Electronic Thesis and Dissertation Repository

9-15-2011 12:00 AM

Structure-Function Analysis of Enzymes of the Polyisoprenyl-Phosphate Hexose-1-Phosphate Transferase family

Kinnari B. Patel
Unversity of Western Ontario

Supervisor
Dr. Miguel Valvano
The University of Western Ontario

© Kinnari B. Patel 2011

Recommended Citation

Patel, Kinnari B., "Structure-Function Analysis of Enzymes of the Polyisoprenyl-Phosphate Hexose-1-Phosphate Transferase family" (2011). *Electronic Thesis and Dissertation Repository*. 269.
https://ir.lib.uwo.ca/etd/269

STRUCTURE-FUNCTION ANALYSIS OF ENZYMES OF THE POLYISOPRENYL-PHOSPHATE HEXOSE-1-PHOSPHATE TRANSFERASE FAMILY

(Spine title: Structure-Function Analysis of PHPT proteins)

(Thesis format: Integrated-Article)

by

Kinnari Bhanuprasad <u>Patel</u>

Graduate Program in Microbiology and Immunology

A thesis submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy

The School of Graduate and Postdoctoral Studies
The University of Western Ontario
London, Ontario, Canada

© Kinnari B. Patel 2011

THE UNIVERSITY OF WESTERN ONTARIO
SCHOOL OF GRADUATE AND POSTDOCTORAL STUDIES

CERTIFICATE OF EXAMINATION

Supervisor	Examiners
_____	_____
Dr. Miguel Valvano	Dr. David Heinrichs
Supervisory Committee	Dr. Martin McGavin
_____	_____
Dr. Carole Creuzenet	Dr. Chris Whitfield
_____	_____
Dr. Stan Dunn	Dr. Brian Shilton

The thesis by

Kinnari B. Patel

entitled:

**Structure-Function Analysis of Enzymes of the Polyisoprenyl-Phosphate
Hexose-1-Phosphate Transferase family**

is accepted in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Date_____   _____
                                      Chair of the Thesis Examination Board

Abstract

Enzymes of the polyisoprenyl-phosphate hexose-1-phosphate transferase (PHPT) family are integral membrane proteins that initiate the synthesis of glycans by catalyzing the transfer of a hexose-1-phosphate sugar from UDP-hexose to the lipid carrier undecaprenyl phosphate (Und-P). These glycans such as O antigen and exopolysaccharide (EPS) provide bacteria with protection and adaptation to the environment and host immune factors. The role of PHPT proteins in initiation and the absence of any eukaryotic homologues make them an attractive target for novel antimicrobials; however study of these proteins is difficult due to the presence of multiple transmembrane helices. A requirement of the C-terminal domain for catalytic activity has been demonstrated *in vitro*, but the importance of specific regions and/or residues for the activity of these enzymes was not understood and was investigated in this work. The galactose-1-phosphate transferase WbaP of *S. enterica* initiates the synthesis of O antigen and was used as a model to study PHPT proteins. *In vivo* and *in vitro* functional assays of WbaP allowed us to show that a soluble loop region affects O antigen chain length distribution and the C-terminal domain (WbaPCT) containing one putative transmembrane (TM) domain and tail region is sufficient for transferase activity *in vivo*. The location of the WbaP tail region was found to be cytoplasmic and to contain highly conserved residues essential for activity. In this work a thioredoxin (TrxA) fusion was used as a tool to improve protein folding, which allowed for the solubilization and purification of WbaPCT. Obtaining a pure WbaPCT sample has allowed for the biochemical characterization of this protein revealing that this region is sufficient to maintain specificity for undecaprenyl phosphate (Und-P). Characterization of two glucose-1-phosphate transferases from *Escherichia coli* K-12 and *Caulobacter crescentus* revealed that some members of the PHPT family are able to utilize a second sugar substrate leading to the synthesis of a different glycan. These studies will pave the way to a detailed structural and mechanistic understanding of these proteins.

Key words: membrane proteins, O antigen, exopolysaccharides, UDP-sugar transferases, undecaprenyl phosphate.

Declaration of Co-authorship

Chapter 2 of this thesis was co-first authored with Dr. Maria Soledad Saldias, a former PhD student in the laboratory of Dr. Inés Contreras, who constructed the mutant strains MSS1, MSS2 and plasmids pSM10, pSM13, pSM18, pSM22, pSM23, pSM28, pSM30 and pSM31. Dr. Saldias proposed the topology of WbaP (Fig. 2.1a) and made the initial observation of the effect of the WbaP soluble loop on O antigen distribution (Fig. 2.4 and 2.5). Cristina Marolda performed the experiments to detect Wzz expression in cells overexpressing the WbaP soluble loop (Fig. 2.6). All other work is that of K.B.P.

In chapter 3 of this thesis, Sarah Furlong assisted with *in vitro* transferase experiments by setting up reactions and performing lipid extractions. All other work is that of K.B.P.

In chapter 4 of this thesis, Ximena Fernandez, an undergraduate student supervised by K.B.P created plasmid pXF1 and strain XBF1 and also performed western blot analysis of HfsE, PssY, and PssZ expression in DH5α cells. Dr. Evelyn Toh, a former PhD student in the laboratory of Dr. Yves Brun, created the plasmids pHfsE, pPssY and pPssY and performed the polystyrene binding experiments in *C. crescentus*. Dr. Anna Hanuszkiewicz assisted with the large scale LPS purification and preparation of neutral sugars and Dr. Mark Bernards performed the GC-MS analysis on these samples. All other work is that of K.B.P.

In chapter 5 of this thesis, Ewa Ciepichal from the laboratory of Dr. Ewa Swiezewska prepared the isoprenoids decaprenyl monophosphate, undecaprenyl monophosphate and dolichyl monophosphate. All other work is that of K.B.P.

Dedication

This work is dedicated to my father Bhanuprasad, whose resilience amidst obstacles has always been an inspiration

Acknowledgements

Table of Contents

List of Tables

List of Figures

List of Abbreviations

Amp	Ampicillin
Ara4N	4-amino-4-deoxy-α L-arabinose
CA	Colanic acid
CA$_{LPS}$	colanic acid attached to lipid A-core OS
CAMP	cationic antimicrobial peptides
Cm	Chloramphenicol
CPS	Capsular polysaccharide
ECA	enterobacterial common antigen
EPS	Exopolysaccharide
Fuc	Fucose
Gal	Galactose
GFP	Green fluorescent protein
Glc	Glucose
GlcNAc	N-acetylglucosamine
IMVs	Inverted membrane vesicles
Kdo	3-deoxy-D-*manno*-oct-2-ulosonic acid
Kn	Kanamycin
LB	Luria-Bertani
LPS	Lipopolysaccharide
LT	Lytic transglycosylase
OAg	O antigen
PCR	Polymerase chain reaction
PHPT	polyisoprenyl-phosphate hexose-1-phosphate transferase
PI	polyisoprenoid
PL2	Periplasmic loop 2
PMSF	Phenylmethylsulfonyl fluoride
PNPT	Polyisoprenyl-phosphate N-acetylaminosugar-1-phosphate transferase
PSI-PRED	Protein structure predication
SAM	Sequence alignment and modeling
SDS-PAGE	Sodium dodecylsulfate polyacrylamide gel electrophoresis
Sp	Spectinomycin
STI	Soybean trypsin inhibitor
TEV	Tobacco etch virus
TLC	Thin layer chromotography
TLR4	Toll like receptor 4
TM	Transmembrane
TrxA	Thioredoxin
UDP	Uridine diphosphate
UMP	Urindine monophosphate
Und-P	Undecaprenyl phosphate
UV	Ultra violet
WbaP$_{CT}$	WbaP C-terminus

Chapter 1

Introduction

## 1.1.	The prokaryotic cell envelope

The bacterial cell envelope is critical to maintaining cell shape, withstanding cell turgor pressure, providing protection against harmful agents, mediating cell adhesion and in interactions with host cells. The cell envelopes of both Gram positive and negative bacteria (Fig. 1.1) contain peptidoglycan which is essential to maintain a rigid structure (see section 1.2.1.1.3). Gram positive cell envelopes are made up of mostly peptidoglycan which accounts for 90% of their content (120). This thick peptidoglycan layer lies above the symmetrical cytoplasmic membrane composed of phospholipids. In some Gram positive bacteria peptidoglycan serves as an attachment point for teichoic acids. Teichoic acids can also be attached to the cytoplasmic membrane (see section 1.2.1.1.4). Many Gram-positive bacteria also produce capsular polysaccharides linked to the cell wall. The proteinaceous surface layer of Gram-positive bacteria (and some Gram negative bacteria) can also be glycosylated forming a paracrystalline sheet covering the surface of the organism (see section 1.2.2.1.6).

The Gram negative cell envelope is more complex and contains an additional membrane with the inner monolayer made up of phospholipids and the outer monolayer composed of lipopolysaccharides (LPS) (see section 1.1.1). These bacteria can also produce capsules (CPS) and exopolysaccharides (EPS) associated with or linked to cell envelope components (see sections 1.2.2.1.3, 1.2.2.1.4, and 1.2.2.1.5). Between this outer membrane and the cytoplasmic membrane is the periplasm, which contains a thin layer of peptidoglycan and proteins. In addition to LPS, the outer membrane also contains proteins including porins, which function as channels for the entrance and exit of small hydrophilic substances. Together, the cytoplasmic membrane, periplasm, peptidoglycan and outer membrane make up the Gram negative cell envelope (20).

### 1.1.1.	Lipopolysaccharide

LPS is a major component of the Gram negative outer membrane (202). A complete LPS consists of the hydrophobic lipid A, embedded in the outer leaflet of the
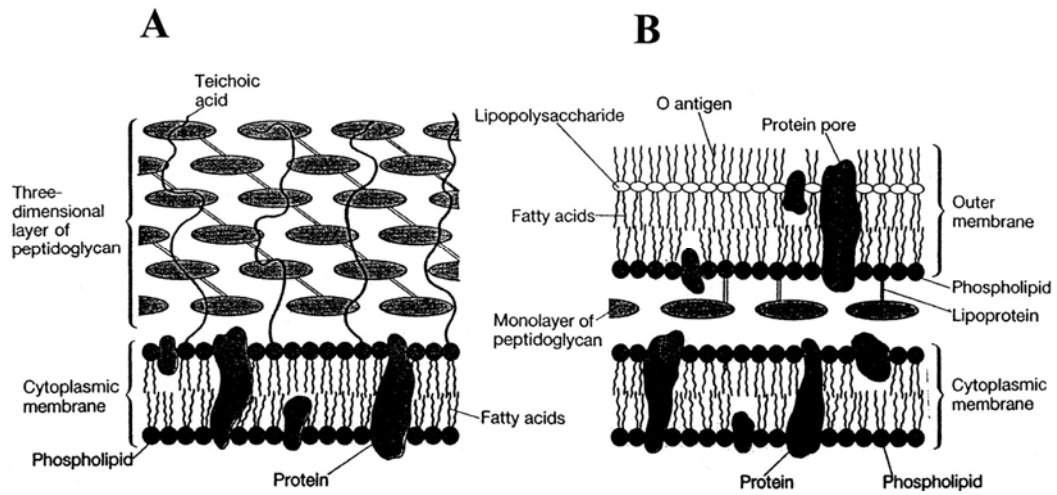
Figure 1.1. Cell envelope of Gram-positive (A) and Gram-negative bacteria (B) from Hamill and Martinac (2001) (80).
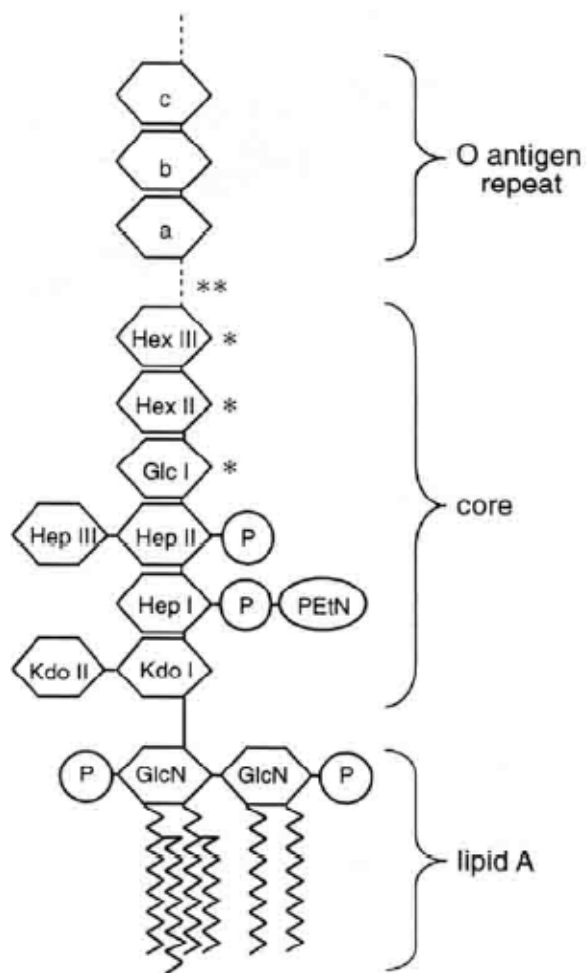
Figure 1.2. General chemical structure of LPS from *E. coli* or *Salmonella enterica*. Figure is from reproduced from Yethon and Whitfield (2001) (225).

outer membrane bilayer, the core oligosaccharide and the polymeric O antigen (141) (Figure. 1.2). As a whole, LPS stabilizes the outer membrane and protects the bacteria against host factors and environmental stresses. Each component of the molecule adds to its functional properties; however lipid A is most crucial for cell viability as it is required for the formation of the outer membrane (157).

1.1.1.1.   Lipid A

Lipid A or endotoxin anchors LPS to the bacterial cell. It consists of two β, 1′-6-linked glucosamine residues bound to a variable number of phosphate or pyrophosphate-linked fatty acids (141, 157). The tight packing of saturated acyl chains linked to the lipid A decreases membrane fluidity preventing entry of hydrophobic molecules (138, 157). In addition, the unique sugar 3-deoxy-D-*manno*-oct-2-ulosonic acid (Kdo), and possibly other sugars attached distal to Kdo, further stabilize the lipid bilayer *via* divalent cation bridges formed between negative charges from the phosphorylated lipid A backbone and the carboxylate of Kdo (138).

 Lipid A is also responsible for the endotoxic effects of LPS including Gram negative sepsis (68). Immune cells detect lipid A through membrane embedded toll like receptor 4 (TLR4), although the degree of activation depends on the number of phosphate groups and the length and number of fatty acyl chains present on the structure (6, 152). For example, lipid A from *Helicobacter pylori* and *Yersinia pestis* are poor activators of TLR4 (197). Upon activation by LPS, TLR4 recruits adapter proteins in the cytoplasm which in turn activate other molecules within the cell to amplify the signal, resulting in induction or suppression of proinflammatory gene expression (212).

Although the biochemical structure and synthesis of lipid A is highly conserved, certain enzymes can modify lipid A, which leads to increased resistance against innate immune responses such as cationic antimicrobial peptides (CAMP) or reduced recognition by TLR4 (212). The addition of positively charged 4-amino-4-deoxy-α-L-arabinose (Ara4N) as substitutions of the phosphate residues linked to the glucosamine dimer of lipid A is one of the best characterized modifications. This was first discovered in *Salmonella enterica* serovar Typhimurium when cells lacking Ara4N were found to be more sensitive to antimicrobial peptides (201). Other components of lipid A that can be

modified are the fatty acyl chains. For example the protein PagP is important for resistance against CAMP by adding palmitate derived from glycerophospholipids to lipid A, resulting in a hepta-acylated structure (4, 89).

Lipid A is a vital component of the outer membrane. Until recently it was believed that the minimal LPS structure required for growth was lipid A and 2 Kdo residues. Meredith *et al* reported selection of a viable suppressor strain that has LPS containing only lipid IV$_A$, a lipid A intermediate missing the secondary acyl chain fatty acids and Kdo (127). Suppressor free mutants have also been created by growth of mutants in heptosyltransferases and late acetyltransferases at low temperatures. Although viable, these strains are severely compromised (123).

1.1.1.1.1. Lipid A synthesis

The lipid A synthesis pathway otherwise known as the Raetz pathway is conserved among Gram negative bacteria and requires the activity of nine enzymes [For a review see (195)]. LpxA catalyzes the first step by acylation of UDP-*N*-acetylglucosamine (GlcNAc) (39, 69). Differences in lipid A structure between bacteria can be attributed to LpxA substrate specificities for the acyl donor such as *R*-3-hydroxymyristate in *E. coli* (195). The second step of this pathway is performed by the deacetylase LpxC that hydrolyses the amide linkage at position 2 of the glucosamine sugar resulting in removal of an acetyl group (14, 103, 226). This allows for the addition of a second fatty acyl chain, by LpxD, in an amide linkage to form UDP-2,3-diacylglucosamine (103). Some of the UDP-2,3-diacylglucosamine is preserved as it is the immediate precursor of the distal sugar of Lipid A. The enzyme LpxH cleaves the remaining UDP-2,3-diacylglucosamine at its pyrophosphate bond to generate 2,3-diacylglucosamine-1-phosphate known as Lipid X, which is the direct precursor of the proximal sugar of Lipid A (15-16).

Next, the disaccharide synthase LpxB condenses Lipid X with a molecule of UDP-2,3-diacylglucosamine (41-42). This results in the loss of UDP, and the transfer of the 2,3-diacylglucosamine to position 6 of the lipid X, generating the β,1′-6 linkage found in all lipid A molecules. Lipid IV$_A$ is generated by the incorporation of a 4′-monophosphate by LpxK, a specific membrane-bound kinase (72-73, 160). The bifunctional Kdo transferase encoded by WaaA then adds two Kdo residues, followed by the addition to

two fatty acyl chains which in *E. coli* are laurate and myristate added by the LpxL and LpxM respectively (38). Recent work has demonstrated that Kdo is present on lipid A to provide the right substrates for the LpxL and LpxM, resulting in the synthesis of penta- and hexaacylated lipid A, which is optimal for the lipid A flippase MsbA (161).

1.1.1.2.   Core oligosaccharide

The core oligosaccharide region is divided into the inner and outer core. The inner core is conserved within a genus and species and always contains Kdo, whereas the outer core is not as conserved and shows more structural diversity (66). The outer core is the attachment site for O antigen and is also the region of LPS that contains many receptors for bacteriophages (86). In *Enterobacteriaceae* the inner core contains Kdo and *glycero-manno*-heptopyranose (heptose) and can contain other sugars as well, depending on the organism and core type. The inner core is important for maintaining outer membrane stability, as demonstrated in *E. coli* and *Salmonella* where "deep rough" mutants that are devoid of the heptose have decreased amounts of outer membrane protein corresponding to an increase in phospholipids (66). This increase in phospholipids also causes a hypersensitivity to hydrophobic compounds, due to the rapid penetration of these compounds through the phospholipid bilayer regions of the membrane (138). These mutants also lose cell surface organelles such as pili and flagella, release periplasmic enzymes and have increased susceptibility to phagocytosis by macrophages. In addition, there is an induction of the Rcs phosphorelay system causing an up regulation of colanic acid synthesis (66, 157).

Phosphorylation of heptose sugars in the inner core has also been shown to affect outer membrane stability. An *E. coli* mutant containing full LPS with O antigen but lacking phosphorylation of the inner core heptoses has increased sensitivity to hydrophobic compounds (224). In *S. enterica* a similar mutant is significantly more sensitive to the CAMP polymyxin and is attenuated for virulence in mouse infection models (223).

1.1.1.2.1.   Core oligosaccharide synthesis

Lipid A-Kdo$_2$ serves as the acceptor for the assembly of the inner core oligosaccharide by sequential glycosyl transfer from nucleotide sugar precursors (158). The requirement for a heptose residue precursor, ADP-L (or D)-glycero-D-manno-heptose, was first identified in *Salmonella enterica* serovar Minnesota and *Shigella sonnei* (108). All the *E. coli* proteins in the pathway for the synthesis of this precursor were purified and the precursor was synthesized *in vitro* (106). The isomerase GmhA converts sedoheptulose 7-phosphate to D-*glycero*-D-*manno*-heptose 7-phosphate (26-27). The first domain of the bifunctional enzyme HldE adds a second phosphate, donated by an ATP molecule, resulting in the formation of D-*glycero*-D-*manno*-heptose 1,7-bisphosphate. The phosphatase GmhB dephosphorylates at position 7, giving rise to D-*glycero*-D-*manno*-heptose-1-phosphate. The HldE adenylyltransferase domain then catalyzes the formation of ADP-D-*glycero*-D-*manno*-heptose (204). HldD, the terminal epimerase enzyme, then converts this product to the preferred heptosyltransferase substrate ADP-L-*glycero*-D-*manno*-heptose (105). The two heptoses in the inner core are added from ADP-L (or D)-*glycero*-D-*manno*-heptose by the heptosyltransferases WaaC and WaaF (78, 227).

The composition of the outer core will depend on the bacteria and therefore the enzymes will differ. In *E. coli*, glucose is the first sugar in the outer core and is added by WaaG (158). Subsequently, two transferases add α1,3- and α1,2-linked hexoses, respectively, and depending on the specific core chemotype (158).

1.1.1.3.   O antigen

The O antigen is the most variable component of LPS giving rise to a large number of O-serotypes (157). These polysaccharides can be either homopolymers or heteropolymers and the O-units may be linear or branched. There is also variation in the glycosidic linkages between the O-units. There are up to 60 O antigens in *S. enterica* and 164 in *E. coli*; however they only share three O antigen structures (212). The length of the O antigen can also differ. In *S. enterica* Typhimurium the length can be ≥ 100 repeating units (133) and in *Pseudomonas aeruginosa* ≥ 50 repeating units for the heteropolymer B-band O antigen (111).

The O antigen extends from the cell surface and is therefore the most surface-exposed portion of LPS. It protects against killing by serum complement by preventing proper insertion of the membrane attack complex in the outer membrane. For example, *S. enterica* cells with a minimum protective length of >4 and <15 repeat units are resistant to complement and an increase in the length of the O antigen has also been correlated with a decrease in uptake by macrophages (133). This is also the case in *Burkholderia cenocepacia* with both macrophages and epithelial cells (168).

1.1.1.3.1.   O antigen biogenesis

O antigens are synthesized in diverse pathways depending on the bacteria and composition of the polysaccharide. The common feature is that they are all built on the lipid carrier undecaprenyl monophosphate (Und-P) (see section 1.3) (Fig. 1.3). An initiating enzyme (see sections 1.2.1 and 1.2.2) catalyzes the transfer of a sugar-1-P onto Und-P resulting in a phosphoanhydride bond in the product of the reaction, Und-P-P-sugar (203). The majority of bacteria utilize the first two pathways described below.

1.1.1.3.1.1.   Wzy/Wzx- dependent pathway

O antigens synthesized in Wzy/Wzx dependent pathway are heteropolymers with synthesis of the O-unit occurring on the cytoplasmic face of the inner membrane. The resulting Und-PP-O antigen is processed by a series of inner membrane proteins. It is translocated to the periplasm by the Wzx flippase and elongated by the Wzy polymerase under the regulation of the Wzz co-polymerase. The flippase activity of Wzx was first identified in a *wzx* mutant of *S. enterica* that accumulated radiolabeled Und-P-P-O antigen in the cytoplasm (118). Wzx activity does not require a complete O-unit as UDP-GlcNAc can be attached to lipid A-core OS in *E. coli* K-12 (62) suggesting that only the first sugar is required as a substrate for Wzx. The specificity of Wzx for this priming sugar was tested in the Valvano laboratory where an *E. coli* K-12 strain expressing genes required for O16 production was tested with Wzx homologues from various bacteria. Wzx proteins from strains utilizing Und-P-P-GlcNAc or Und-PP-GalNAc O antigen
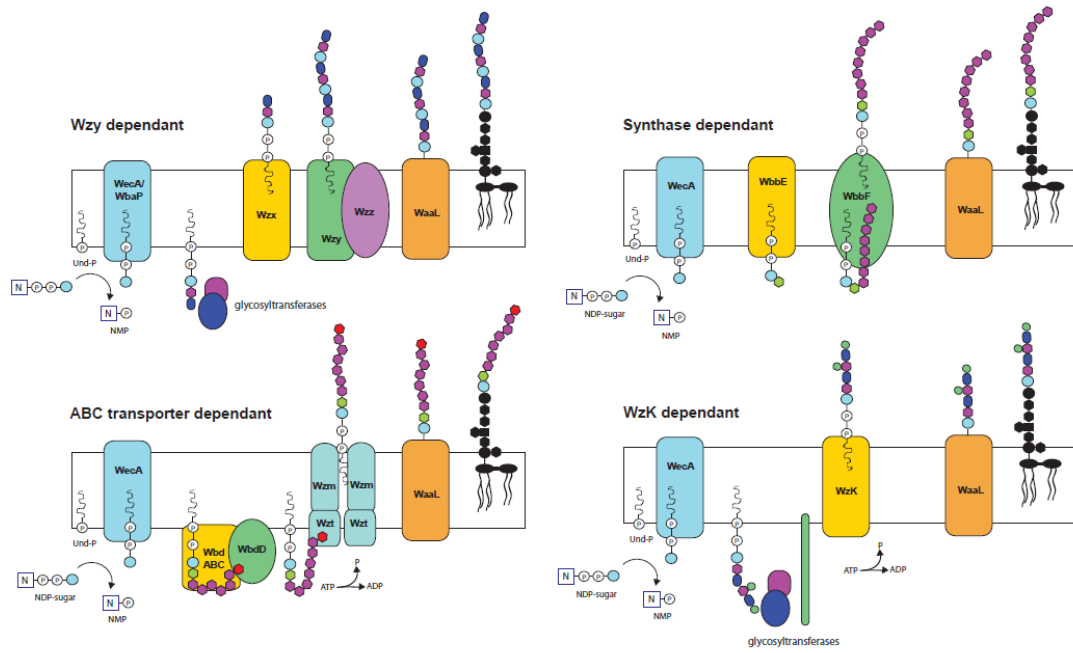
Figure 1. 3. O-antigen biogenesis pathways.

precursors fully complemented O16 LPS whereas Wzx proteins from strains with UDP-PP-FucNAc and Und-PP-Gal resulted in partial and poor complementation (124). These results suggest that Wzx proteins are specific for either *N*-acetylhexosamine or hexose sugars.

In addition to polymerizing the O-unit, Wzy also determines the linkage between the O-units, which can define different serotypes (117). Wzy itself also has high specificity for the O-unit or for glycosidic linkages within it (216-217). Wzz proteins however are not specific for the O-unit they regulate and can be interchanged (28). Certain bacteria such as *Shigella flexneri* contain two Wzz proteins which modulate a long and short O antigen (31). Recently both Wzz and Wzy were purified and their ability to polymerase a O-unit and regulate the chain length was demonstrated *in vitro* (219). Although many theories exist concerning the mechanism of action of these two proteins, this has yet to be elucidated.

1.1.1.3.1.2.   ABC Transporter-Dependent Pathway

ATP-binding cassette (ABC) transporters use the energy of ATP binding and hydrolysis to transport molecules across a membrane. In contrast to the Wzy/Wzx dependent pathway, O antigen synthesized by the ABC transporter pathway is polymerized on the cytoplasmic face of the inner membrane and the complete O antigen is exported. O antigens synthesized in this pathway are often homopolymers, with the exception of the priming and adaptor sugars. Unlike the Wzy/Wzx dependent pathway, the initiating sugar is typically only present once in the polymer as there is no repeating O-unit. Although the O antigen may contain the same repeating sugar, the differences in the linkages in between can still provide antigenicity (117).

Much of the current knowledge on ABC transporters was gained through research in *E. coli* O8 and O9. In these organisms, upon the addition of GlcNAc-1-P onto Und-P, the mannosyl transferases WbdC and WbdB synthesize the disaccharide adaptor followed by the addition of mannose sugars by WbdA and WbdB, which elongate the polymer (104, 202). Phosphorylation and/or methylation modification of the non-reducing end of the nascent Und-P-P-O antigen chain by WbdD terminates the O antigen chain polymerization (44). As there is no copolymerase present in this pathway, WbdD also

controls the length of the O antigen chain (37) and this protein is also required for the export of Und-PP-O antigen by the ABC transporter (44).

ABC-2 transporters comprise a membrane protein transporter, Wzm and an ATP binding protein, Wzt. The ATP binding motif of Wzt is located in the N-terminal part of the protein whereas the C-terminal domain is required for substrate specificity (44-45, 202). This specificity was determined by the inability of Wzt proteins in *E. coli* O8 and O9 to cross-complement (44). ATP hydrolysis is presumably required for the export of the nascent polysaccharide chain through the membrane transporter.

1.1.1.3.1.3. Synthase Pathway

There is only one example of an O antigen synthesized in this pathway which is the plasmid encoded O:54 of *S. enterica* serovar Borreze (99-101). It is a homopolymer *N*-acetylmannosamine (ManNAc). WecA initiates the reaction resulting in Und-PP-GlcNAc after which the first ManNAc residue is added by the non-processive transferase WbbE (98). WbbF, an inner membrane protein and a member of the HasA (hyaluronan synthase) family of glycosaminoglycan glycosyltransferases (49), is a synthase responsible for both the elongation of the O antigen and the transport across the plasma membrane, although little is known about the mechanism of the latter process (48). Other members of the HasA family are involved in biosynthesis of hyaluronic acid capsules in group A streptococci (50, 56), type 3 capsules of *S. pneumoniae* (30, 64) as well as the biosynthesis of cellulose and chitin (48-49).

1.1.1.3.1.4. Wzk Mediated Pathway

Helicobacter pyloris is the only known organism that synthesizes O antigen in this novel pathway recently discovered by Hug *et al* (88). WecA initiates synthesis by transferring GlcNAc-1-phosphate to Und-P after which processive glycosyltransferases alternate between adding galactose (Gal) and GlcNAc residues. Lewis antigen is generated by fucosyltransferases that attach fucose residues at select positions on the O antigen backbone. This unique component of this pathway is the flippase Wzk, a homolog of *Campylobacter jejuni* PglK (209). *C. jejuni* PglK is involved in translocating Und-PP-linked heptasaccharide to the periplasmic space for N-glycosylation of proteins

(102). To determine the role of WzK in *H. pylori*, complementation experiments in *E. coli* were performed where N-glycosylation machinery was expressed in a glycan flippase mutant of *E. coli*. Expression of a *C. jejuni* acceptor protein, AcrA, in *E. coli* was used to test for glycosylation. Wzk was also shown to restore flippase activity in a Wzx mutant of *E. coli.* Together these results indicate that WzK in *H. pylori* translocates the complete Und-P-P- linked O antigen polymer to the periplasmic space. Further studies are required to (88)determine if Wzk has ATPase activity like PglK (9). The O antigen is subsequently ligated to lipid A-core by WaaL (88).

1.1.1.4.   Movement of Lipid A-core across the membrane

The lipid A-core acceptor molecule of the ligation reaction is synthesized in the cytoplasm and moved across the inner membrane prior to ligation. This translocation step is attributed to the MsbA protein, which is also responsible for exporting glycerophospholipids to the periplasmic space (51-52, 231). MsbA is conserved amongst Gram-negative bacteria. Monomers containing transmembrane helices and an ATP-binding domain make up the functioning MsbA homodimer (213). The efficiency of the reaction depends on the lipid A-core molecule with complete molecules given preference over precursors, such as lipid A-Kdo$_2$ from rough core *E. coli* mutants (156). MsbA is essential in *E. coli* as defective MsbA leads to the accumulation of lipid A-core and phospholipids and/or precursors in the cytoplasm (53).

1.1.1.5.   Ligation

The ligation step occurs in the periplasmic space where the complete O antigen is attached to the terminal sugar of the lipid A-core oligosaccharide. The enzyme responsible for the ligation reaction is WaaL, an inner membrane protein with multiple transmembrane helices. The mechanism of the reaction is currently not known. WaaL proteins show specificity for the acceptor lipid-A core molecule (2, 84-85, 173) but not for the O antigen and can attach sugars from Und-P-P-linked oligosaccharides, polysaccharides, and single sugars to lipid A-core. These donor sugars can come from different biosynthesis pathways. For example, in *E. coli* K-12 a small portion of colanic acid (see section 1.2.2.1.5), a polysaccharide mesh which is usually loosely associated

with the cell, can be covalently linked to lipid A-core by WaaL at the same attachment site position for O antigen (128). Recently Tang and Mintz (190) have shown that extracellular matrix protein adhesion (EmaA), an adhesion protein of *Aggregatibacter actinomycetemcomitans*, is glycosylated by a carbohydrate similar to O antigen which they suggest, based on their genetic studies, is ligated to EmaA by WaaL. Glycosylation of other prokaryote components is carried out by functional homologues of WaaL such as *Neisseria meningitides* PglL and *P. aeruginosa* 1244 PilO, which are responsible for O-glycosylation of pili in these organisms (32, 61, 154).

Although WaaL proteins seem to function similarly they differ in their requirement for ATP. The ligation reaction occurs in the periplasm, which is devoid of ATP (155). However, Abeyrathne and Lam (3) found that purified *P. aeruginosa* WaaL has an absolute requirement for ATP *in vitro*, although WaaL does not contain any ATP/binding/hydrolysis motifs. The authors suggest two motifs similar to a Walker A and Walker B motif in the periplasm and cytoplasm respectively. Recently, Hug *et al* (88) demonstrated *in vitro* ligase activity of purified *H. pylori* without the addition of ATP and the addition of 2 mM ATP did not improve ligase activity. Similarly, recent work in the Valvano laboratory has shown that the *E. coli* K-12 WaaL does not require ATP (164).

1.1.1.6. LPS export from the inner to outer membrane

The pathway for the export of LPS from the inner to outer membrane was only recently elucidated. LPS from the inner membrane is extracted by a complex of proteins (137) including inner membrane protein LptC (179), ABC transporter proteins LptF and LptG (165), and the inner membrane associated soluble LptB which functions as an ATPase (178). Upon extraction, LPS associates with LptA, which was initially found to locate to the periplasm (178), but has also been shown to associate with membranes (22). Recent work has shown it interacts with lipid A (194). In addition, structural studies on LptA show that in the presence of LPS it forms fibers (186) and can still be transported to the outer membrane in spheroplasts (192). Due to this finding and the localization of LptA in the cell, there are two schools of thought on the transport of LPS across the periplasm. The soluble intermediate model proposes that LptA is a chaperone that carries the LPS to the outer membrane, and the trans-envelope complex model proposes that

LptA proteins form a bridge across the periplasm (166). Recent work by Sperandeo *et al* supports the latter theory as they found that inner membrane protein LptC and LptA interact and form a stable complex and without this interaction LptA is degraded (180). Upon reaching the outer membrane LPS is assembled by the complex formed by LptD and LptE (35, 220). Both are essential outer membrane proteins but only LptE has been shown to interact with LPS. Recent work showing that all seven Lpt proteins are in a novel outer membrane fraction (OM_L, outer membrane light) that contains both outer and inner membrane components suggesting a complex that spans the periplasm, further supporting the trans-envelope theory (34) .

1.2. Initiating glycosyltransferases

1.2.1. Polyisoprenyl-phosphate *N*-Acetylaminosugar-1-phosphate transferase (PNPT) family

The PNPT family consists of membrane proteins that initiate the synthesis of glycans by utilizing UDP-*N*-acetylaminosugars, which are in most cases UDP-*N*-acetylhexosamines. The eukaryotic members of this family are expressed in the rough endoplasmic reticulum where they transfer *N*-acetylglucosamine-1-P to the lipid carrier dolichol phosphate. This reaction is required to synthesize precursors for N-glycosylation. Prokaryotic PNPT proteins initiate the synthesis of the various polysaccharides described below (203). Others include *Yersinia enterocolitica* O:3 outer core which requires the FucNAc-1-P transferase WbcO (175) and *Streptococcus mutans* rhamnose-glucose polysaccharide requiring the apparent WecA-WbcO hybrid RgpG (222).

1.2.1.1. Bacterial PNPT initiated glycans

1.2.1.1.1. LPS O antigen

In the majority of Gram-negative bacteria, synthesis of O antigen (see section 1.1.1.3) involves an identical initiation step catalyzed by the PNPT WecA, which adds GlcNAc-1-P to Und-P. Exceptions include *S. enterica* utilizes the PHPT WbaP (see section
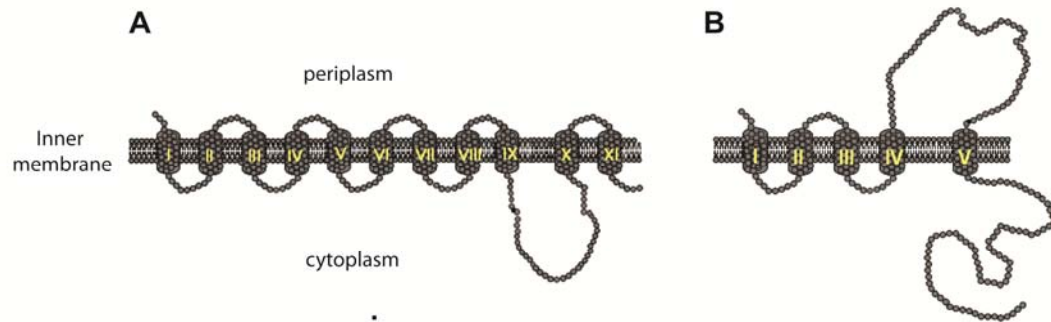
Figure 1.4. Predicted topology of (A) *E. coli* K-12 WecA and (B) *S. enterica* WbaP. Figure adapted from Valvano,(2003) (202).

1.2.2.1) and *P. aeruginosa* WbpL, a bifunctional enzyme utilizing both UDP-GlcNAc and UDP-FucNAc (163) to initiate A-band and B-band O antigen synthesis respectively. A-band O antigen is a homopolymer of rhamnose sugars that is synthesized in the ABC transporter pathway (see section 1.1.1.3.1.2), whereas B-band O antigen is a heteropolymer containing a trisaccharide repeating unit synthesized in the Wzy/Wzx dependent pathway (see section 1.1.1.3.1.1) (163).

The *E. coli* K-12 WecA has been studied extensively by the Valvano laboratory. Topology mapping with thiol labeling of cysteine residues and GFP fusion experiments have revealed that WecA has 11 TMs, five cytoplasmic loops, five periplasmic loops and a cytoplasmic C-terminus (114) (Fig. 1.4). Sequence alignments have identified conserved residues in the cytoplasmic loops that have been experimentally investigated. Mutations of residues presumably involved in catalysis including D90 and D91 in cytoplasmic loop 2 and D156 and D159 in cytoplasmic loop 3 demonstrated they are important for WecA activity (13). Further characterization with Mg^{2+} and Mn^{2+} led to the conclusion the D91 was important for interactions with these divalent cations (114). The drastic defect upon mutation of D156, and its location in the cytoplasm, suggest it has a critical role in WecA activity and may serve as the catalytic nucleophile. Residue D217, previously assigned to a transmembrane domain, is currently being investigated in the Valvano laboratory. Recent work using cysteine labeling has demonstrated that it is located in the cytoplasm and mutation of this residue to an alanine also causes a defect in activity (67). Another region of importance is on cytoplasmic loop 5 containing conserved histidines ($H_{278}IHH_{281}$) which are important in carbohydrate recognition as mutations in these residues lead to a loss in the ability of the enzyme to bind the UDP-GlcNAc analog tunicamycin (12, 83). Al-Dabbagh *et al* (8) have recently purified and performed biochemical characterization of the *Thermatoga maritima* WecA which will aid future studies on this protein to determine a catalytic mechanism.

1.2.1.1.2. Enterobacterial common antigen

Enterobacterial common antigen is found in all members of the *Enterobacteriaceae*. In rough bacteria that do not contain O antigen, this polysaccharide can be linked to lipid A-core (ECA_{LPS}). These bacteria, as well as bacteria producing smooth O antigen also

contain the unligated cyclic form (ECA$_{CYC}$) (18). In addition, ECA polysaccharide chains can also be covalently linked to phosphoglycerides via phosphodiester bonds (ECA$_{PG}$) (17). The ECA repeating unit is a trisaccharide consisting of a linear heteropolysaccharide chain comprising *N*-acetylglucosamine (GlcNAc), *N*-acetyl-D-mannosaminuronic acid (ManNAcA), and 4-acetamido-4,6-dideoxy-D-galactose (Fuc4NAc) (17). Synthesis of ECA is initiated by WecA and the ECA trisaccharide repeat unit is assembled as an Und-PP-linked intermediate, in a process believed to be similar to O antigen assembly, by the Wzy/Wzx-dependent pathway (see section 1.1.1.3.1.1). ECA$_{LPS}$ is ligated to lipid A-core and ECA$_{PG}$ to phosphoglycerides via an unknown mechanism. The export of ECA$_{CYC}$ is believed to require a cyclase that works in coordination with Wzy (94).

1.2.1.1.3. Peptidoglycan

Peptidoglycan is an essential component of the Gram positive and negative cell wall providing rigid cell structure. This mesh like structure is composed of alternating GlcNAc and *N*-acetylmuramic acid (MurNAc) sugars connected by β-1,4 linkages. A short amino acid chain of 5 residues is added to MurNAc sugars and these peptides are cross-linked by transpeptidases. The cross-linking differs between bacteria and alters the strength of the peptidoglycan (120). Peptidoglycan prevents osmotic lysis, maintains bacterial cell shape and is critical in the cell division process (136). It is the target of many antimicrobials such as the β-lactams (74).

The synthesis of peptidoglycan begins at the inner side of the cytoplasmic membrane by the transfer of *N*-acetylmuramyl-pentapeptide-1-P to Und-P by the PNPT MraY, forming the precursor lipid I. MurG catalyzes the transfer of GlcNAc to lipid I forming lipid II which is flipped to the other side of the membrane and polymerized by penicillin binding proteins (136). MraY has been studied to determine its topology, active site(s) and catalytic mechanism. The topology of the *E. coli* and *Staphylococcus aureus* MraY has been experimentally determined by β-lactamase fusions identifying 10 transmembranes segments, five cytoplasmic loops and six periplasmic domains including the N- and C-terminal ends (24). Despite the presence of multiple transmembrane regions, Bouhss *et al* have reported purification of *Bacillus subtilis* MraY and

characterized the protein *in vitro*. However, over six years later there is no structural information on this protein (23).

Llyod *et al* (119) partially purified *E. coli* MraY and investigated conserved putative catalytic residues within cytoplasmic loops, identifying three aspartic acids as critical for *in vitro* activity. These are residues D115 and D116 of cytoplasmic loop 2 and D267 of cytoplasmic loop 4 which correspond to *E. coli* WecA D90, D91 and D217 (section 1.2.1.1.1). Al-Dabbagh *et al* (7) mutated the corresponding residues in the *B. subtilis* MraY and generated similar results with the exception of D231, corresponding to *E. coli* MraY D267 and WecA D217, still maintaining activity. More work is needed to determine which residues make up the active sites of these enzymes.

1.2.1.1.4.   Wall teichoic acids

Gram-positive bacterial cell walls contain teichoic acids (TA), which are anionic polymers covalently bound to peptidoglycan, wall teichoic acid (WTC), or anchored to the cytoplasmic membrane, lipoteichoic acid (LTA). Some Gram-positive bacteria do not contain WTC or LTA but contain other anionic polysaccharide. Many roles for TA have been proposed, including maintenance of cell shape (181, 214), resistance to antimicrobial peptides (1, 109-110), involvement in biofilm formation (79), acid tolerance (25) and the release of secreted proteins (143). The composition and length of both WTA and LTA differ between species and strains (208).

The pathway for WTA synthesis in *B. subtilis* has been well characterized and serves as a model to study TA. WTA is composed predominantly of glycerol phosphate (Gro-P) subunits and in *B. subtilis* 168 these can be up to 60 subunits in length (151). In *B. subtilis*, the majority of genes required for the synthesis of WTA are in operons within the *tag* locus (87, 125, 177). The PNPT TagO initiates synthesis by transfer of GlcNAc-1-P to Und-P on the inner side of the cytoplasmic membrane (177). ManNAc is added by TagA followed by the addition of glycerol phosphate (Gro-P) to the disaccharide by TagB (21) and TagF continues with adding Gro-P subunits (172). TagE adds glucose to the Gro-P subunits (153) and once the polymer is complete it is taken through the membrane by an ABC-transport type system composed of TagG and TagH, comprising the nuclear binding domain and membrane spanning domain respectively (112).

WTA was believed to be critical for viability of the cells that produced it; however the Brown laboratory has shown that cells containing deletion of the initiating enzyme, TagO, are viable (46). The deletion of genes encoding enzymes found latter in the WTA synthetic pathway are lethal due to a build-up of precursors linked to Und-P (see section 1.3) which is an essential component of peptidoglycan biosynthesis (see section 1.2.1.1.3) and hence required for cell viability.

1.2.2. Polyisoprenyl-phosphate hexose-1-phosphate transferase (PHPT) family

The PHPT family consists of membrane proteins that catalyze the synthesis of the prokaryotic glycans described below. Unlike the PNPT family, the PHPT family contains no eukaryotic homologues.

1.2.2.1.   PHPT initiated glycans

1.2.2.1.1. *S. enterica* O antigen

Salmonella enterica is a Gram negative bacterium with six subspecies and over 2500 serovars (5). Two important serovars are *Salmonella enterica* serovar Typhimurium and serovar Typhi which are causative agents of gastroenteritis and typhoid fever respectively. The pathogenesis of *S. enterica* is in part due the O antigen production. Regulation of O antigen polymerization is important, as cells lacking *wzz* genes produce shorter polymers decreasing virulence and increasing susceptibility to complement mediating killing (134-135). The length of the *S*. Typhimurium O antigen is also important in the entry of *S*. Typhimurium into the cell (133).

Within *S. enterica* there are 46 different O antigens with *S*. Typhimurium and *S*. Typhi belonging to Group B and Group D1 respectively (43). They consist of galactose, rhamnose, mannose and either abequose or tyvelose. In other members of the *Enterobacteriaceae* studied to date, the WecA enzyme initiates the synthesis of O antigen (see section 1.2.1.1.1.). In contrast, in *S. enterica* this reaction is catalyzed by the WbaP enzyme, a prototypic member of the PHPT family that transfers galactose-1-phosphate from UDP-galactose to Und-P. It is predicted to contain five transmembrane helices and a large periplasmic loop (202) (Fig. 1.4). The presence of an enzyme that was essential for

O antigen production in *S. enterica* and required UDP-galactose was first discovered in a *galE* (UDP-galactose-4-epimerase) mutant of *S.* Typhimurium (139-140, 146-147). Supplementation of galactose restored the production of wild type LPS. Osborn *et al* (146) further characterized this enzyme and reported a requirement for Mn^{2+} or Mg^{2+}. The *wbaP* gene was initially annotated as *rfbP*, and is found in a O antigen cluster (210). Involvement of the *wb** (previously *rfb*) genes in O antigen synthesis was initially identified in mutants that became sensitive to bacteriophage P22, which utilizes O antigen polymers as a receptor (218). Wang *et al* (210) continued studies on two of these mutants called *wbaP*(T) and *wbaP*(GT) containing mutations in the 5′ and 3′ end of the gene respectively. Through *in vivo* and *in vitro* experiments they assigned the galactose-1-phosphate transferase activity to the C-terminal half of the protein, and a Und-PP-O unit flipping or Und-P-P-galactose releasing function to the N-terminal half of WbaP (210-211). After initiation by WbaP, the biogenesis of the *S. enterica* occurs through the Wzy/Wzx dependent pathway (see section 1.1.1.3.1.1).

1.2.2.1.2. *Aeromonas hydrophila* O antigen

Aeromonas hydrophila O:34 is an aquatic organism that is also found in the gut of animals (198). It is the most common *Aeromonas* serogroup, accounting for 26.4% of all clinical infections including gastroenteritis and septicemia (92). The O antigen subunit of *A. hydrophila* is a tetrasaccharide made up of *N*-acetylgalactosamine (GalNAc), mannose and 6-deoxytalose (107). The gene encoding the *A. hydrophila* O antigen initiating enzyme was originally named *wecA* as the first sugar in the O antigen was GalNAc and mutation of this gene led to an absence of O antigen (93). The *A. hydrophila* 'WecA' protein was recently investigated by Merino *et al* (129) and found to more closely resemble *S. enterica* WbaP both in putative topology and sequence homology. It was therefore renamed WecP (Wec from WecA and P from WbaP). WecA mutants of *E. coli* and *Klebsiella pneumoniae* and a WbaP mutant of *S. enterica* were complemented with WecP and the UDP-*N*-acetylgalactosamine 4-epimerase Gne (which is required to produce GalNAc). Production of O antigen was restored in *E. coli*, and *K. pneumoniae* O1, but not in *S.* Typhimurium. Analysis of the LPS from the first two strains detected the presence of GalNAc. The authors found that in an *in vitro* assay containing total

membranes and ^{14}C labeled substrate, the enzyme was able to transfer GalNAc-1-phosphate from UDP-GalNAc to the lipid fraction. Although the authors conclude that WecP utilizes UDP-GalNAc as a substrate, more work is required to confirm these findings. Activity of pure WecP in the presence of ^{14}C labeled UDP-GalNAc and exogenous Und-P would prove that WecP does truly transfer GalNAc.

1.2.2.1.3. Capsular polysaccharides

Many Gram positive and Gram-negative bacteria produce capsular polysaccharides (CPS), which are associated with the cell via linkages to cellular components. These capsule polymers have a high molecular weight and are longer than O antigen (33). *E. coli* and *Streptococcus pneumoniae* capsules have been extensively studied and have been shown to be a major factor in biofilm formation and virulence. *E. coli* produces over 80 types of capsules (145), referred to as K antigens, as they are serotype specific [reviewed in (215)]. These *E. coli* K antigens fall into groups 1 to 4 and are defined by their composition, assembly and virulence. Group 1 and 4 K antigens are found in *E. coli* that cause intestinal infections. They are acidic in nature and can be ligated to lipid A – core producing K_{LPS}. Biosynthesis of these two groups is Wzy and Wzx dependent (see section 1.1.1.3.1.1). Group 2 and 3 vary extensively and have a different biosynthetic pathway where polymerization of the K antigen occurs in the periplasm however this pathway is not well characterized. Group 1 and 4 K antigens are built on Und-P and in *E. coli* K30, the PHPT WbaP initiates capsule synthesis via transfer of galactose-1-phosphate to Und-P (Fig. 1.5). Instead of the Wzz protein, polymerization is regulated by the tetrameric tyrosine autokinase Wzc (40, 57). This protein contains C-terminal tyrosine residues that are autophosphorylated or are dephosphorylated by the Wzb protein located in the cytoplasm. The complete polymer is then taken through a channel formed by a Wza octamer (58) and assembled on the cell surface by Wzi (159) .

In the Gram-positive bacteria *S. pneumoniae* capsules are essential for virulence. *S. pneumoniae* is a human pathogen that causes pneumonia, bacteremia and meningitis resulting in high mortality (11). Many vaccines have been developed against *S. pneumoniae* CPS, but the efficiency of these vaccines is limited due to antigenic differences in CPS composition and structure (11). There are 90 different *S. pneumoniae*

serotypes which are based on differences between the capsules they produce (19, 205). Most *S. pneumoniae* capsules are synthesized in a pathway resembling that of group 1 capsules in *E. coli* (19, 36) (described above) and their synthesis requires initiation by the a PHPT, which will catalyze the transfer of glucose-1-phosphate to Und-P. Examples include WchA of *S. pneumoniae* serotype 8 and CpsE of serotype 2 (29, 148, 221). As with *S. enterica* WbaP (see section 1.2.2.1), expression of the C-terminal domain of WchA is sufficient for activity *in vitro* (148). An interesting finding made by Xayarath and Yother (221) was that upon deletion of Cps2K, a glycosyltransferase that adds the terminal sugar to the CPS unit, the flippase Cps2J, or the polymerase Cps2H, suppressor mutations were found predominantly in the *cpsE* gene (26 of 29) with 80% in the tail region and 20% in the large loop domain. The mutation of these capsule biosynthesis proteins would lead up to a build-up of Und-P linked precursors depleting Und-P for essential cell wall synthesis. Although suppressor mutations in the C-terminal domain are explained by the enzymatic activity localized in these regions, the mutations in the loop domain are more puzzling. This work would suggest that the loop region also plays a role either directly or indirectly in regulation of catalytic activity.

1.2.2.1.4. Exopolysaccharides

Exopolysaccharide is secreted by bacteria into the extracellular milieu and unlike O antigen and capsule, which are attached to the cell, EPS has no cell anchor. In general, depending on which bacteria produce it, EPS can have major differences in composition and hence chemical and physical properties (187). Therefore the function of EPS as a whole cannot be generalized, however many have been shown to be important in virulence and biofilm formation. They have been studied in human pathogens such as *P. aeruginosa* and *Burkholderia cenocepacia* which produce alginate and cepacian EPS respectively (126, 206). Both organisms can be opportunistic pathogens infecting the lungs of individuals with cystic fibrosis. In *B. cenocepacia* the PHPT BceB initiates the synthesis of EPS and utilizes UDP-glucose. The *bceB* gene from a mucoid *B. cenocepacia* cystic fibrosis isolate was used to study BceB. The authors used the
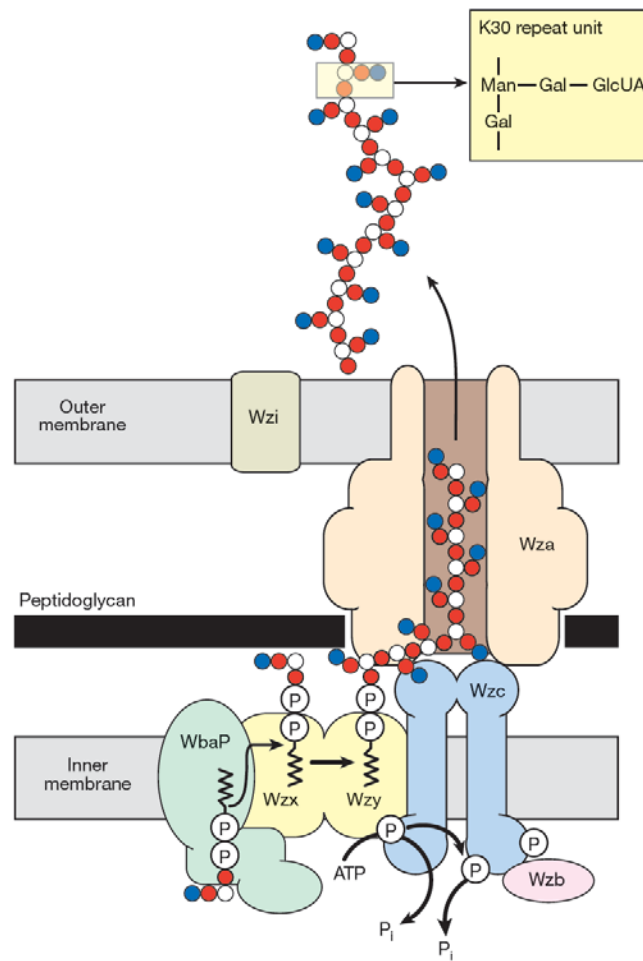
Figure. 1.5. Capsule assembly in *E.coli* K30. Figure is from Dong *et al* (2006)(55).

periplasmic and cytoplasmic reporter fusions PhoA and LacZ respectively. After obtaining results from two PhoA and two LacZ fusions it was proposed that BceB contains 7 transmembrane regions including two large cytoplasmic domains. The authors also found that similar to *S. enterica* WbaP, BceB requires Mg^{2+} for *in vitro* activity. In addition, aspartic acid residues 339, 355 and 439 were shown to be important for *in vitro* activity, however a role for these residues was not identified.

Many plant pathogens such as *Rhizobium meliloti* (succinoglycan production), *Erwinia stewartii* (stewartan synthesis), and *Xanthomonas campestris* (xanthan gum synthesis) also produce EPS (10). The latter also has a positive attribute as a producer of a well known EPS xanthum gum, which is used in many industrial and food applications. Xanthum gum is composed of polymerized pentasaccharide repeating units which contain glucose, mannose, glucuronic acid, and mannose (90). Synthesis of xanthum gum polymer is initiated by the PHPT protein GumD which transfers glucose-1-phosphate to Und-P (96). Similarly, *Erwinia amylovora* and *E. stewartii* contain the PHPT AmsG, which initiates EPS synthesis by transfer of galactose-1-phosphate. *R. meliloti* contains ExoY catalyzing the same reaction (115). As EPS differs greatly between organisms, a common pathway for synthesis of the polysaccharide cannot be outlined, however as describe above they all have a common initiation step. The export of EPS described here has not been elucidated.

1.2.2.1.5. Colanic acid

Colanic acid (or M antigen) is a type of exopolysaccharide that is loosely associated with the cell surface. It was first described by Goebel (76) in 1963 after a laboratory stock of *E. coli* K235 overproduced it, most likely due to mutation(s). Subsequent to its discovery, it was found expressed in other strains of *E. coli*, *S. enterica* and other members of the Enterobacteriaceae (77). In all instances the composition of colanic acid was found to be identical, consisting of the sugars galactose, glucose, glucuronic acid, and fucose (71, 76-77, 128). The structure of colanic acid was determined by mass spectroscopy (71) and was later refined by NMR (128). This analysis revealed that the polymer is made up of repeat units with glucose as the initiating sugar. *In E. coli* K-12,

the genes required for the production of colanic acid are found within the *wca* (formerly *cps*) locus and encode putative glycosyltransferases as well as enzymes that are required for the precursor sugars (184). Transcription of this locus is tightly regulated by the Rcs phosphorelay system comprised of a two-component system with RcsB as the effector and RcsC as the sensor kinase [reviewed in (122)]. The phosphate from RcsC is transferred to the cytoplasmic RcsB. A heterodimer of RcsB-P:RcsA controls the transcription of many genes including upregulation of *wca* genes and downregulation of the flagellar genes *fhIDC*. When not bound to RcsB, RcsA is highly unstable as it is degraded by the Lon protease. Colanic acid is only produced at lower temperatures, but can be induced by introducing extra copies of *rcsA* or by the addition of *p*-fluorophenylalanine (FPA) to the media, which when incorporated into certain proteins alters their activity (65). Meredith *et al* (128) have demonstrated that after induction of *E. coli* K-12 with FPA, the cells add some colanic acid to lipid A-core at the same attachment site as O antigen.

The lack of colanic acid production at physiological temperatures suggests it has a role in survival outside of the host. In *E. coli* K-12, colanic acid is not required for initial attachment to a substrate, but is required in latter stages of biofilm formation for development of a 3-dimensional complex network (47, 54). *In vivo* experiments with *S. enterica* infection in mice found that a colanic acid overproducing strain was actually comprised for survival, as it may mask other surface carbohydrates such as O antigen (70).

WcaJ is likely the glucose-1-phoshate transferase catalyzing the initiation of colanic acid synthesis. This protein shares high sequence homology to the PHPT *S. enterica* WbaP. Colanic acid synthesis and export occurs in a Wzx-Wzy pathway that is similar to Group 1 capsules [see section 1.2.2.1.3], where the completed colanic acid subunit is flipped to the periplasm by Wzx and polymerized by Wzy. Export of the polymer is carried out by Wza, Wzb and Wzc but does not involve Wzi, which is responsible for surface attachment in group 1 CPS.

1.2.2.1.6. S-layer glycans

An S-layer is an organized arrangement of protein or glycosylated protein found on the cell surface [For reviews see (131, 162, 170)]. In contrast to other glycoproteins found within the cell, S-layer glycoproteins are able to assemble into 2-dimensional crystalline structures that cover the whole bacterial surface (169). S-layers were first identified in Gram-positive bacteria and archaea but recently the pathogenic Gram-negative bacteria *Tannerella forsythia* and *Parabacteroides distasonis* have also been found to contain S-layers (63, 113). Although much research has been conducted on their arrangement, composition and assembly, a function for S-layers has not yet been identified (162).

Recently, the assembly of S-layers has been studied in the organisms *Geobacillus stearothermophilus* NRS 2004/3a and *Paenibacillus alvei* CCM 2051. These are glycoproteins whose glycans which contain adapter sugars linked to linear and branch repeating unit polymers respectively (170, 228). Although the repeating units differ in their sugars and arrangements, their initiation is identical; both requires a glycosyltransferase to add the priming galactose. In *Geobacillus stearothermophilus,* this enzyme was identified as WsaP, which shares high sequence homology and predicted topology to the prototype PHPT, WbaP (182). Galactose-1-phosphate transferase activity was confirmed when WsaP expression in WbaP mutants of *S.* Typhimurium and *E. coli* O9:K30 reconstituted O antigen and CPS production respectively. The C-terminal domain of WsaP was also shown to be sufficient for activity (182).

In *P. alvei*, WsfP catalyzes the initiation reaction, as S-layer glycosylation is reconstituted in a *P. alvei* WsfP mutant by complementation with *G. stearothermophilus* WsaP (228). Through biochemical experiments Steiner *et al* (183) have elucidated the S-layer glycoprotein glycan biosynthesis pathway in *Geobacillus stearothermophilus*. Following the addition of the priming galactose by WsaP, biosynthesis of the polymer continues in the cytoplasm by the action of four rhamnosyltransferases. WsaD adds a rhamnose to galactose, followed by the addition of one or two rhamnoses by the WsaC enzyme. Together, these two enzymes form the adapter. Formation of the repeating unit is carried out by the enzymes WsaE, forming α-1,2 and α-1,3 linkages, and WsaF forming β-1,2 linkages. A domain of WsaE by terminates polymerization by mediating 2-O-methylation of the terminal repeating unit (171, 183). The chain is exported via an

ABC transporter system comprising of Wzt and Wzm (144) (See section 1.1.1.3.1.2) with Wzt predicted to bind the 2-O-methylated chain. The addition of the complete glycan is added to the target protein by the oligosaccharide transferase WsaB (183).

1.2.2.1.7. Glycans for N-linked and O-linked glycosylation

Protein glycosylation was first described in S-layers where O-linked glycans were found attached to surface proteins (see section 1.2.2.1.6). Certain bacteria also produce O-glycosylated flagella and pilin, such as *Campylobacter jejuni* and *Neisseria gonorrhoeae* respectively while others (e.g. *P. aeruginosa*) can glycosylate both [For a review see (142)]. In recent years it has been established that certain bacteria can also glycosylate proteins via N-linked glycosylation. *C. jejuni* provides a good example as 130 unique glycopeptides have been found in this organism (174, 189). These glycan modifications lead to increased pathogenesis as a defect in the N-glycosylation pathway *in C. jejuni* causes a defect in adherence to and invasion of human epithelial cells *in vitro,* and a decrease in colonization of chicken *in vivo* (95, 188).

The glycosylation pathways are complex but have been well studied in *C. jejuni* and more recently in *N. gonorrhoeae.* In *C. jejuni* the proteins PglDEF form the UDP precursor UDP-bacillosamine (2,4-diacetamido-2,4,6-trideoxyglucose) from UDP-GalNAc (75, 207). The PHPT PglC then transfers bacillosamine-1-P to Und-P. Purified PglC was shown to utilize UDP-bacillosamine *in vitro* and like other PHPTs, it has a requirement of Mg^{2+}. In addition, it was not inhibited by the WecA inhibitor tunicamycin (75). Upon completion of the oligosaccharide it is taken to the periplasm by an ABC transporter protein PglK and either attached to the target protein or released into the cytoplasm (142). In *N. gonorrhoeae* the putative PHPT is PglB. PglB is bifunctional, with the C-terminal domain (PglB-ATD) catalyzing the amino acetylation of UDP-4-amino to UDP-bacillosamine and the N-terminal domain (PglB-PGTD) possessing the phospho-glycosyltransferase activity transferring bacillosamine-1-P to Und-P (82). PglB-PGTD has 52.3% sequence similarity with *C. jejuni* PglC (82). This is the first description of a PHPT protein that contains catalytic activity at the N-terminus. After completion of the glycan it is flipped to the periplasm by PglF, and either transferred to the pilin by PglL or to other proteins by PilE (142).

1.2.2.1.8. Bacterial holdfast adhesin

The holdfast adhesin (hereafter referred to as holdfast) is produced in bacteria such as *Asticcacaulis excentricus*, *A. biprosthecum* and *Caulobacter crescentus.* Unlike *Asticcacaulis* species, which produce the holdfast in association with the cell body (150), *C. crescentus* adhesin is found at the tip of the holdfast stalk (185) and helps this Gram negative aquatic organism adhere to underwater substrates. *C. crescentus* holdfast has been extensively studied and coined 'nature's strongest glue' when tested by micromanipulation and atomic force measurement experiments (199). This makes it an attractive material for applications requiring underwater adhesives. However, its exact composition is still unknown and therefore it cannot be produced synthetically. The ability of the adhesin to bind wheat germ agglutinin lectin, and its sensitivity to lysozyme and chitinase, imply that it contains polymers of β-1,4 linked GlcNAc (130). These polymers give holdfast its elastic gel-like properties (116). In addition to polysaccharides, holdfast also contains uronic acid (200) and proteins including the outer membrane associated proteins HfaA, HfaB and HfaD, which were recently shown to be required to bind holdfast to the stalk (81).

The genes *hfsEFGH*, required for the biogenesis of holdfast, were initially identified in a transposon mutagenesis screen (132) and were later investigated by Toh *et al* (193). Deletion mutants in these genes identified putative glycosyltransferases HfsG and HfsH as essential for holdfast production. HfsF is not essential and is predicted to encode a flippase, but *C. crescentus* may encode an ABC transporter system compensating for the deletion of *hfsF*. HfsE is also not essential due to two paralogs, PssY and PssZ, found in other parts of the genome. Deletion of all three proteins abolishes holdfast production but expression of just one is sufficient to restore holdfast, as demonstrated by microscopy and polystyrene binding experiments. These three proteins share high sequence homology to the PHPT proteins and are therefore proposed to initiate the synthesis of holdfast. Although holdfast contains GlcNAc this may not be the initiating sugar as members of the PHPT family studied to date do not utilize this substrate. Based on this work and studies on the export genes *hfsDAB* (176), the biogenesis of holdfast is proposed to resemble that of Group 1 capsules, with HfsA, HfsB and HfsD acting as functional

homologs of Wzc, Wzb and Wza respectively (193). More work is needed to determine the composition of holdfast. As the adhesive properties of holdfast make it a difficult material to study, understanding the function of *C. crescentus* glycosyltransferases may shed light on holdfast composition.

1.3.   Undecaprenyl monophosphate

Biogenesis of polysaccharides and glycans in the prokaryotes and eukaryotes share the requirement for a lipid carrier. In eukaryotes, this is the 95-carbon isoprenoid dolichol phosphate and in prokaryotes the 55-carbon isoprenoid Und-P (202). Und-P is the carrier of sugars and glycans involved in the synthesis of diverse products such as cell wall peptidoglycan, O antigen, ECA and other bacterial carbohydrate polymers (191). It is also and the carrier for lipid A modifications in *S*. Typhimurium with aminoarabinose (196). Shorter isoprenoids are also used in certain bacteria. Examples include decaprenyl phosphate which is required for synthesis of cell wall components in mycobacterial species (97, 121) and nonapreyl phosphate providing lipid carrier function in *Paracoccus denitrificans* (91). Biophysical experiments have shown that Und-P and other longer isoprenoids are able to change membrane structure from bilayer to nonbilayer (230). This may partially account for the ability for Und-P linked sugars to move across the membrane, with the assistance of flippase proteins. Zhou *et al* (229) reported the revised length of dolichol phosphate as 33Å. A biological membrane is 40-60Å in width, suggesting that Und-P only spans partway across the membrane (229).

1.3.1. Synthesis and recycling

Und-P is synthesized *de novo* from dephosphorylation of Und-PP by the pyrophosphate phosphatase, UppP, at the inner side of the cytoplasmic membrane. Und-P also comes from the recycling of Und-PP, which is liberated from biosynthetic pathways upon transfer of its carbohydrate cargo. As the cell synthesizes very little Und-P, recycling is crucial in maintaining adequate levels of Und-P for building of essential cell wall peptidoglycan. In *E. coli*, UppP accounts for 75% of pyrophosphatase activity in the cell (59), with the remaining activity a result of YbjG, YeiU and PgpB (60). The Valvano laboratory has demonstrated that the *E. coli* K12 membrane proteins Ybjg and

YeiU, which have periplasmic domains containing acid phosphatase motifs, are important for Und-PP recycling (191).

1.3.2. Und-P recognition by proteins

Although progress has been made in elucidating the biosynthesis of Und-P, the interactions of Und-P with proteins and sugars is still poorly understood. A polyisoprenoid recognition motif has been proposed after sequence analysis of GPT proteins as well as a putative initiating glycosyltransferase and translocase from *E. coli* K1 (229). In addition, sequence analysis of an external loop critical for the activity WaaL ligases has also led to a proposed motif important for recognition of Und-P (149).

1.4. Research objectives and summary of data obtained

Members of the PHPT family of proteins initiates the synthesis of a variety of glycans of clinical and industrial importance. However, very little is known about their structure, substrate binding and catalytic mechanism. This is in part due to the difficult nature of working with membrane proteins and a lack of assays to determine their function. PHPTs are attractive targets for antimicrobials as they are only found in prokaryotes and as they initiate the synthesis of glycans known to cause virulence, targeting them would prevent the formation of any glycan intermediates. The purpose of this work was to characterize PHPT proteins and determine their requirements for activity. The objectives were to determine the membrane topology of PHPTs, investigate the function of their domains, determine the residues important for activity and determine the substrates of members of this family. I also wanted to purify a PHPT protein to study the protein *in vitro* and for future structural analysis.

This work is divided into four results chapters (chapters two to five). In Chapter Two, I show optimized expression and detection of *S. enterica* WbaP constructs and demonstrate that the C-terminus of WbaP is active *in vivo* and *in vitro*. I also demonstrate that in wild type *S.* Typhimurium overexpression of an inactive WbaP encoding the majority of the putative periplasmic loop domain has a dominant negative effect the production of higher molecular weight O antigen polymers. In Chapter Three, I demonstrate that the C-terminus of *S. enterica* WbaP is cytoplasmic and identify seven

highly conserved residues that are critical for activity. In Chapter Four, I describe a purification strategy for the C-terminal domain of WbaP, the first hexose-1-phosphate transferase to be purified to date. I also optimize activity of the pure protein and determine the kinetic parameters. In addition, I also determine that the putative transmembrane helix V is sufficient to maintain specificity for Und-P versus other isoprenoids of varying lengths. In Chapter Five, I investigate the function of four members of the PHPT family, *E. coli* WcaJ and *C. crescentus* HfsE, PssY and Pssz. I show that both WcaJ and PssY are glucose-1-phosphate transferases and have developed an assay in *E. coli* W3110 to test other putative PHPT that also utilize UDP-glucose. In addition, I demonstrate that these glucosyltransferases can also have galactose-1-phosphate activity in a WbaP mutant of *S*. Typhimurium. In Chapter Six, I discuss the contributions my work has made to the field and outline potential experiments to continue this project in the Valvano laboratory.

1.5. Chapter one references

1. **Abachin, E., C. Poyart, E. Pellegrini, E. Milohanic, F. Fiedler, P. Berche, and P. Trieu-Cuot.** 2002. Formation of D-alanyl-lipoteichoic acid is required for adhesion and virulence of *Listeria monocytogenes*. Mol Microbiol **43:**1-14.

2. **Abeyrathne, P. D., C. Daniels, K. K. Poon, M. J. Matewish, and J. S. Lam.** 2005. Functional characterization of WaaL, a ligase associated with linking O-antigen polysaccharide to the core of *Pseudomonas aeruginosa* lipopolysaccharide. J Bacteriol **187:**3002-12.

3. **Abeyrathne, P. D., and J. S. Lam.** 2007. WaaL of *Pseudomonas aeruginosa* utilizes ATP in in vitro ligation of O antigen onto lipid A-core. Mol Microbiol **65:**1345-59.

4. **Ahn, V. E., E. I. Lo, C. K. Engel, L. Chen, P. M. Hwang, L. E. Kay, R. E. Bishop, and G. G. Prive.** 2004. A hydrocarbon ruler measures palmitate in the enzymatic acylation of endotoxin. EMBO J **23:**2931-41.

5. **Akiba, M., M. Kusumoto, and T. Iwata.** 2011. Rapid identification of *Salmonella enterica* serovars, Typhimurium, Choleraesuis, Infantis, Hadar, Enteritidis, Dublin and Gallinarum, by multiplex PCR. J Microbiol Methods **85:**9-15.

6. **Akira, S., S. Uematsu, and O. Takeuchi.** 2006. Pathogen recognition and innate immunity. Cell **124:**783-801.

7. **Al-Dabbagh, B., X. Henry, M. El Ghachi, G. Auger, D. Blanot, C. Parquet, D. Mengin-Lecreulx, and A. Bouhss.** 2008. Active site mapping of MraY, a member of the polyprenyl-phosphate N-acetylhexosamine 1-phosphate transferase superfamily, catalyzing the first membrane step of peptidoglycan biosynthesis. Biochemistry **47:**8919-28.

8. **Al-Dabbagh, B., D. Mengin-Lecreulx, and A. Bouhss.** 2008. Purification and characterization of the bacterial UDP-GlcNAc:undecaprenyl-phosphate GlcNAc-1-phosphate transferase WecA. J Bacteriol **190:**7141-6.

9. **Alaimo, C., I. Catrein, L. Morf, C. L. Marolda, N. Callewaert, M. A. Valvano, M. F. Feldman, and M. Aebi.** 2006. Two distinct but interchangeable mechanisms for flipping of lipid-linked oligosaccharides. Embo J **25:**967-76.

10. **Aldridge, P., F. Bernhard, P. Bugert, D. L. Coplin, and K. Geider.** 1998. Characterization of a gene locus from *Erwinia amylovora* with regulatory functions in exopolysaccharide synthesis of *Erwinia* spp. Can J Microbiol **44:**657-66.

11. **AlonsoDeVelasco, E., A. F. Verheul, J. Verhoef, and H. Snippe.** 1995. *Streptococcus pneumoniae*: virulence factors, pathogenesis, and vaccines. Microbiol Rev **59:**591-603.

12. **Amer, A. O., and M. A. Valvano.** 2001. 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:**3015-25.

13. **Amer, A. O., and M. A. Valvano.** 2002. 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:**571-82.

14. **Anderson, M. S., H. G. Bull, S. M. Galloway, T. M. Kelly, S. Mohan, K. Radika, and C. R. Raetz.** 1993. UDP-N-acetylglucosamine acyltransferase of *Escherichia coli*. The first step of endotoxin biosynthesis is thermodynamically unfavorable. J Biol Chem **268:**19858-65.

15. **Babinski, K. J., S. J. Kanjilal, and C. R. Raetz.** 2002. Accumulation of the lipid A precursor UDP-2,3-diacylglucosamine in an *Escherichia coli* mutant lacking the lpxH gene. J Biol Chem **277:**25947-56.

16. **Babinski, K. J., A. A. Ribeiro, and C. R. Raetz.** 2002. The *Escherichia coli* gene encoding the UDP-2,3-diacylglucosamine pyrophosphatase of lipid A biosynthesis. J Biol Chem **277:**25937-46.

17. **Barr, K., J. Klena, and P. D. Rick.** 1999. The modality of enterobacterial common antigen polysaccharide chain lengths is regulated by o349 of the wec gene cluster of *Escherichia coli* K-12. J Bacteriol **181:**6564-8.

18. **Barr, K., and P. D. Rick.** 1987. Biosynthesis of enterobacterial common antigen in *Escherichia coli*. In vitro synthesis of lipid-linked intermediates. J Biol Chem **262:**7142-50.

19. **Bentley, S. D., D. M. Aanensen, A. Mavroidi, D. Saunders, E. Rabbinowitsch, M. Collins, K. Donohoe, D. Harris, L. Murphy, M. A. Quail, G. Samuel, I. C. Skovsted, M. S. Kaltoft, B. Barrell, P. R. Reeves, J. Parkhill, and B. G. Spratt.** 2006. Genetic analysis of the capsular biosynthetic locus from all 90 pneumococcal serotypes. PLoS Genet **2:**e31.

20. **Beveridge, T. J.** 1999. Structures of gram-negative cell walls and their derived membrane vesicles. J Bacteriol **181:**4725-33.

21. **Bhavsar, A. P., R. Truant, and E. D. Brown.** 2005. The TagB protein in *Bacillus subtilis* 168 is an intracellular peripheral membrane protein that can incorporate glycerol phosphate onto a membrane-bound acceptor in vitro. J Biol Chem **280:**36691-700.

22. **Bos, M. P., V. Robert, and J. Tommassen.** 2007. Biogenesis of the gram-negative bacterial outer membrane. Annu Rev Microbiol **61:**191-214.

23. **Bouhss, A., M. Crouvoisier, D. Blanot, and D. Mengin-Lecreulx.** 2004. Purification and characterization of the bacterial MraY translocase catalyzing the first membrane step of peptidoglycan biosynthesis. J Biol Chem **279:**29974-80.

24. **Bouhss, A., D. Mengin-Lecreulx, D. Le Beller, and J. Van Heijenoort.** 1999. Topological analysis of the MraY protein catalysing the first membrane step of peptidoglycan synthesis. Mol Microbiol **34:**576-85.

25. **Boyd, D. A., D. G. Cvitkovitch, A. S. Bleiweis, M. Y. Kiriukhin, D. V. Debabov, F. C. Neuhaus, and I. R. Hamilton.** 2000. Defects in D-alanyl-lipoteichoic acid synthesis in *Streptococcus mutans* results in acid sensitivity. J Bacteriol **182:**6055-65.

26. **Brooke, J. S.** 1996. Ph.D. Dissertation. Characterization of a phosphoheptose isomerase involved in inner core lipopolysaccharide biosynthesis. Ph.D. Dissertation. University of Western Ontario, London, Ontario.

27. **Brooke, J. S., and M. A. Valvano.** 1996. Molecular cloning of the *Haemophilus influenzae gmhA* (*lpcA*) gene encoding a phosphoheptose isomerase required for lipooligosaccharide biosynthesis. J. Bacteriol. **178:**3339-3341.

28. **Burrows, L. L., D. Chow, and J. S. Lam.** 1997. *Pseudomonas aeruginosa* B-band O-antigen chain length is modulated by Wzz (Ro1). J Bacteriol **179:**1482-9.

29. **Cartee, R. T., W. T. Forsee, M. H. Bender, K. D. Ambrose, and J. Yother.** 2005. CpsE from type 2 *Streptococcus pneumoniae* catalyzes the reversible addition of glucose-1-phosphate to a polyprenyl phosphate acceptor, initiating type 2 capsule repeat unit formation. J Bacteriol **187:**7425-33.

30. **Cartee, R. T., W. T. Forsee, J. S. Schutzbach, and J. Yother.** 2000. Mechanism of type 3 capsular polysaccharide synthesis in *Streptococcus pneumoniae*. J Biol Chem **275:**3907-14.

31. **Carter, J. A., J. C. Jimenez, M. Zaldivar, S. A. Alvarez, C. L. Marolda, M. A. Valvano, and I. Contreras.** 2009. The cellular level of O-antigen polymerase Wzy determines chain length regulation by WzzB and WzzpHS-2 in *Shigella flexneri* 2a. Microbiology **155:**3260-9.

32. **Castric, P.** 1995. pilO, a gene required for glycosylation of *Pseudomonas aeruginosa* 1244 pilin. Microbiology **141 (Pt 5):**1247-54.

33. **Chen, Y., J. Dai, J. G. Morris, Jr., and J. A. Johnson.** 2010. Genetic analysis of the capsule polysaccharide (K antigen) and exopolysaccharide genes in pandemic *Vibrio parahaemolyticus* O3:K6. BMC Microbiol **10:**274.

34.   **Chng, S. S., L. S. Gronenberg, and D. Kahne.** 2010. Proteins required for lipopolysaccharide assembly in *Escherichia coli* form a transenvelope complex. Biochemistry **49:**4565-7.

35.   **Chng, S. S., N. Ruiz, G. Chimalakonda, T. J. Silhavy, and D. Kahne.** 2010. Characterization of the two-protein complex in *Escherichia coli* responsible for lipopolysaccharide assembly at the outer membrane. Proc Natl Acad Sci U S A **107:**5363-8.

36.   **Cieslewicz, M. J., D. L. Kasper, Y. Wang, and M. R. Wessels.** 2001. Functional analysis in type Ia group B *Streptococcus* of a cluster of genes involved in extracellular polysaccharide production by diverse species of streptococci. J Biol Chem **276:**139-46.

37.   **Clarke, B. R., L. Cuthbertson, and C. Whitfield.** 2004. Nonreducing terminal modifications determine the chain length of polymannose O antigens of *Escherichia coli* and couple chain termination to polymer export via an ATP-binding cassette transporter. J Biol Chem **279:**35709-18.

38.   **Clementz, T., J. J. Bednarski, and C. R. Raetz.** 1996. Function of the htrB high temperature requirement gene of *Escherchia coli* in the acylation of lipid A: HtrB catalyzed incorporation of laurate. J Biol Chem **271:**12095-102.

39.   **Coleman, J., and C. R. Raetz.** 1988. First committed step of lipid A biosynthesis in *Escherichia coli*: sequence of the lpxA gene. J Bacteriol **170:**1268-74.

40.   **Cozzone, A. J., C. Grangeasse, P. Doublet, and B. Duclos.** 2004. Protein phosphorylation on tyrosine in bacteria. Arch Microbiol **181:**171-81.

41.   **Crowell, D. N., M. S. Anderson, and C. R. Raetz.** 1986. Molecular cloning of the genes for lipid A disaccharide synthase and UDP-N-acetylglucosamine acyltransferase in *Escherichia coli*. J Bacteriol **168:**152-9.

42.   **Crowell, D. N., W. S. Reznikoff, and C. R. Raetz.** 1987. Nucleotide sequence of the *Escherichia coli* gene for lipid A disaccharide synthase. J Bacteriol **169:**5727-34.

43.   **Curd, H., D. Liu, and P. R. Reeves.** 1998. Relationships among the O-antigen gene clusters of *Salmonella enterica* groups B, D1, D2, and D3. J Bacteriol **180:**1002-7.

44.   **Cuthbertson, L., M. S. Kimber, and C. Whitfield.** 2007. Substrate binding by a bacterial ABC transporter involved in polysaccharide export. Proc Natl Acad Sci U S A **104:**19529-34.

45.   **Cuthbertson, L., J. Powers, and C. Whitfield.** 2005. The C-terminal domain of the nucleotide-binding domain protein Wzt determines substrate specificity in the

ATP-binding cassette transporter for the lipopolysaccharide O-antigens in *Escherichia coli* serotypes O8 and O9a. J Biol Chem **280:**30310-9.

46. **D'Elia, M. A., K. E. Millar, T. J. Beveridge, and E. D. Brown.** 2006. Wall teichoic acid polymers are dispensable for cell viability in *Bacillus subtilis*. J Bacteriol **188:**8313-6.

47. **Danese, P. N., L. A. Pratt, and R. Kolter.** 2000. Exopolysaccharide production is required for development of *Escherichia coli* K-12 biofilm architecture. J Bacteriol **182:**3593-6.

48. **DeAngelis, P. L.** 1999. Hyaluronan synthases: fascinating glycosyltransferases from vertebrates, bacterial pathogens, and algal viruses. Cell Mol Life Sci **56:**670-82.

49. **DeAngelis, P. L.** 2002. Microbial glycosaminoglycan glycosyltransferases. Glycobiology **12:**9R-16R.

50. **DeAngelis, P. L., N. Yang, and P. H. Weigel.** 1994. The *Streptococcus pyogenes* hyaluronan synthase: sequence comparison and conservation among various group A strains. Biochem Biophys Res Commun **199:**1-10.

51. **Doerrler, W. T., H. S. Gibbons, and C. R. Raetz.** 2004. MsbA-dependent translocation of lipids across the inner membrane of *Escherichia coli*. J Biol Chem **279:**45102-9.

52. **Doerrler, W. T., and C. R. Raetz.** 2002. ATPase activity of the MsbA lipid flippase of *Escherichia coli*. J Biol Chem **277:**36697-705.

53. **Doerrler, W. T., M. C. Reedy, and C. R. Raetz.** 2001. An *Escherichia coli* mutant defective in lipid export. J Biol Chem **276:**11461-4.

54. **Domka, J., J. Lee, T. Bansal, and T. K. Wood.** 2007. Temporal gene-expression in *Escherichia coli* K-12 biofilms. Environ Microbiol **9:**332-46.

55. **Dong, C., K. Beis, J. Nesper, A. L. Brunkan-Lamontagne, B. R. Clarke, C. Whitfield, and J. H. Naismith.** 2006. Wza the translocon for *E. coli* capsular polysaccharides defines a new class of membrane protein. Nature **444:**226-9.

56. **Dougherty, B. A., and I. van de Rijn.** 1994. Molecular characterization of hasA from an operon required for hyaluronic acid synthesis in group A streptococci. J Biol Chem **269:**169-75.

57. **Drummelsmith, J., and C. Whitfield.** 1999. Gene products required for surface expression of the capsular form of the group 1 K antigen in *Escherichia coli* (O9a:K30). Mol Microbiol **31:**1321-32.

58.   **Drummelsmith, J., and C. Whitfield.** 2000. Translocation of group 1 capsular polysaccharide to the surface of *Escherichia coli* requires a multimeric complex in the outer membrane. EMBO J **19:**57-66.

59.   **El Ghachi, M., A. Bouhss, D. Blanot, and D. Mengin-Lecreulx.** 2004. The bacA gene of *Escherichia coli* encodes an undecaprenyl pyrophosphate phosphatase activity. J Biol Chem **279:**30106-13.

60.   **El Ghachi, M., A. Derbise, A. Bouhss, and D. Mengin-Lecreulx.** 2005. Identification of multiple genes encoding membrane proteins with undecaprenyl pyrophosphate phosphatase (UppP) activity in *Escherichia coli*. J Biol Chem **280:**18689-95.

61.   **Faridmoayer, A., M. A. Fentabil, M. F. Haurat, W. Yi, R. Woodward, P. G. Wang, and M. F. Feldman.** 2008. Extreme substrate promiscuity of the *Neisseria* oligosaccharyl transferase involved in protein O-glycosylation. J Biol Chem **283:**34596-604.

62.   **Feldman, M. F., C. L. Marolda, M. A. Monteiro, M. B. Perry, A. J. Parodi, and M. A. Valvano.** 1999. The activity of a putative polyisoprenol-linked sugar translocase (Wzx) involved in *Escherichia coli* O antigen assembly is independent of the chemical structure of the O repeat. J Biol Chem **274:**35129-38.

63.   **Fletcher, C. M., M. J. Coyne, D. L. Bentley, O. F. Villa, and L. E. Comstock.** 2007. Phase-variable expression of a family of glycoproteins imparts a dynamic surface to a symbiont in its human intestinal ecosystem. Proc Natl Acad Sci U S A **104:**2413-8.

64.   **Forsee, W. T., R. T. Cartee, and J. Yother.** 2000. Biosynthesis of type 3 capsular polysaccharide in *Streptococcus pneumoniae*. Enzymatic chain release by an abortive translocation process. J Biol Chem **275:**25972-8.

65.   **Fratamico, P. M., S. Bhaduri, and R. L. Buchanan.** 1993. Studies on *Escherichia coli* serotype O157:H7 strains containing a 60-MDa plasmid and on 60-MDa plasmid-cured derivatives. J Med Microbiol **39:**371-81.

66.   **Frirdich, E., and C. Whitfield.** 2005. Lipopolysaccharide inner core oligosaccharide structure and outer membrane stability in human pathogens belonging to the Enterobacteriaceae. J Endotoxin Res **11:**133-44.

67.   **Furlong, S. E., and M. A. Valvano.** Unpublished. The highly conserved VFMGD motif of the sugar-phosphate transferase WecA may contribute to the structure of the enzyme's active site.

68.   **Galanos, C., O. Luderitz, E. T. Rietschel, O. Westphal, H. Brade, L. Brade, M. Freudenberg, U. Schade, M. Imoto, H. Yoshimura, and et al.** 1985. Synthetic and natural *Escherichia coli* free lipid A express identical endotoxic activities. Eur J Biochem **148:**1-5.

69. **Galloway, S. M., and C. R. Raetz.** 1990. A mutant of *Escherichia coli* defective in the first step of endotoxin biosynthesis. J Biol Chem **265:**6394-402.

70. **Garcia-Calderon, C. B., M. Garcia-Quintanilla, J. Casadesus, and F. Ramos-Morales.** 2005. Virulence attenuation in *Salmonella enterica* rcsC mutants with constitutive activation of the Rcs system. Microbiology **151:**579-88.

71. **Garegg, P. J., B. Lindberg, T. Onn, and I. W. Sutherland.** 1971. Comparative structural studies on the M-antigen from *Salmonella typhimurium. Escherichia coli* and *Aerobacter cloacae*. Acta Chem Scand **25:**2103-8.

72. **Garrett, T. A., J. L. Kadrmas, and C. R. Raetz.** 1997. Identification of the gene encoding the *Escherichia coli* lipid A 4'-kinase. Facile phosphorylation of endotoxin analogs with recombinant LpxK. J Biol Chem **272:**21855-64.

73. **Garrett, T. A., N. L. Que, and C. R. Raetz.** 1998. Accumulation of a lipid A precursor lacking the 4'-phosphate following inactivation of the *Escherichia coli* lpxK gene. J Biol Chem **273:**12457-65.

74. **Ghuysen, J. M.** 1977. The concept of the penicillin target from 1965 until today. The thirteenth marjory stephenson memorial lecture. J Gen Microbiol **101:**13-33.

75. **Glover, K. J., E. Weerapana, M. M. Chen, and B. Imperiali.** 2006. Direct biochemical evidence for the utilization of UDP-bacillosamine by PglC, an essential glycosyl-1-phosphate transferase in the *Campylobacter jejuni* N-linked glycosylation pathway. Biochemistry **45:**5343-50.

76. **Goebel, W. F.** 1963. Colanic acid. Proc Natl Acad Sci U S A **49:**464-71.

77. **Grant, W. D., I. W. Sutherland, and J. F. Wilkinson.** 1969. Exopolysaccharide colanic acid and its occurrence in the Enterobacteriaceae. J Bacteriol **100:**1187-93.

78. **Gronow, S., W. Brabetz, and H. Brade.** 2000. Comparative functional characterization in vitro of heptosyltransferase I (WaaC) and II (WaaF) from *Escherichia coli*. Eur J Biochem **267:**6602-11.

79. **Gross, M., S. E. Cramton, F. Gotz, and A. Peschel.** 2001. Key role of teichoic acid net charge in *Staphylococcus aureus* colonization of artificial surfaces. Infect Immun **69:**3423-6.

80. **Hamill, O. P., and B. Martinac.** 2001. Molecular basis of mechanotransduction in living cells. Physiol Rev **81:**685-740.

81. **Hardy, G. G., R. C. Allen, E. Toh, M. Long, P. J. Brown, J. L. Cole-Tobian, and Y. V. Brun.** 2010. A localized multimeric anchor attaches the *Caulobacter* holdfast to the cell pole. Mol Microbiol **76:**409-27.

82. **Hartley, M. D., M. J. Morrison, F. E. Aas, B. Borud, M. Koomey, and B. Imperiali.** 2011. Biochemical Characterization of the O-Linked Glycosylation Pathway in *Neisseria gonorrhoeae* Responsible for Biosynthesis of Protein Glycans Containing N,N'-Diacetylbacillosamine. Biochemistry.

83. **Heifetz, A., R. W. Keenan, and A. D. Elbein.** 1979. Mechanism of action of tunicamycin on the UDP-GlcNAc:dolichyl-phosphate Glc-NAc-1-phosphate transferase. Biochemistry **18:**2186-92.

84. **Heinrichs, D. E., M. A. Monteiro, M. B. Perry, and C. Whitfield.** 1998. The assembly system for the lipopolysaccharide R2 core-type of *Escherichia coli* is a hybrid of those found in *Escherichia coli* K-12 and *Salmonella enterica*. Structure and function of the R2 WaaK and WaaL homologs. J Biol Chem **273:**8849-59.

85. **Heinrichs, D. E., J. A. Yethon, P. A. Amor, and C. Whitfield.** 1998. The assembly system for the outer core portion of R1- and R4-type lipopolysaccharides of *Escherichia coli*. The R1 core-specific beta-glucosyltransferase provides a novel attachment site for O-polysaccharides. J Biol Chem **273:**29497-505.

86. **Heller, K. J.** 1992. Molecular interaction between bacteriophage and the gram-negative cell envelope. Archives of Microbiology **158:**235-248.

87. **Honeyman, A. L., and G. C. Stewart.** 1989. The nucleotide sequence of the rodC operon of *Bacillus subtilis*. Mol Microbiol **3:**1257-68.

88. **Hug, I., M. R. Couturier, M. M. Rooker, D. E. Taylor, M. Stein, and M. F. Feldman.** 2010. Helicobacter pylori lipopolysaccharide is synthesized via a novel pathway with an evolutionary connection to protein N-glycosylation. PLoS Pathog **6:**e1000819.

89. **Hwang, P. M., R. E. Bishop, and L. E. Kay.** 2004. The integral membrane enzyme PagP alternates between two dynamically distinct states. Proc Natl Acad Sci U S A **101:**9618-23.

90. **Ielpi, L., R. O. Couso, and M. A. Dankert.** 1993. Sequential assembly and polymerization of the polyprenol-linked pentasaccharide repeating unit of the xanthan polysaccharide in *Xanthomonas campestris*. J Bacteriol **175:**2490-500.

91. **Ishii, K., H. Sagami, and K. Ogura.** 1986. A novel prenyltransferase from *Paracoccus denitrificans*. Biochem J **233:**773-7.

92. **Janda, J. M., S. L. Abbott, S. Khashe, G. H. Kellogg, and T. Shimada.** 1996. Further studies on biochemical characteristics and serologic properties of the genus *Aeromonas*. J Clin Microbiol **34:**1930-3.

93. **Jimenez, N., R. Canals, M. T. Salo, S. Vilches, S. Merino, and J. M. Tomas.** 2008. The *Aeromonas hydrophila* wb*O34 gene cluster: genetics and temperature regulation. J Bacteriol **190:**4198-209.

94. **Kajimura, J., A. Rahman, and P. D. Rick.** 2005. Assembly of cyclic enterobacterial common antigen in *Escherichia coli* K-12. J Bacteriol **187:**6917-27.

95. **Karlyshev, A. V., J. M. Ketley, and B. W. Wren.** 2005. The *Campylobacter jejuni* glycome. FEMS Microbiol Rev **29:**377-90.

96. **Katzen, F., D. U. Ferreiro, C. G. Oddo, M. V. Ielmini, A. Becker, A. Puhler, and L. Ielpi.** 1998. *Xanthomonas campestris* pv. campestris gum mutants: effects on xanthan biosynthesis and plant virulence. J Bacteriol **180:**1607-17.

97. **Kaur, D., P. J. Brennan, and D. C. Crick.** 2004. Decaprenyl diphosphate synthesis in *Mycobacterium tuberculosis*. J Bacteriol **186:**7564-70.

98. **Keenleyside, W. J., A. J. Clarke, and C. Whitfield.** 2001. Identification of residues involved in catalytic activity of the inverting glycosyl transferase WbbE from *Salmonella enterica* serovar borreze. J Bacteriol **183:**77-85.

99. **Keenleyside, W. J., M. Perry, L. Maclean, C. Poppe, and C. Whitfield.** 1994. A plasmid-encoded rfbO:54 gene cluster is required for biosynthesis of the O:54 antigen in *Salmonella enterica* serovar Borreze. Mol Microbiol **11:**437-48.

100. **Keenleyside, W. J., and C. Whitfield.** 1995. Lateral transfer of rfb genes: a mobilizable ColE1-type plasmid carries the rfbO:54 (O:54 antigen biosynthesis) gene cluster from *Salmonella enterica* serovar Borreze. J Bacteriol **177:**5247-53.

101. **Keenleyside, W. J., and C. Whitfield.** 1996. A novel pathway for O-polysaccharide biosynthesis in *Salmonella enterica* serovar Borreze. J Biol Chem **271:**28581-92.

102. **Kelly, J., H. Jarrell, L. Millar, L. Tessier, L. M. Fiori, P. C. Lau, B. Allan, and C. M. Szymanski.** 2006. Biosynthesis of the N-linked glycan in *Campylobacter jejuni* and addition onto protein through block transfer. J Bacteriol **188:**2427-34.

103. **Kelly, T. M., S. A. Stachula, C. R. Raetz, and M. S. Anderson.** 1993. The firA gene of *Escherichia coli* encodes UDP-3-O-(R-3-hydroxymyristoyl)-glucosamine N-acyltransferase. The third step of endotoxin biosynthesis. J Biol Chem **268:**19866-74.

104. **Kido, N., V. I. Torgov, T. Sugiyama, K. Uchiya, H. Sugihara, T. Komatsu, N. Kato, and K. Jann.** 1995. Expression of the O9 polysaccharide of *Escherichia coli:* sequencing of the *E. coli* O9 rfb gene cluster, characterization of mannosyl

transferases, and evidence for an ATP-binding cassette transport system. J Bacteriol **177:**2178-87.

105.   **Kneidinger, B., M. Graninger, M. Puchberger, P. Kosma, and P. Messner.** 2001. Biosynthesis of nucleotide-activated D-*glycero*-D-*manno*-heptose. J. Biol. Chem. **276:**20935-20944.

106.   **Kneidinger, B., C. L. Marolda, M. Graninger, A. Zamyatina, F. McArthur, P. Kosma, M. A. Valvano, and P. Messner.** 2002. Biosynthesis pathway of ADP-D-*glycero*-L-*manno*-heptose in *Escherichia coli*. J. Bacteriol. **184:**363-369.

107.   **Knirel, Y. A., A. S. Shashkov, S. N. Senchenkova, S. Merino, and J. M. Tomas.** 2002. Structure of the O-polysaccharide of *Aeromonas hydrophila* O:34; a case of random O-acetylation of 6-deoxy-L-talose. Carbohydr Res **337:**1381-6.

108.   **Kontrohr, T., and B. Kocsis.** 1981. Isolation of adenosine 5'-diphosphate-D-glycero-D-mannoheptose. An intermediate in lipopolysaccharide biosynthesis of *Shigella sonnei*. J. Biol. Chem. **256:**7715-7718.

109.   **Kovacs, M., A. Halfmann, I. Fedtke, M. Heintz, A. Peschel, W. Vollmer, R. Hakenbeck, and R. Bruckner.** 2006. A functional dlt operon, encoding proteins required for incorporation of d-alanine in teichoic acids in gram-positive bacteria, confers resistance to cationic antimicrobial peptides in *Streptococcus pneumoniae*. J Bacteriol **188:**5797-805.

110.   **Kristian, S. A., V. Datta, C. Weidenmaier, R. Kansal, I. Fedtke, A. Peschel, R. L. Gallo, and V. Nizet.** 2005. D-alanylation of teichoic acids promotes group a streptococcus antimicrobial peptide resistance, neutrophil survival, and epithelial cell invasion. J Bacteriol **187:**6719-25.

111.   **Lam, J. S., L. L. Graham, J. Lightfoot, T. Dasgupta, and T. J. Beveridge.** 1992. Ultrastructural examination of the lipopolysaccharides of *Pseudomonas aeruginosa* strains and their isogenic rough mutants by freeze-substitution. J Bacteriol **174:**7159-67.

112.   **Lazarevic, V., and D. Karamata.** 1995. The tagGH operon of *Bacillus subtilis* 168 encodes a two-component ABC transporter involved in the metabolism of two wall teichoic acids. Mol Microbiol **16:**345-55.

113.   **Lee, S. W., M. Sabet, H. S. Um, J. Yang, H. C. Kim, and W. Zhu.** 2006. Identification and characterization of the genes encoding a unique surface (S-) layer of *Tannerella forsythia*. Gene **371:**102-11.

114.   **Lehrer, J., K. A. Vigeant, L. D. Tatar, and M. A. Valvano.** 2007. Functional characterization and membrane topology of *Escherichia coli* WecA, a sugar-phosphate transferase initiating the biosynthesis of enterobacterial common antigen and O-antigen lipopolysaccharide. J Bacteriol **189:**2618-28.

115. **Leigh, J. A., and D. L. Coplin.** 1992. Exopolysaccharides in plant-bacterial interactions. Annu Rev Microbiol **46:**307-46.

116. **Li, G., C. S. Smith, Y. V. Brun, and J. X. Tang.** 2005. The elastic properties of the *Caulobacter Crescentus* adhesive holdfast are dependent on oligomers of N-acetylglucosamine. J Bacteriol **187:**257-65.

117. **Liu, B., A. V. Perepelov, D. Li, S. N. Senchenkova, Y. Han, A. S. Shashkov, L. Feng, Y. A. Knirel, and L. Wang.** 2010. Structure of the O-antigen of *Salmonella* O66 and the genetic basis for similarity and differences between the closely related O-antigens of *Escherichia coli* O166 and *Salmonella* O66. Microbiology **156:**1642-9.

118. **Liu, D., R. A. Cole, and P. R. Reeves.** 1996. An O-antigen processing function for Wzx (RfbX): a promising candidate for O-unit flippase. J Bacteriol **178:**2102-7.

119. **Lloyd, A. J., P. E. Brandish, A. M. Gilbey, and T. D. Bugg.** 2004. Phospho-N-acetyl-muramyl-pentapeptide translocase from *Escherichia coli*: catalytic role of conserved aspartic acid residues. J Bacteriol **186:**1747-57.

120. **Madigan, M. T., J. M. Martinko, J. Parker, and T. D. Brock.** 2003. Brock biology of microorganisms, 10th ed. Prentice Hall/Pearson Education, Upper Saddle River, NJ.

121. **Mahapatra, S., T. Yagi, J. T. Belisle, B. J. Espinosa, P. J. Hill, M. R. McNeil, P. J. Brennan, and D. C. Crick.** 2005. Mycobacterial lipid II is composed of a complex mixture of modified muramyl and peptide moieties linked to decaprenyl phosphate. J Bacteriol **187:**2747-57.

122. **Majdalani, N., and S. Gottesman.** 2005. The Rcs phosphorelay: a complex signal transduction system. Annu Rev Microbiol **59:**379-405.

123. **Mamat, U., T. C. Meredith, P. Aggarwal, A. Kuhl, P. Kirchhoff, B. Lindner, A. Hanuszkiewicz, J. Sun, O. Holst, and R. W. Woodard.** 2008. Single amino acid substitutions in either YhjD or MsbA confer viability to 3-deoxy-d-manno-oct-2-ulosonic acid-depleted *Escherichia coli*. Mol Microbiol **67:**633-48.

124. **Marolda, C. L., J. Vicarioli, and M. A. Valvano.** 2004. Wzx proteins involved in biosynthesis of O antigen function in association with the first sugar of the O-specific lipopolysaccharide subunit. Microbiology **150:**4095-105.

125. **Mauel, C., M. Young, and D. Karamata.** 1991. Genes concerned with synthesis of poly(glycerol phosphate), the essential teichoic acid in *Bacillus subtilis* strain 168, are organized in two divergent transcription units. J Gen Microbiol **137:**929-41.

126. **May, T. B., D. Shinabarger, R. Maharaj, J. Kato, L. Chu, J. D. DeVault, S. Roychoudhury, N. A. Zielinski, A. Berry, R. K. Rothmel, and et al.** 1991. Alginate synthesis by *Pseudomonas aeruginosa*: a key pathogenic factor in chronic pulmonary infections of cystic fibrosis patients. Clin Microbiol Rev **4:**191-206.

127. **Meredith, T. C., P. Aggarwal, U. Mamat, B. Lindner, and R. W. Woodard.** 2006. Redefining the requisite lipopolysaccharide structure in *Escherichia coli*. ACS Chem Biol **1:**33-42.

128. **Meredith, T. C., U. Mamat, Z. Kaczynski, B. Lindner, O. Holst, and R. W. Woodard.** 2007. Modification of lipopolysaccharide with colanic acid (M-antigen) repeats in *Escherichia coli*. J Biol Chem **282:**7790-8.

129. **Merino, S., N. Jimenez, R. Molero, L. Bouamama, M. Regue, and J. M. Tomas.** 2011. A UDP-HexNAc:Polyprenol-P GalNAc-1-P Transferase (WecP) Representing a New Subgroup of the Enzyme Family. J Bacteriol **193:**1943-52.

130. **Merker, R. I., and J. Smit.** 1988. Characterization of the adhesive holdfast of marine and freshwater caulobacters. Appl Environ Microbiol **54:**2078-85.

131. **Messner, P., K. Steiner, K. Zarschler, and C. Schaffer.** 2008. S-layer nanoglycobiology of bacteria. Carbohydr Res **343:**1934-51.

132. **Mitchell, D., and J. Smit.** 1990. Identification of genes affecting production of the adhesion organelle of *Caulobacter crescentus* CB2. J Bacteriol **172:**5425-31.

133. **Murray, G. L., S. R. Attridge, and R. Morona.** 2006. Altering the length of the lipopolysaccharide O antigen has an impact on the interaction of *Salmonella enterica* serovar Typhimurium with macrophages and complement. J Bacteriol **188:**2735-9.

134. **Murray, G. L., S. R. Attridge, and R. Morona.** 2005. Inducible serum resistance in *Salmonella* typhimurium is dependent on wzz(fepE)-regulated very long O antigen chains. Microbes Infect **7:**1296-304.

135. **Murray, G. L., S. R. Attridge, and R. Morona.** 2003. Regulation of *Salmonella* typhimurium lipopolysaccharide O antigen chain length is required for virulence; identification of FepE as a second Wzz. Mol Microbiol **47:**1395-406.

136. **Nanninga, N.** 1998. Morphogenesis of Escherichia coli. Microbiol Mol Biol Rev **62:**110-29.

137. **Narita, S., and H. Tokuda.** 2009. Biochemical characterization of an ABC transporter LptBFGC complex required for the outer membrane sorting of lipopolysaccharides. FEBS Lett **583:**2160-4.

138. **Nikaido, H.** 2003. Molecular basis of bacterial outer membrane permeability revisited. Microbiol Mol Biol Rev **67:**593-656.

139. **Nikaido, H.** 1962. Studies on the biosynthesis of cell-wall polysaccharide in mutant strains of *Salmonella*. I. Proc Natl Acad Sci U S A **48:**1337-41.

140. **Nikaido, H.** 1962. Studies on the biosynthesis of cell wall polysaccharide in mutant strains of *Salmonella*. II. Proc Natl Acad Sci U S A **48:**1542-8.

141. **Nikaido, H., and M. Vaara.** 1985. Molecular basis of bacterial outer membrane permeability. Microbiol Rev **49:**1-32.

142. **Nothaft, H., and C. M. Szymanski.** 2010. Protein glycosylation in bacteria: sweeter than ever. Nat Rev Microbiol **8:**765-78.

143. **Nouaille, S., J. Commissaire, J. J. Gratadoux, P. Ravn, A. Bolotin, A. Gruss, Y. Le Loir, and P. Langella.** 2004. Influence of lipoteichoic acid D-alanylation on protein secretion in *Lactococcus lactis* as revealed by random mutagenesis. Appl Environ Microbiol **70:**1600-7.

144. **Novotny, R., C. Schaffer, J. Strauss, and P. Messner.** 2004. S-layer glycan-specific loci on the chromosome of *Geobacillus stearothermophilus* NRS 2004/3a and dTDP-L-rhamnose biosynthesis potential of *G. stearothermophilus* strains. Microbiology **150:**953-65.

145. **Orskov, I., F. Orskov, B. Jann, and K. Jann.** 1977. Serology, chemistry, and genetics of O and K antigens of *Escherichia coli*. Bacteriol Rev **41:**667-710.

146. **Osborn, M. J., S. M. Rosen, L. Rothfield, and B. L. Horecker.** 1962. Biosynthesis of bacterial lipopolysaccharide. I. Enzymatic incorporation of galactose in a mutant strain of *Salmonella*. Proc Natl Acad Sci U S A **48:**1831-8.

147. **Osborn, M. J., S. M. Rosen, L. Rothfield, and B. L. Horecker.** 1962. Galactose Incorporation into Cell-Wall Lipopolysaccharide in Mutant Strain of *Salmonella* typhimurium. Science **136:**328.

148. **Pelosi, L., M. Boumedienne, N. Saksouk, J. Geiselmann, and R. A. Geremia.** 2005. The glucosyl-1-phosphate transferase WchA (Cap8E) primes the capsular polysaccharide repeat unit biosynthesis of *Streptococcus pneumoniae* serotype 8. Biochem Biophys Res Commun **327:**857-65.

149. **Perez, J. M., M. A. McGarry, C. L. Marolda, and M. A. Valvano.** 2008. Functional analysis of the large periplasmic loop of the *Escherichia coli* K-12 WaaL O-antigen ligase. Mol Microbiol **70:**1424-40.

150. **Poindexter, J. S.** 1964. Biological Properties and Classification of the Caulobacter Group. Bacteriol Rev **28:**231-95.

151. **Pollack, J. H., and F. C. Neuhaus.** 1994. Changes in wall teichoic acid during the rod-sphere transition of Bacillus subtilis 168. J Bacteriol **176:**7252-9.

152. **Poltorak, A., X. He, I. Smirnova, M. Y. Liu, C. Van Huffel, X. Du, D. Birdwell, E. Alejos, M. Silva, C. Galanos, M. Freudenberg, P. Ricciardi-Castagnoli, B. Layton, and B. Beutler.** 1998. Defective LPS signaling in C3H/HeJ and C57BL/10ScCr mice: mutations in Tlr4 gene. Science **282:**2085-8.

153. **Pooley, H. M., F. X. Abellan, and D. Karamata.** 1991. A conditional-lethal mutant of bacillus subtilis 168 with a thermosensitive glycerol-3-phosphate cytidylyltransferase, an enzyme specific for the synthesis of the major wall teichoic acid. J Gen Microbiol **137:**921-8.

154. **Power, P. M., K. L. Seib, and M. P. Jennings.** 2006. Pilin glycosylation in *Neisseria meningitidis* occurs by a similar pathway to wzy-dependent O-antigen biosynthesis in *Escherichia coli*. Biochem Biophys Res Commun **347:**904-8.

155. **Pugsley, A. P., and O. Possot.** 1993. The general secretory pathway of *Klebsiella oxytoca*: no evidence for relocalization or assembly of pilin-like PulG protein into a multiprotein complex. Mol Microbiol **10:**665-74.

156. **Raetz, C. R., T. A. Garrett, C. M. Reynolds, W. A. Shaw, J. D. Moore, D. C. Smith, Jr., A. A. Ribeiro, R. C. Murphy, R. J. Ulevitch, C. Fearns, D. Reichart, C. K. Glass, C. Benner, S. Subramaniam, R. Harkewicz, R. C. Bowers-Gentry, M. W. Buczynski, J. A. Cooper, R. A. Deems, and E. A. Dennis.** 2006. Kdo2-Lipid A of *Escherichia coli*, a defined endotoxin that activates macrophages via TLR-4. J Lipid Res **47:**1097-111.

157. **Raetz, C. R., and C. Whitfield.** 2002. Lipopolysaccharide endotoxins. Annu Rev Biochem **71:**635-700.

158. **Raetz, C. R. H., and C. Whitfield.** 2002. Lipopolysaccharide endotoxins. Annu. Rev. Biochem. **71:**635-700.

159. **Rahn, A., K. Beis, J. H. Naismith, and C. Whitfield.** 2003. A novel outer membrane protein, Wzi, is involved in surface assembly of the *Escherichia coli* K30 group 1 capsule. J Bacteriol **185:**5882-90.

160. **Ray, B. L., and C. R. Raetz.** 1987. The biosynthesis of gram-negative endotoxin. A novel kinase in *Escherichia coli* membranes that incorporates the 4'-phosphate of lipid A. J Biol Chem **262:**1122-8.

161. **Reynolds, C. M., and C. R. Raetz.** 2009. Replacement of lipopolysaccharide with free lipid A molecules in *Escherichia coli* mutants lacking all core sugars. Biochemistry **48:**9627-40.

162. **Ristl, R., K. Steiner, K. Zarschler, S. Zayni, P. Messner, and C. Schaffer.** 2011. The s-layer glycome-adding to the sugar coat of bacteria. Int J Microbiol **2011**.

163. **Rocchetta, H. L., L. L. Burrows, J. C. Pacan, and J. S. Lam.** 1998. Three rhamnosyltransferases responsible for assembly of the A-band D-rhamnan polysaccharide in *Pseudomonas aeruginosa*: a fourth transferase, WbpL, is required for the initiation of both A-band and B-band lipopolysaccharide synthesis. Mol Microbiol **28:**1103-19.

164. **Ruan, X., J. M. Perez, C. L. Marolda, and M. A. Valvano.** Unpublished data. Conserved periplasmic-exposed residues of the Waal O-antigen ligase are essential for binding to lipid-linked O-antigen substrates.

165. **Ruiz, N., L. S. Gronenberg, D. Kahne, and T. J. Silhavy.** 2008. Identification of two inner-membrane proteins required for the transport of lipopolysaccharide to the outer membrane of *Escherichia coli*. Proc Natl Acad Sci U S A **105:**5537-42.

166. **Ruiz, N., D. Kahne, and T. J. Silhavy.** 2009. Transport of lipopolysaccharide across the cell envelope: the long road of discovery. Nat Rev Microbiol **7:**677-83.

167. **Rush, J. S., C. Alaimo, R. Robbiani, M. Wacker, and C. J. Waechter.** 2010. A novel epimerase that converts GlcNAc-P-P-undecaprenol to GalNAc-P-P-undecaprenol in *Escherichia coli* O157. J Biol Chem **285:**1671-80.

168. **Saldias, M. S., X. Ortega, and M. A. Valvano.** 2009. *Burkholderia cenocepacia* O antigen lipopolysaccharide prevents phagocytosis by macrophages and adhesion to epithelial cells. J Med Microbiol **58:**1542-8.

169. **Sandercock, L. E., A. M. MacLeod, E. Ong, and R. A. Warren.** 1994. Non-S-layer glycoproteins in eubacteria. FEMS Microbiol Lett **118:**1-7.

170. **Schaffer, C., and P. Messner.** 2004. Surface-layer glycoproteins: an example for the diversity of bacterial glycosylation with promising impacts on nanobiotechnology. Glycobiology **14:**31R-42R.

171. **Schaffer, C., T. Wugeditsch, H. Kahlig, A. Scheberl, S. Zayni, and P. Messner.** 2002. The surface layer (S-layer) glycoprotein of *Geobacillus stearothermophilus* NRS 2004/3a. Analysis of its glycosylation. J Biol Chem **277:**6230-9.

172. **Schertzer, J. W., and E. D. Brown.** 2003. Purified, recombinant TagF protein from *Bacillus subtilis* 168 catalyzes the polymerization of glycerol phosphate onto a membrane acceptor in vitro. J Biol Chem **278:**18002-7.

173. **Schild, S., A. K. Lamprecht, and J. Reidl.** 2005. Molecular and functional characterization of O antigen transfer in *Vibrio cholerae*. J Biol Chem **280:**25936-47.

174. **Scott, N. E., B. L. Parker, A. M. Connolly, J. Paulech, A. V. Edwards, B. Crossett, L. Falconer, D. Kolarich, S. P. Djordjevic, P. Hojrup, N. H. Packer, M. R. Larsen, and S. J. Cordwell.** 2011. Simultaneous glycan-peptide characterization using hydrophilic interaction chromatography and parallel fragmentation by CID, higher energy collisional dissociation, and electron transfer dissociation MS applied to the N-linked glycoproteome of *Campylobacter jejuni*. Mol Cell Proteomics **10:**M000031-MCP201.

175. **Sirisena, D. M., and M. Skurnik.** 2003. Mutations in the genes for synthesis of the outer core region of the lipopolysaccharide of *Yersinia enterocolitica* O:3. J Appl Microbiol **94:**686-92.

176. **Smith, C. S., A. Hinz, D. Bodenmiller, D. E. Larson, and Y. V. Brun.** 2003. Identification of genes required for synthesis of the adhesive holdfast in *Caulobacter crescentus*. J Bacteriol **185:**1432-42.

177. **Soldo, B., V. Lazarevic, and D. Karamata.** 2002. tagO is involved in the synthesis of all anionic cell-wall polymers in *Bacillus subtilis* 168. Microbiology **148:**2079-87.

178. **Sperandeo, P., R. Cescutti, R. Villa, C. Di Benedetto, D. Candia, G. Deho, and A. Polissi.** 2007. Characterization of lptA and lptB, two essential genes implicated in lipopolysaccharide transport to the outer membrane of *Escherichia coli*. J Bacteriol **189:**244-53.

179. **Sperandeo, P., F. K. Lau, A. Carpentieri, C. De Castro, A. Molinaro, G. Deho, T. J. Silhavy, and A. Polissi.** 2008. Functional analysis of the protein machinery required for transport of lipopolysaccharide to the outer membrane of *Escherichia coli*. J Bacteriol **190:**4460-9.

180. **Sperandeo, P., R. Villa, A. M. Martorana, M. Samalikova, R. Grandori, G. Deho, and A. Polissi.** 2011. New insights into the Lpt machinery for lipopolysaccharide transport to the cell surface: LptA-LptC interaction and LptA stability as sensors of a properly assembled transenvelope complex. J Bacteriol **193:**1042-53.

181. **Steen, A., G. Buist, K. J. Leenhouts, M. El Khattabi, F. Grijpstra, A. L. Zomer, G. Venema, O. P. Kuipers, and J. Kok.** 2003. Cell wall attachment of a widely distributed peptidoglycan binding domain is hindered by cell wall constituents. J Biol Chem **278:**23874-81.

182. **Steiner, K., R. Novotny, K. Patel, E. Vinogradov, C. Whitfield, M. A. Valvano, P. Messner, and C. Schaffer.** 2007. Functional characterization of the

initiation enzyme of S-layer glycoprotein glycan biosynthesis in *Geobacillus stearothermophilus* NRS 2004/3a. J Bacteriol **189:**2590-8.

183. **Steiner, K., R. Novotny, D. B. Werz, K. Zarschler, P. H. Seeberger, A. Hofinger, P. Kosma, C. Schaffer, and P. Messner.** 2008. Molecular basis of S-layer glycoprotein glycan biosynthesis in *Geobacillus stearothermophilus*. J Biol Chem **283:**21120-33.

184. **Stevenson, G., K. Andrianopoulos, M. Hobbs, and P. R. Reeves.** 1996. Organization of the *Escherichia coli* K-12 gene cluster responsible for production of the extracellular polysaccharide colanic acid. J Bacteriol **178:**4885-93.

185. **Stovepoindexter, J. L., and G. Cohen-Bazire.** 1964. The Fine Structure of Stalked Bacteria Belonging to the Family Caulobacteraceae. J Cell Biol **23:**587-607.

186. **Suits, M. D., P. Sperandeo, G. Deho, A. Polissi, and Z. Jia.** 2008. Novel structure of the conserved gram-negative lipopolysaccharide transport protein A and mutagenesis analysis. J Mol Biol **380:**476-88.

187. **Sutherland, I.** 2001. Biofilm exopolysaccharides: a strong and sticky framework. Microbiology **147:**3-9.

188. **Szymanski, C. M., D. H. Burr, and P. Guerry.** 2002. Campylobacter protein glycosylation affects host cell interactions. Infect Immun **70:**2242-4.

189. **Szymanski, C. M., R. Yao, C. P. Ewing, T. J. Trust, and P. Guerry.** 1999. Evidence for a system of general protein glycosylation in *Campylobacter jejuni*. Mol Microbiol **32:**1022-30.

190. **Tang, G., and K. P. Mintz.** 2010. Glycosylation of the collagen adhesin EmaA of *Aggregatibacter actinomycetemcomitans* is dependent upon the lipopolysaccharide biosynthetic pathway. J Bacteriol **192:**1395-404.

191. **Tatar, L. D., C. L. Marolda, A. N. Polischuk, D. van Leeuwen, and M. A. Valvano.** 2007. An *Escherichia coli* undecaprenyl-pyrophosphate phosphatase implicated in undecaprenyl phosphate recycling. Microbiology **153:**2518-29.

192. **Tefsen, B., J. Geurtsen, F. Beckers, J. Tommassen, and H. de Cock.** 2005. Lipopolysaccharide transport to the bacterial outer membrane in spheroplasts. J Biol Chem **280:**4504-9.

193. **Toh, E., H. D. Kurtz, Jr., and Y. V. Brun.** 2008. Characterization of the *Caulobacter crescentus* holdfast polysaccharide biosynthesis pathway reveals significant redundancy in the initiating glycosyltransferase and polymerase steps. J Bacteriol **190:**7219-31.

194. **Tran, A. X., M. S. Trent, and C. Whitfield.** 2008. The LptA protein of *Escherichia coli* is a periplasmic lipid A-binding protein involved in the lipopolysaccharide export pathway. J Biol Chem **283:**20342-9.

195. **Trent, M. S.** 2004. Biosynthesis, transport, and modification of lipid A. Biochem Cell Biol **82:**71-86.

196. **Trent, M. S., A. A. Ribeiro, W. T. Doerrler, S. Lin, R. J. Cotter, and C. R. Raetz.** 2001. Accumulation of a polyisoprene-linked amino sugar in polymyxin-resistant *Salmonella* typhimurium and *Escherichia coli:* structural characterization and transfer to lipid A in the periplasm. J Biol Chem **276:**43132-44.

197. **Trent, M. S., C. M. Stead, A. X. Tran, and J. V. Hankins.** 2006. Diversity of endotoxin and its impact on pathogenesis. J Endotoxin Res **12:**205-23.

198. **Trust, T. J., and R. A. Sparrow.** 1974. The bacterial flora in the alimentary tract of freshwater salmonid fishes. Can J Microbiol **20:**1219-28.

199. **Tsang, P. H., G. Li, Y. V. Brun, L. B. Freund, and J. X. Tang.** 2006. Adhesion of single bacterial cells in the micronewton range. Proc Natl Acad Sci U S A **103:**5764-8.

200. **Umbreit, T. H., and J. L. Pate.** 1978. Characterization of the holdfast region of wild-type cells and holdfast mutants of <i>Asticcacaulis biprosthecum</i&gt. Archives of Microbiology **118:**157-168.

201. **Vaara, M., T. Vaara, M. Jensen, I. Helander, M. Nurminen, E. T. Rietschel, and P. H. Makela.** 1981. Characterization of the lipopolysaccharide from the polymyxin-resistant pmrA mutants of *Salmonella* typhimurium. FEBS Lett **129:**145-9.

202. **Valvano, M. A.** 2003. Export of O-specific lipopolysaccharide. Front Biosci **8:**s452-71.

203. **Valvano, M. A., S. E. Furlong, and K. B. Patel.** *In press*. Genetics, biosynthesis and assembly of O-antigen. *In* Y. A. Knirel and M. A. Valvano (ed.), Bacterial Lipopolysaccharides. Structure, Chemical Synthesis, Biogenesis and Interaction with Host Cells.

204. **Valvano, M. A., C. L. Marolda, M. Bittner, M. Glaskin-Clay, T. L. Simon, and J. D. Klena.** 2000. The *rfaE* gene from *Escherichia coli* encodes a bifunctional protein involved in the biosynthesis of the lipopolysaccharide core precursor ADP-L-*glycero*-D-*manno*-heptose. J. Bacteriol. **182:**488-497.

205. **van Dam, J. E., A. Fleer, and H. Snippe.** 1990. Immunogenicity and immunochemistry of *Streptococcus pneumoniae* capsular polysaccharides. Antonie Van Leeuwenhoek **58:**1-47.

206. **Videira, P. A., A. P. Garcia, and I. Sa-Correia.** 2005. Functional and topological analysis of the *Burkholderia cenocepacia* priming glucosyltransferase BceB, involved in the biosynthesis of the cepacian exopolysaccharide. J Bacteriol **187:**5013-8.

207. **Vijayakumar, S., A. Merkx-Jacques, D. B. Ratnayake, I. Gryski, R. K. Obhi, S. Houle, C. M. Dozois, and C. Creuzenet.** 2006. Cj1121c, a novel UDP-4-keto-6-deoxy-GlcNAc C-4 aminotransferase essential for protein glycosylation and virulence in *Campylobacter jejuni.* J Biol Chem **281:**27733-43.

208. **Vollmer, W., and S. J. Seligman.** 2010. Architecture of peptidoglycan: more data and more models. Trends Microbiol **18:**59-66.

209. **Wang, G., P. G. Boulton, N. W. Chan, M. M. Palcic, and D. E. Taylor.** 1999. Novel *Helicobacter pylori* alpha1,2-fucosyltransferase, a key enzyme in the synthesis of Lewis antigens. Microbiology **145 (Pt 11):**3245-53.

210. **Wang, L., D. Liu, and P. R. Reeves.** 1996. C-terminal half of *Salmonella enterica* WbaP (RfbP) is the galactosyl-1-phosphate transferase domain catalyzing the first step of O-antigen synthesis. J Bacteriol **178:**2598-604.

211. **Wang, L., and P. R. Reeves.** 1994. Involvement of the galactosyl-1-phosphate transferase encoded by the *Salmonella enterica* rfbP gene in O-antigen subunit processing. J Bacteriol **176:**4348-56.

212. **Wang, X., and P. J. Quinn.** 2010. Lipopolysaccharide: Biosynthetic pathway and structure modification. Prog Lipid Res **49:**97-107.

213. **Ward, A., C. L. Reyes, J. Yu, C. B. Roth, and G. Chang.** 2007. Flexibility in the ABC transporter MsbA: Alternating access with a twist. Proc Natl Acad Sci U S A **104:**19005-10.

214. **Wecke, J., M. Perego, and W. Fischer.** 1996. D-alanine deprivation of *Bacillus subtilis* teichoic acids is without effect on cell growth and morphology but affects the autolytic activity. Microb Drug Resist **2:**123-9.

215. **Whitfield, C.** 2006. Biosynthesis and assembly of capsular polysaccharides in *Escherichia coli*. Annu Rev Biochem **75:**39-68.

216. **Whitfield, C.** 1995. Biosynthesis of lipopolysaccharide O antigens. Trends Microbiol **3:**178-85.

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

218. **Wilkinson, R. G., P. Gemski, Jr., and B. A. Stocker.** 1972. Non-smooth mutants of *Salmonella* typhimurium: differentiation by phage sensitivity and genetic mapping. J Gen Microbiol **70:**527-54.

219. **Woodward, R., W. Yi, L. Li, G. Zhao, H. Eguchi, P. R. Sridhar, H. Guo, J. K. Song, E. Motari, L. Cai, P. Kelleher, X. Liu, W. Han, W. Zhang, Y. Ding, M. Li, and P. G. Wang.** 2010. In vitro bacterial polysaccharide biosynthesis: defining the functions of Wzy and Wzz. Nat Chem Biol **6:**418-23.

220. **Wu, T., A. C. McCandlish, L. S. Gronenberg, S. S. Chng, T. J. Silhavy, and D. Kahne.** 2006. Identification of a protein complex that assembles lipopolysaccharide in the outer membrane of *Escherichia coli*. Proc Natl Acad Sci U S A **103:**11754-9.

221. **Xayarath, B., and J. Yother.** 2007. Mutations blocking side chain assembly, polymerization, or transport of a Wzy-dependent *Streptococcus pneumoniae* capsule are lethal in the absence of suppressor mutations and can affect polymer transfer to the cell wall. J Bacteriol **189:**3369-81.

222. **Yamashita, Y., Y. Shibata, Y. Nakano, H. Tsuda, N. Kido, M. Ohta, and T. Koga.** 1999. A novel gene required for rhamnose-glucose polysaccharide synthesis in *Streptococcus mutans*. J Bacteriol **181:**6556-9.

223. **Yethon, J. A., J. S. Gunn, R. K. Ernst, S. I. Miller, L. Laroche, D. Malo, and C. Whitfield.** 2000. *Salmonella enterica* serovar typhimurium waaP mutants show increased susceptibility to polymyxin and loss of virulence In vivo. Infect Immun **68:**4485-91.

224. **Yethon, J. A., E. Vinogradov, M. B. Perry, and C. Whitfield.** 2000. Mutation of the lipopolysaccharide core glycosyltransferase encoded by waaG destabilizes the outer membrane of *Escherichia coli* by interfering with core phosphorylation. J Bacteriol **182:**5620-3.

225. **Yethon, J. A., and C. Whitfield.** 2001. Lipopolysaccharide as a target for the development of novel therapeutics in gram-negative bacteria. Curr Drug Targets Infect Disord **1:**91-106.

226. **Young, K., L. L. Silver, D. Bramhill, P. Cameron, S. S. Eveland, C. R. Raetz, S. A. Hyland, and M. S. Anderson.** 1995. The envA permeability/cell division gene of *Escherichia coli* encodes the second enzyme of lipid A biosynthesis. UDP-3-O-(R-3-hydroxymyristoyl)-N-acetylglucosamine deacetylase. J Biol Chem **270:**30384-91.

227. **Zamyatina, A., S. Gronow, C. Oertelt, M. Puchberger, H. Brade, and P. Kosma.** 2000. Efficient Chemical Synthesis of the Two Anomers of ADP-L-glycero- and D-glycero-D-manno-Heptopyranose Allows the Determination of the Substrate Specificities of Bacterial Heptosyltransferases This work was

supported by grants from FWF (grant nos. P11449-MOB and P13843-CHE). Angew Chem Int Ed Engl **39:**4150-4153.

228. **Zarschler, K., B. Janesch, S. Zayni, C. Schaffer, and P. Messner.** 2009. Construction of a gene knockout system for application in *Paenibacillus alvei* CCM 2051T, exemplified by the S-layer glycan biosynthesis initiation enzyme WsfP. Appl Environ Microbiol **75:**3077-85.

229. **Zhou, G. P., and F. A. Troy, 2nd.** 2003. Characterization by NMR and molecular modeling of the binding of polyisoprenols and polyisoprenyl recognition sequence peptides: 3D structure of the complexes reveals sites of specific interactions. Glycobiology **13:**51-71.

230. **Zhou, G. P., and F. A. Troy, 2nd.** 2005. NMR study of the preferred membrane orientation of polyisoprenols (dolichol) and the impact of their complex with polyisoprenyl recognition sequence peptides on membrane structure. Glycobiology **15:**347-59.

231. **Zhou, Z., K. A. White, A. Polissi, C. Georgopoulos, and C. R. Raetz.** 1998. Function of *Escherichia coli* MsbA, an essential ABC family transporter, in lipid A and phospholipid biosynthesis. J Biol Chem **273:**12466-75.

Chapter 2

Distinct functional domains of the *Salmonella enterica* WbaP transferase that is involved in the initiation reaction for synthesis of the O antigen subunit

.

2.1. Introduction

LPS is the most abundant single component of the outer membrane of Gram-negative bacteria. The LPS molecule consists of a lipid A region that is embedded in the outer membrane, the core oligosaccharide, and a surface-exposed O-specific polysaccharide (O antigen) (32, 37). The biosynthesis of LPS is a complex process involving a large number of enzyme activities (32, 37). The core oligosaccharide is assembled on preformed lipid A by the sequential glycosyl transfer of monosaccharides, while the O antigen is assembled onto undecaprenyl phosphate (Und-P) at the cytoplasmic face of the inner membrane. There are two major pathways for the biosynthesis of O antigen, designated the Wzy-dependent and ABC transporter-dependent (Wzy-independent) pathways (32, 37). In the Wzy-dependent mechanism, Und-PP-linked O repeating units are exported to the periplasmic face of the plasma membrane by an unknown mechanism that requires the protein Wzx. Subsequently, the O units are polymerized by the Wzy polymerase, while the Wzz protein regulates, by an unknown mechanism, the length distribution of the polymers, also referred to as modality (9). In the most common ABC transporter-dependent mechanism, the O antigen polymer results from the transfer of glycosyl residues to Und-PP at the cytoplasmic face of the plasma membrane, and its export to the periplasmic face by an ABC transporter. In all of these pathways, the Und-PP-linked O antigen polymers are ultimately ligated to the outer core domain of the lipid A core (32).

The initiation reaction of O antigen biosynthesis involves the formation of a phosphodiester bond between the membrane-associated Und-P and a cytosolic UDP-sugar with the release of UMP. Depending on the specific micro-organism, this reaction is catalysed by two different families of proteins. In *Escherichia coli* and other members of the *Enterobacteriaceae*, initiation of the O antigen requires Und-PP-GlcNAc (3, 6, 17, 37) formed by the action of the UDP-GlcNAc:Und-P GlcNAc-1-P transferase encoded by the *wecA* gene (4-5, 7). WecA belongs to a large family of sugar-phosphate transferases that are widely conserved in both prokaryotes and eukaryotes [the polyisoprenyl-phosphate *N*-acetylhexosamine-1-phosphate transferase (PNPT) family] (7, 22, 31, 37). The other family includes polyisoprenyl-phosphate hexose-1-phosphate transferases (PHPT family), and its prototype member is WbaP from *Salmonella enterica* serovar

Typhimurium (*S.* Typhimurium) (37-39). The O antigen repeat in *S.* Typhimurium consists of a tetrasaccharide made of a linear backbone of galactose, rhamnose and mannose, and a side abequose residue that is linked to the terminal mannose (33). Upon completion of its synthesis, the tetrasaccharide O unit is translocated across the membrane and polymerized by the Wzy-dependent pathway, and the polymer is incorporated into the lipid A–core oligosaccharide by the O antigen ligase WaaL.

WbaP is the initiating UDP-Gal:Und-P Gal-1-P transferase. It has no known homologues in eukaryotic cells and lacks amino acid sequence similarity with members of the PNPT family. WbaP shows high sequence similarity with proteins involved in the synthesis of capsule exopolysaccharides in *Xanthomonas campestris* (21), *Erwinia amylovora* (11), *Streptococcus pneumoniae* (12), *Klebsiella pneumoniae* (8, 16) and *E. coli* K-12 (36) among others, and also with proteins involved in S-layer glycoprotein glycan biosynthesis in *Geobacillus stearothermophilus* (35). The WbaP protein is a large hydrophobic and basic protein with a predicted mass of 56 kDa and a pI of 9. Wang *et al.* (1996) (38) have proposed that the *S.* Typhimurium WbaP is a bifunctional protein that possesses two domains, the C-terminal portion, responsible for the Gal-1-P transferase activity, and the N-terminal region, required for the release of Und-PP-Gal from WbaP (39). According to these authors, the two domains are required for the synthesis of the O antigen. Their conclusions are based on the analysis of two *wbaP*(*T*) mutations, *wbaP4451*(*T*) and *wbaP4452*(*T*), mapping to the 5′ half of the *wbaP* gene, which result in an enzyme retaining Gal-1-P transferase activity but unable to mediate synthesis of O antigen. In this chapter, we propose that WbaP possesses three major domains and assign functional roles to each predicted domain. Our results suggest that the large periplasmic region is probably involved directly or indirectly in modulating O antigen chain length, and that the C-terminal domain carries the Gal-1-P transferase activity.

2.2. Materials and Methods

2.2.1. Bacterial strains, plasmids, media and growth preparation. Bacterial strains and plasmids used in this study are listed in Table 2.1. Bacteria were grown aerobically at 37 °C in Luria–Bertani (LB) medium (Difco). Media were supplemented as appropriate

with ampicillin, chloramphenicol or spectinomycin at final concentrations of 100, 30 and 80 μg ml^{-1}, respectively.

2.2.2. Methods used with nucleic acids. Isolation of plasmid DNA, digestion with restriction enzymes, ligation with T4 ligase and transformation were carried out as described by Maniatis *et al.* (1982) (24). DNA sequences were determined using an automated sequencer at the DNA Sequencing Facility, Robarts Research Institute (London, ON, Canada).

2.2.3. Topological model for WbaP and sequence analysis. To analyse WbaP topology, four topological prediction methods were used: TMHMM 2.0 (http://www.cbs.dtu.dk/services/TMHMM/), MEMSAT (http://bioinf.cs.ucl.ac.uk/psi pred/), TopPred 2.0 (http://bioweb.pasteur.fr/seqanal/interfaces/toppred.html) and PHD 2.1 (www.predictprotein.org/). psi-blast searches were conducted using resources available at NCBI (http://www.ncbi.nlm.nih.gov/BLAST/). Searches were done using the *Salmonella enterica* serovar Typhi (herein *S.* Typhi) WbaP C-terminal region, periplasmic loop and N-terminal domain sequences as a query. Multiple-sequence alignments of WbaP homologues were determined with the program clustal w (http://www.ebi.ac.uk/clustalw/). Protein secondary structure was predicted with PSI-PRED (http://bioinf.cs.ucl.ac.uk/psipred/).

2.2.4. Cloning of WbaP domains. PCR was carried out with *Pwo*I DNA polymerase (Roche Diagnostics) in a Perkin Elmer 2400 GeneAmp PCR system. Refer to Table 2.2 for oligonucleotides used in this study. Plasmid pSM10 was constructed by PCR amplification of a 2.9 kb fragment containing the *wbaP* and *manB* genes using primers manB3 and 592 with *S.* Typhi Ty2 DNA as template. This fragment was ligated into the *Sma*I site of pAA8 (Table 2.1). pSM13 is a pSM10 derivative in which the *manB* gene was deleted by digestion with *Sac*I and the plasmid self-ligated. pSM28 was constructed by PCR amplification of a 610 bp fragment from the 5′ end of *wbaP* using pSM13 as a DNA template, which was obtained using primers 591 and 1136, each incorporating *Sma*I sites. The resulting amplicon was digested with *Sma*I and ligated into the *Sma*I site of pAA8. pSM30 was constructed by PCR amplification of a 630 bp fragment from the 3′

end of *wbaP* using primers 592 and 897 with *S.* Typhi Ty2 DNA as template. The fragment was digested with *Sma*I and ligated into the *Sma*I site of pAA8. Plasmids pSM31 and pSM18 were constructed by inverse PCR using pSM13 as template and primers 1156 and 1155, and primers 897 and 900, respectively. Plasmids pKP10, pKP12, pKP13 and pKP14 were constructed in pBADNTF from *wbaP* derivatives of plasmids pSM18, pSM30, pSM31 and pSM28, respectively, by PCR using primers 1151 and 1152. The resulting amplicons were digested with *Sma*I, treated with polynucleotide kinase and ligated to *Sma*I-digested and alkaline phosphatase-treated pBADNTF. Plasmid pKP1 was constructed from *S.* Typhimurium LT2 DNA by PCR using primers 2158 and 2159. The resulting amplicon was digested with *Sma*I and *Sal*I and ligated to the corresponding sites on pBADFLAG.

2.2.5. Disruption of the *wbaP* gene. Mutagenesis was performed according to the method described by Datsenko and Wanner (2000) (13) to disrupt specific chromosomal genes using PCR products. Thus, *S.* Typhi Ty2 and *S.* Typhimurium LT2 cells carrying pKD46 were transformed by electroporation with a PCR product that was generated using plasmid pKD3 as template and primers 613 and 614. Transformants were plated on LB agar plates containing chloramphenicol, for the selection of *wbaP* mutants.

2.2.6. Oligonucleotide-directed mutagenesis of *wbaP*. Site-directed mutagenesis was performed using the QuikChange Site-Directed Mutagenesis kit from Stratagene, as recommended by the supplier. Plasmid pSM18 was used as template in the PCR reactions for the construction of pSM22 and pSM23, as indicated in Table 2.1. Plasmid pSM23 was constructed using primers 886 and 887. Two consecutive site-directed mutagenesis experiments were performed to construct plasmid pSM22. Primers 882 and 883 were used for the first PCR amplification, and primers 884 and 885 for the second. The resulting plasmids were transformed into *E. coli* DH5α by the calcium chloride method (18). Replacement mutants were confirmed by sequencing the entire *wbaP* gene.

2.2.7. Growth curves. For growth curves, overnight cultures of *E. coli* strains containing the appropriate plasmids were diluted to OD_{600} 0.03 (corresponding to a bacterial density of ~10^4 c.f.u. ml^{-1}), and growth was monitored every 30 min in a 100-

Table 2.1. Bacterial strains and plasmids

Strain or Plasmid	Relevant properties[a]	Source or reference
Strains		
Salmonella Typhi		This study
MSS1	Ty2, *ΔwbaP::cat*, Cm^R	Institute of Public Health, Chile
Ty2	*S. enterica* serovar Typhi, wild-type	
Salmonella Typhimurium		
LT2	*S. enterica* serovar Typhimurium, wild-type	S. Maloy
MSS2	LT2, *ΔwbaP::cat*, Cm^R	This study
SL1196	LT2, *hisD trpA fliC fljB gal xyl wbaP4451*(*T*)	(40)
SL1197	LT2, *hisD trpA fliC fljB gal xyl wbaP4452*(*T*)	(40)
Escherichia coli		
DH5α	F$^-$ ϕ80*lacZ*M15 *endA recA hsdR*(r^-_K m^-_K) *supE thi gyrA relA Δ*(*lacZYA-argF*)*U169*	Laboratory stock
Plasmids		
pAA8	41-bp fragment containing FLAG sequence in pUC18 *Hin*dIII site, Amp^R	(6)
pBADFLAG	pBAD24 for C-terminal FLAG fusions, inducible by arabinose, Amp^R	This study
pBADNTF	pBAD24 for N-terminal FLAG fusions, inducible by arabinose, Amp^R	(26)
pJD132	pBluescript SK derivative containing *wbaP* and flanking sequences from *E. coli* O9a:K30, Amp^R	(34)
pKD46	*bla* P$_{BAD}$ *gam bet exo* pSC101*ori^{ts}*	(13)
pKD3	*bla* FRT *cat* FRT PS1 PS2 *ori$_{R6K}$*	(13)
pKP1	pBADFLAG derivative encoding the *S. enterica* LT2 WbaP$_{M170-R354}$ C-terminally fused to the FLAG epitope	This study
pKP10	Identical insert as in pSM18 cloned in pBADNTF to provide an N-terminal FLAG tag fusion	This study
pKP12	Identical insert as in pSM30 cloned in pBADNTF to provide an N-terminal FLAG tag fusion	This study
pKP13	Identical insert as in pSM31 cloned in pBADNTF to provide an N-terminal FLAG tag fusion	This study
pKP14	Identical insert as in pSM28 cloned in pBADNTF to provide an N-terminal FLAG tag fusion	This study
pMF24	pEXT21 derivative containing the *wzx* from *S. enterica* LT2	(17)
pSM10	pAA8 derivative containing *manB* and a functional *wbaP* from *S. enterica* Ty2 containing a 1-bp deletion	This study

Strain or Plasmid	Relevant properties[a]	Source or reference
	at position 583 and a 2-bp deletion at position 645. This causes a frame shift at WbaP I194 and frame restoration at Y215 (Fig. 2.3)	
pSM13	pSM10 derivative containing only *wbaP* from *S.enterica* Ty2	This study
pSM18	pSM13 derivative encoding *S. enterica* Ty2 WbaP$_{M1\text{-}I144/F258\text{-}Y476}$ that has a deletion of the predicted periplasmic loop	This study
pSM22	pSM18 derivative encoding WbaP$_{H76Y, K81E}$ to recreate the *wbaP4452(T)* mutation	This study
pSM23	pSM18 derivative containing a 1-bp deletion at position 385 of *wbaP* from *S. enterica* Ty2 to recreate the *wbaP4451(T)* mutation	This study
pSM28	pAA8 encoding WbaP$_{M1\text{-}L203}$ from *S. enterica* Ty2	This study
pSM30	pAA8 encoding WbaP$_{Y258\text{-}Y476}$ from *S. enterica* Ty2	This study
pSM31	pSM13 derivative encoding WbaP$_{M1\text{-}T39/K305\text{-}Y476}$ from *S. enterica* Ty2	This study

[a] Cm, chloramphenicol; Amp, ampicillin; Sp, spectinomycin.

Table 2.2. Oligonucleotide primers used in this study

Name	Sequence
manB3	5'-ctggattttcgaaggagtggactaa-3'
NpSM18	5'-ctcggtaccaatggataatattgataataag-3'
CpSM18	5'-tcgctgcagttaatacgcaccatctcgcc-3'
591	5'-ctccccgggaatggataatattgataataag-3'
592	5'-tcggatatcttaatacgcaccatctcgcc-3'
613	5'-atggataatattgataataagtataatccacagctatgtaagtgtaggctggagctgcttcg-3'
614	5'-atcactgccataccgacgacgccgatctgttgcttggacatatgaatatcctcctatg-3'
882	5'-ggttttggattcgtttgagatattatacataccgcaagcc-3'
883	5'-ggcttgcggtatgtataatatcgcaaacgaatccaaaacc-3'
884	5'-cgtttgcgatattatacataccgcgagccattttggtatgagttaaaag-3'
885	5'-cttttaactcataccaaaatggctcgcggtatgtataatatcgcaaacg-3'
886	5'-cttttgccataatcctggtgctttttttcgcgcacttacaaag-3'
887	5'-ctttgtaagtgcgcgaaaaaaagcaccaggattatggcaaaag-3'
897	5'-ctccccgggtttagccatgaagttatgttattaagg-3'
900	5'-ctccccggggatacctagcttgttcaataaatg-3'
1136	5'-ctccccgggttaaatcccaaatagtcccagtg-3'
1151	5'-ctccccggggataatattgataataagtataatccacagc-3'
1152	5'-ctctctagattaatacgcaccatctcgccg-3'
1155	5'-atgcccgggaaagttactcgagatggtggtccg-3'
1156	5'-ttacccgggagtaaaatagacacatcctaatgaaagcc-3'
2158	5'-ctccccgggttggctctgatagcgtttac-3'
2159	5'-gactgtcgactccctggcaataggatcgttag-3'

well microtitre plate using a Bioscreen C automated microbiology growth curve analysis system (MTX Lab Systems). Growth rates were also compared between cultures with and without antibiotic selection to rule out antibiotic effects. Bioscreen C is a computerized bacterial incubator that measures growth continuously by optical density (23). The results were analysed statistically by one-way ANOVA and the Tukey post-test using the Prism 4 software package (GraphPad Software).

2.2.8. LPS analysis. Culture samples were adjusted to OD_{600} 2.0 in a final volume of 100 µl. Then, proteinase-K-digested whole-cell lysates were prepared as described previously (25, 27) and LPS was separated on 14% acrylamide gels using a Tricine–SDS buffer system (25). Gel loadings were normalized so that each sample represented the same number of cells. Each well was loaded with LPS extracted from approximately 1×10^8 c.f.u. Gels were silver-stained as described previously (25). Specific detection of *Salmonella* group D1 (for *S.* Typhi) and group B (for *S.* Typhimurium) O antigens was carried out by Western blotting. Group D1 and B rabbit antisera (Difco) were used at a 1:500 dilution. The reacting O antigen polysaccharides were detected by fluorescence with an Odyssey infrared imaging system (Li-cor Biosciences) using Rdye800-conjugated anti-rabbit affinity-purified secondary antibodies (Rockland). For estimation of the variation in the length of O antigen polysaccharide chains, a Tris–glycine system was used as described previously (27). The O antigen polysaccharide bands in all gels were quantified by densitometry and the pixel densities were measured using ImageJ (1).

2.2.9. *In vitro* transferase assay. Plasmids encoding the empty vector and the WbaP recombinant proteins (pBADNTF, pKP10 and pKP12) were transformed by electroporation into *S.* Typhimurium MSS2, which carries a *wbaP* deletion (Table 2.1). Membranes (containing enzymes and endogenous Und-P) were isolated from these transformants as described by (30). The reaction mixture for the transferase assay contained the membrane fraction (40 µg total protein) and 0.025 µCi radiolabelled UDP-[^{14}C]Gal (GE Healthcare) in 250 µl buffer (5 mM Tris/acetate, pH 8.5, 0.1 mM EDTA, 3 mM $MgCl_2$). After incubation at 37 °C for 30 min, the lipid-associated material was extracted twice with 250 µl 1-butanol. The combined 1-butanol extracts were washed once with 500 µl distilled water and the radioactivity counts of the 1-butanol fraction were determined with a Beckman liquid scintillation counter. Radioactivity counts were

normalized with respect to the amount of total protein in the membrane fraction used for the enzymic assay. For comparisons among the various preparations, relative enzymatic activities were expressed as a percentage of parental WbaP activity in LT2.

2.2.10. Membrane preparation and Western blotting. For visualization of WbaP constructs containing the FLAG epitope, bacterial cultures were grown overnight in 5 ml LB, diluted to an initial OD_{600} of 0.02, and incubated at 37 °C for 2 h until cultures reached OD_{600} 0.5. At this point, arabinose was added to a final concentration of 0.2% (w/v), and cultures were incubated for an additional 3 hrs until reaching OD_{600} ~0.8-1.0. Cells were then harvested by centrifugation at 10000 g for 10 min at 4 °C. The bacterial pellet was suspended in Tris/NaCl (20 mM Tris/HCl, pH 8.5, 300 mM NaCl) and the suspension lysed using a French Press cell. Cell debris was removed by centrifugation (15000 g for 15 min at 4 °C), and the clear supernatant was centrifuged at 30000 g for 30 min at 4 °C. The pellet, containing total membranes, was suspended in Tris/NaCl and frozen at −80 °C until required. The protein concentration was determined by the Bradford assay (Bio-Rad). SDS-PAGE, protein transfer to nitrocellulose membranes and immunoblots with the FLAG M2 mAb were performed as described previously (5), except that the reacting bands were detected by fluorescence with the Odyssey infrared imaging system (Li-cor Biosciences) using Alexa Fluor 680 goat anti-mouse secondary antibody (Molecular Probes). An affinity-purified rabbit antiserum was used for detection of the Wzz protein in similar preparations to those described above, using Alexa Fluor IRDye800 CW affinity-purified anti-rabbit IgG antibody (Rockland) as the secondary antibody for the detection of specific bands by fluorescence.

2.3. Results

2.3.1. WbaP has three unequally conserved predicted domains. Using widely employed algorithms for predicting transmembrane helices and the overall in–out topology of the intervening loops (15, 29), we established that the *Salmonella* WbaP polypeptide (476 aa) possesses three predicted domains: an N-terminal region of 133 aa containing transmembrane helices I–IV (N_{WbaP}), a large central periplasmic loop of 150 aa (P_{WbaP}), and the remaining 193 aa which include transmembrane helix V and a large C-terminal cytosolic tail (C_{WbaP}; Fig. 2.1a). To determine the relative conservation of

each predicted domain we performed psi-blast searches using the amino acid sequences of the N_{WbaP}, P_{WbaP}, and C_{WbaP} regions of the *S.* Typhi WbaP as queries. The WbaP proteins in *S.* Typhi and *S.* Typhimurium are virtually identical (19), except for four non-conservative changes ($WbaP_{STyphi}$/$WbaP_{STyphimurium}$) at positions 31 (Leu/Phe), 39 (Thr/Ile), 198 (Gly/Glu) and 199 (Thr/Ile). Multiple-sequence alignments of WbaP homologues revealed that the C_{WbaP} domain is highly conserved in a large number of different bacterial sugar transferases (pfam02397) involved in the biosynthesis of exopolysaccharides and LPSs from a broad spectrum of Gram-negative, Gram-positive and archaeal species (data not shown). In contrast, the N_{WbaP} and P_{WbaP} domains had amino acid sequence conservation with proteins predominantly involved in O antigen LPS biosynthesis from only a reduced number of species, which included, among others, *Salmonella*, *E. coli*, *Haemophilus influenzae* and *Actinobacillus*. WbaP did not show any sequence homology with members of the WecA family. Therefore, WbaP clearly falls within a different family of transferases that is restricted to prokaryotes only (37).

2.3.2. Overexpression of the *wbaP* gene causes a growth defect in *E.coli*. *S.* Typhi and *S.* Typhimurium mutants with a deletion of the *wbaP* gene were constructed, resulting in the replacement of the parental *wbaP* gene by a *cat* cassette encoding chloramphenicol resistance. As expected, mutants MSS1 (*S.* Typhi Δ*wbaP*::*cat*) and MSS2 (*S.* Typhimurium Δ*wbaP*::*cat*) produced a lipid A–core oligosaccharide band devoid of O antigen that did not react with group B (*S.* Typhimurium) or group D1 (*S.* Typhi) O antigen specific antisera (Fig. 2.2, second lane in all panels). The *wbaP* gene is the last gene of the *S. enterica* O antigen gene cluster and does not appear to contain a promoter immediately upstream (19, 33, 38). Therefore, we attempted to clone this gene as a single ORF downstream of the *lac* promoter region provided by the pAA8 plasmid vector (6) (Table 2.1). Restriction enzyme analysis and DNA sequencing of the recombinant plasmids obtained from these experiments revealed that none of the plasmids carried an intact *wbaP* gene. On a few occasions, an apparently correct clone was obtained, but DNA sequencing revealed gene deletions and/or frame-shift mutations in the coding region of *wbaP*. Failure to clone *wbaP* could not be explained by technical difficulties or PCR polymerase errors, since several cloning experiments on other unrelated *Salmonella* genes done in parallel with the same reagents were successful (data not shown).

Because it was not possible to clone the intact *wbaP* coding region, we first cloned a larger 2.9 kb DNA fragment containing *wbaP* and the upstream *manB* gene (Table 2.1, pSM10), and subsequently deleted most of the *manB* gene (see Methods). The resulting plasmid, pSM13, restored O antigen expression in the Δ*wbaP* mutants MSS2 (*S.* Typhimurium) and MSS1 (*S.* Typhi) (Fig. 2.2, third lane in all panels), and was selected for further study. However, less O antigen was produced by both the Δ*wbaP* mutant strains that contained pSM13, and the O antigen had a different migration pattern from that of the respective parental strains (Fig. 2.2). Transformation of MSS1 and MSS2 with pJD132, carrying an *E. coli* O9:K30 *wbaP* gene homologue and flanking sequences (34), restored O antigen production in both mutant strains (Fig. 2.2, fourth lane in all panels). The migration pattern and length of O antigen polymers mediated by pJD132 were identical to those of each respective parental strain. Therefore, the abnormal O antigen expressed by MSS1 (pSM13) and MSS2 (pSM13) mutants is not due to an intrinsic defect caused during the deletion mutagenesis procedure used to remove the *wbaP* gene in the Ty2 and LT2 strains, suggesting that the WbaP protein encoded by pSM13 is not functionally identical to wild-type WbaP (see next section below). However, the complemented mutants produced O antigens that were immunochemically identical to those of the respective parental strains, as detected by Western blots with group B- and group D1-specific antibodies (Fig. 2.2, third lane in right panels). Conservation of the immunoreactive epitopes indicates that the complete O antigen subunit was properly synthesized in both mutants. Therefore, the differences in O antigen profile are probably due to different processing of the O antigen and not to a defect in the initiation of its synthesis (see below).

The difficulties in obtaining a recombinant clone encoding a full-length WbaP protein in *E. coli* K-12 suggested that the expression of this protein could be deleterious to the host strain. Indeed, the growth rate of *E. coli* DH5α (pSM13) was significantly slower (doubling time 106±3 min, *n*=7) than that of *E. coli* DH5α carrying the vector control (doubling time 76±3.4 min; *P*<0.001, *n*=7). We have previously demonstrated that the *E. coli* Wzx flippase does not recognize the product of the reaction catalysed by the *Salmonella* WbaP protein (Und-PP-Gal) (17, 26). Therefore, the growth defect caused by WbaP expression in *E. coli* could be due to accumulation of Und-PP-Gal in the
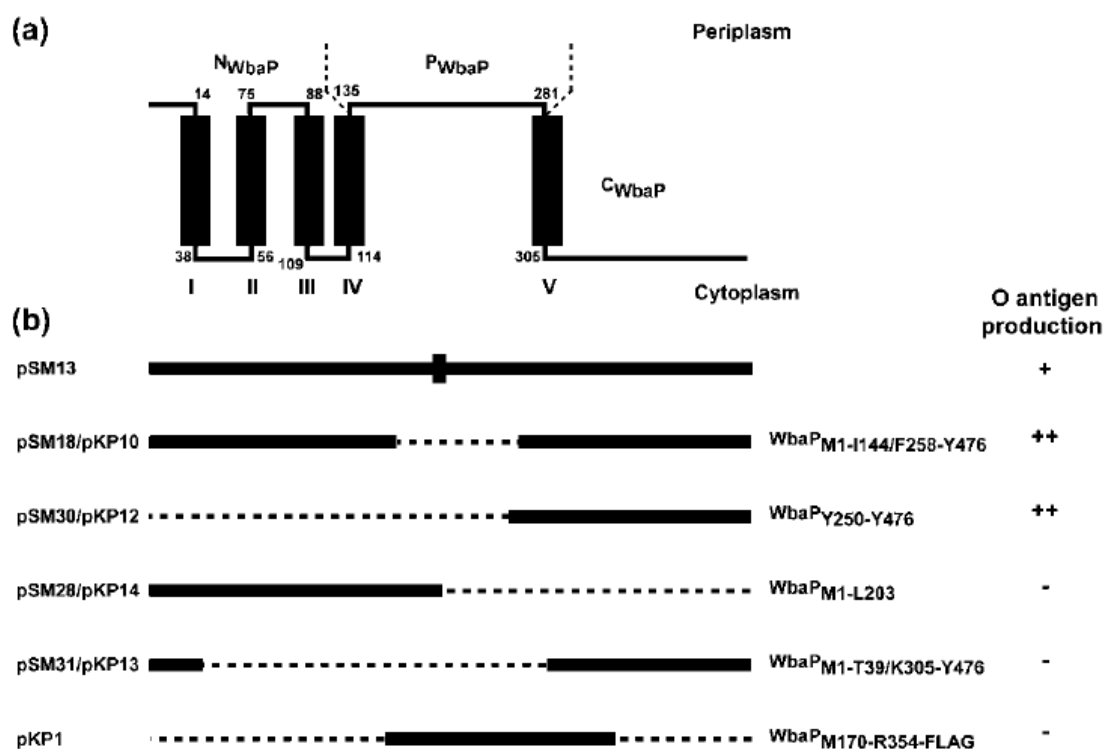
Figure 2.1 (a) Topological model of WbaP. Rectangles represent predicted transmembrane (TM) domains indicated by roman numerals (I–V). The numbers indicate the amino acid positions of the boundary of each TM domain. The dashed lines indicate the boundaries of each of the three predicted domains of WbaP. (b) Diagram representing the WbaP derivatives encoded by the various plasmids constructed in this study. Dashed lines represent regions of the protein absent in truncations. Plasmid pSM13 probably encodes a split functional version of WbaP consisting of two polypeptides (see Results for details). The pKP plasmid series, except for pKP1, encode the same WbaP protein derivatives as the pSM series, except that these proteins have an N-terminally fused FLAG epitope tag. The complementation of O antigen synthesis by each WbaP protein construct is indicated: ++, similar to parental strain; +, reduced O antigen production; −, absence of O antigen.
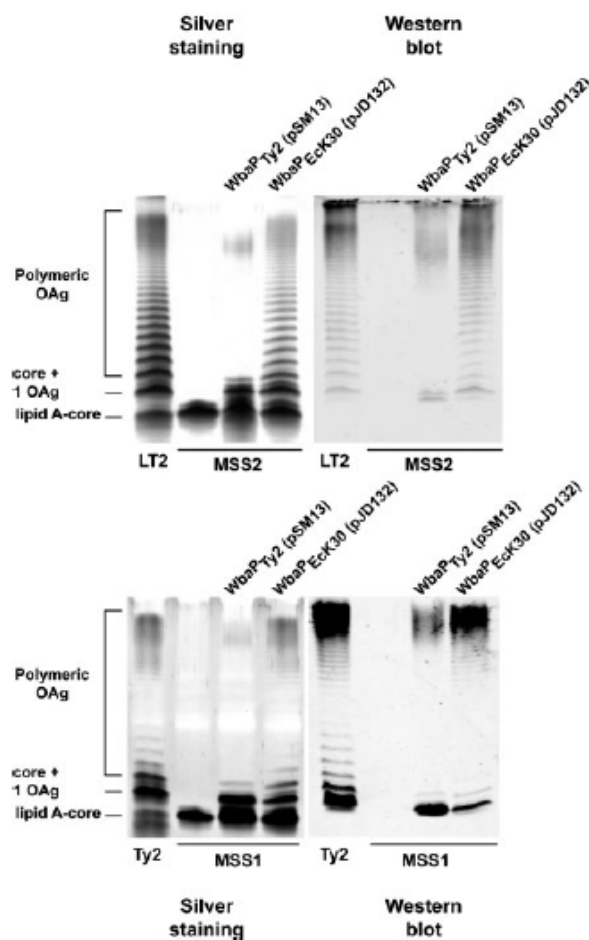
Figure 2.2 Complementation of *wbaP* deletion mutants in *S.* Typhimurium and *S.* Typhi. LPS samples were prepared from the indicated strains, and separated by electrophoresis on SDS–14% polyacrylamide gels. Each lane corresponds to LPS extracted from 10^8 bacteria. Left-hand panels are silver-stained gels; right-hand panels are Western blots of the same gels as on the left, which were reacted with group B (top) and group D1 (bottom) O-specific rabbit antisera, as described in Methods. The strains used for these experiments were LT2, *S.* Typhimurium; MSS2, *S.* Typhimurium Δ*wbaP*::*cat*; Ty2, *S.* Typhi; MSS1, *S.* Typhi Δ*wbaP*::*cat*. MSS1 and MSS2 carried no plasmids, or plasmid pSM13 (encoding WbaP$_{Ty2}$) or pJD132 (encoding WbaP$_{EcK30}$, the WbaP homologue from *E. coli* O9:K30). The bands corresponding to lipid A–core oligosaccharide, lipid A–core oligosaccharide plus one O antigen unit (core+1 OAg), and polymeric O antigen (polymeric OAg) are indicated.

cytoplasmic side of the plasma membrane, which could presumably block the recycling of Und-P and thus compromise cell wall biosynthesis. To test this possibility, we transformed *E. coli* DH5α (pSM13) with plasmid pMF24, carrying the *S*. Typhimurium *wzx* flippase gene. This strain had a doubling time of 67±1 min (*n*=7), which was not significantly different from that of DH5α (pAA8), confirming that introducing the *S*. Typhimurium *wzx* gene relieves the growth delay in *E. coli* DH5α (pSM13).

2.3.3. The N-terminal domain of WbaP is dispensible for *in vivo* and enzymatic activity. The sequence of the DNA inserts in pSM10 and pSM13 revealed a one-base deletion at position 583, producing a frame shift in the *wbaP* reading frame with a stop codon at base 601, and a two-base deletion at position 645 that restored the *wbaP* reading frame (Fig. 2.3). These mutations suggested that the WbaP protein expressed by pSM13 was a truncated form containing the N-terminal domain and part of the predicted periplasmic loop. However, this conclusion was not supported by previous studies demonstrating that *S*. Typhimurium WbaP possesses two functional domains: the C-terminal portion (GT domain) involved in Gal-1-P transferase function, and the N-terminal region (T domain) proposed to act in the release of the Und-PP-Gal from WbaP, prior to the flipping step (38-39). Therefore, various recombinant proteins were constructed to ascertain the functions of these two putative WbaP domains (Fig. 2.1), which could explain our complementation results with pSM13. Plasmid pSM28 contains a 609 bp insert corresponding to a truncated *wbaP* gene encoding the first 198 aa of WbaP (spanning four transmembrane domains and 63 aa of the periplasmic loop), designated WbaP$_{M1-L203}$ (Fig. 2.1b; Table 2.1). This corresponds to the same end point of the frame-shift mutation in pSM13 (Fig. 2.3). Plasmid pSM30 encoded WbaP$_{Y250-Y476}$ (Fig. 2.1b), a derivative of WbaP containing 19 aa of the periplasmic loop, TM-V, and the cytoplasmic C-terminal region. Plasmids pSM28 and pSM30 were separately transformed into MSS2 and the LPS profiles were analysed. The MSS2 (pSM30) strain expressing WbaP$_{Y250-Y476}$ produced O antigen while MSS2 (pSM28) expressing WbaP$_{M1-L203}$ did not (Fig. 2.4). These results indicate that the enzymic activity of WbaP resides in

its C-terminal region, and suggest that this region must also be expressed in pSM13, most likely from an internal translation start site within the 3′ half of the *wbaP* gene (38).

The sequence conservation of the WbaP C-terminal region with other bacterial proteins spans not only the predicted cytosolic 'tail' but also the predicted transmembrane helix V (Fig. 2.1a). To determine whether this transmembrane segment is indeed required for enzyme activity, the predicted C-terminal cytoplasmic domain was fused to the first transmembrane domain of WbaP to yield WbaP$_{M1-T39/K305-Y476}$ (encoded by pSM31; Fig. 2.1b; Table 2.1). Fig. 2.4 shows that pSM31 (WbaP$_{M1-T39/K305-Y476}$) did not restore O antigen production in the MSS2 mutant, suggesting that the predicted transmembrane helix V is required for WbaP activity. Failure to complement O antigen production was not due to lack of protein expression or membrane localization of the mutated WbaP protein, since an N-terminal FLAG-fusion derivative of WbaP$_{M1-T39/K305-Y476}$ (expressed by plasmid pKP13) was detected by Western blot analysis of total membrane preparations (see below).

The results described above suggested that the N-terminal region of WbaP is not required for *in vivo* O antigen synthesis in *S. enterica*. However, earlier work based on the analysis of *wbaP*(T) mutants concluded that the N terminus of WbaP is also required for normal O antigen expression (38). To clarify the contribution of the *wbaP*(T) mutations in WbaP function we performed site-directed mutagenesis of the *wbaP* gene cloned into pSM18 (encoding WbaP$_{M1-I144/F258-Y476}$) to recreate the *wbaP*(T) mutations *wbaP4451*(T) and *wbaP4452*(T) described before (38). Plasmid pSM23 carries a deletion at position 386 of the *wbaP* gene resulting in a frame-shift mutation that recreates *wbaP4451*(T), whereas pSM22 carries a *wbaP* gene with two point mutations within its coding region that result in His76Tyr and Lys81Glu replacements and recreate *wbaP4452*(T). Fig. 2.5 shows that the LPS profiles of MSS2 transformed with each of these plasmids were similar to the LPS profile of MSS2 (pSM18) encoding WbaP$_{M1-I144/F258-Y476}$ (Fig. 2.1b), which carries a deleted periplasmic loop that is not required for O antigen production (see below). In addition, we transformed the original *wbaP*(T) mutant strains of *S.* Typhimurium (SL1196 and SL1197; Table 2.1) with pJD132 that carries the *E. coli* K-30 *wbaP* gene (WbaP$_{EcK30}$). This plasmid complemented O antigen synthesis in MSS2 (Fig. 2.2, fourth lanes) but not in the T mutants (Fig. 2.5). Together, these results

```
505  gagatgatggggtttgatgttatcgcttttttttgatacggatgcgtcagatgctgaaataaatatgttgccg

     E   M   M   G   F   D   V   I   A   F   F   D   T   D   A   S   D   A   E   I   N   M   L   P
    169                                                                                     192

                 ┌a                                                                    ┌tg
577  gtgataaggacactgggactatttgggatttaaatcgtacaggtgatgtccattatatccttgcttaaatac

     V   I   R   T   L   G   L   F   G   I   *               M   S   I   I   S   L   L   K   Y
        194                                                                           215
```

Figure 2.3 Partial sequence of *S.* Typhi *wbaP* gene cloned into pSM13. The deleted bases in the region encoding the periplasmic loop are indicated in larger type above the sequence. Shading denotes the novel amino acids added as a consequence of the frame shift and the subsequent correction of the reading frame of wild-type *wbaP*. Amino acid numbering of the predicted WbaP polypeptide (excluding the new amino acids added as a consequence of the mutations) is indicated below the amino acid sequence.

demonstrate that the mutations in the *wbaP*(T) genes do not affect O antigen synthesis, and suggest that the lack of O antigen production in these strains is probably due to additional mutations in another gene or genes of the *Salmonella* O antigen cluster.

2.3.4. The predicted periplasmic loop of WbaP is involved in the chain-length distribution of the O antigen. A comparison of the first and fourth lanes in Fig. 2.4 reveals that although strain MSS2 expressing $WbaP_{Y250-Y476}$ (pSM30) produces O antigen in similar amounts to LT2, the chain-length distribution of the O antigen polysaccharide is altered in the mutant. Densitometric quantification demonstrated an eightfold increased average intensity of the bands within the region corresponding to 10–20 O-unit repeats (Fig. 2.4, dashed rectangle; data not shown). These results were confirmed using a different gel separation system, based on Tris–glycine, that provides higher resolution of the O antigen polymeric bands (Fig. 2.6a). Since $WbaP_{Y250-Y476}$ lacks most of the predicted periplasmic loop of WbaP (Fig. 2.1), we speculated that this region could be involved in O antigen chain-length distribution, possibly by interacting with the Wzz O chain-length regulator. To confirm this hypothesis we constructed $WbaP_{M1-I144/F258-Y476}$, a deletion derivative of WbaP protein (encoded by pSM18) that only contained 24 aa of the periplasmic loop. MSS2 expressing $WbaP_{M1-I144/F258-Y476}$ produced a smooth LPS, but the O antigen chain-length distribution was different from the wild-type profile (third lane in Figs 2.5 and 2.6a). This result suggests that although the periplasmic region of *Salmonella* WbaP is dispensable for the transferase activity, it is required for normal O antigen chain-length distribution.

Examination of the predicted secondary structure of the periplasmic loop of WbaP using the PSI-PRED program revealed a periodicity of α-helices and β-strands (Fig. 2.7). Most of the residues that are conserved in the alignments with WbaP homologues from other bacterial species are also present within the predicted α-helices and β-strands (Fig. 2.7; data not shown), supporting the idea that the periplasmic loop of WbaP could interact with other proteins involved in O antigen synthesis, with the O antigen itself, or with both. Therefore, we investigated whether expression of a truncated and non-functional WbaP protein with an intact predicted periplasmic loop could affect the function of the parental WbaP. We constructed plasmid pKP1 encoding a non-functional WbaP
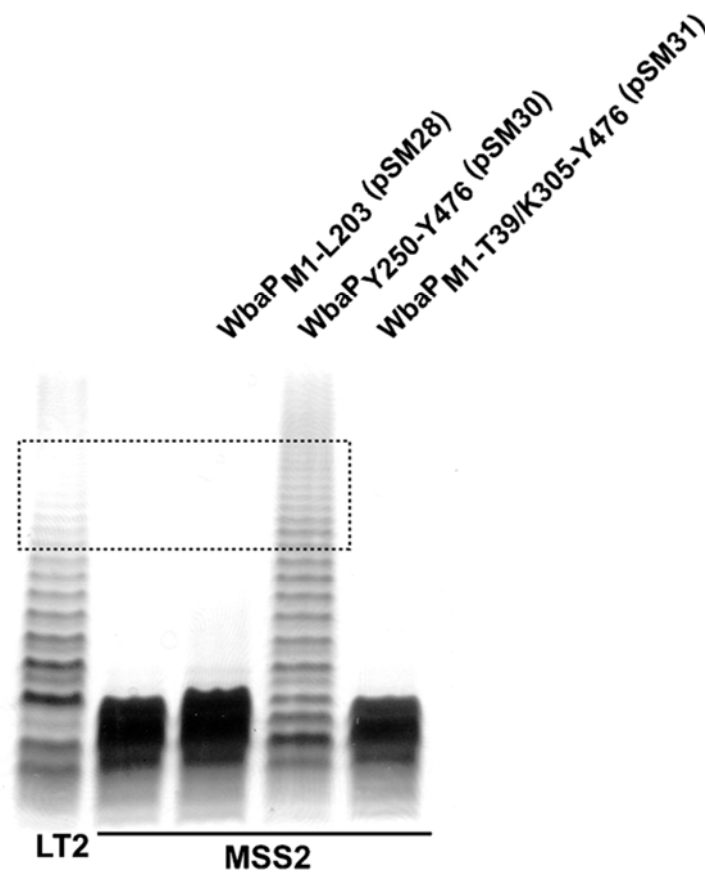
Figure 2.4 The C-terminal and the last transmembrane domains of WbaP are required for *in vivo* enzymatic activity. LPS was prepared from the indicated strains expressing no WbaP or WbaP derivatives with various truncations (described in Fig. 1b), and separated by electrophoresis on an SDS–14% polyacrylamide gel prior to silver staining. Each lane corresponds to LPS from 10^8 bacteria. LT2, *S.* Typhimurium; MSS2, *S.* Typhimurium Δ*wbaP*::*cat*. MSS2 carried no plasmids, or plasmid pSM28 (encoding WbaP$_{M1–L203}$), pSM30 (encoding WbaP$_{Y250–Y476}$) or pSM31 (encoding WbaP$_{M1–T39/K305–Y476}$). The dashed rectangle indicates the region of the O antigen polysaccharide profile (10–20 O-unit repeats) in which the differences in the chain-length distribution of the O antigen polysaccharide between strains LT2 and MSS2 expressing WbaP$_{Y250–Y476}$ are more obvious.
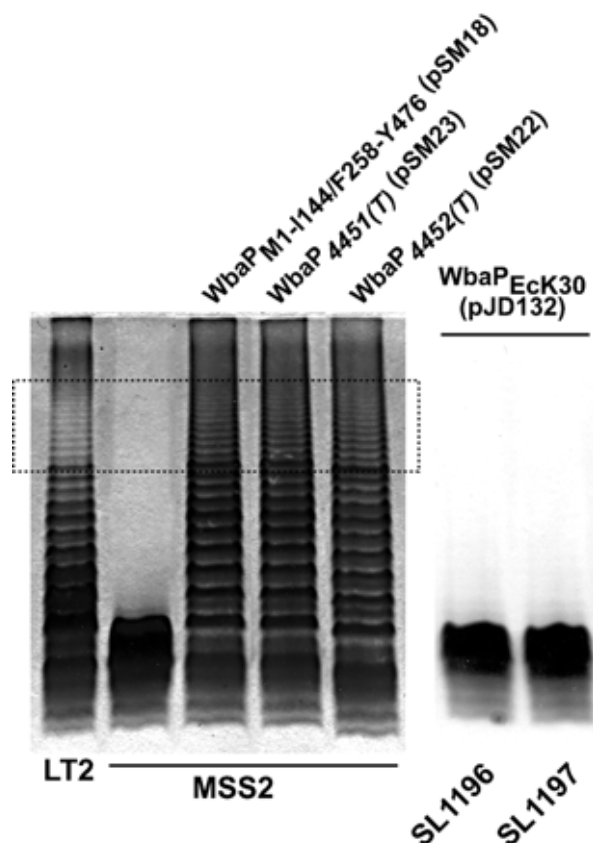
Figure 2.5 The N-terminal domain of WbaP is dispensable for *in vivo* enzymatic activity. LPS was prepared from the indicated strains, separated by electrophoresis on an SDS–14% polyacrylamide gel and silver-stained. Each lane corresponds to LPS from 10^8 bacteria. LT2, *S.* Typhimurium; MSS2, *S.* Typhimurium Δ*wbaP*::*cat*; SL1196, *S.* Typhimurium LT2 *wbaP4451*(*T*); SL1197, *S.* Typhimurium LT2 *wbaP4452*(*T*). MSS2 carried no plasmids, or plasmids pSM18 (encoding WbaP$_{M1-I144/F258-Y476}$), pSM23 [encoding WbaP$_{M1-I144/F258-Y476}$ with additional substitutions recreating the *wbaP4451*(*T*) mutation] and pSM22 [encoding WbaP$_{M1-I144/F258-Y476}$ with additional substitutions recreating the *wbaP4452*(*T*) mutation]. Strains SL1196 and SL1197 carried pJD132 (encoding WbaP$_{EcK30}$, the WbaP homologue from *E. coli* O9:K30). The dashed rectangle indicates the region of the O antigen polysaccharide profile in which the differences in the chain-length distribution of the O antigen polysaccharide between LT2 and MSS2 strains carrying various WbaP constructs are more obvious.

derivative containing most of the amino acids of the periplasmic loop, transmembrane helix V, and a truncated cytoplasmic tail fused to the FLAG epitope under the control of the arabinose-inducible P_{BAD} promoter (Fig. 2.1b). The polypeptide expressed by pKP1 (WbaP$_{M170-R354-FLAG}$) was detected in membrane fractions by immunoblotting with a FLAG mAb (see following section). After introduction of pKP1 into *S*. Typhimurium LT2 and incubation with arabinose, we observed reduced production of high-molecular-mass O antigens (Fig. 2.8a). In contrast, a control experiment on LT2 cells carrying the vector pBADFLAG showed no effect on O antigen production or chain-length distribution (Fig. 2.8b). As an additional control, we used plasmid pKP12 encoding a truncated WbaP protein (WbaP$_{F258-Y476}$) that contains only the last 24 aa of the periplasmic loop, transmembrane helix V, and the entire cytoplasmic tail (Fig. 2.1; Table 2.1). Introduction of pKP12 into strain LT2 (parental *wbaP*$^+$) resulted in an O antigen profile with altered chain-length distribution (Fig. 2.8b), similar to that found in the Δ*wbaP*::*cat* mutant MSS2 complemented with pSM18 (which encodes a WbaP protein derivative with a deletion of most of the periplasmic loop). Together, these results suggest that the WbaP periplasmic loop is to some extent involved in modulating the length of the O antigen chains, thereby affecting the distribution of the O antigen polysaccharide chains. Alternatively, expression of truncated or mutated WbaP proteins could induce stress responses that act non-specifically on Wzz, resulting in reduced production or stability of this protein. To address this possibility we determined the relative amount of Wzz expressed from its chromosomal gene in strains carrying the various recombinant plasmids. The immunoblot was quantitatively analysed by double labelling using FLAG- and Wzz-specific antibodies, to take advantage of an anti-FLAG reacting protein in *Salmonella* that served as an internal control (Fig. 2.6b, lower panel). The results of this experiment clearly indicate that, compared to the wild-type, Wzz production is not reduced by expressing the various recombinant WbaP proteins (Fig. 2.6b; data not shown). Also, there was no correlation between the levels of WbaP recombinant protein expression and the O antigen banding phenotype (see following section). Therefore, stress on *wzz* gene expression or Wzz protein stability cannot explain the abnormal O antigen chain distribution observed with the WbaP recombinant proteins. We conclude that the periplasmic loop of WbaP has a function that is not directly related to Gal-1-P transferase activity and that may involve protein–protein interactions with one or more additional proteins involved in the assembly of the O antigen chain.
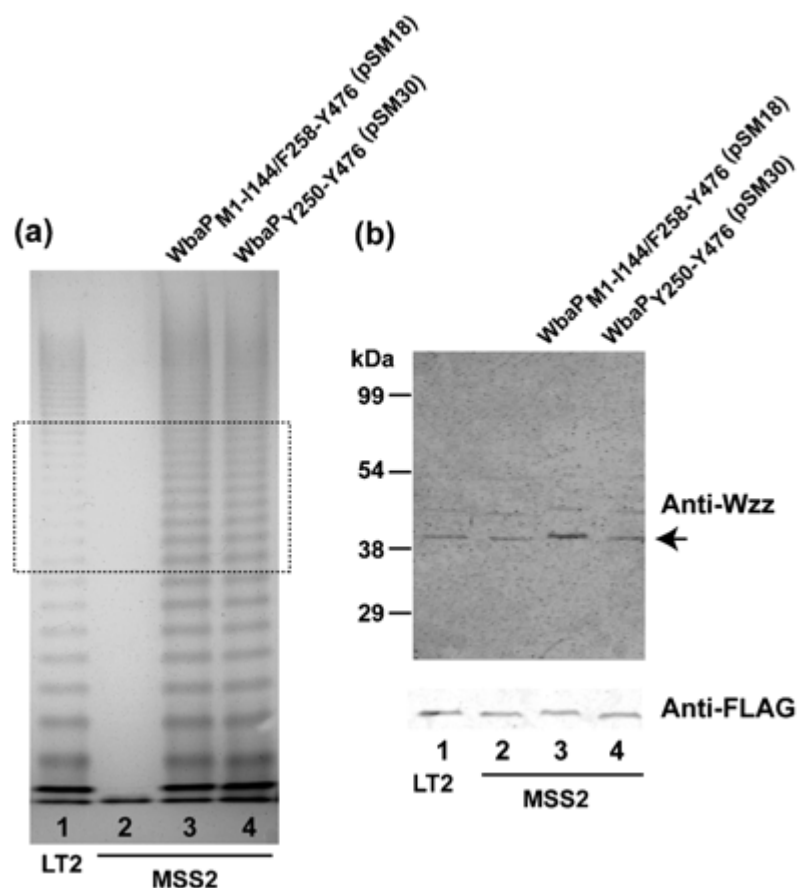
Figure 2.6 (a) Silver-stained LPS gel produced using the Tris-glycine system. LPS was prepared from strains LT2 (*S.* Typhimurium, lane 1) and MSS2 (*S.* Typhimurium Δ*wbaP*::*cat*, lanes 2–4). The dashed rectangle indicates the region of the O antigen polysaccharide profile (10–20 O-unit repeats) in which the differences in the chain-length distribution of the O antigen polysaccharide between strains LT2 and MSS2 expressing WbaP$_{M1–I144/F258–Y476}$ and WbaP$_{Y250–Y476}$ are more obvious. (b) Western blot of a membrane preparation of the strains shown in (a) developed with anti-Wzz and anti-FLAG antibodies, as described in Methods. The arrow indicates the Wzz protein. The positions of molecular mass standards are indicated: BSA (99 kDa), ovalbumin (54 kDa), carbonic anhydrase (38 kDa) and soybean trypsin inhibitor (29 kDa).

2.3.5.   Expression of WbaP protein derivatives as FLAG epitope fusions. To detect WbaP constructs in the bacterial cells, we resorted to the FLAG epitope tag. Numerous attempts to obtain a *wbaP* gene encoding a C-terminal FLAG fusion did not produce any clonable constructs containing an intact *wbaP* gene. However, a FLAG epitope was successfully fused to the C terminus of the truncated version of WbaP cloned in pKP1 (WbaP$_{M170–R354–FLAG}$). The presence of the fusion product was investigated by Western blotting with the FLAG M2 mAb using membrane extracts prepared from *E. coli* DH5α. It should be noted that the doublet present in all the lanes corresponds to minor outer-membrane proteins that cross-react with the FLAG epitope and are usually detected in total membrane preparations from *E. coli*. Heating the protein sample from *E. coli* DH5α (pKP1) at 100 °C for 5 min prior to loading resulted in aggregation, as determined by a new large protein band that barely penetrated the gel matrix (Fig. 2.9a, lane 2, black arrow) and in a polypeptide with an apparent mass of 26 kDa, which was consistent with the predicted mass of WbaP$_{M170–R354–FLAG}$ (23 kDa) (Fig. 2.9a, lane 2, asterisk). Also, a 20 kDa polypeptide was detected (Fig. 2.9a, lane 2, white arrow), which was interpreted as a WbaP$_{M170–R354–FLAG}$ degradation product. Protein aggregates that barely penetrate the gel matrix and anomalous migration in the gel are commonly observed with membrane proteins upon heat denaturation, particularly in proteins with high pIs (20). To avoid this problem, the protein samples were incubated at 45 °C for 30 min, as we have done in previous studies with the integral membrane proteins WecA and Wzx (6, 22, 26). In this case, we also detected a polypeptide with an apparent mass of 26 kDa corresponding to WbaP$_{M170–R354–FLAG}$ (Fig. 2.9a, lane 3, asterisk), and several larger oligomeric forms, most likely due to the mild denaturation conditions (Fig. 2.9a, lane 3, black arrows), as well as the putative degradation product of 20 kDa (Fig. 2.9a, lane 3, white arrow). Because of the difficulty in obtaining C-terminally FLAG-tagged derivatives of WbaP, each of the deleted versions of the *wbaP* gene in pSM18, pSM30, pSM31 and pSM28 was reconstructed as a FLAG N-terminal fusion derivative of WbaP, giving rise to pKP10, pKP12, PKP13 and pKP14, respectively (Fig. 2.1b; Table 2.1). All of these plasmids were derivatives of the vector pBADNTF, such that protein expression was under the control of the arabinose-inducible P$_{BAD}$ promoter. The chimeric proteins were detectable by immunoblotting of total membranes prepared from *S.* Typhimurium MSS2 containing each of the recombinant plasmids, although the levels of each WbaP fusion protein varied from construct to construct (Fig. 2.9b, arrows). Additional bands were also

Figure 2.7 Structural prediction of the WbaP predicted periplasmic loop calculated with PSI-PRED. Conf, confidence of the prediction for each residue as indicated by the height of each rectangle; Pred, graphical representation of the structural prediction that includes *α*-helical regions (black cylinders), *β*-strands (white arrows) and coils (black lines). The amino acid numbers correspond to the position in the wild-type WbaP protein. The grey rectangles represent the regions of residue conservation obtained from clustalw alignments of WbaP homologues identified from psi-blast searches using the *S*. Typhi periplasmic loop of WbaP protein as a query. The alignment includes sequences from (accession numbers in parentheses): *S*. Typhimurium (AAL20986.1), *S. enterica* serovar Choleraesuis (S22623), *Erwinia pyrifoliae* (CAC82926.1), *Erwinia amylovora* (CAA54879.1), *Pantoea stewartii* (AAC27321.1), *E. coli* (AAD21565.1), *Actinobacillus pleuropneumoniae* (AAG45943.1), *H. influenzae* (AAC22530.1), *Aggregatibacter actinomycetemcomitans* (AAG49415.1), *Nostoc punctiforme* (ZP_00110718.1) and *Burkholderia xenovorans* (ZP_00031541.1).

Figure 2.8 Effects on O antigen production caused by the expression of WbaP truncated forms in the context of the wild-type chromosomally encoded protein. Silver-stained gels of LPS preparations from *S.* Typhimurium LT2 and MSS2 are shown. The final concentration of arabinose in the growth medium (w/v) is indicated as 0 (no arabinose), 0.2 and 2%. LT2, *S.* Typhimurium; MSS2, *S.* Typhimurium Δ*wbaP*::*cat*. (a) O antigen migration pattern of LPS extracted from strains LT2, MSS2 and LT2 carrying pKP1 (encoding WbaP$_{M170–R354–FLAG}$). (b) Effect of pKP12 (encoding WbaP$_{Y250–Y476}$) on the O antigen pattern of strain LT2 in comparison with that of the control strains LT2 and LT2 (pBADFLAG).

found in total membranes of MSS2 carrying the pBADNTF vector control (Fig. 2.9b), and correspond to polypeptides that cross-react with the M2 mAb.

***In vitro* transferase activity of WbaP and mutant proteins.** The results obtained from the *in vivo* complementation experiments using the various mutant forms of WbaP suggested that the C-terminal domain is responsible for the transferase activity, and that the periplasmic loop is needed for the normal chain-length distribution of the O antigen. However, we could not identify a functional role for the N-terminal domain of WbaP. Since this domain appears to be predominantly found in WbaP homologues that are involved in O antigen synthesis, we reasoned that it could contribute to maintaining the stability of WbaP in the membrane and perhaps facilitate its association with other protein components involved in O antigen synthesis or assembly. We hypothesized that a defect in protein stability would be accompanied by a reduction of the enzymic activity of WbaP mutants lacking this domain. Thus, we assayed the enzymic activity of WbaP and mutant proteins by measuring the incorporation of radiolabelled [^{14}C]Gal (provided in the reaction as UDP-Gal) into a butanol-extractable fraction, which contains the Und-PP-linked sugar intermediates. The relative enzyme activity of the various mutant proteins was compared to that obtained from reactions using membranes prepared from LT2 (Fig. 2.10). The levels of activity of membranes prepared from the negative control strain MSS2 with plasmid pBADNTF was low, 2.8 ± 0.31 %, however background activity was still detected probably reflecting background levels of enzyme activity most likely due to additional transferases, such as those involved in the terminal steps of lipid A–core biosynthesis, which can also utilize UDP-Gal as a donor substrate. MSS2 with plasmids pKP10 and pKP12 had activities of 12.8 ± 2.1 and 126.1 ± 13.6 %, respectively. Therefore, the results of the enzymatic assays suggest that the mutated forms of WbaP that lack the N-terminal domain are enzymically active *in vitro*.

2.4. Discussion
2.4.1. The present study reports the functional characterization of WbaP from *S. enterica*. This protein is the prototype of the PHPT family, a large family of prokaryotic proteins involved in the initiation reaction for O antigen and exopolysaccharide synthesis (37), and also in the biosynthesis of the glycan moiety attached to bacterial glycoproteins (35). The predicted topology of WbaP revealed important differences from WecA, which
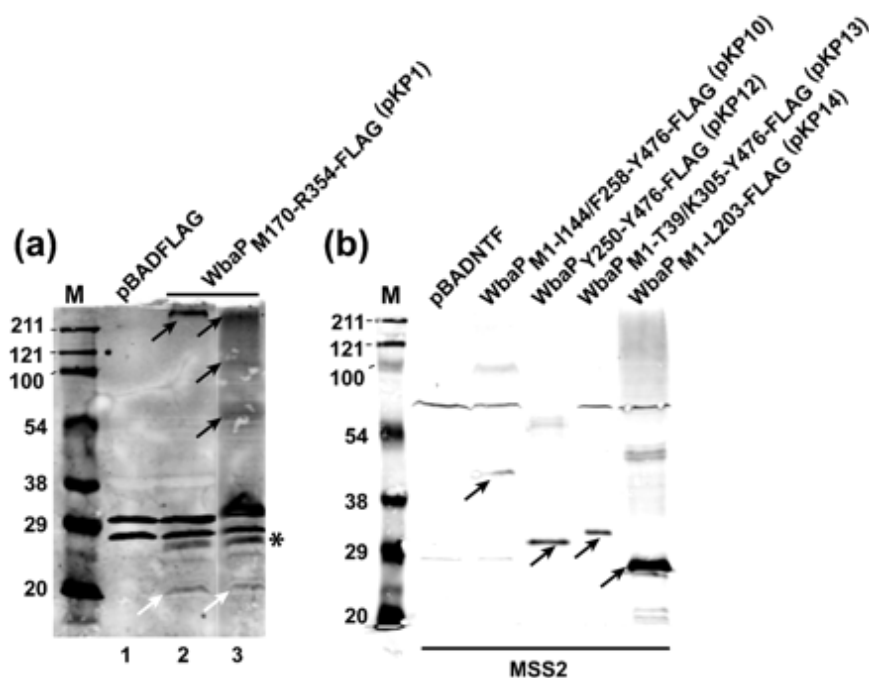
Figure 2.9 Protein expression of WbaP derivatives. After SDS-PAGE, samples were transferred to a nitrocellulose membrane and reacted with anti-FLAG mAbs. Prior to loading, samples were heated at 100 °C for 5 min (a, lanes 1 and 2) or incubated at 45 °C for 30 min (a, lane 3, and b, all lanes). M, molecular mass standards (in kDa): myosin (211), β-galactosidase (121), BSA (100), ovalbumin (54), carbonic anhydrase (38), soybean trypsin inhibitor (29) and lysozyme (20). (a) Total membranes were prepared from *E. coli* DH5α transformed with the control vector pBADFLAG (lane 1) or pKP1 (encoding WbaP$_{M170–R354–FLAG}$) (lanes 2–3). The doublet in the region of 27–29 kDa corresponds to minor outer-membrane proteins that cross-react with the FLAG mAb. The asterisk indicates the WbaP$_{M170–R354–FLAG}$ polypeptide encoded by pKP1. Black arrows indicate multimers of WbaP$_{M170–R354–FLAG}$; white arrows indicate a 20 kDa polypeptide interpreted as a degradation product of WbaP$_{M170–R354–FLAG}$. (b) Total membranes were prepared from *S.* Typhimurium MSS2 ($\Delta wbaP$::*cat*) transformed with the control vector pBADNTF (lane 1) and with various plasmids encoding several versions of WbaP (Fig. 2.1b; Table 2.1). The 70 kDa polypeptide that appears in all the lanes (hardly visible in the lane for the pKP12 sample) corresponds to a protein that cross-reacts with the FLAG mAb. Black arrows indicate the various truncated forms of WbaP.
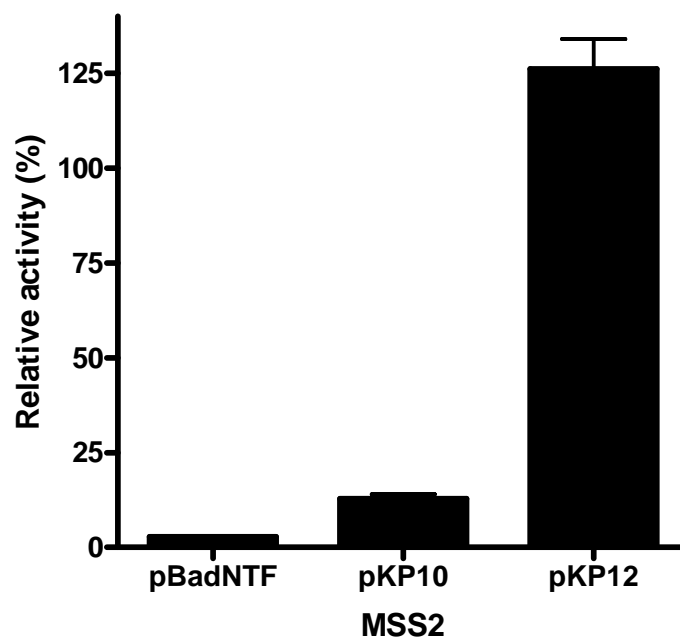
Figure 2.10 Comparative transferase activities of WbaP constructs. Membrane extracts were prepared from LT2 cell or MSS2 cells transformed with plasmids indicated above. The transferase assay measured the incorporation of radioactive Gal into a butanol-extractable fraction, as described in Methods. The results, expressed as a percentage of the wildtype WbaP activity in LT2 at 3 mM $MgCl_2$ concentration using normalized counts, represents the mean $\pm$ SD of the three independent experiments.

belongs to another large family of initiating enzymes in prokaryotes and eukaryotes that catalyse the transfer of *N*-acetylhexosamine-phosphate to isoprenoid lipid-phosphate (37). WecA consists of 11 transmembrane helices, five external and five internal loops (4-5, 7, 22). Whereas specific amino acid residues of functional importance in the transfer of GlcNAc-1-P to Und-P have been identified in cytosolic loops II, III and V of WecA, a domain structure in this protein is not obvious (4, 6, 22). In contrast, the topological model for WbaP revealed three predicted domains that could be clearly delineated: an N-terminal region containing a cluster of four transmembrane helices, a large central periplasmic loop, and a C-terminal cytosolic tail. Previous work by Wang et al. (1996) (38) has shown that the C-terminal half of WbaP carries the Gal-1-P transferase activity. The results of the present study, based on the expression of truncated forms of WbaP, strongly agree with this conclusion and refine the boundaries of the predicted Gal-1-P transferase domain, which contains the last transmembrane helix and a large cytoplasmic tail. The C-terminal domain of WbaP is highly conserved in a broad spectrum of bacterial species, while the N-terminal domain and the periplasmic loop show homology only with a reduced number of proteins, the majority of which are involved in O antigen synthesis.

The unequal conservation of each domain in the bacterial homologues of WbaP prompted us to investigate the potential functional roles of each predicted domain. Despite numerous attempts with various different cloning strategies we could not clone the intact *wbaP* gene from *S.* Typhimurium or *S.* Typhi. The only clone we obtained (plasmid pSM13) had mutations causing a frame shift that would yield a truncated WbaP protein. However, this plasmid could partially complement O antigen synthesis in the *S.* Typhimurium strain carrying a deletion of *wbaP*. Given that the predicted Gal-1-P transferase domain is in the C terminus of WbaP, we concluded that one or more translation start sites within the *wbaP* gene could be responsible for the expression of a polypeptide with Gal-P transferase activity. This is in agreement with earlier observations from *in vivo* radiolabelling showing that the 3′ half of *wbaP* can indeed initiate translation to synthesize a 25 kDa protein with Gal-P transferase activity (38). Several proteins maintain tertiary structure when the peptide backbone is cleaved, such that substrate binding and/or catalytic activity may be retained. For example, two independently cloned fragments containing pieces of the *lacY* and *mdfA* genes in the same plasmid lead to

membrane insertion of functional proteins and the synthesis of polypeptides that correspond to the N- and C-terminal portions of the LacY and MdfA proteins (2, 10). Therefore, although we could not demonstrate it directly, the functional data on WbaP strongly suggest that pSM13 encodes a split functional protein. Bacteria containing pSM13 displayed lower O antigen levels than those containing pSM18 and pSM30, despite all three plasmids having apparently intact C-terminal domains. These differences can be attributed to the split nature of the protein encoded by pSM13, in contrast to the WbaP forms encoded by the other plasmids from optimized promoters in the cloning vectors.

Sequence alignments of WbaP N-terminal domain homologues showed a highly conserved region of 14 aa between helices II and III, which are predicted to face the periplasm. Considering that O antigen polymerization and assembly occur at the periplasmic side of the plasma membrane, this region may be involved in protein–protein interactions with other components of O antigen biosynthetic machinery. In earlier work (38-39), mutations of *S.* Typhimurium LT2 affecting the ligation of polymerized O antigen to core–lipid A were mapped to the 3′ half of *wbaP*. The authors concluded that these mutations, referred to as *wbaP*(T), caused a possible block either in the flipping of the Und-PP-linked O antigen subunits across the plasma membrane or in the release of Und-P-linked Gal from WbaP. One of the O antigen-defective *S.* Typhimurium mutants described earlier (38-39) carries two amino acid substitutions (His76Tyr and Lys81Glu) within the 14 aa conserved region of WbaP. We recreated these mutations in our cloned *wbaP* genes and showed that they have no effect on O antigen production. However, O antigen synthesis in the original T mutants could not be complemented by our plasmids that encode the various forms of WbaP and the WbaP homologue from *E. coli* K-30. We thus conclude that the T mutants of *S.* Typhimurium LT2 carry other defects outside the *wbaP* gene that prevent the synthesis of O antigen. Genomic sequencing of the T mutants would be required to locate small mutations that may explain the defect in O antigen synthesis.

We also demonstrated that the N-terminal region of WbaP is not required for *in vivo* sugar-phosphate transferase activity and complementation of O antigen synthesis or for *in vitro* enzymic activity using crude cell-membrane extracts. In fact, *in vitro* activity

of the C-terminal construct was higher than the one with an intact N-terminal domain, however this may be due to lower protein expression levels in the latter. Further studies, currently under way in our laboratories, are required to determine the precise function of the N-terminal domain of WbaP. The domain corresponding to a predicted periplasmic loop of WbaP was also dispensable for Gal-P transferase activity, since it can be deleted without compromising the function of the protein as determined *in vivo* by complementation of O antigen synthesis and *in vitro* by examining the transfer of radioactive Gal to the endogenous Und-P acceptor. Nevertheless, this region appears to be involved in regulating O antigen polymerization, because mutant proteins that lack the periplasmic loop exhibit an altered chain-length distribution of the O antigen. The chain-length distribution of the O antigen polysaccharide is determined by the Wzz protein (28). This protein can form multimers, and is proposed to interact with the Wzy polymerase and the WaaL ligase in the periplasmic space, although the detailed mechanism that controls the chain-length distribution of O antigens is not well understood. Furthermore, *wzz* gene expression in *Salmonella* is subject to control by several global two-component regulators (14), and we have evidence that the stability of the Wzz protein in *E. coli* is modulated by the extracytoplasmic stress response (C. L. Marolda and others, unpublished data). However, we could not detect differences in Wzz production in any of the strains that contained the various WbaP-expressing recombinant plasmids, ruling out stress as a mechanism to explain the variations in O antigen length distribution that we observed in this study. Therefore, our results suggest that the periplasmic loop of WbaP also interacts with Wzz and/or other proteins involved in O antigen processing, including the Wzy polymerase. This is supported by experiments in which a non-functional WbaP protein with an intact periplasmic domain altered O antigen production in the presence of the parental WbaP protein. We are currently mapping the specific residues from this region in WbaP that are required for the chain-length distribution defect, which may reveal possible sites of contact with Wzz.

Together, the experiments described in this paper demonstrate that WbaP possesses domains with different functional roles. These predicted structural domains carry out several functions that are required not only for the expected transfer of Gal-1-P from UDP-Gal to Und-P, but also for the proper chain-length distribution of the O

antigen. We propose that WbaP acts as a scaffold protein that facilitates the completion of the O antigen subunit on the cytosolic side of the membrane and the downstream steps of O antigen assembly leading to the polymerization of the O antigen.

2.5. Chapter two references

1. **Abramoff, M. D., P. J. Magelhaes, and S. J. Ram.** 2004. Image processing with ImageJ. Biophotonics International **11:**36-42.
2. **Adler, J., and E. Bibi.** 2004. Determinants of substrate recognition by the *Escherichia coli* multidrug transporter MdfA identified on both sides of the membrane. J Biol Chem **279:**8957-65.
3. **Alexander, D. C., and M. A. Valvano.** 1994. Role of the *rfe* gene in the biosynthesis of the *Escherichia coli* O7-specific lipopolysaccharide and other O-specific polysaccharides containing *N*-acetylglucosamine. J Bacteriol **176:**7079-84.
4. **Amer, A. O., and M. A. Valvano.** 2001. 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:**3015-25.
5. **Amer, A. O., and M. A. Valvano.** 2002. 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:**571-82.
6. **Amer, A. O., and M. A. Valvano.** 2000. The N-terminal region of the *Escherichia coli* WecA (Rfe) protein, containing three predicted transmembrane helices, is required for function but not for membrane insertion. J Bacteriol **182:**498-503.
7. **Anderson, M. S., S. S. Eveland, and N. P. Price.** 2000. Conserved cytoplasmic motifs that distinguish sub-groups of the polyprenol phosphate:N-acetylhexosamine-1-phosphate transferase family. FEMS Microbiol Lett **191:**169-75.
8. **Arakawa, Y., R. Wacharotayankun, T. Nagatsuka, H. Ito, N. Kato, and M. Ohta.** 1995. Genomic organization of the *Klebsiella pneumoniae* cps region responsible for serotype K2 capsular polysaccharide synthesis in the virulent strain Chedid. J Bacteriol **177:**1788-96.
9. **Batchelor, R. A., G. E. Haraguchi, R. A. Hull, and S. I. Hull.** 1991. Regulation by a novel protein of the bimodal distribution of lipopolysaccharide in the outer membrane of *Escherichia coli*. J Bacteriol **173:**5699-704.
10. **Bibi, E., and H. R. Kaback.** 1990. In vivo expression of the *lacY* gene in two segments leads to functional lac permease. Proc Natl Acad Sci U S A **87:**4325-9.
11. **Bugert, P., and K. Geider.** 1995. Molecular analysis of the ams operon required for exopolysaccharide synthesis of *Erwinia amylovora*. Mol Microbiol **15:**917-33.
12. **Cartee, R. T., W. T. Forsee, M. H. Bender, K. D. Ambrose, and J. Yother.** 2005. CpsE from type 2 *Streptococcus pneumoniae* catalyzes the reversible addition of glucose-1-phosphate to a polyprenyl phosphate acceptor, initiating type 2 capsule repeat unit formation. J Bacteriol **187:**7425-33.

13. **Datsenko, K. A., and B. L. Wanner.** 2000. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. Proc Natl Acad Sci U S A **97:**6640-5.

14. **Delgado, M. A., C. Mouslim, and E. A. Groisman.** 2006. The PmrA/PmrB and RcsC/YojN/RcsB systems control expression of the *Salmonella* O-antigen chain length determinant. Mol Microbiol **60:**39-50.

15. **Drew, D., D. Sjostrand, J. Nilsson, T. Urbig, C. N. Chin, J. W. de Gier, and G. von Heijne.** 2002. Rapid topology mapping of *Escherichia coli* inner-membrane proteins by prediction and PhoA/GFP fusion analysis. Proc Natl Acad Sci U S A **99:**2690-5.

16. **Drummelsmith, J., and C. Whitfield.** 1999. Gene products required for surface expression of the capsular form of the group 1 K antigen in *Escherichia coli* (O9a:K30). Mol Microbiol **31:**1321-32.

17. **Feldman, M. F., C. L. Marolda, M. A. Monteiro, M. B. Perry, A. J. Parodi, and M. A. Valvano.** 1999. The activity of a putative polyisoprenol-linked sugar translocase (Wzx) involved in *Escherichia coli* O antigen assembly is independent of the chemical structure of the O repeat. J Biol Chem **274:**35129-38.

18. **Hanahan, D.** 1983. Studies on transformation of *Escherichia coli* with plasmids. J Mol Biol **166:**557-80.

19. **Jiang, X. M., B. Neal, F. Santiago, S. J. Lee, L. K. Romana, and P. R. Reeves.** 1991. Structure and sequence of the *rfb* (O antigen) gene cluster of *Salmonella* serovar typhimurium (strain LT2). Mol Microbiol **5:**695-713.

20. **Kashino, Y.** 2003. Separation methods in the analysis of protein membrane complexes. J Chromatogr B Analyt Technol Biomed Life Sci **797:**191-216.

21. **Katzen, F., D. U. Ferreiro, C. G. Oddo, M. V. Ielmini, A. Becker, A. Puhler, and L. Ielpi.** 1998. *Xanthomonas campestris* pv. campestris gum mutants: effects on xanthan biosynthesis and plant virulence. J Bacteriol **180:**1607-17.

22. **Lehrer, J., K. A. Vigeant, L. D. Tatar, and M. A. Valvano.** 2007. Functional characterization and membrane topology of *Escherichia coli* WecA, a sugar-phosphate transferase initiating the biosynthesis of enterobacterial common antigen and O-antigen lipopolysaccharide. J Bacteriol **189:**2618-28.

23. **Lowdin, E., I. Odenholt-Tornqvist, S. Bengtsson, and O. Cars.** 1993. A new method to determine postantibiotic effect and effects of subinhibitory antibiotic concentrations. Antimicrob Agents Chemother **37:**2200-5.

24. **Maniatis, T.** 1982. Molecular cloning : a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

25. **Marolda, C. L., P. Lahiry, E. Vines, S. Saldias, and M. A. Valvano.** 2006. Micromethods for the characterization of lipid A-core and O-antigen lipopolysaccharide. Methods Mol Biol **347:**237-52.

26. **Marolda, C. L., J. Vicarioli, and M. A. Valvano.** 2004. Wzx proteins involved in biosynthesis of O antigen function in association with the first sugar of the O-specific lipopolysaccharide subunit. Microbiology **150:**4095-105.

27. **Marolda, C. L., J. Welsh, L. Dafoe, and M. A. Valvano.** 1990. Genetic analysis of the O7-polysaccharide biosynthesis region from the *Escherichia coli* O7:K1 strain VW187. J Bacteriol **172:**3590-9.

28. **Morona, R., L. Van Den Bosch, and C. Daniels.** 2000. Evaluation of Wzz/MPA1/MPA2 proteins based on the presence of coiled-coil regions. Microbiology **146 (Pt 1):**1-4.

29. **Nilsson, J., B. Persson, and G. von Heijne.** 2000. Consensus predictions of membrane protein topology. FEBS Lett **486:**267-9.

30. **Osborn, M. J., J. E. Gander, E. Parisi, and J. Carson.** 1972. Mechanism of assembly of the outer membrane of *Salmonella* typhimurium. Isolation and characterization of cytoplasmic and outer membrane. J Biol Chem **247:**3962-72.

31. **Price, N. P., and F. A. Momany.** 2005. Modeling bacterial UDP-HexNAc: polyprenol-P HexNAc-1-P transferases. Glycobiology **15:**29R-42R.

32. **Raetz, C. R., and C. Whitfield.** 2002. Lipopolysaccharide endotoxins. Annu Rev Biochem **71:**635-700.

33. **Reeves, P.** 1993. Evolution of *Salmonella* O antigen variation by interspecific gene transfer on a large scale. Trends Genet **9:**17-22.

34. **Schaffer, C., T. Wugeditsch, P. Messner, and C. Whitfield.** 2002. Functional expression of enterobacterial O-polysaccharide biosynthesis enzymes in *Bacillus subtilis*. Appl Environ Microbiol **68:**4722-30.

35. **Steiner, K., R. Novotny, K. Patel, E. Vinogradov, C. Whitfield, M. A. Valvano, P. Messner, and C. Schaffer.** 2007. Functional characterization of the initiation enzyme of S-layer glycoprotein glycan biosynthesis in *Geobacillus stearothermophilus* NRS 2004/3a. J Bacteriol **189:**2590-8.

36. **Stevenson, G., K. Andrianopoulos, M. Hobbs, and P. R. Reeves.** 1996. Organization of the *Escherichia coli* K-12 gene cluster responsible for production of the extracellular polysaccharide colanic acid. J Bacteriol **178:**4885-93.

37. **Valvano, M. A.** 2003. Export of O-specific lipopolysaccharide. Front Biosci **8:**s452-71.

38. **Wang, L., D. Liu, and P. R. Reeves.** 1996. C-terminal half of *Salmonella enterica* WbaP (RfbP) is the galactosyl-1-phosphate transferase domain catalyzing the first step of O-antigen synthesis. J Bacteriol **178:**2598-604.

39. **Wang, L., and P. R. Reeves.** 1994. Involvement of the galactosyl-1-phosphate transferase encoded by the *Salmonella enterica rfbP* gene in O-antigen subunit processing. J Bacteriol **176:**4348-56.

40. **Wilkinson, R. G., P. Gemski, Jr., and B. A. Stocker.** 1972. Non-smooth mutants of *Salmonella* typhimurium: differentiation by phage sensitivity and genetic mapping. J Gen Microbiol **70:**527-54.

Chapter 3

Functional analysis of the C-terminal domain of the WbaP protein that mediates initiation of O antigen synthesis in *Salmonella enterica.*

.

A version of this chapter has been published:

3.1   Introduction

The outer membrane of Gram-negative bacteria is an asymmetric lipid bilayer with an outer leaflet consisting predominantly of lipopolysaccharide (LPS) molecules. LPS has a lipid A portion embedded in the outer membrane, a core oligosaccharide and an O-specific polysaccharide (O antigen) that is exposed to the bacterial surface. O antigen biogenesis is a multi-step process commencing at the cytoplasmic face of the plasma membrane (31, 43). The initiation reaction leads to the formation of a phosphoanhydride bond between a sugar-1-phosphate (commonly delivered as an UDP sugar) and undecaprenyl phosphate (Und-P) with the concomitant release of UMP. Subsequent steps in O antigen synthesis are completed *via* one of two major pathways designated ABC transporter and Wzy-dependent pathways (31, 43). In the Wzy-dependent pathway, glycosyltransferases sequentially add sugars to form a complete Und-P-P-linked O antigen subunit facing the cytosol, which is translocated to the periplasm by a process mediated by the Wzx protein. O antigen subunits are subsequently polymerized by the Wzy polymerase, and they acquire a particular length distribution by an unknown mechanism that requires Wzz, also called chain length determinant or O antigen co-polymerase (26). The polymer is finally ligated to independently synthesized lipid A-core oligosaccharide and the LPS molecule is transported to the outer membrane (31, 43) by the recently discovered Lpt pathway (35).

Enzymes catalyzing the initiation reaction can be grouped into two distinct families: polyisoprenyl-phosphate *N*-acetylhexosamine-1-phosphate (PNPT) and polyisoprenyl-phosphate hexose-1-phosphate (PHPT) transferases (43). PNPTs include prokaryotic and eukaryotic members while PHPTs are restricted to bacteria. WecA, a UDP-GlcNAc:Und-P GlcNAc-1-P transferase commonly found in the Enterobacteriaceae, is a prototypic member of the PNPT family. The *Salmonella enterica* WbaP has UDP-galactose (Gal):Und-P Gal-1-P transferase activity and is a prototypic member of the PHPT family (33, 36, 43-44). Other characterized PHPT members initiate the synthesis of glycans for S-layer protein glycosylation, such as *Geobacillus stearothermophilus* WsaP (39), and exopolysaccharides, such as *Xanthomonas campestris* GumD (18), *Streptococcus pneumonia* Cps2E (7, 47), *Klebsiella pneumoniae* ORF14 (5, 10), and *Escherichia coli K-12* WcaJ (40).

WbaP is a basic membrane protein with five predicted transmembrane helices (36). We have previously demonstrated three domains in WbaP: a hydrophobic, membrane embedded N-terminal region spanning four transmembrane helices, a central periplasmic loop, and a large C-terminal domain comprising the fifth transmembrane helix with the remainder of the protein predicted to face the cytosol (36). The last 219 residues encompassing the tail region are sufficient for galactosyl-phosphate transferase activity *in vivo* (36, 44). In this chapter, we characterize in more detail the C-terminal domain of WbaP. An N-terminal thioredoxin (TrxA) fusion to this domain led to enhanced protein folding and membrane expression. Together, trypsin accessibility and GFP localization experiments demonstrated that the C-terminus is cytoplasmic and analysis of GFP-fused proteins demonstrated that WbaP localizes to the poles of the cells. A targeted amino acid replacement strategy directed at highly conserved charged and polar residues identified three arginines (R319, R377, R401), one lysine (K331), one glutamic acid (E383), and two aspartic acids (D382 and D458) that are critical for *in vivo* and *in vitro* WbaP activity. We also show that residues R377, D331, E383, and D458 are located in three putative a-helical segments, which are highly conserved in other members of the WbaP family, suggesting they may form part of a catalytic center.

3.2   Materials and Methods

3.2.1   Bacterial strains, plasmids, media and growth conditions. Bacterial strains and the plasmids used in this study are listed in Table 3.1. Bacteria were grown aerobically at 37 °C in Luria-Bertani (LB) medium (Difco) (10 mg/ml tryptone; 5 mg/ml yeast extract; 5 mg/ml NaCl). Media were supplemented with 100 μg/ml ampicillin, 30 μg/ml chloramphenicol or 40 μg/ml kanamycin as appropriate.

3.2.2   Recombinant DNA methods, PCR, and cloning strategies. Plasmid DNA was isolated using the Qiagen miniprep kit (Qiagen Inc., Mississauga, Ontario, Canada). Digestion with restriction enzymes, ligation with T4-ligase and transformation were carried out as described by Maniatis *et al*. (1982). DNA sequences were determined using an automated sequencer at the York University Core Molecular Biology and DNA

Sequencing Facility, Toronto, Ontario, Canada. PCR amplifications were carried out with *Pwo*I DNA polymerase (Roche Diagnostics) in a Perkin Elmer 2400 GeneAmp PCR system. The oligonucleotides used for these experiments are listed in Table 3.2. Plasmid pKP2 was constructed by PCR amplification of a 1062-bp fragment using primers 591 and 2159 and *S. enterica* serovar Typhimurium LT2 DNA as template. This fragment was digested with *Sma*I and *Sal*I and ligated to these sites of pBADFLAG (Table 3.1). pSEF4 was constructed by PCR amplification using LT2 DNA as template and primers 3164 and 1152. The 1132-bp fragment was digested with *Sma*I and ligated to this site of pBADNTF. Plasmid pKP41 was constructed by PCR amplification of a 774-bp fragment, using primers 3599 and 258 and plasmid pSEF4 as a template. This fragment was digested with *Hind*III and ligated to this site of pEB-T7. To construct pKP42, we used pKP41 as a DNA template and the primers 3601 and 3602. The 1311-bp fragment obtained by PCR amplification was digested by *Nco*I and ligated to the *Nco*I and *Sma*I sites of pBAD24. The plasmid pKP43 was constructed in a similar way by PCR amplification of a 657-bp fragment using LT2 DNA as template and primers 3674 and 3675. This fragment was digested with *Kpn*I and *BamH*I and ligated to these sites of pWaldo-GFPe. Plasmid pKP80 was constructed using pKP43 as template for PCR with primers 897 and 2991. The 1508-bp fragment was digested with *Sma*I and *Hind*III and ligated to these sites of pBAD24. The plasmid pKP81 was constructed by PCR amplification of the *trxA* gene from the pEB-T7 plasmid using primers 3896 and 3897. The 442-bp *trxA* fragment was digested with *Sma*I and ligated to this site of pKP80. Site-directed mutagenesis was performed using the QuikChange Site-Directed Mutagenesis Kit from Stratagene, as recommended by the supplier. Plasmid pKP42 was used as a template in the PCR reactions for the construction of the plasmids in Table 3.1. Primers were designed with 15-20 nucleotides flanking each side of the targeted mutation. The resulting PCR products were digested with *Dpn*I and introduced into *E. coli* DH5α by transformation using the calcium chloride method (14). WbaP constructs and replacement mutants were confirmed by sequencing.

3.2.3   Isolation of spheroplasts. Bacterial cultures were grown overnight in 5 ml of LB, diluted to an initial OD$_{600}$ of 0.2, and incubated at 37ºC for 2 h until cultures reached an

Table 3.1 Bacterial strains and plasmids

Strain or Plasmid	Relevant properties[a]	Source or reference
Strains		
S. enterica		This study
LT2	*S. enterica* serovar Typhimurium, wild-type	S. Maloy
MSS2	LT2, $\Delta wbaP::cat$, Cm^R	This study
Escherichia coli		
DH5α	F$^-$ ϕ80*lac*ZM15 *endA recA hsdR*(r^-_K m^-_K) *supE thi gyrA relA* Δ(*lacZYA-argF*)*U169*	Laboratory stock
MV501	*E. coli* O7, *wecA*::Tn*10*	
Plasmids		
p0381-FLAG	pDA12 derivative encoding BCAL0381 C-terminally fused to the FLAG epitope	(6)
pBAD24	Cloning vector inducible by arabinose, Amp^R	(13)
pBADFLAG	pBAD24 for C-terminal FLAG fusions	(4)
pBADNTF	pBAD24 for N-terminal FLAG fusions, inducible by arabinose, Amp^R	(24)
pEB-T7	Cloning vector, IPTG-inducible, for N-terminal 6×His fusions, Amp^R	E. H. Ball
pKP2	pBADFLAG derivative encoding the *S. enterica* LT2 WbaP $_{M1-R354}$ C-terminally fused to the FLAG epitope	This study
pKP10	pBADNTF derivative encoding the *S. enterica* Ty2 WbaP$_{M1-I144/F258-Y476}$ N-terminally fused to the FLAG epitope	(36)
pKP12	pBADNTF derivative encoding the *S. enterica* Ty WbaP$_{F258-Y476}$ N-terminally fused to the FLAG epitope	(36)
pKP41	pEB-T7 derivative encoding the S. enterica LT2 WbaP $_{F258-Y476}$ N-terminally fused to 6×His-TrxA	This study
pKP42	pBAD24 derivative encoding the *S. enterica* LT2 WbaP $_{F258-Y476}$ N-terminally fused to 6×His-TrxA	This study
pKP50	pKP42 encoding WbaP-R319A	This study
pKP52	pKP42 encoding WbaP-R377A	This study
pKP54	pKP42 encoding WbaP-R401A	This study
pKP55	pKP42 encoding WbaP-R458A	This study

pKP56	pKP42 encoding WbaP-R377K	This study
pKP57	pKP42 encoding WbaP-R401K	This study
pKP58	pKP42 encoding WbaP-R446A	This study
pKP59	pKP42 encoding WbaP-R449A	This study
pKP60	pKP42 encoding WbaP-R319Q	This study
pKP61	pKP42 encoding WbaP-R401Q	This study
pKP63	pKP42 encoding WbaP-R377Q	This study
pKP64	pKP42 encoding WbaP-Y449F	This study
pKP65	pKP42 encoding WbaP-Y449S	This study
pKP66	pKP42 encoding WbaP-R368A	This study
pKP67	pKP42 encoding WbaP-D458E	This study
pKP68	pKP42 encoding WbaP-Q368A	This study
pKP69	pKP42 encoding WbaP-K331R	This study
pKP70	pKP42 encoding WbaP-R333A	This study
pKP71	pKP42 encoding WbaP-D382A	This study
pKP73	pKP42 encoding WbaP-G393A	This study
pKP74	pKP42 encoding WbaP-E383A	This study
pKP75	pKP42 encoding WbaP-K331A	This study
pKP80	pBAD24 encoding $WbaP_{F258-Y476}$ with an C-terminal GFP	This study
pKP81	pKP80 encoding an N-terminal 6×His-TrxA fusion	This study
pKP84	pKP42 encoding WbaP-R319K	This study
pKP85	pKP42 encoding WbaP-R331Q	This study
pKP87	pKP42 encoding WbaP-D382E	This study
pKP88	pKP42 encoding WbaP-D382N	This study
pKP89	pKP42 encoding WbaP-E383D	This study
pKP90	pKP42 encoding WbaP-E383Q	This study
pKP91	pKP42 encoding WbaP-D458N	This study
pKP92	pKP42 encoding WbaP-I352A	This study
pKP93	pKP42 encoding WbaP-V372A	This study
pKP94	pKP42 encoding WbaP-F388A	This study
pKP95	pKP42 encoding WbaP-E409A	This study
pKP96	pKP42 encoding WbaP-W453A	This study
pKP97	pKP42 encoding WbaP-L462A	This study
pSEF4	pBADNTF derivative encoding *S. enterica* LT2 $WbaP_{I105-Y476}$ N-terminally fused to the FLAG epitope	This study
pWaldo-GFPe	Cloning vector, IPTG-inducible, for C-terminal GFP-6×His fusions, Kn^R	(9)

Table 3.2 Oligonucleotide primers used in this study

Name	Sequence
258	5'-GACCGCTTCTGCGTTCTGAT-3'
591	5'-CTCCCCGGGAATGGATAATATTGATAATAAG-3'
897	5'-CTCCCCGGGTTTAGCCATGAAGTTATGTTATTAAGG-3
1152	5'-CTCTCTAGATTAATACGCACCATCTCGCCG -3
2159	5'-GACTGTCGACTCCCTGGCAATAGGATCGTTAG-3'
2991	5'-GCTAGTTATTGCTCAGCGG-3'
3164	5'-CTACCCGGGATAGCGTTACAAAATGGCAG -3'
3599	5'-CGAAAGCTTGTTTAGCCATGAAGTTAT-3'
3601	5'-AGGGGAATTGTGAGCGGATA-3'
3602	5'-CCTTTCGGGCTTTGTTAGCA-3'
3674	5'-CCAGGTACCTTTAGCCATGAAGTTATGT-3'
3675	5'-CTATAAGGATCCATACGCACCATCTCGCCG-3'
3896	5'-CTACCATGGGCAGCAGCCATCAT-3'
3897	5'-CTACCCGGGTCCCTGAAAATACAG-3'

OD_{600} of 0.6. At this point, arabinose was added to a final concentration of 0.2% (w/v), and cultures were incubated for an additional 3 h until reaching an OD_{600} of approximately 0.8-1.0. The cultures were placed on ice for 20 min and 7.5 ml volumes were centrifuged at 16, 100 xg for 1 min. The pellet was washed in 1 ml of cold sucrose buffer (18% sucrose, 100 mM Tris-HCl, pH 8.5) and resuspended in 1 ml sucrose buffer, 50 μl EDTA (0.5 M) and 50 μl lysozyme (5 mg/ml). Cells were incubated on ice for 1 h and observed using light microscopy after 30 min. Once 90% of cells were observed to be spheroplasts the cells were separated into 250 μl aliquots.

3.2.4   Protease treatment

3.2.4.1 Trypsin digestion. Trypsin accessibility experiments were performed using a protocol adopted from Lee and Manoil (20). Spheroplasts (250 μl) and membrane vesicles (60 μg of total membrane protein) were treated with 25 μl and 7 μl of trypsin (300 μg/ml) respectively for 30 min at room temperature. The reactions were terminated with the addition of one sample volume of soybean trypsin inhibitor (STI) (1 mg/ml) and phenylmethylsulfonyl fluoride (PMSF) (20 mg/ml). Spheroplasts were washed in 250 μl of sucrose buffer containing 25 μl STI and 5 μl PMSF and cells were resuspended in 600 μl of freeze-thaw buffer (10 mM Tris pH 8, 10 μl STI (1 mg/ml), 10 μl PMSF (20 mg/ml)). Cells were lysed with four freeze-thaw cycles using an ethanol-dry ice bath and then treated with 20 μl of $MgCl_2$ (1M), 6 μl of DNase (1mg/ml) and protease inhibitor cocktail (Roche Diagnostics).

3.2.4.2 TEV cleavage. For TEV cleavage experiments spheroplasts were prepared as mentioned above. Samples (150μL) were treated with 2 μL of AcTEV protease (Invitrogen) and 1.5 μL of 0.1 M dithiothreitol. After treatment, cells were centrifuged at 5,900 xg and the supernatant, representing the periplasmic fraction in intact spheroplast samples, was collected for analysis.

3.2.5   Total membrane preparation and immunoblotting. For visualization of epitope tagged WbaP constructs, bacterial cultures were grown and induced with

arabinose as described above. Cells were then harvested by centrifugation at 10,000 x*g* for 10 min at 4ºC. The bacterial pellet was suspended in Tris/NaCl (20 mM Tris-HCl [pH 8.5] + 300 mM NaCl) and protease inhibitor cocktail (Roche diagnostics) and the suspension lysed using a French Press cell. Cell debris were removed by centrifugation (15 000 x*g* for 15 min at 4ºC), and the clear supernatant was centrifuged at 30 000 x*g* for 30 min at 4ºC. The pellet, containing total membranes, was suspended in Tris/NaCl. The protein concentration was determined by the Bradford assay (Bio-Rad). Sodium dodecylsulfate polyacrylamide gel electrophoresis (SDS-PAGE), protein transfers to nitrocellulose membranes, and immunoblots were performed as described (30). For detection of 6×His-, FLAG-, and GFP-tagged proteins, membranes were incubated with a 1:10 000 dilution of anti-His IgG2a monoclonal antibodies (Amersham), anti-FLAG monoclonal antibodies (Sigma), or anti-GFP monoclonal antibodies (Roche) respectively. For quantitative immunoblotting, a 6×His tagged lambda phage lytic transglycosylase (LT) was purified and various amounts (50 to 200 ng) of this standard were loaded on the same gel as test samples. The pixel density of the gel bands was analyzed with ImageJ software (1) (W.S. Rasband, U.S. National Institutes of Health, Bethesda, MD; http//rsb.info.nih.gov/ij/).

3.2.6 Microscopy. Overnight cultures of MSS2 cells containing pKP80 or pKP81, which express WbaP$_{CT}$-GFP or TrxA-WbaP$_{CT}$ -GFP under control of the arabinose-inducible *P$_{BAD}$* promoter (Table 3.1), were diluted in LB medium to obtain an OD$_{600}$ of 0.2, and protein expression was induced with arabinose as described above. After 2 h of induction, the culture was placed on ice for 1 h to facilitate GFP folding. Similar experiments were performed without arabinose in the growth medium. Bacteria were visualized with no fixation using an Axioscope 2 (Carl Zeiss) microscope 24 with an ×100/1.3 numerical aperture Plan-Neofluor objective and a 50-W mercury arc lamp with a GFP band pass emission filter set (Chroma Technology) with excitation at 470 ± 20 nm and emission at 525 ± 25 nm. Images were digitally processed using the Northern Eclipse imaging analysis software (version 6.0; Empix Imaging, Mississauga, Ontario, Canada).

3.2.7 LPS analysis. Culture samples were adjusted to OD_{600} 2.0 in a final volume of 100 μl. Then, proteinase-K digested whole-cell lysates were prepared as described elsewhere (25) and LPS was separated on 14% acrylamide gels using a Tricine/SDS buffer system (22). Gel loading was normalized so that each sample represented the same number of cells. Each well was loaded with approximately $1 \times 10_8$ cfu. Gels were silver stained by a modification of the procedure of Tsai and Frasch (42).

3.2.8 *In vitro* transferase assay. Total membranes (containing WbaP and endogenous Und-P) were isolated from strains expressing WbaP constructs or pBAD24. The reaction mixture for the transferase assay contained the membrane fraction, with 0.025 μCi radiolabeled UDP-^{14}C-galactose (specific activity: 307 mCi/mmol, Amersham Biosciences) in 250 μl buffer (40 mM Tris-HCl, pH 8.5, 0.5 mM EDTA, specified $MgCl_2$). After incubation at 37 ºC for 30 min, the lipid-associated material was extracted twice with 250 μl 1-butanol. The combined 1-butanol extracts were washed once with 500 μl distilled water and the radioactive counts of the 1-butanol fraction were determined with a Beckman liquid scintillation counter. Enzyme activity was expressed as pmol ^{14}Cgalactose incorporated per mg of total membrane protein-1 or radioactive counts were normalized using a 6×His-tagged protein standard (see above) and expressed as pmol ^{14}C-galactose incorporated per mg protein. Assays using exogenous Und-P (Institute of Biochemistry and Biophysics of the Polish Academy of Sciences, Warsaw, Poland) were carried out as described above with 50 μM Und-P and 0.1% 3-[(3-cholamidopropyl)dimethylammonio]-1-propanesulfonate (CHAPS) in the reaction mixture.

3.2.9 Sequence alignment. ClustalW was used for sequence alignments. For amino acid conservation the server SAM TO2 was used (Karplus, K., Karchin, R., et al. 2003) with the input peptide sequence spanning G310-Y476.

3.3   Results

3.3.1   The C-terminus of WbaP resides in the cytoplasm. WbaP consists of five predicted transmembrane helices (TM), a large periplasmic loop between TM-IV and TM-V (PL2), and a large cytoplasmic tail region (Fig. 1) (36). Previous work has demonstrated that only the WbaP C-terminus (WbaP$_{CT}$), from phenylalanine-258 (F258) to tyrosine-476 (Y476), is required for transferase activity (36, 44). To experimentally establish which portion of the C-terminus resides in the cytoplasm we conducted protease accessibility assays on spheroplasts as described in sections 3.2.3 and 3.2.4.1. We opted for this method as an alternative approach to study protein topology due to the poor membrane expression of WbaP constructs fused to reporter proteins such as LacZ (β-galactosidase) and PhoA (alkaline phosphatase). The protease accessibility strategy relies on the principle that after formation of spheroplasts, regions of an integral membrane protein that are exposed to the periplasm would be susceptible to proteolysis, while regions facing the cytoplasm would be protected (8). Trypsin cleaves selectively after lysine and arginine residues (27). Due to its basic nature, WbaP contains multiple trypsin cleavage sites spanning the entire length of the protein. Since we were unable to obtain a recombinant plasmid expressing full length WbaP with C-terminal or N-terminal tags (36) (data not shown), we used truncated WbaP derivatives containing N-terminally (FLAGWbaP$_{M1-I144/F258-Y476}$, encoded by pKP10) and C-terminally (WbaP$_{M1-R354}$-FLAG, encoded by pKP2) fused FLAG epitopes (Fig. 3.1), which enabled us to detect trypsin cleavage fragments from both ends of the protein. *E. coli* DH5α cells expressing the respective constructs were converted to spheroplasts and after incubation with trypsin, soybean trypsin inhibitor (STI) was added before lysis of spheroplasts by freeze thaw. Samples were separated by SDS-PAGE and analyzed by immunoblot. The samples corresponding to untreated FLAG-tagged WbaP constructs localize to the inner membrane, as they are detected in both spheroplasts and membrane vesicles, and each have an apparent mass of approximately 43 kDa (Fig. 3.2A, lanes 1 and 5, and Fig. 3.2B, lanes 1 and 3). A high molecular mass band of approximately 98 kDa was observed also with both proteins, although this band was much more intense in the case of WbaP$_{M1-R354}$-FLAG (Fig. 2A). We interpreted this band as due to oligomerization of WbaP$_{M1-R354}$-
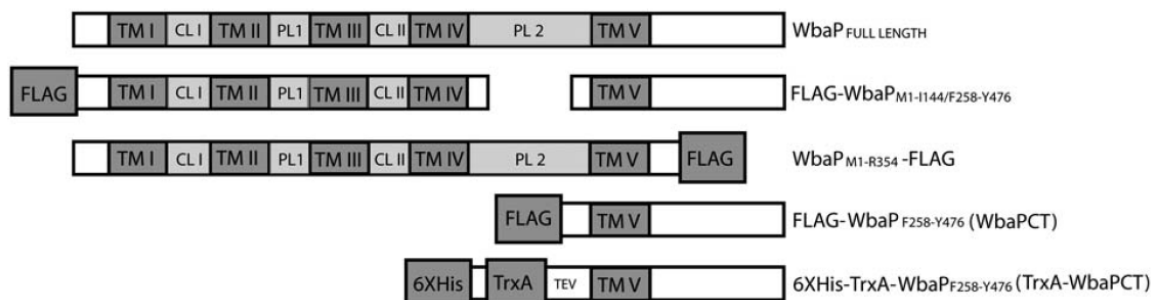
Figure 3.1. Graphic representation of the WbaP parental protein and the various deleted derivatives encoded by plasmids used in this study. All constructs contain either a FLAG or 6×His epitope tag for detection on the N- or C-terminus. Plasmid pKP42 encodes a construct that also has an N-terminal TrxA fusion followed by a TEV cleavage site. Predicted transmembrane (TM), periplasmic loops (PL), and cytoplasmic loops (CL) are indicated.

FLAG either as an artifact of the mild denaturing conditions we routinely use to process membrane proteins by SDS-PAGE (3, 21, 24, 30) (see section 3.2.5) or alternatively, protein-protein interactions mediated by the periplasmic loop that is absent in FLAG-WbaP$_{M1-I144/F258-Y476}$ (Fig. 3.1). Trypsin treatment of spheroplasts expressing WbaP$_{M1-R354}$-FLAG resulted in a band of approximately 36 kDa (Fig. 3.2A, lane 3), which corresponds to the mass of a product cleaved within the first predicted periplasmic loop (PL1; Fig. 3.1). Trypsin treatment of spheroplasts expressing FLAG-WbaP$_{M1-I144/F258-Y476}$ resulted in a band of approximately 9 kDa, which also corresponded to a product cleaved in PL1 (Fig. 3.2A, lane 6). There were no detectable bands corresponding to the predicted cytoplasmic tail region, suggesting that this region was not accessible to proteolysis. To confirm that the C-terminus is susceptible to trypsin cleavage, we isolated inverted membrane vesicles (IMVs) by French press disruption of cells expressing WbaP$_{M1-R354}$-FLAG and FLAG-WbaP$_{M1-I144/F258-Y476}$, and treating the IMVs with trypsin prior to SDS-PAGE. Trypsin treatment of IMVs expressing WbaP$_{M1-R354}$-FLAG resulted in several bands in the region of the gel corresponding approximately to 7-12 kDa (Fig. 3.2B, lane 2), and treatment of IMVs expressing FLAG-WbaP$_{M1-I144/F258-Y476}$ resulted in bands of approximately 30 kDa, 29 kDa, and 22 kDa (Fig. 3.2, lane 4). These bands correspond to the expected products from cleavage sites located within the predicted cytoplasmic tail. Together, these results suggest that amino acids located to at least the C-terminal 20 kDa of WbaP are in the cytoplasm, in agreement with the bioinformatic predictions. In our trypsin experiments using spheroplasts or total membranes vesicles, we observed no proteolytic cleavage products from the large predicted periplasmic loop 2 (PL2). This region is predicted to have complex secondary structure (36), which could prevent access of trypsin.

3.3.2 An N-terminal thioredoxin fusion improves folding of overexpressed WbaP$_{F258-Y476}$ and enhances activity *in vivo*. Overexpressed FLAG-WbaP$_{F258-Y476}$, herein referred to as WbaP$_{CT}$, which contains a 24 residues from PL2, TM-V and the WbaP C terminal cytosolic tail (Fig. 3.1), can restore O antigen surface expression in the *S.* Typhimurium Δ*wbaP* strain MSS2 (Fig. 3.3). However, as compared to the parental strain, WbaP$_{CT}$ does not mediate production of full-length O antigen since only short O
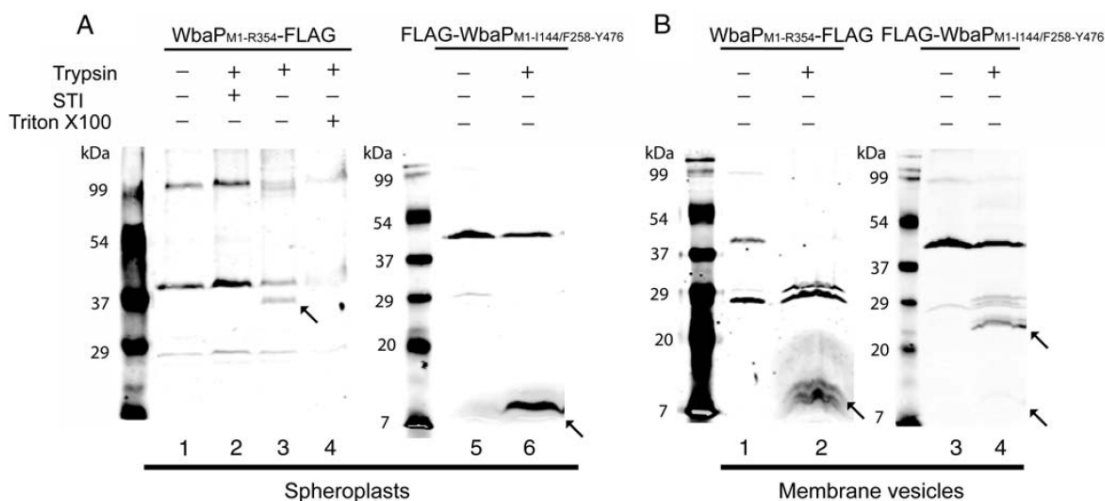
Figure 3.2 Cytoplasmic localization of the C-terminus by trypsin accessibility experiments in spheroplasts and with inverted membrane vesicles. Samples were treated with trypsin, soybean trypsin inhibitor (STI), and/or Triton X-100 as indicated in the panels. Upon treatment, samples were analyzed by immunoblot with anti-FLAG M2 monoclonal antibodies. (A). Spheroplasts of DH5α cells expressing WbaPM1–R354-FLAG or WbaPFLAG–M1–I144/F258–Y476 (both ~43 kDa). Spheroplasts were incubated for 30 min at room temperature. Lanes: 1 and 5, no trypsin; 2, trypsin and STI; 3 and 6, trypsin alone; 4, trypsin and Triton X-100. (B) IMVs were prepared by lysing cells with the French press and were incubated at room temperature for 30 min. Lanes: 1 and 2, no trypsin; 2 and 4, trypsin alone. Samples were separated by 16% SDS–PAGE and arrows indicate trypsin cleavage products containing a FLAG tag.

antigen bands can be visible in the gel. One possible explanation for this result was protein misfolding and possible formation of inclusion bodies upon overexpression of WbaP$_{CT}$.

Thioredoxin (TrxA) is a 12-kDa protein that stabilizes membrane proteins when fused to the N-terminus of a target protein and also prevents the formation of inclusion bodies (19). Therefore, we constructed a WbaP$_{CT}$ derivative with TrxA fused to the N-terminus (TrxA-WbaP$_{CT}$) under the expression of the arabinose-inducible pBAD promoter. To facilitate further purification this construct also has a 6×His N-terminal tag. We investigated the functionality of TrxA-WbaP$_{CT}$ in the strain MSS2 and found that the addition of the N-terminal 6×His-TrxA fusion partner resulted in a protein with the ability to restore O antigen production to similar levels as in the parental strain LT2 (Fig. 3.3). We also used green fluorescent protein (GFP) fusions to determine whether improved activity of overexpressed TrxA-WbaP$_{CT}$ was associated with proper protein folding. Proper folding of GFP fused to the C-terminus of a target protein is highly influenced by the correct folding of the fusion partner such that only folded fusion protein will become fluorescent (12). Membrane proteins fused to GFP migrate as a doublet on SDS-PAGE with the faster migrating bands representing folded proteins that fluoresce in gel under ultraviolet light (UV) and the slower migrating bands representing misfolded proteins unable to fluoresce (12). Since we established that the C-terminus of WbaP is cytoplasmic and GFP can only fluoresce when properly folded in the cytosol (11), we constructed GFP fusions to the C-terminus of WbaP$_{CT}$ derivatives with and without the N-terminal TrxA fusion. Total membranes from MSS2 cells expressing the GFP fusion constructs were separated by SDS-PAGE and detected with anti-GFP antiserum (Fig. 3.4A). In addition to the misfolded and folded forms of protein detected (Fig. 3.4A, white and black arrows respectively) there were also lower molecular weight bands which likely correspond to degraded GFP. When the SDS-PAGE gel was observed under UV light a dense band representing properly folded TrxA-WbaP$_{CT}$-GFP protein and a very faint band representing the properly folded WbaP$_{CT}$-GFP protein were detected (Fig. 3.4B). These results suggest that the fusion of TrxA to WbaP$_{CT}$ improves protein folding.

Next, we used fluorescence microscopy to observe the expression of WbaP$_{CT}$-GFP and TrxAWbaP$_{CT}$-GFP in MSS2 cells. Although cells expressing the two constructs
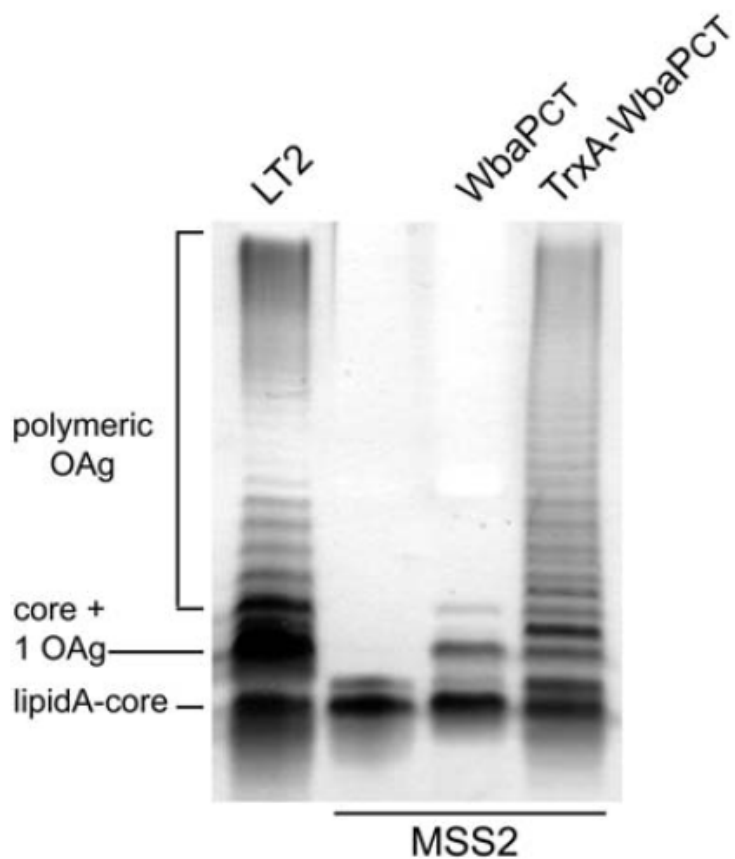
Figure 3.3   Complementation of S. Typhimurium strain MSS2 (ΔwbaP) with overexpressed WbaPCT and TrxA–WbaPCT. LPS samples were prepared and separated by 12% SDS–PAGE. Gels were stained with silver nitrate for O antigen detection. "Polymeric OAg" denotes the region of the gel containing lipid A-core OS bands covalently linked to O antigen polysaccharides of varying molecular mass. "Core + 1 OAg" indicates the band corresponding to lipid A-core OS with one covalently attached O antigen subunit.

fluoresced (Fig. 3.4C and D), dense punctuate expression around the membrane and at the poles were only seen in cells containing TrxA-WbaP$_{CT}$-GFP (Fig. 3.4D). This pattern of localization of WbaP$_{CT}$-GFP in bacterial cell bodies is similar to that observed in *E. coli* K-12 for the initiating glycosyltransferase WecA (21), and further confirms the cytoplasmic localization of the C-terminus of WbaP. Therefore, we concluded that the TrxA fusion improves membrane expression and folding of WbaP$_{CT}$, resulting also in full complementation of O antigen expression in MSS2.

3.3.3 Topological location of the TrxA fusion partner in the membrane. Although TrxA is a cytoplasmic protein, based on the topological model previously described (36), fusion of TrxA to the N-terminus of WbaP$_{CT}$ should result in localization of TrxA to the periplasm. However, rapid folding of TrxA after translation prevents it from crossing the plasma membrane (15). To analyze how TrxA-WbaP$_{CT}$-GFP is situated in the membrane, we took advantage of a tobacco etch virus (TEV) protease site engineered at the fusion endpoint between the TrxA tag and WbaP$_{CT}$. Protease cleavage experiments were carried out on spheroplasts using a 6×His-tagged TEV protease with a mass of 25 kDa, which also served as an internal control. In addition to TrxAWbaP$_{CT}$, the cytosolic FLAG-tagged transcriptional regulator from *Burkholderia cenocepacia*, BCAL0381-FLAG (6), was also expressed in the cells as a lysis control. Spheroplasts were treated with TEV protease and the periplasmic fraction was collected and analyzed by immunoblot. No cleavage was detected in the periplasmic fraction (Fig. 3.5a, lane 3), but cleavage was detectable upon treatment of lysed cells with TEV protease, resulting in a 15 kDa-cleavage product (Fig. 3.5a, lane 4). BCAL0381-FLAG was detected only upon lysis of cells (Fig. 3.5b, lane 4) and not in the periplasmic fractions (Fig. 3.5b, lanes 2 and 4). This suggests that the TrxA partner in the fusion protein is cytoplasmic and that the transmembrane helix in the fusion protein TrxA-WbaP$_{CT}$ must be "pinched in", or loosely associated with the inner membrane.

3.3.4 Functional characterization of TrxA-WbaP$_{CT}$ in vitro. We next investigated whether TrxA influences the galactosyl-phosphate transferase activity of TrxA-WbaP$_{CT}$ in vitro. Total membranes prepared from MSS2, providing both enzyme and endogenous
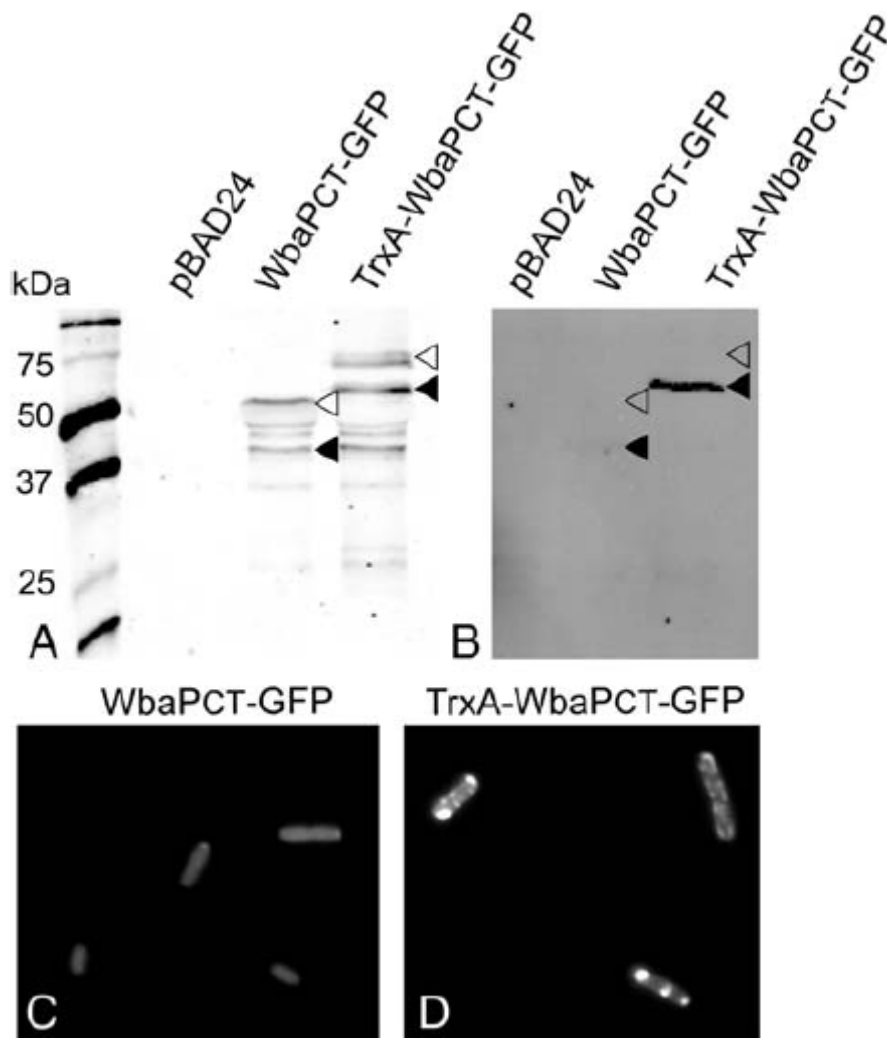
Figure 3.4 Expression and electrophoretic mobility of GFP fusion proteins. Total membranes of MSS2 (ΔwbaP) cells expressing WbaPCT–GFP and TrxA–WbaPCT–GFP were analyzed by immunoblot using anti-GFP antibodies (A) and by in-gel GFP fluorescence using UV light that provides excitation at 395 nm (B). Black and white arrows indicate the position of folded (fluorescent) and misfolded (nonfluorescent) protein, respectively. Samples were separated by 14% SDS–PAGE. Fluorescence microscopy of MSS2 (ΔwbaP) cells expressing WbaPCT–GFP (C) or TrxA–WbaPCT–GFP (D). The micrograph was taken with a ×100 oil immersion objective, resulting in a magnification of 1000-fold.
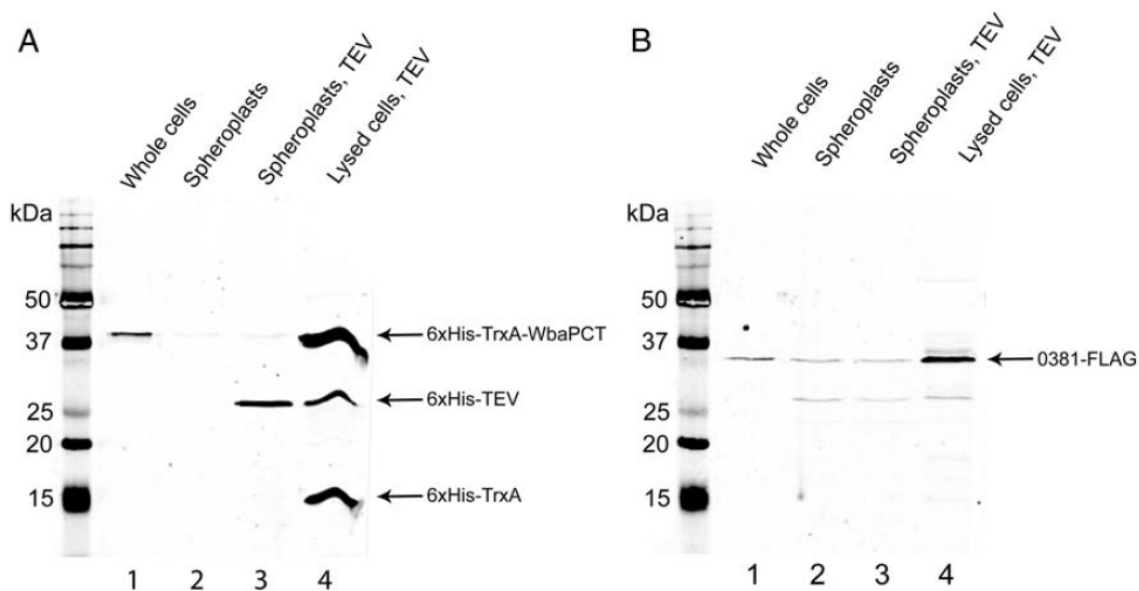
Figure 3.5 Cytoplasmic localization of TrxA in TrxA–WbaPCT. MSS2 (pKP42, p0381-FLAG) cells expressing both TrxA–WbaPCT and the FLAG-tagged cytoplasmic transcriptional regulator BCAL0381-FLAG were incubated for 30 min at 37°C (lane 1). Spheroplasts were incubated alone (lane 2), with TEV protease (lane 3), and with TEV protease after lysis (lane 4). Periplasmic fractions were collected and analyzed by immunoblot. Cleavage of TrxA-WbaPCT was detected in lysed spheroplasts treated with 6×His-tagged TEV protease resulting in a 15 kDa product. (A) Anti-His IgG2a monoclonal antibody. (B) Anti-FLAG M2 monoclonal antibody.

Und-P acceptor, were incubated with ^{14}C-labeled radioactive UDP-Galactose. The lipid fraction containing the Und-P-P-Galactose product was extracted with 1-butanol, and the radioactive counts measured. The assay was performed using 12 mM Mg^{2+} and 80 μg of total membrane protein from MSS2 bacteria expressing WbaP$_{CT}$ (encoded by pKP12), TrxA-WbaP$_{CT}$ (encoded by pKP42) or carrying the pBAD24 vector control. Total membranes expressing WbaP$_{CT}$ or TrxA-WbaP$_{CT}$ resulted in the incorporation of 1.98 ± 0.28 and 2.51 ± 0.28 pmoles of galactose per mg of total membrane protein into the butanol extractable fraction respectively, while membranes prepared from cells with vector control incorporated 0.12 ± 0.01 pmoles of galactose per mg of total membrane protein. Using the analogous sugar-phosphate transferase WecA, we have previously shown that the butanol fraction is highly enriched for undecaprenyl- P-P-sugar species (21). We also tested the activity of TrxA-WbaP$_{CT}$ under varying concentrations of Mg^{2+} and Mn^{2+} and found that the fusion protein had a metal cofactor profile similar to the parental full length protein as described previously (28) (Fig. 3.6). This demonstrates that the TrxA fusion does not disrupt the enzymatic activity of WbaP$_{CT}$.

3.3.5 Identification and *in vivo* functional analysis of conserved amino acids in TrxA-WbaP$_{CT}$. To identify residues within the C-terminus of WbaP that are required for enzyme function we first used bioinformatics comparing several hundred WbaP homologs with the Sequencing Alignment and Modeling system (SAM-T02) (17). This analysis revealed that R319, K331, R333, R377, D382, E383, R401, D446, Y449, and D458 (the numbers denote the position of these residues in the *S.* Typhimurium LT2 WbaP) were the most conserved charged or polar residues in proteins annotated as putative galactosyl-phosphate transferases (Fig. 3.7). These conserved amino acids were individually replaced by alanine in TrxA-WbaP$_{CT}$ and the ability of each mutant protein to complement LPS O antigen surface expression in the Δ*wbaP* mutant MSS2 was evaluated by silver staining. The majority of the alanine replacements abolished O antigen production (Fig. 3.8A), suggesting the mutated WbaP derivatives lose enzymatic activity (see below). In contrast, R333A and Y449A resulted in proteins that mediated partial complementation, as noticed by the formation of lipid A-core OS with one O antigen oligosaccharide unit, while the D446A mutant protein could complement full-

length O antigen (Fig. 3.8A). We also mutated two moderately conserved residues, R368 and Q386. Both mutant proteins mediated the formation of lipid A-core OS with one O antigen unit (Fig. 3.8A), suggesting that WbaP function was reduced but not completely abolished. In addition to the replacement mutants constructed above, an alanine replacement of the highly conserved glycine-393 (G393) was constructed, which resulted in a protein that mediated a full complementation of O antigen synthesis (Fig. 3.8A). All of the mutant proteins were detected in total membrane fractions. However, band shifts were seen upon SDS-PAGE analysis in mutants R401A, D446A, and D458A, as compared to relative migration of the parental TrxA-WbaPCT protein (Fig. 3.8B). This behavior could be attributed to differential binding of SDS by these proteins (32), potentially indicating changes in secondary structure.

To validate the mutagenesis strategy, six non-conserved residues along the C-terminus: isoleucine-352 (I352), valine-372 (V372), phenylalanine -388 (F388), glutamic acid-409 (E409), tryptophan-453 (W453), and leucine-462 (L462) were replaced by alanine and the mutant proteins tested for the ability to complement O antigen production. Cells expressing these WbaP mutant proteins produced full length O antigen (Fig. 3.9), supporting the notion that non-conserved amino acids do not play a role in WbaP function.

To further characterize the alanine replacement mutants with no O antigen production, additional replacements with residues conserving charge and/or structure were constructed. Thus, arginine residues were mutated to lysine and glutamine, aspartic acid residues were mutated to glutamic acid and asparagine, and glutamic acid residues were mutated to aspartic acid and asparagine. WbaP derivatives containing the conserved replacement mutations could not complement O antigen production (Fig. 3.10). In addition, we also further investigated whether a large aromatic residue is required at the position of Y449, since the Y449A mutant mediated only a partial complementation of O antigen production. The Y449S replacement mutant mediated partial complementation while the Y449F replacement led to restoration of full length O-antigen, suggesting that the aromatic group and not the hydroxyl group is required for full enzyme activity (Fig. 3.10). Because maintaining the same net charge with a different amino acid was not
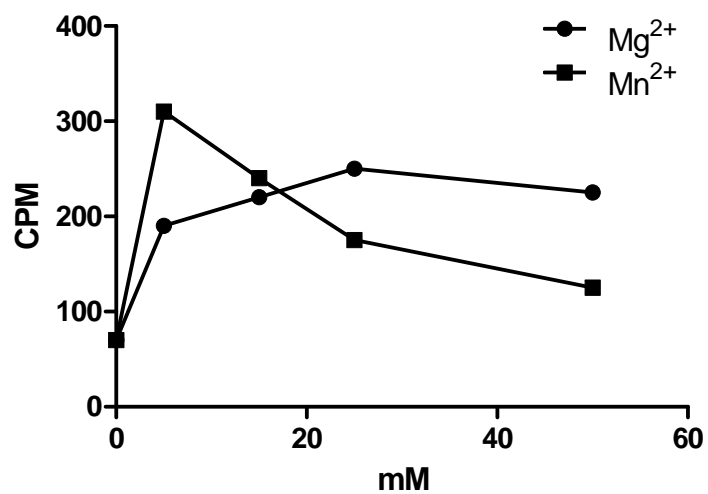
Figure 3.6   Activity of TrxA-WbaPCT under varying concentrations of divalent cations. In vitro assays were carried out as described in section 3.2.8 with 0, 5, 15, 25 or 50 mM $MgCl_2$ or $MnCl_2$. Transferase activity is expressed as counts per minute (CPM)
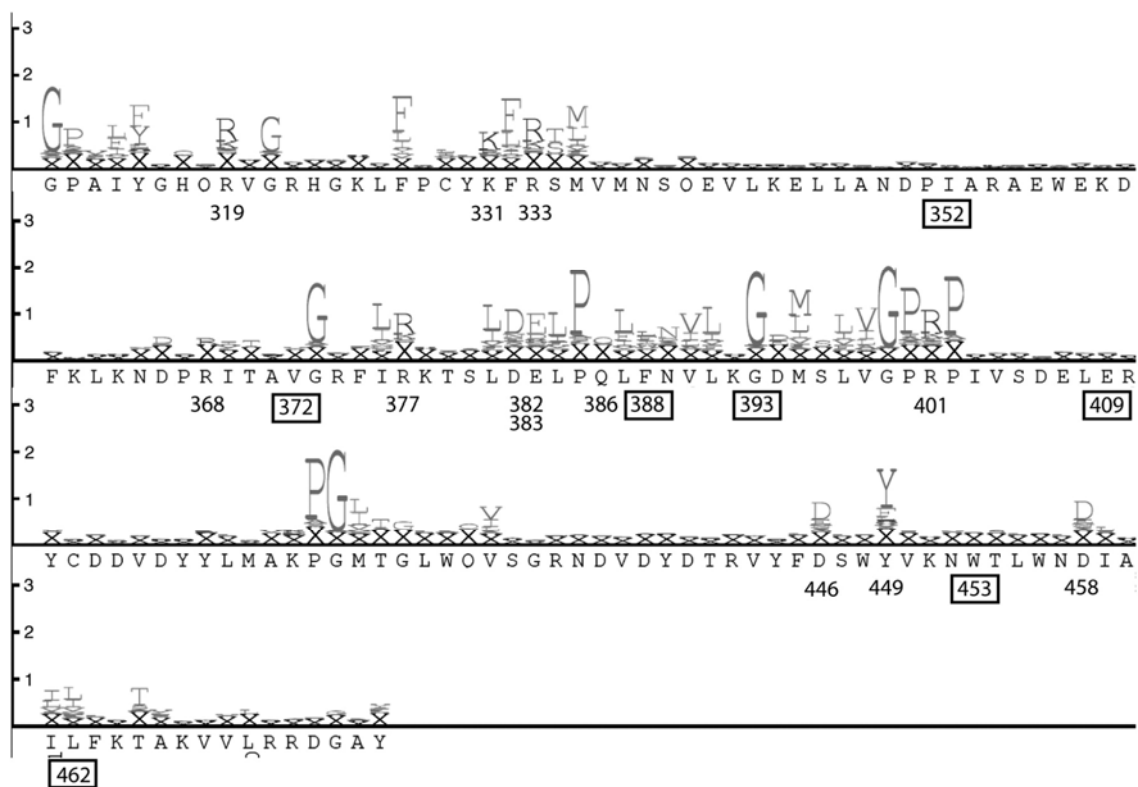
Figure 3.7 Representation of conserved amino acids in WbaP C-terminus obtained by the SAM-T02 protein structure prediction server. The server, using sequences from putative WbaP homologs, identifies conserved amino acids. The size of the letter representing the amino acid is proportional to the level of conservation of this residue on the automatic alignment. The position of residues targeted by mutagenesis is indicated below the letter. Charged and polar residues that are highly conserved (R319, K331, R333, R377, D382, E383, R401, D446, Y449, and D458) and two moderately conserved residues (R368, Q386) were mutated as well as a highly conserved glycine (G393). Numbers in boxes indicate the positions of nonconserved residues (I352, V372, F388, E409, W453, and L462) chosen for replacement mutagenesis.
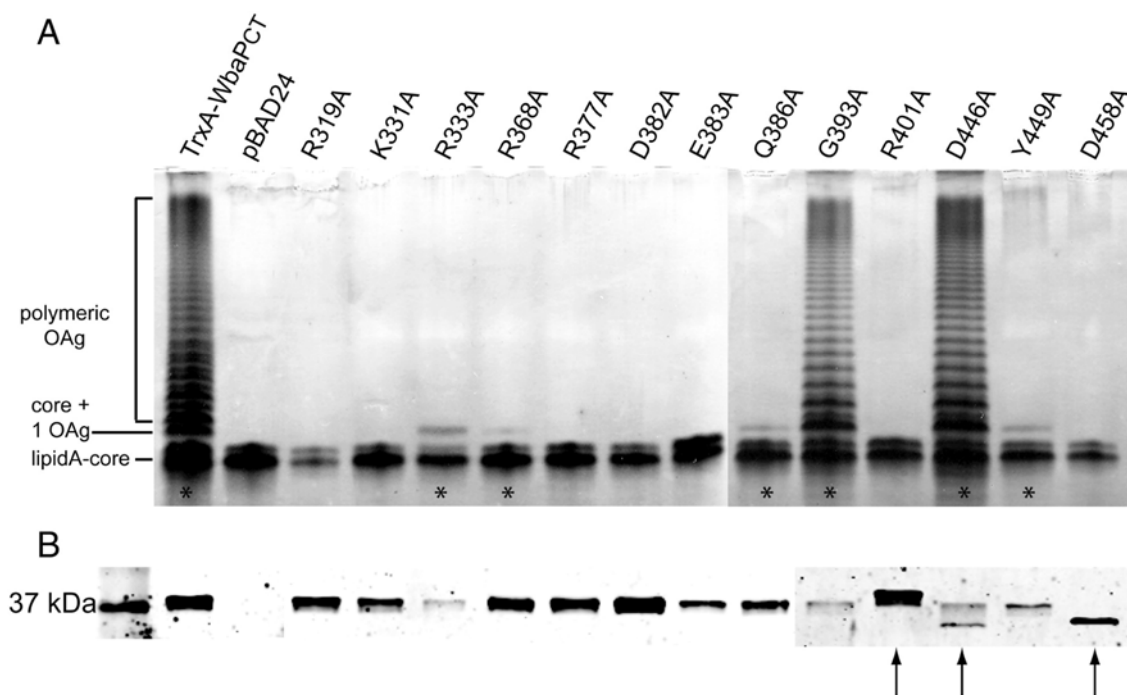
Figure 3.8 Complementation of O antigen synthesis in MSS2 (Δ*wbaP*) by plasmids encoding site-directed alanine mutants of TrxA–WbaP$_{CT}$. (A) Mutant proteins retaining the ability to partially or completely complement O antigen synthesis are marked with an asterisk (*). Only mutations G393A and D446A resulted in proteins with the ability to complement full-length O antigen. LPS samples were prepared and separated by 12% SDS–PAGE. Gels were silver stained and LPS detected. "Polymeric OAg" denotes the region of the gel containing lipid A-core OS bands covalently linked to O antigen polysaccharides of varying molecular mass. "Core + 1 OAg" indicates the band corresponding to lipid A-core OS with one covalently attached O antigen subunit. (B) Total membranes were prepared and separated by 14% SDS–PAGE. Proteins were detected with anti-His$_6$ IgG2a monoclonal antibodies. Band shifting was observed with mutants R401A, D446A, and D458A (arrows).
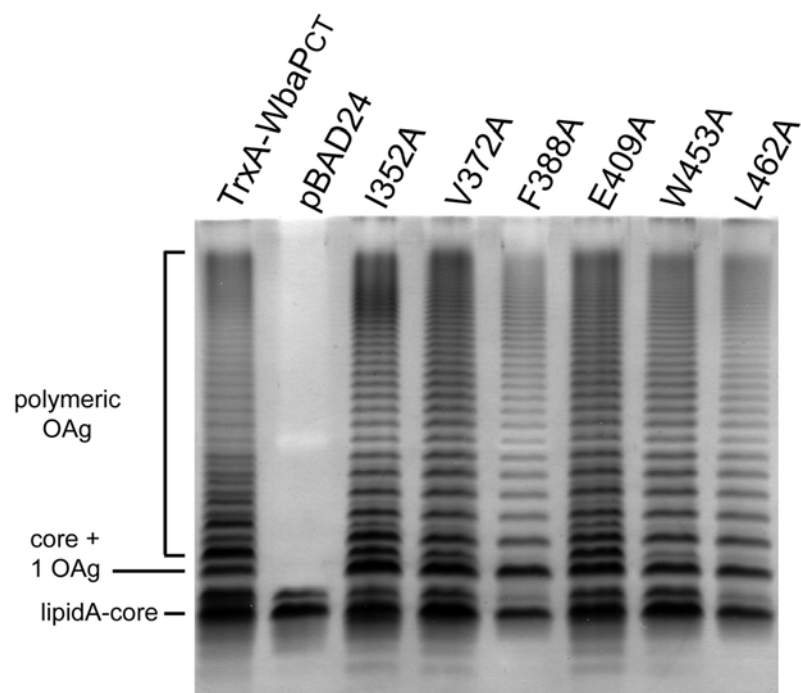
Figure 3.9 Site-directed alanine mutations in nonconserved residues along the C-terminus do not affect function. LPS was extracted and analyzed as described above.

sufficient to restore activity, we concluded that R319, K331, R377, D382, E383, R401, and D458 are absolutely critical for enzyme activity *in vivo*.

3.3.6 Functional characterization of the highly conserved charged or polar residues in the C-terminal domain of WbaP. The enzymatic activity assay, as described above for the parental TrxA-WbaPCT protein, was performed under excess of exogenous Und-P substrate to characterize in more detail the functional defects of the alanine replacement mutants. All mutant proteins were initially expressed in MSS2 (Δ*wbaP*). However, background galactosyl-phosphate transferase activity fluctuated in this strain over time and growth conditions making it difficult to assess the functionality of mutant proteins with low levels of enzymatic activity. This could not be resolved by introducing deletions in the two core galactosyltransferase genes *waaI* and *waaB* (16, 46), as well as in the putative glucosyltransferase gene for colanic acid synthesis *wcaJ* (45). Therefore, we used *E. coli* strain MV501 (*wecA*::Tn*10*), as this strain cannot produce Und-P-P-N-acetylglucosamine (GlcNAc) due to the absence of WecA (2). The galactosyltransferase WbbD acts on the GlcNAc substrate to catalyze the transfer of galactose (34). Therefore, in the absence of WecA galactose cannot be incorporated to the O antigen. Accordingly, this strain demonstrated a low and reproducible background activity in the lipid-associated fraction. To quantify the amounts of parental TrxA-WbaPCT and mutant proteins present in membrane fractions we used quantitative immunoblotting, as described previously (21, 29). For the protein standard, we purified a 6×His tagged lambda phage lytic transglycosylase (LT). An example of this analysis is shown in Fig. 3.11A. The relative intensities of the bands were determined by densitometry, using the program ImageJ, and were used to derive a standard curve to calculate the amount of TrxA-WbaPCT and its derivatives in the total membrane preparations. The mutant proteins were all well expressed in MV501 and total membrane fractions showed banding profiles that were identical to those found in MSS2 (Fig. 3.12). To determine the optimal Mg^{2+} concentration for TrxA-WbaPCT activity in MV501, total membranes were tested with various concentrations of Mg^{2+} and similar to WbaP activity in MSS2 (Fig. 3.6), 25 mM was found to be optimal and was used in the assay. All of the alanine mutants were tested for activity *in vitro*. From these, only G393A and D446A

were active (136.7 ± 16.2 and 53.9 ± 8.7 pmol/mg protein, respectively, compared to 88.7 ± 7.9 pmol/mg protein for the parental construct; Fig. 3.11B). Likewise, upon addition of 50 μM exogenous Und-P only the activity of G393A and D446A increased to 950.5 ± 79.5 and 318.5 ± 37.7 pmol/mg protein respectively, compared to 539.0 ± 125.0 pmol/mg for the parental construct (Fig. 3.11B).

There is no reported structural information on WbaP or any other members of the PHPT family, and solved structures with enough sequence homology to provide templates for protein modeling are not available. However, secondary structure predictions using PSI-PRED indicate that R377, D382, E383, and D458, are located on three putative α-helical regions. Comparisons with other members of the PHPT family indicates similar predictions for the *S. pneumonia* Cps2E, *E. amylovora* AmsG, *E. coli* K-12 WcaJ, *G. stearothermophilus* WsaP, and *C. crescentus* PssY (Fig. 3.13). This suggests these regions are structurally conserved and support the notion that they may contribute to form a putative catalytic center.

3.4   Discussion

Using a combination of protease accessibility and C-terminal GFP fusion experiments we have demonstrated that the C-terminus of WbaP, a prototypic member of the PHPT family, is cytoplasmic. This agrees with computer predictions of the topology of WbaP (36) and other members of the PHPT family, such as *G. stearothermophilus* WsaP (Steiner, K., Novotny, R., et al. 2007) and *S. pneumoniae* Cps2E (47). It was previously established that the C-terminal domain of WbaP is absolutely required for enzymatic activity, while the N-terminus of WbaP is dispensable for the galactosylphosphate transfer function (36, 44). In this study, we observed a partial complementation of O antigen synthesis when the C-terminus was overexpressed in MSS2 under the control of the pBAD promoter, resulting in an O antigen LPS banding pattern consistent with a defect in O antigen polymerization. This defect could be due to instability of the overexpressed protein in the inner membrane or the accumulation of misfolded proteins interfering with O antigen polymerization. Fusions to proteins such as glutathione *S*-transferase (GST), maltose binding protein (MBP) and thioredoxin (TrxA)
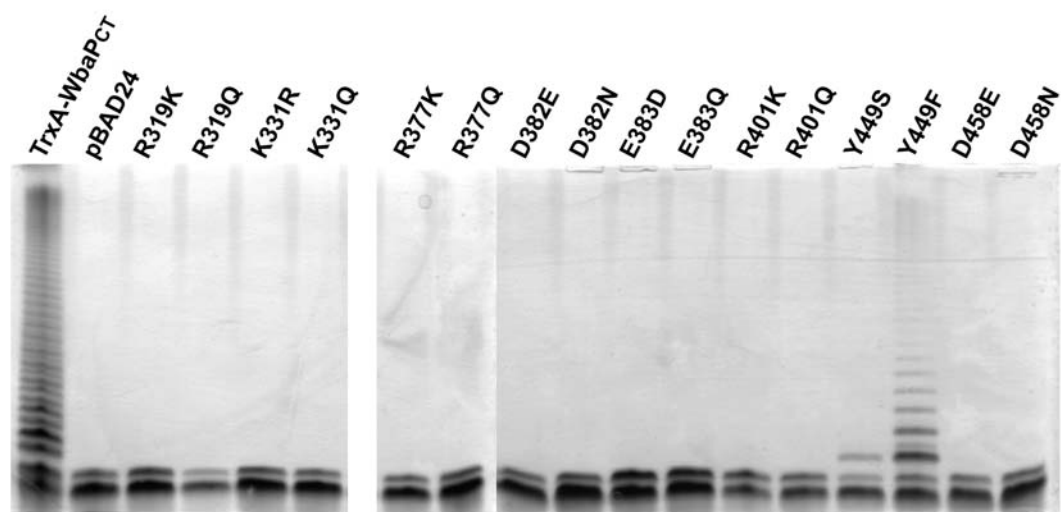
Figure 3.10   Complementation of O antigen synthesis in MSS2 (Δ*wbaP*) by plasmids encoding TrxA–WbaP$_{CT}$ site-directed mutants conserving charge and/or structure of conserved residues.
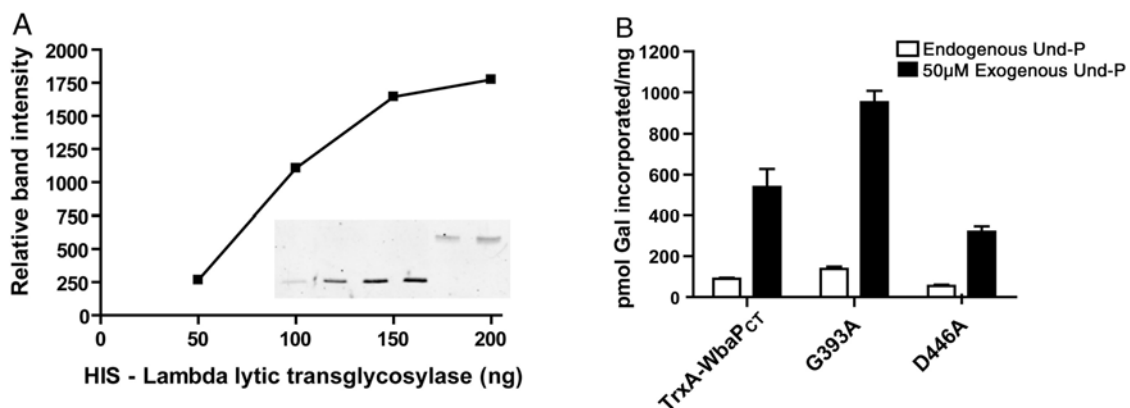
Figure 3.11 *In vitro* transferase activity of site-directed alanine mutants. (**A**) Quantitative immunoblotting to determine the amounts of TrxA–WbaP$_{CT}$ and derivative proteins in total membrane preparations. The relative amounts of 6×His molecules in the samples were determined by comparison with known amounts (50–200 ng) of purified 6×His-lambda phage lytic transglycosylase (LT), which were included on the same gel. The graph is a plot of the relative pixel densities of the 6×His–LT bands, which were calculated with the program ImageJ, versus the amounts of purified protein loaded in the lanes. The inset shows the blot with the different amounts of 6×His–LT (left four lanes) and TrxA–WbaP$_{CT}$ (right two lanes). The amount of WbaP$_{6×His–TrxA–F258–Y476}$ was deduced from a standard curve. (**B**) Enzyme assays were performed with 25 mM Mg^{2+} and ^{14}C-labeled UDP-galactose with both endogenous Und-P (white bars) and 50 μM exogenous Und-P (gray bars). The lipid phase was extracted as described in section 3.2.8. Activity is expressed as incorporation of galactose-1-P per mg of His-tagged protein at 30 min. Total membrane proteins of 40 μg and 50 μg were used for the endogenous and exogenous Und-P experiments, respectively. All assays were carried out at 37°C for 30 min in duplicate.
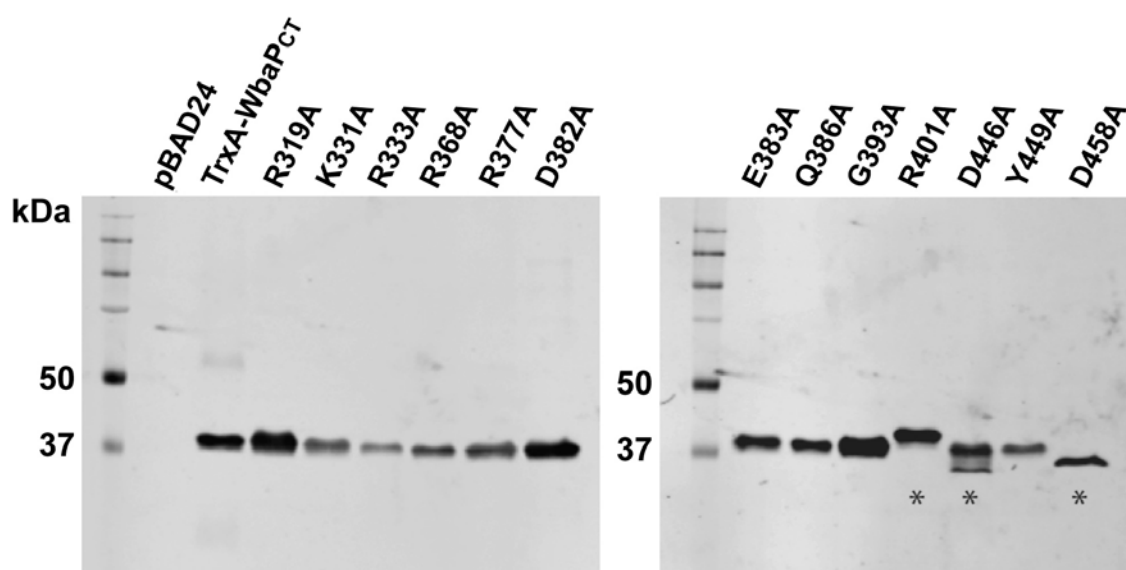
Figure 3.12 Expression of site-directed alanine mutants of TrxA–WbaP$_{CT}$ in strain MV501.

```
WbaP   DPRITAVGRFIRKTSLDELPQLFNVLKGDMSLVGPRP 402   YFDSWYVKNWTLWNDIAILFKTAKVVLRRDGAY 476
Cps2E  DPRITKIGRFIRKTSLDELPQVYNVLKGDMSLVGTRP 381   KLDVAYIDDWTIWKDIEILLKTVKVVFMRDGAK 455
AmsG   DPRITRIGHFIRKTSLDELPQLWNVVRGEMSLVGPRP 403   YFDSWYVKNWSLWNDIAILFKTIGVVLKRDGAY 477
WcaJ   DPRVTKVGNFLRRTSLDELPQFINVLTGGMSIVGPRP 382   EFDLEYIREWSVWFDIKIVFLTVFKGFVNKAAY 464
WsaP   DPRRTKFGAFLRKTSLDELPQFFNVLRGDMSIVGPRP 393   EHDLFYIENWSLWLDIKIILLTITNGLVNKNAY 471
PssY   DDRVTRVGRILRAASIDELPQLLNVLRGDMSLVGPRP 185   AADIDYIQTWSLWRDLKIVLRTVPSLLTTDNAY 267
```

Figure 3.13 ClustalW alignment and PSI-PRED secondary structure prediction of regions of WbaP and homolog proteins containing conserved residues R377, D382, E383, G393, R401, and D458. Conserved residues critical for WbaP activity in vivo and in vitro are indicated with arrowheads and G393 is marked with an asterisk. The sugar transferases compared are as follows: *S. enterica* Typhimurium WbaP (accession number NP_461027), *S. pneumoniae* Cps2E (accession number AAD10174), *E. amylovora* AmsG (accession number YP_003531611), *E. coli* K-12 WcaJ (accession number AP_002647), *G. stearothermophilus* WsaP (accession number AAR99615), and *C. crescentus* PssY (accession number AAK22153). Below the alignment PSI-PRED secondary structure predictions are depicted with α-helices shown as black boxes and β-strands as white arrows. Residues R377, D382, E383, and D458 are predicted to reside in α-helices, whereas G393 and R401 are predicted to be in unstructured regions.

can improve folding of recombinant proteins (19, 37-38). Indeed, overexpression of a TrxAWbaP_{CT} chimeric protein restored normal O antigen production by MSS2, with the same banding pattern as that observed in the parental strain. GFP fusions to WbaP_{CT} and TrxA-WbaP_{CT} allowed us to better understand the folding state of the fusion proteins. These experiments revealed that the addition of TrxA enhanced protein folding and led to a punctuate expression of WbaP_{CT} in the membrane, as detected by fluorescence microscopy. The pattern of distribution of TrxA-WbaP_{CT} in the bacterial membrane mimics that observed with the *E. coli* WecA protein (21), and suggests that this proteins could be localized in microdomains within the membrane.

The TrxA-WbaP_{CT} fusion protein included TM-V, the last predicted transmembrane helix of WbaP (Fig. 3.1). A TEV protease accessibility experiment revealed that, contrary to our expectation, the TrxA portion of the fusion protein resided in the cytoplasm or at cytosolic face of the inner membrane. We know from previous work that TM-V is required for function (36), but this helix may not be required to span the membrane to yield a functional protein as suggested by the topology of TrxA-WbaP_{CT}. Several possibilities may account for a functional TrxA-WbaP_{CT} in association to the inner membrane, but we favor the idea that the TM-V sequences form a membrane hairpin and that distortion of this helix does not affect possible interactions with the Und-P, also present in the membrane, making it possible to retain catalytic activity.

Highly conserved residues have been identified by sequence alignment within proteins of the PHPT family, such as *G. stearothermophilus* WsaP and *Caulobacter crescentus* HfsE, PssY and PssZ (39, 41). To our knowledge, this is the first report investigating these residues in detail to determine which ones play a critical role in enzyme function. Of the ten highly conserved polar or charged residues targeted, seven were important for O antigen production and only mutation D446A resulted in a protein that retain full functionality as determined *in vivo* by complementation of O antigen production and *in vitro* by a galactosyl-phosphate transfer assay. In contrast, alanine replacements in non-conserved residues resulted in functional proteins that were able to mediate production of polymeric O antigen when expressed in MSS2. While some of the identified residues resulting in non-functional proteins may be directly involved in catalysis, substrate binding, and/or metal cofactor coordination, some may also be critical

to retain the structure of a putative catalytic center. The latter is suggested by the observed band shifting of the mutant\ proteins by immunoblotting, which could be attributed to changes in secondary structure as it has been observed for other integral membrane proteins (32).

To further explore a functional role of the critical residues identified in this study, we performed conservative replacements with amino acids that would preserve or reverse polarity without major structural compromise. However, none of these replacements resulted in functional proteins. We also replaced the highly conserved G393 to alanine as an additional control, as we did not expect this residue to be important in catalysis or substrate binding. Although as expected, the mutation did not affect the ability of the protein to complement O antigen synthesis, we reproducibly observed a higher *in vitro* activity of the G393A mutant relative to that of the parental construct. This suggests that this residue may be involved in increasing protein stability or regulating the catalytic reaction. Structural information would be required to better understand the function of G393 in WbaP activity.

Since functionally important residues in the PHPT family of transferases have not been identified it is difficult to assign a role to the residues in this study. In the PNPT family, important residues have been investigated in the two well-characterized *E. coli* members, WecA and MraY. In both proteins a pair of highly conserved sequential aspartic acid residues, D90/D91 in WecA (21) and D115/D116 in MraY (23), are thought to be involved in coordinating Mg^{2+} or Mn^{2+}. Although WecA and WbaP belong to distinct families of sugar transferases, they require divalent cations Mg^{2+} or Mn^{2+} and perform the same biochemical reaction using a UDP-sugar and Und-P, which results in an Und-P-P-sugar product joined by a phosphoanhydride bond formation. Like WecA, WbaP contains two sequential, highly conserved, negatively charged residues, D382/E383, which could be involved in Mg^{2+} coordination.

In summary, we have demonstrated that C-terminal tail of *S. enterica* WbaP is cytoplasmic and have identified critical amino acids in this domain, laying groundwork for further investigation of PHPT proteins and the elucidation of their catalytic mechanism. The ability to stably express the catalytic domain of WbaP will also aid in

studies leading to obtain structural information, which will ultimately facilitate a better understand the mechanism of galactosylphosphate transfer by this enzyme.

3.5   Chapter three references

1.   **Abramoff, M. D., P. J. Magelhaes, and S. J. Ram. .** 2004. Image processing with ImageJ. Biophotonics International **11:**36-42.

2.   **Alexander, D. C., and M. A. Valvano.** 1994. Role of the *rfe* gene in the biosynthesis of the *Escherichia coli* O7-specific lipopolysaccharide and other O-specific polysaccharides containing *N*-acetylglucosamine. J Bacteriol **176:**7079-84.

3.   **Amer, A. O., and M. A. Valvano.** 2002. 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:**571-82.

4.   **Amer, A. O., and M. A. Valvano.** 2000. The N-terminal region of the *Escherichia coli* WecA (Rfe) protein, containing three predicted transmembrane helices, is required for function but not for membrane insertion. J Bacteriol **182:**498-503.

5.   **Arakawa, Y., R. Wacharotayankun, T. Nagatsuka, H. Ito, N. Kato, and M. Ohta.** 1995. Genomic organization of the *Klebsiella pneumoniae* cps region responsible for serotype K2 capsular polysaccharide synthesis in the virulent strain Chedid. J Bacteriol **177:**1788-96.

6.   **Aubert, D. F., R. S. Flannagan, and M. A. Valvano.** 2008. A novel sensor kinase-response regulator hybrid controls biofilm formation and type VI secretion system activity in *Burkholderia cenocepacia*. Infect Immun **76:**1979-91.

7.   **Cartee, R. T., W. T. Forsee, M. H. Bender, K. D. Ambrose, and J. Yother.** 2005. CpsE from type 2 *Streptococcus pneumoniae* catalyzes the reversible addition of glucose-1-phosphate to a polyprenyl phosphate acceptor, initiating type 2 capsule repeat unit formation. J Bacteriol **187:**7425-33.

8.   **Dai, K., Y. Xu, and J. Lutkenhaus.** 1996. Topological characterization of the essential *Escherichia coli* cell division protein FtsN. J Bacteriol **178:**1328-34.

9.   **Drew, D., M. Lerch, E. Kunji, D. J. Slotboom, and J. W. de Gier.** 2006. Optimization of membrane protein overexpression and purification using GFP fusions. Nat Methods **3:**303-13.

10.   **Drummelsmith, J., and C. Whitfield.** 1999. Gene products required for surface expression of the capsular form of the group 1 K antigen in *Escherichia coli* (O9a:K30). Mol Microbiol **31:**1321-32.

11.   **Feilmeier, B. J., G. Iseminger, D. Schroeder, H. Webber, and G. J. Phillips.** 2000. Green fluorescent protein functions as a reporter for protein localization in *Escherichia coli*. J Bacteriol **182:**4068-76.

12. **Geertsma, E. R., M. Groeneveld, D. J. Slotboom, and B. Poolman.** 2008. Quality control of overexpressed membrane proteins. Proc Natl Acad Sci U S A **105:**5722-7.

13. **Guzman, L. M., D. Belin, M. J. Carson, and J. Beckwith.** 1995. Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter. J Bacteriol **177:**4121-30.

14. **Hanahan, D.** 1983. Studies on transformation of *Escherichia coli* with plasmids. J Mol Biol **166:**557-80.

15. **Huber, D., M. I. Cha, L. Debarbieux, A. G. Planson, N. Cruz, G. Lopez, M. L. Tasayco, A. Chaffotte, and J. Beckwith.** 2005. A selection for mutants that interfere with folding of *Escherichia coli* thioredoxin-1 in vivo. Proc Natl Acad Sci U S A **102:**18872-7.

16. **Kadam, S. K., A. Rehemtulla, and K. E. Sanderson.** 1985. Cloning of *rfaG, B, I*, and *J* genes for glycosyltransferase enzymes for synthesis of the lipopolysaccharide core of *Salmonella* typhimurium. J Bacteriol **161:**277-84.

17. **Karchin, R., M. Cline, Y. Mandel-Gutfreund, and K. Karplus.** 2003. Hidden Markov models that use predicted local structure for fold recognition: alphabets of backbone geometry. Proteins **51:**504-14.

18. **Katzen, F., D. U. Ferreiro, C. G. Oddo, M. V. Ielmini, A. Becker, A. Puhler, and L. Ielpi.** 1998. *Xanthomonas campestris* pv. campestris gum mutants: effects on xanthan biosynthesis and plant virulence. J Bacteriol **180:**1607-17.

19. **LaVallie, E. R., E. A. DiBlasio, S. Kovacic, K. L. Grant, P. F. Schendel, and J. M. McCoy.** 1993. A thioredoxin gene fusion expression system that circumvents inclusion body formation in the *E. coli* cytoplasm. Nat. Biotechnol **11:**187-93.

20. **Lee, M. H., and C. Manoil.** 1997. Engineering trypsin-sensitive sites in a membrane transport protein. Protein Eng **10:**715-23.

21. **Lehrer, J., K. A. Vigeant, L. D. Tatar, and M. A. Valvano.** 2007. Functional characterization and membrane topology of *Escherichia coli* WecA, a sugar-phosphate transferase initiating the biosynthesis of enterobacterial common antigen and O-antigen lipopolysaccharide. J Bacteriol **189:**2618-28.

22. **Lesse, A. J., A. A. Campagnari, W. E. Bittner, and M. A. Apicella.** 1990. Increased resolution of lipopolysaccharides and lipooligosaccharides utilizing tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis. J Immunol Methods **126:**109-17.

23. **Lloyd, A. J., P. E. Brandish, A. M. Gilbey, and T. D. Bugg.** 2004. Phospho-N-acetyl-muramyl-pentapeptide translocase from *Escherichia coli*: catalytic role of conserved aspartic acid residues. J Bacteriol **186:**1747-57.

24. **Marolda, C. L., J. Vicarioli, and M. A. Valvano.** 2004. Wzx proteins involved in biosynthesis of O antigen function in association with the first sugar of the O-specific lipopolysaccharide subunit. Microbiology **150:**4095-105.

25. **Marolda, C. L., J. Welsh, L. Dafoe, and M. A. Valvano.** 1990. Genetic analysis of the O7-polysaccharide biosynthesis region from the *Escherichia coli* O7:K1 strain VW187. J Bacteriol **172:**3590-9.

26. **Morona, R., L. Van Den Bosch, and C. Daniels.** 2000. Evaluation of Wzz/MPA1/MPA2 proteins based on the presence of coiled-coil regions. Microbiology **146 (Pt 1):**1-4.

27. **Olsen, J. V., S. E. Ong, and M. Mann.** 2004. Trypsin cleaves exclusively C-terminal to arginine and lysine residues. Mol Cell Proteomics **3:**608-14.

28. **Osborn, M. J., S. M. Rosen, L. Rothfield, and B. L. Horecker.** 1962. Biosynthesis of bacterial lipopolysaccharide. I. Enzymatic incorporation of galactose in a mutant strain of *Salmonella*. Proc Natl Acad Sci U S A **48:**1831-8.

29. **Pedersen, L. B., E. R. Angert, and P. Setlow.** 1999. Septal localization of penicillin-binding protein 1 in *Bacillus subtilis*. J Bacteriol **181:**3201-11.

30. **Perez, J. M., M. A. McGarry, C. L. Marolda, and M. A. Valvano.** 2008. Functional analysis of the large periplasmic loop of the *Escherichia coli* K-12 WaaL O-antigen ligase. Mol Microbiol **70:**1424-40.

31. **Raetz, C. R., and C. Whitfield.** 2002. Lipopolysaccharide endotoxins. Annu Rev Biochem **71:**635-700.

32. **Rath, A., M. Glibowicka, V. G. Nadeau, G. Chen, and C. M. Deber.** 2009. Detergent binding explains anomalous SDS-PAGE migration of membrane proteins. Proc Natl Acad Sci U S A **106:**1760-5.

33. **Reeves, P.** 1993. Evolution of *Salmonella* O antigen variation by interspecific gene transfer on a large scale. Trends Genet **9:**17-22.

34. **Riley, J. G., M. Menggad, P. J. Montoya-Peleaz, W. A. Szarek, C. L. Marolda, M. A. Valvano, J. S. Schutzbach, and I. Brockhausen.** 2005. The *wbbD* gene of *E. coli* strain VW187 (O7:K1) encodes a UDP-Gal: GlcNAc{alpha}-pyrophosphate-R {beta}1,3-galactosyltransferase involved in the biosynthesis of O7-specific lipopolysaccharide. Glycobiology **15:**605-13.

35. **Ruiz, N., D. Kahne, and T. J. Silhavy.** 2009. Transport of lipopolysaccharide across the cell envelope: the long road of discovery. Nat Rev Microbiol **7:**677-83.

36. **Saldías, M. S., K. Patel, C. L. Marolda, M. Bittner, I. Contreras, and M. A. Valvano.** 2008. Distinct functional domains of the *Salmonella enterica* WbaP transferase that is involved in the initiation reaction for synthesis of the O antigen subunit. Microbiology **154:**440-53.

37. **Smith, D. B., and K. S. Johnson.** 1988. Single-step purification of polypeptides expressed in *Escherichia coli* as fusions with glutathione S-transferase. Gene **67:**31-40.

38. **Sonezaki, S., A. Kondo, T. Oba, Y. Ishii, Y. Kato, and H. Nakayama.** 1994. Overproduction and purification of Lon protease from *Escherichia coli* using a maltose-binding protein fusion system. Appl Microbiol Biotechnol **42:**313-8.

39. **Steiner, K., R. Novotny, K. Patel, E. Vinogradov, C. Whitfield, M. A. Valvano, P. Messner, and C. Schaffer.** 2007. Functional characterization of the initiation enzyme of S-layer glycoprotein glycan biosynthesis in *Geobacillus stearothermophilus* NRS 2004/3a. J Bacteriol **189:**2590-8.

40. **Stevenson, G., K. Andrianopoulos, M. Hobbs, and P. R. Reeves.** 1996. Organization of the *Escherichia coli* K-12 gene cluster responsible for production of the extracellular polysaccharide colanic acid. J Bacteriol **178:**4885-93.

41. **Toh, E., H. D. Kurtz, Jr., and Y. V. Brun.** 2008. Characterization of the *Caulobacter crescentus* holdfast polysaccharide biosynthesis pathway reveals significant redundancy in the initiating glycosyltransferase and polymerase steps. J Bacteriol **190:**7219-31.

42. **Tsai, C. M., and C. E. Frasch.** 1982. A sensitive silver stain for detecting lipopolysaccharides in polyacrylamide gels. Anal Biochem **119:**115-9.

43. **Valvano, M. A.** 2003. Export of O-specific lipopolysaccharide. Front Biosci **8:**s452-71.

44. **Wang, L., D. Liu, and P. R. Reeves.** 1996. C-terminal half of *Salmonella enterica* WbaP (RfbP) is the galactosyl-1-phosphate transferase domain catalyzing the first step of O-antigen synthesis. J Bacteriol **178:**2598-604.

45. **White, A. P., D. L. Gibson, S. K. Collinson, P. A. Banser, and W. W. Kay.** 2003. Extracellular polysaccharides associated with thin aggregative fimbriae of *Salmonella enterica* serovar enteritidis. J Bacteriol **185:**5398-407.

46. **Wollin, R., E. S. Creeger, L. I. Rothfield, B. A. Stocker, and A. A. Lindberg.** 1983. *Salmonella* typhimurium mutants defective in UDP-D-galactose:lipopolysaccharide alpha 1,6-D-galactosyltransferase. Structural, immunochemical, and enzymologic studies of *rfaB* mutants. J Biol Chem **258:**3769-74.

47.   **Xayarath, B., and J. Yother.** 2007. Mutations blocking side chain assembly, polymerization, or transport of a Wzy-dependent *Streptococcus pneumoniae* capsule are lethal in the absence of suppressor mutations and can affect polymer transfer to the cell wall. J Bacteriol **189:**3369-81.

Chapter 4

Characterization of UDP-Glucose:Undecaprenyl-Phosphate Glucose-1-Phosphate Transferases of *Escherichia coli* and *Caulobacter crescentus*

.

A version of this chapter has submitted for publication:

4.1. Introduction

Bacterial polysaccharides can either be attached to internal and surface components or be loosely associated to the cell envelope. In many Gram-negative bacteria, the outer membrane contains surface-exposed O antigen polysaccharide, which is a component of the lipopolysaccharide (LPS)[7] (42, 55). O antigen is attached to membrane embedded lipid A via the core oligosaccharide (OS). The O antigen influences bacterial intracellular survival (38), confers resistance to lysis by serum complement (9, 22, 35), and can also influence entry to host cells (13, 45, 59). O antigen and other cell surface polysaccharides also contribute to biofilm formation (5, 57). Polysaccharide biogenesis requires a large number of proteins that include glycosyltransferases, translocases, transporters, polymerases, and ligases. The pathway for assembly of the glycan precursors varies depending on the type of glycan and the bacterial species; however, all polysaccharides share a common initiation step requiring a uridine diphosphate (UDP) sugar and the undecaprenyl-phosphate (Und-P) lipid carrier (36, 56, 62). The initiation reaction forms a phosphoanhydride bond between the 1-phosphate residue of Und-P and the sugar-1-phosphate, resulting in the release of uridine monophosphate (UMP). This reaction is catalyzed by membrane proteins belonging to either the polyisoprenyl-phosphate *N*-acetylaminosugar-1-phosphate transferase (PNPT) or polyisoprenyl-phosphate hexose-1-phosphate transferase (PHPT) family (56). The PNPT family also includes eukaryotic members that initiate the synthesis of glycans involved in N-protein glycosylation. WecA is the prototypic prokaryotic member of the PNPT family that initiates O antigen synthesis in enteric bacteria such as *Escherichia coli* (1, 43, 65), *Klebsiella pneumoniae* O1 (8) and *Shigella dysenteriae* (24). WecA also initiates the synthesis of enterobacterial common antigen (ECA) (32). In contrast, the PHPT family contains only prokaryotic members. One example is WbaP, which initiates O antigen synthesis in *Salmonella enterica* (39, 46, 58) and EPS synthesis in *E. coli* O9a:K30 (12).

E. coli K-12 WcaJ and *Caulobacter crescentus* HfsE, PssY and PssZ, have been assigned to the PHPT family based on amino acid sequence. WcaJ is predicted to initiate the synthesis of colanic acid (CA) by transferring glucose-1-phosphate (Glc-1-P) to Und-P (50). Unlike O antigen, which is attached to the cell surface via lipid A-core OS, CA has no identified lipid anchor, and is loosely associated with the outer membrane. CA is

produced by various species of enteric bacteria and contributes to biofilm formation (51). Production of CA is regulated by a complex system of regulatory proteins that control the transcription of the *cps* (capsular polysaccharide synthesis) gene cluster (2, 18). In the Gram-negative aquatic bacterium *Caulobacter crescentus,* HfsE, PssY and PssZ are predicted to initiate the synthesis of glycans required for the production of the holdfast adhesin (52). During cell division, *C. crescentus* produces a motile swarmer cell containing a flagella and pili at the same pole, and a non-motile cell which adheres to surfaces via a stalk tipped with the holdfast adhesin (28). The swarmer cell retracts its pili, sheds the flagella, and produces a holdfast and stalk (7). Although the exact composition and structure of holdfast is unknown, previous work has determined that it is composed of polysaccharides, proteins and uronic acid (34, 54). To date, the only sugars identified in the holdfast are polymers of β-1-4-linked *N*-acetylglucosamine (GlcNAc) based on its sensitivity to lysozyme and chitinase, both of which cleave β-1,4 linkages in GlcNAc polymers (34), and its specific binding to the lectin wheat germ agglutinin, which recognizes GlcNAc polymers (34).

In this study, we investigated the function of WcaJ, HfsE, PssY and PssZ, and determined that both WcaJ and PssY act as UDP-Glc:Und-P Glc-1-P transferases *in vitro*. *In vivo* complementation assays also demonstrated the requirement of WcaJ for CA production in *E. coli* K-12. Overexpression of WcaJ and PssY in a *wbaP* mutant of *S.* Typhimurium led to the attachment of both O antigen and CA to lipid A-core OS.

4.2. Materials and Methods

4.2.1. Bacterial strains. Table 4.1 lists the bacterial strains and plasmids used in this study. *S. enterica* and *E. coli* grew aerobically at 37 °C in Luria-Bertani (LB) medium (Difco) (10 mg/ml tryptone; 5 mg/ml yeast extract; 5 mg/ml NaCl). *C. crescentus* strains grew in peptone-yeast-extract (PYE) medium (41) at 30°C. Media were supplemented with 100 µg/ml ampicillin, 30 µg/ml chloramphenicol, 40 µg/ml kanamycin or 10µg/ml tetracycline as appropriate.

4.2.2. Molecular cloning and DNA sequencing. Plasmid DNA was isolated using the Qiagen miniprep kit (Qiagen Inc.). Digestion with restriction enzymes, ligation with T4-ligase and transformations were carried out as described (30). DNA sequences were determined at the York University Core Molecular Biology and DNA Sequencing Facility, Toronto, Ontario, Canada. PCR amplifications were carried out with *Pwo*I DNA polymerase (Roche Diagnostics). The oligonucleotides used for these experiments are listed in Table 4.2. Mutagenesis to disrupt specific chromosomal genes was performed by the method of Datsenko and Wanner (11). *E. coli* W3110 and *S.* Typhimurium MSS2 carrying pKD46 were transformed by electroporation with a PCR product that was generated using plasmid pKD4 as template and primers 4589 and 4590 to construct XBF1, and primers 5029 and 5037 to construct KBP10. Plasmid pXF1 was constructed by PCR amplification of a 1429-bp fragment with primers 4587 and 4588 with *E. coli* W3110 DNA as template. This fragment was digested with *Nco*I and *Xba*I and ligated into these sites in pBAD24. Plasmid pKP21 was constructed by PCR amplification of a 1446-bp fragment with primers 3308 and 3309 using pJD132 as template. The amplicon was digested with *Sma*I and ligated into this site in pBADNTF. pHfsE, pPssY and pPssZ were constructed by PCR amplification of 1156-bp, 832-bp and 595-bp fragments using *C. crescentus* CB15 DNA and primers HfsEstartXbaI and HfsEstopHindIII; PssYstartXbaI and PssYstopHindIIIl; and PssZstartXbaI and PssZstopHindIII respectively. The fragments were digested by *Xba*I and *Hin*dIII and ligated into these sites in pBADNTF. Plasmid pKP101 was created by PCR amplification of an 886-bp fragment with primers 4520 and 258 using pPssY DNA as a template. The fragment was digested with *Hin*dIII and cloned into this site in pEB-T7. Plasmid pKP102 was created by PCR amplification of a 649-bp fragment with primers 4522 and 258 using pPssZ as template. The fragment was digested with *Hin*dIII and cloned into this site in pEB-T7. Plasmid pKP103 was constructed by PCR amplification of a 1403-bp fragment with primers 3601 and 3602 using pKP101 as template. The fragment was digested with *Nco*I and ligated into this site and *Sma*I in pBAD24. Plasmid pKP104 was constructed by PCR amplification of a1166-bp fragment with primers 3601 and 3602 using pKP102 as template. The fragment was digested with *Nco*I and ligated into this site and *Sma*I in pBAD24. Plasmid pKP113 was constructed by PCR amplification of a 1476-bp fragment

using primers 4877 and 258 using pXF1 as template. This fragment was digested with *Hin*dIII and ligated into this site in pEB-T7. Plasmid pKP114 was constructed by PCR amplification of a 822-bp fragment using primers 4508 and 258 using pKP12 as template. This fragment was digested with *Xba*I and *Pst*I and ligated into these sites in pBluescript II SK(+). Plasmid pKP115 was constructed by PCR amplification of a 1251-bp fragment using primers 4509 and 258 using pKP42 as template. This fragment was digested with *Xba*I and *Pst*I and ligated into these sites in pBluescript II SK(+). Plasmids pFLAG-WbaPCT and pTrxA-WbaPCT were constructed by digestion of pKP114 and pKP115 respectively with *Xba*I and *Pst*I. The resulting fragments were ligated into these sites in pMR10.

4.2.3. Total membrane preparation and immunoblotting. For visualization of epitope-tagged constructs bacterial cultures grown overnight in 5 ml of LB were diluted to an initial OD_{600} of 0.2, and incubated at 37°C for 2 h until cultures reached an OD_{600} of 0.6. Arabinose or IPTG was added to a final concentration of 0.2% (w/v) or 400 μM respectively. For arabinose induction, cells were incubated for an additional 3 h until reaching an OD_{600} of approximately 0.8–1.0 and for IPTG induction cells were incubated for 5 h at 30°C. Cells were then harvested by centrifugation at $10,000 \times g$ for 10 min at 4°C. The bacterial pellet was suspended in Tris/NaCl buffer (20 mM Tris-HCl [pH 8.5] + 300 mM NaCl) and protease inhibitor cocktail (Roche diagnostics) and passed through a French Press cell. Cell debris were removed by centrifugation ($15,000 \times g$ for 15 min at 4°C), and the clear supernatant was centrifuged at $40,000 \times g$ for 30 min at 4°C. The pellet, containing total membranes, was suspended in Tris/NaCl buffer. The protein concentration was determined by the Bradford assay (Bio-Rad, Hercules, California). SDS–PAGE, protein transfers to nitrocellulose membranes, and immunoblots were performed as described (40). For detection of 6×His- and FLAG-tagged proteins membranes were incubated with a 1:10,000 dilution of anti-His IgG2a monoclonal antibodies (Amersham) and anti-FLAG monoclonal antibodies (Sigma).

Table 4.1. Bacterial strains and plasmids

Strain or Plasmid	Relevant properties*	Source or Reference
Strains		
C. crescentus		
CB15	Wild-type	(41)
YB4997	CB15, $\Delta pssY$ $\Delta pssZ\Delta hfsE$	(52)
NA1000	*syn-1000*, previously called CB15N; a derivative of CB15 that does not synthesize holdfasts	(14)
E. coli		
DH5a	F⁻ ϕ80*lacZ*M15 *endA recA hsdR*(r⁻$_K$ m⁻$_K$) *nupG thi glnV deoR gyrA relA* Δ(*lacZYA-argF*)*U169*	Laboratory stock
W3110	F⁻, λ⁻, IN(rrnD⁻rrnE)1, rph-1	(4)
W3110*rcsA*⁺	W3110 containing pWQ499	This study
XBF1	W3110, $\Delta wcaJ::aph$, Kn^R	This study
XBF1*rcsA*⁺	XBF1 containing pWQ499	This study
XBF1*wbaP*$_{CT}$⁺	XBF1 containing pKP12	This study
MV501	VW187; *wecA*::Tn*10* TcR	(1)
C43(DE3)	F⁻*ompT hsdSB* (r$_{\bar{B}}$ m$_{\bar{B}}$) *gal dcm* (DE3) C43	(3)
Salmonella		
LT2	*S. enterica* serovar Typhimurium, wild-type	S. Maloy
MSS2	LT2, $\Delta wbaP::cat$, Cm^R	(46)
MSS2*wbaP*$_{CT}$⁺	MSS2 containing pKP12	This study
MSS2*wcaJ*⁺	MSS2 containing pXF1	This study
KBP10	MSS2, $\Delta wcaA::aph$, Kn^R	This study
KBP10*rcsA*⁺	KBP10 containing pWQ499	This study
Plasmids		
pBAD24	Cloning vector inducible by arabinose, Amp^R	(19)

pBADNTF	pBAD24 for N-terminal FLAG fusions	(Marolda, Vicarioli et al. 2004)
pEB-T7	Cloning vector, IPTG-inducible, for N-terminal 6×His fusions, Amp^R	E. H. Ball
pFLAG-WbaPCT	pMR10 derivative encoding the *S. enterica* Ty2 WbaP$_{F258-Y476}$ N-terminally fused to the FLAG epitope	This study
pHfsE	pBADNTF derivative encoding *C. crescentus* CB15 HfsE N-terminally fused to the FLAG epitope	This study
pJD132	pBluescript SK derivative containing *wbaP* and flanking sequences from *E.coli* O9a:K30, Amp^R	(48)
pKD4	*bla* FRT *aph* FRT PS1 PS2 *ori$_{R6K}$*	(11)
pKD46	*bla* P$_{BAD}$ *gam bet exo* pSC101 *ori^{ts}*	(11)
pKP12	pBADNTF derivative encoding the *S. enterica* Ty2 WbaP$_{F258-Y476}$ N-terminally fused to the FLAG epitope	(46)
pKP21	pBADNTF derivative encoding the *E. coli* O9a:K30 WbaP N-terminally fused to the FLAG epitope	This study
pKP42	pBAD24 derivative encoding *S. enterica* LT2 WbaP$_{F258-Y476}$ N-terminally fused to 6×His-TrxA	(39)
pKP101	pEB-T7 derivative encoding CB15 PssY N-terminally fused to TrxA	This study
pKP103	pBAD24 derivative encoding *C. crescentus* CB15 PssY N-terminally fused to 6×His-TrxA	This study
pKP104	pEB-T7 derivative encoding W3110 *wcaJ* N-terminally fused to TrxA	This study
pKP113	pEB-T7 derivative encoding *E.coli* CB1W3110 WcaJ N-terminally fused to 6×His-TrxA	This study
pKP114	pBluescript II SK(+) derivative encoding the *S. enterica* Ty2 WbaP$_{F258-Y476}$ N-terminally fused to the FLAG epitope, Amp^R	This study
pKP115	pBluescript II SK(+) derivative encoding encoding the *S. enterica* LT2 WbaP$_{F258-Y476}$ N-terminally fused to 6×His-TrxA	This study
pKV1	pAA26 expressing WecAFLAG tagged with His5 (WecAFLAG-5_His), Amp^R	(26)
pMR10	Mini-RK2 cloning vector; carries RK2 replication and stabilization functions	(52)

pPssY	pBADNTF derivative encoding *C. crescentus* CB15 PssY N-terminally fused to the FLAG epitope	This study
pPssZ	pBADNTF derivative encoding *C. crescentus* CB15 PssZ N-terminally fused to the FLAG epitope	This study
pTrxA-WbapCT	pMR10 derivative encoding the *S. enterica* LT2 WbaP$_{F258-Y476}$ N-terminally fused to 6×His-TrxA	This study
pWQ499	pKV102 derivative containing *rcsA$_{K30}$*, Tc^R	(63)
pXF1	pBAD24 derivative encoding W3110 WcaJ	This study

* Cm, chloramphenicol; Amp, ampicillin; Kn, kanamycin; Tc, tetracycline

Table 4.2. Primers used in this study

Name	Sequence
258	5′-GACCGCTTCTGCGTTCTGAT
3308	5′-CTCCCCGGGACGCATTTGGCAAGAAACGTTAGCTGCAGT
3309	5′-CTCCCCGGGTTAATAAGCACCATCTTTCTTAAGTACAAC
3601	5′-AGGGGAATTGTGAGCGGATA
3602	5′-CCTTTCGGGCTTTCTTAGCA
4508	5′-CGACTGCAGATGGATTACAAGGATGACGAT
4509	5′-CGACTGCAGATGGGCAGCAGCCATCATCA
4520	5′-CTAGCAAGCTTCAAAGAACAAGGTCTA
4522	5′-CTAGCAAGCTTCGATATCGTGGTCGCG
4587	5′-CATGCCATGGAGGAACAACGATGACAAAT
4588	5′-GCTCTAGACTCATATCAATATGCCGC
4589	5′-CATCGTTAATCTCTATGGTGCAACGCTTTTCAGATATCACGTGTAGGCTGGAGCTGCTTCG
4590	5′-CAGGAAAACGATTTTGATATCGAACCAGACGCTCCATTCGCATATGAATATCCTCCTTAG
4877	5′-CGCAAGCTTGACAAATCTAAAAAAGCGC
5029	5′-GCCAACCTGGAACAGGCAGCAGTTGGCTATCCGGGCGATTGTGTAGGCTGGAGCTGCTTCG
5037	5′-TACGCACGGAGAGCAACGTCAGCAACGTGCGCCAGGTCATCATATGAATATCCTCCTTA
HfsEstartXbaI	5′- CGTCAGCCTTCTAGAGGGGATCGGGGC
HfsEstopHindIII	5′- CAGAACATGAAGCTTCTAGCGCACGGC-
PssYstartXbaI	5′- CCGCTAATGTCTAGAGAAAGAACAAGGTCT
PssYstopHindIII	5′- GACCAAGACAAGCTTTCAGTAGGCGTT
PssZstartXbaI	5′- GCT CAAGCGTCTAGAGGATATCGTGGT
PssZstopHindIII	5′- CGCCACCATAAGCTTCTAGTAAGCGTTCTC

4.2.4. Surface binding assay. Polystyrene binding assays were performed as previously described (5), with the following modifications. The cells were allowed to adhere for 45 min, and 1.5 ml of 1 % crystal violet solution was used for staining.

4.2.5. Phage lysis assay. Five-mL LB cultures were started at an OD_{600} of 0.1 and grown to an OD_{600} of 0.4. Gene expression was induced with 0.2% arabinose and cultures were incubated overnight. The next day, 50 µl of the overnight cultures were added to 2 ml of soft agar (0.7%) containing 0.2% arabinose. The inoculated agar was poured onto an equal amount of solidified agar (1.6%). The plates were incubated for 3 h after which 10 µl of 10^6, 10^5 and 10^4 PFU/ml of Ffm phage was spotted on the plates and incubated overnight.

4.2.6. Small-scale LPS extraction and analysis. Overnight culture samples were adjusted to an OD_{600} of 2.0 in a final volume of 100 µl. Then, proteinase-K-digested whole-cell lysates were prepared as described elsewhere (31) and LPS was separated on 14% acrylamide gels using a Tricine/SDS buffer system (27). Gel loading was normalized so that each sample represented the same number of cells. Each well was loaded with approximately 1×10^8 cfu. Gels were silver stained by a modification of the procedure of Tsai and Frasch (53). Detection of *Salmonella* group D1 (for *S.* Typhi) and group B (for *S.* Typhimurium) O antigens was carried out by immunoblot. Group D1 and B rabbit antisera (Difco) were used at a 1:500 dilution. The reacting O antigen polysaccharides were detected by fluorescence with an Odyssey infrared imaging system (Li-cor Biosciences) using Rdye800-conjugated anti-rabbit affinity-purified secondary antibodies (Rockland).

4.2.7. LPS purification and gas chromatography/mass spectrometry (GC-MS) analysis. LPS from strain LT2 was isolated utilizing hot phenol/water extraction method (60). LPS from strains MSS2, MSS2(pXF1), MSS2(pPssY), KBP10(pXF1) and KBP10(pPssY) was extracted with 90% phenol/chloroform/light petroleum (2:5:8, v/v/v) as described by Brade *et al* (6). All samples were purified by ultracentrifugation (105 000 x g, 4 h, 4°C) and the pellet (LPS) was lyophilized and used for analysis. For

determination of neutral sugars, 500 µg of each LPS was hydrolyzed in 250 µl 2 M trifluoroacetic acid at 120°C for 2 h. Next, samples were reduced overnight with 2 mg NaBH₄ in water and peracetylated with acetic anhydride and pyridine (1:1, v/v) at 85°C for 45 min to yield volatile sugar alditol acetate derivatives. Samples were dissolved in chloroform and separated on a CP-Sil 88 column (WCOT fused silica 25 m x 0.25 mm) fitted to a Varian 3800 GC equipped with a 220-MS ion trap mass spectrometer. Samples (1 µl) were injected in split mode (split ratio 1:10) and eluted with the following oven temperature program: initial temperature of 190°C (held for 4 min following injection) raised 4°C/min to 230°C, and held for 11 min. The injection port was heated to 250°C. High purity He was used as a carrier gas. Compounds were identified by their retention time and mass spectrum, as well as comparison to authentic monosaccharide samples that had been converted into their corresponding alditol acetates. The retention times of analyzed sugars were compared with the standards glucose, galactose, mannose, fucose and rhamnose.

4.2.8. ***In vitro*** **transferase activity.** Total membranes were isolated from parental, mutant and tester strains containing vector controls or plasmids expressing proteins. The *in vitro* assay was adopted from Schäffer et al (48) with modifications. The reaction mixture contained 0.2 mg of total membrane protein, with a single substrate comprising 0.025 µCi radiolabeled ¹⁴C labeled UDP-galactose (UDP-Gal) (Amersham Biosciences), ¹⁴C labeled UDP-Glc (PerkinElmer) or ¹⁴C labeled UDP-GlcNAc (Amersham Biosciences) in 100 µl buffer (40 mM Tris-HCl, pH 8.5, 0.5 mM EDTA, 20 mM MgCl₂). After incubation at 37 ºC for 30 min, the reaction was stopped with 500 µl of chloroform-methanol (C/M=3:2). To extract the lipid phase, the mixture was shaken vigorously for 3 min and centrifuged for 2 min at 14,000 × *g*. The lower organic phase was collected and 55 µl of 40 mM MgCl₂ was added followed by 5 min of vigorous shaking. The mixture was centrifuged as before and the top phase was removed. The organic phase was washed twice with 400 µl of pure solvent upper phase (PSUP) (C/M/W/1M MgCl₂=18:294:293:1). For scintillation counting, half the organic phase was added to 5 ml of scintillation fluid. For thin-layer chromatography (TLC), half the organic phase was dried overnight, resuspended in 10 µl of chloroform-methanol (C/M=2:1) and spotted on

a Whatman silica gel (PE SIL G) plate that was pre-run in chloroform-methanol (C/M=3:2). The lipid-linked products were separated using a chloroform-methanol-water (C/M/W=64:25:1) solvent system. After drying and exposure of the plate to a Phosphorimager screen overnight, product formation was detected with a Phosphorimager (Storm 840; Amersham Biosciences) equipped with Image-Quant software.

4.3.   Results

4.3.1. WcaJ is required for CA synthesis in *E. coli* K-12 W3110. The structure information of the CA subunit (16, 33) suggests that the initiating sugar for its synthesis should be glucose (Fig. 4.1*A*), which is donated from UDP-Glc. WcaJ is the only candidate protein with features of an initiating enzyme, and is encoded by the *E. coli* K-12 *cps* cluster (50). WcaJ has a similar predicted topology (Fig. 4.2) and shares amino acid sequence similarity with the Gal-1-P transferase WbaP (Fig. 4.3), including highly conserved C-terminal residues previously shown to be required for WbaP activity (39). To confirm the role of WcaJ in CA synthesis, we deleted *wcaJ* in the *E. coli* K-12 strain W3110 and examined the *wcaJ* null strain (XBF1) for the absence of a mucoid phenotype. As the expression of genes required for CA production is poor in LB at 37°C, we took advantage of RcsA, a positive transcriptional regulator of CA (29). Therefore, pWQ499 encoding the *E. coli* O9a:K30 *rcsA* gene (63) was introduced into W3110 and XBF1(Δ*wcaJ*) strains. Overexpression of RcsA in W3110*rcsA*$^+$ and XBF1*rcsA*$^+$ resulted in mucoid and non-mucoid colonies, respectively (Fig. 4.4). XBF1*rcsA*$^+$ bacteria regained a mucoid phenotype upon introduction of pXF1, which encodes *E. coli* K-12 WcaJ under the control of the arabinose-inducible P_{BAD} promoter (Table 4.3 and Fig. 4.5). These results demonstrate that WcaJ is essential for CA production in *E. coli* K-12. Because of its similarity to WbaP, it is likely that WcaJ is the initiating glucosyltransferase for CA synthesis.

4.3.2. *C. crescentus* PssY can complement CA synthesis in XBF1. The construction of XBF1 provided us with a genetic means to investigate the function of any putative

UDP-Glc:Und-P Glc-1-P transferase that could be expressed in *E. coli* K-12. Therefore, we investigated the role of *C. crescentus* HfsE, PssY and PssZ involved in holdfast synthesis. These proteins vary in size and number of transmembrane helices (Fig. 4.2), but all contain the C-terminal domain previously demonstrated to be required for catalytic activity in PHPT family members (39, 49, 58). HfsE is 512-aa in length. Like WbaP and WcaJ, HfsE also contains five predicted transmembrane helices (39, 46). In contrast, PssY and PssZ are smaller proteins of 267 and 188 aa, respectively, and they each contain one predicted transmembrane helix. To determine whether HfsE, PssY, and PssZ are localized in the membrane, we overexpressed them in *E. coli* DH5α as N-terminal FLAG fusion derivatives under the control of the P_{BAD} promoter. Upon cell lysis, the three proteins were found in membrane fractions. However PssY and PssZ were poorly detected (Fig. 4.6A).

To investigate whether these proteins could complement CA production and hence utilize UDP-Glc, we introduced each of the corresponding recombinant plasmids into XBF1*rcsA*⁺ and the transformants were examined on plates containing no arabinose or with 0.2% arabinose. Although HfsE, PssY and PssZ are redundant for holdfast production in *C. crescentus* (52), only PssY could restore a mucoid phenotype to the XBF1*rcsA*⁺ strain (Table 4.3, and Fig. 4.5). To ensure the lack of complementation by PssZ was not due to low protein levels, we increased membrane expression of this protein by fusing the thioredoxin (TrxA) protein (12 kDa) to the N-terminus of PssZ. We have previously demonstrated that an N-terminal TrxA fusion to the C-terminal domain of WbaP increases membrane expression and activity of the construct (39). TrxA was fused to the N-terminus of PssY and PssZ and tested in XBF1*rcsA*⁺. Although the TrxA fusion increased membrane expression of PssZ (Fig. 4.6*B*), it did not restore CA production (Table 4.3). Together, these data suggest that PssY is an integral membrane Glc-1-P transferase. Since HfsE and PssZ could not restore WcaJ activity in the CA production assay, we reasoned that these proteins utilize UDP-Gal, like the *S.* Typhimurium WbaP or a different nucleotide sugar.

4.3.3. Overexpression of PssY and WcaJ in *S. enterica wbaP*-deficient mutant MSS2 results in the production of O antigen and CA$_{LPS}$. To determine whether HfsE
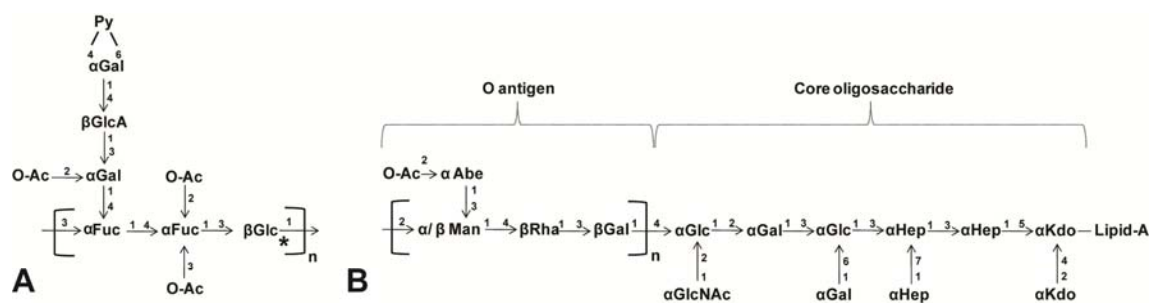
Figure 4.1. . (A) Original chemical structure of *E. coli* K-12 colanic acid (16, 50) derived from mass spectrometry was modified (33) by NMR analysis of one colanic acid subunit attached to lipid A-core OS. This structure would suggest attachment of the repeating subunit via glucose (*) at the carbon-4 position of fucose, the third sugar in the subunit. Fuc, L-fucose; Gal, D-galactose; GlcA, D-glucuronic acid; Glc, D-glucose; OAc, O-acetyl; Pyr, pyruvate. (B) Proposed chemical structure of the *S. enterica* Typhimurium O antigen and core OS. (23, 37)
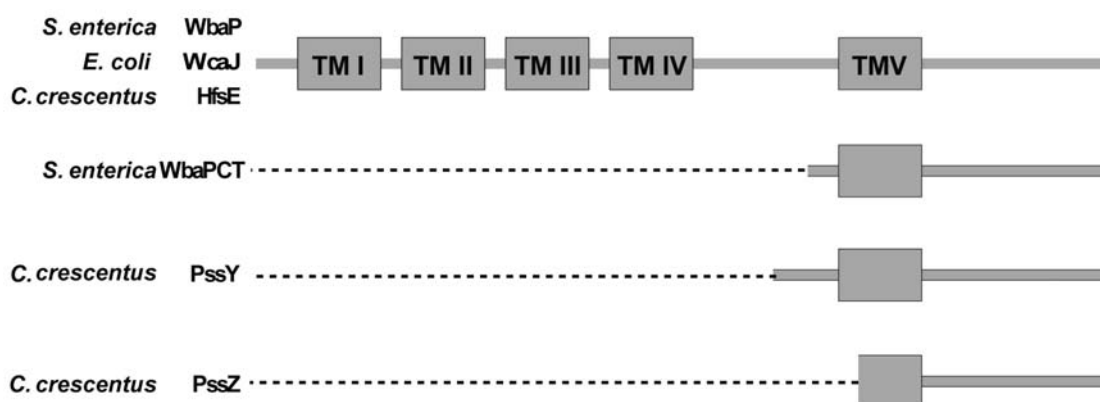
Figure 4.2. Transmembrane (TM) alignment of PHPT proteins. Graphic representation of the *S. enterica* WbaP; *E. coli* WcaJ; and *C. crescentus* HfsE, PssY and PssZ. Transmembrane regions were predicted with the servers TopPred, TMHMM and TMPRED.

```
WbaP : LYNTDMSFIFSHEVMLLRIQNNLAKRSSRFLKRTFDIVCSIMILIIASPLMIYLWYKVTRDG-GPALYGHQR : 319
HfsE : AAEDDAALVRIAAGSEAPLSGSPASTRRALLKRAQDLVVAGLGLLVAGPVMLLVALAVKIDSPGPVFFRQRR : 357
PssY : -----GPLPRVAVD---------------LRKRVLDVVAAALLTAIFAPLLLLAALAIKLESPGPALFRQTR : 117
PssZ : -----------------------------MDIVVAGGALIFFAPLLLLVAILIKLESPGPVLFRQSR :  38
WcaJ : FNILHSRLEEMNGVPVVPLYDTPLSGVNRLLKRAEDIVLATLILLIISPVLCCIALAVKLSSPGPVIFRQTR : 314

WbaP : VGRHGKLFPCYKFRSMVMNSQEVLKELLANDPIARAEWEKDFKLKNDPRITAVGRFIRKTSLDELPQLFNVL : 391
HfsE : HGFNNEAILVWKFRSMRHEMADASAARCVSA--------------DDDRVTRVGKFIRKTSLDELPQLFNVL : 415
PssY : GGLGGAPFQILKLRTMHCR-EDGPDVAQAQR--------------GDDRVTRVGRILRAASIDELPQLLNVL : 174
PssZ : GGLNGQAFTIFKFRSMRCQ-ENGAEVVQAKR--------------DDDRITTIGRFIRKTSIDELPQLLNVL :  95
WcaJ : YGMDGKPIKVWKFRSMKVM-ENDKVVTQATQ--------------NDPRVTKVGNFLRRTSLDELPQFINVL : 371

WbaP : KGDMSLVGPRPIVSDELER------YCDDVDYYLMAKPGMTGLWQVSGRND-----VDYDTRVYFDSWYVKN : 452
HfsE : TGEMSIVGPRPHAIGMKTAGVESAKLVAQYAHRHRMKPGVTGWAAIKGSRGPVDTPESVRLRVALDVEYIER : 487
PssY : RGDMSLVGPRPHATAHDDY---YSARIPEYAARYQARPGLTGLAQVRGLRGGTETVELVRQRIAADIDYIQT : 243
PssZ : RGDMSIVGPRPHALAHDEH---YGALIANYHLRFRTRPGLTGLAQIKGLRGGTSAVDAMAARIDADNEYIER : 164
WcaJ : TGGMSIVGPRPHAVAHNEQ---YRQLIEGYMLRHKVKPGITGWAQINGWRGETDTLEKVEKRVEFDLEYIRE : 440

WbaP : WTLWNDIAILFKTAKVVLRRDGAY- : 476
HfsE : QSFWLDLYIIAMTIPCLLGDRSAVR : 512
PssY : WSLWRDLKIVLRTVPSLLTTDNAY- : 267
PssZ : WTLGGDVKILLMTVPHLLMAENAY- : 188
WcaJ : WSVWFDIKIVFLTVFKGFVNKAAY- : 464
```

Figure 4.3. Comparison of conserved C-terminal region of *S.* Typhimurium LT2 WbaP (accession number NP_461027) with *E. coli* W3110 WcaJ (accession number AP_002647) and *C. crescentus* CB15 HfsE (accession number AAK24396), PssY (accession number AAK22153) and PssZ (accession number AAK24355).
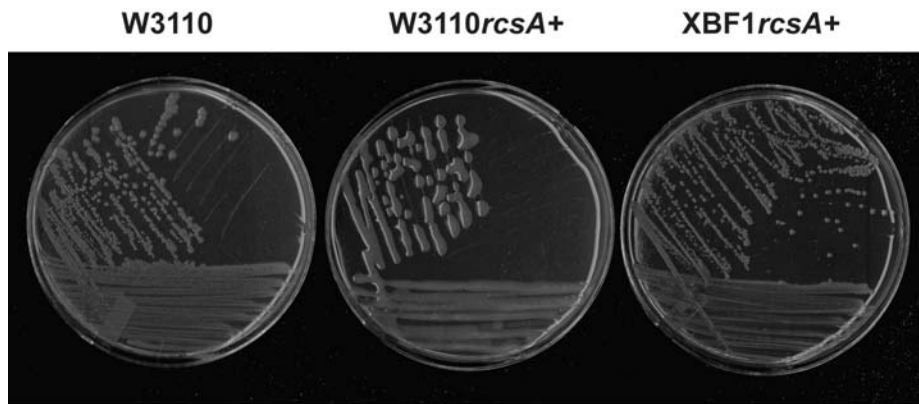
Figure 4.4. Mucoid phenotype of *E. coli* K-12 W3110 and the *wcaJ* mutant XBF1 following introduction of additional copies of *rcsA*. Cells were streaked on LB plates and incubated overnight at 37°C. Cells carrying pWQ499 were streaked on plates containing 10μg/ml tetracycline.

Table 4.3. Mucoid growth on solid media upon colanic acid production

Strain/plasmid	Plate	Mucoid phenotype
XBF1		
pBAD24	LB	Non mucoid
	LB +0.2%arabinose	Non mucoid
pHfsE	LB	Non mucoid
	LB +0.2%arabinose	Non mucoid
pJD132	LB	Non mucoid
	LB +0.2%arabinose	Non mucoid
pKP12	LB	Non mucoid
	LB +0.2%arabinose	Non mucoid
pKP42	LB	Non mucoid
	LB +0.2%arabinose	Non mucoid
pKP103	LB	Non mucoid
	LB +0.2%arabinose	Mucoid
pKP104	LB	Non mucoid
	LB +0.2%arabinose	Non mucoid
pPssY	LB	Non mucoid
	LB +0.2%arabinose	Mucoid
pPssZ	LB	Non mucoid
	LB +0.2%arabinose	Non mucoid
pXF1	LB	Non mucoid
	LB +0.2%arabinose	Mucoid
MSS2		
pWQ499	LB	Mucoid
KBP10		
pWQ499	LB	Non mucoid
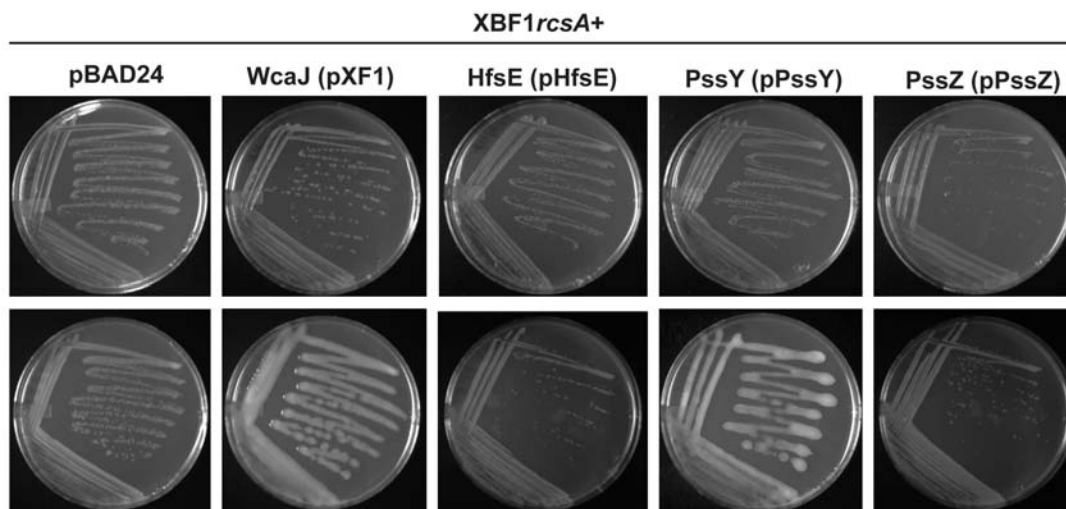
Figure 4.5. *In vivo* complementation of *wcaJ* deletion mutant in *E. coli* K-12 W3110 (XBF1) with *E. coli* WcaJ and putative glucose-1- phosphate transferases HfsE, PssY and PssZ. Cells were streaked on LB agar plates containing no arabinose (top) or induction plates containing 0.2% arabinose (bottom) and incubated overnight at 37°C. Cells carrying pBAD24 or derivatives were streaked on plates containing 100µg/ml ampicillin.

and/or PssZ are functionally similar to WbaP, we introduced plasmids expressing these proteins in the *S.* Typhimurium *wbaP* deficient strain MSS2, and examined LPS O-antigen surface expression. The three proteins were found in total membrane fractions at similar levels to those found in *E. coli* DH5α (Fig. 4.7). Only bacteria expressing PssY produced higher bands in addition to lipid A + core OS that corresponded to O antigen (Fig. 4.8, top). The partial complementation resembled the O antigen complementation of the C-terminal domain of *S.* Typhimurium WbaP (WbaP$_{CT}$) (Fig. 4.8, top). To determine whether HfsE or PssZ activity was hindered by overexpression of the proteins, we repeated the experiment with varying concentrations of arabinose (0.002-2%). Again, only the cells containing PssY produced additional bands (Fig. 4.9). To determine if the absence of PssZ activity was due to low protein expression, we expressed TrxA-PssY and TrxA-PssZ in MSS2, but TrxA-PssZ was still unable to support the *in vivo* production of bands resembling O antigen (Fig. 4.8, top).

Since PssY could initiate the synthesis of lipid A-core OS linked glycans in MSS2 *in vivo,* we reasoned that WcaJ might do the same. MSS2 cells transformed with pXF1 (expressing WcaJ) also produced bands of higher molecular mass than lipid A-core OS (Fig. 4.8, top). *S.* Typhimurium contains an endogenous copy of *wcaJ* in the chromosome. Therefore, we also tested overexpression of *S.* Typhimurium genomic *wcaJ* by introducing the plasmid encoding RcsA in MSS2. This resulted in a similar *in vivo* phenotype to that found by overexpressing PssY and WcaJ in MSS2 (Fig. 4.8, top).

The higher molecular mass silver-stained bands suggested that the lipid A-core OS was substituted with an additional polysaccharide. This was in part confirmed by utilizing Ffm, a bacteriophage that only attacks LPS mutants devoid of O antigen (64). MSS2*wbaP*$_{CT}$$^{+}$ and MSS2*wcaJ*$^{+}$, were more susceptible to the bacteriophage than LT2 and MSS2(pKP42) but less susceptible than MSS2 (Fig. 4.10). These results demonstrated that upon overexpression, both PssY and WcaJ were able to initiate the synthesis of a polysaccharide that is ligated to lipid A-core OS in MSS2.

Three candidates for this polysaccharide were ECA, O antigen, and CA. It was unlikely that the bands represented ECA since this polysaccharide depends on WecA for the initiation of its synthesis. If the polysaccharide attached to these strains was O antigen it would suggest that overexpression of PssY and WcaJ could allow them to utilize both
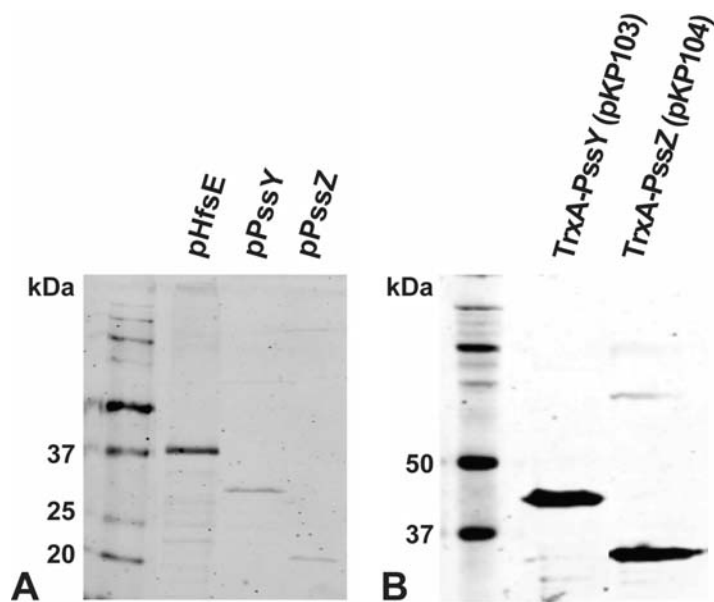
Figure 4.6. Expression of *C. crescentus* proteins in DH5α cells. Total membranes were prepared and 10μg of each sample was separated by 14% SDS-PAGE. Proteins were detected with Anti-FLAG. (A) FLAG tagged HfsE, PssY and PssZ and (B) 6×His tagged TrxA-PssY and TrxA-PssZ migrated at ≈ 40kDa, 30kDa, 22kDa, 45kDa and 34kDa respectively.
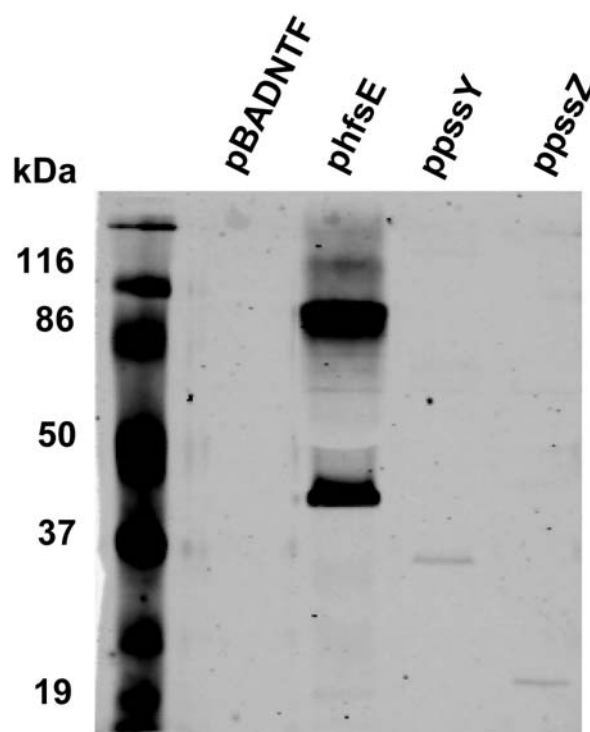
Figure 4.7. Expression of *C. crescentus* proteins in MSS2(Δ*wbaP*) cells. Total membranes were prepared and 10μg of each sample was separated by 14% SDS-PAGE. Proteins were detected with Anti-FLAG. FLAG tagged HfsE, PssY and PssZ migrated at ≈ 40kDa, 30kDa, 22kDa, 45kDa and 34kDa respectively.
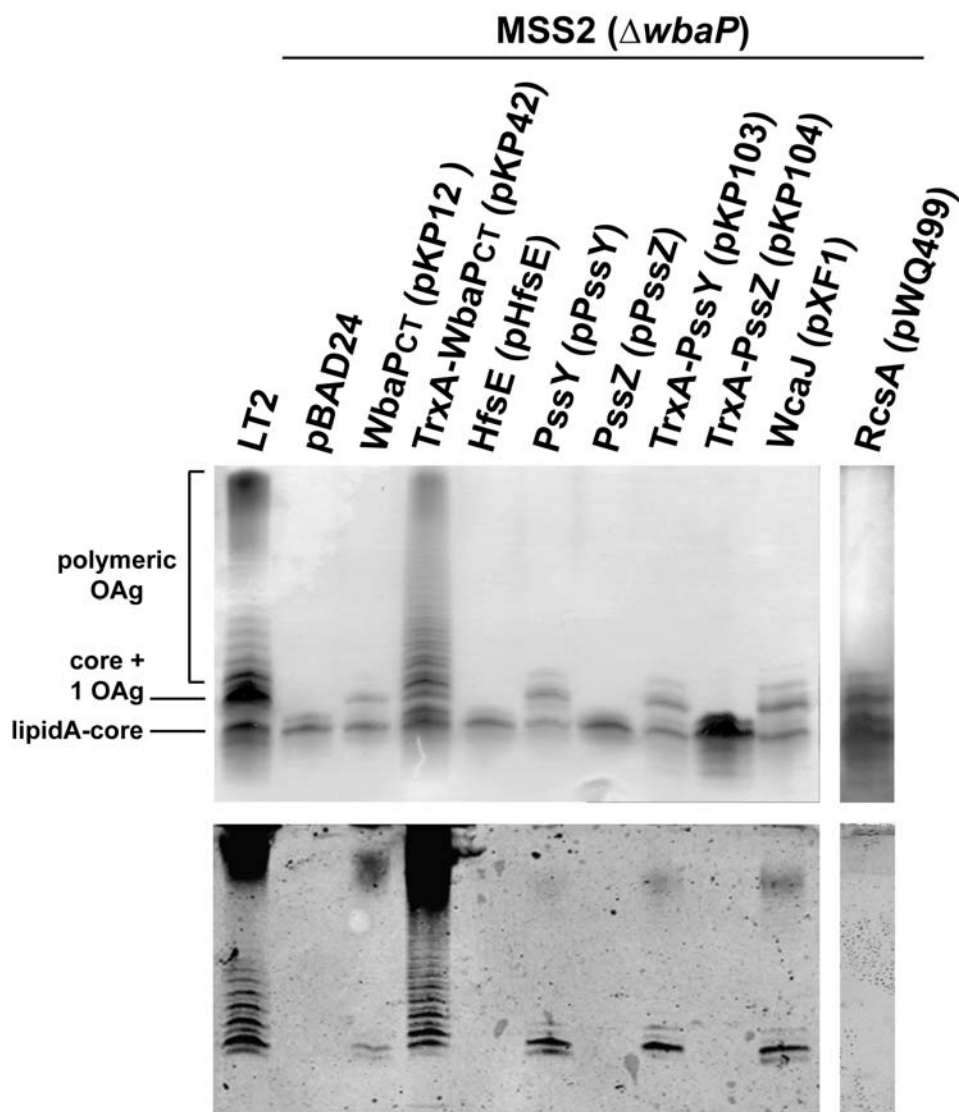
Figure 4.8. *In vivo* complementation of *wbaP* deletion mutant in *S*. Typhimurium LT2 (MSS2) with overexpressed WcaJ, HfsE, PssY, PssZ and RcsA. (Top) LPS samples were prepared and separated by 12% SDS-PAGE. Gels were stained with silver nitrate for O antigen detection. "Polymeric OAg" denotes the region of the gel containing lipid A-core OS bands covalently linked to O antigen polysaccharides of varying molecular mass. "Core + 1 OAg" indicates the band corresponding to lipid A-core OS with one covalently attached O antigen subunit. (Bottom) Immunoblot of same samples as above reacted with Group B O-specific rabbit antiserum.
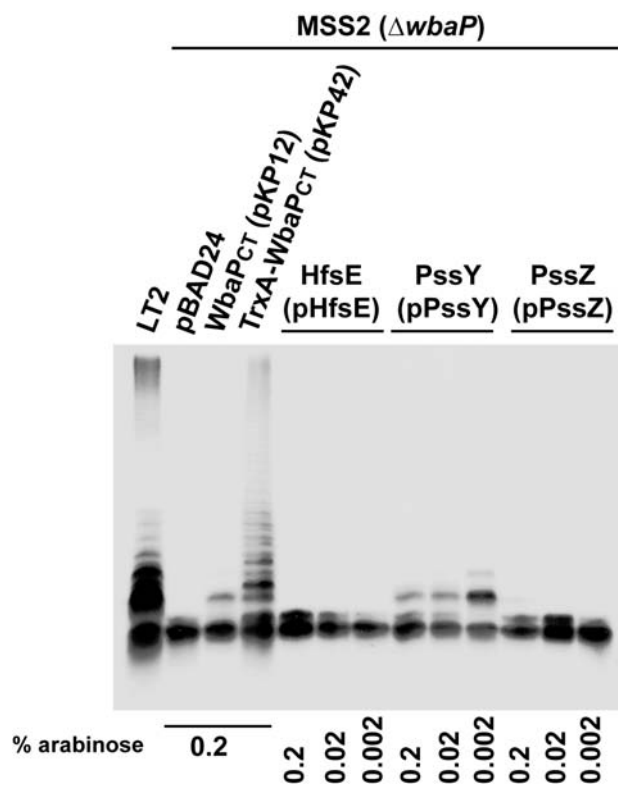
Figure 4.9. *In vivo* complementation of *wbaP* deletion mutant in *S*. Typhimurium LT2 (MSS2) with HfsE, PssY, PssZ induced at varying concentrations of arabinose. LPS samples were prepared and separated by 12% SDS-PAGE. Gels were stained with silver nitrate for O antigen detection.
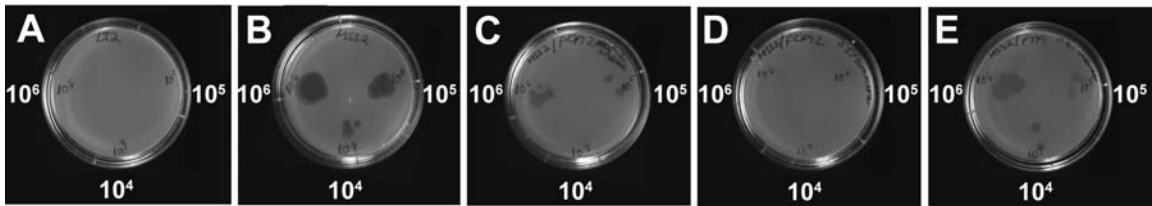
Figure 4.10. Phage Ffm lysis assay. (**A**) LT2 (**B**) MSS2 (**C**) MSS2*wbaP$_{CT}$*$^+$ (**D**) MSS2(KP42) (**E**) MSS2*wcaJ*$^+$. Cells were grown overnight in 0.2% arabinose. 50μl of culture was added to 2ml of soft agar (0.7%) and overlaid on solid agar (1.6%). Agar was dried and 10^6, 10^4 and 10^2 PFU of phage Ffm were spotted on each plate. The plates were incubated overnight at 37°C.

UDP-Glc and UDP-Gal as substrates. CA induction in *E. coli* K-12 can lead to the production of CA_{LPS}, whereby colanic acid is covalently linked to the lipid A-core OS (33). To determine if the unknown glycan was O antigen, we first performed an immunoblot with group B antiserum, which reacts with *S.* Typhimurium O antigen. As a negative control, we also probed with group D1 antiserum, which specifically reacts with *S.* Typhi O antigen. The bands produced upon overexpression of WcaJ or PssY reacted with the group B antiserum (Fig. 4.8, bottom), but not group D1 antiserum (Fig. 4.11). In contrast, the bands produced upon RcsA overexpression did not react with either group B or group D1 antisera. These results suggest that the bands produced by overexpression of WcaJ and PssY correspond to lipid A-core OS substituted with O antigen, while those produced by overexpression of RcsA correspond to CA_{LPS}.

To confirm the attachment of O antigen to lipid A-core OS, we performed GC-MS analysis of purified LPS preparations. As expected, the wild-type LT2 sample contained the sugars abequose, rhamnose and mannose, which are found in the chemical structure of the *S.* Typhimurium O antigen (Fig. 4.1B). These were also detected in the MSS2(pPssY) and MSS2*wcaJ*⁺ samples along with the unique CA sugar fucose. The MSS2 sample did not contain the diagnostic O antigen sugars or fucose. All samples contained glucose and galactose, which are found in the core oligosaccharide (37) (Fig. 4.1*B*). Together, we conclude that overexpression of WcaJ or PssY in MSS2 cause bacterial cells to produce a mixture of lipid A-core OS species with covalently attached O antigen or CA.

4.3.4. Overexpression of WcaJ and PssY in a Δ*wbaP*Δ*wcaA* strain results in the attachment of CA to lipid A-core OS. Next, we aimed to inhibit the production of CA in MSS2 by constructing a mutant with a *wcaA* deletion. *wcaA* is part of the CA *cps* cluster and encodes a predicted glycosyltransferase involved in CA synthesis, which presumably adds a sugar to the repeating CA unit (10, 25, 50, 61). The resulting strain, KBP10(Δ*wbaP*Δ*wcaA*), was examined for CA production in the presence of RcsA. In contrast to MSS2*rcsA*⁺, KBP10*rcsA*⁺ produced nonmucoid colonies (Table 4.3 and Fig. 4.12 A and *B*). Furthermore, overexpression of WcaJ and PssY in KBP10 resulted in the detection of an altered banding pattern, compared to overexpression of these proteins in

MSS2, and loss of reactivity with group B antiserum (Fig. 4.13). We interpreted the change in band migration as due to the loss of O antigen attached to lipid A-core OS concomitantly with the production of a partial CA unit covalently linked to lipid A-core OS. Supporting this conclusion, GC-MS analysis of purified LPS detected fucose but not the O antigen sugars. In the absence of WcaA, incomplete CA units would accumulate and become Und-PP-linked preventing the attachment of O antigen. If this were the case, we should observe decreased O antigen attachment in wild-type *S. enterica* upon the overexpression of WcaJ or PssY. Overexpression of WcaJ and PssY in strain LT2 demonstrated that these proteins, but not RcsA (used as a negative control), mediated less O antigen production (Fig. 4.14).

4.3.5. *In vitro* glycosyltransferase activity of PssY and WcaJ. To biochemically confirm whether WcaJ and PssY utilize UDP-Glc and/ or UDP-Gal, we assayed the *in vitro* glycosyltransferase activity of membrane extracts prepared from *E. coli* C43(DE3) overexpressing TrxA-WcaJ and TrxA-PssY. Aliquots of membrane fractions were incubated with ^{14}C labeled UDP-Gal or ^{14}C labeled UDP-Glc as substrates. As a positive control, membrane fractions from MV501 cells overexpressing WecA were incubated with ^{14}C labeled UDP-GlcNAc. Membrane fractions provided both the enzyme and endogenous Und-P acceptor. Total membranes from *E. coli* C43(DE3) containing pEB-T7 and *E. coli* MV501 containing pBAD24 incubated with the relevant ^{14}C labeled UDP sugars were used as negative controls. The lipid fractions containing the Und-P-P-sugar products were extracted with chloroform/methanol, separated by TLC and exposed on a phosphorimager screen. Membranes overexpressing TrxA-WcaJ and TrxA-PssY incubated with ^{14}C labeled UDP-Glc revealed two products with Rf values of 0.06 and 0.66 (Fig. 4.15). The faster-migrating spot corresponds to spots observed from C43(DE3) cells overexpressing TrxA-WbaP$_{CT}$ and MV501 cells overexpressing WecA. This would suggest that these spots represent Und-P-P linked sugars. Together, these assays demonstrate the formation of a detectable Und-P-P-Glc linked product upon overexpression of TrxA-WcaJ and TrxA-PssY in C43(DE3) membranes.
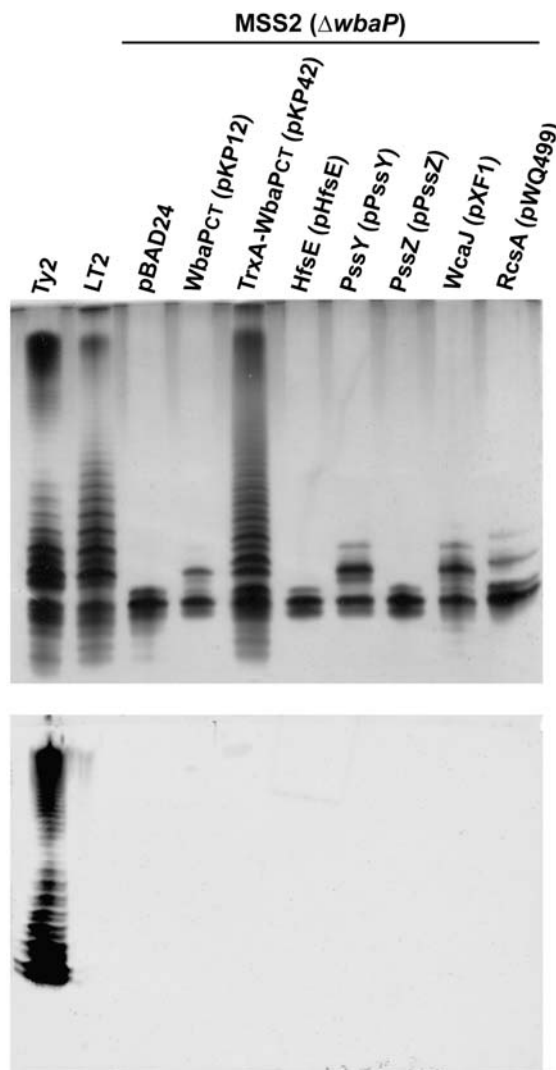
Figure 4.11. *In vivo* complementation of *wbaP* deletion mutant in *S*. Typhimurium LT2 (MSS2) with overexpressed WbaP derivatives and WcaJ, HfsE, PssY, PssZ and RcsA. (Top) LPS samples were prepared and separated by 12% SDS-PAGE. Gels were stained with silver nitrate for O antigen detection. "Polymeric OAg" denotes the region of the gel with lipid A-core OS bands covalently linked to O antigen polysaccharides of varying molecular mass. "Core + 1 OAg" indicates the band corresponding to lipid A-core OS with one covalently attached O antigen subunit. (Bottom) Immunoblot of same samples as above reacted with Group D1 specific rabbit antiserum as described in Experimental Procedures.
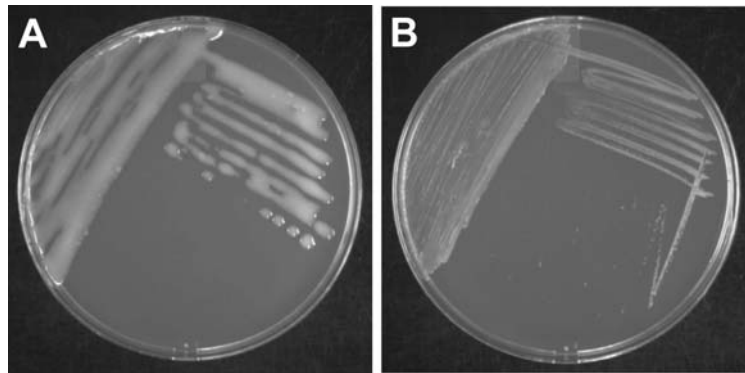
Figure 4.12. Mucoidi phenotype of Δ*wbaP*(MSS2) and Δ*wbaPΔwcaA*(KBP10) deletion mutants in *S.* Typhimurium LT2 following introduction of additional copies of *rcsA*. MSS2(pWQ499) was streaked on an LB plate containing 30μg/ml chloramphenicol and 10μg/ml tetracycline and KBP10(pWQ499) was streaked an LB plate containing and antibiotics above and 40μg/ml kanamycin. Plates were incubated overnight at 37°C.
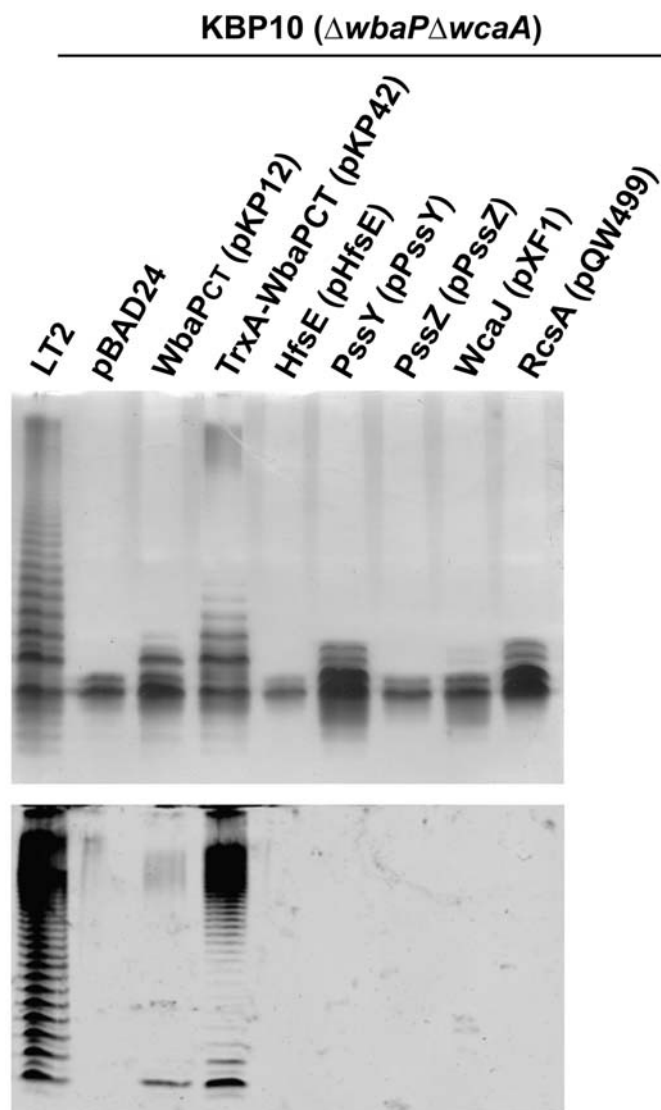
Figure 4.13. *In vivo* complementation of *wbaP* and *wcaA* deletion mutant in *S.* Typhimurium LT2 (KBP10). LPS samples were prepared, separated by 12% SDS-PAGE and (top) silver stained or (bottom) probed by Immunoblot with group B O-specific rabbit antiserum as described in Experimental Procedures.

4.3.6. Gal-1-P transferases do not complement CA synthesis in *E. coli* or holdfast synthesis in C. *crescentus*. As PssY and WcaJ mediated the production of lipid A-core OS linked polysaccharides resembling O antigen in the *wbaP* mutant MSS2, we next explored whether WbaP and other characterized galactose-1-phosphate transferases from *E. coli* and *Geobacillus stearothermophilus* could complement CA synthesis in a *wcaJ* mutant. We expressed WbaP$_{CT}$ of *S. enterica*, WbaP of *E. coli* K30 and WsaP of *G. stearothermophilus* in XBF1*rcsA*$^+$. None of the proteins could restore CA formation in this mutant, suggesting that these proteins do not utilize UDP-Glc (Fig. 4.16). To assess whether WbaP could restore the function of HfsE, PssY or PssZ in *C. crescentus*, we introduced two plasmids harboring the C-terminal of WbaP with or without an N-terminal TrxA fusion into the triple in-frame *hfsE*, *pssY* and *pssZ* deletion mutant and assayed for restoration of surface adhesion. Both of the WbaP$_{CT}$ complemented strains had a similar, reduced ability to bind to a polystyrene surface as the triple deletion mutant of *hfsE*, *pssY*, *pssZ*, suggesting that WbaP cannot rescue the holdfast defect in this mutant strain (Fig. 4.17).

4.4. Discussion

The biosynthesis of CA and the functional assignment of the genes in the *cps* gene cluster are not elucidated. Using *in vivo* complementation experiments combined with *in vitro* transferase assays, we provide here strong evidence that *E. coli* K-12 WcaJ is the UDP-Glc:Und-P Glc-1-P transferase that initiates CA synthesis. We also demonstrate that the *C. crescentus* homolog PssY, which shares high sequence similarity to WcaJ, also has Glc-1-P transferase activity. Indeed, the two proteins can indistinctly restore CA capsule formation in a *wcaJ* deletion mutant of *E. coli* as evidenced by the appearance of mucoid colonies in the presence of the RcsA positive regulator of *cps* gene transcription. When examined in the *S. enterica* MSS2 (Δ*wbaP*) mutant, both WcaJ and PssY mediated the synthesis of lipid A-core OS-linked polysaccharides that were initially detected as silver-stained bands resembling O antigen. Bacteria expressing these bands were less susceptible to the Ffm bacteriophage that recognizes uncapped lipid A-core OS for bacterial cell entry (64). In addition, MSS2 cells displayed similar banding patterns in the presence of RcsA, suggesting that overexpression of the endogenous copy of the *wcaJ*
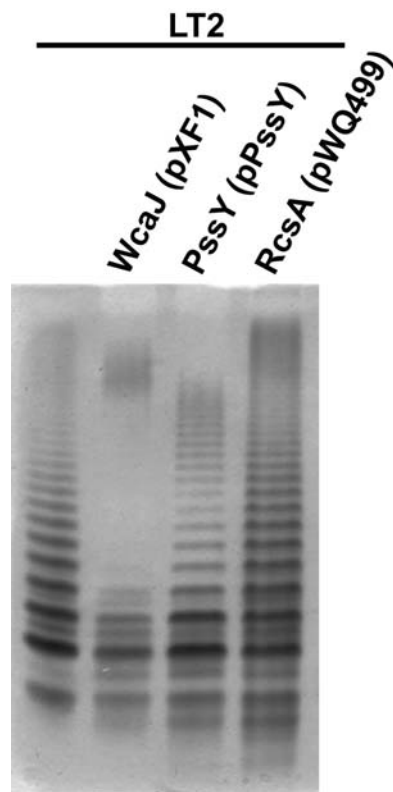
Figure 4.14. Dominant negative effect of WcaJ or PssY expression on O antigen production in wild type *S. enterica* LT2.
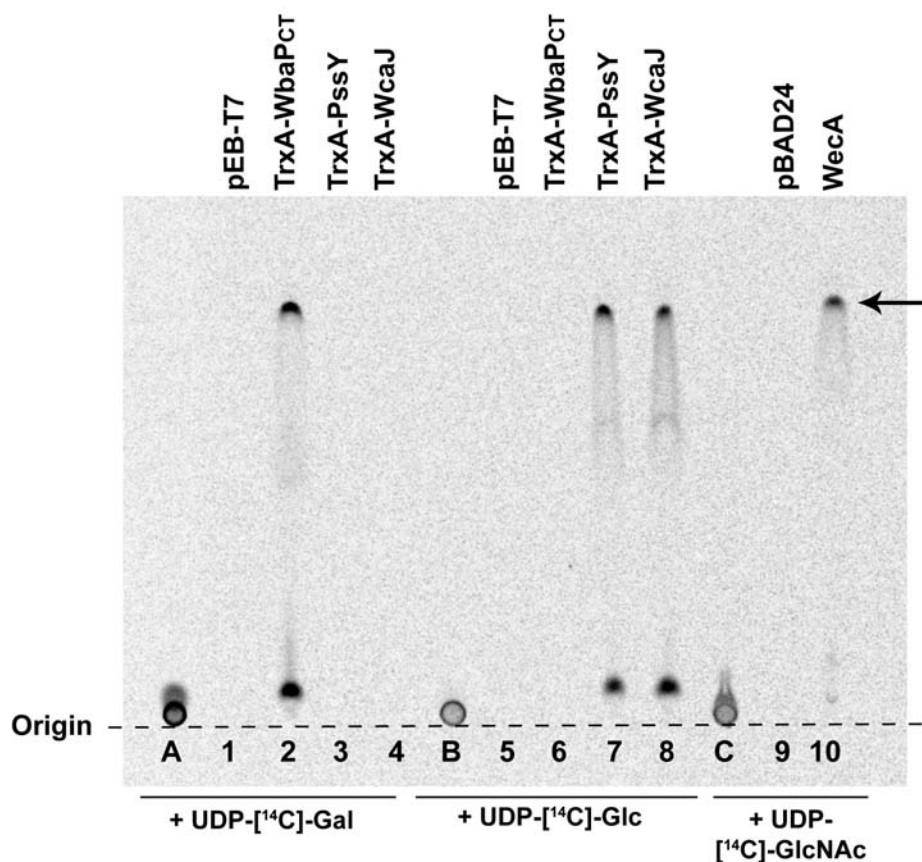
Figure 4.15. Determination of Und-P-P-hexose product formation by TLC. Membranes from C43 cells harboring various plasmids and membranes from additional control strains were assayed for transferase activity with 0.025 µCi of ^{14}C labeled UDP-Gal, ^{14}C labeled UDP-Glc or ^{14}C labeled UDP-GlcNAc as indicated. The lipid associated fraction was extracted with a mixture of chloroform/methanol as outlined in Experimental Procedures. Lane(s):1 and 5, C43(DE3)/pEB-T7; 2 and 6, C43(DE3)/pKP41(TrxA-WbaP$_{CT}$); 3 and 7, C43(DE3)/pKP101(TrxA-PssY); 4 and 8, C43(DE3)/pKP113 (TrxA-WcaJ); 9, MV501; 10, MV501/pKV1(WecA). 0.0025Ci of the radiolabeled UDP sugars were also spotted on the TLC. Lane: A, ^{14}C labeled UDP-Gal; B, ^{14}C labeled UDP-Glc; C, ^{14}C labeled UDP-GlcNAc. The products were separated on silica plates using the solvent system C/M/W (65:25:4).
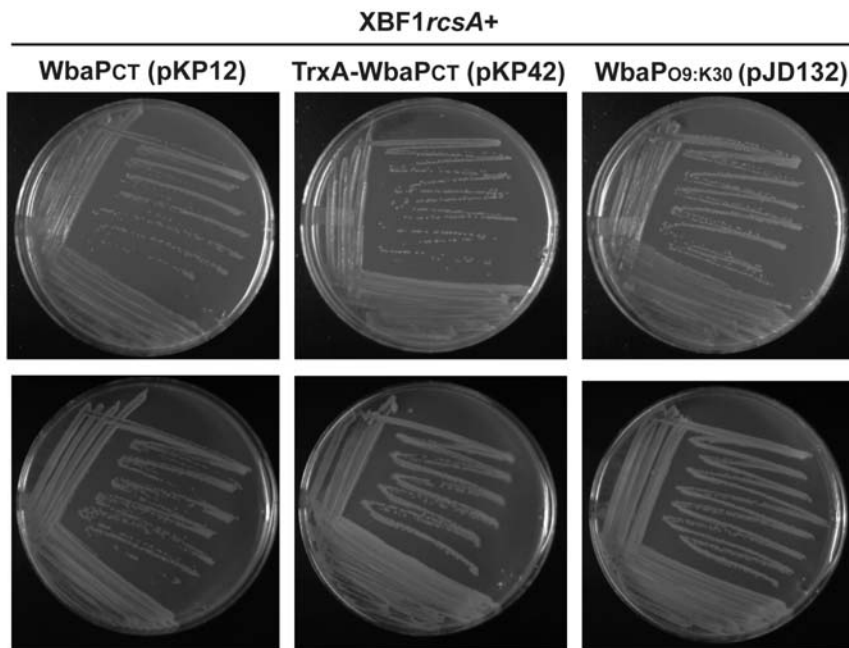
Figure 4.16. Expression of *S.* Typhimurium WbaP derivatives and *E. coli* O9a:K30 WbaP in *wcaJ* deletion mutant in *E. coli* K-12 W3110 (XBF1) does not produce mucoid phenotype. Cells were streaked on LB agar plates containing 100μg/ml ampicillin and no arabinose (top) or induction plates containing 0.2% arabinose (bottom). Plates were incubated overnight at 37°C.
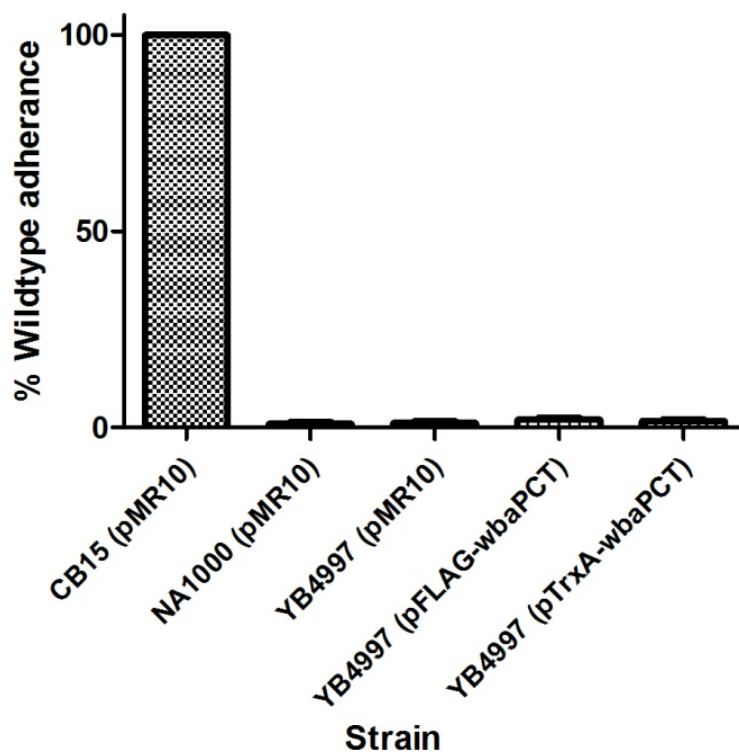
Figure 4.17. Quantification of crystal violet stained cells attached to polystyrene**.** The results represent the means of three replicates. Error bars represent standard deviations.
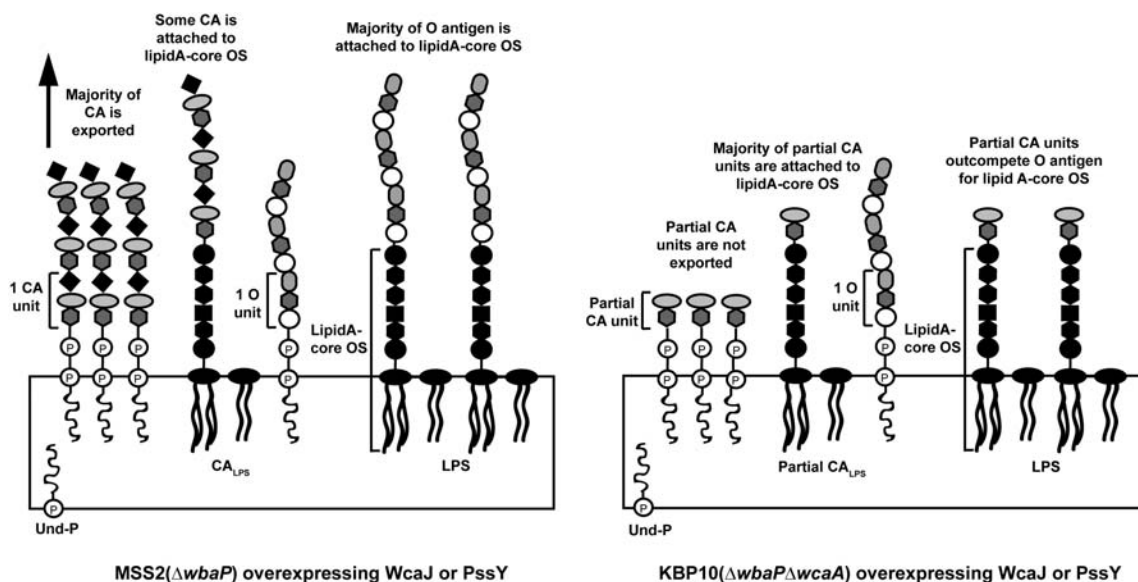
Figure 4.18. Proposed model of LPS and CA$_{LPS}$ production upon overexpression of WcaJ and PssY in MSS2($\Delta wbaP$) and KBP10($\Delta wbaP \Delta wcaA$). (Left) In MSS2$wcaJ^+$ or MSS2(pPssY), the majority of colanic acid produced is exported with a small amount ligated to lipid A-core OS. O antigen is produced at low levels but the majority of it is ligated to lipid A-core OS. (Right) In KBP10$wcaJ^+$ or KBP10(pPssY) the absence of WcaA leads to the formation of partial colanic acid units which are no longer exported and instead ligated to lipid A-core OS, outcompeting O antigen. Rectangles represent the inner membrane with the cytoplasm and periplasm below and above, respectively.

gene caused a similar phenotype. However, reactivity with group B antiserum, which detects O antigen, was observed upon WcaJ and PssY overexpression, but not upon overexpression of RcsA. Since RcsA increases the transcription of the entire *cps* genes in addition to *wcaJ*, we interpreted this result as due to an increase in CA production causing more CA and fewer O antigen molecules attached to lipid A-core OS. The modification of lipid A-core OS with CA is not unique to *Salmonella*. Meredith *et al* (33) have recently demonstrated the attachment of CA to lipid A-core OS in *E. coli* upon overexpression of CA with the amino acid analog *p*-fluorophenylalanine. These investigators also detected several glycoforms but concentrated on determining the structure of the most abundant form, which resulted in one single CA unit attached to a terminal heptose residue in the core OS, which is the normal attachment site recognized by the WaaL ligase (15, 20-21).

The production of lipid A-core OS linked O antigen in MSS2 cells overexpressing WcaJ and PssY was confirmed by GC-MS analysis of the sugars derived from LPS. To our knowledge, this is the first time a Glc-1-P transferase has been shown to transfer Gal-1-P. MSS2 cells overexpressing WcaJ or PssY also contained the unique CA sugar fucose. The structural information provided by Meredith *et al* (33) not only provides a more accurate determination than the earlier structure proposed by Garegg *et al* (16), but it also helps to interpret our results for the *S. enterica wcaA* mutant. The absence of O antigen linked to lipid A-core OS upon overexpression of WcaJ and PssY in *S.* Typhimurium KBP10 and the difference in migration of the silver stained bands suggests that WcaA is required to form a complete CA unit. We propose that in the absence of this protein, CA is no longer produced and exported and instead the partial units are ligated to lipid A-core OS, likely outcompeting the small amount of O antigen present (Fig. 4.18). Our results suggest that WcaA is responsible for the addition of a sugar in the main backbone of the subunit, which includes two L-fucose sugars, in addition to glucose (Figure 4.1A). Based on the change in the migration of the bands, we propose that WcaA adds the second fucose sugar, although this will be confirmed in a future study.

The formation of O antigen upon overexpression of PssY and WcaJ in MSS2 suggests that both proteins have Gal-1-P transferase activity *in vivo*. In contrast, the *in vitro* transferase assay with [^{14}C]-labeled UDP-Gal for WcaJ and PssY did not result in

product formation. It is possible that this assay may not be sensitive enough to capture low levels of labeled Und-PP-Gal product. Supporting this interpretation, site-directed WbaPCT mutants that have *in vivo* activity lack *in vitro* activity when tested under similar assay conditions (39). Both WcaJ and PssY catalyze the synthesis of Und-P-P-Glc *in vitro*. This could be demonstrated by examining the product conversion upon TLC. Two spots were detected; one that corresponds to Und-P-P-Glc and a low migrating spot that has also been reported from assays with WecA and WsaP in *E. coli* (48) and *G. stearothermophilus* (47), respectively. This spot may represent degraded product that did not separate, or the incorporation of ^{14}C-labeled Gal into a different lipid. Further analysis of the low-migrating spot, which is outside the scope of this work, will be required to determine its chemical nature.

The Glc-1-P transferase activity of PssY but not HfsE or PssZ was puzzling, as the activity of these proteins is believed to be redundant based on genetic complementation assays in a *C. crescentus ΔhfsEΔpssYΔpssZ* mutant (52). Adherence assays and visualization of holdfast by microscopy showed that any of these proteins individually expressed is sufficient for holdfast synthesis (52). The inability of HfsE and PssZ to restore CA synthesis in *E. coli* was not due to lack of protein expression, as sufficient HfsE was detected as a FLAG-tagged protein, and PssZ was overexpressed with an N-terminal 6×His-TrxA fusion. Other alternatives could be protein instability in *E. coli* and *S. enterica*, misfolding upon overexpression, posttranslational modifications in *C. crescentus* or the requirement of a different substrate sugar. HfsE, PssY and PssZ may utilize an additional substrate present in *C. crescentus* but absent in *E. coli* and *S. enterica*. Some members of the PHPT family, such as the *Campylobacter jejuni* PglC catalyze the transfer of 2,4-diacetamido-2,4,6-trideoxyglucose (bacillosamine)-1-phosphate to Und-P (17). Therefore, it is possible that PssY has the ability to use more than one substrate, while HfsE and PssZ are more restricted in sugar specificity. A similar case exists in *Pseudomonas aeruginosa* whereby the WbpL initiating transferase acts both in the synthesis of A and B band LPS (44). However, the *E. coli* protein WecA can only restore A band LPS synthesis but not the synthesis of the B band LPS in *wbpL* defective *P. aeruginosa* strain (44). In *C. crescentus,* it is also possible that the holdfast consists of various polysaccharides and HfsE and PssZ may initiate the synthesis of

polysaccharides that differ from those initiated by PssY. In this context, assays conducted in the *C. crescentus ΔhfsEΔpssYΔpssZ* mutant (52) may not be sensitive to establish differences in the qualities of holdfast produced when HfsE, PssY and PssZ are expressed alone.

In summary, we have demonstrated the UDP-Glc:Und-P Glc-1-P transferase activity of *E. coli* K-12 WcaJ and *C. crescentus* PssY, and that WcaJ is the initiating enzyme for CA synthesis. We have also found that when WcaJ or PssY are overexpressed in a mutant of *S.* Typhimurium devoid of the O antigen initiating galactosyltransferase WbaP, the proteins can utilize *in vivo* both UDP-Gal and UDP-Glc as substrates resulting in production of a mixture of lipid A-core OS molecules covalently linked to O antigen and CA.

4.5. Chapter four references

1. **Alexander, D. C., and M. A. Valvano.** 1994. Role of the *rfe* gene in the biosynthesis of the *Escherichia coli* O7-specific lipopolysaccharide and other O-specific polysaccharides containing *N*-acetylglucosamine. J Bacteriol **176:**7079-84.

2. **Allen, P., C. A. Hart, and J. R. Saunders.** 1987. Isolation from *Klebsiella* and characterization of two *rcs* genes that activate colanic acid capsular biosynthesis in *Escherichia coli*. J Gen Microbiol **133:**331-40.

3. **Arechaga, I., B. Miroux, S. Karrasch, R. Huijbregts, B. de Kruijff, M. J. Runswick, and J. E. Walker.** 2000. Characterisation of new intracellular membranes in *Escherichia coli* accompanying large scale over-production of the b subunit of F(1)F(o) ATP synthase. FEBS Lett **482:**215-9.

4. **Bachmann, B. J.** 1972. Pedigrees of some mutant strains of *Escherichia coli* K-12. Bacteriol Rev **36:**525-57.

5. **Bodenmiller, D., E. Toh, and Y. V. Brun.** 2004. Development of surface adhesion in *Caulobacter crescentus*. J Bacteriol **186:**1438-47.

6. **Brade, H., and C. Galanos.** 1982. Isolation, purification, and chemical analysis of the lipopolysaccharide and lipid A of *Acinetobacter calcoaceticus* NCTC 10305. Eur J Biochem **122:**233-7.

7. **Brown, P. J., G. G. Hardy, M. J. Trimble, and Y. V. Brun.** 2009. Complex regulatory pathways coordinate cell-cycle progression and development in *Caulobacter crescentus*. Adv Microb Physiol **54:**1-101.

8. **Clarke, B. R., D. Bronner, W. J. Keenleyside, W. B. Severn, J. C. Richards, and C. Whitfield.** 1995. Role of Rfe and RfbF in the initiation of biosynthesis of D-galactan I, the lipopolysaccharide O antigen from *Klebsiella pneumoniae* serotype O1. J Bacteriol **177:**5411-8.

9. **Clay, C. D., S. Soni, J. S. Gunn, and L. S. Schlesinger.** 2008. Evasion of complement-mediated lysis and complement C3 deposition are regulated by *Francisella tularensis* lipopolysaccharide O antigen. J Immunol **181:**5568-78.

10. **Crawford, R. W., D. L. Gibson, W. W. Kay, and J. S. Gunn.** 2008. Identification of a bile-induced exopolysaccharide required for *Salmonella* biofilm formation on gallstone surfaces. Infect Immun **76:**5341-9.

11. **Datsenko, K. A., and B. L. Wanner.** 2000. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. Proc Natl Acad Sci U S A **97:**6640-5.

12.   **Drummelsmith, J., and C. Whitfield.** 1999. Gene products required for surface expression of the capsular form of the group 1 K antigen in *Escherichia coli* (O9a:K30). Mol Microbiol **31:**1321-32.

13.   **Duerr, C. U., S. F. Zenk, C. Chassin, J. Pott, D. Gutle, M. Hensel, and M. W. Hornef.** 2009. O-antigen delays lipopolysaccharide recognition and impairs antibacterial host defense in murine intestinal epithelial cells. PLoS Pathog **5:**e1000567.

14.   **Evinger, M., and N. Agabian.** 1977. Envelope-associated nucleoid from *Caulobacter crescentus* stalked and swarmer cells. J Bacteriol **132:**294-301.

15.   **Feldman, M. F., C. L. Marolda, M. A. Monteiro, M. B. Perry, A. J. Parodi, and M. A. Valvano.** 1999. The activity of a putative polyisoprenol-linked sugar translocase (Wzx) involved in *Escherichia coli* O antigen assembly is independent of the chemical structure of the O repeat. J Biol Chem **274:**35129-38.

16.   **Garegg, P. J., B. Lindberg, T. Onn, and I. W. Sutherland.** 1971. Comparative structural studies on the M-antigen from *Salmonella typhimurium*. *Escherichia coli* and *Aerobacter cloacae*. Acta Chem Scand **25:**2103-8.

17.   **Glover, K. J., E. Weerapana, M. M. Chen, and B. Imperiali.** 2006. Direct biochemical evidence for the utilization of UDP-bacillosamine by PglC, an essential glycosyl-1-phosphate transferase in the *Campylobacter jejuni N*-linked glycosylation pathway. Biochemistry **45:**5343-50.

18.   **Gottesman, S., and V. Stout.** 1991. Regulation of capsular polysaccharide synthesis in *Escherichia coli* K12. Mol Microbiol **5:**1599-606.

19.   **Guzman, L. M., D. Belin, M. J. Carson, and J. Beckwith.** 1995. Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter. J Bacteriol **177:**4121-30.

20.   **Heinrichs, D. E., M. A. Monteiro, M. B. Perry, and C. Whitfield.** 1998. The assembly system for the lipopolysaccharide R2 core-type of *Escherichia coli* is a hybrid of those found in *Escherichia coli* K-12 and *Salmonella enterica*. Structure and function of the R2 Waak and WaaL. J Biol Chem **273:**8849-59.

21.   **Heinrichs, D. E., J. A. Yethon, P. A. Amor, and C. Whitfield.** 1998. The assembly system for the outer core portion of R1- and R4-type lipopolysaccharides of *Escherichia coli*. The R1 core-specific beta-glucosyltransferase provides a novel attachment site for O-polysaccharides. J Biol Chem **273:**29497-505.

22.   **Ilg, K., K. Endt, B. Misselwitz, B. Stecher, M. Aebi, and W. D. Hardt.** 2009. O-antigen-negative *Salmonella enterica* serovar Typhimurium is attenuated in intestinal colonization but elicits colitis in streptomycin-treated mice. Infect Immun **77:**2568-75.

23. **Jansson, P. E., A. A. Lindberg, B. Lindberg, and R. Wollin.** 1981. Structural studies on the hexose region of the core in lipopolysaccharides from Enterobacteriaceae. Eur J Biochem **115:**571-7.

24. **Klena, J. D., and C. A. Schnaitman.** 1993. Function of the *rfb* gene cluster and the *rfe* gene in the synthesis of O antigen by *Shigella dysenteriae* 1. Mol Microbiol **9:**393-402.

25. **Ledeboer, N. A., and B. D. Jones.** 2005. Exopolysaccharide sugars contribute to biofilm formation by *Salmonella enterica* serovar typhimurium on HEp-2 cells and chicken intestinal epithelium. J Bacteriol **187:**3214-26.

26. **Lehrer, J., K. A. Vigeant, L. D. Tatar, and M. A. Valvano.** 2007. Functional characterization and membrane topology of *Escherichia coli* WecA, a sugar-phosphate transferase initiating the biosynthesis of enterobacterial common antigen and O-antigen lipopolysaccharide. J Bacteriol **189:**2618-28.

27. **Lesse, A. J., A. A. Campagnari, W. E. Bittner, and M. A. Apicella.** 1990. Increased resolution of lipopolysaccharides and lipooligosaccharides utilizing tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis. J Immunol Methods **126:**109-17.

28. **Levi, A., and U. Jenal.** 2006. Holdfast formation in motile swarmer cells optimizes surface attachment during *Caulobacter crescentus* development. J Bacteriol **188:**5315-8.

29. **Majdalani, N., and S. Gottesman.** 2005. The Rcs phosphorelay: a complex signal transduction system. Annu Rev Microbiol **59:**379-405.

30. **Maniatis, T.** 1982. Molecular cloning : a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

31. **Marolda, C. L., J. Welsh, L. Dafoe, and M. A. Valvano.** 1990. Genetic analysis of the O7-polysaccharide biosynthesis region from the *Escherichia coli* O7:K1 strain VW187. J Bacteriol **172:**3590-9.

32. **Meier-Dieter, U., K. Barr, R. Starman, L. Hatch, and P. D. Rick.** 1992. Nucleotide sequence of the *Escherichia coli rfe* gene involved in the synthesis of enterobacterial common antigen. Molecular cloning of the *rfe-rff* gene cluster. J Biol Chem **267:**746-53.

33. **Meredith, T. C., U. Mamat, Z. Kaczynski, B. Lindner, O. Holst, and R. W. Woodard.** 2007. Modification of lipopolysaccharide with colanic acid (M-antigen) repeats in *Escherichia coli*. J Biol Chem **282:**7790-8.

34. **Merker, R. I., and J. Smit.** 1988. Characterization of the adhesive holdfast of marine and freshwater caulobacters. Appl Environ Microbiol **54:**2078-85.

35. **Murray, G. L., S. R. Attridge, and R. Morona.** 2006. Altering the length of the lipopolysaccharide O antigen has an impact on the interaction of *Salmonella enterica* serovar Typhimurium with macrophages and complement. J Bacteriol **188:**2735-9.

36. **Nothaft, H., and C. M. Szymanski.** 2010. Protein glycosylation in bacteria: sweeter than ever. Nat Rev Microbiol **8:**765-78.

37. **Olsthoorn, M. M., B. O. Petersen, J. Duus, J. Haverkamp, J. E. Thomas-Oates, K. Bock, and O. Holst.** 2000. The structure of the linkage between the O-specific polysaccharide and the core region of the lipopolysaccharide from *Salmonella enterica* serovar Typhimurium revisited. Eur J Biochem **267:**2014-27.

38. **Paixao, T. A., C. M. Roux, A. B. den Hartigh, S. Sankaran-Walters, S. Dandekar, R. L. Santos, and R. M. Tsolis.** 2009. Establishment of systemic *Brucella melitensis* infection through the digestive tract requires urease, the type IV secretion system, and lipopolysaccharide O antigen. Infect Immun **77:**4197-208.

39. **Patel, K. B., S. E. Furlong, and M. A. Valvano.** 2010. Functional analysis of the C-terminal domain of the WbaP protein that mediates initiation of O antigen synthesis in *Salmonella enterica*. Glycobiology **20:**1389-401.

40. **Perez, J. M., M. A. McGarry, C. L. Marolda, and M. A. Valvano.** 2008. Functional analysis of the large periplasmic loop of the *Escherichia coli* K-12 WaaL O-antigen ligase. Mol Microbiol **70:**1424-40.

41. **Poindexter, J. S.** 1964. Biological Properties and Classification of the Caulobacter Group. Bacteriol Rev **28:**231-95.

42. **Raetz, C. R., and C. Whitfield.** 2002. Lipopolysaccharide endotoxins. Annu Rev Biochem **71:**635-700.

43. **Rick, P. D., G. L. Hubbard, and K. Barr.** 1994. Role of the *rfe* gene in the synthesis of the O8 antigen in *Escherichia coli* K-12. J Bacteriol **176:**2877-84.

44. **Rocchetta, H. L., L. L. Burrows, J. C. Pacan, and J. S. Lam.** 1998. Three rhamnosyltransferases responsible for assembly of the A-band D-rhamnan polysaccharide in *Pseudomonas aeruginosa*: a fourth transferase, WbpL, is required for the initiation of both A-band and B-band lipopolysaccharide synthesis. Mol Microbiol **28:**1103-19.

45. **Saldias, M. S., X. Ortega, and M. A. Valvano.** 2009. *Burkholderia cenocepacia* O antigen lipopolysaccharide prevents phagocytosis by macrophages and adhesion to epithelial cells. J Med Microbiol **58:**1542-8.

46. **Saldías, M. S., K. Patel, C. L. Marolda, M. Bittner, I. Contreras, and M. A. Valvano.** 2008. Distinct functional domains of the *Salmonella enterica* WbaP

transferase that is involved in the initiation reaction for synthesis of the O antigen subunit. Microbiology **154:**440-53.

47. **Schaffer, C., R. Novotny, S. Kupcu, S. Zayni, A. Scheberl, J. Friedmann, U. B. Sleytr, and P. Messner.** 2007. Novel biocatalysts based on S-layer self-assembly of *Geobacillus stearothermophilus* NRS 2004/3a: a nanobiotechnological approach. Small **3:**1549-59.

48. **Schaffer, C., T. Wugeditsch, P. Messner, and C. Whitfield.** 2002. Functional expression of enterobacterial O-polysaccharide biosynthesis enzymes in *Bacillus subtilis*. Appl Environ Microbiol **68:**4722-30.

49. **Steiner, K., R. Novotny, K. Patel, E. Vinogradov, C. Whitfield, M. A. Valvano, P. Messner, and C. Schaffer.** 2007. Functional characterization of the initiation enzyme of S-layer glycoprotein glycan biosynthesis in *Geobacillus stearothermophilus* NRS 2004/3a. J Bacteriol **189:**2590-8.

50. **Stevenson, G., K. Andrianopoulos, M. Hobbs, and P. R. Reeves.** 1996. Organization of the *Escherichia coli* K-12 gene cluster responsible for production of the extracellular polysaccharide colanic acid. J Bacteriol **178:**4885-93.

51. **Sutherland, I.** 2001. Biofilm exopolysaccharides: a strong and sticky framework. Microbiology **147:**3-9.

52. **Toh, E., H. D. Kurtz, Jr., and Y. V. Brun.** 2008. Characterization of the *Caulobacter crescentus* holdfast polysaccharide biosynthesis pathway reveals significant redundancy in the initiating glycosyltransferase and polymerase steps. J Bacteriol **190:**7219-31.

53. **Tsai, C. M., and C. E. Frasch.** 1982. A sensitive silver stain for detecting lipopolysaccharides in polyacrylamide gels. Anal Biochem **119:**115-9.

54. **Umbreit, T. H., and J. L. Pate.** 1978. Characterization of the holdfast region of wild-type cells and holdfast mutants of *Asticcacaulis biprosthecum*. Archives of Microbiology **118:**157-168.

55. **Valvano, M. A.** 2003. Export of O-specific lipopolysaccharide. Front Biosci **8:**s452-71.

56. **Valvano, M. A., S. E. Furlong, and K. B. Patel.** 2010. O antigen biosynthesis. *In* L. N. Mander and H.-W. Liu (ed.), Comprehensive Natural Products Chemistry, vol. 6. Elsevier Ltd, Oxford.

57. **Vuong, C., J. M. Voyich, E. R. Fischer, K. R. Braughton, A. R. Whitney, F. R. DeLeo, and M. Otto.** 2004. Polysaccharide intercellular adhesin (PIA) protects *Staphylococcus epidermidis* against major components of the human innate immune system. Cell Microbiol **6:**269-75.

58. **Wang, L., D. Liu, and P. R. Reeves.** 1996. C-terminal half of *Salmonella enterica* WbaP (RfbP) is the galactosyl-1-phosphate transferase domain catalyzing the first step of O-antigen synthesis. J Bacteriol **178:**2598-604.

59. **West, N. P., P. Sansonetti, J. Mounier, R. M. Exley, C. Parsot, S. Guadagnini, M. C. Prevost, A. Prochnicka-Chalufour, M. Delepierre, M. Tanguy, and C. M. Tang.** 2005. Optimization of virulence functions through glucosylation of *Shigella* LPS. Science **307:**1313-7.

60. **Westphal, O., and K. Jann.** 1965. Bacterial Lipopolysaccharides, p. 83-91. *In* R. L. Whistler (ed.), Methods in Carbohydrate Chemistry, vol. 5. Academic Press, New York.

61. **White, A. P., D. L. Gibson, S. K. Collinson, P. A. Banser, and W. W. Kay.** 2003. Extracellular polysaccharides associated with thin aggregative fimbriae of *Salmonella enterica* serovar enteritidis. J Bacteriol **185:**5398-407.

62. **Whitfield, C.** 2006. Biosynthesis and assembly of capsular polysaccharides in *Escherichia coli*. Annu Rev Biochem **75:**39-68.

63. **Whitfield, C., W. J. Keenleyside, P. Ronald MacLachlan, P. Jayaratne, and A. J. Clarke.** 1995. [16] Identification of rcs Genes in *Escherichia coli* O9:K30:H12 and Involvement in Regulation of Expression of Group IA K30 Capsular Polysaccharide, p. 301-322. *In* W. A. Kenneth (ed.), Methods in Molecular Genetics, vol. Volume 6. Academic Press.

64. **Wilkinson, R. G., P. Gemski, Jr., and B. A. Stocker.** 1972. Non-smooth mutants of *Salmonella typhimurium*: differentiation by phage sensitivity and genetic mapping. J Gen Microbiol **70:**527-54.

65. **Yao, Z., and M. A. Valvano.** 1994. Genetic analysis of the O-specific lipopolysaccharide biosynthesis region (*rfb*) of *Escherichia coli* K-12 W3110: identification of genes that confer group 6 specificity to *Shigella flexneri* serotypes Y and 4a. J Bacteriol **176:**4133-43.

Chapter 5

The C-terminal domain of the *Salmonella enterica* WbaP (UDP-galactose:Und-P galactose-1-phosphate transferase) is sufficient for catalytic activity and specificity for undecaprenyl monophosphate

5.1.   Introduction

The biogenesis of bacterial cell wall polysaccharides requires membrane-associated and -embedded proteins. One of these proteins in *Salmonella enterica* serovar Typhimurium (*S.* Typhimurium) is WbaP, which initiates O antigen synthesis by catalyzing the transfer of galactose-1-phosphate (Gal-P) from a nucleotide precursor (UDP-Gal), onto the membrane-embedded isoprenoid lipid carrier, undecaprenyl monophosphate (Und-P) yielding Und-P-P-Gal (16, 19, 26-27). WbaP belongs to the polyisoprenyl-phosphate hexose-1-phosphate transferase (PHPT) family (24-25). PHPT members are only found in prokaryotes and initiate the synthesis of a variety of glycans such as colanic acid in *Escherichia coli* K-12 (22), K30 capsule in *E. coli* O9a:K30 (7), type 2 capsule in *Streptococcus pneumoniae* (3), and glycans for S-layer protein glycosylation in *Geobacillus stearothermophilus* (21). Most proteins in this family utilize hexose sugars, but some exceptions exist. For example, PglC in *Campylobacter jejuni* and PglB in *Neisseria* sp transfer 2,4-diacetamido-2,4,6-trideoxyglucose (bacillosamine)-1-P and a 2(4)-acetamido-4(2)-glyceramido-2,4,6-trideoxyhexose-1-P, respectively, to initiate the synthesis of precursors for protein glycosylation pathways (4, 9-10, 18). Another member of this family, WecP, was recently demonstrated to function with UDP-*N*-acetylgalactosamine (13).

WbaP has five predicted transmembrane (TM) helices and three domains (19) (Fig. 5.1). The N-terminal domain, including the first four TMs and their connecting loops, have no assigned function although it may contribute to the overall stability of the protein in the membrane (19). A large predicted soluble loop between TMIV and TMV has been implicated in modulating chain length distribution of O antigen, while the C-terminal domain that includes TMV and the tail is sufficient for the activity of the enzyme *in vivo* and *in vitro* (16, 19, 26). Trypsin cleavage and green-fluorescent protein reporter experiments confirmed that the C-terminus of the protein including approximately the last 20 kDa is cytoplasmic (16).

Although most PHPT proteins studied are predicted to contain five complete TMs, others such as *Caulobacter crescentus* PssY and PssZ only contain the C-terminal domain (23). Previous work determined that some of the highly conserved residues within the cytosolic C-terminal region of WbaP are required for function (16). The

location of the catalytic domain to a defined region of the protein prompted us to purify and biochemically characterize the C-terminal domain of WbaP (spanning residues 258-476, WbaP$_{CT}$) in the absence of contaminating membrane proteins or lipids. Using an *in vitro* assay with chemically defined components we demonstrate that WbaP$_{CT}$ is sufficient for enzymatic activity and also carries high specificity for the Und-P lipid acceptor.

5.2.   Materials and Methods

5.2.1. Bacterial strains and growth conditions. *E. coli* strains DH5α (laboratory stock), BL21(DE3) (Invitrogen) and C43(DE3) (Invitrogen) were used for the overexpression of proteins. *S. enterica* Typhimurium LT2 was used to prepare chromosomal DNA. Bacteria were grown aerobically at 37 °C in Luria-Bertani (LB) medium (Difco) (10 mg/ml tryptone; 5 mg/ml yeast extract; 5 mg/ml NaCl). Media were supplemented with 100 μg/ml ampicillin, 30 μg/ml chloramphenicol or 40 μg/ml kanamycin as appropriate.

5.2.2. Plasmid construction and sequencing. Plasmid DNA was isolated using the Qiagen miniprep kit (Qiagen Inc., Mississauga, Ontario, Canada). Digestion with restriction enzymes, ligation with T4-ligase and transformation were carried out as described by Maniatis *et al* (12). DNA sequences were determined using an automated sequencer at the York University Core Molecular Biology and DNA Sequencing Facility, Toronto, Ontario, Canada. Plasmid pKP18 was constructed by PCR amplification of a 624-bp fragment using primers 2882 (5'-CTGGTCGACATTATTCAGTACTTCTCG-3') and 2907 (5'-CTAGTTAGGATCCAGGTCGTCCCGTTTTCTC-3') and LT2 DNA as template. This fragment was digested with *Sal*I and *Bam*HI and ligated into these sites in pET28a (Novagen, EMD Biosciences, USA). Plasmid pKP41 is described elsewhere (16).

5.2.3. Chemicals. All reagents unless otherwise specified were supplied by Sigma Aldrich. ^{14}C labeled UDP-galactose (Specific activity of 300mCi/mmol) was purchased

from American Radiolabeled Chemicals, Inc (St. Louis, Missouri). Purified neryl monophosphate ($C_{10}H_{25}N_2O_4P$), farnesyl monophosphate ($C_{15}H_{33}N_2O_4P$), geranylgeranyl monophosphate ($C_{20}H_{41}N_2O_4P$), heptaprenyl monophosphate ($C_{35}H_{65}N_2O_4P$), decaprenyl monophosphate ($C_{50}H_{89}N_2O_4P$) undecaprenyl monophosphate ($C_{55}H_{97}N_2O_4P$) and dolichyl monophosphate ($C_{95}H_{163}N_2O_4P$) were chemically phosphorylated using phosphoramidite chemistry (2, 29). The purity of the polyisoprenyl phosphates was evaluated by thin layer chromatography and determined to be >95% in all cases. To prepare the polyisoprenyl phosphates, prenols (C35, C50, C55) were isolated from diverse plant sources (28), while dolichol (C95) was isolated from mammalian liver and purified from the natural extracts by chromatographic methods (6). Oligoprenols (C10, C15 and C20) were from Sigma.

5.2.4. Membrane preparation and immunoblotting. Bacterial cultures were grown overnight in 5 mL of LB, diluted to an initial OD_{600} of 0.2, and incubated at 37°C for 2 h until cultures reached an OD_{600} of 0.6. At this point IPTG was added to a final concentration of 100-400 μM. Cultures were then incubated for 5 h at 30°C. Cells were harvested by centrifugation at $10,000 \times g$ for 10 min at 4°C. The bacterial pellet was suspended in lysis buffer (20 mM Tris–HCl [pH 8.5] + 300 mM NaCl, unless otherwise specified) containing protease inhibitor cocktail (Roche diagnostics), and the cells were lysed using a French Press cell (Thermo Scientific). Cell debris were removed by centrifugation ($15,000 \times g$ for 15 min at 4°C), and the clear supernatant was centrifuged at $40,000 \times g$ for 30 min at 4°C. The pellet, containing total membranes, was suspended in lysis buffer. The protein concentration was determined by the Bradford assay (Bio-Rad, Hercules, CA). Staining was performed with Coomassie brilliant blue R250 (Life Technologies, Inc). SDS–PAGE, protein transfers to nitrocellulose membranes, and immunoblots were performed as described (17). For detection of 6xHis- proteins membranes were incubated with a 1:10,000 dilution of anti-His IgG2a monoclonal antibodies (Amersham, Piscataway, New Jersey).

5.2.5. Membrane protein solubilization. For initial detergent screening, total membranes BL21(DE3) and C43(DE3) cells expressing pKP18 and pKP41 were

solubilized in buffer (20 mM Tris-HCl [pH 8.5], 150 mM NaCl, 2 mM 2-mercaptoethanol, 10% glycerol) containing 2.5% DDM, 7% Triton X-100, or 6% CHAPS in a final volume of 100 µl. The mixture was incubated at 4°C for 2hrs with rotation in a Barnstead Thermolyne LABQUAKE (Barnstead International, Dubuque, Iowa). After centrifugation at 40,000 x g for 30 min at 4°C, the supernatant, representing the soluble fraction, was collected. Aliquots (10 %) of the soluble and pellet fractions were used for SDS-PAGE analysis. For screening to find an optimal detergent for enzymatic activity the equivalent of 200 µg of total membrane protein was solubilized in buffer (25 mM Tris-HCl [pH=8], 150 mM NaCl) and 0.5 % or 1 % of DDM, CHAPS, OG or Triton X-100 to a final volume of 200 µl. The soluble fraction was recovered as described above. Aliquots (2.5 %) of the insoluble and soluble fractions were used for SDS-PAGE analysis and 50 µL of the soluble fraction was used to determine i*n vitro* enzymatic activity. For protein purification, total membranes were resuspended in buffer (20 mM Tris-HCl [pH 8.5], 300 mM NaCl, 10 mM 2-mercaptoethanol, 10% glycerol) with 1% DDM and the soluble fractions was collected as described above.

5.2.6. Purification of 6xHis-TrxA-WbaP$_{CT}$ and cleavage by TEV protease. Ni^{2+}-bound chelating Sepharose Fast Flow resin (GE Healthcare) equilibrated with wash buffer (25mM NaPO$_4$ [pH 7.5], 300 mM NaCl, 10 mM 2-mercaptoethanol, 10 mM imidazole, 10% glycerol and 0.03% DDM) was mixed with the cleared supernatant recovered after solubilization with 1% DDM. After 30 min at 4°C with rotation in a Barnstead Thermolyne LABQUAKE (Barnstead International, Dubuque, Iowa) the resin was centrifuged at 3,000 x g and the flow through collected. The resin was washed with wash buffer containing 30 and 60mM imidazole and the protein was eluted with 250mM. Elutions were combined and dialyzed overnight against 100× dialysis buffer (25 mM NaPO$_4$ [pH 7.5], 150 mM NaCl, 10 mM 2-mercaptoethanol, 10% glycerol and 0.03% DDM). To remove the 6xHis-TrxA fusion, the dialyzed protein was concentrated 5× using a Vivaspin 20 centrifugal concentrator (MWCO = 10,000) (Vivaproducts, Inc, Littleton, MA). 5µL of AcTEV protease (Invitrogen) was added to 500 µl of protein and incubated at 4°C overnight. The next day, the sample was added to Ni^{2+} bound resin and the flow through containing the cleaved protein was collected, concentrated and

exchanged against buffer containing no salts. The cleaved protein was quantified using the bicinchoninic acid (BCA) assay (Pierce, Thermo Scientific).

5.2.7. *In vitro* **transferase assay.** To test the enzymatic activity of detergent solubilized samples, 50 µl of solubilized protein was added to 50 mM Tris-HCl [pH 8], 25 mM MgCl$_2$, 2 mM 2-mercaptoethanol and 0.33µM of ^{14}C labeled UDP-Gal to a final volume of 100 µl and incubated for 30 min at 37°C. For optimization of *in vitro* reaction conditions, effects of pH, salt and were tested with 100, 150 and 50 ng of protein respectively and PIs were tested with 50 ng. Assays were performed using conditions described above with the exception of the addition of 100 µM Und-P and a final volume of 50 µL. Steady state kinetics were determined using optimized assay conditions (pH 8, 25 mM MgCl$_2$, 1% CHAPS). For UDP-Gal, 200 ng of protein was tested in 100 µM Und-P and in the range of 0.2 – 125µM of ^{14}C labeled UDP-Gal. For Und-P, 100 ng of protein was tested in 0.33 µM of ^{14}C labeled UDP-Gal and in the range of 10 – 1000µM for Und-P. Reactions were incubated at 37°C for 15 min. Extraction of the lipid fractions was adopted from Schäffer et al (20) with modifications. After incubation at 37°C, the reactions were stopped with 200-400 µl of chloroform-methanol (C/M=3:2). To extract the lipid phase, the mixture was shaken vigorously for 3 min and centrifuged for 2 min at 14,000 × *g*. The lower organic phase was collected and 55 µl of 40 mM MgCl$_2$ was added followed by 5 min of vigorous shaking. The mixture was centrifuged as before and the top phase was removed. The organic phase was washed twice with 200-400 µl of pure solvent upper phase (PSUP) (C/M/W/1M MgCl$_2$=18:294:293:1). For scintillation counting the organic phase was added to 5 ml of scintillation fluid (Ecolume, MP Biomedical, Solon, OH) and radioactivity was determined by a Beckman liquid scintillation counter (Beckman Coulter Canada, Inc., Mississauga, Ontario, Canada).

5.3.   Results and Discussion

5.3.1. Overproduction and solubilization of the C-terminal domain of WbaP with an N-terminal TrxA fusion. *S*. Typhimurium *wbaP*$_{R274-Y476}$ was amplified from LT2 DNA and cloned into pET28a generating the plasmid pKP18. This construct generated a

gene product with an N-terminal 6xHis tag to facilitate detection by immunoblot and protein purification by Ni^{2+}-affinity chromatography. After induction of BL21(DE3)/pKP18 with IPTG, total membranes were prepared and analyzed by SDS-PAGE. Immunoblotting revealed a 29-kDa band corresponding to the expected molecular weight of 6xHisWbaP$_{CT}$ (Fig. 5.2). We next attempted to solubilize 6xHis-WbaP$_{CT}$ from the total membrane fractions for purification. To retain the activity of the protein we tested non-ionic detergents like Triton X-100 and n-Dodecyl-β-D-maltoside (DDM), and the zwitterionic detergent, 3-[(3-cholamidopropyl)dimethylammonio]-1-propanesulfonate (CHAPS) at over 10× their critical micelle concentrations (CMC). None of the detergents were successful in extracting 6xHis-WbaP$_{CT}$ from the membrane fractions (data not shown). Varying growth and induction conditions such as temperature, time, IPTG concentration, and aeration did not improve solubilization suggesting that BL21(DE3) could produce 6xHisWbaP$_{CT}$ as inclusion bodies. Inclusion bodies are typically described for proteins that accumulate in the periplasm or in the cytosol, but can also be membrane associated and therefore detected in the total membrane fraction (8). An N-terminal fusion of the 12- kDa thioredoxin (TrxA) to a target protein can prevent the formation of inclusion bodies (11). Indeed, we reported previously that an N-terminal TrxA fusion enhanced membrane localization and folding of WbaP$_{CT}$ (16). We reasoned that if this fusion reduced the production of inclusion bodies it would improve solubilization by detergents. Plasmid pKP41 expresses wbaP$_{F258-Y476}$ with an N-

terminal 6xHis tag and TrxA fusion as recently described (16). We solubilized total membranes from BL21(DE3) cells expressing 6xHis-TrxA-WbaP$_{CT}$ and the insoluble 6xHis-WbaP$_{CT}$ in 2.5% DDM. As before, 6xHis-WbaP$_{CT}$ was not solubilized from total membranes (Fig. 5.3, lanes 1-P and 1-S). In contrast, 6xHis-TrxA-WbaP$_{CT}$ was well expressed and readily soluble in the detergent suspension (Fig. 5.3, lanes 2-P and 2-S). To improve the solubilization we repeated the experiments using the C43(DE3) strain. This strain derives from BL21(DE3) and allows for improved expression of membrane proteins (14). Expression and solubilization of 6xHis-TrxA-WbaP$_{CT}$ was more efficient in C43(DE3) bacteria (Fig. 5.3, lanes 4-P and 4-S) whereas 6xHis-WbaPCT was not well expressed in these cells (Fig. 5.3, lanes 3-P and 3-S). Similar results were obtained using 6% CHAPS and 7% Triton X-100.
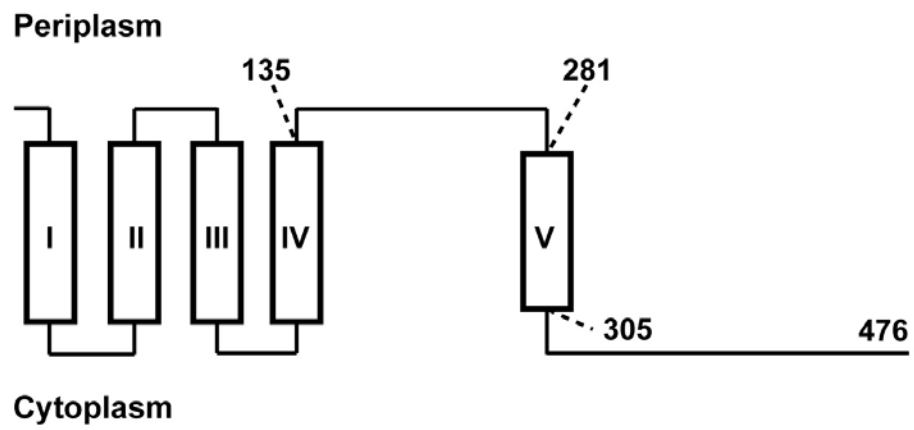
**Figure 5.1.** Predicted topology of *S. enterica* WbaP adapted from Saldias *et al* (2008). (19)
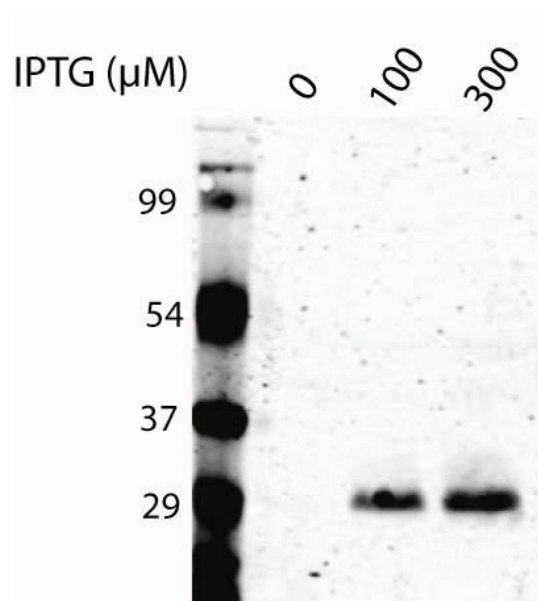
Figure 5.2. Expression of 6xHis-WbaP$_{R274-Y476}$ in total membranes extracted from BL21(DE3) cells. Cells were uninduced, or induced with 100 or 300µM IPTG. Samples were separated by 14% SDS-PAGE and probed with anti-His antibodies. 6xHis-WbaP$_{R274-Y476}$ migrated at ~29kDa.

To find an efficient detergent for solubilization of 6xHis-TrxA-WbaP$_{CT}$ and compatibility with enzymatic activity, we also tested the non-ionic detergent octylglucoside (OG) at 0.5 and 1% concentrations. After solubilization and high-speed centrifugation, 2.5% of the insoluble pellet and the soluble supernatant fractions were analyzed by SDS-PAGE and Western blotting and the soluble fractions were also tested for enzymatic activity with 20 mM MgCl$_2$, 100 µM Und-P and 0.33 µM ^{14}C labeled UDP-Gal. Although DDM was most effective at solubilizing 6xHis-TrxA-WbaP$_{CT}$ (Fig. 5.4), the protein was most active in CHAPS at a concentration of 1% (Table. 1). We concluded that DDM was preferred to solubilize and purify 6xHis-TrxA-WbaP$_{CT}$ and CHAPs would be optimal to assay the protein *in vitro*.

5.3.2. Purification of an active 6xHis-TrxA-WbaP$_{CT}$ protein and cleavage of the 6xHis fusion partner.

6xHis-TrxA-WbaP$_{CT}$ was purified as described in *Materials and Methods*. Total membranes from a 150 mL culture of C43(DE3)/pKP41 were solubilized in buffer containing 1% DDM. Following high-speed centrifugation, the supernatant was recovered and applied to Ni2$^+$-loaded resin. After washing the resin, 6xHis-TrxA-WbaP$_{CT}$ was eluted with 250 µM imidazole. SDS-PAGE and Coomassie staining showed that 6xHis-TrxA-WbaP$_{CT}$ efficiently bound to the column and was eluted only at high levels of imidazole (Fig. 5.5A). After overnight dialysis, the protein concentration was quantified and the purified enzyme remained active as assayed with exogenous Und-P and ^{14}C labeled UDP-Gal.

We next proceeded to remove the 6xHis-TrxA fusion by cleavage at the tobacco etch virus (TEV) site between WbaP$_{CT}$ and the fusion partner. Pure 6xHis-TrxA-WbaP$_{CT}$ was incubated with 6xHis-TEV overnight at 4°C. After proteolysis, the sample was added to Ni^{2+} charged resin and the flow through and all fractions were analyzed by SDS-PAGE. The TEV cleavage was complete (Fig. 5.5B, lanes 3) affording 10% WbaP$_{CT}$ (Fig. 5.5B, lanes 4), while the reminder of the cleaved protein remained associated to the resin (Fig. 5.5B, lanes 5). The protein recovered in the flow through was concentrated and quantified for further analysis (Fig. 5.5C). This protocol provided 0.2 mg/l of highly purified and enzymatically active WbaP$_{CT}$, which was sufficient for the biochemical analysis.
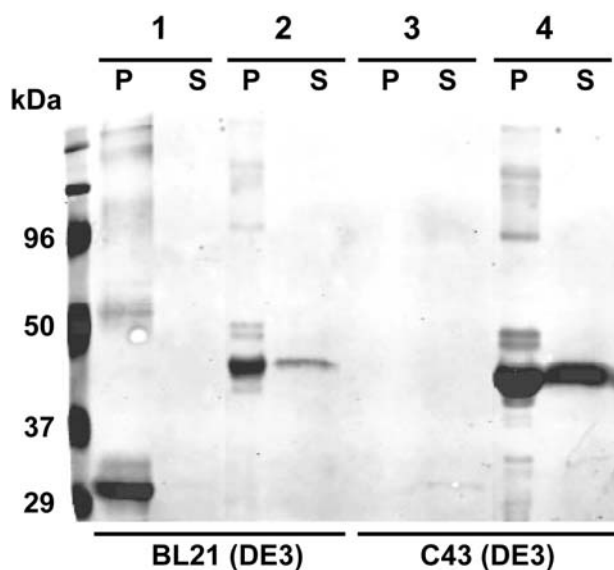
Figure 5.3. Solubilization of 6xHis-WbaP$_{R274-Y476}$ and 6xHis-TrxA-WbaP$_{F258-Y476}$ from total membranes of BL21(DE3) and C43(DE3) cells. After IPTG induction and solubilization with 2.5% DDM the insoluble pellet (P) and soluble supernatant (S) fractions were recovered as described in *Materials and Methods* from 1, BL21(DE3)/ 6xHis-WbaP$_{R274-Y476}$; 2, BL21(DE3)/ 6xHis-TrxA-WbaP$_{F258-Y476}$; 3, C43(DE3)/ 6xHis-WbaP$_{R274-Y476}$ and 4, C43(DE3)/ 6xHis-TrxA-WbaP$_{F258-Y476}$. Samples were separated by 14% SDS-PAGE and probed with anti-His antibodies.
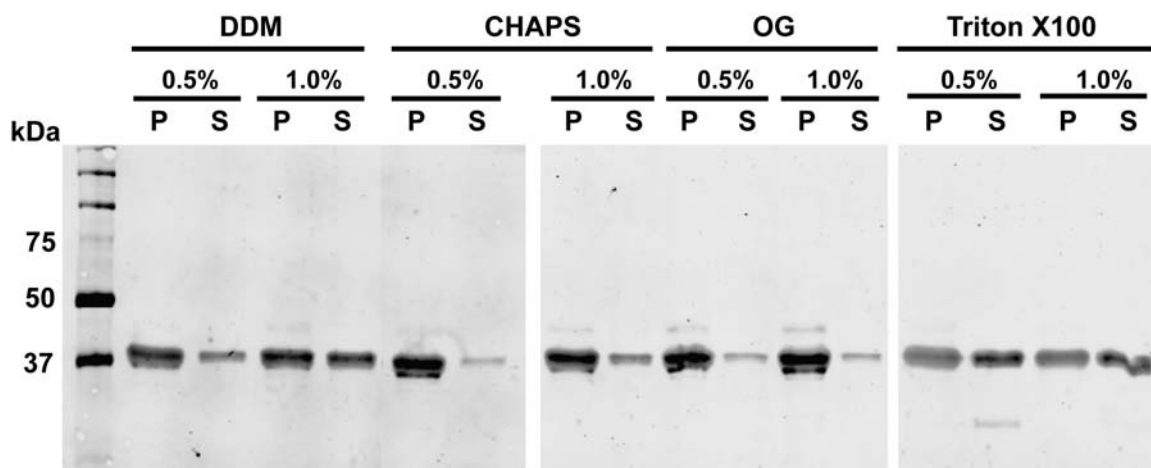
Figure 5.4. Detergent screening for solubilization of 6xHis-TrxA-WbaP$_{F258-Y476}$. Total membranes from IPTG induced C43(DE3)/ 6xHis-WbaP$_{R274-Y476}$ cells were solubilized in 0.5 and 1% DDM, CHAPS, Triton X-100 and OG. The insoluble pellet (P) and soluble supernatant (S) fractions were recovered and 2.5% of each fraction was separated by 14% SDS-PAGE and probed with anti-His antibodies.

Table 5.1. *In vitro* activity of 6xHis-TrxA-WbaP$_{CT}$ solubilized with non-ionic detergents.

| Sample	Activity relative to total enzymatic activity in crude membranes (%)[a]
Crude membranes with no detergent[b]	100.0
Membrane solubilized (detergent - %)	
CHAPS – 0.5	62.4
CHAPS – 1.0	91.8
DDM – 0.5	81.5
DDM – 1.0	44.5
OG – 0.5	65.5
OG – 1.0	76.5
Triton X-100 – 0.5	10.0
Triton X-100 – 1.0	3.5

[a] Total membrane protein (200 µg) was solubilized in buffer with detergents to a final volume of 200 µl. 50 µL of the soluble fraction corresponding to 50 µg of solubilized membrane protein was tested for *in vitro* enzymatic activity as described in *Materials and Methods*.
[b] 50 µg of crude membranes were tested for *in vitro* enzymatic activity

5.3.3. Biochemical properties and kinetic parameters of purified WbaP$_{CT}$. We measured the effect of different variables on the *in vitro* transferase activity of WbaP$_{CT}$. The optimal pH was determined using fresh Tris-HCl buffers ranging from pH 6 to 9.5. Transferase activity was only detected between pH 7 and 9, being optimal at pH of 8.5 (Fig. 5.6A). Tris-HCl buffer at pH 8.5 was therefore used for all subsequent reactions. The purification of 6xHis-TrxA-WbaP$_{CT}$ was performed under high salt conditions (300 mM NaCl) to prevent non-specific interactions. Testing the effect of NaCl and KCl on WbaP$_{CT}$ activity we found that both salts inhibited the reaction resulting in decreased activity between 0 to 1000 mM (Fig. 5.6B). Therefore, upon cleavage of the 6xHis-TrxA, WbaP$_{CT}$ was concentrated by a filtration spin column and exchanged with salt-free buffer to prevent the inhibitory effects of NaCl. *In vitro* assays with crude membrane containing WbaP have previously shown that divalent cations Mg^{2+} or Mn^{2+} are required for activity (15-16). To determine the optimal concentration of metal ion cofactor required for the activity of WbaP$_{CT}$, we assayed the protein at concentrations of 0 to 150 mM of $MgCl_2$. Without added $MgCl_2$ no activity could be detected; activity was optimal at 25 mM, while higher concentrations had an inhibitory effect (Fig. 5.6C).

We next investigated the kinetic parameters of WbaP$_{CT}$. Product formation was assayed over time at various concentrations of protein. The transferase activity profile was linear up to 15 min for concentrations of protein up to 4 ng/µl (Fig 5.7). To determine the steady state kinetics of WbaP$_{CT}$ we utilized optimized reaction conditions (50 mM Tris-HCl [pH 8.5], 25 mM $MgCl_2$, 1% CHAPS) and tested the substrates UDP-Gal and Und-P at concentrations ranging from 0.2 to 2.5 µM and 10 to 1000 µM, respectively. Typical Michaelis-Menten kinetics were observed for both substrates (Fig 5.8A and B) with a K_m of 0.548 ± 0.061 µM, a V_{max} of 2448 ± 93.45 pmol/mg/min and k_{cat} of 4.025 sec^{-1} for UDP-Gal, and a K_m of 78.54 ± 0.015 µM, a V_{max} of 1145 ± 65.06 pmol/mg/min for Und-P.

5.3.4. Substrate specificity of pure WbaP$_{F258-Y476}$ for the lipid carrier. We also determined whether WbaP$_{CT}$ retains specificity for C_{55} Und-P. The *in vitro* activity of purified WbaP$_{CT}$ was tested with phosphoisoprenoids (PI) of varying lengths ranging
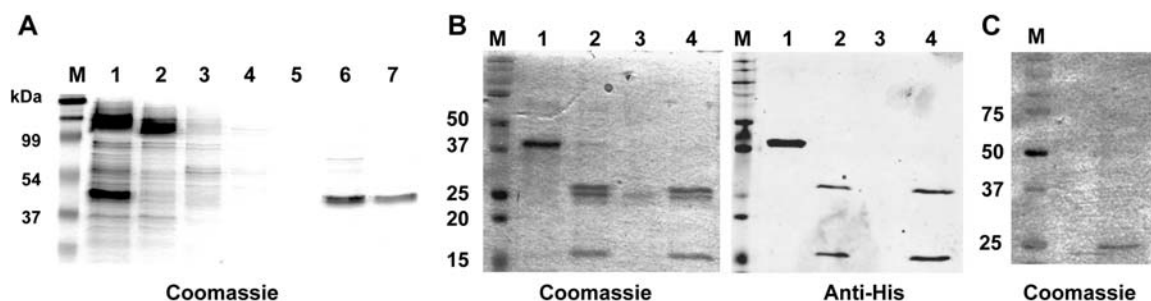
Figure 5.5. Purification of TrxA-WbaP$_{R258-Y476}$. (A) Purification of 6xHis-TrxA-WbaP$_{258-476}$ from C43(DE3) total membranes and analysis by 16% SDS-PAGE and Coomassie staining. Lane 1, DDM solubilization; Lane 2, flowthrough; Lane 3, Wash 1 (30mM imidazole); Lanes 4 and 5, Wash 2 and 3 (60mM imidazole); Lane 6, elution 1 (250mM imidazole); Lane 7, elution 2 (250mM imidazole). Purified protein migrates at 40 kDa (B) Analysis of TEV cleavage of 6xHis-TrxA-WbaP$_{258-476}$ by 16% SDS-PAGE followed by Western blot with anti-His antibodies (right) and by Coomassie staining (left). Lanes 1, purified 6xHis-TrxA-WbaP$_{258-476}$ migrates at 40 kDa ; Lanes 2, 6xHis-TrxA-WbaP$_{258-476}$ after TEV cleavage. Cleaved 6xHis-TrxA migrates at 17 kDa and WbaP$_{258-476}$ migrates at 25 kDa and is not detected by Western blot. The 6xHis TEV protease migrates at 27 kDa. Lanes 3, post TEV cleavage flow through. Lanes 4, elution with 500mM imidazole releases all bound protein. (C) 12% SDS-PAGE and Coomassie staining of concentrated post TEV cleavage flow through.
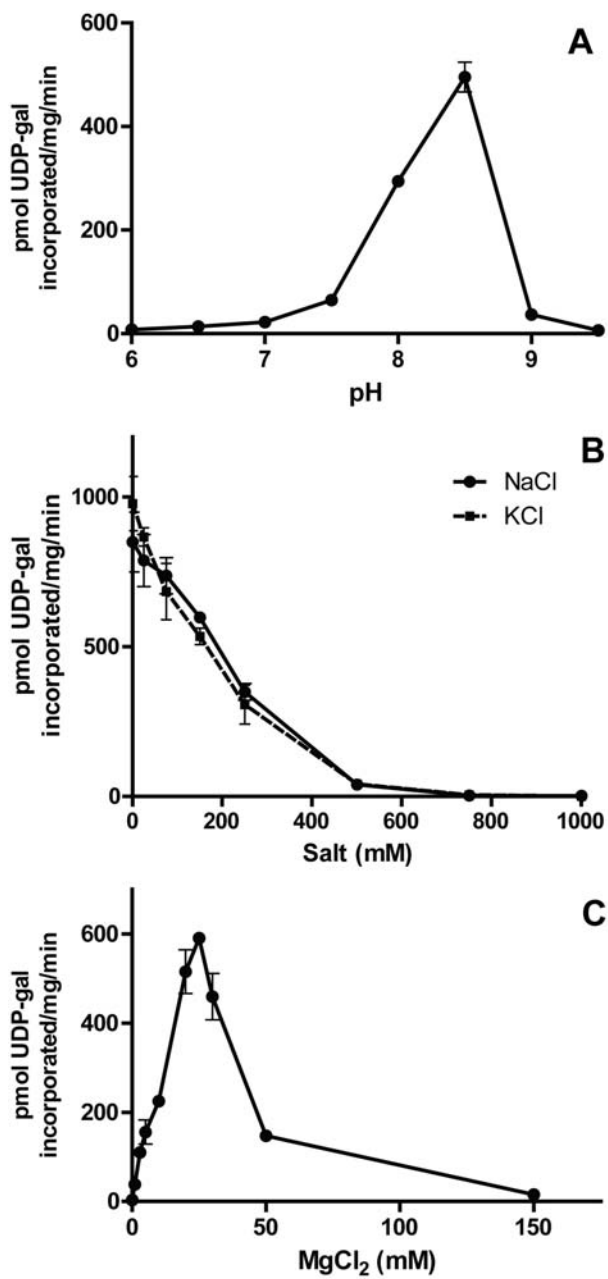
Figure 5.6. Biochemical properties of pure WbaP$_{R258-Y476}$. The effects of (A) pH, (B) the salts NaCl and KCl and (C) MgCl$_2$ were investigated. Each data point represents the mean of three experiments with standard deviations.
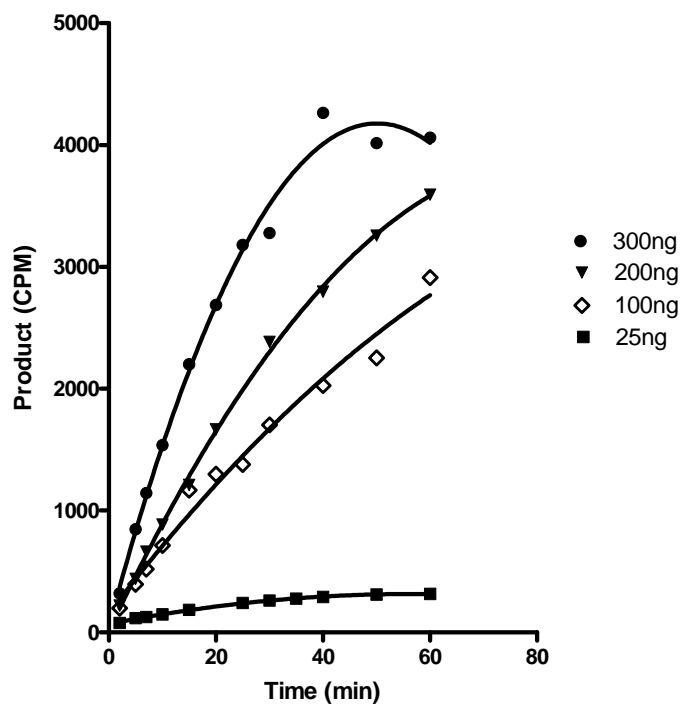
Figure 5.7. Und-P-P-galactose product formation over time. The activity of varying amounts of WbaPCT protein was assayed *in vitro* over time. Activity is reported as counts per minute (CPM).
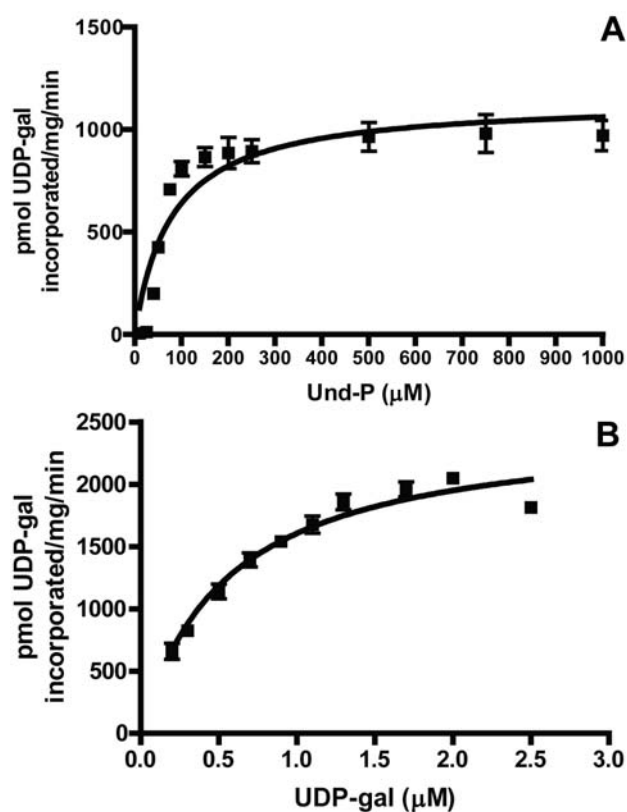
Figure 5.8. Steady state kinetics of WbaP$_{R258-Y476}$. (A) A range of 0.2 to 2.5 μM ^{14}C labeled UDP-gal and (B) 10 to 1000 μM Und-P was assayed under optimized reaction conditions and were carried out for 15 min at 37μC in triplicate as described in *Materials and Methods*. Each point represents the mean of three experiments with standard deviations.
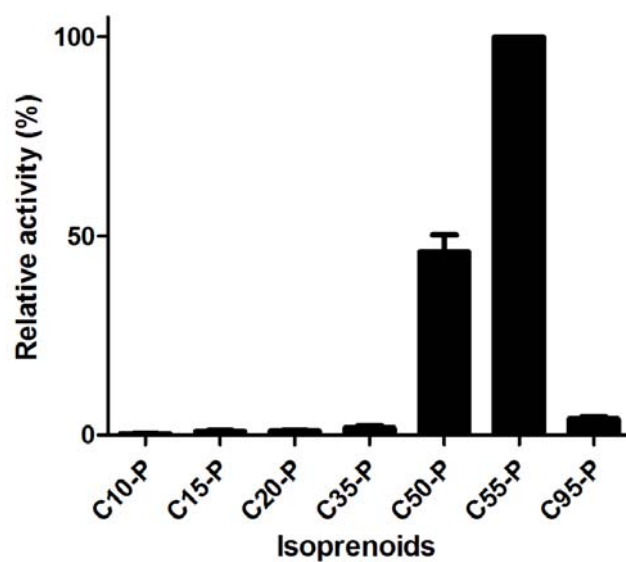
Figure 5.9. Specificity of WbaP$_{R258-Y476}$ for the lipid carrier. A concentration of 100μM was used for each PI tested in optimized *in vitro* assay conditions as described in *Materials and Methods.* Data represent the mean of three independent experiments with standard deviations .

from C10-P to C95-P (Fig. 5.9). Minimal activity was detected with the shorter PIs C10-P, C15-P, C20-P and C35-P at 0.28%, 0.91%, 1% and 1.8% of C55-P activity respectively. A similar result was also obtained with the eukaryotic lipid carrier, C95-P, at 4.02% activity of C55-P. C50-P, however resulted in 49.53% activity of C55-P. These results indicate that WbaP$_{CT}$ retains specificity for Und-P.

5.4.   Concluding remarks

To our knowledge, *in vitro* studies on WbaP as well as other hexose-1-transferases such as CpsE and WsaP have relied on crude membrane preparations (3, 16, 19, 21). Here, we have purified the WbaP$_{CT}$ domain, which has allowed us characterize the transferase reaction *in vitro* assay using chemically defined components. A striking observation in our study was the strong specificity of WbaP$_{CT}$ for Und-P. Kinetic analysis of purified PglC with polyisoprenols of different length, double bond geometry, and degree of saturation have also revealed a preference for the native substrate (5). An interaction of the enzyme with the Und-P (C55-P) acceptor presumably occurs in the inner membrane via regions of the protein within or at the boundaries of transmembrane helices. In previous work, we showed the cytosolic tail of WbaP fused to TMI resulted in a protein without function, suggesting that the TMV is important for the catalytic activity of the enzyme (19). A closer examination of the TMV helix, which is present in the purified WbaP$_{CT}$ protein, reveals the sequence LIIASPLMIYLWY. This sequence is strikingly similar to the 13-aa consensus sequence LL(F/I)IXFXXIPFXFY, which was proposed to be important for the recognition of polyisoprenolphosphates (1, 30). NMR and molecular modeling of peptides containing the consensus sequence suggest that one TM region would be sufficient to interact with Und-P (30-31). Structural studies of this WbaP$_{CT}$ domain, currently underway in our laboratory will provide more detailed information on the mechanism of catalysis as well as the residues that make contact with substrate and cofactor molecules.

5.5.   Chapter 5 References

1.   **Albright, C. F., P. Orlean, and P. W. Robbins.** 1989. A 13-amino acid peptide in three yeast glycosyltransferases may be involved in dolichol recognition. Proc Natl Acad Sci U S A **86:**7366-9.

2.   **Branch, C. L., G. Burton, and S. F. Moss.** 1999. ChemInform Abstract: An Expedient Synthesis of Allylic Polyprenyl Phosphates. ChemInform **30:**no-no.

3.   **Cartee, R. T., W. T. Forsee, M. H. Bender, K. D. Ambrose, and J. Yother.** 2005. CpsE from type 2 *Streptococcus pneumoniae* catalyzes the reversible addition of glucose-1-phosphate to a polyprenyl phosphate acceptor, initiating type 2 capsule repeat unit formation. J Bacteriol **187:**7425-33.

4.   **Chamot-Rooke, J., B. Rousseau, F. Lanternier, G. Mikaty, E. Mairey, C. Malosse, G. Bouchoux, V. Pelicic, L. Camoin, X. Nassif, and G. Dumenil.** 2007. Alternative *Neisseria* spp. type IV pilin glycosylation with a glyceramido acetamido trideoxyhexose residue. Proc Natl Acad Sci U S A **104:**14783-8.

5.   **Chen, M. M., E. Weerapana, E. Ciepichal, J. Stupak, C. W. Reid, E. Swiezewska, and B. Imperiali.** 2007. Polyisoprenol specificity in the *Campylobacter jejuni* N-linked glycosylation pathway. Biochemistry **46:**14342-8.

6.   **Chojnacki, T., W. Jankowski, T. Mankowski, and W. Sasak.** 1975. Preparative separation of naturally occurring mixtures of polyprenols on hydroxyalkoxypropyl-Sephadex. Anal Biochem **69:**114-9.

7.   **Drummelsmith, J., and C. Whitfield.** 1999. Gene products required for surface expression of the capsular form of the group 1 K antigen in *Escherichia coli* (O9a:K30). Mol Microbiol **31:**1321-32.

8.   **Geertsma, E. R., M. Groeneveld, D. J. Slotboom, and B. Poolman.** 2008. Quality control of overexpressed membrane proteins. Proc Natl Acad Sci U S A **105:**5722-7.

9.   **Glover, K. J., E. Weerapana, M. M. Chen, and B. Imperiali.** 2006. Direct biochemical evidence for the utilization of UDP-bacillosamine by PglC, an essential glycosyl-1-phosphate transferase in the *Campylobacter jejuni* N-linked glycosylation pathway. Biochemistry **45:**5343-50.

10.   **Hartley, M. D., M. J. Morrison, F. E. Aas, B. Borud, M. Koomey, and B. Imperiali.** 2011. Biochemical Characterization of the O-Linked Glycosylation Pathway in *Neisseria gonorrhoeae* Responsible for Biosynthesis of Protein Glycans Containing N,N'-Diacetylbacillosamine. Biochemistry.

11. **LaVallie, E. R., E. A. DiBlasio, S. Kovacic, K. L. Grant, P. F. Schendel, and J. M. McCoy.** 1993. A thioredoxin gene fusion expression system that circumvents inclusion body formation in the *E. coli* cytoplasm. Biotechnology (N Y) **11:**187-93.

12. **Maniatis, T.** 1982. Molecular cloning : a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

13. **Merino, S., N. Jimenez, R. Molero, L. Bouamama, M. Regue, and J. M. Tomas.** 2011. A UDP-HexNAc:Polyprenol-P GalNAc-1-P Transferase (WecP) Representing a New Subgroup of the Enzyme Family. J Bacteriol **193:**1943-52.

14. **Miroux, B., and J. E. Walker.** 1996. Over-production of proteins in Escherichia coli: mutant hosts that allow synthesis of some membrane proteins and globular proteins at high levels. J Mol Biol **260:**289-98.

15. **Osborn, M. J., S. M. Rosen, L. Rothfield, and B. L. Horecker.** 1962. Biosynthesis of bacterial lipopolysaccharide. I. Enzymatic incorporation of galactose in a mutant strain of *Salmonella*. Proc Natl Acad Sci U S A **48:**1831-8.

16. **Patel, K. B., S. E. Furlong, and M. A. Valvano.** 2010. Functional analysis of the C-terminal domain of the WbaP protein that mediates initiation of O antigen synthesis in *Salmonella enterica*. Glycobiology **20:**1389-401.

17. **Perez, J. M., M. A. McGarry, C. L. Marolda, and M. A. Valvano.** 2008. Functional analysis of the large periplasmic loop of the *Escherichia coli* K-12 WaaL O-antigen ligase. Mol Microbiol **70:**1424-40.

18. **Power, P. M., L. F. Roddam, M. Dieckelmann, Y. N. Srikhanta, Y. C. Tan, A. W. Berrington, and M. P. Jennings.** 2000. Genetic characterization of pilin glycosylation in Neisseria meningitidis. Microbiology **146 (Pt 4):**967-79.

19. **Saldías, M. S., K. Patel, C. L. Marolda, M. Bittner, I. Contreras, and M. A. Valvano.** 2008. Distinct functional domains of the *Salmonella enterica* WbaP transferase that is involved in the initiation reaction for synthesis of the O antigen subunit. Microbiology **154:**440-53.

20. **Schaffer, C., T. Wugeditsch, P. Messner, and C. Whitfield.** 2002. Functional expression of enterobacterial O-polysaccharide biosynthesis enzymes in *Bacillus subtilis*. Appl Environ Microbiol **68:**4722-30.

21. **Steiner, K., R. Novotny, K. Patel, E. Vinogradov, C. Whitfield, M. A. Valvano, P. Messner, and C. Schaffer.** 2007. Functional characterization of the initiation enzyme of S-layer glycoprotein glycan biosynthesis in *Geobacillus stearothermophilus* NRS 2004/3a. J Bacteriol **189:**2590-8.

22. **Stevenson, G., K. Andrianopoulos, M. Hobbs, and P. R. Reeves.** 1996. Organization of the *Escherichia coli* K-12 gene cluster responsible for production of the extracellular polysaccharide colanic acid. J Bacteriol **178:**4885-93.

23. **Toh, E., H. D. Kurtz, Jr., and Y. V. Brun.** 2008. Characterization of the *Caulobacter crescentus* holdfast polysaccharide biosynthesis pathway reveals significant redundancy in the initiating glycosyltransferase and polymerase steps. J Bacteriol **190:**7219-31.

24. **Valvano, M. A.** 2003. Export of O-specific lipopolysaccharide. Front Biosci **8:**s452-71.

25. **Valvano, M. A., S. E. Furlong, and K. B. Patel.** 2011. Genetics, biosynthesis and assembly of O-antigen. *In* Y. A. Knirel and M. A. Valvano (ed.), Bacterial Lipopolysaccharides. Structure, Chemical Synthesis, Biogenesis and Interaction with Host Cells. Springer

26. **Wang, L., D. Liu, and P. R. Reeves.** 1996. C-terminal half of *Salmonella enterica* WbaP (RfbP) is the galactosyl-1-phosphate transferase domain catalyzing the first step of O-antigen synthesis. J Bacteriol **178:**2598-604.

27. **Wang, L., and P. R. Reeves.** 1994. Involvement of the galactosyl-1-phosphate transferase encoded by the *Salmonella enterica* rfbP gene in O-antigen subunit processing. J Bacteriol **176:**4348-56.

28. **Wellburn, A. R., and F. W. Hemming.** 1966. The occurrence and seasonal distribution of higher isoprenoid alcohols in the plant kingdom. Phytochemistry **5:**969-975.

29. **Ye, X. Y., M. C. Lo, L. Brunner, D. Walker, D. Kahne, and S. Walker.** 2001. Better substrates for bacterial transglycosylases. J Am Chem Soc **123:**3155-6.

30. **Zhou, G. P., and F. A. Troy, 2nd.** 2003. Characterization by NMR and molecular modeling of the binding of polyisoprenols and polyisoprenyl recognition sequence peptides: 3D structure of the complexes reveals sites of specific interactions. Glycobiology **13:**51-71.

31. **Zhou, G. P., and F. A. Troy, 2nd.** 2005. NMR study of the preferred membrane orientation of polyisoprenols (dolichol) and the impact of their complex with polyisoprenyl recognition sequence peptides on membrane structure. Glycobiology **15:**347-59.

Chapter 6

Discussion

6.1. Characterization of PHPT proteins

Proteins of the PHPT family initiate the synthesis of a variety of glycans of medical, agricultural and industrial importance. However the study of these proteins has been hindered by difficulties generally associated with integral membrane proteins. At the initiation of this work the C-terminal domain of WbaP was implicated in harboring the GT (galactosyltransferase) function (13), but little else was known about this protein and other PHPT members. The main goal of this thesis was to understand the various domains of PHPT proteins and their contribution to the activity of the enzymes both *in vivo* and *in vitro*. Within this body of work, a specific aim was to purify a PHPT protein to enable future structural studies and *in vitro* analysis in the absence of contaminating lipids and proteins.

6.1.1. C-terminal domain of WbaP is sufficient for galactose-1-phosphate transferase activity

Work in this thesis has established that only the C-terminal domain of WbaP comprising TM V and the tail region is required for galactose-1-phosphate transferase activity (chapter 2). Although *in vitro* activity of this region was established by Wang el al (1996) (13), the authors could not detect *in vivo* activity suggesting the role of the N-terminus in processing the Und-PP-linked O unit. In chapter 2 we show that this was most likely due to mutations elsewhere as the C-terminal domain of WbaP complements a *wbaP* deficient mutant and is also active *in vitro*. The importance of TM V is also established as a fusion of the tail region to TM I abrogates activity. This thesis has also confirmed that specificity for Und-P is also maintained in this region. In chapter 5 we show that purified WbaPCT has maximal activity with Und-P versus isoprenoids of shorter and longer chain lengths. As this region is only predicted to contain one TM region, it would suggest that this region interacts with Und-P. Together this work demonstrates that only the C-terminal domain of PHPT proteins is required to study hexose-1-transferase function.

6.1.2. WbaP contains multiple conserved residues that are critical for function in the cytoplasmic tail.

The presence of highly conserved residues within PHPT proteins has been identified by other laboratories (9, 11-12). However this thesis is the first study to investigate the majority of these residues. As the location of the C-terminal tail had not been established it was critical to first confirm that the tail region was in fact cytoplasmic, as residues required for catalytic activity would have to reside in the cytoplasm to transfer galactose-1-phosphate to Und-P. In chapter 3 we confirmed cytoplasmic localization of the C-terminus by a GFP fusion and also through protease accessibility experiments. The results revealed the last 20 Kda of the protein are cytoplasmic. Highly conserved residues within this region were identified using the Sequence Alignment and Modeling system (SAM-T02), which aligns hundreds of similar proteins. The observation that seven of the ten highly conserved residues were absolutely required for function suggests that WbaP requires the involvement of many residues for sugar binding, interactions with cofactor, for the catalytic mechanism and/or proper folding. Mutation of these residues conserving charge and or R group structures could not rescue the phenotype demonstrating the importance of the specific residues in specific locations within the protein. The location of critical residue D458 near the end of the protein provides an explanation for the loss of WbaP activity when fusions are made at the C-terminus.

6.1.3. The loop domain of WbaP affects O antigen chain length distribution.

The large putative soluble loop of WbaP was investigated in chapter 2. This region has a large size and predicted secondary structure containing a Rossmann fold, which is a structural motif that binds nucleotides (2) however, this soluble region is predicted by topology prediction programs to reside in the periplasm which does not contain nucleotides (chapter 2). We found that an absence of this region or overexpression of this region led to a change in the distribution of O antigen chain lengths produced by the cell (chapter 2). At the time of the work we thought this may be due to interactions with other O antigen biosynthetic proteins such as Wzz or Wzy whose activities reside in the periplasm. Recently, Woodward *et al* (2010) (14) demonstrated O antigen polymerization and regulation *in vitro* of the *E. coli* O86 repeating unit with

purified Wzz and Wzy proteins from *E. coli* and Und-PP-GlcNAc. However, this does not eliminate the possibility that WbaP may still affect the regulation of chain length in *S. enterica*. In chapter 4 it was interesting to see that although overexpression of both WcaJ and PssY in LT2 led to a decrease in O antigen, the chain length distribution differed between the two proteins. WcaJ has the soluble loop intact and led to the loss of higher O antigen polymers compared to PssY, which does not contain this region. The soluble loop domain has also been of interest in CpsE of *Streptococcus pneumoniae* where it is believed to play a role in regulation of capsule synthesis (15).

To gain a better understanding of the role of this region we attempted to determine its localization. Both spheroplasts and total membrane vesicles from cells expressing WbaP constructs tagged at either the N or C-terminus were treated with trypsin, however we were not able to detect any cleavage within this region, although multiple cleavage sites exist (chapter 3). This may mean that this region assumes a tertiary structure preventing access of trypsin. In addition, TrxA fused to the N-terminus of WbaPCT localized to the cytoplasm, as demonstrated through TEV protease studies, which targeted a TEV site between WbaPCT and the TrxA fusion partner (chapter 3). Together these results suggest that the soluble external loop domain may reside in the cytoplasm. The effect of the chain length modality could be explained by via a regulatory mechanism involving interactions of the UDP-sugar with the Rossmann fold motif, perhaps *via* allosteric interactions

6.1.4. PHPT proteins WcaJ and PssY can accommodate more than one UDP-sugar substrate.

Investigation of putative PHPT proteins from *E. coli* (WcaJ) and *C. crescentus* (HfsE, PssY, PssZ) added to the knowledge gained by work on the PHPT model protein WbaP. These proteins were believed to initiate the synthesis of colanic acid in *E. coli* and the holdfast adhesion in *C. crescentus* (10-11). In chapter 4 we studied these proteins to determine their substrates. *In vitro* assays confirmed that UDP-glucose was a substrate for both WcaJ and Pssy, however substrates for HfsE and PssZ were not identified. An interesting observation was that WcaJ and PssY could also utilize UDP-galactose *in vivo* in a *wbaP* deficient strain however no activity was detected *in vitro*. *In vivo* product

formation during incubation of the cultures overnight leads to the accumulation of product whereas in an *in vitro* reaction, only product formed within 15-30 min is measured. The *in vitro* activity of *S. pneumoniae* WchA was assayed with various UDP sugars and similar to our results, activity was only detected with UDP-glucose (8). As the function of HfsE, PssY and PssZ has been shown to be redundant for holdfast production (11) the lack of activity of HfsE and PssZ in our systems may be due to the requirement of a different sugar substrate. Although UDP-glucose may be the primary substrate of PssY it may have the ability to accommodate a sugar utilized by HfsE and PssZ.

6.1.5. TrxA can be used as a tool to overexpress, solubilize and purify WbaP$_{CT}$.

Purification of PHPT proteins has been limited to PglC of *Campylobacter jejuni* and PglB of *Neisseria* species. Unlike PHPT proteins studied to date which utilize a UDP-hexose sugar, both of these utilize UDP-bacillosamine and initiate the synthesis of precursors for glycosylation pathways in these organisms (3-4). During the course of this thesis a PHPT protein handling a UDP-hexose sugar has not been purified. Although groups have attempted the purification of a hexose-1-phosphate transferase the major set-back has been the production of inclusion bodies (8-9). Even in the absence of the first four predicted TM regions the C-terminal domain of WbaP$_{CT}$ is predominantly misfolded upon overexpression (chapter 3, chapter 5). Although misfolded, these proteins can be found in total membrane fractions. We found that fusion of the 12-kDa oxidoreductase TrxA to WbaP$_{CT}$ resulted in the production of properly folded recombinant protein (chapter 3). This protein was also soluble in the presence of non-ionic detergents and could be purified (chapter 5). Upon purification, the TrxA fusion can be removed via cleavage of a TEV site engineered on the plasmid between the *wbaP$_{CT}$* and *trxA*. This overexpression and purification strategy can be used with other PHPT family proteins as well as other membrane proteins that are difficult to purify.

6.2.   Ongoing and future studies.

Many questions still remain on the characterization of PHPT proteins. Although much information has been gained by taking a genetic approach, the future of this project will require utilization of purified protein and a biochemical approach.

6.2.1. Experimentally determining the topology of PHPT proteins.

Although protease accessibility and GFP experiments have given us an idea of the topology of regions of WbaP, a study of the topology of the whole protein as well as other PHPT proteins is required to determine the cytoplasmic and periplasmic regions of these proteins. This information will help us better understand what roles these regions play based on the presence of other proteins or substrates in their local environment. The topology mapping could use other reporter proteins such as LacZ and PhoA (5), however as work with GFP has shown the majority of the fusion protein is misfolded which may result in false negative results. Another method to determine the topology of membrane proteins is substituted cysteine accessibility method (SCAM) (1), which requires labeling cysteines that have been inserted at desired locations in a cysteine-less version of a membrane protein. This strategy has been successfully used to map the topology of the WecA protein (6). The availability of pure protein would allow for another method of topology determination involving the oxidative labeling of methionines with H_2O_2 and laser to access solvent accessible regions (7).

6.2.2. Determining the role of the essential C-terminal residues in binding UDP-sugars.

In this work seven residues within the cytoplasmic tail region of WbaP were found to be essential for transferase activity, but a role(s) for these has yet to be assigned. As the interaction of PHPT proteins with their UDP-sugar substrates is probably transient, binding assays may not be sensitive enough to detect this interaction. The ability of PHPT proteins to utilize different sugar substrates such as UDP-galactose, UDP-glucose and UDP-GalNAc can aid in identifying the residue(s) in the cytoplasmic C-terminal domains of these proteins that are involved in sugar binding. Chimeras constructed from regions of the cytoplasmic tail from proteins utilizing different primary

sugar substrates can be constructed (e.g. WbaP and WcaJ). The availability of an *in vivo* system to test putative galactosyltransferases (chapter 2) and glucosyltransferases (chapter 4) will allow us to determine the UDP-sugar substrates of putative PHPT members and to study the regions required for UDP-sugar binding.

6.2.3. Structural studies with PHPT proteins.

In the literature membrane protein structures are being solved more frequently. The ability to solve structures of PHPT proteins in association with the lipid carrier, cofactor or sugar together with an understanding of how the proteins are situated in the membrane will enhance our understanding of the mechanism of these proteins and aid in the development of novel inhibitors. The purification of WbaPCT (chapter 5) may facilitate structural studies by crystallography and NMR, as this domain contains all regions required for interaction with substrates for catalytic activity. As the yield of WbaPCT cleaved from the TrxA fusion partner is low, due to association of the cleaved protein with the Ni^{2+} resin, structural studies of the protein with the N-terminal TrxA fusion will also be attempted. This will require the production of a monodisperse sample, which can be assessed by sedimentation velocity centrifugation. Therefore subsequent to purification, the protein will be further purified by a second chromatography step such as ion exchange or size exclusion chromatography. In addition to WbaPCT, purification of other PHPT proteins including those with all domains could be attempted to determine a role for the N-terminal and soluble loop domains.

6.3.   Concluding remarks and significance of this work.

In summary, this work has established that only the C-terminal domains of PHPT proteins are required for the transferase function including specificity for Und-P; however the soluble external loop region contributes to overall activity of the enzyme. This work also demonstrates, for the first time, the utilization of two UDP-sugar substrates by a single PHPT protein and is the first to report purification of an active hexose-1-phosphate transferase. Purification of WbaPCT and its biochemical characterization have laid the groundwork for *in vitro* study of site directed mutants

created in this research and for structural studies. Together these will lead to determination of the enzyme's active site and the production of novel inhibitors.

6.4. Chapter six references

1.   **Bogdanov, M., W. Zhang, J. Xie, and W. Dowhan.** 2005. Transmembrane protein topology mapping by the substituted cysteine accessibility method (SCAM(TM)): application to lipid-specific membrane protein topogenesis. Methods **36:**148-71.

2.   **Buehner, M., G. C. Ford, D. Moras, K. W. Olsen, and M. G. Rossman.** 1973. D-glyceraldehyde-3-phosphate dehydrogenase: three-dimensional structure and evolutionary significance. Proc Natl Acad Sci U S A **70:**3052-4.

3.   **Glover, K. J., E. Weerapana, M. M. Chen, and B. Imperiali.** 2006. Direct biochemical evidence for the utilization of UDP-bacillosamine by PglC, an essential glycosyl-1-phosphate transferase in the *Campylobacter jejuni N*-linked glycosylation pathway. Biochemistry **45:**5343-50.

4.   **Hartley, M. D., M. J. Morrison, F. E. Aas, B. Borud, M. Koomey, and B. Imperiali.** 2011. Biochemical Characterization of the O-Linked Glycosylation Pathway in *Neisseria gonorrhoeae* Responsible for Biosynthesis of Protein Glycans Containing N,N'-Diacetylbacillosamine. Biochemistry.

5.   **Islam, S. T., V. L. Taylor, M. Qi, and J. S. Lam.** 2010. Membrane topology mapping of the O-antigen flippase (Wzx), polymerase (Wzy), and ligase (WaaL) from *Pseudomonas aeruginosa* PAO1 reveals novel domain architectures. MBio **1**.

6.   **Lehrer, J., K. A. Vigeant, L. D. Tatar, and M. A. Valvano.** 2007. Functional characterization and membrane topology of *Escherichia coli* WecA, a sugar-phosphate transferase initiating the biosynthesis of enterobacterial common antigen and O-antigen lipopolysaccharide. J Bacteriol **189:**2618-28.

7.   **Pan, Y., L. Brown, and L. Konermann.** 2010. Site-directed mutagenesis combined with oxidative methionine labeling for probing structural transitions of a membrane protein by mass spectrometry. Journal of The American Society for Mass Spectrometry **21:**1947-1956.

8.   **Pelosi, L., M. Boumedienne, N. Saksouk, J. Geiselmann, and R. A. Geremia.** 2005. The glucosyl-1-phosphate transferase WchA (Cap8E) primes the capsular polysaccharide repeat unit biosynthesis of *Streptococcus pneumoniae* serotype 8. Biochem Biophys Res Commun **327:**857-65.

9.   **Steiner, K., R. Novotny, K. Patel, E. Vinogradov, C. Whitfield, M. A. Valvano, P. Messner, and C. Schaffer.** 2007. Functional characterization of the initiation enzyme of S-layer glycoprotein glycan biosynthesis in *Geobacillus stearothermophilus* NRS 2004/3a. J Bacteriol **189:**2590-8.

10. **Stevenson, G., K. Andrianopoulos, M. Hobbs, and P. R. Reeves.** 1996. Organization of the *Escherichia coli* K-12 gene cluster responsible for production of the extracellular polysaccharide colanic acid. J Bacteriol **178:**4885-93.

11. **Toh, E., H. D. Kurtz, Jr., and Y. V. Brun.** 2008. Characterization of the *Caulobacter crescentus* holdfast polysaccharide biosynthesis pathway reveals significant redundancy in the initiating glycosyltransferase and polymerase steps. J Bacteriol **190:**7219-31.

12. **Videira, P. A., A. P. Garcia, and I. Sa-Correia.** 2005. Functional and topological analysis of the *Burkholderia cenocepacia* priming glucosyltransferase BceB, involved in the biosynthesis of the cepacian exopolysaccharide. J Bacteriol **187:**5013-8.

13. **Wang, L., D. Liu, and P. R. Reeves.** 1996. C-terminal half of *Salmonella enterica* WbaP (RfbP) is the galactosyl-1-phosphate transferase domain catalyzing the first step of O-antigen synthesis. J Bacteriol **178:**2598-604.

14. **Woodward, R., W. Yi, L. Li, G. Zhao, H. Eguchi, P. R. Sridhar, H. Guo, J. K. Song, E. Motari, L. Cai, P. Kelleher, X. Liu, W. Han, W. Zhang, Y. Ding, M. Li, and P. G. Wang.** 2010. In vitro bacterial polysaccharide biosynthesis: defining the functions of Wzy and Wzz. Nat Chem Biol **6:**418-23.

15. **Xayarath, B., and J. Yother.** 2007. Mutations blocking side chain assembly, polymerization, or transport of a Wzy-dependent *Streptococcus pneumoniae* capsule are lethal in the absence of suppressor mutations and can affect polymer transfer to the cell wall. J Bacteriol **189:**3369-81.

Appendix – Copyright releases

Chapter 2: Saldías, M.S., Patel, K., Marolda, C. L., Bittner, M., Contreras, I. & Valvano, M. A. (2008). Distinct functional domains of the *Salmonella enterica* WbaP transferase that is involved in the initiation reaction for synthesis of the O antigen subunit. Microbiology, 154, 440-453.

Received following response:

Thank you for your permission request. SGM no longer responds to individual permission requests; these are now handled by the Copyright Clearance Center.

If you are an author (or an author's representative) of the content for which you are requesting permission:

The copyright transfer agreement/licence to publish that was signed by the authors gives express permission for authors to reuse their own content, provided that the journal is acknowledged as the original source. No further permission is required.

Chapter 3: Patel, K. B., Furlong, S. F. & Valvano, M. A. (2010). Functional analysis of the C-terminal domain of the WbaP protein that mediates initiation of O antigen synthesis in *Salmonella enterica.* Glycobiology, 20, 1389-1401.

AND

Chapter 5: Patel, K. B., Ciepichal, E., Swiezewska, E. & Valvano, M. A. (In Press). The C-terminal domain of the *Salmonella enterica* WbaP (UDP-galactose:Und-P galactose-1-phosphate transferase) is sufficient for catalytic activity and specificity for undecaprenyl monophosphate . Glycobiology.

The following is posted on the Oxford Journals website:

Rights retained by ALL Oxford Journal Authors

- The right, after publication by Oxford Journals, to use all or part of the Article and abstract, for their own personal use, including their own classroom teaching purposes;
- The right, after publication by Oxford Journals, to use all or part of the Article and abstract, in the preparation of derivative works, extension of the article into book-length or in other works, provided that a full acknowledgement is made to the original publication in the journal;
- The right to include the article in full or in part in a thesis or dissertation, provided that this not published commercially;

For the uses specified here, please note that there is no need for you to apply for written permission from Oxford University Press in advance. Please go ahead with the use ensuring that

a full acknowledgment is made to the original source of the material including the journal name, volume, issue, page numbers, year of publication, title of article and to Oxford University Press and/or the learned society.

Curriculum Vitae – Kinnari Patel

Post-secondary Education and Degrees:	University of Waterloo Waterloo, Ontario, Canada Science and Business, Biology option (Co-op), B.Sc 1999-2004 The University of Western Ontario London, Ontario, Canada Microbiology and Immunology, Ph.D candidate 2006-present
Honours and Awards:	FASEB travel award, 2011 Ontario Graduate Scholarship in Science and Technology (OGSST), 2009-2010 UWO Department of Microbiology and Immunology Travel Awards, 2008 and 2009
Relevant Work Experience:	UWO Department of Microbiology and Immunology Teaching Assistant, 2006-2010 Associate, Preclinical Development, Innovative Drug Development, Apotex Inc. , 2004-2005 Pre-clinical assistant, Innovative Drug Development, Apotex Inc., 2003 Research Assistant, Life Sciences Branch, Industry Canada, 2002 Junior Policy Analyst, Therapeutic Products Directorate, Health Canada, 2001

Publications:

Patel, KB., Ciepichal, E., Swiezewska, E. & Valvano, MA. (*In Press*). The C-terminal domain of the *Salmonella enterica* WbaP (UDP-galactose:Und-P galactose-1-phosphate transferase) is sufficient for catalytic activity and specificity for undecaprenyl monophosphate . *Glycobiology*.

Valvano MA, Furlong SE, **Patel KB**. (2011) O antigen biosynthesis. *Comprehensive Natural Products Chemistry*. L. N. Mander and H.-W. Liu. Oxford, Elsevier Ltd. Springer.

Patel KB, Furlong SE, Valvano MA. (2010) Functional analysis of the C-terminal domain of the WbaP protein that mediates initiation of O antigen synthesis in *Salmonella enterica*. *Glycobiology*. 20(11):1389-401.

Saldías MS, **Patel K***, Marolda CL, Bittner M, Contreras I, Valvano MA. (2008) Distinct functional domains of the *Salmonella enterica* WbaP transferase that is involved in the initiation reaction for synthesis of the O antigen subunit. *Microbiology*. 154:440-53.
*Co-first author

Steiner K, Novotny R, **Patel K**, Vinogradov E, Whitfield C, Valvano MA, Messner P, Schäffer C. (2007) Functional characterization of the initiation enzyme of S-layer glycoprotein glycan biosynthesis in *Geobacillus stearothermophilus* NRS 2004/3a. *Journal of Bacteriology*. 189(7):2590-8

Poster Presentations:

Kinnari B. Patel, Evelyn Toh, Ximena B. Fernandez, Anna Hanuszkiewicz, Yves V. Brun, Mark Bernards, and Miguel A. Valvano. Characterization of UDP-Glucose:Undecaprenyl-Phosphate Glucose-1-Phosphate Transferases of *Escherichia coli* and *Caulobacter crescentus.* Microbial Polysaccharides of Medical, Agricultural, and Industrial Importance. Carefree, Arizona. (2011)

Kinnari B. Patel and Miguel A. Valvano. The C-terminal domain of the UDP-galactose:Und-P galactose-1-phosphate transferase WbaP is sufficient for catalytic activity and specificity for undecaprenyl monophosphate. Microbial Polysaccharides of Medical, Agricultural, and Industrial Importance. Carefree, Arizona. (2011)

Kinnari B Patel, M. Soledad Saldias, and Miguel A. Valvano. Investigating the role of HfsE, PssY and PssZ in holdfast synthesis. Infection and Immunity Research Forum. London, Ontario. (2010)

Kinnari B Patel, Sarah E. Furlong, and Miguel A. Valvano. Structure-function of the C-terminal domain of the *Salmonella enterica* undecaprenylphophate:galactose-1-phosphate transferase WbaP. American Society of Microbiology General Meeting. San Diego, California. (2010)

Kinnari B Patel, Sarah E. Furlong, and Miguel A. Valvano**.** Identifying the regions of WbaP critical for function and O-antigen chain length distribution in *Salmonella*

enterica. Microbial Polysaccharides of Medical, Agricultural, and Industrial Importance. Carefree, Arizona. (2008)

Kinnari Patel, M. Soledad Saldias, and Miguel A. Valvano**.** Identifying the regions of WbaP critical for function and interaction with O-antigen biosynthesis proteins in *Salmonella enterica.* Infection and Immunity Research Forum. London, Ontario. (2007)

Oral Presentation:

Kinnari B. Patel, Evelyn Toh, Ximena B. Fernandez, Anna Hanuszkiewicz, Yves V. Brun, Mark Bernards, and Miguel A. Valvano. Characterization of UDP-Glucose:Undecaprenyl-Phosphate Glucose-1-Phosphate Transferases of *Escherichia coli* and *Caulobacter crescentus.* Microbial Polysaccharides of Medical, Agricultural, and Industrial Importance. Carefree, Arizona. (2011)