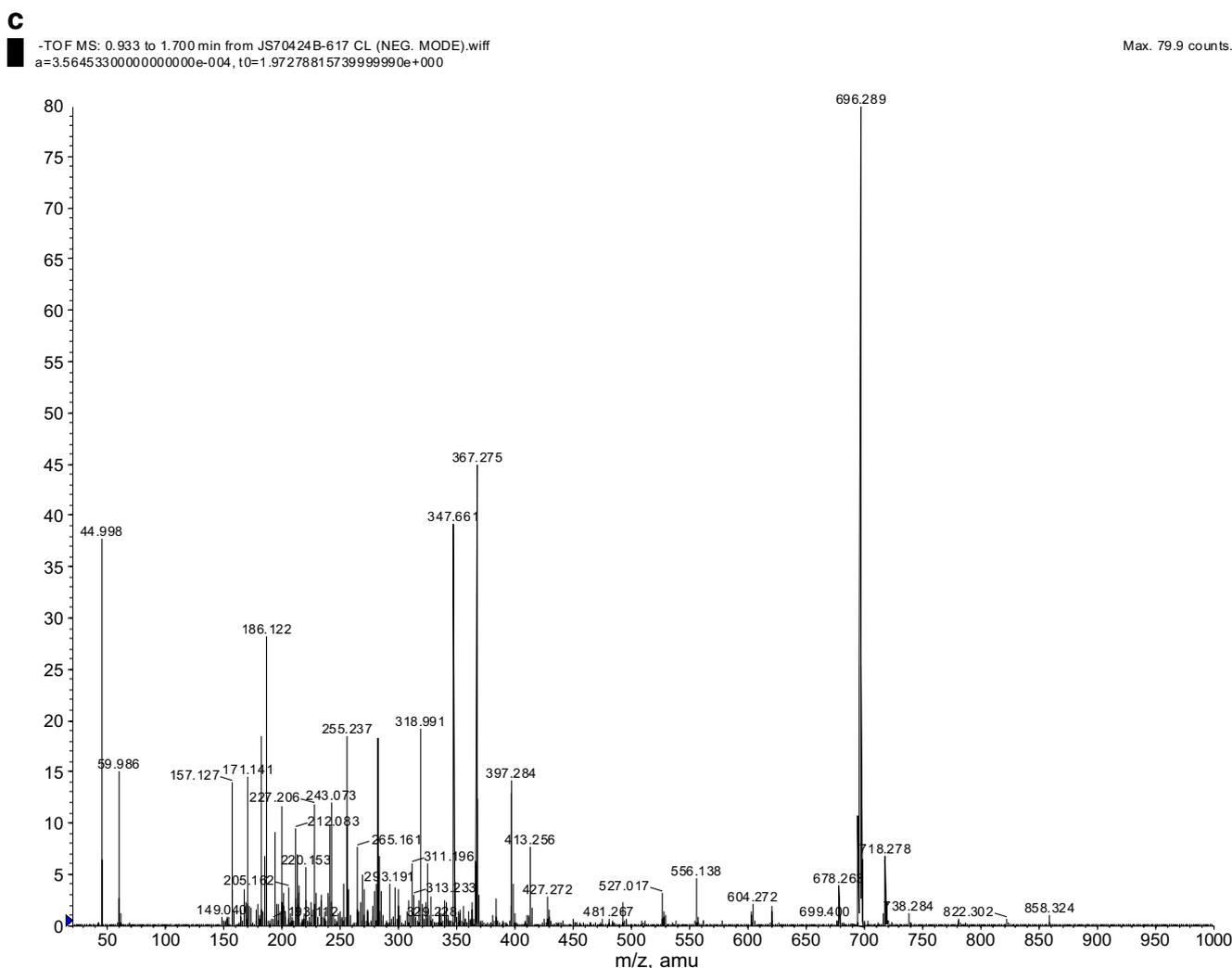
c

-TOF MS: 0.933 to 1.700 min from JS70424B-617 CL (NEG. MODE).wiff
a=3.56453300000000000e-004, t0=1.97278815739999990e+000

Max. 79.9 counts.

Fig. 2 (continued)

value for UDP-Glc was minimally affected, the K_M value for LPS increased. This indicates that the C terminal end of the enzyme is important for binding a hydrophobic substrate. At the C terminus of WbbD, a hydrophobic segment, flanked by positively charged amino acids, may serve to loosely associate the enzyme with negatively charged phospholipids in the inner membrane. The hydrophobic segment may also serve to ensure that the enzyme is optimally positioned and binds to its natural membrane-bound acceptor substrate. Another possible function is that the hydrophobic patches may interact with other proteins to form enzyme complexes that may act at high efficiency.

Several O antigens contain the Galβ1–3GlcNAc linkage at the reducing end of the repeating unit, *i.e.* O64, O114 and O153, while the O5, O104 and O117 antigens have a Galβ1–3GalNAc-linkage. A Galβ1–4GlcNAc linkage is found in the O83 antigen. The Gal-transferases that synthesize these structures are also involved in the second step of O antigen repeating unit synthesis, and may share properties and substrate specificities with WbbD.

The low activity in *E. coli* strain VW187, using UDP-Glc as the donor substrate, may possibly be due to the presence of a Glc-transferase involved in oligosaccharide synthesis. However, a Glc–GlcNAc linkage has not yet been described in the O7 antigen, and it is likely that a 4-epimerase [20] converted UDP-Glc to UDP-Gal which then served as a sugar donor substrate for WbbD. These possibilities need to be further examined.

After Gal transfer in the O7 repeating unit assembly, Man, Rha and VioNAc residues are transferred, presumably in a specific sequence, since the few glycosyltransferases characterized to date that synthesize O antigens in *E. coli* showed distinct substrate specificities [7, 21 and I. Brockhausen *et al.*, in preparation]. The specific substrates for these reactions and the enzymes involved remain to be determined. A number of bacterial glycosyltransferases

involved in the synthesis of terminal carbohydrate structures have been shown to act on low molecular weight acceptor substrates. Examples are the lipooligosaccharide synthesizing β4-Gal-transferase from *Neisseria* that can utilize free GlcNAc as an acceptor [22, 23], and α4-Gal-transferase LgtC that acts on lactose [18]. The *E. coli* serotype O86 has glycosyltransferases that synthesize the human blood group B structure of the O repeating unit. These transferases act on low molecular weight acceptors [24]. In addition, an α1,2-Fuc-transferase involved in the biosynthesis of O128 antigen was active using free Gal as an acceptor substrate [25]. It is therefore possible that a requirement for pyrophosphate-lipid in the acceptor generally applies to the second enzyme of the O antigen assembly sequence, while subsequently acting enzymes have broader specificities, comparable to those of glycosyltransferases from other species including mammals.

GalNAcα-PP-PhU [21] has recently been synthesized and used as a substrate for a GalNAc-transferase (Wbnh) involved in the second step of the biosynthesis of the *E. coli* O86 antigen repeating unit. Like WbbD, Wbnh does not act on small molecule acceptors such as GalNAc-phenyl; it also appears to require pyrophosphate in the acceptor, and does not require the O-phenyl group in the acceptor. Wbnh does not act on UDP-GalNAc which may indicate that it requires a lipid moiety in the substrate. However, the Wbnh enzyme differs from WbbD in that it does not require Mn^{2+} or other divalent metal ion as a cofactor, does not have a DxD motif, and is a member of the CAZy GT1 family with a GT-B fold.

WbbD has virtually no sequence similarity to the ubiquitous mammalian Gal-transferase (β4-Gal-transferase I) or to other mammalian Gal-transferases. The substrate recognition of WbbD and Gal-transferase I is also significantly different [26]. However, the mode of catalysis may be similar. This is based on the observation that both enzymes bind GlcNAc and UDP-Gal, both need divalent cation for activity and have a DxD motif. While the anomeric configuration of GlcNAc in the acceptor is important for the mammalian enzyme, this requirement remains to be determined for the bacterial enzyme. The extended binding site of the mammalian enzyme can accommodate short lipid chains and aromatic groups, as well as glycopeptides and peptides, while long lipid chains, with or without phosphate groups, bind relatively poorly. A major difference in specificity is seen when using glycopeptides as substrates. The mammalian enzyme can act on GlcNAc present at various positions in the peptide, even if it is in the vicinity of large oligosaccharides, while WbbD does not recognize GlcNAc linked to a glycopeptide. WbbD was assayed in bacterial homogenates which could potentially contain other interfering enzymes and substrates, while the mammalian enzyme was purified. However, it is not likely that any other glycosyltransferases

are present in VW187 that also act on compound 30 as a substrate since bacterial enzymes are also very specific for their donor and acceptor substrates. It is possible that modifiers of the activity are present in the crude membrane preparation, and the results will have to be confirmed using a purified enzyme.

In contrast to the mammalian enzyme, the bacterial enzyme prefers a pyrophosphate group in the acceptor and accepts long lipid chains. Compounds that were inactive as substrates did not inhibit the reaction, including potent inhibitors of β4-Gal-transferase, indicating that, in spite of having a long lipid chain, none of these compounds bound in the acceptor binding site of WbbD. Thus, the extended acceptor binding sites of the mammalian and bacterial enzyme must differ significantly, and the bacterial enzyme requires at least one phosphate group attached to GlcNAc.

In conclusion, a useful substrate analog inhibitor for the bacterial enzyme, involved in the second step of O7 antigen repeating unit synthesis, should have a pyrophosphate linked to GlcNAc, or alternatively, a closely related structural analog, and a lipid-like chain. These substrate analogs are expected to poorly bind to the substrate binding sites of mammalian enzymes, and have potential as specific blockers of O antigen synthesis that would render Gram negative bacteria susceptible to the host immune system. The knowledge of glycosyltransferase specificity also allows us to further define mechanisms of O antigen synthesis, and is useful for *in vitro* synthesis of defined O antigen structures for vaccine development or for studies of glycan function and biosynthesis in bacteria.

References

1. L'vov, V.L., Shaskov, A.S., Dmitriev, B.A., Kochetkov, N.K.: Structural studies of the O-specific side chain of the lipopolysaccharide from *Escherichia coli* O:7. Carbohydr. Res. **126**, 249–259 (1984)
2. Raetz, R.H., Whitfield, C.: Lipopolysaccharide endotoxins. Ann. Rev. Biochem. **71**, 635–700 (2002)
3. Samuel, G., Reeves, P.: Biosynthesis of O-antigens: genes and pathways involved in nucleotide sugar precursor synthesis and O-antigen assembly. Carbohydr. Res. **338**, 2503–2519 (2003)
4. Marolda, C.L., Feldman, M.F., Valvano, M.A.: Genetic organization of the O7-specific lipopolysaccharide biosynthesis cluster of *Escherichia coli* VW187 (O7:K1). Microbiology **145**, 2485–2495 (1999)

5. Whitfield, C., Valvano, M.A.: Biosynthesis and expression of cell-surface polysaccharides in gram-negative bacteria. Adv. Microb. Physiol. **35**, 135–246 (1993)

6. Montoya-Peleaz, P.J., Riley, J.G., Szarek, W.A., Valvano, M.A., Schutzbach, J.S., Brockhausen, I.: Identification of a UDP-Gal: GlcNAc-R galactosyltransferase activity in *Escherichia coli* VW187. Bioorg. Med. Chem. Lett. **15**, 1205–1211 (2005)

7. Riley, J.G., Menggad, M., Montoya-Peleaz, P., Szarek, W.A., Marolda, C.L., Valvano, M.A., Schutzbach, J.S., Brockhausen, I.: The *wbbD* gene of *Escherichia coli* strain VW187 (O7:K1) encodes a UDP-Gal: GlcNAcα-pyrophosphate-R β1,3-galactosyltransferase involved in the biosynthesis of O7-specific lipopolysaccharide. Glycoconj. J. **23**, 525–541 (2006)

8. Whitfield, C.: Biosynthesis of lipopolysaccharide O antigens. Trends Microbiol. **3**, 178–185 (1995)

9. Breton, C., Snajdrova, L., Jeanneau, C., Koca, J., Imberty, A.: Structures and mechanisms of glycosyltransferases. Glycobiology **16**, 29R–37R (2006)

10. Brockhausen, I., Benn, M., Bhat, S., Marone, S., Riley, J.G., Montoya-Peleaz, P., Vlahakis, J.Z., Paulsen, H., Schutzbach, J.S., Szarek, W.: UDP-Gal: GlcNAc-R beta1,4 Galactosyltransferase—a target enzyme for drug design. Acceptor specificity and inhibition of the enzyme. Glycoconj. J. **23**, 523–539 (2006)

11. Vavasseur, F., Yang, J., Dole, K., Paulsen, H., Brockhausen, I.: Synthesis of O-glycan core 3: characterization of UDP-GlcNAc: GalNAc-R β3-N-acetyl-glucosaminyltransferase activity from colonic mucosal tissues and lack of the activity in human cancer cell lines. Glycobiology **5**, 351–357 (1995)

12. Amer, A.O., Valvano, M.A.: Conserved amino acid residues found in a predicted cytosolic domain of the lipopolysaccharide biosynthetic protein WecA are implicated in the recognition of UDP-N-acetylglucosamine. Microbiology **147**, 3015–3025 (2001)

13. Amer, A.O., Valvano, M.A.: Conserved aspartic acids are essential for the enzymic activity of the WecA protein initiating the biosynthesis of O-specific lipopolysaccharide and enterobacterial common antigen in *Escherichia coli*. Microbiology **148**, 571–582 (2002)

14. Chen, L., Men, H., Ha, S., Ye, X.-Y., Brunner, L., Hu, Y., Walker, S.: Intrinsic lipid preferences and kinetic mechanism of *Escherichia coli* MurG. Biochemistry **41**, 6824–6833 (2002)

15. Zhang, Y.H., Ginsberg, C., Yuan, Y., Walker, S.: Acceptor substrate selectivity and kinetic mechanism of *Bacillus subtilis* TagA. Biochemistry **45**, 10895–10904 (2006)

16. Brockhausen, I., Larsson, E.A., Hindsgaul, O.: A very simple synthesis of GlcNAc-alpha-pyrophosphoryl-decanol: a substrate for the assay of a bacterial galactosyltransferase. Bioorg. Med. Chem. Lett. **18**, (2), 804–807 (2007)

17. Charnock, S.J., Davies, G.J.: Structure of the nucleotide-diphospho-sugar transferase, SpsA from *Bacillus subtilis*, in native and nucleotide-complexed forms. Biochemistry. **38**, 6380–6385 (1999)

18. Lairson, L.L., Chiu, C.P., Ly, H.D., He, S., Wakarchuk, W.W., Strynadka, N.C., Withers, S.G.: Intermediate trapping on a mutant retaining alpha-galactosyltransferase identifies an unexpected aspartate residue. J. Biol. Chem. **279**, 28339–28344 (2004)

19. Leipold, M.D., Kaniuk, N.A., Whitfield, C.: The C-terminal domain of the *Escherichia coli* WaaJ glycosyltransferase is important for catalytic activity and membrane association. J. Biol. Chem. **282**, 1257–1264 (2007)

20. Guo, H., Yi, W., Li, L., Wang, P.G.: Three UDP-hexose 4-epimerases with overlapping substrate specificity coexist in *E. coli* O86:B7. Biochem. Biophys. Res. Commun. **356**, 604–609 (2007)

21. Yi, W., Yao, Q., Zhang, Y., Motari, E., Lin, S., Wang, P.G.: The wbnH gene of *Escherichia coli* O86:H2 encodes an alpha-1,3-N-acetylgalactosaminyl transferase involved in the O-repeating unit biosynthesis. Biochem. Biophys. Res. Commun. **344**, 631–639 (2006)

22. Wakarchuk, W.W., Cunningham, A., Watson, D.C., Young, N.M.: Role of paired basic residues in the expression of active recombinant galactosyltransferases from the bacterial pathogen *Neisseria meningitidis*. Protein Eng. **11**, 295–302 (1998)

23. Park, J.E., Lee, K.Y., Do, S.I., Olee, S.S.: Expression and characterization of β-1,4-galactosyltransferase from *Neisseria meningitis* and *Neisseria gonorrhoeae*. J. Biochem. Mol. Biol. **35**, 330–336 (2002)

24. Yi, W., Shao, J., Zhu, L., Li, M., Singh, M., Lu, Y., Lin, S., Li, H., Ryu, K., Shen, J., Guo, H., Yao, Q., Bush, C.A., Wang, P.G.: *Escherichia coli* O86 O-antigen biosynthetic gene cluster and stepwise enzymatic synthesis of human blood group B antigen tetrasaccharide. J. Am. Chem. Soc. **127**, 2040–2041 (2005)

25. Shao, J., Li, M., Jia, Q., Lu, Y., Wang, P.G.: Sequence of *Escherichia coli* O128 antigen biosynthesis cluster and functional identification of an α-1,2-fucosyltransferase. FEBS Lett. **553**, 99–103 (2003)

26. Qasba, P.K., Ramakrishnan, B., Boeggeman, E.: Substrate-induced conformational changes in glycosyltransferases. Trends Biochem. Sci. **30**, 53–62 (2005)