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

M. Soledad Saldías,^{1,2†} Kinnari Patel,^{2†} Cristina L. Marolda,²
Mauricio Bittner,^{1‡} Inés Contreras¹ and Miguel A. Valvano^{2,3}

¹Departamento de Bioquímica y Biología Molecular, Facultad de Ciencias Químicas y Farmacéuticas, Universidad de Chile, PO Box 174 Correo 22, Santiago, Chile

²Infectious Diseases Research Group, Siebens–Drake Research Institute, Department of Microbiology and Immunology, University of Western Ontario, London, ON N6A 5C1, Canada

³Infectious Diseases Research Group, Siebens–Drake Research Institute, Department of Medicine, University of Western Ontario, London, ON N6A 5C1, Canada

Correspondence
Miguel A. Valvano
mvalvano@uwo.ca

Inés Contreras
icontrer@uchile.cl

Received 10 September 2007
Revised 11 October 2007
Accepted 19 October 2007

WbaP is a membrane enzyme that initiates O antigen synthesis in *Salmonella enterica* by catalysing the transfer of galactose 1-phosphate (Gal-1-P) onto undecaprenyl phosphate (Und-P). WbaP possesses at least three predicted structural domains: an N-terminal region containing four transmembrane helices, a large central periplasmic loop, and a C-terminal domain containing the last transmembrane helix and a large cytoplasmic tail. In this work, we investigated the contribution of each region to WbaP function by constructing a series of mutant WbaP proteins and using them to complement O antigen synthesis in $\Delta wbaP$ mutants of *S. enterica* serovars Typhi and Typhimurium. Truncated forms of WbaP lacking the periplasmic loop exhibited altered chain-length distributions in O antigen polymerization, suggesting that this central domain is involved in modulating the chain-length distribution of the O polysaccharide. The N-terminal and periplasmic domains were dispensable for complementation of O antigen synthesis *in vivo*, suggesting that the C-terminal domain carries the sugar-phosphate transferase activity. However, despite the fact that they complemented the synthesis of O antigen in the $\Delta wbaP$ mutant *in vivo*, membrane extracts containing WbaP derivatives without the N-terminal domain failed to transfer radioactive Gal from UDP-Gal into a lipid-rich fraction. These results suggest that the N-terminal region of WbaP, which contains four transmembrane domains, is essential for the insertion or stability of the protein in the bacterial membrane. We propose that the domain structure of WbaP enables this protein not only to function in the transfer of Gal-1-P to Und-P but also to establish critical interactions with additional proteins required for the correct assembly of O antigen in *S. enterica*.

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) (Raetz & Whitfield,

2002; Valvano, 2003). The biosynthesis of LPS is a complex process involving a large number of enzyme activities (Raetz & Whitfield, 2002; Valvano, 2003). 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 (Raetz & Whitfield, 2002; Valvano, 2003). 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.

[†]These authors contributed equally to this work.

[‡]**Present address:** Departamento de Ciencias Biológicas, Facultad de Ciencias de la Salud, Universidad Nacional Andrés Bello, Santiago, Chile.

Abbreviations: PHPT, polyisoprenyl-phosphate hexose-1-phosphate transferase; PNPT, polyisoprenyl phosphate *N*-acetylhexosamine-1-phosphate transferase; Und-P, undecaprenyl phosphate.

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 (Batchelor *et al.*, 1991). 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. An additional ABC-transporter mechanism involving a synthase homologue has also been described (Keenleyside & Whitfield, 1996). 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 (Raetz & Whitfield, 2002).

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 (Alexander & Valvano, 1994; Amer & Valvano, 2000; Feldman *et al.*, 1999; Valvano, 2003), formed by the action of the UDP-GlcNAc:Und-P GlcNAc-1-P transferase encoded by the *wecA* gene (Amer & Valvano, 2001, 2002; Anderson *et al.*, 2000). *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] (Anderson *et al.*, 2000; Lehrer *et al.*, 2007; Price & Momany, 2005; Valvano, 2003). 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*) (Valvano, 2003; Wang & Reeves, 1994; Wang *et al.*, 1996). 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 (Reeves, 1993). 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* (Katzen *et al.*, 1998), *Erwinia amylovora* (Bugert & Geider, 1995), *Streptococcus pneumoniae* (Cartee *et al.*, 2005), *Klebsiella pneumoniae* (Arakawa *et al.*, 1995; Drummelsmith & Whitfield, 1999) and *E. coli* K-12 (Stevenson *et al.*, 1996) among others, and also with proteins involved in S-layer glycoprotein glycan biosynthesis in *Geobacillus stearothermophilus* (Steiner *et al.*,

2007). 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) 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 (Wang & Reeves, 1994). 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 (Wang *et al.*, 1996). Here, we propose that WbaP possesses three major domains and assign functional roles to each predicted domain. Our results suggest that the N-terminal domain of WbaP is required for the insertion or stability of the enzyme in the bacterial membrane, 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.

METHODS

Bacterial strains, plasmids, media and growth conditions.

Bacterial strains and plasmids used in this study are listed in Table 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.

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). DNA sequences were determined using an automated sequencer at the DNA Sequencing Facility, Robarts Research Institute (London, ON, Canada).

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/psipred/>), 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/>).

Cloning of WbaP domains. PCR was carried out with *Pwo*I DNA polymerase (Roche Diagnostics) in a Perkin Elmer 2400 GeneAmp PCR system. Plasmid pSM10 was constructed by PCR amplification of a 2.9 kb fragment containing the *wbaP* and *manB* genes using primers manB3 (5'-CTGGATTTTCGAAGGAGTGGACTAA-3') and 592 (5'-TCGGATATCTTAATACGCACCATCTCGCC-3') with *S. Typhi* Ty2 DNA as template. This fragment was ligated into the *Sma*I site of pAA8 (Table 1). pSM13 is a pSM10 derivative in which the *manB* gene was deleted by digestion with *Sac*I and the plasmid

Table 1. Bacterial strains and plasmids

Abbreviations: Cm, chloramphenicol; Amp, ampicillin; Sp, spectinomycin.

Strain or plasmid	Relevant properties	Source or reference
Strains		
S. Typhi		
MSS1	Ty2, $\Delta wbaP::cat$, Cm ^R	This study
Ty2	<i>S. enterica</i> serovar Typhi, wild-type	Institute of Public Health, Chile
S. Typhimurium		
LT2	<i>S. enterica</i> serovar Typhimurium, wild-type	S. Maloy, San Diego State University
MSS2	LT2, $\Delta wbaP::cat$, Cm ^R	This study
SL1196	LT2, <i>hisD trpA fliC fljB gal xyl wbaP4451(T)</i>	Wilkinson <i>et al.</i> (1972)
SL1197	LT2, <i>hisD trpA fliC fljB gal xyl wbaP4452(T)</i>	Wilkinson <i>et al.</i> (1972)
<i>E. coli</i>		
DH5 α	F ⁻ $\phi 80lacZM15 endA recA hsdR(r_K^- m_K^-) supE thi gyrA relA \Delta(lacZYA-argF)U169$	Laboratory stock
Plasmids		
pAA8	41 bp fragment containing FLAG sequence in pUC18 <i>Hind</i> III site, Amp ^R	Amer & Valvano (2000)
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	Marolda <i>et al.</i> (2004)
pJD132	pBluescript SK derivative containing <i>wbaP</i> and flanking sequences from <i>E. coli</i> O9:K30, Amp ^R	Schäffer <i>et al.</i> (2002)
pMF24	pEXT21 derivative containing the <i>wzx</i> from <i>S. enterica</i> LT2, Sp ^R	Feldman <i>et al.</i> (1999)
pKD46	<i>bla</i> P _{BAD} <i>gam bet exo</i> pSC101 <i>ori</i> ^{ts}	Datsenko & Wanner (2000)
pKD3	<i>bla</i> FRT <i>cat</i> FRT PS1 PS2 <i>ori</i> _{R6K}	Datsenko & Wanner (2000)
pKP1	pBADFLAG derivative encoding the <i>S. enterica</i> LT2 WbaP _{M170-R354} C-terminally fused to the FLAG epitope	This study
pKP10	Identical insert as in pSM18 cloned into pBADNTF to provide an N-terminal FLAG tag fusion	This study
pKP12	Identical insert as in pSM30 cloned into pBADNTF to provide an N-terminal FLAG tag	This study
pKP13	Identical insert as in pSM31 cloned into pBADNTF to provide an N-terminal FLAG tag	This study
pKP14	Identical insert as in pSM28 cloned into pBADNTF to provide an N-terminal FLAG tag	This study
pMF24	pEXT21 derivative containing the <i>wzx</i> from <i>S. enterica</i> LT2	Feldman <i>et al.</i> (1999)
pSM10	pAA8 derivative containing <i>manB</i> and a functional <i>wbaP</i> from <i>S. enterica</i> Ty2 containing a 1 bp deletion 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. 3)	This study
pSM13	pSM10 derivative containing only <i>wbaP</i> from <i>S. enterica</i> Ty2	This study
pSM18	pSM13 derivative encoding <i>S. enterica</i> Ty2 WbaP _{M1-I144/F258-Y476} that has a deletion of the predicted periplasmic loop	This study
pSM22	pSM18 derivative encoding WbaP _{H76Y, K81E} to recreate the <i>wbaP4452(T)</i> mutation	This study
pSM23	pSM18 derivative containing a 1 bp deletion at position 385 of <i>wbaP</i> from <i>S. enterica</i> Ty2 to recreate the <i>wbaP4451(T)</i> mutation	This study
pSM28	pAA8 encoding WbaP _{M1-L203} from <i>S. enterica</i> Ty2	This study
pSM30	pAA8 encoding WbaP _{Y250-Y476} from <i>S. enterica</i> Ty2	This study
pSM31	pSM13 derivative encoding WbaP _{M1-T39/K305-Y476} from <i>S. enterica</i> Ty2	This study

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 (5'-CTCCCCGGGAATGGA-TAATATTGATAATAAG-3') and 1136 (5'-CTCCCCGGGTAAAT-CCCAAATAGTCCAGTG-3'), each incorporating *Sma*I sites (underlined). 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 (see above) and 897 (5'-CTCCCCGGGTTTAGCCATG-AAGTTATGTTATTAAGG-3') 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 (5'-TTACCCGGGAGTA-AAATAGACACATCCTAATGAAAGCC-3') and 1155 (5'-ATGCCC-GGGAAAGTTACTCGAGATGGTGGTCCG-3'), and primers 897 (see above) and 900 (5'-CTCCCCGGGATACCTAGCTTGTCAA-TAAATG-3'), 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 (5'-CTCCCCGGGATAATATTGATAATAAG-TATAATCCACAGC-3'; *Sma*I site underlined), and 1152 (5'-CTCTCTAGATTAATACGCACCATCTCGCCG-3'). The resulting

amplicons were digested with *Sma*I, treated with polynucleotide kinase and ligated to *Sma*I-digested and alkaline phosphatase-treated pBADNTE. Plasmid pKP1 was constructed from *S. Typhimurium* LT2 DNA by PCR using primers 2158 (5'-CTCCCCGGGTG-GCTCTGATAGCGTTTAC-3'; *Sma*I site underlined) and 2159 (5'-GACTGTCTGACTCCCTGGCAATAGGATCGTTAG-3'; *Sall* site underlined). The resulting amplicon was digested with *Sma*I and *Sall*I and ligated to the corresponding sites on pBADFLAG.

Disruption of the *wbaP* gene. Mutagenesis was performed according to the method described by Datsenko & Wanner (2000) 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 (5'-ATGGATAATATTGATAATAAGTATAATCCACAGCTATGTAAGT-GTAGGCTGGAGCTGCTTCG-3') and 614 (5'-ATCACTGCCATAC-CGACGACGCCGATCTGTTGCTTGGACATATGAATATCCTCCT-ATG-3'). Transformants were plated on LB agar plates containing chloramphenicol, for the selection of *wbaP* mutants.

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 1. Plasmid pSM23 was constructed using primers 886 (5'-CTTTGCCATAAT-CCTGGTGCTTTTTCGCGCACTTACAAAG-3') and 887 (5'-CTTTGTAAGTGCGCGAAAAAGCACCAGGATTATGGCAAAA-G-3'). Two consecutive site-directed mutagenesis experiments were performed to construct plasmid pSM22. Primers 882 (5'-GG-TTTTGGATTCGTTTGAGATATTATACATACCGCAAGCC-3') and 883 (5'-GGCTTGGCGTATGTATAATATCGCAAACGAATCCAAA-ACC-3') were used for the first PCR amplification, and primers 884 (5'-CGTTTGGGATATTATACATACCGCGAGCCATTTGGT-ATGAGTTAAAAG-3') and 885 (5'-CTTTAACTCATACCAAAAT-GGCTCGCGGTATGTATAATATCGCAAACG-3') for the second. The resulting plasmids were transformed into *E. coli* DH5 α by the calcium chloride method (Hanahan, 1983). Replacement mutants were confirmed by sequencing the entire *wbaP* gene.

Growth curves. For growth curves, overnight cultures of *E. coli* strains containing the appropriate plasmids were diluted to OD₆₀₀ 0.03 (corresponding to a bacterial density of $\sim 10^4$ c.f.u. ml⁻¹), and growth was monitored every 30 min in a 100-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 (Löwdin *et al.*, 1993). The results were analysed statistically by one-way ANOVA and the Tukey post-test using the Prism 4 software package (GraphPad Software).

LPS analysis. Culture samples were adjusted to OD₆₀₀ 2.0 in a final volume of 100 μ l. Then, proteinase-K-digested whole-cell lysates were prepared as described previously (Marolda *et al.*, 1990, 2006) and LPS was separated on 14% acrylamide gels using a Tricine-SDS buffer system (Marolda *et al.*, 2006). Gel loadings were normalized so that each sample represented the same number of cells. Each well was loaded with approximately 1×10^8 c.f.u. Gels were silver-stained as described previously (Marolda *et al.*, 2006). 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 Rdy800-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 (Marolda *et al.*, 1990). The O antigen polysaccharide bands in all gels were quantified by densitometry and the pixel densities were measured using ImageJ (Abramoff *et al.*, 2004).

In vitro transferase assay. Plasmids encoding the different WbaP recombinant proteins (pSM13, pSM18, pSM30 and pJD132) were transformed by electroporation into *S. Typhimurium* MSS2, which carries a *wbaP* deletion (Table 1). Membranes (containing enzymes and endogenous Und-P) were isolated from these transformants as described by Osborn *et al.* (1972). The reaction mixture for the transferase assay contained the membrane fraction (40 μ g total protein) and 0.5 μ Ci (18.5 kBq) radiolabelled UDP-[³H]Gal (Sigma) in 250 μ l buffer (5 mM Tris/acetate, pH 8.5, 0.1 mM EDTA, 3 mM MgCl₂). After incubation at 37 °C for 30 min, the lipid-associated material was extracted twice with 250 μ l 1-butanol. The combined 1-butanol extracts were washed once with 500 μ l 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. Enzyme activity was expressed as pmol [³H]Gal incorporated (mg membrane protein)⁻¹. For comparisons among the various preparations, relative enzymic activities were expressed as a percentage of parental WbaP activity at 1 mM MgCl₂.

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 OD₆₀₀ of 0.02, and incubated at 37 °C for 2 h until cultures reached OD₆₀₀ 0.5. At this point, arabinose was added to a final concentration of 0.2% (w/v), and cells were incubated for an additional 3 h until reaching OD₆₀₀ $\sim$ 0.8–1.0. Cells were then harvested by centrifugation at 10 000 *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 (15 000 *g* for 15 min at 4 °C), and the clear supernatant was centrifuged at 30 000 *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 (Amer & Valvano, 2002), 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.

RESULTS

WbaP has three unequally conserved predicted domains

Using widely employed algorithms for predicting trans-membrane helices and the overall in-out topology of the intervening loops (Drew *et al.*, 2002; Nilsson *et al.*, 2000), 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. 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 (Jiang *et al.*, 1991), 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 (Valvano, 2003).

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* $\Delta wbaP::cat$) and MSS2 (*S. Typhimurium* $\Delta wbaP::cat$) 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, 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 (Jiang *et al.*, 1991; Reeves, 1993; Wang & Reeves, 1994; Wang *et al.*, 1996). Therefore, we attempted to clone this gene as a single ORF downstream of the *lac* promoter region provided by the pAA8 plasmid vector (Amer & Valvano, 2000) (Table 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 1, pSM10), and subsequently deleted most of the *manB* gene (see Methods). The resulting plasmid, pSM13, restored O antigen expression in the $\Delta wbaP$ mutants MSS2 (*S. Typhimurium*) and MSS1 (*S. Typhi*) (Fig. 2, third lane in all panels), and was selected for further study. However,

Fig. 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.

Fig. 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* $\Delta wbaP::cat$; Ty2, *S. Typhi*; MSS1, *S. Typhi* $\Delta wbaP::cat$. 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.

less O antigen was produced by both the $\Delta wbaP$ mutant strains that contained pSM13, and the O antigen had a different migration pattern from that of the respective parental strains (Fig. 2). Transformation of MSS1 and MSS2 with pJD132, carrying an *E. coli* O9:K30 *wbaP* gene homologue and flanking sequences (Schäffer *et al.*, 2002), restored O antigen production in both mutant strains (Fig. 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, 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* DH5 α (pSM13) was significantly slower (doubling time 106 ± 3 min, $n=7$) than that of *E. coli* DH5 α 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) (Feldman *et al.*, 1999; Marolda *et al.*, 2004). Therefore, the growth defect caused by WbaP expression in *E. coli* could be due to accumulation of Und-PP-Gal in the 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* DH5 α (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 DH5 α (pAA8), confirming that introducing the *S. Typhimurium* *wzx* gene relieves the growth delay in *E. coli* DH5 α (pSM13).

The N-terminal domain of WbaP is dispensable for *in vivo* enzymic 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. 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

```

505  gagatgatggggtttgatgttatcgcttttttatacggatgcgtcagatgctgaaataaatatgttgccg
      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
      ra
577  gtgataaggacactgggactatttgggatttaaatcgctacaggtgatgtccattatattccttgcttaataac
      V I R T L G L F G I * M S I I S L L K Y
194                                215

```

Fig. 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.

(Wang & Reeves, 1994; Wang *et al.*, 1996). Therefore, various recombinant proteins were constructed to ascertain the functions of these two putative WbaP domains (Fig. 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. 1b; Table 1). This corresponds to the same end point of the frame-shift mutation in pSM13 (Fig. 3). Plasmid pSM30 encoded WbaP_{Y250-Y476} (Fig. 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. 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 (Wang *et al.*, 1996).

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. 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. 1b; Table 1). Fig. 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 (Wang *et al.*, 1996). 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 *wbaP*4451(T) and *wbaP*4452(T) described before (Wang *et al.*, 1996). Plasmid pSM23 carries a deletion at position 386 of the *wbaP* gene resulting

Fig. 4. The C-terminal and the last transmembrane domains of WbaP are required for *in vivo* enzymic 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* $\Delta wbaP::cat$. 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.

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. 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_{M1-1144/F258-Y476} (Fig. 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 1) with pJD132 that carries the *E. coli* K-30 *wbaP* gene (WbaP_{Eck30}). This plasmid complemented O antigen synthesis in MSS2 (Fig. 2, fourth lanes) but not in the T mutants (Fig. 5). Together, these results 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.

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. 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. 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. 6a). Since WbaP_{Y250-Y476} lacks most of the predicted periplasmic loop of WbaP (Fig. 1), we speculated that this region could be involved in O antigen chain-length

Fig. 5. The N-terminal domain of WbaP is dispensable for *in vivo* enzymic 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* $\Delta wbaP::cat$; SL1196, *S. Typhimurium* LT2 *wbaP4451(T)*; SL1197, *S. Typhimurium* LT2 *wbaP4452(T)*. MSS2 carried no plasmids, or plasmids pSM18 (encoding WbaP_{M1-1144/F258-Y476}), pSM23 [encoding WbaP_{M1-1144/F258-Y476} with additional substitutions recreating the *wbaP4451(T)* mutation] and pSM22 [encoding WbaP_{M1-1144/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.

Fig. 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* $\Delta wbaP::cat$, 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_{M1-1144/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).

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 5 and 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. 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. 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

Fig. 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 actinomycetem-comitans* (AAG49415.1), *Nostoc punctiforme* (ZP_00110718.1) and *Burkholderia xenovorans* (ZP_00031541.1).

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 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. 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. 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. 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. 1; Table 1). Introduction of pKP12 into strain LT2 (parental *wbaP*⁺) resulted in an O antigen profile with altered chain-length distribution (Fig. 8b), similar to that found in the $\Delta wbaP::cat$ mutant MSS2 complemented with pSM18

Fig. 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* $\Delta wbaP::cat$. (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_{F258–Y476}) on the O antigen pattern of strain LT2 in comparison with that of the control strains LT2 and LT2 (pBADFLAG).

(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. 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. 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.

Expression of WbaP protein derivatives as FLAG epitope fusions

To demonstrate that the truncated versions of WbaP are actually expressed 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* DH5 α . 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* DH5 α (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. 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. 9a, lane 2, asterisk). Also, a 20 kDa polypeptide was detected (Fig. 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 (Kashino, 2003). 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 (Amer & Valvano, 2000; Lehrer *et al.*, 2007; Marolda *et al.*, 2004). In this case, we also detected a polypeptide with an apparent mass of 26 kDa corresponding to WbaP_{M170-R354}-FLAG (Fig. 9a, lane 3, asterisk), and several larger oligomeric forms, most likely due to the mild denaturation conditions (Fig. 9a, lane 3, black arrows), as well as the putative degradation product of 20 kDa (Fig. 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. 1b; Table 1). All of these plasmids were derivatives of the vector pBADNTE, 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. 9b, arrows). Additional bands were also found in total membranes of MSS2 carrying the pBADNTE vector control (Fig. 9b), and correspond to polypeptides that cross-react with the M2 mAb. The presence of the FLAG epitope did not alter the results of complementation experiments done with the corresponding untagged proteins, as demonstrated in experiments in which the O antigen expression mediated by tagged and untagged versions of WbaP was assessed side by side (data not shown). Also, the genetic complementation did not correlate with the levels of protein expression, demonstrating that the levels of expressed protein did not cause the phenotypes observed with the various WbaP forms.

In vitro transferase activity of WbaP and mutant proteins reveals a requirement for an intact N-terminal domain

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

Fig. 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* DH5 α 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. 1b; Table 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.

assayed the enzymic activity of WbaP and mutant proteins by measuring the incorporation of radiolabelled [3 H]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 MSS2 (pJD132), which carries the wild-type WbaP protein from *E. coli* K-30. The levels of activity of membranes prepared from the negative control strain MSS2 alone and from MSS2 with plasmids pSM13 and pSM30 were dramatically reduced (Table 2), 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. While, as discussed above, pSM13 probably expresses the WbaP N- and C-terminal domains independently as separate polypeptides, pSM30 can only produce the C-terminal region. In contrast, membranes prepared from MSS2 containing pSM18 showed a similar transferase activity to that of membranes from MSS2 (pJD132). Therefore, the results of the enzymic assays suggest that the mutated forms of WbaP that lack the N-terminal domain

are enzymically inactive *in vitro*, despite being able to complement the synthesis of O antigen *in vivo*.

DISCUSSION

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 (Valvano, 2003), and also in the biosynthesis of the glycan moiety attached to bacterial glycoproteins (Steiner *et al.*, 2007). The predicted topology of WbaP revealed important differences from WecA, which belongs to another large family of initiating enzymes in prokaryotes and eukaryotes that catalyse the transfer of *N*-acetylhexosamine-phosphate onto isoprenoid lipid intermediates (Valvano, 2003). WecA consists of 11 transmembrane helices, five external and five internal loops (Amer & Valvano, 2001, 2002; Anderson *et al.*, 2000; Lehrer *et al.*, 2007). 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

Table 2. Comparative transferase activities of WbaP and its mutated derivatives

Membrane extract*	WbaP protein form	Transfer assay†	Complementation of O antigen synthesis‡
MSS2	None	8.0 ± 2.8	–
MSS2 (pJD132)	Complete	100	++
MSS2 (pSM13)	Independent N- and C-terminal	2.2 ± 0.2	+
MSS2 (pSM30)	C-terminal	9.0 ± 2.5	++ (defective chain-length distribution)
MSS2 (pSM18)	Deleted periplasmic loop	108.0 ± 8.5	++ (defective chain-length distribution)

*Membrane extracts were prepared from MSS2 cells alone or transformed with the plasmids indicated in parentheses (Table 1).

†The transfer assay measured the incorporation of radioactive Gal into a butanol-extractable fraction, as described in Methods. The results, expressed as a percentage of the parental WbaP activity at 1 mM MgCl₂ concentration using normalized counts, represent the mean ± SD of three independent experiments.

‡Results of *in vivo* complementation of O antigen synthesis: ++, similar to parental strain; +, reduced O antigen production; –, absence of O antigen (see Figs 4–6).

domain structure in this protein is not obvious (Amer & Valvano, 2001, 2002; Lehrer *et al.*, 2007). 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) 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 (Wang *et al.*, 1996). 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 (Adler & Bibi, 2004; Bibi & Kaback, 1990). 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 (Wang & Reeves, 1994; Wang *et al.*, 1996), 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 (Wang & Reeves, 1994; Wang *et al.*, 1996) 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, but is essential for *in vitro* enzymic activity using crude cell-membrane extracts. The combined results of our experiments suggest that the N-terminal domain of WbaP plays a role in the stability or correct folding of WbaP in the plasma membrane. 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 (Morona *et al.*, 2000). 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 (Delgado *et al.*, 2006), 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.

ACKNOWLEDGEMENTS

The authors thank the colleagues referred to or mentioned in Table 1 for strains and plasmids, and K. E. Sanderson, University of Calgary, for supplying strains SL1196 and SL1197. This work was supported by grants from the Canadian Institutes of Health Research (to M. A. V.) and grant 1040562 from Fondecyt (to I. C.). M. S. S. was supported by a fellowship from Conicyt and fellowships PG/108/02 and PG/92/2003 from Departamento de Postgrado y Postítulo, Universidad de Chile, and M. B. was supported by Beca de Tesis Doctoral Conicyt. M. A. V. holds a Canada Research Chair in Infectious Diseases and Microbial Pathogenesis.

REFERENCES

- Abramoff, M. D., Magelhaes, P. J. & Ram, S. J. (2004). Image processing with ImageJ. *Biophotonics Int* **11**, 36–43.
- Adler, J. & Bibi, E. (2004). Determinants of substrate recognition by the *Escherichia coli* multidrug transporter MdfA identified on both sides of the membrane. *J Biol Chem* **279**, 8957–8965.
- Alexander, D. C. & Valvano, M. A. (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**, 7079–7084.
- Amer, A. O. & Valvano, M. A. (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.
- Amer, A. O. & Valvano, M. A. (2001). Conserved amino acid residues found in a predicted cytosolic domain of WecA (UDP-*N*-acetylglucosamine: undecaprenol-phosphate *N*-acetylglucosamine-1-phosphate transferase) are implicated in the recognition of UDP-*N*-acetylglucosamine. *Microbiology* **147**, 3015–3025.
- Amer, A. O. & Valvano, M. A. (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**, 571–582.
- Anderson, M. S., Eveland, S. S. & Price, N. P. (2000). Conserved cytoplasmic motifs that distinguish sub-groups of the polyprenol phosphate: *N*-acetylhexosamine-1-phosphate transferase family. *FEMS Microbiol Lett* **191**, 169–175.
- Arakawa, Y., Wacharotayankun, R., Nagatsuka, T., Ito, H., Kato, N. & Ohta, M. (1995). Genomic organization of the *Klebsiella pneumoniae* cps region responsible for serotype K2 capsular polysaccharide synthesis in the virulent strain Chedid. *J Bacteriol* **177**, 1788–1796.
- Batchelor, R. A., Haraguchi, G. E., Hull, R. A. & Hull, S. I. (1991). Regulation by a novel protein of the bimodal distribution of lipopolysaccharide in the outer membrane of *Escherichia coli*. *J Bacteriol* **173**, 5699–5704.
- Bibi, E. & Kaback, H. R. (1990). *In vivo* expression of the *lacY* gene in two segments leads to functional *lac* permease. *Proc Natl Acad Sci U S A* **87**, 4325–4329.
- Bugert, P. & Geider, K. (1995). Molecular analysis of the *ams* operon required for exopolysaccharide synthesis in *Erwinia amylovora*. *Mol Microbiol* **15**, 917–933.

- Cartee, R. T., Forsee, W. T., Bender, M. H., Ambrose, K. D. & Yother, J. (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**, 7425–7433.
- Datsenko, K. A. & Wanner, B. L. (2000). One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc Natl Acad Sci U S A* **97**, 6640–6645.
- Delgado, M. A., Mouslim, C. & Groisman, E. A. (2006). The PmrA/PmrB and RcsC/YojN/RcsB systems control expression of the *Salmonella* O-antigen chain length determinant. *Mol Microbiol* **60**, 39–50.
- Drew, D., Sjöstrand, D., Nilsson, J., Urbig, T., Chin, C.-n., de Gier, J.-W. & von Heijne, G. (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**, 2690–2695.
- Drummlsmith, J. & Whitfield, C. (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**, 1321–1332.
- Feldman, M. F., Marolda, C. L., Monteiro, M. A., Perry, M. B., Parodi, A. J. & Valvano, M. A. (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**, 35129–35138.
- Hanahan, D. (1983). Studies on transformation of *Escherichia coli* with plasmids. *J Mol Biol* **166**, 557–580.
- Jiang, X. M., Neal, B., Santiago, F., Lee, S. J., Romana, L. K. & Reeves, P. R. (1991). Structure and sequence of the *rfb* (O antigen) gene cluster of *Salmonella* serovar typhimurium (strain LT2). *Mol Microbiol* **5**, 695–713.
- Kashino, Y. (2003). Separation methods in the analysis of protein membrane complexes. *J Chromatogr B Analyt Technol Biomed Life Sci* **797**, 191–216.
- Katzen, F., Ferreira, D. U., Oddo, C. G., Ielmini, M. V., Becker, A., Puhler, A. & Ielpi, L. (1998). *Xanthomonas campestris* pv. *campestris* gum mutants: effects on xanthan biosynthesis and plant virulence. *J Bacteriol* **180**, 1607–1617.
- Keenleyside, W. J. & Whitfield, C. (1996). A novel pathway for O-polysaccharide biosynthesis in *Salmonella enterica* serovar Borreze. *J Biol Chem* **271**, 28581–28592.
- Lehrer, J., Vigeant, K. A., Tatar, L. D. & Valvano, M. A. (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**, 2618–2628.
- Löwdin, E., Odenholt-Tornqvist, I., Bengtsson, S. & Cars, O. (1993). A new method to determine postantibiotic effect and effects of subinhibitory antibiotic concentrations. *Antimicrob Agents Chemother* **37**, 2200–2225.
- Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982). *Molecular Cloning: a Laboratory Manual*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.
- Marolda, C. L., Welsh, J., Dafoe, L. & Valvano, M. A. (1990). Genetic analysis of the O7-polysaccharide biosynthesis region from the *Escherichia coli* O7:K1 strain VW187. *J Bacteriol* **172**, 3590–3599.
- Marolda, C. L., Vicarioli, J. & Valvano, M. A. (2004). Wzx proteins involved in O antigen biosynthesis function in association with the first sugar of the O-specific lipopolysaccharide subunit. *Microbiology* **150**, 4095–4105.
- Marolda, C. L., Lahiry, P., Vinés, E., Saldías, S. & Valvano, M. A. (2006). Micromethods for the characterization of lipid A-core and O-antigen lipopolysaccharide. *Methods Mol Biol* **347**, 237–252.
- Morona, R., Van Den Bosch, L. & Daniels, C. (2000). Evaluation of Wzz/MPA1/MPA2 proteins based on the presence of coiled-coil regions. *Microbiology* **146**, 1–4.
- Nilsson, J., Persson, B. & von Heijne, G. (2000). Consensus predictions of membrane protein topology. *FEBS Lett* **486**, 267–269.
- Osborn, M. J., Gander, J. E., Parisi, E. & Carson, J. (1972). Mechanism of assembly of the outer membrane of *Salmonella typhimurium*. Isolation and characterization of cytoplasmic and outer membrane. *J Biol Chem* **247**, 3962–3972.
- Price, N. P. & Momany, F. A. (2005). Modeling bacterial UDP-HexNAc: polyisoprenol-P HexNAc-1-P transferases. *Glycobiology* **15**, 29R–42R.
- Raetz, C. R. H. & Whitfield, C. (2002). Lipopolysaccharide endotoxins. *Annu Rev Biochem* **71**, 635–700.
- Reeves, P. R. (1993). Evolution of *Salmonella* O antigen variation by interspecies gene transfer on a large scale. *Trends Genet* **9**, 17–22.
- Schäffer, C., Wugeditsch, T., Messner, P. & Whitfield, C. (2002). Functional expression of enterobacterial O-polysaccharide biosynthesis enzymes in *Bacillus subtilis*. *Appl Environ Microbiol* **68**, 4722–4730.
- Steiner, K., Novotny, R., Patel, K., Vinogradov, E., Whitfield, C., Valvano, M. A., Messner, P. & Schaffer, C. (2007). Functional characterization of the initiation enzyme of S-layer glycoprotein glycan biosynthesis in *Geobacillus stearothermophilus* NRS 2004/3a. *J Bacteriol* **189**, 2590–2598.
- Stevenson, G., Andrianopoulos, K., Hobbs, M. & Reeves, P. R. (1996). Organization of the *Escherichia coli* K-12 gene cluster responsible for production of the extracellular polysaccharide colanic acid. *J Bacteriol* **178**, 4885–4893.
- Valvano, M. A. (2003). Export of O-specific lipopolysaccharide. *Front Biosci* **8**, s452–s471.
- Wang, L. & Reeves, P. R. (1994). Involvement of the galactosyl-1-phosphate transferase encoded by the *Salmonella enterica rfbP* gene in O-antigen subunit processing. *J Bacteriol* **176**, 4348–4356.
- Wang, L., Liu, D. & Reeves, P. R. (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**, 2598–2604.
- Wilkinson, R. G., Gemski, P. & Stocker, B. A. D. (1972). Non-smooth mutants of *Salmonella typhimurium*: differentiation by phage sensitivity and genetic mapping. *J Gen Microbiol* **70**, 527–554.

Edited by: P. H. Everest